Research Article

Barbara Kaltenbacher, Mostafa Meliani and Vanja Nikolić*

Limiting behavior of quasilinear wave equations with fractional-type dissipation

Abstract: In this work, we investigate a class of quasilinear wave equations of Westervelt type with, in general, nonlocal-in-time dissipation. They arise as models of nonlinear sound propagation through complex media with anomalous diffusion of Gurtin–Pipkin type. Aiming at minimal assumptions on the involved memory kernels – which we allow to be weakly singular – we prove the well-posedness of such wave equations in a general theoretical framework. In particular, the Abel fractional kernels, as well as Mittag-Leffler-type kernels, are covered by our results. The analysis is carried out uniformly with respect to the small involved parameter on which the kernels depend and which can be physically interpreted as the sound diffusivity or the thermal relaxation time. We then analyze the behavior of solutions as this parameter vanishes, and in this way relate the equations to their limiting counterparts. To establish the limiting problems, we distinguish among different classes of kernels and analyze and discuss all ensuing cases.

Keywords: quasilinear wave equations; Westervelt's equation; fractional dissipation; well-posedness; singular limits

MSC 2020: 35L05; 35L72; 35R11

1 Introduction

The classical theory of nonlinear sound propagation is based on employing the Fourier law of heat conduction within the system of governing equations. Using this law has drawbacks as it exhibits the following nonphysical feature: "a disturbance at any point in the body is felt instantly at every other point" [1]. This property is often referred to as the paradox of infinite speed of propagation.

In this work, we investigate quasilinear acoustic equations that originate from the use of the general Gurtin–Pipkin flux law [1] in place of the Fourier one. These equations are given by

$$((1 + 2ku^\epsilon)u_t^\epsilon)_t - c^2\Delta u^\epsilon - \mathfrak{K}_\epsilon * \Delta u_t^\epsilon = f \tag{1.1}$$

and can be understood as nonlocal generalizations of the classical Westervelt wave equation [2]. Here $u^\epsilon = u^\epsilon(x, t)$ represents the acoustic pressure, $c > 0$ the speed of sound, $k \in \mathbb{R}$ the nonlinearity coefficient,

*Corresponding author: Vanja Nikolić, Department of Mathematics, Radboud University, 6525 AJ Nijmegen, The Netherlands,
E-mail: vanja.nikolic@ru.nl
Barbara Kaltenbacher, Department of Mathematics, Alpen-Adria-Universität Klagenfurt, A-9020 Klagenfurt, Austria,
E-mail: barbara.kaltenbacher@aau.at
Mostafa Meliani, Department of Mathematics, Radboud University, 6525 AJ Nijmegen, The Netherlands, E-mail: mostafa.meliani@ru.nl

and $f = f(x, t)$ is the sound source. The function $\mathfrak{K}_\varepsilon = \mathfrak{K}_\varepsilon(t)$ is a memory kernel and $*$ denotes the Laplace convolution in time:

$$(\mathfrak{K}_\varepsilon * u_t^\varepsilon)(t) = \int_0^t \mathfrak{K}_\varepsilon(t - s) u_t^\varepsilon(s)\, ds.$$

The memory kernel $\mathfrak{K}_\varepsilon$ depends on the parameter $\varepsilon \ll 1$ which is in practice rather small, and it is thus important to determine the limiting behavior of solutions to (1.1) as $\varepsilon \searrow 0$. Different forms of the kernel $\mathfrak{K}_\varepsilon$ arise in the literature and will be covered by our investigations. The choice

$$\mathfrak{K}_\varepsilon = \varepsilon \frac{1}{\Gamma(\alpha)} t^{\alpha - 1}, \quad \alpha \in (0, 1), \tag{1.2}$$

where $\Gamma(\cdot)$ is the Gamma function, leads to the Westervelt equation with time-fractional damping of Djrbashian–Caputo-derivative type [3], [4], derived in Ref. [5]. In this case, the parameter ε has the physical meaning of the so-called sound diffusivity [6].

If the underlying flux law involves time-fractional relaxation, the kernels in the resulting acoustic equations have the form

$$\mathfrak{K}_\varepsilon = \left(\frac{\tau_\theta}{\varepsilon} \right)^{a-b} \frac{1}{\varepsilon^b} t^{b-1} E_{a,b}\left(-\left(\frac{t}{\varepsilon} \right)^a \right), \quad a, b \in (0, 1], \tag{1.3}$$

where $E_{a,b}$ is the generalized Mittag-Leffler function; see (2.6) below for its definition. Here τ_θ is a scaling parameter, which we discuss in detail in Section 2. In this setting, ε plays the physical role of the thermal relaxation time. In the physics literature, it is usually denoted by τ.

The goal of the present work is to investigate wave Equations (1.1) in terms of their unique solvability and then perform a singular limit analysis with respect to ε. In particular, we wish to establish in which sense (and, possibly, at which rate) the solutions of (1.1), supplemented by the initial conditions and homogeneous Dirichlet boundary data, converge to the limiting problem as $\varepsilon \searrow 0$. We aim to perform the analysis under minimally restrictive assumptions on the memory kernel $\mathfrak{K}_\varepsilon$ that will allow us to cover the kernels in (1.2) and (1.3).

This analysis constitutes a challenging task because of the interplay of quasilinear and nonlocal evolution, together with the non-restrictive assumptions imposed on the memory kernel $\mathfrak{K}_\varepsilon$. Unlike available results in the literature for the analysis of nonlinear wave equations with memory (see, e.g., [7], [8]), we do not require the kernel to be smooth on $[0, \infty)$ or non-negative. Additionally, the analysis of Westervelt equation must ensure that the leading coefficient does not degenerate; in other words, that

$$1 + 2ku^\varepsilon \geq \underline{\mathrm{m}} > 0$$

holds uniformly in ε. This task puts an additional strain on the analysis, as it requires obtaining ε-uniform bounds on $\|u^\varepsilon\|_{L^\infty(L^\infty(\Omega))}$ and guaranteeing their smallness.

1.1 Related works and novelty

To the best of our knowledge, this is the first rigorous work dealing with the limiting behavior of quasilinear wave equations with dissipation of Djrbashian–Caputo-derivative-type.

The local-in-time nonlinear acoustic models, on the other hand, are by now well-studied. We refer to Refs. [9]–[11] for the analysis of the Westervelt equation with strong damping:

$$((1 + 2ku^\varepsilon)u_t^\varepsilon)_t - c^2 \Delta u^\varepsilon - \varepsilon \Delta u_t^\varepsilon = 0 \tag{1.4}$$

on smooth bounded domains. Note that (1.4) can be cast within the family of equations in (1.1) by choosing $\mathfrak{K}_\varepsilon = \varepsilon \delta_0$, where δ_0 is the Dirac delta distribution. With strong damping in the equation, the corresponding Dirichlet boundary-value problem is known to be globally well-posed and the energy of the system decays exponentially with time; see [9, Theorems 1.2 and 1.3]. In the inviscid case ($\varepsilon = 0$), smooth solutions are expected to exist only until a certain propagation time, after which a blow-up occurs. The local well-posedness analysis of the inviscid

Westervelt equation with Dirichlet boundary conditions follows by the results of [12]. Numerical experiments indicating gradient blowup for inviscid nonlinear acoustic models can be found, for example, in Ref. [13, Ch. 5].

The limiting behavior of the strongly damped Westervelt Equation (1.4) as $\varepsilon \searrow 0$ has been studied in Ref. [14], where sufficient conditions have been determined for solutions $\{u^\varepsilon\}_{\varepsilon \in (0,\bar{\varepsilon})}$ to strongly converge to the solution of the inviscid problem at a linear rate.

We mention that local acoustic models involving more general nonlinearities or higher-order terms have also received plenty of attention in the literature; see, for example [15]–[17], and the review paper [18]. Third-order in time models with regular memory have also been extensively studied; see, e.g., [8], and the references given therein.

The analysis of quasilinear models in nonlinear acoustics involving time-fractional evolution has been initiated only recently and many questions are open. The well-posedness analysis of the Westervelt equation with fractional dissipation involving the kernel given in (1.2) can be found in Ref. [19]; the analysis in Ref. [20] allows for certain Mittag-Leffler kernels as well. We also point out the singular limit analysis for a class of nonlocal differential equations in Refs. [7], [21], from which we adopt certain ideas in the limit analysis with Mittag-Leffler kernels in Section 4.

1.2 Main contributions

Our main results pertain to establishing the limiting behavior of solutions to the family of Equations (1.1) on bounded domains, supplemented by initial conditions and Dirichlet boundary data, as $\varepsilon \searrow 0$. As will be shown, the limiting behavior will depend on the form of the kernel $\mathfrak{K}_\varepsilon$ and its dependence on ε, and so we will distinguish different classes of kernels motivated also by the physical background of sound propagation through media with nonlocal heat flux laws.

In case of the family of kernels given by $\mathfrak{K}_\varepsilon = \varepsilon\mathfrak{K}$, where $\mathfrak{K}$ satisfies suitable regularity and coercivity assumptions and does not depend on ε (this case covers (1.2)), in the limit one obtains the inviscid Westervelt equation:

$$((1 + 2ku)u_t)_t - c^2\Delta u = f.$$

The solutions of (1.1) will then be shown to converge in the standard energy norm

$$\|u\|_{\mathrm{E}} := \left(\|u_t\|^2_{L^\infty(L^2(\Omega))} + \|u\|^2_{L^\infty(H^1(\Omega))} \right)^{1/2} \tag{1.5}$$

to the solution of the inviscid model at a linear rate; see Theorem 4.1 and Corollary 4.2 for details. For this result to hold, sufficient smoothness of initial data is needed, namely

$$(u_0, u_1) \in \{u_0 \in H^3(\Omega) \cap H^1_0(\Omega): \Delta u_0|_{\partial\Omega} = 0\} \times (H^2(\Omega) \cap H^1_0(\Omega)).$$

Additionally smallness of data and short final time are needed to ensure uniform well-posedness; however, smallness of the initial conditions can be imposed in a lower-order norm than that of their regularity space; see Theorem 3.1. This limiting result significantly generalizes [14, Theorem 4.1], where the limiting behavior of the strongly damped Westervelt equation (obtained here by setting $\mathfrak{K} = \varepsilon\delta_0$ in (1.1)) has been studied.

Remarkably, for Mittag-Leffler-type kernels, the limiting dissipation produces a richer family of limiting equations, as we rigorously discuss in Section 4. For example, when $a \leq b$ in (1.3), the limiting equations are given by

$$((1 + 2ku)u_t)_t - c^2\Delta u - \tau^{a-b}_\theta \mathrm{D}^{a-b+1}_t \Delta u = f,$$

where D^{a-b+1}_t denotes the Djrbashian–Caputo fractional derivative of order $a - b + 1$. The details of this result together with the convergence rate can be found in Theorem 4.1 and Proposition 4.2.

1.3 Organization of the presentation

The rest of the presentation is organized as follows. In Section 2, we motivate this study by discussing the physical background of these nonlocal models in the context of nonlinear acoustics and giving examples of relevant

classes of memory kernels $\mathfrak{K}_\varepsilon$. Section 3 tackles the question of the uniform well-posedness of the nonlinear model (1.1) under suitable regularity, uniform boundedness, and coercivity assumptions on the memory kernels. The results of this section form the basis for the rigorous study of the limiting behavior. In Section 4, we then establish the limiting behavior of (1.1) for different classes of $\mathfrak{K}_\varepsilon$. The main results of this part are contained in Theorem 4.1 and Propositions 4.1 and 4.2. Finally, we consider concrete dependencies of $\mathfrak{K}_\varepsilon$ on ε motivated by the physics of nonlinear acoustics and prove in Section 5 that these classes of kernels indeed verify our theoretical assumptions.

1.4 Notation

In the analysis, we use the notation $A \lesssim B$ for $A \leq CB$ when the constant $C > 0$ does not depend on ε. We denote by

$$H_\diamond^2(\Omega) = H_0^1(\Omega) \cap H^2(\Omega),$$

$$H_\diamond^3(\Omega) = \left\{ u \in H^3(\Omega) : \operatorname{tr}_{\partial\Omega} u = 0, \ \operatorname{tr}_{\partial\Omega} \Delta u = 0 \right\},$$

the spaces of functions $H_0^2(\Omega) \subset H_\diamond^2(\Omega) \subset H^2(\Omega)$ and $H_0^3(\Omega) \subset H_\diamond^3(\Omega) \subset H^3(\Omega)$ that satisfy boundary conditions given above.

Given final time $T > 0$ and $p, q \in [1, \infty]$, we use $\|\cdot\|_{L^p(L^q(\Omega))}$ to denote the norm on $L^p(0, T; L^q(\Omega))$ and $\|\cdot\|_{L_t^p(L^q(\Omega))}$ to denote the norm on $L^p(0, t; L^q(\Omega))$ for $t \in (0, T)$. We use $(\cdot, \cdot)_{L^2}$ for the scalar product on $L^2(\Omega)$.

We employ the following short-hand notation for the Abel kernel:

$$g_\alpha(t) := \frac{1}{\Gamma(\alpha)} t^{\alpha-1}, \quad \alpha \in (0, 1) \tag{1.6}$$

and introduce the notational convention

$$g_0 := \delta_0,$$

where δ_0 is the Dirac delta distribution.

Throughout this work, D_t^η denotes the Djrbashian–Caputo fractional derivative, which is for $w \in W^{1,1}(0, t)$ defined by

$$\mathrm{D}_t^\eta w(t) = g_{\lceil \eta \rceil - \eta} * \mathrm{D}_t^{\lceil \eta \rceil} w, \qquad -1 < \eta < 1;$$

see, for example, [22, §1] and [23, §2.4.1]. Here $\lceil \eta \rceil \in \{0, 1\}$ is the integer obtained by rounding up η and $\mathrm{D}_t^{\lceil \eta \rceil}$ is the zeroth or first derivative operator.

2 Physical motivation and examples of relevant classes of kernels

In this section, we motivate different classes of memory kernels $\mathfrak{K}_\varepsilon$ that will be covered by our analysis. These kernels arise in wave models of nonlinear sound propagation through media with nonlocal heat flux laws. The nonlocal flux laws are of the Gurtin–Pipkin type [1] given by

$$\boldsymbol{q}(t) = -\kappa \, \mathfrak{K} * \nabla \theta,$$

where $\boldsymbol{q}$ denotes the flux vector, θ the absolute temperature, and κ is the thermal conductivity. As discussed in Ref. [24], different choices of the kernel $\mathfrak{K}$ lead to a rich family of flux laws that have appeared in the literature.

2.1 Abel kernels

The choice

$$\mathfrak{K} = \tau_\theta^{-\alpha} g_\alpha \tag{2.1}$$

results in the flux law involving the fractional integral:

$$q(t) = -\kappa \tau_\theta^{-\alpha} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \nabla\theta(s)\,ds, \quad 0 < \alpha \le 1; \tag{2.2}$$

see [25, Eq. (2.17)] and [26, Eq. (B.65)] for further modeling details. The constant $\tau_\theta > 0$ in (2.1) and (2.2) serves as a scaling factor to adjust for the dimensional inhomogeneity introduced by the fractional integral, in the way done in Ref. [26], Appendix B.4.1.2]. We mention that some references account for this by changing the units of the thermal conductivity κ.

2.2 Exponential kernels

Choosing an exponential kernel [25, Eq. (2.21)]:

$$\mathfrak{K}_\tau(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right) \tag{2.3}$$

leads to the heat flux law involving short-tail memory

$$q(t) = -\kappa \int_0^t \frac{1}{\tau} \exp\left(-\frac{t-s}{\tau}\right) \nabla\theta(s)\,ds, \tag{2.4}$$

where $\tau \ll 1$ is the intrinsic relaxation time of the heat flux. Assuming $q(0) = 0$, (2.4) can be seen as the solution to the well-known Maxwell–Cattaneo law [27]:

$$q + \tau q_t = -\kappa \nabla\theta.$$

2.3 Mittag-Leffler-type kernels

Exponential kernels (2.3) are, in fact, just a particular case of a large family of Mittag-Leffler memory kernels. In nonlinear acoustics, these are motivated by the Compte–Metzler heat flux laws, put forward and investigated in Ref. [28]:

$$\text{(GFE I)} \qquad (1 + \tau^\alpha D_t^\alpha)q(t) = -\kappa \tau_\theta^{1-\alpha} D_t^{1-\alpha} \nabla\theta;$$

$$\text{(GFE II)} \qquad (1 + \tau^\alpha D_t^\alpha)q(t) = -\kappa \tau_\theta^{\alpha-1} D_t^{\alpha-1} \nabla\theta;$$

$$\text{(GFE III)} \qquad (1 + \tau\partial_t)q(t) = -\kappa \tau_\theta^{1-\alpha} D_t^{1-\alpha} \nabla\theta;$$

$$\text{(GFE)} \qquad (1 + \tau^\alpha D_t^\alpha)q(t) = -\kappa \nabla\theta.$$

Here, as before, τ is the thermal relaxation time and we have introduced the scaling τ_θ to ensure dimensional homogeneity so that κ has the usual dimension of thermal conductivity. These laws can be solved for q and rewritten in the Gurtin–Pipkin form, as noted in Ref. [24]:

$$q(t) = -\kappa \mathfrak{K}_\tau * \nabla\theta$$

where the memory kernels are now given by

$$\mathfrak{K}_\tau = \left(\frac{\tau_\theta}{\tau}\right)^{a-b} \frac{1}{\tau^b} t^{b-1} E_{a,b}\left(-\left(\frac{t}{\tau}\right)^a\right). \tag{2.5}$$

We recall that the generalized Mittag-Leffler function is given by

$$E_{a,b}(t) = \sum_{k=0}^\infty \frac{t^k}{\Gamma(ak+b)}, \qquad a > 0, \ t, b \in \mathbb{R}; \tag{2.6}$$

see, e.g., [22, Ch. 2]. For the Compte–Metzler laws, the parameters (a, b) should be chosen as in Table 1, assuming additionally $\alpha > 1/2$ for the first pair (law GFE I).

Note that setting $a = b = 1$ in (2.5) allows us to cover the exponential kernel in (2.3).

2.4 Resulting acoustic models: Westervelt's equation and beyond

The derivation of nonlinear acoustic models based on the governing system of equations of sound motion that involves the Gurtin–Pipkin flux law can be done following closely the steps taken in Ref. [29, §4]. The modeling at first leads to one of the two approximations of the original governing system: the nonlocal wave equation of Blackstock type [30, p. 20]

$$\psi_{tt} - c^2(1 - 2\tilde{k}\psi_t)\Delta\psi - \delta\mathfrak{K}*\Delta\psi_t + \ell\partial_t|\nabla\psi|^2 = 0 \tag{2.7}$$

or the nonlocal wave equation of Kuznetsov type [31]

$$(1 + 2\tilde{k}\psi_t)\psi_{tt} - c^2\Delta\psi - \delta\mathfrak{K}*\Delta\psi_t + \ell\partial_t|\nabla\psi|^2 = 0, \tag{2.8}$$

where $\delta > 0$ is the sound diffusivity and $\tilde{k}, \ell \in \mathbb{R}$. The equations are expressed in terms of the acoustic velocity potential $\psi = \psi(x, t)$, which is related to the pressure by

$$u = \varrho\psi_t,$$

where ϱ is the medium density. If cumulative nonlinear effects in sound propagation are dominant so that approximation $|\nabla\psi|^2 \approx c^{-2}\psi_t^2$ holds, a nonlocal version of the Westervelt equation [2] in potential form is obtained:

$$(1 + 2\tilde{\tilde{k}}\psi_t)\psi_{tt} - c^2\Delta\psi - \delta\mathfrak{K}*\Delta\psi_t = 0,$$

where $\tilde{\tilde{k}} = \tilde{k} + c^{-2}\ell$. Differentiating in time and using the relation $u = \varrho\psi_t$ leads to the equation in the scope of present work:

$$((1 + 2ku)u_t)_t - c^2\Delta u - \delta\mathfrak{K}*\Delta u_t = \varrho\delta\mathfrak{K}(t)\Delta\psi_1, \tag{2.9}$$

with $k = \tilde{\tilde{k}}/\varrho$, where we have relied on the following differentiation formula:

$$(\mathfrak{K}*\Delta\psi_t)_t = \mathfrak{K}*\Delta\psi_{tt} + \mathfrak{K}(t)\Delta\psi_1.$$

In case of the Mittag-Leffler kernel (2.5), we would simply have $\mathfrak{K}_\tau$ above in place of $\mathfrak{K}$.

2.5 Unifying the physical parameters

Since we wish to investigate the limiting behavior in δ and in τ of the resulting Westervelt equation (where in case of $\tau \searrow 0$, we assume $\delta > 0$ to be fixed) and the uniform well-posedness analysis in both cases is qualitatively the same, we unite δ and τ here into one parameter ε and consider the generalized equation:

$$((1 + 2ku^\varepsilon)u_t^\varepsilon)_t - c^2\Delta u^\varepsilon - \mathfrak{K}_\varepsilon*\Delta u_t^\varepsilon = f$$

with suitable assumptions to be made on $\mathfrak{K}_\varepsilon$ (see (**A₁**) and (**A₂**) below). Setting

$$\mathfrak{K}_\varepsilon = \varepsilon\mathfrak{K} \quad \text{with} \quad \mathfrak{K}(t) = \tau_\theta^{-\alpha}g_\alpha(t)$$

Table 1: Parameters for the Mittag-Leffler kernels motivated by the Compte–Metzler laws.

	GFE I	GFE II	GFE III	GFE
a	α	α	1	α
b	$2\alpha - 1$	1	α	α

will allow us to cover (2.9), whereas choosing

$$\mathfrak{K}_\varepsilon = \delta \left(\frac{\tau_\theta}{\varepsilon} \right)^{a-b} \frac{1}{\varepsilon} \mathfrak{K}\left(\frac{t}{\varepsilon} \right) \quad \text{with} \quad \mathfrak{K}(t) = t^{b-1} E_{a,b}(-t^a)$$

covers the setting of thermal relaxation. In the latter case, we set $\delta = 1$ without loss of generality. We analyze the equation with a general source term f as the regularity assumptions needed for the analysis below are stronger than what we generally have with the right-hand side in (2.9) (unless one assumes that $\psi_1 = 0$).

     We mention that using the Compte–Metzler flux laws stated as fractional ODEs in the derivation of acoustic equations is also possible but leads to qualitatively different, higher-order in time fractional acoustic models. We refer the interested reader to Ref. [32] for their derivation and the well-posedness analysis.

3 Uniform well-posedness analysis

The aim of this section is to establish the well-posedness of quasilinear Westervelt Equation (3.1), uniformly in ε. This result will be the basis for the later study of the limiting behavior. Throughout we assume $\Omega \subset \mathbb{R}^d$, where $d \in \{1, 2, 3\}$, to be a bounded and C^3 regular domain. We consider the following initial boundary-value problem:

$$\begin{cases} ((1 + 2ku^\varepsilon)u_t^\varepsilon)_t - c^2 \Delta u^\varepsilon - \mathfrak{K}_\varepsilon * \Delta u_t^\varepsilon = f & \text{in } \Omega \times (0, T), \\ u^\varepsilon = 0 & \text{on } \partial\Omega \times (0, T), \\ (u^\varepsilon, u_t^\varepsilon) = (u_0, u_1), & \text{in } \Omega \times \{0\}. \end{cases} \qquad (3.1)$$

     Before proceeding to the question of solvability, we need to impose certain uniform regularity and coercivity assumptions on the memory kernel $\mathfrak{K}_\varepsilon$, which should be minimally restrictive.

3.1 Assumptions on the memory kernel in the analysis

Since we wish to study the limiting behavior of the nonlinear models as $\varepsilon \searrow 0$, we may focus our attention in the analysis on an interval $(0, \bar{\varepsilon})$ for some fixed $\bar{\varepsilon} > 0$ without loss of generality. We make the following assumption on the kernel's regularity and boundedness.

> Given $\varepsilon \in (0, \bar{\varepsilon})$, the kernel satisfies
>
> $$\mathfrak{K}_\varepsilon \in L^1(0, T) \cup \{\varepsilon \delta_0\} \qquad (\mathbf{A_1})$$
>
> with the ε-uniform bound:
>
> $$\|\mathfrak{K}_\varepsilon\|_{\mathcal{M}(0,T)} \le C_{\mathbf{A_1}}.$$

     Above we allow for $\mathfrak{K}_\varepsilon = \varepsilon \delta_0$, where δ_0 is the Dirac delta distribution, to cover the case of having strong damping in the equation (i.e., $-\varepsilon \Delta u_t^\varepsilon$), although we primarily focus our attention in the analysis and presentation on $\mathfrak{K}_\varepsilon \in L^1(0, T)$. We use $\| \cdot \|_{\mathcal{M}(0,T)}$ to denote the total variation norm which for the examples considered can be understood as:

$$\|\mathfrak{K}_\varepsilon\|_{\mathcal{M}(0,T)} = \begin{cases} \varepsilon & \text{if } \mathfrak{K}_\varepsilon = \varepsilon \delta_0, \\ \|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} & \text{if } \mathfrak{K}_\varepsilon \in L^1(0, T). \end{cases}$$

     In the well-posedness analysis, we will also need a coercivity assumption in order to achieve sufficient damping from the $\mathfrak{K}_\varepsilon$ term.

For all $y \in L^2(0, T; L^2(\Omega))$, it holds that for all $t \in (0, T)$

$$\int_0^t \int_\Omega (\mathfrak{K}_\varepsilon * y)(s)\, y(s)\, \mathrm{d}x\mathrm{d}s \geq C_{\mathfrak{K}_\varepsilon} \int_0^t \|(\mathfrak{K}_\varepsilon * y)(s)\|_{L^2(\Omega)}^2\, \mathrm{d}s, \tag{$\mathbf{A_2}$}$$

where the constant $C_{\mathfrak{K}_\varepsilon} > 0$ is uniformly bounded from below:

$$C_{\mathfrak{K}_\varepsilon} \geq C_{\mathbf{A_2}},$$

where $C_{\mathbf{A_2}} > 0$ does not depend on $\varepsilon \in (0, \bar{\varepsilon})$.

We mention that the two assumptions are verified for all classes of kernels discussed in the previous section, but postpone the proof of this claim to Section 5.

Strategy in the well-posedness analysis

To analyze (3.1), we introduce the fixed-point mapping

$$\mathcal{T}: \phi \mapsto u^\varepsilon,$$

where ϕ will belong to a ball in a suitable Bochner space and u^ε will solve the linearized problem

$$(\mathfrak{m}(\phi)u_t^\varepsilon)_t - c^2 \Delta u^\varepsilon - \mathfrak{K}_\varepsilon * \Delta u_t^\varepsilon = f$$

with the variable leading coefficient

$$\mathfrak{m}(\phi) = 1 + 2k\phi,$$

supplemented by the initial and boundary conditions given in (3.1). Clearly, a fixed-point of this mapping $\phi = u^\varepsilon$ would solve the nonlinear problem. This general strategy is common in the analysis of nonlinear acoustic models (see, e.g., [9]), however the crucial difference here is that it has to be conducted uniformly in ε.

3.2 Uniform well-posedness of a linear problem with variable principle coefficient

The well-definedness of the mapping as well as the proof of the existence of a unique fixed point rely on the uniform well-posedness of the linear problem, which we tackle first for general smooth and non-degenerate functions $\mathfrak{m} = \mathfrak{m}(\phi)$.

Proposition 3.1. *. Let $\varepsilon \in (0, \bar{\varepsilon})$. Given $T > 0$, let*

$$\phi \in X_\phi := L^\infty(0, T; H_\diamond^3(\Omega)) \cap W^{1,\infty}(0, T; H_\diamond^2(\Omega)).$$

Assume that there exist $\overline{\mathfrak{m}}$ and $\underline{\mathfrak{m}}$, independent of ε, such that

$$0 < \underline{\mathfrak{m}} \leq \mathfrak{m}(\phi) = 1 + 2k\phi(x, t) \leq \overline{\mathfrak{m}} \qquad in \ \ \Omega \times (0, T). \tag{3.2}$$

Let assumptions ($\mathbf{A_1}$) and ($\mathbf{A_2}$) on the kernel $\mathfrak{K}_\varepsilon$ hold. Furthermore, assume that the initial conditions satisfy

$$(u_0, u_1) \in H_\diamond^3(\Omega) \times H_\diamond^2(\Omega)$$

and let $f \in H^1(0, T; H_0^1(\Omega))$. Then there exists a unique solution of the linear problem

$$\begin{cases} (\mathfrak{m}(\phi)u_t^\epsilon)_t - c^2\Delta u^\epsilon - \mathfrak{K}_\epsilon * \Delta u_t^\epsilon = f & in\ \Omega \times (0, T), \\ u^\epsilon = 0 & on\ \partial\Omega \times (0, T), \\ (u^\epsilon, u_t^\epsilon) = (u_0, u_1), & in\ \Omega \times \{0\}, \end{cases} \tag{3.3}$$

such that

$$u^\epsilon \in \mathcal{U} = L^\infty(0, T; H^3_\diamond(\Omega)) \cap W^{1,\infty}(0, T; H^2_\diamond(\Omega)) \cap H^2(0, T; H^1_0(\Omega)). \tag{3.4}$$

The solution fulfills the following estimate:

$$\|\Delta u_t^\epsilon\|^2_{L^\infty(L^2(\Omega))} + \|\nabla\Delta u^\epsilon\|^2_{L^\infty(L^2(\Omega))}$$
$$\leq C(\phi, T)\,(\|\Delta u_1\|^2_{L^2(\Omega)} + \|\nabla\Delta u_0\|^2_{L^2(\Omega)} + \|f\|^2_{H^1(H^1(\Omega))}), \tag{3.5}$$

where the constant has the form

$$C(\phi, T) = C_1\,\exp(C_2\Big(1 + \|\phi\|_{X_\phi} + \cdots + \|\phi\|^6_{X_\phi}\Big)T)$$

for some $C_1, C_2 > 0$ that do not depend on ϵ.

Proof. We perform the analysis using a smooth Galerkin discretization in space; see, e.g., [33, §7] and [14, Section 3] for details on this method. The existence of a unique approximate solution follows by reducing the semi-discrete problem to a system of Volterra integral equations of the second kind and employing relatively standard existence arguments, which we therefore provide in Appendix A.

The key to the proof is a uniform energy bound with respect to the discretization that must, in this context, also be uniform in ϵ. We follow the general testing strategy in Refs. [14], [19] with special attention paid to the uniformity in ϵ. For notational simplicity, we drop the discretization parameter when denoting the approximate solution below.

As we cannot rely much on the dissipation properties of the nonlocal term in the equation, we need sufficient smoothness of the solution to overpower the right-hand side terms arising after testing. For this reason, we test the (semi-discrete) problem with $\Delta^2 u_t^\epsilon$. After integrating over space, we arrive at

$$\Big(\mathfrak{m}u_{tt}^\epsilon - c^2\Delta u^\epsilon - \mathfrak{K}_\epsilon * \Delta u_t^\epsilon + \mathfrak{m}_t u_t^\epsilon,\ \Delta^2 u_t^\epsilon\Big)_{L^2} = (f, \Delta^2 u_t^\epsilon)_{L^2}. \tag{3.6}$$

We note that the following identity holds:

$$(\mathfrak{m}u_{tt}^\epsilon, \Delta^2 u_t^\epsilon)_{L^2} = (\mathfrak{m}\Delta u_{tt}^\epsilon, \Delta u_t^\epsilon)_{L^2} + (u_{tt}^\epsilon\Delta\mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla\mathfrak{m},\ \Delta u_t^\epsilon)_{L^2}$$
$$= \frac{1}{2}\frac{d}{dt}(\mathfrak{m}\Delta u_t^\epsilon, \Delta u_t^\epsilon)_{L^2} - \frac{1}{2}(\mathfrak{m}_t\Delta u_t^\epsilon, \Delta u_t^\epsilon)_{L^2} + (u_{tt}^\epsilon\Delta\mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla\mathfrak{m},\ \Delta u_t^\epsilon)_{L^2}$$

because $u_{tt}^\epsilon|_{\partial\Omega} = \Delta u_t^\epsilon|_{\partial\Omega} = 0$. Similarly, we have

$$\Big(\mathfrak{m}_t u_t^\epsilon, \Delta^2 u_t^\epsilon\Big)_{L^2} = (\Delta(\mathfrak{m}_t u_t^\epsilon), \Delta u_t^\epsilon)_{L^2}$$
$$= (\mathfrak{m}_t\Delta u_t^\epsilon + 2\nabla\mathfrak{m}_t \cdot \nabla u_t^\epsilon + u_t^\epsilon\Delta\mathfrak{m}_t,\ \Delta u_t^\epsilon)_{L^2}.$$

Thus, by integrating by parts in (3.6) in space, integrating over $(0, t)$ for $t \in (0, T)$, and employing coercivity assumption $(\mathbf{A_2})$, we arrive at the following energy inequality:

$$\frac{1}{2}\|\sqrt{\mathfrak{m}}\Delta u_t^\epsilon\|^2_{L^2(\Omega)}\Big|_0^t + \frac{c^2}{2}\|\nabla\Delta u^\epsilon\|^2_{L^2(\Omega)}\Big|_0^t + C_{\mathbf{A_2}}\int_0^t \|(\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon)(s)\|^2_{L^2(\Omega)}\,ds$$

$$\leq -\frac{1}{2}\int_0^t (\mathfrak{m}_t \Delta u_t^\epsilon, \Delta u_t^\epsilon)_{L^2}\, ds - \int_0^t \big(u_{tt}^\epsilon \Delta \mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla \mathfrak{m} + 2\nabla \mathfrak{m}_t \cdot \nabla u_t^\epsilon$$

$$+ u_t^\epsilon \Delta \mathfrak{m}_t, \Delta u_t^\epsilon\big)_{L^2}\, ds + \int_\Omega \nabla f \cdot \nabla \Delta u^\epsilon\, dx \Big|_0^t - \int_0^t \int_\Omega \nabla f_t \cdot \nabla \Delta u^\epsilon\, dx ds, \tag{3.7}$$

where we have also used that $f|_{\partial\Omega} = 0$. We can estimate the first integral on the right-hand side by relying on the embedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$ as follows:

$$-\frac{1}{2}\int_0^t (\mathfrak{m}_t \Delta u_t^\epsilon, \Delta u_t^\epsilon)_{L^2}\, ds \leq |k|\, \|\phi_t\|_{L^\infty(H^2(\Omega))} \|\Delta u_t^\epsilon\|^2_{L_t^2(L^2(\Omega))}$$

$$\lesssim \|\phi\|_{X_\phi} \|\Delta u_t^\epsilon\|^2_{L_t^2(L^2(\Omega))}, \tag{3.8}$$

where the resulting term will be handled by Grönwall's inequality later on. Similarly,

$$\int_0^t (2\nabla \mathfrak{m}_t \cdot \nabla u_t^\epsilon + u_t^\epsilon \Delta \mathfrak{m}_t, \Delta u_t^\epsilon)_{L^2}\, ds \lesssim \|\Delta \mathfrak{m}_t\|_{L^\infty(L^2(\Omega))} \|\Delta u_t^\epsilon\|^2_{L_t^2(L^2(\Omega))}$$

$$\lesssim \|\phi\|_{X_\phi} \|\Delta u_t^\epsilon\|^2_{L_t^2(L^2(\Omega))}.$$

We can use Hölder's and Young's inequalities to estimate the f terms in (3.7):

$$\int_\Omega \nabla f \cdot \nabla \Delta u^\epsilon\, dx \Big|_0^t - \int_0^t \int_\Omega \nabla f_t \cdot \nabla \Delta u^\epsilon\, dx ds$$

$$\leq \frac{1}{c^2}\|\nabla f(t)\|^2_{L^2(\Omega)} + \frac{c^2}{4}\|\nabla \Delta u^\epsilon(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla f(0)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla \Delta u_0\|^2_{L^2(\Omega)}$$

$$+ \|\nabla f_t\|^2_{L^2(L^2(\Omega))} + \|\nabla \Delta u^\epsilon\|^2_{L_t^2(L^2(\Omega))}, \tag{3.9}$$

where the two resulting $\nabla \Delta u^\epsilon$ terms above can be either absorbed by the left-hand side of (3.7) or handled by Grönwall's inequality.

It remains to estimate the u_{tt}^ϵ terms on the right-hand side of (3.7). To this end, we first use Hölder's and Young's inequalities:

$$\int_0^t (u_{tt}^\epsilon \Delta \mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla \mathfrak{m}, \Delta u_t^\epsilon)_{L^2}\, ds$$

$$\leq \big(\|u_{tt}^\epsilon\|_{L^2(L^4(\Omega))} \|\Delta \mathfrak{m}\|_{L^\infty(L^4(\Omega))} + \|\nabla u_{tt}^\epsilon\|_{L^2(L^2(\Omega))} \|\nabla \mathfrak{m}\|_{L^\infty(L^\infty(\Omega))}\big) \|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}.$$

By the embeddings $H^1(\Omega) \hookrightarrow L^4(\Omega)$ and $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$, we further have

$$\|\Delta \mathfrak{m}\|_{L^\infty(L^4(\Omega))} + \|\nabla \mathfrak{m}\|_{L^\infty(L^\infty(\Omega))} \lesssim \|\Delta \phi\|_{L^\infty(H^1(\Omega))} + \|\nabla \phi\|_{L^\infty(H^2(\Omega))}$$

$$\lesssim \|\phi\|_{X_\phi}$$

and thus

$$\int_0^t (u_{tt}^\epsilon \Delta \mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla \mathfrak{m}, \Delta u_t^\epsilon)_{L^2}\, ds \leq \|\phi\|_{X_\phi} \|u_{tt}^\epsilon\|_{L^2(H^1(\Omega))} \|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}. \tag{3.10}$$

From here, we can use the (semi-discrete) PDE to further bound the u_{tt}^ϵ term. We first have by Poincaré's inequality

$$\|u_{tt}^\epsilon\|_{L^2(H^1(\Omega))} \lesssim \|\nabla u_{tt}^\epsilon\|_{L^2(L^2(\Omega))}.$$

Then to estimate the right-hand side term, we use the (semi-discrete) PDE and the identity

$$\mathfrak{m}\nabla u_{tt}^\epsilon + \nabla\mathfrak{m} u_{tt}^\epsilon = \mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon + c^2\nabla\Delta u^\epsilon - \nabla[\mathfrak{m}_t u_t^\epsilon] + \nabla f.$$

The uniform boundedness of $\mathfrak{m}$ in (3.2) allows us to conclude that

$$\|\nabla u_{tt}^\epsilon\|_{L^2(L^2(\Omega))} \lesssim \|\nabla\mathfrak{m}\|_{L^\infty(L^\infty(\Omega))}\|\mathfrak{K}_\epsilon * \Delta u_t^\epsilon + c^2\Delta u^\epsilon - \mathfrak{m}_t u_t^\epsilon + f\|_{L^2(L^2(\Omega))}$$
$$+ \|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon + c^2\nabla\Delta u^\epsilon - \nabla[\mathfrak{m}_t u_t^\epsilon] + \nabla f\|_{L^2(L^2(\Omega))}.$$

Further estimating the right-hand side terms leads to

$$\|\nabla u_{tt}^\epsilon\|_{L^2(L^2(\Omega))} \lesssim \|\phi\|_{X_\phi}\left(\|\Delta u^\epsilon\|_{L^2(L^2(\Omega))} + \|\mathfrak{K}_\epsilon\|_{\mathcal{M}(0,t)}\|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))} + \|\phi\|_{X_\phi}\|u_t^\epsilon\|_{L^2(L^2(\Omega))} + \|f\|_{L^2(L^2(\Omega))}\right)$$
$$+ \|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))} + \|\nabla\Delta u^\epsilon\|_{L^2(L^2(\Omega))} + \|\phi\|_{X_\phi}\|u_t^\epsilon\|_{L^2(L^4(\Omega))} + \|\phi\|_{X_\phi}\|\nabla u_t^\epsilon\|_{L^2(L^2(\Omega))}$$
$$+ \|\nabla f\|_{L^2(L^2(\Omega))}.$$

Thanks to assumption $(\mathbf{A_1})$, we have the uniform bound $\|\mathfrak{K}_\epsilon\|_{\mathcal{M}(0,t)} \leq C_{\mathbf{A_1}}$. Going back to (3.10) and using the estimate on ∇u_{tt}^ϵ thus yields

$$\int_0^t (u_{tt}^\epsilon \Delta\mathfrak{m} + 2\nabla u_{tt}^\epsilon \cdot \nabla\mathfrak{m}, \Delta u_t^\epsilon)_{L^2}\, ds$$

$$\lesssim \left(\|\phi\|_{X_\phi} + \|\phi\|_{X_\phi}^2 + \|\phi\|_{X_\phi}^3\right)\left\{\|\nabla\Delta u^\epsilon\|_{L^2(L^2(\Omega))} + \|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}\right.$$
$$\left. + \|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))} + \|u_t^\epsilon\|_{L^2(H^1(\Omega))} + \|f\|_{L^2(H^1(\Omega))}\right\}\|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}. \tag{3.11}$$

By employing estimates (3.8), (3.9), and (3.11) in (3.7), we arrive at

$$\frac{1}{2}\|\sqrt{\mathfrak{m}}\Delta u_t^\epsilon\|_{L^2(\Omega)}^2\Big|_0^t + \frac{c^2}{2}\|\nabla\Delta u^\epsilon\|_{L^2(\Omega)}^2\Big|_0^t + C_{\mathbf{A_2}}\int_0^t \|(\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon)(s)\|_{L^2(\Omega)}^2\, ds$$

$$\lesssim (1 + \|\phi\|_{X_\phi})\|\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}^2 + \|\nabla\Delta u_0\|_{L^2(\Omega)}^2 + \|f\|_{H^1(H^1(\Omega))}^2$$
$$+ \left(\|\phi\|_{X_\phi} + \|\phi\|_{X_\phi}^2 + \|\phi\|_{X_\phi}^3\right)^2\|\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}^2 + \|\nabla\Delta u^\epsilon\|_{L_t^2(L^2(\Omega))}^2 + \|\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}^2 + \|u_t^\epsilon\|_{L_t^2(H^1(\Omega))}^2$$
$$+ \left(\|\phi\|_{X_\phi} + \|\phi\|_{X_\phi}^2 + \|\phi\|_{X_\phi}^3\right)\|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}\|\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}. \tag{3.12}$$

Other than the last term in (3.12), all other terms on the right-hand side can be tackled using Grönwall's inequality. To treat the last term, we employ Young's inequality:

$$\left(\|\phi\|_{X_\phi} + \|\phi\|_{X_\phi}^2 + \|\phi\|_{X_\phi}^3\right)\|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}\|\Delta u_t^\epsilon\|_{L^2(L^2(\Omega))}$$

$$\leq \gamma\|\mathfrak{K}_\epsilon * \nabla\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}^2 + \frac{1}{4\gamma}\left(\|\phi\|_{X_\phi} + \|\phi\|_{X_\phi}^2 + \|\phi\|_{X_\phi}^3\right)^2\|\Delta u_t^\epsilon\|_{L_t^2(L^2(\Omega))}^2. \tag{3.13}$$

If $\tilde{C} > 0$ is the hidden constant within $\lesssim$ in (3.12), we can choose γ as

$$\gamma = \frac{1}{\tilde{C}}C_{\mathbf{A_2}}/2.$$

The term $\gamma\|\mathfrak{K}_\varepsilon * \nabla\Delta u_t^\varepsilon\|_{L_t^2(L^2(\Omega))}^2$ can be absorbed by the left-hand side of (3.12). We then use Grönwall's inequality for the second term on the right-hand side of (3.13) to arrive at an ε-uniform estimate.

Altogether, starting from estimates (3.12) and (3.13), an application of Grönwall's inequality leads to the final estimate (3.5), at first in a semi-discrete setting. The bound transfers to the continuous setting through standard weak compactness arguments, analogously to, for example, [14, Proposition 3.1]. Uniqueness of the constructed solution can be established by testing the homogeneous problem (where $f = u_0 = u_1 = 0$) with, for example, u_t^ε. We omit these details here. □

3.3 Uniform lower-order estimate

In the fixed-point analysis, we will rely on Agmon's interpolation inequality to prove that the leading factor $\mathfrak{m}(u^\varepsilon) = 1 + 2ku^\varepsilon$ of the nonlinear equation does not degenerate. For $v \in H^2(\Omega)$, the inequality is given by

$$\|v\|_{L^\infty(\Omega)} \leq C_A \|v\|_{L^2(\Omega)}^{1-d/4}\|v\|_{H^2(\Omega)}^{d/4} \quad (d \leq 3), \tag{3.14}$$

see [34, Lemma 13.2, Ch. 13]. Using this estimate on $v = u^\varepsilon(t)$ together with a bound on $\|u^\varepsilon(t)\|_{L^2(\Omega)}$ in terms of data in lower-order topology, will allow us to impose data smallness in that topology instead $H_\diamond^3(\Omega) \times H_\diamond^2(\Omega)$. We derive this bound next.

Under the assumptions of Proposition 3.1, testing (3.3) with u_t^ε and using coercivity assumption ($\mathbf{A_2}$) leads to the inequality

$$\frac{1}{2}\|\sqrt{\mathfrak{m}}\,u_t^\varepsilon(t)\|_{L^2(\Omega)}^2\Big|_0^t + \frac{c^2}{2}\|\nabla u^\varepsilon(t)\|_{L^2(\Omega)}^2\Big|_0^t + C_{\mathbf{A_2}}\int_0^t \|(\mathfrak{K}_\varepsilon * \nabla u_t^\varepsilon)(s)\|_{L^2(\Omega)}^2\,\mathrm{d}s$$

$$\leq \frac{1}{2}\int_0^t (\mathfrak{m}_t u_t^\varepsilon, u_t^\varepsilon)_{L^2}\,\mathrm{d}s + \int_0^t (f, u_t^\varepsilon)_{L^2}\,\mathrm{d}s$$

$$\lesssim \|\phi_t\|_{L^\infty(L^\infty(\Omega))}\|u_t^\varepsilon\|_{L_t^2(L^2(\Omega))}^2 + \|f\|_{L^1(L^2(\Omega))}\|u_t^\varepsilon\|_{L_t^\infty(L^2(\Omega))},$$

where the hidden constant does not depend on ε. From here by using Young's and Grönwall's inequalities, we obtain the following lower-order counterpart of (3.5):

$$\|u_t^\varepsilon\|_{L^\infty(L^2(\Omega))}^2 + \|u^\varepsilon\|_{L^\infty(H^1(\Omega))}^2$$

$$\leq C_3 \exp\left(C_4(1 + \|\phi\|_{X_\phi})T\right)\left(\|u_0\|_{H^1(\Omega)}^2 + \|u_1\|_{L^2(\Omega)}^2 + \|f\|_{L^1(L^2(\Omega))}^2\right), \tag{3.15}$$

for some $C_3, C_4 > 0$ that do not depend on ε. We will use this bound in the fixed-point proof to ensure the smallness of $\|u^\varepsilon\|_{L^\infty(L^2(\Omega))} \lesssim \|u^\varepsilon\|_{L^\infty(H^1(\Omega))}$.

3.4 Fixed-point analysis

We proceed with the analysis of the nonlinear problem, where we will combine our previous results with a fixed-point argument.

Theorem 3.1. *Let $k \in \mathbb{R}$ and $\varepsilon \in (0, \bar{\varepsilon})$. Let $(u_0, u_1) \in H_\diamond^3(\Omega) \times H_\diamond^2(\Omega)$ be such that*

$$\|u_0\|_{H^3(\Omega)}^2 + \|u_1\|_{H^2(\Omega)}^2 + \|f\|_{H^1(H^1(\Omega))}^2 \leq r^2,$$

where r does not depend on ε. Let assumptions ($\mathbf{A_1}$) and ($\mathbf{A_2}$) on the kernel $\mathfrak{K}_\varepsilon$ hold and let $f \in H^1(0, T; H_0^1(\Omega))$. Then there exist data size $r_0 = r_0(r) > 0$ and final time $T = T(r)$, both independent of ε, such that if

$$\|u_0\|_{H^1(\Omega)}^2 + \|u_1\|_{L^2(\Omega)}^2 + \|f\|_{L^1(L^2(\Omega))}^2 \leq r_0^2,$$

then there is a unique solution $u^\epsilon \in \mathcal{U}$ of (3.1), with $\mathcal{U}$ defined in (3.4), which satisfies the following estimate uniformly in ϵ:

$$\|u^\epsilon\|_{\mathcal{U}}^2 \leq C(T)\left(\|u_0\|_{H^3(\Omega)}^2 + \|u_1\|_{H^2(\Omega)}^2 + \|f\|_{H^1(H^1(\Omega))}^2\right).$$

Proof. The proof follows by using the Banach fixed-point theorem on $\mathcal{T}$, in the general spirit of Ref. [14, Theorem 6.1] which considers uniform analysis in a local-in-time setting. The mapping $\mathcal{T}: \phi \mapsto u^\epsilon$ is defined on the ball

$$\phi \in \mathcal{B} = \left\{\phi \in \mathcal{U}: \|\phi\|_{\mathcal{U}} \leq R, \; 4|k|\|\phi\|_{L^\infty(L^\infty(\Omega))} \leq 1, \; (\phi, \phi_t)|_{t=0} = (u_0, u_1)\right\}$$

with a large enough radius $R > 0$, which will be made precise below. The set $\mathcal{B}$ is non-empty as the solution of the linear problem with $\mathfrak{m} = 1$ belongs to it if R is sufficiently large and r_0 sufficiently small.

Let $\phi \in \mathcal{B}$. Then the smoothness assumption on ϕ in Proposition 3.1 holds since $\mathcal{U} \subset X_\phi$. Furthermore, since $4|k|\|\phi\|_{L^\infty(L^\infty(\Omega))} \leq 1$, $\mathfrak{m}$ does not degenerate:

$$\frac{1}{2} = \underline{\mathfrak{m}} \leq \mathfrak{m} = 1 + 2k\phi(x, t) \leq \overline{\mathfrak{m}} = \frac{3}{2}, \quad (x, t) \in \Omega \times (0, T).$$

(1) *The self-mapping property.* We wish to prove that $u^\epsilon := \mathcal{T}\phi \in \mathcal{B}$. Since the assumptions of Proposition 3.1 hold, we know that u^ϵ satisfies

$$\|u^\epsilon\|_{\mathcal{U}}^2 \leq C_1 \exp\left(C_2\left(1 + \|\phi\|_{X_\phi} + \cdots + \|\phi\|_{X_\phi}^6\right)T\right)\left(\|u_0\|_{H^3(\Omega)}^2 + \|u_1\|_{H^2(\Omega)}^2 + \|f\|_{H^1(H^1(\Omega))}^2\right).$$

We can guarantee that $\|u^\epsilon\|_{\mathcal{U}} \leq R$ by choosing $R = R(r)$ and $T = T(r)$ so that

$$C_1 \exp(C_2(1 + R + \cdots + R^6)T)r^2 \leq R^2.$$

We next show that the bound $4|k|\|u^\epsilon\|_{L^\infty(L^\infty(\Omega))} \leq 1$ holds. To this end, we rely on Agmon's interpolation inequality (3.14) and combine it with the uniform lower-order bound in (3.15). We have

$$\|u^\epsilon\|_{L^\infty(H^2(\Omega))} \lesssim \|u^\epsilon\|_{\mathcal{U}} \lesssim R,$$

and so using Agmon's inequality leads to

$$\|u^\epsilon\|_{L^\infty(L^\infty(\Omega))} \leq C_A \|u^\epsilon\|_{L^\infty(L^2(\Omega))}^{1-d/4}\|u^\epsilon\|_{L^\infty(H^2(\Omega))}^{d/4}$$

$$\leq C_A\left\{C_3 \exp(C_4 T(1 + R))\left(\|u_0\|_{H^1(\Omega)}^2 + \|u_1\|_{L^2(\Omega)}^2 + \|f\|_{L^1(L^2(\Omega))}^2\right)\right\}^{1/2-d/8} R^{d/4}.$$

We then choose data size $r_0 = r_0(r) > 0$ small enough, so that

$$4|k|C_A\left\{C_3 \exp(C_4 T(1 + R))r_0^2\right\}^{1/2-d/8} R^{d/4} \leq 1.$$

Then $4|k|\|u^\epsilon\|_{L^\infty(L^\infty(\Omega))} \leq 1$, which was the last condition needed to conclude that $u^\epsilon \in \mathcal{B}$.

(2) *Strict contractivity.* Take $\phi^{(1)}$ and $\phi^{(2)}$ in $\mathcal{B}$, and denote their difference by $\overline{\phi} = \phi^{(1)} - \phi^{(2)}$. Let $u^{\epsilon,(1)} = \mathcal{T}(\phi^{(1)})$ and $u^{\epsilon,(2)} = \mathcal{T}(\phi^{(2)})$. Their difference $\overline{u} = u^{\epsilon,(1)} - u^{\epsilon,(2)} \in \mathcal{B}$ then solves

$$(1 + 2k\phi^{(1)})\overline{u}_{tt} - c^2\Delta\overline{u} - \mathfrak{K}_\epsilon * \Delta\overline{u}_t + 2k\phi_t^{(1)}\overline{u}_t = -2k\overline{\phi}u_{tt}^{\epsilon,(2)} - 2k\overline{\phi}_t u_t^{\epsilon,(2)}$$

with zero initial and boundary data. Since the right-hand side of the above equation does not belong to $H^1(0, T; H^1(\Omega))$, we cannot use estimate (3.5) to prove contractivity of the mapping $\mathcal{T}$ with respect to $\|\cdot\|_{\mathcal{U}}$. Instead, we prove that the mapping $\mathcal{T}$ is strictly contractive in the space

$$X_E := W^{1,\infty}(0, T; L^2(\Omega)) \cap L^\infty(0, T; H_0^1(\Omega)),$$

endowed with the norm defined in (1.5). We can again use the uniform lower-order bound established in (3.15) to conclude that

$$\|\overline{u}_t\|_{L^\infty(L^2(\Omega))} + \|\overline{u}\|_{L^\infty(H^1(\Omega))}$$

$$\lesssim \| -2k\overline{\phi}u_{tt}^{\varepsilon,(2)} - 2k\overline{\phi}_t u_t^{\varepsilon,(2)}\|_{L^1(L^2(\Omega))}$$

$$\lesssim |k|\|\overline{\phi}\|_{L^2(L^4(\Omega))}\|u_{tt}^{\varepsilon,(2)}\|_{L^2(L^4(\Omega))} + |k|\|\overline{\phi}_t\|_{L^2(L^2(\Omega))}\|u_t^{\varepsilon,(2)}\|_{L^2(L^\infty(\Omega))}.$$

By using the embeddings $H^1(\Omega) \hookrightarrow L^4(\Omega)$ and $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$ and the fact that $u^{\varepsilon,(2)} \in \mathcal{B}$, we further infer that

$$\|\overline{u}_t\|_{L^\infty(L^2(\Omega))} + \|\overline{u}\|_{L^\infty(H^1(\Omega))} \lesssim |k|(\|\overline{\phi}\|_{L^2(H^1(\Omega))} + \|\overline{\phi}_t\|_{L^2(L^2(\Omega))})$$

$$\lesssim |k|\sqrt{T}(\|\overline{\phi}\|_{L^\infty(H^1(\Omega))} + \|\overline{\phi}_t\|_{L^\infty(L^2(\Omega))}).$$

Therefore, by reducing T if needed (independently of ε), we can ensure that the mapping $\mathcal{T}$ is strictly contractive with respect to the energy norm (1.5).

The arguments showing that $\mathcal{B}$ is closed with respect to this norm are analogous to those of [14, Theorem 4.1]. By the Banach fixed-point theorem, we therefore obtain a unique fixed point $u^\varepsilon = \mathcal{T}(u^\varepsilon)$ in $\mathcal{B}$, which solves the nonlinear problem. □

We note that as a byproduct of Theorem 3.1, we obtain local well-posedness for the strongly damped ($\mathfrak{K}_\varepsilon = \varepsilon\delta_0$), inviscid ($\mathfrak{K}_\varepsilon \equiv 0$), and fractionally damped ($\mathfrak{K}_\varepsilon = \varepsilon g_\alpha$) Westervelt equations, therefore generalizing the results of Refs. [12], [14], [19].

4 Establishing the limiting behavior for relevant classes of kernels

We next wish to determine the limiting behavior of solutions to (1.1) as $\varepsilon \searrow 0$. We begin by proving (Hölder) continuity of the solution with respect to the memory kernel.

Theorem 4.1. *Let* $\varepsilon_1, \varepsilon_2 \in (0, \overline{\varepsilon})$. *Under the assumptions Theorem* 3.1, *the following estimates hold:*

$$\|u^{\varepsilon_1} - u^{\varepsilon_2}\|_E \lesssim \|(\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2})*1\|_{L^1(0,T)}^{1/2}, \tag{4.1}$$

and

$$\|u^{\varepsilon_1} - u^{\varepsilon_2}\|_E \lesssim \|\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}\|_{\mathcal{M}(0,T)}. \tag{4.2}$$

Proof. To prove the statement, we can see the difference $\overline{u} = u^{\varepsilon_1} - u^{\varepsilon_2}$ as the solution to

$$((1 + 2ku^{\varepsilon_1})\overline{u}_t)_t - c^2\Delta\overline{u} - \mathfrak{K}_{\varepsilon_1}*\Delta\overline{u}_t + 2k(\overline{u}u_t^{\varepsilon_2})_t = (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2})*\Delta u_t^{\varepsilon_2}$$

with zero initial and boundary conditions. We can test this equation with $\overline{u}_t$ and proceed similarly to the proof of contractivity in Theorem 3.1 with now $\mathfrak{m} = 1 + 2ku^{\varepsilon_1}$. The new term compared to before is the convolution on the right-hand side. After testing, this term can be handled using Young's convolution inequality as follows:

$$\int_0^t \int_\Omega (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2})*\Delta u_t^{\varepsilon_2}\overline{u}_t \, dxds$$

$$\lesssim \|\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}\|_{\mathcal{M}(0,T)}\|\Delta u_t^{\varepsilon_2}\|_{L^2(L^2(\Omega))}\|\overline{u}_t\|_{L^2(L^2(\Omega))}$$

$$\lesssim \|\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}\|_{\mathcal{M}(0,T)}^2\|\Delta u_t^{\varepsilon_2}\|_{L^2(L^2(\Omega))}^2 + \|\overline{u}_t\|_{L^2(L^2(\Omega))}^2.$$

Since by Theorem 3.1, $\|\Delta u_t^{\varepsilon_2}\|_{L^2(L^2(\Omega))} \lesssim \|u^{\varepsilon_2}\|_{\mathcal{U}}$ is uniformly bounded, we obtain the claimed estimate by relying on Grönwall's inequality.

The difference of kernels term can also be treated using integration by parts in space:

$$\int_0^t (((\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * \Delta u_t^{\varepsilon_2})(s), \overline{u}_t(s))_{L^2}\, ds$$

$$= \int_0^t \left\{ ((1 * (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * \Delta u_{tt}^{\varepsilon_2})(s), \overline{u}_t(s))_{L^2} + \left((1 * (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}))(s)\Delta u_1, \overline{u}_t(s)\right)_{L^2} \right\} ds$$

$$= \int_0^t \left\{ -((1 * (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * \nabla u_{tt}^{\varepsilon_2})(s), \nabla \overline{u}_t(s))_{L^2} + \left((1 * (\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}))(s)\Delta u_1, \overline{u}_t(s)\right)_{L^2} \right\} ds. \qquad (4.3)$$

Then by Young's convolution inequality

$$\int_0^t (((\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * \Delta u_t^{\varepsilon_2})(s), \overline{u}_t(s))_{L^2}\, ds$$

$$\leq \|(\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * 1\|_{L^1(0,T)} \left(\|\nabla u_{tt}^{\varepsilon_2}\|_{L_t^1(L^2(\Omega))} \|\nabla \overline{u}_t\|_{L_t^\infty(L^2(\Omega))} + \|\Delta u_1\|_{L^2(\Omega)} \|\overline{u}_t\|_{L_t^1(L^2(\Omega))} \right).$$

Note that we would not be able to absorb the term $\|\nabla \overline{u}_t\|^2_{L_t^\infty(L^2(\Omega))}$ by the left-hand side (i.e., the energy norm). However, by Theorem 3.1, we know that the following uniform bound holds:

$$\|\nabla u_{tt}^{\varepsilon_2}\|_{L_t^1(L^2(\Omega))} \|\nabla \overline{u}_t\|_{L_t^\infty(L^2(\Omega))} + \|\Delta u_1\|_{L^2(\Omega)} \|\overline{u}_t\|_{L_t^1(L^2(\Omega))} \leq C.$$

Proceeding otherwise as in the proof of contractivity in Theorem 3.1, we therefore obtain

$$\|u^{\varepsilon_1} - u^{\varepsilon_2}\|_E^2 \leq C \|(\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * 1\|_{L^1(0,T)}$$

with a constant $C > 0$ that is independent of $\varepsilon_{1,2}$, from which the claim follows. $\qquad \square$

Observe that the main culprit for the reduced order of continuity in (4.1) (i.e., having $\|(\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * 1\|_{L^1(0,T)}^{1/2}$ instead of $\|(\mathfrak{K}_{\varepsilon_1} - \mathfrak{K}_{\varepsilon_2}) * 1\|_{L^1(0,T)}$) is the integration by parts with respect to space in (4.3). This approach was forced by a lack of uniform bound on $\Delta u_{tt}^{\varepsilon_2}$. We thus expect that the order can be improved in a more regular setting in terms of data that would lead to a uniform estimate on $\Delta u_{tt}^{\varepsilon}$ in (3.1).

Theorem 4.1 is the key to establishing the limiting behavior of solutions to (3.1) as $\varepsilon \searrow 0$ and, in particular, the convergence rates. As they will depend on the form of the kernel $\mathfrak{K}_\varepsilon$ and, in turn, its dependence on ε, we treat different classes of kernels separately.

4.1 The vanishing sound diffusivity limit with fractional-type kernels

We first discuss the setting $\mathfrak{K}_\varepsilon = \varepsilon \mathfrak{K}$. Recall that a representative example of this class of kernels (up to a constant) is

$$\mathfrak{K}_\varepsilon = \varepsilon g_\alpha,$$

where the Abel kernel is defined in (1.6) for $\alpha \in (0, 1)$ and $g_0 = \delta_0$. We will prove in Section 5 that this kernel indeed verifies assumption (A_2).

Corollary 4.2. *Under the assumptions of Theorem* 3.1 *with the kernel*

$$\mathfrak{K}_\varepsilon = \varepsilon \mathfrak{K}, \quad \varepsilon \in (0, \bar{\varepsilon})$$

*satisfying assumptions (**A₁**) and (**A₂**), the family of solutions* $\{u^\varepsilon\}_{\varepsilon \in (0,\bar\varepsilon)}$ *of* (3.1) *converges in the energy norm to the solution* $u \equiv u^0$ *of the initial boundary-value problem for the inviscid Westervelt equation*

$$\begin{cases} ((1 + 2ku)u_t)_t - c^2\Delta u = f & \text{in } \Omega \times (0,T), \\ u = 0 & \text{on } \partial\Omega \times (0,T), \\ (u, u_t) = (u_0, u_1), & \text{in } \Omega \times \{0\}, \end{cases} \tag{4.4}$$

at a linear rate

$$\|u - u^\varepsilon\|_E \lesssim \varepsilon.$$

Proof. In this setting, the limiting kernel is $\mathfrak{K}_0 = 0$, and it satisfies assumptions (**A₁**) and (**A₂**). By Theorem 4.1 and estimate (4.2), we then immediately have

$$\|u^\varepsilon - u\|_E \leq C\varepsilon\|\mathfrak{K}\|_{\mathcal{M}(0,T)}$$

for some $C > 0$, independent of ε, from which the statement follows. □

The limiting result above largely generalizes the one of Ref. [14, Theorem 4.1], where $\mathfrak{K}$ is the Dirac delta distribution δ_0. Here we allow for a large class of memory kernels, thus establishing the vanishing sound diffusivity limit for the Westervelt equation with general dissipation of time-fractional type.

4.2 The vanishing thermal relaxation time limit with Mittag-Leffler kernels

We now turn our attention to the kernels that were motivated by the presence of thermal relaxation in the heat flux laws of the propagation medium, and have the form

$$\mathfrak{K}_\varepsilon(t) = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b} \frac{1}{\varepsilon^b} t^{b-1} E_{a,b}\left(-\left(\frac{t}{\varepsilon}\right)^a\right), \quad a, b \in (0,1]. \tag{4.5}$$

Before continuing with the singular limit analysis, it is helpful to recall certain properties of the Mittag-Leffler functions, which can be found, for example, in Refs. [35], [36].

4.2.1 Properties of the Mittag-Leffler functions

We recall that the functions

$$t \mapsto E_{a,b}(-\lambda t^a), \ a \in [0,1], \ t > 0, \ \lambda > 0$$

are completely monotone for $b \geq a$ (and, in particular, non-negative); see [36, Corollary 3.2]. Furthermore, the following identity holds:

$$t^{b-1}E_{a,b}(-t^a) = \frac{\mathrm{d}}{\mathrm{d}t}\left(t^b E_{a,b+1}(-t^a)\right). \tag{4.6}$$

We also recall that the Laplace transform of the Mittag-Leffler functions is given by

$$\mathscr{L}\left[t^{b-1}E_{a,b}(-\lambda t^a)\right](z) = \frac{z^{a-b}}{z^a + \lambda}, \quad a, b > 0, \ \text{Re}(z) > 0, \ \lambda \geq 0; \tag{4.7}$$

see [36, Lemma 3.2]. We will additionally rely on the asymptotic behavior of Mittag-Leffler functions:

$$E_{a,b}(-x) \sim \frac{1}{\Gamma(b - a)\,x} \ \text{as } x \to \infty; \tag{4.8}$$

see, e.g., [36, Theorem 3.2].

In what follows, we intend to take the limit $\varepsilon \searrow 0$, while keeping $\tau_\theta > 0$ fixed. We treat the cases $a - b \leq 0$ and $a - b > 0$ (where additionally $\tau_\theta/\varepsilon > 0$ should be fixed) separately when discussing the limiting behavior.

4.2.2 Limiting behavior for $a - b \leq 0$

If $0 < a \leq b \leq 1$, we will prove that solutions u^ε of (3.1) converge to the solution u of the following time-fractional equation:

$$((1 + 2ku)u_t)_t - c^2\Delta u - \tau_\theta^{a-b}D_t^{a-b+1}\Delta u = f,$$

supplemented by the same boundary and initial conditions as in (3.1). Recall that

$$D_t^{a-b+1}\Delta u = g_{b-a}*\Delta u_t.$$

Note also that in case $a = b$, the limiting equation is strongly damped.

Proposition 4.1. *Let $\tau_\theta > 0$ be fixed. Under the assumptions of Theorem* 3.1, *the family of solutions* $\{u^\varepsilon\}_{\varepsilon \in (0,\bar{\varepsilon})}$ *of* (3.1) *with the kernel given by*

$$\mathfrak{K}_\varepsilon(t) = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b}\frac{1}{\varepsilon^b}t^{b-1}E_{a,b}\left(-\left(\frac{t}{\varepsilon}\right)^a\right) \quad \text{where } a - b \leq 0, \;\; a, b \in (0,1],$$

converges to the solution u *of*

$$\begin{cases} ((1 + 2ku)u_t)_t - c^2\Delta u - \mathfrak{K}_0*\Delta u_t = f & \text{in } \Omega \times (0, T), \\ u = 0 & \text{on } \partial\Omega \times (0, T), \\ (u, u_t) = (u_0, u_1), & \text{in } \Omega \times \{0\}, \end{cases} \tag{4.9}$$

with the kernel $\mathfrak{K}_0 = \tau_\theta^{a-b}g_{b-a}$ *in the following sense:*

$$\|u^\varepsilon - u\|_E \lesssim \|(\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1\|_{L^1(0,T)}^{1/2} \sim \varepsilon^{a/2} \quad \text{as } \varepsilon \searrow 0.$$

Proof. By Theorem 4.1 (up to checking that $\mathfrak{K}_0$ satisfies (**A₁**) and (**A₂**)), we have

$$\|u^\varepsilon - u\|_E \leq C\|(\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1\|_{L^1(0,T)}^{1/2}$$

with a constant $C > 0$ that is independent of ε. We next rely on the Laplace transform to further establish the asymptotic behavior of the right-hand side as $\varepsilon \searrow 0$. To this end, we use (4.7) and the formula

$$\mathscr{L}[\mathfrak{K}_0](z) = \tau_\theta^{a-b}\mathscr{L}[g_{b-a}](z) = \tau_\theta^{a-b}z^{a-b}.$$

Without loss of generality, we assume $\tau_\theta = 1$. Consider the Laplace transform of $(\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1$:

$$\mathscr{L}[(\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1](z) = (\mathscr{L}[\mathfrak{K}_\varepsilon] - \mathscr{L}[\mathfrak{K}_0])(z)\mathscr{L}[1](z)$$

$$= \left(\frac{z^{a-b}}{(\varepsilon z)^a + 1} - z^{a-b}\right)\frac{1}{z} = -\varepsilon^a\frac{z^{a-(1+b-a)}}{(\varepsilon z)^a + 1}.$$

From here we conclude that

$$((\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1)(t) = -t^{b-a}E_{a,1+b-a}\left(-\left(\frac{t}{\varepsilon}\right)^a\right).$$

On account of the non-negativity of the function $t \mapsto E_{a,1+b-a}\left(-\left(\frac{t}{\varepsilon}\right)^a\right)$ (which is ensured by $1 + b - a \geq a$ due to the assumptions on a and b), we find that

$$\|(\mathfrak{K}_\varepsilon - \mathfrak{K}_0)*1\|_{L^1(0,T)} = \int_0^T t^{b-a}E_{a,1+b-a}\left(-\left(\frac{t}{\varepsilon}\right)^a\right)dt$$

$$= T^{1+b-a}E_{a,2+b-a}\left(-\left(\frac{T}{\varepsilon}\right)^a\right).$$

Then by asymptotic properties of the Mittag-Leffler functions in (4.8), we have

$$\|(\mathfrak{K}_{\varepsilon} - \mathfrak{K}_0)*1\|_{L^1(0,T)} = T^{1+b-a}E_{a,2+b-a}\left(-\left(\frac{T}{\varepsilon}\right)^a\right)$$

$$\sim \frac{T^{1+b-a}}{\Gamma(2+b-2a)}\left(\frac{T}{\varepsilon}\right)^{-a} \qquad \text{as } \frac{T}{\varepsilon} \to \infty,$$

which concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

4.2.3 Limiting behavior for $a > b$

We next discuss the limiting behavior of solutions to (3.1) in the remaining case for the Mittag-Leffler kernels, which is $1 \geq a > b > 0$. To take the limit $\varepsilon \searrow 0$ of (3.1) with the kernel (4.5) when $a > b$, now we need the following additional assumption on τ_θ:

$$\tau_\theta = \tau_\theta(\varepsilon) \quad \text{and} \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b} = \rho^{a-b} = \text{constant}, \tag{4.10}$$

under which the kernels have the form

$$\mathfrak{K}_{\varepsilon}(t) = \rho^{a-b}\frac{1}{\varepsilon}\left(\frac{t}{\varepsilon}\right)^{b-1}E_{a,b}\left(-\left(\frac{t}{\varepsilon}\right)^a\right), \quad 1 \geq a > b > 0.$$

Proposition 4.2. *Under the assumptions of Theorem* 3.1 *and assumption* (4.10)*, the family of solutions* $\{u^{\varepsilon}\}_{\varepsilon\in(0,\bar{\varepsilon})}$ *of* (3.1) *with*

$$\mathfrak{K}_{\varepsilon}(t) = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b}\frac{1}{\varepsilon^b}t^{b-1}E_{a,b}\left(-\left(\frac{t}{\varepsilon}\right)^a\right), \qquad 1 \geq a > b > 0,$$

converges in the energy norm to the solution u of the inviscid problem (4.4) *at the following rate:*

$$\|u^{\varepsilon} - u\|_E \lesssim \|\mathfrak{K}_{\varepsilon}*1\|_{L^1(0,T)}^{1/2} \sim \varepsilon^{(a-b)/2} \quad \text{as } \varepsilon \searrow 0.$$

The statement of Proposition 4.2 can seem at first inspection unintuitive. Indeed, with the ratio τ_θ/ε constant as assumed in (4.10), we have

$$\mathfrak{K}_{\varepsilon}(t) = \rho^{a-b}\frac{1}{\varepsilon}\mathfrak{K}\left(\frac{t}{\varepsilon}\right),$$

where $\mathfrak{K}$ does not depend on ε. One would naively expect that as the rescaling parameter $\varepsilon \searrow 0$, the kernel would converge in the sense of distributions to a Dirac mass at 0. Results exist in this sense; see [7], [21]. However, in our setting, we do not require $\mathfrak{K} \geq 0$ which allows for the kernel $\mathfrak{K}_{\varepsilon}$ to converge to the zero function even if $\mathfrak{K} \neq 0$.

Proof. The statement follows by Theorem 4.1 and the asymptotic behavior of $\|(\mathfrak{K}_{\varepsilon} - \mathfrak{K}_0)*1\|_{L^1(0,T)} = \|\mathfrak{K}_{\varepsilon}*1\|_{L^1(0,T)}$ as $\varepsilon \searrow 0$. To establish the latter, we note that

$$\mathscr{L}[\mathfrak{K}_{\varepsilon}](z) = \rho^{a-b}\frac{(\varepsilon z)^{a-b}}{(\varepsilon z)^a + 1} \to 0 \qquad \text{as } \varepsilon \searrow 0$$

since $a > b$. Considering the identity

$$\mathscr{L}[\mathfrak{K}_{\varepsilon}*1](z) = \rho^{a-b}\frac{1}{z}\frac{(\varepsilon z)^{a-b}}{(\varepsilon z)^a + 1} = \rho^{a-b}\varepsilon\frac{(\varepsilon z)^{a-(1+b)}}{(\varepsilon z)^a + 1},$$

we find, using the scaling property of the Laplace transform, that

$$(\mathfrak{K}_{\varepsilon}*1)(t) = \rho^{a-b}\left(\frac{t}{\varepsilon}\right)^b E_{a,b+1}\left(-\left(\frac{t}{\varepsilon}\right)^a\right).$$

Now, using a change of variable, we have

$$\|\mathfrak{K}_\varepsilon * 1\|_{L^1(0,T)} = \varepsilon \rho^{a-b} \int\limits_0^{T/\varepsilon} |t^b E_{a,b+1}(-t^a)|\, dt.$$

Since $b + 1 \geq a$, the Mittag-Leffler function is completely monotone and thus non-negative. Furthermore, identity (4.6) and asymptotics (4.8) imply that

$$\varepsilon \rho^{a-b} \int\limits_0^{T/\varepsilon} t^b E_{a,b+1}(-t^a)\, dt = \varepsilon \rho^{a-b} \left(\frac{T}{\varepsilon}\right)^{b+1} E_{a,b+2}\left(-\left(\frac{T}{\varepsilon}\right)^a\right) \sim \rho^{a-b} \frac{T^{1+b-a}}{\Gamma(b+2-a)} \varepsilon^{a-b}, \tag{4.11}$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Remark 1. Assumption (4.10) means that τ_θ should converge to zero at least as fast as ε. Since (4.11) implies that

$$\varepsilon \rho^{a-b} \int\limits_0^{T/\varepsilon} t^b E_{a,b+1}(-t^a)\, dt \sim \tau_\theta^{a-b} \frac{T^{1+b-a}}{\Gamma(b+2-a)},$$

one could remove this assumption by letting $\tau_\theta \searrow 0$ and keeping $\varepsilon > 0$ (that is, the thermal relaxation time) fixed.

5 Verifying the uniform boundedness and coercivity assumptions for different classes of kernels

The remainder of this paper is devoted to proving that the different classes of kernels discussed in Section 2 and the limiting kernels $\mathfrak{K}_0$ from the previous section indeed satisfy assumptions (A_1) and (A_2) imposed in the well-posedness and singular limit analysis.

5.1 How to verify assumption (A_1)

For kernels that have the form $\mathfrak{K}_\varepsilon = \varepsilon \mathfrak{K}$ with $\varepsilon \in (0, \bar{\varepsilon})$ and $\mathfrak{K} \in L^1(0,T) \cup \{\delta_0\}$ independent of ε, we immediately obtain

$$\|\mathfrak{K}_\varepsilon\|_{\mathcal{M}(0,T)} \leq \bar{\varepsilon} \|\mathfrak{K}\|_{\mathcal{M}(0,T)}.$$

The Mittag-Leffler kernels are more interesting from the point of view of obtaining the uniform $L^1(0,T)$ bound. Since the kernels in the Gurtin–Pipkin forms of the Compte–Metzler laws can be written as

$$\mathfrak{K}_\varepsilon(t) = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b} \frac{1}{\varepsilon} \mathfrak{K}\left(\frac{t}{\varepsilon}\right), \tag{5.1}$$

where

$$\mathfrak{K}(t) = t^{b-1} E_{a,b}(-t^a), \tag{5.2}$$

we have

$$\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b} \|\mathfrak{K}\|_{L^1(0,T/\varepsilon)}. \tag{5.3}$$

As in the limiting analysis, we discuss different cases with respect to a and b to determine the behavior of the right-hand side term as $\varepsilon \searrow 0$.

– If $a = b$, then

$$\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} = \|\mathfrak{K}\|_{L^1(0,T/\varepsilon)}.$$

Therefore, as $\varepsilon \searrow 0$, we need to require that $\mathfrak{K} \in L^1(0, \infty)$. Recall that in this case the Mittag-Leffler function is completely monotone and therefore, in particular, non-negative. Furthermore, identity (4.6) implies

$$\|\mathfrak{K}\|_{L^1(0,r)} = \int_0^r \frac{\mathrm{d}}{\mathrm{d}t}\left(t^b E_{a,b+1}(-t^a)\right) \mathrm{d}t = r^b E_{a,b+1}(-r^a).$$

By the asymptotic behavior of Mittag-Leffler functions (4.8), we find that

$$t^b E_{a,b+1}(-t^a) \sim \frac{1}{\Gamma(b+1-a)} t^{b-a} \qquad \text{as } t \to \infty. \tag{5.4}$$

This asymptotic behavior implies that $\|\mathfrak{K}\|_{L^1(0,\infty)}$ with $\mathfrak{K}$ given in (5.2) is finite with $a = b$.

Among the kernels in Section 2 motivated by the physics of nonlinear acoustics, this assumption holds true for the exponential kernel (2.4) where $(a, b) = (1, 1)$ and the kernel coming from the GFE law where $(a, b) = (\alpha, \alpha)$; see Table 1.

– If $a < b$, asymptotics (5.4) is valid but the function tends to infinity. However, note that for $\tau_\theta > 0$ fixed, we have a uniform bound on $\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)}$ on account of

$$\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} = \left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b} \|\mathfrak{K}\|_{L^1(0,T/\varepsilon)} \sim \frac{1}{\Gamma(b+1-a)}\left(\frac{T}{\tau_\theta}\right)^{b-a} \qquad \text{as } \varepsilon \searrow 0.$$

Among the kernels discussed in Section 2, the assumption $a < b$ holds for the GFE II law, where $(a, b) = (\alpha, 1)$; see Table 1.

– If $a > b$, then in view of (5.3), inevitably

$$\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} \to +\infty \qquad \text{as } \varepsilon \searrow 0,$$

except in the trivial case $\mathfrak{K} \equiv 0$. This asymptotic behavior provides a motivation for assuming the ratio τ_θ/ε to be fixed in the limiting analysis when $a > b$ in (5.1), so that we have

$$\|\mathfrak{K}_\varepsilon\|_{L^1(0,T)} = \rho^{a-b} \|\mathfrak{K}\|_{L^1(0,T/\varepsilon)}.$$

The physical interpretation of this assumption is that one needs a scaling parameter, τ_θ, to match τ on the right-hand side of the fractional flux laws GFE I and GFE III; see Table 1. Thus if assumption (4.10) holds, we can make direct use of asymptotics (4.8), from which we conclude

$$\mathfrak{K}(t) \sim \frac{1}{\Gamma(b-a)} t^{b-1-a} \qquad \text{as } t \to \infty.$$

Therefore, $\|\mathfrak{K}\|_{L^1(0,\infty)} < \infty$.

5.2 How to verify assumption (A₂)

There are different ways of verifying coercivity assumption ($\mathbf{A_2}$); we discuss two of them here needed to tackle different classes of kernels we have seen so far.

– **Fourier approach to proving coercivity.** One approach of verifying ($\mathbf{A_2}$), which will be used for Mittag-Leffler kernels when $a \geq b$, is to employ Fourier analysis, similarly to Ref. [37, Lemma 2.3]. To this end, let us denote by $\overline{f}^t$ the extension of a function f by zero outside $\mathbb{R}\backslash(0, t)$, and by $\mathcal{F}$, the Fourier transform. The following identity holds for the Mittag-Leffler functions:

$$\mathcal{F}\left[\overline{x^{b-1}E_{a,b}(\lambda x^a)}^\infty\right](\omega) = \frac{1}{\sqrt{2\pi}}\frac{(\iota\omega)^{a-b}}{(\iota\omega)^a - \lambda}, \qquad \lambda \in \mathbb{R}.$$

Using this formula, we can obtain the Fourier transforms of the Mittag-Leffler kernels in Section 2; see Table 2. Let us also denote by $\star$ the Fourier convolution

$$(f \star g)(t) = \int_{\mathbb{R}} f(t-s)g(s)\,\mathrm{d}s,$$

and by an overline the complex conjugation. Furthermore, let

$$v(\omega) = (\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty)(\omega)(\mathcal{F}\overline{y}^t)(\omega), \qquad m(\omega) = \sqrt{2\pi}/(\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty)(\omega).$$

Then using Plancherel's theorem and the Fourier convolution theorem (see, e.g., [33, Theorems 1, 2, Section 4.3]), we have

$$\int_0^t \left((\mathfrak{K}_\varepsilon * y)(s), y(s)\right)_{L^2}\,\mathrm{d}s = \int_{\mathbb{R}} \left((\overline{\mathfrak{K}_\varepsilon}^\infty \star \overline{y}^t)(s), \overline{y}^t(s)\right)_{L^2}\,\mathrm{d}s$$

$$= \sqrt{2\pi} \int_{\mathbb{R}} \left((\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty)(\omega)(\mathcal{F}\overline{y}^t)(\omega), \overline{(\mathcal{F}\overline{y}^t)(\omega)}\right)_{L^2}\,\mathrm{d}\omega$$

$$= \int_{\mathbb{R}} \overline{m(\omega)}\|v(\omega)\|^2_{L^2(\Omega)}\,\mathrm{d}\omega$$

$$= \int_{\mathbb{R}} \mathrm{Re}(m(\omega))\|v(\omega)\|^2_{L^2(\Omega)}\,\mathrm{d}\omega,$$

where we have also used the fact that the left-hand side of the identity is real valued. Thus, since the real part of the Fourier transform of a real-valued function is even, if

$$\mathrm{Re}(m(\omega)) = \mathrm{Re}(\sqrt{2\pi}/\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty)(\omega) \geq 2\pi\,\tilde{C}_{a,b}, \quad \omega \in (0, \infty), \tag{5.5}$$

by Plancherel's theorem and the Cauchy–Schwarz inequality, we have

$$\int_0^t \left((\mathfrak{K}_\varepsilon * y)(s), y(s)\right)_{L^2}\,\mathrm{d}s \geq \tilde{C}_{a,b}\|\mathfrak{K}_\varepsilon * y\|^2_{L^2_t(L^2(\Omega))}$$

for all $t \in (0, T)$. Therefore, if condition (5.5) holds, the above estimate implies that assumption (**A₂**) holds with the constant

$$C_{\mathfrak{K}_\varepsilon} = \tilde{C}_{a,b}, \tag{5.6}$$

where we highlight the dependence on a and b. An inspection of the last column in Table 2 and using the change of variables $\varepsilon\omega \mapsto \omega$ yields

$$\tilde{C}_{a,b} = \inf_{\omega \in (0,\infty)} \mathrm{Re}(m(\omega)) = 2\pi\left(\frac{\tau_\theta}{\varepsilon}\right)^{b-a} \inf_{\omega \in (0,\infty)} \mathrm{Re}((\iota\omega)^b + (\iota\omega)^{b-a}).$$

The infimum is attained and it is bounded away from zero for $0 < b \leq a \leq 1$. Indeed, for $a = b$, we have

$$\tilde{C}_{a,a} = 2\pi \inf_{\omega \in (0,\infty)} \mathrm{Re}((\iota\omega)^b + 1) = 2\pi\left(\inf_{\omega \in (0,\infty)} \cos(b\pi/2)\omega^b + 1\right) = 2\pi,$$

Table 2: Fourier transforms of the Mittag-Leffler kernels discussed in Section 2.

$\mathfrak{K}_\varepsilon(t)$	$\sqrt{2\pi}\,\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty(\omega)$	$\frac{1}{2\pi}m(\omega) := 1/(\sqrt{2\pi}\mathcal{F}\overline{\mathfrak{K}_\varepsilon}^\infty(\omega))$
$\left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b}\frac{1}{\varepsilon^b}t^{b-1}E_{a,b}\left(-\left(\frac{t}{\varepsilon}\right)^a\right)$	$\left(\frac{\tau_\theta}{\varepsilon}\right)^{a-b}\frac{(\varepsilon\iota\omega)^{a-b}}{(\varepsilon\iota\omega)^a+1}$	$\left(\frac{\tau_\theta}{\varepsilon}\right)^{b-a}\left[(\varepsilon\iota\omega)^b + (\varepsilon\iota\omega)^{b-a}\right]$

while for $0 < b < a \leq 1$,

$$\tilde{C}_{a,b} = 2\pi \left(\frac{\tau_\theta}{\varepsilon} \right)^{b-a} \cos((b-a)\pi/2) \frac{a}{b} \left(\frac{(a-b)\cos((b-a)\pi/2)}{b\cos(b\pi/2)} \right)^{\frac{b-a}{b}}. \tag{5.7}$$

Therefore, we conclude that the Mittag-Leffler kernels (5.1) with $0 < b \leq a \leq 1$ satisfy assumption ($\mathbf{A_2}$) with the coercivity constant (5.6). Recalling our discussion in Section 2, this allows us to cover the kernels originating from the Compte–Metzler laws GFE, GFE I, and GFE III; see Table 1.

To verify ($\mathbf{A_2}$) for the Abel-type kernels and the kernel coming from law GFE II, we need an alternative approach.

–   **Non-Fourier approach for completely monotone kernels.** An alternative approach of verifying ($\mathbf{A_2}$) relies on a coercivity property from Ref. [38, Lemma B.1], combined with a result on resolvents of the first kind, cf. [39]. This approach will be applicable on completely monotone kernels, such as the Abel kernel and the Mittag-Leffler functions when $b \geq a$.

   We first revisit [38, Lemma B.1] to lower the regularity assumption on the memory kernel from $L^p(0, T)$ to $L^1(0, T)$.

Lemma 5.1. (Lemma B.1 in [38] revisited)*Given $T \in (0, \infty]$, let $\mathcal{K} \in L^1(0, T)$. Furthermore, assume that $\mathcal{K} \geq 0$ on $(0, T)$ and that for all $t_0 \in (0, T)$ it holds*

$$\mathcal{K} \in W^{1,1}(t_0, T), \quad \mathcal{K}'|_{[t_0, T]} \leq 0 \quad a.e.$$

Then

$$\int\limits_0^T (\mathcal{K} * y_t, y)_{L^2} \, ds \geq \frac{1}{2} \mathcal{K} * \|y\|_{L^2(\Omega)}^2 (T) - \frac{1}{2} \int\limits_0^T \mathcal{K}(s) \, ds \|y(0)\|_{L^2(\Omega)}^2 \tag{5.8}$$

for all $y \in W^{1,1}(0, T; L^2(\Omega))$.

Proof. Similarly to the proof of [38, Lemma B.1], we use an approximating sequence $\{\mathcal{K}_n\}_{n \in \mathbb{N}} \subseteq W^{1,1}(0, T)$ defined as

$$\mathcal{K}_n(t) = \max\{\tilde{\mathcal{K}}_n(t), w(t)/n\}$$

for some fixed positive weight function $w \in L^1(0, T)$ and

$$\tilde{\mathcal{K}}_n(t) := 1|_{[0,1/n]} \mathcal{K}(1/n) + 1|_{[1/n,T]}(t)\mathcal{K}(t).$$

The identity (5.8) can be shown to hold for $\mathcal{K}_n$ in place of $\mathcal{K}$ as in the proof of [38, Lemma B.1]. It thus remains to prove that $\mathcal{K}_n$ converges to $\mathcal{K}$ in the L^1 norm. To do so, we can rely on Lebesgue's dominated convergence theorem, using pointwise convergence of K_n to K and the fact that for any $n \in \mathbb{N}$, we have $0 < \mathcal{K}_n \leq \mathcal{K} + w \in L^1(0, T)$. □

Having in mind the verification of ($\mathbf{A_2}$), we next extend the reasoning from Ref. [19, Lemma 3.1] to a large family of completely monotone kernels $\mathfrak{K}$. Completely monotone functions belong to $C^\infty(0, \infty) \subset W^{1,1}([t_0, T])$ for all $t_0 > 0$ and thus satisfy the regularity assumptions of Lemma 5.1.

Let $\mathfrak{K} \in L^1_{\text{loc}}(\mathbb{R}^+)$ be completely monotone on $(0, \infty)$ and suppose $\mathfrak{K} > 0$ for almost all $t \in \mathbb{R}^+$. By [39, Theorem 5.5.4], $\mathfrak{K}$ has a resolvent of the first kind which is the sum of a point mass at zero and a completely monotone, locally integrable function:

$$\mathfrak{r} = A\delta_0 + \mathfrak{f} \quad \text{such that} \quad \mathfrak{K} * \mathfrak{r} = 1. \tag{5.9}$$

It is easy to see that A has to be non-negative. Moreover, $A = 0$ if and only if $\lim_{t \to 0} \mathfrak{K}(t) = \infty$. In other words, due to the complete monotonicity, the A-term in (5.9) appears if and only if $\mathfrak{K} \in L^\infty(0, T)$.

Lemma 5.2. *Let $\mathfrak{K} \in L^1_{loc}(\mathbb{R}^+)$ be completely monotone, nonconstant and a.e. positive ($\mathfrak{K} > 0$ for almost all $t \in \mathbb{R}^+$) and let*

$$\mathfrak{r} = A\delta_0 + \mathfrak{f}$$

be its resolvent of the first kind. Then

$$\int_0^t \int_\Omega (\mathfrak{K}*y)(s)\, y(s)\, dxds \geq \frac{1}{2}A\|\mathfrak{K}*y\|^2_{L^2(\Omega)}(t) + \mathfrak{f}(T)\,\frac{1}{2}\int_0^t \|\mathfrak{K}*y\|^2_{L^2(\Omega)}\, ds$$

for all $y \in L^2(0, t;\ L^2(\Omega))$ and $t \in (0, T)$, where for $T < \infty$ we have $\mathfrak{f}(T) > 0$.

Proof. We prove the inequality for $y \in C^\infty([0, T];\ L^2(\Omega))$; the statement follows then by density of $C^\infty([0, T])$ in $L^2(0, T)$. Applying Lemma 5.1 on $\mathcal{K} = \mathfrak{f}$ yields

$$\int_0^t (\mathfrak{r}*v_t, v)_{L^2}\, ds = \int_0^t \left((\mathfrak{f}*v_t, v)_{L^2} + A(v_t, v)_{L^2} \right)\, ds$$

$$\geq \frac{1}{2}A(\|v(t)\|^2_{L^2(\Omega)} - \|v(0)\|^2_{L^2(\Omega)}) + \frac{1}{2}\left(\mathfrak{f}*\|v\|^2_{L^2(\Omega)}(t) - \int_0^t \mathfrak{f}\, ds\|v(0)\|^2_{L^2(\Omega)} \right),$$

for any $v \in W^{1,1}(0, T)$. By picking $v = \mathfrak{K}*y$, we have that $(\mathfrak{K}*y)(0) = 0$ (since $y \in L^\infty(0, T)$), and subsequently that $\mathfrak{r}*v_t = y$. We then conclude that

$$\int_0^t (y, \mathfrak{K}*y)_{L^2}\, ds \geq \frac{1}{2}A\|\mathfrak{K}*y\|^2_{L^2(\Omega)}(t) + \frac{1}{2}\left(\mathfrak{f}*\|\mathfrak{K}*y\|^2_{L^2(\Omega)} \right)(t)$$

$$\geq \frac{1}{2}A\|\mathfrak{K}*y\|^2_{L^2(\Omega)}(t) + \inf_{\tau\in(0,T)}\mathfrak{f}(\tau)\,\frac{1}{2}\int_0^t \|\mathfrak{K}*y\|^2_{L^2(\Omega)}\, ds,$$

where by complete monotonicity $\inf_{\tau\in(0,T)}\mathfrak{f}(\tau) = \mathfrak{f}(T)$.

To prove that $\mathfrak{f}(T) > 0$, we reason by contradiction. Suppose that $\mathfrak{f}(T) = 0$. From the complete monotonicity of $\mathfrak{f}$ and its local absolute continuity it is clear that $\mathfrak{f}(t_1) = 0$ for all $t_1 \geq T$. We define $\tilde{T} = \inf\{t > 0 :\ \mathfrak{f}(t) = 0\}$ and use that then for all $t_1 > \tilde{T}$: $1 = \mathfrak{K}*\mathfrak{r}(t_1) = A\mathfrak{K}(t_1) + \int_0^{\tilde{T}} \mathfrak{K}(t_1 - s)\mathfrak{f}(s)\, ds$, thus

$$0 = \int_0^{\tilde{T}} \left(\mathfrak{K}(\tilde{T} - s) - \mathfrak{K}(t_1 - s) \right)\mathfrak{f}(s)\, ds + A[\mathfrak{K}(\tilde{T}) - \mathfrak{K}(t_1)], \tag{5.10}$$

where by minimality in the definition of $\tilde{T}$ we have $f > 0$ on $(0, \tilde{T})$. Moreover, by complete monotonicity of $\mathfrak{K}$, we have that $A[\mathfrak{K}(\tilde{T}) - \mathfrak{K}(t_1)] \geq 0$. Similarly, the factor $\mathfrak{K}(\tilde{T} - s) - \mathfrak{K}(t_1 - s)$ is non-negative, and we conclude from (5.10) that it actually vanishes. That is, for all $t_1 > \tilde{T}$, $s \in [0, \tilde{T}]$ the identity $\mathfrak{K}(\tilde{T} - s) = \mathfrak{K}(t_1 - s)$ holds, which by setting $s = \tilde{T}$, $t = t_1 - s$ implies $\mathfrak{K}(t) = \mathfrak{K}(0)$ for all $t > 0$. This contradicts our assumption that $\mathfrak{K}$ is nonconstant. □

Lemma 5.2 implies an estimate of the type $(\mathbf{A_2})$ in case $\mathfrak{K}_\varepsilon$ is not constant. In the special case of constant kernels $\mathfrak{K}_\varepsilon \equiv c_\varepsilon$ for some $c_\varepsilon > 0$, it is clear that assumption $(\mathbf{A_2})$ does not hold but instead the straightforward identity

$$\int_0^t \int_\Omega (\mathfrak{K}_\varepsilon*y)(s)\, y(s)\, dxds = \frac{1}{2c_\varepsilon}\|(\mathfrak{K}_\varepsilon*y)(t)\|^2_{L^2(\Omega)}$$

holds for all $y \in L^2(0, t; L^2(\Omega))$ (note that $\|(\mathfrak{K}_\varepsilon * y)(0)\|^2_{L^2(\Omega)} = 0$). The resulting PDE in that case is actually the inviscid Westervelt Equation (4.4) with c^2 and f replaced by $c^2 + c_\varepsilon$ and $f - c_\varepsilon \Delta u(0)$, respectively. Its well-posedness was already mentioned above to be a byproduct of Theorem 3.1. Also letting c_ε converge to some limit, e.g., $c_\varepsilon \to 0$ is technically feasible due to the uniform energy bounds from Theorem 3.1, but probably not of high practical interest.

With $b \geq a$, the Mittag-Leffler function is completely monotone and satisfies assumptions of Lemma 5.2. Thus, recalling the kernels discussed in Section 2, we can use this result to verify the remaining case of the GFE II kernel, where $(a, b) = (\alpha, 1)$. For this particular kernel, the resolvent is given by

$$\mathfrak{r} = \left(\frac{\tau_\theta}{\varepsilon}\right)^{1-\alpha}\left[\varepsilon\delta_0 + \frac{1}{\varepsilon^{\alpha-1}}g_\alpha\right].$$

One may indeed readily check that

$$\mathscr{L}\left[\left(\frac{\tau_\theta}{\varepsilon}\right)^{\alpha-1}\frac{1}{\varepsilon}E_{\alpha,1}\left(-\left(\frac{t}{\varepsilon}\right)^\alpha\right)*\mathfrak{r}\right](z) = 1/z.$$

Similarly, the function $\varepsilon\tau_\theta^{-\alpha}g_\alpha$ satisfies assumptions of Lemma 5.2 and its resolvent is given by $\frac{1}{\varepsilon}\tau_\theta^\alpha g_{1-\alpha}$.

5.2.1 Results of the verification

For convenience, we compile in Table 3 the results of the verification of assumptions (**A₁**) and (**A₂**) for memory kernels arising in the context of nonlinear acoustics. The last column in Table 3 provides the coercivity constant $C_{\mathfrak{K}_\varepsilon}$ in assumption (**A₂**) and another argument in favor of assuming the ratio τ_θ/ε to be constant for laws GFE I and III. In that case, it is straightforward to check using Table 3 that $C_{\mathfrak{K}_\varepsilon}$ can be bounded uniformly from below by some $C_{\mathbf{A_2}} > 0$ for $\varepsilon \in (0, \bar\varepsilon)$.

We point out that the theoretical framework developed in this paper, in particular pertaining to the continuity of the solution with respect to a parameter-dependent kernel, $\mathfrak{K}_\varepsilon$, in Theorem 4.1, can be easily used to study the limiting behavior of the equations with respect to other physical parameters of interest. For instance, Theorem 4.1 can be employed in studying the limiting behavior of Equation (3.1) with the Mittag-Leffler kernels listed in Table 3 as $\alpha \nearrow 1$. One can set $\varepsilon = 1 - \alpha$ to frame the study within the previous theory. The key remaining component of such a limiting analysis would be to prove that assumptions (**A₁**) and (**A₂**) hold uniformly for α in a neighborhood of 1^-.

Table 3: Kernels for flux laws discussed in Section 2 and the assumptions they satisfy; $\tilde{C}_{\alpha,2\alpha-1} > 0$ and $\tilde{C}_{1,\alpha} > 0$ are constants defined in (5.7) that can be made independent of ε (for GFE I and III, provided τ_θ/ε is constant).

Examples of kernels $\mathfrak{K}_\varepsilon$ in nonlinear acoustics	(**A₁**)	(**A₂**)	$C_{\mathfrak{K}_\varepsilon}$
$\varepsilon\tau_\theta^{-\alpha}g_\alpha$	✓	✓	$\frac{1}{\varepsilon}\tau_\theta^\alpha g_{1-\alpha}(T)$
$\frac{1}{\varepsilon}\exp\left(-\frac{t}{\varepsilon}\right)$	✓	✓	2π
GFE I: $\left(\frac{\tau_\theta}{\varepsilon}\right)^{1-\alpha}\frac{1}{\varepsilon^{2\alpha-1}}t^{2\alpha-2}E_{\alpha,2\alpha-1}(-(\frac{t}{\varepsilon})^\alpha)$	✓	✓	$\tilde{C}_{\alpha,2\alpha-1}$
GFE II: $\left(\frac{\tau_\theta}{\varepsilon}\right)^{\alpha-1}\frac{1}{\varepsilon}E_{\alpha,1}(-(\frac{t}{\varepsilon})^\alpha)$	✓	✓	$\tau_\theta^{1-\alpha}g_\alpha(T)$
GFE III: $\left(\frac{\tau_\theta}{\varepsilon}\right)^{1-\alpha}\frac{1}{\varepsilon^\alpha}t^{\alpha-1}E_{1,\alpha}(-(\frac{t}{\varepsilon}))$	✓	✓	$\tilde{C}_{1,\alpha}$
GFE: $\frac{1}{\varepsilon^\alpha}t^{\alpha-1}E_{\alpha,\alpha}(-(\frac{t}{\varepsilon})^\alpha)$	✓	✓	2π

6 Conclusion and outlook

In this work, we have investigated a family of nonlinear Westervelt equations with general dissipation of Djrbashian–Caputo-derivative type. We rigorously studied them in terms of the local well-posedness and their limiting behavior with respect to the parameter ε, which may be physically interpreted as the sound diffusivity or the thermal relaxation time. As we have seen, the limiting behavior is influenced by the dependence of the memory kernel $\mathfrak{K}_\varepsilon$ on ε, and for this reason, we have considered different classes of kernels separately. The framework developed here and leading to Theorem 4.1 allows extending the limiting study to other parameters of interest, as long as the parameter-dependent kernels satisfy assumptions $(\mathbf{A_1})$ and $(\mathbf{A_2})$.

Future work will be concerned with the study of the limiting behavior of quasilinear equations of Kuznetsov and Blackstock type given in (2.7) and (2.8), respectively, where we expect that the kernel assumptions will need to be further tailored to the needs of each of these models. It would also be of interest to extend the limiting analysis to settings with practical boundary conditions, such as Neumann or absorbing boundary conditions. We expect that the main ideas from the current energy arguments can be adjusted but with a higher level of technicality due to the lack of Poincaré-Friedrichs inequality. Finally, the important question of whether global results in time can be established when the limiting problems contain dissipation (such as (4.9)) remains open.

Appendix A: Unique solvability of the nonlocal semi-discrete problem

We present here the proof of existence of a unique semi-discrete approximation of (3.3). As in Ref. [14], we may approximate the solution by

$$u^{\varepsilon,n}(x,t) = \sum_{i=1}^{n} \xi_i^n(t) v_i(x),$$

where $\{v_i\}_{i=1}^{\infty}$ are smooth eigenfunctions of the Dirichlet–Laplacian operator.

With $V_n = \text{span}\{v_1, \ldots, v_n\}$, the semi-discrete problem is given by

$$((1 + 2k\phi)u_{tt}^{\varepsilon,n} - c^2 \Delta u^{\varepsilon,n} - \mathfrak{K}_\varepsilon * \Delta u_t^{\varepsilon,n} + 2k\phi_t u_t^{\varepsilon,n} - f, v_j) = 0$$

for all $j = 1, \ldots, n$, with approximate initial conditions $(u^{\varepsilon,n}, u_t^{\varepsilon,n})|_{t=0} = (u_0^n, u_1^n)$ taken as $L^2(\Omega)$ projections of (u_0, u_1) onto V_n. With $\boldsymbol{\xi} = \begin{bmatrix} \xi_1^n & \xi_2^n & \ldots & \xi_n^n \end{bmatrix}^T$, the approximate problem can be rewritten in matrix form

$$\mathbb{M}_{\mathfrak{m}}(t)\boldsymbol{\xi}_{tt} + c^2 \mathbb{D}\boldsymbol{\xi} + \mathbb{D}\,\mathfrak{K}_\varepsilon * \boldsymbol{\xi}_t + \mathbb{M}_{\mathfrak{m}_t}(t)\boldsymbol{\xi}_t = \boldsymbol{f},$$

where the entries of matrices $\mathbb{M}_{\mathfrak{m}}(t) = [\mathbb{M}_{\mathfrak{m},ij}]$, $\mathbb{M}_{\mathfrak{m}_t}(t) = [\mathbb{M}_{\mathfrak{m}_t,ij}]$, $\mathbb{D} = [\mathbb{D}_{ij}]$, and vector $\boldsymbol{f} = [\boldsymbol{f}_j]$ are given by

$$\mathbb{M}_{\mathfrak{m},ij}(t) = ((1 + 2k\phi)v_i, v_j)_{L^2}, \qquad \mathbb{M}_{\mathfrak{m}_t,ij}(t) = (2k\phi_t v_i, v_j)_{L^2},$$

$$\mathbb{D}_{ij} = -(\Delta v_i, v_j)_{L^2}, \qquad\qquad \boldsymbol{f}_i(t) = (f(t), v_i)_{L^2}.$$

If we introduce the vectors of coordinates of the approximate initial data in the basis:

$$\boldsymbol{\xi}_0 = \begin{bmatrix} \xi_{0,1}^n & \xi_{0,2}^n & \cdots \xi_{0,n}^n \end{bmatrix}^T, \quad \boldsymbol{\xi}_1 = \begin{bmatrix} \xi_{1,1}^n & \xi_{1,2}^n & \cdots \xi_{1,n}^n \end{bmatrix}^T,$$

then by setting $\boldsymbol{\mu} = \boldsymbol{\xi}_{tt}$, we have

$$\boldsymbol{\xi}_t(t) = 1*\boldsymbol{\mu} + \boldsymbol{\xi}_1, \quad \boldsymbol{\xi}(t) = \boldsymbol{\xi}_0 + t\boldsymbol{\xi}_1 + 1*1*\boldsymbol{\mu}.$$

Therefore, the semi-discrete problem can be rewritten as

$$\mathbb{M}_{\mathfrak{m}}(t)\boldsymbol{\mu} + c^2 \mathbb{D}(\boldsymbol{\xi}_0 + t\boldsymbol{\xi}_1 + 1*1*\boldsymbol{\mu}) + \mathbb{D}\,\mathfrak{K}_\varepsilon * (1*\boldsymbol{\mu} + \boldsymbol{\xi}_1) + \mathbb{M}_{\mathfrak{m}_t}(t)(1*\boldsymbol{\mu} + \boldsymbol{\xi}_1) = \boldsymbol{f}.$$

Since $\mathbb{M}_{\mathfrak{m}} \in L^\infty(0, T)$ is positive definite due to assumption (3.2), the semi-discrete problem can be further seen as a system of Volterra integral equations:

$$\boldsymbol{\mu} + \boldsymbol{K}*\boldsymbol{\mu} = \tilde{\boldsymbol{f}},$$

with

$$\boldsymbol{K} = \mathbb{M}_{\mathfrak{m}}^{-1}(t)\big\{c^2 \mathbb{D}1*1 + \mathbb{D}\,\mathfrak{K}_\varepsilon *1 + \mathbb{M}_{\mathfrak{m}_t}(t)\big\}$$

and

$$\tilde{\boldsymbol{f}} = \mathbb{M}_{\mathfrak{m}}^{-1}(t)\big\{-c^2 \mathbb{D}(\boldsymbol{\xi}_0 + t\boldsymbol{\xi}_1) - \mathbb{D}\,\mathfrak{K}_\varepsilon *\boldsymbol{\xi}_1 - \mathbb{M}_{\mathfrak{m}_t}(t)\boldsymbol{\xi}_1 + \boldsymbol{f}(t)\big\}.$$

By the existence theory for systems of Volterra integral equations of the second kind [39, Ch. 2, Theorem 4.5], there is a unique solution $\boldsymbol{\mu} \in L^\infty(0, T)$. Combined with initial data, from $\boldsymbol{\xi}_{tt} = \boldsymbol{\mu}$, we can then conclude that there exists a unique $\boldsymbol{\xi} \in W^{2,\infty}(0, T)$ and thus $u^{n,\varepsilon} \in W^{2,\infty}(0, T; V_n)$.

References

[1] M. E. Gurtin and A. C. Pipkin, "A general theory of heat conduction with finite wave speeds," *Arch. Ration. Mech. Anal.*, vol. 31, pp. 113−126, 1968.

[2] P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Am.*, vol. 35, pp. 535−537, 1963.

[3] M. Caputo, "Linear models of dissipation whose Q is almost frequency independent − II," *Geophys. J. Int.*, vol. 13, pp. 529−539, 1967.

[4] M. M. Djrbashian, *Integral Transformations and Representation of Functions in a Complex Domain [in Russian]*, Moscow, Nauka, 1966.

[5] F. Prieur and S. Holm, "Nonlinear acoustic wave equations with fractional loss operators," *J. Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1125−1132, 2011.

[6] M. J. Lighthill, "Viscosity effects in sound waves of finite amplitude," in *Surveys in Mechanics*, Batchelor and Davies, Eds., London, Cambridge University Press, 1956.

[7] M. Conti, V. Pata, and M. Squassina, "Singular limit of dissipative hyperbolic equations with memory," in *Conference Publications*, vol. 2005, American Institute of Mathematical Sciences, 2005, p. 200.

[8] I. Lasiecka, "Global solvability of Moore−Gibson−Thompson equation with memory arising in nonlinear acoustics," *J. Evol. Equ.*, vol. 17, pp. 411−441, 2017.

[9] B. Kaltenbacher and I. Lasiecka, "Global existence and exponential decay rates for the Westervelt equation," *Discrete Continuous Dyn. Syst.-S*, vol. 2, no. 3, pp. 503−523, 2009.

[10] B. Kaltenbacher, I. Lasiecka, and R. Marchand, "Wellposedness and exponential decay rates for the Moore-Gibson−Thompson equation arising in high intensity ultrasound," *Control Cybern.*, vol. 40, no. 4, pp. 971−988, 2011.

[11] S. Meyer and M. Wilke, "Optimal regularity and long-time behavior of solutions for the Westervelt equation," *Appl. Math. Opt.*, vol. 64, pp. 257−271, 2011.

[12] W. Dörfler, H. Gerner, and R. Schnaubelt, "Local well-posedness of a quasilinear wave equation," *Appl. Anal.*, vol. 95, no. 9, pp. 2110−2123, 2016.

[13] M. Kaltenbacher, *Numerical Simulation of Mechatronic Sensors and Actuators*, vol. 3, Berlin, Springer, 2014.

[14] B. Kaltenbacher and V. Nikolić, "Parabolic approximation of quasilinear wave equations with applications in nonlinear acoustics," *SIAM J. Math. Anal.*, vol. 54, no. 2, pp. 1593−1622, 2022.

[15] B. Kaltenbacher, I. Lasiecka, and M. K. Pospieszalska, "Well-posedness and exponential decay of the energy in the nonlinear Jordan−Moore−Gibson−Thompson equation arising in high intensity ultrasound," *Math. Model Methods Appl. Sci.*, vol. 22, p. 1250035, 2012.

[16] K. Mizohata and S. Ukai, "The global existence of small amplitude solutions to the nonlinear acoustic wave equation," *J. Math. Kyoto Univ.*, vol. 33, no. 2, pp. 505−522, 1993.

[17] A. Tani, "Mathematical analysis in nonlinear acoustics," *AIP Conf. Proc.*, vol. 1907, p. 020003, 2017.

[18] B. Kaltenbacher, "Mathematics of nonlinear acoustics," *Evol. Equ. Control Theor.*, vol. 4, pp. 447−491, 2015.

[19] B. Kaltenbacher and W. Rundell, "On an inverse problem of nonlinear imaging with fractional damping," *Math. Comput.*, vol. 91, pp. 245−276, 2022.

[20] K. Baker, L. Banjai, and M. Ptashnyk, "Numerical analysis of a time-stepping method for the Westervelt equation with time-fractional damping," arXiv preprint arXiv:2210.16349, 2022.

[21] M. Conti, V. Pata, and M. Squassina, *Singular Limit of Differential Systems with Memory*, Indiana University mathematics journal, 2006, pp. 169−215.

[22] A. Kubica, K. Ryszewska, and M. Yamamoto, *Time-Fractional Differential Equations: A Theoretical Introduction*, Singapore, Springer, 2020.

[23] I. Podlubny, *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*, Elsevier, 1998.

[24] Y. Povstenko, "Fractional Cattaneo-type equations and generalized thermoelasticity," *J. Therm. Stresses*, vol. 34, pp. 97−114, 2011.

[25] Y. Povstenko, *Fractional Thermoelasticity*, Springer, 2015.

[26] S. Holm, *Waves with Power-Law Attenuation*, Springer, 2019.

[27] C. Cattaneo, "Sur une forme de l'équation de la chaleur éliminant le paradoxe d'une propagation instantanée," *C. R. Acad. Sci. Paris*, vol. 247, pp. 431−433, 1958.

[28] A. Compte and R. Metzler, "The generalized Cattaneo equation for the description of anomalous transport processes," *J. Phys. A: Math. Gen.*, vol. 30, pp. 7277−7289, 1997.

[29] P. M. Jordan, "Second-sound phenomena in inviscid, thermally relaxing gases," *Discrete Continuous Dyn. Syst.-B*, vol. 19, no. 7, pp. 2189−2205, 2014.

[30] D. T. Blackstock, "Approximate equations governing finite-amplitude sound in thermoviscous fluids," tech. rep., General Dynamics/Electronics Rochester NY, 1963.

[31] V. P. Kuznetsov, "Equations of nonlinear acoustics," *Sov. Phys. Acoust.*, vol. 16, pp. 467−470, 1970.

[32] B. Kaltenbacher and V. Nikolić, "Time-fractional Moore−Gibson−Thompson equations," *Math. Model Methods Appl. Sci.*, vol. 32, pp. 965−1013, 2022.

[33] L. C. Evans, *Partial Differential Equations, Graduate Studies in Mathematics*, vol. 2, AMS, 2010.

[34] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, vol. 369, American Mathematical Soc., 2010.

[35] R. Gorenflo, F. Mainardi, and S. Rogosin, "Mittag-Leffler function: properties and applications," in *Handbook of Fractional Calculus with Applications*, vol. 1, Berlin, Germany, de Gruyter, 2019, pp. 269−296.

[36] B. Jin, *Fractional Differential Equations*, Cham, Switzerland, Springer, 2021.

[37] P. P. B. Eggermont, "On Galerkin methods for Abel-type integral equations," *SIAM J. Numer. Anal.*, vol. 25, no. 5, pp. 1093−1117, 1987.

[38] B. Kaltenbacher, U. Khristenko, V. Nikolić, M. L. Rajendran, and B. Wohlmuth, "Determining kernels in linear viscoelasticity," *J. Comput. Phys.*, vol. 464, p. 111331, 2022.

[39] G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra Integral and Functional Equations*, vol. 34, Cambridge University Press, 1990.