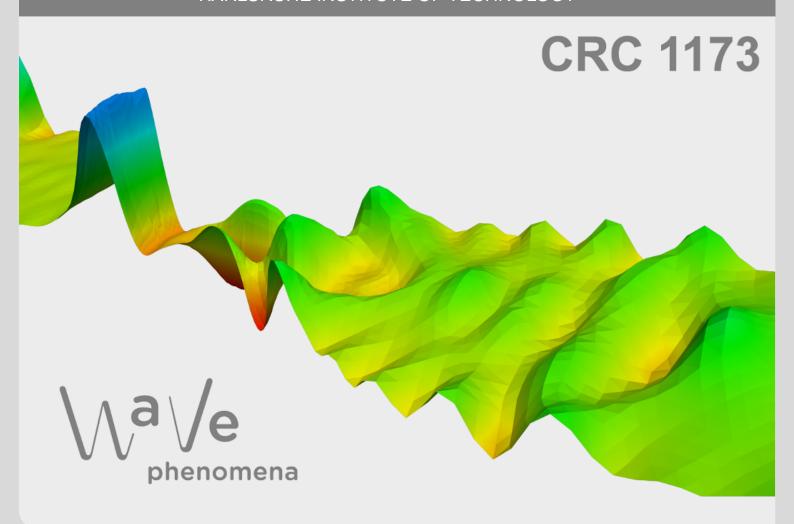# On a Parabolic-Hyperbolic Filter for Multicolor Image Noise Reduction

Valerii Maltsev, Michael Pokojovy

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173

Wave
phenomena

Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Funded by

DFG

On a Parabolic-Hyperbolic Filter for Multicolor Image Noise Reduction

Valerii Maltsev*   Michael Pokojovy†

May 4, 2016

Abstract

We propose a novel PDE-based anisotropic filter for noise reduction in multicolor images. It is a generalization of Nitzberg & Shiota's (1992) model being a hyperbolic relaxation of the well-known parabolic Perona & Malik's filter (1990). First, we consider a 'spatial' molifier-type regularization of our PDE system and exploit the maximal L^2-regularity theory for non-autonomous forms to prove a well-posedness result both in weak and strong settings. Again, using the maximal L^2-regularity theory and Schauder's fixed point theorem, respective solutions for the original quasilinear problem are obtained and the uniqueness of solutions with a bounded gradient is proved. Finally, the long-time behavior of our model is studied.

Key words: image processing, nonlinear partial differential equations, weak solutions, strong solutions, maximal regularity

AMS: 35B30, 35D30, 35D35, 35G61, 35M33, 65J15

1   Introduction

Image processing (also referred to as digital image processing) is one of central tasks in the image science. It includes, but is not limited to denoising, deblurring, decomposition or segmentation of images with appropriate edges (cf. [13, p. 5]). Since the presence of noise is unavoidable due to the image formation process, recording and/or transmission (cf. [32, p. 259]), in practice, a noise reduction technique must be applied before any further processing steps can reasonably be performed.

One of the earlier systematic theories dates back to Marr and Hildreth [25] (cf. [16, p. 182]) and incorporates a low-pass filtering as a noise reduction tool. For a detailed overview of the historical literature, we refer the reader to the comprehensive article by Alvarez et al. [1]. After a decade of gradual improvements and developments by Canny [14], Witkin [38] and many other authors, the field has been revolutionized by Perona

---

*Department of Cybernetics, Kyiv National Taras Shevchenko University, Ukraine maltsev.valerii@gmail.com

†Department of Mathematics, Karlsruhe Institute of Technology, Karlsruhe, Germany michael.pokojovy@kit.edu

& Malik [29] in early 90s, when they proposed their famour anisotropic Perona & Malik image filter. Their development marked a new era in image processing – the era of (time-dependent) partial differential equations (PDEs).

With $G \subset \mathbb{R}^d$ denoting the domain occupied by a monochromic image and $u(t, \mathbf{x})$ standing for the (grayscale) image intensity at time $t \geq 0$ and pixel $\mathbf{x} \in G$, in its modern formulation (cf. [24, 175–190]), Perona & Malik's PDE reads as

$$\begin{aligned} \partial_t u &= \mathrm{div}\left(\mathbf{g}(\nabla u)\nabla u\right) \text{ in } (0, \infty) \times G, \\ \mathbf{g}(\nabla u) \cdot \mathbf{n} &= 0 \text{ on } (0, \infty) \times \Gamma, \\ u(0, \cdot) &= \tilde{u}^0 \text{ in } \Omega, \end{aligned} \tag{1}$$

where $\mathbf{n}$ stands for the outer unit normal vector to $\Gamma := \partial\Omega$, $\mathbf{g}$ is a nonlinear response or diffusivity function (scalar or matrix-valued) and $\tilde{u}^0$ denotes the original noisy image. Whereas, as observed by Witkin [38], Equation (1) leads to a linear Gaussian low-pass filter if g is constant, an appropriate choice of a nonlinear diffusivity turns out to be particularly beneficial for the edge preservation. Selecting $\mathbf{g}$ to vanish as $|\nabla u| \to \infty$, the diffusion slows down at the edges thus preserving their localization. For small values of $|\nabla u|$, the diffusion is active and tends to smoothen around such points (cf. [16, p. 183]). Typical choices of $\mathbf{g}$ can be found in [24, Table 1, p. 178], [37, Section 1.3.3], for example,

$$\mathbf{g}(\mathbf{s}) = \left(1 + \frac{|\mathbf{s}|}{\lambda}\right)^{-1} \mathbf{I} \text{ for some } \lambda > 0.$$

In the analytic sense, it can be observed the Equation (1) is ill-posed due to its connection with the reverse heat equation (cf. [37, pp. 15–19]). Surprisingly, numerical discretizations have been observed to be stable ([4, pp. 20–21], [21]), though undesirable staircaising effects still occur sometimes (cf. [24, p. 176]). Some of the numerical studies have though been critically perceived by other authors (viz. [16, p. 185]). Another major drawback of Perona & Malik's filter is that it can break down if being applied to images contaminated with an irregular noise such as the white noise. Indeed, in such situations, ∇u becomes unbounded almost everywhere in G and the diffusion collapses (see [16, p. 183]). Unfortunately, despite of numerous numerical results, no rigorous analytic theories are known in the literature for Equation (1).

As an alternative for Equation (1), Catté et al. [16] proposed to consider a space-convolution regularization called the 'selecting smoothing' given by

$$\begin{aligned} \partial_t u &= \mathrm{div}\left(\mathbf{g}(\nabla_\sigma u)\nabla u\right) \text{ in } (0, \infty) \times G, \\ \mathbf{g}(\nabla_\sigma u) \cdot \mathbf{n} &= 0 \text{ on } (0, \infty) \times \Gamma, \\ u(0, \cdot) &= \tilde{u}^0 \text{ in } \Omega, \end{aligned} \tag{2}$$

where $\nabla_\sigma u$ ($\sigma > 0$) denotes the gradient operator applied to the convolution of u with a multiple of Gaussian pdf in space (see Section 2.1 below). Heuristically, the convolution is meant to play the role of a low-pass filter which iteratively smoothes the image at the scale of $t^{1/2}$ before recomputing the diffusivity matrix. Under a C^∞-smoothness assumption on the scalar function $\mathbf{g}$, for $\tilde{u}^0 \in L^2(G)$, Catté et al. [16] proved an existence and uniqueness theorem for Equation (2) in the class of weak solutions $H^1\big(0, T; L^2(G)\big) \cap L^2\big(0, T; H^1(G)\big)$

together with a C^∞-regularity of solutions in $(0,T) \times G$ for any $T > 0$. They also developed a finite-difference numerical scheme and presented some illustrations of its performance. As later discovered by Amann [3, p. 1030], Equation (2) results in smoothing of sharp edges and, therefore, produces unwanted blurring effects. A more detailed discussion and some numerical illustrations can be found in [4, Sections 1 and 2].

To overcome the smearing effect of Equation (2), Amann [3, 4] studied a memory-type regularization of Perona & Malik's equation (1) given by

$$\partial_t u - \operatorname{div}\left(g\Big(\int_0^t \theta(t-s) \big| \nabla u(s,\cdot) \big|^2 \mathrm{d}s \Big) \nabla u \right) = 0 \text{ in } (0,\infty) \times G,$$
$$\frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } (0,\infty) \times \Gamma, \tag{3}$$
$$u(0,\cdot) = u^0 \text{ in } \Omega,$$

where $\theta \in L^s_{\mathrm{loc}}(0,\infty; \mathbb{R}_+)$ for some $s > 1$. For C^2-domains G (which rules out rectangular images), $1 < p, q < 1$ such that $\frac{2}{p} + \frac{d}{q} < 1$ and $g \in C^{2-}(\mathbb{R}^d, (0,\infty))$ with C^{2-} denoting the space of functions with locally bounded difference quotients up to order 2, the initial data

$$u^0 \in H^{2,q}_{\mathrm{Neu}} := \Big\{ u \in H^{2,q}(G) \,\big|\, \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \Gamma \Big\}$$

were shown to admit a unique strong solution

$$u \in H^{1,p}\big(0,T; L^q(G)\big) \cap L^p\big(0,T; H^{2,q}_{\mathrm{Neu}}(G)\big)$$

with a maximal existence time $T^* > 0$, which is even infinite if θ has a compact support in $(0,\infty)$. Moreover, it has been proved the solution continuously depends on the data in the respective topologies and a maximum principle for u has been shown, etc. The proof is based on the maximal L^p-regularity theory. A generalization of Equation (3) has also been studied and the results of numerical experiments have been presented.

An abstract linear version of Equation (3) was studied by Prüss in [30] and Zacher in [39]. We refer the reader to the fundamental monograph [31] by Prüss for further details on this and similar problems.

Cottet & El Ayyadi [17] studied the initial-boundary value problem

$$\partial_t u - \operatorname{div}\big(\mathbf{L}\nabla u\big) = 0 \text{ in } (0,\infty) \times G$$
$$\partial_t \mathbf{L} + \mathbf{L} = \mathbf{F}(\nabla_\sigma u) \text{ in } (0,\infty) \times G, \tag{4}$$
$$u(0,\cdot) = u^0, \quad \mathbf{L}(0,\cdot) = \mathbf{L}^0 \text{ in } G$$

together with the periodic boundary conditions for u for the case $d = 2$. Here, $\mathbf{F}$ is a function mapping $\mathbb{R}^d$ into the space of positive semidefinite $(d \times d)$-matrices, $\sigma > 0$ and $\mathbf{L}^0$ is uniformly positive definite. Equation (4) has first been proposed in a similar form and without any mathematial justification by Nitzberg & Shiota in [27]. For the initial data $(u^0, \mathbf{L}^0)^T \in L^\infty(G) \times \big(H^1(G) \cap L^\infty(G)\big)^{d \times d}$, Cottet & El Ayyadi [17] showed the existence of a unique solution

$$u \in L^2\big(0,T; H^1(G)\big) \cap L^\infty\big(0,T; L^\infty(G)\big), \quad \mathbf{L} \in L^\infty\Big(0,T; \big(H^1(G) \cap L^\infty(G)\big)^{d \times d}\Big)$$

for any $T > 0$, which, moreover, continuosly depends on the data in a certain topology. The proof is based on a convolution-like time discretization and *a priori* estimates. The choice of parameters has been discussed. A finite difference scheme together with numerical examples have been presented and a connection to a neural network has been established.

Belahmidi [9, Chapter 4] and Belahmidi & Chambolle [10] studied a modification of Equation (4) reading as

$$
\begin{aligned}
u_t &= \operatorname{div}\big(g(v)\nabla u\big) \text{ in } (0,\infty)\times G,\\
v_t + v &= F\big(|\nabla u|\big) \text{ in } (0,\infty)\times G,\\
\frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } (0,\infty)\times \Gamma,\\
u(0,\cdot) = u^0,\quad v(0,\cdot) &= v^0 \text{ in } \Omega
\end{aligned}
\tag{5}
$$

where g, F are scalar C^1-functions such that g is positive non-increasing and F is bounded together with its first derivative. The main disadvantage of Equation (5) over Equation (4) is that the former is not genuinely anisotropic in sense of [37, Section 1.3.3]. Belahmidi & Chambolle [10] developed a semi-implicit space-time finite difference scheme for Equation (4) and proved a discrete maximum principle for u implying the unconditional stability of their scheme. For the initial data $(u^0, v^0)^T \in \big(H^1(G)\cap L^\infty(G)\big)\cap\big(H^1(G)\cap L^\infty(G)\big)$ with $v^0 \geq 0$, the sequence of numerical solutions was shown to subconverge to a 'weak' solution

$$
(u,v)^T \in \Big(H^1\big((0,T)\times\Omega\big)\cap L^\infty\big((0,T)\times G\big)\Big)^2 \text{ for any } T > 0
$$

in the norm of $\Big(L^2\big(0,T;H^1(G)\big)\Big)^2$ as the lattice size goes to 0, whence an existence theorem for Equation (5) follows. As pointed out by Amann [4, p. 20], their proof is only valid in 2D. For a Hölder-space treatment of Equation (5), we refer the reader to Belahmidi's PhD thesis [9, Chapter 4], for which the author assumes, in particular,

$$
(u^0, v^0)^T \in C^{2,\alpha}(\bar{G})\times C^{1,\alpha}(\bar{G}) \text{ and } \Omega \in C^{2,\alpha} \text{ for } \alpha > 0,
$$

thus ruling out both rectangular images and rough noise patterns.

Equation (5) shares a certain degree of similarity with the equations of compressible and incompressible fluids. Recently, Hieber & Murata [22] studied a fluid-rigid interaction problem for a compressible fluid and used the maximal L^p-regularity theory to prove the local well-posedness. For an overview on the recent developments in the theory of parabolic systems we refer the reader to the same paper [22].

For the sake of completeness, one should also mention the vast literature studying various image filters incorporating the total variation functional as originally proposed by Rudin et al. [32], which turned out to perform particularly well in practice. Omitting the time-independet case (cf. Remark 1 below), the total variation counterpart of Perona & Malik's Equation (1) is given by

$$
u_t = \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) \text{ in } (0,\infty)\times G
$$

together with appropriate boundary conditions, where $\frac{\nabla u}{|\nabla u|}$ formally denotes the gradient/subdifferential of the total variation functional evaluated at u. Without being exhaustive, we refer the reader to the well-posedness and long-time behavior studies [5, 6, 12] and the references therein.

In the present paper, we revisit Equation (4). In Section 2, we derive a multicolor, genuinely anisotropic generalization of Equation (4) and discuss the choice of parameters for our new image filter. In Section 3, we present a well-posedness theory for the multicolor version of Equation (4) for $\sigma > 0$. In contrast to [17], we obtain a more regular solution under a weaker data regularity assumption. In Section 4, we consider the limiting case $\sigma = 0$. First, we prove the existence of mild and/or strong solutions using the classical variational theory for parabolic equations. Again, our approach requires less regularity than in the earlier work [10] and is valid in any space dimension. Under an additional assumption, we further prove the solutions are unique and continuously depend on the data. Next, we study the long-time behavior of our model and prove the exponential stability under a uniform positive definiteness condition on the diffusivity function. In the appendix Section A, we briefly summarize the classical maximal L^2-regularity for non-autonomous forms as well as its recent improvements.

2 Filter Description

In this section, we present a multicolor generalization of the monochromic PDE image filter proposed by Nitzberg & Shiota [27] and further developed by Cottet & El Ayyadi [17]. Our filter is more comprehensive than the monochromic one since it takes into account possible local correclactions between the color components. Besides, we provide some geometric intuition and a connection to diffusion processes to justify the logic of our filter.

2.1 PDE Based Image Filtering

Let G be a bounded domain of $\mathbb{R}^d$ ($d \in \mathbb{N}$) with $\mathbf{n} \colon \partial G \to \mathbb{R}^d$ standing for the unit outer normal vector. Typically, $d = 2$ and $G = (0, L_1) \times (0, L_2)$. Let $\mathbf{u}^0 \colon G \to \mathbb{R}^k$ with $\mathbf{u}^0 = (u_1^0, \ldots, u_k^0)^T$ denote initial color intensity of the image at point $\mathbf{x} = (x_1, x_2, \ldots, x_d)^T \in G$ measured with respect to an additive k-color space (e.g., the RGB space with $k = 3$).

In most practical situations, not the original image $\mathbf{u}^0$ but a corrupted version of it, say, $\tilde{\mathbf{u}}^0$ is known. Various pollution scenarios can occur ranging from noise effects and blurring to missing parts, etc. Here, we want to restrict ourselves to the situation that $\mathbf{u}^0$ is distorted by an additive noise $\boldsymbol{\varepsilon}$, i.e.,

$$\tilde{\mathbf{u}}^0(\mathbf{x}) = \mathbf{u}^0(\mathbf{x}) + \boldsymbol{\varepsilon}(\mathbf{x}) \text{ for } \mathbf{x} \in G. \tag{6}$$

For a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, the noise $\boldsymbol{\varepsilon}$ can be modeled as an $\mathcal{F}$-measurable random variable taking its values in a closed subspace of $L^2(G, \mathbb{R}^k)$, e.g., a Gaussian random field on G. The goal is to reconstruct or at least to 'optimally' estimate the (unknown) original image $\mathbf{u}^0$ based on the noisy observation $\tilde{\mathbf{u}}^0$.

We outline the following abstract approach (known as the scale-space theory) to constructing such estimators based on general techniques of semiparametric statistics (cf. [37, Section 1.2.2]). First, a deterministic semiflow $\big(S(t)\big)_{t\geq 0}$ on $L^2(G,\mathbb{R}^k)$ referred to as a 'scale-space' is introduced. There are various rationales behind a particular selection $\big(S(t)\big)_{t\geq 0}$. For example, $\big(S(t)\big)_{t\geq 0}$ can be designed such that any 'reasonable' unpolluted image $\mathbf{u}^0$ can be approximated by one of the stationary points of $\big(S(t)\big)_{t\geq 0}$. In this case, an estimate $\hat{\mathbf{u}}^0$ of $\mathbf{u}^0$ is given by

$$\hat{\mathbf{u}}^0 = S(T)\tilde{\mathbf{u}}^0 \text{ for an appropriately large } T > 0. \tag{7}$$

Another example is given when $S(\cdot)$ is selected to play the role of a kernel smoothing operator from the nonparametric statistics (cf. [34, Chapter 8]). In this case, the evaluation time T in Equation (7) roughly represents the (reciprocal) bandwidth and is typically selected to minimize the asymptotic mean integrated square error (AMISE) as a function of the design size n (e.g., the number of pixels available).

Remark 1. *Among other popular approaches such as the low-pass filtering, morphological multiscale analysis, neural networks, Bayesian techniques, etc., one should mention the penalized nonparametric regression. Given a noise image $\tilde{\mathbf{u}}^0(\mathbf{x})$ from Equation (6), the filtered image is obtained by minimizing the penalized objective functional*

$$\mathcal{J}(\mathbf{u}) = \frac{1}{2}\int_G \big|\tilde{\mathbf{u}}^0 - \mathbf{u}\big|^2 \mathrm{d}\mathbf{x} + \lambda\mathcal{P}(\mathbf{u}) \tag{8}$$

with a regularization parameter $\lambda > 0$. Here, the first term measures the L^2-goodness of fit between the noisy and the filtered images and can alternatively be replaced with any other L^p-norm. The second represents a Tychonoff-regularization associated with a stronger topology. The typical choices are

$$\mathcal{P}(\mathbf{u}) = \frac{1}{2}\int_G \sum_{|\beta|\leq k} c_\beta\Big(\big(\nabla^\alpha\mathbf{u}\big)_{|\alpha|\leq s}\Big)^\beta \mathrm{d}\mathbf{x} \quad or \quad \mathcal{P}(\mathbf{u}) = \mathrm{TV}(\mathbf{u}),$$

where $\mathrm{TV}(\mathbf{u})$ stands for the total variation of $\mathbf{u}$. Whereas the former choice leads to the classical spline smoothing/De Boor's approach ([34, Section 8.2.2]) or elastic maps/thin plate smoothing splines ([36, Section 4.3]), etc., the latter one is known as ROF-denoising model (cf. [32], [13, pp. 1–70]). The first-order Lagrange optimality condition for the minimum of functional in Equation (8) is typically given as a (parameter-)elliptic partial differential equation or inclusion.

In the following, we use a synthesis of these two approaches to put forth the semiflow $\big(S(t)\big)_{t\geq 0}$. The latter is also referred to as a C_0-operator semigroup (cf., e.g., [8, Chapter 4] or [11, Chapter 5]). As a matter of fact, one can not expect the filter to be able to perfectly reconstruct the original image. At the same time, the filter should be designed the way it performs 'well' on a certain class or set of images. Assuming $\boldsymbol{\varepsilon}$ is only locally autocorrelated, a natural choice is to let the evolution associated with semiflow be driven by a partial differential equation (PDE). In the following, we briefly outline our PDE model.

Motivated by the standard approach adopted in the theory of transport phenomena (cf. [37, p. 2]), let $\mathbf{u}(t, \cdot) \in L^2(G, \mathbb{R}^k)$ denote the color intensity at time $t \geq 0$ after applying the semiflow to the initial noisy measurement $\tilde{\mathbf{u}}^0$. In physical applications, $\mathbf{u}$ is usually a scalar variable representing the heat or material concentration density, etc. Further, let $\mathbf{J}(t, \cdot) \in L^2(G, \mathbb{R}^{k \times d})$ denote the 'color flux' tensor at time $t \geq 0$. Intuitively speaking, $\mathbf{J}(t, \cdot)$ represents the direction the color intensity is flowing into to compensate for local distortions caused by the noise. Assuming that there are no other sources of color distortion, we exploit the divergence theorem to obtain the following conservation or continuity equation

$$\partial_t \mathbf{u} + \operatorname{div} \mathbf{J} = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{9}$$

where $\operatorname{div} \mathbf{J} = \left(\sum_{j=1}^{d} \partial_{x_j} J_{1j}, \sum_{j=1}^{d} \partial_{x_j} J_{2j}, \ldots, \sum_{j=1}^{d} \partial_{x_j} J_{kj} \right)^T$. Since Equation (9) is underdetermined, a so-called constitutive equation establishing a relation between $\mathbf{u}$ and $\mathbf{J}$ is needed. In many applications, one adopts the well-known Fick's law of diffusion, which postulates $\mathbf{P}(t, \cdot)$ to be proportional to $-\nabla \mathbf{u}(t, \cdot)$, i.e.,

$$\mathbf{J}(t, \mathbf{x}) = -\mathbf{H}(t, \mathbf{x}) \nabla \mathbf{u}(t, \mathbf{x}) = -\left(\sum_{I=1}^{d} \sum_{J=1}^{k} H_{ijIJ}(t, \mathbf{x}) \partial_{x_J} u_I(t, \mathbf{x}) \right)_{i=1,\ldots,k}^{j=1,\ldots,d} \text{ for } \mathbf{x} \in G. \tag{10}$$

Here, $\nabla \mathbf{u}$ stands for the Jacobian of $\mathbf{u}$ and the symmetric fourth-order tensor $\mathbf{H}(t, \cdot) \in \mathbb{R}^{(k \times d) \times (k \times d)}$ plays the role of diffusivity tensor and can be interpreted as a symmetric linear mapping from $\mathbb{R}^{k \times d}$ into itself. With this in mind, Equation (9) rewrites as

$$\partial_t \mathbf{u} - \operatorname{div}(\mathbf{H} \nabla \mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G. \tag{11}$$

If $\mathbf{H}$ is a constant tensor, Equation (11) is referred to as the homogeneous diffusion. Otherwise, Equation (11) is still underdetermined and a futher constitutive relation between $\mathbf{H}$ and $\nabla \mathbf{u}$ is indispensable. In physics, this equations models the properties of the medium the diffusion is taking place in and/or the properties of the substance which is diffusing. In image processing, some other principles are adopted. See Section 2.2.1 below for details. Assuming, for example,

$$\mathbf{H} = \mathbf{F}(\nabla \mathbf{u}) \text{ in } (0, \infty) \times G \tag{12}$$

for an appropriate response function $\mathbf{F} \colon \mathbb{R}^{k \times d} \to \mathbb{R}^{(k \times d) \times (k \times d)}$ and plugging Equation (12) into (11) leads to a multicolor anisotropic generalization of Perona & Malik's filter [29]

$$\partial_t \mathbf{u} - \operatorname{div}\big(\mathbf{F}(\nabla \mathbf{u}) \nabla \mathbf{u} \big) = \mathbf{0} \text{ in } (0, \infty) \times G. \tag{13}$$

As a parabolic PDE system, Equation (13) exhibits an infinite signal propagation speed. Since any practically relevant selection of the diffusivity function $\mathbf{F}$ violates the causality principle, the equation even turns out to be ill-posed. Besides, no direct intuition on how to select the stopping time T from Equation (7) is provided. This motivated Nitzberg & Shiota [27] and Cottet & El Ayyadi [17] to consider a hyperbolic relaxation of Equation (12) for the particular case $k = 1$. For a positive relaxation parameter $\tau > 0$, they replaced Equation (12) with the first-order hyperbolic equation

$$\tau \partial_t \mathbf{H} + \mathbf{H} = \mathbf{F}(\nabla \mathbf{u}) \text{ in } (0, \infty) \times G. \tag{14}$$

They called their regularization a 'time-delay', which is, strictly speaking, not correct since it rather has the form of a relaxation. At the same time, Equation (14) can be viewed as a first-order Taylor approximation with respect to τ of the delay equation

$$\mathbf{H}(t + \tau, \mathbf{x}) = \mathbf{F}\big(\nabla\mathbf{u}(t, \mathbf{x})\big) \text{ for } (t, \mathbf{x}) \in (0, \infty) \times G.$$

Equation (14) together with (11) yields a nonlinear PDE system

$$\partial_t\mathbf{u} - \operatorname{div}(\mathbf{H}\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{15}$$

$$\tau\partial_t\mathbf{H} + \mathbf{H} - \mathbf{F}(\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G. \tag{16}$$

Recall that $\mathbf{H}\nabla\mathbf{u}$ stands for the tensor-matrix multiplication (cf. Equation (10)). In fact, Equations (15)–(16) are very much reminiscent of the well-known Cattaneo system (cf. [15]) of relativistic heat conduction. Formally speaking, Equation (14) 'converges' to (12) as $\tau \to 0$. Equations (15)–(16) can be viewed as parabolic-hyperbolic PDE system or a nonlinear Gurtin & Pipkin heat equation (cf. [20]). Indeed, solving Equation (16) for $\mathbf{H}$ and plugging the result into Equation (15), we obtain a memory-type equation

$$\partial_t\mathbf{u} - \operatorname{div}\left(\left(\int_0^{\cdot} \exp\big(-(\cdot - s)/\tau\big)\big(\mathbf{F}(\nabla\mathbf{u})\big)(s)\mathrm{d}s\right)\nabla\mathbf{u}\right) = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{17}$$

which, after being differentiated with respect to t, yields a quasilinear wave equation with a Kelvin & Voigt damping and memory-time coefficients.

Next, appropriate boundary conditions for $(\mathbf{u}, \mathbf{H})^T$ need to be prescribed. Neither Dirichlet, nor periodic boundary conditions seem to be adequate for the most applications. In contrast to that, a nonlinear Neumann boundary condition

$$\mathbf{n}^T(\mathbf{H}\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times \partial G \tag{18}$$

turns out both to be mathematically sound and geometrically intuitive. Equation (18) states that the color flow on the boundary vanishes in the normal direction.

As for the initial conditions, we prescribe

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } G. \tag{19}$$

Here, the fourth-order tensor $\tilde{\mathbf{H}}^0$ can be chosen to be symmetric and positive definite, e.g., $\tilde{\mathbf{H}}^0 = \big(\alpha\delta_{iI}\delta_{jJ}\big)_{i,I=1,\ldots,k}^{j,J=1,\ldots,d}$ for a small parameter $\alpha > 0$.

Collecting Equations (15)–(19), we arrive at an initial-boundary value problem for the quasilinear PDE system

$$\partial_t\mathbf{u} - \operatorname{div}(\mathbf{H}\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{20}$$

$$\tau\partial_t\mathbf{H} + \mathbf{H} - \mathbf{F}(\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G. \tag{21}$$

$$\mathbf{n}^T(\mathbf{H}\nabla\mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times \partial G, \tag{22}$$

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } G. \tag{23}$$

Equations (20)–(23) can be viewed as a mixed parabolic-hyperbolic system. Additionally, the boundary condition in Equation (21) is nonlinear. Hence, neither the classical hyperbolic solution theory (viz. [28] or [35]), nor the classical parabolic solution theory (see, e.g., [8]) are directly applicable.

To make Equations (20)–(23) better feasible, Cottet & El Ayyadi considered in [17] a regularization of $\nabla \mathbf{u}$ in Equation (21) though a spatial mollification. For $\mathbf{u} \in L^2(G, \mathbb{R}^k)$ and a kernel $\rho \in L^\infty_{\text{loc}}(\mathbb{R}^d, \mathbb{R})$, the convolution $\mathbf{u} * \rho$ is given by

$$\big(\mathbf{u} * \rho\big)(\mathbf{x}) := \int_G \rho(\mathbf{x} - \mathbf{y})\mathbf{u}(\mathbf{y})\mathrm{d}\mathbf{y} \text{ for } \mathbf{x} \in G.$$

Selecting now a fixed mollifier $\rho \in W^{1,\infty}\big(\mathbb{R}^d, \mathbb{R}\big)$ with $\rho \geq 0$ a.e. in $\mathbb{R}^d$ and $\int_{\mathbb{R}^d} \rho(\mathbf{x})\mathrm{d}\mathbf{x} = 1$, e.g., ρ can be the Gaussian pdf, as well as a bandwidth $\sigma > 0$, we define for $\mathbf{u} \in L^2(G, \mathbb{R}^k)$ the nonlocal operator

$$\nabla_\sigma \mathbf{u} := \nabla\big(\mathbf{u} * \rho_\sigma\big) \text{ with } \rho_\sigma(\mathbf{x}) := \tfrac{1}{\sigma^d}\rho\big(\tfrac{\mathbf{x}}{\sigma}\big) \text{ for } \mathbf{x} \in G$$

as a regularization of the gradient operator. With $(\rho_\sigma)_{\sigma>0}$ being a delta sequence, ∇_σ is a regular approximation of the ∇-operator. Replacing ∇ with ∇_σ in Equation (21), we arrive at the following system of partial integro-differential equations

$$\partial_t \mathbf{u} - \operatorname{div}(\mathbf{H}\nabla \mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{24}$$

$$\tau \partial_t \mathbf{H} + \mathbf{H} - \mathbf{F}(\nabla_\sigma \mathbf{u}) = \mathbf{0} \text{ in } (0, \infty) \times G, \tag{25}$$

$$(\mathbf{H}\nabla \mathbf{u})^T \mathbf{n} = \mathbf{0} \text{ on } (0, \infty) \times \partial G, \tag{26}$$

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } G. \tag{27}$$

2.2 Parameter Selection

In this subsection, we discuss the choice of parameters τ, $\mathbf{F}$ and $\tilde{\mathbf{H}}^0$. Also, depending on whether the model (20)–(23) or (24)–(27) is adopted, a kernel ρ and a regularization parameter $\sigma > 0$ need or need not to be selected. Here, we restrict ourselves to the limiting case $\sigma = 0$.

2.2.1 Response function F

For $d, k \in \mathbb{N}$, let the space $\mathbb{R}^{k \times d}$ of real $(k \times d)$-matrices be equipped with the Frobenius scalar product

$$\langle \mathbf{D}, \hat{\mathbf{D}} \rangle_{\mathbb{R}^{k \times d}} \equiv \mathbf{D} : \hat{\mathbf{D}} := \sum_{i=1}^k \sum_{j=1}^d D_{ij}\hat{D}_{ij} \text{ for } \mathbf{D}, \hat{\mathbf{D}} \in \mathbb{R}^{k \times d}. \tag{28}$$

For a fixed $\hat{\mathbf{D}} \in \mathbb{R}^{k \times d}$, we can thus define an orthogonal projection operator $\mathbb{P}_{\hat{\mathbf{D}}^\perp} : \mathbb{R}^{k \times d} \to \mathbb{R}^{k \times d}$ onto the orthogonal complement of $\hat{\mathbf{D}}$ via

$$\mathbb{P}_{\hat{\mathbf{D}}^\perp}(\mathbf{D}) = \mathbf{D} - \frac{(\mathbf{D} : \hat{\mathbf{D}})\hat{\mathbf{D}}}{\hat{\mathbf{D}} : \hat{\mathbf{D}}} \text{ for } \mathbf{D} \in \mathbb{R}^{k \times d}.$$

Obviously, $\mathbb{P}_{\hat{\mathbf{D}}^\perp}$ can be viewed as an element of $\mathbb{R}^{(k \times d) \times (k \times d)}$. A first choice of $\mathbf{F}$ could be

$$\mathbf{F}(\mathbf{D}) = \mathbb{P}_{\hat{\mathbf{D}}^\perp}(\mathbf{D}). \tag{29}$$

In addition to being unsmooth at zero and thus possibly leading to technical difficulties when treating Equations (20)–(23) analytically or numerically, this particular choice of the nonlinearity does not seem to meet practical requirements which are desirable for an image filter. Indeed, as reported in [17, Section III.A], the resulting system possesses too many undesirable stationary points and exhibits a too fast convergence speed leading to such a serious drawback that a rather high amount of noise is retained. That is why a regularization of Equation (29) based on a contrast threshold parameter should be adopted. Motivated by [17, Equation (21)], we let

$$
\mathbf{F}_s(\nabla\mathbf{u}) = \begin{cases} \mathbb{P}_{(\nabla\mathbf{u})^\perp}, & (\nabla\mathbf{u}) : (\nabla\mathbf{u}) \geq s^2, \\ \frac{3}{2}\left(1 - \frac{(\nabla\mathbf{u}):(\nabla\mathbf{u})}{s^2}\right) + \frac{(\nabla\mathbf{u}):(\nabla\mathbf{u})}{s^2}\mathbb{P}_{(\nabla\mathbf{u})^\perp}, & \text{otherwise.} \end{cases}
\tag{30}
$$

For a discussion on the particular choice of the coefficient $\frac{3}{2}$ and a connection to a neural network model, we refer to [17, Section IV]. With the color variables being each rescaled to lie in the interval $[-1, 1]$, the contrast threshold s is usually selected as 5% to 10% of the image width or height. Note that the choice of function $\mathbf{F}$ in Equation (30) satisfies Assumption 10 thus complying with our existence theory in Sections 3 and 4. Numerous alternative choices of the response function $\mathbf{F}$ can be found in [24, Table 1, p. 178].

2.3 Tensor $\tilde{\mathbf{H}}^0$

As for the initial diffusivity tensor $\tilde{\mathbf{H}}^0$, assuming the noise $\boldsymbol{\varepsilon}$ in Equation (6) is weakly autocorrelated, we can select

$$
\tilde{\mathbf{H}}^0(\mathbf{x}) = \mathrm{Cov}\big[\nabla\boldsymbol{\varepsilon}(\mathbf{x})\big] \text{ for } \mathbf{x} \in G
\tag{31}
$$

under an additional uniform positive definiteness condition on $\mathrm{Cov}\big[\boldsymbol{\varepsilon}(\cdot)\big]$. Since $\mathrm{Cov}\big[\boldsymbol{\varepsilon}(\cdot)\big]$ is not known in practice, the value of $\mathrm{Cov}\big[\boldsymbol{\varepsilon}(\mathbf{x})\big]$ for a particular $\mathbf{x} \in G$ can be estimated by computing the sample covariance matrix of $\nabla\tilde{\mathbf{u}}^0$ evaluated over an appropriate neighborhood of $\mathbf{x}$. We refer to [14, 25], [34, Chapter 8] for a discussion on the optimal neighborhood size.

2.4 Relaxation time τ

Repeating the calculations in [17, Section III.A] and [38], any particular selection of the parameter τ can be shown to imply that any graphical pattern occurring on scales smaller than $\sqrt{\tau}$ vanishes asymptotically, i.e., converges to its spatial mean. If no prior information on the minimum pattern size is available, statistical methods need to be employed to estimate the former.

3 The regular case $\sigma > 0$

In this section, we provide a well-posedness theory for Equations (24)–(27). In contrast to [17], nonlinear Neumann and not linear periodic boundary conditions are presribed in Equation (26). Being more adequate for practical applications, they are mathematically

more challenging since the evolution is now driven by an operator with a time-varying domain. Thus, a standard application of Faedo & Galerkin method is rather problematic. Therefore, we propose a new solution technique based on the maximum L^2-regularity theory for non-autonomous sesquilinear forms due to Dautray & Lions [18, Chapter 18, §3] as well as its recent improvement by Dier [19]. Another major advantage of our approach over [17] is that we get a much more regular solution under weaker smoothness assumptions on the initial data. Our results have a certain degree of resemblance to [16], where Equations (24)–(27) were studied for $\tau = 0$ and $k = 1$. In contrast to [16], we consider both the weak and strong settings without requiring $\mathbf{F}$ to be a C^∞-function.

Without loss of generality, let $\int_G \tilde{\mathbf{u}}^0 d\mathbf{x} = \mathbf{0}$. Otherwise, replace Equation (27) with

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0 - \frac{1}{|G|} \int_G \tilde{\mathbf{u}}^0 d\mathbf{x}, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } G \tag{32}$$

and solve the resulting system for $\mathbf{u}$. Later, by adding $\frac{1}{|G|} \int_G \tilde{\mathbf{u}}^0 d\mathbf{x}$ to $\mathbf{u}$, a solution to the original system is obtained.

We consider the space $\mathbb{R}^{(k \times d) \times (k \times d)}$ of real fourth-order tensors. Similar to Equation (28), we equip $\mathbb{R}^{(k \times d) \times (k \times d)}$ with the Frobenius inner product

$$\langle \mathbf{H}, \hat{\mathbf{H}} \rangle_{\mathbb{R}^{(k \times d) \times (k \times d)}} \equiv \mathbf{H} : \hat{\mathbf{H}} := \sum_{i,I=1}^{k} \sum_{j,J=1}^{d} H_{ijIJ} \hat{H}_{ijIJ} \text{ for } \mathbf{H}, \hat{\mathbf{H}} \in \mathbb{R}^{(k \times d) \times (k \times d)}.$$

With all norms being equivalent on the finite dimensional space $\mathbb{R}^{(k \times d) \times (k \times d)}$ by virtue of Riesz' theorem, the Frobenius norm $\sqrt{(\cdot) : (\cdot)}$ is equivalent with the operator norm

$$\|\mathbf{H}\|_{L(\mathbb{R}^{k \times d})} = \sup_{\|\mathbf{D}\|_{\mathbb{R}^{k \times d}} = 1} \|\mathbf{H}\mathbf{D}\|_{\mathbb{R}^{k \times d}}.$$

The space $\mathcal{S}(\mathbb{R}^{k \times d})$ of symmetric $(k \times d) \times (k \times d)$-tensors

$$\mathcal{S}(\mathbb{R}^{k \times d}) = \left\{ \mathbf{H} \in \mathbb{R}^{(k \times d) \times (k \times d)} \,|\, H_{ijIJ} = H_{IJij} \text{ for } i, I = 1, \ldots, k \text{ and } j, J = 1, \ldots, d \right\}$$

is a closed subspace of $\mathbb{R}^{(k \times d) \times (k \times d)}$. Obviously, $\mathcal{S}(\mathbb{R}^{k \times d})$ is isomorphic to the space of linear symmetric operators on $\mathbb{R}^{k \times d}$.

For $\kappa \geq 0$, we define a closed subset $\mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d})$ of $\mathcal{S}(\mathbb{R}^{k \times d})$ by the means of

$$\mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d}) := \left\{ \mathbf{H} \in \mathcal{S}(\mathbb{R}^{k \times d}) \,|\, \lambda_{\min}(\mathbf{H}) \geq \kappa \right\},$$

where $\lambda_{\min}(\mathbf{H})$ denotes the smallest eigenvalue of $\mathbf{H}$ viewed as a bounded, linear operator on $\mathbb{R}^{k \times d}$. As we know from the linear algebra, $\lambda_{\min}(\mathbf{H}) \geq \kappa$ is equivalent to

$$(\mathbf{H}\mathbf{D}) : \mathbf{D} \geq \kappa(\mathbf{D} : \mathbf{D}) \text{ for any } \mathbf{D} \in \mathbb{R}^{k \times d}.$$

Further, we define

$$L^\infty\left(G, \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d})\right) := \left\{ \mathbf{H} \in L^\infty\left(G, \mathcal{S}(\mathbb{R}^{k \times d})\right) \,|\, \mathbf{H}(\mathbf{x}) \in \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d}) \text{ for a.e. } \mathbf{x} \in G \right\}$$

$$= \left\{ \mathbf{H} \in L^\infty\left(G, \mathcal{S}(\mathbb{R}^{k \times d})\right) \,|\, \operatorname*{ess\,inf}_{\mathbf{x} \in G} \lambda_{\min}(\mathbf{H}) \geq \kappa \right\}.$$

Throughout this subsection, we require the following assumption on ρ and $\mathbf{F}$.

Assumption 2. *Let the functions* $\mathbf{F}\colon \mathbb{R}^{k\times d} \to \mathcal{S}_{\geq 0}(\mathbb{R}^{k\times d})$ *as well as* $\rho\colon \mathbb{R}^d \to \mathbb{R}$ *be weakly differentiable and let their first-order weak derivatives be essentially bounded by a positive number* $c > 0$.

Remark 3. *Assumption 2 is weaker than the one in [17, p. 293, Equation (4)].*

We introduce the nonlinear mapping $\mathcal{F}_\sigma\colon L^2(G, \mathbb{R}^k) \to L^\infty(G, \mathcal{S}_{\geq 0}(\mathbb{R}^{k\times d}))$ with

$$\mathcal{F}_\sigma(\mathbf{u}) := \mathbf{F}(\nabla_\sigma \mathbf{u}) \text{ for } \mathbf{u} \in L^2(G, \mathbb{R}^k).$$

Lemma 4. *The mapping* $\mathcal{F}$ *is Lipschitz-continuous.*

Proof. Using Hölder's inequality and Assumption 2, we can estimate for any $\mathbf{x} \in G$

$$\left\|(\nabla_\sigma \mathbf{u})(\mathbf{x})\right\|_{\mathbb{R}^{k\times d}} = \left\|\nabla(\mathbf{u} * \rho_\sigma)(\mathbf{x})\right\|_{\mathbb{R}^{k\times d}} = \left\|\nabla\left(\int_G \sigma^{-d}\rho(\sigma^{-1}(\mathbf{x}-\mathbf{y}))\mathbf{u}(\mathbf{y})\mathrm{d}\mathbf{y}\right)\right\|_{\mathbb{R}^{k\times d}}$$

$$\leq \sigma^{-(d+1)}\int_G \left|\rho'(\sigma^{-1}(\mathbf{x}-\mathbf{y}))\right| \|\mathbf{u}(\mathbf{y})\|_{\mathbb{R}^k}\mathrm{d}\mathbf{y} \leq \sigma^{-(d+1)}C(|G|)^{1/2}\|\mathbf{u}\|_{L^2(G,\mathbb{R}^k)},$$

where $|G| < \infty$ is the standard Lebesgue measure of G. Hence, for any $\mathbf{u}, \hat{\mathbf{u}} \in L^2(G, \mathbb{R}^k)$, we get

$$\left\|(\mathbf{F}(\nabla_\sigma \mathbf{u}))(\mathbf{x}) - (\mathbf{F}(\nabla_\sigma \hat{\mathbf{u}}))(\mathbf{x})\right\|_{\mathbb{R}^{(k\times d)\times(k\times d)}} \leq C\left\|(\nabla_\sigma \mathbf{u})(\mathbf{x}) - (\nabla_\sigma \hat{\mathbf{u}})(\mathbf{x})\right\|_{\mathbb{R}^{k\times d}}$$

$$\leq C^2\sigma^{-(d+1)}(|G|)^{1/2}\|\mathbf{u} - \hat{\mathbf{u}}\|_{L^2(G,\mathbb{R}^k)},$$

which finishes the proof. □

We let

$$\mathcal{H} := L^2(G, \mathbb{R}^k)/\{\mathbf{1}\} \equiv \left\{\mathbf{u} \in L^2(G, \mathbb{R}^k) \,\Big|\, \int_G \mathbf{u}\,\mathrm{d}\mathbf{x} = \mathbf{0}\right\}, \quad \mathcal{V} := H^1(G, \mathbb{R}^k) \cap \mathcal{H}.$$

Then $(\mathcal{V}, \mathcal{H}, \mathcal{V}')$ is a Gelfand triple. For a tensor-valued function $\mathbf{H} \in L^\infty(G, \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k\times d}))$ for some $\kappa > 0$, we further consider the bilinear form

$$a(\cdot, \cdot; \mathbf{H})\colon \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \quad (\mathbf{u}, \mathbf{v}) \mapsto \int_G (\mathbf{H}\nabla \mathbf{u}) : (\nabla \mathbf{v})\mathrm{d}\mathbf{x},$$

where

$$\mathcal{V} := (H^1(G, \mathbb{R}^k) \cap \mathcal{H}) \times (H^1(G, \mathbb{R}^k) \cap \mathcal{H}).$$

By virtue of Assumption 2, $a(\cdot, \cdot; \mathbf{H})$ is a symmetric, continuous bilinear form. The associated linear bounded symmetric operator $\mathcal{A}(\mathbf{H})\colon \mathcal{V} \to \mathcal{V}'$ is given as

$$\langle \mathcal{A}(\mathbf{H})\mathbf{u}, \tilde{\mathbf{u}}\rangle_{\mathcal{V}';\mathcal{V}} := a(\mathbf{u}, \tilde{\mathbf{u}}; \mathbf{H}) \text{ for any } \mathbf{u}, \tilde{\mathbf{u}} \in \mathcal{V}.$$

Using Assumption 2 and the second Poincaré's equality, we can estimate

$$a(\mathbf{u}, \mathbf{u}; \mathbf{H}) \geq \kappa\|\nabla \mathbf{u}\|^2_{L^2(G,\mathbb{R}^{k\times d})} \geq \kappa C_P\|\mathbf{u}\|^2_{\mathcal{V}} \text{ for any } \mathbf{u} \in \mathcal{V}, \tag{33}$$

where $C_P = C_P(G) > 0$ is the Poincaré's constant. Hence, $\mathcal{A}(\mathbf{H})$ is continuously invertible and self-adjoint.

From the elliptic theory, we know $\mathcal{A}(\mathbf{H})$ to generalize the Neumann boundary conditions in (26) associated with the PDE in Equation (24). Further, we know that the maximum domain of the strong realization of $\mathcal{A}(\mathbf{H})$

$$D\big(\mathcal{A}(\mathbf{H})\big) := \big\{\mathbf{u} \in \mathcal{V} \,|\, \mathcal{A}(\mathbf{H})\mathbf{u} \in \mathcal{H}\big\}$$

is a dense subspace of $\mathcal{H}$.

Remark 5. *If $\mathbf{H} \in C^1\big(\bar{G}, \mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d})\big)$ for $\kappa > 0$, the elliptic regularity theory implies*

$$D\big(\mathcal{A}(\mathbf{H})\big) \subset H^2(G, \mathbb{R}^k)$$

if $G \in C^2$ (cf. [23, Lemma 3.6] for the case G is a rectangular box). This regularity for $\mathbf{H}$ can be assured by selecting a regular convolution kernel ρ and a smooth nonlinearity $\mathbf{F}$ as well as considering smooth initial data $\mathbf{H}^0$ for $\mathbf{H}$.

With the notation introduced above, Equations (24)–(27) can be written in the following abstract form:

$$\partial_t \mathbf{u} + \mathcal{A}(\mathbf{H})\mathbf{u} = 0 \text{ in } L^2(0, T; \mathcal{V}'), \tag{34}$$

$$\tau\partial_t \mathbf{H} + \mathbf{H} - \mathcal{F}_\sigma(\mathbf{u}) = 0 \text{ in } L^2\big(0, T; L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k\times d}))\big), \tag{35}$$

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0 \text{ in } \mathcal{H}, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k\times d})\big), \tag{36}$$

where the Neumann boundary conditions are now incorporated into the definition of operator $\mathcal{A}(\mathbf{H})$.

Definition 6. *For $T > 0$, a function $(\mathbf{u}, \mathbf{H})^T \colon [0, T] \times \bar{G} \to \mathbb{R}^k \times \mathcal{S}(\mathbb{R}^{k\times d})$ is referred to as a weak solution to Equations (34)–(36) on $[0, T]$ if there exists a number $\kappa > 0$ such that the function pair $(\mathbf{u}, \mathbf{H})^T$ satisfies $\mathbf{H} \in \mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d})$ a.e. in $(0, T) \times G$ and*

$$\mathbf{u} \in H^1(0, T; \mathcal{V}') \cap L^2(0, T; \mathcal{V}), \quad \mathbf{H} \in W^{1,\infty}\big(0, T; L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k\times d}))\big)$$

and fulfils the abstract differential equations (34)–(35) together with the initial conditions (36) in sense of interpolation Equation (75). If $\mathbf{u}$ additionally satisfies

$$\mathbf{u} \in H^1(0, T; \mathcal{H}) \text{ and } \mathrm{div}(\mathbf{H}\nabla\mathbf{u}) \in L^2(0, T; \mathcal{H}),$$

$(\mathbf{u}, \mathbf{H})^T$ is then referred to as a strong solution.

Remark 7. *Note that Definition 6 can easily be generalized to the case $T = \infty$ by replacing the Banach-Sobolev spaces $W^{s,p}$ and L^p with the metric Sobolev spaces $W^{s,p}_{\mathrm{loc}}$ and L^p_{loc}.*

Theorem 8. *Let $(\tilde{\mathbf{u}}^0, \tilde{\mathbf{H}}^0)^T \in \mathcal{H} \times L^\infty\big(G, \mathcal{S}_{\geq\alpha}(\mathbb{R}^{k\times d})\big)$ for some $\alpha > 0$. For any $T > 0$, the initial-boundary value problem (34)–(35) possesses then a unique weak solution on $[0, T]$ satisfying*

$$\mathbf{H}(t, \cdot) \in L^\infty\big(G, \mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d})\big) \text{ for a.e. } t \in [0, T] \text{ with } \kappa := \alpha \exp(-T/\tau).$$

Proof. Equivalent formulation and a priori *estimates:* Solving Equation (35) for $\mathbf{H}$ and plugging the result into Equation (34), Equations (34)–(36) reduce to

$$\partial_t \mathbf{u} + \mathcal{A}\big(\mathbf{H}(\mathbf{u})\big)\mathbf{u} = 0 \text{ in } (0, T), \quad \mathbf{u}(0, \cdot) = \mathbf{u}^0 \tag{37}$$

with

$$\big(\mathbf{H}(\mathbf{u})\big)(t, \cdot) = \exp(-t/\tau)\mathbf{H}_0 + \int_0^t \exp\big(-(t-s)/\tau\big)\mathbf{F}\big(\nabla_\sigma \mathbf{u}(s, \cdot)\big)\mathrm{d}s \text{ for } t \in [0, T]. \tag{38}$$

We prove several *a priori* estimates we later used in the proof. For any $\mathbf{D} \in \mathbb{R}^{k \times d}$, Equation (38) together with Assumption 2 imply

$$\begin{aligned}
\big(\big(\mathbf{H}(\mathbf{u})\big)(t, \cdot)\mathbf{D}\big) : \mathbf{D} &= \exp(-t/\tau)\big((\mathbf{H}^0\mathbf{D}) : \mathbf{D}\big) \\
&\quad + \int_0^t \exp\big(-(t-s)/\tau\big)\Big(\big(\mathbf{F}(\nabla_\sigma \mathbf{u}(s, \cdot))\mathbf{D}\big) : \mathbf{D}\Big)\mathrm{d}s \\
&\geq \kappa\|\mathbf{D}\|_{\mathbb{R}^{k \times d}} \text{ for a.e. } t \in [0, T] \text{ with } \kappa = \alpha\exp(-T/\tau),
\end{aligned} \tag{39}$$

and, therefore, $\mathbf{H}(t, \cdot) \in \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d})$ for a.e. $t \in [0, T]$ a.e. in G. On the other hand, by virtue of Equation (38) and Assumption 2,

$$\begin{aligned}
\big\|\big(\mathbf{H}(\mathbf{u})\big)(t, \cdot)\big\|_{L^\infty(G, \mathcal{S}(\mathbb{R}^{k \times d}))} &\leq \|\mathbf{H}^0\|_{L^\infty(G, \mathcal{S}(\mathbb{R}^{k \times d}))} \\
&\quad + \int_0^t \big\|\mathbf{F}\big(\nabla_\sigma \mathbf{u}(s, \cdot)\big)\big\|_{L^\infty(G, \mathcal{S}(\mathbb{R}^{k \times d}))}\mathrm{d}s \\
&\leq \|\mathbf{H}^0\|_{L^\infty(G, \mathcal{S}(\mathbb{R}^{k \times d}))} + cT \text{ for a.e. } t \in [0, T].
\end{aligned} \tag{40}$$

For $t \in (0, T]$, multiplying Equation (37) with $\mathbf{u}$ in $L^2(0, t; \mathcal{H})$, we obtain

$$\tfrac{1}{2}\big\|\mathbf{u}(t, \cdot)\big\|_{\mathcal{H}}^2 - \tfrac{1}{2}\big\|\mathbf{u}(0, \cdot)\big\|_{\mathcal{H}}^2 + \int_0^t a\big(\mathbf{u}(s, \cdot), \mathbf{u}(s, \cdot); \big(\mathbf{H}(\mathbf{u})\big)(s, \cdot)\big)\mathrm{d}s = 0.$$

Hence, using Equation (39), we arrive at

$$\big\|\mathbf{u}(t, \cdot)\big\|_{\mathcal{H}}^2 + 2\kappa\int_0^t \big\|\nabla\mathbf{u}(s, \cdot)\big\|_{L^2(\mathbb{R}^{k \times d})}^2\mathrm{d}s \leq \|\mathbf{u}^0\|_{\mathcal{H}}^2 \text{ for a.e. } t \in [0, T]. \tag{41}$$

Constructing a fixed point mapping: Consider the Banach space

$$\mathscr{X} := C^0\big([0, T], \mathcal{H}\big)$$

equipped with the standard topology. Now, we define an operator $\mathscr{F} \colon \tilde{\mathscr{X}} \to \tilde{\mathscr{X}}$ which maps each $\tilde{\mathbf{u}} \in \mathscr{X}$ to the unique weak solution

$$\mathbf{u} \in H^1(0, T; \mathcal{V}) \cap L^2(0, T; \mathcal{V}') \hookrightarrow \mathscr{X} \tag{42}$$

of Equations (37)–(38) (See Equation (75) for the embedding above.) By virtue of Equations (39) and (40), the bilinear non-autonomous form $t \mapsto a\big(\cdot, \cdot; \big(\mathbf{H}(\mathbf{u})\big)(t, \cdot)\big)$ is uniformly coercive and bounded. Hence, by Theorem (18), the operator $\mathscr{F}$ is well-defined.

Proving the contraction property of $\mathscr{F}$: To show the contraction property, similar to the classical existence and uniqueness theorem of Picard & Lindelöf, we first equip the Banach space $\mathscr{X}$ with an equivalent norm

$$\|\mathbf{u}\|_{\mathscr{X}} := \max_{t \in [0,T]} e^{-(L+1)t} \|\mathbf{u}(t, \cdot)\|_{\mathcal{H}}$$

with a positive constant L to be selected later. For the sake of simplicity, we keep the same notation for this new isomorphic space.

For $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2 \in \tilde{\mathscr{X}}$, let

$$\mathbf{u}_1 := \mathscr{F}(\tilde{\mathbf{u}}_1) \text{ and } \mathbf{u}_1 := \mathscr{F}(\tilde{\mathbf{u}}_2).$$

By definition, $\bar{\mathbf{u}} := \mathbf{u}_1 - \mathbf{u}_2$ solves then the non-autonomous linear (w.r.t. $\bar{\mathbf{u}}$) problem

$$\partial_t \bar{\mathbf{u}} + \mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}}_1)\big)\bar{\mathbf{u}} = \Big(\mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}}_1)\big) - \mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}}_2)\big)\Big)\mathbf{u}_2 \text{ in } L^2(0,T;\mathcal{V}'), \tag{43}$$

$$\bar{\mathbf{u}}(t, \cdot) = \mathbf{0} \text{ in } \mathcal{H} \tag{44}$$

For $t \in (0,T]$, multiplying Equation (43) with $\bar{\mathbf{u}}$ in $L^2(0,t;\mathcal{H})$, we get

$$
\begin{aligned}
\tfrac{1}{2}\|\bar{\mathbf{u}}(t,\cdot)\|_{\mathcal{H}} &\leq \int_0^t \big\|\mathbf{H}(\tilde{\mathbf{u}}_1(s,\cdot)) - \mathbf{H}(\tilde{\mathbf{u}}_2(s,\cdot))\big\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k \times d}))} \\
&\quad \times \big\|\nabla \mathbf{u}_2(s,\cdot)\big\|_{L^2(G,\mathbb{R}^{k \times d})} \big\|\nabla \bar{\mathbf{u}}(s,\cdot)\big\|_{L^2(G,\mathbb{R}^{k \times d})} \mathrm{d}s \\
&\leq \Big(\max_{s \in [0,t]} \big\|\mathbf{H}(\tilde{\mathbf{u}}_1(s,\cdot)) - \mathbf{H}(\tilde{\mathbf{u}}_2(s,\cdot))\big\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k \times d}))} \Big) \\
&\quad \times \big\|\nabla \mathbf{u}_2\big\|_{L^2((0,T) \times G, \mathbb{R}^{k \times d})} \big\|\nabla \bar{\mathbf{u}}\big\|_{L^2((0,T) \times G, \mathbb{R}^{k \times d})} \\
&\leq \tilde{C}\Big(\max_{s \in [0,t]} \big\|\mathbf{H}(\tilde{\mathbf{u}}_1(s,\cdot)) - \mathbf{H}(\tilde{\mathbf{u}}_2(s,\cdot))\big\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k \times d}))} \Big),
\end{aligned} \tag{45}
$$

where

$$\tilde{C} := \frac{\|\mathbf{u}^0\|_{\mathcal{H}}}{\kappa} \geq \|\nabla \mathbf{u}_2\|_{L^2((0,T) \times G, \mathbb{R}^{k \times d})} \|\nabla \bar{\mathbf{u}}\|_{L^2((0,T) \times G, \mathbb{R}^{k \times d})}$$

by virtue of Equation (41). Using Equation (38), we estimate

$$
\begin{aligned}
\max_{s \in [0,t]} \big\|\mathbf{H}\big(\tilde{\mathbf{u}}_1(s,\cdot)\big) - \mathbf{H}\big(\tilde{\mathbf{u}}_2(s,\cdot)\big)\big\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k \times d}))} & \\
&\hspace{-8em}\leq \int_0^t \big\|\mathbf{F}\big(\nabla_\sigma \tilde{\mathbf{u}}_1(s,\cdot)\big) - \mathbf{F}\big(\nabla_\sigma \tilde{\mathbf{u}}_2(s,\cdot)\big)\big\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k \times d}))} \mathrm{d}s \\
&\hspace{-8em}\leq C_{\mathrm{Lip}} \int_0^t \big\|\tilde{\mathbf{u}}_1(s,\cdot) - \tilde{\mathbf{u}}_2(s,\cdot)\big\|_{\mathcal{H}} \mathrm{d}s,
\end{aligned} \tag{46}
$$

where C_{Lip} is the Lipschitz constant of the mapping $\mathscr{F}$ from Lemma 4. Combining the estimates from Equations (45) and (46), we arrive at

$$\big\|\bar{\mathbf{u}}(t,\cdot)\big\|_{\mathcal{H}} \leq L \int_0^t \big\|\tilde{\mathbf{u}}_1(s,\cdot) - \tilde{\mathbf{u}}_2(s,\cdot)\big\|_{\mathcal{H}} \mathrm{d}s \text{ with } L := 2\tilde{C}C_{\mathrm{Lip}}. \tag{47}$$

Multipying Equation (47) with $\exp(-Lt)$, we estimate

$$
\begin{aligned}
e^{-(L+1)t}\big\|\bar{\mathbf{u}}(t,\cdot)\big\|_{\mathcal{H}} &\le L e^{-(L+1)t}\int_0^t \big\|\tilde{\mathbf{u}}_1(s,\cdot)-\tilde{\mathbf{u}}_2(s,\cdot)\big\|_{\mathcal{H}}\mathrm{d}s \\
&\le L e^{-(L+1)t}\int_0^t e^{(L+1)t}\Big(e^{-(L+1)t}\big\|\tilde{\mathbf{u}}_1(s,\cdot)-\tilde{\mathbf{u}}_2(s,\cdot)\big\|_{\mathcal{H}}\Big)\mathrm{d}s \\
&\le \Big(L e^{-(L+1)t}\int_0^t e^{(L+1)t}\mathrm{d}s\Big)\|\tilde{\mathbf{u}}_1-\tilde{\mathbf{u}}_2\|_{\mathscr{X}} \\
&\le \Big(L e^{-(L+1)t}\frac{e^{(L+1)t}-1}{L+1}\Big)\|\tilde{\mathbf{u}}_1-\tilde{\mathbf{u}}_2\|_{\mathscr{X}} \\
&\le \frac{L}{L+1}\|\tilde{\mathbf{u}}_1-\tilde{\mathbf{u}}_2\|_{\mathscr{X}}.
\end{aligned}
\tag{48}
$$

Hence, taking the maximum over $t\in[0,T]$ on the left-hand side of Equation (48), we find

$$
\big\|\mathscr{F}(\tilde{\mathbf{u}}_1)-\mathscr{F}(\tilde{\mathbf{u}}_2)\big\|_{\mathscr{X}} \le \frac{L}{L+1}\|\tilde{\mathbf{u}}_1-\tilde{\mathbf{u}}_2\|_{\mathscr{X}}.
$$

which implies $\mathscr{X}$ is a contraction. By virtue of Banach's fixed point theorem, $\mathscr{F}$ posseses then a unique fixed point $\mathbf{u}\in\mathscr{X}$. Hence, applying Lemma 4 to Equation (38) and recalling Equation, we further get

$$
\mathbf{H}\in W^{1,\infty}\Big(0,T;L^\infty\big(G,\mathcal{S}(\mathbb{R}^{k\times d})\big)\Big)\ \text{and}\ \mathbf{H}(t,\cdot)\in L^\infty\big(G,\mathcal{S}_{\ge\kappa}(\mathbb{R}^{k\times d})\big)\ \text{for a.e. } t\in[0,T].
$$

Taking into account Equation (42) as well as the equivalence between Equations (34)–(36) and (37)–(38), we deduce $\mathbf{u}$ is the unique weak solution to Equations (24)–(27). $\qquad\square$

Corollary 9. *Under the conditions of Theorem 8, let $\tilde{\mathbf{u}}^0\in\mathcal{V}$. The weak solution $(\mathbf{u},\mathbf{H})^T$ given in the Theorem is then also a strong solution.*

Proof. For the unique weak solution $(\mathbf{u},\mathbf{H})^T$, consider the linear initial value problem

$$
\partial_t\mathbf{u}(t)+\tilde{\mathcal{A}}(t)\mathbf{u}(t)=\mathbf{0}\ \text{for } t\in(0,T),\quad \mathbf{u}(0)=\tilde{\mathbf{u}}^0,
\tag{49}
$$

where

$$
\tilde{\mathcal{A}}(t):=\mathcal{A}\big(\mathbf{H}(t,\cdot)\big)\ \text{for } t\in[0,T].
$$

The associated non-autonomous form is of bounded variation since

$$
\begin{aligned}
|\tilde{a}(\mathbf{u},\mathbf{v};t)-\tilde{a}(\mathbf{u},\mathbf{v};s)| &= \Big|\int_G\big((\mathbf{H}(t,\cdot)-\mathbf{H}(s,\cdot))\nabla\mathbf{u}\big):(\nabla\mathbf{v})\mathrm{d}\mathbf{x}\Big| \\
&= \int_s^t\|\partial_t\mathbf{H}(\xi,\cdot)\|_{L^\infty(G,\mathcal{S}(\mathbb{R}^{k\times d}))}\big\|\mathbf{u}(\xi,\cdot)\big\|_{\mathcal{V}}\big\|\mathbf{v}(\xi,\cdot)\big\|_{\mathcal{V}}\mathrm{d}\xi \\
&\le (t-s)\|\mathbf{H}\|_{W^{1,\infty}(0,T;L^\infty(G,\mathcal{S}(\mathbb{R}^{k\times d})))}\|\mathbf{u}\|_{\mathcal{V}}\|\mathbf{v}\|_{\mathcal{V}}
\end{aligned}
\tag{50}
$$

for $0\le s\le t\le T$ and $\mathbf{u},\mathbf{v}\in\mathcal{V}$. Theorem 19 applied to Equation (49) yields then

$$
\mathbf{u}\in W^{1,2}(0,T;\mathcal{H})\ \text{and}\ \mathrm{div}\big(\mathbf{H}\nabla\mathbf{u}\big)\in L^2(0,T;\mathcal{H}).
$$

Hence, $(\mathbf{u},\mathbf{H})^T$ is also a strong solution. $\qquad\square$

4 Limiting case $\sigma = 0$

In this remaining section, we want to obtain a solution theory for the original PDE system (20)–(23). In contrast to the regularized case, a slightly stronger assumption on the function $\mathbf{F}$ is required here. It is precisely the one used in [17, Equation (4)].

Assumption 10. *Let the function $\mathbf{F} \colon \mathbb{R}^{k \times d} \to \mathcal{S}_{\geq 0}(\mathbb{R}^{k \times d})$ be weakly differentiable such that $\mathbf{F}$ together with its weak Jacobian are essentially bounded by a positive number $c > 0$.*

Now, we introduce the nonlinear mapping $\mathcal{F} \colon H^1(G, \mathbb{R}^k) \to L^\infty(G, \mathcal{S}_{\geq 0}(\mathbb{R}^{k \times d}))$ with

$$\mathcal{F}(\mathbf{u}) := \mathbf{F}(\nabla \mathbf{u}) \text{ for } \mathbf{u} \in H^1(G, \mathbb{R}^d).$$

Obviously, $\mathcal{F}$ is well-defined. Indeed, $\mathbf{F}(\nabla \mathbf{u})$ is strongly measurable as a composition of two strongly measurable functions and essentially bounded by virtue of Assumption 10. Unlike $\mathcal{F}_\sigma$, generally speaking, $\mathcal{F}$ is not Lipschitzian from H^1 and L^∞. At the same time, due to Assumption 10, we trivially have:

Lemma 11. *The mapping $\mathcal{F}$ is Lipschitzian from $H^1(G, \mathbb{R}^d)$ to $L^2(G, \mathbb{R}^d)$.*

With the notations of Section 3, the abstract form of Equations (20)–(23) reads as

$$\partial_t \mathbf{u} + \mathcal{A}(\mathbf{H})\mathbf{u} = 0 \text{ in } L^2(0, T; \mathcal{V}'), \tag{51}$$

$$\tau \partial_t \mathbf{H} + \mathbf{H} - \mathcal{F}(\mathbf{u}) = 0 \text{ in } L^2\big(0, T; L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k \times d})\big)\big), \tag{52}$$

$$\mathbf{u}(0, \cdot) = \tilde{\mathbf{u}}^0 \text{ in } \mathcal{H}, \quad \mathbf{H}(0, \cdot) = \tilde{\mathbf{H}}^0 \text{ in } L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k \times d})\big), \tag{53}$$

We adopt the following solution notions for Equations (51)–(53). Note that the regularity condition on $\mathbf{H}$ differs from the one employed in the regularized case.

Definition 12. *For $T > 0$, we call a function $(\mathbf{u}, \mathbf{H})^T \colon [0, T] \times \bar{G} \to \mathbb{R}^k \times \mathcal{S}(\mathbb{R}^{k \times d})$ a weak solution to Equations (51)–(53) on $[0, T]$ if there exists a number $\kappa > 0$ such that the function pair $(\mathbf{u}, \mathbf{H})^T$ satisfies $\mathbf{H} \in \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d})$ a.e. in $(0, T) \times G$ and*

$$\mathbf{u} \in H^1(0, T; \mathcal{V}') \cap L^2(0, T; \mathcal{V}), \quad \mathbf{H} \in H^1\big(0, T; L^\infty\big(G, \mathcal{S}(\mathbb{R}^{k \times d})\big)\big)$$

and fulfils the abstract differential equations (51)–(52) and the initial conditions (53) in sense of interpolation Equation (75). If $\mathbf{u}$ additionally satisfies

$$\mathbf{u} \in H^1(0, T; \mathcal{H}) \text{ and } \operatorname{div}(\mathbf{H}\nabla \mathbf{u}) \in L^2(0, T; \mathcal{H}),$$

we refer to $(\mathbf{u}, \mathbf{H})^T$ as a strong solution.

Theorem 13. *Let $(\tilde{\mathbf{u}}^0, \tilde{\mathbf{H}}^0)^T \in \mathcal{H} \times L^\infty\big(G, \mathcal{S}_{\geq \alpha}(\mathbb{R}^{k \times d})\big)$ for some $\alpha > 0$. Under Assumption 10, for any $T > 0$, the initial-boundary value problem (51)–(53) possesses then a weak solution $(\mathbf{u}, \mathbf{H})^T$ on $[0, T]$ satisfying*

$$\mathbf{H}(t, \cdot) \in L^\infty\big(G, \mathcal{S}_{\geq \kappa}(\mathbb{R}^{k \times d})\big) \text{ for a.e. } t \in [0, T] \text{ with } \kappa := \alpha \exp(-1/\tau).$$

In addition, weak solutions are globally extendable (not necessarily uniquely).

Proof. Repeating the proof of Equation (39), we get the *a priori* positive definiteness for $\mathbf{H}$, i.e.,

$$\mathbf{H}(t,\cdot) \in L^{\infty}\big(G, \mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d})\big) \text{ for a.e. } t \in [0,T] \text{ with } \kappa := \alpha\exp(-T/\tau).$$

Solving Equation (52) for $\mathbf{H}$ and plugging the result into Equation (51), Equations (51)–(53) reduce to

$$\partial_t\mathbf{u} + \mathcal{A}(\mathbf{H}(\mathbf{u}))\mathbf{u} = 0 \text{ in } (0,T), \quad \mathbf{u}(0,\cdot) = \mathbf{u}^0 \tag{54}$$

with

$$\big(\mathbf{H}(\mathbf{u})\big)(t,\cdot) = \exp(-t/\tau)\mathbf{H}_0 + \int_0^t \exp\big(-(t-s)/\tau\big)\mathbf{F}\big(\nabla\mathbf{u}(s,\cdot)\big)\mathrm{d}s \text{ for } t \in [0,T]. \tag{55}$$

We solve Equation (54) by applying Schauder & Tychonoff's fixed point theorem (see, e.g., [26, p. 165]).

Constructing a fixed point mapping: Consider the convex compact subset

$$\mathscr{Y} := L^2(0,T;\mathcal{V}) \text{ of Hilbert space } \mathscr{X} := W^{-1,2}(0,T;\mathcal{H}),$$

where the compactness is a direct consequence of Rellich & Kondrachov's imbedding theorem and [2, Theorem 5.1]. Let $\mathscr{F}$ map an element $\mathbf{u} \in \mathscr{Y}$ to the (unique) solution $\mathbf{u} \in H^1(0,T;\mathcal{V}')\cap L^2(0,T;\mathcal{V})$ of Equation (54) with $\mathbf{H}$ given in Equation (55). We now show $\mathscr{F}$ is well-defined. For $\tilde{\mathbf{u}} \in \mathscr{Y}$, Assumption 10 implies $\mathbf{F}(\nabla\tilde{\mathbf{u}}) \in L^2\big(0,T;L^{\infty}\big(G,\mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d})\big)\big)$. Hence, by virtue of Equation (55),

$$\mathbf{H} \equiv \mathbf{H}(\mathbf{u}) \in W^{1,2}\big(0,T;L^{\infty}\big(G,\mathcal{S}(\mathbb{R}^{k\times d})\big)\big) \text{ and } \mathbf{H} \in \mathcal{S}_{\geq\kappa}(\mathbb{R}^{k\times d}) \text{ a.e. in } (0,T)\times G. \tag{56}$$

Here, we used the strong measurability of $\mathbf{H}$ and the boundedness of respective norms. Using the fundamental theorem of calculus and Cauchy & Schwarz' inequality, we estimate

$$\big|a(\mathbf{u},\mathbf{v};t)\big| \leq \int_G \Big|\big(\mathbf{H}(t,\mathbf{x})\nabla\mathbf{u}\big):(\nabla\mathbf{v}\Big|\mathrm{d}\mathbf{x} \tag{57}$$

$$\leq \int_G \Big|\Big(\big(\mathbf{H}^0(t,\cdot) + \int_0^t \partial_t\mathbf{H}(\xi,\cdot)\mathrm{d}\xi\big)\nabla\mathbf{u}\Big):(\nabla\mathbf{v}\Big|\mathrm{d}\mathbf{x} \tag{58}$$

$$\leq \Big(\|\mathbf{H}^0\|_{L^{\infty}(G,\mathcal{S}(\mathbb{R}^{k\times d}))} + \sqrt{T}\|\mathbf{H}\|_{W^{1,2}(0,T;L^{\infty}(G,\mathcal{S}(\mathbb{R}^{k\times d})))}\Big)\|\mathbf{u}\|_{\mathcal{V}}\|\mathbf{v}\|_{\mathcal{V}} \tag{59}$$

for $t \in (0,T]$ and $\mathbf{u},\mathbf{v} \in \mathcal{V}$, where

$$a(\mathbf{u},\mathbf{v};t) := \int_G (\mathbf{H}\nabla\mathbf{u}):(\nabla\mathbf{v})\mathrm{d}\mathbf{x}.$$

This together with the fact $\mathbf{u}^0 \in \mathcal{H}$ combined with Theorem 18 yields a unique solution $\mathbf{u} \in \mathscr{Y}$ to Equation (54). Hence, $\mathscr{F}$ is well-defined as a self-mapping on $\mathscr{Y}$.

Showing the continuity of $\mathscr{F}$: For an arbitrary, but fixed $\tilde{\mathbf{u}} \in \mathscr{Y}$ consider a sequence $(\tilde{\mathbf{u}}_n)_{n\in\mathbb{N}} \subset \mathscr{Y}$ such that $\tilde{\mathbf{u}}_n \to \tilde{\mathbf{u}}$ in $\mathscr{Y}$ as $n \to \infty$. Further, let $\mathbf{u} := \mathscr{F}(\tilde{\mathbf{u}})$ and $\mathbf{u}_n := \mathscr{F}(\tilde{\mathbf{u}}_n)$

for $n \in \mathbb{N}$. We want to show $\mathbf{u}_n \to \mathbf{u}$ in $\mathscr{Y}$ as $n \to \infty$. Note that the sequential continuity of $\mathscr{F}$ is equivalent with the regular continuity since $\mathscr{Y}$ is separable.

Let $\bar{\mathbf{u}}_n := \mathbf{u} - \mathbf{u}_n$. By definition, $\bar{\mathbf{u}}_n$ solves the Cauchy problem

$$\partial_t \bar{\mathbf{u}}_n + \mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}})\big)\bar{\mathbf{u}}_n = \mathbf{f}_n \text{ in } (0,T), \quad \mathbf{u}(0,\cdot) = 0 \tag{60}$$

with

$$\mathbf{f}_n := \mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}})\big) - \mathcal{A}\big(\mathbf{H}(\tilde{\mathbf{u}}_n)\big)\tilde{\mathbf{u}}_n = \operatorname{div}\Big(\big(\mathbf{H}(\tilde{\mathbf{u}}) - \mathbf{H}(\tilde{\mathbf{u}}_n)\big)\nabla\tilde{\mathbf{u}}_n\Big) \text{ for } n \in \mathbb{N}.$$

Due to the Lipschitz continuity of $\mathcal{F}$ (cf. Lemma 11), we have

$$\mathbf{F}(\nabla\tilde{\mathbf{u}}_n) \to \mathbf{F}(\nabla\tilde{\mathbf{u}}) \text{ in } L^2\big(0,T;L^2(G,\mathbb{R}^{k\times d})\big) \text{ as } n \to \infty.$$

Hence, by virtue of Equation (55),

$$\mathbf{H}(\tilde{\mathbf{u}}_n) \to \mathbf{H}(\tilde{\mathbf{u}}) \text{ in } H^1\big(0,T;L^2(G,\mathbb{R}^{k\times d})\big) \hookrightarrow L^2\big(0,T;L^2(G,\mathbb{R}^{k\times d})\big) \text{ as } n \to \infty. \tag{61}$$

Using Assumption 10 to verify

$$\sup_{n\in\mathbb{N}} \Big\|\big(\mathbf{H}(\tilde{\mathbf{u}}) - \mathbf{H}(\tilde{\mathbf{u}}_n)\big)\nabla\tilde{\mathbf{u}}_n\Big\|_{L^2((0,T)\times G,\mathbb{R}^{k\times d})} \leq 2C \sup_{n\in\mathbb{N}} \|\nabla\tilde{\mathbf{u}}_n\|_{L^2((0,T)\times G,\mathbb{R}^{k\times d})} < \infty,$$

we apply Lebesgue's dominated convergence theorem to Equation (61) to obtain

$$\big(\mathbf{H}(\tilde{\mathbf{u}}) - \mathbf{H}(\tilde{\mathbf{u}}_n)\big)\nabla\tilde{\mathbf{u}}_n \to \mathbf{0} \text{ in } L^2\big((0,T)\times G,\mathbb{R}^{k\times d}\big) \text{ as } n \to \infty.$$

Hence, since div is a continuous linear mapping between the Hilbert spaces $L^2(G,\mathbb{R}^{k\times d})$ and $\mathcal{V}'$, we find

$$\operatorname{div}\Big(\big(\mathbf{H}(\tilde{\mathbf{u}}) - \mathbf{H}(\tilde{\mathbf{u}}_n)\big)\nabla\tilde{\mathbf{u}}_n\Big) \to \mathbf{0} \text{ in } L^2(0,T;\mathcal{V}') \text{ as } n \to \infty.$$

Therefore, by virtue of Theorem 18,

$$\|\bar{\mathbf{u}}_n\|_{L^2(0,T;\mathcal{H})} \leq \frac{1}{\kappa^2}\Big\|\operatorname{div}\Big(\big(\mathbf{H}(\tilde{\mathbf{u}}) - \mathbf{H}(\tilde{\mathbf{u}}_n)\big)\Big)\Big\|_{L^2(0,T;\mathcal{V}')}^2 \to 0 \text{ as } n \to \infty$$

implying $\mathscr{F}$ is continuous.

Applying the fixed point theorem: Now, by virtue of Schauder & Tychonoff's fixed point theorem, $\mathscr{F}$ posseses a fixed point $\bar{\mathbf{u}} \in L^2\big(0,T;\mathcal{V}\big)$ (not necessarily unique). Using Equations (54) and (56), we finally deduce $\bar{\mathbf{u}} \in H^1(0,T;\mathcal{V}')$. Letting $\bar{\mathbf{H}} := \mathbf{H}(\bar{\mathbf{u}})$, we easily verify $(\bar{\mathbf{u}},\bar{\mathbf{H}})^T$ satisfies Equations (51)–(53), which completes the proof. $\square$

Corollary 14. *Under the conditions of Theorem 13, let $\tilde{\mathbf{u}}^0 \in \mathcal{V}$. Any weak solution $(\mathbf{u},\mathbf{H})^T$ to Equations (51)–(53) given in Theorem 13 is then also a strong solution satisfying*

$$\mathbf{u} \in L^2\big(0,T;W^{1+s,2}(G,\mathbb{R}^k)\big) \text{ for any } s \in [0,1/2).$$

Proof. Strongness: Similar to Equation (50), we use the fundamental theorem of calculus together with Cauchy & Schwarz' inequality to estimate

$$
\left| a(\mathbf{u}, \mathbf{v}; t) - \tilde{a}(\mathbf{u}, \mathbf{v}; s) \right| = \left| \int_G \left(\big(\mathbf{H}(t, \cdot) - \mathbf{H}(s, \cdot) \big) \nabla \mathbf{u} \right) : (\nabla \mathbf{v}) \mathrm{d}\mathbf{x} \right|
$$

$$
= \int_s^t \int_G \| \partial_t \mathbf{H}(\xi, \mathbf{x}) \|_{\mathcal{S}(\mathbb{R}^{k \times d})} \| \mathbf{u} \|_{\mathcal{V}} \| \mathbf{v} \|_{\mathcal{V}} \mathrm{d}\mathbf{x} \mathrm{d}\xi
$$

$$
\leq \sqrt{t - s} \, \| \mathbf{H} \|_{W^{1,2}(0,T; L^\infty(G, \mathcal{S}(\mathbb{R}^{k \times d})))} \| \mathbf{u} \|_{\mathcal{V}} \| \mathbf{v} \|_{\mathcal{V}}
$$

for $0 \leq s \leq t \leq T$ and $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, where

$$
a(\mathbf{u}, \mathbf{v}; t) := \int_G (\mathbf{H} \nabla \mathbf{u}) : (\nabla \mathbf{v}) \mathrm{d}\mathbf{x}.
$$

This together with the assumption $\mathbf{u}^0 \in \mathcal{V}$ enables us to deduce $\mathbf{u} \in \mathcal{MR}_a(\mathcal{H})$ by virtue of Theorem 18.

Extra regularity: Applying [33, Theorem 4] to the following family of elliptic problems

$$
\mathcal{A}\big(\mathbf{H}(t, \cdot)\big) = \mathbf{g}(t, \cdot) \text{ for a.e. } t \in [0, T] \text{ with } \mathbf{g} = \partial_t \mathbf{u} \in L^2(0, T; \mathcal{H}),
$$

the desired regularity follows. □

Concerning the uniqueness for Equations (51)–(53), no existence results are known in the literature both for weak and strong solutions in sense of Definition 12 (cf. [10]). Same is true for the quasilinear heat equation in non-divergence form (see [7]), etc. Nonetheless, under a boundedness condition for $\nabla \mathbf{u}$, the following uniqueness result can be proved.

Theorem 15. *Let* $(\mathbf{u}_1, \mathbf{H}_1)^T$, $(\mathbf{u}_2, \mathbf{H}_2)^T$ *be two weak solutions to Equations (51)–(53) such that* $\nabla \mathbf{u}_1, \nabla \mathbf{u}_2 \in L^\infty\big((0, T) \times G, \mathbb{R}^{k \times d}\big)$. *Then* $\mathbf{u}_1 \equiv \mathbf{u}_2$ *a.e. in* $(0, T) \times G$.

Proof. Letting $\bar{\mathbf{u}} := \mathbf{u}_1 - \mathbf{u}_2$, $\bar{\mathbf{H}} := \mathbf{H}_1 - \mathbf{H}_1$, we observe that $(\bar{\mathbf{u}}, \bar{\mathbf{H}})^T$ satisfies

$$
\partial_t \bar{\mathbf{u}} - \mathrm{div}\left(\mathbf{H}_1 \nabla \bar{\mathbf{u}} \right) - \mathrm{div}\left(\bar{\mathbf{H}} \nabla \mathbf{u}_2 \right) = \mathbf{0} \text{ in } L^2(0, T; \mathcal{V}'), \tag{62}
$$

$$
\tau \partial_t \bar{\mathbf{H}} + \bar{\mathbf{H}} - \left(\mathbf{F}(\nabla \mathbf{u}_1) - \mathbf{F}(\nabla \mathbf{u}_2) \right) = \mathbf{0} \text{ in } L^2\Big(0, T; L^\infty\big(G, \mathcal{S}(\mathbb{R}^{d \times 2}) \big) \Big), \tag{63}
$$

$$
\bar{\mathbf{u}}(0, \cdot) = \mathbf{0} \text{ in } \mathcal{V}, \quad \bar{\mathbf{H}}(0, \cdot) = \mathbf{0} \text{ in } L^\infty\big(G, \mathcal{S}(\mathbb{R}^{d \times 2}) \big). \tag{64}
$$

Multiplying Equation (62) with $\bar{\mathbf{u}}$ in $L^2\big(0, T; L^2(G, \mathbb{R}^d)\big)$, using Green's formula and exploiting the uniform positive definiteness of $\mathbf{H}(t, \cdot)$, we obtain using Hölder's and Young's inequalities

$$
\begin{aligned}
\big\| \bar{\mathbf{u}}(t, \cdot) \big\|_{\mathcal{H}}^2 &\leq -2\kappa \int_0^t \big\| \nabla \bar{\mathbf{u}}(s, \cdot) \big\|_{L^2(G, \mathbb{R}^{k \times d})}^2 \mathrm{d}s + \int_0^t \big\| (\bar{\mathbf{H}} \nabla \mathbf{u}_2) : \nabla \bar{\mathbf{u}} \big\|_{L^1(G)} \mathrm{d}s \\
&\leq -2\kappa \int_0^t \big\| \nabla \bar{\mathbf{u}}(s, \cdot) \big\|_{L^2(G, \mathbb{R}^{k \times d})}^2 \mathrm{d}s + \big\| \nabla \mathbf{u}_2 \big\|_{L^\infty((0,T) \times G, \mathbb{R}^{k \times d})} \times \\
&\quad \times \int_0^t \big\| \bar{\mathbf{H}} \big\|_{L^2(G, \mathcal{S}(\mathbb{R}^{k \times d}))} \big\| \nabla \bar{\mathbf{u}} \big\|_{L^2(G, \mathbb{R}^{k \times d})} \mathrm{d}s \\
&\leq -\kappa \int_0^t \big\| \nabla \bar{\mathbf{u}}(s, \cdot) \big\|_{L^2(G, \mathbb{R}^{k \times d})}^2 \mathrm{d}s + \tilde{C}_1 \int_0^t \big\| \bar{\mathbf{H}}(s, \cdot) \big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 \mathrm{d}s,
\end{aligned} \tag{65}
$$

where $\tilde{C}_1$ depends on the L^∞-norm of $\nabla \mathbf{u}_2$. Further, multiplying Equation (63) with $\bar{\mathbf{H}}(t, \cdot)$ in $L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})$ as well as exploiting Cauchy & Schwarz' and Young's inequalities, we estimate

$$\tau \partial_t \big\| \bar{\mathbf{H}}(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})} \leq \tilde{C}_2 \big\| \bar{\mathbf{H}}(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})} + \kappa \big\| \nabla \bar{\mathbf{u}}(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{k\times d})} \quad (66)$$

for some $\tilde{C}_2 > 0$. Integrating Equation (66) w.r.t. to t and adding the result to Equation (65), we get

$$\big\| \bar{\mathbf{u}}(t, \cdot) \big\|^2_{\mathcal{H}} + \tau \big\| \bar{\mathbf{H}}(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})}$$
$$\leq \tilde{C} \int_0^t \left(\big\| \bar{\mathbf{u}}(s, \cdot) \big\|^2_{\mathcal{H}} + \tau \big\| \bar{\mathbf{H}}(s, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})} \right) \mathrm{d}s$$

for some $\tilde{C} > 0$. Now, the claim follows by virtue of Gronwall's inequality. $\qquad \square$

In a similar fashion, we can prove:

Corollary 16. *Under the conditions of Theorem 15, for any $T > 0$, there exists a constant $\tilde{C} > 0$ such that*

$$\max_{0 \leq t \leq T} \left(\big\| \mathbf{u}_1(t, \cdot) - \mathbf{u}_2(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^k)} + \big\| \mathbf{H}_1(t, \cdot) - \mathbf{H}_2(t, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})} \right)$$
$$\leq \tilde{C} \left(\big\| \mathbf{u}_1(0, \cdot) - \mathbf{u}_2(0, \cdot) \big\|^2_{L^2(G, \mathbb{R}^k)} + \big\| \mathbf{H}_1(0, \cdot) - \mathbf{H}_2(0, \cdot) \big\|^2_{L^2(G, \mathbb{R}^{(k\times d)\times(k\times d)})} \right).$$

For a global weak solution $(\mathbf{u}, \mathbf{H})^T$ to Equations (51)–(53) given in Theorem 13, consider the energy functional

$$\mathcal{E}(t) := \frac{1}{2} \int_G \big\| \mathbf{u}(t, \cdot) \big\|^2_{\mathbb{R}^k} \mathrm{d}\mathbf{x} + \frac{\tau}{2} \int_G \big\| \mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0}) \big\|^2_{\mathbb{R}^{(k\times d)\times(k\times d)}} \mathrm{d}\mathbf{x}. \quad (67)$$

Theorem 17. *In addition to the assumptions of Corollary 14, there may exist a number $\omega > 0$ such that*

$$\mathbf{F}(\mathbf{D}) \in \mathcal{S}_{\geq \omega}(\mathbb{R}^{k\times d}). \quad (68)$$

The energy functional defined in Equation (67) decays then exponentially along any strong solution of Equations (51)–(53), i.e.,

$$\mathcal{E}(t) \leq C \exp(-2\beta t) \mathcal{E}(0) \text{ for a.e. } t \geq 0 \text{ with appropriate } C, \beta > 0,$$

which implies

$$\lim_{t \to 0} (\mathbf{u}, \mathbf{H})^T(t, \cdot) = \big(\mathbf{0}, \mathbf{F}(\mathbf{0})\big)^T \text{ in } L^2(G, \mathbb{R}^k) \times L^2\big(G, \mathbb{R}^{(k\times d)\times(k\times d)}\big).$$

Proof. Solving Equation (52) for $\mathbf{H}$, we obtain

$$\mathbf{H}(t, \cdot) = \exp\big(-t/\tau\big)\mathbf{H}^0 + \int_0^t \exp\big(-(t-s)/\tau\big)\mathbf{F}\big(\nabla \mathbf{u}(s, \cdot)\big)\mathrm{d}s.$$

Using assumption $\mathbf{H}^0 \in \mathcal{S}_{\geq \alpha}(\mathbb{R}^{k \times d})$ of Theorem 13 and Equation (68), this implies

$$
\begin{aligned}
\min \sigma\big(\mathbf{H}(t, \cdot)\big) &\geq \alpha \exp\big(-t/\tau\big) + \omega \int_0^t \exp\big(-(t-s)/\tau\big)\mathrm{d}s \\
&\geq \alpha \exp\big(-t/\tau\big) + \frac{\omega}{\tau}\big(1 - \exp\big(-t/\tau\big)\big) \\
&\geq \min\big\{\alpha \exp(-1), \omega/(2\tau)\big\} =: \kappa > 0
\end{aligned}
\tag{69}
$$

with $\sigma\big(\mathbf{H}(t, \cdot)\big)$ denoting the spectrum of $\mathbf{H}(t, \cdot)$.

For a.e. $t \geq 0$, multiplying Equation (24) in $L^2(G, \mathbb{R}^k)$ with $\mathbf{u}(t, \cdot)$, using Green's formula, utilizing the boundary conditions (26) and taking into account Equation (69), we obtain

$$
\frac{1}{2}\partial_t \big\|\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^k)}^2 \leq -\kappa \big\|\nabla\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^{k \times d})}^2.
$$

By virtue of second Poincaré's inequality, this implies

$$
\frac{1}{2}\partial_t \big\|\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^k)}^2 \leq -\frac{\kappa}{2}\big\|\nabla\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^{k \times d})}^2 - \frac{\kappa C_P}{2}\big\|\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^k)}^2 \text{ for } t > 0.
\tag{70}
$$

Subtracting $\mathbf{F}(\mathbf{0})$ from Equation (25), we get

$$
\tau\partial_t\big(\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big) + \big(\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big) = \mathbf{F}\big(\nabla\mathbf{u}(t, \cdot)\big) - \mathbf{F}(\mathbf{0}) \text{ for a.e. } t > 0.
\tag{71}
$$

Hence, multiplying Equation (71) in $L^2\big(G, \mathbb{R}^{(k \times d) \times (k \times d)}\big)$ with $\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})$ and using Assumption 10, we arrive at

$$
\begin{aligned}
\frac{\tau}{2}\partial_t\big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 &+ \big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 \\
&\leq c\big\|\nabla\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^{k \times d})}\big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}.
\end{aligned}
$$

Now, using Young's inequality, we estimate

$$
\begin{aligned}
\frac{\tau}{2}\partial_t\big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 \\
\leq -\frac{1}{2}\big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 + \frac{c^2}{2}\big\|\nabla\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^{k \times d})}.
\end{aligned}
\tag{72}
$$

Multiplying Equation (72) with $\frac{\kappa C_P}{c^2}$ and adding the result to Equation (70) yields

$$
\begin{aligned}
\partial_t \mathcal{E}(t) &\leq -\frac{\kappa C_P}{2c^2}\big\|\mathbf{u}(t, \cdot)\big\|_{L^2(G, \mathbb{R}^k)}^2 - \frac{1}{2}\big\|\mathbf{H}(t, \cdot) - \mathbf{F}(\mathbf{0})\big\|_{L^2(G, \mathbb{R}^{(k \times d) \times (k \times d)})}^2 \\
&\leq -\min\big\{\kappa C_P/c^2, 1/\tau\big\}\mathcal{E}(t) \text{ for a.e. } t \geq 0.
\end{aligned}
$$

Hence, the exponential decay of $\mathcal{E}$ is a direct consequence of Gronwall's inequality. $\qquad\square$

A   Maximal L^2-Regularity for Non-Autonomous Forms

In this appendix, we briefly summarize the theory of maximal L^2-regularity for non-autonomous forms. We start with the classical theory dating back to Dautray & Lions (cf.

[18]), which furnishes the existence and uniqueness of weak solutions. Further, we present a recent theory developed by Dier in [19] guaranteeing the existence of strong solutions under a boundedness assumption on the variation of non-autononous form associated with the 'elliptic' part of evolution problem.

Let $\mathcal{H}$ and $\mathcal{V}$ be separable Hilbert spaces such that $\mathcal{V}$ is continuously and densely embedded into $\mathcal{H}$. For $T > 0$, we consider the initial value problem

$$\dot{u}(t) + \mathcal{A}(t)u(t) = f(t) \text{ in } L^2(0, T; \mathcal{V}'), \quad u(0) = u^0 \in \mathcal{H}. \tag{73}$$

Further, let

$$a\colon [0, T] \times \mathcal{V} \times \mathcal{V} \to \mathbb{C}, \quad (t, u, v) \mapsto a(u, v; t) \tag{74}$$

be a non-autonomous sesquilinear form, i.e., $a(u, v; \cdot)$ is measurable for all $u, v \in \mathcal{V}$ and $a(\cdot, \cdot; t)$ is sesquilinear for a.e. $t \in [0, T]$. Let a be continuous in u and v uniformly w.r.t. t, i.e., there may exist a number $M > 0$ such that

$$|a(u, v; t)| \leq M\|u\|_{\mathcal{V}}\|v\|_{\mathcal{V}} \text{ for all } u, v \in \mathcal{V} \text{ and a.e. } t \in [0, T].$$

Additionally, let a be uniformly coercive, i.e., there may exist some number $\alpha > 0$ such that

$$\operatorname{Re} a(u, u; t) \geq \alpha\|u\|_{\mathcal{V}}^2 \text{ for any } u \in \mathcal{V} \text{ and a.e. } t \in [0, T].$$

With $\mathcal{V}'$ denoting the antidual of $\mathcal{V}$, the linear bounded operator $\mathcal{A}(t)\colon \mathcal{V} \to \mathcal{V}'$ associated with $a(\cdot, \cdot; t)$ for $t \in [0, T]$ is defined as

$$\langle \mathcal{A}(t)u, v \rangle_{\mathcal{V}';\mathcal{V}} := a(u, v; t) \text{ for } u, v \in \mathcal{V}.$$

A classical result due to Lions states the following well-posedness result in the class of weak solutions.

Theorem 18. *For every $f \in L^2(0, T; \mathcal{V}')$ and $u_0 \in \mathcal{H}$, there exists a unique weak solution*

$$u \in L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{V}')$$

to the initial value problem (73). Moreover, we have the continuous embedding

$$L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{V}') \hookrightarrow C^0([0, T], \mathcal{H}) \tag{75}$$

and the estimate

$$\|u\|_{L^2(0,T;\mathcal{V})}^2 \leq \frac{1}{\alpha^2}\|f\|_{L^2(0,T;\mathcal{V}')}^2 + \frac{1}{\alpha}\|u_0\|_{\mathcal{H}}^2.$$

For the weak solution to be strong, additional assumptions on the non-autonomous form a are required. In the following, let the non-autonomous form a be of bounded variation, i.e., there may exist a nondecreasing function $g\colon [0, T] \to [0, \infty)$ such that

$$|a(u, v; t) - a(u, v; s)| \leq (g(t) - g(s))\|u\|_{\mathcal{V}}\|v\|_{\mathcal{V}} \text{ for all } u, v \in \mathcal{V} \text{ and } 0 \leq s \leq t \leq T.$$

The maximal regularity class for the operator family $(\mathcal{A}(t))_{t \in [0,T]}$ is then defined as

$$\mathcal{MR}_a(\mathcal{H}) := \{u \in L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{H}) \mid \mathcal{A}u \in L^2(0, T; \mathcal{H})\}.$$

Under conditions above, [19, Section 4] provides the following well-posedness result.

Theorem 19. *For every $f \in L^2(0, T; \mathcal{H})$ and $u^0 \in \mathcal{V}$, there exists a unique strong solution $u \in \mathcal{MR}_a(\mathcal{H})$ to the initial value problem (73). Moreover, $\mathcal{MR}_a(\mathcal{H}) \hookrightarrow C^0([0, T], \mathcal{V})$ and*

$$\|u\|_{L^\infty(0,T;\mathcal{V})}^2 \leq \frac{1}{\alpha}\left(\|f\|_{L^2(0,T;\mathcal{H})}^2 + M\|u_0\|_{\mathcal{V}}^2\right)\exp\left(\frac{1}{\alpha}(g(T) - g(0))\right).$$

Acknowledgment

References

[1] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel. Axioms and fundamental equations of image processing. *Archive for Rational Mechanics and Analysis*, 123:199–257, 1993.

[2] H. Amann. Compact embeddings of vector-valued Sobolev and Besov spaces. *Glasnik Matematički*, 35(55):161–177, 2000.

[3] H. Amann. Non-local quasi-linear parabolic equations. *Russian Mathematical Surveys*, 60(6):1021–1033, 2005.

[4] H. Amann. Time-delayed Perona-Malik type problems. *Acta Mathematica Universitatis Comenianae*, LXXVI(1):15–38, 2007.

[5] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variational flow. *Differential and Integral Equations*, 14(3):321–360, 2001.

[6] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Some qualitative properties for the total variation flow. *Journal of Functional Analysis*, 188:516–547, 2002.

[7] W. Arendt and R. Chill. Global existence for quasilinear diffusion equations in isotropic nondivergence form. *Annali della Scuola Normale Superiore di Pisa (5)*, IX:523–539, 2010.

[8] V. Barbu. *Nonlinear differential equations of monotone types in Banach spaces.* Springer Monographs in Mathematics. Springer Science & Business Media, New York Dordrecht Heidelberg London, 2010.

[9] A. Belahmidi. *Équations aux dérivées partielles appliquées à la restauration et à l'agrandissement des images.* PhD thesis, Université de Paris-Dauphine, Paris, 2003.

[10] A. Belahmidi and A. Chambolle. Time-delay regularization of anisotropic diffusion and image processing. *ESAIM: Mathematical Modelling and Numerical Analysis*, 39(2):231–251, 2005.

[11] A. Belleni-Morante and A. C. McBride. *Applied Nonlinear Semigroups: An Introduction.* Wiley Series in Mathematical Methods in Practice. John Wiley & Sons, Chichester, 1998.

[12] G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in $\mathbb{R}^N$. *Journal of Differential Equations*, 184:475–525, 2002.

[13] M. Burger, A. C. G. Menucci, S. Osher, and M. Rumpf, editors. *Level Set and PDE Based Reconstruction Methods in Imaging*, volume 2090 of *Lecture Notes in Mathematics*. Springer International Publishing, Switzerland, 1992.

[14] J. Canny. Finding edges and lines in images. Technical Report 720, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Boston, MA, 1983.

[15] G. R. Cattaneo. Sur une forme de l'équation de la chaleur éliminant le paradoxe d'une propagation instantanée. *Comptes Rendus de l'Académie des Sciences*, 247(4):431–433, 1958.

[16] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.

[17] G. H. Cottet and M. El Ayyadi. A Volterra type model for image processing. *IEEE Transactions on Image Processing*, 7(3):292–303, 1998.

[18] R. Dautray and J.-L. Lions. *Evolution Problems*, volume 5 of *Mathematical Analysis and Numerical Methods for Science and Technology*. Springer-Verlag, Berlin, 1992.

[19] D. Dier. Non-autonomous maximal regularity for forms of bounded variation. *Journal of Mathematical Analysis and Applications*, 425:33–54, 2015.

[20] M. E. Gurtin and A. C. Pipkin. A general theory of heat conduction with finite wave speeds. *Archive for Rational Mechanics and Analysis*, 31:113–126, 1968.

[21] A. Handlovičová, K. Mikula, and F. Sgallari. Variational numerical methods for solving nonlinear diffusion equations arising in image processing. *Journal of Visual Communication and Image Representation*, 13:217–237, 2002.

[22] M. Hieber and M. Murata. The L^p-approach to the fluid-rigid body interaction problem for compressible fluids. *Evolution Equations and Control Theory*, 4:69–87, 2015.

[23] M. Hochbruck, T. Jahnke, and R. Schnaubelt. Convergence of an ADI splitting for Maxwell's equations. *Numerische Mathematik*, 129:535–561, 2015.

[24] S. L. Keeling and R. Stollberger. Nonlinear anisotropic diffusion filtering for multi-scale edge enhancement. *Inverse Problems*, 18:175–190, 2002.

[25] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society B*, 207(1167):187–217, 1980.

[26] S. A. Morris. The Schauder-Tychonoff fixed point theorem and applications. *Matematický Časopis*, 25(2):165–172, 1975.

[27] M. Nitzberg and T. Shiota. Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):826–833, 1992.

[28] T. Ohkubo. Regularity of solutions to hyperbolic mixed problems with uniformly characteristic boundary. *Hokkaido Mathematical Journal*, 10:93–123, 1981.

[29] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Machine Intell.*, 12:629–630, 1990.

[30] J. Prüss. Maximal regularity of linear vector-valued parabolic Volterra equations. *Journal of Integral Equations and Applications*, 3(1):63–83, 1991.

[31] J. Prüss. *Evolutionary Integral Equations and Applications*, volume 87 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1993.

[32] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[33] G. Savaré. Regularity results for elliptic equations in Lipschitz domains. *Journal of Functional Analysis*, 152:176–201, 1998.

[34] D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey, 2nd edition, 2015.

[35] P. Secchi. Well-posedness of characteristic symmetric hyperbolic systems. *Archive for Rational Mechanics and Analysis*, 134:155–197, 1996.

[36] K. Takezawa. *Introduction to Nonparametric Regression.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[37] J. Weickert. *Anisotropic Diffusion in Image Processing.* B. G. Teubner, Stuttgart, 1998.

[38] A. P. Witkin. Scale-space filtering. In *Proceedings of IJCAI*, pages 1019–1021, Karlsruhe, 1983.

[39] R. Zacher. Maximal regularity of type L_p for abstract parabolic Volterra equations. *Journal of Evolution Equations*, 5:79–103, 2005.