



Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational
Mathematics (IMACM)

Preprint BUW-IMACM 22/15

Hanno Gottschalk, Nicolai Rothe und Daniel Siemssen

Cosmological de Sitter Solutions of the Semiclassical Einstein Equation

June 15, 2022

<http://www.imacm.uni-wuppertal.de>

Cosmological de Sitter Solutions of the Semiclassical Einstein Equation

Hanno Gottschalk, Nicolai Rothe and Daniel Siemssen

School of Mathematics and Natural Science & IMACM,
University of Wuppertal, D-42119 Wuppertal, Germany
{gottschalk,rothe,siemssen}@uni-wuppertal.de

June 15, 2022

Exponentially expanding space-times play a central role in contemporary cosmology, most importantly in the theory of inflation and in the Dark Energy driven expansion in the late universe. In this work, we give a complete list of de Sitter solutions of the semiclassical Einstein equation (SCE), where classical gravity is coupled to the expected value of a renormalized stress-energy tensor of a free quantum field. Concretely, we study the stress-energy tensor of a scalar field in the (preferred) Bunch-Davies state on the cosmological coordinate patch using the recently proposed ‘moment approach’. From the energy component of the SCE we thus obtain an analytic consistency equation for the model’s parameters which has to be fulfilled by solutions to the SCE. Using this equation we then investigate the number of solutions and the structure of the solution set in dependency on the coupling parameter of the quantum field to the scalar curvature and renormalization constants using analytic arguments in combination with numerical evidence. We also identify parameter sets where multiple expansion rates separated by several orders of magnitude are possible. Potentially for such parameter settings, a fast (semi-stable) expansion in the early universe could be compatible with a late time ‘Dark Energy-like’ behavior of the universe.

Key words: Semiclassical Einstein equation • de Sitter cosmology • Bunch-Davies vacuum state • inflation in cosmology

MSC (2020): 83C47 • 83C56 • 81T20

1. Introduction

In modern cosmology, the Λ CDM is considered the standard model as it explains a large amount of observational data (see e.g. [48] and references therein). However, one of its predictions is the presence of Dark Energy or, equivalently, a positive cosmological constant. While matter in this model happens to be purely classical, it is one possible option that Dark Energy may naturally emerge from quantum effects, that is, if the matter content is modelled by a quantum field.

Physicists have put a lot of effort into deriving a satisfactory quantum theory of gravity. For an extensive survey on a wide range of quantum cosmological effects we refer to Schander and Thiemann [42]. The semiclassical Einstein equation (SCE) takes a rather moderate approach. In particular, gravity is avoided to be quantized and modelled by a classical metric formalism governed by an Einstein equation. The energy content, on the other hand, is modelled by quantum fields. The latter are then coupled to classical gravity via the expected value of the stress-energy tensor in a quantum state ω and the SCE reads

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa \langle T_{\mu\nu}^{\text{ren}} \rangle_{\omega}. \quad (1)$$

Hereby, $g_{\mu\nu}$ is the metric¹, $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu}$ is the Einstein curvature tensor (consisting of the Ricci curvature tensor $R_{\mu\nu}$ and its trace R) and Λ is the cosmological constant. $T_{\mu\nu}^{\text{ren}}$ is the renormalized quantum stress-energy (QSE) tensor of one or multiple quantum field(s), coupling to the geometry of the underlying space-time with a strength controlled by the parameter κ . We restrict to free scalar fields ϕ governed by the Klein-Gordon equation

$$(\square + \xi R + m^2)\phi = 0 \quad (2)$$

with mass m and curvature coupling ξ . $\square = -g^{\mu\nu}\nabla_{\mu}\nabla_{\nu}$ denotes the d'Alembertian of the metric $g_{\mu\nu}$.

The SCE has been introduced in a series of articles from the late 70's by Davies, Fulling et al. which culminated in [12]. The problem of finding a suitable QSE tensor was then axiomatized by Wald [49] and further refined by Christensen [8, 9], coming up with a properly covariant regularization scheme. The SCE (for cosmological settings) was approached by numerical algorithms and special analytic solutions have been found, e.g. by Anderson [2–5] or Suen & Anderson [46], as well as Starobinski [44]. Another modern view on the physical content of the SCE was provided by Flanagan & Wald in [17]. We refer to the monographs by Birrel & Davies [6], Fulling [20] and Wald [51] for a comprehensive view on the research on the SCE up to the 90's. In 2003, Moretti [37] defined a covariantly conserved QSE tensor as demanded in one of Wald's axioms, whereas in 2004, the same QSE was derived from a completely different point of view in [28].

A mathematical theory of solutions, tailored to cosmological settings, began to be developed in the late 2000's. In their work [10] the authors Dappiaggi, Fredenhagen & Pinamonti observed a distinguished behavior in their solutions which they call de Sitter-type behavior and we will pick up this discussion a bit later. Ongoing, the first

¹We use the signature convention $(-, +, +, +)$

result towards a mathematical solution theory, providing local existence and uniqueness results for the trace of the SCE, was formulated in the seminal article [39] by Pinamonti. This approach was further refined, particularly studying global properties of solutions and their continuability, by Pinamonti & Siemssen in [40, 43]. These works, similarly to many of the older references cited above, focused the conformally coupled case. The review articles by Fredenhagen & Hack [18] and by Hack [24, 26] built bridges between purely mathematical solution theory and the modern physicists’ approaches to cosmology. The results of [40] were recently generalized to non-conformally coupled fields by Meda, Pinamonti & Siemssen in [36]. Therein, the SCE is reformulated as a fixed point equation of a certain operator on a suitable Banach space, which allows to conclude short-time existence and uniqueness of solutions by a fixed-point theorem. Moreover, in the same period of time Eltzner & Gottschalk [13] managed to write the SCE into one dynamical system for both the scaling factor (of the underlying Friedman-Lemaître-Robertson-Walker (FLRW) space-time) and for the expectation values of Wick products of the field and its derivatives. This result was further refined and reformulated in a rigorous framework in [22], where the authors prove (global) existence and uniqueness of solutions. The latter approach, however, has the disadvantage of a somewhat implicit definition of the initial state.

Recently, the SCE has been used to derive special cosmological models. Sanders constructed in [41] maximally symmetric states on a given static (cosmological) space-time of positive (spatial) curvature. Moreover, in [21] the authors of the present article have used the techniques of [22] to study a class of cosmological expansion models driven by a massless scalar field in a Minkowski-like state, motivated by some observation on the Minkowski-vacuum state on Minkowski space. A recent paper by Juárez-Aubry [30] studies the SCE on static and ultrastatic, not necessarily spatially homogeneous space-times as an initial value problem for the state. This work was recently extended to more generic settings and applied to so-called ‘quantum state collapse’ scenarios in [31]. A noteworthy recent result is given in [27], where the author studies the ordinary differential equations for $H = \frac{\dot{a}}{a}$ arising from the conformally coupled and massless SCE and to a certain extent classifies the corresponding dynamical systems in terms of the topological properties of their phase portraits. Another work on special solutions is [29], also by Juárez-Aubry. There the author finds solutions of the SCE which coincide with the classical vacuum solutions for a positive cosmological constant Λ , that is, the de Sitter solution with constant curvature $R = \Lambda$ or, equivalently, $a(t) = \exp(\sqrt{\Lambda/3} t)$. In particular, the expectation value of the QSE tensor is taken with respect to the so-called Bunch-Davies state as first introduced by Bunch & Davies in [7] and further discussed by Allen in [1]. By isometrically embedding a cosmological space-time with pure de Sitter expansion into de Sitter space, this distinguished Bunch-Davies state can be pulled back and indeed yields a global state on the given de Sitter-type cosmological space-time. However, searching for these particular (vacuum) solutions of the SCE corresponds to solving $\langle T_{\mu\nu}^{\text{ren}} \rangle_{\omega} = 0$. (Note that $G_{\mu\nu} + \Lambda g_{\mu\nu} = 0$ vanishes for the currently discussed metric/scale factor a). Thus [29] is studying the parameter set for which the presence of the Bunch-Davies vacuum state has no back-reaction effect to the space-time. While this approach is suitable for the treatment of the so-called *cosmological constant problem* (as

in [29]), it does not cover all cases where a de Sitter expansion and the Bunch-Davies state on the resulting space-time yield a solution to the cosmological SCE. In particular, it omits situations when the presence of the vacuum state does have a back-reaction effect.

Another perspective on studying exponential late-time behavior of cosmological expansions deals with the topic of so-called energy inequalities (EIs). These kinds of considerations are based on an observation by Wald [50] from '83, namely that FLRW solutions to the Einstein equation with positive Λ generally approach the respective exponential vacuum solution with Hubble rate $\sqrt{\Lambda/3}$, provided the stress-energy tensor fulfills these EIs. However, QSE tensors usually do not fulfill the EIs in a pointwise manner and the focus has shifted towards studying whether the observation of [50] remains true if the stress-energy tensor fulfills similar weaker inequalities, for example where now the stress-energy tensor is averaged along time-like geodesics. For further reading we refer to some recent articles on this topic, e.g. to Fewster & Kontou [14], Fewster & Smith [15], Fewster & Verch [16] and, in particular, to the comprehensive introduction by Kontou & Sanders in [32].

Apart from late time de Sitter solutions, many authors also discuss inflation, i.e. solutions with a de Sitter phase at early time. An inflationary phase in the universe's expansion was originally suggested as a solution to the so-called cosmic horizon problem [23, 33, 34]. The main architects of inflationary physics in its modern shape are Guth, Linde and Starobinski, with their most noteworthy articles on that topic [23, 34, 45], respectively, all from the early 1980's. In particular, Starobinski addresses the compatibility of an inflationary phase with semiclassical gravity, and with this purpose notes the existence of pure de Sitter expansion solutions to the SCE as mentioned above. Moreover, considerable progress was made by Mukhanov [38] and others to explain the scale free spectrum of cosmological structures. For comprehensive discussions we refer to reviews of Liddle [33] and of Hack [26], where the latter in particular discusses inflationary models in view of modern algebraic quantum field theory.

Our present work is dedicated to finding *all* solutions to the cosmological SCE whose scaling factor describes a purely exponential expansion, driven by a massless or massive scalar field in the (pullback) Bunch-Davies vacuum state.

The physical motivation is to identify parameter settings, where for both an inflationary phase and a late-time de Sitter phase there are two (or more) exact solutions of the aforementioned kind which approximate the universe's expansion during these phases. Suppose that in future work one can show that the solution with a larger Hubble rate is unstable towards perturbations, that the solution with smaller rate is stable and that these two solutions are connected by a trajectory in phase space. Then one would show that a scalar quantum field is capable to drive both inflation and a late-time Dark Energy-dominated expansion. Note that such stability/instability behavior of de Sitter solutions have been found for simplified semiclassical models in [10] and [27] (cf. the discussion in Section 7).

On the other hand, a complete list of de Sitter solutions is interesting in its own right since these solutions frequently occur in semiclassical cosmology (e.g. [10, 21, 44]). Hereby the pullback Bunch-Davies state can be viewed as distinguished by its symmetry

behavior on (the entire, non-cosmological) de Sitter space. Note that while we are mostly interested in the cosmological setting in order to make sense of any stability features, we also obtain, as a by-product, a list of all parameters for which (the entire) de Sitter space and thereon the (non-pullback) Bunch-Davies solve the SCE.

We summarize our results in the following main theorem:

Theorem 1.1. *For $H > 0$, consider the cosmological space-time with flat spatial sections defined by $a(t) = \exp(Ht)$ and thereon a free scalar quantum field ϕ in the state obtained by pulling back the Bunch-Davies vacuum on de Sitter space with radius $\frac{1}{H}$ along cosmological coordinates. The field dynamics is governed by the Klein-Gordon equation (2) with parameters m and ξ . Then:*

- (i) *The semiclassical Einstein equation (1) with coupling κ and cosmological constant Λ for this field and state breaks down into a (non-dynamic) consistency equation for the parameters $H, m, \xi, \kappa, \Lambda$ and the renormalization constants originating in ϕ 's stress-energy tensor. Of these parameters only four are independent.*
- (ii) *Viewing the consistency equation as a constraint on (ξ, H) -pairs with two (effective) remaining parameters, the solution set can be parameterized by analytic curves in the ξ - H -plane. In numbers, these are one or two curves if $m = 0$, or two or three curves if $m > 0$.*
- (iii) *The large- H and small- H asymptotics of the solution curves can be explicitly worked out.*

Note that, by time-reflection invariance, values $H < 0$ correspond to positive values $-H > 0$. The well known $H = 0$ -Minkowski case is not discussed here. In particular, with the knowledge of part (iii) of the theorem we can conclude:

Corollary 1.2. *There exist parameter settings m, ξ, κ, Λ and renormalization constants, such that multiple H -values solve the consistency equation. Moreover, the model of Theorem 1.1 in the case $m > 0$ is flexible enough such that for any two prescribed positive values of H there exist a set of remaining parameters such that both H -values are solutions.*

While the last statement remains true for an arbitrary triple of positive numbers, we are mostly interested in tuning parameters to obtain two prescribed solutions for the reason discussed above.

Note that while in the following the results are stated more precise and tailored to the respective cases, Theorem 1.1 and Corollary 1.2 immediately follow from Proposition 3.2, Theorems 4.2 and 5.3 as well as the discussions in Sections 2 and 7.

Moreover, note that some intermediate results were merely accessible by numerically evaluating certain functions, for example, such a function's positivity or negativity on the positive reals. In these cases we carefully worked out the asymptotic behavior of a given function before, eventually, evaluating it on a sufficiently dense and sufficiently wide-spread grid. Knowing of such a functions analyticity and identifying any precomputed

asymptotic, we have no doubt that the respective assertions are true, although they are not rigorously proven. In order to stick with the proposition labelling throughout the text, we also call such assertions *Lemma*. However, for clarification we highlighted the respective numerical demonstrations and close them by the rotated ‘q.e.d.-box’ \diamond , while analytical proofs are closed by the usual box \square .

Our paper is organized as follows: The second section contains the derivation of the consistency equation for the special case of cosmological de Sitter space-times and free scalar fields in the (pullback) Bunch-Davies vacuum state. As a main tool, we utilize the ‘moments’ approach to the SCE in [22]. Note that while the consistency equation as such could have been derived faster using the results of [47], we particularly develop a viewpoint in which the de Sitter solution correspond to a phase space trajectory for the SCE as a dynamical system.

The third section is dedicated to the massless case, where the consistency equation simplifies into an explicitly soluble polynomial equation. After solving the equation and plotting the solution sets in the ξ - H -plane we discuss the asymptotics of the solution curves.

In Sections 4, 5 and 6 we study the massive case. First, in Section 4, we exploit the fact that the solution set is the zero set of an analytic function. In particular, we show that any solution to the consistency equation belongs to an analytic solution curve which is extendible into the asymptotics of the equation. Secondly, in Section 5 we explicitly evaluate the asymptotics of said analytic function, completing the argument as it is presented in Theorem 1.1 above. Section 6 then graphically presents the solution set as obtained by numeric evaluation, confirming the results of the previous sections.

Section 7, finally, uses the results of the previous sections in order to show the existence of parameter settings for potential inflationary models as in Corollary 1.2.

In the last section we present our conclusion and discuss some open problems for future research.

2. The semiclassical Einstein equation on de Sitter space-time

In this section we derive the consistency condition for the parameters under which the cosmological SCE admits a solution with a pure de Sitter expansion law, driven by a scalar field in the (pullback) Bunch-Davies state.

2.1. The energy equation as a cosmological model

A priori, (1) is actually a system of 16 equations and by the symmetries of both the Einstein and the stress-energy tensor, these reduce to ten independent equations. For a (flat) cosmological FLRW metric

$$g = -dt^2 + a(t)^2(dy_1^2 + dy_2^2 + dy_3^2) \quad (3)$$

with scaling factor $a(t)$ of these ten equations only two are independent, for example the 00- and one of the jj -components ($j = 1, 2, 3$). These two independent equations can be

captured in the so-called energy and trace equations,

$$G_{00} - \Lambda = \kappa \langle T_{00}^{\text{ren}} \rangle_\omega \quad \text{and} \quad g^{\mu\nu} G_{\mu\nu} + 2\Lambda = \kappa g^{\mu\nu} \langle T_{\mu\nu}^{\text{ren}} \rangle_\omega, \quad (4)$$

respectively. We refer to the discussions in [21, 22] that for the cosmological metric (3) the Einstein tensor's components fulfill²

$$\frac{d}{dt}(a^2 G_{00}) + 2a\dot{a}G_{00} = -a\dot{a} g^{\mu\nu} G_{\mu\nu} \quad (5)$$

(and likewise do the components of $G_{\mu\nu} + \Lambda g_{\mu\nu}$), implying that the energy equation implies the trace equation whenever $\dot{a} \neq 0$. Thus, on any time-interval on which \dot{a} does not vanish, it is sufficient to study the energy equation.

2.2. De Sitter space and the Bunch-Davies state

De Sitter space is the four-dimensional one-sheet hyperboloid of a certain radius as a pseudo-Riemannian submanifold of five-dimensional Minkowski space $\mathbb{M} = \mathbb{R}^5$, oriented around the time axis of the latter. Formally, endow \mathbb{M} with the coordinates $z_0, z_1, z_2, z_3, z_4 : \mathbb{R} \rightarrow \mathbb{M}$ (with time axis along the z_0 coordinate), then de Sitter space is the set

$$\mathfrak{dS}_H = \left\{ (z_0, z_1, z_2, z_3, z_4) \in \mathbb{M} \mid -z_0^2 + \sum_{i=1}^4 z_i^2 = \frac{1}{H^2} \right\}$$

for some parameter $H > 0$, with the pullback Lorentzian metric via the canonical embedding $\mathfrak{dS}_H \rightarrow \mathbb{M}$. The group of isometries of \mathfrak{dS}_H is given by the full Lorentz group $O(4, 1)$ and \mathfrak{dS}_H , as a submanifold of \mathbb{M} , is left invariant under this group's action. In particular, the induced pullback metric is left invariant as well.

We choose the coordinates (t, y_1, y_2, y_3) such that

$$\begin{aligned} z_0 &= \frac{1}{H} \sinh(Ht) + \frac{1}{2} H e^{Ht} (y_1^2 + y_2^2 + y_3^2), \\ z_i &= e^{Ht} y_i \quad (i = 1, 2, 3), \\ z_4 &= \frac{1}{H} \cosh(Ht) - \frac{1}{2} H e^{Ht} (y_1^2 + y_2^2 + y_3^2), \end{aligned}$$

which cover $\widetilde{\mathfrak{dS}}_H = \{ (z_0, z_1, z_2, z_3, z_4) \in \mathfrak{dS}_H \mid z_0 + z_4 \geq 0 \}$, the so-called the cosmological patch of \mathfrak{dS}_H . Pulling the metric of \mathbb{M} back to $\widetilde{\mathfrak{dS}}_H$ through these coordinates we obtain the metric

$$\begin{aligned} g &= -dt^2 + e^{2Ht} (dy_1^2 + dy_2^2 + dy_3^2) \\ &= \frac{1}{H^2 \tau^2} (-d\tau^2 + dy_1^2 + dy_2^2 + dy_3^2) \end{aligned} \quad (6)$$

on \mathbb{R}^4 or $\mathbb{R}_{\tau>0}^4$, respectively, where for the latter representation of the metric we defined

²We denote derivatives of a with respect to cosmological time t by dots, i.e. \dot{a} , \ddot{a} and so on.

the conformal time coordinate

$$\tau(t) = \int_0^t e^{-Ht'} dt'.$$

Hence (the cosmological patch of) de Sitter space can be regarded as flat FLRW-type space and comparing (6) with (3) we identify the scale factor $a(t) = e^{Ht}$ in cosmological time t and $a(\tau) = \frac{1}{H\tau}$ in conformal time τ . In the following we will sloppily speak of $\mathbb{R}_{\tau>0}^4$, endowed with the metric (6), as cosmological de Sitter space or simply as de Sitter space-time.

Generally in algebraic QFT, one major difficulty is to define a state of the field algebra. On de Sitter space \mathfrak{dS}_H there exists a preferred choice of such, namely the Bunch-Davies state [1, 7]. Among all $O(4, 1)$ -invariant states discussed in [1] the Bunch-Davies state is the only Hadamard state, suggesting it as a natural choice of vacuum state. Note that, in order for the Bunch-Davies state to exist we have to assume that the effective de Sitter mass of the field is positive, $m^2 + 12\xi H^2 > 0$ (cf. [1]).

As a quasi-free state [51] the Bunch-Davies state is determined by its two-point function

$$\omega_2^{\text{BD}}(y, z) = \frac{2(6\xi - 1)H^2 + m^2}{8\pi \cos(\pi\nu)} {}_2F_1\left(\frac{3}{2} + \nu, \frac{3}{2} - \nu; 2; \frac{1}{2}(1 + \mathcal{Z}(y, z))\right) \quad (7)$$

($y, z \in \mathfrak{dS}_H$), where ${}_2F_1$ is the hypergeometric function, $\nu = \sqrt{\frac{9}{4} - 12\xi - \frac{m^2}{H^2}}$ and $\mathcal{Z}(y, z)$ is the chord length between y and z ,

$$\mathcal{Z}(y, z) = \mathcal{Z}((\tau_y, \mathbf{y}), (\tau_z, \mathbf{z})) = \frac{\tau_y^2 + \tau_z^2 - (\mathbf{y} - \mathbf{z})^2}{2\tau_y\tau_z}$$

(in conformal-time cosmological coordinates). For this particular representation of the Bunch-Davies state's two-point function we refer to [43], see also [1] for a similar representation.

2.3. The consistency equation in the moment based approach

Starting from the shape of ω_2^{BD} from above, we can, in principle, evaluate all terms constituting $\langle T_{\mu\nu}^{\text{ren}} \rangle_\omega$ following [37, 49]. By inserting the de Sitter expansion $a(\tau) = \frac{1}{H\tau}$ and ω_2^{BD} into the SCE, the dynamic aspect is eliminated and we obtain an equation for the parameters of the model, similarly as in [29]. Note that the latter reference restricts to the case $H = \sqrt{\Lambda/3}$, i.e. to solutions of $\langle T_{\mu\nu}^{\text{ren}} \rangle_\omega = 0$. Moreover, note that while we are mainly interested in the cosmological setting in which a formulation of the SCE as a dynamical system is elaborated, an alternative approach is to use the stress-energy tensors of a scalar field on (the entire) de Sitter space from [47]. Up to some details (cf. Remark 4.10), however, this results in studying the same consistency equations for the parameters. We, however, follow the approach of [22] and view the SCE on a flat FLRW space-time as a dynamical system for both the scaling factor a and a sequence of ‘moments’ derived from the state’s two-point function via a specific ‘cosmological’

parametrix.

In the following we will shortly recapitulate on both the general quantization procedure of a scalar field with the particular goal of a well-defined $\langle T_{\mu\nu}^{\text{ren}} \rangle_\omega$ and on the approach of [22]. This rather serves as an introduction of relevant notation than as a complete description; for details we refer the reader to the pertinent literature cited in Section 1.

We start from the classical stress-energy tensor

$$T_{\mu\nu} = (1 - 2\xi)(\nabla_\mu\phi)(\nabla_\nu\phi) - \frac{1}{2}(1 - 4\xi)g_{\mu\nu}(\nabla^\sigma\phi)(\nabla_\sigma\phi) - \frac{1}{2}g_{\mu\nu}m^2\phi^2 + \xi(G_{\mu\nu}\phi^2 - 2\phi\nabla_\mu\nabla_\nu\phi - 2g_{\mu\nu}\phi\Box\phi)$$

of a classical scalar field ϕ governed by the Klein-Gordon equation (2). The QSE tensor of ϕ after quantization is then obtained by replacing ϕ and its derivatives with their respective quantum counterparts, that is, by the coincidence limit of (derivatives of) the regularized two-point function of a Hadamard state ω . If we denote by $\tilde{H}(y, z)$ the (possibly truncated) distributional kernel of the Hadamard parametrix, the coincidence limit of the regularized two-point function is given by $[\omega_2 - \tilde{H}] := \lim_{z \rightarrow y} (\omega_2(y, z) - \tilde{H}(y, z))$ with the (unregularized) two-point function $\omega_2(y, z) = \omega(\phi(y)\phi(z))$. Accordingly, we have $[(\nabla_\mu \otimes \nabla_\nu)(\omega_2 - \tilde{H})] := \lim_{z \rightarrow y} (\nabla_\mu)_y (\nabla_\nu)_z (\omega_2(y, z) - \tilde{H}(y, z))$ and so on. As usual, a conserved renormalization scheme such as Moretti's [37] is mandatory and thus $\langle T_{\mu\nu}^{\text{ren}} \rangle_\omega$ additionally contains a trace anomaly term $\frac{1}{4\pi^2}g_{\mu\nu}[\nu_1]$ (with the Hadamard coefficient ν_1). By this scheme, the QSE tensor indeed obeys $\nabla^\mu \langle T_{\mu\nu}^{\text{ren}} \rangle_\omega = 0$. Finally, we add the renormalization freedom $c_1 m^4 g_{\mu\nu} + c_2 m^2 G_{\mu\nu} + c_3 I_{\mu\nu} + c_4 J_{\mu\nu}$ in terms of four independent parameters c_1, c_2, c_3, c_4 . We refer to [24] and references therein for precise formulas regarding $\tilde{H}, \nu_1, I_{\mu\nu}$ and $J_{\mu\nu}$. Concluding, the QSE tensor is of the form

$$\begin{aligned} \langle T_{\mu\nu}^{\text{ren}} \rangle_\omega = & \quad (8) \\ & (1 - 2\xi)[(\nabla_\mu \otimes \nabla_\nu)(\omega_2 - \tilde{H})] - \frac{1}{2}(1 - 4\xi)g_{\mu\nu}[(\nabla^\sigma \otimes \nabla_\sigma)(\omega_2 - \tilde{H})] - \frac{1}{2}g_{\mu\nu}m^2[\omega_2 - \tilde{H}] \\ & + \xi(G_{\mu\nu}[\omega_2 - \tilde{H}] - 2[(\mathbb{1} \otimes \nabla_\mu \nabla_\nu)(\omega_2 - \tilde{H})] - 2g_{\mu\nu}[(\mathbb{1} \otimes \Box)(\omega_2 - \tilde{H})]) \\ & + \frac{1}{4\pi^2}g_{\mu\nu}[\nu_1] + c_1 m^4 g_{\mu\nu} + c_2 m^2 G_{\mu\nu} + c_3 I_{\mu\nu} + c_4 J_{\mu\nu}. \end{aligned}$$

Back in the cosmological setting (3), under the assumption that $\omega_2(y, z)$ at points $y = (\tau, \mathbf{y})$ and $z = (\hat{\tau}, \mathbf{z})$ merely depends on $r = |\mathbf{y} - \mathbf{z}|$, the derivatives of ω_2 relevant for (8) are stored in a vector

$$\mathcal{G}(\tau, r) := \begin{pmatrix} \mathcal{G}_{\varphi\varphi}(\tau, r) \\ \mathcal{G}_{(\varphi\pi)}(\tau, r) \\ \mathcal{G}_{\pi\pi}(\tau, r) \end{pmatrix} := \lim_{\hat{\tau} \rightarrow \tau} \begin{pmatrix} \mathbb{1} \\ \frac{1}{2}(\partial_\tau + \partial_{\hat{\tau}}) \\ \partial_\tau \partial_{\hat{\tau}} \end{pmatrix} a(\tau)a(\hat{\tau})\omega_2(\tau, \hat{\tau}, r). \quad (9)$$

The components of \mathcal{G} can be viewed as ‘semi-coincidence limit’ of ω_2 , i.e. the coincidence limit in direction of the τ -coordinate, and these limits exist by the Hadamard property of ω whenever $r \neq 0$. The singular structure of ω_2 is then represented by the singular structure of $\mathcal{G}(\tau, \cdot)$ in the limit $r \rightarrow 0$. The Klein-Gordon equation for ω_2 implies for \mathcal{G}

that

$$\partial_\tau \mathcal{G} = \begin{pmatrix} 0 & 2 & 0 \\ \Delta_r - V & 0 & 1 \\ 0 & 2(\Delta_r - V) & 0 \end{pmatrix} \mathcal{G} \quad (10)$$

with the (spatial) Laplacian $\Delta_r = r^{-2} \partial_r r^2 \partial_r$ and the potential³ $V = (6\xi - 1) \frac{a''}{a} + a^2 m^2$.

For an expression of the (full) coincidence limit of the regularized two-point function $[\omega_2 - \tilde{H}]$ one defines $\tilde{\mathcal{H}}(\tau, r)$ by the analog of formula (9) replacing ω_2 by \tilde{H} . Note that the Hadamard parametrix also depends only on $r = |\mathbf{y} - \mathbf{z}|$. The regularized two-point function and its derivatives from (8) may now be written as certain (linear combinations of) components of $\lim_{r \rightarrow 0} (\mathcal{G}(\cdot, r) - \tilde{\mathcal{H}}(\cdot, r))$, that is, they are obtained by completing the coincidence limit along the spatial coordinate directions.

The main innovation of [22] is now to introduce a new ‘cosmological’ parametrix \mathcal{H} . In principle it is constructed somewhat similar to \tilde{H} (or $\tilde{\mathcal{H}}$, resp.), but adapted to cosmological coordinates. Formally, fix an arbitrary length scale μ and define for $j \in \mathbb{Z}_{\geq -1}$ so-called homogeneous distributions, i.e. the functions $h_{2j} : (0, \infty) \rightarrow \mathbb{R}$,

$$h_{-2}(r) := -\frac{1}{\pi^2 r^4}, \quad h_0(r) := \frac{1}{2\pi^2 r^2}, \quad h_{2j}(r) := \frac{(-1)^j}{2\pi^2} \frac{r^{2(j-1)}}{\Gamma(2j)} \left(\log\left(\frac{r}{\mu}\right) - \psi^{(0)}(2j) \right)$$

with the Digamma function $\psi^{(0)} = \log(\Gamma)'$. Moreover, define

$$\mathcal{H}_n(\tau, r) := \begin{pmatrix} \mathcal{H}_{\varphi\varphi,n}(\tau, r) \\ \mathcal{H}_{(\varphi\pi),n}(\tau, r) \\ \mathcal{H}_{\pi\pi,n}(\tau, r) \end{pmatrix} := \begin{pmatrix} 0 \\ 0 \\ \gamma_{-1}(\tau) \end{pmatrix} h_{-2}(r) + \sum_{j=0}^n \begin{pmatrix} \alpha_j(\tau) \\ \beta_j(\tau) \\ \gamma_j(\tau) \end{pmatrix} h_{2j}(r) \quad (11)$$

with some sequences $(\alpha_j)_{j \geq 0}$, $(\beta_j)_{j \geq 0}$, $(\gamma_j)_{j \geq -1}$ of functions $\alpha_j, \beta_j, \gamma_j : (0, \infty) \rightarrow \mathbb{R}$. Similarly to the Hadamard parametrix the sum may be truncated at a sufficiently large order without affecting the final result, so one can omit questions of convergence. In [22] these functions are found such that \mathcal{H} fulfills the Klein-Gordon system

$$\partial_\tau \mathcal{H}_\infty = \begin{pmatrix} 0 & 2 & 0 \\ \Delta_r - V & 0 & 1 \\ 0 & 2(\Delta_r - V) & 0 \end{pmatrix} \mathcal{H}_\infty \in \mathcal{O}(r^\infty). \quad (12)$$

Hereby, the function class $\mathcal{O}(r^\infty)$ is to be read as that a truncation of (11) at some n yields an error in $\mathcal{O}(r^{m(n)})$ with $m(n) \rightarrow \infty$ as $n \rightarrow \infty$.

As a next step, we replace the regularized expressions for the two-point function and its derivatives in (8), i.e. $[\omega_2 - \tilde{H}]$, $[(\nabla_\mu \otimes \nabla_\nu)(\omega_2 - \tilde{H})]$ and so on, by the respective (linear combinations of) components of

$$\lim_{r \rightarrow 0} (\mathcal{G} - \tilde{\mathcal{H}}) = \lim_{r \rightarrow 0} (\mathcal{G} - \mathcal{H}) + \lim_{r \rightarrow 0} (\mathcal{H} - \tilde{\mathcal{H}}),$$

³Opposed to the convention in footnote 2 we denote derivatives of a with respect to conformal time τ by primes, i.e. a' , a'' and so on. Confusions in higher derivatives $a^{(j)}$ are excluded.

where indeed both limits on the RHS do exist. $\lim_{r \rightarrow 0} (\mathcal{H} - \widetilde{\mathcal{H}})$ does not depend on the precise choice of ω , and can be computed to result in a smooth function on the underlying space-time depending on a and its derivatives only. $\mathcal{G} - \mathcal{H}$ may be called the regularized two-point function in the \mathcal{H} -regularization scheme. Finally, define the moments of ω by

$$m_n := \lim_{r \rightarrow 0} \Delta_r^n (\mathcal{G} - \mathcal{H}),$$

that is, they can be thought of as (even-order, radial) Taylor coefficient of the (radially symmetric) function $\mathcal{G}(\tau, \cdot) - \mathcal{H}(\tau, \cdot)$. With (10) and (12) also $\mathcal{G} - \mathcal{H}$ fulfills the ($\mathcal{O}(r^\infty)$ -approximate) Klein-Gordon system (12) and we can reformulate the latter into a linear evolution equation for the function $\tau \mapsto \mathcal{M}(\tau)$ valued in a suitable Banach space. In this setting the SCE can be written as

$$\begin{cases} A'(\tau) &= V(A(\tau), \mathcal{M}(\tau)) \\ \mathcal{M}'(\tau) &= W(A(\tau)) \cdot \mathcal{M}(\tau) \end{cases} \quad (13)$$

with $A = (a, a', a'', a''')$ and some dynamic vector fields V and W , and the authors of [22] prove existence and uniqueness of the solutions. Note that the second line of (13) is a mere consequence of the Klein-Gordon equation, particularly it is independent of whether a is a solution to any cosmological model or not. The first line of (13) is, usually, obtained from the traced SCE, constrained by the energy equation (4). However, by (5) (or its conformal-time analog, resp.) we can use the energy equation for the first line of (13) in the context of pure de Sitter expansions. In this case A is of the form $A = (a, a', a'')$.

Performing the replacements described above, the energy evaluates to

$$\begin{aligned} 0 = & \left(6(3c_3 + c_4) + \frac{1}{960\pi^2} - \frac{6\xi - 1}{96\pi^2} - \frac{(6\xi - 1)^2}{32\pi^2} \log(a\lambda_0) \right) \\ & \cdot \left(2 \frac{a^{(3)}a'}{a^4} - \frac{(a'')^2}{a^4} - 4 \frac{a''(a')^2}{a^5} \right) \\ & - \frac{(6\xi - 1)^2}{16\pi^2} \frac{a''(a')^2}{a^5} + \frac{1}{960\pi^2} \frac{(a')^4}{a^6} + \left(\frac{\Lambda}{\kappa} - m^4 \left(c_1 + \frac{1}{32\pi^2} \log(a\lambda_0) \right) \right) a^2 \quad (14) \\ & + \left(-\frac{3}{\kappa} + m^2 \left(3c_2 - \frac{1}{96\pi^2} - \frac{6\xi - 1}{16\pi^2} (1 + \log(a\lambda_0)) \right) \right) \frac{(a')^2}{a^2} \\ & + \frac{m^2}{2} m_{\varphi\varphi,0} + (6\xi - 1) \left(-\frac{(a')^2}{2a^4} m_{\varphi\varphi,0} + \frac{a'}{a^3} m_{(\varphi\pi),0} \right) \\ & + \frac{1}{2a^2} (m_{\pi\pi,0} - m_{\varphi\varphi,1}), \end{aligned}$$

where λ_0 is the ratio of the Hadamard length scale and the length scale μ . For more details we refer the interested reader to [22]. Note that to obtain (14) from the latter reference we have merely included a possibly non-vanishing Λ .

As a side remark, note that the representation in (14) nicely shows how precise choices of both the involved length scales do not matter and any changes can be absorbed into

the renormalization constants.

A more or less straightforward computation yields the first moments of the Bunch-Davies state's two-point function (7) on a cosmological de Sitter space-time with $a(\tau) = \frac{1}{H\tau}$ as

$$\begin{aligned}
 m_{\varphi\varphi,0} &= -\frac{2(1-6\xi)H^2 - m^2}{16H^2\pi^2\tau^2} \left[1 + \log\left(\frac{\mu^2}{4\tau^2}\right) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right] \\
 m_{\varphi\varphi,1} &= -\frac{2(1-6\xi)H^2 - m^2}{128H^2\pi^2\tau^4} \\
 &\quad \cdot \left[18 + 84\xi + 7\frac{m^2}{H^2} + 6\left(12\xi + \frac{m^2}{H^2}\right) \left(\log\left(\frac{\mu^2}{4\tau^2}\right) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right) \right] \\
 m_{(\varphi\pi),0} &= \frac{2(1-6\xi)H^2 - m^2}{16H^2\pi^2\tau^3} \left[2 + \log\left(\frac{\mu^2}{4\tau^2}\right) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right] \\
 m_{\pi\pi,0} &= -\frac{2(1-6\xi)H^2 - m^2}{128H^2\pi^2\tau^4} \\
 &\quad \cdot \left[30 + 12\xi + \frac{m^2}{H^2} + 2\left(4 + 12\xi + \frac{m^2}{H^2}\right) \left(\log\left(\frac{\mu^2}{4\tau^2}\right) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right) \right],
 \end{aligned}$$

and for the combination of moments relevant in (14) we compute

$$\begin{aligned}
 &\frac{m^2}{2}m_{\varphi\varphi,0} + (6\xi - 1) \left(-\frac{(a')^2}{2a^4}m_{\varphi\varphi,0} + \frac{a'}{a^3}m_{(\varphi\pi),0} \right) + \frac{1}{2a^2}(m_{\pi\pi,0} - m_{\varphi\varphi,1}) \quad (15) \\
 &= \frac{2(1-6\xi)H^2 - m^2}{128\pi^2H^2\tau^2} \left[6(1-6\xi)H^2 - m^2 - 2m^2 \left(\log\left(\frac{\mu^2}{4\tau^2}\right) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right) \right].
 \end{aligned}$$

Recall that $\nu = \sqrt{\frac{9}{4} - 12\xi - \frac{m^2}{H^2}}$ and note that the Digamma function $\psi^{(0)}$ stems from taking derivatives of the hypergeometric function ${}_2F_1$ in the Bunch-Davies two-point function (7) and not from the occurrence of $\psi^{(0)}$ in the cosmological parametrix via the homogeneous distributions $(h_{2j})_{j \geq -1}$. Moreover, note how all moments and, in particular, the contribution (15) vanish in the massless conformally coupled case $m^2 = \xi - \frac{1}{6} = 0$.

Finally, plugging both the de Sitter Ansatz $a(\tau) = \frac{1}{H\tau}$ and the Bunch-Davies moments (15) into the energy equation (14), we arrive at the following form of the consistency equation

$$\begin{aligned}
 0 &= \left(\frac{1}{960} - \frac{(6\xi - 1)^2}{32} \right) H^4 - 3KH^2 + \Lambda K + m^4 d_1 + m^2 H^2 d_2 \quad (16) \\
 &\quad + \left(\frac{m^4}{64} + \frac{6\xi - 1}{32} m^2 H^2 \right) \left(2 \log(\mu H) + \psi^{(0)}\left(\frac{3}{2} - \nu\right) + \psi^{(0)}\left(\frac{3}{2} + \nu\right) \right).
 \end{aligned}$$

Here we have introduced linearly transformed renormalization constants

$$d_1 := \frac{1}{128} - \frac{1}{32} \log(2\lambda_0) - \pi^2 c_1 \quad \text{as well as} \quad d_2 := 3\pi^2 c_2 - \frac{1}{96} - \frac{6\xi - 1}{16} \log(2\lambda_0)$$

and we have set $K := \frac{\pi^2}{\kappa}$. We note that $I_{00} = J_{00} = 0$ for de Sitter expansions $a(\tau) = \frac{1}{H\tau}$, thus we are left with only two renormalization constants. Moreover, the $\log(\tau)$ -terms of (15) just cancel the $\log(\tau)$ -terms occurring in (14). This is not surprising if we recall that all of them originate in the cosmological parametrix \mathcal{H} . We are left only with the length-scale μ and observe (again) that a different choice of μ leads to additional terms which can be absorbed into the renormalization constants d_1 and d_2 .

3. De Sitter solutions for the massless field

In the massless case $m = 0$, the consistency equation (16) breaks down into the polynomial equation

$$0 = \left(\frac{1}{960} - \frac{(6\xi - 1)^2}{32} \right) H^4 - 3KH^2 + \Lambda K \quad (17)$$

for ξ and H with parameters Λ and K .

Remark 3.1. (i) *Throughout the present section we assume $\xi > 0$ in order to fulfill the existence condition for the Bunch-Davies state.*

(ii) *Some may take the standpoint that the prefactors m^2 (of $G_{\mu\nu}$) and m^4 (of $g_{\mu\nu}$) in the renormalization freedom of $\langle T_{\mu\nu}^{\text{ren}} \rangle_{\omega}$ are only chosen to endow the respective terms with the correct unit in order to obtain unit-free renormalization constants c_1 and c_2 . Consequently, in the massless case $m = 0$ these prefactors should be expressed by some other mass scale \tilde{m} in order to maintain this freedom. However, replacing the coupling constant K and the cosmological constant Λ in equation (17) by their renormalized analogs,*

$$\tilde{K} = K - \frac{\tilde{m}^2 d_2}{3} \quad \text{and} \quad \tilde{\Lambda} = \frac{3\Lambda K + 3\tilde{m}^4 d_1}{3K - \tilde{m}^2 d_2},$$

respectively, we end up with the very same equation with the only difference that \tilde{K} can be a non-positive parameter. Note also how (17) simplifies for $K = 0$ and how for $K < 0$ the analysis of (17) is (in a suitable sense) inverted around the zeros $\frac{1}{6} \pm 1/\sqrt{1080}$ of the H^4 -prefactor. However, we stick to the view that the renormalization freedom compensates ambiguities in the choice of the Hadamard length scale, that is, we always assume $K > 0$.

(iii) *Carrying out the computation to obtain (17) from (14) and (15) one notices that the Bunch-Davies state's moments only yield a contribution into the prefactor of H^4 in (17). Assuming vanishing moments $\mathcal{M} = 0$ instead, the fraction $(6\xi - 1)^2/32$ would be replaced by $(6\xi - 1)^2/8$. Hence, the analysis of the present section can be adjusted for the Minkowski vacuum-like states in [21] by squeezing any graphic by a factor of $1/2$ around $\xi = 1/6$. Note in particular the similarity between the $\Lambda = 0$ -curve in Figure 2 and the respective graphic in [21].*

We introduce

$$\xi_{\text{cc}} := \frac{1}{6}, \quad \xi_{(\pm)} := \xi_{\text{cc}} \pm \frac{1}{\sqrt{1080}} \quad \text{as well as} \quad H_{\text{vac}} := \sqrt{\frac{\Lambda}{3}}$$

(for $\Lambda > 0$) since these particular ξ - and H -values are distinguished by the behavior of the solution set of (17). Here, H_{vac} represents the unique (positive) de Sitter solution of the vacuum Einstein equation $G_{\mu\nu} + \Lambda g_{\mu\nu} = 0$ for $\Lambda > 0$.

We find the following:

Proposition 3.2. *Let $\Lambda \in \mathbb{R}$, $K > 0$. The set of de Sitter solutions of the SCE for these parameters with a scalar field in the Bunch-Davies state can be parameterized for $\Lambda \leq 0$ by one and for $\Lambda > 0$ by two analytic curves in the (ξ, H) -parameter plane $(0, \infty) \times (0, \infty)$. Moreover:*

- (i) *If $\frac{\Lambda}{K} > 2160$ the two solution curves can be globally solved for ξ . Denoting $H_{\min} = \left(\frac{K}{29}(1440^2 + 29 \cdot 960 \frac{\Lambda}{K})^{1/2} - \frac{1440K}{29}\right)^{1/2} \in (0, H_{\text{vac}})$, the solution curves are the (disjoint) graphs of the functions*

$$\Xi_{(+)} : (0, \infty) \rightarrow (0, \infty), \quad \Xi_{(-)} : (H_{\min}, \infty) \rightarrow (0, \infty),$$

$$\Xi_{(\pm)}(H) = \frac{1}{6} \pm \sqrt{\frac{1}{1080} - \frac{8K}{3H^2} + \frac{8\Lambda K}{9H^4}}. \quad (18)$$

In particular, any arbitrary $H > 0$ is the Hubble rate of a de Sitter solution to the SCE for one or two suitable value(s) for ξ , with two possible values if and only if $H > H_{\min}$. On the other hand, for any $\xi \in (\max(\Xi_{(-)}), \min(\Xi_{(+)}))$ ($\ni \xi_{\text{cc}}$) there exist no de Sitter solution H at all.

- (ii) *If $\Lambda \in (0, 2160K)$ the two solution curves can be globally solved for H , that is, they are the (disjoint) graphs of the (analytic continuations of the) functions*

$$H_{(+)} : (\xi_{(-)}, \xi_{(+)}) \rightarrow (0, \infty), \quad H_{(-)} : (0, \infty) \rightarrow (0, \infty),$$

$$H_{(\pm)}(\xi) = \sqrt{1440 K \frac{1 \pm \sqrt{1 - \frac{\Lambda}{2160 K}(1 - 30(6\xi - 1)^2)}}{1 - 30(6\xi - 1)^2}}. \quad (19)$$

In particular, for any $\xi \in (\xi_{(-)}, \xi_{(+)})$ there exist precisely two de Sitter solutions $H_{(-)}(\xi)$ and $H_{(+)}(\xi)$, while for any other $\xi > 0$ there exists precisely one de Sitter solution $H_{(-)}(\xi)$. On the other hand, for any value $H > 0$ with $H \notin (\max(H_{(-)}), \min(H_{(+)})) \neq \emptyset$ there is a ξ -value such that H is the Hubble rate of a de Sitter solution to the SCE.

- (iii) *If $\Lambda/K = 2160$ Equation (17) is equivalent to*

$$\xi - \frac{1}{6} = \pm \frac{1}{\sqrt{1080}} \left(1 - 2 \frac{H_{\text{vac}}^2}{H^2}\right) \quad (20)$$

analog. Note that while the classical and quantum solution branches are clearly separated whenever $\frac{\Lambda}{K} < 2160$ (Cases (ii) and (iv)), they degenerate in ξ_{cc} for $\frac{\Lambda}{K} = 2160$ (Case (iii)) and even annihilate each other around ξ_{cc} if $\frac{\Lambda}{K} > 2160$ (Case (i)). However, around the values $\xi_{(\pm)}$ the separation remain valid in all cases. We will pick up these solutions branches in the discussion of Section 7.

Related results have been observed in the literature: The solutions at $\xi = \xi_{(\pm)}$, namely $H = H_{\text{vac}}$, were previously found in [29] together with the fact that $H = H_{\text{vac}}$ is a solution only for the aforementioned ξ -values. On the other hand, de Sitter solutions at $\xi = \xi_{cc}$ were found before by many authors employing a variety of states or approximations of states. For example, Starobinski [44] found what we called “quantum solution” for $\xi = \xi_{cc}$ using the Bunch-Davies state (synonymously referring to it as “de Sitter state”). Moreover, the authors of [11] found the analogs of both what we called the quantum and the classical solution using approximate KMS states. At third, in [21] the authors of the present work found the quantum solution using massless Minkowski-like vacuum states (i.e. states with $m = 0$), and by introducing a positive cosmological constant also the classical solution would appear (cf. Remark 3.1.(iii)). Note that for a conformally coupled field, however, the state merely contributes geometric terms, hence the precise choice of a (Hadamard) state does not matter. How the two regimes around $\xi_{(\pm)}$ and ξ_{cc} in turn are connected was, to the authors’ knowledge, not observed before.

4. The solution set of de Sitter solutions for the massive field

In this section we also consider the energy equation (16) as a consistency constraint on points $(\xi, H) \in \mathbb{R} \times (0, \infty)$ with parameters m, K, Λ, μ, d_1 and d_2 . In contrast to the previous section this consistency equation is no more an explicitly solvable polynomial equation and one major task is the treatment of the incomparably more complicated dependency of the energy density $\langle T_{00}^{\text{ren}} \rangle_{\omega}$ on the parameters via the Bunch-Davies moments (15). We also remark that negative values for ξ are allowed as long as $m^2 + 12\xi H^2 > 0$ (cf. Section 2.2).

At first we reformulate the consistency equation (16) into a shape tailored to the massive case. In particular we identify two (effective) parameters e_1 and e_2 which influence the shape of the solution set in the (ξ, H) -plane and we point out how the remaining parameters merely rescale the solution set in said plane.

As a next step, in analogy to Proposition 3.2, we prove how the solution set in the (ξ, H) -plane can be parameterized by analytic curves and how each such curve must hit the boundary of admissible (ξ, H) -points at both ends. Partial results are outsourced to Subsections 4.1, 4.2 and 4.3. The asymptotics of the solution set, particularly how many curves constitute the solution set, will be analyzed in the next section, Section 5.

In order to simplify the consistency equation (16), denote

$$f : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \psi^{(0)}\left(\frac{3}{2} - \sqrt{\frac{9}{4} - x}\right) + \psi^{(0)}\left(\frac{3}{2} + \sqrt{\frac{9}{4} - x}\right). \quad (21)$$

A plot of f on the relevant domain is shown in Figure 3.(i) below and a few useful

properties of it are listed in Appendix A. By introducing a shifted curvature coupling

$$x = 12\xi + \frac{m^2}{H^2} \quad (22)$$

(note how the field equation reads $(\square + xH^2)\phi = 0$), we can rewrite the Digamma function terms occurring in the consistency equation (16) into $\psi^{(0)}(\frac{3}{2} - \nu) + \psi^{(0)}(\frac{3}{2} + \nu) = f(x)$. Moreover, regarding (16) as an equation for x instead of ξ simplifies the domain of our problem to such (x, H) -points where both $x > 0$ and $H > 0$. Therefore we note that $x > 0$ is equivalent to the positivity of the effective de Sitter mass $xH^2 = m^2 + 12\xi H^2$, and hence equivalent to the existence of the unique $O(4, 1)$ -invariant Bunch-Davies state on the de Sitter space-time encoded by H , cf. [1].

We recall that changes in the length scale μ can be absorbed into the renormalization constants d_1 and d_2 . Hence, we eliminate the parameters m^2 and μ by setting $\mu = m$ and rewriting the energy equation in terms of $h = \frac{H}{m}$.

Finally, we define new parameters

$$e_1 = 960\left(d_1 + \frac{\Lambda K}{m^4}\right) - \frac{15}{2} \quad \text{and} \quad e_2 = 960\left(d_2 - 3\frac{K}{m^2}\right) \quad (23)$$

and note that e_1 and e_2 are still linear transformations of the original renormalization freedoms c_1 and c_2 , respectively. In particular, they can be arbitrary real numbers.

Using all the above simplifications, we rewrite the consistency equation (16) into

$$F(x, h) = 0 \quad (24)$$

with the functions $F, F_1, F_2 : (0, \infty) \times (0, \infty) \rightarrow \mathbb{R}$,

$$\begin{aligned} F_1(x, h) &= \left(\frac{x^2}{4} - x + \frac{29}{30}\right)h^2 - \left(\frac{x}{2} + \frac{e_2}{30} - 1\right) - \frac{e_1}{30} \frac{1}{h^2} \\ &= \left(\frac{x}{2} - 1\right)^2 h^2 - \left(\frac{x}{2} - 1\right) - \frac{1}{30h^2}(h^4 + e_2 h^2 + e_1) \\ &= \frac{h^2}{4} x^2 - \left(h^2 + \frac{1}{2}\right)x + \frac{29h^2}{30} - \frac{e_2}{30} - \frac{e_1}{30h^2} + 1, \\ F_2(x, h) &= 2\log(h) + f(x), \\ F(x, h) &= F_1(x, h) - \left(\frac{x}{2} - 1\right)F_2(x, h). \end{aligned}$$

Due to the analyticity of f discussed in Appendix A, we have that F, F_1, F_2 are analytic.

We introduce the distinguished x -values

$$x_{(\pm)} := 2 \pm \sqrt{\frac{2}{15}}$$

and note that $x(\xi_{(\pm)}) = 12\xi_{(\pm)} + \frac{1}{h^2} \rightarrow x_{(\pm)}$ in the limit $h \rightarrow \infty$, that is, in said limit these distinguished x -values correspond to the ξ -values distinguished in the massless case,

Section 3.

In the following we study solutions of (24). Denote the analytic variety of solutions by

$$\begin{aligned} \mathcal{S}_{e_1, e_2} := & \{ (x, h) \in (0, \infty) \times (0, \infty) \mid F(x, h) = 0 \text{ for the parameters } e_1, e_2 \} \\ & \subset (0, \infty) \times (0, \infty). \end{aligned}$$

Moreover, inspired by graph theory we introduce the following notion. Note that, whenever we speak of a *connected* set we mean a path-connected set, that is, a set such that any pair of points from that set is connected by a continuous curve contained in that set.

Definition 4.1. *A subset $\mathcal{S} \subset (0, \infty) \times (0, \infty)$ is called tree-like if it is connected and for each $s \in \mathcal{S}$ the set $\mathcal{S} \setminus \{s\}$ is disconnected with finitely many connected components.*

The following theorem characterizes the analytic variety \mathcal{S}_{e_1, e_2} . Note again how, compared to the massless case, the occurrence of the function f prevents a simple, closed expression for solutions in \mathcal{S}_{e_1, e_2} and we approach the equation using certain properties of f such as its analyticity, its asymptotic behavior or certain bounds on its derivatives.

Theorem 4.2. *Let $e_1, e_2 \in \mathbb{R}$. The analytic variety \mathcal{S}_{e_1, e_2} can be parameterized by finitely many inextendible analytic curves. Any such curve*

$$\gamma = (\gamma^{(x)}, \gamma^{(h)}) : I \rightarrow (0, \infty) \times (0, \infty)$$

defined on an open $I \subset \mathbb{R}$ fulfills

$$\gamma^{(x)}(t) \rightarrow 0, \quad \gamma^{(x)}(t) \rightarrow \infty, \quad \gamma^{(h)}(t) \rightarrow 0 \quad \text{or} \quad \gamma^{(h)}(t) \rightarrow \infty$$

whenever t approaches $\inf I$ or $\sup I$. Moreover, \mathcal{S}_{e_1, e_2} is the disjoint union of tree-like subsets.

Remark 4.3. (i) *Note that by the proposed limits we state that any piece of solution curve is analytically continuable until it runs into the “boundary” of the domain $(0, \infty) \times (0, \infty)$ of F . In other words, any inextendible curve leaves any compact subset of $(0, \infty) \times (0, \infty)$, in particular, \mathcal{S}_{e_1, e_2} can have no compact connected component. Hence, all solutions can be found by studying the asymptotics of F in the limits $x \rightarrow 0$, $x \rightarrow \infty$, $h \rightarrow 0$ and $h \rightarrow \infty$ and continuing the arising solution curves. The asymptotic analysis of F is postponed to Section 5.*

(ii) *In the present section we skip the proof that there are at most finitely many solution curves as in the theorem. In Section 5, whenever we use Theorem 4.2, we do not need this particular assertion. In the end, it will follow from the analysis in that section.*

The proof is based on the following lemmas.

Lemma 4.4. *Let $e_1, e_2 \in \mathbb{R}$. For any $s \in \mathcal{S}_{e_1, e_2}$ there exist an open neighborhood $U \ni s$ (open in $(0, \infty) \times (0, \infty)$) such that $U \cap \mathcal{S}_{e_1, e_2}$ can be regularly parameterized by one or two analytic curves $I \rightarrow (0, \infty) \times (0, \infty)$.*

Proof. Let $s \in \mathcal{S}_{e_1, e_2} \subset (0, \infty) \times (0, \infty)$, then either $\nabla F(s) \neq 0$ or $\nabla F(s) = 0$.

In the first case an open neighborhood U and one (unique) curve as in the lemma are provided by the implicit function theorem in its analytic version.

In the second case $\nabla F(s) = 0$ we conclude from the results of Section 4.3 that $\text{Hess } F(s)$ is indefinite and Lemma B.1 (cf. also Remark B.2) provides an open neighborhood U and precisely two curves as in the lemma.

Note that both the implicit function theorem and Lemma B.1 represent the solution curves in a way such that $|\gamma'|$ can be uniformly bounded away from zero, imposing that all curves are regularly parameterizable. \square

Lemma 4.5. *Let $e_1, e_2 \in \mathbb{R}$. \mathcal{S}_{e_1, e_2} is closed in $(0, \infty) \times (0, \infty)$.*

Proof. \mathcal{S}_{e_1, e_2} is the zero set of a continuous function. \square

Lemma 4.6. *Let $e_1, e_2 \in \mathbb{R}$ and suppose that $\gamma : I \rightarrow \mathcal{S}_{e_1, e_2}$ is an inextendible, regularly parameterized analytic solution curve. Then γ is injective.*

Proof. Suppose γ is not injective, then there exist $a, b \in I$, $a < b$, such that $\gamma(a) = \gamma(b)$ and γ is continuous on $[a, b]$ and analytic on $(a, b) \neq \emptyset$.

We first study the case $\nabla F(\gamma(a)) = 0$. Let U be the open neighborhood of $\gamma(a)$ provided by Lemma 4.4. Moreover, let $\eta_j : (-\delta, \delta) \rightarrow U$, $j \in \{1, 2\}$, $\delta > 0$, be the regular parameterizations of $U \cap \mathcal{S}_{e_1, e_2}$ from the aforementioned lemma with $\eta_1(0) = \eta_2(0) = \gamma(a)$. Note that, if necessary, we can regularly reparameterize them to be defined on the same symmetric interval $(-\delta, \delta)$.

γ is a continuous curve with $\gamma'(a) \neq 0$. Thus we can assume that U is small enough such that γ takes at least one value outside U . Consequently, there exists $\varepsilon > 0$ such that $\gamma|_{[a, a+\varepsilon]}$ coincides with precisely one of the four solution branches

$$\eta_1|_{[0, \delta]}, \quad \eta_1|_{(-\delta, 0]}, \quad \eta_2|_{[0, \delta]} \quad \text{or} \quad \eta_2|_{(-\delta, 0]}, \quad (25)$$

up to reparameterization. W.l.o.g. we can label the η_i 's such that $\gamma|_{[a, a+\varepsilon]}$ coincides with $\eta_1|_{[0, \delta]}$. In particular, we can assume that $\gamma(t) \neq \gamma(a)$ for all $t \in (a, b)$, otherwise we replace b by the smallest such point. Therefore note that, since $\gamma|_{[a, a+\varepsilon]}$ coincides with $\eta_1|_{[0, \delta]}$, such points t with $\gamma(t) = \gamma(a)$ do not accumulate in a .

By the same argument as above, there exists $\tilde{\varepsilon} > 0$ such that $\gamma|_{(b-\tilde{\varepsilon}, b]}$ coincides with one of the four solution branches in (25). We go through the cases.

In the first case the curves $\gamma|_{[a, a+\varepsilon]}$ and $\gamma|_{(b-\tilde{\varepsilon}, b]}$ coincide, up to reparameterization. Explicitly, there exist an analytic reparameterization $\theta : (a, a + \varepsilon) \rightarrow (b - \tilde{\varepsilon}, b)$ which, by $\gamma(a) = \gamma(b)$, is monotonously decreasing. Moreover, $\gamma|_{(a, b)}$ represents an analytic continuation of both $\gamma|_{(a, a+\varepsilon)}$ and $\gamma|_{(b-\tilde{\varepsilon}, b)}$, hence the analytic reparameterization θ can be continued to a monotonously decreasing reparameterization $\hat{\theta} : (a, b) \rightarrow (a, b)$. Such a map has a fixed point $t_0 \in (a, b)$, $\hat{\theta}(t_0) = t_0$, in which

$$\gamma'(t_0) = \hat{\theta}'(t_0) \cdot \gamma'(\hat{\theta}(t_0)) = \hat{\theta}'(t_0) \cdot \gamma'(t_0)$$

holds. $\widehat{\theta}'(t_0) < 0$ implies $\gamma'(t_0) = 0$ yielding a contradiction to γ being regular.

In the other three cases of (25), that is, $\gamma|_{[a, a+\varepsilon]}$ coincides with $\eta_1|_{(-\delta, 0]}$, $\eta_2|_{[0, \delta]}$ or $\eta_2|_{(-\delta, 0]}$ up to reparameterization, we can make U and δ smaller, such that $\mathcal{S}_{e_1, e_2} \cap \partial U$ consists precisely of the four points

$$\{\eta_1(\delta) = \gamma(a + \varepsilon), \eta_1(-\delta), \eta_2(\delta), \eta_2(-\delta)\} = \mathcal{S}_{e_1, e_2} \cap \partial U.$$

In particular, $\gamma(t) \notin \overline{U} = U \cup \partial U$ for all $t \in (a + \varepsilon, b - \tilde{\varepsilon})$ and $\gamma(b - \tilde{\varepsilon}) \in \{\eta_1(-\delta), \eta_2(\delta), \eta_2(-\delta)\}$. Thereby we obtain two distinct continuous curves from $\gamma(a + \varepsilon)$ to $\gamma(b - \tilde{\varepsilon})$, one going through U passing $\gamma(0)$ via $\gamma|_{[a, a+\varepsilon]}$ and $\gamma|_{[b-\tilde{\varepsilon}, b]}$ and the other outside of \overline{U} along $\gamma|_{[a+\varepsilon, b-\tilde{\varepsilon}]}$. By concatenating them we obtain a closed continuous curve along which F vanishes and which encloses at least one point in which F does not vanish, say $s_0 \in U$ with $F(s_0) \neq 0$. Thereby, the continuous curve in construction encloses s_0 's whole connected component of non-zeros of F , hence this connected component's closure is a compact subset of $(0, \infty) \times (0, \infty)$. The existence of such a set is excluded in Proposition 4.23.

Concluding, if $\nabla F(\gamma(a)) = 0$, any of the above possibility yields a contradiction. If, on the other hand, $\nabla F(\gamma(a)) \neq 0$ the analytic implicit function theorem provides us an open neighborhood U and a single curve $\eta : (-\delta, \delta) \rightarrow U$ to regularly parameterize $U \cap \mathcal{S}_{e_1, e_2}$. Replacing (25) by the two branches $\eta|_{[0, \delta]}$ and $\eta|_{(-\delta, 0]}$ these cases imply contradictions by same arguments as above. \square

Lemma 4.7. *Let $e_1, e_2 \in \mathbb{R}$ and let*

$$\gamma = (\gamma^{(x)}, \gamma^{(h)}) : (a, b) \rightarrow \mathcal{S}_{e_1, e_2}$$

be continuous and analytic on (a, b) . Moreover, suppose that γ is regularly parameterized, i.e. $\gamma'(t) \neq 0$ for all $t \in (a, b)$. Then either

$$\gamma^{(x)}(t) \rightarrow 0, \quad \gamma^{(x)}(t) \rightarrow \infty, \quad \gamma^{(h)}(t) \rightarrow 0 \quad \text{or} \quad \gamma^{(h)}(t) \rightarrow \infty$$

as $t \rightarrow a$ or the limit $\lim_{t \rightarrow a} \gamma(t)$ exists in $(0, \infty) \times (0, \infty)$. Consequently, in the latter case γ is continuously extendible to $[a, b]$.

Proof. At first, suppose that none of the limits in the lemma holds true, that is, there exists $x_1, x_2, h_1, h_2 \in (0, \infty)$ such that arbitrarily close to a , γ takes a value in $[x_1, x_2] \times [h_1, h_2]$. More precisely, we find a sequence $(a_n)_{n \in \mathbb{N}}$ in (a, b) with $a_n \rightarrow a$ as $n \rightarrow \infty$ and $s_n := \gamma(a_n) \in [x_1, x_2] \times [h_1, h_2]$ for all $n \in \mathbb{N}$. $[x_1, x_2] \times [h_1, h_2]$ is compact, thus $(s_n)_{n \in \mathbb{N}}$ accumulates in some $\widehat{s} \in [x_1, x_2] \times [h_1, h_2]$.

Now let $(a_n)_{n \in \mathbb{N}}$ be an arbitrary sequence in (a, b) with $a_n \rightarrow a$ as $n \rightarrow \infty$ and denote $s_n = \gamma(a_n)$, $n \in \mathbb{N}$. Suppose (s_n) has an accumulation point $\widehat{s} \in (0, \infty) \times (0, \infty)$, but does not converge to \widehat{s} .

In the following we will pass to a subsequence of $(a_n)_{n \in \mathbb{N}}$ several times, but we will not give them a new label each time.

By possibly passing to a subsequence, we can find an open neighborhood $U \subset (0, \infty) \times (0, \infty)$ of \widehat{s} such that $s_{2n} \rightarrow \widehat{s}$ as $n \rightarrow \infty$ and $s_{2n+1} \notin U$ for all n . Moreover, we can achieve that the limit $a_n \rightarrow a$ is strictly monotonous.

Note that, since $s_n \in \mathcal{S}_{e_1, e_2}$ and \mathcal{S}_{e_1, e_2} is closed (Lemma 4.5), also $\widehat{s} \in \mathcal{S}_{e_1, e_2}$, that is, $F(\widehat{s}) = 0$. By Lemma 4.4 and possibly making U smaller we find that $U \cap \mathcal{S}_{e_1, e_2}$ is parameterized by one regular curve $\eta : (-\delta, \delta) \rightarrow U$ with $\eta(0) = \widehat{s}$ (in the case $\nabla F(\widehat{s}) \neq 0$) or by two curves $\eta_i : (-\delta, \delta) \rightarrow U$, $i \in \{1, 2\}$ (in the case $\nabla F(\widehat{s}) = 0$) with $\eta_1(0) = \eta_2(0) = \widehat{s}$.

By possibly passing to a subsequence again we can achieve that $s_{2n} \in U$ for all n and that still $s_{2n} \rightarrow \widehat{s}$ and $s_{2n+1} \notin U$ for all n . Hence, all $(s_{2n})_{n \in \mathbb{N}}$ are contained in the range of η or in the ranges of η_1 and η_2 , respectively, according to the cases of $\nabla F(\widehat{s}) = 0$ or $\nabla F(\widehat{s}) \neq 0$. In the case of $\nabla F(\widehat{s}) = 0$, the range of at least one of η_1 and η_2 contains infinitely many of the $(s_{2n})_{n \in \mathbb{N}}$ and, by relabeling the η_i and possibly passing to a subsequence again, we can achieve that all of the $(s_{2n})_{n \in \mathbb{N}}$ are contained in the range of $\eta := \eta_1$.

Then there exists a sequence $(c_n)_{n \in \mathbb{N}}$ such that $\eta(c_n) = s_{2n} = \gamma(a_{2n})$. Moreover, since both γ and η are regularly parameterized analytic curves, there exist positive sequences $(\varepsilon_{2n}^{(-)})_{n \in \mathbb{N}}$ and $(\varepsilon_{2n}^{(+)})_{n \in \mathbb{N}}$ such that for all n the curves

$$\gamma|_{(a_{2n} - \varepsilon_{2n}^{(-)}, a_{2n} + \varepsilon_{2n}^{(+)})} \quad \text{and} \quad \eta$$

coincide, up to reparameterization. Since U does not contain any s_{2n+1} and since $a_n \rightarrow a$ monotonously, on each interval (a_{2n+2}, a_{2n}) the curve γ also takes values outside of U , namely $\gamma(a_{2n+1})$ with $a_{2n+1} \in (a_{2n+2}, a_{2n})$. Consequently, the intervals

$$(a_{2n} - \varepsilon_{2n}^{(-)}, a_{2n} + \varepsilon_{2n}^{(+)}) \quad \text{with} \quad n \in \mathbb{N}$$

are pairwise disjoint. But then, for some $n \in \mathbb{N}$ the curves

$$\gamma|_{(a_{2n} - \varepsilon_{2n}^{(-)}, a_{2n} + \varepsilon_{2n}^{(+)})}, \quad \gamma|_{(a_{2n+2} - \varepsilon_{2n+2}^{(-)}, a_{2n+2} + \varepsilon_{2n+2}^{(+)})} \quad \text{and} \quad \eta$$

have the same range and, in particular, γ is not injective on (a, b) . This contradicts Lemma 4.6.

Concluding, we have shown, that the values $\gamma(t)$ with $t \rightarrow a$ either approach one of the limits in the lemma, or it accumulates in some $\widehat{s} \in (0, \infty) \times (0, \infty)$. Moreover, we have shown that in the latter case this accumulation point is already a limit. \square

Corollary 4.8. *Let $\gamma : [a, b] \rightarrow \mathcal{S}_{e_1, e_2}$ be the continuous extension of $\gamma : (a, b) \rightarrow \mathcal{S}_{e_1, e_2}$ from the previous lemma with $\gamma(a) = \lim_{t \rightarrow a} \gamma(t) \in (0, \infty) \times (0, \infty)$. Then γ is analytically continuable to $(a - \varepsilon, b]$ for some $\varepsilon > 0$.*

Proof. Analogously to the proof of Lemma 4.6 we find an open neighborhood $U \ni \gamma(a)$ and an analytic curve $\eta : (-\delta, \delta) \rightarrow U$ together with $\tilde{\varepsilon} > 0$ such that $\gamma|_{[a, a+\tilde{\varepsilon})}$ and $\eta|_{[0, \delta)}$ coincide, up to reparameterization. But then γ can be concatenated with $(a$

commensurable reparameterization of) $\eta|_{(-\delta,0]}$ to obtain an analytic continuation of γ beyond the point $\gamma(a)$ on some interval $(a - \varepsilon, b]$. \square

Lemma 4.9. *Let $e_1, e_2 \in \mathbb{R}$. Each connected component of \mathcal{S}_{e_1, e_2} is tree-like.*

Proof. We adjust the argument of Lemma 4.6.

Let \mathcal{S} be a connected component of \mathcal{S}_{e_1, e_2} . If \mathcal{S} is not tree-like, there exists $s \in \mathcal{S}$ such that $\mathcal{S} \setminus \{s\}$ is still connected. Let $U \ni s$ be the open neighborhood provided by Lemma 4.4 and let $\eta : (-\delta, \delta) \rightarrow U$ be the (if $\nabla F(s) \neq 0$) or one of the two (if $\nabla F(s) = 0$) corresponding regular parameterization(s) of $U \cap \mathcal{S}_{e_1, e_2}$ with $\eta(0) = s$. Since the range of η is connected, η is a map $(-\delta, \delta) \rightarrow U \cap \mathcal{S}$.

Since $\mathcal{S} \setminus \{s\}$ is still connected, there exists a continuous curve $\gamma : [a, b] \rightarrow \mathcal{S} \setminus \{s\}$ with $\gamma(a) = \eta(-\delta/2)$ and $\gamma(b) = \eta(\delta/2)$. On the other hand, $\eta|_{[-\delta/2, \delta/2]}$ is a continuous curve in \mathcal{S} connecting $\gamma(a) = \eta(-\delta/2)$ and $\gamma(b) = \eta(\delta/2)$ as well, and since γ cannot take the value s , we have indeed two different continuous curves in \mathcal{S} connecting the aforementioned two points. Hence, by concatenation, we obtain a closed continuous curve in \mathcal{S} which, viewed as curve in $(0, \infty) \times (0, \infty)$, must enclose at least one non-solution (with $F(x, h) \neq 0$) and thus its whole connected component of non-solutions. This connected component cannot exist as we have demonstrated in Lemma 4.23. \square

With the previous lemmas we can finally conclude Theorem 4.2.

Proof of Theorem 4.2. In Lemma 4.4 we have seen that any solution $s \in \mathcal{S}_{e_1, e_2}$ belongs to the graph of an analytic curve. In Lemma 4.7 and Corollary 4.8 we have seen that any such solution curve has an analytic continuation up to one of the four limits in the theorem (or in Lemma 4.7) is approached, in both the infimum and the supremum of the domain. Finally, the tree-like structure is proven in Lemma 4.9. We have not shown that there are only finitely many solution curves. This assertion is postponed to Section 5 (cf. Remark 4.3). \square

Remark 4.10. (i) *Note that Lemma 4.4 may be viewed as a generalized implicit function theorem for functions which, like F does, fulfill that whenever $F(x, h) = 0$ and $\nabla F(x, h) = 0$ for some (x, h) , then $\text{Hess } F(x, h)$ is indefinite. Hence, in the usual fashion of continuing solution curves provided by the implicit function theorem until that theorem is no longer applicable, it is already clear from Lemma 4.4 that \mathcal{S}_{e_1, e_2} consists of only “infinitely long” solution curves. Note that from Lemma B.1 (in particular the power series representation in Remark B.2) it is easy to see that common zeros of both F and ∇F cannot accumulate. However, then one is still burdened with excluding any kind of strange behavior an (even analytic) “infinitely long” curve can show, such as spiraling into a point (with a non-integrable $|\gamma'|$) or a “topologist’s sine curve”-like behavior (e.g. via $x \mapsto \sin(1/x)$). If one, in the end, aims to show that any such behavior must be visible in the asymptotics of F , one is right in the middle of proving Lemma 4.7.*

(ii) *We have remarked before that the consistency equation is the same for the (non-pullback) Bunch-Davies state on (the entire) de Sitter space using the stress-energy*

$F(x, \cdot)$. Taking into account the limits

$$\lim_{h \rightarrow 0} F(x, h) = -\infty \quad \text{and} \quad \lim_{h \rightarrow \infty} F(x, h) = +\infty$$

for $e_1 \geq 0$ we can, for such e_1 , replace “at most” by “exactly”.

If $x = x_{(\pm)}$ the leading coefficient in (26) vanishes. Hence, $\partial_h F(x_{(-)}, \cdot)$ has precisely one zero if $e_1 < 0$ and no zero at all if $e_1 \geq 0$, implying that $F(x_{(-)}, \cdot)$ has at most two or at most one zero, respectively. The very same argument can be applied to $F(x_{(+)}, \cdot)$ by reversing the sign of e_1 . Taking also into account the limits

$$\lim_{h \rightarrow 0} F(x_{(-)}, h) = -\infty \quad \text{and} \quad \lim_{h \rightarrow \infty} F(x_{(-)}, h) = +\infty \quad \text{if } e_1 \geq 0$$

as well as

$$\lim_{h \rightarrow 0} F(x_{(+)}, h) = +\infty \quad \text{and} \quad \lim_{h \rightarrow \infty} F(x_{(+)}, h) = -\infty \quad \text{if } e_1 \leq 0$$

we see that in the latter case “at most” can be replaced by “exactly”.

If $x_{(-)} < x < 2$ the RHS of (26) has precisely one zero if $e_1 \geq 0$, implying that $F(x, \cdot)$ has at most two zeros.

At $x = 2$ the RHS of (26) has no positive zero if $e_1 \leq 0$ and precisely one if $e_1 > 0$. Consequently, $F(2, \cdot)$ has at most one or at most two zeros, respectively. Taking, moreover, into account that

$$\lim_{h \rightarrow 0} F(2, h) = +\infty \quad \text{and} \quad \lim_{h \rightarrow \infty} F(2, h) = -\infty$$

for $e_1 < 0$ in this case “at most” can again be replaced by “exactly”.

For $2 < x < x_{(+)}$ the RHS of (26) has no positive zero if $e_1 \leq 0$ and precisely one positive zero if $e_1 > 0$. Consequently, $F(x, \cdot)$ has at most one or at most two zeros, respectively. Taking into account the limits

$$\lim_{h \rightarrow 0} F(x, h) = +\infty \quad \text{and} \quad \lim_{h \rightarrow \infty} F(x, h) = -\infty$$

for $e_1 < 0$ also here “at most” can be replaced by “exactly”.

Finally, if $x > x_{(+)}$ the RHS of (26) has precisely one positive zero if $e_1 \leq 0$, allowing at most two positive zeros for $F(x, \cdot)$.

This finishes the proof of every entry shown in Table 1. \square

Note that we will continue to study the zeros of $\partial_h F(x, \cdot)$ at fixed x in the subsequent section.

Corollary 4.12. *The solution set \mathcal{S}_{e_1, e_2} of $F(x, h) = 0$ is non-empty for any choice of renormalization constants e_1 and e_2 .*

Proof. For each row of Table 1 we have at least one entry with exactly one solution at the respective x -value. \square

After we have demonstrated how many solutions exist at most if we fix one variable we can give a more accurate location of the three solutions at a fixed h . Note that the solutions of $F(2, h) = 0$ are given by the zeros of the polynomial

$$p(h) = -30h^2F(2, h) = h^4 + e_2h^2 + e_1. \quad (27)$$

Lemma 4.13. *Let $e_1, e_2 \in \mathbb{R}$ and suppose that the polynomial function $p(h) = h^4 + e_2h^2 + e_1$ has a zero h_0 . If $h_0 > \exp(\gamma - 1) \approx 0.6552$ with the Euler-Mascheroni number $\gamma \approx 0.5772$, then $F(x, h_0)$ has at least three solutions $x_1, x_2, x_3 \in (0, \infty)$ with $x_1 \in (0, 2)$, $x_2 = 2$ and $x_3 \in (2, \infty)$.*

Proof. The map $(0, \infty) \rightarrow \mathbb{R}$, $x \mapsto F(x, h_0)$ fulfills

$$\lim_{x \rightarrow 0} F(x, h_0) = -\infty, \quad \lim_{x \rightarrow \infty} F(x, h_0) = +\infty \quad \text{and} \quad F(2, h_0) = 0.$$

Moreover, its derivative in $x = 2$ fulfills

$$\partial_x F(2, h_0) = -\frac{1}{2} - \log(h_0) - \frac{1}{2}f(2) < 0,$$

where we used that $f(2) = 1 - 2\gamma$. Consequently, $F(\cdot, h_0)$ is positive on an interval of the form $(2 - \varepsilon, 2)$ and negative on an interval of the form $(2, 2 + \varepsilon)$. In combination with the limits above, this implies the existence of zeros of $F(\cdot, h_0)$ as asserted in the lemma. \square

Remark 4.14. (i) *In combination with Lemma 4.11.(i) the assertion “at least three” in Lemma 4.13 can be replaced by “exactly three”.*

(ii) *If in Lemma 4.13 we claim $h_0 < \exp(\gamma - 1)$ instead, we still have the solution at $(2, h_0)$, but if two more $h = h_0$ -solutions exist at all, they must be either both larger than 2 or both smaller than 2.*

Complementary to Lemma 4.13 concerning zeros of p we find the following lemma concerning h -values where $p(h) \neq 0$.

Lemma 4.15. *Let $h > 0$, $e_1, e_2 \in \mathbb{R}$ and $p(h) = h^4 + e_2h^2 + e_1$.*

(i) *If $p(h) > 0$, the equation $F(x, h) = 0$ has at least one solution $x \in (0, 2)$.*

(ii) *If $p(h) < 0$, the equation $F(x, h) = 0$ has at least one solution $x \in (2, \infty)$.*

Proof. At $x \neq 2$ the equation $F(x, h) = 0$ is equivalent to the function

$$(0, \infty) \setminus \{2\} \rightarrow \mathbb{R}, \quad x \mapsto \frac{F(x, h)}{\frac{x}{2} - 1}$$

having a zero. For this function we observe

$$\lim_{x \rightarrow 0} \frac{F(x, h)}{\frac{x}{2} - 1} = \lim_{x \rightarrow \infty} \frac{F(x, h)}{\frac{x}{2} - 1} = +\infty.$$

Moreover, this function has a first order pole at $x = 2$, and this poles residue has the opposite sign than the value $p(h)$. Consequently,

$$\lim_{\substack{x \rightarrow 2 \\ x > 2}} \frac{F(x, h)}{\frac{x}{2} - 1} = -\operatorname{sgn} p(h) \cdot \infty \quad \text{and} \quad \lim_{\substack{x \rightarrow 2 \\ x < 2}} \frac{F(x, h)}{\frac{x}{2} - 1} = \operatorname{sgn} p(h) \cdot \infty$$

and the lemma follows. \square

Remark 4.16. *Lemma 4.15 particularly shows that the equation $F(x, h) = 0$ for any fixed $h > 0$ has at least one solution for x (recall that $F(2, h) = 0$ for $p(h) = 0$).*

We postpone a further location of solution curves via the asymptotics of F to Section 5.

4.2. Possible non- h -soluble points

The implicit function theorem (in its analytical version) tells us that, whenever we have a solution $F(x, h) = 0$ such that $\nabla F(x, h) = (\partial_x F(x, h), \partial_h F(x, h)) \neq 0$, then there exists an open neighborhood of (x, h) in which all solutions of $F(x, h) = 0$ are collected in an analytic curve. Moreover, any such ‘‘piece of solution curve’’ can be continued either until it diverges, until it hits the boundary $\{x = 0\} \cup \{h = 0\}$ or until it runs into a point where $\nabla F(x, h) = 0$.

The present section is dedicated to studying (a necessary condition on) points (x, h) in which the gradient of F vanishes. Since $\partial_x F$ involves derivatives of f in a poorly manageable combination, we use $\partial_h F = 0$ as a necessary condition. The latter in turn is (equivalent to) a polynomial equation and is explicitly solvable. Note that by this weaker criterion we also identify points in which the solution curves ‘‘turn around’’, i.e. points where they are not solvable for h , but possibly for x .

Lemma 4.17. *Let $e_1, e_2 \in \mathbb{R}$ and denote*

$$h_{\min} = \begin{cases} 0 & \text{if } e_1 \leq 15/2 \\ (e_1 - 15/2)^{1/4} & \text{if } e_1 > 15/2 \end{cases}.$$

(i) *The mapping*

$$h \mapsto X_{(e_1, \pm)}(h) = 2 + \frac{1}{h^2} \pm \sqrt{\frac{2}{15} + \frac{1}{h^4} \left(1 - \frac{2e_1}{15}\right)}$$

defines real-valued functions $X_{(e_1, \pm)} : (0, \infty) \cap [h_{\min}, \infty) \rightarrow \mathbb{R}$.

(ii) *Any $(x, h) \in (0, \infty) \times (0, \infty)$ with $\nabla F(x, h) = 0$ fulfills $x \in \{X_{(e_1, +)}(h), X_{(e_1, -)}(h)\}$.*

Proof. Finding the zeros of $\partial_h F$ is equivalent to solving the polynomial equation (for x)

$$0 = x^2 - \left(4 + \frac{2}{h^2}\right)x + \frac{58}{15} + \frac{4}{h^2} + \frac{2e_1}{15h^4} \quad (28)$$

and doing so results in the functions $X_{(e_1, \pm)}$ in the lemma. The domain which yields real functions is obtained by requiring the radicand to be non-negative. From this both assertions of the lemma follow immediately. \square

the analytic functions

$$\begin{aligned}
 Y_{(e_1, \pm)} &: \text{Dom}(X_{(e_1, \pm)}) \rightarrow \mathbb{R}, \\
 h \mapsto & \frac{\partial^2 F}{\partial h^2}(X_{(e_1, \pm)}(h), h) \cdot \frac{\partial^2 F}{\partial x^2}(X_{(e_1, \pm)}(h), h) - \left(\frac{\partial^2 F}{\partial h \partial x}(X_{(e_1, \pm)}(h), h) \right)^2 \\
 &= \left[\frac{1}{2} \left(X_{(e_1, \pm)}(h) - 2 + \frac{1}{h^2} \right)^2 - \frac{1}{2h^4} - \frac{e_1}{5h^4} - \frac{1}{15} \right] \\
 &\quad \cdot \left[\frac{h^2}{2} - f' \circ X_{(e_1, \pm)}(h) - \left(\frac{1}{2} X_{(e_1, \pm)}(h) - 1 \right) \cdot f'' \circ X_{(e_1, \pm)}(h) \right] \\
 &\quad - \left[\left(X_{(e_1, \pm)}(h) - 2 \right) h - \frac{1}{h} \right]^2
 \end{aligned} \tag{29}$$

on the (possibly e_1 -dependent) maximal domains of $X_{(e_1, \pm)}$. We study the functions $Y_{(e_1, \pm)}$ in terms of their asymptotics and by numerical means to show that they are mostly negative, and if not, then $\partial_x F(h, X_{(e_1, \pm)}(h)) \neq 0$ and thus the point in question is not critical.

Lemma 4.20. *Let $e_1 \in \mathbb{R}$.*

- (i) $Y_{(e_1, \pm)}(h) < 0$ for sufficiently large h .
- (ii) If $e_1 < \frac{15}{2}$, then $Y_{(e_1, +)}(h) < 0$ for sufficiently small $h > 0$.
- (iii) If $e_1 \in [0, \frac{15}{2}]$, then $Y_{(e_1, -)}(h) < 0$ for sufficiently small $h > 0$.
- (iv) For $e_1 < 0$ denote $h_{\text{crit}} = \left(\left(\frac{15}{29} \right)^2 - \frac{e_1}{29} \right)^{1/2} - \frac{15}{29}$. Then $X_{(e_1, -)}$ is positive on $(h_{\text{crit}}, \infty)$ and there exists $\varepsilon > 0$ such that $Y_{(e_1, -)}$ is negative on $(h_{\text{crit}}, h_{\text{crit}} + \varepsilon)$.

Proof. By Lemma 4.18.(i) we have $X_{(e_1, \pm)}(h) \rightarrow x_{(\pm)}$ as $h \rightarrow \infty$ and since f is smooth (i.e. f' and f'' are continuous) we can read off from (29) that $Y_{(e_1, \pm)} \rightarrow -\infty$ as $h \rightarrow \infty$. More precisely, identifying the dominant terms we find

$$\frac{1}{h^2} Y_{(e_1, \pm)}(h) \rightarrow -\frac{2}{15} \quad \text{as } h \rightarrow \infty.$$

This proves (i).

In order to show (ii), recall from Section 4.2 that both $X_{(e_1, +)}$ for $e_1 < \frac{15}{2}$ and $X_{(e_1, +)}$ for $e_1 \in (0, \frac{15}{2})$ are defined on $(0, \infty)$ and that $X_{(e_1, \pm)}(h) \rightarrow +\infty$ as $h \rightarrow 0$ for the given respective e_1 -values. Expanding each occurrence of $X_{(e_1, \pm)}$ in $Y_{(e_1, \pm)}$ from (29) to a sufficiently high order in h we find that

$$Y_{(e_1, \pm)}(h) \in - \frac{2(1 - \frac{2e_1}{15})}{15(1 \pm \sqrt{1 - \frac{2e_1}{15}})^2} h^2 + \mathcal{O}(h^3).$$

Note that the functions $z \mapsto \frac{z}{(1 \pm \sqrt{z})^2}$ are positive for the relevant domains. This proves (ii) and, moreover, (iii) for $e_1 \in (0, \frac{15}{2})$.

Recall that $X_{(\frac{15}{2}, \pm)}(h) = x_{(\pm)} + \frac{1}{h^2}$ for $h > 0$. By an expansion to sufficiently high order we find that

$$Y_{(\frac{15}{2}, \pm)}(h) \in \left(-\frac{8}{225} \pm \frac{16\sqrt{30}}{1575} \right) h^6 + \mathcal{O}(h^7). \quad (30)$$

In particular, the leading order term in h of $Y_{(\frac{15}{2}, -)}(h)$ is negative, proving (iii) for $e_1 = \frac{15}{2}$.

In order to complete (iii) note that $X_{(0, -)}$ is defined on all of $(0, \infty)$, but now approaches the limit $X_{(0, -)}(h) \rightarrow 2$ as $h \rightarrow 0$. Hence we obtain

$$h^2 Y_{(0, -)}(h) \rightarrow -1$$

in said limit, showing (iii) for $e_1 = 0$.

Finally, if $e_1 < 0$, the function $X_{(e_1, -)}$ is positive only if we restrict it to $(h_{\text{crit}}, \infty)$. In particular, we have $X_{(e_1, -)}(h) \rightarrow 0$ as $h \rightarrow h_{\text{crit}}$. Expanding the respective occurrences of f and its derivatives to sufficient high order in x (cf. Appendix A) we find

$$\lim_{h \rightarrow h_{\text{crit}}} \frac{\partial^2 F}{\partial h^2}(X_{(e_1, -)}(h), h) = \frac{29}{15} - \frac{2}{h_{\text{crit}}^2} - \frac{e_1}{5h_{\text{crit}}^4} > 0.$$

Note that inserting h_{crit} as defined in the lemma the positivity of the latter expression is to be seen in a straightforward computation. Moreover, we find that

$$\frac{\partial^2 F}{\partial x^2}(X_{(e_1, -)}(h), h) = \frac{h^2}{2} - f' \circ X_{(e_1, -)}(h) - \left(\frac{1}{2} X_{(e_1, -)}(h) - 1 \right) f'' \circ X_{(e_1, -)}(h) \rightarrow -\infty,$$

as $h \rightarrow h_{\text{crit}}$, where for the limit we note that $f'(x) \rightarrow +\infty$ and $f''(x) \rightarrow -\infty$ as $x \rightarrow 0$. At last,

$$\lim_{h \rightarrow h_{\text{crit}}} \frac{\partial^2 F}{\partial h \partial x}(X_{(e_1, -)}(h), h) = -2h_{\text{crit}} - \frac{1}{h_{\text{crit}}},$$

in particular, this limit exists. Together these three limits imply $Y_{(e_1, -)}(h) \rightarrow -\infty$ as $h \rightarrow h_{\text{crit}}$, proving (iv). \square

Remark 4.21. (i) Note that $Y_{(\frac{15}{2}, +)}$ is not considered for small h in the lemma. Indeed, due to $-\frac{8}{225} + \frac{16\sqrt{30}}{1575} > 0$, cf. (30), a claim for $Y_{(\frac{15}{2}, +)}$ similar to parts (ii) or (iii) of the lemma is false.

(ii) If $e_1 > \frac{15}{2}$, both $X_{(e_1, \pm)}$ are defined on $[(e_1 - \frac{15}{2})^{1/4}, \infty)$ and bounded, hence $Y_{(e_1, \pm)}$ possesses a limit as $h \rightarrow (e_1 - \frac{15}{2})^{1/4}$ and there is no need to determine an asymptotic behavior.

Lemma 4.22. Let $e_1 \in \mathbb{R}$. For all $h > 0$, either $Y_{(e_1, +)}(h) < 0$ or $\partial_x F(X_{(e_1, +)}(h), h) \neq 0$.

Numerical evidence. In Figures 5.(i)-(iii) we show the functions $(h, e_1) \mapsto Y_{(e_1, \pm)}(h)$ in terms of their level sets, where (iii) is a zoom into (i). Note that we can observe

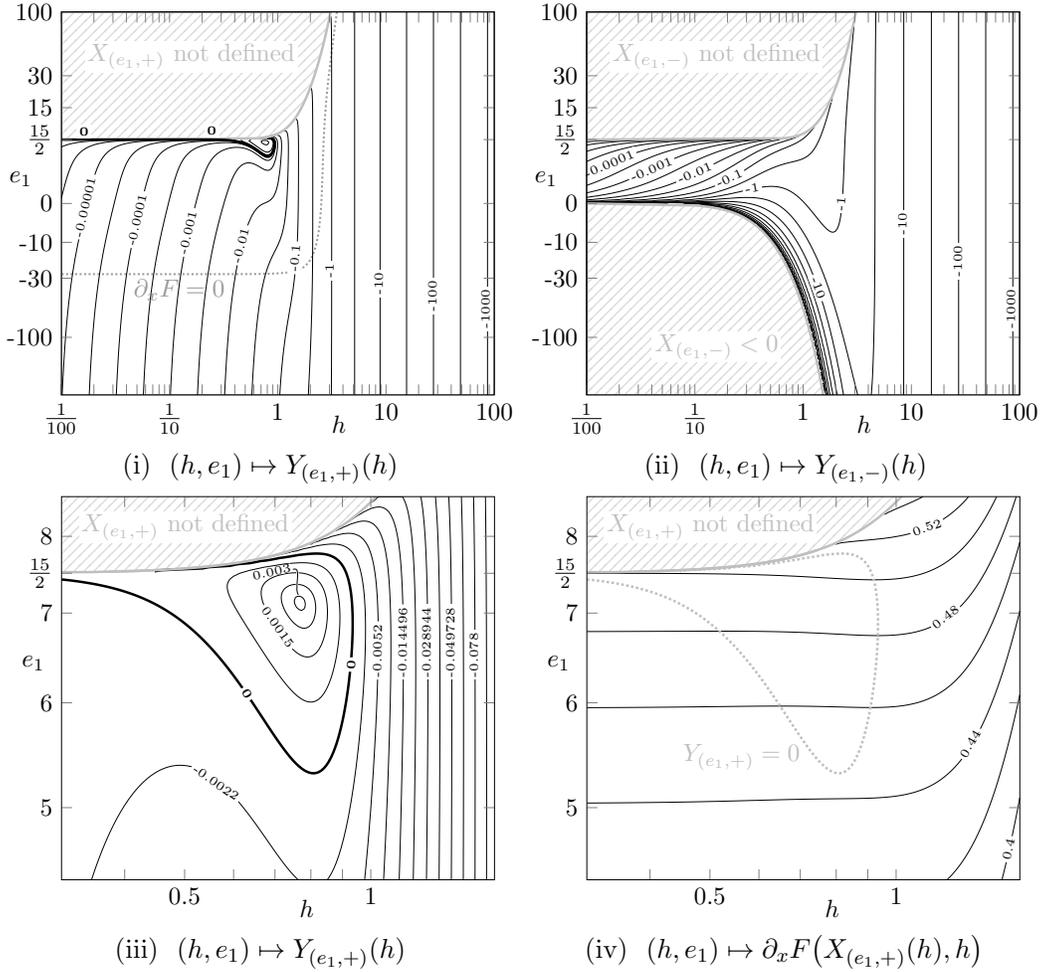


Figure 5: The graphics (i)-(iii) show the level sets and particularly the zero sets (if non-empty as thick lines) of the functions $Y_{(e_1, \pm)}$ in dependence of both e_1 and h , where (iii) is a zoom into (i) at specific values. (iv) shows the level sets of $\partial_x F$ along the graph of $X_{(e_1, +)}$, again in dependence of e_1 and h in the region where $Y_{(e_1, +)}$ takes non-negative values. While we chose a logarithmic scaling for the horizontal axes, the vertical axes are rescaled by a 3rd-order polynomial which is approximately linear around the distinguished values $\{0, \frac{15}{2}\} \ni e_1$, but strongly compresses on the ends of large absolute values. The dotted gray lines in (i) and (iv) mark the zero set of the respective other graphic for orientation.

all analytical assertions of Lemma 4.20. Figure 5.(iv) shows the function $(h, e_1) \mapsto \partial_x F(X_{(e_1,+)}(h), h)$ in terms of its level sets.

Observing that all the asymptotic assertions of Lemma 4.20 are already visible in Figure 5, we have no doubt that $Y_{(e_1,-)}$ is negative wherever it is defined and that the zeros of $Y_{(e_1,+)}$ and $\partial_x F(X_{(e_1,+)}(\cdot), \cdot)$ are widely separated by a considerable margin in the (e_1, h) -plane. Thus we conclude the lemma from numerical evidence. \diamond

To this end, F cannot have a local maximum. As a consequence we obtain the following useful lemma.

Proposition 4.23. *The open set*

$$M = \{ (x, h) \in (0, \infty) \times (0, \infty) \mid F(x, h) \neq 0 \} \subset (0, \infty) \times (0, \infty)$$

of non-solutions to the massive consistency equation $F(x, h) = 0$ possesses no (non-empty) connected component whose closure (w.r.t. \mathbb{R}^2) is a compact subset of $(0, \infty) \times (0, \infty)$.

Proof. Suppose the contrary assertion holds, that is, let N be a connected component of M such that $\overline{N} \subset (0, \infty) \times (0, \infty)$ is compact (in \mathbb{R}^2). As a connected component of the open set M , N is itself open, hence $\overline{N} = N \cup \partial N$ is a disjoint union. On the other hand, $\partial N \subset (0, \infty) \times (0, \infty)$, that is, F is defined on ∂N with $F|_{\partial N} = 0$. Finally, F attains both maximum and minimum on the compact set \overline{N} and assuming that N is non-empty we have either $\max F|_{\overline{N}} > 0$ or $\min F|_{\overline{N}} < 0$. In both cases, F necessarily has a local extremum in N which contradicts Lemma 4.22. \square

5. Asymptotic behavior of the solution set

After we have determined the structure of the solution set to $F(x, h) = 0$ in Theorem 4.2, the present section addresses the identification of asymptotic solution curves. Under an asymptotic solution curve we understand a curve which in a certain region of large or small x or h approximates an actual solution curve to a sufficiently high order. Hereby, sufficiently high means that by this approximation we can isolate solution curves in order to count them. Below we will be more precise about this notion.

By the results of the previous section each solution curve must eventually terminate in such an asymptotic solution curve. Conversely, since there can be no additional compact connected components of solutions we find by continuing the solution curves in the asymptotic regimes already all solution curves constituting the solution set of $F(x, h) = 0$.

Denote for $i \in \{0, -1\}$

$$\beta_i := -\frac{15}{2}W_i(-e^{-2})^2 - 15W_i(-e^{-2}), \quad (31)$$

where W_0 is the principal branch of the Lambert- W -function and W_{-1} is the (-1)-st subprincipal branch. Numerically, $\beta_0 \approx 2.1903$ and $\beta_{-1} \approx -27.046$. The Lambert- W -

function's branches are the piecewise inverses of the map $\mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto te^t$ and occur in the asymptotic behavior of $F(x, h)$ at large x .

We introduce the following manner of speaking in order to characterize the asymptotic behavior of \mathcal{S}_{e_1, e_2} .

Definition 5.1. *An asymptotic solution curve $\gamma : I \rightarrow (0, \infty) \times (0, \infty)$ in some limit*

$$x \rightarrow 0, \quad x \rightarrow \infty, \quad h \rightarrow 0 \quad \text{or} \quad h \rightarrow \infty \quad (32)$$

(or a combination of these limits) is a curve for which at least one of the limits (32) holds and which approximates one curve constituting \mathcal{S}_{e_1, e_2} (according to Theorem 4.2) to sufficiently high order.

Remark 5.2. (i) *Note that F does not necessarily vanish along an asymptotic solution curve, that is, an asymptotic solution curve is not a solution curve in the sense of Section 4.*

(ii) *We intentionally kept the approximation order of an asymptotic solution curve vague in the previous definition. What is a sufficient order depends on the different regimes, in particular, whether the graph of $h \mapsto (X_{(e_1, \pm)}(h), h)$ (cf. Section 4.18) comes amiss in the respective regime.*

For example, at small x we can show for any parameter setting that the graph of $h \mapsto (X_{(e_1, \pm)}(h), h)$ is bounded away from \mathcal{S}_{e_1, e_2} and a rather low order suffices. On the other hand, at large x there are up to three solution curves in \mathcal{S}_{e_1, e_2} , all approximated by $x \mapsto (x, \alpha/\sqrt{x})$ for some values of $\alpha > 0$, and, moreover, the graphs of $h \mapsto (X_{(e_1, \pm)}(h), h)$ are (for some e_1) approximated by such curves as well. Hence, in the latter regime, we need to determine the solution curve's asymptotics to a higher order such that we can separate them both among each other and from the graphs of $h \mapsto (X_{(e_1, \pm)}(h), h)$.

Note that we will mostly express an approximation order employing the Landau-function classes $\mathcal{O}(\cdot)$ and $\mathfrak{o}(\cdot)$.

Theorem 5.3. *Let $e_1, e_2 \in \mathbb{R}$. The solution set \mathcal{S}_{e_1, e_2} of the massive consistency equation $F(x, h) = 0$ in (24) can be parameterized by three analytic curves in the cases*

$$e_1 \in (\beta_{-1}, \beta_0), \quad e_1 = \beta_{-1} \text{ and } e_2 \geq -10 \quad \text{or} \quad e_1 = \beta_0 \text{ and } e_2 \leq -10 \quad (33)$$

and by two analytic curves otherwise. More detailed, we have:

- (i) *Three asymptotic solution curves at large h defined by the lines of constant $x \in \{0, x_{(-)}, x_{(+)}\}$. Each of them approximates an analytic solution curve which can be represented as the graph of an analytic function $x \mapsto h(x)$ on an interval of the form $(0, \varepsilon)$, $(x_{(-)}, x_{(-)} + \varepsilon)$ or $(x_{(+)}, x_{(+)} + \varepsilon)$ (for a sufficiently small $\varepsilon > 0$), respectively, and $h(x) \rightarrow \infty$ as $x \rightarrow 0$, $x \rightarrow x_{(-)}$ or $x \rightarrow x_{(+)}$.*
- (ii) *One asymptotic solution curve at small h in the following cases.*

- (a) If $e_1 < 0$, the line of constant $x = 0$ approximates an analytic solution curve which is representable as the graph of an analytic function $x \mapsto h(x)$ defined on an interval of the form $(0, \varepsilon)$ (for some sufficiently small $\varepsilon > 0$) with $h(x) \rightarrow 0$ as $x \rightarrow 0$.
- (b) If $e_1 = 0$ and $e_2 < 0$, the line of constant $h = 0$ approximates an analytic solution curve which is representable as the graph of an analytic function $x \mapsto h(x)$ defined on an interval of the form $(2 - \varepsilon, 2)$ (for some sufficiently small $\varepsilon > 0$) with $h(x) \rightarrow 0$ as $x \rightarrow 2$.
- (c) If $e_1 = e_2 = 0$, the line of constant $x = 2$ approximates an analytic solution curve which is representable as the graph of an analytic function $x \mapsto h(x)$ defined on an interval of the form $(2, 2 + \varepsilon)$ (for some sufficiently small $\varepsilon > 0$) with $h(x) \rightarrow 0$ as $x \rightarrow 2$.
- (d) If $e_1 = 0$ and $e_2 > 0$, the line of constant $h = 0$ approximates an analytic solution curve which is representable as the graph of an analytic function $x \mapsto h(x)$ defined on an interval of the form $(2, 2 + \varepsilon)$ (for some sufficiently small $\varepsilon > 0$) with $h(x) \rightarrow 0$ as $x \rightarrow 2$.
- (e) If $e_1 > 0$, the map

$$\gamma_\alpha : (x_0, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto \left(x, \frac{\alpha}{\sqrt{x}}\right) \quad (34)$$

for an appropriate $\alpha > 0$ and some $x_0 > 0$ approximates an analytic solution curve at large x which is also representable as the graph of an analytic function $x \mapsto h(x)$ defined on (x_0, ∞) with

$$h(x) \in \frac{\alpha}{\sqrt{x}} + \mathfrak{o}(x^{-1/2}) \quad (35)$$

as $x \rightarrow \infty$.

- (iii) In the cases (33) there are two additional analytic solution curves which are approximated at large x by γ_α from (34) with appropriate (not necessarily different) values of α . Also these solution curves are representable as in Case (ii).(e) with approximation of the form (35).

Consequently, these six (in the cases (33)) or four (otherwise) curves can be extended according to Theorem 4.2 in order to connect into three or two analytic curves, respectively, in a tree-like manner.

Proof. The presence of the solution curves and their asymptotics in the theorem are studied in Sections 5.1, 5.2 and 5.3, treating separately the regimes of large x , of small x and of x -values in between, respectively. The remaining assertions follow from Theorem 4.2. \square

We will, moreover, distinguish cases where there exist solutions at conformal and minimal coupling in Section 5.4 as follows.

Proposition 5.4. *In general, there are at most two solutions of the consistency equation (24) with minimal coupling $\xi = 0$ and at most three solutions with conformal coupling $\xi = \frac{1}{6}$. In dependence of the parameters e_1 and e_2 these bounds can be lowered.*

Proof. This proposition immediately follows from the findings of Section 5.4. \square

This analysis allows to locate the solution curves more accurately and in some cases even shows which of the four or six solution curves in the asymptotics are analytically continuable to one another. For example, suppose we have found two solution curves in the asymptotics for which $\xi < \frac{1}{6}$ holds and two such curves with $\xi > \frac{1}{6}$ for parameters at which there are no solutions with conformal coupling $\xi = \frac{1}{6}$. Then we can already conclude that the two curves with $\xi < \frac{1}{6}$ analytically extend into one another and, analogously, the other two curves extend into one another as well. Note that, however, the analysis of the particular $\xi \in \{0, \frac{1}{6}\}$ -cases is not carried out into every last possible detail and the results should rather be viewed as a rough localization (in some cases) of the solution subvarieties which constitute the solution set \mathcal{S}_{e_1, e_2} in the ambient plane $(0, \infty) \times (0, \infty)$.

5.1. Asymptotics at large x

In the present section we study the asymptotics at large x in the following lemmas. At first, we prove a quite obvious lemma which will be refined afterwards.

Lemma 5.5. *There exists $(x_0, h_0) \in (0, \infty) \times (0, \infty)$ such that F is strictly positive on $(x_0, \infty) \times (h_0, \infty)$. Consequently, any solution of $F(x, h) = 0$ lies in the union of stripes $(0, x_0) \times (0, \infty) \cup (0, \infty) \times (0, h_0)$.*

Proof. This lemma simply follows by noting that for large x and for values of h which are bounded away from 0 the term $\frac{h^2 x^2}{2}$ in F_1 (cf. (24)) is dominant for F . \square

For small h , the situation is more complicated as there are multiple competing terms in F . We study this limit by evaluating F along the curves

$$\gamma_\alpha : (0, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto \left(x, \frac{\alpha}{\sqrt{x}}\right), \quad \alpha > 0. \quad (36)$$

Alternatively, studying F on these curves may be viewed as a reparameterization of $(0, \infty) \times (0, \infty)$ into the coordinates (x, α) . If we then find a solution curve $x \mapsto (x, \alpha(x))$ for which $\alpha(x)$ converges to some α_0 as $x \rightarrow \infty$, then in our original coordinates (x, h) we have a solution curve which to order $\mathcal{O}(x^{-1/2})$ (for $h(x)$ as $x \rightarrow \infty$) is approximated by the asymptotic solution curve γ_{α_0} .

We obtain the equation

$$\begin{aligned} 0 &= F\left(x, \frac{\alpha}{\sqrt{x}}\right) \\ &= s(\alpha)x + \left(1 - \alpha^2 - \frac{e_2}{30} + 2\log(\alpha) - \frac{x}{2}(f(x) - \log(x))\right) + \left(\frac{29\alpha^2}{30x} + f(x) - \log(x)\right), \end{aligned} \quad (37)$$

where we grouped the terms according to their asymptotic behaviour in the limit $x \rightarrow \infty$ into the divergent term, the bounded (but non-zero) terms and the terms vanishing in that limit. For the dominant term we introduced the function

$$s : (0, \infty) \rightarrow \mathbb{R}, \quad \alpha \mapsto \frac{\alpha^2}{4} - \frac{1}{2} - \frac{e_1}{30\alpha^2} - \log(\alpha),$$

For the assignment of $f - \log$ to the terms with the respective behavior we refer to the asymptotic expansion of f in Appendix A.

We are particularly interested in the zeros of the function s and whether s changes sign in its zeros. Note that, given $\alpha_0 \in (0, \infty)$ with $s(\alpha_0) = 0$ and $\varepsilon > 0$ with (w.l.o.g., otherwise the same follows analogously) $s < 0$ on $[\alpha_0 - \varepsilon, \alpha_0)$ and $s > 0$ on $(\alpha_0, \alpha_0 + \varepsilon]$, we observe that

$$F\left(x, \frac{\alpha_0 - \varepsilon}{\sqrt{x}}\right) \rightarrow -\infty \quad \text{as well as} \quad F\left(x, \frac{\alpha_0 + \varepsilon}{\sqrt{x}}\right) \rightarrow \infty,$$

in particular, there exist x_0 large enough such that

$$F\left(x, \frac{\alpha_0 - \varepsilon}{\sqrt{x}}\right) < 0 \quad \text{and} \quad F\left(x, \frac{\alpha_0 + \varepsilon}{\sqrt{x}}\right) > 0$$

for any $x > x_0$. Consequently, for any such x there exists a solution (x, h) of $F(x, h) = 0$ with $\frac{\alpha_0 - \varepsilon}{\sqrt{x}} < h < \frac{\alpha_0 + \varepsilon}{\sqrt{x}}$. Below we will show that each such solution belongs to an analytic curve for which, consequently, γ_{α_0} is the asymptotic approximation. Zeros of s without sign change are a more subtle and will be treated below as well.

We study the function s . For the following lemma recall the definition and approximate values of $\beta_0 \approx 2.1903 \in (0, \frac{15}{2})$ and $\beta_{-1} \approx -27.046 < 0$ from (31).

- Lemma 5.6.** (i) *If $e_1 > \beta_0$ the function s has precisely one zero and changes sign.*
- (ii) *If $e_1 = \beta_0$ the function s has precisely two zeros and does change sign only in the larger one.*
- (iii) *If $e_1 \in (0, \beta_0)$ the function s has precisely three zeros and changes sign in each.*
- (iv) *If $e_1 \in (\beta_{-1}, 0]$ the function s has precisely two zeros and changes sign in both.*
- (v) *If $e_1 = \beta_{-1}$ the function s has precisely one zero and does not change sign.*
- (vi) *If $e_1 < \beta_{-1}$ the function s has no zero.*

Proof. Suppose $e_1 > 0$. Taking three derivatives we find that $\frac{d^3}{d\alpha^3} \alpha s(\alpha)$ is positive, hence $\alpha \mapsto \alpha s(\alpha)$ admits at most three zeros. Note that the subsets of $(0, \infty)$ on which the latter function is positive, negative or zero coincide with the respective sets for s . Thus, also s has at most three zeros. Moreover, noting that

$$\lim_{\alpha \rightarrow 0} s(\alpha) = -\infty \quad \text{and} \quad \lim_{\alpha \rightarrow \infty} s(\alpha) = +\infty, \quad (38)$$

If $e_1 = \beta_0$, the function $\alpha \mapsto \alpha^2 s(\alpha)$ vanishes in the smaller zero $\alpha_{(-)}$ of s' . Hence, s has two zeros, namely $\alpha_{(-)}$ where it consequently does not change sign and the zero in the interval $(\alpha_{(+)}, \infty)$ found above. This shows (ii). Note that s is negative on a pointed neighborhood of $\alpha_{(-)}$.

Finally, if $e_1 \in (0, \beta_0)$, the function $\alpha \mapsto \alpha^2 s(\alpha)$ is positive in the local maximum $\alpha_{(-)}$ of s . Consequently, also $s(\alpha_{(-)}) > 0$ and besides the zero larger than $\alpha_{(+)}$ from above, s has a zero in the interval $(\alpha_{(-)}, \alpha_{(+)})$ and a zero in the interval $(0, \alpha_{(-)})$. In each zero s must change sign. This shows (iii).

To show (iv)-(vi) we first note that s'' is positive for $e_1 \leq 0$. Hence, s is strictly convex and admits at most two zeros. Moreover, s' has at most one zero. Taking into account the limits

$$\lim_{\alpha \rightarrow 0} s(\alpha) = \lim_{\alpha \rightarrow \infty} s(\alpha) = +\infty,$$

we find that s has a global minimum which, consequently, is the only local extremum to be found at the unique positive solution $\alpha_{(+)}$ of (39). Note that the formula for $\alpha_{(-)}$ does not yield a positive real for $e_1 \leq 0$.

Completely analogous to the above analysis of $e_1 \mapsto \alpha_{(-)}^2 s(\alpha_{(-)})$ we study the function

$$(-\infty, 0] \rightarrow \mathbb{R}, \quad e_1 \mapsto \alpha_{(+)}^2 s(\alpha_{(+)})$$

and find an expression for $\frac{d^2}{de_1^2}(\alpha_{(+)}^2 s(\alpha_{(+)}))$ similar to (41) which now is positive, hence also $e_1 \mapsto \alpha_{(+)}^2 s(\alpha_{(+)})$ has at most two zeros. With the limits

$$\lim_{e_1 \rightarrow -\infty} \alpha_{(+)}^2 s(\alpha_{(+)}) = +\infty \quad \text{and} \quad \alpha_{(+)}^2 s(\alpha_{(+)}) \Big|_{e_1=0} = -\log(2) < 0,$$

said function has exactly one zero to be found at

$$e_1 = \beta_{-1} := -\frac{15}{2} W_{-1}(-e^{-2})^2 - 15 W_{-1}(-e^{-2}),$$

in particular, it is positive on $(-\infty, \beta_{-1})$ and negative on $(\beta_{-1}, 0]$.

Consequently, if $e_1 \in (\beta_{-1}, 0]$ the function $\alpha \mapsto \alpha^2 s(\alpha)$ is negative in the global minimum $\alpha_{(+)}$ of s , thus $s(\alpha_{(+)}) < 0$ and s has exactly two zeros, one in the interval $(0, \alpha_{(+)})$ and one in $(\alpha_{(+)}, \infty)$. In both zeros it changes sign and (iv) follows.

If $e_1 = \beta_{-1}$ the function $\alpha \mapsto \alpha^2 s(\alpha)$ vanishes in $\alpha_{(+)}$, hence s vanishes in its global minimum $\alpha_{(+)}$. This proves (v). Note that s is positive everywhere but in $\alpha_{(+)}$.

Finally, if $e_1 \in (-\infty, \beta_{-1})$ the function $e_1 \mapsto \alpha_{(+)}^2 s(\alpha_{(+)})$ is positive in $\alpha_{(+)}$ and thus s is positive in its global minimum. This shows (vi) and completes the proof. \square

Corollary 5.7. *If α_0 is a zero of s with sign change, then $s'(\alpha_0) \neq 0$.*

Proof. In the previous lemma's proof the zeros of s' were labeled $\alpha_{(\pm)}$. Moreover, the values $e_1 \in \{\beta_{-1}, \beta_0\}$ where the only cases in which s vanishes in a zero of s' . In these cases, in turn, s did not change sign in its respective zero. \square

Lemma 5.8. *For any zero α_0 at which s changes sign the curve γ_{α_0} in (36) is an asymptotic solution curve in the limit $x \rightarrow \infty$. The values of γ_{α_0} approximate exactly one analytic solution curve.*

Proof. Above we have noted, if s changes sign in a zero α_0 , then we find $\varepsilon > 0$ such that (w.l.o.g., otherwise interchange ‘>’ and ‘<’ accordingly) $s < 0$ on $[\alpha_0 - \varepsilon, \alpha_0]$ and $s > 0$ on $(\alpha_0, \alpha_0 + \varepsilon]$ and thus there is $x_0 > 0$ such that

$$F\left(x, \frac{\alpha_0 - \varepsilon}{\sqrt{x}}\right) < 0 \quad \text{and} \quad F\left(x, \frac{\alpha_0 + \varepsilon}{\sqrt{x}}\right) > 0$$

for any $x > x_0$. This implies that there is at least one solution (x, h) of $F(x, h) = 0$ with $h \in (\gamma_{\alpha_0 - \varepsilon}(x), \gamma_{\alpha_0 + \varepsilon}(x))$ for such x .

By Corollary 5.7 we can make ε small enough such that s is monotonous on the interval $[\alpha_0 - \varepsilon, \alpha_0 + \varepsilon]$. Consequently, possibly by choosing a larger x_0 , we have that also the map

$$[\alpha_0 - \varepsilon, \alpha_0 + \varepsilon] \rightarrow \mathbb{R}, \quad \alpha \mapsto F\left(x, \frac{\alpha}{\sqrt{x}}\right)$$

is monotonous for each $x > x_0$ and thus, it has precisely one zero. Hence we obtain a map $h : (x_0, \infty) \rightarrow (0, \infty)$ with $F(x, h(x)) = 0$ and $F(x, \tilde{h}) \neq 0$ for all $\tilde{h} \in (\gamma_{\alpha_0 - \varepsilon}(x), \gamma_{\alpha_0 + \varepsilon}(x)) \setminus \{h(x)\}$, at all $x > x_0$.

To this point we do not know whether h is analytic. However, since in the construction of h we can shrink $\varepsilon > 0$ at will, we know that

$$h(x) \in \frac{\alpha_0}{\sqrt{x}} + \mathfrak{o}(x^{-1/2})$$

in the limit $x \rightarrow \infty$ and hence γ_{α_0} approximates the solution curves with an error better than any $\pm \frac{\varepsilon}{\sqrt{x}}$.

Recall the asymptotic expansion of $X_{(e_1, \pm)}$ stated in Lemma 4.18. The leading order was given by

$$X_{(e_1, \pm)}(\tilde{h}) = \frac{\alpha_{(\pm)}^2}{\tilde{h}^2} + \mathcal{O}(1) \quad \text{as } \tilde{h} \rightarrow 0 \quad (42)$$

(compare Lemma 4.18.(ii) and Equation (39)). In the cases where $X_{(e_1, \pm)}(\tilde{h}) \rightarrow +\infty$ as $\tilde{h} \rightarrow 0$ (i.e. $e_1 \in (0, \frac{15}{2}]$ for $X_{(e_1, -)}$ and $e_1 \leq \frac{15}{2}$ for $X_{(e_1, +)}$) we can solve the curves

$$(0, \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad \tilde{h} \mapsto (X_{(e_1, \pm)}(\tilde{h}), \tilde{h})$$

($\varepsilon > 0$) for \tilde{h} and obtain

$$(x_1, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto (x, h_{(\pm)}(x))$$

for a sufficiently large x_1 and some functions $h_{(\pm)} : (x_1, \infty) \rightarrow (0, \infty)$. By (42) we have

$$h_{(\pm)}(x) = \frac{\alpha_{(\pm)}}{\sqrt{x}} + \mathcal{o}(x^{-1/2})$$

in the limit $x \rightarrow \infty$. Since now α_0 was assumed to be a zero with sign change, it coincides in particular neither with $\alpha_{(+)}$ nor with $\alpha_{(-)}$. Hence, the solutions $h(x)$ which we found above and the points of the form $(X_{(e_1, \pm)}(h), h)$ in which the solution set of $F(x, h) = 0$ is not locally soluble for h are asymptotically, at large x , bounded away from each other. Consequently, h is analytic. \square

Remark 5.9. (i) *Note that to this point we have found all solutions in the limit $x \rightarrow \infty$ for $e_1 \in (\beta_{-1}, \beta_0)$. By “all solutions in the limit” we mean that there exists x_0 such that any solution at $x > 0$ belongs to one of the curves provided by Lemma 5.6, Parts (iii),(iv) and (v) via Lemma 5.8. That there can be no more solutions can be concluded using the upper bounds in Lemma 4.11, matching the amount of zeros of s .*

(ii) *Also, we have found all solution curves for $e_1 > \frac{15}{2}$ using the same argument. That is, s in that case has one zero α_0 and we have one solution curve approximated by the respective γ_{α_0} . On the other hand, by Lemma 4.18.(vi) the functions $X_{(e_1, \pm)}$ are bounded from above, in particular, $\partial_h F(x, \cdot)$ has no zeros for x above such an upper bound on $X_{(e_1, +)}$ and thus, at any x above that bound, there exist at most one solution.*

The following lemma shows that any solution in the limit of large x is approximated by γ_α with $s(\alpha) = 0$.

Lemma 5.10. *Any solution of $F(x, h) = 0$ in the limit $x \rightarrow \infty$ is approximated in that limit by a curve of the form γ_α from (36), with a zero α of s . The approximation is valid in the function class $\mathcal{o}(x^{-1/2})$.*

Proof. Note that, we have

$$\lim_{\alpha \rightarrow \infty} s(\alpha) = +\infty, \quad \lim_{\alpha \rightarrow \infty} s'(\alpha) = +\infty \quad \text{as well as} \quad \lim_{\alpha \rightarrow \infty} s''(\alpha) = \frac{1}{2},$$

hence we can find $\alpha_1 \in (0, \infty)$ such that

$$s(\alpha_1) > 0, \quad s'(\alpha_1) > 0 \quad \text{as well as} \quad s''(\alpha) \geq \frac{1}{4} \quad \text{for all } \alpha \geq \alpha_1$$

holds. By applying the fundamental theorem of calculus we also have

$$s'(\alpha) \geq s'(\alpha_1) \quad \text{and} \quad s(\alpha) \geq s(\alpha_1)$$

for all $\alpha \geq \alpha_1$ and thus we can find $x_1 \in (0, \infty)$ such that

$$\frac{\partial^2}{\partial \alpha^2} F\left(x, \frac{\alpha}{\sqrt{x}}\right) = s''(\alpha)x - 2 - \frac{2}{\alpha^2} + \frac{29}{15x} > 0$$

for all $x \geq x_1$ and all $\alpha \geq \alpha_1$.

From (37) we can read off that

$$\lim_{x \rightarrow \infty} F\left(x, \frac{\alpha_1}{\sqrt{x}}\right) = +\infty.$$

Moreover, by taking one x -derivative of (37) we can also read off that

$$\lim_{x \rightarrow \infty} \frac{\partial}{\partial x} F\left(x, \frac{\alpha_1}{\sqrt{x}}\right) = s(\alpha_1),$$

where we used that $f - \log$ possesses an asymptotic Puiseux expansion allowing term-wise differentiation and, consequently, showing that $\frac{\partial}{\partial x} \frac{x}{2}(f(x) - \log(x)) \in \mathcal{O}(\frac{1}{x})$ and $\frac{\partial}{\partial x}(f(x) - \log(x)) \in \mathcal{O}(\frac{1}{x^2})$, cf. Appendix A. By enlarging x_1 , if necessary, we can guarantee that

$$F\left(x_1, \frac{\alpha_1}{\sqrt{x_1}}\right) > 0 \quad \text{and} \quad \frac{\partial}{\partial x} F\left(x, \frac{\alpha_1}{\sqrt{x}}\right) \geq \frac{s(\alpha_1)}{2} \quad \text{for all } x \geq x_1.$$

Finally, for any $(x, \alpha) \in [x_1, \infty) \times [\alpha_1, \infty)$ we conclude that

$$\begin{aligned} F\left(x, \frac{\alpha}{\sqrt{x}}\right) &= F\left(x, \frac{\alpha_1}{\sqrt{x}}\right) + \frac{\partial}{\partial \alpha} F\left(x, \frac{\alpha}{\sqrt{x}}\right) \Big|_{\alpha=\alpha_1} (\alpha - \alpha_1) + \int_{\alpha_1}^{\alpha} d\tilde{\alpha} \int_{\alpha_1}^{\tilde{\alpha}} d\tilde{\alpha} \frac{\partial^2}{\partial \tilde{\alpha}^2} F\left(x, \frac{\tilde{\alpha}}{\sqrt{x}}\right) \\ &\geq F\left(x, \frac{\alpha_1}{\sqrt{x}}\right) \\ &= F\left(x_1, \frac{\alpha_1}{\sqrt{x_1}}\right) + \int_{x_1}^x d\tilde{x} \frac{\partial}{\partial \tilde{x}} F\left(\tilde{x}, \frac{\alpha_1}{\sqrt{\tilde{x}}}\right) \\ &\geq F\left(x_1, \frac{\alpha_1}{\sqrt{x_1}}\right) \\ &> 0, \end{aligned}$$

where we applied the fundamental theorem of calculus several times and used the above estimates on the derivatives of F in the coordinates (x, α) in the respective regime. Consequently, such $(x, \alpha) \in [x_1, \infty) \times [\alpha_1, \infty)$ cannot be a solution to (37).

We work out a similar argument at small α . In this way we distinguish the cases of positive, negative or zero e_1 .

At first, in the case $e_1 < 0$ we have a similar argument as above. Therefore we reparameterize (37) via $\alpha \mapsto \frac{1}{\alpha}$, that is, we consider the equation

$$\begin{aligned} 0 &= F\left(x, \frac{1}{\alpha\sqrt{x}}\right) \\ &= s\left(\frac{1}{\alpha}\right)x + \left(1 - \frac{1}{\alpha^2} - \frac{e_2}{30} - 2\log(\alpha) - \frac{x}{2}(f(x) - \log(x))\right) + \left(\frac{29}{30x\alpha^2} + f(x) - \log(x)\right). \end{aligned}$$

Since $e_1 < 0$ we have

$$\lim_{\alpha \rightarrow \infty} s\left(\frac{1}{\alpha}\right) = +\infty \quad \text{as well as} \quad \lim_{\alpha \rightarrow \infty} \frac{\partial}{\partial \alpha} s\left(\frac{1}{\alpha}\right) = \lim_{\alpha \rightarrow \infty} \left(-\frac{1}{2\alpha^3} - \frac{e_1 \alpha}{15} + \frac{1}{\alpha}\right) = +\infty,$$

hence we can find $\alpha_2 \in (0, \infty)$ such that

$$s\left(\frac{1}{\alpha}\right) \geq s(\alpha_2) > 0 \quad \text{as well as} \quad \frac{\partial}{\partial \alpha} s\left(\frac{1}{\alpha}\right) \geq 1$$

for all $\alpha \geq \frac{1}{\alpha_2}$. In this way we choose 1 as an arbitrary, but still positive, lower bound on $\frac{\partial}{\partial \alpha} s\left(\frac{1}{\alpha}\right)$ above $\frac{1}{\alpha_2}$.

Now we can find $x_2 \in (0, \infty)$ such that

$$\frac{\partial}{\partial \alpha} F\left(x, \frac{1}{\alpha\sqrt{x}}\right) = \frac{\partial}{\partial \alpha} s\left(\frac{1}{\alpha}\right)x + \frac{2}{\alpha^3} - \frac{2}{\alpha} - \frac{29}{15x\alpha^3} > 0$$

for all $x \geq x_2$ and all $\alpha \geq \frac{1}{\alpha_2}$. Again, we observe that

$$\lim_{x \rightarrow \infty} F\left(x, \frac{\alpha_2}{\sqrt{x}}\right) = +\infty \quad \text{as well as} \quad \lim_{x \rightarrow \infty} \frac{\partial}{\partial x} F\left(x, \frac{\alpha_2}{\sqrt{x}}\right) = s(\alpha_2),$$

and by possibly enlarging x_2 we can guarantee that

$$F\left(x_2, \frac{\alpha_2}{\sqrt{x_2}}\right) > 0 \quad \text{and} \quad \frac{\partial}{\partial x} F\left(x, \frac{\alpha_2}{\sqrt{x}}\right) \geq \frac{s(\alpha_2)}{2} \quad \text{for all } x \geq x_2,$$

again using the asymptotic expansion of $f - \log$.

Finally, we compute for any $(x, \alpha) \in [x_2, \infty) \times [\frac{1}{\alpha_2}, \infty)$ that

$$\begin{aligned} F\left(x, \frac{1}{\alpha\sqrt{x}}\right) &= F\left(x, \frac{\alpha_2}{\sqrt{x}}\right) + \int_{1/\alpha_2}^{\alpha} d\tilde{\alpha} \frac{\partial}{\partial \tilde{\alpha}} F\left(x, \frac{1}{\tilde{\alpha}\sqrt{x}}\right) \\ &\geq F\left(x, \frac{\alpha_2}{\sqrt{x}}\right) \\ &= F\left(x_2, \frac{\alpha_2}{\sqrt{x_2}}\right) + \int_{x_2}^x d\tilde{x} F\left(\tilde{x}, \frac{\alpha_2}{\sqrt{\tilde{x}}}\right) \\ &\geq F\left(x_2, \frac{\alpha_2}{\sqrt{x_2}}\right) \\ &> 0, \end{aligned}$$

showing that the equation $F\left(x, \frac{1}{\alpha\sqrt{x}}\right) = 0$ has no solution in $[x_2, \infty) \times [\frac{1}{\alpha_2}, \infty) \ni (x, \alpha)$.

Reversing our reparameterization $\alpha \mapsto \frac{1}{\alpha}$ from above, this means that the equation $F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0$ has no solution in $[x_2, \infty) \times (0, \alpha_2] \ni (x, \alpha)$.

To find α_2 and x_2 in the case $e_1 > 0$ is much easier. At first, observe that

$$\lim_{\alpha \rightarrow 0} s(\alpha) = -\infty \quad \text{as well as} \quad \lim_{\alpha \rightarrow 0} s'(\alpha) = +\infty,$$

hence there exist $\alpha_2 \in (0, \infty)$ such that

$$s(\alpha) \leq s(\alpha_2) < 0 \quad \text{for all } \alpha < \alpha_2.$$

From (37) we can read off that there exist $M \in \mathbb{R}$ such that

$$F\left(x, \frac{\alpha}{\sqrt{x}}\right) - s(\alpha)x \leq M$$

for all $\alpha \in (0, \alpha_2)$ and all $x \geq \frac{9}{4}$ (we employ this lower bound on x in favour of the explicit bound in (58), Appendix A). In particular, we can find $x_2 \in (0, \infty)$ such that for all $x \geq x_2$

$$F\left(x, \frac{\alpha}{\sqrt{x}}\right) \leq s(\alpha)x + M < 0,$$

for all $\alpha \in (0, \alpha_2)$. Consequently, $[x_1, \infty) \times (0, \alpha_2]$ contains no solution of $F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0$.

As a side remark, note that M is indeed a uniform bound. Such uniform bound cannot be found in the cases we have treated before. Moreover, note that we comment on the case $e_1 = 0$ at the end of the proof.

To this point, if $e_1 \neq 0$ we have found x_1, x_2, α_1 and α_2 , such that any solution of $F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0$ with $x \geq \max\{x_1, x_2\}$ must fulfill $\alpha \in [\alpha_2, \alpha_1]$.

Now let $\varepsilon > 0$ and suppose we have $\alpha \in [\alpha_2, \alpha_1]$ such that $|s(\alpha)| \geq \varepsilon$. We assume that ε is small enough, such that $s(\alpha_1) > \varepsilon$ and $s(\alpha_2) > \varepsilon$. Then, similar as above, we read off from (37) that there exists $M \in \mathbb{R}$ such that

$$\left|F\left(x, \frac{\alpha}{\sqrt{x}}\right) - s(\alpha)x\right| \leq M$$

for all $\alpha \in [\alpha_2, \alpha_1]$ and all $x \geq \frac{9}{4}$. That this bound is now uniform in α is a consequence of restricting the α -values to the compact interval $[\alpha_2, \alpha_1]$. However, we can find $x_3 \in (\max\{x_1, x_2\}, \infty)$, such that for all $x \geq x_3$ and all $\alpha \in [\alpha_2, \alpha_1]$

$$\left|F\left(x, \frac{\alpha}{\sqrt{x}}\right)\right| \geq |s(\alpha)x| - \left|F\left(x, \frac{\alpha}{\sqrt{x}}\right) - s(\alpha)x\right| \geq \varepsilon x - M > 0$$

holds. In particular, for any solution of $F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0$ with $x \geq x_3$ we have shown that necessarily $s(\alpha) < \varepsilon$ holds. As $\varepsilon > 0$ can be chosen arbitrarily small, the claim of the lemma follows for our present case $e_1 \neq 0$.

Finally, if $e_1 = 0$ one could work out a similar method. However, the occurring log-terms (which are dominant in the $e_1 = 0$ -case) prevent from analog estimates on the derivatives of s and $\alpha \mapsto F\left(x, \frac{\alpha}{\sqrt{x}}\right)$ and an adaption is not straightforward. On the other hand, a more profound method is not necessary. As we have already noted in

Remark 5.9, we have found all solutions at large x in the case $e_1 = 0$ together with the correct approximation using Lemmata 5.6.(iv) and 5.8 as well as the respective upper bound in Lemma 4.11. \square

Remark 5.11. *At this point, only the cases $e_1 \in \{\beta_0, \beta_{-1}\}$ remain open. Actually, in addition to the cases commented on in Remark 5.9 (i.e. for $e_1 \in (\beta_0, \beta_{-1})$), we have shown that for $e_1 > \beta_0$ there exists precisely one asymptotic solution curve γ_α with the single zero α of s (cf. Lemma 5.6.(i)), whereas for $e_1 < \beta_{-1}$ there exists no solution at all above a certain bound on x as s has no zero in that case (cf. Lemma 5.6.(vi)).*

Lemma 5.12. (i) *Let $e_1 = \beta_0$ and let $\alpha_{(-)}$ be the smaller zero of s provided by Lemma 5.6.(ii), defined by Equation (39). If $e_2 \leq -10$, there exist two analytic solution curves at large x approximated by $\gamma_{\alpha_{(-)}}$, whereas if $e_2 > -10$ there exist no solution curves approximated by said curve.*

(ii) *Let $e_1 = \beta_{-1}$ and let $\alpha_{(+)}$ be the only zero of s provided by Lemma 5.6.(ii), defined by Equation (39). If $e_2 \geq -10$, there exist two analytic solution curves at large x approximated by $\gamma_{\alpha_{(+)}}$, whereas if $e_2 < -10$ there exist no solution curves approximated by said curve.*

In all assertions the approximation order is given by $\frac{\alpha}{\sqrt{x}} + \mathcal{O}\left(\frac{1}{\sqrt{x}}\right)$.

Proof. For $e_1 = \beta_0$ the function $\alpha \mapsto F\left(x, \frac{\alpha}{\sqrt{x}}\right)$ has, at sufficiently large x , a local maximum at the smaller solution of

$$\frac{\partial}{\partial \alpha} F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0. \quad (43)$$

On the other hand, for $e_1 = \beta_{-1}$ said function has a local minimum at the only real (positive) solution of (43). For these assertions we recall the shape of the dominating function s from previous lemmas.

Note that (43) is a polynomial equation (in x , as (28) was, too) and we denote its solutions by

$$\begin{aligned} \tilde{\alpha}_{(\pm)} : (x_0, \infty) &\rightarrow (0, \infty), \quad \tilde{\alpha}_{(\pm)}(x) := \frac{x\left(\frac{x}{2}-1\right)}{2\left(\frac{x^2}{4}-x+\frac{29}{30}\right)} \left[1 \pm \sqrt{1 - \frac{2e_1}{15} + \frac{4e_1}{225(x-2)^2}} \right] \\ &= \alpha_{(\pm)} + \mathcal{O}\left(\frac{1}{x}\right) \end{aligned} \quad (44)$$

with some sufficiently large x_0 and the zeros $\alpha_{(\pm)}$ of s' . A lengthy computation shows the surprisingly simple result that

$$F\left(x, \frac{\tilde{\alpha}_{(\pm)}(x)}{\sqrt{x}}\right) = -\frac{e_2 + 10}{30} + \frac{1 + W_j(-e^{-2})}{30} \frac{1}{x} + \mathcal{O}\left(\frac{1}{x^2}\right) \quad (45)$$

for the presently relevant cases $(\pm, j, e_1) = (-, 0, \beta_0)$ or $(\pm, j, e_1) = (+, -1, \beta_{-1})$.

Now let $e_1 = \beta_0$. In Lemma 5.6.(ii) we have shown that s does not change sign at its zero $\alpha_{(-)}$. As we, moreover, have remarked in the proof of said lemma, s is negative on a pointed neighborhood of $\alpha_{(-)}$. Consequently, for any sufficiently small $\varepsilon > 0$, we have

$$\lim_{x \rightarrow \infty} F\left(x, \frac{\alpha_{(-)} - \varepsilon}{\sqrt{x}}\right) = \lim_{x \rightarrow \infty} F\left(x, \frac{\alpha_{(-)} + \varepsilon}{\sqrt{x}}\right) = \lim_{x \rightarrow \infty} s(\alpha_{(-)} \pm \varepsilon) x = -\infty,$$

independently of e_2 , in particular, the respective expressions are negative for x above some certain common lower bound $x_1 \in (0, \infty)$. On the other hand, from the asymptotic expansion (45) we read off for $e_2 \leq -10$ that

$$F\left(x, \frac{\tilde{\alpha}_{(-)}(x)}{\sqrt{x}}\right) > 0$$

for sufficiently large x , w.l.o.g. (if necessary enlarge x_1) we assume that the latter inequality holds for all $x \geq x_1$. For the boundary case $e_2 = -10$ the sub-leading order in (45) is decisive and we note that $\frac{1}{30}(1 + W_0(-e^{-2})) \approx 0.02805 > 0$.

By possibly enlarging x_1 again and by recalling the asymptotic expansion (44) we can, moreover, guarantee that

$$\tilde{\alpha}_{(-)}(x) \in (\alpha_{(-)} - \varepsilon, \alpha_{(-)} + \varepsilon)$$

for all $x \geq x_1$. Hence, for all $x \geq x_1$ the equation $F\left(x, \frac{\alpha}{\sqrt{x}}\right) = 0$ has solutions α in both the open intervals

$$(\alpha_{(-)} - \varepsilon, \tilde{\alpha}_{(-)}(x)) \quad \text{and} \quad (\tilde{\alpha}_{(-)}(x), \alpha_{(-)} + \varepsilon). \quad (46)$$

Together with the third solution at such sufficiently large x -values (if necessary, enlarge x_1 once again) provided by Lemma 5.8 around the larger (sign-changing) zero of s , we already exhaust the upper bound given in Lemma 4.11. Consequently, we can choose x_1 such that each of the intervals (46) contains precisely one solution, for all $x \geq x_1$.

Note that, not only does the curve defined by $\tilde{\alpha}_{(-)}$ provide points in which F is positive, it also parameterizes all points (around $\alpha_{(-)}$, particularly, bounded away from $\alpha_{(+)}$) in which the hypothesis of the implicit function theorem fails. Since we have found our two solutions in the *open* intervals (46), the latter theorem's hypothesis fails in none of these solutions and we obtain, for now in the coordinates (x, α) , indeed two analytic curves as the graphs of analytic functions of x which are, to order $\mathfrak{o}(1)$ (recall that ε can be chosen arbitrarily small), approximated by $\alpha_{(-)}$. Inverting the reparameterization $(x, h) \mapsto (x, \alpha)$, finally, provides us with the analytic solution curves in our original coordinates (x, h) , approximated in $\mathfrak{o}\left(\frac{1}{\sqrt{x}}\right)$ as asserted in the lemma.

For $e_2 > -10$ we read off from the asymptotic expansion (45) that there exists some $x_1 \in (0, \infty)$ such that

$$F\left(x, \frac{\tilde{\alpha}_{(\pm)}(x)}{\sqrt{x}}\right) < 0$$

Case		#	α
$e_1 < \beta_{-1}$	e_2 arb.	0	–
$e_1 = \beta_{-1}$	$e_2 < -10$	0	–
$e_1 = \beta_{-1}$	$e_2 \geq -10$	2	$\alpha_1 = \alpha_2 = \alpha_{(+)}$
$e_1 \in (\beta_{-1}, 0]$	e_2 arb.	2	$\alpha_1 < \alpha_{(+)} < \alpha_2$
$e_1 \in (0, \beta_0)$	e_2 arb.	3	$\alpha_1 < \alpha_{(-)} < \alpha_2 < \alpha_{(+)} < \alpha_3$
$e_1 = \beta_0$	$e_2 \leq -10$	3	$\alpha_1 = \alpha_2 = \alpha_{(-)} < \alpha_{(+)} < \alpha_3$
$e_1 = \beta_0$	$e_2 > -10$	1	$\alpha > \alpha_{(+)}$
$e_1 \in (\beta_0, \frac{15}{2})$	e_2 arb.	1	$\alpha > \alpha_{(+)}$
$e_1 = \frac{15}{2}$	e_2 arb.	1	$\alpha > 1$ ($= \alpha_{(-)} = \alpha_{(+)}$)
$e_1 > \frac{15}{2}$	e_2 arb.	1	$\alpha > 1$ (and $\alpha_{(\pm)} \notin \mathbb{R}$)

Table 2: The possible cases which can occur for solution curves in the limit $x \rightarrow \infty$, distinguished by value ranges of the parameters e_1 and e_2 . For each case we have listed how many branches the solution set \mathcal{S}_{e_1, e_2} of $F(x, h) = 0$ has in the limit $x \rightarrow \infty$ (column #), together with rough bounds on the coefficient α in terms of $\alpha_{(\pm)}$ (if possible) such that the solutions of the respective branch are approximated by $\gamma_\alpha(x) = (x, \frac{\alpha}{\sqrt{x}})$ (column α). If necessary, the α -values are labeled increasingly.

for all $x \geq x_1$. In particular, for some given (sufficiently small) $\varepsilon > 0$ the function

$$(\alpha_{(-)} - \varepsilon, \alpha_{(-)} + \varepsilon) \rightarrow \mathbb{R}, \quad \alpha \mapsto F\left(x, \frac{\alpha}{\sqrt{x}}\right) \tag{47}$$

is negative in its local maximum, for each $x \geq x_1$. Since (43) has precisely the two solutions (44), we can choose ε small enough such that $\tilde{\alpha}_{(+)}(x) \notin (\alpha_{(-)} - \varepsilon, \alpha_{(-)} + \varepsilon)$ for all $x \geq x_1$, in other words, choose ε small enough such that said negative local maximum serves as a uniform upper bound on the interval $(\alpha_{(-)} - \varepsilon, \alpha_{(-)} + \varepsilon)$. From the asymptotic expansion (45) we read off that this uniform upper bound on (47) can be estimated from above by $-\frac{1}{2} \frac{e_2 + 10}{30} < 0$ for all sufficiently large x , w.l.o.g. for $x \geq x_1$. Since at this point we have found a uniform, x -independent, negative upper bound on the values of $F(x, \frac{\alpha}{\sqrt{x}})$, we conclude that our equation $F(x, \frac{\alpha}{\sqrt{x}}) = 0$ has no solution with $x \geq x_1$ and $\alpha \in (\alpha_{(-)} - \varepsilon, \alpha_{(-)} + \varepsilon)$. In our original coordinates (x, h) this imposes what we have claimed in the lemma, finishing the proof of (i).

The proof of (ii) works in principle the same. One merely has to reverse inequalities and take into account that now the upper bound on the number of solution from Lemma 4.11 is 2 and the two solutions in the intervals analog to (46) already exhaust said upper bound. Moreover, for the boundary case $e_2 = -10$ we note that now $\frac{1}{30}(1 + W_{-1}(-e^{-2})) \approx -0.07154 < 0$, though the main argument remains the same. This finishes the proof. \square

Finally, we collect the results of the present section in a list in Table 2. Recall that for

$0 < e_1 < \frac{15}{2}$ we have $\alpha_{(-)} < 1 < \alpha_{(+)}$, that for $e_1 = \frac{15}{2}$ we have $\alpha_{(\pm)} = 1$, that for $e_1 \leq 0$ only $\alpha_{(+)}$ is positive and for $e_1 > \frac{15}{2}$ none of the $\alpha_{(\pm)}$ is a real number. Moreover, recall that in each possible case we can find $x_0 > 0$ such that either $F(x, h) = 0$ admits no solution at all for $x > x_0$ or such that each solution of $F(x, h) = 0$ is approximated by a curve $x \mapsto \gamma_\alpha(x) = (x, \frac{\alpha}{\sqrt{x}})$ in the limit $x \rightarrow \infty$ with some $\alpha \in (0, \infty)$ for which $s(\alpha) = 0$.

5.2. Asymptotics at small x

At first we study the situation of h away from the limits of small or large h .

Lemma 5.13. *For any compact interval $[a, b] \subset (0, \infty)$ there is $\varepsilon > 0$ such that $F(x, h)$ is uniformly bounded away from 0 on $(0, \varepsilon) \times [a, b] \ni (x, h)$. Consequently, the equation $F(x, h) = 0$ has no solution in $(0, \varepsilon) \times [a, b]$.*

Proof. Note that for any fixed $h \in [a, b]$ we have $F(x, h) \rightarrow -\infty$ as $x \rightarrow 0$ since the dominant term in this limit is $-\left(\frac{x}{2} - 1\right)f(x)$ in F_2 . Recall that $f(x) = -\frac{3}{x} + \mathcal{O}(1)$ in said limit (cf. Appendix A). Any other term in F can be continued to $x = 0$, that is, there exists a continuous function

$$\tilde{F} : [0, 1] \times [a, b] \rightarrow \mathbb{R} \quad \text{s.t.} \quad \tilde{F}(x, h) = F(x, h) + \left(\frac{x}{2} - 1\right)f(x)$$

for all $x > 0$ and all $h \in [a, b]$. Let M be an upper bound for the continuous function \tilde{F} on its compact domain. By the above limit we find $\varepsilon > 0$ such that

$$-\left(\frac{x}{2} - 1\right)f(x) < -M$$

for all $x < \varepsilon$. Therefore, we have for all $h \in [a, b]$ and all $x \in (0, \varepsilon)$

$$F(x, h) = \tilde{F}(x, h) - \left(\frac{x}{2} - 1\right)f(x) < 0$$

and the lemma follows. \square

By the previous lemma it remains to study the limits $h \rightarrow \infty$ and $h \rightarrow 0$ in the following two lemmata.

Lemma 5.14. *For any $e_1 \in \mathbb{R}$ we have an asymptotic solution curve in the limit $h \rightarrow \infty$ parameterized by*

$$(0, \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto \left(x, \sqrt{\frac{90}{29x}}\right)$$

for some $\varepsilon > 0$.

Proof. In the limit of large h and small x the terms

$$\frac{29}{30}h^2 + f(x)$$

are dominant. Studying when they mutually compensate (i.e. setting them to zero) and solving for x yields the lemma. Note that approximating $f(x) \approx -\frac{3}{x}$ is sufficient for this claim. \square

Remark 5.15. (i) *Using the results from Section 4.2 we can make ε small enough such that the graphs of $X_{(e_1, \pm)}$ are bounded away from our solution curve. Hence, indeed the solution curve approximated by our asymptotic solution curve in the previous lemma is solvable to be the graph of an analytic function of x on $(0, \varepsilon)$.*

(ii) *Note that, without taking into account also terms of lower order than the one considered in Lemma 5.14, we can only conclude that the solution curve is in $\sqrt{90/29x} + \mathcal{O}(x^{-1/2})$. However, this suffices to conclude from $\sqrt{90/29} > 1$ that the solution curve in consideration provides solutions at positive $\xi = \frac{1}{12}(x - \frac{1}{h^2})$, at least at sufficiently small x or sufficiently large h , respectively.*

Lemma 5.16. *If $e_1 < 0$ the map*

$$(0, \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto \left(x, \sqrt{-\frac{e_1}{90}x}\right)$$

is an asymptotic solution curve at small x with some sufficiently small $\varepsilon > 0$. If $e_1 \geq 0$ there exists $\varepsilon > 0$ such that $(0, \varepsilon) \times (0, \varepsilon)$ contains no solution of $F(x, h) = 0$.

Proof. The relevant terms in the present setting are

$$-\frac{e_1}{30} \frac{1}{h^2} + \log(h) + f(x). \quad (48)$$

Let $e_1 < 0$. Equating the first and last term of (48) (and approximating $f(x) \approx -\frac{3}{x}$, cf. Appendix A) yields the asymptotic solution curve in the lemma. In this scenario the log-term only contributes in higher orders.

Note that the graphs of $X_{(e_1, \pm)}$ cannot come amiss to the present consideration and we indeed obtain an analytic solution curve approximated by the asymptotic solution curve in the lemma.

If $e_1 \geq 0$, all the dominant terms diverge to $-\infty$ as $h \rightarrow 0$ or $x \rightarrow 0$ and since the remaining contributions into F that are left out in (48) are bounded in this limit we can find a negative upper bound on F on some $(0, \varepsilon) \times (0, \varepsilon)$, $\varepsilon > 0$. \square

5.3. Asymptotics at finite x

At first we study the regime of small h .

Lemma 5.17. (i) *Let $e_1 \neq 0$. For any compact interval $[a, b] \subset (0, \infty)$ there exists $\varepsilon > 0$ such that F is bounded away from 0 on $[a, b] \times (0, \varepsilon)$. Consequently, the equation $F(x, h) = 0$ has no solution in $[a, b] \times (0, \varepsilon)$.*

(ii) *Let $e_1 = 0$. For any compact interval $[a, b] \subset (0, \infty)$ with $2 \notin [a, b]$ there exists $\varepsilon > 0$ such that F is bounded away from 0 on $[a, b] \times (0, \varepsilon)$. Consequently, the equation $F(x, h) = 0$ has no solution in $[a, b] \times (0, \varepsilon)$.*

Proof. The proof of (i) goes very similar to the proof of Lemma 5.13 and we skip the details. Note that the term $-\frac{e_1}{30} \frac{1}{h^2}$ in F_1 is dominant in this regime and, in particular, suppresses the influence of the log-term in F_2 .

If, on the other hand, $e_1 = 0$, the dominant term in the limit $h \rightarrow 0$ is the log-term in F_2 . This term does come with the prefactor $-(\frac{x}{2} - 1)$, that is, with a prefactor which changes sign at $x = 2$. If we, however, stay away from $x = 2$ and assume that $2 \notin [a, b]$, then all terms other than the log-term are continuable to $h = 0$ on $[a, b]$. By the same argument as for Lemma 5.13 we conclude (ii). \square

Remark 5.18. *Using Lemma 5.16 we can formulate part (i) of the previous lemma (i.e. for the case $e_1 > 0$) for intervals of the form $(0, b]$.*

Next we study the limit $(x, h) \rightarrow (2, 0)$ for $e_1 = 0$ which leads to cases sensible to e_2 .

Lemma 5.19. *Let $e_1 = 0$.*

- (i) *If $e_2 > 0$, there is precisely one solution curve in a sufficiently small neighborhood of $(2, 0) \in (0, \infty) \times (0, \infty)$. This solution curve is of the form*

$$(2, 2 + \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto (x, h(x)) \quad \text{and fulfills} \quad h(x) \in \mathfrak{o}(x - 2),$$

for some $\varepsilon > 0$.

- (ii) *If $e_2 < 0$, there is precisely one solution curve in a sufficiently small neighborhood of $(2, 0) \in (0, \infty) \times (0, \infty)$. This solution curve is of the form*

$$(2 - \varepsilon, 2) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto (x, h(x)) \quad \text{and fulfills} \quad h(x) \in \mathfrak{o}(2 - x),$$

for some $\varepsilon > 0$.

- (iii) *If $e_2 = 0$, there is precisely one solution curve in a sufficiently small neighborhood of $(2, 0) \in (0, \infty) \times (0, \infty)$. This solution curve is of the form*

$$(0, \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad h \mapsto (x(h), h) \quad \text{and fulfills} \quad x(h) \in 2 + \mathcal{O}(h^2),$$

for some $\varepsilon > 0$. Moreover, $x(h) > 2$ for all $h \in (0, \delta)$.

Proof. We use polar coordinates around $(x, h) = (2, 0)$ and evaluate

$$F(2 + \varrho \sin \psi, \varrho \cos \psi) \quad \text{with} \quad \psi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad \text{and} \quad \varrho \in (0, 2).$$

The dominant terms in the limit $\varrho \rightarrow 0$ stem from the log-term in F_2 as well as from the e_2 -term which does not depend on ϱ or ψ . More precisely, evaluating the lengthy expression for $F(2 + \varrho \sin \psi, \varrho \cos \psi)$ we find that

$$F(2 + \varrho \sin \psi, \varrho \cos \psi) + \varrho \sin \psi \log(\varrho \cos \psi) + \frac{e_2}{30} \rightarrow 0 \quad \text{as} \quad \varrho \rightarrow 0,$$

uniformly in $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and at least of order $\mathcal{O}(\varrho)$. Consequently, for sufficiently small ϱ the solutions of $F(2 + \varrho \sin \psi, \varrho \cos \psi) = 0$ are approximated by solutions of $\varrho \sin \psi \log(\varrho \cos \psi) + \frac{e_2}{30} = 0$.

At fixed $\varrho \in (0, 1)$ the map $\psi \mapsto \varrho \sin \psi \log(\varrho \cos \psi)$ is monotonously decreasing on $(-\frac{\pi}{2}, \frac{\pi}{2})$ with $\varrho \sin \psi \log(\varrho \cos \psi) \rightarrow \mp\infty$ as $\psi \rightarrow \pm\frac{\pi}{2}$ and a zero at $\psi = 0$. Since now

$$\varrho \sin \psi \log(\varrho \cos \psi) + \frac{e_2}{30} \rightarrow \frac{e_2}{30} \quad \text{as } \varrho \rightarrow 0$$

pointwise in $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and uniformly in $\psi \in (-\frac{\pi}{2} + \delta, \frac{\pi}{2} - \delta)$ for any $\delta \in (0, \frac{\pi}{2})$, we conclude that at sufficiently small ϱ we have precisely one solution ψ , and if $e_2 > 0$ we have $\psi \rightarrow \frac{\pi}{2}$, whereas if $e_2 < 0$ we have $\psi \rightarrow -\frac{\pi}{2}$. If $e_2 = 0$ we have $\psi \rightarrow 0$.

Back in the coordinates (x, h) , for $e_2 > 0$ we obtain a solution curve as stated in the lemma, part (i), where $\psi \rightarrow \frac{\pi}{2}$ implies that $h \rightarrow 0$ faster than linear. Analogously we obtain the assertion (ii). Note that the graph of $h \mapsto (X_{0,-})(h), h$ is parameterized in the (ϱ, ψ) -coordinates by a curve with $\psi \rightarrow 0$ and $\psi < 0$ as $\varepsilon \rightarrow 0$. In particular, the set where $F(x, h) = 0$ is not soluble for h is bounded away from our solution curves, thus they are indeed representable as the graph of a function of x .

Finally, for the assertion (iii) note that we have $\psi \rightarrow 0$ from above as $\varrho \rightarrow 0$, in particular, the graph of $h \mapsto (X_{0,-})(h), h$ is also bounded away from the solution curve for $e_2 = 0$. Moreover, since the map $\psi \mapsto \varrho \sin \psi \log(\varrho \cos \psi)$ has a linear zero in $\psi = 0$ for sufficiently small ϱ , that is, we have

$$\varrho \sin \psi \log(\varrho \cos \psi) \in \psi \varrho \log \psi + \mathcal{O}(\psi^3),$$

the limit $\psi \rightarrow 0$ as $\varrho \rightarrow 0$ is also achieved at least linearly. Transformation back into the coordinates (x, h) adds another power and we obtain assertion (iii). \square

Next we study the situation at large h .

Lemma 5.20. *For any choice of $e_1, e_2 \in \mathbb{R}$ the straight lines defined by $x \in \{x_{(\pm)}\}$ are asymptotic solution curves. Moreover, the solution curves approximated by these asymptotic solution curves are representable as graphs of analytic functions*

$$(x_{(\pm)}, x_{(\pm)} + \varepsilon) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto (x, h(x))$$

for some $\varepsilon > 0$.

Proof. We evaluate the function F at $x_{(\pm)}$ and at $x_{(\pm)} + 2\delta$ for some $\delta \in (0, 1/\sqrt{30})$ and obtain

$$F(x_{(\pm)}, h) = \mp \frac{1}{\sqrt{30}} - \frac{e_2}{30} - \frac{e_1}{30} \frac{1}{h^2} \mp \frac{1}{\sqrt{30}} (2 \log(h) + f(x_{(\pm)})) \quad (49)$$

as well as

$$\begin{aligned}
 & F(x_{(\pm)} + 2\delta, h) \tag{50} \\
 &= \left(\delta \pm \sqrt{\frac{2}{15}} \right) \delta h^2 \mp \frac{1}{\sqrt{30}} - \delta - \frac{e_2}{30} - \frac{e_1}{30} \frac{1}{h^2} - \left(\delta \pm \frac{1}{\sqrt{30}} \right) (2 \log(h) + f(x_{(\pm)} + 2\delta)),
 \end{aligned}$$

respectively.

Note that, for our given bound on δ both the prefactors $\delta \pm \sqrt{2/15}$ and $\delta \pm \frac{1}{\sqrt{30}}$ of the divergent terms in the limit $h \rightarrow \infty$ in (50) do not change sign in $(0, 1/\sqrt{30}) \ni \delta$.

From (49) and (50) we can read off that

$$F(x_{(\pm)}, h) \rightarrow \mp \infty \quad \text{as well as} \quad F(x_{(\pm)} + 2\delta, h) \rightarrow \pm \infty$$

as $h \rightarrow \infty$ and, consequently, there exists $h_0 \in (0, \infty)$ such that for all $h > h_0$ there exists a solution of $F(x, h) = 0$ in both the intervals

$$(x_{(-)}, x_{(-)} + 2\delta) \quad \text{and} \quad (x_{(+)}, x_{(+)} + 2\delta). \tag{51}$$

By possibly enlarging h_0 we find that there is a third solution in some interval $(0, \tilde{\delta})$ provided by Lemma 5.14 and with Lemma 4.11 it is indeed *the* third solution. Hence, each of the intervals (51) contains precisely one solution.

Since the above argument works for any (sufficiently small) $\delta > 0$, and since the solutions we have found are in the open intervals in (51) (in particular they do not coincide with these interval's lower bounds), we can already conclude that the solutions “come closer to $x_{(\pm)}$ ” as $h \rightarrow \infty$, more precisely, we have the asymptotic solution curves as stated in the lemma.

Finally, evaluating $F(X_{(e_1, \pm)}(h), h)$ (with $X_{(e_1, \pm)}(h)$ from Lemma 4.17) we find the dominant term in the limit $h \rightarrow \infty$ to be the log-term in F_2 . In particular, we find that also $F(X_{(e_1, \pm)}(h), h) \rightarrow \mp \infty$ as $h \rightarrow \infty$, and we can replace (51) by the refined intervals

$$(X_{(e_1, -)}(h), x_{(-)} + 2\delta) \quad \text{and} \quad (X_{(e_1, +)}(h), x_{(+)} + 2\delta)$$

at sufficiently large h . Again the solutions do not coincide with these refined interval's lower bounds. Hence, in every solution point found above we can solve the solution set to yield the graph of an analytic function of x on some interval $(x_{(\pm)}, x_{(\pm)} + \varepsilon)$. \square

Remark 5.21. Recall that the values $x_{(\pm)}$ correspond, in the limit of large h , to the values $\xi_{(\pm)}$ which we have distinguished in the massless case, where we also have a divergent ($H \rightarrow \infty$) solution branch around these values.

5.4. Minimal and conformal coupling

In this section we want to study the physically distinguished ξ -values of minimal coupling $\xi = 0$ and conformal coupling $\xi = \frac{1}{6}$. This allows us to exclude solutions or to specify a

number of solutions along the curves defined by these ξ -values for certain choices of e_1 and e_2 . For this analysis we use similar arguments as in Section 4.1.

Minimal and conformal coupling are realized in (x, h) coordinates by the curves

$$(0, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad x \mapsto (x, h_{\text{mc}}(x)) := \left(x, \frac{1}{\sqrt{x}}\right) \quad (52)$$

and

$$(2, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad y \mapsto (y, h_{\text{cc}}(y)) := \left(y, \frac{1}{\sqrt{y-2}}\right), \quad (53)$$

respectively.

Lemma 5.22. (i) For $e_1 < -\frac{15}{2}$ the consistency equation (24) has precisely one solution with minimal coupling $\xi = 0$ for any value of e_2 .

(ii) For $e_1 = -\frac{15}{2}$ the consistency equation has precisely one solution with minimal coupling $\xi = 0$ for $e_2 < 20$ and no solution with minimal coupling for $e_2 \geq 20$.

(iii) For $e_1 > -\frac{15}{2}$ there exists $\widehat{e}_2 \in \mathbb{R}$ such that the consistency equation has

- precisely two solutions along minimal coupling if $e_2 < \widehat{e}_2$,
- precisely one solution along minimal coupling if $e_2 = \widehat{e}_2$ and
- no solution along minimal coupling if $e_2 > \widehat{e}_2$.

Numerical evidence. Evaluating the function F from (24) along the curve of minimal coupling (52) results in

$$F(x, h_{\text{mc}}(x)) = -\left(\frac{1}{4} + \frac{e_1}{30}\right)x - \left(\frac{e_2}{30} + \frac{x}{2}(f(x) - \log(x))\right) + \frac{29}{30x} + f(x) - \log(x) \quad (54)$$

for $x > 0$, where we grouped the terms according to their relevance in the different regimes $x \rightarrow 0$, finite x or $x \rightarrow \infty$.

For the second derivative of (54) we compute that

$$\frac{\partial^2}{\partial x^2} F(x, h_{\text{mc}}(x)) \in -\frac{61}{15x^3} + \mathcal{O}\left(\frac{1}{x^2}\right) \quad \text{as } x \rightarrow 0$$

as well as

$$\frac{\partial^2}{\partial x^2} F(x, h_{\text{mc}}(x)) \in -\frac{74}{21x^4} + \mathcal{O}\left(\frac{1}{x^5}\right) \quad \text{as } x \rightarrow \infty,$$

for which we have employed the Puiseux expansion of $f - \log$ from Appendix A.

Consequently, at sufficiently small and at sufficiently large x the latter second derivative is negative. That it is also negative in the intermediate regime is numerically justified in Figure 6.(a), where we have plotted said second derivative in log-log-scaling together with its asymptotics we have just computed.

Lemma 5.23. (i) If $e_1 < -\frac{15}{2}$ the consistency equation (24) has at most three solutions with conformal coupling $\xi = \frac{1}{6}$.

(ii) Let $e_1 = -\frac{15}{2}$. If $e_2 \leq -10$ the consistency equation has exactly one solution with conformal coupling. Moreover, there exists $\widehat{e}_2 > -10$ such that the consistency equation has

- precisely two solutions with conformal coupling if $e_2 \in (-10, \widehat{e}_2)$,
- precisely one solution with conformal coupling if $e_2 = \widehat{e}_2$ and
- no solution with conformal coupling if $e_2 > \widehat{e}_2$.

(iii) If $e_1 > -\frac{15}{2}$ there exists $\widehat{e}_2 \in \mathbb{R}$ such that the consistency equation has

- precisely two solutions with conformal coupling if $e_2 < \widehat{e}_2$,
- precisely one solution with conformal coupling if $e_2 = \widehat{e}_2$ and
- no solution with conformal coupling if $e_2 > \widehat{e}_2$.

Numerical evidence. We proceed similar to Lemma 5.22, although we need to adjust the arguments at some points.

For $x > 2$ and $y = x - 2$ we obtain

$$\begin{aligned} F(x, h_{\text{cc}}(x)) &= F\left(y + 2, \frac{1}{\sqrt{y}}\right) \\ &= -\left(\frac{1}{4} + \frac{e_1}{30}\right)y - \frac{e_2}{30} - \frac{y}{2}(f(y+2) - \log(y)) - \frac{1}{30y} \end{aligned} \quad (55)$$

for the values of F along the curve of conformal coupling (53). Note that, employing some Puiseux series arithmetics⁴ one can compute that

$$f(y+2) \in \log(y) + \frac{2}{3y} - \frac{1}{15y^2} - \frac{8}{315y^3} + \mathcal{O}\left(\frac{1}{y^4}\right)$$

at large y and thus obtain the asymptotic expansions

$$F\left(y + 2, \frac{1}{\sqrt{y}}\right) \in -\frac{1}{30y} + \mathcal{O}(1) \quad \text{as } y \rightarrow 0$$

and

$$F\left(y + 2, \frac{1}{\sqrt{y}}\right) \in -\left(\frac{1}{4} + \frac{e_1}{30}\right)y - \left(\frac{e_2}{30} + \frac{1}{3}\right) + \mathcal{O}\left(\frac{1}{y}\right) \quad \text{as } y \rightarrow \infty.$$

On the other hand, for the second derivative (w.r.t. y) of (55) we obtain the asymptotic expansions

$$\frac{\partial^2}{\partial y^2} F\left(y + 2, \frac{1}{\sqrt{y}}\right) \in -\frac{1}{15y^3} + \mathcal{O}\left(\frac{1}{y^2}\right) \quad \text{as } x \rightarrow 0$$

⁴After we expanded $f(y+2)$ we, moreover, need to expand $\frac{1}{(y+2)^k}$ as well as $\log\left(\frac{y}{y+2}\right)$ in terms of $\frac{1}{x^l}$.

as well as

$$\frac{\partial^2}{\partial y^2} F\left(y + 2, \frac{1}{\sqrt{y}}\right) \in \frac{24}{315y^4} + \mathcal{O}\left(\frac{1}{y^5}\right) \quad \text{as } x \rightarrow \infty.$$

We read off that, above some bound on y -values the map $y \mapsto \frac{\partial^2}{\partial y^2} F\left(y + 2, \frac{1}{\sqrt{y}}\right)$ is positive and below some bound it is negative. By numeric evaluation as shown in Figure 6.(b) we conclude that in the intermediate regime this second derivative has precisely one zero (indicated by the dip in the log-log plot) and hence, the map $y \mapsto F\left(y + 2, \frac{1}{\sqrt{y}}\right)$ has precisely one inflection point. In particular, it has independently of any parameters, at most three zeros. This completes part (i) of the lemma.

For $e_1 < -\frac{15}{2}$ we omit to improve the upper bound of three on the number of solutions at conformal coupling.

If $e_1 = -\frac{15}{2}$, we read off from the asymptotic expansions that $F\left(y + 2, \frac{1}{\sqrt{y}}\right) \rightarrow -\infty$ as $y \rightarrow 0$ and $F\left(y + 2, \frac{1}{\sqrt{y}}\right) \rightarrow -\frac{e_2}{30} - \frac{1}{3}$ as $y \rightarrow \infty$. Consequently, if $e_2 \leq -10$ the latter large- x -limit is non-negative and having only one inflection point the map $y \mapsto F\left(y + 2, \frac{1}{\sqrt{y}}\right)$ must have precisely one zero. Therefore note that the second derivative is indeed positive at large y .

On the other hand, from the limiting behavior we can read off that the map $y \mapsto F\left(y + 2, \frac{1}{\sqrt{y}}\right)$ has a global maximum, in which (as in Lemma 5.22) the maximum value depends linearly on e_2 . Hence there exists \hat{e}_2 such that this maximum value is zero, where in the context of the aforementioned argument necessarily $\hat{e}_2 > -10$ holds. The remaining claims of part (ii) in the lemma follow immediately.

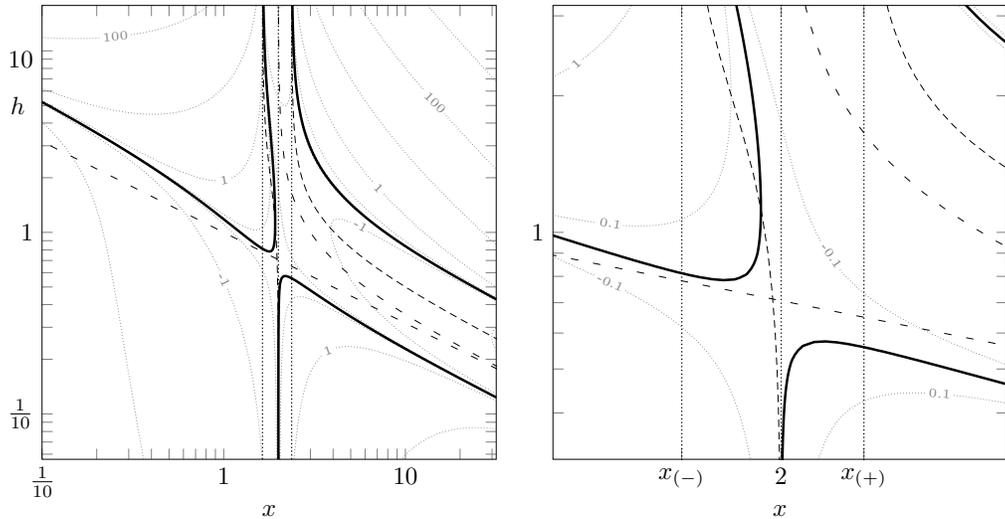
Finally, for $e_1 > -\frac{15}{2}$ we have $F\left(y + 2, \frac{1}{\sqrt{y}}\right) \rightarrow -\infty$ in both limits $y \rightarrow 0$ and $y \rightarrow \infty$, consequently the function $y \mapsto F\left(y + 2, \frac{1}{\sqrt{y}}\right)$ has a global maximum in which its value, again, linearly depends on e_2 . By the same arguments as above and in Lemma 5.22 we obtain a $\hat{e}_2 \in \mathbb{R}$ fulfilling (iii) of the lemma. \diamond

6. Numerical treatment of the solution set

In the previous sections we discussed properties of the set \mathcal{S}_{e_1, e_2} of de Sitter solutions of the energy equation in the massive case, in particular, the structure of \mathcal{S}_{e_1, e_2} and its asymptotic behavior. The non-polynomial terms of the energy equation prevented an explicit solution. In the present section have a look at the solution set by determining the zeros of F numerically. In particular, we study at the behavior of the solution set around distinguished parameter sets that were found in the above analysis. Hereby we give attention to the topological changes of the solution set.

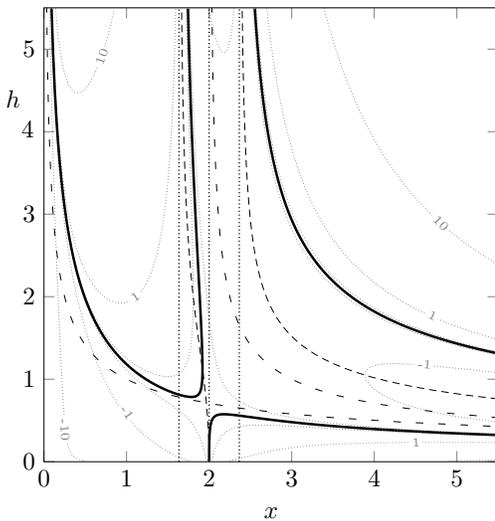
At first, in Figure 7, we look at the parameter pair $e_1 = e_2 = 0$. There, and throughout this section, we use the following graphical conventions:

- Thick solid lines mark the solution set \mathcal{S}_{e_1, e_2} of the consistency equation (24).
- Densely dashed lines mark the graphs of $X_{(e_1, \pm)}$, where the equation $F(x, h) = 0$ is



(a) log-log scaling

(b) log-log scaling on a smaller section



(c) normal scaling

Figure 7: The graphics show the solution sets of the energy equation for $e_1 = e_2 = 0$ in log-log and in normal scaling as the thick, black lines. The dotted black lines mark the values $x \in \{x_{(-)}, 2, x_{(+)}\}$, the densely dashed lines mark the points (x, h) where $x = X_{(0,\pm)}(h)$ and the loosely dashed black lines mark minimal and conformal coupling. The dotted gray lines finally mark some more level sets of F with the given values. In the following we mostly stick with the representation in log-log scaling as the straight outgoing lines show off the asymptotics of our solution curves. Note that in (a), although it seems slightly curved, the left branch of solutions also ends in a straight line with (logarithmic) slope $-\frac{1}{2}$.

not soluble for h .

- Loosely dashed lines mark the curves from (52) and (53) of minimal ($\xi = 0$) and conformal ($\xi = \frac{1}{6}$) coupling, respectively.
- Dotted (black) lines mark the distinguished values of $x \in \{x_{(-)}, 2, x_{(+)}\}$.

- If suitable, we include some additional level sets of the function F in gray dotted. As e_2 enters the function F as an offset, these are solution curves for some other value of e_2 .

Note that we display the solution sets in log-log scaling since by our results of Section 5 all solution curves are, in that scaling, asymptotically equivalent with straight lines, either vertical or with slope $\pm\frac{1}{2}$.

In Figure 7 we identify most of the analytic assertions we have made in the previous sections. These are, we have a solution curve running into $(x, h) = (2, 0)$ (cf. Section 5.3), three solution curves which are asymptotically equivalent with $x \mapsto \frac{\alpha}{\sqrt{x}}$, one of them at small x (cf. Section 5.2) and two at large x (cf. Section 5.1) and, finally, two solution curves approaching the asymptotes at $x = x_{(\pm)}$ for large h (cf. also Section 5.3). Moreover, we have exactly one point where the equation is not locally solvable for h , precisely where the curve of solutions crosses the curve defined by the graph of $X_{(0,-)}$.

We have shown in Section 5.4 that the function F , restricted to the curve of minimal coupling $\xi = 0$ is concave and divergent to $-\infty$ both at small and large x . However, as we can see in Figure 7, the maximum of the function F is indeed negative and, consequently, there is no solution for the minimally coupled model. Moreover, there is also no solution for the conformally coupled case. Thereby we conclude that the six open ends of the present case constitute three connected components, namely the connected component below the minimal coupling curve, the one above the conformal coupling curve and the one in between these two curves. Moreover, we can separate the connected components by the line $x = 2$, since for $e_1 = e_2 = 0$ we have $F(2, h) = -\frac{h^2}{30} < 0$ for all $h > 0$. Finally, we can also observe the precise amount of solutions that we have asserted in Table 1 on the lines at $x = x_{(-)}$ and $x = x_{(+)}$, namely exactly one solution for each x -value.

At next, we discuss the behavior of the solution set around the parameter point $e_1 = e_2 = 0$, particularly we look at the asymptotic solution curve running into $(x, h) = (2, 0)$. Recall that there exists such a curve only for $e_1 = 0$, where its precise shape is determined by the parameter e_2 (cf. Section 5.3). We have shown how for $e_2 = 0$ the solution curve running into $(x, h) = (2, 0)$ is tangent to the $x = 2$ -line in the limit, whereas the for $e_2 > 0$ and $e_2 < 0$ it is tangent to the $h = 0$ -axis, starting in positive or negative direction, respectively. This is what we observe in Figure 8.(f).

If we tune e_1 away from zero to values $e_1 < 0$, we observe that this curve, instead of running into $(x, h) = (2, 0)$, now runs into $(x, h) = (0, 0)$, and ends in the asymptotic solution curve at small x we expect for such e_1 values. If we choose, however, values $e_1 > 0$ we observe that now this branch adds to a third asymptotic solution curve at large x . This behavior is shown in Figures 8.(c)-(e).

Moreover, in Figure 8 we capture how a saddle of the function F can be tuned to coincide with a zero of F . While the value of F at its saddle is positive or negative in Figure 8.(a) or 8.(c), respectively, the value of e_1 in Figure 8.(b) was chosen such that we indeed observe crossing solution curves, reducing the number of connected components of our solution set by one. That in such a case the zero set of F can be locally decomposed into two analytic curves was shown in Section 4.3

Furthermore, we want to look at the parameters nearby $e_1 = \beta_0$ and $e_2 = -10$. We

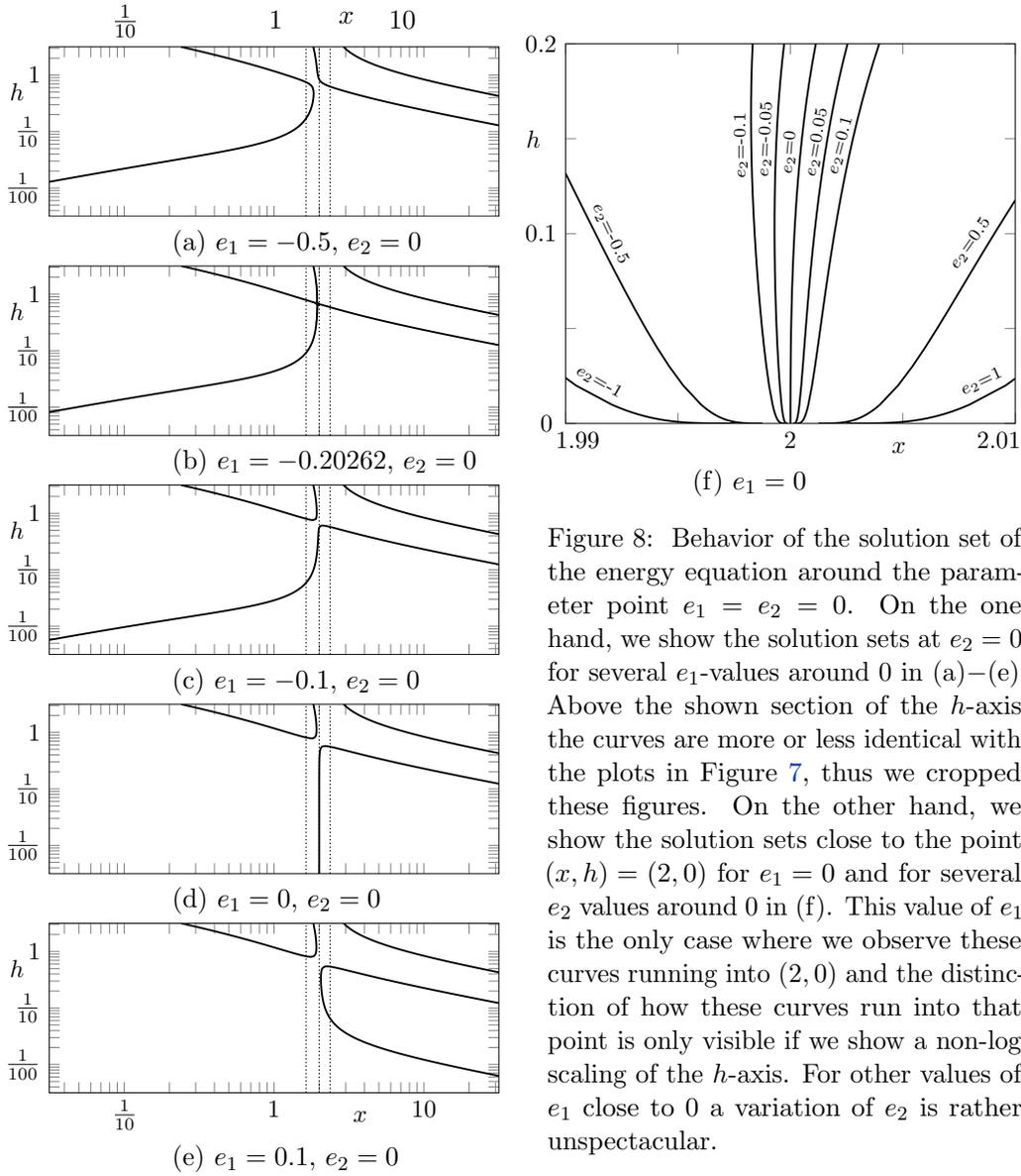


Figure 8: Behavior of the solution set of the energy equation around the parameter point $e_1 = e_2 = 0$. On the one hand, we show the solution sets at $e_2 = 0$ for several e_1 -values around 0 in (a)–(e). Above the shown section of the h -axis the curves are more or less identical with the plots in Figure 7, thus we cropped these figures. On the other hand, we show the solution sets close to the point $(x, h) = (2, 0)$ for $e_1 = 0$ and for several e_2 values around 0 in (f). This value of e_1 is the only case where we observe these curves running into $(2, 0)$ and the distinction of how these curves run into that point is only visible if we show a non-log scaling of the h -axis. For other values of e_1 close to 0 a variation of e_2 is rather unspectacular.

have studied this setting in Section 5.1. The solution sets for certain choices of parameters are shown in Figure 9. At the value $e_1 = \beta_0$ the solution set transitions from having three asymptotic solution curves to having only one. The lower two of these three curves for $e_1 < \beta_0$ are connected to each other in Figure 9.(a).(i) and are connected to the two curves which are left at large h without crossing in 9.(a).(ii). In between these values there must, consequently, be an e_2 -value at which they are connected to the same two curves, but now interchanged, and we can observe them to cross in the arising saddle of

Hence we can, similarly as in Figure 9, observe that the two asymptotic solution curves come closer to each other as $e_1 \rightarrow \beta_{-1}$ from above, have the same asymptotic (first order) approximation at $e_1 = \beta_{-1}$ and form a sling (now around the curve of $X_{(e_1,+)}$) which is more and more pulled to smaller values of x as e_1 further decreases.

7. Parameter choices for potential inflationary models

So far we mostly considered the consistency equation in the massive case as an equation of x instead of the physical parameter ξ in order to simplify the analysis of the function f . However, we can identify the curves of constant ξ as parameterized by

$$(0, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad h \mapsto (12\xi + \frac{1}{h^2}, h), \quad (56)$$

where the curves of minimal coupling (52) and conformal coupling (53) are special cases. If $\xi < 0$ these curves leave the domain $(0, \infty) \times (0, \infty)$ above a certain h -value, since the Bunch-Davies state exists for negative ξ only if h is sufficiently small.

In the present section we want to identify parameter settings of potential inflationary models. By a potential inflationary model we mean the following:

According to ‘standard’ (i.e. Λ CDM) cosmology our universe will, eventually at late times, expand in a Dark Energy-dominated exponential manner. In other words, it will be well approximated by a cosmological de Sitter solution, say with Hubble rate H_{DE} . On the other hand, assuming that the universe went through an inflationary early phase solves various problems of Λ CDM physics, most importantly, the so-called cosmic horizon problem. Also an inflationary phase is modelled by an (approximately) exponential expansion, with a much larger Hubble rate. Denote it by H_{I} ($\gg H_{\text{DE}}$).

The aim of the present section is to identify parameter settings (e_1, e_2, ξ) in which there exist multiple solutions of our model, say two values h_1 and h_2 with $h_1 \gg h_2$, which restore the ratio $\frac{h_1}{h_2} = \frac{H_{\text{I}}}{H_{\text{DE}}}$ of some given physical data H_{I} and H_{DE} . More advanced, we will show that by parameter tuning one is more or less free to produce an arbitrary ratio $\frac{h_1}{h_2}$ with a (more or less) arbitrary smaller solution h_2 . To establish a realistic magnitude for the rates h_1 and h_2 , we note that for today’s Hubble rate and for the Higgs mass (i.e. the only mass of a scalar field occurring in the standard model of particle physics) one can compute

$$h_2 = \frac{H_{\text{today}}}{M_{\text{Higgs}}} \approx 10^{-44}$$

in our unit system (i.e. where $\hbar = c = 1$). On the other hand, supposing that the inflationary phase lasted about 10^{-34} s and caused an expansion by a factor of 10^{26} we can compute $H_{\text{I}} \approx 10^{34} \frac{1}{\text{s}}$ and thus a magnitude of

$$h_1 = \frac{H_{\text{I}}}{M_{\text{Higgs}}} \approx 10^7.$$

Suppose that in a subsequent step, for a parameter setting as above, one is able to show how the dynamical system (13) is unstable towards perturbations around the de

Sitter solution defined by the larger value h_1 and stable around the solution defined by the smaller h_2 . Moreover, suppose that one is able to show that the dynamical system (13) (with the aforementioned stability properties) indeed possesses a solution starting close to the unstable de Sitter solution with rate h_1 and, after some intermediate phase, eventually approaches the stable solution with rate h_2 . In that case one would provide a physical model for our universe in which the driving forces for both the inflationary and the late-time expansion are modelled by the presence of a quantum field. In particular, such a model omits introducing the new and unknown Dark Energy, that is, a form of energy which evades any observation other than enforcing an exponential late time expansion. In such an approach, the present section makes the first step of providing suitable parameter settings.

Note that similar to our hypothetical outline above, such effects have indeed been observed among solutions of the semiclassical Einstein equation. At first we refer to M. Hänsel [27] who found a parameter regime for the SCE of a massless, conformally coupled scalar field (with the conformal vacuum) in which precisely the stability properties of two (different) de Sitter solutions as described above are present. Moreover, the corresponding phase diagram (Figure 5.6.(a) of [27]) shows how solutions that start close to the larger de Sitter point approach the smaller de Sitter point under certain conditions. In another article [10] the authors found two different de Sitter solutions of which the one with a larger rate is unstable and the one with a smaller rate is stable. They also work with a conformally coupled scalar field, using approximate KMS states. As a third reference, the authors of the present article found in [21] a similar scenario in which a de Sitter solution is present and appears attractive towards perturbations. In that latter article, solutions are constructed using Minkowski-like states for a massless field. Note that while the analysis of this scenario is carried out explicitly only for $\xi = \frac{1}{6}$, it is also stated how this can be generalized for ξ close but not equal to $\frac{1}{6}$.

In the following we provide a few examples of parameter settings in which there exist multiple solutions. Moreover, we show how to tune parameters in order to control the h -values of these solutions (particularly their ratio) by exploiting the knowledge acquired in Section 5.

In Figure 10 we can see how the solution set behaves if $e_2 \rightarrow -\infty$. We started in (a) with a parameter point $e_1 = 2$ and $e_2 = -10$ close to the parameters studied in Figure 9, such that e_1 lies in the interval (β_1, β_0) where we have three asymptotic solution curves at small h , and therefrom lowered e_2 .

In Section 5.4 we have seen that F , restricted to the curve of $\xi = \xi_{cc} = \frac{1}{6}$, parameterized as in (56), diverges to $-\infty$ for both $h \rightarrow 0$ and $h \rightarrow \infty$ at our choice of e_1 . Moreover, by Lemma 5.23, it has two zeros if e_2 lies below a certain bound and no zero above that bound. We have not computed this bound, but apparently in Figure 10.(a) we are above this bound whereas in Figure 10.(b) we are below it. Since then e_2 enters F simply as a (sign-reversed) offset, it is clear that, if we lower e_2 , the two corresponding h -values of the two zeros of F drift more and more apart, and the smaller solution's h converges to 0, whereas the larger solution's h diverges to $+\infty$ in the limit $e_2 \rightarrow -\infty$.

What we have done above works not only for $\xi = \xi_{cc} = \frac{1}{6}$, but for any ξ with $|\xi - \frac{1}{6}| \leq 1/\sqrt{1080}$, that is, for all ξ such that the vertical asymptote of the corresponding

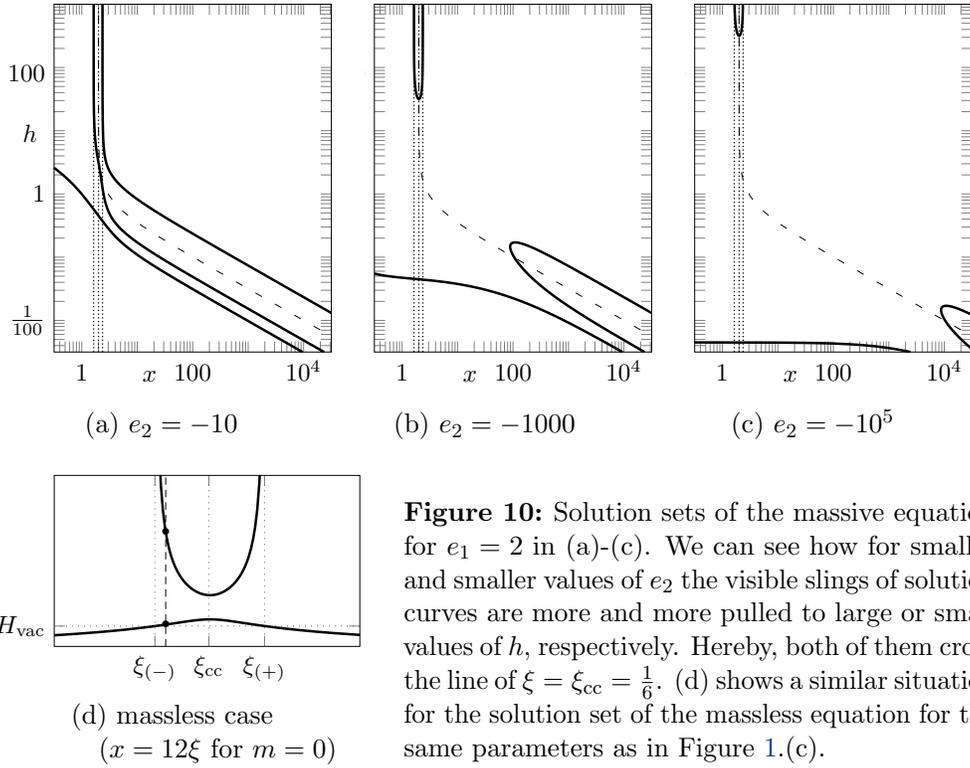


Figure 10: Solution sets of the massive equation for $e_1 = 2$ in (a)-(c). We can see how for smaller and smaller values of e_2 the visible slings of solution curves are more and more pulled to large or small values of h , respectively. Hereby, both of them cross the line of $\xi = \xi_{cc} = \frac{1}{6}$. (d) shows a similar situation for the solution set of the massless equation for the same parameters as in Figure 1.(c).

curve (56) lies in between the values $x_{(\pm)} = 2 \pm \sqrt{2/15}$.

To see this, we observe that all constant- ξ -curves are asymptotically equivalent to $x \mapsto 1/\sqrt{x}$ at large x and that the function s from Section 5.1 has one zero larger than 1 and two zeros smaller than 1 for our present value of e_1 . In particular, all these curves of constant ξ lie, asymptotically at large x , in between the upper two solution curves, and for sufficiently small e_2 intersect the solution set at small h . On the other hand, for the ξ -values specified above the solution set intersects all these curves at large h as well. By lowering e_2 the h -values of these two solutions then drift more and more apart.

We have a similar situation for this particular interval of ξ -values in the massless case. We have included a graphic for this situation in Figure 10.(d). If the parameter ratio $\frac{\Lambda}{K}$ tends to zero, the two branches of the solution set drift more and more apart. More precisely, while for a fixed value $\xi \in (\xi_{(-)}, \xi_{(+)})$ the larger solution diverges, the smaller solution remains close to $H_{vac} = \sqrt{\Lambda/3}$. Both these limiting behaviors can easily be read off from (19) computing $H_{(\pm)}/H_{vac}$.

As a next example we want to study solutions for ξ -values at and around minimal coupling $\xi_{mc} = 0$. Figure 11 shows the solution sets for the same parameters e_1 and e_2 as in Figure 10. In Section 5.4 we have shown that also for minimal coupling we have basically the same situation as above, namely that F along the minimal coupling curve diverges to $-\infty$ at both small and large h . Now already the value $e_2 = -10$ in Figure 11

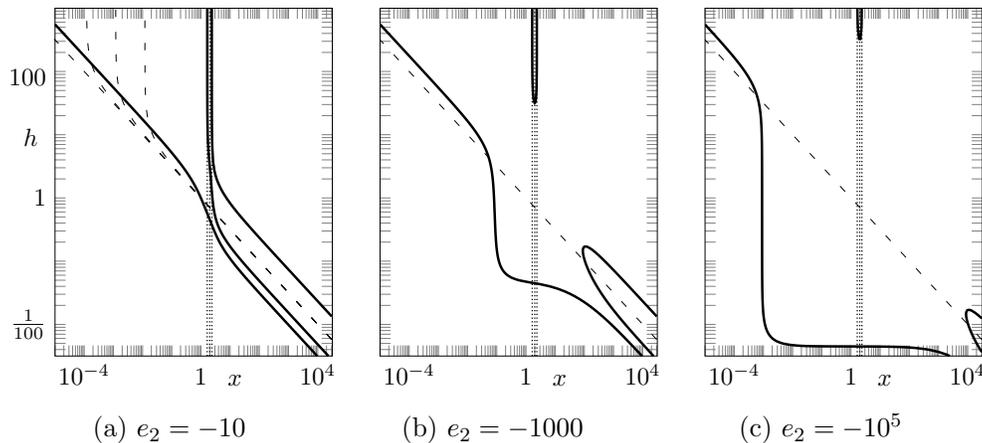


Figure 11: Solution sets of the massive equation for $e_1 = 2$ in (a)-(c), i.e. for the same parameters as in Figure 10, but on a larger interval of x -values. Moreover, in this figure we included the curves of constant $\xi = 0$, as well as, of constant $\xi \in \{10^{-3}, 10^{-4}, 10^{-5}\}$ in Subfigure (a).

lies below the upper bound on e_2 -values from Lemma 5.22.(iii), consequently, in all three graphics of Figure 11 we have two solutions with minimal coupling. The h -values of these solutions drift apart as $e_2 \rightarrow -\infty$.

In this scenario, however, we can tune our parameter ξ to positive small values such that the h -values of two solutions described above stay approximately stationary, but a third solution with a large h -value comes into play. Therefore, Figure 11.(a) includes the constant- ξ -curves for some such positive values. While it is obvious from the figure, we can also conclude the existence of such a solution with $h \rightarrow \infty$ as $\xi \rightarrow 0$ from Section 5.2. Therefore we reparameterize the (small- x -) asymptotic solution curve from Lemma 5.14 into

$$(0, \infty) \rightarrow (0, \infty) \times (0, \infty), \quad h \mapsto \left(\sqrt{\frac{29}{90} \frac{1}{h^2}}, h \right)$$

and note that this curve lies above the line of $\xi = 0$ in our x - h -parameter plane. On the other hand, the constant- ξ -curve for any $\xi > 0$ has a vertical asymptote and, consequently, crosses the asymptotic solution curve above. For this crossing we have $x \rightarrow 0$ and $h \rightarrow \infty$ as $\xi \rightarrow 0$. Consequently, if ξ is small enough (such that this crossing is at sufficiently small x -values for a good approximation of the actual solution curve by the asymptotic solution curve), we observe a third solution with the behavior as claimed above. Note that the smaller two solutions remain more or less stationary, obviously their h -values are continuous in ξ around $\xi = 0$.

Finally, we want to show the solution sets to a family of parameters with the reversed behavior as in the previous case, that is, we prescribe a certain (approximate) larger h -value and tune parameters so that the smaller solution's h -value tends to 0.

Recalling the shape of the function s from Section 5.1, we find that $\alpha = 1$ being a zero

Besides our given examples one can find several other parameter regimes with a similar behavior.

To conclude, our model is sufficiently flexible to guarantee, by a suitable choice of parameters, that for any two values of h there exist de Sitter solutions with these h values as the corresponding de Sitter rates. In particular, these two values can be prescribed using an arbitrary mass for the field.

8. Conclusion and outlook

Significantly improving upon earlier results in [29] on semiclassical de Sitter solutions, which were limited to vacuum solutions with vanishing QSE tensor, we have given an extensive and complete analysis of the consistency equation (16) to obtain a complete picture of cosmological de Sitter / Bunch-Davies solutions to the SCE.

Our main result is the description of the solution set of cosmological de Sitter / Bunch-Davies solutions as a union of analytic curves in the parameter space of the coupling to the scalar curvature ξ and the expansion rate H as a function of renormalization parameters. Also, as a function of the renormalization parameters, we have obtained a description of the structure of solution set as the union of up to three analytic curves in the plane of (ξ, H) values or in the related (x, h) -plane. The techniques applied to generate this map range from elementary solutions of algebraic equations, over the asymptotic analysis of certain analytic functions, continuity arguments based on the implicit function theorem to the reduction of analytic varieties. Note that a large number of solutions do not require a positive cosmological constant $\Lambda > 0$ and even for particular cases with $\Lambda < 0$ there exist solutions to the SCE with a positive rate of expansion, due to the nature of the QSE tensor. This is particularly interesting in the massless case, where this effect cannot simply be blamed on a positive *renormalized* cosmological constant.

Based on these findings, in particular on the explicit asymptotic expansions of the solution curves, we have identified parameter settings which are compatible with multiple de Sitter solutions, both with very large and very small rates of expansion.

In such settings, studying the Lyapunov stability of the de Sitter solutions is a natural next step. While for the case of conformal coupling with massless fields the dissertation of [27] clarifies the situation to a certain extent, the situation in the general case seems to be largely open.

Note that the question of Lyapunov stability can be answered on several levels. For special cases with a decoupling of the state degrees of freedom from the SCE like in [21, 27] this can be answered by the standard analysis of a dynamical system in finite dimension. Whenever the state dynamics couples non-trivially to the SCE, stability can either be answered in a reduced, cosmological setting [22, 36] or in the setting of the full SCE. See [35] for some investigations on stability in the case of a toy model and [19, 25] for linearization techniques of the full SCE system.

It would be especially attractive to find unstable directions for de Sitter solutions with high expansion rates and stability for de Sitter solutions with low expansion rate in the situations described in Section 7. Whether this is achievable or not, at present remains

an open research question.

Finally, the inclusion of more general kinds of matter, modeled by fermionic fields or gauge fields, as well as the inclusion of positive or negative spatial curvature certainly is also of interest.

Acknowledgement. Robert L. Bryant is gratefully acknowledged for providing the argument in Appendix B. The authors thank Paolo Meda, Nicola Pinamonti and Hendrik Herrmann for interesting discussions.

A. Properties of the Bunch-Davies Digamma terms f

In this appendix we collect a few properties of the function

$$f : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \psi^{(0)}\left(\frac{3}{2} - \sqrt{\frac{9}{4} - x}\right) + \psi^{(0)}\left(\frac{3}{2} + \sqrt{\frac{9}{4} - x}\right) \quad (57)$$

as defined in Section 4, with the Digamma function $\psi^{(0)} = \Gamma'/\Gamma$.

At first, we want to be a bit more precise on its definition. Note that a priori the mapping in (57) defines a meromorphic function \tilde{f} on a slit plane, that is, on the complex plane \mathbb{C} from which a half-ray starting in $\frac{9}{4}$ was removed. However, one can show that the particular values of that function do not depend on the choice of which ray was removed, hence (57) defines a meromorphic function on $\mathbb{C} \setminus \{\frac{9}{4}\}$. In a final step, one can show that $\tilde{f}(\frac{9}{4}) := 2\psi^{(0)}$ defines a holomorphic continuation to the point in question and (57) defines a meromorphic function on \mathbb{C} whose poles lie at $\{-n^2 - 3n \mid n \in \mathbb{Z}_{\geq 0}\}$. In particular, by restricting $f = \tilde{f}|_{(0, \infty)}$ (57) defines a (real) analytic function on the positive real axis.

Studying the poles and residues of the Gamma function we find that \tilde{f} has a pole of order 1 in $0 \in \mathbb{C}$. Employing the chain rule for residues we find that $\text{Res}_0 \tilde{f} = -3$ and we conclude the asymptotic equivalence

$$f(x) \sim -\frac{3}{x} \quad \text{as } x \rightarrow 0.$$

On the other hand, we have asymptotically

$$f(x) \sim \log(x) \quad \text{as } x \rightarrow \infty$$

which immediately follows from Lemma 1 in the appendix of Juárez-Aubry's article [29]. More precisely, the latter lemma shows that

$$|f(x) - \log(x)| \leq \frac{3}{x}$$

for all $x > \frac{9}{4}$. The proof of the latter lemma can easily be extended to see that

$f(x) < \log(x)$ and thus

$$f(x) - \log(x) \in [-\frac{3}{x}, 0) \quad (58)$$

for all $x > \frac{9}{4}$.

Without proof we state the first few Puiseux coefficients of $f - \log$ in the limit $x \rightarrow \infty$ as

$$f(x) \in \log(x) - \frac{4}{3x} - \frac{11}{15x^2} - \frac{92}{315x^3} + \mathcal{O}(\frac{1}{x^4}).$$

At last, we can show that f is strictly increasing by estimating its derivative and, moreover, by comparing the values $f(2) = 1 - 2\gamma < 0$ and $f(\frac{9}{4}) = 4 + 2\psi^{(0)}(\frac{1}{2}) = 4 - 4\log(2) - 2\gamma > 0$ (with the Euler-Mascheroni number γ) we find that it must have its only zero in the interval $(2, \frac{9}{4})$ (numerically ≈ 2.1646).

As an orientation, Figure 3.(a) in the text shows a plot of f together with its asymptotics from above.

B. Lemma on the reduction of analytic varieties

The following result is used in Theorem 4.2 to prove that de Sitter solutions to the SCE form a union of analytic curves in the x - h -plane. The argument was provided by Robert L. Bryant from Duke University, Durham, North Carolina by private communication.

Lemma B.1. *Let $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ be analytic such that $F(0,0) = 0$ and let F_n be the lowest-order non-vanishing homogeneous term in F 's Taylor expansion around $(0,0)$, say F_n is a homogeneous polynomial of degree $n \in \mathbb{N}$. Suppose that all linear polynomials occurring in the factorization of F_n into irreducibles (over \mathbb{R}) are pairwise distinct, say these are $m \leq n$ in number. Then there exists a neighborhood $U \ni (0,0)$ such that*

$$U \cap \{ (x, y) \mid F(x, y) = 0 \} = \bigcup_{i=1}^m \gamma_i(J_i)$$

with regularly parameterized analytic curves $\gamma_i : J_i \rightarrow U$ (defined on some intervals $J_i \subset \mathbb{R}$), $i = 1, \dots, m$, which only intersect in $(0,0)$ and which are each linearly approximated around $(0,0)$ by the zero set of one of the linear factors of F_n .

Remark B.2. (i) *In other words, the analytic variety defined as the zero set of F can, locally around $(0,0)$, be decomposed into m non-singular analytic subvarieties.*

(ii) *The assumption of the lemma is clearly imposed by the (partially numeric) results of section 4.3. That is, an indefinite Hessian of F has two distinct eigenvectors, which allows an affine linear coordinate transformation such that*

$$F(x, y) \in xy + \mathcal{O}(\|(x, y)\|^3).$$

This representation of F , moreover, shows that there is an open neighborhood U of $(0,0)$ in the given coordinates such that $\nabla F \neq 0$ on $U \setminus \{(0,0)\}$, that is, the zero $(0,0)$ of ∇F is isolated.

(iii) Note that the claim of distinct linear factors is necessary, otherwise the analytic variety defined by $x^2 = y^3$ provides a counter example.

Proof. By the assumptions of the lemma, one finds a linear coordinate transformation after which F takes the form

$$F(x, y) = x \cdot F_{n-1}(x, y) + \sum_{k=n+1}^{\infty} F_k(x, y),$$

where each F_k is a homogeneous polynomial of order k , $k = n - 1$ or $k \geq n + 1$, and where x does not divide F_{n-1} in $\mathbb{R}[x, y]$. Moreover, this power series converges on a neighborhood of the origin.

By a blow-up substitution $x \mapsto xy$ we obtain

$$F(xy, y) = y^n \cdot \left(x \cdot G_{n-1}(x, y) + \sum_{k=n+1}^{\infty} y^{k-n} G_k(x, y) \right) =: y^n \cdot G(x, y),$$

where $G_k(x, y) := \frac{F_k(xy, y)}{y^n}$ are again polynomials of two variables, $k = n - 1$ or $k \geq n + 1$. The map $(x, y) \mapsto G(x, y)$ defines another analytic function around the origin with $G(0, 0) = 0$ and the assumption that x does not divide $F_{n-1}(x, y)$ implies that $\partial_x G(0, 0) \neq 0$. Consequently, the zero set of G is, locally around $(0, 0)$, given by the graph of an analytic curve of the form

$$J_1 \mapsto \mathbb{R}^2, \quad y \mapsto (g(y), y),$$

$J_1 \subset \mathbb{R}$ and $g : J_1 \rightarrow \mathbb{R}$ analytic. Moreover, inverting the blow-up substitution yields that F vanishes on the graph of the analytic curve

$$\gamma_1 : J_1 \mapsto \mathbb{R}^2, \quad y \mapsto (y \cdot g(y), y), \tag{59}$$

hence the analytic function

$$L_1 : (x, y) \mapsto x - y \cdot g(y)$$

is a prime factor of F in the ring

$$\mathcal{R} := \{ H \in \mathbb{R}[[x, y]] \mid H \text{ converges on some open neighborhood of } (0, 0) \}$$

of formal power series which converge around $(0, 0)$. Note that $G(0, 0) = 0$ implies $g(0) = 0$ and thus γ is, up to order $\mathcal{O}(y^2)$, approximated by the zero set of $(x, y) \mapsto x$, that is, of the factor of F 's lowest-order (non-vanishing) homogeneous term (LOHT) in consideration. Moreover, note that by (59) we have $|\gamma'| \geq 1$, i.e. $|\gamma'|$ is bounded away from zero.

However, since L_1 is a factor of F we can decomposed F into a product

$$F(x, y) = L_1(x, y) \cdot H_1(x, y),$$

for some $H_1 \in \mathcal{R}$. By this product representation, multiplying the LOHTs of L_1 and H_1 must result in the LOHT of F (in $\mathbb{R}[x, y]$). Since the LOHT of L_1 is just x , the LOHT of H_1 consequently equals $F_{n-1}(x, y)$.

Finally, one can linearly transform the coordinates x and y to single out one of the remaining linear factors of $F_{n-1}(x, y)$ as x again and repeat the above factorization procedure. Note that in the step specifying G 's zero set around $(0, 0)$ we particularly restricted our considerations to the stripe defined by $y \in J_1$ (and even just an open subset of this stripe around $(0, 0)$), and this new domain for the subsequent factorization step is linearly transformed as well.

Eventually, after m steps we end up with a factorization

$$F(x, y) = H(x, y) \cdot \prod_{j=1}^m L_j(x, y) \quad (60)$$

($H = H_m$), where each L_i vanishes on an analytic curve $\gamma_i : J_i \rightarrow \mathbb{R}^2$. Hereby, we can find a sufficiently small open neighborhood U of $(0, 0)$ such that each of the power series in (60) converges on U and such that for each $i \in \{1, \dots, m\}$ and each $(x, y) \in U$

$$L_i(x, y) = 0 \quad \text{if and only if} \quad (x, y) \in \gamma_i(J_i).$$

Again, multiplying the LOHTs of each factor on the RHS of (60) must result in the LOHT of F , called F_n in the lemma. By construction of the L_i , each of them has a LOHT equal to the respective linear prime factor of F_n . Consequently, the LOHT of H is the product of the remaining non-linear prime factors of F_n . Hence, either F_n has no second-order prime factors, then H is a unit in $\mathbb{R}[[x, y]]$ and as such is non-zero on a neighborhood of $(0, 0)$, w.l.o.g. on U , or F_n has second-order prime factors, then the LOHT of H is precisely the product of these. In the latter case, H is non-zero on a $(0, 0)$ -pointed neighborhood of $(0, 0)$, w.l.o.g. on $U \setminus \{(0, 0)\}$, where we used that each second-order prime factor is non-zero on, w.l.o.g., $U \setminus \{(0, 0)\}$. In any case, H is non-zero on $U \setminus \{(0, 0)\}$.

Concluding, by the factorization representation (60) of F and the aforementioned properties of the factors, the zeros of F in the open neighborhood U are precisely the images of the analytic curves $\gamma_1, \dots, \gamma_m$. \square

References

- [1] Allen, B.: Vacuum states in de Sitter space. *Phys. Rev. D* **32**, 3136–3149 (1985)
- [2] Anderson, P. R.: Effects of quantum fields on singularities and particle horizons in the early universe. *Phys. Rev. D* **28**, 271–285 (1983)
- [3] Anderson, P. R.: Effects of quantum fields on singularities and particle horizons in the early universe. II. *Phys. Rev. D* **29**, 615–627 (1984)

- [4] Anderson, P. R.: Effects of quantum fields on singularities and particle horizons in the early universe. III. The conformally coupled massive scalar field. *Phys. Rev. D* **32**, 1302–1315 (1985)
- [5] Anderson, P. R.: Effects of quantum fields on singularities and particle horizons in the early universe. IV. Initially empty universes. *Phys. Rev. D* **33**, 1567–1575 (1986)
- [6] Birrell, N. D., Davies, P. C. W.: *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge, UK: Cambridge University Press, (1984)
- [7] Bunch, T. S., Davies, P. C. W.: Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting. *Proc. Roy. Soc. Lond. A* **360**, 117–134 (1978)
- [8] Christensen, S. M.: Vacuum expectation value of the stress tensor in an arbitrary curved background: The covariant point-separation method. *Phys. Rev. D* **14**, 2490–2501 (1976)
- [9] Christensen, S. M.: Regularization, renormalization, and covariant geodesic point separation. *Phys. Rev. D* **17**, 946–963 (1978)
- [10] Dappiaggi, C., Fredenhagen, K., Pinamonti, N.: Stable cosmological models driven by a free quantum scalar field. *Phys. Rev. D* **77**, 104015 (2008). [arXiv:0801.2850 \[gr-qc\]](#)
- [11] Dappiaggi, C., Hack, T.-P., Möller, J., Pinamonti, N.: Dark Energy from Quantum Matter. *arXiv Preprint* (2010). [arXiv:1007.5009 \[astro-ph.CO\]](#)
- [12] Davies, P. C. W., Fulling, S. A., Christensen, S. M., Bunch, T. S.: Energy Momentum Tensor of a Massless Scalar Quantum Field in a Robertson-Walker Universe. *Ann. Phys. (N. Y.)* **109**, 108–142 (1977)
- [13] Eltzner, B., Gottschalk, H.: Dynamical Backreaction in Robertson-Walker Spacetime. *Rev. Math. Phys.* **23**, 531–551 (2011). [arXiv:1003.3630 \[math-ph\]](#)
- [14] Fewster, C. J., Kontou, E.-A.: Quantum strong energy inequalities. *Phys. Rev. D* **99**, 045001 (2019). [arXiv:1809.05047 \[gr-qc\]](#)
- [15] Fewster, C. J., Smith, C. J.: Absolute quantum energy inequalities in curved spacetime. *Ann. Henri Poincaré* **9**, 425–455 (2008). [arXiv:gr-qc/0702056](#)
- [16] Fewster, C. J., Verch, R.: A quantum weak energy inequality for Dirac fields in curved space-time. *Commun. Math. Phys.* **225**, 331–359 (2002). [arXiv:math-ph/0105027](#)
- [17] Flanagan, E. E., Wald, R. M.: Does back reaction enforce the averaged null energy condition in semiclassical gravity? *Phys. Rev. D* **54**, 6233–6283 (1996). [arXiv:gr-qc/9602052](#)

- [18] Fredenhagen, K., Hack, T.-P.: Quantum field theory on curved spacetime and the standard cosmological model. *Lect. Notes Phys.* **899**, 113–129 (2013). [arXiv:1308.6773 \[math-ph\]](#)
- [19] Fröb, M. B., Hack, T.-P., Higuchi, A.: Compactly supported linearised observables in single-field inflation. *J. Cosmol. Astropart. Phys.* **2017**, 043 (2017). [arXiv:1703.01158 \[gr-qc\]](#)
- [20] Fulling, S. A.: *Aspects of Quantum Field Theory in Curved Spacetime*. Cambridge University Press, (1989)
- [21] Gottschalk, H., Rothe, N., Siemssen, D.: Special cosmological models derived from the semiclassical Einstein equation on flat FLRW space-times. *Class. Quantum Gravity* **39**, 125004 (2022). [arXiv:2112.15050 \[gr-qc\]](#)
- [22] Gottschalk, H., Siemssen, D.: The Cosmological Semiclassical Einstein Equation as an Infinite-Dimensional Dynamical System. *Ann. Henri Poincaré* **22**, 3915–3964 (2021). [arXiv:1809.03812 \[math-ph\]](#)
- [23] Guth, A. H.: Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D* **23**, 347–356 (1981)
- [24] Hack, T.-P.: On the Backreaction of Scalar and Spinor Quantum Fields in Curved Spacetimes. Ph.D. thesis, Hamburg University, Institut für Theoretische Physik II, (2010). <http://www-library.desy.de/cgi-bin/showprep.pl?thesis10-042>. [arXiv:1008.1776 \[gr-qc\]](#)
- [25] Hack, T.-P.: Quantization of the linearized Einstein-Klein-Gordon system on arbitrary backgrounds and the special case of perturbations in inflation. *Class. Quantum Gravity* **31**, 215004 (2014). [arXiv:1403.3957 \[gr-qc\]](#)
- [26] Hack, T.-P.: *Cosmological Applications of Algebraic Quantum Field Theory in Curved Spacetimes*. SpringerBriefs in Mathematical Physics. Springer International Publishing, (2015). [arXiv:1506.01869 \[gr-qc\]](#)
- [27] Hänsel, M.: Qualitative Analysis of Solutions to the Semiclassical Einstein Equation in homogeneous and isotropic Spacetimes. Ph.D. thesis, Universität Leipzig, (2019). <https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-325983>
- [28] Hollands, S., Wald, R. M.: Conservation of the stress tensor in interacting quantum field theory in curved spacetimes. *Rev. Math. Phys.* **17**, 227–312 (2005). [arXiv:gr-qc/0404074](#)
- [29] Juárez-Aubry, B. A.: Semi-classical gravity in de Sitter spacetime and the cosmological constant. *Phys. Lett. B* **797**, 134912 (2019). [arXiv:1903.03924 \[gr-qc\]](#)
- [30] Juárez-Aubry, B. A.: Semiclassical Gravity in Static Spacetimes as a Constrained Initial Value Problem. *Ann. Henri Poincaré* **23**, 1451–1487 (2022). [arXiv:2011.05947 \[gr-qc\]](#)

- [31] Juárez-Aubry, B. A., Kay, B. S., Miramontes, T., Sudarsky, D.: On the initial value problem for semiclassical gravity without and with quantum state collapses. arXiv Preprint (2022). [arXiv:2205.11671 \[gr-qc\]](#)
- [32] Kontou, E.-A., Sanders, K.: Energy conditions in general relativity and quantum field theory. *Class. Quantum Gravity* **37**, 193001 (2020). [arXiv:2003.01815 \[gr-qc\]](#)
- [33] Liddle, A. R.: *An Introduction to Cosmological Inflation*. In: A. Masiero, G. Senjanović, A. Smirnov (eds.) *ICTP Summer School in High-Energy Physics and Cosmology*. World Scientific, (1999). [arXiv:astro-ph/9901124](#)
- [34] Linde, A. D.: A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Phys. Lett. B* **108**, 389–393 (1982)
- [35] Meda, P., Pinamonti, N.: Linear stability of semiclassical theories of gravity. arXiv Preprint (2022). [arXiv:2201.10288 \[math-ph\]](#)
- [36] Meda, P., Pinamonti, N., Siemssen, D.: Existence and uniqueness of solutions of the semiclassical Einstein equation in cosmological models. *Ann. Henri Poincaré* **22**, 3965–4015 (2021). [arXiv:2007.14665 \[math-ph\]](#)
- [37] Moretti, V.: Comments on the stress energy tensor operator in curved space-time. *Commun. Math. Phys.* **232**, 189–221 (2003). [arXiv:gr-qc/0109048 \[gr-qc\]](#)
- [38] Mukhanov, V.: CMB, Quantum Fluctuations and the Predictive Power of Inflation. arXiv Preprint (2003). [arXiv:astro-ph/0303077](#)
- [39] Pinamonti, N.: On the Initial Conditions and Solutions of the Semiclassical Einstein Equations in a Cosmological Scenario. *Commun. Math. Phys.* **305**, 563 (2010). [arXiv:1001.0864 \[gr-qc\]](#)
- [40] Pinamonti, N., Siemssen, D.: Global Existence of Solutions of the Semiclassical Einstein Equation for Cosmological Spacetimes. *Commun. Math. Phys.* **334**, 171–191 (2015). [arXiv:1309.6303 \[math-ph\]](#)
- [41] Sanders, K.: Static symmetric solutions of the semi-classical Einstein-Klein-Gordon system. *Ann. Henri Poincaré* **23**, 1321–1358 (2020). [arXiv:2007.14311 \[math-ph\]](#)
- [42] Schander, S., Thiemann, T.: Backreaction in Cosmology. *Front. Astron. Space Sci.* **8**, 692198 (2021). [arXiv:2106.06043 \[gr-qc\]](#)
- [43] Siemssen, D.: The Semiclassical Einstein Equation on Cosmological Spacetimes. Ph.D. thesis, University of Genova, (2015). http://www.infn.it/thesis/thesis_dettaglio.php?tid=10652. [arXiv:1503.01826 \[math-ph\]](#)
- [44] Starobinsky, A. A.: A New Type of Isotropic Cosmological Models Without Singularity. *Phys. Lett. B* **91**, 99–102 (1980)

- [45] Starobinsky, A. A.: Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Phys. Lett. B* **117**, 175–178 (1982)
- [46] Suen, W.-M., Anderson, P. R.: Reheating in the higher-derivative inflationary models. *Phys. Rev. D* **35**, 2940–2954 (1987)
- [47] Tadaki, S.-I.: Stress Tensor in De Sitter Space. *Prog. Theor. Exp. Phys.* **80**, 654–662 (1988)
- [48] Tanabashi, M., *et al.* (Particle Data Group): Review of Particle Physics. *Phys. Rev. D* **98**, 030001 (2018) and 2019 update. <https://pdg.lbl.gov/2019/>
- [49] Wald, R. M.: The Back Reaction Effect in Particle Creation in Curved Space-Time. *Commun. Math. Phys.* **54**, 1–19 (1977)
- [50] Wald, R. M.: Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant. *Phys. Rev. D* **28**, 2118–2120 (1983)
- [51] Wald, R. M.: *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. University of Chicago Press, (1994)