www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Optimal control of an energy-critical semilinear wave equation in 3D with spatially integrated control constraints

K. Kunisch, H. Meinlschmidt

RICAM-Report 2019-09

www.ricam.oeaw.ac.at

Optimal control of an energy-critical semilinear wave equation in 3D with spatially integrated control constraints

Karl Kunisch[a,b], Hannes Meinlschmidt[b,*]

[a] *Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, 8010 Graz, Austria*

[b] *Johann Radon Institute for Computational and Applied Mathematics (RICAM), Altenberger Straße 69, 4040 Linz, Austria*

Abstract

This paper is concerned with an optimal control problem subject to the H^1-critical defocusing semilinear wave equation on a smooth and bounded domain in three spatial dimensions. Due to the criticality of the nonlinearity in the wave equation, unique solutions to the PDE obeying energy bounds are only obtained in special function spaces related to Strichartz estimates and the nonlinearity. The optimal control problem is complemented by pointwise-in-time constraints of Trust-Region type $\|u(t)\|_{\mathrm{L}^2(\Omega)} \leq \omega(t)$. We prove existence of globally optimal solutions to the optimal control problem and give optimality conditions of both first- and second order necessary as well as second order sufficient type. A nonsmooth regularization term for the natural control space $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$, which also promotes sparsity in time of an optimal control, is used in the objective functional.

Keywords: Optimal control of PDEs, Critical Wave Equation, Second-order optimality conditions, Nonsmooth regularization,
2010 MSC: 35L05, 35L71, 49J20, 49K20

*Corresponding author
Email addresses: `karl.kunisch@uni-graz.at` (Karl Kunisch), `hannes.meinlschmidt@ricam.oeaw.ac.at` (Hannes Meinlschmidt)

1. Introduction

We consider the optimal control problem

$$\begin{aligned} &\min_{y,u} \quad \ell(y,u) \\ &\text{s.t.} \quad \begin{cases} u \in \mathcal{U}_{\mathrm{ad}}, \\ y \text{ is the solution to (CWE)}, \end{cases} \end{aligned} \tag{OCP}$$

where the underlying partial differential equation is the H^1-critical defocusing wave equation on a bounded domain Ω with smooth boundary in three spatial dimensions over a finite interval $(0, \mathsf{T})$, complemented with homogeneous Dirichlet boundary conditions, in the prototype form

$$\left.\begin{aligned} \partial_t^2 y - \Delta y + y^5 &= u && \text{in } (0,\mathsf{T}) \times \Omega, \\ y &= 0 && \text{on } (0,\mathsf{T}) \times \partial\Omega, \\ \big(y(0), \partial_t y(0)\big) &= (y_0, y_1) && \text{in } \Omega. \end{aligned}\right\} \tag{CWE}$$

We suppose that $\xi_0 := (y_0, y_1) \in \mathrm{H}_0^1(\Omega) \times \mathrm{L}^2(\Omega) =: \mathcal{E}$ and $u \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$, which is the natural H^1-setting for the wave equation. The performance index ℓ for (CWE) is chosen to be

$$\ell(y,u) := \frac{1}{2}\|y(\mathsf{T}) - y_d\|^2_{\mathrm{L}^2(\Omega)} + \frac{\gamma}{4}\|y\|^4_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))} \\ + \beta_1 \|u\|_{\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))} + \frac{\beta_2}{2}\|u\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}$$

for $y_d \in \mathrm{L}^2(\Omega)$ and scaling parameters γ, β_1, β_2. The objective in (OCP) is thus to find a control $u \in \mathcal{U}_{\mathrm{ad}}$ such that the associated solution to (CWE) $y(\mathsf{T})$ at time T matches a given profile y_d as well as possible in the L^2-sense. The parameters γ, β_1 and β_2 in ℓ are nonnegative. It will be specifically mentioned if their positivity is required. While the $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$ term describes a quadratic control cost, the $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ term is known to be sparsity enhancing. The purpose of the $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$ norm term will become clear later. The constraint set $\mathcal{U}_{\mathrm{ad}}$ is of the form

$$\mathcal{U}_{\mathrm{ad}} := \Big\{ v\colon (0,\mathsf{T}) \to \mathrm{L}^2(\Omega)\colon \|v(t)\|_{\mathrm{L}^2(\Omega)} \le \omega(t) \text{ f.a.a. } t \in (0,\mathsf{T}) \Big\}$$

for a measurable function ω which is nonnegative almost everywhere on $(0,\mathsf{T})$. It models a maximum overall input power at every time $t \in (0,\mathsf{T})$ for the controls u. We emphasize that ω is not assumed to be bounded away from 0 uniformly almost everywhere. It is thus possible to model e.g. a forced soft "shut-off" or decay to zero of $\|u(t)\|_{\mathrm{L}^2(\Omega)}$ as $t \searrow T$ for some $T \in (0,\mathsf{T}]$.

Context

The state equation (CWE) is a semilinear wave equation. Such equations are of interest in several areas of natural sciences, in particular in relation to mathematical physics [20, 23, 26, 32, 39, 40], in nonlinear elasticity [29], and the theory of vibrating strings [35].

The exponent 5 in the power-law nonlinearity y^5 in (CWE) is the H^1-critical one since it satisfies $5 = \frac{n+2}{n-2}$, with $n = 3$ being the space dimension. This terminology stems from the case $\Omega = \mathbb{R}^n$ where "critical" implies that (classical) solutions y to (CWE) with $u = 0$ are invariant under the scaling $\lambda \mapsto \lambda^{-\frac{n-2}{2}} y(\frac{t}{\lambda}, \frac{x}{\lambda})$ and thereby preserve $\|\xi_0\|_{\mathcal{E}}$, cf. e.g. [41, Ch. 3.1]. A major difficulty here is that one does not obtain a bound on $y \in \mathrm{L}^5(0, \mathsf{T}; \mathrm{L}^{10}(\Omega))$ and thus on $y^5 \in \mathrm{L}^1(0, \mathsf{T}; \mathrm{L}^2(\Omega))$, which makes global-in-time existence difficult to prove. It turns out that a uniform bound on $y \in \mathrm{L}^4(0, \mathsf{T}; \mathrm{L}^{12}(\Omega))$ is also sufficient due to interpolation and $\mathrm{L}^\infty(0, \mathsf{T}; \mathrm{L}^6(\Omega))$ energy conservation. The nonlinearity is defocusing due to its sign, which does not play a role for local existence of solutions to (CWE), but is crucial for long-time, i.e., global-in-time, existence.

There is a large body of rather recent work about the analysis of solutions to critical wave equations and their global-in-time existence, including monographs (partially) dedicated to the topic such as [41] or [38]. We focus on the works on space dimension 3. Historically, global existence was proven first for $\Omega = \mathbb{R}^3$ by Grillakis [19] in 1992 for smooth solutions and by Shatah and Struwe [33, 34] for energy space solutions in 1993/1994. Smith and Sogge were able to extend this result to the exterior of convex obstacles [37], and finally the case of a bounded domain Ω with Dirichlet conditions was treated by Burq, Lebeau, and Planchon [3] in 2008. This was followed by the treatment of the Neumann conditions case by Burq and Planchon [4]. By now, improved Strichartz estimates compared to the ones which were available in [3] are proven by Blair, Smith and Sogge [2] and these allow for a slightly more convenient existence proof as outlined in [38, Ch. IV]. Of course, this is all for the defocusing case since the mentioned works establish global existence for initial data of arbitrary size. Extensions of the mentioned results and especially the focusing case are in the focus of current research; we mention exemplarily [11, 13, 14, 21, 25].

Semilinear wave equations have attracted significant interest in the classical control theory community and are a subject of ongoing research. Let us just mention for example [8, 9, 22, 27], where subcritical nonlinearities are considered, and the more recent works [6, 24] for the critical case. On the other hand, the literature regarding optimal control of semilinear wave equations and especially the one about stronger nonlinearities appears to be rather scarce. We refer to [31] and related works, where the focus lies more on the state constraints imposed on the system, or to [16] with a mild nonlinearity. We are not aware of any work related to the optimal control of a *critical* semilinear wave equation.

The contributions of this work are thus threefold:

- Up to now, the proofs of global existence of solutions to (CWE) mentioned above do not incorporate forcing terms or controls u. We thus explicitly revisit the proof in [3] to obtain a global existence result including the control. Due to the lack of a uniform bound on the nonlinearity, as mentioned above, the proof is quite sophisticated and taylored to the critical nature of the problem.

- We show existence of globally optimal solutions to (OCP) and derive optimality conditions of both first-and second order type. Again, a particular point is that there is no uniform bound on the $\mathrm{L}^1(0, \mathsf{T}; \mathrm{L}^2(\Omega))$ norm of y^5 for varying controls u from the PDE (CWE) which we thus have to enforce using the cost functional ℓ with $\gamma > 0$. The derivative of the associated term then enters the optimality conditions.

- We consider the constraint set of pointwise-in-time Trust Region type $\mathcal{U}_{\mathrm{ad}}$ for the optimal control problem which seems not to have received much attention in the available literature so far. This plays a most prominent and demanding role in the derivation of second order necessary optimality conditions for (OCP). It is of independent interest and a novel contribution in its own right.

Let us also point out that we could also consider more general nonlinearities $f(y)$ in place of y^5 in (CWE), as long as they exhibit comparable growth and continuity properties. We have chosen to omit the technical details for the sake of exposition.

Overview

We first establish several basic but fundamental results about solutions to (CWE) in Section 2. This includes the concept of mild and weak solutions to (CWE) and several important estimates for solutions to wave equations. It ends with local-in-time wellposedness of (CWE). As announced above, the solution regularity will then be $(y, y') \in \mathrm{C}([0, T^\bullet); \mathcal{E})$ with $y \in \mathrm{L}^4_{\mathrm{loc}}([0, T^\bullet); \mathrm{L}^{12}(\Omega))$ for some $T^\bullet \in (0, \mathsf{T}]$ but there will be no uniform bound in the latter space. In Section 3, we establish that such local-in-time solutions in fact exist globally in time in the energy space class by incorporating the inhomogeneity u into the related proof in [3]. The optimal control problem is treated in Section 4. After some preparatory differentiability results, we prove existence of globally optimal controls for (OCP) as well as necessary optimality conditions of first and second order, and also second order sufficient conditions. For the latter, we need to assume that $\beta_2 > 0$.

Notation and conventions

We already mentioned above that we often consider the energy space $\mathcal{E} := \mathrm{H}^1_0(\Omega) \times \mathrm{L}^2(\Omega)$ and throughout equip $\mathrm{H}^1_0(\Omega)$ with the norm $\|y\|_{\mathrm{H}^1_0(\Omega)} := \|\nabla y\|_{\mathrm{L}^2(\Omega)}$. Moreover, for a time-dependent function y, we write

$$\xi_y(t) := \big(y(t), \partial_t y(t)\big) \quad \text{and} \quad \big\|\xi_y(t)\big\|^2_{\mathcal{E}} := \big\|\nabla y(t)\big\|^2_{\mathrm{L}^2(\Omega)} + \big\|\partial_t y(t)\big\|^2_{\mathrm{L}^2(\Omega)}.$$

We use $(\cdot,\cdot)_\Omega$ for the $\mathrm{L}^2(\Omega)$ inner product and write $A \lesssim B$ if there is a constant $C > 0$ such that $A \leq C \cdot B$. If necessary, a dependency of the constant C on another quantity D will be denoted by $A \lesssim_D B$. All other notation will be standard. We consider all function spaces to be real ones.

Lastly, the global time interval length T is given and fixed for this work, but we will sometimes have to deal with the theory of functions or solutions on intervals other than $[0, \mathsf{T}]$. In these cases, let $T \in (0, \mathsf{T}]$, and consider, if necessary, implicitly a dilation of $[t_0, t_1] \subseteq [0, \mathsf{T}]$ to $[0, T]$ via $T := t_1 - t_0$.

2. Existence and uniqueness of local solutions

The classical notion of a solution in the energy space $\mathcal{E}$ is that of a *mild solution.* For this, let us introduce the Laplacian operator Δ in $\mathrm{L}^2(\Omega)$ given by

$$\begin{aligned} D(\Delta) &:= \Big\{\varphi \in \mathrm{H}^1_0(\Omega) \colon \exists f \in \mathrm{L}^2(\Omega) \colon \big(\nabla\varphi, \nabla\psi\big)_\Omega = \big(f, \psi\big)_\Omega \text{ for all } \psi \in \mathrm{H}^1_0(\Omega)\Big\}, \\ \Delta\varphi &:= -f. \end{aligned}$$

Definition 2.1 (Mild solution)**.** We say that the function y is a *mild solution* to (CWE) on $[0,T]$ if $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$ and $\xi_y \in \mathrm{C}([0,T];\mathcal{E})$ with $\int_0^t \xi_y(s)\,\mathrm{d}s \in D(\Delta)\times \mathrm{H}_0^1(\Omega)$ for all $t \in (0,T]$ and

$$\xi_y(t) = \xi_0 + \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta & 0 \end{pmatrix} \int_0^t \xi_y(s)\,\mathrm{d}s + \int_0^t \begin{pmatrix} 0 \\ u(s) - y^5(s) \end{pmatrix} \mathrm{d}s \quad \text{for all } t \in [0,T] \tag{2.1}$$

is satisfied.

Remark 2.2. Let us briefly comment on the—at first glance—somewhat curious requirement $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$ in the definition of a mild solution. By Sobolev embedding, we have $\mathrm{H}_0^1(\Omega) \hookrightarrow \mathrm{L}^6(\Omega)$, so a mild solution y can be considered as an element of $\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega)) \cap \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$. The Hölder inequality shows that

$$\|f\|^5_{\mathrm{L}^5(0,T;\mathrm{L}^{10}(\Omega))} \le \|f\|^4_{\mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))} \|f\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))} \tag{2.2}$$

and thus

$$\|f^5\|_{\mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))} \le \|f\|^4_{\mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))} \|f\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))} \tag{2.3}$$

for all $f \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega)) \cap \mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))$. This shows that for a mild solution y on $[0,T]$ we have $y^5 \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$ such that the second row in (2.1) is in fact self-consistent. Of course, we could have required the weaker condition $y \in \mathrm{L}^5(0,T;\mathrm{L}^{10}(\Omega))$ instead at this point; we will see the benefit of the stronger requirement later.

Following [24], we moreover introduce the following concept of solution for the problem (CWE) which is more suited to the optimal control problem. It is named after the authors of [33, 34].

Definition 2.3 (Shatah-Struwe solution)**.** We say that the function y is a *Shatah-Struwe solution* to (CWE) on $[0,T]$ if $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$ and $\xi_y \in \mathrm{L}^\infty(0,T;\mathcal{E})$ with $\xi_y(0) = \xi_0$ satisfies the weak formulation

$$\begin{aligned} -\int_0^T \big(\partial_t y(t), \partial_t \varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^T \big(\nabla y(t), \nabla\varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^T \big(y^5(t), \varphi(t)\big)_\Omega \,\mathrm{d}t \\ = \int_0^T \big(u(t), \varphi(t)\big)_\Omega \,\mathrm{d}t \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty\big((0,T)\times\Omega\big). \end{aligned}$$

Note that the notion of a Shatah-Struwe solution in fact also makes sense if y does not have the additional integrability property. It is also meaningful for initial data only from $\mathrm{H}^{-1}(\Omega)\times\mathrm{L}^2(\Omega)$, since the solution will be continuous with values in $\mathrm{H}^{-1}(\Omega)\times\mathrm{L}^2(\Omega)$, and for $u \in \mathrm{L}^1(0,T;\mathrm{H}^{-1}(\Omega)$. For our purpose, we had supposed $\xi_0 \in \mathcal{E} = \mathrm{H}_0^1(\Omega)\times\mathrm{L}^2(\Omega)$ and $u \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$; thus, we can in fact show that Shatah-Struwe and mild solutions coincide and are unique. The additional integrability property $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$ is also important here. The first step is to establish conservation of energy for Shatah-Struwe solutions, from which we immediately obtain *continuity*, and later also *uniqueness* of such solutions:

Proposition 2.4 (Energy conservation, [24, Prop. 3.3])**.** *Let y be a Shatah-Struwe solution of* (CWE) *on* $[0,T]$*. Then the* energy function E_y *associated to y given by*

$$E_y(t) \coloneqq \frac{1}{2}\|\xi_y(t)\|^2_{\mathcal{E}} + \frac{1}{6}\|y(t)\|^6_{\mathrm{L}^6(\Omega)} - \int_0^t \big(u(s), \partial_t y(s)\big)_\Omega \,\mathrm{d}s$$

is absolutely continuous. Moreover, $E_y(t) = E_y(0) = \frac{1}{2}\|\xi_0\|_{\mathcal{E}}^2 + \frac{1}{6}\|y_0\|_{\mathrm{L}^6(\Omega)}^6$ *for all* $t \in [0,T]$*, and we have* $\xi_y \in \mathrm{C}([0,T];\mathcal{E})$*.*

Here, the continuity of ξ_y follows from the continuity of E_y, cf. [30, Ch. 3, Thm. 8.2]. The energy conservation gives us an *a priori* bound on the $\mathrm{C}([0,T];\mathcal{E})$ norm of ξ_y and on $\|y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))}$ which will prove very useful when proving that solutions exist globally-in-time:

Lemma 2.5. *Let* y *be a Shatah-Struwe solution to* (CWE) *on* $[0,T]$*. Then we have*

$$\|\xi_y\|_{\mathrm{C}([0,T];\mathcal{E})}^2 + \|y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))}^6 \lesssim \|\xi_0\|_{\mathcal{E}}^2 + \|y_0\|_{\mathrm{L}^6(\Omega)}^6 + \|u\|_{\mathrm{L}^1(0,\mathrm{T};\mathrm{L}^2(\Omega))}^2 =: E_0. \tag{2.4}$$

Proof. From $E_y(t) = E_y(0)$, we clearly find

$$\frac{1}{2}\|\xi_y(t)\|_{\mathcal{E}}^2 + \frac{1}{6}\|y(t)\|_{\mathrm{L}^6(\Omega)}^6 - \|\partial_t y\|_{\mathrm{L}^\infty(0,t;\mathrm{L}^2(\Omega))}\|u\|_{\mathrm{L}^1(0,t;\mathrm{L}^2(\Omega))} \le E_y(t) = E_y(0) = \frac{1}{2}\|\xi_0\|_{\mathcal{E}}^2 + \frac{1}{6}\|y_0\|_{\mathrm{L}^6(\Omega)}^6,$$

so for every $\varepsilon > 0$

$$\sup_{t\in[0,T]} \left(\frac{1}{2}\|\xi_y(t)\|_{\mathcal{E}}^2 + \frac{1}{6}\|y(t)\|_{\mathrm{L}^6(\Omega)}^6\right) \le E_y(0) + \varepsilon\|\partial_t y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}^2 + \frac{1}{4\varepsilon}\|u\|_{\mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))}^2.$$

This implies

$$\frac{1}{2}\|\xi_y\|_{\mathrm{C}([0,T];\mathcal{E})}^2 + \frac{1}{6}\|y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^6(\Omega))}^6 \lesssim E_y(0) + \varepsilon\|\partial_t y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}^2 + \frac{1}{4\varepsilon}\|u\|_{\mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))}^2,$$

and with ε small enough we can absorb the $\varepsilon\|\partial_t y\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}^2$ term in the left-hand side to obtain (2.4). ☐

Next we show that Shatah-Struwe and mild solutions coincide.

Lemma 2.6. *A function* y *is a mild solution to* (CWE) *if and only if it is a Shatah-Struwe solution.*

Proof. Let y be a mild solution to (CWE). Testing the second row in (2.1) with $\psi \otimes \zeta \in \mathrm{C}_c^\infty(0,T) \otimes \mathrm{C}_c^\infty(\Omega)$, we find

$$\int_0^T \big(\partial_t y(t), \partial_t \psi(t)\zeta\big)_\Omega \,\mathrm{d}t = \big(y_1, \zeta\big)_\Omega \int_0^T \partial_t\psi(t)\,\mathrm{d}t + \int_0^T \left(\Delta \int_0^t y(s)\,\mathrm{d}s, \zeta\right)_\Omega \partial_t\psi(t)\,\mathrm{d}t + \int_0^T \left(\int_0^t u(s) - y^5(s)\,\mathrm{d}s, \zeta\right)_\Omega \partial_t\psi(t)\,\mathrm{d}t.$$

Using selfadjointness of the Laplacian on $\mathrm{L}^2(\Omega)$ and Fubini's theorem, we continue with

$$= \int_0^T \left(\int_0^t y(s)\,\mathrm{d}s, \Delta\zeta \right)_\Omega \partial_t\psi(t)\,\mathrm{d}t + \left(\int_0^T \int_s^T \big(u(s) - y^5(s)\big)\partial_t\psi(t)\,\mathrm{d}t, \zeta \right)_\Omega \mathrm{d}s$$
$$= \int_0^T \big(y(s), -\Delta\zeta\big)_\Omega \psi(t)\,\mathrm{d}t - \int_0^T \big(u(s) - y^5(s), \zeta\big)_\Omega \psi(s)\,\mathrm{d}s.$$

It remains to observe that $\big(y(s), -\Delta\zeta\big)_\Omega = \big(\nabla y(s), \nabla\zeta\big)_\Omega$, due to $\zeta \in \mathrm{C}_c^\infty(\Omega) \subset D(\Delta)$ and the definition of the Laplacian in $\mathrm{L}^2(\Omega)$, and that $\mathrm{C}_c^\infty(0,T) \otimes \mathrm{C}_c^\infty(\Omega)$ is dense in $\mathrm{C}_c^\infty((0,T)\times\Omega)$ ([17, Thm. 4.3.1]) to conclude that y is a weak solution to (CWE).

For the reverse assertion, let y be a Shatah-Struwe solution to (CWE). We need to show that the first row in (2.1) is satisfied in $\mathrm{H}_0^1(\Omega)$. Writing

$$y(t) - y_0 = \int_0^t \partial_t y(s)\,\mathrm{d}s,$$

at first in $\mathrm{L}^2(\Omega)$, we observe that the left-hand side is in fact a continuous function in $\mathrm{H}_0^1(\Omega)$ due to Proposition 2.4. This gives the assertion.

For the second row and $\int_0^t y(s)\,\mathrm{d}s \in D(\Delta)$, observe that from the definition of a weak solution, we have

$$\big\langle \partial_t^2 y(t), \zeta \big\rangle_{\mathrm{H}^{-1}(\Omega),\mathrm{H}_0^1(\Omega)} = \big(\nabla y(t), \nabla\zeta\big)_\Omega + \big(u(t) - y^5(t), \zeta\big)_\Omega \quad \text{for all } \zeta \in \mathrm{C}_c^\infty(\Omega) \tag{2.5}$$

for almost all $t \in (0,T)$. With $\partial_t y(t) \in \mathrm{L}^2(\Omega)$ for almost all $t \in (0,T)$ together with $y_1 \in \mathrm{L}^2(\Omega)$, we thus find

$$\begin{aligned}\big(\partial_t y(t) - y_1, \zeta\big)_\Omega &= \big\langle \partial_t y(t) - y_1, \zeta \big\rangle_{\mathrm{H}^{-1}(\Omega),\mathrm{H}_0^1(\Omega)} \\ &= \int_0^t \big\langle \partial_t^2 y(s), \zeta \big\rangle_{\mathrm{H}^{-1}(\Omega),\mathrm{H}_0^1(\Omega)}\,\mathrm{d}s \\ &= \int_0^t \big(\nabla y(s), \nabla\zeta\big)_\Omega + \big(u(s) - y^5(s), \zeta\big)_\Omega\,\mathrm{d}s\end{aligned}$$

and so

$$\big(\partial_t y(t) - y_1 + \int_0^t y^5(s) - u(s)\,\mathrm{d}s, \zeta\big)_\Omega = \left(\int_0^t y(s)\,\mathrm{d}s, \Delta\zeta \right)_\Omega.$$

Since this equality extends from all ζ in $\mathrm{C}_c^\infty(\Omega)$ to all $\zeta \in D(\Delta)$, we have by definition of the adjoint operator and selfadjointness of Δ:

$$\int_0^t y(s)\,\mathrm{d}s \in D(\Delta) \quad \text{and} \quad \Delta \int_0^t y(s)\,\mathrm{d}s = \partial_t y(t) - y_1 + \int_0^t y^5(s) - u(s)\,\mathrm{d}s.$$

This is the second row in (2.1). Continuity of the Shatah-Struwe solution follows from the energy inequality as noted in Proposition 2.4. □

Another consequence of the energy conservation for Shatah-Struwe and, per Proposition 2.4, mild solutions, is *uniqueness*:

Proposition 2.7 (Shatah-Struwe uniqueness, [24, Cor. 3.4])**.** *The Shatah-Struwe solution of* (CWE) *is unique, if it exists.*

Corollary 2.8. *If there exists a mild solution y to* (CWE)*, then it is unique and coincides with the Shatah-Struwe solution.*

Next, we establish some important estimates and finally local-in-time existence of solutions. For this purpose, consider the block operator in Definition 2.1 (mild solution) as a closed operator in $\mathrm{H}_0^1(\Omega) \times \mathrm{L}^2(\Omega)$ via

$$\mathcal{A} \coloneqq \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta & 0 \end{pmatrix}, \quad D(\mathcal{A}) \coloneqq D(\Delta) \times \mathrm{H}_0^1(\Omega).$$

It can be shown that the operator $\mathcal{A}$ generates a C_0-semigroup $t \mapsto e^{\mathcal{A}t}$ on $\mathcal{E}$, cf. [7, Ch. XVII §3 Sect. 3.4]. Thus, for $\xi_{0,z} = (z_1, z_0) \in \mathcal{E}$ and $f \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$, the usual *variation-of-constants formula*

$$\xi_z(t) = e^{\mathcal{A}t}\xi_{0,z} + \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ f(s) \end{pmatrix} \,\mathrm{d}s. \tag{2.6}$$

is well defined and gives the unique mild solution $z \in \mathrm{C}([0,T];\mathcal{E})$ to the linear wave equation

$$\left.\begin{aligned} \partial_t^2 z - \Delta z &= f && \text{in } (0,T) \times \Omega, \\ z &= 0 && \text{on } (0,T) \times \partial\Omega, \\ \big(z(0), \partial_t z(0)\big) &= (z_0, z_1) && \text{in } \Omega, \end{aligned}\right\} \tag{LWE}$$

cf. [1, Prop. 3.1.16].

We state the fundamental estimates for z satisfying (2.6). While the first one is the standard energy space estimate which follows immediately from (2.6) and Sobolev embedding, the second one is a nontrivial Strichartz estimate as proven in [2]:

Lemma 2.9 (Energy- and Strichartz estimates)**.** *Let $f \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$ as well as $\xi_{0,z} \in \mathcal{E}$, and let ξ_z be given by* (2.6) *for $t \in [0,T]$. Then there exist constants $C_e(T), C_s(T)$ such that the* energy estimate

$$\big\|\xi_z(t)\big\|_{\mathcal{E}} + \|y(t)\|_{\mathrm{L}^6(\Omega)} \le C_e(T)\Big(\big\|\xi_{0,z}\big\|_{\mathcal{E}} + \big\|f\big\|_{\mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))}\Big) \tag{2.7}$$

for all $t \in [0,T]$, and the Strichartz estimate

$$\big\|z\big\|_{\mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))} \le C_s(T)\Big(\big\|\xi_{0,z}\big\|_{\mathcal{E}} + \big\|f\big\|_{\mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))}\Big) \tag{2.8}$$

are satisfied.

Remark 2.10. Suppose that $\xi_{0,z} = 0$. Then it is easy to see that the constants $C_e(\tau)$ and $C_s(\tau)$ associated to the estimates (2.7) and (2.8) for solutions on the intervals $[0,\tau]$ are monotonously increasing in τ. In other words, given a function $f \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$, we have $C_e(\tau) \le C_e(T)$ and $C_s(\tau) \le C_s(T)$ for all $\tau \in [0,T]$.

Now, if ξ_y with $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$ is given by (2.6) with $f = u - y^5$ and initial data ξ_0, so that

$$\xi_y(t) = e^{\mathcal{A}t}\xi_0 + \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ u(s) - y^5(s) \end{pmatrix} \, \mathrm{d}s, \tag{2.9}$$

then y is in fact the unique mild solution to (CWE) in the sense of Definition 2.1 on $[0,T]$ (see again [1, Prop. 3.1.16]) and thus also the Shatah-Struwe solution, recall Corollary 2.8. In particular, y satisfies the estimates in Lemma 2.9 for $f = u - y^5$ and initial data ξ_0. As explained in Remark 2.2, the integrability $y^5 \in \mathrm{L}^1(0,T;\mathrm{L}^2(\Omega))$ follows from the additional dispersion information $y \in \mathrm{L}^4(0,T;\mathrm{L}^{12}(\Omega))$.

Using the linear estimates in Lemma 2.9 together with the interpolation inequalities as in Remark 2.2, local existence and uniqueness of a function y satisfying (2.9)—which is then also the unique mild and Shatah-Struwe solution on the interval of existence—depending continuously on the given data follows from a standard fixed point argument. (See also [24, Prop. 3.1] for another explicit proof.)

Theorem 2.11 (Local-in-time existence)**.** *There exists a maximal unique mild and Shatah-Struwe solution to* (CWE) *such that* (CWE) *is well posed with respect to* $\mathcal{E}$ *and* $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$. *More precisely, there exists a maximal time* $T^\bullet \in (0,\mathsf{T}]$ *depending on* $\xi_0 \in \mathcal{E}$ *and* $u \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ *and a unique function* y *on* $[0,T^\bullet)$ *with* ξ_y *given by* (2.9) *such that*

$$\xi_y \in \mathrm{C}\big([0,T^\bullet);\mathcal{E}\big) \quad \text{and} \quad y \in \mathrm{L}^4_{loc}\big([0,T^\bullet);\mathrm{L}^{12}(\Omega)\big).$$

This function y *is the unique mild and Shatah-Struwe solution to* (CWE) *on* $[0,T]$ *for every* $T \in (0,T^\bullet)$. *Moreover, there exists* $\varepsilon > 0$ *such that for every initial value* $\zeta_0 \in B_\varepsilon(\xi_0) \subset \mathcal{E}$ *and every right-hand side* $v \in B_\varepsilon(u) \subset \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ *there exists a unique solution* $\bar{y}$ *in the foregoing sense on the same intervals of existence and the mapping* $B_\varepsilon(\xi_0) \times B_\varepsilon(u) \ni (\zeta_0, v) \to \bar{y}$ *is continuous.*

Remark 2.12. The wellposedness of (CWE) allows to obtain certain classical properties of more regular solutions to (CWE) and related equations also for Shatah-Struwe solutions by approximation. This includes in particular *finite speed of propagation* of such solutions, cf. e.g. [15, Ch. 2.4.3, Thm. 6].

3. Global solutions

In this section, we establish that there indeed exists a unique *global-in-time* mild and Shatah-Struwe solution to (CWE) on $[0,\mathsf{T}]$ given by (2.9). The argument follows [3] which in turn builds upon [37], see also [38, Ch. IV, §3] or [41, Ch. 5.1]. We mostly outline the strategy and give a minimally invasive modification of the proof in [3]. The modification is necessary in the first place because on the one hand, only the case $u = 0$ is treated in [3], and on the other hand, the improved Strichartz estimates in [2] as stated in Lemma 2.9 allow to simplify the proof at some places compared to [3]. The global existence result is also stated in [24, Thm. 3.8], however, without a proof.

The proof consists in principle of an ordinary extension argument via the energy space $\mathcal{E}$: Let y be the maximal solution to (CWE) as in Theorem 2.11. We show that the limit $\lim_{t\nearrow T^\bullet} \xi_y(t) =: \xi_y(T^\bullet)$ exists in $\mathcal{E}$. Then we either have already $T^\bullet = \mathsf{T}$, or we can extend the solution by re-applying Theorem 2.11 starting from $T^\bullet$ with initial data $\xi_y(T^\bullet)$ until we have a solution on the whole $[0,\mathsf{T}]$.

For this purpose, it is imperative to observe that, by the energy estimate (2.7), the limit of $\xi_y(t)$ as $t \nearrow T^\bullet$ in (2.9) only fails to exist if $y \notin \mathrm{L}^5(0,T^\bullet;\mathrm{L}^{10}(\Omega))$; so we want to prove that in fact $y \in \mathrm{L}^5(0,T^\bullet;\mathrm{L}^{10}(\Omega))$. Due to energy conservation (cf. Lemma 2.5) and (2.2), it is moreover sufficient to show that $y \in \mathrm{L}^4(0,T^\bullet;\mathrm{L}^{12}(\Omega))$.

In the introduction it was mentioned several times that there is no direct bound on a local solution y in $\mathrm{L}^4(0,T^\bullet;\mathrm{L}^{12}(\Omega))$ in the critical case. This is in contrast to the subcritical case with a nonlinearity y^p with $1<p<5$. We give a quick demonstration of this and how it does not work for the critical case (cf. [38, Ch. IV.2]). Let y be the local-in-time solution of (CWE) as in Theorem 2.11 and let $t \in [0,T^\bullet)$. The Hölder inequality yields the more general form of (2.3)

$$\left\|y^p\right\|_{\mathrm{L}^1(t,T^\bullet;\mathrm{L}^2(\Omega))} \le \|y\|^{p-1}_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))}\|y\|_{\mathrm{L}^{\frac{4}{5-p}}(t,T^\bullet;\mathrm{L}^{\frac{12}{7-p}}(\Omega))}.$$

Now note that $\frac{12}{7-p} < p+1$ if and only if $1<p<5$. Hence

$$\|y\|_{\mathrm{L}^{\frac{4}{5-p}}(t,T^\bullet;\mathrm{L}^{\frac{12}{7-p}}(\Omega))} \lesssim_\Omega \left(T^\bullet - t\right)^{\frac{5-p}{4}}\|y\|_{\mathrm{L}^\infty(t,T^\bullet;\mathrm{L}^{p+1}(\Omega))}.$$

If y^p is the nonlinearity in (CWE), then $\|y\|_{\mathrm{L}^\infty(t,T^\bullet;\mathrm{L}^{p+1}(\Omega))}$ is present in Proposition 2.4 instead of the corresponding $\mathrm{L}^6(\Omega)$ terms. Accordingly, it is uniformly bounded by a power of E_0 and we obtain from the above

$$\left\|y^p\right\|_{\mathrm{L}^1(t,T^\bullet;\mathrm{L}^2(\Omega))} \lesssim_{\Omega,E_0} \left(T^\bullet - t\right)^{\frac{5-p}{4}}\|y\|^{p-1}_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))}.$$

The Strichartz estimate (2.8) then shows that

$$\|y\|_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))} \lesssim_{\Omega,E_0} \sqrt{E_0} + \left(T^\bullet - t\right)^{\frac{5-p}{4}}\|y\|^{p-1}_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))}.$$

Via Lemma A.1, this implies that there is $t^\star \in (0,T^\bullet)$ such that $\|y\|_{\mathrm{L}^4(t^\star,T^\bullet;\mathrm{L}^{12}(\Omega))}$ is finite. Unfortunately, the foregoing proof breaks down completely for the critical value $p=5$ since we obtain

$$\|y\|_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))} \lesssim_\Omega \sqrt{E_0} + \|y\|_{\mathrm{L}^\infty(t,T^\bullet;\mathrm{L}^6(\Omega))}\|y\|^4_{\mathrm{L}^4(t,T^\bullet;\mathrm{L}^{12}(\Omega))}$$

and $\|y\|_{\mathrm{L}^\infty(t,T^\bullet;\mathrm{L}^6(\Omega))}$, although bounded by $E_0^{1/6}$, does not go to zero as $t \searrow T^\bullet$ in general. It is thus necessary to proceed differently. The idea is to replace the "full-spacetime" norm $\|y\|_{\mathrm{L}^\infty(t,T^\bullet;\mathrm{L}^6(\Omega))}$ by a localized one which indeed goes to zero as $t \searrow T^\bullet$. This is done as follows.

We first show an L^6-*non-concentration effect* in $T^\bullet$, namely that the L^6 norm of the solution cannot concentrate in a single point x_0, i.e., be greater than 0. This is the most involved and nontrivial result, but luckily we only need to make appropriate modifications to incorporate the inhomogeneity u compared to the proof in [3, Prop. 3.3]; see Proposition 3.1. The non-concentration effect allows to prove that the solution must be in $\mathrm{L}^4_t\mathrm{L}^{12}_\mathrm{x}$-integrable on a backwards light cone through $(T^\bullet,\mathrm{x}_0)$. (For precise definitions, see below.) This is done in Proposition 3.3. We then show in Proposition 3.6 that this $\mathrm{L}^4_t\mathrm{L}^{12}_\mathrm{x}$-integrability enables us to prove that the $\mathrm{L}^\infty_t\mathrm{L}^6_\mathrm{x}$-norm of y becomes arbitrarily small on a *slightly larger* light cone as we approach $T^\bullet$. This allows to employ an argument similar to the one displayed for the subcritical case above which then finally leads to boundedness of $\mathrm{L}^4(t^\star,T^\bullet;\mathrm{L}^{12}(\Omega))$ for some $t^\star$ close to $T^\bullet$ and thus finishes the proof of the main result, Theorem 3.7.

3.1. Global existence

We fix $\mathrm{x}_0 \in \overline{\Omega}$ for the following, if not stated otherwise, as well as the blowup time $T^\bullet > 0$. Frequently needed objects are the δ-enlarged *backwards light cone* through (s, x_0) for $0 \le t_0 \le s \le T^\bullet$ and $\delta \ge 0$ given by

$$\Lambda(\delta; t_0, s) \coloneqq \Big\{(t, \mathrm{x}) \in [t_0, s] \times \overline{\Omega} \colon |\mathrm{x} - \mathrm{x}_0| \le \delta + T^\bullet - t\Big\}.$$

Moreover, we need its "time slice" at time level τ

$$D_\tau^\delta \coloneqq \Big\{(\tau, \mathrm{x}) \colon \mathrm{x} \in \overline{\Omega},\ |\mathrm{x} - \mathrm{x}_0| \le \delta + T^\bullet - \tau\Big\}$$

with $D_\tau \coloneqq D_\tau^0$. Clearly, the sets D_τ^δ live in the four-dimensional space $\mathbb{R} \times \mathbb{R}^3$, and we use $P_\mathrm{x} D_\tau^\delta$ to denote the projection of this set onto the second coordinate block, so in $\mathbb{R}^3$. We moreover use the mixed Lebesgue norm notation

$$\|w\|^p_{\mathrm{L}^p_t \mathrm{L}^q_\mathrm{x}(\Lambda(\delta; t_0, s))} \coloneqq \int_{t_0}^s \left(\int_{P_\mathrm{x} D_t^\delta} |w(t, \mathrm{x})|^q \,\mathrm{dx} \right)^{p/q} \mathrm{d}t,$$

with the usual modification for $p = \infty$. Finally, let us define the local energy of a function v on D_t^δ for $0 \le t < T^\bullet$ by

$$E_v(\delta; t) \coloneqq \int_{P_\mathrm{x} D_t^\delta} \frac{|\nabla v(t, \mathrm{x})|^2 + |\partial_t v(t, \mathrm{x})|^2}{2} + \frac{|v(t, \mathrm{x})|^6}{6} \,\mathrm{dx}.$$

The first result is the L^6-non-concentration effect:

Proposition 3.1 (L^6-nonconcentration)**.** *There holds*

$$\lim_{t \nearrow T^\bullet} \int_{P_\mathrm{x} D_t} |y(t, \mathrm{x})|^6 \,\mathrm{dx} = 0.$$

Proof. We need only make appropriate modifications in the proof in [3, Prop. 3.3], whose strategy follows [37, Lem. 3.3] or [38, Ch. V, Prop. 3.2], to incorporate the inhomogeneity u. There are essentially three aspects:

1. The estimate
$$\|\partial_\nu y\|^2_{\mathrm{L}^2(0,s;\mathrm{L}^2(\partial\Omega))} \lesssim E_0$$
is still satisfied uniformly for every $0 \le s < T^\bullet$, where ∂_ν is the trace of the outer unit normal on $\partial\Omega$. This follows as in [3, Prop. 3.2] by taking care of the estimate
$$\int_0^s \int_\Omega \Big[(Zu)(t, \mathrm{x}) \cdot y(t, \mathrm{x}) - (Zy)(t, \mathrm{x}) \cdot u(t, \mathrm{x})\Big] \,\mathrm{dx}\,\mathrm{d}t \lesssim E_0$$
uniformly in s, where Z is a smooth scalar field on Ω which coincides with ∂_ν on $\partial\Omega$. (For this argument, we suppose u and y to be smooth and refer to the wellposedness of the equation, recall Remark 2.12.) Such an estimate follows immediately using integration by parts and the energy conservation (2.4).

2. We refer to [3, Sect. 3.1], [37, Lem. 3.2] or [38, Ch. IV, §3] for the derivation of

$$E_y(0,s) + \mathrm{Flux}(y;\tau_0,s) + \int_{\Lambda(0;\tau_0,s)} \partial_t y(t,\mathrm{x}) \cdot u(t,\mathrm{x}) \,\mathrm{d}(t,\mathrm{x}) = E_y(0,\tau_0) \tag{3.1}$$

for $0 \le \tau_0 \le s < T^\bullet$. Here,

$$\mathrm{Flux}(y;\tau_0,s) \coloneqq \int_{M^s_{\tau_0}} e(t,\mathrm{x}) \cdot \nu(t,\mathrm{x}) \,\mathrm{d}\sigma(t,\mathrm{x}),$$

where $M^s_{\tau_0} \coloneqq \big\{(t,\mathrm{x}) \in \Lambda(0;\tau_0,s) \colon |\mathrm{x}-\mathrm{x}_0| = T^\bullet - t\big\}$ is the "mantle" and ν the unit outer normal to $\Lambda(0;\tau_0,s)$, and the vector field e is given by

$$e(t,\mathrm{x}) \coloneqq \left(\frac{|\partial_t y(t,\mathrm{x})|^2 + |\nabla y(t,\mathrm{x})|^2}{2} + \frac{|y(t,\mathrm{x})|^6}{6}, -\partial_t y(t,\mathrm{x}) \nabla y(t,\mathrm{x}) \right).$$

Thus, $\mathrm{Flux}(y;\tau_0,s)$ is the energy transferred across $M^s_{\tau_0}$ during transition from D_{τ_0} to D_s, and we have $\mathrm{Flux}(y;\tau_0,s) \ge 0$. As in the references for the proof, we show that $\lim_{\tau_0 \nearrow T^\bullet} \mathrm{Flux}(y;\tau_0,s) = 0$.

Estimating the integral involving u in (3.1) from below and using the energy bound (2.7) we derive

$$E_y(0,s) + \mathrm{Flux}(y;\tau_0,s) \le E_y(0,\tau_0) + E_0 \|u\|_{\mathrm{L}^1(\tau_0,s;\mathrm{L}^2(\Omega))}. \tag{3.2}$$

The nonnegativity of the flux now implies that the function f defined by $t \mapsto E_y(0,t) + E_0\|u\|_{\mathrm{L}^1(t,T^\bullet;\mathrm{L}^2(\Omega))}$ is nonincreasing. Due to the energy bound (2.4), it is moreover uniformly bounded for $t \in (0,T^\bullet)$, and thus admits a limit $\lim_{t\nearrow T^\bullet} f(t)$. Back in (3.2), we now have

$$0 \le \mathrm{Flux}(y;\tau_0,s) \le f(\tau_0) - f(s) \xrightarrow{\tau_0 \nearrow T^\bullet} 0,$$

so indeed $\lim_{\tau_0\nearrow T^\bullet} \mathrm{Flux}(y;\tau_0,s) = 0$.

3. Lastly, in the proof in [3, Prop. 3.3], a Morawetz identity is used which can be formally derived by multiplying the state equation with $(t \cdot \partial_t y(t,\mathrm{x}) + \mathrm{x} \cdot \nabla y(t,\mathrm{x}) + y(t,\mathrm{x}))$. The identity is then integrated over $\Lambda(0;\tau_0,s)$ and a bound on the $\mathrm{L}^6(P_\mathrm{x} D_{\tau_0})$-norm of $y(\tau_0)$ is derived; this is of course again for smooth solutions of the equation and the claim for mild solutions follows by approximation. To comply with the line of proof in [37] or [3], we need to make sure that

$$\int_{\Lambda(0;\tau_0,s)} u(t,\mathrm{x}) \cdot \big(t \cdot \partial_t y(t,\mathrm{x}) + \mathrm{x} \cdot \nabla y(t,\mathrm{x}) + y(t,\mathrm{x})\big) \,\mathrm{d}(t,\mathrm{x}) \xrightarrow{\tau_0 \nearrow 0} 0.$$

(Note that, in order to stay close to the referred works, we have shifted $(T^\bullet, \mathrm{x}_0)$ to $(0,0)$ here, so now $\tau_0 \le s < 0$.) This however follows quite immediately from Hölder's inequality and the energy bound (2.4), as the absolute value of the left-hand side can be estimated by

$$\begin{aligned} &\int_{\Lambda(0;\tau_0,s)} \big|u(t,\mathrm{x}) \cdot \big(t \cdot \partial_t y(t,\mathrm{x}) + \mathrm{x} \cdot \nabla y(t,\mathrm{x}) + y(t,\mathrm{x})\big)\big| \,\mathrm{d}(t,\mathrm{x}) \\ &\qquad \le \int_{\tau_0}^s \|u\|_{\mathrm{L}^2(\Omega)} \big(|\tau_0| \|\partial_t y\|_{\mathrm{L}^2(\Omega)} + \mathrm{diam}(\Omega)\|\nabla y\|_{\mathrm{L}^2(\Omega)} + \|y\|_{\mathrm{L}^2(\Omega)}\big) \,\mathrm{d}t \\ &\qquad\qquad \lesssim \|u\|_{\mathrm{L}^1(\tau_0,0;\mathrm{L}^2(\Omega))}. \end{aligned}$$

With these three modifications, we can now repeat the proof of L^6-non-concentration verbatim as in [3, Prop. 3.3] with u inserted at the appropriate places. □

To make use of the foregoing Proposition 3.1, we next establish a series of preliminary results. The first one is a technical result which allows us to localize functions to the "slices" $P_{\mathrm{x}} D_t^\delta$.

Lemma 3.2 (Localization ([3, Lem. 3.3]))**.** *Let $1 \le p \le \infty$. For every $\mathrm{x}_0 \in \overline{\Omega}$ there exist numbers $r_{ext} > 0$ and $C_{ext} \ge 0$ with the following significance: For $\delta < r_{ext}$, there exists $t_0 \in (0 \vee T^\bullet + \delta - r_{ext}, T^\bullet)$ such that given $v \in \mathrm{L}^1_{loc}(0, T^\bullet; \mathrm{L}^p(\Omega))$, there exists a function $\check{v} \in \mathrm{L}^1_{loc}(0, T^\bullet; \mathrm{L}^p(\Omega))$ such that*

$$v(t) = \check{v}(t) \quad \text{a.e. on } P_{\mathrm{x}} D_t^\delta$$

and

$$\|\check{v}(t)\|_{\mathrm{L}^p(\Omega)} \le C_{ext} \|v(t)\|_{\mathrm{L}^p(P_{\mathrm{x}} D_t^\delta)} \tag{3.3}$$

for all $t \in (t_0, T^\bullet)$.

The number r_{ext} in the next proposition is the one from Lemma 3.2.

Proposition 3.3. *Let $0 \le \delta < r_{ext}$ and assume that for every $\varepsilon > 0$ there exists $\tau_0 \in (0, T^\bullet)$ such that*

$$\|y\|_{\mathrm{L}^\infty_t \mathrm{L}^6_{\mathrm{x}}(\Lambda(\delta; \tau_0, T^\bullet))} < \varepsilon$$

Then there is $t_0 \in (0, T^\bullet)$ such that $y \in \mathrm{L}^4_t \mathrm{L}^{12}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))$.

Proof. Let $t_0 \in (0 \vee T^\bullet + \delta - r_{\mathrm{ext}}, T^\bullet)$ be fixed for now, to be chosen later. We use the assumption in conjunction with the Strichartz estimate (2.8). Let $\check{y}$ be the function from Lemma 3.2 coinciding with y on $\Lambda(\delta; t_0, T^\bullet)$ and let w be given by

$$\xi_w(t) = e^{\mathcal{A}t} \xi_y(t_0) + \int_{t_0}^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ u(s) - \check{y}^5(s) \end{pmatrix} \mathrm{d}s,$$

so the mild solution to the linear wave equation (LWE) on $[t_0, T^\bullet)$ with $f = u - \check{y}$ and initial data $\xi_y(t_0)$. Then w coincides with y on $\Lambda(\delta; t_0, T^\bullet)$ due to finite speed of propagation, cf. Remark 2.12. Using the Strichartz estimate (2.8) for this linear equation we obtain

$$\begin{aligned} \|y\|_{\mathrm{L}^4_t \mathrm{L}^{12}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))} &\le \|w\|_{\mathrm{L}^4(t_0, T^\bullet; \mathrm{L}^{12}(\Omega))} \\ &\lesssim_{T^\bullet} \|\xi_y(t_0)\|_{\mathcal{E}} + \|\check{y}^5\|_{\mathrm{L}^1(t_0, T^\bullet; \mathrm{L}^2(\Omega))} + \|u\|_{\mathrm{L}^1(t_0, T^\bullet; \mathrm{L}^2(\Omega))}. \end{aligned}$$

The extension estimate (3.3) and the interpolation inequality (2.2) further yield

$$\begin{aligned} \|\check{y}^5\|_{\mathrm{L}^1(t_0, T^\bullet; \mathrm{L}^2(\Omega))} &\lesssim \|y\|_{\mathrm{L}^5_t \mathrm{L}^{10}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))} \\ &\lesssim \|y\|^4_{\mathrm{L}^4_t \mathrm{L}^{12}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))} \|y\|_{\mathrm{L}^\infty_t \mathrm{L}^6_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))}. \end{aligned}$$

Hence, choosing ε small enough (cf. Lemma A.1) and if necessary enlarging t_0 to $\tau_0(\varepsilon)$, we obtain

$$\|y\|_{\mathrm{L}^4_t \mathrm{L}^{12}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))} \lesssim \|\xi_y(t_0)\|_{\mathcal{E}} + \|u\|_{\mathrm{L}^1(t_0, T^\bullet; \mathrm{L}^2(\Omega))}.$$

An application of the energy bound (2.4) then yields the claim. □

An immediate consequence of Proposition 3.3 and its proof together with the interpolation inequality (2.2) is the following:

Corollary 3.4. *Let the assumption of Proposition 3.3 hold true for some δ satisfying $0 \le \delta < r_{ext}$. Then, for every $\varepsilon > 0$, there exists $t_0 \in (0, T^\bullet)$ such that*

$$\|y\|_{\mathrm{L}^5_t \mathrm{L}^{10}_{\mathrm{x}}(\Lambda(\delta; t_0, T^\bullet))} < \varepsilon.$$

We will need a bound for the energy transfer from one time level to another in the light cones when we come close enough to $T^\bullet$. The following lemma states that this is possible and, crucially, even uniformly in δ.

Lemma 3.5. *For every $\varepsilon > 0$ there is $t_0 \in [0, T^\bullet)$ such that*

$$E_y(\delta; s) \le E_y(\delta; \tau_0) + \varepsilon$$

for all τ_0, s satisfying $t_0 \le \tau_0 \le s < T^\bullet$, and all $\delta \ge 0$.

Proof. Let $0 \le t_0 \le \tau_0 \le s < T^\bullet$. As in the proof of Proposition 3.6, we obtain for *every* $\delta \ge 0$

$$E_y(\delta, s) + \int_{\Lambda(\delta; \tau_0, s)} \partial_t y(t, \mathrm{x}) \cdot u(t, \mathrm{x}) \, \mathrm{d}(t, \mathrm{x}) \le E_y(\delta, \tau_0).$$

and so

$$E_y(\delta, s) \le E_y(\delta, \tau_0) + E_0 \|u\|_{\mathrm{L}^1(t_0, T^\bullet; \mathrm{L}^2(\Omega))}.$$

Choosing t_0 sufficiently close to $T^\bullet$, this gives the claim. □

Finally, the next proposition shows that the L^6-non-concentration effect as proven in Proposition 3.1 in fact also holds in α-enlarged light cones for $\alpha > 0$ sufficiently small. This will then immediately imply the main Theorem 3.7 below.

Proposition 3.6. *Let the assumption of Proposition 3.3 hold true for $\delta = 0$. Then, for every $\varepsilon > 0$ there exist $t_0 \in (0, T^\bullet)$ and $0 < \alpha < r_{ext}$ such that*

$$\|y\|_{\mathrm{L}^\infty_t \mathrm{L}^6_{\mathrm{x}}(\Lambda(\alpha; t_0, T^\bullet))} < \varepsilon.$$

Proof. Let $\varepsilon > 0$. We do explicit estimates to demonstrate that there are no implicit dependencies on the choice of t_0 along the proof. Via Corollary 3.4 and Lemma 3.5, choose $\tau_0 \in (0 \vee T^\bullet - r_{\mathrm{ext}}, T^\bullet)$ such that

$$\|y\|_{\mathrm{L}^5_t \mathrm{L}^{10}_{\mathrm{x}}(\Lambda(0; s, T^\bullet))} < \varepsilon$$

and

$$E_y(\delta; s) \le E_y(\delta; \eta_0) + C_e(T^\bullet) C_{\mathrm{ext}}^5 \varepsilon^5,$$

as well as

$$\|u\|_{\mathrm{L}^1(s, T^\bullet; \mathrm{L}^2(\Omega))} \le C_{\mathrm{ext}}^5 \varepsilon^5,$$

all for all $\tau_0 \le \eta_0 \le s < T^\bullet$, and all $\delta \ge 0$, where C_e was the constant from Lemma 2.9.

Let again $\check{y}$ be the function from Lemma 3.2 coinciding with y on $\Lambda(0;\tau_0,T^\bullet)$. We split the local solution y on $[\tau_0,T^\bullet)$ into a homogeneous part y_h and an inhomogeneous part y_i by

$$\xi_{y_h}(t) := e^{\mathcal{A}t}\xi_y(\tau_0), \qquad \xi_{y_i}(t) := \int_{\tau_0}^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ u(s) - y^5(s) \end{pmatrix} \mathrm{d}s.$$

With w_i defined by

$$\xi_{w_i}(t) := \int_{\tau_0}^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ u(s) - \check{y}^5(s) \end{pmatrix} \mathrm{d}s$$

on $[\tau_0,T^\bullet]$, we have $w_i = y_i$ on $\Lambda(0;\tau_0,T^\bullet)$. Thus, the estimates in Lemma 2.9 together with Remark 2.10, the choice of τ_0, and (3.3) imply

$$\begin{aligned}\|\nabla y_i\|_{\mathrm{L}^\infty_t\mathrm{L}^2_\mathrm{x}(\Lambda(0;\tau_0,T^\bullet))} &+ \|\partial_t y_i\|_{\mathrm{L}^\infty_t\mathrm{L}^2_\mathrm{x}(\Lambda(0;\tau_0,T^\bullet))} + \|y_i\|_{\mathrm{L}^\infty_t\mathrm{L}^6_\mathrm{x}(\Lambda(0;\tau_0,T^\bullet))} \\ &\le \|w_i\|_{\mathrm{C}([\tau_0,T^\bullet];\mathcal{E})} + \|w_i\|_{\mathrm{L}^\infty(\tau_0,T^\bullet;\mathrm{L}^6(\Omega))} \\ &\le C_e(T^\bullet)\left(\|\check{y}\|^5_{\mathrm{L}^5(\tau_0,T^\bullet);\mathrm{L}^{10}(\Omega)} + \|u\|_{\mathrm{L}^1(\tau_0,T^\bullet;\mathrm{L}^2(\Omega))}\right) < 2C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5.\end{aligned}$$

For the local energy of the homogeneous part y_h of y, we find by $y_h \in \mathrm{C}([0,T^\bullet];\mathcal{E})$ and conservation of energy

$$\begin{aligned}&\frac{1}{2}\int_{P_\mathrm{x}D_t} |\nabla y_h(t,\mathrm{x})|^2\,\mathrm{dx} \\ &\qquad \lesssim \|\nabla y_h(t) - \nabla y_h(T^\bullet)\|^2_{\mathrm{L}^2(\Omega)} + \int_{P_\mathrm{x}D_t} |\nabla y_h(T^\bullet,\mathrm{x})|\,\mathrm{dx} \xrightarrow{t\nearrow T^\bullet} 0.\end{aligned}$$

Treating the $\partial_t y_h$ term in $E_{y_h}(0,t)$ analogously, we thus obtain

$$\lim_{t\nearrow T^\bullet} E_{y_h}(0;t) = 0.$$

On the other hand, the inhomogeneous part was already estimated by

$$E_{y_i}(0;t) < 2C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5$$

for all $t \in [\tau_0,T^\bullet]$, hence we can choose $t_0 \in [\tau_0,T^\bullet)$ to obtain

$$E_y(0;t) \le 2^6\big(E_{y_h}(0;t) + E_{y_i}(0;t)\big) < 129C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5$$

for all $t \in [t_0,T^\bullet]$. The "homogeneous energy" $\|\xi_y(t_0)\|_\mathcal{E} + \|y(t_0)\|_{\mathrm{L}^6(\Omega)}$ is an upper bound for $E_y(\delta;t_0)$ for every δ, and finite by Lemma 2.5. Thus, the dominated convergence theorem, used with respect to δ, yields $\alpha = \alpha(t_0) > 0$ such that

$$E_y(\alpha;t_0) < 130C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5$$

and $\alpha < r_{\mathrm{ext}}$. We can then finally make use of the choice of τ_0 done at the beginning of the proof and its uniformity w.r.t. δ to find

$$\begin{aligned}\|y\|_{\mathrm{L}^\infty_t\mathrm{L}^6_\mathrm{x}(\Lambda(\alpha;t_0,T^\bullet))} &\le \sup_{t_0\le t<T^\bullet} E_y(\alpha;t) \\ &\le E_y(\alpha;t_0) + C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5 < 131C_e(T^\bullet)C_{\mathrm{ext}}^5\varepsilon^5.\end{aligned}$$

This completes the proof. □

Theorem 3.7 (Global existence)**.** *For every $u \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$, the local solution y to* (CWE) *as given in Theorem 2.11 exists globally in time on the interval* $[0,\mathsf{T}]$ *and satisfies $\xi_y \in \mathrm{C}([0,\mathsf{T}];\mathcal{E})$ and $y \in \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$.*

Proof. We had already noted that it is sufficient to show that the local solution y satisfies $y \in \mathrm{L}^4(0,T^\bullet;\mathrm{L}^{12}(\Omega))$ since this allows to show that $\lim_{t\nearrow T^\bullet}\xi_y(t)$ exists in $\mathcal{E}$ via the variation of constants formula and the estimates as in Lemma 2.9 together with the energy conservation (2.4). Now, for this purpose, let $\mathrm{x}_0 \in \overline{\Omega}$ be fixed. Proposition 3.1 tells us that the premise of Proposition 3.3 is satisfied for $\delta = 0$. From there, Proposition 3.6 implies, again via Proposition 3.3, that there are $t_0 \in (0,T^\bullet)$ and $\alpha > 0$ such that $y \in \mathrm{L}^4_t\mathrm{L}^{12}_\mathrm{x}(\Lambda(\alpha;t_0,T^\bullet))$.

This can be done for every $\mathrm{x}_0 \in \overline{\Omega}$. Then, the collection of sets $\big(P_\mathrm{x} D^{\alpha(\mathrm{x}_0)}_{T^\bullet}\big)_{\mathrm{x}_0\in\overline{\Omega}}$ is a (relatively) open covering of $\overline{\Omega}$. Compactness of the latter gives a finite set of points $\mathrm{x}_i \in \overline{\Omega}$, $i = 1,\dots,n$, such that $\big(P_\mathrm{x} D^{\alpha(\mathrm{x}_i)}_{T^\bullet}\big)_{i=1,\dots,n}$ is still a (relatively) open covering of $\overline{\Omega}$. Setting $t_0^\star \coloneqq \max_{i=1,\dots,n} t_0(\mathrm{x}_i)$, we find $y \in \mathrm{L}^4(t_0^\star,T^\bullet;\mathrm{L}^{12}(\Omega))$, and since we already knew that $y \in \mathrm{L}^4_{\mathrm{loc}}([0,T^\bullet);\mathrm{L}^{12}(\Omega))$, this gives $y \in \mathrm{L}^4(0,T^\bullet;\mathrm{L}^{12}(\Omega))$ as desired. □

4. Optimal control

We recall the setup of the optimal control problem. Let $y_d \in \mathrm{L}^2(\Omega)$ and nonnegative scaling parameters γ,β_1,β_2 be given. We had

$$\ell(y,u) \coloneqq \frac{1}{2}\|y(\mathsf{T}) - y_d\|^2_{\mathrm{L}^2(\Omega)} + \frac{\gamma}{4}\|y\|^4_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))} + \beta_1\|u\|_{\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))} + \frac{\beta_2}{2}\|u\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}$$

for $y \in \mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega)) \cap \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$ and $u \in \mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$, where $r = 1$ if $\beta_2 = 0$ and $r = 2$ if $\beta_2 > 0$. We consider ℓ as a cost functional or performance index for (CWE), resulting in the associated optimal control problem

$$\min_{y,u} \quad \ell(y,u) \qquad \text{s.t.} \quad \begin{cases} u \in \mathcal{U}_{\mathrm{ad}}, \\ y \text{ is the solution to (CWE).} \end{cases} \tag{OCP}$$

Here, $\mathcal{U}_{\mathrm{ad}}$ is a closed and convex, and thus weakly closed, nonempty set of the form

$$\mathcal{U}_{\mathrm{ad}} \coloneqq \Big\{v \in \mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))\colon \|v(t)\|_{\mathrm{L}^2(\Omega)} \le \omega(t) \text{ f.a.a. } t \in (0,\mathsf{T})\Big\}$$

for a measurable function ω which is nonnegative almost everywhere on $(0,\mathsf{T})$. We emphasize once more that ω is not assumed to be bounded away from 0 uniformly almost everywhere. Further, of course, the solution y in (OCP) is meant in the sense of Theorem 3.7.

We will proceed to establish existence of globally optimal solutions to (OCP) in the following. Moreover, we will give necessary optimality conditions of both first and second order, and also second order sufficient conditions.

4.1. Existence of globally optimal controls

It now becomes convenient that a solution y associated to u—which we denote by y_u from now on—in the sense of Theorem 3.7 is a Shatah-Struwe solution as noted in Lemma 2.6:

Theorem 4.1 (Existence of optimal controls)**.** *Let* $\beta_2 > 0$ *or let* $\omega \in \mathrm{L}^1(0,\mathsf{T})$*, and let* $\gamma > 0$*. Then the optimal control problem* (OCP) *admits at least one globally optimal pair* $(y_{\bar{u}}, \bar{u})$ *with* $\bar{u} \in \mathcal{U}_{ad}$ *such that the state* $y_{\bar{u}}$ *is the unique global Shatah-Struwe solution to* (CWE) *for the right-hand side* $\bar{u}$*.*

Proof. Since $\mathcal{U}_{\mathrm{ad}} \neq \emptyset$, and ℓ is bounded from below by zero, we obtain an infimal sequence (y_k, u_k) with $(u_k) \subseteq \mathcal{U}_{\mathrm{ad}}$, such that $\ell(y_k, u_k)$ tends to $\inf_{u\in\mathcal{U}_{\mathrm{ad}}} \ell(y_u, u) > -\infty$ as k goes to infinity, where $y_k \coloneqq y_{u_k}$. Due to the assumptions, the sequence (u_k) admits a subsequence, denoted by the same name, which converges weakly in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ to some limit $\bar{u}$. (We will keep the notation for all convergent subsequences in the following.) Indeed, if $\beta_2 > 0$, this is true because then the sequence (u_k) is bounded in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$. If $\omega \in \mathrm{L}^1(0,\mathsf{T})$, then $\mathcal{U}_{\mathrm{ad}}$ is in fact weakly compact in $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$, cf. [10, Cor. 2.6]. Due to weak closedness of $\mathcal{U}_{\mathrm{ad}}$, we also have $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$.

We turn to (y_k): The boundedness of (u_k) implies that (ξ_{y_k}) must be bounded in $\mathrm{L}^\infty(0,\mathsf{T};\mathcal{E})$ by the energy bound (2.4). This gives a weakly-$*$ convergent subsequence of (ξ_{y_k}) with the weak-$*$ limit denoted by $\bar{y} \in \mathrm{L}^\infty(0,\mathsf{T};\mathcal{E})$. We need to show that $\bar{y} = y_{\bar{u}}$.

Looking at the definition of a Shatah-Struwe solution

$$
\begin{aligned}
-\int_0^{\mathsf{T}} \big(\partial_t y_k(t), \partial_t \varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^{\mathsf{T}} \big(\nabla y_k(t), \nabla \varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^{\mathsf{T}} \big(y_k^5(t), \varphi(t)\big)_\Omega \,\mathrm{d}t \\
= \int_0^{\mathsf{T}} \big(u_k(t), \varphi(t)\big)_\Omega \,\mathrm{d}t \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty\big((0,\mathsf{T}) \times \Omega\big), \quad (4.1)
\end{aligned}
$$

we observe that the linear terms are already dealt with. It remains to show that in fact $\bar{y} \in \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$, that $\xi_{\bar{y}}(0) = \xi_0$, and that the nonlinear term involving y_k^5 converges to the correct one involving $\bar{y}^5$.

For the latter, let $5 \le p < 6$. Then boundedness of (ξ_{y_k}) in $\mathrm{L}^\infty(0,\mathsf{T};\mathcal{E})$ and compactness of the embedding $\mathrm{H}_0^1(\Omega) \hookrightarrow \mathrm{L}^p(\Omega)$ implies that (y_k) is in fact a pre*compact* set in $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^p(\Omega))$, see [36, Cor. 4]. Accordingly, there is yet another subsequence of (y_k) denoted by the same name such that (y_k) tends to $\bar{y}$ in that space. This further means that (y_k^5) converges to $\bar{y}^5$ in $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^{p/5}(\Omega))$.

The next step is to show that $\xi_{\bar{y}}(0) = \xi_0$. We already know that (ξ_{y_k}) is bounded in $\mathrm{L}^\infty(0,\mathsf{T};\mathcal{E})$. Further, (2.5) and boundedness of (u_k) in $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ show that $(\partial_t^2 y_k)$ is bounded in $\mathrm{L}^1(0,\mathsf{T};\mathrm{H}^{-1}(\Omega))$. Using again [36, Cor. 4], we infer that (ξ_{y_k}) is precompact in $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega) \times \mathrm{H}^{-1}(\Omega))$ and thus admits a convergent subsequence in that space, with the same name and the limit $\bar{y}$. Thus, in particular,

$$
\xi_0 = \xi_{y_k}(0) \longrightarrow \xi_{\bar{y}}(0) \quad \text{in } \mathrm{H}^{-1}(\Omega) \times \mathrm{L}^2(\Omega)
$$

and we obtain $\xi_{\bar{y}}(0) = \xi_0$, which by assumption even lies in $\mathcal{E}$.

We are now prepared to taking limits in (4.1) and obtain that $\xi_{\bar{y}} \in \mathrm{L}^\infty(0,\mathsf{T};\mathcal{E})$ satisfies $\xi_{\bar{y}}(0) = \xi_0$ and

$$-\int_0^{\mathsf{T}} \big(\partial_t \bar{y}(t), \partial_t \varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^{\mathsf{T}} \big(\nabla \bar{y}(t), \nabla \varphi(t)\big)_\Omega \,\mathrm{d}t + \int_0^{\mathsf{T}} \big(\bar{y}^5(t), \varphi(t)\big)_\Omega \,\mathrm{d}t$$
$$= \int_0^{\mathsf{T}} \big(\bar{u}(t), \varphi(t)\big)_\Omega \,\mathrm{d}t \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty\big((0,\mathsf{T}) \times \Omega\big). \quad (4.2)$$

It only remains to show that $\bar{y} \in \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$. From $\gamma > 0$ we infer that (y_k) is also bounded in $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$. Hence, there exists a weakly convergent subsequence with the limit $\hat{y}$ in $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$. This means that we have

$$\int_0^{\mathsf{T}} \big(\bar{y}(t), \phi(t)\big)_\Omega \,\mathrm{d}t = \int_0^{\mathsf{T}} \big(\hat{y}(t), \phi(t)\big)_\Omega \,\mathrm{d}t$$
$$\text{for all } \phi \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^{6/5}(\Omega)) \cap \mathrm{L}^{4/3}(0,\mathsf{T};\mathrm{L}^{12/11}(\Omega))$$

and thus

$$\bar{y} = \hat{y} \quad \text{in} \quad \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^6(\Omega)) + \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega)) \hookrightarrow \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^6(\Omega)).$$

But then $\bar{y} = \hat{y}$ almost everywhere in $(0,\mathsf{T}) \times \Omega$ and in fact $\bar{y} \in \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$.

Since $\bar{y}$ has now been shown to be a Shatah-Struwe solution to (CWE), uniqueness of such solutions as established in Corollary 2.8 then finally implies that indeed $\bar{y} = y_{\bar{u}}$.

It is now standard to use the convergences $(y_k) \to \bar{y}$ in $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega))$ and $(y_k) \rightharpoonup \bar{y}$ in $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$ as well as $(u_k) \rightharpoonup \bar{u}$ in $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ together with lower semicontinuity of norms in the objective ℓ to infer that

$$\ell(\bar{y},\bar{u}) = \inf_{u \in \mathcal{U}_{\mathrm{ad}}} \ell(y_u, u).$$

This shows that there indeed exists a globally optimal control to (OCP). □

From the proof of Theorem 4.1 we obtain the following auxiliary result for the case $\gamma = 0$, so the case where there is no additional $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$ norm term in the objective. It underlines the role of this term in upgrading weak solutions to (unique) mild solutions:

Proposition 4.2. *Let $\beta_2 > 0$ or $\omega \in \mathrm{L}^1(0,\mathsf{T})$. Then the optimal control problem* (OCP) *admits at least one globally optimal pair $(\bar{y},\bar{u})$ with $\bar{u} \in \mathcal{U}_{ad}$ such that $\bar{y}$ is a—possibly non-unique—weak solution to* (CWE) *with right-hand side $\bar{u}$. That means we have $\xi_{\bar{y}}(0) = \xi_0$ and $\xi_{\bar{y}} \in \mathrm{L}^\infty([0,\mathsf{T}];\mathcal{E})$, and $\bar{y}$ satisfies the weak formulation* (4.2).

4.2. Optimality conditions

Let us set

$$\mathcal{Y} := \Big\{ y \colon \xi_y \in \mathrm{C}([0,\mathsf{T}];\mathcal{E}) \Big\} \quad \text{and} \quad \mathcal{Y}_+ := \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega)) \cap \mathcal{Y}$$

with

$$\|y\|_{\mathcal{Y}} := \|\xi_y\|_{\mathrm{C}([0,\mathsf{T}];\mathcal{E})} \quad \text{and} \quad \|y\|_{\mathcal{Y}_+} := \|y\|_{\mathcal{Y}} + \|y\|_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))}.$$

As a first step towards optimality conditions, we show that the control-to-state mapping $u \mapsto y_u$ is twice continuously differentiable from $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ to $\mathcal{Y}_+$. We recall the definition of y_u in terms of the variation-of-constants (or Duhamel) formula in (2.9) and define the mapping

$$e\colon \mathcal{Y}_+ \times \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega)) \to \mathrm{C}([0,\mathsf{T}];\mathcal{E})$$

as follows:

$$\big[e(y,u)\big](t) := e^{\mathcal{A}t}\xi_0 + \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ u(s) - y^5(s) \end{pmatrix} \mathrm{d}s - \xi_y(t).$$

By construction it is clear that $e(y_u,u)=0$ and of course it is our first goal to use the implicit function theorem to show the following:

Theorem 4.3. *The control-to-state operator $\mathcal{S}\colon u \mapsto y_u$ is twice continuously differentiable from $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ to $\mathcal{Y}_+$. Its derivative $\mathcal{S}'(\bar u)h$ in $\bar u$ in direction $h \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ is given by the mild solution $z_h \in \mathcal{Y}_+$ of*

$$\begin{aligned} \partial_t^2 z - \Delta z + 5\bar y_{\bar u}^4 z &= h && \text{in } (0,T)\times\Omega,\\ z &= 0 && \text{on } (0,T)\times\partial\Omega,\\ \big(z(0),\partial_t z(0)\big) &= (0,0) && \text{in } \Omega, \end{aligned}$$

on $[0,\mathsf{T}]$, i.e., z_h satisfies

$$\xi_z(t) = \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ h(s) - 5y_{\bar u}^4(s)z(s) \end{pmatrix} \mathrm{d}s$$

for all $t \in [0,\mathsf{T}]$. Its second derivative $\mathcal{S}''(\bar u)(h_1,h_2)$ in $\bar u$ in directions $h_1,h_2 \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ is given by the mild solution $z_{h_1,h_2} \in \mathcal{Y}_+$ of

$$\begin{aligned} \partial_t^2 z - \Delta z + 5\bar y_{\bar u}^4 z &= -20\bar y_{\bar u}^3 z_{h_1} z_{h_2} && \text{in } (0,T)\times\Omega,\\ z &= 0 && \text{on } (0,T)\times\partial\Omega,\\ \big(z(0),\partial_t z(0)\big) &= (0,0) && \text{in } \Omega, \end{aligned}$$

on $[0,\mathsf{T}]$, where $z_{h_i} = \mathcal{S}'(\bar u)h_i$ for $i=1,2$, i.e., z_{h_1,h_2} satisfies

$$\xi_z(t) = \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ -20\bar y_{\bar u}^3(s) z_{h_1}(s) z_{h_2}(s) - 5y_{\bar u}^4(s)z(s) \end{pmatrix} \mathrm{d}s$$

for all $t \in [0,\mathsf{T}]$.

Proof. We begin by showing that e is twice continuously differentiable. Clearly, $y \mapsto \xi_y$ is a continuous linear mapping from $\mathcal{Y}$ into $\mathrm{C}([0,\mathsf{T}];\mathcal{E})$, just as

$$v \mapsto \left[t \mapsto \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ v(s) \end{pmatrix} \mathrm{d}s \right] \tag{4.3}$$

is a continuous linear mapping from $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ to $\mathrm{C}([0,\mathsf{T}];\mathcal{E})$. It is thus sufficient to show that $y \mapsto y^5$ is twice continuously differentiable considered as a mapping from $\mathcal{Y}_+$

to $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$. This, however, follows immediately from the interpolation inequality (2.2), which implies that

$$\mathcal{Y}_+ \hookrightarrow \mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega)) \cap \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^6(\Omega)) \hookrightarrow \mathrm{L}^5(0,\mathsf{T};\mathrm{L}^{10}(\Omega)), \tag{4.4}$$

together with twice continuous differentiability of the Nemytskii operator induced by the real function $x \mapsto x^5$ between $\mathrm{L}^5(0,\mathsf{T};\mathrm{L}^{10}(\Omega))$ and $\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$, cf. [18, Thms. 7&9]. Altogether e is twice continuously differentiable.

In order to use the implicit function theorem, it remains to show that $e_y(y_u,u)$ is continuously invertible as a linear operator between $\mathcal{Y}_+$ and $\mathrm{C}([0,\mathsf{T}];\mathcal{E})$ for every $u \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$. From the foregoing considerations and the chain rule, we obtain

$$\big[e_y(y_u,u)z\big](t) = -\int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ 5y_u^4(s)z(s) \end{pmatrix} \,\mathrm{d}s - \xi_z(t).$$

Thanks to the open mapping theorem, it will be sufficient to prove that for every $F \in \mathrm{C}([0,\mathsf{T}];\mathcal{E})$ there is a unique $z \in \mathcal{Y}_+$ such that $e_y(y_u,u)z = F$. We can again use a fixed point theorem to show that this is the case: Choose a partition $0 = t_0 < t_1 < \cdots < t_n = \mathsf{T}$ such that

$$5\big(C_e(\mathsf{T}) + C_s(\mathsf{T})\big)C_{\mathrm{em}}\|y_u\|^4_{\mathrm{L}^4(t_i,t_{i+1};\mathrm{L}^{12}(\Omega))} < \frac{1}{2} \quad \text{for all } i = 0,\ldots,n-1,$$

where C_{em} is the embedding constant of $\mathrm{H}^1_0(\Omega) \hookrightarrow \mathrm{L}^6(\Omega)$; for the constants C_s, C_e, see Lemma 2.9. Let $\mathcal{Y}(t_i,t_{i+1})$ and $\mathcal{Y}_+(t_i,t_{i+1})$ be the spaces $\mathcal{Y}$ and $\mathcal{Y}_+$ on the interval $[t_i,t_{i+1}]$, *mutatis mutandis*. Let further $\xi^i \in \mathcal{E}$ for $i = 0,\ldots,n-1$ be given and consider the mappings $\mathcal{T}_i\colon \mathcal{Y}(t_i,t_{i+1}) \to \mathcal{Y}(t_i,t_{i+1})$ defined by $\mathcal{T}_i h = z$ such that

$$\xi_z(t) = e^{\mathcal{A}(t-t_i)}\xi^i - \int_{t_i}^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ 5y_u^4(s)h(s) \end{pmatrix} \,\mathrm{d}s - F(t) \quad \text{for } t \in [t_i,t_{i+1}].$$

Then, by the estimates (2.7) and (2.8), semigroup properties and Hölder's inequality,

$$\begin{aligned} &\big\|\mathcal{T}_i h_1 - \mathcal{T}_i h_2\big\|_{\mathcal{Y}_+(t_i,t_{i+1})} \\ &\qquad \le 5\big(C_e(\mathsf{T}) + C_s(\mathsf{T})\big)C_{\mathrm{em}}\|y_u\|^4_{\mathrm{L}^4(t_i,t_{i+1};\mathrm{L}^{12}(\Omega))}\big\|h_1 - h_2\big\|_{\mathrm{C}([t_i,t_{i+1}];\mathrm{H}^1_0(\Omega))}. \end{aligned}$$

The choice of the partition (t_i) and Banach's fixed point theorem tell us that every mapping $\mathcal{T}_i$ possesses a unique fixed point $z_i \in \mathcal{Y}_+(t_i,t_{i+1})$, still depending on ξ^i. If we iteratively choose $\xi^0 = 0$ and $\xi^i = z_{i-1}(t_i)$ for $i = 1,\ldots,n-1$ and glue together the resulting functions z_i to a function $z \in \mathcal{Y}_+$, then we obtain

$$\xi_z(t) + \int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ 5y_u^4(s)z(s) \end{pmatrix} \,\mathrm{d}s = -F(t) \quad \text{for all } t \in [0,\mathsf{T}], \tag{4.5}$$

i.e., $z \in \mathcal{Y}_+$ satisfies $e_y(y_u,u)z = F$. Thus, $e_y(y_u,u)^{-1} \in \mathcal{L}(\mathrm{C}([0,\mathsf{T}];\mathcal{E});\mathcal{Y}_+)$.

Finally, the expression for the derivative $\mathcal{S}'(\bar{u})$ comes from the well known formula

$$\mathcal{S}'(\bar{u})h = -e_y(\mathcal{S}(\bar{u}),\bar{u})^{-1}e_u(\mathcal{S}(\bar{u}),\bar{u})h,$$

the observation that $e_u(\mathcal{S}(\bar{u}), \bar{u})$ is given exactly by (4.3), and plugging this into (4.5) for F. For the second derivative $\mathcal{S}''(\bar{u})$, we take another derivative in the foregoing expression and use

$$\big[e_{yy}(y_u, u)(z_1, z_2)\big](t) = -\int_0^t e^{\mathcal{A}(t-s)} \begin{pmatrix} 0 \\ 20 y_u^3(s) z_1(s) z_2(s) \end{pmatrix} \mathrm{d}s.$$

This gives the claim. □

As usual, the control-to-state operator $\mathcal{S}$ allows us to define the *reduced problem* which we consider from now on:

$$\min_{u \in \mathcal{U}_{\mathrm{ad}}} \ell_r(u), \tag{ROCP}$$

where we set $\ell_r(u) \coloneqq \ell(y_u, u)$. We decompose the objective function ℓ_r further into

$$\ell_r(u) \coloneqq F(u) + \beta_1 \mathsf{j}(u) \tag{4.6}$$

with

$$F(u) = \frac{1}{2}\|y_u(\mathsf{T}) - y_d\|^2_{\mathrm{L}^2(\Omega)} + \frac{\gamma}{4}\|y_u\|^4_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))} + \frac{\beta_2}{2}\|u\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))},$$

which is smooth as we see below, and the non-differentiable part

$$\mathsf{j}(u) \coloneqq \|u\|_{\mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))}.$$

For the following derivation of necessary and sufficient optimality conditions, we use several ideas and results from [5]. The main difference between this work and the present one is that the constraints on u in [5] are classical box constraints.

We first quote the following result; it is a characterization of the subdifferential $\partial\mathsf{j}$ and a formula for the directional derivative of j:

Proposition 4.4 ([5, Prop. 3.8])**.** *Let $u \in \mathcal{U}_{ad}$ and $\lambda \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$. Then the following equivalence holds true:*

$$\lambda \in \partial\mathsf{j}(u) \quad \iff \quad \textit{f.a.a. } t \in [0,\mathsf{T}]\colon \begin{cases} \lambda(t) = \frac{u(t)}{\|u(t)\|_{\mathrm{L}^2(\Omega)}} & \textit{if } \|u(t)\|_{\mathrm{L}^2(\Omega)} \neq 0, \\ \|\lambda(t)\|_{\mathrm{L}^2(\Omega)} \le 1 & \textit{if } \|u(t)\|_{\mathrm{L}^2(\Omega)} = 0. \end{cases}$$

Moreover, the directional derivative of j *in $u \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ exists in every direction $v \in \mathrm{L}^1(0,\mathsf{T};\mathrm{L}^2(\Omega))$ and is given by*

$$\mathsf{j}'(u;v) = \int_{\left[\|u\|_{\mathrm{L}^2(\Omega)}=0\right]} \|v(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t + \int_{\left[\|u\|_{\mathrm{L}^2(\Omega)}\neq 0\right]} \frac{\big(v(t), u(t)\big)_\Omega}{\|u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t.$$

We next establish that F is twice continuously differentiable. For a concise form of its derivatives, it will be useful to define the adjoint state:

Definition 4.5 (Adjoint state)**.** Given $\bar u \in \mathcal{U}_{\mathrm{ad}}$, we denote by $\bar p$ the *adjoint state* defined by

$$\bar p := \mathcal{S}'(\bar u)^* \big(\delta_{\mathsf{T}}^*(y_{\bar u}(\mathsf{T}) - y_d) + \gamma \psi_{\bar u}\big) \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega)).$$

Here $\psi_{\bar u} \in \mathrm{L}^{4/3}(0,\mathsf{T};\mathrm{L}^{12/11}(\Omega))$ is given by

$$\psi_{\bar u}(t) := \|y_{\bar u}(t)\|_{\mathrm{L}^{12}(\Omega)}^{-8} |y_{\bar u}(t)|^{10} y_{\bar u}(t) \tag{4.7}$$

and δ_{T}^* is the adjoint operator of the (continuous linear) point evaluation δ_{T} from $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega))$ to $\mathrm{L}^2(\Omega)$.

The function $\psi_{\bar u}$ is the L^2-gradient of $\Psi(y) := \frac{1}{4}\|y\|^4_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega)}$ in $y_{\bar u}$, that is,

$$\Psi'(y_{\bar u})h = \int_0^{\mathsf{T}} \big(\psi_{\bar u}(t), h(t)\big)_\Omega \,\mathrm{d}t \quad \text{for all } h \in \mathrm{L}^4\big(0,\mathsf{T};\mathrm{L}^{12}(\Omega)\big). \tag{4.8}$$

This is shown in Corollary A.4 in the appendix. This corollary is also important in the next and final result for this preparationary subsection:

Lemma 4.6. *The first summand F of the reduced objective function ℓ_r as in* (4.6) *is twice continuously differentiable from* $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ *to* $\mathbb{R}$*. Its derivatives in $\bar u$ are given by*

$$F'(\bar u)v = \int_0^{\mathsf{T}} \big(\bar p(t) + \beta_2 \bar u(t), v(t)\big)_\Omega \,\mathrm{d}t$$

and

$$F''(\bar u)v^2 = \big\|z_v(\mathsf{T})\big\|^2_{\mathrm{L}^2(\Omega)} - \int_0^{\mathsf{T}} \big(\bar p(t), 20 y_{\bar u}^3(t) z_v^2(t)\big)_\Omega \,\mathrm{d}t + \gamma \Psi''(y_{\bar u}) z_v^2 + \beta_2 \big\|v\big\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}$$

for all $v \in \mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$, where $\bar p$ is the adjoint state and $z_v = \mathcal{S}'(\bar u)v$.

Moreover, the quadratic form $v \mapsto F''(\bar u)v^2$ is weakly lower semicontinuous.

We recall that $r = 1$ if $\beta_2 = 0$ and $r = 2$ otherwise.

Proof of Lemma 4.6. Corollary A.4 in the appendix shows that Ψ is twice continuously differentiable on the whole $\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))$, with the first derivative as in (4.8). The remaining differentiability assertions for F and the formula for $F''(\bar u)$ are derived by routine calculations.

The remainder of this proof is devoted to verifying the weak lower semicontinuity of the quadratic form induced by $F''(\bar u)$. Let $v_k \rightharpoonup v$ in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$. It is clear that the quadratic $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$ norm term in $F''(\bar u)v^2$ is weakly lower semicontinuous in v. We show that the remaining terms are even weakly continuous as functions in v. From $v_k \rightharpoonup v$ in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ it follows that $z_k := \mathcal{S}'(\bar u)v_k \rightharpoonup z_v =: \mathcal{S}'(\bar u)v$ in $\mathcal{Y}_+$ by Theorem 4.3. This implies that, cf. (4.4),

$$\|z_k\|_{\mathrm{L}^5(0,\mathsf{T};\mathrm{L}^{10}(\Omega))} \lesssim \|z_k\|_{\mathcal{Y}} \lesssim 1. \tag{4.9}$$

By the compact embedding $\mathcal{Y}_+ \hookrightarrow \mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega))$ as derived from [36, Cor. 4]), we further have $z_k \to z_v$ in $\mathrm{C}([0,\mathsf{T}];\mathrm{L}^2(\Omega))$. Hence, $z_k(t) \to z_v(t)$ in $\mathrm{L}^2(\Omega)$ for every $t \in [0,\mathsf{T}]$, and in particular $\|z_k(\mathsf{T})\|^2_{\mathrm{L}^2(\Omega)} \to \|z_v(\mathsf{T})\|^2_{\mathrm{L}^2(\Omega}$ as $k \to \infty$. Moreover, there is a subsequence (z_{k_ℓ}) of (z_k) such that $z^2_{k_\ell}$ converges to z_v^2 pointwise almost everywhere on $(0,\mathsf{T}) \times \Omega$. Hölder's inequality yields

$$\int_{E_t}\int_{E_\mathrm{x}} \left|\bar{p}(t,\mathrm{x})\, y^3_{\bar{u}}(t,\mathrm{x})\, z^2_{k_\ell}(t,\mathrm{x})\right| \,\mathrm{d}\mathrm{x}\,\mathrm{d}t \\ \leq \big\|p\big\|_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))}\big\|y_{\bar{u}}\big\|^3_{\mathrm{L}^5(E_t;\mathrm{L}^{10}(E_\mathrm{x}))}\big\|z_{k_\ell}\big\|^2_{\mathrm{L}^5(0,\mathsf{T};\mathrm{L}^{10}(\Omega))}$$

for every measurable subset $E_t \times E_\mathrm{x}$ of $(0,\mathsf{T}) \times \Omega$. Due to (4.9), the integral on the left-hand side thus goes to zero uniformly in ℓ as $|E_t \times E_\mathrm{x}| \to 0$. We infer that the functions $(\bar{p}y^3_{\bar{u}}z^2_{k_\ell})$ are uniformly integrable and the Vitali convergence theorem ([12, Thm. III.3.6.6]) implies that

$$-\int_0^\mathsf{T} \big(\bar{p}(t), 20y^3_{\bar{u}}(t)z^2_{k_\ell}(t)\big)_\Omega \,\mathrm{d}t \quad \xrightarrow{\ell\to\infty} \quad -\int_0^\mathsf{T} \big(\bar{p}(t), 20y^3_{\bar{u}}(t)z^2_v(t)\big)_\Omega \,\mathrm{d}t.$$

A subsequence-subsequence argument shows that this convergence in fact holds true for the whole sequence (z_k).

We next turn to the sequence $(\Psi''(y_{\bar{u}})z_k^2)$ which is the last term in $F''(\bar{u})v_k^2$. According to Corollary A.4 with $p = 4$ and $q = 12$, we have

$$\Psi''(y_{\bar{u}})z_k^2 = 11\int_0^\mathsf{T} \|y_{\bar{u}}(t)\|^{-8}_{\mathrm{L}^{12}(\Omega)}\big(|y_{\bar{u}}(t)|^{10}, z_k^2(t)\big)_\Omega \,\mathrm{d}t \\ - 8\int_0^\mathsf{T} \|y_{\bar{u}}(t)\|^{-20}_{\mathrm{L}^{12}(\Omega)}\big(|y_{\bar{u}}(t)|^{10}y_{\bar{u}}(t), z_k(t)\big)^2_\Omega \,\mathrm{d}t. \tag{4.10}$$

The limit for the first term

$$\lim_{k\to\infty}\int_0^\mathsf{T} \|y_{\bar{u}}(t)\|^{-8}_{\mathrm{L}^{12}(\Omega)}\big(|y_{\bar{u}}(t)|^{10}, z_k^2(t)\big)_\Omega \,\mathrm{d}t = \int_0^\mathsf{T} \|y_{\bar{u}}(t)\|^{-8}_{\mathrm{L}^{12}(\Omega)}\big(|y_{\bar{u}}(t)|^{10}, z_v^2(t)\big)_\Omega \,\mathrm{d}t$$

can be proven analogously to the above with Hölder's inequality and the Vitali theorem due to boundedness of (z_k) in $\mathcal{Y}_+$. For the second term in (4.10), we first show that $\big(|y_{\bar{u}}(t)|^{10}y_{\bar{u}}(t), z_k(t)\big)_\Omega$ converges towards $\big(|y_{\bar{u}}(t)|^{10}y_{\bar{u}}(t), z_v(t)\big)_\Omega$ pointwise a.e. on $[0,\mathsf{T}]$ as $k \to \infty$. Then convergence of the overall integral follows from Hölder's inequality

$$\int_{E_t} \|y_{\bar{u}}(t)\|^{-20}_{\mathrm{L}^{12}(\Omega)}\big(|y_{\bar{u}}(t)|^{10}y_{\bar{u}}(t), z_k(t)\big)^2_\Omega \,\mathrm{d}t \\ \leq \|y_{\bar{u}}(t)\|^2_{\mathrm{L}^4(E_t;\mathrm{L}^{12}(\Omega))}\|z_k(t)\|^2_{\mathrm{L}^4(0,\mathsf{T};\mathrm{L}^{12}(\Omega))}$$

for every measurable subset $E_t \subseteq (0,\mathsf{T})$ and yet another application of the Vitali theorem, using boundedness of (z_k) in $\mathcal{Y}_+$.

Pointwise convergence of $(\big(|y_{\bar{u}}(t)|^{10}y_{\bar{u}}(t), z_k(t)\big)_\Omega)$ can be obtained as follows: We had seen that $z_k(t) \to z_v(t)$ in $\mathrm{L}^2(\Omega)$ for every $t \in [0,\mathsf{T}]$. Thus for every such t there exists a subsequence $(z_{k_m}(t))$ such that $z_{k_m}(t) \to z_v(t)$ almost everywhere on Ω. Moreover, we estimate for every measurable subset $E_\mathrm{x} \subseteq \Omega$

$$\int_{E_\mathrm{x}} |y_{\bar{u}}(t)|^{11} |z_{k_m}(t)| \,\mathrm{dx} \leq \|y_{\bar{u}}(t)\|^{11}_{\mathrm{L}^{\frac{66}{5}}(E_\mathrm{x})} \|z_{k_m}(t)\|_{\mathrm{L}^6(\Omega)} \lesssim \|y_{\bar{u}}(t)\|^{11}_{\mathrm{L}^{\frac{66}{5}}(E_\mathrm{x})},$$

since (z_{k_m}) is bounded in $\mathcal{Y}_+ \hookrightarrow \mathrm{C}([0,\mathsf{T}];\mathrm{L}^6(\Omega))$. The general form of the Strichartz estimates (2.8) as in [2] shows that in fact $y_{\bar{u}} \in \mathrm{L}^{\frac{11}{3}}(0,\mathsf{T};\mathrm{L}^{\frac{66}{5}}(\Omega))$. Since t was fixed, the foregoing expression thus goes to zero uniformly in m as $|E_\mathrm{x}|$ goes to zero. Hence, again the Vitali convergence theorem together with a subsequence-subsequence argument shows that

$$\int_\Omega |y_{\bar{u}}(t)|^{10} y_{\bar{u}}(t) z_k(t) \,\mathrm{dx} \quad \xrightarrow{k\to\infty} \quad \int_\Omega |y_{\bar{u}}(t)|^{10} y_{\bar{u}}(t) z_v(t) \,\mathrm{dx}.$$

This finally implies that overall $\Psi''(y_{\bar{u}}) z_k^2 \to \Psi''(y_{\bar{u}}) z_v^2$ and finishes the proof. □

4.2.1. First-order necessary conditions

The *tangent cone* to $\mathcal{U}_\mathrm{ad}$ in a point $u \in \mathcal{U}_\mathrm{ad}$ is

$$\mathcal{T}(u) := \Big\{ v \in \mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega)) \colon \big(v(t), u(t)\big)_\Omega \leq 0 \text{ if } \|u(t)\|_{\mathrm{L}^2(\Omega)} = \omega(t) > 0, \\ v(t) = 0 \text{ if } \|u(t)\|_{\mathrm{L}^2(\Omega)} = \omega(t) = 0 \Big\}.$$

We use it to state the basic first-order necessary optimality condition for (ROCP) in a concise form. We refer to [5, Thm. 3.1] for the routine proof.

Theorem 4.7 (First-order necessary optimality condition). *Let $\bar{u} \in \mathcal{U}_{ad}$ be a locally optimal solution of* (ROCP). *Then there exists $\bar{\lambda} \in \partial \mathsf{j}(\bar{u}) \subset \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ such that*

$$\int_0^\mathsf{T} \big(\bar{p}(t) + \beta_1 \bar{\lambda}(t) + \beta_2 \bar{u}(t), v(t)\big)_\Omega \,\mathrm{d}t \geq 0 \quad \textit{for all } v \in \mathcal{T}(\bar{u}). \tag{4.11}$$

Note that since j is Lipschitz continuous and convex, $\bar{\lambda} \in \partial \mathsf{j}(\bar{u})$ and (4.11) imply that (cf. [5, Lem. 4.2])

$$F'(\bar{u})v + \beta_1 \mathsf{j}'(\bar{u}; v) \geq 0 \quad \text{for all } v \in \mathcal{T}(\bar{u}). \tag{4.12}$$

We henceforth always consider a fixed locally optimal control $\bar{u} \in \mathcal{U}_\mathrm{ad}$ and the optimality condition (4.11) as given. Let us further subdivide $(0,\mathsf{T})$ into active and inactive regions w.r.t. the constraint in (ROCP) by defining the following sets:

$$\mathcal{I} := \big[\|\bar{u}\|_{\mathrm{L}^2(\Omega)} < \omega\big], \quad \mathcal{A}_+ := \big[\|\bar{u}\|_{\mathrm{L}^2(\Omega)} = \omega > 0\big], \quad \mathcal{A}_0 := \big[\omega = 0\big].$$

We first show that the integrated optimality condition (4.11) is equivalent to the pointwise one.

Corollary 4.8. *Condition* (4.11) *is equivalent to*

$$\big(\bar{p}(t) + \beta_1 \bar{\lambda}(t) + \beta_2 \bar{u}(t), v(t)\big)_\Omega \geq 0 \quad \textit{for almost all } t \in (0,\mathsf{T}) \tag{4.13}$$

for all $v \in \mathcal{T}(\bar{u})$. It moreover follows that

$$\bar{p}(t) + \beta_1 \bar{\lambda}(t) + \beta_2 \bar{u}(t) = 0 \quad \textit{for almost all } t \in \mathcal{I}. \tag{4.14}$$

Proof. It is obvious that (4.13) implies (4.11). For the other way around, it suffices to observe that if $v \in \mathcal{T}(\bar{u})$, then also $\chi_N v \in \mathcal{T}(\bar{u})$ for every measurable set $N \subseteq (0,\mathsf{T})$, so one can do the usual proof by contradiction. Quite similarly, (4.14) follows from inserting $\pm\chi_{\mathcal{I}}\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega)) \subset \mathcal{T}(\bar{u})$ into (4.13). □

The next result is then an observation regarding *sparsity* and regularity of an optimal control $\bar{u}$, as well as uniqueness of the subgradient $\bar{\lambda}$. The proof is analogous to the one in [5, Cor. 3.9] using (4.13) and (4.14).

Corollary 4.9. *The following properties hold true:*

- *If* $\beta_2 > 0$*: We have* $\bar{u} \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ *and thus* $\omega \in \mathrm{L}^\infty(\mathcal{A}_+)$*. Moreover, for almost all* $t \in \mathcal{I}$*, the following equivalence holds true:*

$$\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} = 0 \quad \iff \quad \|\bar{p}(t)\|_{\mathrm{L}^2(\Omega)} \leq \beta_1.$$

- *If* $\beta_2 = 0$*: For almost all* $t \in \mathcal{I}$*, we have the implications*

$$\|\bar{p}(t)\|_{\mathrm{L}^2(\Omega)} < \beta_1 \quad \Longrightarrow \quad \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} = 0 \quad \Longrightarrow \quad \|\bar{p}(t)\|_{\mathrm{L}^2(\Omega)} \leq \beta_1.$$

In both cases, $\bar{\lambda} \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ *satisfies*

$$\bar{\lambda}(t) = \begin{cases} \frac{\bar{u}(t)}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}} & \text{if } \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} \neq 0, \\ -\frac{1}{\beta_1}\bar{p}(t) & \text{if } \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} = 0 \end{cases} \qquad \text{f.a.a. } t \in \mathcal{A}_+ \cup \mathcal{I}.$$

It is thus unique on $\mathcal{A}_+ \cup \mathcal{I}$*.*

Using Corollaries 4.8 and 4.9 we can show that a unique bounded Lagrange multiplier associated to $\mathcal{U}_{\mathrm{ad}}$ and the locally optimal control $\bar{u}$ exists. We use the convention that $\frac{0}{0} = 0$.

Definition 4.10 (Lagrange multiplier)**.** We say that a measurable function $\bar{\mu}\colon \mathcal{A}_+ \cup \mathcal{I} \to [0,\infty)$ is a *Lagrange multiplier associated to* $\mathcal{U}_{ad}$ if $\bar{\mu}(t)(\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} - \omega(t)) = 0$ for almost all $t \in \mathcal{A}_+ \cup \mathcal{I}$ is satisfied (complementarity), and the gradient equation

$$\bar{p}(t) + \beta_1\bar{\lambda}(t) + \beta_2\bar{u}(t) + \bar{\mu}(t)\frac{\bar{u}(t)}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}} = 0 \quad \text{for almost all } t \in \mathcal{A}_+ \cup \mathcal{I} \tag{4.15}$$

holds true.

Lemma 4.11. *There exists a unique Lagrange multiplier* $\bar{\mu} \in \mathrm{L}^\infty(\mathcal{A}_+ \cup \mathcal{I})$ *associated to* $\mathcal{U}_{ad}$*.*

Proof. We set, of course, necessarily $\bar{\mu}(t) = 0$ for $t \in \mathcal{I}$. Then the complementarity condition and (4.15) on $\mathcal{I}$ are already satisfied, the latter due to (4.14).

The next step is to show that there exists a Lagrange multiplier $\bar{\mu} \in \mathrm{L}^1(\mathcal{A}_+)$ associated to $\mathcal{U}_{\mathrm{ad}}$. This $\bar{\mu}$ is then necessarily already an element of $\mathrm{L}^\infty(\mathcal{A}_+)$ and moreover unique, which we see as follows: From taking $\mathrm{L}^2(\Omega)$ norms in (4.15) for $t \in \mathcal{A}_+$ it follows that

$$\left\|\bar{p}(t) + \beta_1\bar{\lambda}(t) + \beta_2\bar{u}(t)\right\|_{\mathrm{L}^2(\Omega)} = \bar{\mu}(t) \quad \text{for almost all } t \in \mathcal{A}_+.$$

The left-hand side is an $\mathrm{L}^\infty(\mathcal{A}_+)$ function in t and unique due to Corollary 4.9.

It thus remains to show that the $\mathrm{L}^1(\mathcal{A}_+)$ Lagrange multiplier exists in the first place. Suppose the contrary, i.e., that $\bar{p} + \beta_1\bar{\lambda} + \beta_2\bar{u} \neq -\mu\|\bar{u}\|_{\mathrm{L}^2(\Omega)}^{-1}\bar{u}$ in $\mathrm{L}^1(\mathcal{A}_+;\mathrm{L}^2(\Omega))$ for all $\mu \in \mathrm{L}^1(\mathcal{A}_+)$ with $\mu \geq 0$ a.e.. Then the Hahn-Banach theorem yields a function $\varphi \in \mathrm{L}^\infty(\mathcal{A}_+)$ such that

$$\int_{\mathcal{A}_+} \frac{\big(\varphi(t), -\mu(t)\bar{u}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t \leq 0 < \int_{\mathcal{A}_+} \big(\varphi(t), \bar{p}(t) + \beta_1\bar{\lambda}(t) + \beta_2\bar{u}(t)\big)_\Omega\,\mathrm{d}t.$$

From the first inequality it follows that $-\chi_{\mathcal{A}_+}\varphi \in \mathcal{T}(\bar{u})$ (proof by contradiction) which however is incompatible with the second one by the first order necessary condition (4.11). Hence, there exists the searched-for $\bar{\mu} \in \mathrm{L}^1(\mathcal{A}_+)$ satisfying $\bar{\mu} \geq 0$ and (4.15) on $\mathcal{A}_+$. This finishes the proof ☐

Henceforth, $\bar{\mu}$ will denote the unique Lagrange multiplier associated to $\mathcal{U}_{\mathrm{ad}}$ for the locally optimal control $\bar{u}$.

4.2.2. Second order necessary conditions

We define the *critical cone* $C(\bar{u})$ associated to $\mathcal{U}_{\mathrm{ad}}$ in a locally optimal control $\bar{u}$ to consist of tangential directions along which the directional derivative of ℓ_r vanishes, so

$$C(\bar{u}) := \Big\{v \in \mathcal{T}(\bar{u})\colon F'(\bar{u})v + \beta_1\mathsf{j}'(\bar{u};v) = 0\Big\}.$$

Due to Lipschitz continuity of j, it is straightforward to show that $C(\bar{u})$ is a closed convex cone in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$.

A formal computation shows that the second derivatives of j in u in directions (v,v) should be given by

$$\mathsf{j}''(u;v^2) := \int_{\big[\|u\|_{\mathrm{L}^2(\Omega)}\neq 0\big]} \|u(t)\|_{\mathrm{L}^2(\Omega)}^{-1} \left[\|v(t)\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(u(t),v(t)\big)_\Omega}{\|u(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t,$$

where we consider the whole expression as 0 if $u = 0$. Clearly, this expression is always nonnegative, but there may be directions $v \in \mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ for which $\mathsf{j}''(u;v^2) = \infty$. In this sense, $\mathsf{j}''(u;\cdot)$ should not be seen as a traditional derivative of j' at u. We still set $\ell_r''(u;v^2) := F''(u)v^2 + \beta_1\mathsf{j}''(u;v^2)$ which is, any way, a useful object, as the following second order necessary conditions shows. Its proof will occupy the rest of this subsection:

Theorem 4.12 (Second order necessary conditions)**.** *Assume that* $\omega \in \mathrm{L}^1(0,\mathsf{T})$. *Let* $\bar{u} \in \mathcal{U}_{ad}$ *be a locally optimal solution to* (ROCP) *and let* $\bar{\mu} \in \mathrm{L}^\infty(\mathcal{A}_+\cup\mathcal{I})$ *be the associated Lagrange multiplier. Then there holds*

$$\ell_r''(\bar{u};v^2) + \int_{\mathcal{A}_+} \bar{\mu}(t)\big\|\bar{u}(t)\big\|_{\mathrm{L}^2(\Omega)}^{-1} \left[\|v(t)\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t),v(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t \geq 0$$

for all $v \in C(\bar{u})$.

The expression in Theorem 4.12 corresponds to $\partial_u^2 L(\bar{u}, \bar{\mu}; v^2) \geq 0$ with the Lagrangian

$$L(u, \mu) := \ell_r(u) + \int_0^{\mathsf{T}} \mu(t)\big(\|u(t)\|_{\mathrm{L}^2(\Omega)} - \omega(t)\big)\,\mathrm{d}t,$$

where in the theorem we have already inserted $\bar{\mu} = 0$ a.e. on $\mathcal{I}$. Both $\mathsf{j}''(\bar{u}; v^2)$ and the explicit integral in the substitute for the second derivative of the Lagrange penalty term in Theorem 4.12 may be infinite. We emphasize once more that we do not require ω to be bounded away from zero. In the case $\bar{u} \equiv 0$, the condition in Theorem 4.12 collapses to $F''(0)v^2 \geq 0$ for all $v \in C(0)$.

Remark 4.13. It is also possible to obtain an analogous result for the Lagrangian with a quadratic penalty term

$$L_2(u, \mu) := \ell_r(u) + \int_0^{\mathsf{T}} \mu(t)\big(\|u(t)\|^2_{\mathrm{L}^2(\Omega)} - \omega(t)^2\big)\,\mathrm{d}t.$$

Then the necessary condition in Theorem 4.12 becomes

$$\ell_r''(\bar{u}_2; v^2) + \int_{\mathcal{A}_+} \bar{\mu}(t) \frac{\|v(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t \geq 0$$

with the same multiplier $\bar{\mu}$ as before. The integral in the foregoing expression may also be infinite. The proof works nearly exactly as the one for Theorem 4.12 presented below.

We next prepare for the proof of Theorem 4.12 with some auxiliary results. From [5, Prop. 4.1/Lem. 4.2] together with the pointwise first-order necessary condition (4.13) and the Lagrange gradient equation (4.15) we obtain the first lemma for critical directions:

Lemma 4.14. *For all $v \in C(\bar{u})$, there holds*

$$\mathsf{j}'(\bar{u}; v) = \int_0^{\mathsf{T}} \big(\bar{\lambda}(t), v(t)\big)_\Omega\,\mathrm{d}t$$

and thus

$$0 = \big(\bar{p}(t) + \beta_1\bar{\lambda}(t) + \beta_2\bar{u}(t), v(t)\big)_\Omega = -\bar{\mu}(t)\frac{\big(\bar{u}(t), v(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}} \tag{4.16}$$

for almost all $t \in \mathcal{A}_+ \cup \mathcal{I}$.

Equation (4.16) also shows that if $t \in \mathcal{A}_+$ and $(\bar{u}(t), v(t))_\Omega < 0$ for some $v \in C(\bar{u})$, then $\bar{\mu}(t) = 0$ follows.

In the proof of Theorem 4.12, we will need properties of $\ell_r'(\bar{u}; w)$ with directions w which are possibly not in the critical cone, but derived from some $v \in C(\bar{u})$. The next lemma gives the required results.

Lemma 4.15. *Let $v \in C(\bar{u})$ be given and let $w \in \mathrm{L}^r(0, \mathsf{T}; \mathrm{L}^2(\Omega))$ be another function such that for almost all $t \in [\|\bar{u}\|_{\mathrm{L}^2(\Omega)} = 0]$, $w(t)$ is either $v(t)$ or zero. Then*

$$F'(\bar{u})w + \beta_1\mathsf{j}'(\bar{u}; w) = -\int_{\mathcal{A}_+} \bar{\mu}(t)\frac{\big(\bar{u}(t), w(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t.$$

Further, there holds $\chi_M C(\bar{u}) \subseteq C(\bar{u})$ for any measurable set $M \subseteq (0, \mathsf{T})$.

Proof. Let $v \in C(\bar{u})$. Arguing as for [5, (4.12)], we obtain that

$$\bar{\lambda}(t) = \frac{v(t)}{\|v(t)\|_{L^2(\Omega)}} \quad \text{f.a.a. } t \in \big[\|\bar{u}\|_{L^2(\Omega)} = 0\big] \cap \big[\|v\|_{L^2(\Omega)} \neq 0\big]. \tag{4.17}$$

Now let $w(t)$ be either $v(t)$ or zero for almost all $t \in [\|\bar{u}\|_{L^2(\Omega)} = 0]$. Then

$$\big(v(t), w(t)\big)_\Omega = \|v(t)\|_{L^2(\Omega)} \|w(t)\|_{L^2(\Omega)} \quad \text{f.a.a. } t \in \big[\|\bar{u}\|_{L^2(\Omega)} = 0\big].$$

Using this together with (4.17), we find

$$\int_{[\|\bar{u}\|_{L^2(\Omega)}=0]} \big(\bar{\lambda}(t), w(t)\big)_\Omega \,\mathrm{d}t = \int_{[\|\bar{u}\|_{L^2(\Omega)}=0]} \|w(t)\|_{L^2(\Omega)} \,\mathrm{d}t$$

and thus, with (4.14) and (4.15),

$$\begin{aligned} F'(\bar{u})w + \beta_1 \mathrm{j}'(\bar{u}; w) &= \int_0^{\mathsf{T}} \big(\bar{p}(t) + \beta_2 \bar{u}(t), w(t)\big)_\Omega \,\mathrm{d}t + \beta_1 \int_{[\|\bar{u}\|_{L^2(\Omega)}=0]} \|w(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\ &\quad + \beta_1 \int_{[\|\bar{u}\|_{L^2(\Omega)}\neq 0]} \big(\bar{\lambda}(t), w(t)\big)_\Omega \,\mathrm{d}t \\ &= \int_{\mathcal{A}_+} \big(\bar{p}(t) + \beta_1 \bar{\lambda}(t) + \beta_2 \bar{u}(t), w(t)\big)_\Omega \,\mathrm{d}t \\ &= -\int_{\mathcal{A}_+} \bar{\mu}(t) \frac{\big(\bar{u}(t), w(t)\big)_\Omega}{\|\bar{u}(t)\|_{L^2(\Omega)}} \,\mathrm{d}t. \end{aligned} \tag{4.18}$$

(See Proposition 4.4 for the derivative formula for $\mathrm{j}'(\bar{u}; w)$.) This was the first claim.

Let now $w = \chi_M v$ for some measurable set $M \subseteq (0, \mathsf{T})$. Then (4.18) holds true. Moreover, $\bar{\mu}(t)\|\bar{u}(t)\|^{-1}_{L^2(\Omega)}(\bar{u}(t), w(t))_\Omega = 0$ for almost all $t \in \mathcal{A}_+$ by (4.17) which again in (4.18) shows that $F'(\bar{u})w + \beta_1 \mathrm{j}'(\bar{u}; w) = 0$, so $w \in C(\bar{u})$. □

We further want to use second-order Taylor approximations for j. These are not immediate since we have already seen that the substitute for the second order derivative $\mathrm{j}''(\bar{u}; v^2)$ may be infinite for some directions v.

Consider $\Upsilon_2(f) := \|f\|_{L^2(\Omega)}$. We have

$$\begin{aligned} \Upsilon_2'(f)h &= \|f\|^{-1}_{L^2(\Omega)} \big(f, h\big)_\Omega, \\ \Upsilon_2''(f)h^2 &= \|f\|^{-1}_{L^2(\Omega)} \|h\|^2_{L^2(\Omega)} - \|f\|^{-3}_{L^2(\Omega)} \big(f, h\big)^2_\Omega \end{aligned}$$

for $f, h \in L^2(\Omega)$ with $f \neq 0$, cf. the appendix, and we will also need

$$\Upsilon_2'''(f)h^3 = 3\,\|f\|^{-3}_{L^2(\Omega)} \Big[\|f\|^{-2}_{L^2(\Omega)} \big(f, h\big)^3_\Omega - \|h\|^2_{L^2(\Omega)} \big(f, h\big)_\Omega\Big]$$

now; this is obtained by the chain rule since Υ_2'' is composed of continuously differentiable functions away from zero.

Lemma 4.16. *Let $M \subseteq (0, \mathsf{T})$ be a measurable set.*

1. *Let $f, h \in \mathrm{L}^1(M; \mathrm{L}^2(\Omega))$. Then*

$$\int_M \Big(\Upsilon_2\big(f(t)+h(t)\big) - \Upsilon_2\big(f(t)\big)\Big) \,\mathrm{d}t \geq \int_M \Upsilon_2'(f(t))h(t)\,\mathrm{d}t.$$

2. *Let moreover $\eta \in \mathrm{L}^\infty(M)$. Suppose that there is a number $\alpha > 0$ such that $\|f(t)\|_{\mathrm{L}^2(\Omega)} \geq \alpha$ for almost all $t \in M$. Then $h \mapsto \int_M \eta(t))\Upsilon_2''(f(t))h(t)^2\,\mathrm{d}t$ defines a continuous quadratic form on $\mathrm{L}^2(M; \mathrm{L}^2(\Omega))$. If $\eta \geq 0$ a.e. on M, then the quadratic form is convex.*

3. *Consider further $h \in \mathrm{L}^3(M; \mathrm{L}^2(\Omega))$. If for all functions $\theta \colon M \to [0,1]$ there is $\alpha_\theta > 0$ such that $\|f(t) + \theta(t)h(t)\|_{\mathrm{L}^2(\Omega)} \geq \alpha_\theta$ for almost all $t \in M$, then we have the Taylor expansion*

$$\begin{aligned}\int_M &\eta(t)\Big(\Upsilon_2\big(f(t)+h(t)\big) - \Upsilon_2\big(f(t)\big)\Big)\,\mathrm{d}t \\ &= \int_M \eta(t)\Big(\Upsilon_2'\big(f(t)\big)h(t) + \frac{1}{2}\Upsilon_2''\big(f(t)\big)h(t)^2\Big)\,\mathrm{d}t + \mathcal{O}\big(\|h\|^3_{\mathrm{L}^3(M;\mathrm{L}^2(\Omega))}\big). \quad (4.19)\end{aligned}$$

Proof. 1. We consider the Taylor expansion for $\Upsilon_2(f(t)+h(t))$ for almost every $t \in M$, with a function $\vartheta \colon M \to [0,1]$:

$$\Upsilon_2\big(f(t)+h(t)\big) - \Upsilon_2\big(f(t)\big) = \Upsilon_2'\big(f(t)\big)h(t) + \frac{1}{2}\Upsilon_2''\big(f(t)+\vartheta(t)h(t)\big)h(t)^2.$$

Since $\Upsilon_2''(g)w^2 \geq 0$ for all $g, w \in \mathrm{L}^2(\Omega)$, the claim follows from inserting this in the foregoing inequality and integrating over M. The integrals are finite due to $f, h \in \mathrm{L}^1(M; \mathrm{L}^2(\Omega))$.

2. Under the assumptions on f, we find

$$\int_M \big|\eta(t)\Upsilon_2''(f(t))h(t)^2\big|\,\mathrm{d}t \leq 2\alpha^{-1}\|\eta\|_{\mathrm{L}^\infty(M)} \int_M \big\|h(t)\big\|^2_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t.$$

This implies the continuity assertion. Moreover, a quadratic form is convex if and only if it is nonnegative, and the latter is ensured by $\eta \geq 0$ a.e. on M.

3. For the Taylor expansion for the integrated Υ_2, we again have from Taylor expansion for $\Upsilon_2(f(t)+h(t))$ for almost every $t \in M$, with a function $\vartheta \colon M \to [0,1]$:

$$\begin{aligned}\Upsilon_2&\big(f(t)+h(t)\big) - \Upsilon_2\big(f(t)\big) \\ &= \Upsilon_2'\big(f(t)\big)h(t) + \frac{1}{2}\Upsilon_2''\big(f(t)\big)h(t)^2 + \frac{1}{6}\Upsilon_2'''\big(f(t)+\vartheta(t)h(t)\big)h(t)^3.\end{aligned}$$

If $\big\|f(t)+\vartheta(t)h(t)\big\|_{\mathrm{L}^2(\Omega)} \geq \alpha_\vartheta > 0$ for almost all $t \in M$, then

$$\begin{aligned}\int_M &\big|\eta(t)\Upsilon_2'''\big(f(t)+\vartheta(t)h(t)\big)h(t)^3\big|\,\mathrm{d}t \\ &\leq 6\alpha_\vartheta^{-2}\|\eta\|_{\mathrm{L}^\infty(M)} \int_M \big\|h(t)\big\|^3_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t \in \mathcal{O}\big(\|h\|^3_{\mathrm{L}^3(M;\mathrm{L}^2(\Omega))}\big).\end{aligned}$$

The claim thus follows from multiplying the Taylor expansion for $\Upsilon_2(f(t)+h(t))$ by $\eta(t)$ and integrating over M. □

We next give the proof of Theorem 4.12. The principal idea is to approximate the critical direction $v \in C(\bar{u})$ in multiple stages.

Proof of Theorem 4.12. Let $v \in C(\bar{u})$. The proof is achieved as follows: We first suppose that $v \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ and that

$$\int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} \frac{\|v(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}} < \infty. \tag{4.20}$$

Since a multiple of this integral is an upper bound for $\mathrm{j}''(\bar{u}; v^2)$, (4.20) implies that $\mathrm{j}''(\bar{u}; v^2)$ is finite. We then construct two-staged approximations $u_{\rho,k}$ of $\bar{u}$ such that $u_{\rho,k} \to \bar{u}$ uniformly as $\rho \searrow 0$, as well as $u_{\rho,k} \in \mathcal{U}_{\mathrm{ad}}$ for $\rho > 0$ small enough and k fixed. Another property we need later is that $\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} = \omega(t)$ for $t \in \mathcal{A}_+$ with $(\bar{u}(t), v(t))_\Omega = 0$. Since such constructed $u_{\rho,k}$ is feasible and close to $\bar{u}$ for $\rho > 0$ sufficiently small, we then make the ansatz $0 \leq L(u_{\rho,k}, \bar{\mu}) - L(\bar{u}, \bar{\mu})$ and pass to the limit in ρ and k in the second order Taylor expansions there which gives the claim. For this, we need and establish that the derivative $v_k = \lim_{\rho\searrow 0} \rho^{-1}(u_{\rho,k} - \bar{u})$ exists in $\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ and satisfies $v_k \to v$ in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ as $k \to \infty$. Finally, we remove the assumptions on v above.

Step 1: *Construction of $u_{\rho,k}$.* Let α_k, ω_k be arbitrary positive sequences converging monotonically to zero. We define

$$N_k := \Big\{ t \in (0,\mathsf{T})\colon 0 < \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} < \alpha_k \ \text{ or } \ (1-\alpha_k)\omega(t) < \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} < \omega(t) \\ \text{or } \ 0 \leq \omega(t) < \omega_k \ \text{ or } \ \omega(t) > \omega_k^{-1} \Big\}.$$

Note that $\mathcal{A}_0 \subset N_k$ and $|N_k \setminus \mathcal{A}_0| \to 0$ as $k \to \infty$. For $t \notin N_k$, we moreover have $\omega_k \leq \omega(t) \leq \omega_k^{-1}$. Set

$$u_{\rho,k}(t) = \begin{cases} \bar{u}(t) & \text{if } t \in N_k, \\ \big(1 - \omega_t^k(\rho)\big)\bar{u}(t) + \rho v(t) & \text{if } t \in \mathcal{A}_+ \cap N_k^c \text{ and } (u(t), v(t))_\Omega = 0, \\ (1 - \rho\alpha_k)\bar{u}(t) + \rho v(t) & \text{if } t \in \mathcal{A}_+ \cap N_k^c \text{ and } (u(t), v(t))_\Omega < 0, \\ \bar{u}(t) + \rho v(t) & \text{elsewhere,} \end{cases}$$

with

$$\omega_t^k(\rho) := 1 - \sqrt{1 - \frac{\rho^2 \|v(t)\|^2_{\mathrm{L}^2(\Omega)}}{\omega(t)^2}} \quad \text{where } |\rho| < \tfrac{1}{2}\omega_k \|v\|^{-1}_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))}.$$

The function ω_t^k is chosen exactly such that $\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} = \omega(t)$ for $t \in \mathcal{A}_+ \cap N_k^c$ with $(\bar{u}(t), v(t))_\Omega = 0$.

We have $u_{\rho,k} \in \mathcal{U}_{\mathrm{ad}}$ for k fixed and ρ sufficiently small as we observe as follows:

- if $t \in N_k$, then $u_{\rho,k}(t) = \bar{u}(t)$ which is feasible,
- if $t \notin N_k$ and $\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} \leq (1-\alpha_k)\omega(t)$, then for $\rho \leq \alpha_k\omega_k\|v\|^{-1}_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))}$:

$$\big\|u_{\rho,k}(t)\big\|_{\mathrm{L}^2(\Omega)} = \big\|\bar{u}(t) + \rho v(t)\big\|_{\mathrm{L}^2(\Omega)} \leq (1-\alpha_k)\omega(t) + \rho\|v(t)\|_{\mathrm{L}^2(\Omega)} \leq \omega(t),$$

- if $t \in \mathcal{A}_+ \cap N_k^c$ and $(u(t), v(t))_\Omega = 0$, then $\|u_{\rho,k}(t)\|_{L^2(\Omega)} = \omega(t)$,
- if $t \in \mathcal{A}_+ \cap N_k^c$ and $(u(t), v(t))_\Omega < 0$, then

$$\big\|u_{\rho,k}(t)\big\|_{L^2(\Omega)}^2 \le \omega(t)^2 \quad \iff \quad \big(1 - \rho\alpha_k\big)^2 \omega(t)^2 + \rho^2 \|v(t)\|_{L^2(\Omega)}^2 \le \omega(t)^2$$
$$\iff \quad \big(-2\alpha_k + \rho\alpha_k^2\big)\omega(t)^2 + \rho\|v(t)\|_2^2 \le 0$$

and the latter is satisfied uniformly in t for this case if

$$\rho \le \frac{2\alpha_k \omega_k}{\alpha_k^2 \omega_k^{-2} + \|v\|_{L^\infty(0,T;L^2(\Omega))}}.$$

Step 2: *Limits as $\rho \searrow 0$ and $k \to \infty$.* It is clear that $u_{\rho,k}(t) \to \bar{u}(t)$ in $L^2(\Omega)$ as $\rho \searrow 0$ for almost every $t \in (0, \mathsf{T})$. We show that this convergence is in fact uniform. First, note that due to $\|\bar{u}(t)\|_{L^2(\Omega)} = \omega(t) \le \omega_k^{-1}$ on $\mathcal{A}_+ \cap N_k^c$, we have $\bar{u} \in L^\infty(\mathcal{A}_+ \cap N_k^c; L^2(\Omega))$. It follows that $w_{\rho,k} := u_{\rho,k} - \bar{u} \in L^\infty(0, \mathsf{T}; L^2(\Omega))$. For $t \in \mathcal{A}_+ \cap N_k^c$, we have

$$0 \le \omega_t^k(\rho) \le 1 - \sqrt{1 - \frac{\rho^2 \|v\|_{L^\infty(0,\mathsf{T};L^2(\Omega))}^2}{\omega_k^2}} \le \frac{\rho^2 \|v\|_{L^\infty(0,\mathsf{T};L^2(\Omega))}^2}{\omega_k^2}, \tag{4.21}$$

so $\omega_t^k(\rho)$ tends to zero uniformly in $t \in \mathcal{A}_+ \cap N_k^c$ as $\rho \searrow 0$ and hence

$$\big\|w_{\rho,k}\big\|_{L^\infty(0,\mathsf{T};L^2(\Omega))} = \big\|u_{\rho,k} - \bar{u}\big\|_{L^\infty(0,\mathsf{T};L^2(\Omega))} \in \mathcal{O}(\rho). \tag{4.22}$$

Next, set

$$v_k(t) := \begin{cases} 0 & \text{if } t \in N_k, \\ v(t) - \alpha_k \bar{u}(t) & \text{if } t \in \mathcal{A}_+ \cap N_k^c \text{ and } (\bar{u}(t), v(t))_\Omega < 0, \\ v(t) & \text{elsewhere.} \end{cases}$$

Then

$$v_k \to v \quad \text{in } L^r\big(0, \mathsf{T}; L^2(\Omega)\big) \quad \text{as } k \to \infty. \tag{4.23}$$

Further, $\omega_t^k(\cdot)$ is continuously differentiable with $(\omega_t^k)'(0) = 0 = \omega_t^k(0)$. Thus

$$v_k(t) = \lim_{\rho \searrow 0} \rho^{-1} w_{\rho,k}(t) \quad \text{for almost every } t \in (0, \mathsf{T}).$$

We again show that this convergence is uniform. In fact, for $t \in \mathcal{A}_+ \cap N_k^c$, we have $\omega_t^k(\rho) = \omega_t^k(0) + (\omega_t^k)'(0)\rho + \omega_t^k(\rho)$ and, via (4.21),

$$0 \le \frac{\omega_t^k(\rho) - \omega_t^k(0)}{\rho} = \frac{\omega_t^k(\rho)}{\rho} \le \frac{\rho \|v\|_{L^\infty(0,\mathsf{T};L^2(\Omega))}}{\omega_k^2}.$$

Hence the difference quotients for $(\omega_t^k)'(0) = \lim_{\rho \searrow 0} \rho^{-1}(\omega_t^k(\rho) - \omega_t^k(0))$ converge uniformly in $t \in \mathcal{A}_+ \cap N_k^c$. We obtain that

$$v_k = \lim_{\rho \searrow 0} \rho^{-1} w_{\rho,k} \quad \text{in } L^\infty(0, \mathsf{T}; L^2(\Omega)). \tag{4.24}$$

Step 3: *Core of the proof.* We first check that, for $t \in \mathcal{A}_+$,

$$\bar{\mu}(t)\big(\|u_{\rho,k}(t)\|_{L^2(\Omega)} - \|\bar{u}(t)\|_{L^2(\Omega)}\big) = 0.$$

This is true because of the following:

- for $t \in \mathcal{A}_+ \cap N_k$, we have $u_{\rho,k}(t) = \bar{u}(t)$,
- for $t \in \mathcal{A}_+ \cap N_k^c$ with $(u(t), v(t))_\Omega = 0$, we have $\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} = \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}$ by construction,
- and for $t \in \mathcal{A}_+ \cap N_k^c$ with $(u(t), v(t))_\Omega < 0$ we had already seen that $\bar{\mu}(t) = 0$ follows from (4.16).

For ρ small enough and fixed k, we have $u_{\rho,k} \in \mathcal{U}_{\mathrm{ad}}$ and $\ell_r(\bar{u}) \le \ell_r(u_{\rho,k})$ due to local optimality of $\bar{u}$ and (4.22). We thus make the ansatz

$$0 \le \ell_r(u_{\rho,k}) - \ell_r(\bar{u}) + \int_{\mathcal{A}_+} \bar{\mu}(t)\big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t. \tag{4.25}$$

We want to employ Taylor expansions for $\ell_r = F + \beta_1 \mathsf{j}$ and the multiplier term. The direction will be $w_{\rho,k} := u_{\rho,k} - \bar{u}$. For F, this is easily done since F is twice continuously differentiable on $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ by Lemma 4.6, but the nonsmooth terms require some justification in order to use Lemma 4.16. For both the reference point is $\bar{u}$. Consider

$$\begin{aligned}\mathsf{j}(u_{\rho,k}) - \mathsf{j}(\bar{u}) = \int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}=0]} &\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t \\ &+ \int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} \big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t. \end{aligned}\tag{4.26}$$

We focus on the second integral. By construction, $u_{\rho,k}(t) = \bar{u}(t)$ if $t \in [0 < \|\bar{u}\|_{\mathrm{L}^2(\Omega)} < \alpha_k]$. Hence

$$\begin{aligned}\int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} &\big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t \\ &= \int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\ge \alpha_k]} \big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t.\end{aligned}$$

Let $t \in [\|\bar{u}\|_{\mathrm{L}^2(\Omega)} \ge \alpha_k]$ and let $\theta_{\rho,k}(t) \in [0,1]$. Then, for ρ sufficiently small we estimate

$$\begin{aligned}\big\|u_{\rho,k}(t) + \theta_{\rho,k}(t) w_{\rho,k}(t)\big\|_{\mathrm{L}^2(\Omega)} &\ge \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)} - \theta_{\rho,k}(t)\|w_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} \\ &\ge \alpha_k - \|w_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} \ge \frac{\alpha_k}{2},\end{aligned}$$

since we had $\|w_{\rho,k}\|_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))} \in \mathcal{O}(\rho)$, cf. (4.22). Thus the prerequisites for the Taylor expansion in Lemma 4.16 are satisfied and we obtain, using (4.22),

$$\begin{aligned}\int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} \big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t &= \int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} \frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t \\ + \frac{1}{2}\int_{[\|\bar{u}\|_{\mathrm{L}^2(\Omega)}\neq 0]} \|\bar{u}(t)\|^{-1}_{\mathrm{L}^2(\Omega)} &\left[\big\|w_{\rho,k}(t)\big\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{\big(\bar{u}(t)\, w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t + \mathcal{O}\big(\rho^3\big).\end{aligned}$$

Re-inserting into (4.26), we finally get

$$\mathsf{j}(u_{\rho,k}) - \mathsf{j}(\bar{u}) = \mathsf{j}'(\bar{u}; w_{\rho,k}) + \mathsf{j}''(\bar{u}; w_{\rho,k}^2) + \mathcal{O}(\rho^3).$$

For the multiplier term, we argue analogously (thereby using $\mu \in \mathrm{L}^\infty(\mathcal{A}_+)$ for Lemma 4.16) to show that

$$\int_{\mathcal{A}_+} \bar{\mu}(t)\big(\|u_{\rho,k}(t)\|_{\mathrm{L}^2(\Omega)} - \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}\big)\,\mathrm{d}t = \int_{\mathcal{A}_+} \bar{\mu}(t)\frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t$$
$$+ \frac{1}{2}\int_{\mathcal{A}_+} \bar{\mu}(t)\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^{-1}\left[\big\|w_{\rho,k}(t)\big\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t + \mathcal{O}\big(\rho^3\big).$$

We thus obtain from the ansatz (4.25), with some function $\vartheta_{\rho,k}\colon (0,\mathsf{T}) \to [0,1]$:

$$0 \le F'(\bar{u})w_{\rho,k} + \beta_1 \mathsf{j}'(\bar{u}; w_{\rho,k}) + \int_{\mathcal{A}_+} \bar{\mu}(t)\frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t$$
$$+ \frac{1}{2}\Big(F''\big(\bar{u} + \vartheta_{\rho,k}w_{\rho,k}\big)w_{\rho,k}^2 + \beta_1\mathsf{j}''(\bar{u}; w_{\rho,k}^2)\Big)$$
$$+ \frac{1}{2}\int_{\mathcal{A}_+} \bar{\mu}(t)\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^{-1}\left[\big\|w_{\rho,k}(t)\big\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t + \mathcal{O}\big(\rho^3\big).$$

But for $t \in [\|\bar{u}\|_{\mathrm{L}^2(\Omega)} = 0]$, $w_{\rho,k}(t)$ equals either $v(t)$ or 0, hence we have seen in (4.18) that

$$F'(\bar{u})(w_{\rho,k}) + \beta_1\mathsf{j}'(\bar{u}; w_{\rho,k}) = -\int_{\mathcal{A}_+} \bar{\mu}(t)\frac{\big(\bar{u}(t), w_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t.$$

Inserting in the foregoing inequality and dividing by $\rho^2/2$, we find

$$0 \le F''\big(\bar{u} + \vartheta_{\rho,k}w_{\rho,k}\big)v_{\rho,k}^2 + \beta_1\mathsf{j}''(\bar{u}; v_{\rho,k}^2)$$
$$+ \int_{\mathcal{A}_+} \bar{\mu}(t)\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^{-1}\left[\big\|v_{\rho,k}(t)\big\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t), v_{\rho,k}(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t + \mathcal{O}\big(\rho\big),$$

where we have set $v_{\rho,k}(t) := \rho^{-1}w_{\rho,k}(t)$.

We let $\rho \searrow 0$. Recall that $\lim_{\rho\searrow 0} w_{\rho,k} = 0$ and $\lim_{\rho\searrow 0} v_{\rho,k} = v_k$, both in $\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ by (4.24) and (4.23). It is thus immediate from Lemma 4.6 that

$$F''(\bar{u} + \vartheta_{\rho,k}w_{\rho,k})v_{\rho,k}^2 \to F''(\bar{u})v_k^2 \quad \text{as } \rho \searrow 0.$$

For the two other terms, we use again that if $t \in [0 < \|\bar{u}\|_{\mathrm{L}^2(\Omega)} < \alpha_k]$, then by construction $v_{\rho,k}(t) = 0 = v_k(t)$. Hence it suffices to consider the integrals on $[\|\bar{u}\|_{\mathrm{L}^2(\Omega)} \ge \alpha_k]$. Lemma 4.16 (2) shows that the substitutes for the second derivative induce continuous quadratic forms on $\mathrm{L}^2([\|\bar{u}\|_{\mathrm{L}^2(\Omega)} \ge \alpha_k];\mathrm{L}^2(\Omega))$ and $\mathrm{L}^2([\|\bar{u}\|_{\mathrm{L}^2(\Omega)} \ge \alpha_k] \cap \mathcal{A}_+;\mathrm{L}^2(\Omega))$, respectively. Using $\lim_{\rho\searrow 0} v_{\rho,k} = v_k$, we thus obtain

$$0 \le F''(\bar{u})v_k^2 + \beta_1\mathsf{j}''(\bar{u}; v_k^2)$$
$$+ \int_{\mathcal{A}_+} \bar{\mu}(t)\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^{-1}\left[\|v_k(t)\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t), v_k(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t. \quad (4.27)$$

We next pass to the limit in (4.27) as $k \to \infty$, so $v_k \to v$. Now the preliminary assumption from (4.20) becomes important. Since $v_k \to v$ in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$, we know that the integrand $\xi_k(t)$ in $\mathsf{j}''(\bar u; v_k^2) = \int_0^{\mathsf{T}} \xi_k(t)$ converges to the integrand in $\mathsf{j}''(\bar u; v^2)$ pointwise almost everywhere on $(0,\mathsf{T})$. It is moreover nonnegative and bounded by

$$\begin{aligned}
\xi_k(t) &\le \chi_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\neq 0]}(t)\frac{\|v_k(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}} \\
&\le \chi_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\neq 0]}(t)\frac{\left(\|v(t)\|_{\mathrm{L}^2(\Omega)} + \chi_{[(\bar u(t),v(t))_\Omega<0]\,\cap\,\mathcal{A}}(t)\,\cdot\,\omega(t)\right)^2}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}} \\
&\le 2\chi_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\neq 0]}(t)\left(\frac{\|v(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}} + \omega(t)\right).
\end{aligned}$$

By (4.20), the right-hand side is integrable over $(0,\mathsf{T})$. Thus we can use the dominated convergence theorem to infer that $\mathsf{j}''(\bar u; v_k^2) \to \mathsf{j}''(\bar u; v^2)$ as $k \to \infty$. We again argue analogously for the multiplier term in (4.27). Since $F''(\bar u)$ is a continuous bilinear form on $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$, we also have $F''(\bar u)v_k^2 \to F''(\bar u)v^2$ as $k \to \infty$. Overall we obtain

$$\begin{aligned}
0 \le\ & F''(\bar u)v^2 + \beta_1 \mathsf{j}''(\bar u; v^2) \\
&+ \int_{\mathcal{A}_+} \bar\mu(t)\|\bar u(t)\|^{-1}_{\mathrm{L}^2(\Omega)}\left[\|v(t)\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{\big(\bar u(t), v(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right] \mathrm{d}t
\end{aligned}$$

for all $v \in C(\bar u) \cap \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ satisfying (4.20).

Step 4: *Removing the additional assumptions.* To finally remove the assumptions on $v \in C(\bar u)$, we do another approximation. Let $v \in C(\bar u)$ and let ν_ℓ be a positive sequence converging monotonically to zero. Define

$$N_\ell := \Big\{ t \in \big[\|\bar u\|_{\mathrm{L}^2(\Omega)} \neq 0\big] \colon \|v(t)\|_{\mathrm{L}^2(\Omega)} > \nu_\ell^{-1}\min(\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}, 1)\Big\}$$

and set $v_\ell := \chi_{N_\ell^c} v$. Then for every ℓ, we have $v_\ell \in \mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ and

$$\int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\neq 0]} \frac{\|v_\ell(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t = \int_{N_\ell^c} \frac{\|v(t)\|^2_{\mathrm{L}^2(\Omega)}}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t \le \nu_\ell^{-1}\int_{N_\ell^c}\|v(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t < \infty.$$

Moreover, by Lemma 4.15, $v_\ell \in C(\bar u)$, so by the above

$$\begin{aligned}
0 \le\ & F''(\bar u)v_\ell^2 + \beta_1 \mathsf{j}''(\bar u; v_\ell^2) \\
&+ \int_{\mathcal{A}_+} \bar\mu(t)\|\bar u(t)\|^{-1}_{\mathrm{L}^2(\Omega)}\left[\|v_\ell(t)\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{\big(\bar u(t), v_\ell(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right] \mathrm{d}t.
\end{aligned}$$

We have $v_\ell \to v$ in $\mathrm{L}^r(0,\mathsf{T};\mathrm{L}^2(\Omega))$ as $\ell \to \infty$ due to $|N_\ell| \to 0$. So $F''(\bar u)v_\ell^2 \to F''(\bar u)v^2$ as $\ell \to \infty$. Set moreover

$$\xi_\ell(t) := \chi_{N_\ell^c}(t)\,\cdot\,\|\bar u(t)\|^{-1}_{\mathrm{L}^2(\Omega)}\left[\|v_\ell(t)\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{\big(\bar u(t), v_\ell(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right] \mathrm{d}t.$$

Then $0 \leq \xi_\ell(t) \leq \xi_{\ell+1}(t)$ for every $t \in (0, \mathsf{T})$ due to $N_{\ell+1} \subseteq N_\ell$. Thus, the monotone convergence theorem yields

$$\lim_{\ell\to\infty} \mathsf{j}''(\bar{u}; v_\ell^2) = \lim_{\ell\to\infty} \int_0^{\mathsf{T}} \xi_\ell(t) = \int_0^{\mathsf{T}} \lim_{\ell\to\infty} \xi_\ell(t) = \mathsf{j}''(\bar{u}; v^2).$$

Again, we argue analogously for the multiplier term and finally obtain

$$0 \leq F''(\bar{u})v^2 + \beta_1 \mathsf{j}''(\bar{u}; v^2) \\ + \int_{\mathcal{A}_+} \bar{\mu}(t) \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^{-1} \left[\|v(t)\|_{\mathrm{L}^2(\Omega)}^2 - \left(\frac{\big(\bar{u}(t), v(t)\big)_\Omega}{\|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}} \right)^2 \right] \mathrm{d}t$$

for *all* $v \in C(\bar{u})$. This was the claim. ☐

4.2.3. Second order sufficient conditions

We finally prove no-gap second order sufficient conditions for (ROCP) for strong local solutions, i.e., in the $\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$-sense. The proof is fairly standard, but we again have to circumvent the possible singularity of the substitute for the second derivative $\mathsf{j}''(\bar{u}; v^2)$. We assume $\beta_2 > 0$ to enforce coercivity of the problem. The case $\beta_2 = 0$ is an open problem.

Theorem 4.17 (Second order sufficient conditions)**.** *Let $\beta_2 > 0$. Assume that $\bar{u} \in \mathcal{U}_{ad}$ satisfies*

$$\ell_r''(\bar{u}; v^2) > 0 \quad \textit{for all } v \in C(\bar{u}) \setminus \{0\}. \tag{4.28}$$

Then $\bar{u}$ is a strong local minimum, that is, there are $\varepsilon, \delta > 0$ such that

$$\ell_r(\bar{u}) + \frac{\delta}{2}\big\|u - \bar{u}\big\|_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}^2 \leq \ell_r(u) \\ \textit{for all } u \in \mathcal{U}_{ad} \textit{ with } \big\|u - \bar{u}\big\|_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))} < \varepsilon. \tag{4.29}$$

Proof. Suppose not. Then there are positive nonincreasing sequences $\alpha_k, \eta_k \searrow 0$ and feasible controls $(u_k) \subset \mathcal{U}_{\mathrm{ad}}$ such that

$$\big\|u_k - \bar{u}\big\|_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))} \leq \alpha_k \quad \text{and} \quad \ell_r(u_k) < \ell_r(\bar{u}) + \frac{\eta_k}{2}\big\|u_k - \bar{u}\big\|_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}^2. \tag{4.30}$$

We set $\rho_k := \|u_k - \bar{u}\|_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))} > 0$ and $v_k := \rho_k^{-1}\big(u_k - \bar{u}\big)$. Then clearly $\rho_k \searrow 0$. Since v_k is normalized in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$, we have $v_k \rightharpoonup v \in \mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$, possibly after going over to a subsequence. The rest of the proof will consist of showing that $v = 0$ which then will lead to a contradiction with $\beta_2 > 0$. In order to show $v = 0$, we establish that $v \in C(\bar{u})$ and $\ell_r''(\bar{u}; v^2) \leq 0$ and then conclude from (4.28).

Step 1: $v \in C(\bar{u})$. We first show that $v \in \mathcal{T}(\bar{u})$. Since $u_k \in \mathcal{U}_{\mathrm{ad}}$ for each k, we have $\|u_k(t)\|_{\mathrm{L}^2(\Omega)}^2 \leq \omega(t)^2 = \|\bar{u}(t)\|_{\mathrm{L}^2(\Omega)}^2$ for almost all $t \in \mathcal{A}$. Hence $(u_k(t) - \bar{u}(t), \bar{u}(t))_\Omega \leq 0$ for $t \in \mathcal{A}_+$ and $u_k(t) = \bar{u}(t) = 0$ for $t \in \mathcal{A}_0$. This implies $v_k \in \mathcal{T}(\bar{u})$ and thus $v \in \mathcal{T}(\bar{u})$, since $\mathcal{T}(\bar{u})$ is weakly closed in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$.

It remains to show that $F'(\bar u)v+\beta_1 \mathsf{j}'(\bar u;v)=0$. We argue as in [5, Proof of Thm. 5.12]. The function j is Lipschitz continuous and convex, thus

$$\mathsf{j}'(\bar u;v) \le \liminf_{k\to\infty} \frac{\mathsf{j}(u_k)-\mathsf{j}(\bar u)}{\rho_k}.$$

This inequality and (4.30) then show that

$$\begin{aligned} F'(\bar u)v + \beta_1 \mathsf{j}'(\bar u;v) &\le \liminf_{k\to\infty} \frac{\ell_r(u_k)-\ell_r(\bar u)}{\rho_k} \\ &\le \liminf_{k\to\infty} \frac{\eta_k}{2\rho_k}\big\|u_k-\bar u\big\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))} = \liminf_{k\to\infty}\frac{\rho_k\eta_k}{2} = 0. \end{aligned}$$

From the reverse inequality from the first order necessary optimality condition (4.12) we infer $v\in C(\bar u)$.

Step 2: $\ell_r''(\bar u;v^2)\le 0$: We again define approximations to v and v_k to cope with the possible unboundedness of $\mathsf{j}''(\bar u;v^2)$. Let $\kappa_\ell>0$ be a nonincreasing sequence converging to zero. Set

$$v_{k,\ell}(t) := \begin{cases} 0 & \text{if } t\in \big[0<\|\bar u\|_{\mathrm{L}^2(\Omega)}<\kappa_\ell\big], \\ v_k(t) & \text{elsewhere} \end{cases}$$

and v_ℓ analogously. Then $v_{k,\ell}\rightharpoonup v_\ell$ for ℓ fixed and $k\to\infty$, and $v_\ell\to v$ if $\ell\to\infty$, both in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$. We find for almost all $t\in(0,\mathsf{T})$

$$\big\|\bar u(t)+\rho_k v_{k,\ell}(t)\big\|_{\mathrm{L}^2(\Omega)} \ge \kappa_\ell-\alpha_k \ge \frac{\kappa_\ell}{2},$$

for ℓ fixed and k large enough, since from (4.30)

$$\rho_k\|v_{k,\ell}(t)\|_{\mathrm{L}^2(\Omega)} \le \rho_k\|v_k(t)\|_{\mathrm{L}^2(\Omega)} \le \alpha_k.$$

Hence we have integrated Taylor expansions of $\|\bar u+\rho_k v_{k,\ell}\|_{\mathrm{L}^2(\Omega)}$ from Lemma 4.16 (3) at hand. Further, from Lemma 4.16 (1),

$$\begin{aligned} \int_{[0<\|\bar u\|_{\mathrm{L}^2(\Omega)}<\kappa_\ell]} \Big(\big\|\bar u(t)+\rho_k v_k(t)\big\|_{\mathrm{L}^2(\Omega)} - \big\|\bar u(t)\big\|_{\mathrm{L}^2(\Omega)}\Big)\,\mathrm{d}t& \\ \ge \rho_k \int_{[0<\|\bar u\|_{\mathrm{L}^2(\Omega)}<\kappa_\ell]} \frac{\big(\bar u(t),v_k(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t.& \end{aligned} \tag{4.31}$$

Using (4.30) and the definition of $v_{k,\ell}$, we expand

$$\begin{aligned} \frac{\eta_k\rho_k^2}{2} &> \ell_r(u_k)-\ell_r(\bar u) \\ &= F(u_k)-F(\bar u)+\beta_1\big(\mathsf{j}(u_k)-\mathsf{j}(\bar u)\big) \\ &= F(u_k)-F(\bar u)+\beta_1\rho_k\int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}=0]} \|v_k(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t \\ &\quad + \beta_1\int_{[0<\|\bar u\|_{\mathrm{L}^2(\Omega)}<\kappa_\ell]}\Big(\big\|\bar u(t)+\rho_k v_k(t)\big\|_{\mathrm{L}^2(\Omega)}-\big\|\bar u(t)\big\|_{\mathrm{L}^2(\Omega)}\Big)\,\mathrm{d}t \\ &\quad + \beta_1\Big(\mathsf{j}(\bar u+\rho_k v_{k,\ell}) - \int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\ge\kappa_\ell]}\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t\Big). \end{aligned}$$

We insert Taylor expansions for F and $\mathsf{j}(\bar u+\rho_k v_{k,\ell})$ (cf. Lemma 4.16 (3)) as well as (4.31) to obtain, with a function $\vartheta_k\colon [0,\mathsf{T}] \to [0,1]$,

$$\begin{aligned}
\frac{\eta_k\rho_k}{2} &> \rho_k F'(\bar u)v_k + \frac{\rho_k^2}{2}F''\big(\bar u + \vartheta_k\rho_k v_k\big)v_k^2 + \beta_1\rho_k \int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}=0]} \|v_k(t)\|_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t \\
&\quad + \beta_1\rho_k\left(\int_{[0<\|\bar u\|_{\mathrm{L}^2(\Omega)}<\kappa_\ell]} \frac{\big(\bar u(t), v_k(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t + \int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\geq\kappa_\ell]} \frac{\big(\bar u(t), v_k(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\,\mathrm{d}t\right) \\
&\quad + \frac{\beta_1\rho_k^2}{2}\int_{[\|\bar u\|_{\mathrm{L}^2(\Omega)}\geq\kappa_\ell]} \|\bar u(t)\|^{-1}_{\mathrm{L}^2(\Omega)} \left[\big\|v_{k,\ell}(t)\big\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{\big(\bar u(t), v_{k,\ell}(t)\big)_\Omega}{\|\bar u(t)\|_{\mathrm{L}^2(\Omega)}}\right)^2\right]\mathrm{d}t \\
&\quad + \mathcal{O}\big(\rho_k^3\|v_k\|^3_{\mathrm{L}^3(0,\mathsf{T};\mathrm{L}^2(\Omega))}\big).
\end{aligned}$$

Note that $\rho_k\|v_k\|^3_{\mathrm{L}^3(0,\mathsf{T};\mathrm{L}^2(\Omega))} \leq \rho_k\|v_k\|_{\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))} \in \mathcal{O}(\alpha_k)$ by (4.30). Inserting this and rearranging, we arrive at

$$\begin{aligned}
\frac{\eta_k\rho_k^2}{2} &> \rho_k\big(F'(\bar u)v_k + \beta_1\mathsf{j}'(\bar u; v_k)\big) \\
&\qquad + \frac{\rho_k^2}{2}\left(F''\big(\bar u + \vartheta_k\rho_k v_k\big)v_k^2 + \beta_1\mathsf{j}''(\bar u; v_{k,\ell}^2)\right) + \mathcal{O}\left(\alpha_k\rho_k^2\right).
\end{aligned}$$

Since $v_k \in \mathcal{T}(\bar u)$, we have $F'(\bar u)v_k + \beta_1\mathsf{j}'(\bar u; v_k) \geq 0$ by first order optimality (4.12). It is thus practical to divide by $\rho_k^2/2$ and insert a zero to obtain

$$F''(\bar u)v_k^2 + \beta_1\mathsf{j}''(\bar u; v_{k,\ell}^2) \leq \big|F''(\bar u)v_k^2 - F''\big(\bar u + \vartheta_k\rho_k v_k\big)v_k^2\big| + \mathcal{O}(\alpha_k) + \eta_k. \tag{4.32}$$

We had $\rho_k v_k \to 0$ in $\mathrm{L}^\infty(0,\mathsf{T};\mathrm{L}^2(\Omega))$ as $k\to\infty$ by construction. Hence, since F is twice continuously differentiable by Lemma 4.6 and v_k is normalized in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$,

$$\big|F''(\bar u)v_k^2 - F''\big(\bar u + \vartheta_k\rho_k v_k\big)v_k^2\big| \to 0 \quad \text{as } k\to\infty.$$

It follows from Lemma 4.6 and Lemma 4.16 (2) that the second derivatives on the left-hand side in (4.32) are both weakly lower semicontinuous with respect to their directions. We thus obtain from (4.32)

$$F''(\bar u)v^2 + \beta_1\mathsf{j}''(\bar u; v_\ell^2) \leq \liminf_{k\to\infty}\big(F''(\bar u)v_k^2 + \beta_1\mathsf{j}''(\bar u; v_{k,\ell}^2)\big) = 0.$$

Finally, a monotone convergence argument as in the proof of Theorem 4.12 shows that

$$\ell_r''(\bar u; v^2) = F''(\bar u)v^2 + \beta_1\lim_{\ell\to\infty}\mathsf{j}''(\bar u; v_\ell^2) \leq 0$$

so that $v = 0$ by (4.28).

Step 3: Contradiction to $\beta_2 > 0$. From $v_k \rightharpoonup v = 0$, it follows that $z_{v_k} \rightharpoonup z_v = 0$ in $\mathcal{Y}$ due to Theorem 4.3. Hence, setting $G(u) \coloneqq F(u) - \frac{\beta_2}{2}\|u\|^2_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))}$ and using Lemma 4.6 for $\beta_2 = 0$, we find

$$0 = G''(\bar u)z_v^2 \leq \liminf_{k\to\infty} G''(\bar u)z_{v_k}^2.$$

Since v_k was normalized in $\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))$ and $\mathsf{j}''(\bar{u};v_{k,\ell}^2) \geq 0$ for every ℓ, we thus have

$$\begin{aligned}
\beta_2 &\leq \liminf_{k\to\infty} G''(\bar{u})z_{v_k}^2 + \beta_2 \\
&= \liminf_{k\to\infty} G''(\bar{u})z_{v_k}^2 + \beta_2 \liminf_{k\to\infty} \|v_k\|_{\mathrm{L}^2(0,\mathsf{T};\mathrm{L}^2(\Omega))} \\
&\leq \liminf_{k\to\infty} F''(\bar{u})v_k^2 \\
&\leq \liminf_{k\to\infty} F''(\bar{u})v_k^2 + \beta_1 \liminf_{k\to\infty} \mathsf{j}''(\bar{u};v_{k,\ell}^2) \\
&\leq \liminf_{k\to\infty} \left(F''(\bar{u})v_k^2 + \beta_1 \mathsf{j}''(\bar{u};v_{k,\ell}^2)\right) = 0,
\end{aligned}$$

where again (4.32) was used. This is the final contradiction and completes the proof. □

Appendix A. Auxiliary results

At several points we need the following growth lemma.

Lemma A.1 ([38, Ch. IV, Lem. 2.2])**.** *Let $C_0 > 0$ and suppose that $f \in \mathrm{C}([a,b))$ satisfies $f(s) \geq 0$ for all $s \in [a,b)$ and $f(a) = 0$ as well as*

$$f(s) \leq C_0 + \varepsilon f(s)^\sigma \quad \text{for all } s \in [a,b)$$

for some $\sigma > 0$. If $\varepsilon < 2^{-\sigma}C_0^{1-\sigma}$, then it follows that

$$f(s) \leq 2C_0 \quad \text{for all } s \in [a,b).$$

The remaining assertions are about the first and second derivatives of (powers of) norms on Banach spaces. We need the results for Lebesgue spaces for which the following result is a basic one.

Proposition A.2 ([28, Thms. 3.3&3.9])**.** *Let $\Upsilon \subset \mathbb{R}^d$ for some $d \in \mathbb{N}$ and let X be a Banach space. Then the norm on $\mathrm{L}^p(\Upsilon;X)$ for $2 < p < \infty$ is twice continuously differentiable away from 0 if and only if the norm on X is twice continuously differentiable away from 0 and the second derivative is uniformly bounded on the unit sphere in X. The norm on $\mathrm{L}^2(\Upsilon;X)$ is twice continuously differentiable away from 0 if and only if X is a Hilbert space.*

The requirements in Proposition A.2 are clearly satisfied in the case $X = \mathbb{R}$. Setting $\Upsilon_p(f) := \|f\|_{\mathrm{L}^p(\Omega)}$ for $2 \leq p < \infty$, we have the following derivatives in $f \in \mathrm{L}^p(\Omega) \setminus \{0\}$:

$$\begin{aligned}
\Upsilon_p'(f)h &= \|f\|_{\mathrm{L}^p(\Omega)}^{1-p}\big(|f|^{p-2}f,h\big)_\Omega, \\
\Upsilon_p''(f)h_1h_2 &= (p-1)\left(\|f\|_{\mathrm{L}^p(\Omega)}^{1-p}\big(|f|^{p-2}h_1,h_2\big)_\Omega - \|f\|_{\mathrm{L}^p(\Omega)}^{-1}\prod_{i=1}^{2}\big(\Upsilon_p'(f)h_i\big)\right)
\end{aligned}$$

The Hölder inequality shows that $|\Upsilon_p''(f)h_1h_2| \lesssim \|f\|_{\mathrm{L}^p(\Omega)}^{-1}\|h_1\|_{\mathrm{L}^p(\Omega)}\|h_2\|_{\mathrm{L}^p(\Omega)}$, so Υ_p'' is bounded on the unit sphere in $\mathrm{L}^p(\Omega)$. This allows to invoke Proposition A.2 again for

$\Phi_{p,q}(f) := \|f\|_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))}$ with $2 < p,q < \infty$ or $p = q = 2$, and a straight forward calculation with the chain rule gives the derivatives in $f \neq 0$ as follows:

$$\begin{aligned}
\Phi'_{p,q}(f)h &= \|f\|^{1-p}_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))} \int_0^{\mathsf{T}} \|f(t)\|^{p-1}_{\mathrm{L}^q(\Omega)} \Upsilon'_q(f(t))h(t)\,\mathrm{d}t,\\
\Phi''_{p,q}(f)h_1h_2 &= (p-1)\|f\|^{1-p}_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))} \int_0^{\mathsf{T}} \|f(t)\|^{p-2}_{\mathrm{L}^q(\Omega)} \prod_{i=1}^{2}\Big(\Upsilon'_q(f(t))h_i(t)\Big)\,\mathrm{d}t\\
&\quad - (p-1)\|f\|^{1-2p}_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))} \prod_{i=1}^{2} \int_0^{\mathsf{T}} \|f(t)\|^{p-1}_{\mathrm{L}^q(\Omega)} \Upsilon'_q(f(t))h_i(t)\,\mathrm{d}t\\
&\quad + \|f\|^{1-p}_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))} \int_0^{\mathsf{T}} \|f(t)\|^{p-1}_{\mathrm{L}^q(\Omega)} \Upsilon''_q(f(t))h_1(t)h_2(t)\,\mathrm{d}t.
\end{aligned}$$

We next consider powers of the norms $\Phi_{p,q}$. The general result is as follows. It shows that with a sufficiently large power one can overcome the nondifferentiability in 0 of $\Phi_{p,q}$.

Lemma A.3. *Let X be a Banach space and $\nu > 1$. Consider $f\colon X \to \mathbb{R}$ and set $g(x) := |f(x)|^{\nu-1}f(x)$.*

1. *Suppose that f is locally Lipschitz continuous in 0 and continuously differentiable on $X \setminus \{0\}$. Then g is continuously differentiable on X with $g'(0) = 0$.*
2. *Suppose in addition that $\nu > 2$ and that f is twice continuously differentiable on $X \setminus \{0\}$ with $\|f''(x)\|_{\mathcal{L}(X\times X;\mathbb{R})} \lesssim \|x\|_X^{-1}$ as $x \to 0$. Then g is twice continuously differentiable on X with $g''(0) = 0$.*

Proof. We can without loss of generality assume that $f(0) = 0$.

1. It is clear that g is continuously differentiable on $X\setminus\{0\}$ with $g'(x)h = \nu|f(x)|^{\nu-1}f'(x)h$. Moreover, let $R > 0$ be such that f is Lipschitz continuous on a ball around 0 with radius R. We denote the Lipschitz constant by L_R. Then we have for all h with $\|h\|_X \le R$

$$\big|g(h) - g(0) - 0\cdot h\big| = |f(h)|^\nu \le L_R^\nu \|h\|_X^\nu \in o\big(\|h\|_X\big),$$

so g is differentiable in 0 with derivative $g'(0) = 0$. Since $\|f'(x)\|_{X'} \le 1$ for all $X \setminus \{0\}$, we further find

$$|g'(x)h| \le \nu L_R^{\nu-1}\|x\|_X^{\nu-1}\|h\|_X$$

for all x with $\|x\|_X \le R$. This shows that $\|g'(x)\|_{X'} \to 0$ as $x \to 0$ in X, so g' is continuous in 0.

2. Now g is even twice continuously differentiable on $X \setminus \{0\}$ with

$$g''(x)h_1h_2 = \nu(\nu-1)|f(x)|^{\nu-2}\big(f'(x)h_1\big)\big(f'(x)h_2\big) + \nu|f(x)|^{\nu-1}f''(x)h_1h_2.$$

As above, we obtain, for all $h_2 \in X$ and h_1 with $\big\|h_1\big\|_X \le R$,

$$\big|g'(h_1)h_2 - g'(0)h_2 - 0\cdot h_1h_2\big| \le \nu L_R^{\nu-1}\big\|h_1\big\|_X^{\nu-1}\big\|h_2\big\|_X,$$

hence $\|g'(h_1) - g'(0) - 0\|_{X'} \in o\big(\|h_1\|_X\big)$ and g' is differentiable in 0 with $g''(0) = 0$. Here we have used that $\nu > 2$ now. The claim that g'' is continuous in 0 follows from the observation that

$$\big|g''(x)h_1h_2\big| \lesssim \|x\|^{\nu-2}\big\|h_1\big\|_X\big\|h_2\big\|_X$$

for all $h_1, h_2 \in X \setminus \{0\}$ as $x \to 0$ by the assumption on f'', and again $\nu > 2$. □

We set $\Psi_{p,q}(y) \coloneqq \frac{1}{p}\|y\|^p_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))} = \frac{1}{p}\Phi_{p,q}(y)^p$ and want to show that it is twice continuously differentiable for $p > 2$ using Lemma A.3 with $f = \Phi_{p,q}$ and $\nu = p$. The Lipschitz condition in Lemma A.3 is clearly satisfied since $\Phi_{p,q}$ is a norm, and we had already seen that Proposition A.2 implies the differentiability assumptions. Moreover, repeated use of Hölder's inequality in the foregoing formula for $\Phi''_{p,q}$ shows that

$$\big|\Phi''_{p,q}(f)h_1h_2\big| \lesssim \|f\|^{-1}_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))}\|h_1\|_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))}\|h_2\|_{\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))}.$$

Hence we can indeed apply Lemma A.3 with $f = \Phi_{p,q}$ and $\nu = p$ to obtain the following:

Corollary A.4. *For $p > 2$, the function $\Psi_{p,q}$ is twice continuously differentiable as a mapping from $\mathrm{L}^p(0,\mathsf{T};\mathrm{L}^q(\Omega))$ to $\mathbb{R}$. Its derivatives are given by*

$$\Psi'_{p,q}(y)h = \int_0^{\mathsf{T}} \|y(t)\|^{p-1}_{\mathrm{L}^q(\Omega)}\Upsilon'_q(y(t))h(t)\,\mathrm{d}t$$
$$= \int_0^{\mathsf{T}} \|y(t)\|^{p-q}_{\mathrm{L}^q(\Omega)}\big(|y(t)|^{q-2}y(t), h(t)\big)_\Omega\,\mathrm{d}t$$

and

$$\Psi''_{p,q}(y)h^2 = (p-q)\int_0^{\mathsf{T}} \|y(t)\|^{p-2}_{\mathrm{L}^q(\Omega)}\big(\Upsilon'_q(y(t))h(t)\big)^2\,\mathrm{d}t$$
$$+ (q-1)\int_0^{\mathsf{T}} \|y(t)\|^{p-q}_{\mathrm{L}^q(\Omega)}\big(|y(t)|^{q-2}, h^2(t)\big)_\Omega\,\mathrm{d}t.$$

Acknowledgements

The work of Karl Kunisch was partly supported by the ERC advanced grant 668998 (OCLOC) under the EU H2020 research program.

References

[1] W. Arendt, C. J. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Springer Science + Business Media, 2011.

[2] M. D. Blair, H. F. Smith, and C. D. Sogge, *Strichartz estimates for the wave equation on manifolds with boundary*, Annales de l'Institut Henri Poincaré (C) Non Linear Analysis, 26 (2009), pp. 1817–1829.

[3] N. Burq, G. Lebeau, and F. Planchon, *Global Existence for Energy Critical Waves in 3-D Domains*, Journal of the American Mathematical Society, 21 (2008), pp. 831–845.

[4] N. Burq and F. Planchon, *Global existence for energy critical waves in 3-d domains: Neumann boundary conditions*, American Journal of Mathematics, 131 (2009), pp. 1715–1742.

[5] E. Casas, R. Herzog, and G. Wachsmuth, *Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations*, ESAIM: Control, Optimisation and Calculus of Variations, 23 (2016), pp. 263–295.

[6] M. Cavalcanti, L. Fatori, and T. Ma, *Attractors for wave equations with degenerate memory*, Journal of Differential Equations, 260 (2016), pp. 56–83.

[7] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 5*, Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.

[8] B. Dehman and G. Lebeau, *Analysis of the HUM control operator and exact controllability for semilinear waves in uniform time*, SIAM Journal on Control and Optimization, 48 (2009), pp. 521–550.

[9] B. Dehman, G. Lebeau, and E. Zuazua, *Stabilization and control for the subcritical semilinear wave equation*, Annales Scientifiques de l'École Normale Supérieure, 36 (2003), pp. 525–551.

[10] J. Diestel, W. M. Ruess, and W. Schachermayer, *Weak compactness in $L^1(\mu, X)$*, Proceedings of the American Mathematical Society, 118 (1993), p. 447.

[11] R. Donninger, *Strichartz estimates in similarity coordinates and stable blowup for the critical wave equation*, Duke Mathematical Journal, 166 (2017), pp. 1627–1683.

[12] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958.

[13] T. Duyckaerts, C. Kenig, and F. Merle, *Concentration-compactness and universal profiles for the non-radial energy critical wave equation*, Nonlinear Analysis, 138 (2016), pp. 44–82.

[14] ———, *Global existence for solutions of the focusing wave equation with the compactness property*, Annales de l'Institut Henri Poincaré (C) Non Linear Analysis, 33 (2016), pp. 1675–1690.

[15] L. C. Evans, *Partial differential equations*, Graduate studies in mathematics, American Mathematical Society, Providence (R.I.), 1998.

[16] M. Farahi, J. Rubio, and D. Wilson, *The global control of a nonlinear wave equation*, International Journal of Control, 65 (1996), pp. 1–15.

[17] F. Friedlander and M. Joshi, *Introduction to the Theory of Distributions*, Cambridge University Press, 1998.

[18] H. Goldberg, W. Kampowsky, and F. Tröltzsch, *On Nemytskij operators in L^p-spaces of abstract functions*, Mathematische Nachrichten, 155 (1992), pp. 127–140.

[19] M. G. Grillakis, *Regularity for the wave equation with a critical nonlinearity*, Communications on Pure and Applied Mathematics, 45 (1992), pp. 749–774.

[20] W. Heisenberg, *Mesonenerzeugung als Stoßwellenproblem*, Zeitschrift für Physik A Hadrons and Nuclei, 133 (1952), pp. 65–79.

[21] H. Jia, B. Liu, W. Schlag, and G. Xu, *Generic and non-generic behavior of solutions to defocusing energy critical wave equation with potential in the radial case*, International Mathematics Research Notices, (2016), p. rnw181.

[22] R. Joly and C. Laurent, *Stabilization for the semilinear wave equation with geometric control condition*, Analysis & PDE, 6 (2013), pp. 1089–1119.

[23] K. Jörgens, *Über die nichtlinearen Wellengleichungen der mathematischen Physik*, Mathematische Annalen, 138 (1959), pp. 179–202.

[24] V. Kalantarov, A. Savostianov, and S. Zelik, *Attractors for damped quintic wave equations in bounded domains*, Annales Henri Poincaré, 17 (2016), pp. 2555–2584.

[25] C. Kenig, *The focusing energy-critical wave equation*, in Harmonic Analysis, Partial Differential Equations and Applications, Springer International Publishing, 2017, pp. 97–107.

[26] N.-A. Lai, *Blow up of critical semilinear wave equations on Schwarzschild spacetime*, Journal of Mathematical Analysis and Applications, 409 (2014), pp. 716–721.

[27] I. Lasiecka and D. Tataru, *Uniform boundary stabilization of semilinear wave equations with nonlinear boundary damping*, Differential Integral Equations, 6 (1993), pp. 507–533.

[28] I. E. Leonard and K. Sundaresan, *Geometry of Lebesgue-Bochner function spaces–smoothness*, Trans. Amer. Math. Soc., 198 (1974), pp. 229–229.

[29] H. A. Levine, *Minimal periods for solutions of semilinear wave equations in exterior domains and for solutions of the equations of nonlinear elasticity*, Journal of Mathematical Analysis and Applications, 135 (1988), pp. 297–308.

[30] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Springer Berlin Heidelberg, 1972.

[31] B. S. Mordukhovich and J.-P. Raymond, *Neumann boundary control of hyperbolic equations with pointwise state constraints*, SIAM Journal on Control and Optimization, 43 (2004), pp. 1354–1372.

[32] L. I. Schiff, *Nonlinear meson theory of nuclear forces. I. neutral scalar mesons with point-contact repulsion*, Physical Review, 84 (1951), pp. 1–9.

[33] J. Shatah and M. Struwe, *Regularity results for nonlinear wave equations*, The Annals of Mathematics, 138 (1993), p. 503.

[34] ———, *Well-posedness in the energy space for semilinear wave equations with critical growth*, International Mathematics Research Notices, 1994 (1994), p. 303.

[35] A. Shaw, T. Hill, S. Neild, and M. Friswell, *Periodic responses of a structure with 3:1 internal resonance*, Mechanical Systems and Signal Processing, 81 (2016), pp. 19–34.

[36] J. Simon, *Compact sets in the space* $L^p(0,T;B)$, Ann. Mat. Pura Appl. (4), 146 (1986), pp. 65–96.

[37] H. F. Smith and C. D. Sogge, *On the critical semilinear wave equation outside convex obstacles*, Journal of the American Mathematical Society, 8 (1995), pp. 879–879.

[38] C. D. Sogge, *Lectures on nonlinear wave equations*, no. Bd. 2 in Monographs in analysis, International Press, 1995.

[39] D. M. Stuart, *The geodesic hypothesis and non-topological solitons on pseudo-Riemannian manifolds*, Annales Scientifiques de l'École Normale Supérieure, 37 (2004), pp. 312–362.

[40] ———, *Geodesics and the Einstein nonlinear wave system*, Journal de Mathématiques Pures et Appliquées, 83 (2004), pp. 541–587.

[41] T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*, no. Nr. 106 in Conference Board of the Mathematical Sciences. Regional conference series in mathematics, American Mathematical Soc., 2006.