Adaptive Finite Elements for State-Constrained Optimal Control Problems
-
Convergence Analysis and A Posteriori Error Estimation

Von der Fakultät Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von
Simeon Steinig
aus Oberhausen

Hauptberichter: Prof. Dr. K.G. Siebert
Mitberichter: Prof. Dr. A. Rösch
2. Mitberichter: Prof. Dr. B. von Harrach

Tag der mündlichen Prüfung: 29.10.2014

Institut für angewandte Analysis und numerische Simulation der Universität Stuttgart

2014

Contents

List of Figures

Zusammenfassung der Dissertation

Optimalsteuerprobleme und besonders zustandsbeschränkte Optimalsteuerprobleme treten in vielen verschiedenen Wissenschaftsgebieten auf, so beispielweise in der Aeronautik, der Robotik, der Prozessteuerung und der Simulationstechnik im Autobau. Vor diesem Hintergrund ist es von Interesse, solche Probleme effizient zu lösen.

Kennzeichnend für Optimalsteuerprobleme ist die Existenz einer Steuerung u, die auf einen Zustand y einwirkt. Letzterer ist besimmt durch eine (gewöhnliche/partielle/stochastische) Differentialgleichung. In dieser Dissertation beschäftigten wir uns mit linearen, stationären partiellen Differentialgleichungen (PDE), d.h. insbesondere, dass der Zustand y eine lineare Funktion der Steuerung u ist, $y = Su$. Nun ist es so, dass das Lösen dieser Optimalsteuerprobleme das numerische Lösen von zwei linearen PDEs in jedem Schitt eines Optimierungsalgorithmus nach sich zieht. In den letzten Jahrzehnten wurde intensiv über das effiziente Lösen solcher linearen PDEs geforscht, insbesondere adaptive Finite Elemente Methoden haben sich dabei als besonders nützlich herauskristallisiert. U.a. deswegen lag es nahe, diese adaptiven Finite Elemente Methoden auch auf das spezielle Feld der zustandsbeschränkten Optimalsteuerprobleme anzuwenden:

In dieser Dissertation gab es zwei Ziele:

1. Das erste Ziel war es, ein **Basis-Konvergenzresultat** zu beweisen, d.h.: die Folge der diskreten Lösungen, die man durch die Diskretisierung des Optimalsteuerproblems mit Finiten Elementen gewinnt, $\bar{U}_k$, konvergiert gegen die eigentliche Lösung des undiskretisierten Optimalsteuerproblems $\bar{u}$: $\bar{U}_k \to \bar{u}$.

2. Das zweite Ziel war es, einen **zuverlässigen Fehlerschätzer** herzuleiten, d.h. eine obere Schranke für die Differenz $\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|$, die nur aus bekannten diskreten und kontinuierlichen Funktionen und *linearen Fehlern* besteht, wobei

   - $\bar{U}_k^\varepsilon$ die diskrete Lösung zu einem regularisierten Problem bezeichnet - mit Parameter $\varepsilon > 0$ - , welche die diskrete Lösung ist, die man tatsächlich berechnet, denn sie ist eine Lösung, die man - im Gegensatz zu $\bar{U}_k$ - durch Newton-artige Optimierungsmethoden gewinnen kann.

- *lineare Fehler* solche Fehler sind, die man durch etablierte a posteriori Fehlerschätzungstechniken der reinen PDE-Welt abschätzen kann, d.h. dies sind gerade die Fehler, die aus der Differenz zwischen einer Finite Element Lösung und der tatsächlichen Lösung einer PDE mit *bekannter* rechter Seite bestehen.

1. Ziel: Wir konnten erfolgreich die Konvergenz $\bar{U}_k \to \bar{u}$ exakt charakterisieren, Theorem 3.3.8 und Theorem 3.3.10, d.h. wir haben eine notwendige und hinreichende Bedingung für Konvergenz $\bar{U}_k \to \bar{u}$ hergeleitet. Diese Bedingung wurde mit Hilfe einer diskreten Größe formuliert, welche potenziell zum Steuern eines Algorithmus eingesetzt werden kann, s. Abschnitt 6.3. Wir konnten kein Beispiel dafür finden, dass diese Bedingung tatsächlich erfüllt ist; nichtsdestotrotz, da dieses Resultat bewiesen wurde ohne irgendeine zusätzliche Regularität für die Folge der Triangulierungen oder das Problem zu fordern, stellt es einen bedeutenden Beitrag zur Konvergenzanalysis von adaptiven Finite-Elemente-Methoden für zustandsbeschränkte Optimalsteuerprobleme da.

2. Ziel: Das zweite Ziel, der a posteriori Fehlerschätzer, wurde in Theorem 4.2.12 und Theorem 4.2.13 erreicht. Tatsächlich gelang es sogar nachzuweisen, dass der hergeleitete Fehlerschätzer unter milden Annahmen konvergiert, Theorem 4.3.14.

In den abschließenden Kapiteln dieser Dissertation konstruierten wir auf der Basis unseres a posteriori Fehlerschätzers einen **adaptiven Algorithmus**, Kapitel 5, bevor wir diesen erfolgreich an zwei Beispielen austesten, Kapitel 6.

Summary of PhD Thesis

Optimal control problems and in particular state-constrained optimal control problems frequently occur in all sorts of fields of science, from aerospace engineering to robotics, from process engineering to vehicle simulations. Against this backdrop, it is of interest to solve these kinds of problems in an efficient manner.
Optimal control problems are characterised by the existence of a control u acting on a state y which is governed by a (ordinary/partial/stochastic) differential equation. In this PhD thesis, we considered linear, stationary partial differential equations (PDE); in particular, the state y is a linear function of the control u, $y = Su$. Now, solving such optimal control problems numerically involves solving two linear PDEs in each iterate of an optimisation algorithm. Over the last decades much research has been undertaken to numerically solve such linear PDEs efficiently, especially discretisations with adaptive finite elements have been proven to be highly useful for such a task. Thus, trying to apply these adaptive finite element methods to the specific setting of state-constrained optimal control problems suggested itself as an appropriate approach:

The aim of this thesis was twofold:

1. The first goal was to prove a **basic convergence result**, i.e.: the sequence of discrete solutions obtained by discretising the optimal control problem with finite elements, $\bar{U}_k$, converges to the true solution of the undiscretised problem $\bar{u}$: $\bar{U}_k \to \bar{u}$.

2. The second goal was to derive a **reliable a posteriori error estimator**, i.e. an upper bound for the difference $\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|$ containing only known discrete and continuous functions and *linear errors*, where

   - $\bar{U}_k^\varepsilon$ denotes the discrete solution to a regularised problem - with parameter $\varepsilon > 0$ - which is the discrete solution actually computed, because unlike the unregularised solution $\bar{U}_k$, it is a solution which can be obtained by Newton-type optimisation algorithms.
   - *linear errors* are those errors which can be estimated by established a posteriori error estimation technique from the pure PDE world, i.e. these are the errors

consisting of the difference between a finite element solution and the true solution to a PDE with a *known* right hand side.

1st aim: We succeeded in characterising convergence $\bar{U}_k \to \bar{u}$ exactly, Theorem 3.3.8 and Theorem 3.3.10, i.e. we derived a necessary and sufficient condition for convergence $\bar{U}_k \to \bar{u}$, in terms of a discrete quantity which can potentially be used to steer a numerical algorithm, as we did in Section 6.3. We could not find an example, where this condition is fulfilled; nevertheless, because this result was achieved without assuming any additional regularity for the sequence of triangulations or the problem itself, it constitutes a major contribution to the convergence analysis for adaptive finite element methods for state-constrained optimal control problems.

2nd aim: The second goal, the a posteriori error estimator, was achieved in Theorem 4.2.12 and Theorem 4.2.13. Remarkably, the derived a posteriori estimator was proved to *converge* under relatively mild assumptions, Theorem 4.3.14.

In the concluding chapters of this thesis, we constructed an **adaptive algorithm** on the basis of our a posteriori error estimator, Chapter 5, before successfully testing it for two problems, Chapter 6.

Acknowledgements

I spent more than three years on this thesis, starting at the University of Duisburg-Essen, then moving on to the University of Stuttgart before finally completing it at the Technical University Dortmund.
I want to thank my advisors Prof. Dr. K.G. Siebert and Prof. Dr. A. Rösch for their support during this time. A big thank you goes out to my colleague Dr. F. Gaspoz at the University of Stuttgart for always taking time to discuss my research topics and being a patient listener.
I am also extremely grateful for the support of my family and close friends during the often lengthy dry spells and set-backs, which seem to plague mathematical research.
I also want to thank the Deutsche Forschungsgemeinschaft for 3 years of funding within the priority program SPP1253 'Optimization with Partial Differential Equations'.
Last but not least, I especially want to thank Prof. Dr. C. Meyer for giving me the opportunity to work at the TU Dortmund during the last phase of my PhD and for frequently taking time off to discuss my research.

Chapter 1

Motivation

Modelling and optimising physical processes naturally lead to mathematical optimal control problems with state constraints. Let us illustrate this with the help of two examples:

- In problems of heat conduction a typical goal would be to find an optimally adjusted heat source to come as close as possible to a desired temperature distribution in a given workpiece. Here, the control is the regulation of the heat source and the state is the temperature distribution. A partial differential equation, namely the (stationary) heat equation links these to variables. A typical state constraint would be to force the temperature distribution to stay below a certain threshold, for instance to prevent the material from melting.

- In the optimisation of diffusion processes the latter being modelled e.g. by Fick's diffusion, compare [22], Chapter 1, a characteristic aim would be to achieve a desired concentration of a chemical substance by optimally adjusting the chemical sources by e.g. decreasing or increasing inflow. In this setting the control is represented by the calibration of the chemical source while the state itself is represented by the concentration of the substance. A natural state constraint here would be to demand that the concentration do not surpass a certain critical threshold, e.g. possibly for health reasons.

In these two examples we already discern the structure of state-constrained problems: A control u governs a state y, determined by a partial differential equation, with which we want to come close to a desired state y_d subject to a constraint on the state y. In mathematical terms this problem represents an infinite-dimensional optimisation problem for which the solution is in general not known.

In view of the fact that the solution is in general unkown, it would be highly desirable to solve these problems efficiently numerically. As the first step towards solving such a problem numerically we have to discretise it, i.e. we have to break it down to a finite-dimensional optimisation problem from an infinite-dimensional one. In this setting, this is usually done by using finite element spaces (FE-spaces). The effort that we now have to put in to solve this problem is heavily influenced by the dimension of our finite element spaces, i.e. by the amount of degrees of freedom (DOFs) that are available to us. Clearly, it would now be of advantage to be able to use those DOFs smartly, i.e. in a problem-dependent, adaptive manner, while also being certain that adding DOFs and thus increasing the computational effort actually gets us closer to the unknown solution. The latter aspect is the one of **convergence** of this method, the focus of Chapter 3, while the former demands the derivation of a **reliable a posteriori error estimator**, the subject of Chapter 4, on whose basis we can judge the quality and thus the smartness of our finite element approximation. To put it in a brief mathematical term: We want to build a *convergent adaptive finite element method.* This was the ultimate aim of this thesis, one which is challenging both from an analytical and numerical persepctive.

Adaptive finite element methods have already been succesfully applied to PDEs, e.g. [84], [34],[76] and many many more and purely control constrained optimal control problems, e.g. [50], to name just one. In the a posteriori error analysis of state-constrained problems, research has also been underway, mostly focussing on estimates of 'quantities of interest' which do not provide a **reliable** bound - an upper bound up to constants depending on data - of the error between the current discrete solution and the true solution [86], [6], while others have not taken into account certain error sources, [46]. Besides, in general, the regularisation error, which is a natural part of the numerical solution of state-constrained optimal control problems, compare Section 2.2.2, was also neglected. In addition, it is not clear whether the sequence of finite element solutions actually converge to the true one.

The fundamental difficulty of state-constrained optimal control problems is their **lack of regularity**: Once a uniform mesh is no longer demanded, convergence properties of FE-solutions to PDEs, such as convergence in L_∞, can no longer be presupposed. In this situation, the inherent difficulties of state constraints, chiefly the singular nature of the associated Lagrange multiplier, see Section 2.2.1, strike with full ferocity. Yet, these difficulties also formed part of my motivation because even though it proved to be a formidable challenge to come up with a whole new set of tools for the analysis of state-constrained optimal control problems - tools which could potentially be used in many other branches - it was precisely this challenge which provided me with the opportunity to explore branches of mathematics, especially functional analysis, such as the interpolation of spaces, but also optimisation in Banach spaces, whose rich applications and powerful theory offered a truly fascinating study.

Chapter 2

General Framework

This chapter introduces the reader to the general mathematical framework of optimal control problems and the adaptive finite element method. We will try to give a concise overview, referring the reader - whenever it is needed - for more detailed information to additional literature.
We will start by collecting some important notions and results respecting function spaces such as embeddings, dual spaces and separation of convex sets, before briefly describing the function spaces which one naturally deals with when tackling optimal control problems.

2.1 Function Spaces

Especially in the context of optimisation in Banach spaces, notions such as duality, reflexivity, weak compactness as well as properties of convex sets will naturally come into play. It is therefore advantageous to briefly gather important definitions and theorems, not least because the reader will be able to follow this thesis more easily.

2.1.1 Banach Spaces and Convex Sets

Throughout this section, X is a real Banach space with norm $\|\cdot\|_X$, in short $(X, \|\cdot\|_X)$. Its *dual space*, denoted by X^*, consists of all linear and continuous mappings - referred to as linear functionals - $f : X \to \mathbb{R}$ and is endowed with the canonical norm

$$\|f\|_{X^*} := \sup_{\|x\|_X = 1} |f(x)| \,.$$

It is a Banach space itself. Sometimes we write

$$\langle f, x \rangle_{X^*,X} = \langle f, x \rangle := f(x)$$

where $\langle \cdot, \cdot \rangle_{X^*,X}$ is the duality product, which bears certain similarities to a scalar product. If the bi-dual of X, $X^{**} = (X^*)^*$, can be identified with X by an isometric isomorphism, we say that X is *reflexive.*

A special class of Banach spaces are Hilbert spaces. Hilbert spaces H possess a scalar product

$$(u, v)_H \ \ \forall u, v \in H$$

and are normed with the canonical norm induced by the scalar product, i.e.:

$$\|u\|_H = \sqrt{(u,u)_H}$$

Hilbert spaces have the property that their dual space H^* can be isometrically isomorphically identified with the space H itself. As a consequence, they are always reflexive. The isomorphism is referred to as the Riesz-isomorphism, and the images in H of functionals in H^* are called Riesz representatives. Sometimes, though, it can still be advantageous to treat the dual H^* as a separate space.

n-tuples of Hilbert spaces $H = H_1 \times H_2 \times ... \times H_n$ are Hilbert spaces themselves equipped with the canonical scalar product

$$(u, v)_H := \sum_{i=1}^{n} (u^i, v^i)_{H_i}$$

and induced norm.

Dual spaces induce a topology that is usually referred to as the *weak topology.* In particular, we are interested in the notion of weak convergence, which will be used on several occasions in this thesis.

Definition 2.1.1 (weak convergence)**.** *A sequence $\{x_k\} \subset X$ converges weakly to an element $x \in X$, denoted by $x_k \rightharpoonup x$, $k \to \infty$, if*

$$\langle f, x_k \rangle_{X^*,X} \to \langle f, x \rangle_{X^*,X}, \ k \to \infty \quad \forall f \in X^*.$$

The notion of weak convergence is a crucial tool in deriving existence results for optimisation problems because - in some sense - it replaces the classic principle of finite dimensional analysis that every bounded sequence contains a convergent subsequence, which plays a key role in proofs of existence of minima. First of all, though, we need some separation theorems for convex sets: The first can be found in [58], Section 5.12 Theorem 1:

Theorem 2.1.2 (Mazur's Theorem/Geometric Hahn-Banach)**.** *Let C be a convex subset of X with non-empty interior. Suppose V is an affine subspace in X with $V \cap \operatorname{int}(C) = \emptyset$. Then* *there exists $x^* \in X^*$ such that the hyperplane H*

$$H = \{x \in X \ : \ \langle x^*, x\rangle_{X^*,X} = c, \ c \in \mathbb{R}\}$$

fulfils

$$V \subset H, \ H \cap \operatorname{int}(C) = \emptyset, \ \langle x^*, x\rangle_{X^*,X} < c \, \forall x \in \operatorname{int}(C).$$

Another separation theorem is the following, which can be found in [3], Theorem 6.11.

Theorem 2.1.3. *Let X be a Banach space, $C \subset X$ non-empty, convex and closed. Besides, let $x_0 \in X$ with $x_0 \notin C$.Then there exists $f \in X^*$ and $\alpha \in \mathbb{R}$ such that*

$$\langle f, x\rangle_{X^*,X} \leq \alpha \, \forall x \in C$$

and

$$\langle f, x_0\rangle_{X^*,X} > \alpha.$$

Obviously, $f \neq 0$ and

$$\{x \in X \ : \ \langle f, x\rangle_{X^*,X} = \alpha\}$$

is a hyperplane in X.

The next theorem is to a certain extent the infinite-dimensional equivalent of the Bolzano-Weierstrass principle in finite dimension formulated in terms of weak convergence. A proof can e.g. be found in [87], Section V.2, Theorem 1.

Theorem 2.1.4. *Let X be a reflexive Banach space and $\{x_k\}$ be a bounded sequence. Then $\{x_k\}$ possesses a weakly convergent subsequence.*

This result will be used frequently throughout this thesis. Next, let us prove a Lemma offering a way to make the step from (weak) convergence of subsequence to (weak) convergence of the entire sequence.

Lemma 2.1.5. *Let X be a reflexive Banach space and $\{x_k\} \subset X$ be a bounded sequence. Suppose that every weakly convergent subsequence of $\{x_k\}$ converges to the same $x \in X$. Then the entire sequence weakly converges to x:*

$$x_k \rightharpoonup x, \ k \to \infty$$

Proof. Suppose the contrary. Then there exists a subsequence $\left\{x_{k_j}\right\}$, an $\varepsilon > 0$ and $f \in X^*$ such that

$$|\langle f, x_{k_j}\rangle - \langle f, x\rangle| \geq \varepsilon \, \forall j \in \mathbb{N} \tag{2.1.1}$$

However, since x_{k_j} is bounded by assumption, there exists a subsequence of $\left\{x_{k_j}\right\}$ denoted by $\left\{x_{k_{j_l}}\right\}$. $\left\{x_{k_{j_l}}\right\}$ being a subsequence of $\{x_k\}$, it weakly converges to x by assumption. Thus, for some $L = L(\varepsilon) \in \mathbb{N}$,

$$|\langle f, x_{k_{j_l}}\rangle - \langle f, x\rangle| < \varepsilon \,\forall l \geq L.$$

This is the desired contradiction to (2.1.1) which completes the proof. □

We will often encounter a situation where the bounded sequence belongs to a certain subset of X, and we need the weak limit to belong to this subset, too.

Theorem 2.1.6. *Let C be a convex and closed subset of X. Then C is weakly compact, i.e. every bounded sequence $\{x_k\}$ contains a weakly convergent subsequence $\{x_{k_n}\}$ weakly converging to an element $x \in C$ as $n \to \infty$.*

The proof of this theorem can be found in [3], Theorem 6.12.

Another important property of convex sets is the fact that they allow for the definition of a projection operator onto them satisfying a variational inequality, which is very useful for interpreting first-order optimality conditions for the optimisation problems we will consider later:

Theorem 2.1.7 (projection on convex sets)**.** *Let H be a Hilbert space and C a convex, closed and non-empty subset of H. Then, for every $x \in H$, there exists a unique element $\Pi_C(x) \in C$, the projection of x on C, solving the minimisation problem*

$$\inf_{v\in C} \frac{1}{2} \|v - x\|_H^2 , \tag{2.1.2}$$

and satisfying

$$(\Pi_C(x) - x, v - \Pi_C(x))_H \geq 0 \;\; \forall v \in C. \tag{2.1.3}$$

Conversely, if an element $\tilde{v}$ satisfies

$$(\tilde{v} - x, v - \tilde{v})_H \geq 0 \;\; \forall v \in C, \tag{2.1.4}$$

then $\tilde{v} \in C$ solves (2.1.2) *and thus $\tilde{v} = \Pi_C(x)$.*
In addition, the projection Π_C is Lipschitz continuous with Lipschitz constant 1, *i.e.*

$$\|\Pi_C(x) - \Pi_C(y)\|_H \leq \|x - y\|_H \tag{2.1.5}$$

Proof. The existence of $\Pi_C(x)$ can be easily transferred from the case where C is a closed subspace of H (see e.g. [4] Theorem 10.5), because in essence, everything that is needed is that for the infimal sequence $\{v_n\}_{n\in\mathbb{N}} \subset C$ we have

$$\frac{1}{2}(v_n + v_m) \in C,$$

which is true for convex sets.
As to the uniqueness of the solution to (2.1.2): Let us suppose there exist two solutions v_1, v_2 to (2.1.2) with $v_1 \neq v_2$ and

$$d = \inf_{v\in C} \|v - x\|_H^2 = \|v_1 - x\|_H^2 = \|v_2 - x\|_H^2$$

Using the parallelogram identity, we obtain:

$$\begin{aligned}\frac{1}{2}\left\|\frac{v_1 + v_2}{2} - x\right\|_H^2 &= \frac{1}{4}\|v_1 - x\|_H^2 + \frac{1}{4}\|v_2 - x\|_H^2 - \frac{1}{8}\|v_1 - v_2\|_H^2 \\ &< \frac{1}{4}\|v_1 - x\|_H^2 + \frac{1}{4}\|v_2 - x\|_H^2 \\ &= \frac{d}{2} + \frac{d}{2} = d,\end{aligned}$$

which is a contradiction, since v_1 and v_2 solve (2.1.2). Hence, the solution to (2.1.2) is unique, and we denote it by $\Pi_C(x)$.
The solution $\Pi_C x$ satisfies (2.1.3) because for all $t \in (0, 1]$ we can deduce

$$\begin{aligned}0 &\leq \frac{1}{2t}(\|(1 - t)\Pi_C(x) + tv - x\|_H^2 - \|\Pi_C(x) - x\|_H^2) \\ &= \frac{1}{2t}((\Pi_C(x) - x, t(v - \Pi_C(x))_H + \|tv\|_H^2).\end{aligned}$$

Drawing the limit $t \to 0$ yields the assertion (2.1.3).
Now, suppose $\tilde{v}$ fulfills (2.1.4), then for all $v \in C$ we can estimate in the ensuing way:

$$\begin{aligned}\frac{1}{2}\|v - x\|_H^2 &= (v - \tilde{v}, \tilde{v} - x)_H + \frac{1}{2}\|v - \tilde{v}\|_H^2 + \frac{1}{2}\|x - \tilde{v}\|_H^2 \\ &\geq \frac{1}{2}\|x - \tilde{v}\|_H^2.\end{aligned}$$

Hence, $\tilde{v}$ solves (2.1.2).
Let us now turn to (2.1.5): (2.1.3) yields

$$\begin{aligned}(\Pi_C(x) - x, \Pi_C(y) - \Pi_C(x)) \geq 0 \\ (\Pi_C(y) - y, \Pi_C(x) - \Pi_C(y)) \geq 0.\end{aligned}$$

Adding and rearranging these two inequalities, we can conclude harnessing Cauchy-Schwarz's inequality

$$\begin{aligned}\|\Pi_C(x)-\Pi_C(y)\|_H^2 &\leq (x-y,\Pi_C(y)-\Pi_C(x))\\ &\leq \|x-y\|_H\,\|\Pi_C(y)-\Pi_C(x)\|_H\end{aligned}$$

Divinding by $\|\Pi_C(y)-\Pi_C(x)\|_H$ completes the proof. □

A subclass of convex sets are convex cones, which are important in optimisation because they induce a pre-order relation that provides an extension to the $\leq$ on the real numbers in general vector spaces.

Definition 2.1.8 (Cones)**.** *Let $C \subset X$ be a convex set such that*

$$x \in C \Rightarrow \lambda x \in C \ \forall \lambda \geq 0.$$

Then C is called a convex cone*.*
A cone induces a pre-ordering $\leq_C$ by the relation

$$x \leq_C y \ \Leftrightarrow \ y-x \in C.$$

The relation $\leq_C$ is compatible with vector space operations, i.e.

$$\begin{aligned}\forall x,y,z \in X \ &: x \leq_C y \ \Rightarrow \ x+z \leq_C y+z\\ \forall \lambda \geq 0 \ &: x \leq_C y \ \Rightarrow \ \lambda x \leq_C \lambda y\end{aligned}$$

The polar cone C^- to C is given by:

$$C^- := \{f \in X^* \ : \ \langle f,x\rangle \leq 0 \ \forall x \in C\}.$$

If we do not consider C and its polar cone w.r.t to the canonical norm-topology in X but w.r.t to the norm-topology of another Banach space $Z \subset X$, we specifically write:

$$C_Z = C \cap Z$$

and

$$C_Z^- = \{\varphi \in Z^* \ : \ \langle \varphi,z\rangle_{Z^*,Z} \leq 0 \ \forall z \in C_Z\}.$$

2.1.2 Linear & More General Mappings, Notions of Differentiabilty

We have already encountered the dual space X^* of an arbitrary real Banach space X. Sometimes it is also necessary to treat more general mappings. We will now list some results pertaining to such mappings, starting with linear mappings $S : X \mapsto Y$ between two Banach spaces X and Y.

Theorem 2.1.9. *Let X and Y be two Banach spaces. Then the space $\mathcal{L}(X,Y)$ consisting of all linear and continuous operators $S : X \mapsto Y$ is a Banach space itself endowed with the norm*

$$\|S\|_{\mathcal{L}(X,Y)} := \sup_{\|x\|_X=1} \|Sx\|_Y = \sup_{\|x\|_X \leq 1} \|Sx\|_Y = \sup_{x\in X\setminus\{0\}} \frac{\|Sx\|_Y}{\|x\|_X}$$

For every $S \in \mathcal{L}(X,Y)$ there exists an adjoint operator S^* mapping Y^* to X^*. It maps $y \in Y^*$ linearly and continuously on the element $x = y(S) \in X^*$. In the special case of Hilbert spaces, the adjoint operator S^* to an operator $S \in \mathcal{L}(H_1, H_2)$, where H_1, H_2 are Hilbert spaces, maps H_2 to H_1 and is defined by the relation

$$(Su, v)_{H_2} = (u, S^*v)_{H_1} \;\; \forall u \in H_1,\; v \in H_2.$$

For existence proofs for Lagrange multipliers, the following characterisation of surjective linear mappings, the famous open mapping theorem, is very useful. A proof can be found in [72], Theorem 2.11.

Theorem 2.1.10 (Open Mapping Theorem)**.** *Let X, Y be two Banach spaces and $S : X \to Y$ a continuous surjective linear mapping. Then S is an open mapping, i.e. the image of every open subset $V \subset X$, $S(V)$, is open in Y.*

A special class of linear and continuous operators are embedding operators. At several points in this thesis such operators will be important. We will specify this notion in the ensuing definition which can be found in [1], Definition 1.25.

Definition 2.1.11 (Embeddings)**.** *Let Y be a Banach space with norm $\|\cdot\|_Y$ and $Y \subset X$. We say that Y* embeds *(continuously) into X, $Y \hookrightarrow X$, if Y is a vector subspace of X and the mapping $I : Y \to X$ defined by*

$$Iy = y \; \forall y \in Y$$

is continuous.
If I is also compact, i.e.

$$x_k \rightharpoonup x \;\in Y \;\Rightarrow\; Ix_k \to Ix,\; k \to \infty \;\in X,$$

we say that $Y \hookrightarrow X$ compactly*.*

In optimisation one invariably encounters *goal functions*, i.e. mappings $g : V \to \mathbb{R}$, where V is a convex subset of a Banach space X. In view of Theorem 2.1.6, the question arises how they act on weakly convergent sequences.

First of all, let us define the appropriate notions:

Definition 2.1.12 (weak upper/lower semicontinuity). *A function $g : X \to \mathbb{R}$ where X is a Banach space is weakly lower semicontinuous if $x_k \rightharpoonup x$ implies*

$$\liminf_{k\to\infty} g(x_k) \geq g(x)$$

It is called weakly upper semicontinuous if $-g$ is weakly lower semicontinuous.

Obviously, a question to ask is what kind of functions fulfil the properties mentioned in Definition 2.1.12. Is mere continuity of g in the strong, norm-topology enough? For affine functions this is obviously the case; for nonlinear functions, though, an extra (sufficient) ingredient is needed, namely convexity:

Definition 2.1.13 ((Strictly) Convex and Concave Functions). *Let $g : V \subset X \to \mathbb{R}$ be a function defined on a convex subset V of X. g is* convex *if*

$$g(\lambda u + (1-\lambda)v) \leq \lambda g(u) + (1-\lambda)g(v) \quad \forall u, v \in V,\ \lambda \in (0,1).$$

It is called strictly convex *if*

$$g(\lambda u + (1-\lambda)v) < \lambda g(u) + (1-\lambda)g(v) \quad \forall u, v \in V,\ u \neq v,\ \lambda \in (0,1).$$

g is (strictly) concave *if $-g$ is (strictly) convex.*

With this definition we can return to the subject of weakly continuous functions:

Theorem 2.1.14. *Let $g : X \to \mathbb{R}$ be a convex, continuous function and X a reflexive Banach space. Then g is weakly lower semicontinuous. Conversely, if g is concave and continuous, it is weakly upper semicontinuous.*

Proof. For a convex and continuous function g, the epigraph

$$\mathrm{epi}(g) := \{(x,a) \in X \times \mathbb{R} \,:\, g(x) \leq a\}$$

is closed and convex, compare [29], Proposition 2.1. Proposition 2.3. and Corollary 2.2 in [29] now yield that every convex and continuous function is weakly lower semicontinuous. The second part of the theorem is a consequence of the first. If g is concave, then $-g$ is convex, thus:

$$\limsup_{k\to\infty} g(x_k) = -\liminf_{k\to\infty} -g(x_k) \leq -(-g(x)) = g(x).$$

□

In the context of existence results for minimisation problems, two properties are often needed as prerequisites: radial unboundedness and boundedness from below:

Definition 2.1.15. *Let $g : V \subset X \to \mathbb{R}$ be a function defined on a subset V of X. The function g is radially unbounded on V if for every sequence $\{x_k\} \subset V$ with $\|x_k\|_X \to \infty$ $g(x_k) \to +\infty$ follows as $k \to \infty$.*
The function g is said to be bounded from below on V if there exists a real number b such that

$$g(u) \geq b \quad \forall u \in V.$$

Notions of differentiability naturally come into play when one wants to formulate optimality conditions. In the course of this paper we will employ two: That of Fréchet-differentiabilty, which is perhaps the strictest, and that of semi-smoothness, which is a comparatively weak one, but nevertheless a very valuable tool in analysing superlinearly convergent optimisation algorithms.
Let us commence with the notion of Fréchet-differentiability; the following definition can be found in [45], Definition 1.29:

Definition 2.1.16 (Fréchet-differentiability)**.** *An operator $G : X \to Y$, where X, Y are Banach spaces, is called Fréchet differentiable at $x \in X$, if there exists a linear operator $G^{'}(x) \in \mathcal{L}(X,Y)$ such that*

$$\left\|G(x+h) - G(x) - G^{'}(x)h\right\|_Y = o(\|h\|_X) \ \text{for} \ \ \|h\|_X \to 0. \tag{2.1.6}$$

If G is Fréchet-differentiable at every $x \in V$, where V is any open subset of X, then G is Fréchet-differentiable on V.

Especially in the context of proving certain convergence rates for optimisation algorithms, it will also be helpful to consider non-linear operators G which need not be Fréchet-differentiable, but still possess a certain smoothness that can be compared to the smoothness of Fréchet-differentiable operators (2.1.6). These operators will be called semismooth. Following Section 3.2. in [82], we define the notion of semismoothness in the following fashion:

Definition 2.1.17 (Semismoothness)**.** *Let $G : V \subset X \to Y$ be defined on an open subset V of a Banach space X with images in the Banach space Y. Furthermore, let a set-valued mapping $\partial G : V \to \mathcal{L}(X,Y)$ be given with non-empty images, i.e. $\partial G(x) \neq \emptyset$ for all $x \in V$:*

- *G is ∂G-semismooth at x if G is continuous in a neighbourhood of x and*

$$\sup_{M \in \partial G(x+h)} \|G(x+h) - G(x) - Mh\|_Y = o(\|h\|_X) \ \text{for} \ \ \|h\|_X \to 0.$$

- *G is α-order ∂G-semismooth at x, $0 < \alpha \leq 1$ if G is continuous in a neighbourhood of* *x and*

$$\sup_{M \in \partial G(x+h)} \|G(x+h) - G(x) - Mh\|_Y = o(\|h\|_X^{1+\alpha}) \textit{ for } \|h\|_X \to 0.$$

The multifunction ∂G will be called generalised differential of G and the non-emptiness of ∂G will always be assumed. In particular, ∂G-semismoothness of G will always entail that $\partial G(v) \neq \emptyset$ for all $v \in V$.

The importance of the concept of semismoothness will be illustrated in Section 2.2.3, where we will discuss a q-superlinearly convergent method for solving non-linear equations, the Semismooth Newton Method, which will be highly useful for solving optimality systems. At this point, let us merely specify what superlinear convergence means. The definition below can be found in [45], Section 2.1:

Definition 2.1.18 (superlinear convergence)**.** *Let X be a Banach spaces and $\{x_k\} \subset X$ a sequence with $x_k \to x$, $x \in X$.*
The sequence x_k converges q-superlinearly to x if $x_k \to x$ as $k \to \infty$ and

$$\|x_{k+1} - x\|_X = o(\|x_k - x\|_X).$$

If for some $\alpha > 0$

$$\|x_{k+1} - x\|_X = O(\|x_k - x\|_X^{1+\alpha}),$$

then $x_k \to x$ converges q-superlinearly with order $1 + \alpha$.

In the next section, we will list some crucial results pertaining to the solvability of variational equalities because they will form a key part in the analysis of optimal control problems.

2.1.3 Linear Equations

In this section we will analyse a variational equality of the following general type:

$$y \in Y \ : \quad \mathcal{B}[y, w] = \langle f, w \rangle \ \ \forall w \in Y, \tag{2.1.7}$$

where Y is a Hilbert space and $f \in Y^*$. $\mathcal{B}$ is a *continuous bilinear form* on Y, a notion which we will specify in the next definition

Definition 2.1.19 (Continuous Bilinear Form)**.** *Suppose Y is a Hilbert space. A mapping $\mathcal{B} : Y \times Y \to \mathbb{R}$ is a* ***continuous bilinear form*** *if it is linear in each component and if there* *exists a constant c*

$$|\mathcal{B}[y,w]| \leq c \|y\|_Y \|w\|_Y \quad \forall y, w \in Y.$$

The norm $\|\mathcal{B}\|$ of B is defined by

$$\|\mathcal{B}\| := \inf \left\{ c \, : \, |\mathcal{B}[y,w]| \leq c \|y\|_Y \|w\|_Y \quad \forall y, w \in Y \right\}.$$

In case

$$\mathcal{B}[y,w] = \mathcal{B}[w,y] \quad \forall y, w \in Y,$$

$\mathcal{B}$ is called ***symmetric****.*
If

$$\mathcal{B}[w,w] \geq \beta \|w\|_Y^2, \quad \forall w \in Y,$$

$\mathcal{B}$ is called ***coercive****.*

We are interested in conditions under which (2.1.7) is uniquely solvable, that is conditions, which safeguard that for every $f \in Y^*$ there exists a unique solution $y = Sf$ depending continuously on the data, i.e. for some constant c_S

The key condition which is necessary and sufficient for the solvability of (2.1.7) is the inf − sup-condition, which is the subject of the next theorem, taken from [61], Theorem 3.3 or [65], Section 2.3. Theorem 2.

Theorem 2.1.20 (Nečas Theorem)**.** *Let Y be a Hilbert space and $\mathcal{B} : Y \times Y \to \mathbb{R}$ be a continuous bilinear form. Then the variational problem* (2.1.7) *admits a unique solution $y = Sf$ if and only if*

$$\inf_{w \in Y} \sup_{z \in Y} \frac{\mathcal{B}[w,z]}{\|w\|_Y \|z\|_Y} = \inf_{z \in Y} \sup_{w \in Y} \frac{\mathcal{B}[w,z]}{\|w\|_Y \|z\|_Y} = \alpha > 0 \tag{2.1.8}$$

or equivalently:

$$\begin{gathered} \exists \alpha > 0 \, : \, \sup_{z \in Y} \frac{\mathcal{B}[w,z]}{\|z\|_Y} \geq \alpha \|w\|_Y \\ \textit{and for every } 0 \neq z \in Y \textit{ there exists } w \in Y \textit{ such that } \mathcal{B}[w,z] \neq 0. \end{gathered} \tag{2.1.9}$$

In addition, y satisfies

$$\|y\|_Y = \|Sf\|_Y \leq \frac{1}{\alpha} \|f\|_{Y^*} .$$

In the example section, Section 2.4, the reader will encounter continuous bilinear forms which are symmetric and coercive. The following corollary ensures that these properties imply the $\inf - \sup$-conditions of Theorem 2.1.20:

Corollary 2.1.21. *Let Y be a Hilbert space and $\mathcal{B} : Y \times Y \to \mathbb{R}$ be a symmetric and coercive bilinear form. Then $\mathcal{B}$ satisfies the condition* (2.1.9) *of Theorem 2.1.20 with $\alpha \geq \beta$, where β is the coercivity constant.*

Proof. We can estimate in the following fashion:

$$\begin{aligned} \sup_{z\in Y} \frac{\mathcal{B}[w,z]}{\|z\|_Y} &\geq \frac{\mathcal{B}[w,w]}{\|w\|_Y} \\ &\geq \beta \|w\|_Y . \end{aligned}$$

In addition, for $0 \neq z \in Y$, we obtain

$$\mathcal{B}[z,z] \geq \beta \|z\|_Y^2 > 0,$$

which yields (2.1.9) and $\alpha \geq \beta$. □

One important aspect of coercivity is that it is immediately inherited by (closed) subspaces $Z \subset Y$ and thus, in particular by finite-dimensional subspaces, which is very helpful for proving existence and stability results for discretisations of linear equations. We tackle the following problem.

$$y_Z \in Z \ : \quad \mathcal{B}[y_Z, w_Z] = \langle f, w_Z \rangle \ \ \forall w_Z \in Z, \tag{2.1.10}$$

where Z is a closed subspace of Y.

Corollary 2.1.22. *Suppose problem* (2.1.10) *is given with a bilinear form $\mathcal{B} : Y \times Y \to \mathbb{R}$ that is coercive and continuous on Y with coercivity constant $\beta > 0$. Suppose further that Z is a closed subspace of the Hilbert space Y with the same norm, i.e. $\|\cdot\|_Z = \|\cdot\|_Y$. Then there exists a unique solution $S_Z f = y_Z$ of* (2.1.10) *with*

$$\|S_Z f\|_Y = \|y_Z\|_Y \leq \frac{1}{\beta} \|f\|_{Y^*} . \tag{2.1.11}$$

In particular, the estimate above does not depend on the subspace Z.

Proof. Coercivity and continuity of $\mathcal{B}$ on Y imply coercivity and continuity on Z of $\mathcal{B}$, since the same norm is used. Corollary 2.1.21 then yields condition (2.1.9) in the following modified way:

$$\exists \alpha > 0 \ : \ \sup_{w_Z \in Z} \frac{\mathcal{B}[q_Z, w_Z]}{\|w_Z\|_Y} \geq \alpha \, \|q_Z\|_Y \, ,$$

and for every $0 \neq w_Z \in Z$, there exists $q_Z \in Z$ such that $\mathcal{B}[q_Z, w_Z] \neq 0$.

That in turn yields the existence of a solution $S_Z f = y_Z \in Z$ thanks to Theorem 2.1.20. Corollary 2.1.21 yields:

$$\|S_Z f\|_Y \leq \frac{1}{\beta} \, \|f\|_{Y^*}$$

where β is the coercivity constant of $\mathcal{B}$ on Y. This is estimate (2.1.11) which completes the proof. □

We will conclude this section with a remark on the right-hand side in (2.1.7).

Remark 2.1.23. *Suppose that a Hilbert space U with $U \hookrightarrow Y^*$ and the following modified form of* (2.1.7) *are given with $u \in U$:*

$$y \in Y \ : \quad \mathcal{B}[y, w] = (u, w)_U \ \ \forall w \in Y.$$

Then all the preceding results Theorem 2.1.20, Corollaries 2.1.21 and 2.1.22 can be immediately transferred to this setting because

$$f_u(w) := (u, w)_U \ w \in Y$$

is a linear functional on Y. In addition, the embedding $U \hookrightarrow Y^*$ yields

$$\|f_u\|_{Y^*} \leq c \, \|u\|_U \, ,$$

and the stability results of Theorem 2.1.20 and (2.1.11) *also hold in the following sense*

$$\|Su\|_Y \leq \frac{c}{\beta} \, \|u\|_U \, .$$

2.1.4 Spaces of Classically Differentiable and Continuous Functions

In this section, we will specify the notion of (Hölder-) continuous and continuously differentiable spaces of functions, which we will often come across in the course of the thesis.
Those functions are defined on a set Ω which throughout this thesis is a bounded domain in $\mathbb{R}^d$ with $d = 2$ or $d = 3$ with closure $\bar{\Omega}$.
We now introduce the following spaces, compare also [31], Section 5.1. Theorem 1.

Definition 2.1.24. *The space $C^{l,\gamma}(\bar{\Omega})$, $0 \leq l \leq \infty$, $0 \leq \gamma \leq 1$, consists of all functions g which are l-times classically differentiable and whose derivatives of order l are Hölder conti-*

nous with exponent γ. In case $0 \leq l < \infty$ the norm is defined by

$$\|g\|_{C^{l,\gamma}(\bar{\Omega})} := \sum_{|\alpha|\leq l} \max_{x\in\bar{\Omega}} |D^\alpha g(x)| + \sum_{|\alpha|=l} \sup_{x,y\in\bar{\Omega},x\neq y} \Big(\frac{|g(x)-g(y)|}{|x-y|^\gamma}\Big).$$

It is a Banach space.

At several points throughout this thesis, we will encounter the space $C_0^\infty(\Omega)$, which we will now define:

Definition 2.1.25. *The space $C_0^\infty(\Omega)$ consists of all functions v with compact support - i.e. the set*

$$\text{supp}(v) := \overline{\{x \in \Omega \,:\, |v(x)| > 0\}}$$

is compact - which also satisfy $v \in C^\infty(\Omega)$. It is not metrisable, hence, in particular it does not possess any norm.

2.1.5 Regularity of the Domain Ω

In the context of regularity results for the (weak) solutions of partial differential equations, one often encounters conditions on the boundary of the domain, the boundary has to fulfil certain smoothness conditions. The notion of boundary smoothness is defined below, the definition itself can be found in [36], Section 6.2.

Definition 2.1.26. *A bounded domain Ω in $\mathbb{R}^d$ and its boundary $\partial\Omega$ are of class $C^{k,\alpha}$, $0 \leq \alpha \leq 1$, $0 \leq k \leq \infty$ if at each point $x_0 \in \partial\Omega$ there exists a ball $B_r(x_0)$ with radius $r > 0$ and a one-to-one mapping Ψ of B onto $D \subset \mathbb{R}^d$ such that:*

- $\Psi(B \cap \Omega) \subset \mathbb{R}^d_+$,
- $\Psi(B \cap \partial\Omega) \subset \partial\mathbb{R}^d_+$,
- $\Psi \in C^{k,\alpha}(B)$, $\Psi^{-1} \in C^{k,\alpha}(D)$,

where

$$\mathbb{R}^d_+ := \Big\{x = (x_1, ..., x_d) \in \mathbb{R}^d :\ x_d > 0\Big\},$$

and $\partial\mathbb{R}^d_+$ denotes its boundary.

In particular, compare again [36], Section 6.2, a domain Ω is in $C^{k,\alpha}$ if for each $x_0 \in \partial\Omega$ there exists a neighbourhood of x_0 in which $\partial\Omega$ is the graph of a $C^{k,\alpha}$ function of $d-1$ of the coordinates $x_1, ..., x_d$. The latter characterisation of a $C^{k,\alpha}$-domain can be found in [33], Definition 1.18.

We can now turn our attention to Lebesgue L_p-spaces and Sobolev W_p^k-spaces, which, in the context of optimal control problems and in that of weak solutions for partial differential equations, naturally come into play.

2.1.6 Lebesgue and Sobolev Spaces

Let us start with the Lebesgue spaces $L_p(\Omega)$ with $1 \leq p < \infty$:

Definition 2.1.27. *Let $f : \Omega \mapsto \mathbb{R}$ be a function measurable with respect to the standard d-dimensional Lebesgue measure $d\Omega$. $f \in L_p(\Omega)$ iff*

$$\int\limits_\Omega |f(x)|^p \; d\Omega \;\; < \infty.$$

The associated norm is defined by:

$$\|f\|_p := \big(\int\limits_\Omega |f(x)|^p \; d\Omega \big)^{\frac{1}{p}}.$$

A more general definiton can be found in [3], Section 1.13. There, we also find that the definitions above naturally and readily extend to the case of the spaces $L_p(\partial\Omega)$, $1 \leq p < \infty$ where $\Omega \in C^{0,1}$ and the measures is the standard Hausdorff measure on $\partial\Omega$, for a precise description with the help of the local boundary descriptions in Definition 2.1.26 we refer to [62] and [1], Sections 5.34 and 5.35.
We will now introduce the space $L_\infty(\Omega)$:

Definition 2.1.28. *Let $f : \Omega \mapsto \mathbb{R}$ be a function measurable with respect to the standard d-dimensional Lebesgue measure. $f \in L_\infty(\Omega)$ iff*

$$\sup_{x \in \Omega \setminus N} |f(x)| < \infty,$$

where N is a null set with respect to the Lebesgue measure.
The norm is defined and denoted by

$$\|f\|_{L_\infty(\Omega)} := \inf \{\alpha > 0 \; : \; | \{x \in \Omega \; : \; |f(x)| > \alpha\} \, | = 0\}$$

$| \cdot |$ denotes the Lebesgue measure of the described set. The expression on the right is called the essential supremum.

All $L_p(\Omega)$ spaces are Banach spaces, the spaces $L_p(\Omega)$ with $1 < p < \infty$ are reflexive, and the space $L_2(\Omega)$ is a Hilbert space with scalar product

$$(f,g) := \int\limits_{\Omega} f(x)g(x)\,dx \quad f,g \in L_2(\Omega).$$

The analogous results are valid for the spaces $L_p(\partial\Omega)$.

Let us now turn to the Sobolev spaces $W_p^k(\Omega)$. First we need to define the notion of weak differentiability. The following definition is taken from [36], Section 7.3.

Definition 2.1.29. *Let $f \in L_{1,loc}(\Omega)$, i.e. $f \in L_1(K)$ for all compact subsets $K \subset \Omega$ with* $\operatorname{dist}(\partial\Omega, K) > 0$*. The weak derivative of f of order $|\alpha|$, $D^\alpha f$, is a function $g \in L_{1,loc}(\Omega)$ fulfilling*

$$\int\limits_{\Omega} f(x)D^\alpha \phi(x)\,dx = (-1)^{|\alpha|} \int\limits_{\Omega} g(x)\phi(x)\,dx, \ \ \forall \phi \in C_0^\infty(\Omega).$$

Having clarified the notion of weak differentiability, we can now define the spaces $W_p^k(\Omega)$:

Definition 2.1.30. *The space $W_p^k(\Omega)$, $k \in \mathbb{N} \cup \{0\}$, $1 \leq p \leq \infty$, consists of all measurable functions f whose weak derivatives of order $|\alpha|$, $D^\alpha f$, belong to $L_p(\Omega)$ for all $0 \leq |\alpha| \leq k$. The associated norm in case $1 \leq p < \infty$ is given by:*

$$\|f\|_{W_p^k(\Omega)} := \big(\sum_{0 \leq |\alpha| \leq k} \|D^\alpha f\|_p^p \big)^{\frac{1}{p}}.$$

The corresponding semi-norm for $1 \leq p < \infty$is defined by

$$|f|_{W_p^k(\Omega)} := \big(\sum_{|\alpha| = k} \|D^\alpha f\|_p^p \big)^{\frac{1}{p}}.$$

In case $p = \infty$ the norm on $W_\infty^k(\Omega)$ is defined by:

$$\|f\|_{W_\infty^k(\Omega)} := \max_{0 \leq |\alpha| \leq k} \|D^\alpha f\|_{L_\infty(\Omega)}$$

For a detailed introduction to the notion of weak derivatives, we refer to [3], Section 1.25. The spaces $W_p^k(\Omega)$ are Banach spaces, in particular the spaces $W_2^k(\Omega) =: H^k(\Omega)$ are Hilbert spaces with scalar product:

$$(f,g)_{H^k(\Omega)} := \sum_{0 \leq |\alpha| \leq k} (D^\alpha f, D^\alpha g).$$

In the context of partial differential equations, it is crucial to be able to assign boundary values to functions $f \in W_p^k(\Omega)$, $k \geq 1$. Nominally, they do not exist, because functions belonging to $L_p(\Omega)$ or $W_p^k(\Omega)$ are strictly speaking not functions but equivalence classes of functions whose members are equal to each other up to sets of measure zero. The boundary, however, is a set of measure zero; hence, it is not clear if boundary values are actually defined for such functions. The trace theorem, though, offers a way around this dilemma, compare [62], Theorem 5.5 and Theorem 5.7.

Theorem 2.1.31 (Trace Theorem)**.** *Let Ω be a bounded domain in $\mathbb{R}^d$ of class $C^{0,1}$ and $1 < p < \infty$. Then there exists a unique linear, continuous and **surjective** mapping T, the trace operator, with*

$$T : W_p^1(\Omega) \mapsto W_p^{1-1/p}(\partial\Omega)$$

and

$$Tf = f\big|_{\partial\Omega} \quad \forall f \in W_p^1(\Omega) \cap C(\bar{\Omega}).$$

The space $W_p^{1-1/p}(\partial\Omega)$ is defined as the space of those functions $f \in L_p(\partial\Omega)$ for which the norm

$$\|f\|_{W_p^{1-1/p}(\partial\Omega)} := \Bigg(\|f\|^p_{L_p(\partial\Omega)} + \int\limits_{\partial\Omega}\int\limits_{\partial\Omega} \frac{|f(x)-f(y)|^p}{|x-y|^{d+p}} d\partial\Omega(x) d\partial\Omega(y) \Bigg)^{1/p}$$

is finite. Here, $d\partial\Omega(x)$ denotes the Hausdorff measure with respect to the x-variable.
For $p = 2$ we again use the familiar abbreviation $H^{1/2}(\partial\Omega) := W_p^{1-1/p}(\partial\Omega)$.

The definition of the norm of $W_p^{1-1/p}(\partial\Omega)$ with fractional exponent can be found in [62], Section 3.8. and Section 5. For a more detailed discussion of traces we refer to [1], Section 7. With the help of the trace theorem we can define zero boundary values for functions belonging to certain Sobolev spaces $W_p^k(\Omega)$, $k \geq 1$:

Definition 2.1.32. *Let Ω be a domain of class $C^{0,1}$. The space $\mathring{W}_p^1(\Omega)$ consists of all functions $f \in W_p^1(\Omega)$ which fulfil*

$$Tf = 0$$

where T is the trace operator form Theorem 2.1.31.
In case $p = 2$: $\mathring{W}_2^1(\Omega) =: \mathring{H}^1(\Omega)$.

To derive existence and uniqueness results for variational formulations of second order partial differential equations, where zero boundary values are given, one often has to work with the space $\mathring{H}^1(\Omega)$. In this context, it is crucial that $\mathring{H}^1(\Omega)$ is a Banach space with the semi-norm $|\cdot|_{H^1(\Omega)}$. First, though, we need a very valuable auxiliary result, see e.g. in [31] Theorem 3, Chapter 5.

Theorem 2.1.33 (Poincaré - Friedrich's inequality)**.** *Let Ω be a domain of class $C^{0,1}$ and let $f \in \mathring{W}^1_p(\Omega)$ be arbitrary. Then there exists a constant $C = C(\Omega) > 0$ such that*

$$\|f\|_{L_p(\Omega)} \leq C\, |f|_{W^1_p(\Omega)} \,.$$

As a consequence, there exist constants $c_1, c_2 > 0$ such that

$$c_1 \|f\|_{W^1_p(\Omega)} \leq |f|_{W^1_p(\Omega)} \leq c_2 \|f\|_{W^1_p(\Omega)} \ \forall f \in \mathring{W}^1_p(\Omega) \tag{2.1.12}$$

The Poincaré-Friedrich inequality leads to the following result:

Theorem 2.1.34. *Let Ω be a bounded domain of class $C^{0,1}$. The space $\mathring{W}^1_p(\Omega)$, $p \geq 1$, normed by $|\cdot|_{W^1_p(\Omega)}$ is a Banach space.*

Proof. The key to the proof is the Poincaré - Friedrich's inequality, Theorem 2.1.33: There exists a constant $C = C(\Omega)$ such that

$$\|v\|_{L_p(\Omega)} \leq C\, |v|_{W^1_p(\Omega)} \,.$$

As a consequence, the semi-norm $|\cdot|_{W^1_p(\Omega)}$ and the full norm $\|\cdot\|_{W^1_p(\Omega)}$ are equivalent on $\mathring{W}^1_p(\Omega)$, see (2.1.12), which yields all the results of the theorem above. □

At different points of the thesis we will use embedding results for Sobolev spaces. The ones necessary for this thesis are recorded in the theorem below, which can be found in [36], Theorem 7.26,

Theorem 2.1.35 (Embedding Results)**.** *Assume Ω is a bounded domain in $\mathbb{R}^d$ of class $C^{0,1}$. Then the following embedding results are valid:*

- *If $0 < k < \frac{d}{p}$, the space $W^k_p(\Omega)$ is continuously embedded in $L_{p*}(\Omega)$ with $p^* = \frac{dp}{d-kp}$ and compactly embedded in $L_q(\Omega)$ for any $q < p^*$.*
- *If $0 \leq m < k - \frac{d}{p} < m+1$, the space $W^k_p(\Omega)$ is continuously embedded in $C^{m,\alpha}(\bar{\Omega})$, $\alpha = k - \frac{d}{p} - m$ and compactly embedded in $C^{m,\beta}(\bar{\Omega})$ for any $\beta < \alpha$.*
- *If $d = 1$, then $W^1_1(a,b)$ is continuously embedded in $C[a,b]$ for any interval $(a,b) \subset \mathbb{R}$.*

Throughout this thesis it will sometimes be necessary to consider Lebesgue and Sobolev spaces of functions f with $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}^m$. We will now extend the definitions above to this more general setting in a natural, 'component-wise' way:

Definition 2.1.36 ($(L_p)^m$ and $(W^k_p)^m$-spaces)**.** *The space $L_p(\Omega, \mathbb{R}^m)$, $1 \leq p \leq \infty$, $m \in \mathbb{N}$, consists of all measurable functions $f : \Omega \to \mathbb{R}^m$ for which the norm*

$$\|f\|_{L_p(\Omega,\mathbb{R}^m)} := (\sum_{i=1}^{m} \|f_i\|^p_{L_p(\Omega)})^{1/p}$$

is finite, f_i denoting the i-th component function of f.
In particular, the space $L_2(\Omega, \mathbb{R}^m)$ is a Hilbert space with scalar product

$$(f, g)_{L_2(\Omega,\mathbb{R}^m)} := \sum_{i=1}^{m} (f_i, g_i)_{L_2(\Omega)}.$$

The space $W_p^k(\Omega, \mathbb{R}^m)$ consists of all measurable functions f with $f : \Omega \to \mathbb{R}^m$ for which the i-th component function f_i is k-times weakly differentiable and the norm

$$\|f\|_{W_p^k(\Omega,\mathbb{R}^m)} := (\sum_{i=1}^{m} \|f_i\|^p_{W_p^k(\Omega)})^{1/p}$$

is finite.

Sometimes it is also worth considering spaces for which just some weak partial derivatives exist, first and foremost the space $H(\mathrm{div}, \Omega)$ defined below, compare also [78], Definition 20.1:

Definition 2.1.37. *The space $H(\mathrm{div}, \Omega)$ defined by*

$$H(\mathrm{div}, \Omega) := \left\{ f \in L_2(\Omega, \mathbb{R}^d) \; : \; \mathrm{div}\, f = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i} \in L_2(\Omega) \right\},$$

where div f *is understood in the weak sense, i.e a function $f \in L_2(\Omega, \mathbb{R}^d)$ has a divergence* div $f \in L_2(\Omega)$, *if*

$$\int_\Omega (\mathrm{div}\, f)(x)\phi(x)\, dx = -\int_\Omega f(x) \cdot \nabla\phi(x)\, dx \;\; \forall \phi \in C_0^\infty(\Omega).$$

$H(\mathrm{div}, \Omega)$ is a Banach space with the norm

$$\|f\|_{H(\mathrm{div},\Omega)} := (\|f\|^2_{L_2(\Omega,\mathbb{R}^d)} + \|\mathrm{div} f\|^2_{L_2(\Omega)})^{1/2}.$$

2.1.7 Interpolation Spaces and Lorentz Spaces

Sometimes it is also of importance to deal with intermediate spaces, in this thesis we are solely concerned with spaces intermediate between different $L_p(\Omega)$, the Lorentz spaces. We will specify the notion of 'intermediateness' and introduce the method of interpolation of Banach spaces to obtain such intermediate spaces in this section.
For a much more general introduction to the interpolation of spaces and their many powerful applications, we refer to the books [7],[79] and Section 7 of [1].
First, we will record an embedding Theorem for $L_p(\Omega)$:

Theorem 2.1.38. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. For $1 \leq p \leq q \leq \infty$ we have*

$$L_q(\Omega) \hookrightarrow L_p(\Omega)$$

Proof. For $q/p \geq 1$ the dual exponent p' defined by

$$\frac{p}{q} + \frac{1}{p'} = 1$$

is given by $p' = \frac{p+1}{q}$. Using Hölder's inequality, we obtain for all $g \in L_q(\Omega)$.

$$\int\limits_\Omega |g|^p \leq (\int\limits_\Omega 1^{p'})^{1/p'} (\int\limits_\Omega |g|^q)^{p/q}$$

Drawing the p-th root on each side gives the desired result, since Ω is bounded. □

Following the approach of [1], Definitions 7.7, 7.8. and 7.9. and Theorem 7.10, we define the following space obtained by the *K-method* of real interpolation:

Definition 2.1.39. *Let X_0, X_1 be two Banach spaces, with $X_1 \hookrightarrow X_0$. Let $t > 0$ be fixed. The K-functional given by*

$$K(t;x) := \inf \left\{ \|x_0\|_{X_0} + t\,\|x_1\|_{X_1} \ : \ x = x_0 + x_1,\, x_0 \in X_0,\, x_1 \in X_1 \right\}$$

defines a norm on X_0 which is equivalent to $\|\cdot\|_{X_0}$.
The space

$$X = (X_0, X_1)_{\theta,q}, \ \ 0 < \theta < 1, \ 1 \leq q < \infty,$$

consists of all functions $x \in X_0 + X_1 = X_0$ such that

$$\int\limits_0^\infty (t^{-\theta} K(t;x))^q \frac{dt}{t})^{1/q} < \infty.$$

Here, $\frac{dt}{t}$ denotes the Haar measure, which is translation-invariant.
The space X is a Banach space itself with the property that it is ***intermediate*** *between X_0 and X_1 in the sense that*

$$X_1 \hookrightarrow X \hookrightarrow X_0. \tag{2.1.13}$$

A highly useful property of such intermediate spaces obtained by interpolation is the fact that they enable us to obtain estimates for the norms of linear operators acting on these spaces. The following theorem is a combination of Definition 7.22 and Theorem 7.23 in [1]:

Theorem 2.1.40. *Suppose that $f \in X_i^*$, $i = 0, 1$ and X as in Definition 2.1.39. Then $f \in X^*$, too, and we have the estimate*

$$\|f\|_{X^*} \leq \|f\|_{X_0^*}^{1-\theta} \|f\|_{X_1^*}^{\theta} . \tag{2.1.14}$$

We remark that this theorem can be extended to more general linear operators; however, in this thesis we only need to make use of this property for linear functionals.
We now intend to apply this definition to the setting of L_p spaces. For the validity of the definition we refer to [1], Theorem 7.26, and [1], Corollary 7.27.

Definition 2.1.41. *Let $1 \leq p_1 < p < p_2 \leq \infty$ and $\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$. Then the Lorentz space $L_{p,q}(\Omega)$ is defined by*

$$L_{p,q}(\Omega) = (L_{p_1}(\Omega), L_{p_2}(\Omega))_{\theta,q} ,$$

and we have the property that for $1 < p < \infty$

$$L_{p,p}(\Omega) = L_p(\Omega)$$

with equivalent norms.

Lorentz spaces can also be obtained by a more direct way, see Definition 7.25 in [1] or [66], Example 2, Chapter 2. However, as it is not significantly faster and the estimate (2.1.14) cannot be obtained in a direct way, it is more convenient to work with the interpolation spaces approach utilised here.

We can now turn to introducing the optimal control problem and presenting its relevant properties.

2.2 Optimal Control and Optimisation in Banach Spaces

2.2.1 Problem Setting and Existence Results

Let us introduce the following continuous state-constrained optimal control problem:

$$\left.\begin{array}{c}\min_{u\in\mathbb{U},y\in\mathbb{Y}}\frac{1}{2}\|y-y_d\|_{\mathbb{W}}^2+\frac{\nu}{2}\|u\|_{\mathbb{U}}^2\\ \text{s.t.}\\ \mathcal{B}[y,w]=(u,w)_{\mathbb{U}}\quad\forall w\in\mathbb{Y}\\ \text{and}\\ u\in\mathcal{U}\subset\mathbb{U}\\ y_c-y\in C\subset L_2(\Omega,\mathbb{R}^m)\end{array}\right\}\qquad(P)$$

To formulate the required properties for (P), we will often use the notation $x \lesssim y$ for real numbers x, y so as to do without constants: $x \lesssim y$ means that there exists a constant c independent of x, y and solely depending on data such as the domain $\Omega, S, y_c, y_d,$, etc. such that

$$x \leq cy.$$

For the spaces $\mathbb{Y}, \mathbb{U}$ and $\mathbb{W}$ we require the following properties:

Properties of the spaces $\mathbb{Y}, \mathbb{U}, \mathbb{W}$:

Pr1. $\mathbb{U}$, $\mathbb{Y}$ and $\mathbb{W}$ are Hilbert spaces with norms $\|\cdot\|_{\mathbb{U}}$, $\|\cdot\|_{\mathbb{Y}}$ and $\|\cdot\|_{\mathbb{W}}$ and scalar products $(\cdot,\cdot)_{\mathbb{U}}$, $(\cdot,\cdot)_{\mathbb{Y}}$ and $(\cdot,\cdot)_{\mathbb{W}}$. Each space $\mathbb{U}, \mathbb{Y}$ and $\mathbb{W}$ is a subset of a suitable L_2-space. To be more precise, for meshable (compare Definition 2.3.2) subsets $\Gamma, \Omega \subset \mathbb{R}^d$ with $\Gamma \subset \bar{\Omega}$ we assume $\mathbb{W} \hookrightarrow L_2(\Omega, \mathbb{R}^m)$, $\mathbb{Y} \hookrightarrow L_2(\Omega, \mathbb{R}^m)$ and $\mathbb{U} \hookrightarrow L_2(\Gamma)$, where $L_2(\Omega, \mathbb{R}^m)$ is equipped with its canonical norm $\|\cdot\|_{L_2(\Omega,\mathbb{R}^m)} = \|\cdot\|$ and scalar product $(\cdot,\cdot)_{L_2(\Omega,\mathbb{R}^m)} = (\cdot,\cdot)$. In addition, we demand that the squares of the all the $\mathbb{W}, \mathbb{Y}, \mathbb{U}$-norms are additive, i.e. for $\mathbb{Y}$ we have for $\omega_1, \omega_2 \subset \bar{\Omega}$ with $|\omega_1 \cap \omega_2| = 0$ that

$$\|y|_{\omega_1\cup\omega_2}\|^2 =: \|y\|_{\mathbb{Y}(\omega_1\cup\omega_2)}^2 = \|y\|_{\mathbb{Y}(\omega_1)}^2 + \|y\|_{\mathbb{Y}(\omega_2)}^2 = \|y|_{\omega_1}\|^2 + \|y|_{\omega_2}\|^2$$

and similarly for $\mathbb{W}$ and $\mathbb{U}$. Here, $y|_\gamma$ denotes the restriction of y on a subset $\gamma \subset \bar{\Omega}$. Lastly, we assume that $\mathbb{U}$ embeds into $\mathbb{Y}^*$, $\mathbb{U} \hookrightarrow \mathbb{Y}^*$ with the associate embedding operator denoted by E, and $\mathbb{Y} \hookrightarrow \mathbb{U}$.

Property guaranteeing the solvability of the equation $\mathcal{B}[u,w] = (u,w)_{\mathbb{U}}$:

Pr2. The bilinear form $\mathcal{B}$ is continuous on $\mathbb{Y} \times \mathbb{Y}$ and satisfies an $\inf - \sup$ condition

$$\inf_{z\in\mathbb{Y}}\sup_{w\in\mathbb{Y}}\frac{\mathcal{B}[z,w]}{\|z\|_{\mathbb{Y}}\|w\|_{\mathbb{Y}}} = \inf_{w\in\mathbb{Y}}\sup_{z\in\mathbb{Y}}\frac{\mathcal{B}[z,w]}{\|z\|_{\mathbb{Y}}\|w\|_{\mathbb{Y}}} = \alpha > 0.$$

This is tantamount (compare Theorem 2.1.20 and Remark 2.1.23) to the fact that for every $u \in \mathbb{U}$ there exists a unique solution $y = Su \in \mathbb{Y}$ of

$$\mathcal{B}[y, w] = (u, w)_{\mathbb{U}} \quad \forall w \in \mathbb{Y}$$

with $S \in \mathcal{L}(\mathbb{U}, \mathbb{Y})$. Thus, due to the embeddings in (Pr1), $S \in \mathcal{L}(\mathbb{U}, \mathbb{W})$ and $S \in \mathcal{L}(\mathbb{U}, L_2(\Omega, \mathbb{R}^m))$, too.

Throughout this paper, we will not distinguish between these nominally different operators. It will be clear from the context which one we refer to in the specific setting.

Properties ensuring the existence of a unique minimiser:

Pr3. $\mathcal{U}$ is a convex and closed subset of $\mathbb{U}$.

Pr4. C is a convex, closed cone in $L_2(\Omega, \mathbb{R}^m)$.

Pr5. $y_c \in L_2(\Omega, \mathbb{R}^m)$ represents the state constraint and, $y_d \in \mathbb{W}$ the desired state that one wants to reach.

Pr6. There exists $u \in \mathbb{U}$ with $u \in \mathcal{U}$ and $y_c - Su \in C$, i.e. the set of admissible functions for (P)

$$\mathbb{U}^{ad} := \{u \in \mathbb{U} : \; u \in \mathcal{U}, \; y_c - Su \in C\}$$

is non-empty.

Let us add some explanatory comments to these properties:

- (Pr1) gives a general Hilbert space setting in which to analyse an optimal control problem. A typical choice for a second order elliptic distributed optimal control problem would be $\mathbb{U} = \mathbb{W} = L_2(\Omega)$, $\mathbb{Y} = \overset{\circ}{H}{}^1(\Omega)$, compare with (MP^{ε}).

- (Pr2) is a characterisation (see Theorem 2.1.20) of the property that the partial differential equation represented in its variational formulation by the bilinear form $\mathcal{B}$ possesses a unique solution which depends continuously on the data. Thus, basically, we merely demand that the partial differential equation is well-posed.

- (Pr4) defines the cone used to formulate the state constraint. It is defined with respect to $L_2(\Omega, \mathbb{R}^m)$, which, since $\mathbb{Y} \hookrightarrow L_2(\Omega, \mathbb{R}^m)$, makes it the natural choice in the present setting where *no higher regularity* than the $\mathbb{Y}$-regularity for the solution to the state equation in (P) is assumed. Typical choices, such as $L_\infty(\Omega, \mathbb{R}^m)$, are not well-defined in this setting.

  Sometimes, though, we will still need the cone C and its polar cone to be defined with respect to different topologies of Banach spaces $Z \subset L_2(\Omega, \mathbb{R}^m)$. However, in this case, this will always be indicated in the following way:

$$C_Z := C \cap Z$$

The corresponding polar cone is then defined as

$$C_Z^- := \{\phi \in Z^* \ : \ \langle \phi, z\rangle_{Z^*,Z} \leq 0 \,\forall z \in C_Z\}\,.$$

If there is no lower case index Z as above, then C and C^- are always understood with respect to the topology in $L_2(\Omega, \mathbb{R}^m)$, compare also Definition 2.1.8.

- (Pr6) ensures that there exist feasible points for (P). If this were not fulfilled, then problem (P) would be tantamount to optimising over the empty set.

Examples fitting this setting will be given in Section 2.4.

Thanks to Property (Pr2) the problem can be transformed to a reduced formulation:

$$\min_{u\in\mathbb{U}^{ad}} f(u) \tag{2.2.1}$$

where

$$f : \mathbb{U} \ni u \mapsto \frac{1}{2}\left\|Su - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|u\right\|_{\mathbb{U}}^2 . \tag{2.2.2}$$

Naturally, one wonders if there exists a solution to this problem and, if so, whether it is unique. A crucial tool to derive existence and uniqueness results is the following theorem, which we want to apply to problem (2.2.1):

Theorem 2.2.1 (existence & uniqueness)**.** *Suppose V is a non-empty, convex and closed subset of the Hilbert space H. Suppose further that $g : H \to \mathbb{R}$ is a weakly lower semicontinuous, strictly convex function which is bounded from below and radially unbounded.*
Let the following optimisation problem be given:

$$\min_{u\in V} g(u) \tag{2.2.3}$$

Then there exists a unique solution to (2.2.3)*, denoted by $\bar{u}$.*
If, in addition, g is Fréchet-differentiable on an open subset O containing V, then $\bar{u}$ is the solution to (2.2.3) *iff*

$$(\nabla g(\bar{u}), u - \bar{u})_H \geq 0 \ \ \forall u \in V \tag{2.2.4}$$

where $\nabla g(\bar{u})$ denotes the Riesz-representative of the Fréchet derivative $g'(\bar{u})$.

Proof. Since g is bounded from below, there exists $j = \inf\limits_{u\in V} g(u)$ and $\{u_k\} \subset V$ with $g(u_k) \to j$ as $k \to \infty$ For $\left\|u_k\right\|_H \to \infty$, we would obtain

$$g(u_k) \to \infty,\ k \to \infty$$

since g is radially unbounded. That means that the infimal sequence $\{u_k\}$ is bounded. Theorem 2.1.4 ensures the existence of a weakly convergent subsequence (w.l.o.g. the sequence itself), and Theorem 2.1.6 guarantees that the limit $\tilde{u}$ belongs to V. Since g is weakly lower semicontinuous, we find:

$$j = \lim_{k\to\infty} g(u_k) \geq g(\tilde{u}).$$

Hence, $g(\tilde{u}) = j$, and $\bar{u} := \tilde{u}$ solves the optimisation problem. It is the only solution, because if there were another one $\bar{v} \neq \bar{u}$, convexity of V and strict convexity of g would imply

$$j \leq g(\lambda\bar{u} + (1-\lambda)\bar{v}) < \lambda g(\bar{u}) + (1-\lambda)g(\bar{v}) = j,$$

which is obviously a contradiction.

Let us now prove (2.2.4). If $\bar{u}$ solves (2.2.3), then (recall (2.1.6)) for any $u \in V$ we have

$$0 \leq g(\bar{u} + t(u-\bar{u})) - g(\bar{u}) = (\nabla g(\bar{u}), t(u-\bar{u}))_H + o(\|t(u-\bar{u})\|_H) \ \ \forall t \in [0,1]$$

This implies in particular that

$$0 \leq t(\nabla g(\bar{u}), u-\bar{u})_H + o(\|t(u-\bar{u})\|_H) \ \ \forall t \in (0,1].$$

Dividing by t and letting $t \to 0$, we can deduce that (again recall (2.1.6))

$$0 \leq (\nabla g(\bar{u}), u-\bar{u})_H.$$

Since this is true for all $u \in V$, we can conclude

$$(\nabla g(\bar{u}), u-\bar{u})_H \geq 0 \ \ \forall u \in V.$$

Let us now prove the other inclusion.

First of all, let $u, v \in V$ be arbitrary, then we discover:

$$g(u) - g(v) \geq (\nabla g(v), u-v)_H \tag{2.2.5}$$

Let us prove this inequality: For all $t \in (0,1]$ we obtain

$$\begin{aligned} t(g(u)-g(v)) &\geq g(v+t(u-v)) - g(v) \\ &= (\nabla g(v), t(u-v))_H + o(\|t(u-v)\|_H) \end{aligned}$$

and thus in particular after dividing by t

$$g(u) - g(v) \geq (\nabla g(v), u-v)_H.$$

Suppose now that (2.2.4) is true for all $u \in V$, but $\bar{u}$ does not solve (2.2.3). As we have already demonstrated, there exists a unique solution $\tilde{u}$ to (2.2.3) and thus utilising (2.2.5), we deduce

$$0 \geq g(\tilde{u}) - g(\bar{u}) \geq (\nabla g(\bar{u}), \bar{u} - \tilde{u})_H.$$

However, because

$$(\nabla g(\bar{u}), u - \bar{u})_H \geq 0 \ \ \forall u \in V,$$

we can conclude

$$g(\tilde{u}) = g(\bar{u})$$

and due to the uniqueness of the solution $\bar{u} = \tilde{u}$. This is the desired contradiction. The proof is now complete. □

We now want to apply this abstract existence and uniqueness result to the setting of (P), deriving also necessary and sufficient first-order optimality conditions.
The ensuing theorem will provide this application:

Theorem 2.2.2 (existence, uniqueness, first-order optimality for (P))**.** *There exists a unique solution $\bar{u}$ to (P) with corresponding state $\bar{y} = S\bar{u}$. Furthermore, the following necessary and sufficient optimality condition holds*

$$(\bar{p} + \nu\bar{u}, u - \bar{u})_{\mathbb{U}} \geq 0 \quad \forall u \in \mathbb{U}^{ad}. \tag{2.2.6}$$

Here, $\bar{p} = S^*(\bar{y} - y_d)$ and S^* denotes the adjoint operator $S^* : \mathbb{W} \to \mathbb{U}$.

Proof. First of all, we observe that the functional f in (2.2.1) is Fréchet-differentiable, a proof can be found e.g. in [80], Section 2.6. It is also radially unbounded because

$$f(u) \geq \frac{\nu}{2} \|u\|_{\mathbb{U}}^2,$$

and weakly lower semicontinuous because $\|\cdot\|_H^2$ is weakly lower semicontinuous in any Hilbert space H and because S is weakly continuous, as it is a linear and continuous operator. Furthermore, as a short computation shows, it is also strictly convex. Due to Property (Pr6), the feasible set $\mathbb{U}^{ad}$ is non-empty. Consequently, we can apply Theorem 2.2.1 with $H = \mathbb{U}$, $V = \mathbb{U}^{ad}$ and $g = f$ and obtain a unique solution $\bar{u}$ to (P) with corresponding state $\bar{y} = S\bar{u}$. Thanks to (2.2.4) there holds:

$$(\nabla f(\bar{u}), u - \bar{u})_{\mathbb{U}} \geq 0 \ \ \forall u \in \mathbb{U}^{ad}$$

Let us now prove that the Riesz-representative of the Fréchet derivative is indeed $\nabla f(\bar{u}) =$

$\bar{p} + \nu\bar{u}$.
For the Fréchet-derivative f' we obtain:

$$\langle f'(\bar{u}), u - \bar{u}\rangle_{\mathbb{U}^*,\mathbb{U}} = (\bar{y} - y_d, S(u - \bar{u}))_{\mathbb{W}} + (\nu\bar{u}, u - \bar{u})_{\mathbb{U}}$$

Using the adjoint operator, we can simplify the expression above to

$$(\nabla f(\bar{u}), u - \bar{u})_{\mathbb{U}} = (S^*(\bar{y} - y_d) + \nu\bar{u}, u - \bar{u})_{\mathbb{U}}$$

Now, we can just plug in the definition of $\bar{p}$, and using the optimality condition (2.2.4), we obtain the desired result. □

The optimality condition (2.2.6) can also be interpreted as an optimality condition for the projection on the convex set $\mathbb{U}^{ad}$. This is the subject of the next lemma.

Lemma 2.2.3. *The optimality condition* (2.2.6) *is equivalent to*

$$\bar{u} = \Pi_{\mathbb{U}^{ad}}(-\frac{1}{\nu}\bar{p}),$$

where $\Pi_{\mathbb{U}^{ad}}$ *denotes the projection on the convex set* $\mathbb{U}^{ad}$.

Proof. Apply Theorem 2.1.7 to (2.2.6). □

Remark 2.2.4. *Lemma 2.2.3 reduces the optimality condition* (2.2.6) *to a fix point equation*

$$\bar{u} - F(\bar{u}) = 0 \tag{2.2.7}$$

with

$$F : \mathbb{U} \ni u \mapsto \Pi_{\mathbb{U}^{ad}}(-\frac{1}{\nu}S^*(Su - y_d)).$$

For some classes of convex sets, such as convex sets defined by box constraints of the type

$$\mathcal{U} = \{u \in \mathbb{U} \,:\, a \leq u \leq b\}\,, \quad a, b \in \mathbb{R} \cup \{-\infty, +\infty\}\,,$$

the function F is a semismooth function (compare Definition 2.1.17); thus the fix point equation (2.2.7) *can be solved by semismooth Newton methods, compare Section 2.2.3. Indeed, in many applications, $\mathcal{U}$ is precisely of the above structure. However, in the case of state constraints, this is not the case, a fact that constitutes one of the main difficulties for treating state-constrained problems numerically.*

Theorem 2.2.6 is a 'multiplier-free' formulation of a first-order necessary and sufficient optimality condition. 'Multiplier-free' stands for an approach doing without Lagrange multipliers which are used to eliminate certain constraints. In the rest of the section, we will explain how such a Lagrange multiplier ansatz would formally work for the state constraint in (P) - tacitly assuming that the control constraints in (Pr3) can be dealt with in an 'easy' manner - before giving two existence results for Lagrange multipliers. Let us start by specifying the notion of Lagrange multipliers in the present setting, compare also [56] and [58], Section 8.

Definition 2.2.5. *Suppose the following general optimisation problem in a Hilbert space H is given:*

$$\min_{u \in M,\, Lu \in K_Z} g(u) \tag{2.2.8}$$

with $M \subset H$ closed and convex and $K_Z \subset Z$ a convex cone defined w.r.t the topology of the Banach space Z, compare Definition 2.1.8. Furthermore, $g : H \to \mathbb{R}$ is Fréchet-differentiable and additionally fulfills the prerequisites for g of Theorem 2.2.1 (strictly convex, bounded from below, radially unbounded). The feasible set is given by

$$F^{ad} := \{u \in H \;:\; u \in M, Lu \in K_Z\}\,.$$

Besides, let $L : H \to Z$ be an affine, continuous mapping.
*A **Lagrange multiplier** $\bar{\mu} \in K_Z^-$ for the constraint $Lu \in K$ is an element for which we have with the unique solution $\bar{u}$ of* (2.2.8)*:*

$$\min_{u \in M} \{g(u) + \langle \bar{\mu}, Lu \rangle_{Z^*,Z}\} = g(\bar{u}) \tag{2.2.9}$$

$$(\nabla g(\bar{u}), u - \bar{u}) + \langle \bar{\mu}, L'(u - \bar{u}) \rangle_{Z^*,Z} \geq 0 \;\forall u \in M \tag{2.2.10}$$

$$\langle \bar{\mu}, L\bar{u} \rangle_{Z^*,Z} = 0. \tag{2.2.11}$$

Interestingly, any couple $(\tilde{u}, \tilde{\mu}) \in M \times K_Z^-$ fulfilling (2.2.9), (2.2.10) and (2.2.11) is optimal for (2.2.8). This is the subject of the next theorem; its proof can be found in [58], Section 8.4 Theorem 1.

Theorem 2.2.6 (Lagrange optimality condition)**.** *Suppose the setting of Definition 2.2.5 is given. Suppose further that there exists a couple $(\tilde{u}, \tilde{\mu}) \in \mathcal{U} \times K_Z^-$ fulfilling* (2.2.9)*,* (2.2.10) *and* (2.2.11)*. Then $\tilde{u} = \bar{u}$ and $\tilde{\mu}$ is a Lagrange multiplier for the constraint $Lu \in K_Z$.*

For (P) the setting of Definition 2.2.5 is reflected by the choices $M = \mathcal{U}$, $Lu = y_c - Su$ and $K_Z = C_Z$. Apart from the condition that $Z \subseteq L_2(\Omega, \mathbb{R}^m)$ we still have some leeway in the selection of Z. Since the Lagrange multiplier $\bar{\mu}$ is an element of the dual space Z^*, it would be favourable for both theoretical analysis and numerical applications that Z^* is not 'too difficult to handle'. Suppose e.g. that $Z = L_2(\Omega)$, then $Z^* = L_2(\Omega)$ and $\bar{\mu}$ is a proper function with values pointwise almost everywhere. However, if $Z = C(\bar{\Omega})$, then Z^* is a space of measures and both analytically and numerically not easy to treat.

The trouble with state constraints is that the topology of spaces such as L_2-type spaces is in general too weak to allow for the existence of Lagrange multipliers. Let us elucidate this point: A classic existence result for a Lagrange multiplier is the following, which can be found e.g. in [58] Section 8.3 Theorem 1. For extensions and additional conditions for the existence of Lagrange multipliers, we also refer to [56] and [55].

Theorem 2.2.7. *Suppose the setting of Definition 2.2.5 is given, i.e.*

$$\min_{u \in M,\, Lu \in K_Z} g(u)$$

with g Fréchet-differentiable, radially unbounded, bounded from below and strictly convex. Let the set LM be defined by:

$$LM := \{Lu \,:\, u \in M\} \subset Z$$

If

$$LM \cap \operatorname{int}(K_Z) \neq \emptyset, \tag{2.2.12}$$

where the topological interior int *is taken with respect to the topology in Z, then there exists a Lagrange multiplier $\bar{\mu} \in Z^*$ fulfilling* (2.2.9)*,* (2.2.10) *and* (2.2.11)*.*

In view of the preceding theorem it is advantageous for K_Z to have a non-empty topological interior. In the setting of (P^ε) we have $C_Z = K_Z$ and the obvious (first) choice would be $Z = L_2(\Omega, \mathbb{R}^m)$, specifically in the case $m = 1$ $Z = L_2(\Omega)$.
The trouble with choosing $Z = L_2(\Omega)$ (or even $Z = H^1(\Omega)$ for $d > 1$) is that the most important cones in applications, cones defined by pointwise almost everywhere inequality constraints of the type

$$f(x) \leq 0 \ \text{ f.a.a. } x \in \Omega,$$

have an empty interior illustrated by the example below

Example 2.2.8. *Consider the cone*

$$K = \{f \in L_2(0,1) \,:\, f(x) \geq 0 \, f.a.a. \ x \in (0,1)\}$$

in $L_2(0,1)$. Naturally, one would think that the function $f(x) \equiv 1$ is an interior point. However, the functions

$$g_n(x) = \begin{cases} 1 & x \in (\frac{1}{n}, 1) \\ -\frac{1}{n} & x \in (0, \frac{1}{n}) \end{cases}$$

are arbitrarily close to f w.r.t the $L_2(0,1)$-topology as a short computation shows. Yet, $g_n \notin K$ for all n, in fact int$(K) = \emptyset$ with respect to the topology in $L_2(0,1)$.
If we were now to consider the topology of $C[0,1]$, however, then f would be an interior point, since for any $\delta > 0$

$$\|f - g\|_{C[0,1]} \leq \delta$$

in particular implies

$$g(x) \geq 1 - \delta \;\; \forall x \in (0,1),$$

and hence, provided δ is small enough, $g \in K$. Thus, K does not have an empty interior in this topoplogy.

To circumvent this obstacle, one has to switch to 'stronger' topologies more compatible with pointwise constraints (compare Example 2.2.8) such as the topology of $C(\bar{\Omega})$ or $L_\infty(\Omega)$. The downside is that the dual spaces, as mentioned before, are very irregular, both being spaces of measures. Indeed, the space $C(\bar{\Omega})^*$ can be isometrically identified with the space $\mathcal{M}(\Omega)$, consisting of all Borel measures defined on the Borel σ-algebra on $\bar{\Omega}$, the smallest σ-algebra containing all closed subsets of $\bar{\Omega}$, see e.g. [53], Theorem 1.7.2.
Let us - despite this lack of regularity - record the theorem below, which applies the general existence result of Theorem 2.2.7 to the specific setting of problem (P).

Theorem 2.2.9. *Suppose the cone C in* (Pr4) *is defined w.r.t the toplogy of $C(\bar{\Omega}, \mathbb{R}^m)$, i.e.*

$$C = C_{C(\bar{\Omega},\mathbb{R}^m)} := \left\{ f \in C(\bar{\Omega}, \mathbb{R}^m) \; : \; f_i \leq 0, \; i \in I \right\}.$$

Suppose further that $S : \mathbb{U} \to C(\bar{\Omega}, \mathbb{R}^m)$ continuously. Besides, let there exist an element u_s such that $y_c - Su_s \in \mathrm{int}(C_{C(\bar{\Omega},\mathbb{R}^m)})$.
Then there exists $\bar{\mu} \in C^-_{C(\bar{\Omega},\mathbb{R}^m)}$ such that the unique solution $\bar{u}$ to (P) and $\bar{\mu}$ fulfil the following optimality system:

$$\begin{aligned} (S^*(S\bar{u} - y_d) + \nu\bar{u}, u - \bar{u})_{\mathbb{U}} - \langle \bar{\mu}, Su - S\bar{u} \rangle_{C(\bar{\Omega},\mathbb{R}^m)^*, C(\bar{\Omega},\mathbb{R}^m)} &\geq 0 \;\; \forall u \in \mathcal{U} \\ \langle \bar{\mu}, y_c - S\bar{u} \rangle_{C(\bar{\Omega},\mathbb{R}^m)^*, C(\bar{\Omega},\mathbb{R}^m)} &= 0 \end{aligned} \tag{2.2.13}$$

Proof. Applying Theorem 2.2.7 with $M = \mathcal{U}$, $Lu = y_c - Su$ and $g = f$ to problem (P) yields the desired result. □

Theorem 2.2.9 ensures the existence of a Lagrange multiplier under certain conditions. Taking the adjoint S^*, we can transform (2.2.13) in the following way, taking into account Lemma

2.2.3:

$$\begin{aligned} \bar{u} - \Pi_{\mathcal{U}}(-\frac{1}{\nu}(S^*(S\bar{u} - y_d - \bar{\mu}))) &= 0 \\ \langle \bar{\mu}, y_c - S\bar{u} \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} &= 0. \end{aligned}$$

If we assume in the vein of Remark 2.2.4 that the projection on $\mathcal{U}$ is 'easily' computable, the reformulation above, especially the equation

$$\bar{u} - \Pi_{\mathcal{U}}(-\frac{1}{\nu}(S^*(S\bar{u} - y_d - \bar{\mu}))) = 0,$$

offers a way to gain additional regularity for the optimal control $\bar{u}$ provided S^* possesses some smoothing property. Algorithmically, though, despite arriving at a situation where $\bar{u}$ can be expressed as an accessible projection ($\Pi_{\mathcal{U}}$ instead of $\Pi_{\mathbb{U}^{ad}}$, i.e. a pointwise nonsmooth equation (compare also Section 2.2.3)), we still face the obstacle of the possibly measure-valued multiplier, for which it is not possible to derive an equation similar to the projection equation for $\bar{u}$ above guaranteeing uniqueness and also additional regularity. That is why to apply fast optimisation methods such as Newton methods, one inevitably has to do something else. Here, we present the technique of relaxation of the state constraint coupled with a penalisation of its violation in the goal functional. The goal is to obtain unique Lagrange multipliers in $L_2(\Omega, \mathbb{R}^m)$ and an optimality system comparable to (2.2.13) that can be solved by Newton-type methods. Targeting these goals, we are stuck with the topology of $L_2(\Omega, \mathbb{R}^m)$ for the state constraint, since Lagrange multipliers naturally belong to the dual space, and, thus, constraint qualifications of the type (2.2.12), where the topological interior is used, are not conducive to obtaining existence results for Lagrange multipliers in $L_2(\Omega, \mathbb{R}^m)$ in the setting of pointwise inequality constraints, as Example 2.2.8 all too clearly shows. Fortunately, there are other constraint qualifications of which the most helpful for our purposes will be one which states that $L_2(\Omega, \mathbb{R}^m)$ Lagrange multipliers exist if the range of the constraint mapping $y_c - Su$ is rich enough, in particular, if it is surjective. If you think of S as the solution operator of the Poisson equation with an L_2-control on the right-hand side, however, it is clear that S is not surjective as a mapping from L_2 to L_2, since the solution of the PDE is at the very least in H^1. Considering a different space pairing, L_2-H^1 for instance, does not help either, since at the heart of the problem is the fact that the solution operator has a smoothing effect on the right-hand side. Nevertheless, we have not reached the end of the line, since by introducing a virtual control $v \in L_2(\Omega, \mathbb{R}^m)$ and an additional (not strictly necessary) parameter ε into the constraint mapping, i.e.

$$y_c - Su \quad \rightsquigarrow \quad y_c - Su - \varepsilon v, \tag{2.2.14}$$

one forces the constraint mapping to be surjective. As a consequence, we can apply the ensuing existence result for Lagrange multipliers, which can be found in e.g. [55], Thm 4.3.(ii), or as a corollary to a more general result in [56], Thm 4.1. As it is central to our later analyses, we will present an overview of the proof.

Theorem 2.2.10. *Suppose the following general optimisation problem in a Hilbert space H is given:*

$$\min_{u\in M,\, Lu\in K} g(u) \tag{2.2.15}$$

with $M \subset H$ closed and convex, Z a Banach space and $K_Z = K \subset Z$ a convex cone. Furthermore, $g : H \to \mathbb{R}$ is Fréchet-differentiable and additionally fulfills the prerequisites for g of Theorem 2.2.1. The feasible set is given by

$$F^{ad} := \{u \in H \;:\; u \in M, Lu \in K\}\,.$$

Besides, let $L : H \to Z$ be an affine, continuous mapping with

$$\mathrm{ran}(L(M)) = Z \tag{2.2.16}$$

where ran *denotes the range of a mapping.*
Then there exists a Lagrange multiplier $\bar{\mu} \in K^- = K_Z^-$ for the constraint $Lu \in K$, i.e. for the solution $\bar{u}$ to (2.2.15) *we have:*

$$\begin{aligned} \min_{u\in M} \{g(u) + \langle\bar{\mu}, Lu\rangle_{Z^*,Z}\} &= g(\bar{u}) \\ (\nabla g(\bar{u}), u-\bar{u})_{\mathbb{U}} + \langle\bar{\mu}, L'(u-\bar{u})\rangle_{Z^*,Z} &\geq 0 \quad \forall u \in M \\ \langle\bar{\mu}, L\bar{u}\rangle_{Z^*,Z} &= 0. \end{aligned}$$

To prove this result, we first have to harness a fundamental result by [68], Theorem 1. In this paper, the result is formulated in terms of set-valued mappings; we will, however, restrict it to our single-valued case.

Theorem 2.2.11. *Let* (2.2.15) *as in Theorem 2.2.10. Suppose further that*

$$0 \in \mathrm{int}\,\{LM - K\} \tag{2.2.17}$$

Then for any $u_0 \in F^{ad}$, there exist $\gamma = \gamma(u_0)$ and $\rho = \rho(u_0) > 0$ such that we have

$$\mathrm{dist}(u, F^{ad}) \leq \gamma\,\mathrm{dist}(Lu, K) \;\; \forall u \in M \cap B_\rho(u_0). \tag{2.2.18}$$

Observe that (2.2.18) can also be interpreted as a perturbation estimate. In the proof of Theorem 2.2.10 we will see that Robinson's famous constraint qualification (2.2.17) is actu-

ally sufficient for the existence of Lagrange multipliers highlighting how closely the topics of perturbation analysis and existence of Lagrange mutlipliers are linked. Combining this with the lack of regularity for Lagrange multipliers mentioned previously and with the unavailability of perturbation estimates of the type (2.2.18), we are starkly reminded of the serious difficulties we face when analysing discretisations because in some sense, discretisations can be interpreted as perturbations of the original problem.
We can now turn to the proof of Theorem 2.2.10

Proof of Theorem 2.2.10. The proof traces the arguments of [56].
Let $\bar{u}$ be the unique solution to (2.2.15). First of all, we introduce the following cones:

$$\begin{aligned}
C(\bar{u}) &= \{\lambda(u-\bar{u}) \,:\, \lambda \geq 0,\, u \in M\} \\
K(L\bar{u}) &= K - \lambda L\bar{u},\ \lambda \geq 0 \\
T(F^{ad},\bar{u}) &= \left\{u \in H \,:\, u = \lim_{n\to\infty} \frac{1}{t_n}(u_n - \bar{u}),\, t_n \to 0_+, u_n \in F^{ad}\right\} \\
L(F^{ad},\bar{u}) &= \left\{u \in H \,:\, u \in C(\bar{u}), L'u \in K(L\bar{u})\right\}.
\end{aligned}$$

The last two cones are called sequential tangent cone and linearising cone of M respectively. The key inclusion now is

$$L(F^{ad},\bar{u}) \subset T(F^{ad},\bar{u}).$$

To realise that, pick an arbitrary $s \in L(F^{ad},\bar{u})$. By definition $\exists \lambda_1, \lambda_2, u \in M, k \in K$ such that:

$$s = \lambda_1(u-\bar{u}) \text{ and } L's = \lambda_2(k - L\bar{u}). \tag{2.2.19}$$

Here, remember that K is a cone.
In case $\lambda_1 = 0$, we have $s = 0 \in T(F^{ad},\bar{u})$ (pick $u_n = \bar{u}$ in the definition of $T(F^{ad},\bar{u})$). Thus, we can assume $\lambda_1 > 0$. Presently, we define:

$$\tilde{u}_n := \bar{u} + t_n(u-\bar{u}),\ k_n := L(\bar{u}) + \frac{\lambda_2}{\lambda_1}t_n(k - L\bar{u}).$$

Due to convexity of M and $u, \bar{u} \in M$ $\tilde{u}_n \in M$ if $t_n \leq 1$. Likewise, convexity of K yields $k_n \in K$ provided $\frac{\lambda_2}{\lambda_1}t_n \leq 1$. Since $t_n \to 0_+$, we can thus pick n large enough such that $\tilde{u}_n \in M$ and $k_n \in K$.
Now pick n large enough such that $\tilde{u}_n \in B_\rho(\bar{u})$, compare (2.2.18). Then, with the help of (2.2.18), we deduce

$$\operatorname{dist}(\tilde{u}_n, F^{ad}) \leq \gamma \operatorname{dist}(L\tilde{u}_n, K).$$

Since $k_n \in K$, we can estimate in the following fashion:

$$\begin{aligned}\operatorname{dist}(L\tilde{u}_n, K) &\leq \|L\bar{u} + L(t_n(u - \bar{u})) - k_n\|_Z \\ &= \left\| L\bar{u} + \frac{t_n}{\lambda_1} L's - k_n \right\|_Z \\ &= \|L\bar{u} + (k_n - L\bar{u}) - k_n\|_Z \\ &= 0.\end{aligned}$$

Hence $\tilde{u}_n \in F^{ad}$. Defining $\tau_n := \frac{t_n}{\lambda_1} \to 0_+$, we obtain

$$\lim_{n\to\infty} \frac{1}{\tau_n}(\tilde{u}_n - \bar{u}) = s \in T(F^{ad}, \bar{u})$$

Thus

$$L(F^{ad}, \bar{u}) \subset T(F^{ad}, \bar{u}). \tag{2.2.20}$$

Now, due to optimality of $\bar{u}$ and Fréchet differentiability of g, we have

$$(\nabla g(\bar{u}), s)_H \geq 0 \;\; \forall s \in T(F^{ad}, \bar{u}). \tag{2.2.21}$$

(2.2.20) then yields

$$(\nabla g(\bar{u}), s)_H \geq 0 \;\; \forall s \in L(F^{ad}, \bar{u}).$$

We now define the convex cone $Q \subset Z \times \mathbb{R}$ by

$$Q := \left\{(L's - y, (\nabla g(\bar{u}), s)_H + \alpha) \;:\; s \in C(\bar{u}),\, y \in K(L\bar{u}),\, \alpha \geq 0\right\}.$$

Surjectivity of L yields surjectivity of L', since L is affine. The open mapping theorem, Theorem 2.1.10, then ensures the existence of a suitable γ such that

$$\{(z, \alpha) \;:\; z \in B_\gamma(0),\, \alpha \geq \max\{(\nabla g(\bar{u}), s)_H \;:\; s \in \{C - \bar{u}\} \cap B_1(0)\}\} \subset Q. \tag{2.2.22}$$

Why does this hold? First of all, we observe that thanks to surjectivity of L and thus L' as a mapping $L : M \to Z$, respectively $L' : M \to Z$, we know that:

$$\forall z \in Z,\; \exists u \in M,\; \lambda \geq 0 \text{ s.t. } \lambda L'u = z + \lambda L\bar{u}$$

Recalling the definition of $C(\bar{u})$, we can deduce from the equation above that

$$L'(C - \bar{u}) = L'C(\bar{u}) = Z.$$

In particular, $L'(C-\bar{u})$ contains an open ball denoted by $B_\gamma(0)$. Combining this with the definition of Q and (2.2.22), we deduce

$$B_\gamma(0) \times \{\alpha\} \subset Q \;\; \forall \alpha \geq \max\left\{(\nabla g(\bar{u}), s)_H \; : s \in \{C - \bar{u}\} \cap B_1(0)\right\}.$$

This means that $\mathrm{int}(Q)$ is non-empty.
The origin $(0,0)$ is a boundary point of Q due to (2.2.21) and $0 \in T(F^{ad}, \bar{u})$. Theorem 2.1.2 ensures the existence of a non-trivial hyperplane $(y^*, \beta) \in Z^* \times \mathbb{R}$ supporting Q at $(0,0)$, i.e.

$$-\langle y^*, L's - y\rangle_{Z^*,Z} + \beta((\nabla g(\bar{u}), s)_H + \alpha) \geq 0 \; \forall s \in C(\bar{u}),\, y \in K(L\bar{u}),\, \alpha \geq 0.$$

Inserting $s = 0, \alpha = 0$, we observe

$$\langle y^*, y\rangle_{Z^*,Z} \geq 0 \; \forall y \in K(L\bar{u})$$

Recalling the definition of $K(L\bar{u})$, we discern that

$$\langle y^*, k - \lambda L\bar{u}\rangle_{Z^*,Z} \geq 0 \; \forall \lambda \geq 0.$$

Hence, inserting $\lambda = 0$ and in another step $L\bar{u} = k, \lambda = 1 \pm \varepsilon, \varepsilon > 0$, we find

$$y^* \in -K^- \text{ and } \langle y^*, L\bar{u}\rangle_{Z^*,Z} = 0. \tag{2.2.23}$$

Besides, $\beta > 0$ because if β were zero, affine linearity of L and the definition of $C(\bar{u})$ and $K(L\bar{u})$ would yield

$$-\langle y^*, Lu - L\bar{u} + \lambda L\bar{u}\rangle_{Z^*,Z} \geq 0, \; \forall u \in M, \; \lambda \geq 0.$$

Inserting $\lambda = 1$, we would be able to deduce that

$$-\langle y^*, Lu\rangle_{Z^*,Z} = 0 \; \forall u \in M$$

which immediately results in $y^* = 0$, since L is surjective. This is the desired contradiction, thus $\beta > 0$.
Setting $\alpha, y = 0$ and $K^- \ni \bar{\mu} := -\frac{1}{\beta} y^*$ and recalling (2.2.23), we obtain

$$\begin{aligned}(\nabla g(\bar{u}), u - \bar{u})_H + \langle \bar{\mu}, L'(u - \bar{u})\rangle_{Z^*,Z} \geq 0 \; \forall u \in M,\; y \\ \langle \bar{\mu}, L\bar{u}\rangle_{Z^*,Z} = 0.\end{aligned}$$

These are the properties (2.2.10) and (2.2.11) formulated in Definition 2.2.5. Property (2.2.9) follows by straightforward duality arguments. □

We will now apply this valuable result to problem (P), relaxing the state constraint in the way indicated by (2.2.14) and penalising the new control v in the functional. This regularisation technique will be presented in the next section.

2.2.2 The Regularisation Approach

Let us introduce the continuous regularised problem:

$$\left.\begin{aligned} \min_{u\in\mathbb{U},y\in\mathbb{Y},v\in L_2(\Omega,\mathbb{R}^m)} \frac{1}{2}\,\|y-y_d\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon}\,\|v\|_{L_2(\Omega,\mathbb{R}^m)}^2 \\ \text{s.t.} \\ \mathcal{B}[y,w] = (u,w)_{\mathbb{U}} \quad \forall w\in\mathbb{Y} \\ \text{and} \\ u\in\mathcal{U} \\ y_c - y - \varepsilon v \in C \end{aligned}\right\} \qquad (P^\varepsilon)$$

Using the solution operator S for the state equation in (Pr2), we can again transfer the problem above to a reduced formulation. Defining

$$f^\varepsilon : \mathbb{U}\times L_2(\Omega,\mathbb{R}^m) \ni (u,v) \mapsto \frac{1}{2}\,\|Su-y_d\|_{\mathbb{W}}^2 + \frac{\nu}{2}\,\|u\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon}\,\|v\|^2$$

and the admissible set

$$\mathbb{U}^{\varepsilon,ad} = \left\{(u,v)\in\mathcal{U}\times L_2(\Omega,\mathbb{R}^m) \;:\; y_c - Su - \varepsilon v \in C\right\},$$

we can lay out the reduced formulation of (P^ε)

$$\min_{(u,v)\in\mathbb{U}^{\varepsilon,ad}} f^\varepsilon(u,v). \tag{2.2.24}$$

The first question we want to answer is whether the existence and uniqueness results of the original problem (P) are inherited by the regularised problem. The good news is that they are because Theorem 2.2.1 can be readily applied to the new setting of the regularised problem.

Theorem 2.2.12 (existence & uniquenss for the regularised problem)**.** *For every fixed $\varepsilon > 0$, there exists a unique solution $(\bar{u}^\varepsilon,\bar{v}^\varepsilon)\in\mathbb{U}^{\varepsilon,ad}$ such that the following necessary and sufficient optimality condition is fulfilled*

$$(\bar{p}^\varepsilon + \nu\bar{u}^\varepsilon, u-\bar{u}^\varepsilon)_{\mathbb{U}} + \frac{1}{\varepsilon}(\bar{v}^\varepsilon, v-\bar{v}^\varepsilon) \geq 0 \quad \forall (u,v)\in\mathbb{U}^{\varepsilon,ad}, \tag{2.2.25}$$

with $\bar{p}^\varepsilon = S^*(\bar{y}^\varepsilon - y_d)$ being the adjoint state.

Proof. As in the proof of Theorem 2.2.2 we want to apply the general result of Theorem 2.2.1

to the present, specific setting.

First of all, let us remark that f^ε is a Fréchet-differentiable, radially unbounded, strictly convex and continuous functional on the Hilbert space $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$ endowed with the norm

$$\|(u,v)\|_{\mathbb{U}\times L_2(\Omega,\mathbb{R}^m)} := (\|u\|_{\mathbb{U}}^2 + \|v\|^2)^{1/2}. \tag{2.2.26}$$

For the Fréchet-differentiability property we refer to Section 2.6 in [80]. Besides the set $\mathbb{U}^{ad,\varepsilon}$ is non-empty, since $(\bar{u}, 0) \in \mathbb{U}^{\varepsilon,ad}$. Thus, we can apply Theorem 2.2.1 with $H = \mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$, $V = \mathbb{U}^{\varepsilon,ad}$ and $g = f^\varepsilon$ to deduce that there exists a unique solution couple $\bar{u}^\varepsilon, \bar{v}^\varepsilon$ with corresponding state $S\bar{u}^\varepsilon = \bar{y}^\varepsilon$.

To derive the first-order necessary and sufficient condition, we observe that - similar to the proof of Theorem 2.2.2 - $\nabla f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$ can be expressed as

$$\nabla f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) = \begin{bmatrix} S^*(\bar{y}^\varepsilon - y_d) + \nu\bar{u}^\varepsilon & \frac{1}{\varepsilon}\bar{v}^\varepsilon \end{bmatrix} \in \mathbb{U} \times L_2(\Omega, \mathbb{R}^m).$$

Consequently, inserting the definition of $\bar{p}^\varepsilon$ and employing (2.2.4), we obtain (2.2.25). The fact that (2.2.25) is sufficient follows straight from Theorem 2.2.1. □

The key reason for the regularisation and the introduction of the new control v was to obtain existence of Lagrange multipliers in more favourable spaces, here $L_2(\Omega, \mathbb{R}^m)$. The next theorem shows that these troubles have not been in vain. In fact, we will also gain uniqueness of the multiplier, which will turn out to be very helpful for employing efficient optimisation algorithms as we will explain later. To formulate this theorem, let us, however, first introduce the Lagrangian and the dual and primal problem for (P^ε).

First, let us define the Lagrangian $\mathcal{L}^\varepsilon : \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times L_2(\Omega, \mathbb{R}^m) \to \mathbb{R}$:

$$\mathcal{L}^\varepsilon(u, v, \theta) := f^\varepsilon(u, v) + (\theta, y_c - Su - \varepsilon v).$$

With the Lagrangian we can define the primal problem

$$\inf_{(u,v)\in\mathcal{U}\times L_2(\Omega,\mathbb{R}^m)} \sup_{\theta\in C^-} \mathcal{L}^\varepsilon(u, v, \theta) \tag{2.2.27}$$

and the dual problem

$$\sup_{\theta\in C^-} \inf_{(u,v)\in\mathcal{U}\times L_2(\Omega,\mathbb{R}^m)} \mathcal{L}^\varepsilon(u, v, \theta). \tag{2.2.28}$$

Recall that in our notation C^- signifies that the polar cone is taken with respect to the topology induced by the standard $L_2(\Omega, \mathbb{R}^m)$-norm.

Now, we can turn to the question of existence of a Lagrange multiplier:

Theorem 2.2.13 (existence & uniqueness of Lagrange multiplier)**.** *Let (P^ε) for a fixed $\varepsilon > 0$ be given. Then there exists a unique element $\bar{\theta}^\varepsilon \in C^-$ such that $\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon$ solve the following*

Karush-Kuhn-Tucker (KKT) system:

$$\begin{aligned}(S^*(\bar{y}^\varepsilon - y_d) + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon)_{\mathbb{U}} - (\bar{\theta}^\varepsilon, Su - \bar{y}^\varepsilon) &\geq 0 \quad \forall u \in \mathcal{U}\\ -\varepsilon^2\bar{\theta}^\varepsilon + \bar{v}^\varepsilon &= 0 \\ (\bar{\theta}^\varepsilon, \bar{y}^\varepsilon - y_c + \varepsilon\bar{v}^\varepsilon) &= 0.\end{aligned} \tag{2.2.29}$$

Furthermore, $(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon)$ *solve the dual and primal problem, for which the following equality holds:*

$$f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon) = \mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon). \tag{2.2.30}$$

Besides, if $(\tilde{u}, \tilde{v}, \tilde{\theta}) \in \mathcal{U} \times L_2(\Omega, \mathbb{R}^m) \times L_2(\Omega, \mathbb{R}^m)$ *solves* (2.2.29) *then*

$$(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon) = (\tilde{u}, \tilde{v}, \tilde{\theta})$$

Proof. First of all, the existence of a Lagrange multiplier $\bar{\theta}^\varepsilon \in C^-$ is guaranteed by the fact that the constraint mapping $M^\varepsilon : \mathcal{U} \times L_2(\Omega, \mathbb{R}^m) \to L_2(\Omega, \mathbb{R}^m)$ with

$$M^\varepsilon(u, v) := y_c - \varepsilon v - Su$$

is surjective, after all, in this setting, we can apply Theorem 2.2.10 with $H = \mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$, $L = M^\varepsilon$, $Z = L_2(\Omega, \mathbb{R}^m)$ and $K_Z = C$. The KKT system (2.2.29) then readily follows, compare also (2.2.11) and (2.2.10). The fact that the Lagrange multiplier $\bar{\theta}^\varepsilon$ is unique is a consequence of the equation

$$\varepsilon^2\theta^\varepsilon = \bar{v}^\varepsilon$$

in (2.2.29) and the fact that $\bar{v}^\varepsilon$ is unique, see Theorem 2.2.12.
The definition of a Lagrange multiplier Definition 2.2.5 entails the solvability of the dual and primal problems and the fact that their solutions are indeed the triples $(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon)$. (2.2.30) then follows from (2.2.11). □

Naturally, analytically, it is more convenient to work with Lagrange multipliers which are proper functions and not just measures. The uniqueness result of the Lagrange multiplier will also be very helpful. However, at this stage, we want to focus more on the numerical aspect of the results of Theorem 2.2.13. Under certain conditions on the control constraints represented by the set $\mathcal{U}$, it is possible to transform (2.2.29) into a non-linear equation for which generalised Newton methods such as the Semismooth Newton Method are applicable. The fact that this method can be applied in a function space setting opens up the possibility (and indeed, this is observed, see e.g. [54], [41] and [43]) of convergence of the Newton method independent of the mesh used later for defining the discrete spaces. This is especially important because for algorithms tackling the discrete problem with iterative methods from finite-dimensional nonlinear optimisation it was observed that the number of iterations increases linearly with the degrees of freedom, see e.g. [9]. Given that in 3D problems 500,000 degrees of freedom are not unusual, it is clear that those methods are not practically applicable.

Another important property of the Semismooth Newton Method is the fact that - under certain conditions - it generates a sequence of iterates that converge (locally) q-superlinearly to the solution of the optimsation problem, cf. Definition 2.1.18. This is of course a very desirable property for an optimisation algorithm, since essentially, it makes it very efficient in the sense that the iterates converge fast to the true solution and not too many steps are needed to get a 'good' approximation of the true solution. After all, as we will show in the course of this thesis, every step of an optimisation algorithm involves solving two partial differential equations - something which one does not want to have repeated too often.

We will elucidate some of these aspects in the next section, which is intended to give an explanatory overview without delving too much into the mathematical details. Detailed proofs will either be omitted or postponed; we would simply like to give the reader an idea as to how Semismooth Newton Methods can be applied in a setting with regularised state constraints with the help of a simple model problem.

2.2.3 The Semismooth Newton Method

In this section, we will present an application of the semismooth Newton method to an optimal control problem.

Let us first, for the sake of simplicity and just for this section, assume that $\Omega \subset \mathbb{R}^2$, $\mathbb{U} = L_2(\Omega)$ and the set $\mathcal{U}$ is given by

$$\mathcal{U} = \{u \in L_2(\Omega) \ : \ a \leq u < \infty\}$$

with a real number a. Besides, let $\mathbb{W} = L_2(\Omega)$ and $\mathbb{Y} = \mathring{H}^1(\Omega)$.

Let us define the following linear-quadratic elliptic optimal control model problem with a regularised state constraint:

$$\left.\begin{array}{c} \min\limits_{u\in\mathbb{U},y\in\mathbb{Y}} \frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L_2(\Omega)} + \frac{1}{2\varepsilon}\|\bar{v}^\varepsilon\|^2 \\ \text{s.t.} \\ \int\limits_\Omega \nabla y \cdot \nabla w \, d\Omega = \int\limits_\Omega uw \, d\Omega. \ \forall w \in \mathring{H}^1(\Omega) \\ \text{and} \\ u \in \mathcal{U} \subset \mathbb{U} \\ y_c - y - \varepsilon v \leq 0 \ \text{ a.e. in } \Omega \end{array}\right\} \qquad (MP^\varepsilon)$$

We assume that $y_c \in H^1(\Omega)$ and $\Omega \in C^{0,1}$.

Here, the state equation (compare (Pr2)) simply is the variational formulation of the Poisson equation with Dirichlet boundary data. In this particular case, the adjoint operator S^* :

$L_2(\Omega) \ni z \mapsto p \in L_2(\Omega)$ can be represented by the solution operator to

$$\int_\Omega \nabla p \cdot \nabla w \, d\Omega = \int_\Omega zw \, d\Omega. \ \forall w \in \mathring{H}^1(\Omega). \tag{2.2.31}$$

We will postpone the proof until Chapter 4 of the thesis. Instead, let us point out that S is self-adjoint and, more crucially, as we will later explain, S maps into better spaces, i.e. for every $w \in L_2(\Omega)$ $p = S^*w$ is actually an element of (at least) $\mathring{H}^1(\Omega)$.

Bearing this in mind, we can now combine Theorems 2.2.1 and 2.2.12 as well as the existence result for Lagrange multipliers Theorems 2.2.10 and 2.2.13 to obtain the following Karush Kuhn Tucker system for the solution couple $(\bar{u}_m^\varepsilon, \bar{v}_m^\varepsilon)$ with corresponding state $\bar{y}_m^\varepsilon = S\bar{u}_m^\varepsilon$ and unique Lagrange mutliplier $\bar{\theta}_m^\varepsilon \geq 0$ a.e. in Ω:

$$\begin{aligned} (S^*(\bar{y}_m^\varepsilon - y_d) + \nu \bar{u}_m^\varepsilon, u - \bar{u}_m^\varepsilon)_{L_2(\Omega)} - (\bar{\theta}_m^\varepsilon, Su - \bar{y}_m^\varepsilon)_{L_2(\Omega)} &\geq 0 \quad \forall u \in \mathcal{U} \\ -\varepsilon^2 \bar{\theta}_m^\varepsilon + \bar{v}_m^\varepsilon &= 0 \\ (\bar{\theta}_m^\varepsilon, \bar{y}_m^\varepsilon - y_c + \varepsilon \bar{v}_m^\varepsilon)_{L_2(\Omega)} &= 0. \end{aligned} \tag{2.2.32}$$

$\bar{\theta}_m^\varepsilon$ is an $L_2(\Omega)$-function. Hence, we can reformulate (2.2.32) by defining the adjoint state $\bar{p}_m^\varepsilon := S^*(\bar{y}^\varepsilon - y_d - \bar{\theta}_m^\varepsilon)$. The optimality condition for $\bar{u}_m^\varepsilon$ in (2.2.32)

$$(S^*(\bar{y}_m^\varepsilon - y_d) + \nu \bar{u}_m^\varepsilon, u - \bar{u}_m^\varepsilon)_{L_2(\Omega)} - (\bar{\theta}_m^\varepsilon, Su - \bar{y}_m^\varepsilon)_{L_2(\Omega)} \geq 0 \quad \forall u \in \mathcal{U}$$

can then be transformed to

$$(\bar{p}_m^\varepsilon + \nu \bar{u}_m^\varepsilon, u - \bar{u}_m^\varepsilon)_{L_2(\Omega)} \geq 0 \ \ \forall \, a \leq u \leq b.$$

The good news is that the variational inequality above can be reformulated as a non-smooth equation when we use pointwise min and max operators

$$\begin{aligned} \bar{u}_m^\varepsilon(x) &= a - \min(0, \frac{1}{\nu}\bar{p}_m^\varepsilon(x) + a) && (2.2.33) \\ &= a - \min(0, \frac{1}{\nu} S^*(S\bar{u}_m^\varepsilon - y_d) + a) \ \text{ f.a.a. } x \in \Omega. && (2.2.34) \end{aligned}$$

For $\bar{v}_m^\varepsilon$ we can deduce the following equation

$$\begin{aligned} \bar{v}_m^\varepsilon(x) = \varepsilon^2 \bar{\theta}_m^\varepsilon(x) &= -\frac{1}{\varepsilon} \min(0, \bar{y}_m^\varepsilon(x) - y_c(x)) && (2.2.35) \\ &= -\frac{1}{\varepsilon} \min(0, S\bar{u}_m^\varepsilon(x) - y_c(x)) \ \text{ f.a.a. } x \in \Omega. && (2.2.36) \end{aligned}$$

Thus, all in all, using (2.2.33) and (2.2.35), we can conclude that (2.2.32), and at the same time solving (MP^ε) is actually equivalent to finding a solution to the fix-point equation

$$\begin{bmatrix} u & v \end{bmatrix}^T - F_m(u,v) = 0 \tag{2.2.37}$$

with $F_m : L_2(\Omega) \times L_2(\Omega) \to L_2(\Omega) \times L_2(\Omega)$ and

$$F_m(u,v) = \begin{bmatrix} a - \min(0, \frac{1}{\nu} S^*(Su - y_d) + a) & -\frac{1}{\varepsilon}\min(0, Su - y_c) \end{bmatrix}^T .$$

Here, we want to stress that the reformulation of the optimality system (2.2.32) in terms of a fix-point equation with a pointwise superposition operator F_m, (2.2.37), is not possible for the optimality system (2.2.13). The key difficulty here is the fact that the lack of regularity for the multiplier $\bar{\mu}$ does not permit the transformation of (2.2.11) into a non-smooth equation with a pointwisely define operator. Furthermore, the multiplier $\bar{\mu}$ might not be unique, which would also lead to issues pertaining to the solvability of the KKT system (2.2.13). As we have shown, it is possible to circumvent these problems by investigating a regularised problem, thereby highlighting the importance of regularisation in the context of optimal control with constraints on the state.

To solve (2.2.37), one naturally wants to apply Newton-type methods. However, in this case, one is hampered by the fact that the min operator is not classically differentiable due to its kinks.

The good news is that it is still semismooth. To prove this result, it is crucial that S and S^* map to 'better' spaces, as the following theorem, which is a slight reformulation of the (more general) Theorem 2.14 in [45], clearly shows:

Theorem 2.2.14. *Let ϕ be given by*

$$\phi(u,v) = \begin{bmatrix} u & v \end{bmatrix}^T - F_m(u,v).$$

Then the operator $\phi : L_2(\Omega) \to L_2(\Omega)$ is $\partial\phi$-semismooth in the sense of Definition 2.1.17 provided

$$S^*S : L_2(\Omega) \to L_2(\Omega)$$

is Fréchet-differentiable and

$$S^*S : L_2(\Omega) \to L_p(\Omega),\ p > 2 \tag{2.2.38}$$

as well as

$$S^* : L_2(\Omega) \to L_p(\Omega),\ p > 2 \tag{2.2.39}$$

are locally Lipschitz-continuous.

At this stage, we want to quickly point out that S and S^*, the latter defined by the solution of (2.2.31), fulfil the prerequisites of Theorem 2.2.14. After all, S and S^* both map $L_2(\Omega)$-functions to $H^1(\Omega)$-functions. As a consequence of Theorem 2.1.35, $H^1(\Omega) \hookrightarrow L_p(\Omega)$ for all $1 \leq p < \infty$ (recall that we assumed that $d = 2$ in this section) and, hence, conditions (2.2.38) and (2.2.39) are always fulfilled.

Remark 2.2.15. *Conditions* (2.2.38) *and* (2.2.39) *are critical because superposition operators such as the pointwise* min *in* (2.2.33) *and* (2.2.35) *are in general not semismooth as mappings from L_q to L_q, $1 \leq q < \infty$, see e.g. Lemma 2.7 in [45] or Example 3.57 in [82]. Therefore, the smoothing done by the solution operator S and its adjoint S^* is essential.*

We can now formulate the semismooth Newton method for my model regularised problem (MP^ε). A general version of this algorithm can be found in [45], Algorithm 2.11.

Algorithm 2.2.1 Semismooth Newton Method

```
Choose $(u_0, v_0) \in L_2(\Omega) \times L_2(\Omega)$.
for $k = 0, 1, 2, ...$ do
    Choose $M_k \in \partial\phi((u_k, v_k))$
    Solve $M_k s_k = -\phi((u_k, v_k))$
    Set $(u_{k+1}, v_{k+1}) = (u_k, v_k) + s_k$
end for
```

Algorithm 2.2.1 involves the solution of an equation

$$M_k s_k = -\phi((u_k, v_k)),$$

which obviously makes it imperative that $M_k \in \partial\phi((u_k, v_k))$ be invertible. The following regularity condition, which can be found in [45], Equation 2.20, ensures just that:

$$\begin{aligned} &\exists C > 0,\ \delta > 0 \ \text{ s.t. } \left\|M^{-1}\right\|_{\mathcal{L}(L_2(\Omega),L_2(\Omega))} \leq C,\ \forall M \in \partial\phi(u, v), \\ &\qquad\qquad \forall (u, v) \in L_2(\Omega)^2 :\ \|(u, v) - (\bar{u}_m^\varepsilon, \bar{v}_m^\varepsilon)\|_{L_2(\Omega)^2} < \delta \qquad (2.2.40) \end{aligned}$$

Let us remark that Theorem 4.8, [82], provides a more accessible approach to condition (2.2.40). For more detailed information regarding this condition, we also want to refer to [83] and [81].

We now conclude this chapter by citing the central theorem below, which states that Algorithm 2.2.1 generates a q-superlinearly convergent sequence of iterates provided the assumptions of Theorem 2.2.14 and condition (2.2.40) are fulfilled.

Theorem 2.2.16. *Suppose that the Assumptions of Theorem 2.2.14 and condition* (2.2.40) *are fulfilled. Suppose further that the starting point* (u_0, v_0) *of Algorithm 2.2.1 is chosen such that*

$$\|(u_0, v_0) - (\bar{u}_m^\varepsilon, \bar{v}_m^\varepsilon)\|_{L_2(\Omega)^2} < \delta$$

where δ *is the* δ *of condition* (2.2.40)*.*
Then the sequence of iterates generated by Algorithm 2.2.1 converges q-superlinearly to the solution $(\bar{u}_m^\varepsilon, \bar{v}_m^\varepsilon)$*.*

This theorem and its proof can be found in [45], Theorem 2.12.
As yet, we have not addressed the question of globalisation of convergence, which is a very helpful property. After all, Theorem 2.2.16 requires us to start 'somewhere in the vicinity' of the true solution, possibly already quite close if $\delta > 0$ is small. Such issues have been investigated e.g. in [41], Theorem 3.2, and [47], Section 3.

2.3 Adaptive Finite Element Method

In this section, we will give a brief introduction to the adaptive finite element method, in short AFEM, in the context of optimal control problems. Central to the AFEM is the

$$\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}$$

cycle or loop.
To explain it, one first has to introduce the notion of 'triangulation' and 'finite element space', which is the subject of the next section.

2.3.1 Triangulations

To define finite element spaces, one first has to clarify what is meant by a triangulation because the spaces themselves are defined on triangulations. Triangulations consist of simplices which we will definie first. The defintion itself is taken from [65], Definition 5 and 6, Lemma 1, Section 3.2.

Definition 2.3.1 (Simplex and Subsimplex)**.** *Let* $d \in \mathbb{N}$*. A subset* T *of* $\mathbb{R}^d$ *is an* n*-simplex in* $\mathbb{R}^d$ *if there exist* $n+1$ *points* $z_0, z_1, ..., z_n \in \mathbb{R}^d$ *such that*

$$T = conv\ hull\{z_0, ..., z_n\} = \left\{ \sum_{i=0}^{n} \ : \ \lambda_i \geq 0 \, \forall i, \ \sum_{i=0}^{n} \lambda_i = 1 \right\}$$

and $z_1 - z_0, ..., z_n - z_0$ *are linearly independent vectors in* $\mathbb{R}^d$*. Individual points are* 0*-simplices.*

*A subset $T^{'}$ of T is a (proper) k-*subsimplex *of T if $T^{'}$ is a k-simplex such that*

$$T^{'} = conv\ hull\left\{ z_0^{'}, ..., z_k^{'} \right\} \subset \partial T$$

with $k < n$ and $z_0^{'}, ..., z_k^{'} \in \{z_0, ..., z_n\}$.
Additionally, the following quantities define the diameter, inball diameter *and* scaled volume *of T:*

$$\begin{aligned} d_T &:= \sup\{|x-y| \ : \ x, y \in T\} \\ r_T &:= \sup\{2r \ : \ B_r \subset T \ \ is\ a\ ball\ of\ radius\ r\} \\ h_T &:= |T|^{1/d}. \end{aligned}$$

The shape coefficient *is the ratio of the diameter and the inball diameter:*

$$\sigma_T := \frac{d_T}{r_T}.$$

Having settled the question of what constitutes a simplex, we can now proceed to define a triangulation of a domain Υ, compare [65], Definition 7, Section 3.2.

Definition 2.3.2 (Triangulation)**.** *Let $\Upsilon \subset \mathbb{R}^d$ be a bounded set. A finite set $\mathcal{T}$ of d-simplices in $\mathbb{R}^d$ with*

$$\bar{\Upsilon} = \bigcup_{T \in \mathcal{T}} T \quad and \quad |\Upsilon| = \sum_{T \in \mathcal{T}} |T|$$

*is called a triangulation of Υ. The set of 0-simplices of a triangulation $\mathcal{T}$ are called **nodes**, the set of $d-1$-simplices **faces**. A set Υ which admits such a triangulation is called **meshable**.*
A triangulation $\mathcal{T}$ is conforming *if it satisfies the following property: If any two simplices $T_1, T_2 \in \mathcal{T}$ have a non-empty intersection S, then S is k-subsimplex of both T_1 and T_2 with $k \in \{0, ..., d\}$.*
A sequence of triangulations $\{\mathcal{T}_k\}_{k \geq 0}$ is shape-regular *if*

$$\sup_{k \in \mathbb{N}} \sup_{T \in \mathcal{T}_k} \sigma_T \leq C.$$

Every triangulation $\mathcal{T}$ of a domain Υ is associated with a piecewise constant mesh-size function $h_{\mathcal{T}} : \Upsilon \to \mathbb{R}^+$, $h_{\mathcal{T}} \in L_\infty(\Upsilon)$ defined by:

$$h_{\mathcal{T}}(x) := |T|^{1/d} \quad x \in \mathrm{int}(T). \tag{2.3.1}$$

A restrictive condition on a triangulation is the quasi-uniformity condition, which, in essence, demands that every element of a triangulation be roughly about the same size. We will specify this notion below because we will need it for comparison purposes throughout this thesis:

Definition 2.3.3 (Quasi-Uniform Triangulations)**.** *Let Υ and $\mathcal{T}_k$ be as in Definition 2.3.2. The sequence of triangulations $\mathcal{T}_k$ is called* quasi-uniform *if there exists a constant independent of k such that*

$$\max_{T\in\mathcal{T}_k} h_{\mathcal{T}_k} \lesssim \min_{T\in\mathcal{T}_k} h_{\mathcal{T}_k}$$

In this case, there exists $h_k \in \mathbb{R}^+$ such that

$$h_k \lesssim h_{\mathcal{T}_k}(x) \lesssim h_k \;\; f.a.a. \;\, x \in \Omega,$$

where the hidden constants are independent of k. Hence, the local mesh-size functions $h_{\mathcal{T}_k}$ is equivalent to a global mesh-size parameter h_k.

From the definition above, it is clear that quasi-uniformity demands that the local mesh-size function be equivalent to a global mesh-size parameter. That makes it impossible to locally refine in some area of the domain Υ without (or just 'moderately') changing the mesh in other areas where the error may already be quite small. In a quasi-uniform setting, refinement is always global, potentially leading to a case where a very fine mesh-size is 'wasted' in parts of the domain where the numerical approximation to the true solution had already been quite accurate on a coarser grid. This is numerically a severe disadvantage and explains why in an adaptive setting like ours one always does without such a condition.

We have now collected several geometric properties of triangulations. All these are important for rigourously defining finite element spaces, the task we will turn to now.

2.3.2 Finite Element Spaces

Finite element (FE) spaces are essentially spaces of piecewise polynomial functions. Following [19], Chapter 2, we define a finite element space $\mathbb{V}(\mathcal{T}) \subset \mathbb{Y}$ in the ensuing way:

Definition 2.3.4 (Conforming FE Space)**.** *Let $\mathbb{V}$ be a Banach space, $\mathcal{T}$ a conforming triangulation of a set $\Upsilon \subset \mathbb{R}^d$ and $m \geq 0$. The FE space $\mathbf{FES}(\mathcal{T}, \mathbb{P}_m, \mathbb{V})$ equipped - unless explicitly stated otherwise - with the norm $\|\cdot\|_{\mathbb{V}}$ is then defined by*

$$\mathbf{FES}(\mathcal{T}, \mathbb{P}_m, \mathbb{V}) := \{V \in \mathbb{V} \;:\; V|_T \in \mathbb{P}_m(T) \; \forall T \in \mathcal{T}\}.$$

Here, $\mathbb{P}_m(T)$ denotes the space of all polynomials up to degree m on a single element T, i.e.

$$\mathbb{P}_m(T) := \left\{ p \in \mathbb{V} \;:\; p(x_1, x_2, ...x_d)|_T = \sum_{|\beta|\leq m} \alpha_\beta x_1^{\beta_1} x_2^{\beta_2} ... x_d^{\beta_d} \right\},$$

with multi-index $\beta = (\beta_1, \beta_2, ..., \beta_d)$ *(*β *is a d-tupel of numbers in* $\mathbb{N} \cup \{0\}$*).*

We now give two examples of FE spaces which are the most important applications of the theoretical framework of this thesis. They are the piecewise constant and piecewise linear ones:

Example 2.3.5 ($\mathbb{P}_0$ and $\mathbb{P}_1$ FE spaces)**.** *Let* Ω *be a bounded domain in* $\mathbb{R}^d$ *and* $\mathcal{T}$ *a triangulation of* Ω*.*
The space

$$\mathbf{FES}(\mathcal{T}, \mathbb{P}_0, L_2(\Omega)) := \{V \in L_2(\Omega) \ : \ V|_T \in \mathbb{P}_0 \, \forall T \in \mathcal{T}\}$$

consists of all functions $V \in L_2(\Omega)$ *which are constant on every element* $T \in \mathcal{T}$*.*
The space

$$\mathbf{FES}(\mathcal{T}, \mathbb{P}_1, H^1(\Omega)) := \left\{V \in H^1(\Omega) \ : \ V|_T \in \mathbb{P}_1 \, \forall T \in \mathcal{T}\right\}$$

contains all $H^1(\Omega)$*-functions which are linear on every element* $T \in \mathcal{T}$*. Due to Theorem 2.1.1., [19], this space is identical to the finite element space defined by*

$$\mathbf{FES}(\mathcal{T}, \mathbb{P}_1, C(\bar{\Omega})) := \left\{V \in C(\bar{\Omega}) \ : \ V|_T \in \mathbb{P}_1 \, \forall T \in \mathcal{T}\right\}.$$

We now want to apply this general setting to the optimal control problem (P), specifically, we want to discretise it. This will be the subject of the next section.

2.3.3 Discretisation

To define a finite element discretisation for (P), we first have to introduce triangulations $\mathcal{T}$ and $\mathcal{S}$ of Ω and Γ respectively.

Let $\mathcal{T}$ now be a conforming, shape-regular triangulation of the domain Ω and $\mathcal{S}$ be a conforming, shape-regular one of Γ.
To these initial triangulations $\mathcal{T}$ and $\mathcal{S}$ we assign the index $k = 0$, i.e

$$\mathcal{T} =: \mathcal{T}_0 \quad \mathcal{S} =: \mathcal{S}_0.$$

We now assume that a sequence of **conforming** and **shape-regular** $\mathcal{T}_k$ and $\mathcal{S}_k$ is generated, starting with the initial triangulation by a suitable refinement algorithm - in Section 2.3.5 we will clarify the notion of 'suitability' and give examples of methods producing such sequences of triangulations.
Having introduced the sequences of triangulations $\mathcal{T}_k$ and $\mathcal{S}_k$, we can now proceed to define the discrete counterparts to the control space $\mathbb{U}$ and state space $\mathbb{Y}$:

Definition 2.3.6 (the spaces $\mathbb{U}_k$ and $\mathbb{Y}_k$)**.** *Let* $m, n, k \in \mathbb{N} \cup \{0\}$ *and* $\mathcal{T}_k$ *and* $\mathcal{S}_k$ *be conforming and shape-regular triangulations of* Ω *and* Γ *respectively. Then* $\mathbb{U}_k$ *is either defined as the* *finite element space*

$$\mathbb{U}_k = \mathbb{U}_k(\mathcal{S}_k) := \mathbf{FES}(\mathcal{S}_k, \mathbb{P}_m, \mathbb{U}),$$

which is the ***full discretisation*** *approach, or it is the entire space* $\mathbb{U}$*, i.e.*

$$\mathbb{U} = \mathbb{U}_k,$$

which is the ***variational discretisation*** *technique.*
The space $\mathbb{Y}_k$ *is in both cases defined as*

$$\mathbb{Y}_k = \mathbb{Y}_k(\mathcal{T}_k) := \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_n, \mathbb{Y}).$$

Remark 2.3.7. *The choice* $\mathbb{U}_k = \mathbb{U}$ *in Definition 2.3.6 is the* ***variational discretisation*** *approach pioneered in [44]. Occasionally, in this thesis, we will refer to it and add some explanatory comments.*

For the state constraint we define another discrete space $\mathbb{V}_k$:

Definition 2.3.8. *The space* $\mathbb{V}_k$ *is either the space* $\mathbb{Y}_k$ ***equipped with the norm*** $\|\cdot\|_{L_2(\Omega,\mathbb{R}^m)}$*, in short*

$$\mathbb{V}_k = (\mathbb{Y}_k, \|\cdot\|_{L_2(\Omega,\mathbb{R}^m)}).$$

or - mirroring the variational discretisation approach - the entire space $L_2(\Omega, \mathbb{R}^m)$*, i.e.*

$$\mathbb{V}_k = L_2(\Omega, \mathbb{R}^m).$$

equipped with the $L_2(\Omega, \mathbb{R}^m)$*-norm.*

Having introduced the discrete spaces $\mathbb{U}_k$, $\mathbb{Y}_k$ and $\mathbb{V}_k$, we can now define a discretisation of (P):

$$\left.\begin{array}{c} \min\limits_{U\in\mathbb{U}_k, Y\in\mathbb{Y}_k} \dfrac{1}{2}\|Y - y_d\|_{\mathbb{W}}^2 + \dfrac{\nu}{2}\|U\|_{\mathbb{U}}^2 \\ \text{s.t.} \\ \mathcal{B}[Y,W] = (U,W)_{\mathbb{U}} \quad \forall W \in \mathbb{Y}_k \\ \text{and} \\ U \in \mathcal{U}_k \\ I_k y_c - Y \in C_{\mathbb{V}_k}, \end{array}\right\} \qquad (P_k)$$

At this stage, we want to stress that on the discrete level we treat the cone C w.r.t the topology of the space $\mathbb{V}_k$. In terms of sets we still have

$$W \in C_{\mathbb{V}_k} \Rightarrow W \in C \tag{2.3.2}$$

After all, roughly speaking, we do not change the set C just the topology.
However, for the polar cone $C_{\mathbb{V}_k}^-$ things are different. Let us first observe that

$$C_{\mathbb{V}_k}^- = \left\{ \phi_k \in \mathbb{V}_k^* \, : \, \langle \phi_k, W \rangle_{\mathbb{V}_k^*, \mathbb{V}_k} \leq 0 \; \forall W \in C_{\mathbb{V}_k} \right\} = \{ F \in \mathbb{V}_k \, : \, (F, W) \leq 0 \; \forall W \in C_{\mathbb{V}_k} \} \tag{2.3.3}$$

The latter equality is due to the fact that $\mathbb{V}_k$ is a Hilbert space itself with the norm $\|\cdot\|$ and associated scalar product $(\cdot, \cdot)$. However in contrast to (2.3.2)

$$F \in C_{\mathbb{V}_k} \not\Rightarrow F \in C^-,$$

if $\mathbb{V}_k \neq L_2(\Omega, \mathbb{R}^m)$.
This is important to bear in mind for out future analysis. To conclude these remarks about the conce $C_{\mathbb{V}_k}$, let us another short one:

Remark 2.3.9. *Especially for higher order finite elements it can be helpful to dispense with the assumption that $C_{\mathbb{V}_k} \subset C$. The reason for this lies in the fact that pointwise a.e. constraints for quadratic or cubic polynomials are hard to verify due to oscillation. In this setting state constraints could e.g. be transferred to formulations of the type*

$$C_{\mathbb{V}_k} = \left\{ V \in \mathbb{Y}_k \, : \, \frac{1}{|T|} \int_T V \, dT \geq 0, \; \forall T \in \mathcal{T}_k \right\}.$$

Let us therefore remark that the results of this thesis remain valid even if $C_{\mathbb{V}_k} \not\subset C$ as long as the approximation is consistent, i.e.

$$V_k \in C_{\mathbb{V}_k}, V_k \rightharpoonup v \text{ in } L_2(\Omega, \mathbb{R}^m), k \to \infty \Rightarrow v \in C.$$

In addition to the properties (Pr1)-(Pr6) required of the continuous problem (P), we have to list two more properties and make additional assumptions for the discrete problem (P_k).

Property guaranteeing solvability of the equation $\mathcal{B}[Y, W] = (u, W)_{\mathbb{U}}$

Pr7. The bilinear form $\mathcal{B}$ satisfies a stable $\inf - \sup$ condition on $\mathbb{Y}_k$, i.e.

$$\inf_{Z \in \mathbb{Y}_k} \sup_{W \in \mathbb{Y}_k} \frac{\mathcal{B}[Z, W]}{\|Z\|_{\mathbb{Y}} \|W\|_{\mathbb{Y}}} = \inf_{W \in \mathbb{Y}_k} \sup_{Z \in \mathbb{Y}_k} \frac{\mathcal{B}[Z, W]}{\|Z\|_{\mathbb{Y}} \|W\|_{\mathbb{Y}}} = \alpha_k \geq \tilde{\alpha} > 0,$$

compare e.g. [65], Section 3.1 Theorem 4 and Section 3.1.2.
This is equivalent to the fact that there exists a unique solution $Y = S_k u \in \mathbb{Y}_k$ for all $u \in \mathbb{U}$ of

$$\mathcal{B}[Y, W] = (u, W)_{\mathbb{U}} \quad \forall W \in \mathbb{Y}_k$$

with $S_k \in \mathcal{L}(\mathbb{U}, \mathbb{Y})$ and

$$\|S_k u\|_{\mathbb{Y}} \lesssim \|u\|_{\mathbb{U}} \tag{2.3.4}$$

independent of k. Due to inequality (2.3.4) and (Pr1), in particular the embeddings, the operator norm of S_k is also uniformly bounded if interpreted as an operator $S_k : \mathbb{U} \to \mathbb{W}$ and $S_k : \mathbb{U} \to L_2(\Omega, \mathbb{R}^m)$. As in the continuous case, we will not distinguish between these nominally different operators, it being clear from the context which one we refer to.

Properties of the operator I_k of the state constraint in (P_k)

Pr8. I_k is an operator defined on a dense subspace $\mathbb{D}$ of $\mathbb{Y}$ with the property that $I_k y_c \in \mathbb{V}_k$, $I_k y_c \to y_c$ in $L_2(\Omega, \mathbb{R}^m)$ as $k \to \infty$ and

$$\|I_k y_c\| \lesssim \|y_c\|$$

independent of k.

To analyse convergence of the discrete solutions, we have to make the following assumptions:

Assumption ensuring existence of a bounded sequence of discrete solution of (P_k)

A1. There exists a bounded sequence $\left\{\hat{U}_k\right\}_{k\geq 0}$ such that for some fixed $N \in \mathbb{N}$

$$\hat{U}_k \in \mathbb{U}_k^{ad} := \{U \in \mathbb{U}_k \ : \ U \in \mathcal{U}_k,\ I_k y_c - S_k U \in C_{\mathbb{V}_k}\} \quad \forall k \geq N.$$

Assumptions needed to analyse convergence of discrete solutions

A2. For the sequence of discrete spaces $\mathbb{U}_k$, $\mathbb{Y}_k$ and $\mathbb{V}_k$, we assume that

$$\mathbb{U} = \overline{\bigcup_{k\geq 0} \mathbb{U}_k}^{\|\cdot\|_{\mathbb{U}}}, \quad \mathbb{Y} = \overline{\bigcup_{k\geq 0} \mathbb{Y}_k}^{\|\cdot\|_{\mathbb{Y}}}, \quad L_2(\Omega, \mathbb{R}^m) = \overline{\bigcup_{k\geq 0} \mathbb{V}_k}^{\|\cdot\|}$$

This is tantamount to the associated mesh size functions for the triangulations $\mathcal{S}_k$ and $\mathcal{T}_k$, $h_{\mathcal{S}_k}$ $h_{\mathcal{T}_k}$, see (2.3.1), converging to 0 pointwise almost everywhere in Γ and Ω respectively, see [60], Lemma 4.3.

A3. $\{\mathcal{U}_k\}_{k\geq 0}$ is a sequence of closed and convex subsets of $\mathbb{U}$ such that in $\mathbb{U}$ as $k \to \infty$

$$\forall u \in \mathcal{U} \ \ \exists P_k u \in \mathcal{U}_k \ \text{ s.t. } \ P_k u \to u \tag{2.3.5}$$

and

$$U_k \in \mathcal{U}_k \text{ and } U_k \rightharpoonup \tilde{u} \quad \Rightarrow \quad \tilde{u} \in \mathcal{U}. \tag{2.3.6}$$

A4. Similar to (2.3.5), for every $w \in C$ there exists $H_k w \in \mathbb{V}_k \cap C_{\mathbb{V}_k}$ such that as $k \to \infty$

$$H_k w \to w \ \text{ in } L_2(\Omega, \mathbb{R}^m). \tag{2.3.7}$$

We should explain the property (Pr7) as well as the technical assumptions (A1)-(A4):

- The stable $\inf - \sup$ condition in (Pr7) ensures - among else - that the operator norm $\|S_k\|_{\mathcal{L}(\mathbb{U},\mathbb{Y})}$ is uniformly bounded. This is crucial for proving convergence results of the type $S_k U_k \to Su$ as $U_k \to u$ and $k \to \infty$.

- (A1) - among else - safeguards that we are not optimising over the empty set, a necessary condition. Besides, the existence of the bounded sequence $\left\{\hat{U}_k\right\}$ ensures that the sequence of discrete solution stays bounded in $\mathbb{U}$. The reader should note that **almost all constants in the estimates for discrete functions** depend on the existence of such a norm-bound for the sequence $\left\{\hat{U}_k\right\}$.

- (A3): In (2.3.5) we have assumed that basically every function $u \in \mathcal{U}$ can be approximated by discrete functions $P_k u \in \mathcal{U}_k$. At this stage we want to emphasise that these functions do not have to fulfil the state constraint, i.e. in general $I_k y_c - S_k P_k u \notin C$. Provided (A2) holds, density always safeguards the existence of function $P_k u \in \mathbb{U}_k$ such that $P_k u \to u$ for all $u \in \mathbb{U}$. Thus, we have enforced the additional condition that these functions also belong to $\mathcal{U}_k$. In Section 2.4, we will give several examples of problems for which this condition holds, the easiest case being $\mathcal{U} = \mathbb{U}$ and $\mathcal{U}_k = \mathbb{U}_k$.
  Condition (2.3.6) guarantees that - in a sense - the closure w.r.t the weak topology of the sequence of sets $\mathcal{U}_k$ is contained in $\mathcal{U}$. It is trivially fulfilled if $\mathcal{U}_k \subset \mathcal{U}$, for all k. Again, we refer to Section 2.4 for some helpful examples.

- (A4) guarantees that every $w \in C$ can be approximated by a sequence of discrete functions $H_k w \in C_{\mathbb{V}_k}$, a property mirroring (A3). To the best of the authors' knowledge it is always fulfilled in case C is given by pointwise inequality constraints, the most important application for this theoretical framework. It is very technical to prove, though, especially for higher order finite elements. Compare also Section 2.4 for some examples of cases where it is fulfilled.

Before we move on, let us shortly record an important consequence of the Property (Pr7) and Assumption (A2):

Theorem 2.3.10 (Convergence of Discrete Solutions)**.** *For every $u \in \mathbb{U}$ we have $S_k u \to Su$ as $k \to \infty$ in $\mathbb{Y}$.*

Proof. First, we observe that the sequence $\{S_k u\}$ is uniformly bounded in $\mathbb{Y}$ thanks to (2.3.4). Thus, there exists a weakly convergent subsequence with limit $\tilde{y} \in \mathbb{Y}$. Now, pick an arbitrary $w \in \mathbb{Y}$. Due to (A2) there exists a sequence $\{W_k\}$ with $W_k \to w$ strongly in $\mathbb{Y}$ and $W_k \in \mathbb{Y}_k$ for all k. For the weakly convergent subsequence of $\{S_k u\}$ we can thus now conclude

$$\mathcal{B}[S_k u, W_k] \to \mathcal{B}[\tilde{y}, w], \ \ k \to \infty.$$

In addition, we have:

$$\begin{aligned} 0 = \mathcal{B}[S_k u - Su, W_k] = \mathcal{B}[S_k u, W_k] - \mathcal{B}[Su, W_k] &\to \mathcal{B}[\tilde{y}, w] - \mathcal{B}[Su, w], \ k \to \infty \\ &\Leftrightarrow \mathcal{B}[Su, w] = \mathcal{B}[\tilde{y}, w]. \end{aligned}$$

Since this result is valid for all $w \in \mathbb{Y}$, we immediately deduce $\tilde{y} = Su$ due to uniqueness of the solution Su.

Now, we realise that these deductions are true for *every* weakly convergent subsequence of $\{S_k u\}$ with the limit Su being unique. Lemma 2.1.5 then gives the desired result. □

Naturally, one is interested in existence and uniqueness results for the discrete problems (P_k) and optimality conditions. The good news is that the results of Theorem 2.2.2 can be readily transferred. First though, let us introduce the reduced formulation of (P_k) similar to (2.2.1). To this end, we define

$$f_k : \mathbb{U}_k \ni U \mapsto \frac{1}{2} \left\| S_k U - y_d \right\|_{\mathbb{W}}^2 + \frac{\nu}{2} \left\| U \right\|_{\mathbb{U}}^2 .$$

The reduced formulation now reads:

$$\min_{U \in \mathbb{U}_k^{ad}} f_k(U). \tag{2.3.8}$$

Applying Theorem 2.2.1 to (2.3.8) yields:

Theorem 2.3.11. *For every k there exists a unique solution $\bar{U}_k$ and corresponding state $\bar{Y}_k = S_k \bar{U}_k$ to (P_k) satisfying the following necessary and sufficient optimality condition*

$$(\bar{P}_k + \nu \bar{U}_k, U - \bar{U}_k)_{\mathbb{U}} \geq 0 \ \ \forall U \in \mathbb{U}_k^{ad}, \tag{2.3.9}$$

where $\bar{P}_k = S_k^*(\bar{Y}_k - y_d)$ with $S_k^* : \mathbb{W} \to \mathbb{U}$.

Furthermore, the sequences $\left\{\bar{U}_k\right\}_{k \in \mathbb{N}}$ and $\left\{f_k(\bar{U}_k)\right\}$ are bounded independently of k, and any weak limit $\tilde{u}$ of a subsequence of $\left\{\bar{U}_k\right\}$ fulfils

$$\tilde{u} \in \mathbb{U}^{ad}$$

Proof. Assumption (A1) is assumed to hold. In particular, this implies that the discrete admissible set $\mathbb{U}_k^{ad}$ is non-empty. Besides, f_k is radially unbounded, strictly convex, bounded from below and Fréchet differentiable, the arguments are exactly the same as for f, compare the proof of Theorem 2.2.2. Thus, we can apply Theorem 2.2.1 with $H = \mathbb{U}_k$, $V = \mathbb{U}_k^{ad}$ and $g = f^k$ to prove the existence of a unique solution for (P_k) and the optimality condition (2.3.9), where

$$0 \leq (\nabla f_k(\bar{U}_k), U - \bar{U}_k)_{\mathbb{U}} = (\underbrace{S_k^*(\bar{Y}_k - y_d)}_{=\bar{P}_k} + \nu \bar{U}_k, U - \bar{U}_k)_{\mathbb{U}} \ \ \forall U \in \mathbb{U}_k^{ad}.$$

Let us now turn to the remaining claims of the theorem.
To prove the boundedness property of the sequence $\left\{\bar{U}_k\right\}_{k\in\mathbb{N}}$, we take advantage of (A1) and optimality of $\bar{U}_k$ to estimate in the ensuing way:

$$\begin{aligned}
\frac{\nu}{2}\left\|\bar{U}_k\right\|_{\mathbb{U}}^2 &\leq \frac{1}{2}\left\|\bar{Y}_k - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k\right\|_{\mathbb{U}}^2 = f_k(\bar{U}_k) \\
&\leq \frac{1}{2}\left\|S_k\hat{U}_k - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\hat{U}_k\right\|^2 \\
&\lesssim \frac{1}{2}\left\|S_k\right\|_{\mathcal{L}(\mathbb{U},\mathbb{W})}\left\|\hat{U}_k\right\|^2 + \left\|S_k\right\|_{\mathcal{L}(\mathbb{U},\mathbb{W})}\left\|\hat{U}_k\right\|_{\mathbb{U}}\left\|y_d\right\|_{\mathbb{W}} + \left\|y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\hat{U}_k\right\|^2 \\
&\lesssim (1 + \frac{\nu}{2})\left\|\hat{U}_k\right\|_{\mathbb{U}}^2 + \left\|y_d\right\|_{\mathbb{W}}\left\|\hat{U}_k\right\|_{\mathbb{U}} + \left\|y_d\right\|_{\mathbb{W}}^2.
\end{aligned}$$

In the second to last line, we took advantage of continuity of S_k and Cauchy-Schwarz' inequality and in the last line, we additionally employed Assumption (Pr7), specifically the fact that $\|S_k\|_{\mathcal{L}(\mathbb{U},\mathbb{W})} \lesssim 1$ thanks to (2.3.4) and $\mathbb{Y} \hookrightarrow \mathbb{W}$. This yields boundedness of both $\left\{\bar{U}_k\right\}$ and $\left\{f_k(\bar{U}_k)\right\}$ which completes the proof; after all, $\left\{\hat{U}_k\right\}_{k\in\mathbb{N}}$ is bounded by Assumption (A1).
Let us now turn to the question of the feasibility of the weak limit $\tilde{u}$ of a convergent subsequence of $\left\{\bar{U}_k\right\}$ w.l.o.g. - for the sake of convenience - not distinguished from the entire sequence by notation.
First of all, the corresponding sequence of states $\bar{Y}_k$ is bounded thanks to (2.3.4). Since $\mathbb{Y}$ is a Hilbert space, it possesses a weakly convergent subsequence $\bar{Y}_{k_l}$ with limit $\tilde{y}$. Fixing an arbitrary $v \in \mathbb{Y}$, we choose a strongly convergent sequence $\{V_{k_l}\} \subset \mathbb{Y}$ with $V_{k_l} \in \mathbb{Y}_{k_l}$ whose existence is ensured by the density Assumption (A2) and observe that

$$\mathcal{B}[\tilde{y}, v] \leftarrow \mathcal{B}[\bar{Y}_{k_l}, V_{k_l}] = (\bar{U}_{k_l}, V_{k_l})_{\mathbb{U}} \rightarrow (\tilde{u}, v),\ l \rightarrow \infty$$

where we used strong convergence of $V_k \rightarrow v$ and weak convergence $\bar{Y}_k \rightharpoonup \tilde{y}$ in $\mathbb{Y}$ and continuity of $\mathcal{B}$ on $\mathbb{Y} \times \mathbb{Y}$ on the left-hand side and strong convergence $V_k \rightarrow v$ in $\mathbb{U}$, due to $\mathbb{Y} \hookrightarrow \mathbb{U}$, (Pr1), and weak convergence $\bar{U}_k \rightharpoonup \tilde{u}$ in $\mathbb{U}$ on the right-hand side.

Hence, because v is an arbitrary element of $\mathbb{Y}$

$$\tilde{y} = S\tilde{u} \leftharpoonup S_{k_l}\bar{U}_{k_l} = \bar{Y}_{k_l},\ l \to \infty \tag{2.3.10}$$

This is true for **every** subsequence of $\{\bar{Y}_k\}$ with the limit $S\tilde{u}$ being unique. Lemma 2.1.5 now guarantees that the entire sequence $\{\bar{Y}_k\}$ converges to $S\tilde{u}$.
Using Assumption (A3), in particular (2.3.6), we obtain

$$\mathcal{U}_k \ni \bar{U}_k \rightharpoonup \tilde{u} \in \mathcal{U}.$$

First of all, thanks to (2.3.2) we have

$$I_k y_c - \bar{Y}_k \in C_{\mathbb{V}_k} \Rightarrow I_k y_c - \bar{Y}_k \in C,\ k \to \infty$$

Now, harnessing the fact that the convex cone C is convex and closed and thus weakly closed, we can conclude

$$C \ni I_k y_c - \bar{Y}_k \rightharpoonup y_c - \tilde{y} = y_c - S\tilde{u} \in C,\ k \to \infty$$

As a consequence, we obtain

$$\tilde{u} \in \mathbb{U}^{ad},$$

which completes the proof. □

At this stage, we do not delve into the question of existence of Lagrange multipliers for the discrete problem or the reformulation of (2.3.9) as a projection equation as we did in the continuous case. Instead, we will move on to the discrete counterparts of the continuous regularised problem (P^ε). This analysis will form a centrepiece of this thesis. In fact, as already discussed in Section 2.2.3, these are the problems that are actually numerically solved, since these are the problems that can be tackled by efficient optimisation algorithms. We will present these regularised problems in the next section and also record several important properties.

2.3.4 The Discrete Regularised Problems

The discrete counterparts to (P^ε) are then defined in the following way:

$$\left.\begin{array}{c}\displaystyle\min_{U\in\mathbb{U}_k, Y\in\mathbb{Y}_k, V\in\mathbb{V}_k} \frac{1}{2}\left\|Y-y_d\right\|_{\mathbb{W}}^2+\frac{\nu}{2}\|U\|_{\mathbb{U}}^2+\frac{1}{2}\left\|V\right\|^2 \\ \text{s.t.} \\ \mathcal{B}[Y,W]=(U,W)_{\mathbb{U}} \quad \forall W\in\mathbb{Y}_k \\ \text{and} \\ U\in\mathcal{U}_k \\ I_k y_c - Y - \varepsilon V \in C_{\mathbb{V}_k}, \end{array}\right\} \qquad (P_k^\varepsilon)$$

First and foremost, we want to prove an existence and uniqueness result for (P_k^ε). Again, as in the unregularised case, the existence and uniqueness result for the continuous case of Theorem 2.2.12 readily finds its counterpart in the discrete case as the next theorem shows. To apply those results, we first transform (P_k^ε) into its reduced formulation by eliminating the state with the help of the discrete solution operator S_k, cf. (Pr7). We define

$$f_k^\varepsilon : \mathbb{U}_k \times \mathbb{V}_k \ni (U,V) \mapsto \frac{1}{2}\left\|S_kU-y_d\right\|_{\mathbb{W}}^2+\frac{\nu}{2}\left\|U\right\|_{\mathbb{U}}^2+\frac{1}{2\varepsilon}\left\|V\right\|^2$$

and the admissible set

$$\mathbb{U}_k^{\varepsilon,ad} := \{(U,V)\in\mathcal{U}_k\times\mathbb{V}_k \,:\, I_k y_c - S_kU - \varepsilon V \in C_{\mathbb{V}_k}\}$$

to put forward the reduced formulation for (P_k^ε):

$$\min_{(U,V)\in\mathbb{U}_k^{\varepsilon,ad}} f_k^\varepsilon(U,V). \tag{2.3.11}$$

To this constrained strictly convex optimisation problem we can apply Theorem 2.2.1 to obtain:

Theorem 2.3.12. *For every k and every fixed ε, there exists a unique solution couple $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\in\mathbb{U}_k^{\varepsilon,ad}$ of (P_k^ε) and corresponding state $\bar{Y}_k^\varepsilon = S_k\bar{U}_k^\varepsilon$ such that the following first-order necessary and sufficient optimality condition is fulfilled:*

$$(\bar{P}_k^\varepsilon+\nu\bar{U}_k^\varepsilon, U-\bar{U}_k^\varepsilon)_{\mathbb{U}}+\frac{1}{\varepsilon}(\bar{V}_k^\varepsilon, V-\bar{V}_k^\varepsilon)\geq 0 \quad \forall (U,V)\in\mathbb{U}_k^{\varepsilon,ad} \tag{2.3.12}$$

with $\bar{P}_k = S_k^*(\bar{Y}_k^\varepsilon - y_d)$.
Furthermore, the sequence $\left\{\bar{U}_k^\varepsilon, \frac{1}{\sqrt{\varepsilon}}\bar{V}_k^\varepsilon\right\}_{k\in\mathbb{N}}$ is bounded independently of ε and k in $\mathbb{U}\times L_2(\Omega,\mathbb{R}^m)$.

Proof. Once again, we apply Theorem 2.2.1 to problem (2.3.11) to obtain existence of a unique solution and the first-order optimality condition (2.3.12). We note that like f^ε, f_k^ε is

Fréchet-differentiable, radially unbounded, bounded from below and strictly convex. Besides, $\mathbb{U}_k^{\varepsilon,ad}$ is non-empty, because $(\bar{U}_k, 0) \in \mathbb{U}_k^{\varepsilon,ad}$. Hence, we can take advantage of Theorem 2.2.1 with $H = \mathbb{U}_k \times \mathbb{V}_k$, $V = \mathbb{U}_k^{\varepsilon,ad}$ and $g = f_\varepsilon^k$ to gain the existence of a unique solution and the first order optimality condition (2.3.12), where

$$\begin{aligned} 0 &\leq (\nabla f_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon), (U - \bar{U}_k^\varepsilon, V - \bar{V}_k^\varepsilon))_{\mathbb{U} \times L_2(\Omega, \mathbb{R}^m)} \\ &= (\bar{P}_k^\varepsilon + \nu \bar{U}_k^\varepsilon, U - \bar{U}_k^\varepsilon)_{\mathbb{U}} + \frac{1}{\varepsilon}(\bar{V}_k^\varepsilon, V - \bar{V}_k^\varepsilon) \; \forall (U, V) \in \mathbb{U}_k^{\varepsilon,ad}. \end{aligned}$$

To derive the boundedness result for the sequence $\left\{\bar{U}_k^\varepsilon, \frac{1}{\sqrt{\varepsilon}} \bar{V}_k^\varepsilon\right\}_{k \in \mathbb{N}}$, we first observe that

$$(U, 0) \in \mathbb{U}_k^{\varepsilon,ad}, \; \forall U \in \mathbb{U}_k^{ad}$$

and hence

$$f_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \leq f_k(\bar{U}_k),$$

which in particular yields

$$\frac{\nu}{2} \left\|\bar{U}_k^\varepsilon\right\|_{\mathbb{U}}^2 + \left\|\frac{1}{\sqrt{\varepsilon}} \bar{V}_k^\varepsilon\right\|^2 \leq f_k(\bar{U}_k)$$

which thanks to the boundedness results of Theorem 2.3.11 implies the assertion. □

As we have already remarked, it is (P_k^ε) that will be solved numerically not the unregularised problem (P_k). That is why when we discuss an adaptive finite element method for the discretisation of (P) it is natural that a lot of effort goes into studying the properties of the discretised regularised problem (P_k^ε): It will be the information extracted from its solution which steers the adaptive algorithm, a brief introduction of which we will now given in the next section.

2.3.5 The Different Modules of the AFEM

In this section, we will briefly discuss the four modules of the adaptive cycle in this state-constrained optimal control setting. This section is not geared towards a rigorous mathematical analysis; rather, the goal is to give the reader an overview of what we aim for and what we mean when discussing an adaptive algorithm.

To avoid certain technicalities, we assume that the spaces $\mathbb{U}_k$ and $\mathbb{Y}_k$ are defined on the same triangulations, i.e. $\mathcal{T}_k = \mathcal{S}_k$.

First of all, let us recall the adaptive cycle:

$$\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}$$

In Chapter 4 and Chapter 5 we will derive an estimator which up to constants depending solely on data (=Ω,y_c,y_d,...) provides the following upper bound:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{\mathbb{U}}^2 \lesssim \varepsilon^{\gamma N} + \mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \tag{2.3.13}$$

with $0 < \gamma < 1$ determined by properties of the continuous problem, a parameter N which can be choosen freely and expressions $\mathfrak{E}_r = \mathfrak{E}_k(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\mathfrak{E}_s = \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ of which we assume that we can compute them exactly (in truth, they are evaluated with numerical integration, but we will not address the issues we face in this case in this thesis). Besides, we demand that both $\mathfrak{E}_r$ and $\mathfrak{E}_s$ can be **localised** in the following sense:

$$\begin{aligned}\mathfrak{E}_r^2 &= \sum_{T\in\mathcal{T}_k} \mathfrak{e}_r^2(T)\\ \mathfrak{E}_s^2 &\lesssim \sum_{T\in\mathcal{T}_k} \mathfrak{e}_s^2(T)\end{aligned} \tag{2.3.14}$$

with element contributions $\mathfrak{e}_r(T)$ and $\mathfrak{e}_s(T)$. We observe that there is a $\lesssim$ in the second line in (2.3.14). Indeed, as we will find out in Chapter 5, we will pay for a localisation of $\mathfrak{E}_s$ by additional constants.

Furthermore, note that in (2.3.13) we estimate the distance between the solution of the discrete regularised problem $\bar{U}_k^\varepsilon$ and the true solution, because - as already explained before - it is the discrete regularised problem (P_k^ε) which is solved numerically and not the unregularised one (P_k). The reason for this is that only the regularised problem can be treated efficiently by Newton-type methods, a behaviour which we have already explored on the continuous level in Section 2.2.2 and Section 2.2.3.

We are now given a certain tolerance $TOL > 0$ and - for simplicity - assume that we fix N in (2.3.13) in such a way that

$$\varepsilon^{\gamma N} \le \frac{TOL}{2} \tag{2.3.15}$$

After these preliminaries we can now take a brief course through the different modules of our adaptive algorithm.

- 'SOLVE': In this module we solve (P_k^ε). For the presentation of the algorithm that we will use, we refer to Chapter 5. Here, it is important that upon completing this module, we assume that we possess the *exact* solution $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ of the problem (P_k^ε).

- 'ESTIMATE': In this module, the error estimators $\mathfrak{E}_r^2$ and $\mathfrak{E}_s$ as well as the element contributions $\mathfrak{e}_r(T)$ and $\mathfrak{e}_s(T)$ (compare (2.3.14)) are computed. If

$$\mathfrak{E}_r^2 + \mathfrak{E}_s \le \frac{TOL}{2},$$

we terminate the algorithm, because in view of (2.3.13) and (2.3.15) the current solution $\bar{U}_k^\varepsilon$ fulfills:

$$\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|_{\mathbb{U}}^2 \lesssim TOL.$$

If not, we continue with 'MARK'.

- 'MARK': In this step, a set of elements $\mathcal{M}_T \subset \mathcal{T}_k$ is marked. As we will explain in greater detail in Chapter 5, we will treat both local indicators $E_r^2(T)$ and $E_s^2(T)$ separately with the help of a maximum strategy: First, let

  $$\begin{aligned}\mathfrak{e}_r^{\max} &:= \max\{\mathfrak{e}_r(T) \;:\; T \in \mathcal{T}_k\}\\ \mathfrak{e}_s^{\max} &:= \max\{\mathfrak{e}_s(T) \;:\; T \in \mathcal{T}_k\}\end{aligned}$$

  Given fixed parameters $\sigma_r, \sigma_s \in [0, 1]$, the following set of elements will then be marked for refinement

  $$\mathcal{M}_T = \{T \in \mathcal{T}_k \;:\; \mathfrak{e}_r(T) \geq \sigma_r \mathfrak{e}_s^{\max} \vee \mathfrak{e}_s(T) \geq \sigma_s \mathfrak{e}_s^{\max}\}$$

  One could also adapt different marking strategies (equidistribution strategy,...) to the setting of (2.3.13).

- 'REFINE': During this step the marked elements are refined. We demand of any refinement algorithm that it generate a new conforming triangulation $\mathcal{T}_{k+1}$ and a sequence of triangulations that is shape-regular, cf. Definition 2.3.2. In this thesis, we operate solely in the context of **bisectional refinement**. One example of such a bisectional refinement technique is the so-called recursive refinement, [51], another iterative refinement, see e.g. [5], compare also [65], Section 4.3. We do not go into great detail here, because it would impede the cogent presentation of the results of this thesis. Let us merely mention that in the process of refinement not only the questions of shape-regularity and conformity have to be settled, it is also crucial that the algorithm addresses the question of refinement staying local. As the image below demonstrates, a conforming closure ultimately leads to refinement of possibly unmarked neighbouring elements indicated by the dotted line:

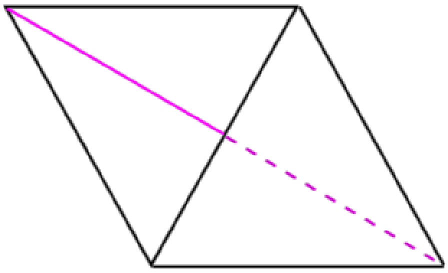

Figure 2.1: Conforming Closure

However, this necessary overhead needs to be limited in a way that assures that local refinement and ensuing conforming closure do not spill over into an almost global refinement obliterating the advantages of the adaptive finite element method. In [51], Theorem 2, it is shown that for the recursive refinement algorithm an appropriate bound safeguarding against such an effect can be obtained.
As a final note, let us also mention that there are of course other refinement techniques apart from bisectional refinement.

Having sketched the different modules of the AFEM, we can now present some examples which fit the abstract setting of Sections 2.2.1 and 2.3.3.

2.4 Examples

In this section, we will present several applications that fit into the framework of the preceding sections, i.e applications for which the Properties (Pr1)-(Pr6) and the Properties (Pr7)-(Pr8) as well as Assumptions (A1)-(A4) are fulfilled.

2.4.1 Distributed Elliptic Optimal Control Problems

Let us introduce the following elliptic optimal control problem with distributed control:

$$\begin{aligned} \min_{u\in L_2(\Omega),y\in H^1(\Omega)} \frac{1}{2}\,\|y-y_d\|^2_{\mathbb{W}} + \frac{\nu}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \\ -\mathrm{div}(A\cdot\nabla y)+cy = u \quad \text{in } \Omega \\ y = 0 \quad \text{in } \partial\Omega \\ \text{and} \\ a \le u \le b \text{ a.e. in } \Omega \\ y_c - y \le 0 \text{ a.e. in } \Omega \end{aligned} \tag{2.4.1}$$

Here Ω is a bounded domain in $\mathbb{R}^d$ which is assumed to be meshable, compare Definition 2.3.2. Besides, $\mathbb{W} = L_2(\Omega)$ or $\mathbb{W} = H^1(\Omega)$. $\mathbb{U}$ is chosen as $L_2(\Omega)$. Besides, let $y_d \in \mathbb{W}$ and $y_c \in W^1_p(\Omega)$, $p > d$.
In addition, $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$ and

$$\mathcal{U} = \{u \in L_2(\Omega) \, : \, a \leq u \leq b \text{ a.e. in } \Omega\}.$$

C is the following cone:

$$C = \{f \in L_2(\Omega) \, : \, f \leq 0 \text{ a.e. in } \Omega\}.$$

Following [36], Chapter 8, we make the following assumptions for the mappings A and c:

$$\begin{aligned} A &\in L_\infty(\Omega, \mathbb{R}^m) \\ c &\in L_\infty(\Omega) \end{aligned}$$

Besides, we demand that there exist $\lambda > 0$ such that

$$\eta^T A(x) \eta \geq \lambda |\eta|^2 \text{ f.a.a. } x \in \Omega, \ \forall \eta \in \mathbb{R}^d \tag{2.4.2}$$

and that there hold

$$c \geq 0 \text{ a.e. in } \Omega. \tag{2.4.3}$$

The bilinear form $\mathcal{B}$ of (Pr2) is given by

$$\mathcal{B} : \mathring{H}^1(\Omega) \times \mathring{H}^1(\Omega) \ni (y, w) \to \int\limits_\Omega A \nabla y \cdot \nabla w \, d\Omega \in \mathbb{R}.$$

The state equation then reads

$$\int\limits_\Omega A \nabla y \cdot \nabla w \, d\Omega = \int\limits_\Omega u w \, d\Omega \, \forall w \in \mathring{H}^1(\Omega). \tag{2.4.4}$$

In the next theorem, we want to collect some properties of the bilinear form $\mathcal{B}$, which will turn out to be very valuable in demonstrating that (2.4.4) is uniquely solvable

Theorem 2.4.1. *The bilinear form $\mathcal{B}$ is continuous on $\mathring{H}^1(\Omega) \times \mathring{H}^1(\Omega)$, i.e.*

$$|\mathcal{B}[y, w]| \leq \|A\|_{L_\infty(\Omega, \mathbb{R}^{d \times d})} |y|_{H^1(\Omega)} |w|_{H^1(\Omega)}$$

and coercive: For some $\alpha > 0$, there holds

$$\mathcal{B}[w, w] \geq \alpha |w|^2_{H^1(\Omega)}.$$

Proof. First of all, Theorem 2, Chapter 6 in [31], yields the continuity of the bilinear form $\mathcal{B}$. To prove that $\mathcal{B}$ is coercive, we use (2.4.2) and the fact that $c \geq 0$ a.e, compare (2.4.3) to conclude that

$$\begin{aligned}\lambda\, |w|^2_{H^1(\Omega)} &\leq \int\limits_\Omega A\nabla w \cdot \nabla w \; d\Omega \\ &\leq \mathcal{B}[w,w].\end{aligned}$$

□

Corollary 2.1.21 now yields the $\inf - \sup$ condition of (Pr2).
Having collected the properties of $\mathcal{B}$, we are now in the position to prove that problem (2.4.1) possesses the Properties (Pr1)-(Pr5):

Theorem 2.4.2 (elliptic model problem)**.** *Suppose* (2.4.1) *is given and the conditions listed above hold. Then Properties* (Pr1)-(Pr5) *are fulfilled.*

Proof. It is clear that the spaces $\mathbb{U} = L_2(\Omega)$, $\mathbb{W} = L_2(\Omega)$ or $\mathbb{W} = H^1(\Omega)$ and $\mathbb{Y} = \overset{\circ}{H}{}^1(\Omega)$ fit the prerequisites of (Pr1). Let us therefore immediately turn to (Pr2).
Theorem 2.1.20 and Corollary 2.1.21 now ensure that there exists a unique weak solution to (2.4.4), fulfilling Property (Pr2); in particular, the solution operator S maps $L_2(\Omega)$ continuously to $\overset{\circ}{H}{}^1(\Omega)$. The cone C of functions in $L_2(\Omega)$ which are non-positive on Ω, is a convex and closed cone. The convexity property is straightforward. For the closedness property, choose $\{f_n\} \subset C$ with $f_n \to f$ in $L_2(\Omega)$. As a consequence, there exists a subsequence $\{f_{n_k}\}$ which converges pointwise almost everywhere to f in Ω, see e.g. [39], Theorem 11.31, in combination with Theorem 11.26. Since $f_{n_k}(x) \leq 0$ pointwise almost everyhwere on Ω, pointwise a.e convergence implies $f(x) \leq 0$ for almost all $x \in \Omega$. Thus, (Pr4) is also fulfilled.
The same argument can be made to prove closedness of the set $\mathcal{U}$. Convexity of $\mathcal{U}$ is straightforward; thus, all in all, $\mathcal{U}$ satisifies (Pr3). Since $y_d \in \mathbb{W}$ and $y_c \in W^1_p(\Omega) \subset H^1(\Omega)$, $p > d$, (Pr5) is fulfilled, too. □

Before we turn to a discretisation approach for (2.4.1), let us briefly go into the question whether (Pr6) is fulfilled: First of all, we want to point out that there is no general way to ensure that there exist feasible points for (2.4.1). However, there are certain cases where (Pr6) is satisfied, most of which hinge on the application of maximum principles for elliptic equations. Sometimes it is also possible to explicitly construct feasible points. In the following remark, we will present one case where (Pr6) is fulfilled:

Remark 2.4.3. *Let us suppose that $b \geq 0$ and $y_c \leq 0$ a.e. in Ω. Then there exist feasible points for* (2.4.1)*. The reason for this is the fact that $u \geq 0$ implies $Su \geq 0$. This is a consequence of well-known maximum principles for elliptic equations, compare e.g. [67], Chapter 5.*

Let us now verify whether the Assumptions (Pr7)-(A4) are fulfilled for the discretisation we are now going to consider. For $\mathbb{U}_k$ we choose piecewise polynomials of up to degree $l \geq 0$ defined on a sequence of shape-regular, conforming triangulations $\mathcal{T}_k$ of Ω, compare Definition 2.3.4 and Example 2.3.5:

$$\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_l, L_2(\Omega))$$

For $\mathbb{Y}_k$ we choose the space of $\mathring{H}^1(\Omega)$-conforming piecewise polynomials of up to degree $m \geq 1$, i.e.

$$\mathbb{Y}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_m, \mathring{H}^1(\Omega)) \tag{2.4.5}$$

For $\mathbb{V}_k$ we choose

$$\mathbb{V}_k = (\mathbb{Y}_k, \|\cdot\|_{L_2(\Omega)}) \tag{2.4.6}$$

The admissible set $\mathcal{U}_k$ is defined by

$$\mathcal{U}_k = \{U \in \mathbb{U}_k \,:\, a \leq U \leq b \text{ a.e. in } \Omega\}.$$

For the operator I_k in (Pr8), we choose any of the Clément, Scott-Zhang or nodal interpolant. For a definition of these interpolants and further information, we refer e.g. to [20], Chapter 2, Section 6, [75], [30], Section 1.6. An alternative approach would be to just take the best-approximation of y_c in $\mathbb{Y}_k$.
The state equation on the discrete level is then given by

$$\int\limits_\Omega A\nabla Y \cdot \nabla W : d\Omega = \int\limits_\Omega UW \, d\Omega \;\; \forall W \in \mathbb{Y}_k. \tag{2.4.7}$$

If we use this discretisation scheme, then Properties (Pr7), (Pr8) and Assumptions (A2)-(A4) are fulfilled. This is the subject of the next theorem.

Theorem 2.4.4. *For the discretisation setting detailed above,* (Pr7),(Pr8) *and* (A2)-(A4) *are fulfilled.*

Proof. Let us tackle (Pr7) first: At this stage, we want to refer to Corollary 2.1.22, which ensures that (Pr7) is fulfilled, since $\mathbb{Y}_k \subset \mathring{H}^1(\Omega)$ is a closed subspace of $\mathring{H}^1(\Omega)$.

By continuing refinement and letting the mesh-size tend to 0 we ensure that (A2) is fullfilled. Let us check (A3) now. First of all, let us observe that the L_2-projection of a function $u \in \mathcal{U}$

on the space of piecewise constant functions (compare Example 2.3.5) defined by

$$\int\limits_{\Omega} (P_k u - u) W \, d\Omega = 0 \; \forall W \in \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$$

belongs to $\mathbb{U}_k$. We can now choose for all $T \in \mathcal{T}_k$ $V = \frac{\chi_T}{|T|}$ and obtain (observe that $P_k u$ is constant on every element T)

$$P_k u|_T = \frac{1}{|T|} \int\limits_T u \, dT \;\; \forall T \in \mathcal{T}_k.$$

Since $a \le u \le b$, so

$$a \le \frac{1}{|T|} \int\limits_T u \, dT \le b,$$

and thus

$$a \le P_k u|_T \le b \;\; \forall T \in \mathcal{T}_k,$$

which leads to $a \le P_k u \le b$ and thus $P_k u \in \mathcal{U}_k$. $P_k u$ is the best-approximation of u in $\mathbf{FES}(L_2(\Omega), \mathbb{P}_0, \mathcal{T}_k)$; hence, the density relation (A2) allows us to conclude that $P_k u \to u$ for all $u \in \mathcal{U}$. Thus, Assumption (2.3.5) is satisfied. Assumption (2.3.6) is also fulfilled, since the set $\mathcal{U}$ is convex and closed in $\mathbb{U}$ and, as a consequence, weakly closed, and there holds:

$$\mathcal{U}_k \subset \mathcal{U}.$$

To show that (Pr8) is fulfilled, we employ standard interpolation results, e.g. [14], Theorem 4.4.4., and more generally [20], Theorems 16.1 and 16.2, which yield stability, i.e.

$$\|I_k y_c\| \lesssim \|y_c\|$$

with the constant depending - among else - on the shape regularity of $\mathcal{T}_k$, and the following estimates

$$\begin{aligned} \|y_c - I_k y_c\|_{H^1(\Omega)} &\lesssim \sum_{T \in \mathcal{T}_k} h_T^{d(\frac{1}{2} - \frac{1}{p})} \, |y_c|_{W_p^1(T)} \\ &\le \sum_{|\alpha|=1} \left\| h_{\mathcal{T}_k}^{d(p/2-1)} D^\alpha y_c \right\|_{L_p(\Omega)} \end{aligned}$$

where $h_{\mathcal{T}_k}$ is the local mesh-size function of the triangulation $\mathcal{T}_k$. Here, we took advantage of the higher regularity of $W_p^1(\Omega)$. Density, cf. (A2), then yields pointwise a.e. convergence of $h_{\mathcal{T}_k} \to 0$ and, thus, since $\|h_{\mathcal{T}_k}\|_{L_\infty} \lesssim 1$, where the constant solely depends on the initial mesh-size, an application of the dominated convergence theorem, compare [85], Theorem 5.36, yields:

$$h_{\mathcal{T}_k} \to 0 \;\text{ in } L_p(\Omega) \;\; \forall 1 \leq p \leq \infty.$$

Using Hölder's inequality, we can then deduce that

$$\sum_{|\alpha|=1} \left\| h_{\mathcal{T}_k}^{d(p/2-1)} D^\alpha y_c \right\|_{L_p(\Omega)} \to 0$$

and thus $I_k y_c \to y_c$ in $H^1(\Omega)$ (and $\overset{\circ}{H}{}^1(\Omega)$) and hence, in particular, $I_k y_c \to y_c$ in $L_2(\Omega)$.
Lastly, let us investigate (A4). For the operator H_k demanded by (A4), we use a positivity preserving finite element approximation described in [64]. In general, such an operator does not give an optimal approximation rate, but in our case it is enough that $H_k w \to w$ for all $w \in L_2(\Omega)$ which is ensured by the operator given in [64]. This completes the proof. □

Let us conclude this section with some remarks on the question of whether (A1) is fulfilled: As in the continuous case, generally it is not clear a priori that the discrete admissible set $\mathbb{U}_k^{ad}$ is non-empty. If it were, though, the existence of the bounded sequence $\left\{\hat{U}_k\right\}$ would immediately be ensured, since for any feasbile point $U \in \mathbb{U}_k^{ad}$ we have $a \leq U \leq b$.
In certain special cases, this non-emptiness is ensured by additional properties of the mesh and the solution operator. One is similar to Remark 2.4.3, namely if a maximum principle holds for the discrete solution operator S_k, then the setting of Remark 2.4.3 can immediately be transferred to the discrete case. For further reading regarding the conditions that need to be satisfied for the discrete maximum principle to hold, we refer to [21] and [57] . The second case that we want to mention here is strongly related to $L_\infty(\Omega)$-convergence of the states $S_k u \to Su$ for all $u \in \mathbb{U}$. The convergence is given in terms of the mesh size,

$$\|Su - S_k u\|_{L_\infty(\Omega)} \lesssim h_k^\gamma \;\; \gamma > 0,$$

where the sequence of triangulations has to be uniform, compare Definition 2.3.3. Provided condition (2.2.12) holds, there are ways to prove that for sufficiently fine meshes (A1) is fulfilled compare [69], Remark 3.1.

2.4.2 Neumann Elliptic Optimal Control Problems

My second example is a Neumann elliptic optimal control problem with pointwise state constraints. We consider the following model problem

$$\begin{aligned} \min_{u\in L_2(\Omega),y\in H^1(\Omega)} \frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L_2(\partial\Omega)} \\ \text{s.t.} \\ -\Delta y + y = 0 \quad \text{in } \Omega \\ \partial_n y = u \quad \text{on } \partial\Omega \\ \text{and} \\ a \leq u \leq b \text{ a.e. in } \partial\Omega \\ y_c - y \leq 0 \text{ a.e. in } \Omega \end{aligned} \tag{2.4.8}$$

Here, $\mathbb{U} = L_2(\partial\Omega)$. The set $\mathcal{U}$ is defined by:

$$\mathcal{U} := \{u \in L_2(\partial\Omega) \, : \, a \leq u \leq b\}, \quad a, b \in \mathbb{R} \cup \{-\infty, \infty\}$$

$y_c \in W^1_p(\Omega)$, $p > d$, and the state constraint is given by

$$y_c - y \leq 0 \text{ a.e. in } \Omega.$$

The cone C is given by

$$C := \{f \in L_2(\Omega) \, : \, f(x) \leq 0 \text{ f.a.a. } x \in \Omega\}.$$

As in the previous section, we first want to check whether the properties of the continuous problem, Properties (Pr1)-(Pr6), are fulfilled. Evidently, the setting of (2.4.8) satisfies (Pr1), since $L_2(\partial\Omega) \hookrightarrow H^1(\Omega)^*$, so we immediately turn to (Pr2). To check whether there exists a unique solution to the state equation in (2.4.8) we first define the weak formulation:

$$\int\limits_{\Omega} \nabla y \cdot \nabla w \, d\Omega + \int\limits_{\Omega} yw \, d\Omega = \int\limits_{\partial\Omega} uv \, d\partial\Omega \; \forall w \in H^1(\Omega) \tag{2.4.9}$$

This variational formulation is well-defined, since functions belonging to $H^1(\Omega)$ possess boundary values in the sense of traces, compare Theorem 2.1.31. That is why using the embedding operator $E : L_2(\partial\Omega) \to H^1(\Omega)^*$, we can ensure that the right-hand side in (2.4.9) defines a functional on $H^1(\Omega)$ by virtue of:

$$\langle Eu, w\rangle_{(H^1)^*,H^1} := \int\limits_{\partial\Omega} uw \, d\partial\Omega.$$

The next theorem now ensures that the state equation in 2.4.8 possess a unique solution depending continuously on the data:

Theorem 2.4.5. *The bilinear form given by the left-hand side in* (2.4.9)

$$\mathcal{B}[y,w] = \int\limits_{\Omega} \nabla y \cdot \nabla w \, d\Omega + \int\limits_{\Omega} yw \, d\Omega$$

is continuous and coercive on $H^1(\Omega) \times H^1(\Omega)$, i.e.

$$|\mathcal{B}[y,w]| \lesssim \|y\|_{H^1(\Omega)} \|w\|_{H^1(\Omega)} ,$$

and there exists an $\alpha > 0$ such that

$$\mathcal{B}[w,w] \geq \alpha \|w\|^2_{H^1(\Omega)} .$$

Thus, for every $u \in L_2(\partial\Omega)$ there exists a unique solution $y = Su$ to (2.4.9) *with*

$$\|Su\|_{H^1(\Omega)} \lesssim \|u\|_{L_2(\partial\Omega)} .$$

Proof. Continuity of $\mathcal{B}$ is a consequence of a straightfoward application of Cauchy-Schwarz's inequality. The equation

$$\mathcal{B}[w,w] = \|w\|^2_{H^1(\Omega)}$$

immediately yields coercivity.
An application of Corollary 2.1.21 yields the $\inf - \sup$ condition of $\mathcal{B}$ and thus ensures that (2.4.9) is uniquely solvable with the solution depending continuously on the data. □

The next theorem states that (Pr1)-(Pr5) are fulfilled:

Theorem 2.4.6. *Let* (2.4.8) *be given. Then Properties* (Pr1)-(Pr5) *are fulfilled*

Proof. That (Pr1) is fulfilled is evident, (Pr2) being satisfied is a consequence of Theorem 2.4.5. (Pr3), (Pr4) and (Pr5) hold too, in the case of (Pr3) and (Pr4) the proof of Theorem 2.4.2 can be readily adapted, while (Pr5) is clear from the problem setting. □

(Pr6) can only be checked a priori in certain special cases. This problem has already been discussed in the previous section; that's why we move on to a possible way to discretise (2.4.8). Since the discretisation of $\mathbb{U} = L_2(\partial\Omega)$ requires a discretisation of space defined on the boundary $\partial\Omega$, things are a bit more technical compared to the distributed control case of the previous section. Therefore, we will focus on $\mathbb{P}_0$-elements for the control and $H^1(\Omega)$-conforming $\mathbb{P}_1$-elements for the state. In addition, we choose $\mathbb{V}_k$ as in the distributed case, compare (2.4.6).

We start with a sequence of shape-regular and conforming triangulations $\mathcal{T}_k$ of Ω with $h_{\mathcal{T}_k} \to 0$. The space $\mathbb{Y}_k$ is then defined completely analogously to (2.4.5) with $n = 1$.
To define $\mathbb{U}_k$, we first examine the *skeleton* $\hat{\mathcal{S}}_k$ of $\mathcal{T}_k$

$$\hat{\mathcal{S}}_k = \bigcup_{T \in \mathcal{T}_k} \partial T$$

However, we are not interested in the entire skeleton, but only in the segments lying on the boundary. Thus, we define the following triangulation $\mathcal{S}_k$ of $\partial\Omega$, where $\partial\Omega$ is now viewed as a $d-1$ dimensional domain so that the notions of Definition 2.3.2 and Definition 2.3.4 are immediately transferable:

$$\mathcal{S}_k = \bigcup_{S \in \hat{\mathcal{S}}_k} S \cap \partial\Omega.$$

At this stage, we want to point out that such a definition is only possible if Ω is meshable, cf Definition 2.3.2. Having defined a triangulation of $\partial\Omega$, we can now proceed to define $\mathbb{U}_k$:

$$\mathbb{U}_k = \mathbf{FES}(\mathcal{S}_k, \mathbb{P}_0, L_2(\partial\Omega)).$$

Compare also Definition 2.3.4 and Example 2.3.5.
The set $\mathcal{U}_k$ is given by

$$\mathcal{U}_k := \{U \in \mathbb{U}_k \;:\; a \le U \le b\}$$

and as in the case of the distributed control of the previous section, we observe that

$$\mathcal{U}_k \subset \mathcal{U}.$$

The discretised state equation reads

$$\int_\Omega \nabla Y \cdot \nabla W \, d\Omega + \int_\Omega Y W \, d\Omega = \int_{\partial\Omega} U W \, d\partial\Omega \; \forall W \in \mathbb{Y}_k.$$

The state constraint on the continuous level finds its counterpart on the discrete level with

$$I_k y_c - Y \le 0,$$

For the operator I_k the definitions and arguments of the previous section can be immediately transferred to this setting. Having introduced a discretisation, we can now turn to verifying the Properties (Pr7),(Pr8) and (A2)-(A4), which is done in the ensuing theorem:

Theorem 2.4.7. *For the discretisation setting defined above, the Properties* (Pr7)*,*(Pr8) *and*

Assumptions (A2)-(A4) *hold.*

Proof. To verify Property (Pr7), we first observe that the continuity and coercivity results of Theorem 2.4.5 are inherited by the discrete space pairing $\mathbb{Y}_k \times \mathbb{Y}_k$. As detailed in the proof of Theorem 2.4.4, the bilinear form $\mathcal{B}$ is $\inf-\sup$ stable on $\mathbb{Y}_k \times \mathbb{Y}_k$, which yields the existence of a discrete solution operator S_k with all the properties demanded in (Pr7).
Due to mesh size function for $\mathcal{T}_k$, $h_{\mathcal{T}_k}$, tending to 0 almost everywhere in Ω as $k \to \infty$ we have

$$\mathbb{Y} = \overline{\bigcup_{k\geq 0} \mathbb{Y}_k}^{\|\cdot\|_{H^1(\Omega)}},\ L_2(\Omega) = \overline{\bigcup_{k\geq 0} \mathbb{V}_k}^{\|\cdot\|}$$

For $\mathbb{U}_k$ things are not immediately clear, because $\mathbb{U}_k$ is just defined on the boundary. However, shape regularity ensures that for the $d-1$-dimensional Hausdorff measure $|S|_{d-1}$ of any $S \in \mathcal{S}_k$ contained in an element $T \in \mathcal{T}_k$ we have up to constants independent of k

$$|S|_{d-1} = |T|^{\frac{1}{d-1}} = h_{\mathcal{T}_k}(x)^{\frac{d-1}{d}},\ x \in T$$

Since $h_{\mathcal{T}_k} \to 0$ as $k \to \infty$ a.e., we also have that $|S|_{d-1} \to 0$ for all $S \in \mathcal{S}_k$ as $k \to \infty$. This in turn ensures that

$$\mathbb{U} = \overline{\bigcup_{k\geq 0} \mathbb{U}_k}^{\|\cdot\|_{L_2(\partial\Omega)}}.$$

Thus, Assumption (A2) holds.
To realise that (A3) is fulfilled, the arguments from Theorem 2.4.4 can be immediately transferred, the same holds for (Pr8).
To realise that (A4) is fulfilled, we again employ the positivity preserving finite element of [64]. □

As in the case of the distributed control setting of the previous section, it is only clear in certain special cases whether (A1) is fulfilled. We do not want to list them again here, so we just refer to the discussion of said previous section. For maximum principles for elliptic equations with Neumann data, we refer to [15].

2.4.3 Stationary Stokes Model Problem

We now want to depart from the elliptic setting of the previous two sections and present another application: the control of a stationary Stokes model problem:

$$\begin{aligned} \min_{u\in L_2(\Omega,\mathbb{R}^d),(y,p)\in H^1(\Omega,\mathbb{R}^d)\times L_{2,0}(\Omega)} \frac{1}{2}\left\|(y,p)-y_d\right\|^2_{\mathbb{W}} &+ \frac{\nu}{2}\|u\|^2_{L_2(\Omega,\mathbb{R}^d)} \\ &\text{s.t.} \\ -\Delta y + \nabla p &= u \quad \text{in } \Omega \\ \operatorname{div}\, y &= 0 \\ y &= 0 \quad \text{in } \partial\Omega \\ &\text{and} \\ p_c - p \leq 0 &\text{ a.e. in } \Omega \end{aligned} \tag{2.4.10}$$

In this setting $\mathbb{Y} = \mathbb{W} = \mathring{H}^1(\Omega,\mathbb{R}^d)\times L_{2,0}(\Omega)$, where

$$L_{2,0}(\Omega) := \left\{ p \in L_2(\Omega) \ : \ \int_\Omega p \, d\Omega = 0 \right\}$$

is the space of L_2-functions with zero mean value.
Here, we picked a setting where there are no constraints on the control and the velocity v, though both can be included, too. Hence,

$$\mathcal{U} = L_2(\Omega,\mathbb{R}^d)$$

and (Pr3) is trivially satisfied.
However, there is a constraint on the pressure p. The cone C here again is the cone of non-positive function in $L_2(\Omega)$ given by

$$C := \{ f \in L_2(\Omega) \ : \ f(x) \leq 0 \text{ f.a.a. } x \in \Omega \},$$

which fulfils (Pr4). Besides, we assume that $p_c \in W^1_p(\Omega)$, $p > d$. Defining the bilinear forms

$$a : \mathring{H}^1(\Omega,\mathbb{R}^d)\times \mathring{H}^1(\Omega,\mathbb{R}^d) \ni (w,z) \mapsto \int_\Omega \nabla w : \nabla z \, d\Omega \in \mathbb{R},$$

where, in a slight abuse of notation,

$$\nabla w : \nabla z := \sum_{i=1}^d \nabla w_i \cdot \nabla z_i,$$

and

$$b : L_{2,0}(\Omega) \times \mathring{H}^1(\Omega) \ni (q, w) \mapsto \int_\Omega \mathrm{div} w \, q \, d\Omega,$$

we can specify $\mathcal{B} : \mathbb{Y} \times \mathbb{Y} \to \mathbb{R}$ in (Pr2), which is given by

$$\mathcal{B}[y, w] := a[y_1, w_1] + b[y_2, w_1] + b[w_2, y_1] \tag{2.4.11}$$

with

$$y = (y_1, y_2) \in \mathring{H}^1(\Omega, \mathbb{R}^d) \times L_{2,0}(\Omega), \quad w = (w_1, w_2) \in \mathring{H}^1(\Omega, \mathbb{R}^d) \times L_{2,0}(\Omega).$$

The right-hand side in the state equation in (2.4.10) is defined by:

$$(u, w)_{L_2(\Omega,\mathbb{R}^d)} = \int_\Omega u \cdot w \, d\Omega \quad w \in H^1(\Omega, \mathbb{R}^d)$$

The next theorem states that (Pr2) holds. A proof can be found in [13], Section 6, in particular Remark 6.5, and for general saddle point problems, we refer to [65], Section 2.4.2.

Theorem 2.4.8. *For every $u \in L_2(\Omega, \mathbb{R}^d)$, there exists a unique weak solution $(y, p) = Su$ to the state equation in* (2.4.10) *with*

$$(\|y\|^2_{H^1(\Omega,\mathbb{R}^d)} + \|p\|^2_{L_2(\Omega)})^{1/2} \lesssim \|u\|_{L_2(\Omega,\mathbb{R}^d)} \, .$$

Since there are no control constraints in (2.4.10), it is not difficult to construct a feasible point. One starts with choosing a smooth p such that $p \geq p_c$. Then picking a smooth velocity field y satisfying $\mathrm{div} y = 0$ and $y = 0$ on $\partial\Omega$, one can simply set

$$u = \Delta y - \nabla p$$

in the classical sense. This u is a feasible point. Hence (Pr6) holds.

Discretising the stationary Stokes problem, one has to be careful to obtain a disretisation scheme that fulfils the stable $\inf - \sup$ condition of (Pr7). As it turns out, the straightforward $\mathbb{P}_1$ elements of the previous two sections for the velocity field coupled with piecewise constant elements for the pressure do not satisfy this stability conditon, see. e.g. [11], Section 2.1. Thus, it is crucial that one turns to other, stable methods; one of which is the Taylor-Hood-Element, see e.g. [13], Section 7, and [84]. For a conforming and shape-regular sequence of triangulations $\mathcal{T}_k$ of Ω, we define

$$\mathbb{W}_k := \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_2, C(\bar{\Omega}, \mathbb{R}^d) \cap \mathring{H}^1(\Omega, \mathbb{R}^d))$$

and

$$\mathbb{Q}_k := \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, C(\Omega) \cap L_{2,0}(\Omega))$$

and

$$\mathbb{Y}_k = \mathbb{W}_k \times \mathbb{Q}_k.$$

For the space $\mathbb{U}_k$ we choose

$$\mathbb{U}_k = \mathbb{W}_k.$$

Consequently,

$$\mathcal{U}_k = \mathbb{U}_k.$$

Indeed other choices are possible for the discretisation of the control u (piecewise constant, piecewise linear, etc..), since for $\mathbb{U}_k$ we do not have to take into account any stability issues as for the state.
Against the backdrop (2.4.11), the variational formulation of the state equation in (2.4.10) reads

$$\mathcal{B}[Y, W] = (U, W)_{L_2(\Omega,\mathbb{R}^d)} \;\; \forall W = (W_1, W_2) \in \mathbb{Y}_k, \; Y = (Y_1, Y_2) \in \mathbb{Y}_k. \tag{2.4.12}$$

For the proof of the stable inf − sup condition, we again refer to [84]. We will summarise this result in the next theorem:

Theorem 2.4.9. *For every $U \in \mathbb{U}_k$, there exists a unique solution $(Y_1, Y_2) = Y = S_k U$ such that*

$$\|Y\|_{H^1(\Omega,\mathbb{R}^d)\times L_{2,0}(\Omega)} \lesssim \|U\|_{L_2(\Omega,\mathbb{R}^d)},$$

where the hidden constant is independent of k. Hence, for this discretisation setting, Property (Pr7) *holds.*

Let us now verify the other properties and assumptions for this discretisation scheme:

Theorem 2.4.10. *For the discretisation scheme detailed above, the Property* (Pr8) *and Assumptions* (A2)-(A4) *are satisfied.*

Proof. By continuing refinement and letting the mesh-size tend to 0 we ensure that (A2) is fulfilled.
The density relation (A2) for the sequence of discrete spaces $\mathbb{U}_k$ ensures that (A3) is fulfilled. For P_k one can just take the best-approximation of an arbitrary function $u \in L_2(\Omega, \mathbb{R}^d)$ in $\mathbb{U}_k$.
For the definition of an appropriate operator $I_k p_c$ fulfilling (Pr8), we refer to the previous sections, especially to the proof of Theorem 2.4.2.
Theorem 2.4.9 ensures that (Pr7) holds. For (A4) we again refer to the positivity preserving finite element approximation of [64]. □

For the verification of (A1), we again refer to the discussions of the previous sections on the same matter.

Chapter 3

A Basic Convergence Result

In this chapter, we will prove a basic convergence result for the sequence of discrete solutions $\bar{U}_k$ of (P_k). In essence, this result states, without the assumption of any additional regularity for the bilinear form $\mathcal{B}$ in (P) and the sequence of triangulation $\mathcal{T}_k$, that given a certain condition

$$\bar{U}_k \to \bar{u}, k \to \infty \text{ in } \mathbb{U},$$

where $\bar{u}$ denotes the solution to (P). This condition will turn out to be both necessary and sufficient for convergence, i.e. in this chapter we derive an exact characterisation of convergence $\bar{U}_k \to \bar{u}$ as $k \to \infty$.

The first question one has to ask is: What is the worth of such a 'low-regularity' convergence result? To answer this question, let us first recall the density assumption we made for the sequence of discrete spaces, in particular the one for $\mathbb{U}_k$, (A2):

$$\mathbb{U} = \overline{\bigcup_{k \geq 0} \mathbb{U}_k}^{\|\cdot\|_{\mathbb{U}}}. \tag{3.0.1}$$

Now, let us also once again state the adaptive cycle:

SOLVE → ESTIMATE → MARK → REFINE

Basically, (A2) and (3.0.1) imply that we continue to refine according to our estimator, marking and refinement strategy, cf Section 2.3.5, - the 'ESTIMATE', 'MARK' and 'REFINE' modules. We now pose the natural question whether this process of continuing refinement actually gets us any closer to the true solution $\bar{u}$. This, however, is only guaranteed if we know that $\bar{U}_k \to \bar{u}$, otherwise we could refine and refine and still would not get any closer to the true solution. This means that without convergence $\bar{U}_k \to \bar{u}$ continuing refinement may become completely pointless making such a basic convergence result an absolutely necessary ingredient for a working adaptive algorithm.

The key obstacle we now face is that convergence $\bar{U}_k \to \bar{u}$ should remain valid regardless of how we continue to refine our mesh. After all, the hallmark of an adaptive algorithm precisely is that one does not know a priori how the mesh looks like at iterate k, because information from the discrete solution is a posteriori extracted to guide refinement in a problem-dependent manner. That is why one is for better or for worse confined to using solely density information, such as (A2) and (3.0.1), and properties of the discrete problem (P_k) which are ensured irrespective of regularity properties of the mesh, such as quasi-uniformity, cf Definition 2.3.3. This is the fundamental difference to the convergence results for state-constrained optimal control problems which have already been proven, e.g. [16],[26],[25], [59], [69] and [24]. Here, quasi-uniformity was assumed a priori.

Having made these introductory remarks, we can now turn to the actual results of this section:
The condition for convergence we briefly mentioned before is given in terms of a smoothness property of the sequence of continuous and discrete regularised problems. Therefore, we will proceed as follows: First, we will prove major properties of the continuous and discrete regularised problems. Next, we will prove a theorem linking the question of convergence $\bar{U}_k \to \bar{u}$ to a smoothness property of the regularised problems. To conclude this section, we will derive a necessary and sufficient condition for which this smoothness property holds and prove the central convergence result of this section.

Before we tackle the convergent results, let us list three technical lemmata, which we will often make use of throughout this chapter:

3.1 Three Auxiliary Results

The first lemma provides a way to make the step from weak to strong convergence:

Lemma 3.1.1. *Suppose H_i, $i = 1, .., n$, are Hilbert spaces and*

$$H = \prod_{i=1}^{n} H_i$$

is the product space with norm

$$\|g\|_H = (\sum_{i=1} \alpha_i \left\|g^i\right\|_{H_i}^2)^{1/2},$$

with $\alpha_i > 0$ *for all* $i = 1, ...n$*. Suppose further that the sequence* $\{g_k\} \subset H$ *fulfils*

$$g_k = (g_k^1, ..., g_k^n) \rightharpoonup (g^1, ..., g^n) = g, \ k \to \infty \ in \ H$$

and

$$\|g_k\|_H^2 = \left\|(g_k^1, ..., g_k^n)\right\|_H^2 \to \left\|(g^1, ..., g^n)\right\|_H^2 = \|g\|_H^2 \ \ k \to \infty.$$

Then

$$g_k = (g_k^1, ..., g_k^n) \to (g^1, ..., g^n) = g, \ k \to \infty \ in \ H$$

and

$$g_k^i \to g^i, \ k \to \infty, \ in \ H_i, \ \forall i$$

Proof. Since $g_k \rightharpoonup g$ and $\|g_k\|_H^2 \to \|g\|_H^2$ as $k \to \infty$, we can estimate in the following fashion

$$\begin{aligned}\|g_k - g\|_H^2 &= \sum_{i=1}^n \alpha_i (g_k^i - g^i, g_k^i - g^i)_{H_i} \\ &= \sum_{i=1}^n \alpha_i (\left\|g_k^i\right\|_{H_i}^2 - 2(g^i, g_k^i)_{H_i} + \left\|g^i\right\|^2) \\ &= \|g_k\|_H^2 - 2(g_k, g)_H + \|g\|_H^2 \\ &\to 0, \ \ k \to \infty.\end{aligned}$$

Since $g_k \to g$ implies $g_k^i \to g^i$ for all $i = 1, ..., n$ as $k \to \infty$, we can conclude this proof. □

The second lemma of this section offers a way to deduce convergence of a sum of sequences of real numbers from the convergence of the entire sum. The proof is trivial.

Lemma 3.1.2. *Suppose that* $\{x_k\}$ *and* $\{y_k\}$ *are sequences of real numbers. Furthermore, suppose that* $y_k \to y$*,*

$$\lim_{k\to\infty} (x_k + y_k) = x + y$$

and

$$\liminf_{k\to\infty} x_k \geq x.$$

Then

$$\lim_{k\to\infty} x_k = x.$$

The third and last lemma deals with weak convergence of discrete states:

Lemma 3.1.3. *Suppose the sequence $\{U_k\}$ with $U_k \in \mathbb{U}_k$ converges weakly to $u \in \mathbb{U}$, i.e. $U_k \rightharpoonup u$. Then, we also have $S_k U_k \rightharpoonup Su$ in $\mathbb{Y}$.*

Proof. First of all, we observe that thanks to (A2) we have for every $v \in \mathbb{Y}$ a sequence $\{V_k\}$ with $V_k \in \mathbb{Y}_k$ and $V_k \to v$ strongly in $\mathbb{Y}$ as $k \to \infty$. Besides, since

$$\|S_k U_k\|_{\mathbb{Y}} \lesssim \|U_k\|_{\mathbb{U}}\,,$$

compare our stability assumption (Pr7), the sequence $\{S_k U_k\}$ is bounded in $\mathbb{Y}$. After all, $\{U_k\}$ is bounded, because it is weakly convergent. Thus, we can pick a subsequence $\{S_{k_l} U_{k_l}\}$ which is weakly convergent to an element $\tilde{y} \in \mathbb{Y}$. Basically, we now have to show that $\tilde{y} = Su$: Taking the previously introduced sequence $\{V_k\}$, we can conclude

$$\mathcal{B}[\tilde{y}, v] \leftarrow \mathcal{B}[S_{k_l} U_{k_l}, V_{k_l}] = (U_{k_l}, V_{k_l})_{\mathbb{U}} \to (u, v)_{\mathbb{U}},\ l \to \infty$$

Here we used strong convergence of $V_k \to v$ as well as weak convergence $U_{k_l} \rightharpoonup u$ and $S_{k_l} U_{k_l} \rightharpoonup \tilde{y}$ as $l \to \infty$. Consequently, because $v \in \mathbb{Y}$ is arbitrary, we have

$$\mathcal{B}[\tilde{y}, v] = (u, v)_{\mathbb{U}}\ \forall v \in \mathbb{Y}.$$

This is tantamount to $\tilde{y} = Su$. However, as yet, we only have convergence of a subsequence $S_{k_l} U_{k_l} \rightharpoonup Su$, $l \to \infty$. To extend convergence to the entire sequence, we again take advantage of the fact that the limit Su is unique. Hence, harnessing the property that every weakly convergent subsequence of $\{S_k U_k\}$ converges to the same limit Su we deduce weak convergence for the entire sequence $\{S_k U_k\}$ as $k \to \infty$ with limit Su, compare again the arguments of Lemma 2.1.5. □

3.2 Properties of the Regularised Problems

We will start by listing and proving a number of important properties of the continuous regularised problem (P^{ε}).

3.2.1 The Continuous Regularised Problem

First, let us recall the continuous regularised problem:

$$\begin{aligned} \min_{u\in\mathbb{U},y\in\mathbb{Y},v\in L_2(\Omega,\mathbb{R}^m)} \frac{1}{2}\|y-y_d\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon}\|v\|_{L_2(\Omega,\mathbb{R}^m)}^2 \\ \text{s.t.} \\ \mathcal{B}[y,w] = (u,w)_{\mathbb{U}} \quad \forall w\in\mathbb{Y} \\ \text{and} \\ u\in\mathcal{U} \\ y_c - y - \varepsilon v \in C \end{aligned}$$

For the continuous regularised and unregularised problems, we define an optimal value function a linking both, which - as the reader will find out later - will play a central role in deriving and formulating conditions for the convergence of $\bar{U}_k \to \bar{u}$.
The optimal value function $a : [0,1] \to \mathbb{R}^{\geq 0}$ is given by:

$$a(\varepsilon) := \min_{(u,v)\in\mathbb{U}^{\varepsilon,ad}} f^\varepsilon(u,v),\ \varepsilon > 0$$

and

$$a(0) := \min_{u\in\mathbb{U}^{ad}} f(u)$$

Naturally, we are interested in the properties of this function, the most important of which are recorded in the theorem below:

Theorem 3.2.1 (properties of the optimal value function)**.** *The optimal value function a is uniformly bounded, i.e. $a(\varepsilon) \leq C$ for all $\varepsilon \in [0,1]$ with C independent of ε, continuous and monotonically decreasing on $[0,1]$. Besides, for every $\varepsilon \in [0,1]$ we have as $\varepsilon_k \to \varepsilon$:*

$$(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k}, \bar{y}^{\varepsilon_k}) \to (\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon}, \bar{y}^{\varepsilon}) \ \text{ in } \mathbb{U}\times L_2(\Omega,\mathbb{R}^m)\times\mathbb{Y} \ \ \forall\varepsilon\in[0,1],\ \varepsilon_k\to\varepsilon \tag{3.2.1}$$

with the convention that $\bar{v}^0 := 0$.
It is differentiable on $(0,1)$ with the derivative $a'(\varepsilon)$ given by

$$a'(\varepsilon) = -\frac{3}{2\varepsilon^2}\|\bar{v}^\varepsilon\|^2$$

Furthermore, a is an element of the Sobolev space $W_1^1(0,1)$.

The proof is rather lengthy and technical, therefore we will split it into several lemmata in the following way:

- boundedness of a on $[0,1]$, continuity of a on $(0,1]$ and relation (3.2.1) on $(0,1]$, **Lemma 3.2.2**

- continuity of a at 0 and relation (3.2.1) at 0, **Lemma 3.2.3**
- differentiability of a on $(0,1)$ and $a \in W_1^1(0,1)$, **Lemma 3.2.4**

The lemmata Lemma 3.2.2, 3.2.3 and 3.2.4 below combined will then give Theorem 3.2.1. We will tackle the proof of continuity of a on $(0,1]$ and the uniform boundedness of a first:

Lemma 3.2.2 (continuity of a)**.** *The optimal value function a is continuous on $(0,1]$ and uniformly bounded on $[0,1]$. Besides, for every $\varepsilon \in (0,1]$ we have as $\varepsilon_k \to \varepsilon$:*

$$(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k}, \bar{y}^{\varepsilon_k}) \to (\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon}, \bar{y}^{\varepsilon}) \ \text{ in } \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times \mathbb{Y},\ \varepsilon_k \to \varepsilon.$$

Proof. We observe that there holds

$$a(\varepsilon) \le a(0) \ \ \forall \varepsilon \in [0,1], \tag{3.2.2}$$

because $(\bar{u}, 0) \in \mathbb{U}^{\varepsilon,ad}$.
This is the uniform boundedness property postulated in Theorem 3.2.1 and Lemma 3.2.2. Let us now tackle continuity of a on $(0,1]$: Taking an arbitrary point $\varepsilon \in (0,1)$ and a sequence $\{\varepsilon_k\} \subset (0,1)$ with $\varepsilon_k \to \varepsilon$, we can use (3.2.2) to deduce

$$a(\varepsilon_k) \le a(0) \ \ \forall k,$$

which, in particular, yields

$$\frac{\nu}{2} \|\bar{u}^{\varepsilon_k}\|_{\mathbb{U}}^2,\ \frac{1}{2\varepsilon_k} \|\bar{v}^{\varepsilon_k}\|^2 \le a(0).$$

Thus, there exists a weakly convergent subsequence of $\{(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k})\}$ in $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$ with weak limit $(\tilde{u}, \tilde{v})$. For notational convenience we will not distinguish between the subsequence and the sequence. Later, we will show that in fact this can be done w.l.o.g.
The proof now takes the following steps:

1. Show $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$.
2. Deduce $(\tilde{u}, \tilde{v}) = (\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon})$
3. Prove strong convergence for the entire sequence: $(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k}) \to (\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon})$ as $\varepsilon_k \to \varepsilon$.

Step 1: Let us now show that $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$: For the weak limit $\tilde{u}$ we know that $\tilde{u} \in \mathcal{U}$, because $\{\bar{u}^{\varepsilon_k}\} \subset \mathcal{U}$ and $\mathcal{U}$ is weakly closed. We also have that $S\bar{u}_k^{\varepsilon} \rightharpoonup S\tilde{u}$ (weak continuity of S) and because C is weakly closed, too, we realise that

$$C \ni y_c - S\bar{u}^{\varepsilon_k} - \varepsilon_k \bar{v}^{\varepsilon_k} \rightharpoonup y_c - S\tilde{u} - \varepsilon\tilde{v} \in C,\ \varepsilon_k \to \varepsilon,$$

Consequently, $\tilde{u} \in \mathcal{U}$ and $y_c - S\tilde{u} - \tilde{v} \in C$, hence $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$.

Step 2: $(\tilde{u},\tilde{v}) = (\bar{u}^\varepsilon,\bar{v}^\varepsilon)$: We now know that

$$\|(y,u,v)\|_\nu := (\frac{1}{2}\|y\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2 + \frac{1}{2}\|v\|^2)^{1/2}$$

defines a Hilbert space norm on $\mathbb{W}\times\mathbb{U}\times L_2(\Omega,\mathbb{R}^m)$ which is equivalent to the canonical norm given by

$$\|(y,u,v)\|_{\mathbb{W}\times\mathbb{U}\times L_2(\Omega,\mathbb{R}^m)} := (\|y\|_{\mathbb{W}}^2 + \|u\|_{\mathbb{U}}^2 + \|v\|^2)^{1/2}.$$

Harnessing weak lower semi-continuity for any Hilbert space norm and thus in particular for $\|\cdot\|_\nu^2$ as well as $\frac{1}{\sqrt{\varepsilon_k}}\bar{v}^{\varepsilon_k} \rightharpoonup \frac{1}{\sqrt{\varepsilon}}\tilde{v}$, we obtain:

$$\liminf_{\varepsilon_k\to\varepsilon} \frac{1}{2}\|\bar{y}^{\varepsilon_k} - y_d\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|\bar{u}^{\varepsilon_k}\|_{\mathbb{U}}^2 + \frac{1}{2}\left\|\frac{1}{\sqrt{\varepsilon^k}}\bar{v}^{\varepsilon_k}\right\|^2 \geq f^\varepsilon(\tilde{u},\tilde{v}). \tag{3.2.3}$$

Thanks to $(\tilde{u},\tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$ we immediately discern

$$\liminf_{\varepsilon_k\to\varepsilon} a(\varepsilon_k) \geq f^\varepsilon(\tilde{u},\tilde{v}) \geq f^\varepsilon(\bar{u}^\varepsilon,\bar{v}^\varepsilon) = a(\varepsilon).$$

Furthermore, we know that

$$(\bar{u}^\varepsilon, \frac{\varepsilon}{\varepsilon_k}\bar{v}^\varepsilon) \in \mathbb{U}^{\varepsilon_k,ad}$$

and as a consequence

$$a(\varepsilon_k) \leq f^{\varepsilon_k}(\bar{u}^\varepsilon, \frac{\varepsilon}{\varepsilon_k}\bar{v}^\varepsilon) = a(\varepsilon) + \frac{\varepsilon^3 - \varepsilon_k^3}{2\varepsilon\varepsilon_k^3}\|\bar{v}^\varepsilon\|^2. \tag{3.2.4}$$

Drawing the lim inf on each side in (3.2.4), we realise

$$\liminf_{\varepsilon_k\to\varepsilon} a(\varepsilon_k) \leq \liminf_{\varepsilon_k\to\varepsilon}(a(\varepsilon) + \frac{\varepsilon^3 - \varepsilon_k^3}{2\varepsilon\varepsilon_k^3}\|\bar{v}^\varepsilon\|^2) = a(\varepsilon).$$

Thus, recalling (3.2.3), we gain

$$\liminf_{\varepsilon_k\to\varepsilon} a(\varepsilon_k) = a(\varepsilon).$$

Ultimately, we deduce:

$$f^\varepsilon(\bar{u}^\varepsilon,\bar{v}^\varepsilon) = a(\varepsilon) = \liminf_{\varepsilon_k\to\varepsilon} a(\varepsilon_k) \geq f^\varepsilon(\tilde{u},\tilde{v})$$

Optimality and uniqueness of $(\bar{u}^\varepsilon,\bar{v}^\varepsilon)$ then yields $(\tilde{u},\tilde{v}) = (\bar{u}^\varepsilon,\bar{v}^\varepsilon)$.

Step 3: strong convergence of the entire sequence: Let us prove weak convergence first.

The arguments detailed in steps 1 and 2 are valid for any subsequence of $\{(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k})\}$. Besides, the limit $(\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon})$ is unique. Consequently, employing Lemma 2.1.5, we deduce weak convergence of the entire sequence, i.e.

$$\bar{u}^{\varepsilon_k} \rightharpoonup \bar{u}^{\varepsilon},\ \bar{v}^{\varepsilon_k} \rightharpoonup \bar{v}^{\varepsilon},\ \varepsilon_k \to \varepsilon.$$

To prove strong convergence of the control u and virtual control v, we employ the optimality condition (2.2.25) and the fact that

$$(\bar{u}^{\varepsilon}, \frac{\varepsilon}{\varepsilon_k}\bar{v}^{\varepsilon}) \in \mathbb{U}^{\varepsilon_k,ad}, \;\; (\bar{u}^{\varepsilon_k}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon_k}) \in \mathbb{U}^{\varepsilon,ad}$$

to obtain

$$\begin{aligned}
(\bar{p}^{\varepsilon} + \nu\bar{u}^{\varepsilon}, \bar{u}^{\varepsilon_k} - \bar{u}^{\varepsilon})_{\mathbb{U}} + \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon}{\varepsilon_k}\bar{v}^{\varepsilon_k} - \bar{v}^{\varepsilon}) \geq 0\\
(\bar{p}^{\varepsilon_k} + \nu\bar{u}^{\varepsilon_k}, \bar{u}^{\varepsilon} - \bar{u}^{\varepsilon_k})_{\mathbb{U}} + \frac{1}{\varepsilon_k}(\bar{v}^{\varepsilon_k}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k}) \geq 0.
\end{aligned}$$

Adding and rearranging these inequalities, we derive

$$\begin{aligned}
0 \leq (\bar{p}^{\varepsilon} - \bar{p}^{\varepsilon_k}, \bar{u}^{\varepsilon_k} - \bar{u}^{\varepsilon})_{\mathbb{U}} + \nu(\bar{u}^{\varepsilon} - \bar{u}^{\varepsilon_k}, \bar{u}^{\varepsilon_k} - \bar{u}^{\varepsilon})_{\mathbb{U}}\\
+ \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon}{\varepsilon_k}v^{\varepsilon_k} - \bar{v}^{\varepsilon}) + \frac{1}{\varepsilon_k}(\bar{v}^{\varepsilon_k}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k})
\end{aligned}$$

Using $\bar{p}^{\varepsilon(k)} = S^*(\bar{u}^{\varepsilon(k)} - y_d)$ and doing further calculations, we arrive at

$$\begin{aligned}
\nu\left\|\bar{u}^{\varepsilon} - \bar{u}^{\varepsilon_k}\right\|^2_{\mathbb{U}} + \left\|\bar{y}^{\varepsilon_k} - \bar{y}^{\varepsilon}\right\|^2_{\mathbb{W}} &\leq \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon}{\varepsilon_k}v^{\varepsilon_k} - \bar{v}^{\varepsilon}) + \frac{1}{\varepsilon_k}(\bar{v}^{\varepsilon_k}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k})\\
&= \frac{1}{\varepsilon_k}(\bar{v}^{\varepsilon_k} - \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k}) + \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k})\\
&+ \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon}{\varepsilon_k}v^{\varepsilon_k} - \bar{v}^{\varepsilon})\\
&= -\frac{1}{\varepsilon_k}\left\|\bar{v}^{\varepsilon_k} - \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon}\right\|^2 + \frac{1}{\varepsilon}(\bar{v}^{\varepsilon}, \frac{\varepsilon - \varepsilon_k}{\varepsilon_k}\bar{v}^{\varepsilon_k} - \frac{\varepsilon_k - \varepsilon}{\varepsilon}\bar{v}^{\varepsilon})
\end{aligned}$$

Hence

$$\nu\left\|\bar{u}^{\varepsilon} - \bar{u}^{\varepsilon_k}\right\|^2_{\mathbb{U}} + \left\|\bar{y}^{\varepsilon_k} - \bar{y}^{\varepsilon}\right\|^2_{\mathbb{W}} + \frac{1}{\varepsilon_k}\left\|\bar{v}^{\varepsilon_k} - \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon}\right\|^2 \leq \frac{\varepsilon_k - \varepsilon}{\varepsilon\varepsilon_k}(\bar{v}^{\varepsilon}, \frac{\varepsilon_k}{\varepsilon}\bar{v}^{\varepsilon} - \bar{v}^{\varepsilon_k}).$$

Due to weak convergence $\bar{v}^{\varepsilon_k} \rightharpoonup \bar{v}^{\varepsilon}$ and $\varepsilon_k \to \varepsilon$, the right-hand side in the inequality above tends to zero. This yields strong convergence $\bar{u}^{\varepsilon_k} \to \bar{u}^{\varepsilon}$ and $\bar{v}^{\varepsilon_k} \to \bar{v}^{\varepsilon}$ as $\varepsilon_k \to \varepsilon$. □

Let us now extend this continuity result to the full closed interval $[0, 1]$.

Lemma 3.2.3 (continuity of a at 0). *The optimal value function a is continuous at* 0. *Besides,* *as* $\varepsilon \to 0$ *we have*

$$(\bar{u}^{\varepsilon}, \bar{v}^{\varepsilon}) \to (\bar{u}, 0) \ \ in\ \mathbb{U} \times L_2(\Omega, \mathbb{R}^m),\ \varepsilon \to 0$$

Proof. Let $\{\varepsilon_k\}$ be an arbitrary null sequence with $\varepsilon_k > 0$ for all k. As in the proof of Lemma 3.2.2 we use (3.2.2) to deduce boundedness of the sequence $\{(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k})\}$ in $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$. Thus, there exists a weakly convergent subsequence of $\{(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k})\}$ with weak limit $(\tilde{u},\ \tilde{v})$. For notational convenience, we will not distinguish between a subsequence of $\{(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k})\}$ and the entire sequence. Later, we will demonstrate that this can in fact be done w.l.o.g.
Since $\mathcal{U}$ is closed and convex and hence weakly closed, we observe that $\tilde{u} \in \mathcal{U}$. $\bar{v}^{\varepsilon_k}$ is bounded and thus $\varepsilon_k \bar{v}^{\varepsilon_k} \to 0$ as $\varepsilon_k \to 0$. Taking advantage of the fact that C is closed and convex and as a consequence weakly closed and also of S being linear and continuous and hence weakly continuous, we can conclude

$$C \ni y_c - S\tilde{u} \leftharpoonup y_c - S\bar{u}^{\varepsilon_k} - \varepsilon_k \bar{v}^{\varepsilon_k} \in C,\ \varepsilon_k \to 0$$

This means $\tilde{u} \in \mathbb{U}^{ad}$. Combining this with (3.2.2) and the fact that f is weakly lower-semicontinuous, we can conclude:

$$\begin{aligned} f(\bar{u}) \le f(\tilde{u}) &\le \liminf_{\varepsilon_k \to 0} f(\bar{u}^{\varepsilon_k}) \le \liminf_{\varepsilon_k \to 0} f^{\varepsilon_k}(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k}) \\ \le \limsup_{\varepsilon_k \to 0} f^{\varepsilon_k}(\bar{u}^{\varepsilon_k}, \bar{v}^{\varepsilon_k}) &= \limsup_{\varepsilon_k \to 0} a(\varepsilon_k) \le a(0) = f(\bar{u}). \end{aligned} \tag{3.2.5}$$

This entails $f(\tilde{u}) = f(\bar{u})$, and because $\bar{u}$ is the unique solution to (P), we immediately deduce $\tilde{u} = \bar{u}$.
The conclusions above are true for any subsequence of $\bar{u}^{\varepsilon_k}$. Together with the fact that the limit $\tilde{u} = \bar{u}$ is unique, we obtain weak convergence for the entire sequence $\bar{u}^{\varepsilon_k} \rightharpoonup \bar{u}$ as $\varepsilon_k \to 0$, compare again Lemma 2.1.5.
Besides, we have

$$\frac{1}{2\varepsilon_k} \|\bar{v}^{\varepsilon_k}\|^2 \le a(0),$$

which in turn implies strong convergence $\bar{v}^{\varepsilon_k} \to 0$ as $\varepsilon_k \to 0$ for every subsequence and thus, by arguments completely analogous to the ones used for for $\bar{u}^{\varepsilon_k}$, we obtain strong convergence of the entire sequence $\bar{v}^{\varepsilon_k} \to 0$ as $\varepsilon_k \to 0$. Estimate (3.2.5) can also be interpreted in the following fashion (again in combination with (3.2.2))

$$a(0) \le \liminf_{\varepsilon_k \to 0} a(\varepsilon_k) \le \limsup_{\varepsilon_k \to 0} a(\varepsilon_k) \le a(0).$$

Hence $\limsup\limits_{\varepsilon_k \to 0} a(\varepsilon_k)$ and $\liminf\limits_{\varepsilon_k \to 0} a(\varepsilon_k)$ coincide and are equal to $a(0)$. We obtain:

$$\lim_{\varepsilon_k \to 0} a(\varepsilon_k) = a(0).$$

Therefore a is continuous at 0.
Let us now show that apart from $\bar{v}^{\varepsilon_k} \to 0$ there also holds $\bar{u}^{\varepsilon_k} \to \bar{u}$ as $\varepsilon_k \to 0$. To do so, we recall the optimality condition (2.2.25). Since $(\bar{u}, 0) \in \mathbb{U}^{\varepsilon,ad}$, there holds

$$(\bar{p}^{\varepsilon_k} + \nu\bar{u}^{\varepsilon_k}, \bar{u} - \bar{u}^{\varepsilon_k})_{\mathbb{U}} - \frac{1}{\varepsilon_k} \|\bar{v}^{\varepsilon_k}\|^2 \geq 0 = (\bar{p} + \nu\bar{u}, \bar{u}^{\varepsilon_k} - \bar{u})_{\mathbb{U}} + (\bar{p} + \nu\bar{u}, \bar{u} - \bar{u}^{\varepsilon_k})_{\mathbb{U}}.$$

Rearranging this inequality yields:

$$\nu \|\bar{u} - \bar{u}^{\varepsilon_k}\|_{\mathbb{U}}^2 + \|\bar{y}^{\varepsilon_k} - \bar{y}\|_{\mathbb{W}}^2 + \frac{1}{\varepsilon_k} \|\bar{v}^{\varepsilon_k}\|^2 \leq |(\bar{p} + \nu\bar{u}, \bar{u} - \bar{u}^{\varepsilon_k})_{\mathbb{U}}|$$

Since $\bar{u}^{\varepsilon_k} \rightharpoonup \bar{u}$ as $\varepsilon_k \to 0$, the right-hand side tends to 0. This completes the proof. □

Let us now turn to the question of differentiability of a:

Lemma 3.2.4 (differentiability of a)**.** *The optimal value function a is differentiable on* $(0, 1)$ *with the derivative given by*

$$a'(\varepsilon) = -\frac{3}{2\varepsilon^2} \|\bar{v}^\varepsilon\|^2$$

As a consequence, a is monotonically decreasing on $[0, 1]$.
Furthermore, $a^{'} \in L_1(0, 1)$, and thus $a \in W_1^1(0, 1)$.

Proof. We remark that for $h \in \mathbb{R}$ with $\varepsilon + h, \varepsilon \in (0, 1)$ there holds:

$$(\bar{u}^\varepsilon, \frac{\varepsilon}{\varepsilon + h}\bar{v}^\varepsilon) \in \mathbb{U}^{\varepsilon+h,ad}$$

and

$$(\bar{u}^{\varepsilon+h}, \frac{\varepsilon + h}{\varepsilon}\bar{v}^{\varepsilon+h}) \in \mathbb{U}^{\varepsilon,ad}.$$

This allows us to estimate in the following fashion for arbitrary $h > 0$:

$$\begin{aligned} \frac{1}{h}(a(\varepsilon + h) - a(\varepsilon)) &= \frac{1}{h}(f^{\varepsilon+h}(\bar{u}^{\varepsilon+h}, \bar{v}^{\varepsilon+h}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)) \\ &\leq \frac{1}{h}(f^{\varepsilon+h}(\bar{u}^\varepsilon, \frac{\varepsilon}{\varepsilon + h}\bar{v}^\varepsilon) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)) \\ &= \frac{1}{h}(\frac{\varepsilon^2}{2(\varepsilon + h)^3} - \frac{1}{2\varepsilon}) \|\bar{v}^\varepsilon\|^2 \\ &= (\frac{-3\varepsilon h - 3\varepsilon^2 - h^2}{2(\varepsilon + h)^3\varepsilon}) \|\bar{v}^\varepsilon\|^2, \end{aligned} \tag{3.2.6}$$

which allows us to conclude:

$$\limsup_{h \searrow 0}(\frac{1}{h}(a(\varepsilon + h) - a(\varepsilon))) \leq -\frac{3}{2\varepsilon^2} \|\bar{v}^\varepsilon\|^2. \tag{3.2.7}$$

Conversely, we can estimate in the following way, again for $h > 0$:

$$\begin{aligned}\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon)) &= \frac{1}{h}(f^{\varepsilon+h}(\bar{u}^{\varepsilon+h},\bar{v}^{\varepsilon+h})-f^{\varepsilon}(\bar{u}^{\varepsilon},\bar{v}^{\varepsilon}))\\ &\geq \frac{1}{h}(f^{\varepsilon+h}(\bar{u}^{\varepsilon+h},\bar{v}^{\varepsilon+h})-f^{\varepsilon}(\bar{u}^{\varepsilon+h},\frac{\varepsilon+h}{\varepsilon}\bar{v}^{\varepsilon+h}))\\ &= \frac{1}{h}(\frac{1}{2(\varepsilon+h)}-\frac{(\varepsilon+h)^2}{2\varepsilon^3})\left\|\bar{v}^{\varepsilon+h}\right\|^2\\ &= (\frac{-3\varepsilon^2-3\varepsilon h-h^2}{2\varepsilon^3(\varepsilon+h)})\left\|\bar{v}^{\varepsilon+h}\right\|^2\end{aligned} \tag{3.2.8}$$

Hence, employing Lemma 3.2.2, especially $\bar{v}^{\varepsilon+h}\to\bar{v}^{\varepsilon}$ as $h\to 0$, we obtain

$$\liminf_{h\searrow 0}(\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon))) \geq -\frac{3}{2\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|^2 \tag{3.2.9}$$

Combining (3.2.7) and (3.2.9), we derive

$$\lim_{h\searrow 0}(\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon))) = -\frac{3}{2\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|^2 \tag{3.2.10}$$

Now, let us tackle the case $h < 0$. Proceeding as in (3.2.6), we obtain

$$\liminf_{h\nearrow 0}(\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon))) \geq -\frac{3}{2\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|^2.$$

Observe that in step 2 in (3.2.6) we now estimate from below, because $h < 0$. Likewise, mirroring (3.2.8), we gain

$$\limsup_{h\nearrow 0}(\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon))) \leq -\frac{3}{2\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|^2.$$

Hence

$$\lim_{h\nearrow 0}(\frac{1}{h}(a(\varepsilon+h)-a(\varepsilon))) = -\frac{3}{2\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|^2 \tag{3.2.11}$$

Combining (3.2.10) and (3.2.11), we can conclude that a is differentiable on $(0,1)$. The monotonicity of a follows from the fact that $a' \leq 0$ on $(0,1)$.
We still have to show that a belongs to $W_1^1(0,1)$. Since a is continuous on $[0,1]$, a is an element of $L_1(0,1)$. Besides, $a \in C^1(0,1)$ and $a^{'} \in L_{1,loc}(0,1)$; hence $a^{'}$ is the weak derivative of a. Thus, the only thing we have to show is that $a' \in L_1(0,1)$. We first investigate the interval $(\delta,1]$ with an arbitrary, but fixed $\delta > 0$. Thanks to the differentiability of a, we can invoke the mean value theorem and write:

$$a(1)-a(\delta) = \int_{\delta}^{1} -\frac{3}{2s^2}\left\|\bar{v}^{s}\right\|^2\, ds$$

Now we can draw the limit $\delta \to 0$, utilising continuity of a at 0. Thus, we obtain:

$$\begin{aligned} a(1) - a(0) &= \lim_{\delta \to 0} a(1) - a(\delta) \\ &= \lim_{\delta \to 0} \int\limits_{\delta}^{1} -\frac{3}{2s^2} \left\| \bar{v}^s \right\|^2 \, ds. \end{aligned}$$

Thus, by definition of the improper integral, there holds

$$a(1) - a(0) = \lim_{\delta \to 0} \int\limits_{\delta}^{1} -\frac{3}{2s^2} \left\| \bar{v}^s \right\|^2 \, ds = \int\limits_{0}^{1} -\frac{3}{2s^2} \left\| \bar{v}^s \right\|^2 \, ds. \tag{3.2.12}$$

Since the derivate $a^{'}$ never changes sign on $[0, 1]$, this is tantamount to $a^{'}$ belonging to $L_1(0, 1)$. Hence, $a \in W_1^1(0, 1)$. □

Combining the lemmata Lemma 3.2.2, 3.2.3 and 3.2.4 yields the proof of Theorem 3.2.1. Having collected these important properties, we can now turn to the discrete regularised problem (P_k^ε), where we prove a result completely analogous to Theorem 3.2.1.

3.2.2 The Discrete Regularised Problem

First, let us recall the discrete regularised problem (P_k^ε) for fixed k and ε.

$$\begin{aligned} \min_{U \in \mathbb{U}_k, Y \in \mathbb{Y}_k, V \in \mathbb{V}_k} & \frac{1}{2} \left\| Y - y_d \right\|_{\mathbb{W}}^2 + \frac{\nu}{2} \|U\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon} \left\| V \right\|^2 \\ & \text{s.t.} \\ \mathcal{B}[Y, W] &= (U, W)_{\mathbb{U}} \quad \forall W \in \mathbb{Y}_k \\ & \text{and} \\ U &\in \mathcal{U}_k \\ I_k y_c - Y - \varepsilon V &\in C_{\mathbb{V}_k}, \end{aligned}$$

As in the continuous case, we first define an optimal value function $a_k : [0, 1] \to [0, \infty)$ for every fixed k:

$$a_k(\varepsilon) := \min_{(U,V) \in \mathbb{U}_k^{\varepsilon, ad}} f_k^\varepsilon(U, V), \quad \varepsilon > 0$$

and

$$a_k(0) := \min_{U \in \mathbb{U}_k^{ad}} f_k(U).$$

The properties of a, cf Theorem 3.2.1, are reflected by its discrete counterpart a_k as the next theorem shows:

Theorem 3.2.5 (properties of the discrete optimal value function)**.** *On* $[0,1]$ *the discrete optimal value function* a_k *is uniformly bounded, i.e.* $a_k(\varepsilon) \leq C$, $\varepsilon \in [0,1]$*, with* C *independent of* ε *and* k*, continuous and monotonically decreasing. Besides, for* $\varepsilon \in [0,1]$ *we have as* $\varepsilon_l \to \varepsilon$*:*

$$(\bar{U}_k^{\varepsilon_l}, \bar{V}_k^{\varepsilon_l}, \bar{Y}_k^{\varepsilon_l}) \to (\bar{U}_k^{\varepsilon}, \bar{V}_k^{\varepsilon}, \bar{Y}_k^{\varepsilon}) \;\; in \; \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times \mathbb{Y} \;\; \forall \varepsilon \in [0,1], \; \varepsilon_l \to \varepsilon \tag{3.2.13}$$

with the convention that $\bar{V}_k^0 := 0$*.*
In addition, a_k *is differentiable on* $(0,1)$*. The derivative* $a_k'(\varepsilon)$ *is given by*

$$a_k^{'}(\varepsilon) = -\frac{3}{2\varepsilon^2} \left\| \bar{V}_k^{\varepsilon} \right\|^2$$

a_k *is also an element of the Sobolev space* $W_1^1(0,1)$*. Besides, the norm* $\|a_k\|_{W_1^1(0,1)}$ *is bounded independently of* k *and for all* $r > 0$ *we have*

$$|a_k'(\varepsilon)| \lesssim \frac{1}{r} \, \forall \varepsilon \in [r,1] \tag{3.2.14}$$

with a constant independent of k*.*

Proof. Proving for every fixed k that a_k is continuous on $[0,1]$, (3.2.13) holds, and that a_k differentiable on $(0,1)$ as well as an element of $W_1^1(0,1)$ can be done exactly in the same way as on the continuous level, Theorem 3.2.1 which we had split up into the lemmata Lemma 3.2.2, 3.2.3 and 3.2.4. Thus, we now want to tackle the uniform bounds of a_k.
We observe that since $(\bar{U}_k, 0) \in \mathbb{U}_k^{\varepsilon,ad}$ for all $\varepsilon > 0$

$$|a_k(\varepsilon)| \leq f_k(\bar{U}_k) \;\; \forall \varepsilon \in [0,1]$$

Since $f_k(\bar{U}_k)$ is uniformly bounded thanks to Theorem 2.3.11, we have

$$|a_k(\varepsilon)| \lesssim 1, \;\; \forall \varepsilon \in [0,1]. \tag{3.2.15}$$

This immediately results in a uniform bound

$$\|a_k\|_{L_1(0,1)} \lesssim 1.$$

Let us now bound the L_1-norm of the derivative. The mean value theorem and continuity of

a_k imply for every $1 > \delta > 0$

$$\begin{aligned}|a_k(0) - a_k(1)| &= \lim_{\delta \to 0} (\int_{\delta}^{1} \frac{1}{\varepsilon^2} \left\|\bar{V}_k^\varepsilon\right\|^2 \, d\varepsilon) \\ &= \lim_{\delta \to 0} \left\|a_k'\right\|_{L_1(\delta,1)} \\ &= \left\|a_k'\right\|_{L_1(0,1)}\end{aligned}$$

Combining this with (3.2.15) yields the uniform bound for $\|a_k\|_{W_1^1(0,1)}$.
For the bound (3.2.14) we first recall that

$$\frac{1}{\varepsilon} \left\|\bar{V}_k^\varepsilon\right\|^2 \leq a_k(\varepsilon) \lesssim 1 \; \forall \varepsilon \in [0,1]$$

uniformly. Hence for all $\varepsilon \in [r, 1]$ with an arbitrary $r > 0$ we deduce

$$|a_k'(\varepsilon)| = \frac{3}{2\varepsilon^2} \left\|\bar{V}_k^\varepsilon\right\|^2 \lesssim \frac{1}{r}$$

□

3.2.3 Convergence Analysis for the Regularised Problems

At the start of this chapter we explained that the question of $\bar{U}_k \to \bar{u}$ is closely linked to a smoothness property of the regularised problems. To be more precise, in Theorem 3.3.1 we will demonstrate that

$$\bar{U}_k \to \bar{u} \;\Leftrightarrow\; a_k \to a \text{ in } W_1^1(0,1), \;\; k \to \infty.$$

We will not be able to prove this result straightaway, first we need to collect several pointwise convergence results for the sequence of optimal value functions $\{a_k\}$, where by pointwise convergence results we mean convergence results for arbitrary but fixed regularisation parameters $\varepsilon \in (0,1)$. This is the main aim of this section captured by the following main result:

Theorem 3.2.6 (pointwise convergence)**.** *For all fixed $\varepsilon \in (0,1]$ there holds*

$$\bar{U}_k^\varepsilon \to \bar{u}^\varepsilon \textit{ in } \mathbb{U} \quad \bar{V}_k^\varepsilon \to \bar{v}^\varepsilon \textit{ in } L_2(\Omega, \mathbb{R}^m), \;\; k \to \infty$$

As a consequence, the following ***pointwise convergence*** *results hold for all fixed but arbitrary $\varepsilon \in (0,1]$:*

$$a_k(\varepsilon) \to a(\varepsilon), \; k \to \infty$$

and

$$a_k'(\varepsilon) \to a'(\varepsilon), \; k \to \infty$$

In fact both convergences are equivalent, i.e.

$$a_k \to a,\, a_k' \to a' \text{ ptw. on } (0,1] \;\Leftrightarrow$$
$$(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \to (\bar{u}^\varepsilon, \bar{v}^\varepsilon) \text{ in } \mathbb{U} \times L_2(\Omega, \mathbb{R}^m),\ \varepsilon \in (0,1],\ k \to \infty \qquad (3.2.16)$$

The proof is again fairly technical. For the reader's convenience we will therefore split it up into several lemmata in the following way:

- Prove the existence of a weakly convergent subsequence of $\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\}$ with weak limit $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$, **Lemma 3.2.9**.
- Prove weak convergence of the entire sequence $\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\}$ to $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$, **Lemma 3.2.10**.
- Prove strong convergence of the entire sequence $\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\}$ to $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$. All the (other) results of Theorem 3.2.6 then follow. This is done in the actual **'Proof of Theorem 3.2.6'** further below.

starting with some observations about the dual problems to (P^ε) and (P_k^ε) respectively, compare also (2.2.28), which will be of great assistance in said proof.
First of all, we recall the definition of the continuous Lagrangian:

$$\mathcal{L}^\varepsilon : \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times L_2(\Omega, \mathbb{R}^m) \ni (u, v, \theta) \mapsto f^\varepsilon(u, v) - (\theta, Su - \varepsilon v - y_c) \qquad (3.2.17)$$

and its discrete counterpart

$$\mathcal{L}_k^\varepsilon : \mathbb{U}_k \times \mathbb{V}_k \times \mathbb{V}_k \ni (U, V, \theta) \mapsto f^\varepsilon(U, V) - (\theta, S_k U - \varepsilon V - I_k y_c) \qquad (3.2.18)$$

With their help we can define the continuous and discrete dual problems. Let us commence with the continuous one:

$$\sup_{\theta \in C^-} \inf_{(u,v) \in \mathcal{U} \times L_2(\Omega, \mathbb{R}^m)} \mathcal{L}^\varepsilon(u, v, \theta). \qquad (DP_c^\varepsilon)$$

The discrete problem can be formulated in an analogous fashion:

$$\sup_{\theta \in C_{\mathbb{V}_k}^-} \inf_{(U,V) \in \mathcal{U}_k \times \mathbb{V}_k} \mathcal{L}_k^\varepsilon(U, V, \theta). \qquad (DP_d^\varepsilon)$$

An important aspect of regularisation was that the regularised problems guarantee the existence of an $L_2(\Omega, \mathbb{R}^m)$-multiplier as we already discussed in Section 2.2.2 and Theorem 2.2.13. To make the results of this section easy to follow we repeat Theorem 2.2.13 before transferring these existence result to the discrete level (DP_d^ε):

Theorem 3.2.7. *Let (P^ε) for a fixed $\varepsilon > 0$ be given. Then there exists a unique element*

$\bar\theta^\varepsilon \in C^-$ such that $\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon$ solves the following Karush-Kuhn-Tucker (KKT) system:

$$\begin{aligned}(S^*(\bar y^\varepsilon - y_d) + \nu\bar u^\varepsilon, u - \bar u^\varepsilon)_{\mathbb{U}} - (\bar\theta^\varepsilon, Su - \bar y^\varepsilon) &\ge 0 \quad \forall u \in \mathcal{U}\\ -\varepsilon^2\bar\theta^\varepsilon + \bar v^\varepsilon &= 0 \\ (\bar\theta^\varepsilon, \bar y^\varepsilon - y_c + \varepsilon\bar v^\varepsilon) &= 0.\end{aligned} \tag{3.2.19}$$

Furthermore, $(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon)$ solve the dual problem (DP_c^ε), for which the following equality holds:

$$f^\varepsilon(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon) = \mathcal{L}^\varepsilon(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon). \tag{3.2.20}$$

Besides, if $(\tilde u, \tilde v, \tilde\theta) \in \mathcal{U} \times L_2(\Omega, \mathbb{R}^m) \times L_2(\Omega, \mathbb{R}^m)$ solves (3.2.19) then

$$(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon) = (\tilde u, \tilde v, \tilde\theta)$$

Lastly, we have

$$\mathcal{L}^\varepsilon(u, v, \bar\theta^\varepsilon) \ge \mathcal{L}^\varepsilon(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon) \;\; \forall (u, v) \in \mathcal{U} \times L_2(\Omega, \mathbb{R}^m). \tag{3.2.21}$$

Proof. The only thing which has not yet been proven in Theorem 2.2.13 is (3.2.21). To realise this, note that for $\bar\theta^\varepsilon$, $(\bar u^\varepsilon, \bar v^\varepsilon)$ solve the inner minimisation in dual problem (DP_c^ε), i.e.

$$(\bar u^\varepsilon, \bar v^\varepsilon) = \operatorname*{arg\,min}_{(u,v)\in\mathcal{U}\times L_2(\Omega,\mathbb{R}^m)} \mathcal{L}^\varepsilon(u, v, \bar\theta^\varepsilon).$$

Since for the fixed $\bar\theta^\varepsilon$ the Lagrangian $\mathcal{L}^\varepsilon(\cdot,\cdot,\bar\theta^\varepsilon)$ is a convex function, the minimum is global. Thus:

$$\mathcal{L}^\varepsilon(u, v, \bar\theta^\varepsilon) \ge \mathcal{L}^\varepsilon(\bar u^\varepsilon, \bar v^\varepsilon, \bar\theta^\varepsilon)$$

□

Crucially, this result finds a ready counterpart on the discrete level:

Theorem 3.2.8. *Let (P_k^ε) be given. Then, for every fixed $\varepsilon > 0$ there exists a unique Lagrange multiplier $\bar\theta_k^\varepsilon \in C_{\mathbb{V}_k}^-$, which is also bounded in $L_2(\Omega, \mathbb{R}^m)$ independent of k (but **not** of ε), such that*

$$\begin{aligned}(S_k^*(\bar Y_k^\varepsilon - y_d) + \nu\bar U_k^\varepsilon, U - \bar U_k^\varepsilon)_{\mathbb{U}} - (\bar\theta_k^\varepsilon, S_kU - \bar Y_k^\varepsilon) &\ge 0 \quad \forall U \in \mathcal{U}_k\\ -\varepsilon^2\bar\theta_k^\varepsilon + \bar V_k^\varepsilon &= 0 \\ (\bar\theta_k^\varepsilon, \bar Y_k^\varepsilon - I_k y_c + \varepsilon\bar V_k^\varepsilon) &= 0.\end{aligned} \tag{3.2.22}$$

As in the continuous case, the triple $(\bar U_k^\varepsilon, \bar V_k^\varepsilon, \bar\theta_k^\varepsilon)$ solves the dual problem (DP_d^ε). In particular,

$$f_k^\varepsilon(\bar U_k^\varepsilon, \bar V_k^\varepsilon) = \mathcal{L}_k^\varepsilon(\bar U_k^\varepsilon, \bar V_k^\varepsilon, \bar\theta_k^\varepsilon) \tag{3.2.23}$$

Lastly, we have for every fixed k and $\varepsilon > 0$

$$\mathcal{L}_k^\varepsilon(U, V, \bar{\theta}_k^\varepsilon) \geq \mathcal{L}_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon) \;\; \forall (U, V) \in \mathcal{U}_k \times \mathbb{V}_k \tag{3.2.24}$$

Proof. The proof runs along the same lines as that of Theorem 2.2.13.
First of all, we define the discrete constraint mapping $M_k^\varepsilon : \mathcal{U}_k \times \mathbb{V}_k \to \mathbb{V}_k$ by

$$M_k^\varepsilon(U, V) := I_k y_c - \varepsilon V - S_k U.$$

Obviously, it is surjective as a mapping

$$M_k^\varepsilon : \mathcal{U}_k \times \mathbb{V}_k \to \mathbb{V}_k$$

In this setting, we can apply Theorem 2.2.10 as in the continuous case. The multiplier $\bar{\theta}_k^\varepsilon$ then belongs to $\mathbb{V}_k^*$ which, it being a Hilbert space with the standard $L_2(\Omega, \mathbb{R}^m)$ scalar product, can be identified with $\mathbb{V}_k$.
The KKT system (3.2.22) then readily follows, compare also (2.2.11) and (2.2.10). The fact that the Lagrange multiplier $\bar{\theta}_k^\varepsilon$ is unique is a consequence of the equation

$$\varepsilon^2 \theta_k^\varepsilon = \bar{V}_k^\varepsilon$$

in (3.2.22) and the fact that $\bar{V}_k^\varepsilon$ is unique, see Theorem 2.3.12.
The definition of the Lagrange multiplier Definition 2.2.5 entails that the triple $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon)$ does solve the dual problem (DP_d^ε). The relation (3.2.23) then follows from the complementary slackness condition

$$(\bar{\theta}_k^\varepsilon, S_k \bar{U}_k^\varepsilon + \varepsilon \bar{V}_k^\varepsilon - I_k y_c) = 0.$$

The relation (3.2.24) is proven completely analogously to (3.2.21). □

Let us now return to the proof of Theorem 3.2.6: As a quick reminder, here is our course of action:

- Prove the existence of a weakly convergent subsequence of $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\}$ with weak limit $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$, **Lemma 3.2.9**.

- Prove weak convergence of the entire sequence $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\}$ to $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$, Lemma **3.2.10**.

- Prove strong convergence of the entire sequence $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\}$ to $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$. All the (other) results of Theorem 3.2.6 then follow. This is done in the actual **'Proof of Theorem 3.2.6'** further below.

So let us now start by proving the existence of weakly convergent subsequences of $\left\{\bar{U}_k^\varepsilon\right\}$ and

$\{\bar{V}_k^\varepsilon\}$ and the feasibility of their weak limits for the continuous problem (P^ε). This is the subject of the next lemma:

Lemma 3.2.9. *The sequences* $\{\bar{U}_k^\varepsilon\}$ *and* $\{\bar{V}_k^\varepsilon\}$ *are bounded independent of* ε *and* k *in* $\mathbb{U}$ *and* $L_2(\Omega, \mathbb{R}^m)$ *respectively. Thus, there exist weakly convergent subsequences whose weak limits* $\tilde{u}$ *and* $\tilde{v}$ *fulfil*

$$(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}.$$

In particular, there holds

$$f^\varepsilon(\tilde{u}, \tilde{v}) \geq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon). \tag{3.2.25}$$

In addition, the sequence of Lagrange multipliers $\{\bar{\theta}_k^\varepsilon\}$ *is bounded independent of* k *in* $L_2(\Omega, \mathbb{R}^m)$*. In particular, there exists a weakly convergent subsequence with weak limit* $\tilde{\theta} \in C^-$*.*

Proof. The uniform boundedness (independent of ε and k) property of a_k in Theorem 3.2.5, compare also (2.3.12), immediately yields uniform boundedness of

$$\|\bar{U}_k^\varepsilon\|_{\mathbb{U}}^2, \|\bar{V}_k^\varepsilon\|^2 \lesssim 1.$$

Since $\mathbb{U}$ and $L_2(\Omega, \mathbb{R}^m)$ are Hilbert spaces there exist weakly convergent subsequences with weak limits $\tilde{u}$ and $\tilde{v}$ respectively. The corresponding sequence of states $\{\bar{Y}_k^\varepsilon\}$ is also bounded in $\mathbb{Y}$ thanks to continuity of S_k and Assumption (Pr7). The weak limit of this sequence is denoted by $\tilde{y}$. Thanks to Lemma 3.1.3 we gain $\tilde{y} = S\tilde{u} \leftharpoonup S_k\bar{U}_k^\varepsilon$. Employing (A3), we obtain $\tilde{u} \in \mathcal{U}$ and utilising weak closedness of C, we deduce

$$C \supset C_{\mathbb{V}_k} \ni I_k y_c - \varepsilon\bar{V}_k^\varepsilon - S_k\bar{U}_k^\varepsilon \rightharpoonup y_c - \varepsilon\tilde{v} - S\tilde{u},\ k \to \infty$$

Hence $(\tilde{u}, \tilde{v}) \in \mathbb{U}^{\varepsilon,ad}$. The relation

$$f^\varepsilon(\tilde{u}, \tilde{v}) \geq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$$

is a consequence of optimality of $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$.
The fact that $\{\bar{\theta}_k^\varepsilon\}$ is bounded independently of k follows from

$$-\frac{1}{\varepsilon^2}\bar{V}_k^\varepsilon + \bar{\theta}_k^\varepsilon = 0$$

in (3.2.22) and the fact that $\bar{V}_k^\varepsilon$ is bounded independently of k and ε. Thus, for every fixed ε there exists a weakly convergent subsequence with weak limit $\tilde{\theta} \in L_2(\Omega, \mathbb{R}^m)$. Let us now prove that, in fact, $\tilde{\theta} \in C^-$. To this end, we take for every function $y \in C$ its approximation

$H_k y \in C_{\mathbb{V}_k}$ with $H_k y \to y$, recall (A4). We obtain

$$0 \geq (\bar{\theta}_k^\varepsilon, H_k y) \to (\tilde{\theta}, y) \;\; \forall y \in C, \; k \to \infty,$$

because $\bar{\theta}_k^\varepsilon \rightharpoonup \tilde{\theta}$ and $H_k y \to y$ in $L_2(\Omega, \mathbb{R}^m)$ strongly as $k \to \infty$. Since $y \in C$ was arbitrary, we can conclude that $\tilde{\theta} \in C^-$. This completes the proof.

□

Having ascertained these important boundedness and feasibility results, we can turn our attention back to the proof of the claims of Theorem 3.2.6. The next step is to show weak convergence of the discrete solution couple to the continuous one:

Lemma 3.2.10. *For every fixed $\varepsilon \in (0,1]$ there holds:*

$$\bar{U}_k^\varepsilon \rightharpoonup \bar{u}^\varepsilon \text{ in } \mathbb{U}, \; \bar{V}_k^\varepsilon \rightharpoonup \bar{v}^\varepsilon \text{ in } L_2(\Omega, \mathbb{R}^m), \; \bar{\theta}_k^\varepsilon \rightharpoonup \bar{\theta}^\varepsilon \text{ in } L_2(\Omega, \mathbb{R}^m), \; k \to \infty$$

Besides, for every fixed $\varepsilon \in (0,1]$

$$\bar{Y}_k^\varepsilon \rightharpoonup \bar{y}^\varepsilon \text{ in } \mathbb{Y} \text{ and } \mathbb{W}, \; k \to \infty. \tag{3.2.26}$$

Proof. Choose a weakly convergent subsequence of $\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon)\}$ with weak limit $(\tilde{u}, \tilde{v}, \tilde{\theta})$, cf Lemma 3.2.9. We have already proven in Lemma 3.2.9 that the weak limit satisfies $(\tilde{u}, \tilde{v}, \tilde{\theta}) \in \mathbb{U}^{\varepsilon,ad} \times C^-$ and $f^\varepsilon(\tilde{u}, \tilde{v}) \geq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$. We will now show that in fact

$$f^\varepsilon(\tilde{u}, \tilde{v}) = f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon),$$

which will entail the postulated weak convergence.
We denote the best-approximation of $\bar{v}^\varepsilon$ in $\mathbb{V}_k$ w.r.t $\|\cdot\|$ by $B_k \bar{v}^\varepsilon$, i.e.

$$B_k \bar{v}^\varepsilon := \operatorname*{arg\,min}_{W \in \mathbb{V}_k} \|W - \bar{v}^\varepsilon\|^2 . \tag{3.2.27}$$

Thanks to (A2) we have

$$B_k \bar{v}^\varepsilon \to \bar{v}^\varepsilon \text{ in } L_2(\Omega, \mathbb{R}^m), \; k \to \infty \tag{3.2.28}$$

With the help of Assumption (A3) we can deduce the existence of $P_k \bar{u}^\varepsilon \in \mathcal{U}_k$ with $P_k \bar{u}^\varepsilon \to \bar{u}^\varepsilon$ as $k \to \infty$ such that for $\bar{\theta}_k^\varepsilon \rightharpoonup \tilde{\theta}$, $k \to \infty$:

$$\mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \tilde{\theta}) = \lim_{k \to \infty} \mathcal{L}_k^\varepsilon(P_k \bar{u}^\varepsilon, B_k \bar{v}^\varepsilon, \bar{\theta}_k^\varepsilon). \tag{3.2.29}$$

Besides, we know that because $\tilde{\theta} \in C^-$:

$$f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \geq \mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \tilde{\theta}). \tag{3.2.30}$$

Combining relations (3.2.29) and (3.2.30) with (3.2.24) and (3.2.23) then yields:

$$\begin{aligned}
f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) &\geq \mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \tilde{\theta}) && \text{cf (3.2.30)}\\
&= \lim_{k\to\infty} \mathcal{L}_k^\varepsilon(P_k\bar{u}^\varepsilon, B_k\bar{v}^\varepsilon, \bar{\theta}_k^\varepsilon) && \text{cf (3.2.29)}\\
&= \liminf_{k\to\infty} \mathcal{L}_k^\varepsilon(P_k\bar{u}^\varepsilon, B_k\bar{v}^\varepsilon, \bar{\theta}_k^\varepsilon)\\
&\geq \liminf_{k\to\infty} \mathcal{L}_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon) && \text{cf (3.2.24)}\\
&= \liminf_{k\to\infty} f_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) && \text{cf (3.2.23)}\\
&\geq f^\varepsilon(\tilde{u}, \tilde{v}),
\end{aligned}$$

where in the last line we have used weak lower semi continuity of f^ε for every fixed $\varepsilon > 0$. Combining this with (3.2.25) we obtain:

$$f^\varepsilon(\tilde{u}, \tilde{v}) = f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) = \mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \tilde{\theta})$$

Since $\bar{u}^\varepsilon$ and $\bar{v}^\varepsilon$ as well as the associated Lagrange multiplier $\bar{\theta}^\varepsilon$ are unique this means that

$$\tilde{u} = \bar{u}^\varepsilon, \;\; \tilde{v} = \bar{v}^\varepsilon, \;\; \tilde{\theta} = \bar{\theta}^\varepsilon.$$

At this stage, we have to emphasise that this is still only true for one weakly convergent subsequence of $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon)\right\}$. We have to extend this result to the entire sequence: The arguments above are valid for every weakly convergent subsequence of $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon)\right\}$, hence, utilising the fact that the limit $(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon)$ is unique, we can conclude that in fact the entire sequence $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon)\right\}$ weakly converges, i.e.:

$$\bar{U}_k^\varepsilon \rightharpoonup \bar{u}^\varepsilon, \; \bar{V}_k^\varepsilon \rightharpoonup \bar{v}^\varepsilon, \; \bar{\theta}_k^\varepsilon \rightharpoonup \bar{\theta}^\varepsilon, \;\; k \to \infty.$$

The detailed arguments for this step of the proof are recorded in Lemma 2.1.5.
The relation (3.2.26) is a direct consequence of weak convergence $\bar{U}_k^\varepsilon \rightharpoonup \bar{u}^\varepsilon$ and Lemma 3.1.3 and $\mathbb{Y} \hookrightarrow \mathbb{W}$. □

As yet, we have only demonstrated that the sequence $\left\{(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\}$ converges weakly to $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$, we now intend to make the step to strong convergence of the sequence. More or less on the way the results of the pointwise convergence theorem, Theorem 3.2.6, will follow:

Proof of Theorem 3.2.6. In this proof, we will take the following steps:

1. Show $a_k(\varepsilon) \to a(\varepsilon)$ pointwise for all $\varepsilon \in (0,1]$ as $k \to \infty$.
2. Prove that this implies $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{Y}_k^\varepsilon) \to (\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{y}^\varepsilon)$ in $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times \mathbb{Y}$ as $k \to \infty$.
3. Demonstrate that $a_k'(\varepsilon) \to a'(\varepsilon)$ pointwise for all $\varepsilon \in (0,1]$ as $k \to \infty$.
4. Prove the equivalence relation (3.2.16).

Step 1: Recalling the results of Lemma 3.2.10, we can then estimate in the following way taking once again the best-approximation B_k, compare (3.2.27). The steps taken are explained below:

$$\begin{aligned}
f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) = \mathcal{L}^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{\theta}^\varepsilon) &= \lim_{k\to\infty} \mathcal{L}_k^\varepsilon(P_k\bar{u}^\varepsilon, B_k\bar{v}^\varepsilon, \bar{\theta}_k^\varepsilon) \\
&= \limsup_{k\to\infty} \mathcal{L}_k^\varepsilon(P_k\bar{u}^\varepsilon, B_k\bar{v}^\varepsilon, \bar{\theta}_k^\varepsilon) \\
&\geq \limsup_{k\to\infty} \mathcal{L}_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{\theta}_k^\varepsilon) \\
&= \limsup_{k\to\infty} f_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \\
&\geq \liminf_{k\to\infty} f_k^\varepsilon(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \\
&\geq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon),
\end{aligned}$$

In the first line we used the complimentary slackness equality, then we employed the fact that by Assumption (A3) there exists $P_k u \in \mathcal{U}_k$ for every $u \in \mathcal{U}$ with $P_k u \to u$ in $\mathbb{U}$ as well as $\bar{\theta}_k^\varepsilon \rightharpoonup \bar{\theta}^\varepsilon$, Lemma 3.2.10, while in the third line we took advantage of the interior minimisation in (DP_d^ε), compare (3.2.24). The fourth is a consequence of (3.2.23) and the sixth of weak lower semi-continuity of f^ε for every fixed $\varepsilon > 0$.
Hence,

$$a(\varepsilon) \geq \limsup_{k\to\infty} a_k(\varepsilon) \geq \liminf_{k\to\infty} a_k(\varepsilon) \geq a(\varepsilon)$$

and as a consequence - $\limsup$ and $\liminf$ coincide -

$$a_k(\varepsilon) \to a(\varepsilon),\ \forall \varepsilon > 0,\ k \to \infty. \tag{3.2.31}$$

which is one of the pointwise convergence relations stated in Theorem 3.2.6.

Step 2: Strong convergence $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \bar{Y}_k^\varepsilon) \to (\bar{u}^\varepsilon, \bar{v}^\varepsilon, \bar{y}^\varepsilon)$ in $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times \mathbb{Y}$ as $k \to \infty$: We define for every fixed $\varepsilon > 0$ the norm

$$\|(y,u,v)\|_{**} := (\frac{1}{2}\|y\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon}\|v\|^2)^{1/2}$$

which is equivalent to the canonical norm on $\mathbb{W} \times \mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$ defined by

$$\|(y,u,v)\|_{\mathbb{W}\times\mathbb{U}\times L_2(\Omega,\mathbb{R}^m)} := (\|y\|_{\mathbb{W}}^2 + \|u\|_{\mathbb{U}}^2 + \|v\|^2)^{1/2}$$

for every fixed $\varepsilon > 0$.

In addition, a short computation yields

$$\|(\bar{y}^\varepsilon, \bar{u}^\varepsilon, \bar{v}^\varepsilon)\|_{**}^2 = a(\varepsilon) - (\bar{y}^\varepsilon, y_d)_{\mathbb{W}} + \frac{1}{2}\|y_d\|_{\mathbb{W}}^2 \tag{3.2.32}$$

and the corresponding relation on the discrete level:

$$\left\|(\bar{Y}_k^\varepsilon, \bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\|_{**}^2 = a_k(\varepsilon) - (\bar{Y}_k^\varepsilon, y_d)_{\mathbb{W}} + \frac{1}{2}\|y_d\|_{\mathbb{W}}^2 \tag{3.2.33}$$

We now observe that the bounded (it is weakly convergent thanks to Lemma 3.2.10) sequence $\left\|(\bar{Y}_k^\varepsilon, \bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\|_{**}^2$ fulfils all the prerequisites of the sequence x_k of Lemma 3.1.2 with

$$\liminf_{k\to\infty} \left\|(\bar{Y}_k^\varepsilon, \bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\|_{**}^2 \geq \|(\bar{y}^\varepsilon, \bar{u}^\varepsilon, \bar{v}^\varepsilon)\|_{**}^2$$

because of weak lower semi-continuity of a squared Hilbert space norm. The convergent sequence

$$-(\bar{Y}_k^\varepsilon, y_d)_{\mathbb{W}} + \frac{1}{2}\|y_d\|_{\mathbb{W}}^2$$

plays the role of y_k from Lemma 3.1.2.

Using (3.2.31), (3.2.32) and (3.2.33), we realise

$$\begin{aligned}\lim_{k\to\infty} a_k(\varepsilon) &= \lim_{k\to\infty} \left\|(\bar{Y}_k^\varepsilon, \bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\|_{**}^2 + (\bar{Y}_k^\varepsilon, y_d)_{\mathbb{W}} + \frac{1}{2}\|y_d\|_{\mathbb{W}}^2 \\ &= a(\varepsilon) \\ &= \|(\bar{y}^\varepsilon, \bar{u}^\varepsilon, \bar{v}^\varepsilon)\|_{**}^2 + (\bar{y}^\varepsilon, y_d)_{\mathbb{W}} + \frac{1}{2}\|y_d\|_{\mathbb{W}}^2\end{aligned}$$

Utilising the results of Lemma 3.1.2, we can then deduce

$$\lim_{k\to\infty} \left\|(\bar{Y}_k^\varepsilon, \bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\right\|_{**}^2 = \|(\bar{y}^\varepsilon, \bar{u}^\varepsilon, \bar{v}^\varepsilon)\|_{**}^2 .$$

Lemma 3.1.1 then ensures that

$$\bar{U}_k^\varepsilon \to \bar{u}^\varepsilon,\ \bar{V}_k^\varepsilon \to \bar{v}^\varepsilon,\ \bar{Y}_k^\varepsilon \to \bar{y}^\varepsilon,\ \text{in } \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \times \mathbb{W},\ k \to \infty.$$

To prove $\bar{Y}_k^\varepsilon \to \bar{y}^\varepsilon$ in $\mathbb{Y}$, we observe that thanks to Assumption (Pr7)and convergence $\bar{U}_k^\varepsilon \to \bar{u}^\varepsilon$ in $\mathbb{U}$ as $k \to \infty$:

$$\left\|\bar{Y}_k^\varepsilon - \bar{y}^\varepsilon\right\|_{\mathbb{Y}} \leq \underbrace{\|S\bar{u}^\varepsilon - S_k\bar{u}^\varepsilon\|_{\mathbb{Y}}}_{\to 0} + \underbrace{\|S_k\|_{\mathcal{L}(\mathbb{U},\mathbb{Y})} \left\|\bar{u}^\varepsilon - \bar{U}_k^\varepsilon\right\|_{\mathbb{U}}}_{\to 0} .$$

This gives convergence $\bar{Y}_k^\varepsilon \to \bar{y}^\varepsilon$ in $\mathbb{Y}$ as $k \to \infty$.

Step 3: Pointwise convergence $a_k'(\varepsilon) \to a'(\varepsilon)$ is now a simple consequence of $\bar{V}_k^\varepsilon \to \bar{v}^\varepsilon$ in $L_2(\Omega, \mathbb{R}^m)$ as $k \to \infty$, the formulas for $a'(\varepsilon)$ and $a_k'(\varepsilon)$, compare Theorem 3.2.1 and Theorem 3.2.5, and continuity of $\|\cdot\|^2$.

Step 4: Now, we can move on to the fourth and last step of the proof, the equivalence result (3.2.16):

$$\begin{aligned} a_k \to a,\, a_k' \to a' \text{ ptw. on } (0,1) \;\;\Leftrightarrow& \\ &(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \to (\bar{u}^\varepsilon, \bar{v}^\varepsilon) \text{ in } \mathbb{U} \times L_2(\Omega, \mathbb{R}^m),\ \varepsilon \in (0,1],\ k \to \infty \end{aligned}$$

First of all,

$$\bar{U}_k^\varepsilon \to \bar{u}^\varepsilon,\, \bar{V}_k^\varepsilon \to \bar{v}^\varepsilon \text{ in } \mathbb{U} \times L_2(\Omega, \mathbb{R}^m) \;\Rightarrow a_k(\varepsilon) \to a(\varepsilon),\, a_k'(\varepsilon) \to a'(\varepsilon)$$

is a consequence of f^ε and $-\frac{3}{2\varepsilon}\|\cdot\|^2$ being continuous for every fixed $\varepsilon > 0$.
Let us therefore turn to the other inclusion: Combining (3.2.32), (3.2.33) and Lemma 3.1.1 as we did before in this proof, namely in Step 2, we obtain

$$a_k(\varepsilon) \to a(\varepsilon) \Rightarrow \bar{U}_k^\varepsilon \to \bar{u}^\varepsilon,\, \bar{V}_k^\varepsilon \to \bar{v}^\varepsilon.$$

$a_k'(\varepsilon) \to a'(\varepsilon)$ is then a consequence of continuity of $-\frac{3}{2\varepsilon}\|\cdot\|^2$. □

This proof concludes this section. We have now collected every necessary ingredient to prove the previously mentioned equivalence relation:

$$\bar{U}_k \to \bar{u} \text{ in } \mathbb{U} \;\Leftrightarrow\; a_k \to a \text{ in } W_1^1(0,1),\;\; k \to \infty.$$

3.3 Convergence Analysis for the Unregularised Problems

In this section we will accomplish two things. First, we will prove the equivalence relation:

$$\bar{U}_k \to \bar{u} \text{ in } \mathbb{U} \;\Leftrightarrow\; a_k \to a \text{ in } W_1^1(0,1),\;\; k \to \infty.$$

Then, having ascertained this crucial result, we will search for a condition for which $a_k \to a$ in $W_1^1(0,1)$ as $k \to \infty$ holds. This condition will turn out to be both necessary and sufficient, thus, this will be the 'exact' characterisation of convergence $\bar{U}_k \to \bar{u}$ we had set out to gain at the start of this chapter.
At the end of this chapter we will then list some consequences of $\bar{U}_k$ converging to $\bar{u}$, the most striking of which will be that for *every* null sequence $\varepsilon_k \to 0$ we have $\bar{U}_k^{\varepsilon_k} \to \bar{u}$ as $k \to \infty$,

Corollary 3.3.11, a remarkably strong result.
Let us start with the equivalence relation:

3.3.1 Equivalence of Convergence

The next theorem constitutes our equivalence relation:

Theorem 3.3.1 (equivalence of convergence)**.** *The following statements are equivalent:*

1.

$$\bar{U}_k \to \bar{u} \textit{ in } \mathbb{U},\ k \to \infty$$

2.

$$a_k \to a. \textit{ in } W_1^1(0,1),\ k \to \infty$$

Proof. We will prove the first implication $\bar{U}_k \to \bar{u} \;\Rightarrow\; a_k \to a$ with respect to $W_1^1(0,1)$ as $k \to \infty$ first:
Let us start by demonstrating that $a_k \to a$ in $L_1(0,1)$ as $k \to \infty$. From Theorem 3.2.5 we know that a_k is uniformly bounded which means that

$$|a_k - a| \leq |a_k| + |a| \leq C + |a| \quad \text{a.e. on } (0,1)$$

Since $|a_k - a| \to 0$ pointwise everywhere (cf Theorem 3.2.6) as $k \to \infty$, an application of Lebesgue's dominated convergence theorem, [85], Theorem 5.36, results in $a_k \to a$ in $L_1(0,1)$. We still have to show that $a_k' \to a'$ in $L_1(0,1)$. First, we remark that $S_k\bar{U}_k \to S\bar{u}$ in both $\mathbb{Y}$ and $\mathbb{W}$ and continuity of the norm $\|\cdot\|^2$ for any Hilbert space yields

$$\begin{aligned}\bar{U}_k \to \bar{u} \Rightarrow\ & \frac{1}{2}\left\|S_k\bar{U}_k - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k\right\|_{\mathbb{U}}^2 \to \frac{1}{2}\left\|S\bar{u} - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{u}\right\|_{\mathbb{U}}^2 \\ \Leftrightarrow\ & a_k(0) \to a(0),\ k \to \infty\end{aligned}$$

Utilising the mean value theorem, continuity of a_k and a and pointwise convergence, we deduce as $k \to \infty$

$$\left\|a_k'\right\|_{L_1(0,1)} = -\int\limits_0^1 a_k'(\varepsilon)\, d\varepsilon = a_k(0) - a_k(1) \to a(0) - a(1) = -\int\limits_0^1 a'(\varepsilon)\, d\varepsilon = \left\|a'\right\|_{L_1(0,1)},$$

because $a_k(0) \to a(0)$ as $k \to \infty$.
Consequently,

$$\lim_{k\to\infty} \left\|a_k'\right\|_{L_1(0,1)} = \left\|a'\right\|_{L_1(0,1)}.$$

All in all, we have pointwise convergence of a_k' on $(0,1)$ and convergence with respect to the norm. Applying a slight generalisation of the dominated convergence theorem, compare [32],

Section 1.3. Theorem 4, we obtain convergence $a'_k \to a'$ in $L_1(0,1)$.
We now tackle the other implication, i.e.

$$a_k \to a \text{ in } W_1^1(0,1) \ \ \Rightarrow \bar{U}_k \to \bar{u},\ k \to \infty$$

Thanks to the embedding $W_1^1(0,1) \hookrightarrow C[0,1]$, compare Theorem 2.1.35, we also have convergence $a_k \to a$ in $C[0,1]$. $a_k \to a$ w.r.t $C[0,1]$ in particular implies that $a_k(0) \to a(0)$. Uniform boundedness of a_k on $[0,1]$ in particular implies:

$$\|\bar{U}_k\|_{\mathbb{U}} \lesssim 1.$$

Thus there exist a weakly convergent subsequence of $\{\bar{U}_k\}$. For every such weakly convergent subsequence of $\bar{U}_k$ with limit $\tilde{u} \in \mathbb{U}^{ad}$ Theorem 2.3.11 yields

$$\bar{U}_k \rightharpoonup \tilde{u},\ \bar{Y}_k = S_k\bar{U}_k \rightharpoonup S\tilde{u},\ k \to \infty,$$

compare also Lemma 3.1.3 for the weak convergence of the states $S_k\bar{U}_k$.
This in turn implies

$$a(0) = f(\bar{u}) \leq f(\tilde{u}) \leq \liminf_{k\to\infty} f_k(\bar{U}_k) = \liminf_{k\to\infty} a_k(0) = \lim_{k\to\infty} a_k(0) = a(0)$$

due to weak lower semi-continuity of $J(S_k u, u) = f_k(u) = \frac{1}{2}\|S_k u - y_d\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2$.
All in all, $f(\bar{u}) = f(\tilde{u})$, hence $\tilde{u} = \bar{u}$ due to the uniqueness of the solution to (P). These arguments apply to every weakly convergent subsequence with the limit $\bar{u} = \tilde{u}$ being unique, as a consequence, again compare Lemma 2.1.5,

$$\bar{U}_k \rightharpoonup \bar{u},\ k \to \infty$$

for the entire sequence $\{\bar{U}_k\}$. This entails weak convergence

$$\bar{Y}_k \rightharpoonup \bar{y},\ k \to \infty,$$

compare once again Lemma 3.1.3.
To prove strong convergence $\bar{U}_k \to \bar{u}$, we make similar arguments as in the proof of Theorem 3.2.6:
First, we define the norm

$$\|(y,u)\|_* := (\frac{1}{2}\|y\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2)^{1/2}, \tag{3.3.1}$$

which is equivalent to the canonical norm on $\mathbb{W} \times \mathbb{U}$ given by

$$\|(y,u)\|_{\mathbb{W}\times\mathbb{U}} = (\|y\|_{\mathbb{W}}^2 + \|u\|_{\mathbb{U}}^2)^{1/2}.$$

We then discern that

$$a_k(0) = \left\|(\bar{Y}_k, \bar{U}_k)\right\|_*^2 - 2(\bar{Y}_k, y_d)_{\mathbb{W}} + \|y_d\|_{\mathbb{W}}^2$$

and that

$$a(0) = \|(\bar{y}, \bar{u})\|_*^2 - 2(\bar{y}, y_d)_{\mathbb{W}} + \|y_d\|_{\mathbb{W}}^2 .$$

Besides, weak lower semi-continuity of any squared Hilbert space norm yields

$$\liminf_{k\to\infty} \left\|(\bar{Y}_k, \bar{U}_k)\right\|_*^2 \geq \|(\bar{y}, \bar{u})\|_*^2 .$$

Furthermore, the sequence $\left\|(\bar{Y}_k, \bar{U}_k)\right\|_*^2$ is bounded because $(\bar{Y}_k, \bar{U}_k)$ is weakly convergent. In this setting, we can apply the results of Lemma 3.1.2 with $\left\|(\bar{Y}_k, \bar{U}_k)\right\|_*^2$ playing the role of x_k of Lemma 3.1.2 and $2(\bar{Y}_k, y_d)_{\mathbb{W}} + \|y_d\|_{\mathbb{W}}^2$ that of the convergent sequence y_k of Lemma 3.1.2. After all,

$$\lim_{k\to\infty} 2(\bar{Y}_k, y_d)_{\mathbb{W}} + \|y_d\|_{\mathbb{W}}^2 = 2(\bar{y}, y_d)_{\mathbb{W}} + \|y_d\|_{\mathbb{W}}^2$$

due to weak convergence $\bar{Y}_k \rightharpoonup \bar{y}$. Thus, thanks to Lemma 3.1.2

$$\left\|(\bar{Y}_k, \bar{U}_k)\right\|_*^2 \to \|(\bar{y}, \bar{u})\|_*^2 ,\ k \to \infty$$

Together with weak convergence $\bar{U}_k \rightharpoonup \bar{u}$ Lemma 3.1.1 ensures strong convergence $\bar{U}_k \to \bar{u}$. □

In view of Theorem 3.3.1 we would be best advised to search for conditions under which $a_k \to a$ in $W_1^1(0,1)$. First of all, let us prove a lemma demonstrating that in fact the question of whether $a_k \to a$ in $W_1^1(0,1)$ boils down to ensuring that $a_k' \to a'$ in $L_1(0,1)$, since the functions a_k converge in $L_1(0,1)$.

Lemma 3.3.2. *We have*

$$a_k \to a \text{ in } L_1(0,1).$$

Proof. Employing the uniform boundedness of a_k, compare Theorem 3.2.5, we obtain almost everywhere on $(0,1)$

$$|a_k - a| \leq |a_k| + |a| \lesssim 1 + |a|$$

The right hand side is integrable, hence using the fact that $a_k \to a$ pointwise a.e. on $(0,1)$ compare Theorem 3.2.6, and dominated convergence theorem we obtain

$$\|a_k - a\|_{L_1(0,1)} \to 0$$

□

We will now tackle the question of convergence $a'_k \to a'$.
Theorem 3.2.6 ensures that a'_k converges pointwise on $(0,1)$. The trouble is that this is not enough to guarantee convergence in $L_1(0,1)$ as the example below demonstrates:

Example 3.3.3. *The sequence $g_k : [0,1] \to \mathbb{R}$ with*

$$g_k(x) := \begin{cases} 0 & \textit{if } x \in (\frac{1}{k}, 1) \\ k & \textit{else} \end{cases}$$

converges pointwise to 0 on $(0,1)$, but since $\|g_k\|_{L_1(0,1)} = 1$ for all k

$$g_k \not\to 0 \textit{ in } L_1(0,1).$$

To derive criteria for convergence $a'_k \to a'$ in $L_1(0,1)$ we will first introduce the notion of equi-integrability, compare Theorem 1.3 (b), Section VII in [29].

Definition 3.3.4 (equi-integrability)**.** *Let M be a subset of $L_1(0,1)$. We say that all $g \in M$ are equi-integrable if for all $\delta > 0$ there exists $\lambda = \lambda(\delta) > 0$ such that*

$$\int\limits_{|g(x)|>\lambda} |g(x)|\, dx \le \delta \quad \forall g \in M. \tag{3.3.2}$$

It is evident that condition (3.3.2) exerts some measure of control on the behaviour of functions g on sets of small measure. Thus, it is not surprising that (3.3.2) is exactly the condition we need to exclude the pathological case of Example 3.3.3. The following theorem confirms this view. It can be found in [29], Corollary 1.3, Section VIII.

Theorem 3.3.5. *Suppose that $\{g_k\}$ is an equi-integrable sequence in $L_1(0,1)$, i.e. for all $\delta > 0$ there exists $\lambda > 0$ such that independent of k we have*

$$\int\limits_{|g_k(x)|>\lambda} |g_k(x)|\, dx \le \delta.$$

Suppose further that $g_k(x) \to g(x)$ pointwise a.e. on $(0,1)$. Then $g \in L_1(0,1)$ and $g_k \to g$ in $L_1(0,1)$.

With the help of the notion of equi-integrability and Theorem 3.3.5 we can formulate the following crucial auxiliary lemma to prove convergence $\bar{U}_k \to \bar{u}$.

Lemma 3.3.6. *Suppose that the sequence* $\{a_k'\}$ *is equi-integrable, i.e. for all* $\delta > 0$ *there exists* $\lambda > 0$ *such that independent of* k

$$\int\limits_{|a_k'(\varepsilon)|>\lambda} |a_k'(\varepsilon)| \, d\varepsilon \leq \delta. \tag{3.3.3}$$

Then $a_k' \to a'$ *in* $L_1(0,1)$.

Proof. Since $a_k' \to a'$ pointwise on $(0,1)$ we can apply Theorem 3.3.5 to obtain the desired result. □

These auxiliary results are the key ingredients to prove the convergence $\bar{U}_k \to \bar{u}$ which we will do in the next section

3.3.2 Convergence Theorem

Before we formulate the central convergence theorem of this section, let us first introduce a slightly more accessible notion of equi-integrability for the sequence $\{a_k'\}$:

Lemma 3.3.7. *The sequence* $\{a_k'\}$ *is equi-integrable in the sense of* (3.3.3) *iff for all* $\delta > 0$ *there exists* $\xi = \xi(\delta) > 0$ *such that independent of* k

$$\int\limits_0^\xi |a_k'(\varepsilon)| \, d\varepsilon = \int\limits_0^\xi \frac{3}{2\varepsilon^2} \left\| \bar{V}_k^\varepsilon \right\|^2 \, d\varepsilon \leq \delta. \tag{3.3.4}$$

Proof. Let us prove ⇐ first: Suppose that (3.3.4) holds. Then for given $\delta > 0$ pick $\xi = \xi(\delta) > 0$ such that (3.3.4) holds. Besides choose $\lambda > 0$ large enough so that

$$\left\{ \varepsilon \in (0,1) \ : \ |a_k'(\varepsilon)| > \lambda \right\} \cap (\xi, 1) = \emptyset$$

This is possible, since by Theorem 3.2.5, compare (3.2.14), a_k' is bounded independent of k and ε on $[r,1]$ for every fixed $r > 0$. For this λ we now have

$$\int\limits_{|a_k'(\varepsilon)|>\lambda} |a_k'(\varepsilon)| \, d\varepsilon \leq \int\limits_0^\xi |a_k'(\varepsilon)| \, d\varepsilon \leq \delta$$

independent of k. Hence $\{a_k'\}$ is equi-integrable according to (3.3.3).
Let us now turn to ⇒: For given $\delta > 0$ we can choose $\lambda > 0$ such that (3.3.3) holds. Again

due to boundedness of a_k' on any compact subset of $(0,1)$, (3.2.14), we can choose $\tilde{\lambda} \geq \lambda > 0$ and $\xi > 0$ such that

$$(0,\xi) \subset \left\{\varepsilon \in (0,1) \;:\; |a_k'(\varepsilon)| > \tilde{\lambda}\right\}$$

This now yields

$$\begin{aligned} \int_0^\xi |a_k'(\varepsilon)|\, d\varepsilon &\leq \int_{\left\{|a_k'(\varepsilon)|>\tilde{\lambda}\right\}} |a_k'(\varepsilon)|\, d\varepsilon \\ &\leq \int_{\left\{|a_k'(\varepsilon)|>\lambda\right\}} |a_k'(\varepsilon)|\, d\varepsilon \\ &\leq \delta. \end{aligned}$$

This gives the desired result completing the proof. □

Combining Lemma 3.3.2 and Lemma 3.3.6 we can now prove the central convergence theorem of this section:

Theorem 3.3.8 (convergence theorem). *Suppose that* (3.3.4) *holds, i.e.: For all* $\delta > 0$ *there exists* $\xi = \xi(\delta) > 0$ *such that independent of* k

$$\int_0^\xi |a_k'(\varepsilon)|\, d\varepsilon = \int_0^\xi \frac{3}{2\varepsilon^2} \left\|\bar{V}_k^\varepsilon\right\|^2 \, d\varepsilon \leq \delta.$$

Then

$$\bar{U}_k \to \bar{u} \text{ in } \mathbb{U},\ k \to \infty.$$

Proof. First of all, we recall Theorem 3.2.6 which yielded $a_k \to a$ and $a_k' \to a'$ in $(0,1)$ pointwise everywhere. Lemma 3.3.2 yields convergence $a_k \to a$ in $L_1(0,1)$. Now, observe that thanks to Lemma 3.3.7, (3.3.4) is tantamount to (3.3.2) in Definition 3.3.4. Theorem 3.3.5 with $g_k = a_k'$ then implies convergence $a_k' \to a'$ in $L_1(0,1)$. Thus, all in all $a_k \to a$ in $W_1^1(0,1)$. Theorem 3.3.1 then yields convergence $\bar{U}_k \to \bar{u}$ as $k \to \infty$. □

At this stage, it is important to point out that (3.3.4) is an additional condition enforced on the behaviour of a_k'. However, as the next theorem states, we did not lose anything on the way, i.e. convergence $\bar{U}_k \to \bar{u}$ implies condition (3.3.4), i.e (3.3.4) is an **exact characterisation of convergence**. If (3.3.4) does not hold, then $\bar{U}_k$ does not converge to the true solution $\bar{u}$.

To prove this result we first need an auxiliary lemma, the 'Dunford-Pettis compactness criterion'. It can be found in [28], Theorem IV.8.9 and Corollary IV.8.11. Compare also [29], Chapter VIII, Theorem 1.3.

Lemma 3.3.9. *Let a' and a'_k be the derivatives of the continuous and discrete optimal value functions and $M := \{\{a'_k\}, a'\} \subset L_1(0,1)$. Then the following equivalence relations hold*

$$M \textit{ weakly compact } \Leftrightarrow (3.3.4) \Leftrightarrow a'_k \to a' \textit{ in } L_1(0,1)$$

Proof. The inclusion M weakly compact $\Leftrightarrow$ (3.3.4) is the so called 'Dunford-Pettis compactness criterion' which can be found in [28], Corollary IV.8.11., compare also [27]. Here, observe also that as shown in the proof of Theorem 3.3.8 the notion of equi-integrability (3.3.4) is equivalent to the definition of equi-integrability, Definition 3.3.4 specifically formula (3.3.2). The implication (3.3.4) $\Rightarrow a'_k \to a'$ was shown in Theorem 3.3.8. The inclusion $a'_k \to a' \Rightarrow M$ weakly compact is trivial. After all, since $a'_k \to a'$ every subsequence converges strongly in $L_1(0,1)$ and thus also weakly. □

Now we can turn to the aforementioned exact characterisation of convergence.

Theorem 3.3.10 (exact characterisation of convergence)**.** *The following equivalence is valid: $\bar{U}_k \to \bar{u}$ as $k \to \infty$ if and only if the sequence $\{a'_k\}$ is equi-integrable in the sense of* (3.3.4)*.*

Proof. We start with the direction $\Rightarrow$ first. As Theorem 3.3.1 demonstrates, $\bar{U}_k \to \bar{u}$ implies $a_k \to a$ in $W^1_1(0,1)$ and in particular $a'_k \to a'$ in $L_1(0,1)$. The latter trivially implies $a'_k \rightharpoonup a'$ in $L_1(0,1)$. Hence, the set $M = \{\{a'_k\}, a'\}$ is weakly compact. Lemma 3.3.9 then implies (3.3.4).
Conversely, (3.3.4) implies $a'_k \to a'$ in $L_1(0,1)$ due to Theorem 3.3.8 and together with Lemma 3.3.2 $a_k \to a$ in $W^1_1(0,1)$. Theorem 3.3.1 now yields the desired result. □

Let us conclude this section by recording a result stating that provided $\bar{U}_k \to \bar{u}$ regularisation and discretisation can be decoupled:

Corollary 3.3.11. *Let $k \to \infty$ and $\varepsilon_k \to 0$ as $k \to \infty$. Suppose that $\bar{U}_k \to \bar{u}$ or equivalently,* (3.3.4) *holds. Then*

$$\bar{U}_k^{\varepsilon_k} \to \bar{u},\ k \to \infty.$$

Proof. If either $\bar{U}_k \to \bar{u}$ or (3.3.4) then Theorem 3.3.8 and Theorem 3.3.1 ensure that $a_k \to a$ in $W^1_1(0,1)$ and due to the embedding $W^1_1(0,1) \hookrightarrow C[0,1]$, compare Theorem 2.1.35, we also have convergence $a_k \to a$ in $C[0,1]$. This implies that the set $M = \{\{a_k\}, a\}$ is compact w.r.t to the canonical norm of $C[0,1]$. Thus, thanks to the famous Theorem of Arzelá-Ascoli, cf [71], Section 8, Theorem 33 and Corollary 34, the sequence of functions $\{a_k\}$ is equi-continuous, i.e. for every $\varepsilon_0 \in [0,1]$ and $\eta > 0$ there exists a $\delta > 0$ solely depending on η such that

$$|a_k(\varepsilon) - a_k(\varepsilon_0)| < \eta \quad \forall k, \ \ \forall \varepsilon : \ |\varepsilon_0 - \varepsilon| < \delta.$$

In particular, this means that for every null sequence $\{\varepsilon_k\} \subset [0,1]$ we have

$$a_k(\varepsilon_k) \to a(0), \ k \to \infty. \tag{3.3.5}$$

The reason for this is that, after all, thanks to equicontinuity and pointwise convergence $a_k(0) \to a(0)$ for every $\eta > 0$ we can choose $\delta = \delta(\eta) > 0$ and K large enough such that

$$\begin{aligned} |a_k(\varepsilon_k) - a(0)| &\leq |a_k(\varepsilon_k) - a_k(0)| + |a_k(0) - a(0)| \\ &\leq \frac{\eta}{2} + \frac{\eta}{2} = \eta \ \ \forall |\varepsilon_k| < \delta, \ k \geq K. \end{aligned}$$

Uniform boundedness of $\{a_k\}$, compare Theorem 3.2.5, yields that $\left\{(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k})\right\}$ is uniformly bounded in $\mathbb{U} \times L_2(\Omega, \mathbb{R}^m)$ and thus possesses a weakly convergent subsequence with weak limit $(\tilde{u}, \tilde{v})$.
Since

$$I_k y_c - S_k \bar{U}_k^{\varepsilon_k} - \varepsilon_k \bar{V}_k^{\varepsilon_k} \rightharpoonup y_c - S\tilde{u}, \ k \to \infty$$

and C is weakly closed, we can conclude that $\tilde{u} \in \mathbb{U}^{ad}$. Here, recall that weak convergence of $S_k \bar{U}_k^{\varepsilon_k} \rightharpoonup S\tilde{u}$ is ensured thanks to Lemma 3.1.3. We can now estimate in the following way using properties of the lim inf and weak lower semicontinuity of

$$f_k^{\varepsilon}(U,V) = J(S_k U, U, \frac{1}{2\sqrt{\varepsilon_k}} V) = \frac{1}{2} \left\| S_k U - y_d \right\|_{\mathbb{W}}^2 + \frac{\nu}{2} \left\| V \right\|_{\mathbb{U}}^2 + \left\| \frac{1}{2\sqrt{\varepsilon_k}} V \right\|^2$$

and our previous observation (3.3.5)

$$\begin{aligned} a(0) = \lim_{k\to\infty} a_k(0) = \lim_{k\to\infty} a_k(\varepsilon_k) &= \liminf_{k\to\infty} a_k(\varepsilon_k) \\ &\geq \liminf_{k\to\infty} (\frac{1}{2} \left\| S_k \bar{U}_k^{\varepsilon_k} - y_d \right\|_{\mathbb{W}}^2 + \frac{\nu}{2} \left\| \bar{U}_k^{\varepsilon_k} \right\|_{\mathbb{U}}^2) + \liminf_{k\to\infty} \frac{1}{2\varepsilon_k} \left\| \bar{V}_k^{\varepsilon_k} \right\|^2 \\ &\geq (\frac{1}{2} \left\| S\tilde{u} - y_d \right\|_{\mathbb{W}}^2 + \frac{\nu}{2} \left\| \tilde{u} \right\|_{\mathbb{U}}^2). \end{aligned}$$

Thus,

$$a(0) = f(\bar{u}) \geq f(\tilde{u}).$$

From the estimate above we can deduce that $\tilde{u} = \bar{u}$, because $\bar{u}$ is the unique solution to (P) and $\tilde{u} \in \mathbb{U}^{ad}$. As this is true for every subsequence, the entire sequence $\bar{U}_k^{\varepsilon_k}$ converges weakly to $\bar{u}$, compare Lemma 2.1.5.

Let us now prove strong convergence $\bar{U}_k^{\varepsilon_k} \to \bar{u}$ as $k \to \infty$. Due to (recall that $a_k(0) \to a(0)$ and $a_k(\varepsilon_k) \to a_k(0)$!)

$$\begin{aligned} a(0) = \lim_{k\to\infty} a_k(0) = \limsup_{k\to\infty} a_k(0) = \limsup_{k\to\infty} a_k(\varepsilon_k) \\ \underbrace{\geq \limsup_{k\to\infty} f_k(\bar{U}_k^{\varepsilon_k})}_{f_k^{\varepsilon_k}(\bar{U}_k^{\varepsilon_k},\bar{V}_k^{\varepsilon_k})\geq f_k(\bar{U}_k^{\varepsilon_k})} \geq \liminf_{k\to\infty} f_k(\bar{U}_k^{\varepsilon_k}) \geq a(0), \end{aligned}$$

where in the last step we again used lower semicontinuity, we obtain $f_k(\bar{U}_k^{\varepsilon_k}) \to f(\bar{u}) = a(0)$. Let us now take a closer look at the functional f_k:

$$\begin{aligned} f_k(\bar{U}_k^{\varepsilon_k}) &= \frac{1}{2}\left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k^{\varepsilon_k}\right\|_{\mathbb{U}}^2 \\ &= \frac{1}{2}\left\|\bar{Y}_k^{\varepsilon_k}\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k^{\varepsilon_k}\right\|_{\mathbb{U}}^2 - (y_d, \bar{Y}_k^{\varepsilon_k})_{\mathbb{U}} + \|y_d\|_{\mathbb{W}}^2 . \end{aligned}$$

We recall the definition of (3.3.1)

$$\|(u,y)\|_* := (\frac{1}{2}\|y\|_{\mathbb{W}}^2 + \frac{\nu}{2}\|u\|_{\mathbb{U}}^2)^{1/2}$$

which defines a norm on $\mathbb{W} \times \mathbb{U}$ that is equivalent to the canonical norm

$$\|(u,y)\|_{\mathbb{W}\times\mathbb{U}} = (\|y\|_{\mathbb{W}}^2 + \|u\|_{\mathbb{U}}^2)^{1/2}.$$

We can then estimate in the following fashion using weak lower semi-continuity of the norm $\|(u,y)\|_*^2$ and properties of lim inf and lim sup as well as weak convergence $\bar{U}_k^{\varepsilon_k} \rightharpoonup \bar{u}$:

$$\begin{aligned} a(0) = \lim_{k\to\infty} a_k(0) &= \limsup_{k\to\infty} a_k(0) \\ &= \limsup_{k\to\infty}(\frac{1}{2}\left\|\bar{Y}_k^{\varepsilon_k}\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k^{\varepsilon_k}\right\|_{\mathbb{U}}^2 - (y_d, \bar{Y}_k^{\varepsilon_k})_{\mathbb{U}} + \|y_d\|_{\mathbb{W}}^2) \\ &= \limsup_{k\to\infty}\left\|(\bar{U}_k^{\varepsilon_k}, \bar{Y}_k^{\varepsilon_k})\right\|_*^2 + \limsup_{k\to\infty}(-(y_d, \bar{Y}_k^{\varepsilon_k})_{\mathbb{U}} + \|y_d\|_{\mathbb{W}}^2) \\ &\geq \liminf_{k\to\infty}\left\|(\bar{U}_k^{\varepsilon_k}, \bar{Y}_k)\right\|_*^2 + \lim_{k\to\infty}(-(y_d, \bar{Y}_k^{\varepsilon_k})_{\mathbb{U}} + \|y_d\|_{\mathbb{W}}^2) \geq a(0). \end{aligned}$$

Evidently,

$$\limsup_{k\to\infty}\left\|(\bar{U}_k^{\varepsilon_k}, \bar{Y}_k^{\varepsilon_k})\right\|_*^2 = \liminf_{k\to\infty}\left\|(\bar{U}_k^{\varepsilon_k}, \bar{Y}_k^{\varepsilon_k})\right\|_*^2 = \|(\bar{u}, \bar{y})\|_*^2 ,$$

recall also Lemma 3.1.2.
This results in

$$\left\|(\bar{Y}_k^{\varepsilon_k}, \bar{U}_k^{\varepsilon_k})\right\|_* \to \|(\bar{y}, \bar{u})\|_* ,\ k \to \infty$$

Consequently, $\bar{U}_k^{\varepsilon_k} \to \bar{u}$ (weak convergence and norm convergence as detailed in Lemma 3.1.1) which completes the proof. □

As the last corollary of this section and last result of this chapter we obtain a 'diagonal convergence' property for the sequence of virtual controls $\left\{\bar{V}_k^{\varepsilon_k}\right\}$.

Corollary 3.3.12. *Suppose that the convergence condition* (3.3.4) *holds. Then for all* $\varepsilon_k \to 0$

$$\frac{1}{2\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2 \to 0, \;\; k \to \infty$$

Proof. Thanks to Corollary 3.3.11 we have $\bar{U}_k^{\varepsilon_k} \to \bar{u}$. This immediately entails:

$$\bar{Y}_k^{\varepsilon_k} = S_k \bar{U}_k^{\varepsilon_k} \to S\bar{u} = \bar{y}, k \to \infty$$

in $\mathbb{Y}$ and by assumption also in $\mathbb{W}$.
As in the proof of Corollary 3.3.11 we gain convergence

$$a(\varepsilon_k) \to a(0), \;\; k \to \infty.$$

Let us now take a closer look at the function $a_k(\varepsilon_k)$:

$$a_k(\varepsilon_k) = \frac{1}{2}\left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k^{\varepsilon_k}\right\|_{\mathbb{U}}^2 + \frac{1}{2\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2.$$

The first two terms converge since $\bar{U}_k^{\varepsilon_k} \to \bar{u}$ in $\mathbb{U}$ and $\bar{Y}_k^{\varepsilon_k} \to \bar{y}$ in $\mathbb{Y}$ and thus also in $\mathbb{W}$. Employing once again Lemma 3.1.2 with $\frac{1}{2}\left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{\mathbb{W}}^2 + \frac{\nu}{2}\left\|\bar{U}_k^{\varepsilon_k}\right\|_{\mathbb{U}}^2$ playing the role of the convergent sequence y_k in Lemma 3.1.2 and $\frac{1}{2\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2$ that of x_k with $x = 0$, we then deduce the desired result. □

In this chapter we have characterised convergence of $\bar{U}_k \to \bar{u}$ exactly. We can now turn to deriving an a posteriori error estimator steering an adaptive algorithm. This will be the subject of the next Chapter.

Chapter 4

The Estimator

In this chapter we will derive expressions $\mathcal{E}_r = \mathcal{E}_r(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \varepsilon, \text{data})$ and $\mathcal{E}_s = \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon, \varepsilon, \text{data})$ consisting of

- **computable quantities**, i.e. quantities of the following kind:

  $$\|f - g\|^2 \ ; (f, g - h)$$

  with **known** continuous or discrete functions f, g, h, of which we assume that we can evaluate them exactly. Needless to say, in our actual numerical experiments, we have to use numerical quadrature rules for terms with continuous functions. In certain special cases the errors arising from numerical integration can be included in a rigorous a posteriori analysis, e.g. [63].

- **linear errors** of the type

  $$\|(S - S_k)g_k\|^2$$

  with a known right hand side g_k which in turn can estimated by an a posteriori error estimator for FE solutions to linear PDE.

With these quantities we estimate the error

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{\mathbb{U}}^2$$

in the following **reliable** way, i.e. the estimator provides an upper bound up to constants:

$$\begin{aligned} \left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{\mathbb{U}}^2 &\lesssim \left\|\bar{u}^{\varepsilon^N} - \bar{u}\right\|_{\mathbb{U}}^2 + \left\|\bar{U}_k^\varepsilon - \bar{u}^{\varepsilon^N}\right\|_{\mathbb{U}}^2 \\ &\lesssim \varepsilon^{\gamma N} + \mathcal{E}_r^2 + \mathcal{E}_s \end{aligned} \tag{4.0.1}$$

with $\gamma < 1$ and $N \geq 1$.
Let us stress that to prove this reliable upper bound (4.0.1) **we can dispense with Assumptions** (A1)-(A4)**!**. In particular, this means that even if we face a situation where the set of admissible functions for the discrete unregularised problem is empty, we can still prove the bound (4.0.1), because we solely use the regularised problem, for which there are always admissible functions.
Only at the end of this chapter, namely in Section 4.3, where we will then prove - under certain conditions - that the derived estimator converges as $\varepsilon \to 0$ and $k \to \infty$, will we again require Assumptions (A1)-(A4).
As we will see in Theorem 4.1.10 the term

$$\left\| \bar{u}^{\varepsilon^N} - \bar{u} \right\|_{\mathbb{U}}^2$$

is estimated a priori in terms of the regularisation parameter ε^N, in Theorem 4.1.10, where the N gives us greater leeway in pushing the error $\left\| \bar{u}^{\varepsilon^N} - \bar{u} \right\|_{\mathbb{U}}^2$ a priori below some tolerance TOL: N can e.g. be chosen in such a way that

$$\left\| \bar{u}^{\varepsilon^N} - \bar{u} \right\|_{\mathbb{U}}^2 \lesssim \varepsilon^{\gamma N} \leq TOL^2.$$

However, as the reader will realise in Section 4.2.2, in three space dimensions increasing N will be paid for by a factor $\varepsilon^{-\delta N}$, $\delta \leq 1$, scaling one of the terms in $\mathcal{E}_s$.
Existing results have either focused on estimating the difference in the goal functionals $f(\bar{u}) - f(\bar{U}_k)$, see e.g. [6] and [86] and/or neglected to estimate terms that are related to the (regularised) Lagrange multiplier in providing a bound for the error $\left\| \bar{U}_k - \bar{u} \right\|_{\mathbb{U}}$, cf e.g. [46] and in the gradient-constrained case [42]. Others, such as [70] have primarily worked with L_∞ a posteriori estimators. Existing L_∞-error estimators, however, come with a scaling by $\ln h_{\min}$-terms, where $h_{\min}$ denotes the minimal mesh-size and demand higher W_p^1-regularity $p > d$, see e.g [63].
Thus, our work which provides an upper bound up to constants depending solely on the data in the sense of (4.0.1) constitutes a genuine extension and improvement to existing results.

In this chapter, we will deal with a linear-quadratic elliptic model problem and its discretisation, which we will introduce in the next section. **Throughout this chapter**, we assume that the space dimension d is either $d = 2$ or $d = 3$.

4.1 Problem Setting

In Chapter 3 we have investigated a fairly general optimal control problem. In this and the following sections and chapters, we will deal with a model Poisson linear-quadratic problem. In the course of the first few sections we will thus focus on formulating the results proven for the general problems in the specific setting of our model problem, also adding a number of extending results on the way.

4.1.1 The Continuous Model Problem

Besides, we demand that $\Omega \in C^{0,1}$ and that $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, as assumed, be meshable, cf Definition 2.1.26 and Definition 2.3.2.

In the previous chapters we have nearly always operated within the framework of abstract Hilbert spaces $\mathbb{U}, \mathbb{Y}$ and $\mathbb{W}$. We now make concrete choices setting:

$$\mathbb{U} = L_2(\Omega),\ \mathbb{Y} = \mathring{H}^1(\Omega),\ \mathbb{W} = L_2(\Omega).$$

For notational convenience the norm $\|\cdot\|_{L_2(\Omega)}$ and corresponding scalar product $(\cdot,\cdot)_{L_2(\Omega)}$ will be shortened to $\|\cdot\|$ and $(\cdot,\cdot)$ respectively.

$\mathcal{U}$ is given by

$$\mathcal{U} = \{u \in L_2(\Omega)\ :\ a \leq u \leq b\},\ a, b \in \mathbb{R},\ b - a > 0.$$

A short computation yields that all functions $u \in \mathcal{U}$ are uniformly bounded with

$$\|u\| \leq \max(|a|, |b|)|\Omega|^{1/2}. \tag{4.1.1}$$

Remark 4.1.1. *We remark that the results of this chapter do not hinge on the additional enforcement of box-constraints. The bounds and estimates remain valid with most constants now depending on the continuous solution $\bar{u}$ (see below) and the bound on the sequence in* (A1)*.*

The following continuous unregularised problem is given:

$$\left.\begin{array}{c} \min\limits_{u\in L_2(\Omega),y\in \mathring{H}^1(\Omega)} \frac{1}{2}\,\|y-y_d\|^2+\frac{\nu}{2}\|u\|^2 \\ \text{s.t.} \\ \int\limits_\Omega \nabla y\cdot\nabla w\,d\Omega = \int\limits_\Omega uw\,d\Omega\ \forall w\in \mathring{H}^1(\Omega) \\ \text{and} \\ u\in\mathcal{U} \\ y_c-y\leq 0\ \text{ a.e. on } \Omega \end{array}\right\} \qquad (CMP)$$

Here, we also suppose $y_c \in H^1(\Omega)$ and $\nabla y_c \in H(\text{div}, \Omega)$ as well as $y_c|_{\partial\Omega} < 0$ and, as always, a fixed $\nu > 0$.

For the sake of abbreviation we define a continuous bilinear form b in a slight abuse of notation by

$$b[y,w] := (\nabla y, \nabla w) := \int\limits_\Omega \nabla y\cdot\nabla w\,d\Omega\ \ \forall y,w\in \mathring{H}^1(\Omega). \tag{4.1.2}$$

Since the bilinear form b is also coercive, compare Theorem 2.4.1, we immediately obtain the existence of a solution operator mapping $S : L_2(\Omega) \to \mathring{H}^1(\Omega)$, compare also Corollary 2.1.21. In this special case, we even get additional regularity, i.e. the solution operator maps linearly and continuously to better spaces. To prove this, we first have to record a classic regularity theorem on the interior of a domain for elliptic equations which can e.g. be found in [36], Theorem 8.9.:

Theorem 4.1.2. *For any $\Omega' \subset\subset \Omega$ and $u \in L_2(\Omega)$ the solution $y \in \mathring{H}^1(\Omega)$ of*

$$\int\limits_\Omega \nabla y\cdot\nabla w\,d\Omega = \int\limits_\Omega uw\,d\Omega\ \forall w\in \mathring{H}^1(\Omega)$$

satisfies $y \in \mathring{H}^1(\Omega) \cap H^2(\Omega')$.

Having listed this result, we can now turn to the main regularity result:

Lemma 4.1.3. *Suppose that for fixed $u \in L_2(\Omega)$, $y \in \mathring{H}^1(\Omega)$ solves*

$$\int\limits_\Omega \nabla y\cdot\nabla w\,d\Omega = \int\limits_\Omega uw\,d\Omega\ \forall w\in \mathring{H}^1(\Omega)$$

Then, in addition $\nabla y \in H(\mathrm{div}, \Omega)$ *and*

$$\|\nabla y\|_{H(\mathrm{div},\Omega)} \lesssim \|u\|_{L_2(\Omega)} .$$

Proof. Pick an arbitrary $w \in C_0^\infty(\Omega)$. Since $\mathrm{supp}(w) \subset\subset \Omega$, Theorem 4.1.2 ensures $y \in H^2(\mathrm{supp}(w))$. Then by Green's formula, cf [37], Theorem 1.5.3.1 and Lemma 1.5.3.2, we can deduce that

$$\int\limits_{\mathrm{supp}(w)} \nabla y \cdot \nabla w \, d\Omega = \int\limits_{\Omega} \nabla y \cdot \nabla w \, d\Omega = -\int\limits_{\Omega} \mathrm{div} \nabla y w \, d\Omega = \int\limits_{\Omega} u w \, d\Omega,$$

Since $v \in C_0^\infty(\Omega)$ was arbitrary, we obtain

$$-\int\limits_{\Omega} \mathrm{div} \nabla y w \, d\Omega = \int\limits_{\Omega} u w \, d\Omega \ \ \forall w \in C_0^\infty(\Omega)$$

Harnessing the fundamental lemma of the calculus of variations, compare e.g [35], Chapter 1, Lemma 3, we are then able to deduce

$$-\mathrm{div} \nabla y = u \ \text{ a.e. on } \Omega.$$

u is an element of $L_2(\Omega)$, hence, we can conclude

$$\mathrm{div} \nabla y \in L_2(\Omega).$$

Evidently,

$$\|\mathrm{div} \nabla y\| = \|u\| .$$

Continuity of S and the definition of the $H(\mathrm{div}, \Omega)$-norm then yields:

$$\|\nabla y\|_{H(\mathrm{div},\Omega)} \lesssim \|u\|$$

which completes the proof. □

After this regularity detour, let us now return to the optimal control problem itself. Existence and uniqueness results as well as optimality conditions are our first focus. To this end, we can merely transfer the results of Section 2.2.1 to our specific model problem setting:
The fact that problem (CMP) satisfies Properties (Pr1)-(Pr5) has already been discussed for an even more general problem in Section 2.4.1. To ensure that we are not optimising over the empty set, we make the following Slater-type assumption:

Assumption 4.1.4. *There exists $u^s \in \mathcal{U}$ such that*

$$Su_s - y_c \geq \tau > 0, \text{ a.e. on } \Omega$$

Evidently, (Pr6) is fulfilled, too, hence, we can now deduce the existence of a unique solution $\bar{u}$ and corresponding state $\bar{y} = S\bar{u}$ (compare Theorem 2.2.1) such that

$$(\bar{p} + \nu\bar{u}, u - \bar{u}) \geq 0 \ \ \forall u \in \mathbb{U}^{ad} \tag{4.1.3}$$

with $\bar{p} = S^*(\bar{y} - y_d)$ with $S^* : L_2(\Omega) \to L_2(\Omega)$ and

$$\mathbb{U}^{ad} := \{u \in \mathcal{U} \ : \ y_c - Su \leq 0\}\,.$$

As before, we define the reduced functional by

$$f(u) = \frac{1}{2}\left\|Su - y_d\right\|^2 + \frac{\nu}{2}\left\|u\right\|^2$$

Having settled existence and uniqueness questions, we now aim to characterise S^* as the solution operator to another partial differential equation, in fact we obtain $S = S^*$. This will be important for both analytical and numerical reasons:

Theorem 4.1.5. *Given $q \in L_2(\Omega)$, $z = S^*q$ solves the variational problem*

$$\int\limits_{\Omega} \nabla z \cdot \nabla w \, d\Omega = \int\limits_{\Omega} qw \, d\Omega \ \ \forall w \in \mathring{H}^1(\Omega) \tag{4.1.4}$$

Thus, $z = S^*q \in \mathring{H}^1(\Omega)$, $\nabla z \in H(\mathrm{div}, \Omega)$ and

$$\|z\|_{\mathring{H}^1(\Omega)} + \|\nabla z\|_{H(\mathrm{div},\Omega)} \lesssim \|q\|\,.$$

Proof. By definition of the adjoint operator, compare Section 2.1.2, S^* satisfies

$$(h, z) = (h, S^*q) = (Sh, q), \ \forall h, q \in L_2(\Omega)$$

The variational problem (4.1.4) possesses a unique solution $z \in \mathring{H}^1(\Omega), \nabla z \in H(\mathrm{div}, \Omega)$ for every right-hand side $q \in L_2(\Omega)$, again compare Theorem 2.4.1 and Lemma 4.1.3. Thus, utilising the fact that $\mathring{H}^1(\Omega) \hookrightarrow L_2(\Omega)$, compare the Poincaré-Friedrich inequality Theorem 2.1.33, we can define a solution operator $\hat{S} : L_2(\Omega) \to L_2(\Omega)$ to (4.1.4). Employing (4.1.4) and the fact that $\hat{S}q \in \mathring{H}^1(\Omega)$, we deduce:

$$(h, \hat{S}q) = (\nabla Sh, \nabla\hat{S}q) = (q, Sh) \ \ \forall q, h \in L_2(\Omega).$$

As a consequence $\hat{S} = S = S^*$.
Corollary 2.1.21, Theorem 4.1.2 and Lemma 4.1.3 now yield the bound for $z = S^*q$. □

Let us now turn to the regularisation of (CMP).

4.1.2 The Continuous Model Regularised Problem

In this section, the focus lies very much on deriving an a priori estimate in terms of the regularisation parameter for the difference between the continuous regularised and unregularised solution $\bar{u}^\varepsilon, \bar{u}$. This will be achieved in Theorem 4.1.10 and 4.1.11. Before, though, we have to lay some notational and theoretical groundwork:

We tackle the following regularised model problem.

$$\left.\begin{array}{c} \min\limits_{u\in L_2(\Omega), y\in \mathring{H}^1(\Omega), v\in L_2(\Omega)} \frac{1}{2}\|y-y_d\|^2 + \frac{\nu}{2}\|u\|^2 + \frac{1}{2\varepsilon}\|v\|^2 \\ \text{s.t.} \\ \int\limits_\Omega \nabla y \cdot \nabla v \, d\Omega = \int\limits_\Omega uv \, d\Omega. \ \forall v \in \mathring{H}^1(\Omega) \\ \text{and} \\ u \in \mathcal{U} \\ y_c - y - \varepsilon v \leq 0 \ \text{ a.e. on } \Omega \end{array}\right\} \qquad (CMP^\varepsilon)$$

Recalling Theorem 2.2.12, we obtain the following necessary and sufficient optimality condition for the unique solution couple $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$:

$$(\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) + \frac{1}{\varepsilon}(\bar{v}^\varepsilon, v - \bar{v}^\varepsilon) \geq 0 \ \ \forall (u,v) \in \mathbb{U}^{\varepsilon,ad}. \tag{4.1.5}$$

where $\bar{p}_r^\varepsilon$ is the regular adjoint state defined by $\bar{p}_r^\varepsilon = S^*(\bar{y}^\varepsilon - y_d)$ and

$$\mathbb{U}^{\varepsilon,ad} := \left\{(u,v) \in L_2(\Omega) \times L_2(\Omega) \ : \ a \leq u \leq b, \ Su + \varepsilon v \geq y_c\right\}.$$

The reduced functional is defined by

$$f^\varepsilon(u,v) := \frac{1}{2}\|Su - y_d\|^2 + \frac{\nu}{2}\|u\|^2 + \frac{1}{2\varepsilon}\|v\|^2.$$

As in Chapter 3 we introduce the optimal value function

$$a : (0,1] \ni \varepsilon \mapsto f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$$

and extend it to the unregularised case by setting

$$a(0) := f(\bar{u}).$$

At this stage, the reader may want to recall the results of Chapter 3, especially Theorem 3.2.1.

We now apply the results of Section 2.2.2 and Theorem 3.2.7 to the present, less general setting. The key results are subsumed in the following theorem:

Theorem 4.1.6. *Let (CMP^ε) be given. Then, for each $\varepsilon > 0$ the unique solution couple $(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \in L_2(\Omega) \times L_2(\Omega)$ and corresponding state $\bar{y}^\varepsilon = S\bar{u}^\varepsilon$ and the unique Lagrange multiplier $\bar{\theta}^\varepsilon \in L_2(\Omega)$, $\bar{\theta}^\varepsilon \geq 0$ a.e. in Ω, fulfil the following Karush-Kuhn-Tucker system:*

$$\begin{aligned} (\bar{p}^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) \geq 0 \quad \forall u \in \mathcal{U} \\ -\varepsilon^2\bar{\theta}^\varepsilon + \bar{v}^\varepsilon = 0 \\ (\bar{\theta}^\varepsilon, \bar{y}^\varepsilon - y_c + \varepsilon\bar{v}^\varepsilon) = 0. \end{aligned} \tag{4.1.6}$$

where the ***full adjoint state*** *is defined by $\bar{p}^\varepsilon = S^*(\bar{y}^\varepsilon - y_d - \bar{\theta}^\varepsilon)$ and its* ***regular*** *and* ***singular*** *part by $\bar{p}_r^\varepsilon := S^*(\bar{y}^\varepsilon - y_d)$ and $\bar{p}_s^\varepsilon := -S^*\bar{\theta}^\varepsilon$ respectively.*

The optimality condition

$$(\bar{p}^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon)_{L_2(\Omega)} \geq 0 \quad \forall u \in \mathcal{U}$$

can be reformulated in the following pointwise fashion

$$\bar{u}^\varepsilon(x) = \min(\max(-\frac{1}{\nu}\bar{p}^\varepsilon(x), a), b) =: \Pi(\bar{p}^\varepsilon) \ \textit{f.a.a. } x \in \Omega. \tag{4.1.7}$$

Besides:

$$\bar{v}^\varepsilon(x) = -\frac{1}{\varepsilon}\min(\bar{y}^\varepsilon(x) - y_c(x), 0) \ \textit{f.a.a } x \in \Omega \tag{4.1.8}$$

Thus, $\bar{v}^\varepsilon \in \mathring{H}^1(\Omega)$ for all $\varepsilon > 0$ and in addition, $\bar{\theta}^\varepsilon \in \mathring{H}^1(\Omega)$ for all $\varepsilon > 0$.

Furthermore, the slackness condition

$$(\bar{\theta}^\varepsilon, \bar{y}^\varepsilon - y_c + \varepsilon\bar{v}^\varepsilon) = 0$$

is equivalent to the pointwise formulation

$$\bar{\theta}^\varepsilon(x)(\bar{y}^\varepsilon(x) - y_c(x) + \varepsilon\bar{v}^\varepsilon(x)) = 0 \ \textit{f.a.a } x \in \Omega \tag{4.1.9}$$

Proof. In Theorem 2.2.13 we have already proven that $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$ fulfil the KKT system

$$\begin{aligned}(\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) - (\bar{\theta}^\varepsilon, Su - S\bar{u}^\varepsilon) &\geq 0 \quad \forall u \in \mathcal{U}\\ -\varepsilon^2\bar{\theta}^\varepsilon + \bar{v}^\varepsilon &= 0\\ (\bar{\theta}^\varepsilon, \bar{y}^\varepsilon - y_c + \varepsilon\bar{v}^\varepsilon) &= 0.\end{aligned}$$

Now using the adjoint operator S^*, compare Theorem 4.1.5, we can reformulate it in the desired fashion to gain (4.1.6).

The proof of formula (4.1.7) can be found in [80], Theorem 2.33.

Thus, we can now tackle the proof of the penalty structure (4.1.8). Here, the basic idea is to prove that the function $\hat{v}$ defined by (4.1.8), i.e.

$$\hat{v}(x) = -\frac{1}{\varepsilon}\min(\bar{y}^\varepsilon(x) - y_c(x), 0),$$

satisfies $(\bar{u}^\varepsilon, \hat{v}) \in \mathbb{U}^{\varepsilon,ad}$ and

$$f(\bar{u}^\varepsilon, \hat{v}) \leq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon).$$

Due to uniqueness of the optimal solution this would entail $\bar{v}^\varepsilon = \hat{v}$ and thus all that we set out to prove. So let us now tackle this proof:

First of all, we observe that $\hat{v} \geq 0$ a.e. on Ω and by construction $(\bar{u}^\varepsilon, \hat{v}) \in \mathbb{U}^{\varepsilon,ad}$. We now still have to prove that

$$f^\varepsilon(\bar{u}^\varepsilon, \hat{v}) \leq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon).$$

We discern that thanks to admissibility of $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$ we have

$$\bar{v}^\varepsilon(x) \geq -\frac{1}{\varepsilon}(\bar{y}^\varepsilon(x) - y_c(x)), \text{ f.a.a. } x \in \Omega. \tag{4.1.10}$$

On the set

$$M^- := \{x \in \Omega \,:\, \bar{y}^\varepsilon(x) - y_c(x) < 0\},$$

we have $\bar{v}^\varepsilon(x) \geq 0$ f.a.a. $x \in M^-$. Thanks to (4.1.10) we furthermore deduce f.a.a $x \in M^-$:

$$\bar{v}^\varepsilon(x) \geq -\frac{1}{\varepsilon}(\bar{y}^\varepsilon(x) - y_c(x)) = -\frac{1}{\varepsilon}\min(\bar{y}^\varepsilon(x) - y_c(x)) = \hat{v}(x) \geq 0 \; x \in M^-.$$

As a consequence we deduce $|\hat{v}| \leq |\bar{v}^\varepsilon|$ a.e. on M^-.

By construction of $\hat{v}$ we have $\hat{v} = 0$ a.e. on $\Omega \setminus M^-$ and hence

$$|\hat{v}(x)| \leq |\bar{v}^\varepsilon(x)| \text{ f.a.a. } x \in \Omega.$$

Standard properties of the $L_2(\Omega)$-norm then enable us to conclude

$$\frac{1}{2\varepsilon} \|\hat{v}\|^2 \leq \frac{1}{2\varepsilon} \|\bar{v}^\varepsilon\|^2 .$$

Thus

$$f^\varepsilon(\bar{u}^\varepsilon, \hat{v}) \leq f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon).$$

As explained before, uniqueness of the solution $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$ now entails $\hat{v} = \bar{v}^\varepsilon$ - which is the desired penalty structure (4.1.8).
Thanks to the penalty structure (4.1.8) we know that $\bar{v}^\varepsilon \in \overset{\circ}{H}{}^1(\Omega)$ for all $\varepsilon > 0$. After all, $\min(y^\varepsilon - y_c, 0) \in \overset{\circ}{H}{}^1(\Omega)$. The improved regularity for $\bar{\theta}^\varepsilon$ then readily follows as a consequence of the equation for $\bar{v}^\varepsilon$ and $\bar{\theta}^\varepsilon$ in (4.1.6).
The fact that the pointwise slackness condition (4.1.9) holds is a consequence of $\bar{\theta}^\varepsilon \in C^-$ (which means $\bar{\theta}^\varepsilon \geq 0$ a.e. on Ω), $\bar{y}^\varepsilon - y_c + \varepsilon\bar{v}^\varepsilon \geq 0$ a.e. on Ω and standard Lebesgue integration theory. After all, $L_2(\Omega) \ni f, g \geq 0$ and $(f, g) = 0$ imply $f(x)g(x) = 0$ f.a.a. $x \in \Omega$. □

Our aim now is to prove an a priori estimate

$$\left\|\bar{u}^{\varepsilon^N} - \bar{u}\right\|^2 \lesssim \varepsilon^{\gamma N},\ \gamma > 0,$$

compare also (4.0.1). To do so, we need a couple of auxiliary results: The starting point will be an improved bound for the Lagrange multiplier $\bar{\theta}^\varepsilon$.

Lemma 4.1.7. *Let $1 \leq p \leq 2$ and p' be its dual exponent, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$. Then, for $\bar{\theta}^\varepsilon$ the following bound is valid:*

$$\|\bar{\theta}^\varepsilon\|_{L_p(\Omega)} \lesssim (\frac{1}{\tau})^{1-2/p'} \varepsilon^{-3/p'},$$

where τ is the constant from the continuous Slater point in Assumption 4.1.4.
In particular, there holds:

$$\|\bar{\theta}^\varepsilon\|_{L_1(\Omega)} \lesssim \frac{1}{\tau}.$$

Both constants are independent of p'.

Proof. The basic idea of the proof is to first prove a uniform $L_1(\Omega)$ bound for the multiplier

$\bar{\theta}^\varepsilon$ and then demonstrate the improved bound for the $\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)}$, $1 \le p \le 2$ by interpolation arguments.

Thus, let us commence with the uniform $L_1(\Omega)$-bound. First of all, we observe that as an L_2-function for every fixed ε, $\bar{\theta}^\varepsilon \in L_p(\Omega)$ for all $1 \le p \le 2$.

Let us now recall the optimality condition for $\bar{u}^\varepsilon$ in (4.1.6) in the following slightly reformulated fashion where we split the full adjoint state into its regular and singular part:

$$\begin{aligned}0 \le (\bar{p}^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) &= (\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) - (\bar{p}_s^\varepsilon, u - \bar{u}^\varepsilon)\\ &= (\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) - (S^*\bar{\theta}^\varepsilon, u - \bar{u}^\varepsilon)\\ &= (\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u - \bar{u}^\varepsilon) - (\bar{\theta}^\varepsilon, Su - S\bar{u}^\varepsilon) \ \ \forall u \in \mathcal{U}.\end{aligned}$$

Inserting $u = u^s$ from Assumption 4.1.4 into the inequality above, we can now deduce after a short rearrangement:

$$\begin{aligned}(\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u^s - \bar{u}^\varepsilon) &\ge (\bar{\theta}^\varepsilon, Su^s - S\bar{u}^\varepsilon)\\ &= (\bar{\theta}^\varepsilon, Su^s - y_c) + (\bar{\theta}^\varepsilon, y_c - S\bar{u}^\varepsilon)\end{aligned}$$

Using the slackness equation in (4.1.6) and the fact that $\bar{\theta}^\varepsilon \ge 0$ a.e. on Ω since $\bar{\theta}^\varepsilon \in C^-$, we can proceed in the following way:

$$\begin{aligned}(\bar{\theta}^\varepsilon, Su^s - y_c) + (\bar{\theta}^\varepsilon, y_c - S\bar{u}^\varepsilon) &= \tau\left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)} + (\bar{\theta}^\varepsilon, y_c - S\bar{u}^\varepsilon)\\ &= \tau\left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)} - \frac{1}{\varepsilon}\left\|\bar{v}^\varepsilon\right\|^2\end{aligned}$$

All in all, we thus gain:

$$\left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)} \le \frac{1}{\tau}((\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u^s - \bar{u}^\varepsilon) + \frac{1}{\varepsilon}\left\|\bar{v}^\varepsilon\right\|^2)$$

Now, recall that due to Theorem 3.2.1 $a(\varepsilon) \le a(0)$ and thus in particular

$$\frac{1}{\varepsilon}\left\|\bar{v}^\varepsilon\right\|^2 \lesssim 1 \tag{4.1.11}$$

Presently, we can harness continuity of S and S^* as well as the uniform bound of functions belonging to $\mathcal{U}$ derived in (4.1.1) to obtain:

$$\begin{aligned}\left\|\bar{\theta}^\varepsilon\right\|_{L_1(\Omega)} &\le \frac{1}{\tau}|(\bar{p}_r^\varepsilon + \nu\bar{u}^\varepsilon, u^s - \bar{u}^\varepsilon) + \frac{1}{\varepsilon}\left\|\bar{v}^\varepsilon\right\|^2|\\ &\lesssim \frac{1}{\tau}(\|S^*\|\,\|\bar{y}^\varepsilon - y_d\| + 1)\\ &\le \frac{1}{\tau}(\|S^*\|\,(\|S\|\,\|u^\varepsilon\| + \|y_d\| + 1)) \lesssim \frac{1}{\tau}\end{aligned}$$

Thus, we have ascertained the uniform L_1-bound

$$\left\|\bar{\theta}^{\varepsilon}\right\|_{L_1(\Omega)} \lesssim \frac{1}{\tau} \tag{4.1.12}$$

We can now utilise the interpolation theory arguments of Section 2.1.7 to derive the desired result.

We observe that the $L_p(\Omega)$-spaces have the property that

$$(L_p(\Omega))^* \cong L_{p'}(\Omega),\ \frac{1}{p}+\frac{1}{p'}=1,\ 1<p<\infty, \tag{4.1.13}$$

see e.g. [85], Theorem 10.44, and

$$L_1(\Omega) \hookrightarrow L_\infty(\Omega)^* \tag{4.1.14}$$

by Hölder's inequality, compare [85], Theorem 10.43 or [40].

Since $\bar{\theta}^{\varepsilon} \in L_2(\Omega)$ for all fixed $\varepsilon > 0$, $\bar{\theta}^{\varepsilon}$ thus can also be interpreted as a functional on $L_{p'}(\Omega)$, $2 \le p' \le \infty$.

From the equation for $\bar{\theta}^{\varepsilon}$ and $\bar{v}^{\varepsilon}$ in (4.1.6) we gain

$$\left\|\bar{\theta}^{\varepsilon}\right\| = \frac{1}{\varepsilon^2}\left\|\bar{v}^{\varepsilon}\right\|.$$

Combining this with the bound (4.1.11), we deduce:

$$\left\|\bar{\theta}^{\varepsilon}\right\|_{L_2(\Omega)} \lesssim \varepsilon^{-3/2}. \tag{4.1.15}$$

Setting σ according to $\frac{1}{p'} = \frac{1-\sigma}{2}$, $0<\sigma<1$, i.e. $\sigma = 1-\frac{2}{p'}$ we have by Definition 2.1.41

$$L_{p'}(\Omega) = L_{p',p'}(\Omega) = (L_2(\Omega), L_\infty(\Omega))_{\sigma,p'},\ 1<p'<\infty$$

Presently, using the properties of Lorentz spaces, Definition 2.1.41, the interpolation estimate (2.1.14) from Theorem 2.1.40 and $L_2(\Omega) \cong L_2(\Omega)^*$, we obtain for $1 < p \le 2$ and its dual exponent p':

$$\begin{aligned}\left\|\bar{\theta}^{\varepsilon}\right\|_{L_{p,p}(\Omega)} &= \left\|\bar{\theta}^{\varepsilon}\right\|_{L_p(\Omega)}\\ &= \left\|\bar{\theta}^{\varepsilon}\right\|_{L_{p'}(\Omega)^*}\\ &\le \left\|\bar{\theta}^{\varepsilon}\right\|_{L_2(\Omega)^*}^{1-\sigma}\left\|\bar{\theta}^{\varepsilon}\right\|_{L_\infty(\Omega)^*}^{\sigma}\\ &= \left\|\bar{\theta}^{\varepsilon}\right\|_{L_2(\Omega)}^{1-\sigma}\left\|\bar{\theta}^{\varepsilon}\right\|_{L_\infty(\Omega)^*}^{\sigma},\ \ \sigma = 1-\frac{2}{p'}\end{aligned} \tag{4.1.16}$$

Now recall (4.1.13) and (4.1.14). Together with the bounds (4.1.12) and (4.1.15) these relations allow us to continue our estimates for $1 < p \le 2$ and its dual exponent p' in the following way

$$\begin{aligned}\|\bar{\theta}^{\varepsilon}\|_{L_p(\Omega)} &\leq \|\bar{\theta}^{\varepsilon}\|^{1-\sigma}_{L_2(\Omega)}\|\bar{\theta}^{\varepsilon}\|^{\sigma}_{L_\infty(\Omega)^*}\\ &\leq \|\bar{\theta}^{\varepsilon}\|^{1-\sigma}_{L_2(\Omega)}\|\bar{\theta}^{\varepsilon}\|^{\sigma}_{L_1(\Omega)}\\ &\lesssim \varepsilon^{-\frac{3}{2}+\frac{3\sigma}{2}}(\frac{1}{\tau})^{\sigma} = (\frac{1}{\tau})^{1-2/p'}\varepsilon^{-3/p'} \;\; \sigma = 1-\frac{2}{p'}.\end{aligned}$$

Combining this with our uniform bound $L_1(\Omega)$-bound for $\bar{\theta}^{\varepsilon}$, (4.1.12), we can include the case $p = 1$ and its dual exponent $p' = \infty$ in the above estimate and thus finally deduce the desired result:

$$\|\bar{\theta}^{\varepsilon}\|_{L_p(\Omega)} \lesssim (\frac{1}{\tau})^{1-2/p'}\varepsilon^{-3/p'}, \; 1 \leq p \leq 2.$$

□

Before we move on to the next lemma, let us for notational convenience introduce a generic constant $s(\tau)$ which is assigned to indicate that in those estimates where it appears **negative powers** of τ enter. The negative powers themselves usually depend on p' and the specific setting, hence the attribute 'generic'. The estimate in Lemma 4.1.7 can in this way be shortened to

$$\|\bar{\theta}^{\varepsilon}\|_{L_p(\Omega)} \lesssim s(\tau)\varepsilon^{-3/p'}, \;\text{ with } s(\tau) := (\frac{1}{\tau})^{1-2/p'}.$$

Likewise, for the embedding constant of the embedding $\overset{\circ}{H}{}^1(\Omega) \hookrightarrow L_{p'}(\Omega)$, $1 \leq p' < \infty$ in 2d, $1 \leq p' \leq 6$ in 3d we introduce in the same spirit as for $s(\tau)$ another generic constant $c(p')$ which is assigned to indicate that the embedding constant enters with some positive power depending on the specific setting. In particular, the appearance of $c(p')$ indicates that in 2d, $c(p') \to \infty$ as $p' \to \infty$. In a setting restricted to 3d we will do without explicitly stating the constant as we are restricted to $p' \leq 6$ anyway and do not have to investigate the case $p' \to \infty$.

After this notational detour, let us now continue our stability estimates.

The next lemma provides a bound for the H^1-semi norm of the violation of the state constraint:

Lemma 4.1.8. *Let $2 \leq p' < \infty$ in case $d = 2$ and $2 \leq p' \leq 6$ in case $d = 3$. Furthermore, let p denote its dual exponent, i.e. $\frac{1}{p'} + \frac{1}{p} = 1$. For the difference $(\bar{y}^{\varepsilon} - y_c)^-$ we have the following error bound*

$$|(\bar{y}^{\varepsilon} - y_c)^-|_{H^1(\Omega)} \lesssim c(p')s(\tau)\varepsilon^{1-1/p'}. \tag{4.1.17}$$

Proof. Evidently, $(\bar{y}^{\varepsilon} - y_c)^- \in \overset{\circ}{H}{}^1(\Omega)$, since $\bar{y}^{\varepsilon} = 0$ on $\partial\Omega$ and $y_c|_{\partial\Omega} < 0$ by assumption!.

Thus, we can use it as a test function for the bilinear form b, (4.1.2), to obtain

$$\begin{aligned}|(\bar{y}^\varepsilon - y_c)^-|^2_{H^1(\Omega)} &= |(\nabla(\bar{y}^\varepsilon - y_c), \nabla(\bar{y}^\varepsilon - y_c)^-| \\ &= |(\bar{u}^\varepsilon + \Delta y_c, (\bar{y}^\varepsilon - y_c)^-|\end{aligned}$$

Now we can take advantage of the penalty structure of $\bar{v}^\varepsilon$, compare (4.1.8), to get

$$\begin{aligned}|(\bar{y}^\varepsilon - y_c)^-|^2_{H^1(\Omega)} &= |(\bar{u}^\varepsilon + \Delta y_c, -\varepsilon\bar{v}^\varepsilon)| \\ &\leq \|\bar{u}^\varepsilon + \Delta y_c\| \, \|\varepsilon\bar{v}^\varepsilon\| \, .\end{aligned} \tag{4.1.18}$$

Presently, let us estimate $\|\bar{v}^\varepsilon\|$:
Recalling the complimentary slackness condition (4.1.9), $\bar{\theta}^\varepsilon \geq 0$ a.e. in Ω and the embedding $H^1(\Omega) \hookrightarrow L_{p'}(\Omega)$, $1 \leq p' < \infty$ in case $d = 2$ and $1 \leq p' \leq 6$ in case $d = 3$, compare Theorem 2.1.35, we obtain

$$\begin{aligned}\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2 = (\bar{\theta}^\varepsilon, y_c - \bar{y}^\varepsilon) &\leq \|\bar{\theta}^\varepsilon\|_{L_p(\Omega)} \left\|(\bar{y}^\varepsilon - y_c)^-\right\|_{L_{p'}(\Omega)} \\ &\lesssim c(p') \|\bar{\theta}^\varepsilon\|_{L_p(\Omega)} \left\|(\bar{y}^\varepsilon - y_c)^-\right\|_{H^1(\Omega)} .\end{aligned}$$

Using the Poincaré-Friedrich-inequality Theorem 2.1.33 ($(\bar{y} - y_c)^- \in \overset{\circ}{H}{}^1(\Omega)$!) and the results of Lemma 4.1.7, we can pursue our estimates in the following way:

$$\begin{aligned}\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2 &\lesssim c(p') \|\bar{\theta}^\varepsilon\|_{L_p(\Omega)} \left\|(\bar{y}^\varepsilon - y_c)^-\right\|_{H^1(\Omega)} \\ &\lesssim \|\bar{\theta}^\varepsilon\|_{L_p(\Omega)} |(\bar{y}^\varepsilon - y_c)^-|_{H^1(\Omega)} \\ &\lesssim c(p') s(\tau) \varepsilon^{-3/p'} |(\bar{y}^\varepsilon - y_c)^-|_{H^1(\Omega)}\end{aligned}$$

Thus, we have gained:

$$\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2 \lesssim c(p') s(\tau) \varepsilon^{-3/p'} |(\bar{y}^\varepsilon - y_c)^-|_{H^1(\Omega)}.$$

A short rearrangement of this bound yields

$$\|\bar{v}^\varepsilon\| \lesssim c(p') s(\tau) \varepsilon^{\frac{1}{2} - \frac{3}{2p'}} |(\bar{y}^\varepsilon - y_c)^-|^{\frac{1}{2}}_{H^1(\Omega)}.$$

Let us now insert this bound in (4.1.18) bearing in mind that $\|\bar{u}^\varepsilon + \Delta y_c\|$ is uniformly bounded in $L_2(\Omega)$:

$$|(\bar{y}^\varepsilon - y_c)^-|^2_{H^1(\Omega)} \lesssim \varepsilon \|\bar{v}^\varepsilon\| \lesssim c(p') s(\tau) \varepsilon^{\frac{3}{2} - \frac{3}{2p'}} |(\bar{y}^\varepsilon - y_c)^-|^{\frac{1}{2}}_{H^1(\Omega)}$$

Dividing by $|(\bar{y}^\varepsilon - y_c)^-|^{1/2}_{H^1(\Omega)}$ (if this term were 0, the postulated bound (4.1.17) trivially holds) yields:

$$|(\bar{y}^\varepsilon - y_c)^-|_{H^1(\Omega)}^{\frac{3}{2}} \lesssim c(p')s(\tau)\varepsilon^{\frac{3}{2}-\frac{3}{2p'}}.$$

(4.1.17) then readily follows. □

A consequence of the previous lemma is the following corollary:

Corollary 4.1.9. *For $2 \le p' < \infty$ in case $d = 2$ and $2 \le p' \le 6$ in case $d = 3$ we have the a priori bound*

$$\|\bar{v}^\varepsilon\| \lesssim \min(\varepsilon^{1/2}, c(p')s(\tau)\varepsilon^{1-\frac{2}{p'}})$$

Proof. First of all, we observe that boundedness of the continuous optimal value function a, compare Theorem 3.2.1, in particular yields

$$\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2 \le a(\varepsilon) = f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \le f(\bar{u}) \lesssim 1.$$

Thus

$$\|\bar{v}^\varepsilon\| \lesssim \varepsilon^{1/2},$$

which gives the first estimate.
Conversely, recall the slackness equation (4.1.9). We obtain as in the proof of Lemma 4.1.8

$$\frac{1}{\varepsilon}\|\bar{v}^\varepsilon\|^2 = |(\bar{\theta}^\varepsilon, \bar{y}^\varepsilon - y_c)| \lesssim c(p')\left\|\bar{\theta}^\varepsilon\right\|_{L_p(\Omega)} |(\bar{y}^\varepsilon - y_c)^-|_{H^1(\Omega)}, \quad \frac{1}{p} + \frac{1}{p'} = 1$$

Now we can just plug in the bounds for $\bar{\theta}^\varepsilon$, Lemma 4.1.7, and the energy norm of $(\bar{y}^\varepsilon - y_c)^-$ from Lemma 4.1.8 to derive the desired result. □

Having collected these auxiliary results we are now in a position to prove an a priori estimate

$$\|\bar{u} - \bar{u}^\varepsilon\| \lesssim c(p')s(\tau)\varepsilon^\gamma, \gamma > 0$$

into which we will insert - as the last step - $\varepsilon = \varepsilon^N$ to gain the desired estimate for (4.0.1).

Theorem 4.1.10. *Let $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. The following a priori estimates hold true:*

$$\|\bar{u}^\varepsilon - \bar{u}\|^2 \lesssim \begin{cases} c(p')s(\tau)\varepsilon^{1-4/p'} & \text{if } d = 2,\ 4 < p' < \infty \\ s(\tau)\varepsilon^{1/3} & \text{if } d = 3, (p' = 6) \end{cases} \tag{4.1.19}$$

Proof. The proof uses Taylor expansion of f^ε at $(\bar{u}^\varepsilon, \bar{v}^\varepsilon)$. We refer to Chapter 8 in [48] and especially Theorem 8.16 in [48]. Since f^ε is quadratic, we gain

$$\begin{aligned} f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) &= f^\varepsilon(\bar{u}, 0) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \\ &= D^1 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \Big[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\Big] + D^2 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \Big[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\Big]^2. \end{aligned}$$

Differentiating now yields

$$D^1 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \Big[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\Big] = (\bar{p}_r^\varepsilon + \nu \bar{u}^\varepsilon, \bar{u} - \bar{u}^\varepsilon) - \frac{1}{\varepsilon} \|\bar{v}^\varepsilon\|^2 \geq 0,$$

where we have also used (4.1.5).
For the second derivatives we can conclude

$$D^2 f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \cdot \Big[\bar{u} - \bar{u}^\varepsilon, \quad -\bar{v}^\varepsilon\Big]^2 = \|\bar{y} - \bar{y}^\varepsilon\|^2 + \nu \|\bar{u} - \bar{u}^\varepsilon\|^2 + \frac{1}{\varepsilon} \|\bar{v}^\varepsilon\|^2.$$

Thus

$$f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) \geq \nu \|\bar{u} - \bar{u}^\varepsilon\|^2.$$

Using the differentiability of the optimal value function, compare again Theorem 3.2.1, we deduce

$$\begin{aligned} \nu \|\bar{u} - \bar{u}^\varepsilon\|^2_{L_2(\Omega)} &\leq f(\bar{u}) - f^\varepsilon(\bar{u}^\varepsilon, \bar{v}^\varepsilon) = a(0) - a(\varepsilon) \\ &= \int_0^\varepsilon -a'(t)\, dt = \int_0^\varepsilon \frac{1}{t^2} \left\|\bar{v}^t\right\|^2\, dt. \end{aligned} \tag{4.1.20}$$

Harnessing $\bar{\theta}^t \geq 0$ for all $t > 0$, Hölder's inequality and the complimentary slackness condition in (4.1.6) with $\varepsilon = t$, we deduce:

$$\frac{1}{t^2} \left\|\bar{v}^t\right\|^2 = \frac{1}{t}(\bar{\theta}^t, y_c - \bar{y}^t) \leq \frac{1}{t} \left\|\bar{\theta}^t\right\|_{L_p(\Omega)} \left\|(\bar{y}^t - y_c)^-\right\|_{L_{p'}(\Omega)}, \quad \frac{1}{p} + \frac{1}{p'} = 1$$

The embedding $\overset{\circ}{H}{}^1(\Omega) \hookrightarrow L_{p'}(\Omega)$, $1 \leq p' < \infty$ if $d = 2$ and $1 \leq p' \leq 6$ if $d = 3$, and the results of Lemma 4.1.7 and Lemma 4.1.8 for $\varepsilon = t$ enable us to continue our estimates in the ensuing vein:

$$\begin{aligned} \frac{1}{t^2} \left\|\bar{v}^t\right\|^2 &\leq \frac{1}{t} \left\|\bar{\theta}^t\right\|_{L_p(\Omega)} \left\|(\bar{y}^t - y_c)^-\right\|_{L_{p'}(\Omega)} \\ &\leq c(p')s(\tau)\frac{1}{t}(t^{1-4/p'}) = c(p')s(\tau)t^{-4/p'}. \end{aligned}$$

Presently, we can pick up the thread we left off in (4.1.20) and continue our estimates taking advantage of the estimates above (remember that we assumed $p' > 4$ if $d = 2$ and $p' = 6$ if

$d = 3!$).

$$
\begin{aligned}
\nu \left\|\bar{u} - \bar{u}^{\varepsilon}\right\|^2_{L_2(\Omega)} &\leq \int\limits_0^{\varepsilon} \frac{1}{t^2} \left\|\bar{v}^t\right\|^2 \\
&\leq c(p')s(\tau) \int\limits_0^{\varepsilon} t^{-4/p'}\, dt \\
&= c(p')s(\tau)\frac{1}{1-4/p'}\varepsilon^{1-4/p'} \lesssim c(p')s(\tau)\varepsilon^{1-4/p'}, \;\; p' > 4
\end{aligned}
$$

In case $d = 2$, this gives the first estimate in (4.1.19). If $d = 3$ we just have to insert $p' = 6$ into the estimates above. Here, note that the dependence on the embedding constant $c(p')$ can be neglected since we fix it to $c(p') = c(6)$. Ultimately, we deduce in case $d = 3$

$$\nu \left\|\bar{u} - \bar{u}^{\varepsilon}\right\|^2_{L_2(\Omega)} \lesssim 3s(\tau)\varepsilon^{1/3} \lesssim s(\tau)\varepsilon^{1/3}.$$

This gives the second estimate in (4.1.19). □

As a corollary we now obtain the desired estimate for the difference $\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2$ in (4.0.1) by inserting $\varepsilon = \varepsilon^N$ in the bounds of Theorem 4.1.10:

Corollary 4.1.11. *Let* $4 < p' < \infty$ *in case* $d = 2$ *and* $p' = 6$ *in case* $d = 3$. *Then the following a priori upper bounds are valid:*

$$\left\|\bar{u}^{\varepsilon^N} - \bar{u}\right\|^2 \lesssim \begin{cases} c(p')s(\tau)\varepsilon^{(1-4/p')N} & \textit{if } d = 2,\; 4 < p' < \infty \\ s(\tau)\varepsilon^{N/3} & \textit{if } d = 3, (p' = 6) \end{cases} \tag{4.1.21}$$

Hence, the γ *in the estimator* (4.0.1) *is defined by*

$$\gamma := \begin{cases} (1 - 4/p') & \textit{if } d = 2,\; 4 < p' < \infty \textit{ fixed} \\ 1/3 & \textit{if } d = 3. \end{cases} \tag{4.1.22}$$

Proof. Since the constants in the estimates of Theorem 4.1.10 do not depend on ε, we can merely insert $\varepsilon = \varepsilon^N$ in (4.1.19) to obtain the desired result. □

In deriving the a priori result in terms of the error in regularisation, Theorem 4.1.10, we used a different approach compared to the techniques employed in [52] and [18], where higher regularity of the solution operator - it maps to $C^{0,\alpha}(\bar{\Omega})$ - was used to gain estimates on the maximal violation $\|(y_c - \bar{y}^{\varepsilon})^+\|_{L_\infty(\Omega)}$. The reasons for pursuing this alternative path are twofold: First, we wanted to gain improved bounds for the multiplier $\bar{\theta}^{\varepsilon}$ in $L_p(\Omega)$, $1 \leq p \leq 2$, compare Lemma 4.1.7, and secondly, and more importantly, we wanted to prove the a priori bound **without any addtional regularity** for the PDE making them potentially applicable

in settings with less regularity such as constraints on the pressure as in Section 2.4.3.
Let us also stress that to the best of author's knowledge Theorem 4.1.10 is the first result where Hölder-stability for the regularised problems has been proven without any such additional regularity. In fact, the results - at least for $d = 2$ - even almost match the already cited existing ones such as [52], Theorem 11, where in case $d = 2$, exactly $\varepsilon^{1/2}$ was proven for more general elliptic PDE, but also under stricter regularity assumptions, namely $S \in \mathcal{L}(L_2(\Omega), C^{0,1}(\bar{\Omega}))$.

At this stage let us shortly recall our aim (4.0.1):

$$\begin{aligned}\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|_{\mathbb{U}}^2 &\lesssim \left\|\bar{u}^{\varepsilon^N} - \bar{u}\right\|_{\mathbb{U}}^2 + \left\|\bar{U}_k^\varepsilon - \bar{u}^{\varepsilon^N}\right\|_{\mathbb{U}}^2 \\ &\lesssim \varepsilon^{\gamma N} + \mathcal{E}_r^2 + \mathcal{E}_s^2.\end{aligned}$$

For the term $\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2$ we have proven the desired a priori estimate in Theorem 4.1.10 and Corollary 4.1.11. Now, we have to tackle the term $\left\|\bar{U}_k^\varepsilon - \bar{u}^{\varepsilon^N}\right\|^2$. To this end, though, we first have to introduce a discretisation of the original problem (CMP) and a corresponding regularisation which we will do in the next two sections.

4.1.3 The Unregularised Discretised Problem

We introduce a series of triangulations of Ω, $\mathcal{T}_k$ such that:

$$\bar{\Omega} = \bigcup_{T \in \mathcal{T}_k} \bar{T}.$$

This enables us to define the following spaces:
The control space $\mathbb{U}_k$ will be either left undiscretised, this is the so called **variational discretisation** approach, compare [44], i.e.

$$\mathbb{U}_k = \mathbb{U}$$

or discretised by piecewise constant functions, the **full discretisation** approach, i.e

$$\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega)).$$

We will treat both types of ansatz spaces simultaneously in this and the following section.
For the discretisation of the state space $\mathbb{Y}$ we choose (regardless of the control discretisation)

$$\mathbb{Y}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, \mathring{H}^1(\Omega)),$$

compare also Definition 2.3.4. For $\mathbb{V}_k$ we choose (again irrespective of the control discretisation)

$$\mathbb{V}_k = (\mathbb{Y}_k, L_2(\Omega)).$$

We introduce the set

$$\mathcal{U}_k := \{U \in \mathbb{U}_k \ : \ a \le U \le b \text{ a.e. in } \Omega\},$$

for which (4.1.1) is also valid. Despite the fact that $\mathbb{U}_k \subset \mathbb{U}_{k+1}$ (irrespective of the control discretisation) we do in general have that

$$\mathbb{U}_k^{ad} \not\subset \mathbb{U}_{k+1}^{ad}.$$

Let us now lay out the discrete unregularised problem:

$$\left.\begin{array}{c} \min\limits_{U \in \mathbb{U}_k, Y \in \mathbb{Y}_k} \frac{1}{2}\|Y - y_d\|^2 + \frac{\nu}{2}\|U\|^2 \\ \text{s.t.} \\ \int\limits_\Omega \nabla Y \cdot \nabla W \, d\Omega = \int\limits_\Omega UW \, d\Omega \ \ \forall W \in \mathbb{Y}_k \\ \text{and} \\ U \in \mathcal{U}_k \\ I_k y_c - Y \le 0 \ \text{ a.e. on } \Omega. \end{array}\right\} \qquad (DMP_k)$$

For the verification of Assumptions (A2)-(A4) we again refer to the discussions of Section 2.4.1. Note that we do not assume anything apart from Property (Pr8) for the operator I_k. To ensure that (A1) is fulfilled, we next assume that the feasible set

$$\mathbb{U}_k^{ad} := \{U \in \mathcal{U}_k \ : \ I_k y_c - S_k U \le 0\}$$

is non-empty for all k. We observe that every sequence $\left\{\hat{U}_k\right\} \subset \mathbb{U}$ with $\hat{U}_k \in \mathbb{U}_k^{ad}$ is uniformly bounded thanks to (4.1.1).

The discrete solution operator S_k fulfils a Galerkin-orthogonality property which is the subject of the next lemma:

Lemma 4.1.12. *For any $q \in L_2(\Omega)$ we have*

$$(\nabla(Sq - S_k q), \nabla W) = 0 \ \ \forall W \in \mathbb{Y}_k. \qquad (4.1.23)$$

Proof. Since $\mathbb{Y}_k \subset \overset{\circ}{H}{}^1(\Omega)$ we can deduce that

$$(\nabla(Sq - S_k q), \nabla W) = (q, W) - (q, W) = 0.$$

□

For the discrete adjoint operator S_k^* we are able to prove a result mirroring Theorem 4.1.5:

Theorem 4.1.13. *Suppose we interpret S_k as an operator $S_k : L_2(\Omega) \to L_2(\Omega)$. Given $q \in L_2(\Omega)$, $Z = S_k^* q \in \mathbb{Y}_k$ solves the variational problem*

$$\int\limits_{\Omega} \nabla Z \cdot \nabla W \, d\Omega = \int\limits_{\Omega} qW \, d\Omega \;\; \forall W \in \mathbb{Y}_k. \tag{4.1.24}$$

Thus

$$\|S_k^* q\|_{\overset{\circ}{H}{}^1(\Omega)} = \|Z\|_{\overset{\circ}{H}{}^1(\Omega)} \lesssim \|q\| \, .$$

Proof. By definition of the adjoint operator, compare Section 2.1.2, S_k^* satisfies

$$(h, S_k^* q) = (S_k h, q), \; \forall h, q \in L_2(\Omega).$$

Defining Z as the unique solution $Z \in \mathbb{Y}_k \subset \overset{\circ}{H}{}^1(\Omega)$ of (4.1.24), we obtain by employing (4.1.24), the properties of the solution operator S and Galerkin orthogonality (4.1.23):

$$(h, Z) = (\nabla Sh, \nabla Z) = (\nabla S_k h, \nabla Z) = (q, S_k h) \;\; \forall q, h \in L_2(\Omega).$$

Thus

$$(h, S_k q) = (h, Z) = (q, S_k h) \;\; \forall q, h \in L_2(\Omega)$$

Hence $Z = S_k^* q$ which completes the proof. □

The discrete solution $\bar{U}_k$ and its corresponding optimal state $\bar{Y}_k = S_k \bar{U}_k$ and adjoint state $\bar{P}_k := S_k^*(\bar{Y}_k - y_d)$ satisfy the following familiar first order necessary and sufficient optimality condition:

$$(\bar{P}_k + \nu \bar{U}_k, U - \bar{U}_k) \geq 0 \;\; \forall U \in \mathbb{U}_k^{ad}.$$

We next turn to the regularised discretised problem whose solution will be the one actually computed by the algorithm presented in Chapter 5.

4.1.4 The Regularised Discretised Problem

We tackle the following discrete regularised problem:

$$\left.\begin{array}{c} \displaystyle\min_{U\in\mathbb{U}_k, Y\in\mathbb{Y}_k, V\in\mathbb{V}_k} \frac{1}{2}\|Y-y_d\|^2+\frac{\nu}{2}\|U\|^2+\frac{1}{2\varepsilon}\|V\|^2 \\ \text{s.t.} \\ \displaystyle\int_\Omega \nabla Y\cdot\nabla W\,d\Omega = \int_\Omega UW\,d\Omega\ \forall W\in\mathbb{Y}_k \\ \text{and} \\ U\in\mathcal{U}_k \\ I_k y_c - Y - \varepsilon V \le 0 \ \text{ a.e. on } \Omega. \end{array}\right\} \qquad (DMP_k^\varepsilon)$$

The next theorem is an application of the results of Theorem 3.2.8 to the present setting:

Theorem 4.1.14. *The unique discrete solution couple $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ fulfils the following necessary and sufficient optimality system:*

$$\begin{aligned} (S_k^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon) + \nu\bar{U}_k^\varepsilon, U-\bar{U}_k^\varepsilon) &\ge 0 \quad \forall U\in\mathcal{U}_k \\ -\varepsilon^2\bar{\theta}_k^\varepsilon + \bar{V}_k^\varepsilon &= 0 \\ (\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c + \varepsilon\bar{V}_k^\varepsilon) &= 0 \end{aligned} \tag{4.1.25}$$

*As in the continuous case, we define the discrete **full adjoint state** by $\bar{P}_k^\varepsilon := S_k^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)$, its **regular** part by $\bar{P}_{k,r}^\varepsilon := S_k^*(\bar{Y}_k^\varepsilon - y_d)$ and its **singular** part by $\bar{P}_{k,s}^\varepsilon := -S_k^*\bar{\theta}_k^\varepsilon$.*

Proof. The KKT system (4.1.25) is an immediate consequence of the adjoint representation (4.1.24) and Theorem 3.2.8. □

Before we turn the derivation of the estimator for the difference $\left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|^2$, let us prove a projection relation for $\bar{U}_k^\varepsilon$ mirroring (4.1.7). To do so - for future use in a different setting - we will first formulate a fairly general result for the projection of functions on box-constrained convex sets:

Lemma 4.1.15. *Let $\mathbb{L}$ either be a closed subspace of $L_2(\Omega)$ or $\mathbb{L} = L_2(\Omega)$. Furthermore, let the convex and closed set $W \subset \mathbb{L}$ be defined by*

$$W := \{w \in \mathbb{L} \,:\, \omega_1 \le w \le \omega_2 \ a.e. \ in\ \Omega\}$$

with $\omega_1, \omega_2 \in \mathbb{R}\cup\{-\infty,\infty\}$, $\omega_1 \le \omega_2$ and let for $g \in L_2(\Omega)$

$$(\Pi_{\omega_2}^{\omega_2} g)(x) = \min(\max(g(x),\omega_1)),\omega_2) \ \ x\in\Omega. \tag{4.1.26}$$

denote the pointwise cut-off of the function g which is in addition assumed to fulfil

$$\Pi_{\omega_2}^{\omega_2}(z) \in \mathbb{L} \ \ \forall z \in \mathbb{L}. \tag{4.1.27}$$

Besides, we denote by $P_{\mathbb{L}}$ *the orthogonal projection on* $\mathbb{L}$ *defined for arbitrary* $g \in L_2(\Omega)$ *by*

$$(P_{\mathbb{L}}(g) - g, z) = 0 \ \ \forall z \in \mathbb{L}.$$

Then the projection $\Pi_W(g)$ *of a function* $g \in L_2(\Omega)$ *on* W *defined by (compare Theorem 2.1.7)*

$$(\Pi_W(g) - g, w - \Pi_W(g)) \geq 0 \ \ \forall w \in W \tag{4.1.28}$$

satisfies:

$$(\Pi_W g)(x) = (\Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g))(x) \ \ f.a.a\ x \in \Omega.$$

Proof. Let us first investigate the case $\mathbb{L} = L_2(\Omega)$ before we turn to the general case $\mathbb{L} \subset L_2(\Omega)$:
Since $\mathbb{L} = L_2(\Omega)$ $P_{\mathbb{L}} = I$, where I is the identity operator. Recalling Theorem 2.1.7, we gain for the projection $\Pi_W(g)$ of an arbitrary function $g \in L_2(\Omega)$ on W:

$$(\Pi_W(g) - g, w - \Pi_W(g)) \geq 0 \ \ \forall w \in W$$

Inserting $w = \Pi_{\omega_2}^{\omega_2}(g) \in W$, we deduce

$$\begin{aligned} 0 \leq (\Pi_W(g) - g, \Pi_{\omega_2}^{\omega_2}(g) - \Pi_W(g)) &= (\Pi_W(g) - \Pi_{\omega_1}^{\omega_2} g + \Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_2}^{\omega_2}(g) - \Pi_W(g)) \\ &= -\left\|\Pi_W g - \Pi_{\omega_2}^{\omega_2} g\right\|^2 + (\Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_1}^{\omega_2} g - \Pi_W g) \end{aligned}$$

Rearranging this inequality, we derive

$$\left\|\Pi_W g - \Pi_{\omega_2}^{\omega_2} g\right\|^2 \leq (\Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_1}^{\omega_2} g - \Pi_W g). \tag{4.1.29}$$

Now we distinguish between three (not necessarily non-empty) sets:

$$\begin{aligned} \Omega^+ &:= \{x \in \Omega \ : \ g(x) > \omega_2 \text{ a.e}\} \\ \Omega^- &:= \{x \in \Omega \ : \ g(x) < \omega_1 \text{ a.e. }\} \\ \Omega^0 &:= \Omega \setminus (\Omega^+ \cup \Omega^-). \end{aligned}$$

Distinguishing between these different sets on the right hand side in (4.1.29), we derive by definition of $\Pi_{\omega_1}^{\omega_2} g$:

$$\begin{aligned}\left\|\Pi_W g - \Pi_{\omega_2}^{\omega_2} g\right\|^2 &\leq (\Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_1}^{\omega_2} g - \Pi_W g)_{L_2(\Omega^+)} + (\Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_1}^{\omega_2} g - \Pi_W g)_{L_2(\Omega^-)} \\ &+ (\Pi_{\omega_1}^{\omega_2} g - g, \Pi_{\omega_1}^{\omega_2} g - \Pi_W g)_{L_2(\Omega^0)} \\ &= (\underbrace{\omega_2 - g}_{\leq 0}, \underbrace{\omega_2 - \Pi_W g}_{\geq 0,\, \Pi_W g \in W})_{L_2(\Omega^+)} + (\underbrace{\omega_1 - g}_{\geq 0}, \underbrace{\omega_1 - \Pi_W g}_{\leq 0,\, \Pi_W g \in W})_{L_2(\Omega^-)} \\ &+ \underbrace{(g - g, g - \Pi_W g)_{L_2(\Omega^0)}}_{=0} \\ &\leq 0\end{aligned}$$

Hence

$$\Pi_W g = \Pi_{\omega_1}^{\omega_2} g, \tag{4.1.30}$$

because the projection on W is unique, compare Theorem 2.1.7.
Having settled the case $\mathbb{L} = L_2(\Omega)$, we can now turn to the general setting $\mathbb{L} \subset L_2(\Omega)$: Here, we investigate

$$\begin{aligned}\left\|\Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)\right\|^2 &= (\Pi_W(g) - g, \Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)) \\ &+ (g - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g), \Pi_W(g) - \Pi_{\omega_1}^{\omega_2} P_{\mathbb{L}}(g)).\end{aligned} \tag{4.1.31}$$

Due to (4.1.27) we have $\Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g) \in W$. Thus, (4.1.28) yields for the first term on the right above:

$$(\Pi_W(g) - g, \Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)) \leq 0.$$

Hence, continuing our estimates in (4.1.31), we obtain

$$\begin{aligned}\left\|\Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)\right\|^2 &\leq (g - P_{\mathbb{L}}(g), \Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)) \\ &+ (P_{\mathbb{L}}(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g), \Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)).\end{aligned}$$

For the first term on the right in the inequality above we use the definition of the orthogonal projection and once again (4.1.27) to deduce

$$(g - P_{\mathbb{L}}(g), \underbrace{\Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)}_{=0, \Pi_W g, \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g) \in \mathbb{L}}) = 0$$

For the second term we take advantage of (4.1.30) for $g = P_{\mathbb{L}} g$ to obtain

$$(P_{\mathbb{L}}(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g), \Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)) \leq 0,$$

because $\omega_1 \leq \Pi_W g \leq \omega_2$ a.e. by definition of W.
Thus, all in all

$$\left\|\Pi_W(g) - \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g)\right\|^2 \leq 0$$

and hence

$$\Pi_W g = \Pi_{\omega_1}^{\omega_2}(P_{\mathbb{L}} g).$$

This completes the proof. □

This result can now be transferred to the optimality condition for $\bar{U}_k^\varepsilon$ in both the variational and full discretisation setting. This is the subject of the next theorem:

Theorem 4.1.16. *For the variational discretisation setting, i.e.* $\mathbb{U}_k = \mathbb{U}$*, in the notation of* (4.1.7)*, the following relation is valid:*

$$\bar{U}_k^\varepsilon(x) = \min(\max(-\frac{1}{\nu}\bar{P}_k^\varepsilon(x), a), b) = (\Pi_a^b(-\frac{1}{\nu}\bar{P}_k^\varepsilon))(x) = (\Pi(\bar{P}_k^\varepsilon))(x) \;\; f.a.a. \;\; x \in \Omega \quad (4.1.32)$$

For the full discretisation setting we have with the L_2*-orthogonal projection on* $\mathbb{U}_k$*,* $P_{\mathbb{U}_k}$

$$\begin{aligned} \bar{U}_k^\varepsilon(x) &= \min(\max(P_{\mathbb{U}_k}(-\frac{1}{\nu}\bar{P}_k^\varepsilon)(x), a), b) \\ &= (\Pi_a^b(P_{\mathbb{U}_k}(-\frac{1}{\nu}(\bar{P}_k^\varepsilon))(x) =: (\Pi_k(\bar{P}_k^\varepsilon))(x) \;\; f.a.a \;\; x \in \Omega \end{aligned} \quad (4.1.33)$$

holds.

Proof. Let us recall the optimality condition for $\bar{U}_k^\varepsilon$ in the following slightly reformulated way:

$$(\bar{U}_k^\varepsilon + \frac{1}{\nu}\bar{P}_k^\varepsilon, U - \bar{U}_k^\varepsilon) \geq 0 \;\; \forall U \in \mathcal{U}_k.$$

In the variational discretisation setting we can thus immediately apply Lemma 4.1.15 with $W = \mathcal{U}_k$ and $\omega_1 = a$, $\omega_2 = b$ to deduce:

$$\bar{U}_k^\varepsilon = \Pi_a^b(-\frac{1}{\nu}\bar{P}_k^\varepsilon).$$

This gives (4.1.32).
To derive (4.1.33), we again intend to apply Lemma 4.1.15, however, we have to verify (4.1.27) for $\mathbb{U}_k$ first. Naturally, though, it is clear that the pointwise cut-off (4.1.26) of a piecewise constant function is still a piecewise constant function. Hence (4.1.27) holds. This in turn

allows us to harness the results of Lemma 4.1.15 to gain

$$\bar{U}_k^\varepsilon = \Pi_a^b(P_{\mathbb{U}_k}(-\frac{1}{\nu}(\bar{P}_k^\varepsilon)))$$

This completes the proof. □

Let us now finally turn to constructing an estimator for

$$\left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|^2.$$

4.2 Derivation of the Estimator

We will not be able to derive an estimator for $\left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|^2$ all at once in one step. For better readability, we will thus shortly list the necessary steps to achieve this aim:
Recalling the definition of $\Pi(\cdot)$, (4.1.7), i.e.

$$(\Pi(v))(x) := \min(\max(v(x), a), b) \;\; v \in L_2(\Omega),$$

(4.1.32), we can lay out the broad strategy to derive the desired estimator:

1. Estimate for arbitrary $P \in \overset{\circ}{H}{}^1(\Omega)$
$$\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2 \tag{4.2.1}$$

2. Estimate the resulting terms involving the multipliers $\bar{\theta}^{\varepsilon^N}, \bar{\theta}_k^\varepsilon$

3. Combining both, derive an estimate for the variational discretisation approach $\bar{U}_k^\varepsilon = \Pi(\bar{P}_k^\varepsilon)$ by inserting $P = \bar{P}_k^\varepsilon$ in (4.2.1)

4. Deduce estimate for the full discretisation by splitting the difference and using (4.2.1):
$$\left\|\bar{U}_k^\varepsilon - \bar{u}^{\varepsilon^N}\right\|^2 \leq 2\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon\right\|^2 + 2\left\|\Pi(\bar{P}_k^\varepsilon) - \bar{u}^{\varepsilon^N}\right\|^2$$
The first term on the left can be evaluated exactly, compare [49], Remark 4.3, the second one can be dealt with as in the variational discretisation setting.

As we see, (4.2.1) is the starting point for our analyses. This 'basic' estimate is therefore our first goal:

4.2.1 The Basic Estimator

Before, we turn to the actual estimator, let us first recall the well-known Young's inequality:

$$ab \leq \frac{1}{2\delta}a^2 + \frac{\delta}{2}b^2,\ a, b, \delta > 0 \tag{4.2.2}$$

On several occasions, we will also use the following application of Young's inequality. The proof is trivial.

Lemma 4.2.1. *Let a_i, b_i, $i = 1, ...m$ be non-negative real numbers. Then*

$$(\sum_{i=1}^{m} a_i b_i)^2 \leq m \sum_{i=1}^{m} a_i^2 b_i^2$$

We are now in a position to prove the first crucial estimate

Lemma 4.2.2. *Let $(P, Y) \in \overset{\circ}{H}{}^1(\Omega) \times \overset{\circ}{H}{}^1(\Omega)$ be arbitrary. Furthermore, we define $\hat{y} := S\Pi(P), \hat{p} := S^*(Y - y_d - \bar{\theta}_k^\varepsilon)$ and $U := \Pi(P)$. Then, with a fixed $N \geq 1$, we have*

$$\begin{aligned} \left\| \Pi(P) - \bar{u}^{\varepsilon^N} \right\|^2 \leq \frac{1}{2\nu} \|\hat{y} - Y\|^2 + \frac{1}{\nu^2} \|\hat{p} - P\|^2 \\ + \frac{2}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - \hat{y}) + \frac{2}{\nu}(\bar{\theta}_k^\varepsilon, \hat{y} - \bar{y}^{\varepsilon^N}), \end{aligned} \tag{4.2.3}$$

Proof. We start by remarking that for the projection on the closed and convex set $\mathcal{U}$ of $-\frac{1}{\nu}$, $\Pi_{\mathcal{U}}(-\frac{1}{\nu}P)$ we have thanks to Lemma 4.1.15

$$(\Pi_{\mathcal{U}}(-\frac{1}{\nu}P))(x) = \min(\max(-\frac{1}{\nu}\bar{P}_k^\varepsilon(x), a), b) = (\Pi(P))(x) \text{ f.a.a. } x \in \Omega.$$

Using the variational inequality for the projection on a closed and convex set, compare Theorem 2.1.7 or (4.1.28), we then obtain:

$$(P + \nu\Pi(P), u - \Pi(P)) \geq 0 \quad \forall u \in \mathcal{U}. \tag{4.2.4}$$

In particular, we thus have

$$(P + \nu\Pi(P), \Pi(P) - \bar{u}^{\varepsilon^N}) \leq 0,$$

because $\bar{u}^{\varepsilon^N}$. Similarly, employing the optimality condition for $\bar{u}^{\varepsilon^N}$ in (4.1.6) into which we insert $u = \Pi(P) \in \mathcal{U}$, we obtain

$$-(\bar{p}^{\varepsilon^N} + \nu\bar{u}^{\varepsilon^N}, \Pi(P) - \bar{u}^{\varepsilon^N}) \leq 0$$

From these observations we can infer that:

$$\begin{aligned}\nu \left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2 &= \underbrace{(P + \nu\Pi(P), \Pi(P) - \bar{u}^{\varepsilon^N})_{L_2(\Omega)}}_{\leq 0} \\ &\underbrace{-(\bar{p}^{\varepsilon^N} + \nu\bar{u}^{\varepsilon^N}, \Pi(P) - \bar{u}^{\varepsilon^N})}_{\leq 0} + (\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N}) \\ &\leq (\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N})\end{aligned}$$

Inserting $\hat{p}$, we can split the last term in the following fashion:

$$(\bar{p}^{\varepsilon^N} - P, \Pi(P) - \bar{u}^{\varepsilon^N}) = (\bar{p}^{\varepsilon^N} - \hat{p}, \Pi(P) - \bar{u}^{\varepsilon^N}) + (\hat{p} - P, \Pi(P) - \bar{u}^{\varepsilon^N}) \tag{4.2.5}$$

For the second term on the right hand side we use Cauchy-Schwarz and then Young's inequality (4.2.2) with $\delta = \nu$ to obtain:

$$(\hat{p} - P, \Pi(P) - \bar{u}^{\varepsilon^N}) \leq \frac{1}{2\nu}\left\|\hat{p} - P\right\|^2 + \frac{\nu}{2}\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2. \tag{4.2.6}$$

For the first term on the right hand side in (4.2.5) we use $\bar{p}^{\varepsilon^N}, \hat{p} \in \mathring{H}^1(\Omega)$ for every fixed $\varepsilon > 0$, as well as $\hat{y} = S\Pi(P)$ and $\hat{p} = S^*(Y - y_d - \bar{\theta}_k^\varepsilon)$ to deduce:

$$\begin{aligned}(\bar{p}^{\varepsilon^N} - \hat{p}, \Pi(P) - \bar{u}^{\varepsilon^N}) &= (\nabla(\hat{y} - \bar{y}^{\varepsilon^N}), \nabla(\bar{p}^{\varepsilon^N} - \hat{p})) \\ &= (\bar{y}^{\varepsilon^N} - Y, \hat{y} - \bar{y}^{\varepsilon^N}) + (\bar{\theta}_k^\varepsilon - \bar{\theta}^{\varepsilon^N}, \hat{y} - \bar{y}^{\varepsilon^N})\end{aligned}$$

The second term already forms a part of (4.2.3), thus, at this stage, we content ourselves with estimating the first using Cauchy-Schwarz's and then Young's inequality (4.2.2) with $\delta = 2$:

$$\begin{aligned}(\bar{y}^{\varepsilon^N} - Y, \hat{y} - \bar{y}^{\varepsilon^N}) &= (\bar{y}^{\varepsilon^N} - \hat{y}, \hat{y} - \bar{y}^{\varepsilon^N}) + (\hat{y} - Y, \hat{y} - \bar{y}^{\varepsilon^N}) \\ &= -\left\|\hat{y} - \bar{y}^{\varepsilon^N}\right\|^2 + \left\|\hat{y} - Y\right\|\left\|\hat{y} - \bar{y}^{\varepsilon^N}\right\| \\ &\leq -\left\|\hat{y} - \bar{y}^{\varepsilon^N}\right\|^2 + \frac{1}{4}\left\|\hat{y} - Y\right\|^2 + \left\|\bar{y}^{\varepsilon^N} - \hat{y}\right\|^2 \\ &= \frac{1}{4}\left\|\hat{y} - Y\right\|^2\end{aligned} \tag{4.2.7}$$

Combining (4.2.6) and (4.2.7), we discern

$$\begin{aligned}\nu\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2 \leq{}& \frac{1}{2\nu}\left\|\hat{p} - P\right\|^2 + \frac{\nu}{2}\left\|\Pi(P) - \bar{u}^{\varepsilon^N}\right\|^2 + \frac{1}{4}\left\|\hat{y} - Y\right\|^2 \\ &+ (\bar{\theta}_k^\varepsilon - \bar{\theta}^{\varepsilon^N}, \hat{y} - \bar{y}^{\varepsilon^N})\end{aligned}$$

We can now subtract the term $\frac{\nu}{2}\left\|\Pi(P)-\bar{u}^{\varepsilon^N}\right\|^2$ in the inequality above. Then, we obtain:

$$\begin{aligned}\frac{\nu}{2}\left\|\Pi(P)-\bar{u}^{\varepsilon^N}\right\|^2 &\le \frac{1}{4}\|\hat{y}-Y\|^2+\frac{1}{2\nu}\|\hat{p}-P\|^2\\ &\quad+(\bar{\theta}^{\varepsilon^N},\bar{y}^{\varepsilon^N}-\hat{y})+(\bar{\theta}^{\varepsilon}_k,\hat{y}-\bar{y}^{\varepsilon^N}).\end{aligned}$$

Multiplying the inequality by $\frac{2}{\nu}$ then yields the desired result. □

A similar estimate was also derived in [46]. However, we will now extend this result by providing an upper bound for the terms

$$(\bar{\theta}^{\varepsilon},\bar{y}^{\varepsilon^N}-\hat{y}),\ (\bar{\theta}^{\varepsilon}_k,\hat{y}-\bar{y}^{\varepsilon^N}),\ \hat{y}=S\Pi(P), P\in\mathring{H}^1(\Omega) \tag{4.2.8}$$

This is the second step of our 'roadmap' to derive an estimator for the difference $\left\|\bar{u}^{\varepsilon^N}-\bar{U}^{\varepsilon}_k\right\|^2$ which we presented at the beginning of this section.
Before, though, we need a couple of auxiliary results. The first one provides a 'monotonicity' property of the continuous solution operator S, see e.g. [36], Theorem 8.1. Its proof is rooted in the maximum principle for elliptic differential operators:

Lemma 4.2.3. *Suppose that $q\in L_2(\Omega)$ with $q\ge 0$ a.e. in Ω is given. Then $Sq\ge 0$ a.e., too, and since $S=S^*$ by Theorem 4.1.5, we also have $S^*q\ge 0$.*

Secondly, we need a special projection:

Definition 4.2.4. *For an arbitrary function $z\in L_2(\Omega)$ we define the projection $P^{0+}_k : L_2(\Omega)\to\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))$ by*

$$P^{0+}_k z=\operatorname*{arg\,min}_{W\in\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega)),W\ge 0\ a.e.}\frac{1}{2}\|W-z\|^2, \tag{4.2.9}$$

i.e. it is the projection on the closed and convex subset of the space of piecewise constant functions

$$F:=\{W\in\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))\ :\ 0\le W<\infty\ a.e.\}.$$

P^{0+}_k perfectly fits into the setting of Lemma 4.1.15. The following lemma summarises the properties we will often use throughout our analyses:

Lemma 4.2.5. *For any $v\in L_2(\Omega)$, P^{0+}_k satisfies*

$$\left\|P^{0+}_k v\right\|_{L_2(\Omega)}\le\|v\|_{L_2(\Omega)}. \tag{4.2.10}$$

Besides, on every element $T \in \mathcal{T}_k$, P_k^{0+} is defined by

$$P_k^{0+}v|_T = \max(\frac{1}{|T|}\int\limits_T v(x)\,dT, 0) = \max((P_{\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))}v)(x), 0), \tag{4.2.11}$$

where $P_{\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))}$ denotes the L_2-orthogonal projection on the space $\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))$.

Proof. Let us first turn to the stability estimate (4.2.10): First of all, we recall the definition of P_k^{0+}, Definition 4.2.4, where we observed that P_k^{0+} is the projection on the closed and convex set

$$F := \{W \in \mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))\ :\ 0 \le W < \infty \text{ a.e.}\}\,.$$

Thanks to Theorem 2.1.7 it is uniquely defined - hence, the operator P_k^{0+} is well-defined. Due to $P_k^{0+}0 = 0$ Lipschitz continuity, compare again Theorem 2.1.7, now yields for all $v \in L_2(\Omega)$:

$$\left\|P_k^{0+}v\right\| = \left\|P_k^{0+}v - 0\right\| = \left\|P_k^{0+}v - P_k^{0+}0\right\| \le \|v - 0\| = \|v\|\,.$$

This gives (4.2.10).
Let us now prove (4.2.11): To do so, we simply apply Lemma 4.1.15. First of all, we are again projecting on the space of piecewise constant functions, hence, (4.1.27) holds. With $\omega_1 = 0$, $\omega_2 = \infty$ and $P_C := P_{\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))}$ as the $L_2(\Omega)$-orthogonal projection on the space of piecewise constant functions, Lemma 4.1.15 yields for any $v \in L_2(\Omega)$:

$$(P_k^{0+}v)(x) = \max((P_C v)(x), 0) = (\Pi_0^\infty(P_C v))(x) \ \text{ f.a.a } x \in \Omega. \tag{4.2.12}$$

Finally, in the proof of Theorem 2.4.4 we have already demonstrated that the $L_2(\Omega)$-orthogonal projection on the space of piece constant functions satisfies:

$$P_C v|_T := \frac{1}{|T|}\int\limits_T v(x)\,dT \ \ \forall T \in \mathcal{T}_k.$$

Combining this with formula (4.2.12) gives (4.2.11).

□

We can now return to the question of estimating (4.2.8), starting with the term

$$(\bar{\theta}_k^\varepsilon, \hat{y} - \bar{y}^{\varepsilon^N}),\ \hat{y} = S\Pi(P),\ P \in \overset{\circ}{H}{}^1(\Omega)$$

Lemma 4.2.6. *Let $(U,P,Y) \in L_2(\Omega) \times \overset{\circ}{H}{}^1(\Omega) \times \mathbb{Y}_k$ be arbitrary. Furthermore, let $\hat{y} = S\Pi(P)$*

and $\hat{p} = S^*(Y - y_d - \bar{\theta}_k^\varepsilon)$ *be given. Finally, let* R_k *denote the Ritz-projection on* $\mathbb{Y}_k$ *defined by*

$$(\nabla(R_k q - q), \nabla W) = 0 \ \ \forall W \in \mathbb{Y}_k, \ q \in \mathring{H}^1(\Omega). \tag{4.2.13}$$

Then

$$\begin{aligned}(\bar{\theta}_k^\varepsilon, \hat{y} - \bar{y}^{\varepsilon^N}) \leq\ & (\Pi(P), P - \hat{p}) - (Y - y_d, R_k(SU) - \hat{y}) \\ & + (U - \Pi(P), P) + (\bar{\theta}_k^\varepsilon, R_k(SU) - Y) + (\bar{\theta}_k^\varepsilon, Y - I_k y_c) \\ & + (\bar{\theta}_k^\varepsilon, I_k y_c - y_c) + (\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}).\end{aligned} \tag{4.2.14}$$

Let $4 < p' < \infty$ *in case* $d = 2$ *and* $p' = 6$ *in case* $d = 3$*. Then, for the last term in 4.2.14, we have the estimate*

$$(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^\varepsilon) \lesssim \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| + \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{3N/2}, c(p')s(\tau)\varepsilon^{2N(1-1/p')}), \tag{4.2.15}$$

where P_k^{0+} *is defined as in Definition 4.2.4.*
The constants in (4.2.15) *depend on* $a, b, \|S\|, \Omega, y_d, y_c$*.*

Proof. Let us first split the left hand side in (4.2.14) in the following way

$$\begin{aligned}(\bar{\theta}_k^\varepsilon, \hat{y} - \bar{y}^{\varepsilon^N}) &= (\bar{\theta}_k^\varepsilon, \hat{y} - y_c + y_c - \bar{y}^{\varepsilon^N}) \\ &= (\bar{\theta}_k^\varepsilon, \hat{y} - y_c) + (\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}).\end{aligned} \tag{4.2.16}$$

The proof is now divided into two parts: First, we will prove (4.2.15) for the second term on the right in the equation above which already appears in (4.2.14). Secondly, we will derive the rest of the bound in (4.2.14) by investigating the first term on the right in (4.2.16):

1st part of the proof:
We turn our focus now towards the second term on the right hand side above, which already appears on the right as the last term in (4.2.14): To prove its additional property (4.2.15), we first split it in the following way:

$$(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) = (\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) + (P_k^{0+}\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}). \tag{4.2.17}$$

Let us first tackle the second term on the right in the equation above. Since $(P_k^{0+}\bar{\theta}_k^\varepsilon)(x) \geq 0$ by construction, we obtain

$$(P_k^{0+}\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) \leq (P_k^{0+}\bar{\theta}_k^\varepsilon, (y_c - \bar{y}^{\varepsilon^N})^+) = (P_k^{0+}\bar{\theta}_k^\varepsilon, -(\bar{y}^{\varepsilon^N} - y_c)^-)$$

The penalty structure (4.1.8) then yields:

$$(P_k^{0+}\bar{\theta}_k^\varepsilon, -(\bar{y}^{\varepsilon^N} - y_c)^-) = \varepsilon^N (P_k^{0+}\bar{\theta}_k^\varepsilon, \bar{v}^{\varepsilon^N}).$$

Corollary 4.1.9 presently allows us to conclude that

$$\begin{aligned}(P_k^{0+}\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) &\leq \varepsilon^N (P_k^{0+}\bar{\theta}_k^\varepsilon, \bar{v}^{\varepsilon^N}) \\ &\leq \varepsilon^N \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \left\|\bar{v}^{\varepsilon^N}\right\| \\ &\lesssim \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{3N/2}, c(p')s(\tau)\varepsilon^{2(N-1/p')}).\end{aligned}$$

Going back to (4.2.17), we have thus gained the bound

$$(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) \lesssim (\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) + \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{3N/2}, c(p')s(\tau)\varepsilon^{2(N-1/p')}).$$

Employing Cauchy-Schwarz's inequality for the first term on the right in the inequality above, we can pursue our estimates to obtain

$$(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) \lesssim \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \left\|y_c - \bar{y}^{\varepsilon^N}\right\| + \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{3N/2}, c(p')s(\tau)\varepsilon^{2(N-1/p')}). \tag{4.2.18}$$

Now, observe that thanks to continuity of S and the uniform bound on $\bar{u}^{\varepsilon^N} \in \mathcal{U}$, (4.1.1), we have

$$\left\|y_c - \bar{y}^{\varepsilon^N}\right\| \leq \|S\|_{\mathcal{L}(L_2(\Omega), L_2(\Omega))} \left\|\bar{u}^{\varepsilon^N}\right\| \|y_c\| \lesssim 1.$$

Thus, recalling (4.2.18), we can then finally derive (4.2.15):

$$(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N}) \lesssim \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| + \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{3N/2}, c(p')s(\tau)\varepsilon^{2(N-1/p')}).$$

This completes the first part of the proof, let us now tackle the second.

2nd part of the proof:
Let us have a look at the following term, its significance will become evident later on:

$$(\nabla(R_k(\hat{p}) - \hat{p}), \nabla(Y - \hat{y})). \tag{4.2.19}$$

Using the Ritz-projection R_k, (4.2.13), $\hat{y} = S\Pi(P)$ and $\hat{p} = S^*(Y - y_d - \bar{\theta}_k^\varepsilon)$ we can conclude that:

$$
\begin{aligned}
(\nabla(R_k(\hat{p})-\hat{p}),\nabla(Y-\hat{y})) &= (\nabla(R_k(\hat{p})-\hat{p}),\nabla(Y-R_k(SU)+R_k(SU)-\hat{y}))\\
&= \underbrace{(\nabla(R_k(\hat{p})-\hat{p}),\nabla(Y-R_k(SU)))}_{=0,Y,R_k(SU)\in\mathbb{Y}_k}\\
&+ (\nabla(R_k(\hat{p})-\hat{p}),\nabla(R_k(SU)-\hat{y}))\\
&= (\nabla(R_k(\hat{p})-\hat{p}),\nabla(R_k(SU)-\hat{y}))\\
&= (\nabla R_k(\hat{p}),\nabla(R_k(SU)-\hat{y}))\\
&- (\nabla\hat{p},\nabla(R_k(SU)-\hat{y}))\\
&= (\nabla R_k(\hat{p}),\nabla(SU-\hat{y})) - (\nabla\hat{p},\nabla(R_k(SU)-\hat{y}))\\
&= (R_k(\hat{p}),U) - (R_k(\hat{p}),\Pi(P)) - (\nabla\hat{p},\nabla(R_k(SU)-\hat{y}))\\
&= (R_k(\hat{p}),U) - (R_k(\hat{p}),\Pi(P))\\
&- (Y-y_d-\bar{\theta}_k^\varepsilon, R_k(SU)-\hat{y})
\end{aligned}
\tag{4.2.20}
$$

Now, let us look at (4.2.19) from another point of view. Harnessing once again the Ritz projection R_k, (4.2.13), and $\hat{y}=S\Pi(P)$ we immediately arrive at

$$
(\nabla(R_k(\hat{p})-\hat{p}),\nabla(Y-\hat{y})) = -(\nabla(R_k(\hat{p})-\hat{p}),\nabla\hat{y}) = -(\Pi(P),R_k(\hat{p})-\hat{p}).
$$

Rearranging (4.2.20), we derive

$$
(\bar{\theta}_k^\varepsilon,\hat{y}-R_k(SU)) = (\Pi(P),R_k(\hat{p})-\hat{p}) - (Y-y_d,R_k(SU)-\hat{y}) + (U-\Pi(P),R_k(\hat{p}))
$$

and thus, combining this with our previous deductions, we obtain

$$
\begin{aligned}
(\bar{\theta}_k^\varepsilon,\hat{y}-y_c) &= (\bar{\theta}_k^\varepsilon,\hat{y}-R_k(SU)+R_k(SU)-Y+Y-I_ky_c+I_ky_c-y_c)\\
&= (\bar{\theta}_k^\varepsilon,\hat{y}-R_k(SU)) + (\bar{\theta}_k^\varepsilon,R_k(SU)-Y) + (\bar{\theta}_k^\varepsilon,Y-I_ky_c) + (\bar{\theta}_k^\varepsilon,I_ky_c-y_c)\\
&= (\Pi(P),R_k(\hat{p})-\hat{p}) - (Y-y_d,R_k(SU)-\hat{y}) + (U-\Pi(P),R_k(\hat{p}))\\
&+ (\bar{\theta}_k^\varepsilon,R_k(SU)-Y) + (\bar{\theta}_k^\varepsilon,Y-I_ky_c) + (\bar{\theta}_k^\varepsilon,I_ky_c-y_c).
\end{aligned}
$$

Recalling (4.2.16), we are now in possession of the desired result. □

Remark 4.2.7. *We remark that the computation of P_k^{0+} requires the evaluation of element mean values, see* (4.2.11)*. Computationally, this requires limited effort and is thus acceptable as a part of an a posteriori error estimator.*

Remark 4.2.8. *Instead of using the space of piecewise constant functions* $\mathbf{FES}(\mathcal{T}_k,\mathbb{P}_0,L_2(\Omega))$ *as the image space for the projector P_k^{0+} in Definition 4.2.4, we could also project on the subset of a.e. non-negative functions in a space of discontinuous piecewise linear finite elements, such* *as*

$$\mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, L_2(\Omega)).$$

Computationally, this would require inverting a $(d+1) \times (d+1)$ *matrix on every element* $T \in \mathcal{T}_k$*. Again, this is a computationally justifiable effort.*

We will now tackle the second term in (4.2.8)

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - \hat{y})$$

Before, though, we again need some auxiliary results beginning with the definition of the *harmonic extension*:

Definition 4.2.9 (Harmonic Extension)**.** *Let* $z \in H^{1/2}(\partial\Omega)$ *(compare Theorem 2.1.31) be given. Then the* harmonic extension $Hz \in H^1(\Omega), \nabla Hz \in H(\mathrm{div}, \Omega)$ *of* z *is the unique solution to the boundary value problem:*

$$\begin{aligned} -\Delta Hz &= 0 \;\; \textit{in } \Omega \\ Hz &= z \;\; \textit{on } \partial\Omega \end{aligned}$$

Let us shortly explain that the harmonic extension is well-defined. Surjectivity of the trace mapping, compare Theorem 2.1.31, immediately allows us to conclude that there exists a function $\phi \in H^1(\Omega)$ such that $\phi = z$ in the sense of traces on $\partial\Omega$. The unique solvability of the boundary value problem in Definition 4.2.9 then is a standard result, which can e.g. be found in [36], Theorem 8.3.

Using the harmonic extension, we gain the following lemma, which is a consequence of Theorem 8.1 in [36] and the fact that $y_c|_{\partial\Omega} < 0$ by assumption.

Lemma 4.2.10. *Define* $\iota \in H^{1/2}(\partial\Omega)$ *by*

$$\iota(x) := y_c(x), \; x \in \Gamma$$

Then $H\iota \leq 0$ *a.e.,* $y_c - H\iota \in \mathring{H}^1(\Omega)$ *and* $\nabla y_c, \nabla\iota \in H(\mathrm{div}, \Omega)$ *and* $\Delta H\iota = 0$ *a.e. in* Ω*.*

We can now turn to estimating the continuous multiplier term:

Lemma 4.2.11. *Let* $(Y, P) \in \mathring{H}^1(\Omega) \times \mathring{H}^1(\Omega)$ *be arbitrary,* $\hat{y} = S\Pi(P)$ *and* $4 < p' < \infty$ *in* *case $d = 2$ and $p' = 6$ in case $d = 3$. Then the following estimate is valid*

$$\begin{aligned}(\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - \hat{y}) \lesssim \min \Big\{ &\|\bar{p}_s^{\varepsilon^N}\| \left\|(\Pi(P) + \Delta y_c)^-\right\|, \\ &+ c(p')s(\tau)\varepsilon^{-3N/p'}(|\hat{y} - Y|_{H^1(\Omega)} + |(Y - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)})\Big\} \\ &- \frac{1}{\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2.\end{aligned} \tag{4.2.21}$$

Proof. First of all, we take $H\iota$ as in Lemma 4.2.10 and split the critical term in the following way.

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - \hat{y}) = (\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - (y_c - H\iota)) + (\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y}).$$

Investigating the first term, we take advantage of the complimentary slackness condition (4.1.9) and the sign on the harmonic extension $H\iota$, see Lemma 4.2.10, to deduce

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^\varepsilon - y_c + H\iota) = -\frac{1}{\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2 + \underbrace{(\bar{\theta}^{\varepsilon^N}, H\iota)}_{\leq 0} \leq 0.$$

We still need to estimate the term

$$(\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y}),$$

though:
Using the singular part of the adjoint state $\bar{p}_s^{\varepsilon^N} = -S^*\bar{\theta}^{\varepsilon^N}$, compare Theorem 4.1.6, $\nabla\hat{y}, \nabla y_c, \nabla H\iota \in H(\mathrm{div}, \Omega)$, $y_c - H\iota \in \mathring{H}^1(\Omega)$ and Green's formula, we deduce:

$$\begin{aligned}(\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y}) &= (-\nabla\bar{p}_s^{\varepsilon^N}, \nabla(y_c - H\iota - \hat{y})) \\ &= (\bar{p}_s^{\varepsilon^N}, \Delta y_c - \Delta H\iota - \Delta\hat{y})..\end{aligned}$$

Now, observe that $\Delta H\iota = 0$ and $-\Delta\hat{y} = \Pi(P)$ a.e in Ω. This allows us to conclude:

$$(\bar{\theta}^{\varepsilon^N}, y_c - H\iota - \hat{y}) = (\bar{p}_s^{\varepsilon^N}, \Delta y_c - \Delta H\iota + \Delta\hat{y}) = (\bar{p}^{\varepsilon^N}, \Delta y_c + \Pi(P)).$$

Thanks to $-\bar{\theta}^{\varepsilon^N} \leq 0$ a.e. and Lemma 4.2.3 we know that $\bar{p}^{\varepsilon^N} = -S^*\bar{\theta}^{\varepsilon^N} \leq 0$ a.e. Thus, we discern:

$$\begin{aligned}(\underbrace{\bar{p}_s^{\varepsilon^N}}_{\leq 0}, \Delta y_c + \Pi(P)) &\leq (\bar{p}^{\varepsilon^N}, (\Delta y_c + \Pi(P))^-) \\ &\leq \left\|\bar{p}_s^{\varepsilon^N}\right\| \left\|(\Delta y_c + \Pi(P))^-\right\|.\end{aligned} \tag{4.2.22}$$

All in all, we have proven the bound:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y}) \leq \left\| \bar{p}_s^{\varepsilon^N} \right\| \left\| (\Delta y_c + \Pi(P))^- \right\| - \frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 =: min_1. \tag{4.2.23}$$

This gives the first argument in the min operation in (4.2.21).
To derive the other bound given by the second argument in the min operation in (4.2.21), we return to the start and split the critical term in the following way:

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y}) = (\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - y_c) + (\bar{\theta}^{\varepsilon^N}, y_c - \hat{y}) \tag{4.2.24}$$

For the first term, we once again take advantage of the complimentary slackness condition (4.1.9):

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - y_c) = -\frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 .$$

For the other term on the right in (4.2.24) we discern that

$$(\bar{\theta}^{\varepsilon^N}, y_c - \hat{y}) = (\bar{\theta}^{\varepsilon^N}, y_c - I_k y_c) + (\bar{\theta}^{\varepsilon^N}, I_k y_c - Y) + (\bar{\theta}^{\varepsilon^N}, Y - \hat{y})$$

Utilising the bound for $\bar{\theta}^{\varepsilon^N}$ derived in Lemma 4.1.7, $\bar{\theta}^{\varepsilon^N} \geq 0$ and the embedding $\overset{\circ}{H}{}^1(\Omega) \hookrightarrow L_p(\Omega)$, see Theorem 2.1.35 and the Poincaré-Friedrich inequality, Theorem 2.1.33, we obtain for $1 \leq p \leq 2$ and $\frac{1}{p} + \frac{1}{p'} = 1$:

$$\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, y_c - I_k y_c) &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \| y_c - I_k y_c \|_{L_{p'}(\Omega)} \lesssim c(p') s(\tau) \varepsilon^{-3N/p'} \| y_c - I_k y_c \|_{H^1(\Omega)} \\
(\bar{\theta}^{\varepsilon^N}, I_k y_c - Y) &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \left\| (Y - I_k y_c)^- \right\|_{L_{p'}(\Omega)} \lesssim c(p') s(\tau) \varepsilon^{-3N/p'} |(Y - I_k y_c)^-|_{H^1(\Omega)} \\
(\bar{\theta}^{\varepsilon^N}, Y - \hat{y}) &\leq \left\| \bar{\theta}^{\varepsilon^N} \right\|_{L_p(\Omega)} \| Y - \hat{y} \|_{L_{p'}(\Omega)} \lesssim c(p') s(\tau) \varepsilon^{-3N/p'} |Y - \hat{y}|_{H^1(\Omega)},
\end{aligned}$$

where the hidden constants depend on $a, b, \Omega, \nu, y_d, \|S\|$.
Reviewing the previous estimates, we have proven the bound given by the second argument of the min argument in (4.2.21):

$$\begin{aligned}
(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y}) \lesssim\ & c(p') s(\tau) \varepsilon^{-3N/p'} \big(\| y_c - I_k y_c \|_{H^1(\Omega)} + |(Y - I_k y_c)^-|_{H^1(\Omega)} + |Y - \hat{y}|_{H^1(\Omega)} \big) \\
& - \frac{1}{\varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 \\
& =: min_2
\end{aligned}$$

Combining this with (4.2.23), we obtain

$$(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon} - \hat{y}) \lesssim \min(min_1, min_2)$$

which is the desired result. □

Let us shortly recapitulate our 'roadmap' for deriving an estimator which we presented at the beginning of this section:

1. Estimate for arbitrary $P \in \mathring{H}^1(\Omega)$
$$\left\| \Pi(P) - \bar{u}^{\varepsilon^N} \right\|^2$$
2. Estimate the resulting terms involving the multipliers $\bar{\theta}^{\varepsilon^N}, \bar{\theta}^\varepsilon_k$
3. Combining both, derive an estimate for the variational discretisation approach $\bar{U}^\varepsilon_k = \Pi(\bar{P}^\varepsilon_k)$ by inserting $P = \bar{P}^\varepsilon_k$ in (4.2.1)
4. Deduce estimate for the full discretisation by splitting the difference and using (4.2.1):
$$\left\| \bar{U}^\varepsilon_k - \bar{u}^{\varepsilon^N} \right\|^2 \leq 2 \left\| \Pi(\bar{P}^\varepsilon_k) - \bar{U}^\varepsilon_k \right\|^2 + 2 \left\| \Pi(\bar{P}^\varepsilon_k) - \bar{u}^{\varepsilon^N} \right\|^2$$

The first and second step we have completed with the Lemmata 4.2.2, 4.2.6 and 4.2.11. What remains to be done are steps 3 and 4, to which we now turn presently:

4.2.2 Estimators for the Semi and Fully Discrete Problem

In the case of the variational discretisation technique, we have thanks to Theorem 4.1.16, in particular formula (4.1.32), $\bar{U}^\varepsilon_k = \Pi(\bar{P}^\varepsilon_k)$. As described in our 'roadmap', in this case, Lemmata 4.2.2, Lemma 4.2.6 and Lemma 4.2.11 already provide us with a full estimator:

Theorem 4.2.12. *Let $N \geq 1$ be fixed and $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. In the case of variational discretisation, i.e $\Pi(\bar{P}^\varepsilon_k) = \bar{U}^\varepsilon_k$ we gain the error bound*

$$\left\| \bar{U}^\varepsilon_k - \bar{u} \right\|^2 + \frac{4}{\nu \varepsilon^N} \left\| \bar{v}^{\varepsilon^N} \right\|^2 \lesssim c(p') s(\tau) \varepsilon^{\gamma N} + \mathcal{E}^2_r(\bar{U}^\varepsilon_k, \bar{V}^\varepsilon_k) + \mathcal{E}_s(\bar{U}^\varepsilon_k, \bar{V}^\varepsilon_k) \tag{4.2.25}$$

with γ as in (4.1.22) and

$$\begin{aligned} \mathcal{E}^2_r(\bar{U}^\varepsilon_k, \bar{V}^\varepsilon_k) = {} & \frac{1}{\nu} \left\| (S - S_k) \bar{U}^\varepsilon_k \right\|^2 + \frac{2}{\nu^2} \left\| (S - S_k)^* (\bar{Y}^\varepsilon_k - y_d - \bar{\theta}^\varepsilon_k) \right\|^2 \\ & + \frac{4}{\nu} (\bar{\theta}^\varepsilon_k, \bar{Y}^\varepsilon_k - I_k y_c) + \frac{4}{\nu} (\bar{\theta}^\varepsilon_k, I_k y_c - y_c) \end{aligned} \tag{4.2.26}$$

and

$$\begin{aligned}\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = {} & \frac{4}{\nu}\left\|\bar{Y}_k^\varepsilon - y_d\right\| \left\|(S-S_k)\bar{U}_k^\varepsilon\right\| + \frac{4}{\nu}\left\|\bar{U}_k^\varepsilon\right\| \left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\| \\ & + \frac{4}{\nu}\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \\ & + \frac{4}{\nu}c(p')s(\tau)\left\|P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\| \min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2}) \\ & + \frac{4}{\nu}\min\left\{ \left\|\bar{p}_s^{\varepsilon^N}\right\| \left\|(\bar{U}_k^\varepsilon + \Delta y_c)^-\right\|, \right. \\ & c(p')s(\tau)\varepsilon^{-3N/p'}\Bigg(|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)} \\ & \left. + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\Bigg)\right\}, \end{aligned} \tag{4.2.27}$$

where P_k^{0+} is the projection defined in Definition 4.2.4.

Proof. First of all, we split the error $\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|^2$ with the help of the triangle and Young's inequality with $\delta = 1$, (4.2.2), in the following way

$$\left\|\bar{u} - \bar{U}_k^\varepsilon\right\|^2 \leq 2\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2 + 2\left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|^2. \tag{4.2.28}$$

For the first term on the right in the inequality above we can use Theorem 4.1.10 to deduce

$$\left\|\bar{u} - \bar{u}^{\varepsilon^N}\right\|^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N}.$$

with γ defined in (4.1.22).

For the second term on the right in (4.2.28) we employ Lemma 4.2.2 with $U = \bar{U}_k^\varepsilon = \Pi(\bar{P}_k^\varepsilon)$, $P = \bar{P}_k^\varepsilon$ and $Y = \bar{Y}_k^\varepsilon$ to gain

$$\begin{aligned} 2\left\|\bar{u}^{\varepsilon^N} - \bar{U}_k^\varepsilon\right\|^2 \leq {} & \frac{1}{\nu}\left\|S\bar{U}_k^\varepsilon - \bar{Y}_k^\varepsilon\right\|^2 + \frac{2}{\nu^2}\left\|S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon) - \bar{P}_k^\varepsilon\right\|^2 \\ & + \frac{4}{\nu}(\bar{\theta}_k^\varepsilon, S\bar{U}_k^\varepsilon - \bar{y}^{\varepsilon^N}) + \frac{4}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\bar{U}_k^\varepsilon). \end{aligned} \tag{4.2.29}$$

Investigating the second to last term on the right, we harness Lemma 4.2.6 and $R_k(S\bar{U}_k^\varepsilon) = \bar{Y}_k^\varepsilon$

thanks to Lemma 4.1.12 to conclude with the additional help of (4.2.15):

$$\begin{aligned}(\bar\theta_k^\varepsilon, S\bar U_k^\varepsilon - \bar y^{\varepsilon^N}) &\le \underbrace{(\bar U_k^\varepsilon, \bar P_k^\varepsilon - S^*(\bar Y_k^\varepsilon - y_d - \bar\theta_k^\varepsilon))}_{Cauchy-Schwarz} - \underbrace{(\bar Y_k^\varepsilon - y_d, \bar Y_k^\varepsilon - \hat y)}_{Cauchy-Schwarz}\\ &+ (\bar\theta_k^\varepsilon, \bar Y_k^\varepsilon - I_k y_c) + (\bar\theta_k^\varepsilon, I_k y_c - y_c) + \underbrace{(\bar\theta_k^\varepsilon, y_c - \bar y^{\varepsilon^N})}_{(4.2.15)}\\ &\lesssim \|\bar U_k^\varepsilon\| \, \|\bar P_k^\varepsilon - S^*(\bar Y_k^\varepsilon - y_d - \bar\theta_k^\varepsilon)\| + \|\bar Y_k^\varepsilon - y_d\| \, \|\bar Y_k^\varepsilon - \hat y\|\\ &+ (\bar\theta_k^\varepsilon, \bar Y_k^\varepsilon - I_k y_c) + \|\bar\theta_k^\varepsilon - P_k^{0+}(\bar\theta_k^\varepsilon)\|\\ &+ c(p')s(\tau)\,\|P_k^{0+}(\bar\theta_k^\varepsilon)\| \min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2})\end{aligned} \tag{4.2.30}$$

Combining (4.2.29) and (4.2.30) and we obtain

$$\begin{aligned}\left\|\bar U_k^\varepsilon - \bar u^{\varepsilon^N}\right\|^2 &\le \mathcal{E}_r^2(\bar U_k^\varepsilon, \bar V_k^\varepsilon) + \frac{4}{\nu}\|\bar U_k^\varepsilon\| \, \|\bar P_k^\varepsilon - S^*(\bar Y_k^\varepsilon - y_d - \bar\theta_k^\varepsilon)\|\\ &+ \frac{4}{\nu}\|\bar Y_k^\varepsilon - y_d\| \, \|\bar Y_k^\varepsilon - \hat y\|\\ &+ \frac{4}{\nu}\|\bar\theta_k^\varepsilon - P_k^{0+}(\bar\theta_k^\varepsilon)\|\\ &+ \frac{4}{\nu}c(p')s(\tau)\,\|P_k^{0+}(\bar\theta_k^\varepsilon)\| \min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2})\\ &+ \frac{4}{\nu}(\bar\theta^{\varepsilon^N}, \bar y^{\varepsilon^N} - S\bar U_k^\varepsilon)\end{aligned} \tag{4.2.31}$$

The last term remains to be estimated.
Recalling Lemma 4.2.11 with $Y = \bar Y_k^\varepsilon$, $P = \bar P_k^\varepsilon$ and $\bar U_k^\varepsilon = \Pi(\bar P_k^\varepsilon)$, we deduce

$$\begin{aligned}\frac{4}{\nu}(\bar\theta^{\varepsilon^N}, \bar y^{\varepsilon^N} - S\bar U_k^\varepsilon) \lesssim \frac{4}{\nu}\min\Bigg\{ &\|\bar p_s^{\varepsilon^N}\| \, \|(\bar U_k^\varepsilon + \Delta y_c)^-\|,\\ &+ c(p')s(\tau)\varepsilon^{-3N/p'}\Bigg(|S\bar U_k^\varepsilon - \bar Y_k^\varepsilon|_{H^1(\Omega)} + |(\bar Y_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\\ &+ \|y_c - I_k y_c\|_{H^1(\Omega)}\Bigg)\Bigg\}\\ &- \frac{4}{\nu\varepsilon^N}\left\|\bar v^{\varepsilon^N}\right\|^2\end{aligned}$$

The term $\frac{4}{\nu\varepsilon^N}\left\|\bar v^{\varepsilon^N}\right\|^2$ can be shifted to the left in (4.2.29) to complete the left hand side in (4.2.25). Presently, inserting the estimate above in (4.2.31) and recalling the definition of $\mathcal{E}_s(\bar U_k^\varepsilon, \bar V_k^\varepsilon)$, (4.2.27), we deduce the the right hand side in (4.2.25), i.e.:

$$\|\bar U_k^\varepsilon - \bar u\|^2 + \frac{4}{\nu\varepsilon^N}\left\|\bar v^{\varepsilon^N}\right\|^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar U_k^\varepsilon, \bar V_k^\varepsilon) + \mathcal{E}_s(\bar U_k^\varepsilon, \bar V_k^\varepsilon)$$

□

As the final step of our 'roadmap', we now naturally want to extend the results of Theorem 4.2.12 to the setting of the **full discretisation approach** where the controls are discretised by piecewise constant functions. In this setting, in general $\Pi(\bar{P}_k^\varepsilon) \neq \bar{U}_k^\varepsilon$, hence, we encounter an additional error which has to be taken into account leading to estimators which are slightly different to the estimators $\mathcal{E}_r$ and $\mathcal{E}_s$ in Theorem 4.2.12.
This is the subject of the next theorem:

Theorem 4.2.13. *Let $\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon$ be the solution to (DMP_k^ε) and $N \geq 1$ be fixed. Besides, let $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$. Then the following estimates hold:*

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 + \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \tag{4.2.32}$$

with

$$\begin{aligned}\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \frac{4}{\nu}(&\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2 + (2 + \frac{4\left\|S\right\|^2}{\nu})\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2 \\ &+ \frac{4}{\nu^2}\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2 \\ &+ \frac{8}{\nu}(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) + \frac{8}{\nu}(\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c) \\ &+ \frac{8}{\nu}(\bar{\theta}_k^\varepsilon, I_k y_c - y_c),\end{aligned} \tag{4.2.33}$$

where $\|S\| = \|S\|_{\mathcal{L}(L_2(\Omega), L_2(\Omega))}$, and

$$\begin{aligned}\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = &\frac{8}{\nu}\left\|\bar{Y}_k^\varepsilon - y_d\right\|(\left\|(S - S_k)\bar{U}_k^\varepsilon\right\| + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|) \\ &+ \frac{8}{\nu}\left\|\Pi(\bar{P}_k^\varepsilon)\right\|\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\| \\ &+ \frac{8}{\nu}c(p')s(\tau)\left\|P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\|\min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2}) \\ &+ \frac{8}{\nu}\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \\ &+ \frac{8}{\nu}\min\left\{\left\|\bar{p}_s^{\varepsilon^N}\right\|\left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, c(p')s(\tau)\varepsilon^{-3N/p'}\big(|(S - S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)}\right. \\ &+ |S|\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\| + \left\|y_c - I_k y_c\right\|_{H^1(\Omega)} \\ &\left.+ |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\big)\right\}\end{aligned} \tag{4.2.34}$$

where $|S| = \|S\|_{\mathcal{L}(L_2(\Omega), \mathring{H}^1(\Omega))}$.

Proof. As in the proof of Theorem 4.2.12 we harness the results of Lemmas 4.2.2, 4.2.6 and 4.2.11 to prove the bound (4.2.32).

Utilising Lemma 4.2.1 and Theorem 4.1.10, we observe that

$$\begin{aligned}\left\|\bar U_k^\varepsilon - \bar u\right\|^2 &\leq 2\left\|\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon)\right\|^2 + 2\left\|\Pi(\bar P_k^\varepsilon) - \bar u\right\|^2\\ &\leq 2\left\|\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon)\right\|^2 + 4\left\|\Pi(\bar P_k^\varepsilon) - \bar u^{\varepsilon^N}\right\|^2 + 4\left\|\bar u - \bar u^{\varepsilon^N}\right\|^2\\ &\lesssim 2\left\|\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon)\right\|^2 + 4\left\|\Pi(\bar P_k^\varepsilon) - \bar u^{\varepsilon^N}\right\|^2 + 4c(p')s(\tau)\varepsilon^{\gamma N}.\end{aligned} \tag{4.2.35}$$

The first term $\left\|\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon)\right\|^2$ can be evaluated by numerical integration, compare again [49], Remark 4.3. Thus, the aim now is to control the term $\left\|\Pi(\bar P_k^\varepsilon) - \bar u^{\varepsilon^N}\right\|^2$.
Inserting $(\bar Y_k^\varepsilon, \bar P_k^\varepsilon)$ for (Y, P) in Lemma 4.2.2 with $Y = \bar Y_k^\varepsilon$ and $P = \bar P_k^\varepsilon$, we obtain

$$\begin{aligned}\left\|\Pi(\bar P_k^\varepsilon) - \bar u^{\varepsilon^N}\right\|^2 \leq \frac{1}{2\nu}\left\|S\Pi(\bar P_k^\varepsilon) - \bar Y_k^\varepsilon\right\|^2 + \frac{1}{\nu^2}\left\|S^*(\bar Y_k^\varepsilon - y_d - \bar\theta_k^\varepsilon) - \bar P_k^\varepsilon\right\|^2\\ + \frac{2}{\nu}(\bar\theta^{\varepsilon^N}, \bar y^{\varepsilon^N} - S\Pi(\bar P_k^\varepsilon)) + \frac{2}{\nu}(\bar\theta_k^\varepsilon, S\Pi(\bar P_k^\varepsilon) - \bar y^{\varepsilon^N})\end{aligned} \tag{4.2.36}$$

We will derive (4.2.33) first. To begin with, we discern with $\|S\| = \|S\|_{\mathcal{L}(L_2(\Omega), L_2(\Omega))}$ and $\bar Y_k^\varepsilon = S_k \bar U_k^\varepsilon$ that

$$\begin{aligned}\left\|S\Pi(\bar P_k^\varepsilon) - \bar Y_k^\varepsilon\right\|^2 &= \left\|S\Pi(\bar P_k^\varepsilon) - S\bar U_k^\varepsilon + S\bar U_k^\varepsilon - \bar Y_k^\varepsilon\right\|^2\\ &\leq 2\left\|S(\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon))\right\|^2 + 2\left\|(S - S_k)\bar U_k^\varepsilon\right\|^2\\ &\leq 2\left\|S\right\|^2\left\|\bar U_k^\varepsilon - \Pi(\bar P_k^\varepsilon)\right\|^2 + 2\left\|(S - S_k)\bar U_k^\varepsilon\right\|^2\end{aligned} \tag{4.2.37}$$

Let us now tackle the remaining terms in (4.2.36) beginning with $(\bar\theta_k^\varepsilon, S\Pi(\bar P_k^\varepsilon) - \bar y^{\varepsilon^N})$:
First of all we discern that $(\bar Y_k^\varepsilon = S_k \bar U_k^\varepsilon)$:

$$\left\|\bar Y_k^\varepsilon - S\Pi(\bar P_k^\varepsilon)\right\| \leq \|S\|\left\|\bar U_k^\varepsilon - \Pi(\bar U_k^\varepsilon)\right\| + \left\|(S - S_k)\bar U_k^\varepsilon\right\|. \tag{4.2.38}$$

Presently, to gain an estimate for $(\bar\theta_k^\varepsilon, S\Pi(\bar P_k^\varepsilon) - \bar y^{\varepsilon^N})$, we employ Lemma 4.2.6 with $U = \bar U_k^\varepsilon$,

$P = \bar{P}_k^\varepsilon$ and $Y = \bar{Y}_k^\varepsilon$; also recollect that $R_k(S\bar{U}_k^\varepsilon) = \bar{Y}_k^\varepsilon$:

$$\begin{aligned}
(\bar{\theta}_k^\varepsilon, S\Pi(\bar{P}_k^\varepsilon) - \bar{y}^{\varepsilon^N}) &\leq \underbrace{(\Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon - S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon))}_{Cauchy-Schwarz} - \underbrace{(\bar{Y}_k^\varepsilon - y_d, \bar{Y}_k^\varepsilon - S\Pi(\bar{P}_k^\varepsilon))}_{Cauchy-Schwarz+(4.2.38)} \\
&+ \underbrace{(\bar{\theta}_k^\varepsilon, y_c - \bar{y}^{\varepsilon^N})}_{(4.2.15)} + (\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) \\
&+ (\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c) + (\bar{\theta}_k^\varepsilon, I_k y_c - y_c) \\
&\lesssim \left\|\Pi(\bar{P}_k^\varepsilon)\right\| \left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\| \\
&+ \left\|\bar{Y}_k^\varepsilon - y_d\right\| (\left\|(S - S_k)\bar{U}_k^\varepsilon\right\| + \|S\| \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|) \\
&+ \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\| \\
&+ c(p')s(\tau) \left\|P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\| \min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2}) \\
&+ (\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) + (\bar{\theta}_k^\varepsilon, I_k y_c - y_c) + (\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)
\end{aligned} \tag{4.2.39}$$

This bound can now be used in (4.2.36): We set

$$\begin{aligned}
\mathcal{X} := &\left\|\Pi(\bar{P}_k^\varepsilon)\right\| \left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\| \\
&+ \left\|\bar{Y}_k^\varepsilon - y_d\right\| (\left\|(S - S_k)\bar{U}_k^\varepsilon\right\| + \|S\| \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|) \\
&+ \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\| + c(p')s(\tau) \left\|P_k^{0+}(\bar{\theta}_k^\varepsilon)\right\| \min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2})
\end{aligned} \tag{4.2.40}$$

and observe that $\mathcal{X}$ already contains many of the terms given in (4.2.34). Combining the estimate (4.2.39) with (4.2.35), (4.2.36) and (4.2.37), we obtain

$$\begin{aligned}
\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 \lesssim\, &c(p')s(\tau)\varepsilon^{\gamma N} + (2 + \frac{4\left\|S\right\|^2}{\nu}) \left\|\Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon\right\|^2 \\
&+ \frac{4}{\nu} \left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2 + \frac{4}{\nu^2} \left\|S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon) - \bar{P}_k^\varepsilon\right\|^2 \\
&+ \frac{8}{\nu}(\bar{\theta}_k^\varepsilon, I_k y_c - y_c) + \frac{8}{\nu}(\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c) + \frac{8}{\nu}(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) \\
&+ \frac{8}{\nu}\mathcal{A} + \frac{8}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon))
\end{aligned}$$

In short:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \frac{8}{\nu}\mathcal{A} + \frac{8}{\nu}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon)) \tag{4.2.41}$$

The last term remains to be estimated:

We employ Lemma 4.2.11 with $Y = \bar{Y}_k^\varepsilon$ and $P = \bar{P}_k^\varepsilon$ to obtain:

$$\begin{aligned}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon)) \lesssim \min\Big\{ & \|\bar{p}_s^{\varepsilon^N}\| \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, \\ & + c(p')s(\tau)\varepsilon^{-3N/p'}\big(\left|S\Pi(\bar{P}_k^\varepsilon) - S_k\bar{U}_k^\varepsilon\right|_{H^1(\Omega)} \\ & + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)}\big)\Big\} \\ & - \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2.\end{aligned} \tag{4.2.42}$$

Except for $\left|S\Pi(\bar{P}_k^\varepsilon) - S_k\bar{U}_k^\varepsilon\right|_{H^1(\Omega)}$ all terms already appear in (4.2.34). Let us therefore further estimate this term. Using continuity of S and setting $|S| = \|S\|_{\mathcal{L}(L_2(\Omega),H^1(\Omega))}$, we derive

$$|\bar{S}_k\bar{U}_k^\varepsilon - S\Pi(\bar{P}_k^\varepsilon)|_{H^1(\Omega)} \leq |S| \left\|\Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon\right\| + \left|(S - S_k)\bar{U}_k^\varepsilon\right|_{H^1(\Omega)}$$

With the help of this bound we can further estimate the left hand side in (4.2.42) to obtain the bound:

$$\begin{aligned}(\bar{\theta}^{\varepsilon^N}, \bar{y}^{\varepsilon^N} - S\Pi(\bar{P}_k^\varepsilon)) \lesssim \min\Big\{ & \|\bar{p}_s^{\varepsilon^N}\| \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, c(p')s(\tau)\varepsilon^{-3N/p'}\big(\left|(S - S_k)\bar{U}_k^\varepsilon\right|_{H^1(\Omega)} \\ & + |S| \left\|\Pi(\bar{P}_k^\varepsilon) - \bar{U}_k^\varepsilon\right\| + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)} \\ & + \|y_c - I_k y_c\|_{H^1(\Omega)}\big)\Big\} - \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2 \\ & =: \mathcal{Y}\end{aligned} \tag{4.2.43}$$

We observe that by definition, compare (4.2.40):

$$\mathcal{X} + \mathcal{Y} = \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) - \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2.$$

Bearing this relation in mind, we can now insert the bound derived in (4.2.43) in (4.2.41) to deduce:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 \lesssim c(p')s(\tau)\varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) - \frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2.$$

Shifting the term $-\frac{8}{\nu\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2$ to the left we obtain the desired inequality (4.2.32). This completes the proof.

□

4.2.3 Boundedness of the Explicit Constants

In the inequalities (4.2.27), (4.2.26), (4.2.34) and (4.2.33) there still appear the quantities

$$\left\|\bar{p}^{\varepsilon^N}\right\|, \left\|\Pi(\bar{P}_k^\varepsilon)\right\|.$$

In this section, our aim is to bound them uniformly to justify treating them as constants in the error estimators of Section 5.2.2. To prove this result we have to take advantage of the powerful machinery of solution notions of PDE, where the right hand side is merely a measure. Limiting the scope of this thesis, we will merely cite the most basic result which goes back to [77], Theorem 9.1. For additional information, we refer to the instructive paper [23].

Theorem 4.2.14 (Existence of Solutions for L_1 RHS)**.** *For every $\mu \in L_1(\Omega)$ there exists a unique $S^*\mu \in \mathring{W}_s^1(\Omega)$, $s < \frac{d}{d-1}$, $d = 2$ or $d = 3$, such that the following equation is fulfilled:*

$$(\nabla S^*\mu, \nabla z) = (\mu, z) \ \ \forall z \in \mathring{W}_{s'}^1(\Omega),\ \frac{1}{s} + \frac{1}{s'} = 1.$$

Besides

$$\|S^*\mu\|_{W_s^1(\Omega)} \lesssim \|\mu\|_{L_1(\Omega)}.$$

We can now tackle the main result of this section which is the following theorem:

Theorem 4.2.15. *The following bound is valid*

$$\left\|\Pi(\bar{P}_k^\varepsilon)\right\| \lesssim 1 \tag{4.2.44}$$

with a hidden constant depending solely on a, b, Ω.
Suppose further that Assumption 4.1.4 holds. Then for all $s < \frac{d}{d-1}$, $d = 2$ or $d = 3$

$$\left\|\bar{p}_s^{\varepsilon^N}\right\|, \left\|\bar{p}_s^{\varepsilon^N}\right\|_{W_s^1(\Omega)} \lesssim s(\tau). \tag{4.2.45}$$

Proof. Let us tackle (4.2.44) first. Since

$$a \le \Pi(\bar{P}_k^\varepsilon) \le b,$$

this bound is trivial. For the second bound (4.2.45) we have to utilise Theorem 4.2.14, which provides us with the following estimate:

$$\left\|\bar{p}_s^{\varepsilon^N}\right\|_{W_s^1(\Omega)} = \left\|S^*\bar{\theta}^{\varepsilon^N}\right\|_{W_s^1(\Omega)} \lesssim \left\|\bar{\theta}^{\varepsilon^N}\right\|_{L_1(\Omega)}$$

with $s < \frac{d}{d-1}$. Combining this result with Lemma 4.1.7, we obtain the desired result. After all, thanks to Theorem 2.1.35 $W^1_s(\Omega) \hookrightarrow L_p(\Omega)$ compactly for all $p < \infty$, $d = 2$ and $p < 3$ for $d = 3$, we have

$$\left\|\bar{p}_s^{\varepsilon^N}\right\|_{L_2(\Omega)} \lesssim \left\|\bar{p}_s^{\varepsilon^N}\right\|_{W^1_s(\Omega)}.$$

□

Remark 4.2.16. *Theorem 4.2.15 provides the justification for shifting*

$$\left\|\bar{p}^{\varepsilon^N}\right\| \left\|\Pi(\bar{P}_k^{\varepsilon})\right\|$$

into the hidden constant $\lesssim$ which then depends on data $S, \Omega, a, b, y_d, \nu, ..$ and the generic embedding constant $c(p')$ as well as the generic Slater point related constant $s(\tau)$.

Let us finish this section with a remark about the properties and structure of the estimators derived in Theorem 4.2.12 and Theorem 4.2.13:

Remark 4.2.17. *The estimators of Theorem 4.2.12 and Theorem 4.2.13 have two main issues: The first one is the fact that they still contain linear errors such as $\left\|(S - S_k)\bar{U}_k^{\varepsilon}\right\|$ which need to be estimated further, because the function $S\bar{U}_k^{\varepsilon}$ is in general not known. To overcome this problem, we will introduce residual type error estimators providing an upper bound for these terms in Section 4.3.3.*
The second issue is centred around the fact that the term $\mathcal{E}_s$ in both the variational and full discretisation setting does not lend itself to localisation*, because, in essence, $\mathcal{E}_s$ does not contain squared L_2- and H^1-(semi)norms, just plain norms. To remedy this disadvantage, we will describe a way to estimate the term(s) $\mathcal{E}_s$ further to allow precisely for a localisation by elementwise contributions. This will be done in Section 5.2 and Section 5.3.*

We will now conclude this chapter by examining the issues of convergence of the estimator.

4.3 Convergence Properties of the Estimator

In this section the onus lies on giving the reader an impression of how the estimators derived in Theorems 4.2.12 and 4.2.13 behave as $k \to \infty$ and $\varepsilon_k \to 0$. Naturally, the desired property would be

$$c(p')s(\tau)\varepsilon_k^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) + \mathcal{E}_s(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) \to 0 \text{ as } \varepsilon_k \to 0, k \to \infty$$

without any further conditions.
However, as we will soon discover, we will not be able to prove such a result without enforcing

further **sufficient** conditions. Apart from certain technical assumptions, see (CA1)-(CA5) below, the key condition for convergence will take the shape of

$$\varepsilon_k^{-r}(h_k^{\max})^s, \; r, s > 0, \tag{4.3.1}$$

where

$$h_k^{\max} := \max_{T \in \mathcal{T}_k} h_T. \tag{4.3.2}$$

defines the maximal mesh size at iterate k. We emphasise that conditions of the type 4.3.1 should be viewed in a way that eventually successive refinement, i.e. $h_k^{\max} \to 0$, will make the adaptive algorithm converge provided we are careful in choosing the regularisation parameter ε and N.

To be more specific, we will present the reader with our **convergence theorem**, the key result of this section, which we will prove step by step over the next few pages:

Theorem (Convergence of Estimator)**.** *Let $\varepsilon_k \to 0$, $k \to \infty$, $N \geq 1$ and $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$ be chosen such that*

$$\varepsilon_k^{3/4-3N/p'}, \; \varepsilon_k^{-3N/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{q}}, \; \varepsilon_k^{-\frac{3N+1}{p'}} h_k^{\max} \to 0,$$
$$\varepsilon_k^{-9/4}(h_k^{\max})^{1/2} \to 0, \varepsilon_k^{-3} h_k^{\max} \to 0 \;\; \min(\varepsilon_k^{2(N-1/p')-3/2}, \varepsilon_k^{\frac{3}{2}(N-1)}) \to 0$$

Then

$$\varepsilon_k^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) + \mathcal{E}_s(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) \to 0 \;\; \textit{as } \varepsilon_k \to 0, k \to \infty.$$

At this stage we want to stress that the error estimators (4.2.32) and (4.2.25) still contain **linear errors**, which still need to be estimated. We will sketch some of the issues and present residual type estimators in Section 4.3.3, which also converge as $\varepsilon_k \to 0$ and $k \to \infty$.

The focus in this section lies very much on the error estimators derived for the full discretisation setting, Theorem 4.2.13. Convergence of the error estimators for the variational discretisation approach, Theorem 4.2.12, is then just an easy consequence of convergence of the error estimators in the full discretisation case.

For notational convenience, we will **now drop explicitly stating the constants** $c(p'), s(\tau)$. The reader should bear in mind that in certain estimates they arise.

Let us now give a list of the **additional assumptions** we make to prove the convergence theorem:

CA1. The Slater-type assumption, Assumption 4.1.4, is fulfilled.

CA2. Let $\bar{q} > \frac{d}{2}$ be a fixed real number: Then y_c has the following properties: $y_c \in \mathring{H}^1(\Omega) \cap W^2_{\bar{q}}(\Omega)$ and $\nabla y_c \in H(\mathrm{div}, \Omega)$.

CA3. The convergence condition (3.3.4) is fulfilled.

CA4. The operator I_k is the Lagrange interpolant. It is well-defined for y_c since $W^2_{\bar{q}}(\Omega) \hookrightarrow C(\bar{\Omega})$, $\bar{q} > \frac{d}{2}$, compare Theorem 2.1.35.

CA5. For all $k \geq N$ there exists a constant independent of k and ε, such that

$$\|\bar{\theta}^{\varepsilon}_k\|_{L_1(\Omega)} \lesssim 1$$

CA6. Let $4 < p' < \infty$ in case $d = 2$ and $p' = 6$ in case $d = 3$ and $\frac{1}{p} + \frac{1}{p'} = 1$

It is worthwhile to add some explanatory remarks to these assumptions:

- (CA1): We already needed this assumption to derive the estimators (4.2.25) and (4.2.32). Thus, it is only natural that we need it again in this setting.
- (CA2): The higher regularity of y_c is needed to use standard interpolation estimates. However, it is not overly restrictive at all, since - as we will see in Theorem 4.3.1 - it merely reflects the generic regularity of the solution to the PDE in (CMP).
- (CA4): We want to stress here that interpolation operators such as the Scott-Zhang or Clément operators could be used as well for the operator I_k. The crucial thing is that I_k provides us with certain rates of convergence in terms of the mesh-size: $\|y_c - I_k y_c\| \lesssim (h^{\max}_k)^{\rho}$, ρ large enough.
- (CA5): This bound is crucial for providing estimates for the L_p-norm of the discrete multiplier, $\|\bar{\theta}^{\varepsilon}_k\|_{L_p(\Omega)}$, with $1 \leq p \leq 2$ which in turn are need to bound quantities such as $|S^*\bar{\theta}^{\varepsilon}_k|_{H^1(\Omega)}$, a key tool to prove convergence as we will soon discover.
  However, the reader should note that we only enforce a uniform bound in the relatively weak $L_1(\Omega)$-norm. On the continuous level we have such a bound, compare Lemma 4.1.7.
- (CA6): This is the by now familiar convention for p' and p which enables us to use the a priori estimates of Theorem 4.1.10 and which we already demanded in Theorem 4.2.12 and Theorem 4.2.13, where we derived the estimators for the variational and full discretisation respectively.

Before we explore some consequence of these assumptions, let us first - for future use - record an additional regularity result for the solution of the PDE in (CMP). The proof of the W_p^1-regularity can be found in [73], Section 4, Theorem 2, the proofs of the W_q^2-regularity are in Chapter 4, [37], in case $d = 2$ and Chapter 2, [38], if $d = 3$:

Theorem 4.3.1. *The solution operator S is a linear and continuous mapping $S : L_2(\Omega) \to \mathring{W}_{\bar{p}}^1(\Omega) \cap W_{\bar{q}}^2(\Omega)$ for some $\bar{p} > d$ and $\bar{q} > \frac{d}{2}$ from Assumption* (CA2)*.*

Remark 4.3.2. *The $\bar{q}$ of the theorem above is the same $\bar{q}$ as in* (CA2) *and **will always remain the same in the next two sections!***

As a corollary of Assumption (CA5) we obtain the following bound

Corollary 4.3.3. *Suppose that Assumption* (CA5) *holds. Then with p', p from Assumption* (CA6)*:*

$$\|\bar{\theta}_k^{\varepsilon}\|_{L_p(\Omega)} \lesssim \varepsilon^{-3/p'}.$$

Proof. The proof merely constitutes tracing the arguments of Lemma 4.1.7. □

Now, we can actually commence the task of proving our convergence theorem. The terms we will focus upon first are the linear errors $S - S_k$, this will be the subject of the next section:

4.3.1 Convergence of the Linear Errors

To start with, we observe that condition (CA3) ensures that (3.3.4) is fulfilled, which, in turn, guarantees, compare Corollary 3.3.11 and Corollary 3.3.12, that for any null sequence ε_k and $k \to \infty$:

$$\begin{aligned} \bar{U}_k^{\varepsilon_k} &\to \bar{u} \text{ in } L_2(\Omega) \\ \bar{Y}_k^{\varepsilon_k} &\to \bar{y} \text{ in } H^1(\Omega) \\ \frac{1}{\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2 &\to 0. \end{aligned} \tag{4.3.3}$$

These are the necessary ingredients to prove the following theorem:

Theorem 4.3.4 (Convergence of Linear Errors)**.** *Let $\varepsilon_k \to 0$ be a null sequence and $k \to \infty$. Then*

$$\left\|(S - S_k)\bar{U}_k^{\varepsilon_k}\right\|, |(S - S_k)\bar{U}_k^{\varepsilon_k}|_{H^1(\Omega)} \to 0,\ k \to \infty \tag{4.3.4}$$

Suppose further that ε_k and p' from Assumption (CA6) *are chosen in such a way that*

$$\varepsilon_k^{-\frac{3}{p'}}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{q}} \to 0,\ k \to \infty, \tag{4.3.5}$$

with $\bar{q} > \frac{d}{2}$ *from* (CA2)*, then*

$$\left\|(S-S_k)^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon_k})\right\| \to 0,\ k \to \infty. \tag{4.3.6}$$

Proof. We prove (4.3.4) first: We observe with $\|S\| = \|S\|_{\mathcal{L}(L_2(\Omega),L_2(\Omega))}$ and $|S| = \|S\|_{\mathcal{L}(L_2(\Omega),\mathring{H}^1(\Omega)}$ and the same convention for $\|S_k\|$ and $|S_k|$ that

$$\begin{aligned}\left\|(S-S_k)\bar{U}_k^{\varepsilon_k}\right\| &\le \|S\| \left\|\bar{u} - \bar{U}_k^{\varepsilon_k}\right\| + \|(S-S_k)\bar{u}\| + \|S_k\| \left\|\bar{u} - \bar{U}_k^{\varepsilon_k}\right\| \\ \left|(S-S_k)\bar{U}_k^{\varepsilon_k}\right|_{H^1(\Omega)} &\le |S| \left\|\bar{u} - \bar{U}_k^{\varepsilon_k}\right\| + |(S-S_k)\bar{u}|_{H^1(\Omega)} + |S_k| \left\|\bar{u} - \bar{U}_k^{\varepsilon_k}\right\|.\end{aligned}$$

Since $S_k g \to Sg$ for all $g \in L_2(\Omega)$ - see Theorem 2.3.10 - and (4.3.3) holds, uniform boundedness of $\|S_k\|, |S_k|$, compare Corollary 2.1.22, ensures that the right hand sides in the inequalities above converge. This gives (4.3.4).

Let us now tackle (4.3.6): With arguments completely analogous to those above, we immediately obtain

$$S_k^*(\bar{Y}_k^{\varepsilon_k} - y_d) \to S^*(\bar{y} - y_d), \ \text{ in } \mathring{H}^1(\Omega),\ k \to \infty$$

Thanks to the Poincaré-Friedrich inequality, Theorem 2.1.33, we thus immediately deduce:

$$(S-S_k)^*(\bar{Y}_k^{\varepsilon_k} - y_d) \to 0 \ \text{ in } L_2(\Omega),\ k \to \infty. \tag{4.3.7}$$

Hence, to prove (4.3.6), we have to show $L_2(\Omega)$-convergence of $(S-S_k)^*\bar{\theta}_k^{\varepsilon_k} \to 0$. Here, though, things are not that straightforward, since $\bar{\theta}_k^{\varepsilon_k}$ need not strongly converge. In the remaining part of the proof, we will demonstrate how to overcome this obstacle:

We note that

$$\left\|(S-S_k)^*(-\bar{\theta}_k^{\varepsilon_k})\right\| = \sup_{g \in L_2(\Omega)\setminus\{0\}} \frac{((S-S_k)^*(-\bar{\theta}_k^{\varepsilon_k}), g)}{\|g\|}. \tag{4.3.8}$$

We define the space W by

$$W := W_{\bar{q}}^2(\Omega) \cap \left\{\psi \in \mathring{H}^1(\Omega) \ : \ \nabla\psi \in H(\operatorname{div}, \Omega)\right\}$$

and observe that thanks to Theorem 4.3.1 and Lemma 4.1.3 we have for an arbitrary $g \in L_2(\Omega)$ a $\psi \in W$ with $-\Delta\psi = g$ a.e in Ω. Consequently (4.3.8) is equivalent to

$$\left\|(S-S_k)^*(-\bar{\theta}_k^{\varepsilon_k})\right\| = \sup_{\psi \in W\setminus\{0\}} \frac{((S-S_k)^*(-\bar{\theta}_k^{\varepsilon_k}), -\Delta\psi)}{\|\Delta\psi\|}. \tag{4.3.9}$$

Let us therefore investigate the term $((S-S_k)^*(-\bar{\theta}_k^{\varepsilon_k}), -\Delta\psi)$ for arbitrary $\psi \in W$. Using

Green's formula and $\psi \in \mathring{H}^1(\Omega)$, we deduce

$$((S - S_k)^*(-\bar\theta_k^{\varepsilon_k}), -\Delta\psi) = -(\nabla S\bar\theta_k^{\varepsilon_k}, \nabla\psi) + (\nabla S_k\bar\theta_k^{\varepsilon_k}, \nabla\psi). \tag{4.3.10}$$

Using the Ritz projection R_k, (4.2.13), and the fact that $R_k(S^*\bar\theta_k^{\varepsilon_k}) = S_k^*\bar\theta_k^{\varepsilon_k}$, we can continue to rearrange (4.3.10) in the following way:

$$\begin{aligned}((S - S_k)^*(-\bar\theta_k^{\varepsilon_k}), -\Delta\psi) &= -(\nabla S^*\bar\theta_k^{\varepsilon_k}, \nabla\psi) + (\nabla S_k^*\bar\theta_k^{\varepsilon_k}, \nabla\psi)\\ &= -(\nabla S^*\bar\theta_k^{\varepsilon_k}, \nabla\psi) + (\nabla \bar S_k^*\bar\theta_k^{\varepsilon_k}, \nabla R_k\psi)\\ &= -(\nabla S^*\bar\theta_k^{\varepsilon_k}, \nabla\psi) + (\nabla \bar S^*\bar\theta_k^{\varepsilon_k}, \nabla R_k\psi)\\ &= -(\nabla S^*\bar\theta_k^{\varepsilon_k}, \nabla(\psi - R_k\psi)).\end{aligned}$$

An application of Cauchy-Schwarz's inequality on the right then yields:

$$((S - S_k)^*(-\bar\theta_k^{\varepsilon_k}), -\Delta\psi) \le \left|S^*\bar\theta_k^{\varepsilon_k}\right|_{H^1(\Omega)} |\psi - R_k\psi|_{H^1(\Omega)}. \tag{4.3.11}$$

Since $S^* : H^{-1}(\Omega) \to \mathring{H}^1(\Omega)$ and $L_p(\Omega) \hookrightarrow H^{-1}(\Omega)$, $\frac{1}{p} + \frac{1}{p'} = 1$ thanks to Assumption (CA6), we obtain using Corollary 4.3.3:

$$|S^*\bar\theta_k^{\varepsilon_k}|_{H^1(\Omega)} \lesssim \left\|\bar\theta_k^{\varepsilon_k}\right\|_{L_p(\Omega)} \lesssim \varepsilon_k^{-3/p'}. \tag{4.3.12}$$

This can now be inserted in (4.3.11) so that together with standard interpolation estimates for the Ritz projection, see e.g. [14], Sections 4.4 and 4.8, we are able to conclude :

$$\begin{aligned}((S - S_k)^*(-\bar\theta_k^{\varepsilon_k}), -\Delta\psi) &\lesssim \varepsilon_k^{-3/p'} |\psi - R_k\psi|_{H^1(\Omega)}\\ &\lesssim \varepsilon_k^{-3/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\tilde q}} \|\psi\|_{W^2_{\tilde q}(\Omega)}\end{aligned}$$

Now, observe that due to Theorem 4.3.1 and the definition of ψ, $\|\psi\|_{W^2_{\tilde q}(\Omega)} \lesssim \|g\| = \|-\Delta\psi\|$. Thus for any such $0 \neq \psi \in W$:

$$\frac{((S - S_k)^*(-\bar\theta_k^{\varepsilon_k}), -\Delta\psi)}{\|\Delta\psi\|} \lesssim \varepsilon_k^{-3/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\tilde q}} \frac{\|\psi\|_{W^2_{\tilde q}(\Omega)}}{\|-\Delta\psi\|} \lesssim \varepsilon_k^{-3/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\tilde q}}.$$

Plugging this bound into (4.3.9) yields:

$$\left\|(S - S_k)^*(-\bar\theta_k^{\varepsilon_k})\right\| \lesssim \varepsilon_k^{-3/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\tilde q}}$$

Due to (4.3.6) the right hand side converges and together with (4.3.7) we obtain the desired result.

□

In the singular parts of the error estimators (4.2.27) and (4.2.34) 'weighted' terms of the type:

$$\varepsilon^{-3N/p'} \left|(S - S_k)\bar{U}_k^{\varepsilon}\right|_{H^1(\Omega)}$$

occur. With the help of a condition similar to (4.3.6) we can prove convergence for those terms, too:

Theorem 4.3.5. *Suppose that* $\varepsilon_k \to 0$*,* N *and* p' *from Assumption* (CA6) *are chosen in such a way that*

$$\varepsilon_k^{-3N/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\bar{q}}} \to 0,\ k \to \infty. \tag{4.3.13}$$

with $\bar{q}$ *from Assumption* (CA2)*. Then*

$$\varepsilon_k^{-3N/p'}|(S - S_k)\bar{U}_k^{\varepsilon_k}|_{H^1(\Omega)} \to 0,,\ k \to \infty. \tag{4.3.14}$$

Proof. Standard interpolation estimates yield

$$|(S - S_k)\bar{U}_k^{\varepsilon_k}|_{H^1(\Omega)} \lesssim (h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\bar{q}}} \left\|S\bar{U}_k^{\varepsilon_k}\right\|_{W_{\bar{q}}^2(\Omega)}$$

Multiplying this term with $\varepsilon_k^{-3N/p'}$ yields:

$$\varepsilon_k^{-3N/p'}|(S - S_k)\bar{U}_k^{\varepsilon_k}|_{H^1(\Omega)} \lesssim \varepsilon_k^{-3N/p'}(h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\bar{q}}} \left\|S\bar{U}_k^{\varepsilon_k}\right\|_{W_{\bar{q}}^2(\Omega)}$$

Theorem 4.3.1 now allows us to conclude uniform boundedness of $\left\|S\bar{U}_k^{\varepsilon_k}\right\|_{W_{\bar{q}}^2(\Omega)}$ and condition (4.3.14) guarantees convergence of the right hand side and thus the desired result (4.3.14). □

Remark 4.3.6. *We observe that for* $N \geq 1$ *condition* (4.3.14) *is a stronger condition than* (4.3.6)*. Thus, it is condition* (4.3.14) *which appears in our convergence theorem not its weaker counterpart* (4.3.6)*.*

Having ascertained convergence of the linear error terms we can now focus on those terms in (4.2.25) and (4.2.32) which contain only known discrete and continuous functions, i.e. those quantities which we dubbed 'computable' at the beginning of this chapter.

4.3.2 Convergence of the Computable Quantities

Theorem 4.3.7 (Convergence of Projection Error)**.** *Let* $\varepsilon_k \to 0$ *and* p' *from Assumption* (CA6) *are chosen in such a way that*

$$\varepsilon_k^{-3/p'} h_k^{\max} \to 0,\ k \to \infty. \tag{4.3.15}$$

Then

$$\left\|\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k})\right\| \to 0,\ k \to \infty.$$

In particular,

$$\Pi(\bar{P}_k^{\varepsilon_k}) \to \bar{u},\ k \to 0$$

Proof. The key estimate that we prove in the course of this proof is

$$\left\|\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k})\right\| \lesssim \varepsilon_k^{-3/p'} h_k^{\max}.$$

To derive this estimate, we recall Lemma 4.1.15. There, we demonstrated that $\Pi(\cdot)$ is the best-approximation of $-\frac{1}{\nu}\cdot$ in the convex and closed set

$$\mathcal{U} = \{u \in L_2(\Omega)\ :\ a \leq u \leq b \text{ a.e in } \Omega\}.$$

Thanks to Theorem 2.1.7 we know that Π is Lipschitz-continuous. Together with the projection relation (4.1.33) and standard estimates for the L_2-projection of an H^1-function, we then obtain:

$$\begin{aligned}\left\|\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k})\right\| &= || \underbrace{\Pi(P_{\mathbb{U}_k}(\bar{P}_k^{\varepsilon_k}))}_{=\bar{U}_k^{\varepsilon_k}} - \Pi(\bar{P}_k^{\varepsilon_k})|| \\ &\leq \frac{1}{\nu}\left\|P_{\mathbb{U}_k}(\bar{P}_k^{\varepsilon_k}) - \bar{P}_k^{\varepsilon_k}\right\| \\ &\lesssim h_k^{\max}|\bar{P}_k^{\varepsilon_k}|_{H^1(\Omega)}.\end{aligned} \tag{4.3.16}$$

We now observe that $\bar{P}_k^{\varepsilon}$ is the Ritz-projection (4.2.13) of $S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon_k})$. Harnessing its stability in the H^1 semi-norm, we can pursue our estimates with the help of Assumption (CA5) and (4.3.12) in the following way:

$$\begin{aligned}h_k^{\max}|\bar{P}_k^{\varepsilon_k}|_{H^1(\Omega)} &\leq h_k^{\max}|\bar{P}_k^{\varepsilon_k} - S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon_k})|_{H^1(\Omega)} + h_k^{\max}|S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon_k})|_{H^1(\Omega)} \\ &\lesssim h_k^{\max}|S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon_k})|_{H^1(\Omega)} \\ &\lesssim h_k^{\max}\left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\| + h_k^{\max}\left\|\bar{\theta}_k^{\varepsilon_k}\right\|_{L_p(\Omega)},\quad \frac{1}{p} + \frac{1}{p'} = 1 \\ &\lesssim \varepsilon_k^{-3/p'} h_k^{\max}.\end{aligned}$$

Inserting this bound in (4.3.16), we obtain:

$$\left\|\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k})\right\| \lesssim \varepsilon_k^{-3/p'} h_k^{\max}.$$

If (4.3.15) holds, then the right hand side in the inequality above converges and hence, the left side does, too. Due to (4.3.3) we know in addition that

$$\lim_{k\to\infty} \Pi(\bar{P}_k^{\varepsilon_k}) = \lim_{k\to\infty} \bar{U}_k^{\varepsilon_k} = \bar{u}.$$

This is the desired result. □

Let us recollect that in the singular part of the error estimator (4.2.34) the following term enters:

$$\varepsilon_k^{-3N/p'} \left\| \bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}) \right\|$$

To prove convergence of this term, we need to strengthen condition (4.3.15) in a similar way as in Theorem 4.3.5. This is done in the next theorem:

Theorem 4.3.8. *Suppose that $\varepsilon_k \to 0$, p' from Assumption* (CA6) *and $N \geq 1$ are chosen such that the following condition is fulfilled*

$$\varepsilon_k^{\frac{-3(N+1)}{p'}} h_k^{\max} \to 0,\ k \to \infty. \tag{4.3.17}$$

Then,

$$\varepsilon_k^{\frac{-3N}{p'}} \left\| \bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}) \right\| \to 0,\ k \to \infty. \tag{4.3.18}$$

Proof. From Theorem 4.3.7 we know that

$$\left\| \bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}) \right\| \lesssim \varepsilon_k^{-3/p'} h_k^{\max}.$$

Multiplying the inequality above with $\varepsilon^{3N/p'}$ yields:

$$\varepsilon_k^{-3N/p'} \left\| \bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}) \right\| \lesssim \varepsilon_k^{-3(N+1)/p'} h_k^{\max}.$$

Condition (4.3.17) then implies the desired result. □

Remark 4.3.9. *We observe that* (4.3.17) *is a stronger condition compared to* (4.3.15)*, thus it is the former which enters as a condition in our convergence theorem.*

We can now turn to the remaining explicit quantities. To tackle them, we need the ensuing auxiliary estimate:

Lemma 4.3.10. *For each $\varepsilon > 0$ and $\bar{q}$ from Assumption* (CA2) *we have:*

$$|(\bar{Y}_k^{\varepsilon} - I_k y_c)^-|_{H^1(\Omega)} \lesssim (h_k^{\max})^{1+\frac{d}{2}-\frac{d}{\bar{q}}} + \varepsilon^{3/4}$$

Proof. Define $\hat{y} := S\bar{U}_k^\varepsilon$ be given. Then we can estimate in the following fashion:

$$\begin{aligned}
|(\bar{Y}_k^\varepsilon - I_k y_c)^-|^2_{H^1(\Omega)} &= \int\limits_\Omega \nabla(\bar{Y}_k^\varepsilon - I_k y_c)^- \nabla(\bar{Y}_k^\varepsilon - I_k y_c)\, d\Omega \\
&= \int\limits_\Omega \nabla(\bar{Y}_k^\varepsilon - I_k y_c)^- \nabla(\bar{Y}_k^\varepsilon - \hat{y})\, d\Omega + \int\limits_\Omega \nabla(\bar{Y}_k^\varepsilon - I_k y_c)^- \nabla(\hat{y} - y_c)\, d\Omega \\
&+ \int \nabla(\bar{Y}_k^\varepsilon - I_k y_c)^- \nabla(y_c - I_k y_c)\, d\Omega.
\end{aligned}$$

Using Young's inequality, (4.2.2), and Green's formula, we can pursue our estimates above to derive:

$$\begin{aligned}
|(\bar{Y}_k^\varepsilon - I_k y_c)^-|^2_{H^1(\Omega)} \le{}& \frac{1}{2}|(\bar{Y}_k^\varepsilon - I_k y_c)^-|^2_{H^1(\Omega)} + |\hat{y} - \bar{Y}_k^\varepsilon|^2_{H^1(\Omega)} + \|y_c - I_k y_c\|^2_{H^1(\Omega)} \\
&+ \int\limits_\Omega (-\Delta\hat{y} + \Delta y_c)(\bar{Y}_k^\varepsilon - I_k y_c)^-\, d\Omega
\end{aligned} \tag{4.3.19}$$

Let us have a closer look at the term $(\bar{Y}_k^\varepsilon - I_k y_c)^-$. In case $\bar{Y}_k^\varepsilon - I_k y_c \le 0$ feasibility of $(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ yields

$$\varepsilon\bar{V}_k^\varepsilon \ge I_k y_c - \bar{Y}_k^\varepsilon.$$

Thus,

$$\left\|(\bar{Y}_k^\varepsilon - I_k y_c)^-\right\| \le \varepsilon \left\|\bar{V}_k^\varepsilon\right\|.$$

Observe now that $-\Delta\hat{y} = \bar{U}_k^\varepsilon$ a.e. in Ω. Employing our deductions above, we gain for the last integral on the right in (4.3.19):

$$\begin{aligned}
\int\limits_\Omega (-\Delta\hat{y} + \Delta y_c)(\bar{Y}_k^\varepsilon - I_k y_c)^-\, d\Omega &\le ||(\underbrace{\bar{U}_k^\varepsilon}_{=-\Delta\hat{y}} + \Delta y_c)||\left\|(\bar{Y}_k^\varepsilon - I_k y_c)^-\right\| \\
&\le \left\|\bar{U}_k^\varepsilon + \Delta y_c\right\| \varepsilon \left\|\bar{V}_k^\varepsilon\right\| \lesssim \varepsilon^{3/2}.
\end{aligned}$$

Let us now insert this estimate in (4.3.19). Harnessing standard interpolation estimates, Assumption (CA2) and the fact that $\hat{y} \in W^2_{\bar{q}}(\Omega)$ thanks to Theorem 4.3.1, we can then conclude for (4.3.19):

$$\begin{aligned}
|(\bar{Y}_k^\varepsilon - I_k y_c)^-|^2_{H^1(\Omega)} &\lesssim |\underbrace{\hat{y} - \bar{Y}_k^\varepsilon}_{=(S-S_k)\bar{U}_k^\varepsilon}|^2_{H^1(\Omega)} + \|y_c - I_k y_c\|^2_{H^1(\Omega)} + \varepsilon^{3/2} \\
&\lesssim (h_k^{\max})^{2+d-\bar{q}} \|\hat{y}\|^2_{W^2_{\bar{q}}(\Omega)} (h_k^{\max})^{2+d-\bar{q}} \|y_c\|^2_{W^2_{\bar{q}}(\Omega)} + \varepsilon^{3/2}.
\end{aligned}$$

Drawing the square root then yields the postulated assertion. □

After this slight detour, we are now in the position to prove the following theorem:

Theorem 4.3.11 (Convergence of Explicit Quantities)**.** *Let* $\varepsilon_k \to 0$ *be a null sequence and* $k \to \infty$*. Then*

$$(\bar{\theta}_k^{\varepsilon}, \bar{Y}_k^{\varepsilon} - I_k y_c) \to 0,\ k \to \infty \tag{4.3.20}$$

Suppose now that (4.3.15) *is satisfied. Then,*

$$\left\|(\Pi(\bar{P}_k^{\varepsilon_k}) + \Delta y_c)^-\right\| \to \left\|(\bar{u} + \Delta y_c)^-\right\|,\ k \to \infty. \tag{4.3.21}$$

Besides, let (4.3.5) *be fulfilled. Then*

$$(\bar{\theta}_k^{\varepsilon}, I_k y_c - y_c) \to 0,\ k \to \infty \tag{4.3.22}$$

and

$$(\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}), \bar{P}_k^{\varepsilon_k}),\ k \to \infty. \tag{4.3.23}$$

Assume further that $N \geq 1$*,* $\varepsilon_k \to 0$ *and* p' *from Assumption* (CA6) *are chosen in such a way that conditions* (4.3.13)*,* (4.3.17) *as well as*

$$\varepsilon_k^{\frac{3}{4} - \frac{3N}{p'}} \to 0,\ k \to \infty, \tag{4.3.24}$$

are fulfilled. Then

$$\begin{aligned} \varepsilon_k^{-3N/p'} \big(\left\|\Pi(\bar{P}_k^{\varepsilon_k}) - \bar{U}_k^{\varepsilon_k}\right\| + |(\bar{Y}_k^{\varepsilon_k} - I_k y_c)^-|_{H^1(\Omega)} \\ + \|y_c - I_k y_c\|_{H^1(\Omega)} \big) \to 0,\ k \to \infty \end{aligned} \tag{4.3.25}$$

Proof. The slackness identity in (4.1.25) immediately yields

$$(\bar{\theta}_k^{\varepsilon}, \bar{Y}_k^{\varepsilon} - I_k y_c) = -\frac{1}{\varepsilon_k} \left\|\bar{V}_k^{\varepsilon_k}\right\|^2.$$

The fact that we demanded that the basic convergence condition (3.3.4) be fulfilled, Assumption (CA3), then allows us to conclude, compare Corollary 3.3.12:

$$\frac{1}{\varepsilon_k} \left\|\bar{V}_k^{\varepsilon_k}\right\|^2 \to 0,\ k \to \infty.$$

Thus, (4.3.20) follows.

Let us tackle (4.3.21). Since we enforced condition (4.3.15), we know thanks to Theorem 4.3.7

that

$$\Pi(\bar{P}_k^{\varepsilon_k}) \to \bar{u},\ k \to \infty.$$

Continuity of $\|\cdot\|$ then gives the desired result (4.3.21).
Let us now examine (4.3.22): For p' and p from Assumption (CA6) we have the following estimate (compare also Corollary 4.3.3):

$$\begin{aligned}(\bar{\theta}_k^{\varepsilon}, I_k y_c - y_c) &\le c(p') \left\|\bar{\theta}_k^{\varepsilon_k}\right\|_{L_p(\Omega)} \|I_k y_c - y_c\|_{L_{p'}(\Omega)}\\ &\lesssim c(p') \left\|\bar{\theta}_k^{\varepsilon_k}\right\|_{L_p(\Omega)} \|I_k y_c - y_c\|_{H^1(\Omega)}\\ &\lesssim \varepsilon_k^{-3/p'} (h_k^{\max})^{1+\frac{d}{2}-\frac{d}{q}} \|y_c\|_{W_{\bar{q}}^2(\Omega)},\end{aligned}$$

Condition (4.3.5) now implies convergence.
Turning to (4.3.23), we first realise that thanks to (CA5) and Theorem 4.2.14 we know that

$$\left\|S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon})\right\| \lesssim 1. \tag{4.3.26}$$

We now have the standard estimate:

$$\left\|S_k^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon})\right\| - \left\|S^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon})\right\| \le \left\|(S - S_k)^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon})\right\|.$$

Condition (4.3.5) guarantees convergence of the right-hand side above which in turn - coupled with (4.3.26) - ensures boundedness

$$\left\|\bar{P}_k^{\varepsilon_k}\right\| = \left\|S_k^*(\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon})\right\| \lesssim 1.$$

Now, Cauchy-Schwarz's inequality, (4.3.15) and the previous deductions together imply

$$(\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k}), \bar{P}_k^{\varepsilon_k}) \le \left\|\bar{U}_k^{\varepsilon_k} - \Pi(\bar{P}_k^{\varepsilon_k})\right\| \left\|\bar{P}_k^{\varepsilon_k}\right\| \to 0,\ k \to \infty$$

which is (4.3.23).
Let us now investigate (4.3.25) term by term:
For $|(\bar{Y}_k^{\varepsilon_k} - I_k y_c)^-|_{H^1(\Omega)}$ we employ Lemma 4.3.10 to deduce:

$$\varepsilon_k^{-3N/p'} |(\bar{Y}_k^{\varepsilon_k} - I_k y_c)^-|_{H^1(\Omega)} \lesssim \varepsilon_k^{-3N/p'} (h_k^{\max})^{1+\frac{d}{2}-\frac{d}{q}} + \varepsilon_k^{3/4-3N/p'}.$$

Conditions (4.3.13) and (4.3.24) then yield convergence.
Taking a look at $\|I_k y_c - y_c\|_{H^1(\Omega)}$, we are able to conclude with the help of standard interpo-

lation estimates:

$$\varepsilon_k^{-3N/p'} \left\| I_k y_c - y_c \right\|_{H^1(\Omega)} \lesssim \varepsilon_k^{-3N/p'} (h_k^{\max})^{1+\frac{d}{2}-\frac{d}{q}} \left\| y_c \right\|_{W_q^2(\Omega)} .$$

Lastly, for the term

$$\varepsilon_k^{-3N/p'} \left\| \Pi(\bar{P}_k^{\varepsilon_k}) - \bar{U}_k^{\varepsilon_k} \right\|$$

we take advantage of Theorem 4.3.8, in particular (4.3.17) to deduce convergence of the entire term (4.3.25). □

Remark 4.3.12. *In case $d = 3$, where $p' = 6$ by Assumption* (CA6)*, condition* (4.3.24) *poses a restriction on the choice of N with the help of which we can decrease the regularisation error on the continuous level. Naturally, this is not something for which we would wish. However, it should be pointed out that numerically we often observe a faster convergence of the regularisation error $\left\| \bar{V}_k^{\varepsilon} \right\|$, which is the source of the restriction* (4.3.24)*, compare the proof of Lemma 4.3.10. Thus, condition* (4.3.24) *is in effect not as severe it might strike the reader.*

Let us now tackle the term $\left\| \bar{\theta}_k^{\varepsilon_k} - P_k^{0+} \bar{\theta}_k^{\varepsilon_k} \right\|$:

Theorem 4.3.13. *Suppose that $\varepsilon_k \to 0$ as $k \to \infty$ satisfies*

$$\varepsilon_k^{-9/4} (h_k^{\max})^{1/2} \to 0, \;\; \varepsilon_k^{-3} h_k^{\max} \to 0, \; k \to \infty. \tag{4.3.27}$$

Then

$$\left\| \bar{\theta}_k^{\varepsilon_k} - P_k^{0+}(\bar{\theta}_k^{\varepsilon_k}) \right\| \to 0.$$

Proof. The proof is fairly technical, thus we will restrict ourselves to the most important steps here. The proof itself is based on the comparison of the problem (DMP_k^{ε}) with a semi-discrete problem of the type:

$$\left.\begin{array}{c} \displaystyle\min_{U \in \mathbb{U}_k, Y \in \mathbb{Y}_k, V \in L_2(\Omega)} \frac{1}{2} \left\| Y - y_d \right\|_{L_2(\Omega)}^2 + \frac{\nu}{2} \|U\|_{L_2(\Omega)}^2 + \frac{1}{2\varepsilon} \left\| V \right\|_{L_2(\Omega)}^2 \\ \text{s.t.} \\ \displaystyle\int_\Omega \nabla Y \cdot \nabla W \, d\Omega = \int_\Omega UW \, d\Omega. \; \forall W \in \mathbb{Y}_k \\ \text{and} \\ U \in \mathcal{U}_k \\ I_k y_c - Y - \varepsilon V \leq 0 \;\text{ a.e. on } \Omega \end{array}\right\} \qquad (SDMP_k^{\varepsilon})$$

Here, $(\hat{U}_k^{\varepsilon_k}, \hat{Y}_k^{\varepsilon_k}, \hat{V}_k^{\varepsilon_k}, \hat{\theta}_k^{\varepsilon_k})$ denotes the solution and associated Lagrange multiplier $\hat{\theta}_k^{\varepsilon_k}$ for problem $(SDMP_k^{\varepsilon})$.

The crucial theoretical advantage of $(SDMP_k^{\varepsilon})$ is a penalty structure for the virtual control mirroring the one in the continuous case, (4.1.8), i.e:

$$\hat{V}_k^{\varepsilon_k}(x) = -\frac{1}{\varepsilon_k}(\min((\hat{Y}_k^{\varepsilon_k}(x) - (I_k y_c)(x), 0)). \tag{4.3.28}$$

Note that thanks to (4.1.1) and $\|S_k\| \lesssim 1$ due to (Pr7), we have

$$\left\|\hat{Y}_k^{\varepsilon_k}\right\|_{H^1(\Omega)} \lesssim 1.$$

Hence, we immediately deduce thanks to (4.3.28)

$$\left\|\varepsilon_k \hat{V}_k^{\varepsilon_k}\right\|_{H^1(\Omega)} \lesssim 1. \tag{4.3.29}$$

Besides, using the Lagrange interpolant I_k, we realise at once that

$$(I_k \hat{V}_k^{\varepsilon_k})(x) \geq \hat{V}_k^{\varepsilon_k}(x) \;\text{ f.a.a. } x \in \Omega.$$

After all, $\hat{V}_k^{\varepsilon_k}$ is just the cut-off of the piecewise affine function $\hat{Y}_k^{\varepsilon_k} - I_k y_c$. Consequently, we discern $(\hat{U}_k^{\varepsilon_k}, I_k \hat{V}_k^{\varepsilon_k}) \in \mathbb{U}_k^{\varepsilon,ad}$.

Presently, testing the respective optimality conditions of (DMP_k^{ε}) and $(SDMP_k^{\varepsilon})$, we can then proceed in the following way:

$$\begin{aligned} \frac{1}{\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k} - \hat{V}_k^{\varepsilon_k}\right\|^2 &\leq \frac{1}{\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\| \left\|I_k(\hat{V}_k^{\varepsilon_k}) - \hat{V}_k^{\varepsilon_k}\right\| \\ &\leq \varepsilon_k^{-2}\left\|\bar{V}_k^{\varepsilon_k}\right\| \left\|I_k(\varepsilon_k \hat{V}_k^{\varepsilon}) - (\varepsilon_k \hat{V}_k^{\varepsilon_k})\right\| \\ &\leq \varepsilon_k^{-3/2} h_k^{\max} \left\|\varepsilon_k \hat{V}_k^{\varepsilon_k}\right\|_{H^1(\Omega)}. \end{aligned} \tag{4.3.30}$$

Here, we also used the bound, compare Theorem 3.2.5

$$\frac{1}{\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2 \lesssim 1.$$

Taking advantage of (4.3.29), we ultimately obtain for (4.3.30):

$$\left\|\bar{V}_k^{\varepsilon_k} - \hat{V}_k^{\varepsilon_k}\right\| \lesssim \varepsilon_k^{-1/4}(h_k^{\max})^{1/2}. \tag{4.3.31}$$

The KKT system of the semi-discrete problem also yields an equation for $\hat{\theta}_k^{\varepsilon_k}$ and $\hat{V}_k^{\varepsilon_k}$:

$$\hat{\theta}_k^{\varepsilon_k} = \frac{1}{\varepsilon_k^2} \hat{V}_k^{\varepsilon_k}. \tag{4.3.32}$$

Combining this with the same relation for $\bar{V}_k^{\varepsilon_k}$ and $\bar{\theta}_k^{\varepsilon_k}$ in the KKT system for (DMP_k^{ε}), (4.1.25), (4.3.31) allows to deduce:

$$\left\| \bar{\theta}_k^{\varepsilon_k} - \hat{\theta}_k^{\varepsilon_k} \right\| \lesssim \varepsilon_k^{-9/4} (h_k^{\max})^{1/2}. \tag{4.3.33}$$

Besides, we note that $\hat{\theta}_k^{\varepsilon_k} \geq 0$, hence Lemma 4.2.5 and in particular (4.2.11) yield:

$$P_k^{0+} \hat{\theta}_k^{\varepsilon_k} = P_{\mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))} \hat{\theta}_k^{\varepsilon_k} =: P_k^0 \hat{\theta}_k^{\varepsilon_k}.$$

Here, $P_{\mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))}$ denotes the L_2-orthogonal projection on the space $\mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))$. Standard interpolation estimates, (4.3.29) and (4.3.32) enable us to conclude:

$$\left\| P_k^0 \hat{\theta}_k^{\varepsilon_k} \right\| \lesssim \varepsilon_k^{-3} h_k^{\max} \left\| \varepsilon_k \hat{V}_k^{\varepsilon_k} \right\|_{H^1(\Omega)}. \tag{4.3.34}$$

With this at hand we can now use Lipschitz continuity of P_k^{0+} with constant 1 - compare Lemma 4.1.15 and Theorem 2.1.7 - (4.3.33) and (4.3.34) to arrive at:

$$\begin{aligned}
\| \bar{\theta}_k^{\varepsilon_k} - P_k^{0+}(\bar{\theta}_k^{\varepsilon_k}) \| &\leq \left\| \bar{\theta}_k^{\varepsilon_k} - \hat{\theta}_k^{\varepsilon_k} \right\| + \left\| \hat{\theta}_k^{\varepsilon_k} - P_k^{0+}(\hat{\theta}_k^{\varepsilon_k}) \right\| + \left\| P_k^{0+}(\hat{\theta}_k^{\varepsilon_k}) - P_k^{0+}(\bar{\theta}_k^{\varepsilon_k}) \right\| \\
&\leq 2 \left\| \bar{\theta}_k^{\varepsilon_k} - \hat{\theta}_k^{\varepsilon_k} \right\| + \left\| \hat{\theta}_k^{\varepsilon_k} - P_k^0(\hat{\theta}_k^{\varepsilon_k}) \right\| \\
&\lesssim \varepsilon_k^{-9/4} (h_k^{\max})^{1/2} + \varepsilon_k^{-3} h_k^{\max} \left\| \varepsilon_k \hat{V}_k^{\varepsilon_k} \right\|_{H^1(\Omega)}.
\end{aligned}$$

Recalling (4.3.29), we know that the right hand side of the inequality above converges to 0 provided (4.3.27) is satisfied. This is the desired result. □

We are now finally in the position to prove our convergence theorem from the beginning of the section:

Theorem 4.3.14 (Convergence of Estimator)**.** *Let $\varepsilon_k \to 0$, $k \to \infty$, $N \geq 1$ and p' from Assumption* (CA6) *be chosen such that as $k \to \infty$*

$$\begin{gathered}
\varepsilon_k^{3/4 - 3N/p'}, \; \varepsilon_k^{-3N/p'} (h_k^{\max})^{1 + \frac{d}{2} - \frac{d}{q}}, \; \varepsilon_k^{-\frac{3N+1}{p'}} h_k^{\max} \to 0, \\
\varepsilon_k^{-9/4} (h_k^{\max})^{1/2} \to 0, \varepsilon_k^{-3} h_k^{\max} \to 0 \;\; \min(\varepsilon_k^{2(N - 1/p') - 3/2}, \varepsilon_k^{\frac{3}{2}(N-1)}) \to 0.
\end{gathered} \tag{4.3.35}$$

Then

$$\varepsilon_k^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) + \mathcal{E}_s(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) \to 0 \;\; \textit{as } \varepsilon_k \to 0, k \to \infty.$$

Proof. We will start with the regular parts (4.2.26) and (4.2.33) and proceed term by term:

The regular part:
Theorem 4.3.4 yields convergence of $\left\|(S-S_k)\bar{U}_k^{\varepsilon_k}\right\|$ as $k\to\infty$ without any further condition. Thus, let us tackle $\left\|\bar{U}_k^{\varepsilon_k}-\Pi(\bar{P}_k^{\varepsilon_k})\right\|\to 0$: Here, (4.3.35) implies (4.3.15) for $N\geq 1$. With the help of Theorem 4.3.7 we can deduce convergence of $\left\|\bar{U}_k^{\varepsilon_k}-\Pi(\bar{P}_k^{\varepsilon_k})\right\|\to 0$ as $k\to\infty$. Since for $N\geq 1$ (4.3.35) additionally ensures that (4.3.6) holds, Theorem 4.3.11, in particular (4.3.23), allows us to conclude:

$$(\bar{U}_k^{\varepsilon_k}-\Pi(\bar{P}_k^{\varepsilon_k}),\bar{P}_k^{\varepsilon_k})\to 0,\ k\to\infty.$$

As already remarked, for $N\geq 1$ (4.3.35) is stronger than (4.3.5). Thus, thanks to Theorem 4.3.4, this allows us to conclude:

$$\left\|(S-S_k)^*(\bar{Y}_k^{\varepsilon_k}-y_d-\bar{\theta}_k^{\varepsilon_k})\right\|\to 0,\ k\to\infty$$

Condition (4.3.35) also ensures (4.3.5), which in turn thanks to Theorem 4.3.11 also guarantees convergence of $(\bar{\theta}_k^{\varepsilon_k},I_ky_c-y_c)\to 0$ and $(\bar{\theta}_k^{\varepsilon_k},\bar{Y}_k^{\varepsilon_k}-I_ky_c)\to 0$. All in all, we can thus conclude:

$$\mathcal{E}_r^2(\bar{U}_k^{\varepsilon_k},\bar{V}_k^{\varepsilon_k})\to 0,\ k\to\infty$$

Let us now turn to $\mathcal{E}_s$:

The singular part:
Again, we will proceed term-by-term. Convergence of the terms

$$\left\|(S-S_k)^*(\bar{Y}_k^{\varepsilon_k}-y_d-\bar{\theta}_k^{\varepsilon_k})\right\|,\left\|\bar{U}_k^{\varepsilon_k}-\Pi(\bar{P}_k^{\varepsilon_k})\right\|,\left\|(S-S_k)\bar{U}_k^{\varepsilon_k}\right\|$$

has already been established. Therefore, let us tackle the term

$$\left\|P_k^{0+}\bar{\theta}_k^{\varepsilon_k}\right\|\min(\varepsilon_k^{2N(1-1/p')},\varepsilon_k^{3N/2}).$$

We observe that due to Theorem 3.2.5 and (4.1.25):

$$\frac{1}{\varepsilon_k}\left\|\bar{V}_k^{\varepsilon_k}\right\|^2\lesssim 1\ \Leftrightarrow\ \left\|\bar{\theta}_k^{\varepsilon_k}\right\|\lesssim\varepsilon_k^{-3/2}.$$

Lemma 4.2.5 then provides the following bound:

$$\left\|P_k^{0+}\bar{\theta}_k^{\varepsilon_k}\right\|\leq\left\|\bar{\theta}_k^{\varepsilon_k}\right\|\lesssim\varepsilon^{-3/2}.$$

Consequently,

$$\|P_k^{0+}\bar{\theta}_k^{\varepsilon_k}\| \min(\varepsilon_k^{2N(1-1/p')}, \varepsilon_k^{3N/2}) \lesssim \min(\varepsilon_k^{2N(1-1/p')-3/2}, \varepsilon_k^{\frac{3}{2}(N-1)}).$$

Condition (4.3.35) then ensures convergence.
In addition, (4.3.35) also implies (4.3.27), which in turn allows us to conclude harnessing Theorem 4.3.13:

$$\|\bar{\theta}_k^{\varepsilon_k} - P_k^{0+}(\bar{\theta}_k^{\varepsilon_k})\| \to 0,\ k \to \infty$$

Besides, (4.3.35) ensures that (4.3.24),(4.3.13) and (4.3.17) all hold. Theorem 4.3.11, in particular (4.3.25), in combination with Theorem 4.3.5 then immediately result in:

$$\begin{aligned} \min\Bigg(\|(\Pi(\bar{P}_k^{\varepsilon_k}) + \Delta y_c)^-\|, \\ \varepsilon_k^{-3N/p'}(\|\Pi(\bar{P}_k^{\varepsilon_k}) - \bar{U}_k^{\varepsilon_k}\| + |(S - S_k)\bar{U}_k^{\varepsilon_k}|_{H^1(\Omega)} \\ +|(\bar{Y}_k^{\varepsilon_k} - I_k y_c)^-|_{H^1(\Omega)} + \|y_c - I_k y_c\|_{H^1(\Omega)})\Bigg) \to 0. \end{aligned}$$

All in all, we thus have

$$\mathcal{E}_s(\bar{U}_k^{\varepsilon_k}, \bar{V}_k^{\varepsilon_k}) \to 0,\ k \to \infty$$

which completes the proof. □

We can now turn to the last section of this chapter in which we will derive residual error estimators for the linear error terms $(S - S_k)\cdot$ which still appear in the error estimators $\mathcal{E}_r$ and $\mathcal{E}_r$.

4.3.3 Convergent Residual Type Estimators for the Linear Errors

In this section we derive residual estimators for the linear errors in (4.2.32) and (4.2.25). For detailed information on definitions and further reading we refer to [2], Section 2.2. and [65]. To begin with, we make the following definition

Definition 4.3.15 (skeleton and jump residual)**.** *Let $\mathcal{T}_k$ be given. We define the* skeleton $\mathcal{S}_k$ *by*

$$\mathcal{S}_k := (\bigcup_{T \in \mathcal{T}_k} \partial T) \setminus \partial\Omega$$

For a given finite element function $V \in \mathbb{Y}_k$ *we define the* jump residual *by:*

$$[\![V]\!]_S := \mathbf{n}^+ \cdot \nabla V + \mathbf{n}^- \cdot \nabla V,$$

for all $S \in \mathcal{S}_k$*, where* $\mathbf{n}^+$ *and* $\mathbf{n}^-$ *are the outer unit normal vectors pointing towards* T^+ *and* T^- *respectively with* T^+ *and* T^- *being the elements meeting at the side* S*.*

In the ensuing theorem, we will now derive a reliable L_2- error estimator for the linear error in the adjoint state $(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)$.

Theorem 4.3.16. *Let* $\bar{q}$ *be as in Assumption* (CA2) *and* $\bar{q}'$ *be its dual exponent, i.e.* $\frac{1}{\bar{q}} + \frac{1}{\bar{q}'} = 1$*. Then the following bound is valid:*

$$\begin{aligned}
\big\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\big\|^2 \lesssim \Big(&\sum_{T \in \mathcal{T}_k} h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \big\|\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon\big\|_{L_2(T)}^{\bar{q}'} \\
&+ h_T^{\bar{q}'(\frac{3}{2}+\frac{d}{2}-\frac{d}{\bar{q}})} \big\|[\![\nabla \bar{P}_k^\varepsilon]\!]\big\|_{L_2(\partial T)}^{\bar{q}'} \Big)^{2/\bar{q}'} \\
&:= \mathbf{EP}_k^2(\bar{Y}_k^\varepsilon, \bar{\theta}_k^\varepsilon, \mathcal{T}_k, \bar{q}),
\end{aligned} \tag{4.3.36}$$

Proof. For the sake of abbreviation we use the notation:

$$\hat{p}^\varepsilon := S^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon).$$

Theorem 4.3.1 is crucial to the proof, it guarantees that for each $g \in L_2(\Omega)$ we obtain a $\psi \in W_{\bar{q}}^2(\Omega) \cap \mathring{W}_{\bar{p}}^1(\Omega)$, $\bar{p} > d$, and $\nabla\psi \in H(\mathrm{div}, \Omega)$ with $-\Delta\psi = g$ a.e. in Ω such that

$$\begin{aligned}
\big\|\hat{p} - \bar{P}_k^\varepsilon\big\| &= \sup_{\|g\|=1} (\hat{p}^\varepsilon - \bar{P}_k^\varepsilon, g) \\
&\lesssim \sup_{\|\Delta\psi\| \le 1} (\hat{p}^\varepsilon - \bar{P}_k^\varepsilon, -\Delta\psi).
\end{aligned}$$

up to a constant depending solely on $\|S\|_{\mathcal{L}(L_2(\Omega), W_{\bar{q}}^2(\Omega))}$.

We can now estimate the term on the right using Green's formula, Galerkin orthogonality (Lemma 4.1.12) and standard interpolation estimates for the Lagrange interpolant $I_k : \mathring{W}_{\bar{p}}^1(\Omega) \cap W_{\bar{q}}^2(\Omega) \to \mathbb{Y}_k$:

$$\begin{aligned}
(\hat{p}^\varepsilon - \bar{P}_k^\varepsilon, -\Delta\psi) &= (\nabla(\hat{p}^\varepsilon - \bar{P}_k^\varepsilon), \nabla\psi) \\
&= (\nabla(\hat{p}^\varepsilon - \bar{P}_k^\varepsilon), \psi - I_k\psi) \\
&= \sum_{T \in \mathcal{T}_k} (\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon, \psi - I_k\psi)_{L_2(T)} + \sum_{S \in \mathcal{S}_k} ([\![\bar{P}_k^\varepsilon]\!], \psi - I_k\psi)_{L_2(S)} \\
&\lesssim \sum_{T \in \mathcal{T}_k} |\psi|_{W_{\bar{q}}^2(T)} \big(h_T^{2+\frac{d}{2}-\frac{d}{\bar{q}}} \big\|\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon\big\|_{L_2(T)} + h_T^{\frac{3}{2}+\frac{d}{2}-\frac{d}{\bar{q}}} \big\|[\![\bar{P}_k^\varepsilon]\!]\big\|_{L_2(\partial T)}\big),
\end{aligned}$$

Using Hölder's inequality for sums, we can continue in the following fashion:

$$\begin{aligned}(\hat{p}^{\varepsilon} - \bar{P}_k^{\varepsilon}, -\Delta\psi) \leq \Big(\sum_{T \in \mathcal{T}_k} & h_T^{2\bar{q}' + \bar{q}'(\frac{d}{2} - \frac{d}{q})} \left\| \bar{Y}_k^{\varepsilon} - y_d - \bar{\theta}_k^{\varepsilon} \right\|_{L_2(\Omega)}^{\bar{q}'} \\ & + h_T^{\bar{q}'(\frac{3}{2} + \frac{d}{2} - \frac{d}{\bar{q}})} \left\| [\![\bar{P}_k^{\varepsilon}]\!] \right\|_{L_s(\partial T)}^{\bar{q}'} \Big)^{1/\bar{q}'} \\ & \cdot \Big(\sum_{T \in \mathcal{T}_k} |\psi|_{W_{\bar{q}}^2(T)}^{\bar{q}} \Big)^{1/\bar{q}} \\ & = \Big(\sum_{T \in \mathcal{T}_k} h_T^{\bar{q}'(2 + \frac{d}{2} - \frac{d}{\bar{q}})} \left\| \bar{Y}_k^{\varepsilon} - y_d - \bar{\theta}_k^{\varepsilon} \right\|_{L_2(\Omega)}^{\bar{q}'} \\ & + h_T^{\bar{q}'(\frac{3}{2} + \frac{d}{2} - \frac{d}{\bar{q}})} \left\| [\![\bar{P}_k^{\varepsilon}]\!] \right\|_{L_2(\partial T)}^{\bar{q}'} \Big)^{1/\bar{q}'} \cdot \|\psi\|_{W_{\bar{q}}^2(\Omega)}\end{aligned}$$

with $\frac{1}{\bar{q}'} + \frac{1}{\bar{q}} = 1$. Since $\|\psi\|_{W_{\bar{q}}^2(\Omega)} \lesssim 1$ we obtain the desired result by taking the supremum over all ψ on each side. □

As an easy consequence we obtain the following L_2 residual type estimator for the linear error in the state:

Theorem 4.3.17. *Let $\bar{q}$ be given and $\bar{q}'$ denote its dual exponent. Then the following bound is valid:*

$$\begin{aligned}\left\| (S - S_k)\bar{U}_k^{\varepsilon} \right\|^2 \lesssim & \Big(\sum_{T \in \mathcal{T}_k} h_T^{\bar{q}'(2 + \frac{d}{2} - \frac{d}{\bar{q}'})} \|U_k^{\varepsilon}\|_{L_2(T)}^{\bar{q}'} + \quad h_T^{\bar{q}'(\frac{3}{2} + \frac{d}{2} - \frac{d}{q})} \left\| [\![\nabla \bar{Y}_k^{\varepsilon}]\!] \right\|_{L_2(\partial T)}^{\bar{q}'} \Big)^{2/\bar{q}'} \\ & := \mathbf{EYL2}_k^2(\bar{U}_k^{\varepsilon}, \mathcal{T}_k, \bar{q}),\end{aligned} \tag{4.3.37}$$

Proof. The proof is an application of the techniques presented in the proof of Theorem 4.3.16. □

Finally, we present an H^1 residual type estimator. The proof of the bound given below can e.g. be found in [65], Section 6.

Theorem 4.3.18. *The following bound holds:*

$$\begin{aligned}|(S - S_k)\bar{U}_k^{\varepsilon}|_{H^1(\Omega)}^2 \lesssim & \sum_{T \in \mathcal{T}_k} h_T^2 \, \|U_k^{\varepsilon}\|_{L_2(T)}^2 + \quad h_T \left\| [\![\nabla \bar{Y}_k^{\varepsilon}]\!] \right\|_{L_2(\partial T)}^{\bar{q}'} \\ & := \mathbf{EYH1}_k^2(\bar{U}_k^{\varepsilon}, \mathcal{T}_k),\end{aligned} \tag{4.3.38}$$

We now conclude this chapter by demonstrating that $\mathbf{EY}_k^2(\bar{U}_k^{\varepsilon_k}, q)$ and $\mathbf{EY}_k^2(\bar{Y}_k^{\varepsilon_k}, \bar{\theta}_k^{\varepsilon_k}, \mathcal{T}_k, q)$ converge to 0 as $\varepsilon_k \to 0$ and $k \to \infty$ **irrespective of the refinement strategy** employed. The convergence is independent of the particular refinement strategy because we demanded that $h_k^{\max}$ tend to 0 a.e, something which we will use frequently in the proof of convergence. Needless to say, in the actual implementation one should opt for a smart marking strategy, e.g. one that at least captures the maximal error indicator, and not just a random one. In Section 5.3 we will present a maximum strategy adapted to our setting.

Theorem 4.3.19. *Let $\varepsilon_k \to 0$ and $k \to \infty$ such that condition* (4.3.5) *is fulfilled. Then*

$$\begin{aligned} \mathbf{EP}_k^2(\bar{Y}_k^{\varepsilon_k}, \bar{\theta}_k^{\varepsilon_k}, \mathcal{T}_k, q) &\to 0 \\ \mathbf{EYL2}_k^2(\bar{U}_k^{\varepsilon_k}, \mathcal{T}_k, q) &\to 0 \\ \mathbf{EYH1}_k^2(\bar{U}_k^{\varepsilon}, \mathcal{T}_k) &\to 0 \end{aligned}$$

Proof. First of all, $\mathbf{EYH1}_k^2(\bar{U}_k^{\varepsilon}, \mathcal{T}_k) \to 0$ is a consequence of the fact that for this estimator there exists a local lower bound up to oscillation. The detailed arguments and proofs can be looked up in [76],[60],[17] and [65], Sections 6 and 7.
Let us therefore tackle $\mathbf{EP}_k^2(\bar{Y}_k^{\varepsilon_k}, \bar{\theta}_k^{\varepsilon_k}, \mathcal{T}_k, \bar{q}) \to 0$. We recall the well known inequality for $1 \leq p \leq q < \infty$.

$$\big(\sum_n |a_n|^q\big)^{1/q} \leq \big(\sum_n |a_n|^p\big)^{1/p} \tag{4.3.39}$$

We can then estimate in the following way:

$$\begin{aligned} \sum_{T\in\mathcal{T}_k} h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \left\|\bar{Y}_k^{\varepsilon_k} - y_d - \bar{\theta}_k^{\varepsilon}\right\|_{L_2(T)}^{\bar{q}'} &\lesssim \sum_{T\in\mathcal{T}_k} h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{L_2(T)}^{\bar{q}'} \\ &\quad + h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \left\|\bar{\theta}_k^{\varepsilon_k}\right\|_{L_2(T)}^{\bar{q}'} \\ &\lesssim \sum_{T\in\mathcal{T}_k} h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{L_2(T)}^{\bar{q}'} \\ &\quad + h_T^{\bar{q}'(2-\frac{d}{\bar{q}})} \left\|\bar{\theta}_k^{\varepsilon}\right\|_{L_1(T)}^{\bar{q}'}. \end{aligned}$$

In the last line, we have used inverse estimates, compare [14], Section 4.5.
For the first term in the sum above, we can then proceed in the following way:

$$\begin{aligned} \sum_{T\in\mathcal{T}_k} h_T^{\bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}})} \left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{L_2(T)}^{\bar{q}'} &\leq h_{\max}^{\gamma}\Big(\sum_{T\in\mathcal{T}_k} \left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|_{L_2(T)}^{2}\Big)^{\bar{q}'/2} \\ &\leq h_{\max}^{\gamma} \left\|\bar{Y}_k^{\varepsilon_k} - y_d\right\|^{\bar{q}'} \end{aligned}$$

with $\gamma = \bar{q}'(2+\frac{d}{2}-\frac{d}{\bar{q}}) > 0$, after all $\bar{q} > \frac{d}{2}$. Hence, the term converges to 0, after all, $h_k^{\max} \to 0$. The same estimates (with different exponents) can be made for the term

$$\sum_{T\in\mathcal{T}_k} h_T^{\bar{q}'(2-\frac{d}{\bar{q}})} \left\|\bar{\theta}_k^{\varepsilon}\right\|_{L_1(T)}^{\bar{q}'}.$$

Taking advantage of Assumption CA5, we can conclude convergence, too.

Let us now tackle the jump residual. For any side S, where two elements $T^+, T^- \in \mathcal{T}_k$ meet we can deduce using the fact that the gradient is piecewise constant on each $T \in \mathcal{T}_k$:

$$\begin{aligned}\int\limits_S |\nabla \bar{P}_k^{\varepsilon_k} \cdot n^+ + \nabla \bar{P}_k^{\varepsilon_k} \cdot n^-|^2 &\lesssim \int\limits_S |\nabla \bar{P}_k^{\varepsilon_k}|_{T^+}|^2 + |\nabla \bar{P}_k^{\varepsilon_k}|_{T^-}|^2 \\ &= |S|(|\nabla \bar{P}_k^{\varepsilon_k}|_{T^+}|^2 + |\nabla \bar{P}_k^{\varepsilon_k}|_{T^-}|^2) \\ &= \frac{|S|}{|T^+|} \int\limits_{T^+} |\nabla \bar{P}_k^{\varepsilon_k}|_{T^+}|^2 + \frac{|S|}{|T^-|} \int\limits_{T^-} |\nabla \bar{P}_k^{\varepsilon_k}|_{T^-}|^2 \\ &\lesssim h_{T^+}^{-1} \left\| \nabla P_k^{\varepsilon_k} \right\|_{L_2(T^+)}^2 + h_{T^-}^{-1} \left\| \nabla P_k^{\varepsilon_k} \right\|_{L_2(T^-)}^2 .\end{aligned}$$

This in turn implies:

$$h_T^{\bar{q}'(\frac{3}{2}+\frac{d}{2}-\frac{d}{\bar{q}})} \left\| [\![\bar{P}_k^{\varepsilon_k}]\!] \right\|_{L_2(S)}^{\bar{q}'} \lesssim \sum_{T \in \{T^+, T^-\}} h_T^{\bar{q}'(1+\frac{d}{2}-\frac{d}{\bar{q}})} \left\| \nabla \bar{P}_k^{\varepsilon_k} \right\|_{L_2(T)}^{\bar{q}'} .$$

At this stage we need to stress that $\nabla \bar{P}_k^{\varepsilon_k}$ is in general not bounded uniformly in L_2. However, completely analogous to (4.3.12), we can conclude that

$$\varepsilon_k^{-3/p'} |\bar{P}_k^{\varepsilon_k}|_{H^1(\Omega)} \lesssim 1$$

which implies the following estimates for the jump residual, compare also (4.3.39):

$$\begin{aligned}\sum_{T \in \mathcal{T}_k} h_T^{\bar{q}'(\frac{3}{2}+\frac{d}{2}-\frac{d}{\bar{q}})} \left\| [\![\bar{P}_k^{\varepsilon_k}]\!] \right\|_{L_2(\partial T)}^{\bar{q}'} &\lesssim \Big(\sum_{T \in \mathcal{T}_k} h_T^{2(1+\frac{d}{2}-\frac{d}{q})} |\bar{P}_k^{\varepsilon_k}|_{H^1(T)}^2 \Big)^{\bar{q}'/2} \\ &\lesssim (h_k^{\max})^{\bar{q}'(1+\frac{d}{2}-\frac{d}{\bar{q}})} |\bar{P}_k^{\varepsilon_k}|_{H^1(\Omega)}^{\bar{q}'} \\ &\lesssim (h_k^{\max})^{\bar{q}'(1+\frac{d}{2}-\frac{d}{\bar{q}})} \varepsilon^{-3\bar{q}'/p'} \\ &\to 0\end{aligned}$$

thanks to condition (4.3.5).

The result $\mathbf{EYL2}_k^2(\bar{U}_k^{\varepsilon_k}, \mathcal{T}_k, \bar{q}) \to 0$ is a straightforward application of the arguments made for the term $\mathbf{EP}_k^2(\bar{Y}_k^{\varepsilon_k}, \bar{\theta}_k^{\varepsilon_k}, \mathcal{T}_k, \bar{q})$ and is valid regardless of condition (4.3.5). □

Chapter 5

The Adaptive Algorithm

In this chapter we will describe the adaptive algorithm used for the numerical experiments evaluated in Chapter 6. Theoretically, the framework is that of the previous chapter, Chapter 4, specifically we deal with the model problem (CMP), its regularisation (CMP^{ε}) and the corresponding discrete problems (DMP_k) and (DMP_k^{ε}). Structurally, we will follow the adaptive loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}$$

and explain the different modules in detail on the way. At the end of this chapter we will then present our adaptive algorithm in a compact way, Algorithm 5.4.1.
An crucial aspect of this chapter will be the **localisation** of the estimators derived in Theorem 4.2.12 and Theorem 4.2.13. We will describe how to obtain a localisable error estimator in detail in the 'ESTIMATE' and 'MARK' sections, Sections 5.2 and 5.3.

The reader should note that in this Chapter we focus solely on the **full discretisation technique**, i.e.

$$\mathbb{U}_k = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega)),$$

as this is the discretisation method we employed for our numerical experiments.

5.1 'SOLVE'

In this section we will explain the optimisation algorithm used to compute a solution to (DMP_k^{ε}) which - as usual - is denoted by $(\bar{U}_k^{\varepsilon}, \bar{V}_k^{\varepsilon})$. The optimisation method which we use is the primal-dual active set strategy (PDAS) which has been proven to be equivalent to a semismooth Newton method introduced on the continuous level in Section 2.2.3, cf [41]. This method ensures fast convergence, which is also demonstrated by our results in Chapter 6 and theoretically underpinned by Theorem 2.2.16. Before taking the reader step by step through the algorithm, we first have to lay some notational groundwork, which revolves around transferring the equations and inequalities of the discrete KKT system (4.1.25) expressed in terms of discrete functions into a matrix vector equation setting.

5.1.1 Deriving a Matrix-Vector Setting

First recall that the finite element spaces for the control and state in the full discretisation setting of Chapter 4 were given by

$$\begin{aligned}\mathbb{U}_k &= \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_0, L_2(\Omega))\\ \mathbb{Y}_k &= \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, \mathring{H}^1(\Omega)),\end{aligned}$$

i.e piecewise constant functions for the control and piecewise linear finite elements for the state. Here, $\mathcal{T}_k$ is a shape-regular, conforming triangulation of Ω.
For a fixed k we introduce bases for both spaces, first for the control space.

$$\mathbb{U}_k = \operatorname{span}\left\{\psi_k^1, \psi_k^2, ...\psi_k^{n_u}\right\}.$$

Numbering the elements of the triangulation $\mathcal{T}_k$ by $E_1, E_2, ..., E_{n_u}$, we choose a basis for $\mathbb{U}_k$ by defining the basis functions ψ in the following way:

$$\psi_k^i|_{E_j} = \delta_{ij} = \chi_{E_i}, \tag{5.1.1}$$

where δ_{ij} is the Kronecker symbol and χ_{E_i} the characteristic function of the element E_i.
This allows us to expand functions in $\mathbb{U}_k$, i.e. $U \in \mathbb{U}_k$ can be written as

$$U = \sum_{i=1}^{n_u} u^i \psi_k^i$$

with the coefficient vector $\mathbf{u} = (u^i)_{i=1}^{n_u}$. Thus, a function $U \in \mathbb{U}_k$ is uniquely determined by its associated coeffcient vector $\mathbf{u}$. Notationally, we will stick to a bold face notation for the corresponding vector of coefficients of a discrete function.
An important consequence of (5.1.1) is that

$$a \leq U \leq b \;\Leftrightarrow\; a \leq u^i \leq b \;\forall i \in \{1, ..., n_u\}.$$

Consequently,

$$\mathcal{U}_k \Leftrightarrow \left\{\mathbf{u} \in \mathbb{R}^{n_u} \;:\; a \leq u^i \leq b\right\} =: \mathcal{U}_k(\mathbb{R}^{n_u}),$$

where the $\Leftrightarrow$ is understood in the sense that every $U \in \mathcal{U}_k$ is uniquely associated with a coefficient vector $\mathbf{u} \in \mathcal{U}_k(\mathbb{R}^{n_u})$ and vice versa. This will be highly useful for the PDAS algorithm of the SOLVE module.

To construct a basis for the state space $\mathbb{Y}_k$ we first introduce the larger finite element space $\mathbb{Y}_k^f$ defined by

$$\mathbb{Y}_k^f = \mathbf{FES}(\mathcal{T}_k, \mathbb{P}_1, H^1(\Omega)).$$

Denoting the set of nodes of the triangulation $\mathcal{T}_k$ by $\mathcal{N}_k$, compare Definition 2.3.2, we define a basis for $\mathbb{Y}_k^f$ by setting for all $n_j \in \mathcal{N}_k$

$$\varphi_k^i(n_j) = \delta_{ij},$$

where δ_{ij} is again the Kronecker symbol.

Then $\mathbb{Y}_k^f$ possesses the following basis:

$$\mathbb{Y}_k = \operatorname{span}\left\{\varphi_k^1, \varphi_k^2, ...\varphi_k^{n_y}\right\} \subset H^1(\Omega).$$

Presently, for the basis of $\mathbb{Y}_k$ we now take those φ_k^i which vanish on the boundary $\partial\Omega$:

$$\mathring{\mathbb{Y}}_k = \operatorname{span}\left\{\varphi_k^1, ..., \varphi_k^{n_d}\right\} \subset \mathring{H}^1(\Omega),\ n_d < n_y.$$

The basis expansion allows us to write the discretised state equation $S_k \bar{U}_k^\varepsilon = \bar{Y}_k^\varepsilon$ and the adjoint equation $S_k^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon) = \bar{P}_k^\varepsilon$ in the following matrix vector fashion:

$$\mathbf{A}\mathbf{y}_k^\varepsilon = \mathbf{M}^{SC}\mathbf{u}_k^\varepsilon \tag{5.1.2}$$

for the state equation and

$$\mathbf{A}\mathbf{p}_k^\varepsilon = \mathbf{M}^{SS}(\mathbf{y}_k^\varepsilon - \frac{1}{\varepsilon^2}\mathbf{v}_k^\varepsilon) - \mathbf{y_d}, \tag{5.1.3}$$

for the adjoint equation, where we also used the relation $\frac{1}{\varepsilon^2}\bar{V}_k^\varepsilon = \bar{\theta}_k^\varepsilon$ of the discrete KKT system (4.1.25). As before, the bold face letters denote the coefficient vectors to the discrete function $\bar{Y}_k^\varepsilon, \bar{P}_k^\varepsilon, \bar{U}_k^\varepsilon$. For the stiffness matrix $\mathbf{A} = (a_{ij})_{ij}^{n_y}$ we have

$$a_{ij} = \begin{cases} = (\nabla\varphi_k^i, \nabla\varphi_k^j) & 1 \le i \le n_y,\ 1 \le j \le n_d \\ = \delta_{ij} & i > n_d. \end{cases}$$

The mass matrices $\mathbf{M}^{SS}$ and $\mathbf{M}^{SC}$ are given by

$$\begin{aligned}\mathbf{M}^{SS} &= ((\varphi_k^i, \varphi_k^j))_{i,j}^{n_y}\\ \mathbf{M}^{SC} &= ((\varphi_k^i, \psi_k^j))_{i,j} \in \mathbb{R}^{n_y \times n_u}.\end{aligned}$$

The vector $\mathbf{y_d} = (y_d^i)_i^{n_y}$ is given by

$$(y_d^i) = \begin{cases} (y_d, \varphi_k^i) & 1 \leq i \leq n_d \\ 0 & \text{else.} \end{cases}$$

We now intend to rewrite the optimality condition for $\bar{U}_k^\varepsilon$ and $\bar{V}_k^\varepsilon$, (4.1.25) in terms of the associated coefficient vectors $\mathbf{u}_k^\varepsilon$ and $\mathbf{v}_k^\varepsilon$ starting with the optimality condition for $\bar{U}_k^\varepsilon$. Basis expansion allows us to write the optimality condition for $\bar{U}_k^\varepsilon$ in (4.1.25),

$$(\bar{P}_k^\varepsilon + \nu \bar{U}_k^\varepsilon, U - \bar{U}_k^\varepsilon) \geq 0 \ \ \forall U \in \mathcal{U}_k, \tag{5.1.4}$$

in an indicewise way:
Introducing an additional mass matrix $\mathbf{M}^{CC}$ defined by

$$\mathbf{M}^{CC} = ((\psi_k^i, \psi_k^j))_{i,j},$$

we claim that (5.1.4) is equivalent to the following indicewise projection formula

$$\bar{u}^{\varepsilon,i} = \min(b, \max(a, \bar{u}^{\varepsilon,i} - ((\mathbf{M}^{SC})^t \mathbf{p}_\mathbf{k}^\varepsilon)^i - (\nu \mathbf{M}^{CC} \mathbf{u}_\mathbf{k}^\varepsilon)^i). \tag{5.1.5}$$

The proof of this assertion is given in the ensuing lemma:

Lemma 5.1.1. *The optimality conditions in the form* (5.1.4) *and* (5.1.5) *are equivalent.*

Proof. As an intermediate step we observe that (5.1.4) can be reformulated in the following way by simply writing the involved functions $\bar{U}_k^\varepsilon, \bar{P}_k^\varepsilon$ in their basis representation:

$$(u^i - \bar{u}^{\varepsilon,i})((\mathbf{M}^{SC})^t \mathbf{p}_k^\varepsilon + \nu \mathbf{M}^{CC} \mathbf{u}_k^\varepsilon)^i \geq 0 \ \forall 1 \leq i \leq n_u, \ u^i \in \mathbb{R}, \ a \leq u^i \leq b. \tag{5.1.6}$$

Investigating the three cases $\bar{u}^{\varepsilon,i} = a$, $a < \bar{u}^{\varepsilon,i} < b$ and $\bar{u}^{\varepsilon,i} = b$ we immediately deduce:

$$((\mathbf{M}^{SC})^t \mathbf{p}_k^\varepsilon + \nu \mathbf{M}^{CC} \mathbf{u}_k^\varepsilon)^i \begin{cases} = 0 & \text{if } a < \bar{u}^{\varepsilon,i} < b \\ \geq 0 & \text{if } \bar{u}^{\varepsilon,i} = a \\ \leq 0 & \text{if } \bar{u}^{\varepsilon,i} = b \end{cases}$$

Now, (5.1.5) readily follows.

Conversely, (5.1.5) implies (5.1.6) and thus in turn (5.1.4). □

The relation (5.1.5) constitutes the bedrock of our PDAS algorithm to solve the optimal control problem it being the matrix-vector equivalent of the projection formula of Theorem 4.1.16. At this stage, we also want to stress that it was precisely such a projection formula which lay at the heart of the semismooth Newton method on the continuous level presented in Section 2.2.3.

As we will shortly discover, the algorithm hinges on sets of active or inactive indices A, which are subsets of either $\{1,...,n_u\}$ or $\{1,...,n_y\}$, and to which are associated matrices $\mathbf{P}_A = (p_{ij})_{ij}$

$$p_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } i \in A \\ 0 & \text{else.} \end{cases} \tag{5.1.7}$$

Defining

$$\begin{aligned} A_u^a &:= \big\{i \in \{1,...,n_u\} \ : \ \bar{u}^{\varepsilon,i} = a\big\} \\ A_u^b &:= \big\{i \in \{1,...,n_u\} \ : \ \bar{u}^{\varepsilon,i} = b\big\} \\ I_u &:= \{1,...,n_u\} \setminus (A_u^a \cup A_u^b) \end{aligned}$$

with associated matrices $\mathbf{P}_{A_u^a}, \mathbf{P}_{A_u^b}$ and $\mathbf{P}_{I_u}$, (5.1.5) can be expressed by the equation

$$\mathbf{u}_k^\varepsilon = \mathbf{P}_{I_u}(-(\mathbf{M}^{CS})^t\mathbf{p}_k^\varepsilon + (\mathbf{I} - \nu\mathbf{M}^{CC})\mathbf{u}_k^\varepsilon) + \mathbf{P}_{A_u^b}(\mathbf{b}) + \mathbf{P}_{A_u^a}(\mathbf{a}), \tag{5.1.8}$$

where $\mathbf{a}$ and $\mathbf{b}$ are vectors of a's and b's of length n_u respectively and $\mathbf{I}$ denotes the identity matrix. The rather forced relation (5.1.8) is important for understanding the steps of the PDAS properly, which we will present in a compact way at the end of this section. Before, though, let us first derive a relation similar to (5.1.5) and (5.1.8) for the virtual control $\bar{V}_k^\varepsilon$. Employing the usual notation for the coefficient vectors associated with $I_k y_c$, $\bar{Y}_k^\varepsilon$, $\bar{V}_k^\varepsilon$ and $\bar{\theta}_k^\varepsilon$, we derive from (4.1.25):

$$\bar{v}^{\varepsilon,i} = \frac{1}{\varepsilon^2}\bar{\theta}^{\varepsilon,i}. \tag{5.1.9}$$

The condition $\bar{\theta}_k^\varepsilon \in C_{\mathbb{V}_k}^-$ can be transferred to a matrix vector setting by demanding

$$(\mathbf{M}^{SS}\boldsymbol{\theta}_k^\varepsilon)^i \geq 0 \ \forall i = 1,...,n_y.$$

or equivalently, compare (5.1.9)

$$(\mathbf{M}^{SS}\mathbf{v}_k^\varepsilon)^i \geq 0 \ \forall i = 1,...,n_y$$

Standard nonlinear programming theory guarantees that the slackness equation in (4.1.25) can be reformulated with the familiar min-NCP-function in the following way:

$$\min\left((\mathbf{M}^{SS}\mathbf{v}_k^\varepsilon)^i, I_k y_c^i - \bar{y}^{\varepsilon,i} - \varepsilon\bar{v}^{\varepsilon,i}\right) = 0 \;\forall i = 1, ..., n_y$$

Defining an active and an inactive set by

$$\begin{aligned} A_v &:= \left\{ i \in \{1, ..., n_y\} \;:\; \bar{y}^{\varepsilon,i} + \varepsilon\bar{v}^{\varepsilon,i} - \frac{1}{\varepsilon^2}(\mathbf{M}^{SS}\mathbf{v}_k^\varepsilon)^i < I_k y_c^i \right\} \\ I_v &:= \{1, ..., n_y\} \setminus A_v, \end{aligned}$$

we obtain a relation analogous to (5.1.8)

$$\mathbf{P}_{I_v}\frac{1}{\varepsilon^2}\mathbf{M}^{SS}\mathbf{v}_k^\varepsilon + \mathbf{P}_{A_v}\varepsilon\mathbf{v}_k^\varepsilon = -\mathbf{P}_{A_v}(\mathbf{y}_k^\varepsilon + \mathbf{I_k y_c}), \tag{5.1.10}$$

Combining (5.1.8),(5.1.10), (5.1.2) and (5.1.3) we realise that the discrete KKT system (4.1.25) is equivalent to the following system of matrix vector equations.

$$\begin{pmatrix} \mathbf{A} & \mathbf{M}^{SC} & 0 & 0 \\ 0 & \mathbf{I} - \mathbf{P}_{I_u}(\mathbf{I} - \nu\mathbf{M}^{CC}) & \mathbf{P}_{I_u}(\mathbf{M}^{SC})^t & 0 \\ -\mathbf{M}^{SS} & 0 & \mathbf{A} & \frac{1}{\varepsilon^2}\mathbf{M}^{SS} \\ \mathbf{P}_{A_v} & 0 & 0 & \mathbf{P}_{I_v}\frac{1}{\varepsilon^2}\mathbf{M}^{SS} + \mathbf{P}_{A_v}\varepsilon \end{pmatrix} \begin{pmatrix} \mathbf{y}_k^\varepsilon \\ \mathbf{u}_k^\varepsilon \\ \mathbf{p}_k^\varepsilon \\ \mathbf{v}_k^\varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{P}_{A_u^a}\mathbf{a} + \mathbf{P}_{A_u^b}\mathbf{b} \\ -\mathbf{y_d} \\ \mathbf{P}_{A_v}\mathbf{I_k y_c}. \end{pmatrix} \tag{5.1.11}$$

We can now write down the PDAS algorithm in a compact way:

5.1.2 PDAS Algorithm

(5.1.11) constitutes the core of the PDAS - algorithm through which we will now take the reader:

Algorithm 5.1.1 PDAS

```
1: Set $n = 0$.
2: Choose sets of active and inactive indices $\mathcal{A}_u^{a,n}, \mathcal{A}_u^{b,n}, \mathcal{I}_u^n \subset \{1, ..., n_u\}$, $\mathcal{I}_u^n = \{1, ..., n_u\} \setminus (\mathcal{A}_u^{a,n} \cup \mathcal{A}_u^{b,n})$, and $\mathcal{A}_v^n, \mathcal{I}_v^n \subset \{1, ..., n_y\}$, $\mathcal{I}_v^n = \{1, ..., n_u\} \setminus \mathcal{A}_v^n$.
3: Build the associated matrices $\mathbf{P}_{\mathcal{A}_u^{a,n}}, ...$ according to (5.1.7).
4: Solve
   $$\begin{pmatrix} \mathbf{A} & \mathbf{M}^{SC} & 0 & 0 \\ 0 & \mathbf{I} - \mathbf{P}_{\mathcal{I}_u^n}(\mathbf{I} - \nu \mathbf{M}^{CC}) & \mathbf{P}_{\mathcal{I}_u^n}(\mathbf{M}^{SC})^t & 0 \\ -\mathbf{M}^{SS} & 0 & \mathbf{A} & \frac{1}{\varepsilon^2}\mathbf{M}^{SS} \\ \mathbf{P}_{\mathcal{A}_v^n} & 0 & 0 & \mathbf{P}_{\mathcal{I}_v^n} \frac{1}{\varepsilon^2}\mathbf{M}^{SS} + \mathbf{P}_{\mathcal{A}_v^n}\varepsilon \end{pmatrix} \begin{pmatrix} \mathbf{y}^n \\ \mathbf{u}^n \\ \mathbf{p}^n \\ \mathbf{v}^n \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{P}_{\mathcal{A}_u^{a,n}}\mathbf{a} + \mathbf{P}_{\mathcal{A}_u^{b,n}}\mathbf{b} \\ -\mathbf{y_d} \\ \mathbf{P}_{\mathcal{A}_v^n}\mathbf{I_k}\mathbf{y_c}. \end{pmatrix}$$
5: Compute the new active and inactive sets for the control $\mathcal{A}_u^{a,n+1}, \mathcal{A}_u^{b,n+1}, \mathcal{I}_u^{n+1}$ according to the ensuing formula
   $$u^{i,n} - ((\mathbf{M}^{SC})^t\mathbf{p}^n + \nu\mathbf{M}^{CC}\mathbf{u}^n)^i \begin{cases} < a & \Rightarrow i \in \mathcal{A}_u^{a,n+1} \\ > b & \Rightarrow i \in \mathcal{A}_u^{b,n+1} \\ \text{else} & \Rightarrow i \in \mathcal{I}_u^{n+1}. \end{cases}$$
6: Compute the new active and inactive set for the virtual control $\mathcal{A}_v^{n+1}, \mathcal{I}_v^{n+1}$ in the following fashion
   $$y^{n,i} + \varepsilon v^{n,i} - \frac{1}{\varepsilon^2}\mathbf{M}^{SS}\mathbf{v}^n \begin{cases} < I_k y_c^i & \Rightarrow i \in \mathcal{A}_v^{n+1} \\ \text{else} & \Rightarrow i \in \mathcal{I}_v^{n+1}. \end{cases}$$
7: If $\mathcal{A}_u^{a,n+1} = \mathcal{A}_u^{a,n}$, $\mathcal{A}_u^{b,n+1} = \mathcal{A}_u^{b,n}$ and $\mathcal{A}_v^{n+1} = \mathcal{A}_v^n$ then $\mathbf{u}^n, \mathbf{v}^n$ is optimal, since it solves (5.1.11).
   Else Go to Step 3.
```

This algorithm can be interpreted as a semismooth Newton method, see. e.g [41]. Convergence properties are also investigated in this paper. Let us finish this section with a remark on computational aspects:

Remark 5.1.2. *From a computational point of view it is often advantageous, especially for small regularisation parameters ε, to use lumped masses instead of the mass matrix $\mathbf{M}_{SS}$ in* *the virtual control equation in the KKT system. We used lumped masses for regularisation parameters* $\varepsilon < 0.02$ *for both the smooth example, Section 6.1, and the Dirac example, Section 6.2.*

We can now turn to the 'ESTIMATE' module.

5.2 'ESTIMATE'

In this and the next section we will explain how to turn the error estimator of Theorem 4.2.13, which still contains linear errors $(S - S_k)\cdot$ into a localisable, numerically evaluable estimator, i.e. an estimator which only contains known discrete and continuous functions. To be more precise, in this and the next section we will derive quantities $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, $\mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ which **contain only known continuous and discrete functions** such that

$$\begin{aligned}
\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) &\lesssim \mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \sum_{T\in\mathcal{T}_k} (\mathfrak{e}_k^r)^2(T)\\
\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) &\lesssim \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)\\
\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)^2 &\lesssim \hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \sum_{T\in\mathcal{T}_k} (\mathfrak{e}_k^s)^2(T)
\end{aligned}$$

with local indicators $(\mathfrak{e}_k^r)^2(T)$ and $(\mathfrak{e}_k^s)^2(T)$.
First of all, let us recall the estimator for the full discretisation derived in Theorem 4.2.13, compare also Theorem 4.2.15 and Remark 4.2.16. In this section, we will stop explicitly stating the generic constants $s(\tau)$ and $c(p')$ as well as $\|S\|$ and fix $\nu = 1$:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 + \frac{8}{\varepsilon^N}\left\|\bar{v}^{\varepsilon^N}\right\|^2 \lesssim \varepsilon^{\gamma N} + \mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \tag{5.2.1}$$

with

$$\begin{aligned}
\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = {}& 4\left\|(S - S_k)\bar{U}_k^\varepsilon\right\|^2 + 6\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2\\
&+ 4\left\|(S - S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2\\
&+ 8(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) + 8(\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)\\
&+ 8(\bar{\theta}_k^\varepsilon, I_k y_c - y_c)
\end{aligned} \tag{5.2.2}$$

and

$$\begin{aligned}\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = 8 &\left\|\bar{Y}_k^\varepsilon - y_d\right\| \left(\left\|(S-S_k)\bar{U}_k^\varepsilon\right\| + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|\right)\\ &+ 8\left\|\Pi(\bar{P}_k^\varepsilon)\right\| \left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\| + 8\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\|\\ &+ 8\left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\| \min(\varepsilon^{2N(1-1/p'),\varepsilon^{3N/2}})\\ &+ 8\min\Big\{ \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, \varepsilon^{-3N/p'}\big(|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)} + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|\\ &+ \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\big)\Big\}.\end{aligned} \tag{5.2.3}$$

As we have already mentioned, this estimator has the drawback that it cannot be localised on each element. Let us expound on this point a bit: $\mathcal{E}_r$ does not pose any problem in this aspect, since it solely contains squared L_2-norms and scalar products which can be evaluated on each element $T \in \mathcal{T}_k$ and then added to obtain the bound $\mathcal{E}_r$. In addition, the local indicator is stored on the element as an indicator for the 'MARK'-procedure.

However, $\mathcal{E}_s$ cannot be localised in this manner, because it contains solely norms not squared norms. This is not an issue if one just wants to compute the global error estimator $\mathcal{E}_s$, yet, we do not get any quantities on each T which can be interpreted as local error indicators as the foundation for a marking algorithm. Therefore, the basic idea now is to square $\mathcal{E}_s$ to get squared norms - which can then be computed on each $T \in \mathcal{T}_k$ and stored on each T to get local indicators for the marking algorithm. This is the subject of the next theorem:

Theorem 5.2.1. *For $\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ we have the following estimate:*

$$\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \leq (\hat{\mathcal{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon))^{1/2}$$

with

$$\begin{aligned}\hat{\mathcal{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) := 48 &\left\|\bar{Y}_k^\varepsilon - y_d\right\|^2 \left(\left\|(S-S_k)\bar{U}_k^\varepsilon\right\|^2 + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2\right)\\ &+ 48\left\|\Pi(\bar{P}_k^\varepsilon)\right\|^2 \left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2 + 48\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2\\ &+ 48\left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2 \min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N})\\ &+ 48\min\Big\{ \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|^2, 4\varepsilon^{-6N/p'}\big(|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)}^2\\ &+ \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2 + \|y_c - I_k y_c\|_{H^1(\Omega)}^2 + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}^2\big)\Big\}\end{aligned} \tag{5.2.4}$$

Proof. To prove the bound (5.2.4), we merely square $\mathcal{E}_s$ and repeatedly apply Young's inequality, (4.2.2) and then draw the root:

Since $\mathcal{E}_s \geq 0$ we know that $(\mathcal{E}_s^2)^{1/2} = \mathcal{E}_s$. Squaring $\mathcal{E}_s$ and using Lemma 4.2.1, we obtain:

$$\begin{aligned}\frac{1}{48}(\mathcal{E}_s)^2 \leq{}& \left\|\bar{Y}_k^\varepsilon - y_d\right\|^2 \left(\left\|(S-S_k)\bar{U}_k^\varepsilon\right\|^2 + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2\right)\\ &+ \left\|\Pi(\bar{P}_k^\varepsilon)\right\|^2 \left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2 + \left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2\\ &+ \left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2 \min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N})\\ &+ \Bigg(\min\Bigg\{\left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, \varepsilon^{-3N/p'}\big(|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)}\\ &+ \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\| + \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\big)\Bigg\}\Bigg)^2.\end{aligned} \tag{5.2.5}$$

We now merely have to tackle the min term, all other terms already appear in (5.2.4). Setting

$$\begin{aligned}\mathcal{X} &:= \left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|\\ \mathcal{Y} &:= \varepsilon^{-3N/p'}\big(|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)}\\ &\quad + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\| + \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\big)\end{aligned}$$

we deduce employing standard properties of the min operator

$$\mathcal{Z}^2 := (\min(\mathcal{X}, \mathcal{Y}))^2 = \min(\mathcal{X}^2, \mathcal{Y}^2).$$

Let us now as the final step estimate $\mathcal{Y}^2$. Harnessing once again Lemma 4.2.1, we gain:

$$\begin{aligned}\mathcal{Y}^2 \leq{}& 4\varepsilon^{-6N/p'}\big(|(S-S_k)\bar{U}_k^\varepsilon|^2_{H^1(\Omega)}\\ &+ \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2 + \|y_c - I_k y_c\|^2_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|^2_{H^1(\Omega)}\big).\end{aligned}$$

The estimates for $\mathcal{Z}$ and $\mathcal{Y}$ can now be inserted in (5.2.5) to get the desired result. □

The quantities $\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon), \mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\hat{\mathcal{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ still contain linear errors $(S-S_k)\cdot$ which need to be estimated. This will be done by the residual-type error estimators of Section 4.3.3. Let us thus therefore briefly recapitulate them:

5.2.1 Estimators for the Linear Errors

Let us for simplicity assume that Ω is regular enough to admit an $H^2(\Omega)$ solution, i.e. for $\bar{q}$ from Theorem 4.3.1 we have $\bar{q} = 2$.
Theorem 4.3.17 then provides an L_2-residual estimator which gives the following global upper

bound for the linear error in the state $\left\|(S-S_k)\bar{U}_k^\varepsilon\right\|^2$:

$$\begin{aligned}\left\|(S-S_k)\bar{U}_k^\varepsilon\right\|^2 &\lesssim \sum_{T\in\mathcal{T}_k} h_T^4\left\|\bar{U}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T^3\left\|[\![\bar{Y}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2 \\ &:= \mathbf{EYL2}_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, \mathcal{T}_k)\end{aligned} \tag{5.2.6}$$

The global upper bound $\mathbf{EYL2}_k^2(\bar{U}_k^\varepsilon, \mathcal{T}_k)$ is the sum of local indicators $EYL2_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T)$, compare (5.2.6)

$$\mathbf{EYL2}_k^2 = \mathbf{EYL2}_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, \mathcal{T}_k) = \sum_{T\in\mathcal{T}_k} EYL2_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T), \tag{5.2.7}$$

where

$$EYL2_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T) = EYL2_k^2 := h_T^4\left\|\bar{U}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T^3\left\|[\![\nabla\bar{Y}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2$$

The local indicators $EYL2_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T)$ can then be used as the basis of a marking strategy as explained in the next section. That is why a **localisation** of the type (5.2.7) and (5.2.6) is so crucial.

For the $H^1(\Omega)$-error we recall Theorem 4.3.18 to gain the following upper bound:

$$\begin{aligned}|(S-S_k)\bar{U}_k^\varepsilon|_{H^1(\Omega)}^2 &\lesssim \sum_{T\in\mathcal{T}_k} h_T^2\left\|\bar{U}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T\left\|[\![\nabla\bar{Y}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2 \\ &:= \mathbf{EYH1}_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, \mathcal{T}_k),\end{aligned} \tag{5.2.8}$$

Again, as in (5.2.7) we can localise $\mathbf{E}_{y,1}^2$ in the following way:

$$\mathbf{EYH1}_k^2 = \mathbf{EYH1}_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, \mathcal{T}_k) = \sum_{T\in\mathcal{T}_k} EYH1_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T) \tag{5.2.9}$$

with

$$EYH1_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T) = EYH1_k^2 := h_T^2\left\|\bar{U}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T\left\|[\![\nabla\bar{Y}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2.$$

The local indicators $EYH1_k^2(\bar{U}_k^\varepsilon, \bar{Y}_k^\varepsilon, T)$ can then again be used as the foundation for a marking strategy.

The linear error which remains to be dealt with is the one in the adjoint state:

$$\left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2.$$

Theorem 4.3.16 supplies the global upper bound

$$\begin{aligned}\left\|(S-S_k)^*(\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon)\right\|^2 &\lesssim \sum_{T\in\mathcal{T}_k} h_T^4 \left\|\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T^3 \left\|[\![\nabla \bar{P}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2 \\ &:= \mathbf{EP}_k^2(\bar{Y}_k^\varepsilon, \bar{\theta}_k^\varepsilon, \bar{P}_k^\varepsilon, \mathcal{T}_k).\end{aligned} \tag{5.2.10}$$

As before, we can localise (5.2.10) in the following way:

$$\mathbf{EP}_k^2 = \mathbf{EP}_k^2(\bar{P}_k^\varepsilon, \bar{Y}_k^\varepsilon, \bar{\theta}_k^\varepsilon, \mathcal{T}_k) = \sum_{T\in\mathcal{T}_k} EP_k^2(\bar{Y}_k^\varepsilon, \bar{\theta}_k^\varepsilon, \bar{P}_k^\varepsilon, T) \tag{5.2.11}$$

with

$$EP_k^2(\bar{Y}_k^\varepsilon, \bar{\theta}_k^\varepsilon, \bar{P}_k^\varepsilon, T) = EP_k^2 := h_T^4 \left\|\bar{Y}_k^\varepsilon - y_d - \bar{\theta}_k^\varepsilon\right\|_{L_2(T)}^2 + h_T^3 \left\|[\![\nabla \bar{P}_k^\varepsilon]\!]\right\|_{L_2(\partial T)}^2.$$

We are now in the position to finally derive our estimators $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, $\mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ which only contains known discrete or continuous functions. This is the subject of the next section:

5.2.2 Collecting the Global Estimate

Inserting the error estimators of Section 5.2.1 into the definitions of $\mathcal{E}_r^2$, (5.2.2), $\mathcal{E}_s$, (5.2.3), and $\hat{\mathcal{E}}_s$, (5.2.4), we obtain the following bounds:

$$\mathcal{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \lesssim \mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$$

with

$$\begin{aligned}\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) := {}& 4\mathbf{EYL2}_k^2 + 6\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2 + 4\mathbf{EP}_k^2 \\ &+ 8(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon) + 8(\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c) \\ &+ 8(\bar{\theta}_k^\varepsilon, I_k y_c - y_c).\end{aligned} \tag{5.2.12}$$

Secondly,

$$\mathcal{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \lesssim \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$$

with

$$\begin{aligned}\mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) :=\; & 8\left\|\bar{Y}_k^\varepsilon - y_d\right\|\left(\mathbf{EYL2}_k + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|\right) \\ & + 8\left\|\Pi(\bar{P}_k^\varepsilon)\right\|\mathbf{EP}_k + 8\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\| + 8\left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\|\min(\varepsilon^{2N(1-1/p')}, \varepsilon^{3N/2}) \\ & + 8\min\Bigg\{\left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|, \varepsilon^{-3N/p'}\Big(\mathbf{EYH1}_k + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\| \\ & + \|y_c - I_k y_c\|_{H^1(\Omega)} + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}\Big)\Bigg\}\end{aligned} \tag{5.2.13}$$

And lastly,

$$\hat{\mathcal{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \lesssim \hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$$

with

$$\begin{aligned}\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) :=\; & 48\left\|\bar{Y}_k^\varepsilon - y_d\right\|^2\left(\mathbf{EYL2}_k^2 + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2\right) \\ & + 48\left\|\Pi(\bar{P}_k^\varepsilon)\right\|^2\mathbf{EP}_k^2 + 48\left\|\bar{\theta}_k^\varepsilon - P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2 + 48\left\|P_k^{0+}\bar{\theta}_k^\varepsilon\right\|^2\min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N}) \\ & + 48\min\Bigg\{\left\|(\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^-\right\|^2, 4\varepsilon^{-6N/p'}\Big(\mathbf{EYH1}_k^2 \\ & + \left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|^2 + \|y_c - I_k y_c\|_{H^1(\Omega)}^2 + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}^2\Big)\Bigg\}.\end{aligned} \tag{5.2.14}$$

As we aimed for, the right hand sides in (5.2.12), (5.2.13) and (5.2.14) contain only known discrete or continuous functions. Combining the bounds above with the result of Theorem 4.2.13, we obtain the estimator:

$$\left\|\bar{U}_k^\varepsilon - \bar{u}\right\|^2 \lesssim \varepsilon^{\gamma N} + \mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon),$$

where, as observed before, the estimators $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ contain only known discrete or continuous functions

5.3 'MARK'

In this section we will explain how to obtain local indicators from our estimators $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\hat{\mathfrak{E}}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ which we can then use for our marking strategy, Algorithm 5.3.1:
For the regular part $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ we gain:

$$\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \sum_{T \in \mathcal{T}_k} (\mathfrak{e}_k^r)^2(T)$$

with local indicator (compare (5.2.6) and (5.2.10) for the definition of $EY2_k^2$ and EP_k^2)

$$\begin{aligned}(\mathfrak{e}_k^r)^2(T) := {} & 8(\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon)_{L_2(T)} + 8(\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(T)} \\ & + 8(\bar{\theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(T)} + 6 \left\| \bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon) \right\|_{L_2(T)}^2 \\ & + 4EP_k^2 + 4EY2_k^2.\end{aligned}$$

Due to the min-bracket in the localisable singular part $\hat{\mathfrak{E}}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, (5.2.14), things are a bit more complicated:
Let us first define:

$$c_k^t := \left\| \bar{Y}_k^\varepsilon - y_d \right\|^2 \text{ and } d_k^t := \left\| \Pi(\bar{P}_k^\varepsilon) \right\|^2$$

and recall the localisable singular part $\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$:

$$\begin{aligned}\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) := {} & 48 c_k^t \left\| \bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon) \right\|^2 + 48 d_k^t \mathbf{EP}_k^2 + 48 c_k^t \mathbf{EYL2}_k^2 \\ & + 48 \left\| \bar{\theta}_k^\varepsilon - P_k^{0+}(\bar{\theta}_k^\varepsilon) \right\|^2 + 48 \left\| P_k^{0+}(\bar{\theta}_k^\varepsilon) \right\|^2 \min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N}) \\ & + 48 \min \Bigg\{ \left\| (\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^- \right\|^2, 4\varepsilon^{-6N/p'} (\mathbf{EYH1}_k^2 + \left\| \bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon) \right\|^2 \\ & + \| y_c - I_k y_c \|_{H^1(\Omega)}^2 + |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}^2) \Bigg\}.\end{aligned}$$

In case the minimum in the definition above is attained by $\left\| (\bar{U}_k^\varepsilon + \Delta y_c)^- \right\|^2$ the local indicators are given by

$$\begin{aligned}(\mathfrak{e}_k^s)^2(T) := {} & 48 c_k^t \left\| \bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon) \right\|_{L_2(T)}^2 + 48 \left\| (\Pi(\bar{P}_k^\varepsilon) + \Delta y_c)^- \right\|_{L_2(T)}^2 + 48 \left\| \bar{\theta}_k^\varepsilon - P_k^{0+}(\bar{\theta}_k^\varepsilon) \right\|_{L_2(T)}^2 \\ & + 48 \left\| P_k^{0+}(\bar{\theta}_k^\varepsilon) \right\|_{L_2(T)}^2 \min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N}) + 48 d_k^t EP_k^2 + 48 c_k^t EYL2_k^2.\end{aligned}$$

If not, the local indicators take the following shape (compare (5.2.9) for the definition of $EYH1_k^2$):

$$\begin{aligned}(\mathfrak{e}_k^s)^2(T) := {} & (48 + 4\varepsilon^{-6N/p'}) c_k^t \left\| \bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon) \right\|_{L_2(T)}^2 + 48 d_k^t EP_k^2 + c_k^t EYL2_k^2 \\ & + 48 \left\| \bar{\theta}_k^\varepsilon - P_k^{0+} \bar{\theta}_k^\varepsilon \right\|_{L_2(T)}^2 + 48 \left\| P_k^{0+}(\bar{\theta}_k^\varepsilon) \right\|_{L_2(T)}^2 \min(\varepsilon^{4N(1-1/p')}, \varepsilon^{3N}) \\ & + 48\varepsilon^{-6N/p'} (4 \| y_c - I_k y_c \|_{H^1(\Omega)}^2 + 4 |(\bar{Y}_k^\varepsilon - I_k y_c)^-|_{H^1(\Omega)}^2 + 4EYH1_k^2).\end{aligned}$$

Thus, all in all

$$\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \sum_{T \in \mathcal{T}_k} (\mathfrak{e}_k^s)^2(T).$$

These observations now enable us to present our marking algorithm:

Algorithm 5.3.1 Marking

```
Choose parameters $\eta_r \in (0,1]$ and $\eta_s \in (0,1]$.
Let $\mathfrak{e}_k^{\max,r} := \max_{T\in\mathcal{T}_k} \{\mathfrak{e}_k^r\}$ and $\mathfrak{e}^{\max,s} := \max_{T\in\mathcal{T}_k} \{\mathfrak{e}_k^s\}$
for $T \in \mathcal{T}_k$ do MARK $T$ if
    $$\mathfrak{e}_k^r \geq \eta_r \mathfrak{e}_k^{\max,r} \vee \mathfrak{e}_k^s \geq \eta_s \mathfrak{e}_k^{\max,s}$$
end for
```

Algorithm 5.3.1 generates a set of marked elements $\mathcal{M} \subset \mathcal{T}_k$ which is then refined with the help of a refinement algorithm generating a new shape-regular and conforming triangulation $\mathcal{T}_{k+1}$, compare also our remarks in Section 2.3.5. On the new grid $\mathcal{T}_{k+1}$ we again start with 'SOLVE', Section 5.1.

As a brief summary of this chapter we can now lay out the complete adaptive algorithm:

5.4 The Complete Adaptive Algorithm

Algorithm 5.4.1 The Adaptive Algorithm

```
: Choose a tolerance $TOL$ and parameters $N \geq 1$, in case $d = 2\ p' > 4$ and $\eta_r, \eta s \in (0,1]$,
  the latter for the marking algorithm, Algorithm 5.3.1.
SOLVE: Perform the PDAS-algorithm, Algorithm 5.1.1
ESTIMATE: Compute $\mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$, $\mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ and $\hat{\mathfrak{E}}_s^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ as well as the local indicators $\mathfrak{e}_k^r$ and $\mathfrak{e}_k^s$.
  If
    $$\varepsilon^{\gamma N} + \mathfrak{E}_r^2(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) + \mathfrak{E}_s(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) \leq TOL^2$$
  with $\gamma$ from (4.1.22), then break.
  Else: go to MARK
MARK: Perform the marking algorithm, Algorithm 5.3.1.
REFINE: Refine marked elements and generate a new shape-regular and conforming
  triangulation $\mathcal{T}_{k+1}$, go to SOLVE
```

Chapter 6

Numerical Experiments

In this chapter, we will present two numerical examples of the successful implementation of our adaptive algorithm. The first is a smooth example, the second is an example with a Dirac δ-distribution as the multiplier to the continuous unregularised problem, compare Theorem 2.2.7. All computations were done with the help of the finite element software library ALBERTA, cf. [74].

To check the performance of our adaptive algorithm, we constructed an analytic solution of a state-constrained optimal control problem with the help of an additional source term g, which does not change the analyses of the previous sections. The true solution satisfies the following necessary and sufficient optimality system:

$$\begin{aligned} -\Delta\bar{y} &= \bar{u} + g && \text{in } \Omega \\ \bar{y} &= 0 && \text{on } \partial\Omega \\ -\Delta\bar{p} &= \bar{y} - y_d - \bar{\mu} && \text{in } \Omega \\ \bar{p} &= 0 && \text{on } \partial\Omega \\ \bar{u} &= \Pi(\bar{p}) \\ \bar{\mu} &\geq 0 \\ \langle\bar{\mu}, \bar{y} - y_c\rangle &= 0. \end{aligned} \tag{6.0.1}$$

Here, we deliberately leave some ambiguity as to the duality product in the last line. In the smooth case of Section 6.1 the duality product simply is the $L_2(\Omega)$ scalar product, but in the minimum regularity setting of Section 6.2 we have

$$\langle\bar{\mu}, \bar{y} - y_c\rangle = \langle\bar{\mu}, \bar{y} - y_c\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})}.$$

Let us now turn to the actual examples:

6.1 The Smooth Example

For this example we choose $\Omega = [0,1]^2$ and a smooth solution to the optimal control problem. The goal was to verify that the adaptive algorithm deploys degrees of freedom smartly and does not waste them. As we will see, the adaptive algorithm performs reasonably well.

The following functions were given with $x = (x_0, x_1)$:

$$\begin{aligned}
\bar{y}(x) &= \sin(\pi x_0)\sin(\pi x_1) \\
\bar{p}(x) &= 100x_0(x_0-1)x_1(x_1-1)e^{-\frac{1}{100}(x_0-\frac{1}{10})} \\
y_c(x) &= \begin{cases} \bar{y}(x) & \text{if } |(x_0,x_1)-(\frac{1}{2},\frac{1}{2})|^2 \le 0.125 \\ \bar{y}(x) - (|(x_0,x_1)-(\frac{1}{2},\frac{1}{2})|^2 - 0.125)^2 & \text{else} \end{cases} \\
\bar{u}(x) &= \Pi(\bar{p}(x)) \\
\bar{\mu}(x) &= \begin{cases} 10 & \text{if } |(x_0,x_1)-(\frac{1}{2},\frac{1}{2})|^2 \le 0.07 \\ 0 & \text{else.} \end{cases}
\end{aligned}$$

y_d and g were adjusted to solve the adjoint and state equation in (6.0.1). Besides, $a = -5$ and $b = 5$ was chosen.

Though the problem is 'smooth' in the sense that the Lagrange multiplier to the state constraint $\bar{\mu}$ is a regular L_2-function, the problem is of interest numerically, because *strict complementarity* is lacking, i.e.

$$\bar{\mu}(x) = 0 \not\Rightarrow \bar{y}(x) > y_c(x),\ x \in \Omega.$$

From an optimisation point of view this is a disadvantage, because lack of strict complementarity can lead to 'chattering' of active sets, compare [8], Example 5.2 and [10], Section 6.1.5, which is reflected on the finite-dimensional level by the same index being flagged active, then - in the next iterate of the PDAS loop - being flagged inactive again and then active again and so on. This is the dreaded circling behaviour of active set strategies. Fortunately, we did not encounter many numerical issues in our example.

The regularisation parameter ε was fixed to $\varepsilon = 0.015$. For the adaptive algorithm, Algorithm 5.4.1, we chose $N = 3$, $p' = 18$ and $TOL = 10^{-2}$. The parameters for the marking strategy, compare the marking algorithm, Algorithm 5.3.1, were fixed to $\eta_r = 0.65$, $\eta_s = 0.75$. Besides, we observed that by using the modified regular part $\tilde{\mathfrak{E}}(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon)$ defined by

$$\tilde{\mathfrak{E}}(\bar{U}_k^\varepsilon, \bar{V}_k^\varepsilon) = \sum_{T\in\mathcal{T}_k} (\tilde{\mathfrak{e}}_k^r)^2(T) \tag{6.1.1}$$

with local indicator

$$
\begin{aligned}
(\tilde{\mathfrak{e}}_k^r)^2(T) := {} & 8\max((\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon), \bar{P}_k^\varepsilon)_{L_2(T)}, 0) + 8\max((\bar{\theta}_k^\varepsilon, \bar{Y}_k^\varepsilon - I_k y_c)_{L_2(T)}, 0) \\
& + 8\max(\bar{\theta}_k^\varepsilon, I_k y_c - y_c)_{L_2(T)}, 0) + 6\left\|\bar{U}_k^\varepsilon - \Pi(\bar{P}_k^\varepsilon)\right\|_{L_2(T)}^2 \\
& + 4EP_k^2 + 4EY2_k^2
\end{aligned}
$$

we achieved a better performance of the adaptive algorithm.
As the following figure shows, the adaptive strategy achieves a higher computational precision for a given number of degrees of freedom (DOFs), where DOFs= $\dim(\mathbb{U}_k)$. This is of course the desired effect:

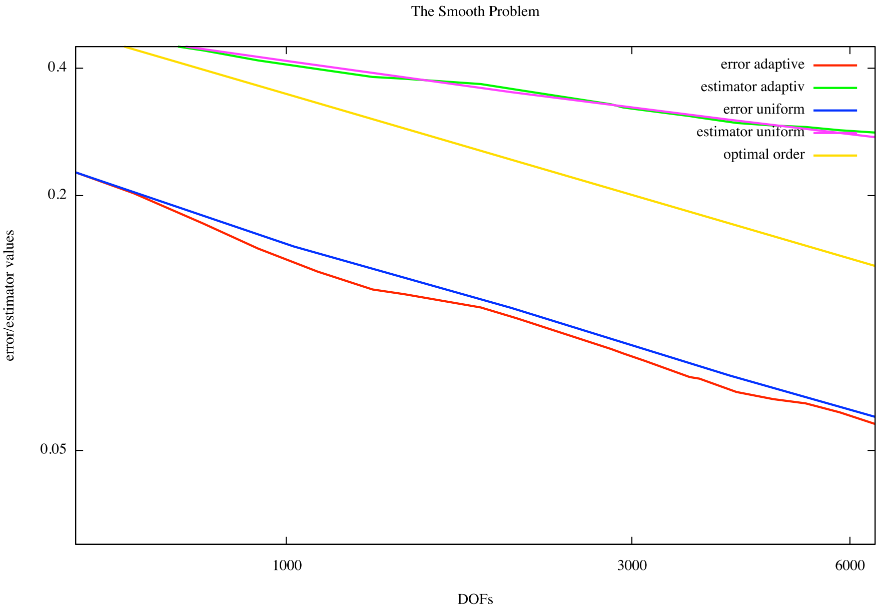

Figure 6.1: The Smooth Example

The figure above has logarithmic scale for both axes.
Let us add some remarks to Figure 6.1:

- For comparison purposes we have included a gold 'optimal order' curve in Figure 6.1. It is motivated by the fact that for the **best-approximation** in $L_2(\Omega)$ of a function $u \in H^1(\Omega)$ by a piecewise constant function $P_k u \in \mathbb{U}_k$ we have the optimal estimate, compare Theorem 2.3, [34] and [12]:

$$\|u - P_k u\| \leq (\dim(\mathbb{U}_k))^{-1/2}. \tag{6.1.2}$$

Thus, we cannot expect the discrete solution $\bar{U}_k^\varepsilon$ to approximate $\bar{u}$ better than the rate $(\dim(\mathbb{U}_k))^{-1/2}$ given by the inequality above. As we see in Figure 6.1, both the adaptive and uniform refinement strategies achieve this order - as would be expected in such a smooth case.

- We observe that the actual error decays slightly faster compared to the estimator, that is, the actual error curve possesses a steeper slope. This is partly explained by our modified estimator (6.1.1) and - in addition - by the fact that the estimator is built for the worst-case, i.e. the least regular case as exemplified by the Dirac example of Section 6.2, where we will see that in this case the estimator reflects the true decay of the error almost exactly. However, in the present case which is significantly smoother, because the Lagrange multiplier for the state constraint $\bar{\mu}$ is a regular L_2-function, some of the estimates used for deriving the estimator are too conservative and too pessimistic. This explains the slightly worse slope.

Next, we want to tackle a more singular example for which the gain of using an adaptive refinement strategy is more obvious.

6.2 The Dirac Example

We choose the ball $B_1(0) \subset \mathbb{R}^2$ as our domain Ω. Even though it is not meshable, the adaptive algorithm resolves the boundary quite well. Thus, at least in this example, the adaptive algorithm even performs well in a setting where there is an additional error coming from the resolution of the curved boundary.
As in the first example and Chapter 4 we treat the model problem with an additional function g as a source term. The following true solution solving the optimality system (6.0.1) was given

$$\begin{aligned}
\bar{y}(x) &= \sin(\pi\,|x|^2)\\
\bar{p}(x) &= 35\ln(|x|)\\
y_c(x) &= \bar{y}(x) - |x|^2\\
\bar{u}(x) &= \Pi(\bar{p}(x))\\
y_d(x) &= \bar{y}(x)
\end{aligned}$$

The Lagrange multiplier for the continuous unregularised problem is given by $35\delta(0)$, the Dirac source at $x = 0$, and $\bar{p}(x)$ is the (scaled) fundamental solution in $2d$ solving

$$\begin{aligned}
-\Delta\bar{p} &= 35\delta(0) \text{ in } B_1(0)\\
\bar{p} &= 0 \qquad \text{on } \partial B_1(0).
\end{aligned}$$

This setting represents the minimum regularity, worst-case setting, because, in this case, $\bar{p}$ is not an $H^1(\Omega)$-function. In choosing $a = -1e12$ and $b = 1e12$, we ensured that the singularity is not completely nullified by the cut off with Π and is making itself felt in the optimal control

$\bar{u}$, too.

As parameters for the adaptive algorithm, recall Algorithm 5.4.1, we chose $N = 3$, $p' = 18$, $TOL = 10^{-2}$ and a fixed regularisation parameter $\varepsilon = 0.018$. The parameters for the marking algorithm, Algorithm 5.3.1, were taken to be $\eta_r = 0.7$ and $\eta_s = 0.8$.

As the following chart shows, the discrete solutions generated by the adaptive algorithm of Chapter 5 are a superior approximation to the true solution compared to uniform refinement. Again DOFs= $\dim(\mathbb{U}_k)$ and again, we employed the modified estimator (6.1.1):

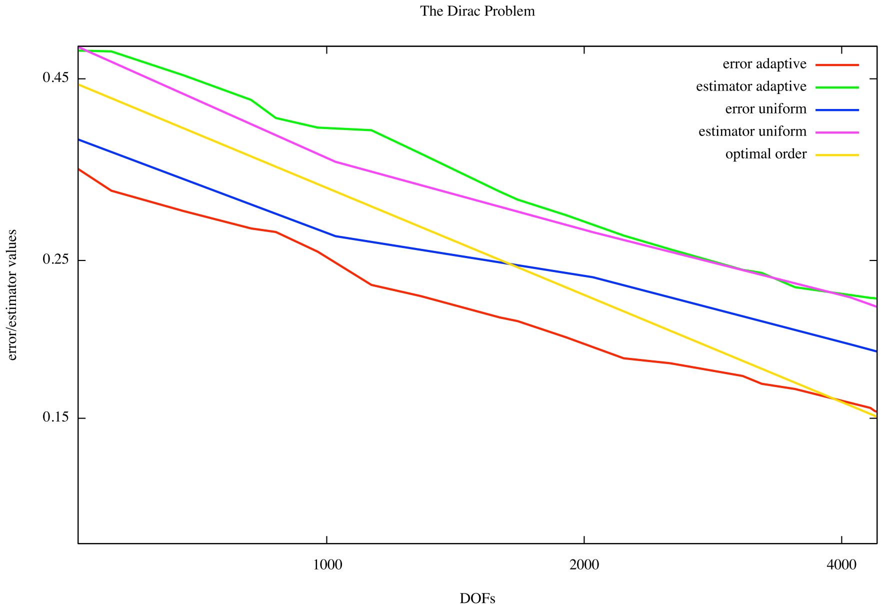

Figure 6.2: The Dirac Example

Again, we employed a logarithmic scale for both axes.

We discern that the adaptive refinement strategy deploys degrees of freedom smartly, since for a given number of DOFs its output is a more accurate solution. In addition, the behaviour of the true error is captured well by the estimator for both the adaptive and uniform refinement case. However, the optimal order of $(\dim(\mathbb{U}_k))^{-1/2}$ - represented by the gold 'optimal order' curve, compare (6.1.2), is not reached. The reason for this is that if the regularisation parameter ε is fixed, eventually the error due to regularisation in this low-regularity setting, which is highly sensitive to changes in ε, dominates. This is highlighted by both the adaptive and uniform error curves flattening out for higher degrees of freedom: The discrete solutions converge to a continuous (smooth) solution $\bar{u}^\varepsilon$ with

$$\|\bar{u} - \bar{u}^\varepsilon\|_{L_2(\Omega)} \approx \varepsilon^\rho, \text{ for some } \rho > 0.$$

That is of course not a satisfiying state of affairs. Therefore, in the next section, we will present an heuristic steering approach for the regularisation parameter ε to achieve the 'optimal order' which - as we will see - provides a remedy to this conundrum.

6.3 An Heuristic Steering Approach for Discretisation and Regularisation

In this section, we will give an example of an heuristic approach to steer both regularisation and discretisation simultaneously. Let us first dwell a bit on the motivation:
Suppose that the discrete unregularised problem (DMP_k) admits the existence of a Lagrange multiplier $\bar{\theta}_k \in C^-_{\mathbb{V}_k}$ for the state constraint in the vein of Definition 2.2.5 which is furthermore uniformly bounded in $L_1(\Omega)$. The following optimality condition for the unique solution $\bar{U}_k$ holds:

$$(\bar{P}_k + \nu\bar{U}_k, U - \bar{U}_k) - (\bar{\theta}_k, S_k U - S_k\bar{U}_k) \geq 0 \ \ \forall U \in \mathcal{U}_k.$$

Then, testing the optimality condition above with $\bar{U}^\varepsilon_k$ and the optimality condition in the KKT system for the regularised discrete problem (4.1.25) with $\bar{U}_k$, we deduce - after a short computation and also harnessing the improved bounds for the discrete multiplier $\bar{\theta}_k$ which are merely an application of the results of Corollary 4.3.3 to the unregularised setting:

$$\nu \left\|\bar{U}_k - \bar{U}^\varepsilon_k\right\|^2 + \frac{1}{\varepsilon}\left\|\bar{V}^\varepsilon_k\right\|^2 \lesssim \varepsilon^{1-3/p'}|\bar{V}^\varepsilon_k|_{H^1(\Omega)}.$$

This gives rise to the interpretation of $|\bar{V}^\varepsilon_k|_{H^1(\Omega)}$ as an 'indicator' for the overall error which is *solely generated by regularisation.* Needless to say, this is not a solid mathematical basis for a simultaneous steering of regularisation and discretisation, but as Figure 6.3 below demonstrates, it is not without its merits.
Based on this indicator, we then performed the following ε-adaption given a tolerance TOL for the adaptive algorithm and an initial regularisation parameter ε:

Algorithm 6.3.1 ε-adaption

```
Choose parameters $0 < \rho < 1$, $0 < \gamma < 1$ and $C_s > 0$ as well as $\varepsilon_{\min} > 0$.
Compute $|V^\varepsilon_k|_{H^1(\Omega)}$
if $\varepsilon^{\rho-3/p'}|V^\varepsilon_k|_{H^1(\Omega)} > C_s TOL^2$ then set
    $\varepsilon = \max(\gamma_1\varepsilon, \varepsilon_{\min})$
else set
    Do not change $\varepsilon$
end if
```

We realise that strictly speaking, we have not utilised $\varepsilon^{1-3/p'}|\bar{V}^\varepsilon_k|_{H^1(\Omega)}$ as an indicator but $\varepsilon^{\rho-3/p'}|\bar{V}^\varepsilon_k|_{H^1(\Omega)}$ with $\rho < 1$. This modification has its roots in the observation that a uniform (independent of ε and k) bound on the term $\varepsilon^{\rho-3/p'}|\bar{V}_k^\varepsilon|_{H^1(\Omega)}$ with $\rho < 1$ coupled with a uniform $L_1(\Omega)$-bound for the Lagrange multiplier $\bar{\theta}_k^\varepsilon$ for the regularised problem (DMP_k) ensures that the central convergence condition (3.3.4) is fulfilled. We will shortly sketch why: A uniform bound for $\|\bar{\theta}_k^\varepsilon\|_{L_1(\Omega)}$ immediately leads to the bound - compare the arguments of Lemma 4.1.7 -

$$\|\bar{\theta}_k^\varepsilon\|_{L_p(\Omega)} \lesssim \varepsilon^{-3/p'},\ 1 \leq p \leq 2,\ \frac{1}{p} + \frac{1}{p'} = 1.$$

This in turn allows us to estimate:

$$|a_k'(\varepsilon)| = \frac{3}{2\varepsilon^2}\left\|\bar{V}_k^\varepsilon\right\|^2 = \frac{3}{2\varepsilon^2}|(\bar{V}_k^\varepsilon, \bar{V}_k^\varepsilon)| \lesssim |(\bar{\theta}_k^\varepsilon, \bar{V}_k^\varepsilon)| \lesssim \varepsilon^{-3/p'}\left|\varepsilon\bar{V}_k^\varepsilon\right|_{H^1(\Omega)} \lesssim C_s\varepsilon^{-\rho}TOL^2.$$

Now, because $0 < \rho < 1$, we have an integrable function on $(0,1)$ bounding $|a_k'(\varepsilon)|$ which means that the convergence condition (3.3.4), where equi-integrability of the sequence a_k' was demanded, is met.

We employed the following paramters: $TOL = 10^{-2}$, $N = 3$, $p' = 18$ and for the marking strategy $\eta_r = 0.65$ and $\eta_s = 0.75$. The parameters of Algorithm 6.3.1 were chosen as $\rho = 0.91$, $C_s = 6.5$, $\gamma_1 = 0.85$ and $\varepsilon_{\min} = 0.004$. and obtained the following results:

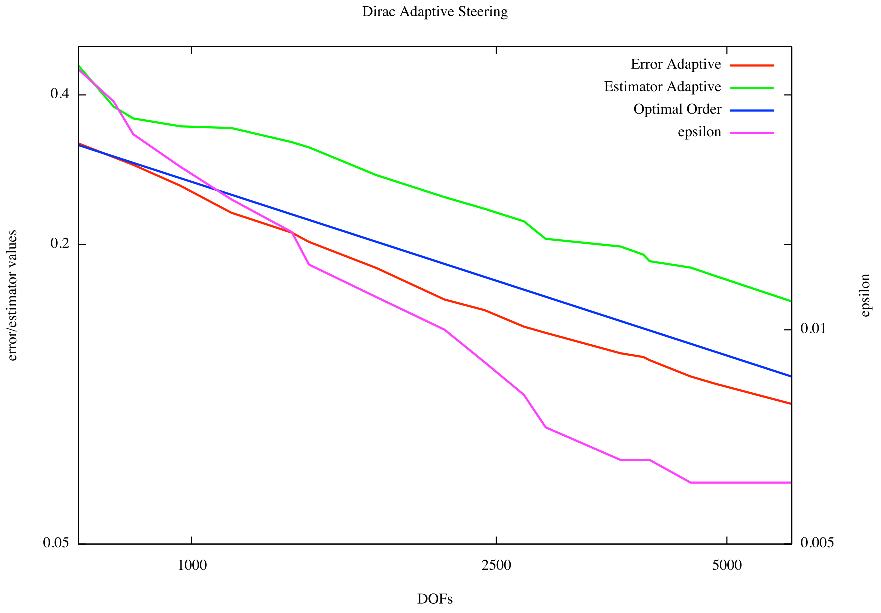

Figure 6.3: Simultaneous Steering

As before, a logarithmic scale for both axes was used.

We realise that employing the steering strategy for the regularisation parameter ε proposed by Algorithm 6.3.1 combined with adaptive refinement according to our estimator, we recover the optimal order of $(\dim(\mathbb{U}_k))^{-1/2}$ that can be expected for a piecewise constant ansatz, in fact there is even a slight superconvergence effect.

For a uniform strategy such an approach is at least not immediately applicable, because there are fewer iterates and the mesh is refined globally such that the degrees of freedom double in each step. Due to the high ε-sensitivity of this Dirac example this would mean that ε would have to be decreased quite swiftly otherwise we are (almost) in the same situation as in the fixed regularisation parameter setting, where eventually the regularisation error dominates. In particular γ_1 in Algorithm 6.3.1 would have to be chosen much closer to 0. This, however, leads to numerical difficulties because the driving down the regularisation parameter to 0 will lead to a problem for which the PDAS algorithm fails to converge. Thus, Algorithm 6.3.1 lends itself better to a situation where the grid is refined adaptively - which is why we implemented it merely in an adaptive setting.

Conclusions & Outlook

To assess the results of this thesis, it is perhaps best to recall its title: 'AFEM for State-Constrained Optimal Control - Convergence Analysis and A Posteriori Error Estimation'. Reflecting upon the theorems and proofs of Chapter 3 and Chapter 4, we realise that what we set out to do, namely to provide a rigorous convergence analysis without imposing any regularity conditions on the mesh and/or the solution and to derive a reliable a posteriori estimator, we have achieved.

Needless to say, there are directions for possible expansions of the results presented in this thesis. First and foremost, perhaps, one would like to measure the efficiency of the a posteriori error estimator. After all, efficiency is a property highly coveted in a posteriori error analysis and unsurprisingly so, since this is the notion that perhaps best captures the inherent advantage of adaptive methods compared to the strategy of uniform refinement. In the context of a posteriori error estimation for (linear elliptic) PDE, efficiency is defined by the existence of a *a local lower bound up to oscillation* in the following vein, compare [65], Theorem 6.2:

$$E_T \lesssim \left\|\bar{u} - \bar{U}_k\right\|_{L_2(\omega(T))} + \operatorname{osc}(\bar{U}_k)$$

Here, $\omega(T)$ denotes a patch of elements

$$\omega_T := \left\{T' \in \mathcal{T}_k \,:\, \bar{T} \cap \bar{T}' \neq \emptyset\right\},$$

E_T is the local error indicator, $\bar{u} - \bar{U}_k$ the error between the true and discrete solution and osc denotes a data oscillation term that converges faster than the true error. However, such a notion is unsuited to the unusual structure of the a posteriori error estimator derived in this setting, after all the estimator derived in Chapter 4 has one term, $\mathcal{E}_r^2$ squared, and the other, $\mathcal{E}_s$ entering without. Therefore, another notion would have to be developed.

Also, it would of course be desirable to possess an efficient strategy to steer discretisation and regularisation simultaneously. Unfortunately, this is a very intricate question and the subject of a posteriori error estimates for the regularisation would have to be addressed in this setting - a subject whose complexity must not be underestimated.

Besides, it is also of interest to find out if there are settings in which the necessary and sufficient condition of convergence in Chapter 3 is not fulfilled. In the setting of the control of a Poisson equation with a constant state constraint $y_c \equiv \beta < 0$, a sufficient condition for convergence would be the existence of a sequence of functions $z_k^\varepsilon \in H^1(\Omega)$ with $\nabla z_k^\varepsilon \in H(\mathrm{div}, \Omega)$ such that

$$|\bar{Y}_k^\varepsilon - I_k y_c - z_k^\varepsilon|_{H^1(\Omega)} \lesssim \varepsilon^\gamma, \;\; \|\Delta z_k^\varepsilon\| \lesssim \varepsilon^{-\rho}, \;\; \gamma, \rho > 0 \;\; \gamma > \frac{3}{p'}, \; \rho < 3(\frac{1}{2} - \frac{1}{p'}),$$

where p' is chosen in such a way that $H^1(\Omega) \hookrightarrow L_{p'}(\Omega)$.
At the end, though, let us with these final thoughts stress that it is fair to say that in this thesis, basic techniques were developed with which state-constrained optimal control problems can be analysed without the machinery of maximum norm error estimates, be it on a discrete or on the continuous level. That is a genuine novelty in this particular branch of mathematics - and in the authors' humble view a novelty with some merit - , hopefully, perhaps, one which in the future will come to serve at least as a little brick in a large, striking and impressive edifice of future results on adaptive methods in the context of state-constrained optimal control problems.

Bibliography

[1] R.A. Adams and J.J.F. Fournier. *Sobolev spaces.* Academic Press, San Diego, 2007.

[2] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2000.

[3] H.W. Alt. *Lineare Funktionalanalysis.* Springer-Verlag, Berlin, 2006.

[4] G. Bachmann and L. Narici. *Functional Analysis.* Academic Press, 3rd edition, 1968.

[5] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *IMPACT Comput. Sci. Eng.*, 3:181–191, 1991.

[6] O. Benedix and B. Vexler. A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Computational Optimization and Applications*, 44(1):3–25, 2009.

[7] J. Bergh and J. Löfström. *Interpolation spaces.* Springer, Berlin, 1976.

[8] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37(4):1176–1194, 1999.

[9] M. Bergounioux and K. Kunisch. Primal-dual strategy for state-constrained optimal control problems. *Comput. Optim. Appl.*, 22(2):193–224, 2002.

[10] Maïtine Bergounioux, Mounir Haddou, Michael Hintermüller, and Karl Kunisch. A comparison of a moreau–yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM Journal on Optimization*, 11(2):495–521, 2000.

[11] Daniele Boffi, Franco Brezzi, and Michel Fortin. Finite elements for the Stokes problem. In Daniele Boffi and Lucia Gastaldi, editors, *Mixed Finite Elements, Compatibility Conditions, and Applications*, volume 1939 of *Lecture Notes in Mathematics*, pages 45–100. Springer Berlin Heidelberg, 2008.

[12] A. Bonito and R. Nochetto. Quasi-optimal convergence rate of an adaptive discontinuous galerkin method. *SIAM J. Numer. Anal.*, 48:734–771, 2010.

[13] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* Cambridge University Press, 2007.

[14] S.C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods.* Texts in Applied Mathematics. Springer, 2008.

[15] J. Campos, J. Mahwin, and R. Ortega. Maximum principles around an eigenvalue with constant eigenfunctions. *Communications in Contemporary Mathematics*, 10(06):1243–1259, 2008.

[16] E. Casas and M. Mateos. Error estimates for the numerical approximation of Neumann control problems. *COAP*, 39(3):265–295, 2008.

[17] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Siebert. Quasi-optimal convergence rates for an adaptive finite element method. *SIAM Journal on Numerical Analysis*, 46(5):2524–2550, 2008.

[18] S. Cherednichenko, K. Krumbiegel, and A. Rösch. Error estimates for the Lavrentiev regularization of elliptic optimal control problems. *Inverse Problems*, 24(6), 2008.

[19] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems.* SIAM Classics In Applied Mathematics, Philadelphia, 2002.

[20] P.G Ciarlet and Lions J.-L. *Handbook of Numerical Analysis: Finite Element Methods (Part I)*, volume 2. Elsevier, Amsterdam New York, 2006.

[21] P.G. Ciarlet and P. Raviart. Maximum principle and uniform convergence. *Computational Methods in Applied Mechanics and Engineering*, 2:17–31, 1973.

[22] J. Crank. *The Mathematics of Diffusion.* Clarendon Press, Oxford, 2nd edition edition, 1975.

[23] G. Dal Maso, F. Murat, L. Orsina, and A. Prignet. Renormalized solutions of elliptic equations with general measure data. *Annali della Scuola Normale Superiore di Pisa*, 28(4):741–808, 1999.

[24] J.C. de los Reyes, C. Meyer, and B. Vexler. Finite element error analysis for state-constrained optimal control of the Stokes equations. *Control and Cybernetics*, 37(2):251–284, 2008.

[25] K. Deckelnick, A. Günther, and M. Hinze. Finite element approximation of elliptic control problems with constraints on the gradient. *Numer. Math.*, to appear, 2008.

[26] K. Deckelnick and M. Hinze. Numerical analysis of a control and state constrained elliptic control problem with piecewise constant control approximations. In K. Kunisch, G. Of, and O. Steinbach, editors, *Numerical Mathematics and Advanced Applications*, pages 597–604, Berlin Heidelberg, 2008. Springer-Verlag.

[27] N. Dunford and B.J. Pettis. Linear operations on summable functions. *Transactions of the American Mathematical Society*, 47(3):323–392, 1940.

[28] N. Dunford and J.T. Schwartz. *Linear Operators: Part I.* Wiley Interscience, New York, 1957.

[29] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* SIAM Classics In Applied Mathematics, 1999.

[30] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159. Springer, 2004.

[31] Lawrence C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, Rhode Island, 2002.

[32] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton, FL, 1992.

[33] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen.* Akademie–Verlag, Berlin, 1974.

[34] Fernando D. Gaspoz and Pedro Morin. Convergence rates for adaptive finite elements. *IMA J. Numer. Anal.*, 29(4):917–936, 2009.

[35] M. Giaquinta and S. Hildebrandt. *Calculus of Variations I: The Lagrangian Formalism.* Calculus of Variations. Springer, 1996.

[36] D. Gilbarg and Trudinger N.S. *Elliptic Partial Differential Equations of Second Order.* Springer–Verlag, Berlin Heidelberg, 2001.

[37] P. Grisvard. *Elliptic Problems in Nonsmooth Domains.* Pitman, Boston, 1985.

[38] P. Grisvard. *Singularities in Boundary Value Problems.* Recherches en mathématiques appliquées. Masson, 1992.

[39] E. Hewitt and K. Stromberg. *Real and Abstract Analysis.* Springer-Verlag, Berlin Heidelberg New York, 3rd edition, 1965.

[40] E. Hewitt and K. Yosida. Finitely additive measures. *Transactions of the American Mathematical Society*, 72:46–66, 1952.

[41] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM Optim.*, 13(3):865–888, 2003.

[42] M. Hintermüller, Hinze M., and R. H. W. Hoppe. Weak-duality based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state constraints. *Journal of Computational Mathematics*, 30(2):101–123, 2012.

[43] M. Hintermüller, F. Tröltzsch, and I. Yousept. Mesh-independence of semismooth Newton methods for Lavrentiev-regularized state constrained nonlinear optimal control problems. *Numer. Math.*, 108(4):571–603, 2008.

[44] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Computational Optimization and Applications*, 30(1):45–61, 2005.

[45] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer-Verlag, Berlin, 2009.

[46] R.H.W. Hoppe and M. Kieweg. Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems. *Computational Optimization and Applications*, 46(3):511–533, 2010.

[47] Kazufumi Ito and Karl Kunisch. On a semi-smooth Newton method and its globalization. *Mathematical Programming*, 118:347–370, 2009.

[48] J. Jost. *Postmodern Analysis.* Universitext (1979). Springer, 2005.

[49] K. Kohls, A. Rösch, and K. Siebert. A posteriori error analysis of optimal control problems with control constraints. *SIAM Journal on Control and Optimization*, 52(3):1832–1861, 2014.

[50] K. Kohls, A. Rösch, and K. G. Siebert. A posteriori error estimators for control constrained optimal control problems. In Günter Leugering, Sebastian Engell, Andreas Griewank, Michael Hinze, Rolf Rannacher, Volker Schulz, Michael Ulbrich, and Stefan Ulbrich, editors, *Constrained Optimization and Optimal Control for Partial Differential Equations*, volume 160 of *International Series of Numerical Mathematics*, pages 431–443. Springer Basel, 2012.

[51] I. Kossacký. A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, 55:275–288, 1994.

[52] K. Krumbiegel, I. Neitzel, and A. Rösch. Regularization error estimates for semilinear elliptic optimal control problems with pointwise state and control constraints. *Comput. Optim. Appl.*, 52:181–207, 2012.

[53] A. Kufner, O. John, and S. Fučik. *Function Spaces.* Noordhoff International Publishing, Leyden, Netherlands, 1977.

[54] K. Kunisch and A. Rösch. Primal-dual active set strategy for a general class of constrained optimal control problems. *SIAM Journal Optimization*, 13(2):321–334, 2002.

[55] S. Kurcyusz. On the existence and nonexistence of Lagrange multipliers in Banach spaces. *Journal of Optimization Theory and Applications*, 20(1):81 – 110, September 1976.

[56] S. Kurcyusz and J. Zowe. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.

[57] R. Liska and M. Shashkov. Enforcing the discrete maximum principle for linear finite element solutions of second-order elliptic problems. *Communications in Computational Physics*, 3(4):852–877, 2008.

[58] D.G. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.

[59] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control and Cybernetics*, 37:51–85, 2008.

[60] P. Morin, K. G. Siebert, and A Veeser. A basic convergence result for conforming adaptive finite elements. *Mathematical Models and Methods in Applied Sciences*, 18(5):707–737, 2008.

[61] Jindřich Nečas. Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 16(4):305–326, 1962.

[62] J. Necǎs. *Les méthodes directes en théorie des équations elliptiques.* Masson, Paris, 1967.

[63] R. H. Nochetto, A. Schmidt, K. G. Siebert, and A. Veeser. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numerische Mathematik*, 104(4):515–538, 2006.

[64] R. H. Nochetto and L. Wahlbin. Positivity preserving finite element approximation. *Mathematics of Computation*, 71(240):1405–1419, 2001.

[65] Ricardo H. Nochetto, Kunibert G. Siebert, and Andreas Veeser. Theory of adaptive finite element methods: An introduction. In Ronald DeVore and Angela Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 409–542. Springer Berlin Heidelberg, 2009.

[66] J. Peetre and Duke University. Mathematics Dept. *New thoughts on Besov spaces.* Duke University mathematics series. Mathematics Dept., Duke University, 1976.

[67] P. Pucci and J. Serrin. *The Maximum Principle.* Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, Basel, 2007.

[68] S.M. Robinson. Stability theory for systems of inequalities, part II: Differentiable nonlinear systems. *SIAM J. Numer. Anal.*, 13(4):497–513, 1976.

[69] A. Rösch and S Steinig. A priori error estimates for a state-constrained elliptic optimal control problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(5):1107–1120, 2012.

[70] A. Rösch and D. Wachsmuth. A posteriori error estimates for optimal control problems with state and control constraints. *Numerische Mathematik*, 120(4):733–762, 2012.

[71] Halsey L. Royden. *Real Analysis.* Macmillan, New York, 3rd edition, 1988.

[72] W. Rudin. *Functional Analysis.* International series in pure and applied mathematics. McGraw-Hill, 1991.

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

[74] Alfred Schmidt and Kunibert G. Siebert. *Design of Adaptive Finite Element Software. The Finite Element Toolbox ALBERTA*, volume 42. Springer, 2005.

[75] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483–493, 1990.

[76] K.G Siebert. A convergence proof for adaptive finite elements without lower bounds. *IMA Journal of Numerical Analysis*, 31(3):947–970, 2011.

[77] G. Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second order à coefficients discontinus. *Annales de l'Institut Fourier*, 15(1):189–257, 1965.

[78] L. Tartar. The space H(div;Ω). In *An Introduction to Sobolev Spaces and Interpolation Spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*, pages 99–101. Springer Berlin Heidelberg, 2007.

[79] H Triebel. *Interpolation Theory, Function Spaces, Differential Operators.* J. A. Barth Verlag, Heidelberg-Leipzig, 1995.

[80] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications.* Graduate Studies in Mathematics. American Mathematical Society, 2010.

[81] M. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim*, 13:805–842, 2003.

[82] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces.* MOS-SIAM Series on Optimization, Philadelphia, PA, 2011.

[83] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. *SIAM Journal on Control and Optimization*, 38(6):1938–1984, 2000.

[84] R. Verfürth. Error estimates for a mixed finite element approximation of the Stokes equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 18(2):175–182, 1984.

[85] R. Wheeden and A. Zygmund. *Measure and Integral.* Marcel Dekker, Inc., New York Basel, 1st edition, 1977.

[86] W. Wollner. A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints. *Computational Optimization and Applications*, 47(1):133–159, 2010.

[87] K. Yosida. *Functional Analysis.* Springer, Berlin Heidelberg, 6th edition, 1995.

Lebenslauf

1992 - 1996	Besuch der Städtischen Grundschule Königsschule Oberhausen
1996 - 2005	Besuch des Freiherr-vom-Stein Gymnasiums Oberhausen
2005	Abitur am Freiherr-vom-Stein Gymnasium (Notendurchschnitt: 1,2)
2005 - 2010	Studium der Mathematik und Volkswirtschaftslehre an der Universität Duisburg-Essen
2010	Diplomprüfung an der Universität Duisburg-Essen (Gesamtnote: 1,0 mit Auszeichung), Diplomarbeit: '*A Priori Error Estimates for State-Constrained Optimal Control Problems'* (Betreuer: Prof. Dr. Arnd Rösch)
2010 - Juni 2011	wissenschaflicher Angestellter am Lehrstuhl Numerische Mathematik der Universität Duisburg-Essen, Leiter: Prof. Dr. K.G. Siebert
Mai 2011	Ehrung für die beste Diplomarbeit der Fakultät für Mathematik der Universität Duisburg-Essen des Jahrgnags im Rahmen des Dies Academicus
Juni 2011 - August 2013	wissenschaftlicher Angestellter am Lehrstuhl Numerik für Höchstleistungsrechner der Universität Stuttgart, Leiter: Prof. Dr. K.G. Siebert
September 2013 - jetzt	wissenschaftlicher Angestellter am Lehrstuhl X, AG Kontinuierliche Optimierung der TU Dortmund, Leiter: Prof. Dr. C. Meyer
Juli 2014	Einreichung der Dissertation: '*Adaptive Finite Elements for State-Constrained Optimal Control Problems - Convergence Analysis and A Posteriori Error Estimation'* (Hauptberichter: Prof. Dr. K.G. Siebert)
29.10.2014	Datum der mündlichen Prüfung. Gesamtnote der Dissertation: summa cum laude (mit Auszeichnung)