Analysis and Optimal Control of Quasilinear Parabolic Evolution Equations in Divergence Form on Rough Domains

Hannes Meinlschmidt

Contents

Preface

The central objects in this thesis are quasilinear parabolic evolution equations in divergence form. Parabolic evolution equations are the driving force behind the majority of irreversible processes in natural sciences, provided there exists a suitable mathematical model of the process in the first place. We will be concerned with the special type of quasilinear parabolic evolution equations in divergence form, the abstract prototype being

$$y'(t) - \nabla \cdot \sigma(y)(t)\rho\nabla y(t) = F(y)(t) \quad \text{for a.a. } t \in J, \quad y(T_0) = y_0. \qquad (1)$$

Here, J is a finite time interval with left end point T_0 and the equation is supposed to hold already in a function space of functions living on a spatial domain $\Omega \subset \mathbb{R}^d$. The latter indicates that we will stay in a very abstract setting which will allow to handle a large class of problems at once, including problems subject to mixed boundary conditions and inhomogeneous Robin– or Neumann boundary data. The function space of choice will then be the dual space $\mathrm{W}_D^{-1,q}(\Omega)$ of a Sobolev space $\mathrm{W}_D^{1,q'}(\Omega)$ whose elements satisfy $u \restriction D = 0$ in a suitable sense. This choice corresponds to homogeneous Dirichlet-data on the subset D of $\partial\Omega$ and Robin– or Neumann boundary data on $N \coloneqq \partial\Omega \setminus D$, where the extreme cases $D = \emptyset$ and $D = \partial\Omega$ are allowed. As a general rule of thumb, the reader may imagine $y(t)$ to be the spatial temperature profile in a workpiece Ω whose temporal evolution is described by (1). In this case, the boundary conditions posed correspond to cooling at the Dirichlet part D and to insulation or "insertion" or "extraction" of heat at N, depending on the actual form of the Robin– or Neumann conditions.

While there are multiple ways to study such a quasilinear parabolic evolution equation, for instance Galerkin approximation approaches, monotone operator theory, or Leray-Schauder techniques based on the fixed point theorem of the same name, our approach will be via maximal parabolic regularity as a so-to-say "sub-approach" to that of analytic semigroups. We refer to [3] for more general information and to the monograph of Ladyzhenskaya et al. [101] for a showcase of most of the other mentioned techniques. This maximal regularity– or analytic semigroup approach has been thoroughly studied in the last two decades by, among others, Prüss [127] and Clément and Li [41], Lunardi [111] and Amann [3, 6, 8, 10], building upon the works of Clément and Simonett [42], Da Prato and Grisvard [44], as well as Angenent [14], and may possibly also be called "classical" nowadays.

The fundamental idea is to use maximal parabolic regularity to construct fixed point mappings in the maximal regularity spaces and works roughly as follows. Suppose we want to show existence and possibly uniqueness of solutions to the quasilinear model equations as above, and thereby assume that we are able to solve the corresponding *linear* equation, where the nonlinearities in (1) are "frozen", in a satisfactory way; that is, the equation

$$y'(t) - \nabla \cdot \sigma(w)(t)\rho\nabla y(t) = F(w)(t) \quad \text{for a.a. } t \in J, \quad y(T_0) = y_0, \tag{2}$$

has a unique solution $y = y_w$ for every w from a suitable class of functions. Seeing it this way, it "only" remains to show existence of a fixed point of the mapping $w \mapsto y_w$, including continuity– or compactness properties of this mapping, to obtain a solution to the original nonlinear equation (1).

The maximal regularity approach has the advantage that it is very versatile and works in an abstract setting, in particular for the function spaces encoding mixed boundary conditions. It is however apparent that this general procedure depends strongly on the properties of the differential operator on these function spaces. There are some rather delicate technicalities to deal with: On the one hand, the domains of the differential operators $-\nabla \cdot \sigma(w)(t)\rho\nabla$ may vary already with respect to t for fixed w but of course also with respect to w itself. This means that the dependence of the divergence-gradient operator on w must be well-behaved, and this is true not only with respect to the domains, but also in a sense of Lipschitz-continuity. On the other hand, it is known already from the theory of ordinary differential equations that solutions to nonlinear differential equations may blow-up after finite time, as the standard example

$$\mathfrak{u}'(t) = \mathfrak{u}^2(t), \quad \mathfrak{u}(0) = 1 \tag{3}$$

shows, whose unique solution $\mathfrak{u}(t) = \frac{1}{1-t}$ will blow up for $t \nearrow 1$ in the sense that $\mathfrak{u}(t)$ goes to infinity. This means we cannot in general expect global-in-time solutions for our much more general class of differential equations.

On top of the already difficult analytic structure, the ultimate goal is to do optimal control of quasilinear parabolic evolution equations. That is, we assume that we are able to manipulate a control function or a parameter u inside the equation (1). We will consider being able to do so only via the inhomogeneity F. Since we are in a very abstract $\mathrm{W}^{-1,q}_D(\Omega)$-setting for the equations, this includes many interesting cases, in particular boundary control. The determining equation is thus given by

$$y'(t) - \nabla \cdot \sigma(y)(t)\rho\nabla y(t) = F(y,u)(t) \quad \text{for a.a. } t \in J, \quad y(T_0) = y_0. \tag{4}$$

The aim is now to find a control u such that it, together with its associated state, performs best possible with respect to a cost– or objective function. More mathematically spoken, we want to prove that, given such an objective function and the controlled quasilinear equation (4), there exists a control u which minimizes the given objective functional, and to characterize such an optimal control.

Of course, finding an optimal control is intimately related to finding the optimal state corresponding to this equation and the properties of the associated states are often the defining subject of the objective functions. A typical objective would be to drive the associated state y to a preferred state depending on the context, so-called tracking. Thereby, there might be additional constraints which further narrow the possible states. These might be both control– and state constraints, so limitations on the control u which we are able to impose on the system and limitations on the class of states among which we want to find an optimal one. Returning to the visual example of $y(t)$ being the spatial temperature profile in a workpiece Ω, the state constraints require that y does not exceed certain critical temperatures, for example a melting point. In particular, both state– and control constraints are often critical to a correct modeling and an actually usable result.

While linear and semilinear, both elliptic and parabolic, and quasilinear elliptic optimal control problems are rather well-understood nowadays, the available literature concerning quasilinear parabolic optimal control problems is still surprisingly scarce. We refer to the excellent textbooks of Tröltzsch [148] and Hinze, Pinnau, M. Ulbrich and S. Ulbrich [86] for a comprehensive treatment of the mentioned well-understood classes of optimal control problems. Quasilinear parabolic optimal control problems seem to have attracted some interest at different points in the last decades, starting from the sixties with Lions [107], continuing up to the present day [35, 62, 121, 123–125, 134, 141, 145]. The classical ansatz relies on monotone operator theory and the techniques of Ladyzhenskaya et al. [101]. Although the Hilbert space setting is notably absent nearly already in the very first works, we were unable to locate a paper in which the maximal regularity ansatz, possibly in a nonsmooth setting, is used, not to speak of mixed boundary conditions.

Overview

We obtain solutions to the quasilinear parabolic evolution equation as in (1) under minimal assumptions by appealing to fundamental general "solution theorems" of Amann [10] or Prüss [127] for this class of equations. The very general data allowed in these theorems allows to also treat systems of equations which are subordinated to (1) in a suitable sense, by solving the remaining equations in dependency of y and insert the dependence into the right-hand side F. Since the theorem in [10] also includes nonlocal-in-time mappings F and σ, this in particular also works for coupled systems of evolution equations (see [89] for an example of this technique in the L^p setting). To verify the assumptions of these theorems, we build upon recent advances regarding the Kato square root property [20, 58] together with elliptic– and parabolic regularity results for the divergence-gradient operators in function spaces related to mixed boundary conditions [52, 77, 80]. These give the opportunity to tackle the problem of varying domains and a well-behaved dependence on the coefficient function in the model equations (1) and (2) under very weak assumptions on the spatial domain Ω. In particular, we can leave the already quite general class of (non-strong) Lipschitz domains behind, adopted for mixed boundary conditions by GRÖGER and his *regular sets* [73].

The problem of possible blow-up however is still present in both abstract solution theorems and it is not to be expected of such general results to be able to exclude this possibility. We thus establish a new Hölder-regularity result for nonautonomous parabolic evolution equations under weak assumptions, from which we are able to deduce a *global* existence result for equations of type (1), however under much stronger assumptions on the Lipschitz-continuity of F including a global boundedness condition. This should not be surprising as it is already well-known from semilinear theory, where σ is constant, that one has to pose growth– or absolute bounds on the inhomogeneities F to obtain global solutions (we refer to the standard textbook of PAZY [126]).

These considerations are laid out in Chapter 2, building upon a significant amount of work in which we set up a functional-analytic basis for the treatment of quasilinear parabolic evolution equations in Chapter 1: We give a short overview of interpolation theory in Chapter 1.1, together with a summary of function spaces and relevant properties in Chapter 1.2. Then we define the geometric framework for the spatial domains Ω in Chapter 1.3, whereas we turn to maximal parabolic regularity for parabolic differential operators in Chapter 1.4. Finally, the most important properties of the divergence-gradient operators on spaces of type $\mathrm{W}_D^{-1,q}(\Omega)$ are collected in Chapter 1.5. The "switch" to real Banach spaces is explained in Chapter 1.6. Many of these general results are already known in various, sometimes isolated, parts in the literature, so we have decided to compile them for comfortable use. We have extended or generalized many of these results to more general geometries, again based on recent results for Sobolev spaces with vanishing traces for very general sets Ω [28,57,59].

In Chapter 2, we first show the Hölder regularity result, which says that solutions to the nonautonomous parabolic evolution equation whose coefficient matrix function μ is merely measurable, coercive, and bounded, are Hölder continuous on $J \times \Omega$ and that the set of solutions corresponding to a bounded set of right-hand sides f is bounded in the Hölder space uniformly with respect to the coercivity– and upper bound of μ. This is included in Chapter 2.1, whereas we consider the quasilinear problem in Chapter 2.2, including the global solutions result. The results in this chapter have been published together with Joachim Rehberg in the article "Hölder-estimates for non-autonomous parabolic problems with rough data" [116].

Turning to the optimal control problem, at first sight it seems like a strange idea to try to optimally control a system for which one does not even know that the associated states do not blow up. It will however turn out that we are able to circumvent this problem by reducing the optimal control problem to the set U_g of controls u whose associated solutions y_u exist globally in time in a suitable sense. While this sounds rather radical, almost all classical choices of the objective functional for the state y require y either to be defined on the whole underlying time interval J or in a fixed time point $T \in J$ and are thus incompatible with only local-in-time solutions.

The most difficult part for the optimal control reduced to U_g is to ensure that there exists a globally optimal control for the system. Existence of optimal controls for nonlinear partial differential equations is fundamentally a hard problem because one needs to pass to the limit in the partial differential equation starting only from weak convergence in a control function space and there are very little *a priori* bounds from which one could infer further convergence properties of the associated states. On top of that, we now need to make sure that we do not leave the set of controls U_g admitting global solutions. We propose to resolve this problem by using boundedness of the objective functional, together with the admissible set, for the sequence of controls and states under consideration (a similar idea was pursued in [12]). Turning our attention to first order necessary conditions, it turns out that the restriction to the set of "global controls" is nearly irrelevant in their formulation because we are able to show that this set is *open* as a byproduct of standard techniques. This will allow to formulate first order necessary optimality conditions which take exactly the same form as one would have obtained if the problem of non-global solutions never existed.

We give these results in a rather abstract form in Chapter 3. To validate the practical usability of the abstract results, we consider the optimal control of the quasilinear *thermistor problem* in two and three spatial dimensions in Chapter 4 as a real-world example where the full strength of the previous considerations and results has to be used. The thermistor problem consists of a coupled system of a parabolic– and an elliptic equation and we use the full generality of the quasilinear existence theorems of Amann or Prüss by solving the elliptic equation in dependence of the searched-for variable in the parabolic equation and insert this dependency into the right-hand side in the parabolic equation. This way, we are able to show existence of globally optimal controls and derive classical first order necessary optimality conditions, where we even obtain provably global solutions in case of space dimension $d = 2$. Distinguishing between space dimension two and three further allows to make the difference in the quasilinear existence theorems apparent. The results for space dimension three will be published in joint articles together with Christian Meyer and Joachim Rehberg [113, 114]. In Chapter 3.1.1, we complement these considerations with the proposal of a suitable control function space for nonlinear optimal control problems which admits crucial compactness properties while being functional-analytically "nice". We use such a space also already in the treatment of the thermistor problem, and there is a paper in preparation about this subject, again together with Christian Meyer and Joachim Rehberg [115].

Since we will amass a quite large collection of objects and different notation, there is a list of symbols provided at the end of this thesis for easier reference. This also includes a bare minimum of standard notation which we do not introduce formally, such as $\mathbb{N}, \mathbb{R}$ and the likes.

Acknowledgment

After five years in the optimization research group at TU Darmstadt and various collaborations all over Germany, I am naturally indebted to a large group of people who in one way or the other contributed to the existence of this thesis.

Of course, the first and foremost thanks need to go to Stefan Ulbrich who willingly took me in his research group. Stefan, you have always taken your time when I needed support and your responses to my, more often than not rather involved and ad-hoc, questions were always very helpful. At the same time, you always gave me a feeling of having trust in my work which I appreciate. You moreover enabled, allowed and encouraged me to pursue collaborations with other researchers and to travel the world visiting conferences. This wasn't a given and I am very grateful for that.

Further "scientific" acknowledgments go to Christian Meyer and Joachim Rehberg for taking me on board for the three-dimensional thermistor problem. The collaboration on this topic has opened my eyes for the underlying issues in many subjects concerning both analysis and optimal control of quasilinear evolution equations. Your steady enthusiasm and encouragement for doing mathematics and the connections on a personal level have played a big role for me.

Moreover, I thank Robert Haller-Dintelmann for sparking my interest in evolution equations and introducing me to the ISEMs, and of course for acting as a referee for this thesis. Speaking of which, I appreciate the support of Winnifried Wollner and Herbert Egger, both for acting as members of my examination committee and for taking their times when I appeared in their doorframe with an urgent question, as well as the support of Jens Lang for acting as the chair of the committee.

I have enjoyed being part of the optimization research group at TU Darmstadt which has produced many memorable moments, for which I thank everyone who was part of it at some point. Particular thanks go to Sebastian Pfaff who was my go-to guy for discussions about mathematics and everything else, office neighbor, and "teammate" for many teaching assis-

tances during the years, and to Michael Fischer who filled that gap when Sebastian had left, in particular in the last days of writing this thesis. I am indebted to Anne-Therese Rauls and Cedric Sehrt for dealing with my not-quite-final-form scribblings which I urged them to read.

Last but in no way least, my friends and parents had my back anytime I needed them. I especially want to thank my parents and Anja, for everything and more. There are no words to describe my gratefulness to you.

Darmstadt, August 2017

Zusammenfassung in deutscher Sprache

Das zentrale Objekt dieser Arbeit sind quasilineare parabolische Evolutionsgleichungen in Divergenzform. Die Mehrzahl aller durch mathematische Modelle erfassten irreversiblen Prozesse in den Naturwissenschaften werden durch parabolische Evolutionsgleichungen beschrieben, und der Protoyp einer solchen Gleichung quasilinearer Natur in Divergenzform ist

$$y'(t) - \nabla \cdot \sigma(y)(t)\rho\nabla y(t) = F(y)(t) \quad \text{für fast alle } t \in J, \quad y(T_0) = y_0, \tag{1}$$

wobei J ein endliches Zeitintervall mit linkem Endpunkt T_0 ist. Wir betrachten solche Gleichungen als abstrakte Operatordifferentialgleichungen in Funktionenräumen, deren Elemente Funktionen auf einem Orts-Gebiet $\Omega \subset \mathbb{R}^d$ sind. Auf diese Weise können wir große Klassen von konkreten Problemstellungen auf einmal bearbeiten. Wir wählen den auf *maximaler parabolischer Regularität* beruhenden Ansatz, der eine große Flexibilität und optimale Regularität der erhaltenen Lösungen verspricht. Quasilineare parabolische Gleichungen sind generell schwierig zu behandeln, da eine Reihe unangenehmer Phänomene auftreten: Die Operatoren $-\nabla \cdot \sigma(y)(t)\rho\nabla$ können unterschiedliche Definitionsmengen haben, sowohl in Bezug auf t für festes y als auch in Bezug auf *alle* in Frage kommenden Funktionen y, und es ist *a priori* nicht klar, wie diese Definitionsmengen von y abhängen. Weiterhin ist bereits aus der Theorie gewöhnlicher Differentialgleichungen bekannt, dass nichtlineare Differentialgleichungen Lösungen haben können, die nach endlicher Zeit explodieren. Da wir erheblich allgemeinere Differentialgleichungen betrachten, muss man von einem solchen Phänomen ebenso ausgehen.

Zu diesen Schwierigkeiten fügen wir noch weitere hinzu, da wir nicht nur an der Analysis quasilinearer parabolischer Evolutionsgleichungen sondern auch an deren optimaler Steuerung interessiert sind. Dabei nehmen wir an, dass wir in der Lage sind, eine Kontrolle bzw. einen Parameter u in der rechten Seite F der Gleichung zu manipulieren, und wollen dies so tun, dass die zugehörige Lösung ein von uns aufgesetztes sogenanntes Kostenfunktional minimiert. Dieses Funktional „bewertet" das Paar von Steuerung und Zustand. Ein häufig verfolgtes Ziel ist dabei, den Zustand y zu einer gewissen Zeit, oder über einen gewissen Zeithorizont, einem vorgegebenen Ziel-Zustand möglichst nahe zu bringen. Es ist dabei klar, dass ein solches Ziel potenziell mit der Problematik des eventuellen Explodierens der zugehörigen Zustände unvereinbar erscheint. Weiterhin beinhalten Optimalsteuerungsprobleme quasilinear parabolischer Natur die grundlegende Schwierigkeit, dass man kaum Vorab-Informationen über die Zustände zu gegebenen Steuerungen erhält. Dies gestaltet dann bereits den Nachweis der Existenz einer optimalen Lösung der Optimalsteuerungsaufgabe sehr schwierig.

Um diesen Schwierigkeiten beizukommen, verfolgen wir folgendes Programm: Da wir, wie angekündigt, in einem recht abstrakten Rahmen argumentieren, setzen wir zunächst in Kapitel 1 ein allgemeines funktionalanalytisches Fundament. Dieser beinhaltet Interpolationstheorie sowie eine systematische Einführung klassischer Funktionenräume mit und ohne verallgemeinerten Null-Randwerten auf Gebieten. Zudem sammeln wir einige Aussagen über maximale parabolische Regularität, die unser zentrales Werkzeug darstellt, und den Hauptprotagonisten der Gleichungen: den Differentialoperator in Divergenz-Form. Dabei werden viele bereits bekannte Resultate zusammengestellt, aber teilweise auch erweitert und angepasst.

In Kapitel 2 widmen wir uns quasilinearen Gleichungen der Form (1). Dabei beweisen wir zunächst ein Hilfsresultat von eigenem Interesse über gleichmässige Hölder-Stetigkeit von Lösungen nichtautonomer linearer parabolischer Evolutionsgleichungen. Solche treten auf natürliche Art und Weise bei der Analysis quasilinearer Probleme auf. Wir beweisen anschließend, dass der Divergenzform-Operator in (1) unter schwachen Voraussetzungen die Annahmen sehr allgemeiner Existenz– und Eindeutigkeitsaussagen über quasilineare parabolische Evolutionsgleichungen erfüllt. Diese liefern allerdings nur, wie erwartet, Lösungen, die lokal in der Zeit existieren, also nicht garantierterweise global. Leider erscheint es auch nicht möglich, die Resultate und Beweise auf globale Existenz zu verallgemeinern. Mittels des Resultats über die Hölder-Regularität nichtautonomer Gleichungen können wir allerdings einen eigenen Satz über Existenz und Eindeutigkeit *globaler* Lösungen der quasilinearen Gleichung (1) herleiten, dessen Annahmen aber dann entsprechend stärker gestellt werden müssen.

Aufbauend auf den Resultaten in Kapitel 2 wird in Kapitel 3 die Gleichung (1) mit einem Parameter bzw. einer Kontrolle u versehen und in ein Optimalsteuerungsproblem eingebettet. Wir kommen dem Problem eventueller Explosion der Lösungen bei, indem wir die Menge der zulässigen Steuerungen auf implizite Weise auf solche einschränken, die globale Lösungen liefern. Dies fügt allerdings dem ohnehin schon schweren Nachweis der Existenz optimaler Steuerungen eine weitere Facette hinzu, da nun noch beachtet werden muss, dass eine das Kostenfunktional minimierende Folge von Steuerungen zu globalen Lösungen als Grenzwert auch wieder eine solche haben sollte. Wir schlagen vor, dieser Problematik durch geeignete Zusatzinformationen über diese minimierende Folge aus dem Kostenfunktional sowie weiterer Einschränkungen im Optimalsteuerungsproblem beizukommen, was allerdings je nach konkreter Ausgestaltung des Optimalsteuerungsproblems vorgenommen werden muss. Diese Schwierigkeiten werden bei der Betrachtung von Optimalitätsbedingungen erster Ordnung für das Optimalsteuerungsproblem wieder ausgeglichen, da es sich hier herausstellt, dass die Einschränkung auf die Menge der Steuerungen, die zu globalen Lösungen führen, praktisch keine Rücksicht genommen werden muss: Man erhält die gleichen Optimalitätsbedingungen, die man auch erhalten würde, wenn die Problematik eventueller Lösungs-Explosion gar nicht aufgetreten wäre. Zudem schlagen wir einen geeigneten Funktionenraum für die Behandlung nichtlinearer Optimalsteuerungsprobleme vor.

Da die Ausführungen in Kapitel 3 recht abstrakter Natur sind, behandeln wir in Kapitel 4 ein praktisches Anwendungsproblem, das sogenannte Thermistor-Problem. Dabei handelt sich um ein gekoppeltes System aus einer Gleichung vom Typ (1) und einer elliptischen Gleichung, das die Wärme-Evolution in einem Bauteil beschreibt, welches durch Stromdurchfluss aufgeheizt wird. Das Ziel ist es, die Temperatur zum Endzeitpunkt durch Anpassen der Stromintensität auf ein gegebenes Temperaturniveau zu bringen. Dabei ist es essenziell, dass der Schmelzpunkt des Materials zu keinem Zeitpunkt des Prozesses überschritten wird. Wir zeigen, dass sich dieses System durch die in Kapitel 3 abstrakt aufgesetzten Ergebnisse behandeln lässt. Dies ist möglich, da die Ergebnisse in Kapitel 2 auch rechte Seiten F in (1) erlauben, die durch Auflösen der elliptischen Gleichung nach der Temperatur und Rück-Einsetzen in die parabolische Gleichung entstehen (tatsächlich erlauben die Ergebnisse sogar eine solche Prozedur für weitere Evolutionsgleichungen). Wir beenden unsere Ausführungen mit numerischen Ergebnissen und Beobachtungen zu einer konkreten Ausprägung des Thermistor-Problems unter realistischen Daten. Dabei verifizieren wir insbesondere, dass die gestellten Bedingungen an das Optimalsteuerungsproblem auch tatsächlich notwendig sind, um einen realistischen, aber auch durchführbaren Prozess zu beschreiben.

CHAPTER 1

Functional-analytic framework

This first chapter serves as a foundation and collection of the concepts, conventions and terminology used in the following chapters to treat quasilinear optimal control problems on irregular domains. We begin with a brief overview of interpolation theory before we continue by introducing function spaces. This allows us, later on, to stay abstract first and then use interpolation theory for the actual function spaces. Now, why bother with interpolation theory at all? On the one hand, it is a rather beautiful and fascinating theory. On the other hand, it exhibits relationships between function spaces, namely, *embeddings* and *interpolation identities*, which are of particular interest for us.

Definition 1.0.1 (Embedding)**.** Let X, Y be topological vector spaces. We say that Y is *continuously embedded* into X and write $Y \hookrightarrow X$ if Y is a linear subspace of X and the natural injection $i\colon Y \to X$, $i(y) = y$, is a continuous linear mapping . If X, Y are Banach spaces with $Y \hookrightarrow X$ and the natural injection i is even a compact mapping, we say that Y is *compactly embedded* into X and write $Y \hookrightarrow\hookrightarrow X$. If $Y \hookrightarrow X$ and $i(Y)$ is dense in X, we say that Y is *densely embedded into* X and write $Y \hookrightarrow_{\mathrm{d}} X$.

The existence of an embedding of Y into X tells us that Y is both algebraically and topologically contained in X in the sense that $Y \subseteq X$ and the topology on Y is finer than the one induced on X. This will allow us to derive properties of Y in terms of X.

It will be helpful to consider "adjoint embeddings" in the following sense (see [3, P. 271]). Let X, Y be locally convex vector spaces such that $Y \hookrightarrow_{\mathrm{d}} X$ via the embedding $i \in \mathscr{L}(Y; X)$. Then the set

$$Y'_X \coloneqq \Big\{ y' \in Y' \colon y' \text{ is continuous w.r.t. the } X\text{-topology of } Y \Big\}$$

can be identified with X' via the observation that $i'(X') = Y'_X$ with the *injective* mapping i' given by $i'(x') = x' \circ i$, because of the denseness of Y in X. In this spirit, we have the following formal statement:

Proposition 1.0.2 ([3, Ch. V, Prop. 1.4.8])**.** *Let X, Y be locally convex vector spaces such that $Y \hookrightarrow_{\mathrm{d}} X$. Then $X' \hookrightarrow Y'$ and*

$$\langle x', y \rangle_{Y',Y} = \langle x', y \rangle_{X',X} \quad \text{for all } x' \in X',\ y \in Y$$

via the identification above. If Y is reflexive, then we even have $X' \hookrightarrow_{\mathrm{d}} Y'$.

Interpolation on the other hand allows us to construct seemingly complicated spaces Z by "superimposing" two others, X and Y, in a topologically useful way; a common visualization is that of a convex combination of spaces X and Y. This gives us a tool to infer properties of Z from properties of X and Y.

1.1 A brief overview of interpolation theory

We give a brief summary of interpolation theory, touching at least the bits relevant for this thesis. The monographs of Triebel [146] and Bergh and Löfström [24] provide a more in-depth treatment and proofs of the claims below.

The idea is to derive function spaces—or generally, Banach spaces—from given, well-understood spaces by superimposing their characteristic properties, or to identify function spaces, already obtained by some other construction, as interpolation spaces. Fundamentally, we need that the Banach spaces under consideration to be interpolated are compatible in an algebraic and topological way in the following sense.

Definition 1.1.1. Let A_0 and A_1 be Banach spaces. We say that (A_0, A_1) is an *interpolation couple* if there exists a topological vector space $\mathscr{A}$ with the Hausdorff property such that

$$A_0 \hookrightarrow \mathscr{A} \quad \text{and} \quad A_1 \hookrightarrow \mathscr{A}.$$

For an interpolation couple (A_0, A_1), the spaces $A_0 \cap A_1$ and $A_0 + A_1$ equipped with their standard norms are also Banach spaces with the property that

$$A_0 \cap A_1 \hookrightarrow A_i \hookrightarrow A_0 + A_1, \quad \text{for } i = 0, 1.$$

Now let (A_0, A_1) and (B_0, B_1) be two interpolation couples with the superspaces $\mathscr{A}$ and $\mathscr{B}$. Interpolation theory asks for Banach spaces A such that $A \hookrightarrow \mathscr{A}$ and B such that $B \hookrightarrow \mathscr{B}$ with the following interpolation property: For every linear continuous operator $T\colon \mathscr{A} \to \mathscr{B}$ which satisfies $T \restriction A_0 \in \mathscr{L}(A_0, B_0)$ and $T \restriction A_1 \in \mathscr{L}(A_1, B_1)$, the restriction of T to A is a continuous linear operator from A to B, i.e., $T \restriction A \in \mathscr{L}(A, B)$.

The mathematically sound foundation for the construction of the spaces A, B from (A_0, A_1) and (B_0, B_1) is that of *categories* and their associated *functors*:

Definition 1.1.2 (Categories and functors)**.**

(i) A *category* $\mathfrak{C}$ consists of a collection of *objects* $A, B, C, \ldots$ and *morphisms* $R, S, T, \ldots$, were every morphism T is associated to an ordered tuple $[A, B]$ of objects in an unique way, for which we write $T \sim [A, B]$. We say that T is a morphism *from* A *into* B. If

$T \sim [A, B]$ and $S \sim [B, C]$ are two morphisms in $\mathfrak{C}$, there also exists the composition morphism $R := ST \sim [A, C]$ in $\mathfrak{C}$. This composition satisfies the associative law, i.e. $T(SR) = (TS)R$ for every three compatible morphisms T, S, R. For each object A there exists the identity morphism $I_A \sim [A, A]$ with the property $I_A T = T$ for every $T \sim [B, A]$ and $SI_A = S$ for every $S \sim [A, B]$, both for any object B.

(ii) Let $\mathfrak{C}_1, \mathfrak{C}_2$ be two categories. A (covariant) *functor* is a mapping $\mathcal{F}$ from $\mathfrak{C}_1$ into $\mathfrak{C}_2$ such that $\mathcal{F}(A)$ is an object of $\mathfrak{C}_2$ for every object A of $\mathfrak{C}_1$ and such that $\mathcal{F}(T) \sim [\mathcal{F}(A), \mathcal{F}(B)]$ is a morphism of $\mathfrak{C}_2$ for every morphism $T \sim [A, B]$ of $\mathfrak{C}_1$. This means $\mathcal{F}(I_A) = I_{\mathcal{F}(A)}$ for every object A of $\mathfrak{C}_1$ and $\mathcal{F}(TS) = \mathcal{F}(T)\mathcal{F}(S)$ for every two compatible morphisms T, S of $\mathfrak{C}_1$.

The two categories that will be used for interpolation theory are

- the category $\mathfrak{C}_{\mathrm{B}}$ consisting of all complex Banach spaces $A, B, \ldots$ as objects and the set of linear continuous operators between these Banach spaces as morphisms, i.e., $T \sim [A, B]$ means $T \in \mathscr{L}(A; B)$,
- the category $\mathfrak{C}_{\mathrm{I}}$ consisting of all interpolation couples $(A_0, A_1), (B_0, B_1), \ldots$ with the morphisms in $\mathfrak{C}_{\mathrm{I}}$ defined by $T \sim [(A_0, A_1), (B_0, B_1)]$ iff $T \in \mathscr{L}(A_0 + A_1; B_0 + B_1)$ and $T \restriction A_i \in \mathscr{L}(A_i; B_i)$ for $i = 0, 1$.

Definition 1.1.3 (Interpolation functor and interpolation space)**.** Let $\mathcal{F}$ be a functor from $\mathfrak{C}_{\mathrm{I}}$ to $\mathfrak{C}_{\mathrm{B}}$. We say that $\mathcal{F}$ is an *interpolation functor* if

(i) $A_0 \cap A_1 \hookrightarrow \mathcal{F}((A_0, A_1)) \hookrightarrow A_0 + A_1$ for every interpolation couple (A_0, A_1),

(ii) for every morphism $T \sim [(A_0, A_1), (B_0, B_1)]$ of $\mathfrak{C}_{\mathrm{I}}$, we have $\mathcal{F}(T) = T \restriction \mathcal{F}((A_0, A_1))$.

The interpolation functor is said to be *of type* θ, $\theta \in [0, 1]$, if there exists $C \geq 1$

$$\|T\|_{\mathscr{L}(\mathcal{F}(A_0,A_1);\mathcal{F}(B_0,B_1))} \leq C \|T\|^{1-\theta}_{\mathscr{L}(A_0;B_0)} \|T\|^{\theta}_{\mathscr{L}(A_1;B_1)}$$

for all morphisms $T \sim [(A_0, A_1), (B_0, B_1)]$ of $\mathfrak{C}_{\mathrm{I}}$. It is said to be *exact* if we may choose $C = 1$.

An interpolation functor produces exactly what we asked for an interpolation space above. For a given interpolation couple (A_0, A_1) we thus call the set $\{\mathcal{F}((A_0, A_1))\colon \mathcal{F} \text{ is an interpolation functor}\}$ the set of interpolation spaces between A_0 and A_1. Gagliardo and Aronszajn have shown that the concept of producing interpolation spaces via functors is not an obstruction in finding *all* possible spaces with the interpolation property as explained above – they showed that for each such space there exists an exact functor which produces it [24, Thm. 2.5.1]. Rather obvious interpolation functors are given by $\mathcal{F}_i((A_0, A_1)) = A_i$ for $i = 0, 1$ and by $\mathcal{F}((A_0, A_1)) = A_0 \cap A_1$ and $\mathcal{F}((A_0, A_1)) = A_0 + A_1$. Before we introduce nontrivial interpolation functors, we quickly insert a general interpolation principle which will prove exceptionally useful in the following.

Definition 1.1.4 (Retraction-coretraction)**.** Let A, B be Banach spaces. We say that an operator $R \in \mathscr{L}(A; B)$ is a *retraction* if there exists an operator $E \in \mathscr{L}(B; A)$ such that $RE = I_B$, the identity operator on B. The operator E is called *coretraction* (to R).

Theorem 1.1.5 ([146, Ch. 1.2.4])**.** *Let (A_0, A_1) and (B_0, B_1) be interpolation couples and let $E \sim [(B_0, B_1), (A_0, A_1)]$ and $R \sim [(A_0, A_1), (B_0, B_1)]$ be given, with the property that $E \restriction B_i \in \mathscr{L}(B_i; A_i)$ is a coretraction and $R \restriction A_i \in \mathscr{L}(A_i; B_i)$ is the corresponding retraction for $i = 0, 1$. Let $\mathcal{F}$ be an interpolation functor. Then*

$$E \restriction \mathcal{F}((B_0, B_1)) \in \mathscr{L}_{\mathrm{iso}}(\mathcal{F}((B_0, B_1)); ER\mathcal{F}((A_0, A_1))),$$

where $ER\mathcal{F}((A_0, A_1))$ is a complemented subspace of $\mathcal{F}((A_0, A_1))$ with its inherited norm and ER acts as a projection on it.

The actual form of Theorem 1.1.5 which we need is the following:

Corollary 1.1.6 ([56, Cor. 1.3.7]). *Let the assumptions of Theorem 1.1.5 be given. If the space $R\mathcal{F}((A_0, A_1))$ is equipped with the quotient norm*

$$\|f\|_{R\mathcal{F}((A_0,A_1))} := \inf_{\substack{g\in\mathcal{F}((A_0,A_1))\\ Rg=f}} \|g\|_{\mathcal{F}((A_0,A_1))},$$

then $\mathcal{F}((B_0, B_1)) \doteq R\mathcal{F}((A_0, A_1))$.

Let us now turn to the actual interpolation functors. The two functor families we use are

- the *real interpolation* functor $\mathcal{F}_{\mathbb{R}}((A_0, A_1)) = (A_0, A_1)_{\theta,p}$ for the parameters $\theta \in [0, 1]$, $1 \le p \le \infty$,
- the *complex interpolation* functor $\mathcal{F}_{\mathbb{C}}((A_0, A_1)) = [A_0, A_1]_\theta$ for the parameter $\theta \in [0, 1]$.

The real interpolation functor/method may be obtained by a zoo of different methods (K– and J-functionals, mean– or trace methods, ...), most of them due to Peetre or J.L. Lions. We will briefly introduce its construction by means of the J-functional, taken from [146, Ch. 1.6.1], since we need a few details from "under the hood" later in Lemma 1.1.9. The complex interpolation functor/method on the other hand is based on the Three Lines theorem from complex analysis. The method is due to Caldéron, J.L. Lions and Krejn, see e.g. [32]. A particular famous use of complex interpolation before the invention of complex interpolation is the Riesz-Thorin theorem. We will, however, not go into detail of the construction of the complex interpolation spaces. Note that the real and complex interpolation functor are both exact of type θ [146, Ch. 1.3.3 and 1.9.3], but applying them to a fixed interpolation couple (A_0, A_1) generally gives different spaces.

Definition 1.1.7 (Real interpolation by J-functional). Let (A_0, A_1) be an interpolation couple, $0 < \theta < 1$ and $1 \le p \le \infty$. Define the *J-functional* $J\colon \mathbb{R}^+ \times (A_0 \cap A_1) \to \mathbb{R}_0^+$ by $J(t, f) := \max(\|f\|_{A_0}, t\|f\|_{A_1})$. Then the space $(A_0, A_1)_{\theta,p}$ consists of those $f \in A_0 + A_1$ such that there exists a continuous function $u\colon \mathbb{R}^+ \to A_0 \cap A_1$ satisfying

$$\int_0^\infty \left(t^{-\theta} J(t, u(t))\right)^p \frac{\mathrm{d}t}{t} < \infty, \tag{1.1}$$

such that f is given by

$$f = \int_0^\infty u(t) \frac{\mathrm{d}t}{t} \quad \text{in} \quad A_0 + A_1, \tag{1.2}$$

and we set

$$\|f\|_{(A_0,A_1)_{\theta,p}} \coloneqq \inf_{u \text{ as in (1.1), (1.2)}} \Big(\int_0^\infty \left(t^{-\theta} J(t, u(t))\right)^p \frac{\mathrm{d}t}{t} \Big)^{\frac{1}{p}}. \tag{1.3}$$

For $p = \infty$, we replace integration over $(0, \infty)$ of $t^{-\theta} J(t, u(t))$ by taking the essential supremum over $(0, \infty)$ of the same function.

Let us next collect often-needed properties of the interpolation spaces obtained by the real and complex method.

Lemma 1.1.8 ([146, Ch. 1.3.3, 1.6.2 and 1.9.3], [24, Ch. 3.8])**.** *Let (A_0, A_1) be an interpolation couple and let $0 < \theta < 1$ and $1 \leq p \leq \infty$.*

(i) $(A_0, A_1)_{\theta,p} = (A_1, A_0)_{1-\theta,p}$.

(ii) $A_0 \cap A_1$ *is dense in* $(A_0, A_1)_{\theta,p}$ *for* $p \neq \infty$.

(iii) There exists $C = C_{\theta,p} > 0$ *such that*

$$\|f\|_{(A_0,A_1)_{\theta,p}} \leq C \|f\|_{A_0}^{1-\theta} \|f\|_{A_1}^{\theta} \quad \textit{for all } f \in A_0 \cap A_1. \tag{1.4}$$

(iv) If $p \leq q \leq \infty$, *then*

$$(A_0, A_1)_{\theta,1} \hookrightarrow (A_0, A_1)_{\theta,p} \hookrightarrow (A_0, A_1)_{\theta,q} \hookrightarrow (A_0, A_1)_{\theta,\infty}. \tag{1.5}$$

(v) If $A_0 \hookrightarrow A_1$ *and* $\theta < \zeta < 1$, *then*

$$A_0 \hookrightarrow (A_0, A_1)_{\theta,p} \hookrightarrow (A_0, A_1)_{\zeta,q} \hookrightarrow A_1 \tag{1.6}$$

for all $1 \leq q \leq \infty$. *If even* $A_0 \hookrightarrow\hookrightarrow A_1$*, then the embedding in the middle of* (1.6) *is compact, too.*

All the above properties are also true if the real interpolation spaces are replaced by the complex ones, thereby ignoring the second parameter.

Most properties in Lemma 1.1.8 are intrinsic for the interpolation property and show that the interpolation functors are "working as intended". A particular case of this phenomenon is the following lemma, which will find widespread use since it gives a simple tool to establish embeddings of interpolation spaces without directly messing with interpolation theory. The assertion should be compared with equations (1.4) and (1.5), see also [146, Ch. 1.10.1].

Lemma 1.1.9. *Let* (A_0, A_1) *be an interpolation couple,* $0 < \theta < 1$*, and let* E *be a Banach space such that* $A_0 \cap A_1 \hookrightarrow E \hookrightarrow A_0 + A_1$*. Suppose that there is a constant* $C \geq 0$ *such that*

$$\|f\|_E \leq C\|f\|_{A_0}^{1-\theta}\|f\|_{A_1}^{\theta} \quad \textit{for all } f \in A_0 \cap A_1. \tag{1.7}$$

Then $(A_0, A_1)_{\theta,1} \hookrightarrow E$*.*

Proof. First of all, observe that (1.7) implies that, for all $f \in A_0 \cap A_1$,

$$\|f\|_E \leq C\tau^{-\theta}\|f\|_{A_0}^{1-\theta}\big(\tau\|f\|_{A_1}\big)^{\theta} \leq C\tau^{-\theta}\max\big(\|f\|_{A_0}, \tau\|f\|_{A_1}\big) \tag{1.8}$$

uniformly for $\tau \in \mathbb{R}^+$. In particular, $A_0 \cap A_1 \hookrightarrow E$ follows again from setting $\tau = 1$. Now let $f \in (A_0, A_1)_{\theta,1}$. According to the construction of the real interpolation space by means of the J-method, see Definition 1.1.7, there exists a continuous function $u\colon \mathbb{R}^+ \to A_0 \cap A_1$ such that

$$f = \int_0^\infty u(t)\,\frac{\mathrm{d}t}{t} \quad \text{and} \quad \int_0^\infty t^{-\theta}\max\big(\|u(t)\|_{A_0}, t\|u(t)\|_{A_1}\big)\,\frac{\mathrm{d}t}{t} < \infty.$$

Now using (1.8) for $u(t) \in A_0 \cap A_1 \hookrightarrow E$ and $\tau = t$ shows that

$$\int_0^\infty \|u(t)\|_E\,\frac{\mathrm{d}t}{t} \leq C\int_0^\infty t^{-\theta}\max\big(\|u(t)\|_{A_0}, t\|u(t)\|_{A_1}\big)\,\frac{\mathrm{d}t}{t} < \infty,$$

hence
$$g_u := \int_0^\infty u(t)\,\frac{\mathrm{d}t}{t} \in E$$
with
$$\|g_u\|_E \le C \int_0^\infty t^{-\theta} \max\bigl(\|u(t)\|_{A_0}, t\|u(t)\|_{A_1}\bigr)\,\frac{\mathrm{d}t}{t}. \tag{1.9}$$
It is crucial to observe that the function $g_u \in E$ is in fact *independent* of the function u used to represent f in $A_0 + A_1$, since by assumption $E \hookrightarrow A_0 + A_1$ and for every f-representative u, g_u in $A_0 + A_1$ is exactly f. Since embeddings are supposed to be injective, this means that g_u must in fact be independent of u and we identify it with f. This allows to take the infimum over all functions u associated to f in (1.9), and using the definition of the $(A_0, A_1)_{\theta,1}$-norm as in (1.3) we finally infer that $\|f\|_E \le C\|f\|_{(A_0,A_1)_{\theta,1}}$ for all $f \in (A_0, A_1)_{\theta,1}$. This was the claim. $\square$

Together with its analogue for an upper embedding $E \hookrightarrow (A_0, A_1)_{\theta,\infty}$, cf. [146, Ch. 1.10.3], Lemma 1.1.9 shows that
$$(A_0, A_1)_{\theta,1} \hookrightarrow [A_0, A_1]_\theta \hookrightarrow (A_0, A_1)_{\theta,\infty}, \tag{1.10}$$
where $0 < \theta < 1$, for every interpolation couple (A_0, A_1). The lemma moreover admits another useful technique of obtaining embeddings between different interpolation spaces:

Corollary 1.1.10. *Let (A_0, A_1) be an interpolation couple, let B_0 be another Banach space such that $A_0 \hookrightarrow B_0 \hookrightarrow A_1$, and let $0 < \theta < 1$ and $1 \le p \le \infty$. Then $(A_0, A_1)_{\theta,1} \hookrightarrow (B_0, A_1)_{\theta,p}$ and $(A_0, A_1)_{\theta,1} \hookrightarrow [B_0, A_1]_\theta$.*

Proof. From the assumption on B_0, we observe that (B_0, A_1) is also an interpolation couple and that
$$A_0 \cap A_1 \doteq A_0 \hookrightarrow B_0 \doteq B_0 \cap A_1 \hookrightarrow (B_0, A_1)_{\theta,p} \hookrightarrow B_0 + A_1 \doteq A_1 \doteq A_0 + A_1,$$
Using (1.4) and the presumed embedding $A_0 \hookrightarrow B_0$, we moreover estimate
$$\|f\|_{(B_0,A_1)_{\theta,p}} \le C\|f\|_{B_0}^{1-\theta}\|f\|_{A_1}^{\theta} \le C\|f\|_{A_0}^{1-\theta}\|f\|_{A_1}^{\theta}$$

for all $f \in (B_0 \cap A_1) \cap (A_0 \cap A_1) = A_0 \cap A_1$. Thus, Lemma 1.1.9 implies the claim. The case for the complex interpolation method works of course analogously. $\square$

A reoccurring motif in interpolation theory is that of convexity. We have already encountered the condition for interpolation functors to be of type θ and, using this property for the real– and complex method, the norm-inequality (1.4), both of which are logarithmically convex in the parameter θ. A common way to imagine the process of interpolation is taking a θ-convex combination of the interpolation couple (A_0, A_1). This is made particularly clear in the next theorem. Apart from the intuition it gives for the convexity roots of interpolation, it also states that both interpolation methods are stable under repeated application of their interpolation functor. We refer to [146, Ch. 1.10.2] or [24, Thm. 4.6.1] for a proof.

Theorem 1.1.11 (Reiteration theorem)**.** *Let (A_0, A_1) be an interpolation couple, $0 \leq \theta_0 \leq \theta_1 \leq 1$ and let $1 \leq p, q_0, q_1 \leq \infty$ and $0 < \lambda < 1$. Then*

$$\big((A_0, A_1)_{\theta_0,q_0}, (A_0, A_1)_{\theta_1,q_1}\big)_{\lambda,p} \doteq (A_0, A_1)_{(1-\lambda)\theta_0+\lambda\theta_1,p}.$$

Moreover, if $A_0 \cap A_1$ is dense in A_0, A_1 and $[A_0, A_1]_{\theta_0} \cap [A_0, A_1]_{\theta_1}$, then

$$\big[[A_0, A_1]_{\theta_0}, [A_0, A_1]_{\theta_1}\big]_{\lambda} \doteq [A_0, A_1]_{(1-\lambda)\theta_0+\lambda\theta_1}.$$

Note that Theorem 1.1.11 is not cited in its whole generality, cf. [146, Ch. 1.10.2]. Another particular reiteration theorem which we will use is that for domains of fractional powers of *positive operators*.

Definition 1.1.12 (Positive operator)**.** Let X be a Banach space and let A be a linear closed operator on X with dense domain $\operatorname{dom} A$. We say that A is *positive* if $(-\infty, 0] \subset \rho(A)$, the resolvent set of A, and

$$\sup_{\lambda\in[0,\infty)} (1+\lambda)\big\|(A+\lambda)^{-1}\big\|_{\mathscr{L}(X)} < \infty$$

holds true.

If A is a positive operator, its fractional powers A^z for $z \in \mathbb{C}$ are meaningful and often also very useful operators on X. We refer to [146, Ch. 1.15] and [3, Ch. III.4.6] for the precise constructions and properties, which we will not need explicitly. Having a positive operator at hand, we obtain the following reiteration theorem anchored at X for the domains of its fractional powers. We state it only for real powers, since we will not need complex ones. See [146, Ch. 1.15.4].

Theorem 1.1.13 (Reiteration theorem for fractional powers)**.** *Let A be a positive operator on X and let $0 \le \alpha < \beta$ as well as $0 < \theta < 1$ and $1 \le p \le \infty$. Then*

$$\left(\operatorname{dom} A^\alpha, \operatorname{dom} A^\beta\right)_{\theta,p} \doteq \left(X, \operatorname{dom} A^\beta\right)_{\zeta,p} \quad \textit{with} \quad \zeta = \frac{\alpha(1-\theta)+\beta\theta}{\beta}$$

holds true.

We close this brief exposure to interpolation theory by duality properties. In order for this to work properly, we need that (A_0', A_1') is again an interpolation couple if (A_0, A_1) is so. The simple solution is to require that $A_0 \cap A_1$ is dense in both A_0 and A_1. Then we indeed obtain

$$(A_0 + A_1)' \hookrightarrow A_i' \hookrightarrow (A_0 \cap A_1)' \quad \text{for } i = 0, 1,$$

where the “upper” embedding is the critical one, and the following duality property holds true:

Lemma 1.1.14 ([146, Ch. 1.11.2/1.11.3])**.** *Let (A_0, A_1) be an interpolation couple such that $A_0 \cap A_1 \hookrightarrow_{\mathrm{d}} A_i$ for $i = 0, 1$ and let $0 < \theta < 1$ and $1 \le p < \infty$. Then*

$$(A_0, A_1)_{\theta,p}' \doteq (A_0', A_1')_{\theta,p'}. \tag{1.11}$$

If A_0 or A_1 is in addition reflexive, (1.11) *remains true for the complex interpolation method, ignoring the parameter p.*

From the identity (1.11), we directly obtain a condition for the reflexivity of interpolation spaces:

Corollary 1.1.15 ([146, Ch. 1.11.3])**.** *Let A_0, A_1 be reflexive and let $A_0 \cap A_1 \hookrightarrow_{\mathrm{d}} A_i$ and $A_0' \cap A_1' \hookrightarrow_{\mathrm{d}} A_i'$, each for $i = 0, 1$. Then, for $1 < p < \infty$ and $0 < \theta < 1$, the interpolation spaces $(A_0, A_1)_{\theta,p}$ and $[A_0, A_1]_\theta$ are reflexive as well.*

1.2 Function spaces

The study of partial differential equations via functional analysis relies critically on the properties of the underlying function spaces chosen to represent the equations in. We thus introduce the function spaces used in the following, thereby starting with the general basic spaces such as of the Lebesgue spaces and the spaces of continuous or continuously differentiable functions. They have the advantage that they are simple to define on general sets Υ in Euclidean space, whereas we will have to draw distinctions between function spaces of functions defined on $\mathbb{R}^d$ and on (bounded) domains $\Lambda \subset \mathbb{R}^d$ later on. We take the opportunity and introduce the spaces already in their vector-valued forms.

Definition 1.2.1 (Lebesgue space)**.** Let $\Upsilon \subset \mathbb{R}^d$ with an associated measure space $(\Upsilon, \mathfrak{A}, \mu)$, and let $1 \leq p \leq \infty$.

(i) The space $\mathrm{L}^p(\Upsilon; \mu)$ consists of all μ-measurable functions $f\colon \Upsilon \to \mathbb{C}$ modulo equivalence μ-almost everywhere such that

$$\|f\|_{\mathrm{L}^p(\Upsilon;\mu)} := \left(\int_\Upsilon |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} \quad \text{if } 1 \leq p < \infty$$

or

$$\|f\|_{\mathrm{L}^\infty(\Upsilon;\mu)} := \operatorname*{ess\,sup}_{\mathrm{x}\in\Upsilon} |f(\mathrm{x})| \quad \text{if } p = \infty$$

is finite.

(ii) Let X be a Banach space. We define $\mathrm{L}^p(\Upsilon;\mu,X)$ to be the space of all μ-measurable functions $g\colon \Upsilon \to X$ such that $(\|\cdot\|_X \circ g) \in \mathrm{L}^p(\Upsilon;\mu)$.

We refer to [51, Ch. II] for a comprehensive treatment of general measurable vector-valued functions, see also [17, Ch. 1]. Note that if $X = \mathscr{L}(Y;Z)$ for some Banach spaces Y, Z, we generally suppose *strong* measurability, that is, measurability of the mapping $\Upsilon \ni \mathfrak{y} \mapsto f(\mathfrak{y})y \in Z$ for each $y \in Y$. This terminology is analogous to the distinction between uniform and strong continuity for operator-valued mappings. Of course, we follow the usual slight abuse of notation and call elements of the Lebesgue spaces *functions* instead of *equivalence classes of functions with respect to equality μ-almost everywhere*. If the measure space is clear from the context, we drop the reference to the measure. This is in particular the case if Υ is an open subset of Euclidean space $\mathbb{R}^d$, for which we always use the Lebesgue measure λ^d, including $\Upsilon = \mathbb{R}^d$.
It is well-known that $\left(\mathrm{L}^p(\Upsilon;\mu)\right)' = \mathrm{L}^{p'}(\Upsilon;\mu)$ for $1 \le p < \infty$, thus the $\mathrm{L}^p(\Upsilon;\mu)$ spaces are reflexive if $1 < p < \infty$. A similar result

$$\left(\mathrm{L}^p(\Upsilon;\mu,X)\right)' = \mathrm{L}^{p'}(\Upsilon;\mu,X') \tag{1.12}$$

is true for the vector-valued case if X' has the *Radon-Nikodym property*, cf. [51, Ch. IV, Thm. 1]. We quote the following corollary, first obtained by PHILLIPS, which is sufficient for our needs:

Theorem 1.2.2 ([51, Ch. IV.1, Cor. 2])**.** *Let $\Upsilon \subset \mathbb{R}^d$ with an associated* finite *measure space $(\Upsilon, \mathfrak{A}, \mu)$, let X be a Banach space and let $1 \le p \le \infty$. The spaces $\mathrm{L}^p(\Upsilon;\mu,X)$ are reflexive if and only if both $\mathrm{L}^p(\Upsilon;\mu)$ and X are reflexive, and in this case* (1.12) *is true.*

We next turn to continuous functions. A particularly interesting species of (uniformly) continuous functions are those which are Hölder-continuous, that is, functions f satisfying an estimate of the form

$$\|f(\mathrm{x}) - f(\mathrm{y})\| \leq C\|\mathrm{x} - \mathrm{y}\|^\alpha$$

with a constant C for some $\alpha \in (0, 1]$ and all x, y from their domain of definition Υ.

Definition 1.2.3 (Continuous functions and Hölder spaces)**.** Let Υ be a subset of $\mathbb{R}^d$ and let X be a Banach space. We define

$$\mathrm{C}(\Upsilon; X) := \Big\{ f\colon \Upsilon \to X\colon f \text{ is continuous and } \sup_{\mathrm{x}\in\Upsilon} \|f(\mathrm{x})\|_X < \infty \Big\}$$

with $\|\cdot\|_{\mathrm{C}(\Upsilon;X)} := \sup_{\mathrm{x}\in\Upsilon} \|f(\mathrm{x})\|_X$. For $\alpha \in (0, 1)$ we set the *Hölder spaces* to be

$$\mathrm{C}^\alpha(\Upsilon; X) := \Big\{ f \in \mathrm{C}(\Upsilon; X)\colon [f]_{\alpha,\Upsilon,X} := \sup_{\substack{\mathrm{x},\mathrm{y}\in\Upsilon \\ \mathrm{x}\neq\mathrm{y}}} \frac{\|f(\mathrm{x}) - f(\mathrm{y})\|_X}{\|\mathrm{x} - \mathrm{y}\|^\alpha} < \infty \Big\},$$

equipped with the norm $\|f\|_{\mathrm{C}^\alpha(\Upsilon;X)} := \|f\|_{\mathrm{C}(\Upsilon;X)} + [f]_{\alpha,\Upsilon,X}$. The definition for $\mathrm{C}^\alpha(\Upsilon; X)$ also makes sense for $\alpha = 1$ for which we obtain the space of Lipschitz-continuous bounded functions on Υ. We write $\mathrm{C}^{1-}(\Upsilon; X)$ instead of $\mathrm{C}^1(\Upsilon; X)$ for this case in order to not create confusion with the set of continuously differentiable functions introduced below.

The spaces $\mathrm{C}(\Upsilon; X)$ and $\mathrm{C}^\alpha(\Upsilon; X)$ for $\alpha \in (0, 1) \cup \{1-\}$ are Banach spaces [3, Ch. II.1.1].

Remark 1.2.4.

(i) It is clear that Hölder-continuous functions are uniformly continuous, and that all f satisfying $[f]_{\alpha,\Upsilon,X} \leq C$ for some fixed $C > 0$ share a common modulus of continuity. Thus, for $\Upsilon \neq \mathbb{R}^d$, every Hölder-continuous function $f \in \mathrm{C}^\alpha(\Upsilon; X)$ admits a unique (uniformly) continuous extension $\bar{f}$ to $\overline{\Upsilon}$ with $\|f\|_{\mathrm{C}^\alpha(\Upsilon;X)} = \|\bar{f}\|_{\mathrm{C}^\alpha(\overline{\Upsilon};X)}$. Vice-versa, for $\Upsilon_1, \Upsilon_2 \subset \mathbb{R}^d$ with $\Upsilon_1 \supseteq \Upsilon_2$ and a function $f \in \mathrm{C}^\alpha(\Upsilon_1; X)$, the restriction $g := f \restriction \Upsilon_2$ of f to Υ_2 is an element of $\mathrm{C}^\alpha(\Upsilon_2; X)$ with $\|g\|_{\mathrm{C}^\alpha(\Upsilon_2;X)} \leq \|f\|_{C^\alpha(\Upsilon_1;X)}$. We thus do not distinguish between $\mathrm{C}^\alpha(\Upsilon; X)$ and $\mathrm{C}^\alpha(\overline{\Upsilon}; X)$.

(ii) Since we have defined $\mathrm{C}^\alpha(\Upsilon; X)$ to consist of *bounded* functions, Hölder-continuity is in fact a *local* property. In other words, one may equivalently replace $[f]_{\alpha,\Upsilon,X}$ in Definition 1.2.3 by

$$[f]^*_{\alpha,\Upsilon,X} := \sup_{\substack{\mathrm{x},\mathrm{y}\in\Upsilon \\ 0<\|\mathrm{x}-\mathrm{y}\|<\varepsilon}} \frac{\|f(\mathrm{x}) - f(\mathrm{y})\|_X}{\|\mathrm{x}-\mathrm{y}\|^\alpha}$$

for some fixed $\varepsilon > 0$. Indeed, we have

$$\frac{\|f(\mathrm{x}) - f(\mathrm{y})\|_X}{\|\mathrm{x}-\mathrm{y}\|^\alpha} \leq 2\varepsilon^{-\alpha}\|f\|_{\mathrm{C}(\Upsilon;X)}$$

for all $\mathrm{x}, \mathrm{y} \in \Upsilon$ satisfying $\|\mathrm{x} - \mathrm{y}\| \geq \varepsilon$, so it suffices to have the behavior of f in small scales under control.

At this point we obtain the first important embedding result. Let X, Y be Banach spaces with $X \hookrightarrow Y$. While it is obvious that $\mathrm{C}^\alpha(\Upsilon; X) \hookrightarrow \mathrm{C}(\Upsilon; Y)$ for any set Υ and that $\mathrm{C}^\alpha(\Upsilon; X) \hookrightarrow \mathrm{C}^\beta(\Upsilon; Y)$ for $\alpha > \beta$ for bounded sets Υ, the following ingenious characterization of compact subsets of $\mathrm{C}(\Upsilon; Y)$ for *compact* sets Υ by Arzelà and Ascoli allows to derive even a compact embedding.

Theorem 1.2.5 (Arzelà-Ascoli Theorem, [103, Ch. III, §3])**.** *Let $\Upsilon \subset \mathbb{R}^d$ be compact and let X be a Banach space. Let Φ be a subset of $\mathrm{C}(\Upsilon; X)$. Then Φ is relatively compact in $\mathrm{C}(\Upsilon; X)$ if and only if the following two conditions are satisfied:*

(i) Φ *is* equicontinuous*, that is, for given $\varepsilon > 0$ and $\mathrm{x}_0 \in \Upsilon$ there exists $\delta > 0$ such that whenever $\mathrm{x} \in \mathbb{B}(\mathrm{x}_0, \delta) \cap \Upsilon$, then*

$$\|f(\mathrm{x}) - f(\mathrm{x}_0)\|_X < \varepsilon$$

for all $f \in \Phi$.

(ii) *For each* $\mathrm{x} \in \Upsilon$*, the set* $\Phi(\mathrm{x}) = \{f(\mathrm{x})\colon f \in \Phi\}$ *is relatively compact in* X*.*

If X is finite-dimensional, the second condition in Theorem 1.2.5 reduces to the requirement that the sets $\Phi(\mathrm{x})$ are merely bounded.

Corollary 1.2.6. *Assume that* X, Y *are Banach spaces with* $Y \hookrightarrow X$*. Let* $\Upsilon \subset \mathbb{R}^d$ *be compact and let* $\alpha, \beta \in [0,1) \cup \{1-\}$ *with* $\alpha > \beta$ *and* $\mathrm{C}^0(\Upsilon; X) \coloneqq \mathrm{C}(\Upsilon; X)$*. Then*

$$\mathrm{C}^\alpha(\Upsilon; Y) \hookrightarrow \mathrm{C}^\beta(\Upsilon; X).$$

Proof. We show that every sequence in the unit ball $\mathbb{B}(0,1)$ in $\mathrm{C}^\alpha(\Upsilon; Y)$ admits a subsequence converging in $\mathrm{C}^\beta(\Upsilon; X)$. The case $\beta = 0$ follows immediately from the definition of Hölder continuity for $\Phi = \mathbb{B}(0,1)$ in Theorem 1.2.5. For $\beta > 0$, let $(f_n) \subset \mathbb{B}(0,1)$. First, we extract a subsequence (f_{n_k}) which converges in $\mathrm{C}(\Upsilon; X)$. Then it remains to observe that

$$\begin{aligned}
[f_{n_k} - f_{n_\ell}]_{\beta,\Upsilon,X} &\le [f_{n_k} - f_{n_\ell}]_{\alpha,\Upsilon,X}^{\frac{\beta}{\alpha}} \cdot \left(2\|f_{n_k} - f_{n_\ell}\|_{\mathrm{C}(\Upsilon;X)}\right)^{1-\frac{\beta}{\alpha}} \\
&\le 2\|f_{n_k} - f_{n_\ell}\|_{\mathrm{C}(\Upsilon;X)}^{1-\frac{\beta}{\alpha}}
\end{aligned}$$

for all $k, \ell \in \mathbb{N}$, since $f_{n_k}, f_{n_\ell} \in \mathbb{B}(0,1)$. Hence, (f_{n_k}) is a Cauchy sequence in $\mathrm{C}^\beta(\Upsilon; X)$. □

Lastly, we state a duality theorem for the space $\mathrm{C}(\Upsilon) \coloneqq \mathrm{C}(\Upsilon; \mathbb{C})$ for compact Υ. We thus introduce the space of regular Borel measures on Υ, cf. [103, Ch. VII, §3, Ch. IX, §2, Rem. 3] or [43, Appendix C].

Definition 1.2.7 (Total variation (measure) and regular measures)**.** Let $\Upsilon \subset \mathbb{R}^d$ be compact. Let further $\mathfrak{B}_\Upsilon$ be the σ-algebra of Borel sets of Υ and consider a measure $\mu\colon \mathfrak{B}_\Upsilon \to \mathbb{C}$.

(i) We define the *total variation of* μ by $|\mu|\colon \mathfrak{B}_\Upsilon \to \mathbb{R}^+$ with

$$|\mu|(A) := \sup\Big\{ \sum_i |\mu(A_i)| \colon A_i \in \mathfrak{B}_\Upsilon,\ \bigcup_i A_i = A, \\ A_j \cap A_k = \emptyset \text{ when } j \neq k \Big\}.$$

This defines a positive measure $|\mu|$ on $\mathfrak{B}_\Upsilon$.

(ii) We say that μ is a *regular* measure if $|\mu(K)| < \infty$ for every compact set $K \in \mathfrak{B}_\Upsilon$ and in addition for all $A \in \mathfrak{B}_\Upsilon$ we have
   (i) $|\mu|(A) = \inf\{|\mu(V)| \colon V \in \mathfrak{B}_\Upsilon,\ V \text{ open with } V \supset A\}$ (outer regularity), and
   (ii) $|\mu|(A) = \sup\{|\mu(K)| \colon K \in \mathfrak{B}_\Upsilon,\ K \text{ compact with } K \subset A\}$ (inner regularity).

(iii) We lastly define

$$\mathcal{M}(\Upsilon) := \Big\{ \mu \colon \mathfrak{B}_\Upsilon \to \mathbb{C} \colon \mu \text{ is a regular measure} \Big\},$$
$$\|\mu\|_{\mathcal{M}(\Upsilon)} := |\mu|(\Upsilon).$$

The space $\mathcal{M}(\Upsilon)$ is a Banach space [103, Ch. IX, Thm. 4.1] with the following significance, a result which is also known as the Riesz *representation theorem*:

Theorem 1.2.8 ([103, Ch. IX, Thm. 4.2]). *Let $\Upsilon \subset \mathbb{R}^d$ be compact. Then the mapping*

$$\mathcal{M}(\Upsilon) \ni \mu \mapsto \Big[f \mapsto \int_\Upsilon f \,\mathrm{d}\mu \Big] \in \mathrm{C}(\Upsilon)',$$

defines an isometric isomorphism between $\mathcal{M}(\Upsilon)$ and $\mathrm{C}(\Upsilon)'$. Here, $\mathrm{d}\mu$ is to be understood via $\mathrm{d}\mu = h\,\mathrm{d}|\mu|$ for some $h \in \mathrm{L}^1(\Upsilon; |\mu|)$ with $|h| = 1$, $|\mu|$-a.e. on Υ.

It will be sufficient to consider the scalar-valued case in Theorem 1.2.8, since we will need the dual space of vector-valued continuous functions only for spaces of the form $\mathrm{C}(\Upsilon_1; \mathrm{C}(\Upsilon_2))$ for $\Upsilon_i \subset \mathbb{R}^{d_i}$, $i = 1, 2$. Due to $\mathrm{C}(\Upsilon_1 \times \Upsilon_2) \doteq \mathrm{C}(\Upsilon_1; \mathrm{C}(\Upsilon_2))$ we are able to fall back to Theorem 1.2.8 in this case, cf. also [5, Sect. 2], [103, Ch. IX, §6]. For the full vector-valued Riesz representation theorem, we refer to [83] and the references there.

Next, we add classical derivatives and thus turn to the definition of classically differentiable functions. We use standard multiindex notation for this, i.e., D^s for a multiindex $s = (s_1, \dots, s_d) \in \mathbb{N}_0^d$ stands for the iterated differential operator taking s_i times the partial derivative in the i-th coordinate direction, with $|s| := s_1 + \dots + s_d$. The order of taking partial derivatives will not be of importance since we only apply such differential operators to sufficiently smooth functions for which Schwarz's theorem holds.

Definition 1.2.9 (Continuously differentiable & smooth functions)**.** Let Υ be an *open* subset of $\mathbb{R}^d$ and let X be a Banach space. We define for $k \in \mathbb{N}_0$

$$\mathrm{C}^k(\Upsilon; X) := \Big\{ f \colon \Upsilon \to X : D^s f \in C(\Upsilon; X) \text{ for all } s \in \mathbb{N}_0^d,\ |s| \le k \Big\},$$

equipped with the norm $\|f\|_{\mathrm{C}^k(\Upsilon;X)} = \sum_{|s|\le k} \|D^s f\|_{\mathrm{C}(\Upsilon;X)}$. This is consistent with $\mathrm{C}^0(\Upsilon; X) = \mathrm{C}(\Upsilon; X)$. We moreover set, for $k \neq 0$,

$$\begin{aligned} &\mathrm{C}^{k-}(\Upsilon; X) \\ &\quad := \Big\{ f \in \mathrm{C}^{k-1}(\Upsilon; X) \colon D^s f \in \mathrm{C}^{1-}(\Upsilon; X) \text{ for all } s \in \mathbb{N}_0^d,\ |s| = k-1 \Big\} \end{aligned}$$

and

$$\begin{aligned} &\mathrm{C}^{r}(\Upsilon; X) \\ &\quad := \Big\{ f \in \mathrm{C}^{\lfloor r \rfloor}(\Upsilon; X) \colon D^s f \in \mathrm{C}^{r-\lfloor r \rfloor}(\Upsilon; X) \text{ for all } s \in \mathbb{N}_0^d,\ |s| = \lfloor r \rfloor \Big\} \end{aligned}$$

for $r > 0$ with $r \notin \mathbb{N}$. Lastly, let us set

$$\mathrm{C}^\infty(\Upsilon; X) := \bigcap_{k\in\mathbb{N}} \mathrm{C}^k(\Upsilon; X).$$

and

$$\mathrm{C}_c^k(\Upsilon; X) := \left\{ f \in \mathrm{C}^k(\Upsilon; X) \colon \operatorname{supp} f \Subset \Upsilon \right\},$$

for $k \in \mathbb{N}_0 \cup \{\infty\}$.

It is known that $\mathrm{C}^s(\Upsilon; X)$ for $s \in \mathbb{R}_0^+$ and $\mathrm{C}^{k-}(\Upsilon; X)$ for $k \in \mathbb{N}$ are Banach spaces [2, Ch. 1.29], where the proof works completely analogous to the scalar-valued case.

Remark 1.2.10.

(i) As above, we will write $\mathrm{C}^s(\Upsilon)$ and $\mathrm{C}_c^k(\Upsilon)$ instead of $\mathrm{C}^s(\Upsilon; \mathbb{C})$ and $\mathrm{C}_c^k(\Upsilon; \mathbb{C})$ for the case $X = \mathbb{C}$.

(ii) It also makes sense to talk about the above function spaces on the set $\overline{\Upsilon}$. We then mean that the functions are elements of the corresponding space on Υ and all involved derivatives are uniformly continuous on Υ and thus admit a unique extension to $\overline{\Upsilon}$. This is, in fact, only an assumption on the highest derivatives due to the boundedness of all occurring derivatives built into the definition.

The space of *test functions* $\mathrm{C}_c^\infty(\Upsilon)$ is of particular interest and turns out to be a locally convex space with various good properties ([153, Ch. I.1]) often denoted by $\mathscr{D}(\Upsilon)$ in literature, a notation which we adopt for this short paragraph. It owes its rise to fame to the systematic Fields-medal-earning exploration of *distributions*, the topological dual space $\mathscr{D}'(\Upsilon) := \mathscr{L}(\mathscr{D}(\Upsilon); \mathbb{C})$, by Schwartz [132,133]. Since $\mathscr{D}'(\Upsilon)$ will serve as the common superset for the objects introduced below, we briefly agree on the following, where Υ is an arbitrary open set:

- The *restriction* $\mathfrak{R}_\Upsilon$ of a distribution $\varphi \in \mathscr{D}'(\mathbb{R}^d)$ to $\mathscr{D}'(\Upsilon)$ is defined as the adjoint operator of the extension by zero $\mathfrak{E}^0 \colon \mathscr{D}(\Upsilon) \to \mathscr{D}(\mathbb{R}^d)$, i.e.,

  $$\langle \mathfrak{R}_\Upsilon \varphi, f \rangle := \langle \varphi, \mathfrak{E}^0 f \rangle \quad \text{for all } \varphi \in \mathscr{D}'(\mathbb{R}^d),\ f \in \mathscr{D}(\Upsilon).$$

  Note that $\mathfrak{R}_\Upsilon \mathscr{D}'(\mathbb{R}^d) \subsetneq \mathscr{D}'(\Upsilon)$, see e.g. [142, Ch. 6.7].

- The expression D^s for some multiindex $s \in \mathbb{N}_0^d$ is to be understood in the distributional sense.
- We abbreviate the tedious explanation that a distribution f on Υ is regular and the inducing function (which we also call f) is an element of $\mathrm{L}^p(\Upsilon)$ by the simple statement $\|f\|_{\mathrm{L}^p(\Upsilon)} < \infty$.

We refer to [131, Ch. 6] and [142] as complimentary references about $\mathscr{D}(\Upsilon)$ and $\mathscr{D}'(\Upsilon)$.

Remark 1.2.11. We will encounter, albeit seldom, situations in which we need sets of functions on sets $\Upsilon \subseteq \mathbb{R}^d$ as defined in Definitions 1.2.3 (Hölder-continuous functions) and 1.2.9 (Smooth functions) *without* the boundedness properties holding globally. Of course, we merely talk about vector spaces instead of normed spaces in this case. The notation of choice to refer to these sets is $\mathrm{C}^s_{\mathrm{loc}}(\Upsilon; X)$ for $s \in \mathbb{R}_0^+$, which is compliant with the usual interpretation of the *loc* subscript: A function f is in $\mathrm{C}^s_{\mathrm{loc}}(\Upsilon; X)$ whenever $f \in \mathrm{C}^s(K; X)$ for every compact subset $K \Subset \Upsilon$.

1.2.1 Function spaces on $\mathbb{R}^d$

Next, we introduce function spaces of functions exhibiting weak differentiability properties, both of integer and noninteger orders of smoothness. For a concise treatment, we begin with the versions of the spaces defined on Euclidean space. The first definition is a preliminary one, serving as a foundation for the coming ones.

Definition 1.2.12 (Schwartz space)**.** We define the *Schwartz space* of rapidly decaying functions

$$\mathscr{S}(\mathbb{R}^d) := \left\{ f \in \mathrm{C}^\infty(\mathbb{R}^d) \colon \forall \alpha, \beta \in \mathbb{N}_0^n \colon \sup_{\mathrm{x} \in \mathbb{R}^d} \left| \mathrm{x}^\alpha D^\beta f(\mathrm{x}) \right| < \infty \right\}.$$

Of course, the fundamental property of the Schwartz space is that the Fourier transform $\mathscr{F}$ acts as an automorphism on this space. However, we do not need its precise topological structure. We only point out that it shares many nice properties with $\mathrm{C}_c^\infty(\mathbb{R}^d)$, cf. [153, Ch. VI.1] or [131, Ch. 7], and is related to the previously introduced spaces of smooth functions via

$$\mathrm{C}_c^\infty(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d) \subset \mathrm{C}^\infty(\mathbb{R}^d).$$

From the Schwartz space we derive its topological dual space $\mathscr{S}'(\mathbb{R}^d) = \mathscr{L}(\mathscr{S}(\mathbb{R}^d); \mathbb{C})$, also known as the set of *tempered distributions*, as a common superset for the set of admissible objects in the following definitions. These follow [24] and [146].

Definition 1.2.13 (Sobolev spaces on $\mathbb{R}^d$)**.** Let $1 \le p \le \infty$ and $k \in \mathbb{N}_0$. We define

$$\mathrm{W}^{k,p}(\mathbb{R}^d) := \left\{ f \in \mathscr{S}'(\mathbb{R}^d) \colon \|f\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} := \Bigg(\sum_{|s| \le k} \|D^s f\|_{\mathrm{L}^p(\mathbb{R}^d)}^p \Bigg)^{\frac{1}{p}} < \infty \right\}$$

for $p < \infty$, and

$$\mathrm{W}^{k,\infty}(\mathbb{R}^d) := \left\{ f \in \mathscr{S}'(\mathbb{R}^d) \colon \|f\|_{\mathrm{W}^{k,\infty}(\mathbb{R}^d)} := \max_{|s| \le k} \|D^s f\|_{\mathrm{L}^\infty(\mathbb{R}^d)} < \infty \right\}.$$

In particular, we re-obtain $\mathrm{L}^p(\mathbb{R}^d) = \mathrm{W}^{0,p}(\mathbb{R}^d)$ for $1 \le p \le \infty$.

The next type of spaces introduces non-integer orders of smoothness, formulated via the Fourier transform $\mathscr{F}$.

Definition 1.2.14 (Bessel potential spaces on $\mathbb{R}^d$)**.** Let $-\infty < s < \infty$ and $1 < p < \infty$. We define

$$\begin{aligned} \mathrm{H}^{s,p}(\mathbb{R}^d) &:= \left\{ f \in \mathscr{S}'(\mathbb{R}^d) \colon \mathscr{F}^{-1}\big((1 + |\cdot|^2)^{\frac{s}{2}} \mathscr{F} f\big) \in \mathrm{L}^p(\mathbb{R}^d) \right\}, \\ \|f\|_{\mathrm{H}^{s,p}(\mathbb{R}^d)} &:= \big\|\mathscr{F}^{-1}\big((1 + |\cdot|^2)^{\frac{s}{2}} \mathscr{F} f\big)\big\|_{\mathrm{L}^p(\mathbb{R}^d)}. \end{aligned}$$

The Bessel potential spaces coincide with the Sobolev spaces up to equivalent norms if $s \in \mathbb{N}_0$, that is, $\mathrm{H}^{k,p}(\mathbb{R}^d) \doteq \mathrm{W}^{k,p}(\mathbb{R}^d)$ for all $k \in \mathbb{N}_0$

and $1 < p < \infty$, see [146, Ch. 2.3.3]. We note that the above spaces are Banach spaces and that both $\mathrm{C}_c^\infty(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$ are dense subsets in both scales [146, Ch. 2.3.2]. Moreover, the Bessel scale is stable under duality, i.e., for all admissible s and p as in Definition 1.2.14, we have [146, Ch. 2.6.1]

$$\big(\mathrm{H}^{s,p}(\mathbb{R}^d)\big)' = \mathrm{H}^{-s,p'}(\mathbb{R}^d). \tag{1.13}$$

In particular, each space $\mathrm{H}^{s,p}(\mathbb{R}^d)$ is reflexive. We next give embedding results for the freshly introduced spaces.

Theorem 1.2.15 ([146, Ch. 2.8.1])**.** *Let $1 < p \leq q < \infty$. Then we have the following continuous embeddings:*

$$\mathrm{H}^{s,p}(\mathbb{R}^d) \hookrightarrow \mathrm{H}^{t,q}(\mathbb{R}^d) \quad \text{if} \quad s - \frac{d}{p} \geq t - \frac{d}{q}$$

and

$$\mathrm{H}^{\frac{d}{p}+\alpha,p}(\mathbb{R}^d) \hookrightarrow \mathrm{C}^\alpha(\mathbb{R}^d) \quad \text{for} \quad \alpha \notin \mathbb{N}_0.$$

Although we formulated, at least for the second embedding, only a special case of the ones in [146], the importance of these embeddings can hardly be overstated. Essentially, the first one allows to trade smoothness for integrability, while the second one shows a sufficiently high degree of weak differentiability in fact implies strong differentiability in a good sense, albeit of course only to a smaller degree. Both results will be of fundamental importance in the later chapters of this work. We collect two special cases of particular interest which immediately follow from Theorem 1.2.15 and $\mathrm{H}^{k,p}(\mathbb{R}^d) = \mathrm{W}^{k,p}(\mathbb{R}^d)$ for $k \in \mathbb{N}_0$ and $1 < p < \infty$. Let us set the *ℓ-Sobolev conjugate* for $\ell \in \mathbb{N}$ and $\ell p < d$ to

$$p^\star(\ell) := \frac{dp}{d - \ell p}$$

with the convention that $p^\star := p^\star(1)$.

Corollary 1.2.16. *Let* $1 < p \leq q < \infty$ *and* $k, \ell \in \mathbb{N}$ *with* $0 < \ell \leq k$. *Then*

$$\mathrm{W}^{k,p}(\mathbb{R}^d) \hookrightarrow \mathrm{W}^{k-\ell,q}(\mathbb{R}^d) \quad \textit{for} \quad \begin{cases} p \leq q \leq p^\star(\ell) & \textit{if } \ell p < d, \\ p \leq q < \infty & \textit{if } \ell p = d. \end{cases} \tag{1.14}$$

and

$$\mathrm{W}^{k,p}(\mathbb{R}^d) \hookrightarrow \mathrm{C}^\alpha(\mathbb{R}^d) \quad \textit{for} \quad \alpha = k - \frac{d}{p} \notin \mathbb{N} \quad \textit{if } kp > d. \tag{1.15}$$

Both embeddings, here obtained as corollaries of more abstract results, have a rich history and may, of course, also be derived on their own. The case $k = 1$ is of special interest. There, the Hölder embedding (1.15) in its original form is due to Morrey [120], with the famous *Morrey inequality*

$$|f(\mathrm{x}) - f(\mathrm{y})| \leq C|\mathrm{x} - \mathrm{y}|^{1-\frac{d}{p}} \|\nabla f\|_{L^p(\mathbb{R}^d)} \quad \text{for all } f \in \mathrm{C}^1(\mathbb{R}^d) \cap \mathrm{L}^p(\mathbb{R}^d).$$

The embedding (1.14) for $k = 1$ on the other hand follows from the equally famous Gagliardo-Nirenberg-Sobolev *inequality*

$$\|f\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq C\|\nabla f\|_{\mathrm{L}^p(\mathbb{R}^d)} \quad \text{for all } f \in \mathrm{C}^1_c(\mathbb{R}^d)$$

whose proof as given by Nirenberg [122] is remarkable because of its simplicity: using essentially only the fundamental theorem of calculus and elementary inequalities, one first establishes

$$\|f\|_{\mathrm{L}^{\frac{d}{d-1}}(\mathbb{R}^d)} \leq C\|\nabla f\|_{\mathrm{L}^1(\mathbb{R}^d)}.$$

The general case then follows from inserting the function $|f|^\gamma$ for suitable $\gamma > 1$ into this estimate (an analogous technique also allows to derive the Gagliardo-Nirenberg *interpolation inequality*). This “proof” moreover shows that (1.14) is in fact also true for $p = 1$. We did not include this case here because we arrived via Theorem 1.2.15 and, unfortunately, $\mathrm{H}^{k,1}(\mathbb{R}^d) \neq \mathrm{W}^{k,1}(\mathbb{R}^d)$ in general, cf. [146, Ch. 2.3.3 Rem. 5].

Remark 1.2.17. The dutiful reader may have noticed that the Hölder spaces on $\mathbb{R}^d$ in [146, Ch. 2.7.1] are defined via a space of continuous functions on $\mathbb{R}^d$ obtained by the closure of the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ w.r.t. $\|\cdot\|_{\infty,\mathbb{R}^d}$. But this closure yields $\mathrm{C}_0(\mathbb{R}^d)$, the bounded continuous functions on $\mathbb{R}^d$ *vanishing at infinity*, i.e., Hölder functions on $\mathbb{R}^d$ in [146] are defined to vanish at infinity. In this sense, Theorem 1.2.15 and Corollary 1.2.16 are not formulated as strictly as they could be. However, since we will not need the behavior at infinity of functions on $\mathbb{R}^d$, we decided to stick with the common Hölder spaces as in Definition 1.2.3. In view of the previous historical explanations above, note that it is indeed well-known that $\mathrm{W}^{1,p}(\mathbb{R}^d) \hookrightarrow \mathrm{C}_0(\mathbb{R}^d)$ for $p > d$, cf. for example [143, Ch. 9].

Lastly we consider interpolation for the $\mathrm{H}^{s,p}(\mathbb{R}^d)$ scale, which includes Sobolev– and Lebesgue spaces.

Theorem 1.2.18 ([146, Ch. 2.4.2])**.** *Let $\theta \in (0,1)$ and let $s_0, s_1 \in \mathbb{R}$ and $1 < p_0, p_1 < \infty$. Then*

$$\big[\mathrm{H}^{s_0,p_0}(\mathbb{R}^d), \mathrm{H}^{s_1,p_1}(\mathbb{R}^d)\big]_\theta = \mathrm{H}^{s,p}(\mathbb{R}^d)$$

for $s = (1-\theta)s_0 + \theta s_1$ and $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

Particular cases of Theorem 1.2.18 are, for $\theta \in (0,1)$ and $1 < p_0, p_1 < \infty$,

$$\big[\mathrm{L}^{p_0}(\mathbb{R}^d), \mathrm{L}^{p_1}(\mathbb{R}^d)\big]_\theta = \mathrm{L}^{p}(\mathbb{R}^d) \quad \text{for} \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \tag{1.16}$$

and, if $k, \ell \in \mathbb{N}_0$,

$$\big[\mathrm{W}^{k,p_0}(\mathbb{R}^d), \mathrm{W}^{k+\ell,p_1}(\mathbb{R}^d)\big]_\theta = \mathrm{H}^{k+\ell\theta,p}(\mathbb{R}^d) \quad \text{for } \theta \in (0,1)$$

and p as in (1.16).

1.2.2 Function spaces on a domain

We transfer the function spaces defined in the previous section to their versions on a domain $\Lambda \subset \mathbb{R}^d$.

Definition 1.2.19. We call a set $\Lambda \subset \mathbb{R}^d$ a *domain* if it is open, connected and nonempty.

We fix $\Lambda \subset \mathbb{R}^d$ to be a domain for the rest of this section. Generally, a domain Λ may be of more or less arbitrary unpleasantness regarding smoothness or regularity of its boundary, which ultimately poses serious difficulties when working with function spaces defined on Λ. Since we strive for generality, we try to put as little restriction on Λ as possible, thereby following recent developments in the PDE community. To start with the most fundamental problem, the definition for the $\mathrm{W}^{k,p}(\Lambda)$ scale one encounters most often is the following:

Definition 1.2.20. Let $1 \leq p \leq \infty$ and $k \in \mathbb{N}_0$. We define

$$\mathrm{W}^{k,p}(\Lambda) := \Big\{ f \in \mathrm{L}^p(\Lambda) \colon \|f\|_{\mathrm{W}^{k,p}(\Lambda)} := \Big(\sum_{|s| \leq k} \|D^s f\|_{\mathrm{L}^p(\Lambda)}^p \Big)^{\frac{1}{p}} < \infty \Big\}$$

for $p < \infty$, and

$$\mathrm{W}^{k,\infty}(\Lambda) := \Big\{ f \in \mathrm{L}^\infty(\Lambda) \colon \|f\|_{\mathrm{W}^{k,\infty}(\Lambda)} := \max_{|s| \leq k} \|D^s f\|_{\mathrm{L}^\infty(\Lambda)} < \infty \Big\}.$$

In particular, we re-obtain $\mathrm{L}^p(\Lambda) = \mathrm{W}^{0,p}(\Lambda)$.

Since $\mathrm{W}^{k,p}(\Lambda)$ is isometrically isomorphic to a finite copy of $\mathrm{L}^p(\Lambda)$ spaces via the identification $f \mapsto (f, (D^s f)_{|s| \leq 1}, \ldots, (D^s f)_{|s| \leq k})$, the space is both a Banach space and reflexive.
In view of the definition of the Bessel potential spaces $\mathrm{H}^{s,p}(\mathbb{R}^d)$, it is clear that such an intrinsic definition for $\mathrm{H}^{s,p}(\Lambda)$ will be difficult and generally

impossible to obtain.[1] Hence, we define these spaces by restriction:

Definition 1.2.21. Let $-\infty < s < \infty$ and $1 < p < \infty$. We define

$$\mathrm{H}^{s,p}(\Lambda) \coloneqq \mathfrak{R}_\Lambda \mathrm{H}^{s,p}(\mathbb{R}^d), \quad \|f\|_{\mathrm{H}^{s,p}(\Lambda)} \coloneqq \inf_{\substack{g \in \mathrm{H}^{s,p}(\mathbb{R}^d) \\ f = \mathfrak{R}_\Lambda g}} \|g\|_{\mathrm{H}^{s,p}(\mathbb{R}^d)}.$$

From the definition one identifies $\mathrm{H}^{s,p}(\Lambda)$ as quotient spaces, which makes them already Banach spaces and reflexive due to $\mathrm{H}^{s,p}(\mathbb{R}^d)$ being so.

We have noted above that $\mathrm{H}^{k,p}(\mathbb{R}^d) = \mathrm{W}^{k,p}(\mathbb{R}^d)$ for $k \in \mathbb{N}_0$ and $1 < p < \infty$. A major point to note now is that $\mathrm{H}^{k,p}(\Lambda) \neq \mathrm{W}^{k,p}(\Lambda)$ in general when $d > 1$, that is, $\mathrm{W}^{k,p}(\Lambda) \neq \mathfrak{R}_\Lambda \mathrm{W}^{k,p}(\mathbb{R}^d)$ if we do not pose any further assumption on Λ. If $\mathrm{H}^{k,p}(\Lambda)$ was equal to $\mathrm{W}^{k,p}(\Lambda)$ for any domain Λ, we would have to be able to approximate a function in $\mathrm{W}^{k,p}(\Lambda)$ by (restrictions to Λ of) a sequence of $\mathrm{C}_c^\infty(\mathbb{R}^d)$ functions, since $\mathrm{C}_c^\infty(\mathbb{R}^d)$ is dense in $\mathrm{H}^{k,p}(\mathbb{R}^d)$. But then, due to the nature of the $\mathrm{W}^{k,p}(\Lambda)$-norm, defects in Λ of measure zero will be "looked over" by the smooth approximations. The usual counterexample is that of a sliced disk $\mathbb{B}(0,1) \setminus \{[0,1) \times \{0\}\}$ in $\mathbb{R}^2$, see [56, Ex. 1.1.10] for an explicit calculation.

Of course, we will need to make sure that indeed $\mathrm{H}^{k,p}(\Lambda) = \mathrm{W}^{k,p}(\Lambda)$. By the above example, we see that the obstacle to overcome is to enforce that every $\mathrm{W}^{k,p}(\Lambda)$ function is indeed the restriction of a $\mathrm{W}^{k,p}(\mathbb{R}^d)$ one. The abstract condition to pose, which will in fact also open the door to transferring Theorem 1.2.15 and many other properties to function spaces on Λ, is the following.

Definition 1.2.22. Let $k \in \mathbb{N}$ and $1 \leq p \leq \infty$. A domain $\Lambda \subset \mathbb{R}^d$ is called $\mathrm{W}^{k,p}$-*extension domain* if there exists a linear continuous extension operator $\mathfrak{E}_\Lambda \colon \mathrm{W}^{k,p}(\Lambda) \to \mathrm{W}^{k,p}(\mathbb{R}^d)$, that is, an operator which acts as a right-inverse for the restriction operation $\mathfrak{R}_\Lambda$ on $\mathrm{W}^{k,p}(\mathbb{R}^d)$. We say that an extension operator $\mathfrak{E}_\Lambda$ is *universal* if it is an extension operator for all

[1] The situation is different for the "other" scale of Sobolev-type spaces with noninteger smoothness, the Slobodeckij spaces $\mathrm{W}^{s,p}(\Lambda)$, cf. [94, Thm. 1.1, Ch. 5].

$k \in \mathbb{N}_0$ and all $1 < p < \infty$ and call Λ a *universal extension domain* in this case.

Note that we have required only the range $1 < p < \infty$ for an extension operator to be called *universal*, mainly because this will be the range with which we work later on. In any case, coming back to the considerations above, *if* Λ is a $\mathrm{W}^{k,p}$-extension domain, then $\mathrm{W}^{k,p}(\Lambda)$ indeed agrees with the quotient space obtained by the restriction $\mathfrak{R}_\Lambda \mathrm{W}^{k,p}(\mathbb{R}^d)$:

Lemma 1.2.23. *Let $k \in \mathbb{N}$ and $1 < p < \infty$ and let $\Lambda \subset \mathbb{R}^d$ be a $\mathrm{W}^{k,p}$-extension domain. Then the spaces $\mathfrak{R}_\Lambda \mathrm{W}^{k,p}(\mathbb{R}^d) = \mathrm{H}^{k,p}(\Lambda)$ in the sense of Definition 1.2.21 and $\mathrm{W}^{k,p}(\Lambda)$ coincide up to equivalent norms.*

Proof. We show this as a brief application for the use of the extension operator $\mathfrak{E}_\Lambda$. Let $f \in \mathrm{W}^{k,p}(\Lambda)$ be given. Then f is contained in the image $\mathfrak{R}_\Lambda \mathrm{W}^{k,p}(\mathbb{R}^d) = \mathrm{H}^{k,p}(\Lambda)$ by definition of the extension operator: $f = \mathfrak{R}_\Lambda \mathfrak{E}_\Lambda f$ with $\mathfrak{E}_\Lambda f \in \mathrm{W}^{k,p}(\mathbb{R}^d)$. The definition of the quotient norm on $\mathrm{H}^{k,p}(\Lambda)$ as in Definition 1.2.21 now shows that

$$\|f\|_{\mathrm{H}^{k,p}(\Lambda)} \leq \|\mathfrak{E}_\Lambda f\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \leq \|\mathfrak{E}_\Lambda\|_{\mathscr{L}(\mathrm{W}^{k,p}(\Lambda);\mathrm{W}^{k,p}(\mathbb{R}^d))} \|f\|_{\mathrm{W}^{k,p}(\Lambda)}.$$

Now let f be from $\mathrm{H}^{k,p}(\Lambda)$. Then there exists at least one function $g \in \mathrm{W}^{k,p}(\mathbb{R}^d)$ such that $f = \mathfrak{R}_\Lambda g$. For every such function, the definitions of the $\mathrm{W}^{k,p}(\mathbb{R}^d)$- and $\mathrm{W}^{k,p}(\Lambda)$-norms immediately show that $\|f\|_{\mathrm{W}^{k,p}(\Lambda)} \leq \|g\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)}$. Hence, again by the definition of the quotient norm on $\mathrm{H}^{k,p}(\Lambda)$, we have $\|f\|_{\mathrm{W}^{k,p}(\Lambda)} \leq \|f\|_{\mathrm{H}^{k,p}(\Lambda)}$. □

Remark 1.2.24. Extension operators for various classes of function spaces and their properties constitute an active and rather fascinating field of research. Concerning *cutting-edge*, Rogers [130], building upon the work of Jones in [93], found an *universal* extension operator for so-called (ε, δ)-domains. These domains, also called *uniform domains* if they are bounded, are a very general class of domains which seems to be close to best-possible in the sense that in space dimension $d = 2$ every simply connected $\mathrm{W}^{k,p}$-extension domain must be a (ε, δ)-domain. It is, however, known that for each space dimension $d \geq 3$ there are bounded $\mathrm{W}^{1,p}$-extension domains which are not uniform (see [151]). We make use of extension operators from now on, and in the later Chapters 1.2.4, 1.2.3 and 1.3, but will not go into much more details and refer to [2, Ch. 5], [31], [30, Ch. 6 and 7] and [112, Ch. 1.5.1] for more general information.

Let us now turn to further consequences of the existence of extension operators. The next theorem transfers interpolation properties to the spaces on Λ and relates the spaces $\mathrm{H}^{s,p}(\Lambda)$ for noninteger s and $\mathrm{W}^{k,p}(\Lambda)$ for $k \in \mathbb{N}$ via an universal extension operator $\mathfrak{E}_\Lambda$. This works by virtue of a double application of Corollary 1.1.6, using the universal extension operator acting as a coretraction. The first application is done in the following lemma, which justifies the name *universal* extension operator further:

Lemma 1.2.25. *Let $\Lambda \subset \mathbb{R}^d$ be a domain and let $\mathfrak{E}_\Lambda$ be a universal extension operator. Then $\mathfrak{E}_\Lambda$ is also an $\mathrm{H}^{s,p}$-extension operator for $s \in \mathbb{R}_0^+$ and $1 < p < \infty$.*

Proof. We already know by assumption that $\mathfrak{E}_\Lambda$ is an $\mathrm{H}^{k,p}$-extension operator for every $k \in \mathbb{N}_0$, so let $s \in \mathbb{R}^+ \setminus \mathbb{N}$ and set $\theta = s - \lfloor s \rfloor$. By Theorem 1.2.18 and Corollary 1.1.6 applied to the retraction-coretraction pair $\mathfrak{R}_\Lambda$, $\mathfrak{E}_\Lambda$, we have

$$\begin{aligned}\left[\mathrm{W}^{\lfloor s \rfloor,p}(\Lambda), \mathrm{W}^{\lceil s \rceil,p}(\Lambda)\right]_\theta &\doteq \mathfrak{R}_\Lambda\left[\mathrm{W}^{\lfloor s \rfloor,p}(\mathbb{R}^d), \mathrm{W}^{\lceil s \rceil,p}(\mathbb{R}^d)\right]_\theta \\ &= \mathfrak{R}_\Lambda \mathrm{H}^{s,p}(\mathbb{R}^d) = \mathrm{H}^{s,p}(\Lambda).\end{aligned}$$

But then, by the fundamental properties of the interpolation functor $\mathcal{F}_\mathbb{C}$, the operators $\mathfrak{E}_\Lambda$ and $\mathfrak{R}_\Lambda$ extend to continuous linear mappings $\mathcal{F}_\mathbb{C}(\mathfrak{E}_\Lambda) \in \mathscr{L}(\mathrm{H}^{s,p}(\Lambda); \mathrm{H}^{s,p}(\mathbb{R}^d))$ and $\mathcal{F}_\mathbb{C}(\mathfrak{R}_\Lambda) \in \mathscr{L}(\mathrm{H}^{s,p}(\mathbb{R}^d); \mathrm{H}^{s,p}(\Lambda))$ such that $\mathcal{F}_\mathbb{C}(\mathfrak{E}_\Lambda)$ is still a right-inverse for $\mathcal{F}_\mathbb{C}(\mathfrak{R}_\Lambda)$. In other words: $\mathfrak{E}_\Lambda$ extends to an $\mathrm{H}^{s,p}$-extension operator. □

Now the general interpolation theorem for our function spaces on domains follows immediately from the second strike of Corollary 1.1.6 and Theorem 1.2.18.

Theorem 1.2.26. *Let $\Lambda \subset \mathbb{R}^d$ be a universal extension domain, let $\theta \in (0,1)$ as well as $s_0, s_1 \in \mathbb{R}_0^+$ and $1 < p_0, p_1 < \infty$. Then*

$$\left[\mathrm{H}^{s_0,p_0}(\Lambda), \mathrm{H}^{s_1,p_1}(\Lambda)\right]_\theta = \mathrm{H}^{s,p}(\Lambda)$$

for $s = (1-\theta)s_0 + \theta s_1$ and $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

Next, we establish the analogues of the embeddings in Theorem 1.2.15 for domains Λ. We directly state the results for bounded domains. A new phenomenon occurring here is that smoothness can not only be traded for integrability, but "giving" just an arbitrary amount of smoothness yields even *compactness* of the embeddings. The order of results is turned around, compared to Theorem 1.2.15 and Corollary 1.2.16, recall also the definition of the Sobolev conjugate $p^\star$ there.

Theorem 1.2.27. *Let Λ be a* bounded *universal extension domain and let $1 < p, q < \infty$ and $k, \ell \in \mathbb{N}$ with $0 < \ell \leq k$. Then we have the following continuous embeddings:*

$$\mathrm{W}^{k,p}(\Lambda) \hookrightarrow \mathrm{W}^{k-\ell,q}(\Lambda) \quad \textit{for} \quad \begin{cases} q \leq p^\star(\ell) & \textit{if } \ell p < d, \\ q < \infty & \textit{if } \ell p = d, \end{cases} \tag{1.17}$$

and

$$\mathrm{W}^{k,p}(\Lambda) \hookrightarrow \mathrm{C}^\alpha(\Lambda) \quad \textit{for} \quad \alpha = k - \frac{d}{p} \notin \mathbb{N} \quad \textit{if} \quad kp > d.$$

Moreover, the embedding in (1.17) *is* compact *for $q < p^\star(\ell)$ if $\ell p < d$ and for all $q < \infty$ if $\ell p = d$.*

Proof. The embeddings are transferred from Corollary 1.2.16 via their factorization by virtue of the extension property, here exemplarily displayed for the first one:

$$\mathrm{W}^{k,p}(\Lambda) \xrightarrow{\mathfrak{E}_\Lambda} \mathrm{W}^{k,p}(\mathbb{R}^d) \hookrightarrow \mathrm{W}^{k-\ell,q}(\mathbb{R}^d) \xrightarrow{\mathfrak{R}_\Lambda} \mathrm{W}^{k-\ell,q}(\Lambda).$$

Since we have assumed Λ to be bounded, we can dispose of the strict limits for the integrability exponents. Lastly, the compactness properties of the embeddings for the non-limit cases are an instance of the Rellich-Kondrachov theorem. We refer to [112, Ch. 1.4.6 Thm. 2]. □

Corollary 1.2.28. *Let $\Lambda \subset \mathbb{R}^d$ be a universal extension domain and let $1 < p, q < \infty$ and $s, t \in \mathbb{R}_0^+$ with $s \geq t$. Then*

$$\mathrm{H}^{s,p}(\Lambda) \hookrightarrow \mathrm{H}^{t,q}(\Lambda) \quad \textit{for} \quad s - \frac{d}{p} \geq t - \frac{d}{q} \tag{1.18}$$

and

$$\mathrm{H}^{\frac{d}{p}+\alpha,p}(\Lambda) \hookrightarrow \mathrm{C}^\alpha(\Lambda) \quad \textit{for} \quad \alpha \notin \mathbb{N}_0.$$

If $s > t$, the embedding (1.18) *is even* compact.

Proof. The embeddings follow from Lemma 1.2.25, analogously to Theorem 1.2.27. It remains to show the compactness property of (1.18) for $s > t$, for which we of course use the already established compact embedding from Theorem 1.2.27. Similarly to the proof of Lemma 1.2.25, we write $\mathrm{H}^{s,p}(\Lambda)$ as an interpolation space between Sobolev spaces which are compactly embedded into each other. Thus, Lemma 1.1.8 shows that

$$\mathrm{H}^{s,p}(\Lambda) = \left[\mathrm{L}^p(\Lambda), \mathrm{W}^{\lceil s\rceil,p}(\Lambda)\right]_{\frac{s}{\lceil s\rceil}} \overset{c}{\hookrightarrow} \left[\mathrm{L}^p(\Lambda), \mathrm{W}^{\lceil s\rceil,p}(\Lambda)\right]_{\frac{r}{\lceil s\rceil}} = \mathrm{H}^{r,p}(\Lambda)$$

for all $0 \leq r < s$. Now given $t < s$, it remains to apply the embedding (1.18) with $\mathrm{H}^{r,p}(\Lambda)$ for $t \leq r < s$ instead of $\mathrm{H}^{s,p}(\Lambda)$. □

Remark 1.2.29.

(i) The actual definition of regularity classes for domains $\Lambda \subset \mathbb{R}^d$ and the validation that they have the extension property will be done in Chapter 1.3.

(ii) In order to achieve the complete analogue of the results on $\mathbb{R}^d$, so far we have assumed the domains Λ to be *universal* extension domains. However, Lemma 1.2.25, Theorem 1.2.26 and Theorem 1.2.27 also hold true for an appropriate restricted set of admissible parameters k, ℓ, s, p, q, if Λ is only an $\mathrm{W}^{k,p}$-extension domain for k from some index set which is not $\mathbb{N}_0$ and p from a subset of $(1, \infty)$. This is clear from Lemma 1.2.23 and the proofs of the mentioned theorems.

We close this section with the remark that another useful result obtained by the $\mathrm{W}^{k,p}$-extension property of a domain Λ is that for $k \in \mathbb{N}_0$ and $1 \leq p < \infty$, not only $\mathrm{C}^\infty(\Lambda) \cap \mathrm{W}^{k,p}(\Lambda)$ is dense in $\mathrm{W}^{k,p}(\Lambda)$ (see [2, Thm. 3.17])[2], but also $\mathfrak{R}_\Lambda \mathrm{C}_c^\infty(\mathbb{R}^d)$.

1.2.3 Sobolev functions with partially vanishing trace

We define and investigate closed subspaces $\mathrm{W}^{k,p}_\Xi(\Lambda)$ of the spaces $\mathrm{W}^{k,p}(\Lambda)$ for domains $\Lambda \subset \mathbb{R}^d$, with $\Xi \subseteq \partial\Lambda$, where functions $f \in \mathrm{W}^{k,p}_\Xi(\Lambda)$ are supposed to satisfy $D^s f \restriction \Xi = 0$ for $|s| \leq k-1$ in some sense, where we call $D^s f \restriction \Xi$ the *trace* of $D^s f$. These spaces naturally arise in the studies of partial differential equations with mixed boundary conditions, where we suppose that a homogeneous Dirichlet condition is posed on Ξ, whereas $\partial\Lambda \setminus \Xi$ is given other conditions (Robin- or Neumann-conditions), hence it is of interest to consider them systematically in view of the well-known properties of their counterparts $\mathrm{W}^{k,p}(\Lambda)$.

Let us mention that we require Ξ to be a subset of the boundary of Λ, but the results in this subchapter are also true for $\Xi \subset \overline{\Lambda}$ generally. Since we use them only for $\Xi \subseteq \partial\Lambda$, we decided to focus on this case only. We also use $\Gamma := \partial\Lambda \setminus \Xi$ for the complement of Ξ within $\partial\Lambda$.

[2]Interestingly, the denseness of $\mathrm{C}^\infty(\Lambda) \cap \mathrm{W}^{k,p}(\Lambda)$ in $\mathrm{W}^{k,p}(\Lambda)$ was not known for quite some time until in 1964 the paper with the convincing name $H = W$ by Meyers and Serrin [118] put an end to the confusion.

Definition 1.2.30 (Sobolev functions w. partially vanishing trace)**.** Let $\Lambda \subset \mathbb{R}^d$ be a domain and let $\Xi \subseteq \partial\Lambda$ be closed. We define

$$\mathrm{C}^\infty_\Xi(\mathbb{R}^d) \coloneqq \Big\{ f \in \mathrm{C}^\infty_c(\mathbb{R}^d) \colon \operatorname{supp} f \cap \Xi = \emptyset \Big\}, \quad \mathrm{C}^\infty_\Xi(\Lambda) \coloneqq \mathfrak{R}_\Lambda \mathrm{C}^\infty_\Xi(\mathbb{R}^d)$$

and, for $k \in \mathbb{N}$,

$$\mathrm{W}^{k,p}_\Xi(\Lambda) \coloneqq \overline{\mathrm{C}^\infty_\Xi(\Lambda)}^{\|\cdot\|_{\mathrm{W}^{k,p}(\Lambda)}}.$$

It is clear that $\mathrm{W}^{k,p}_\Xi(\Lambda)$ is a closed subspace of $\mathrm{W}^{k,p}(\Lambda)$ and thus also a (reflexive) Banach space, and that $\mathrm{W}^{k,2}_\Xi(\Lambda)$ is a Hilbert space. Since $\mathfrak{R}_\Lambda \mathrm{C}^\infty_{\partial\Lambda}(\mathbb{R}^d) = \mathrm{C}^\infty_c(\Lambda)$, we find that $\mathrm{W}^{k,p}_{\partial\Lambda}(\Lambda)$ coincides with the closure of $\mathrm{C}^\infty_c(\Lambda)$ w.r.t. the $\mathrm{W}^{k,p}(\Lambda)$-norm, which is usually called $\mathrm{W}^{k,p}_0(\Lambda)$.

At the other end of the spectrum for $\Xi \subseteq \partial\Lambda$, we have already learned above that $\mathrm{W}^{k,p}_\emptyset(\Lambda) \neq \mathrm{W}^{k,p}(\Lambda)$ in general, since the $\mathrm{W}^{k,p}_\emptyset(\Lambda)$ in fact coincides with $\mathfrak{R}_\Lambda \mathrm{W}^{k,p}(\mathbb{R}^d)$. The remedy to recover equality of these two spaces and the key to many properties for the function spaces on Λ was the $\mathrm{W}^{k,p}$-extension property for Λ as in Theorems 1.2.26 and 1.2.27. Indeed, if Λ is a bounded universal extension domain, the spaces $\mathrm{W}^{k,p}_\Xi(\Lambda)$ already inherit all the embeddings as in Theorem 1.2.26. However, it turns out that the zero values of functions from $\mathrm{W}^{k,p}_\Xi(\Lambda)$ on Ξ allow to weaken this requirement significantly. We concentrate on the case $\Xi \subseteq \partial\Lambda$, but point out that the cited results are also true if merely $\Xi \subseteq \overline{\Lambda}$. See also [77, Thm. 4.5].

Theorem 1.2.31 ([57, Thm. 6.9])**.** *Let $k \in \mathbb{N}$ and $1 < p < \infty$. Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $\Xi \subseteq \partial\Lambda$ be closed. Assume that for every $\mathrm{x} \in \overline{\Gamma}$ there is an open neighborhood U_x of x such that $\Lambda \cap U_\mathrm{x}$ is a $\mathrm{W}^{k,p}$-extension domain. Then Λ is a $\mathrm{W}^{k,p}_\Xi$-extension domain, i.e., there exists a linear bounded extension operator $\mathfrak{E}_\Lambda \colon \mathrm{W}^{k,p}_\Xi(\Lambda) \to \mathrm{W}^{k,p}_\Xi(\mathbb{R}^d)$.*

Remark 1.2.32. Remarkably, the property of $\mathfrak{E}_\Lambda$ in the previous Theorem 1.2.31 to take its values in $\mathrm{W}^{k,p}_\Xi(\mathbb{R}^d)$ instead of $\mathrm{W}^{k,p}(\mathbb{R}^d)$ is "for free" in the setting of the theorem. Indeed, in [57, Prop. 6.5] it is firstly shown that the assumptions of the theorem allow to construct an extension operator from $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ to $\mathrm{W}^{k,p}(\mathbb{R}^d)$ and the authors give sufficient conditions under which such an extension operator actually maps into $\mathrm{W}^{k,p}_{\Xi}(\mathbb{R}^d)$ [57, Lem. 6.7]. However, it turns out that one of these sufficient conditions is always satisfied in the setting of Theorem 1.2.31/[57, Prop. 6.5] by the result that every $\mathrm{W}^{k,p}$-extension domain is already a d-set [76, Thm. 2] (cf. Definition 1.2.36 and Proposition 1.2.38 below). Thus the theorem above follows.

The case where Λ itself is already a $\mathrm{W}^{k,p}$-extension domain is now done easily by choosing the neighborhood $U_{\mathrm{x}} = \mathbb{R}^d$ for each $\mathrm{x} \in \overline{\Gamma}$:

Corollary 1.2.33 ([56, Cor. 2.2.13])**.** *Let $k \in \mathbb{N}$ and $1 < p < \infty$. Let $\Lambda \subseteq \mathbb{R}^d$ be a $\mathrm{W}^{k,p}$-extension domain and let $\Xi \subseteq \partial\Lambda$ be closed. Then Λ is a $\mathrm{W}^{k,p}_{\Xi}$-extension domain with the same extension operator.*

So far, we have avoided to talk about the more precise meaning of "$D^s f \restriction \Xi = 0$" for a function $f \in \mathrm{W}^{k,p}_{\Xi}(\Lambda)$. We will give a precise sense to this now, requiring some definitions first, starting with the Hausdorff measure as defined in [61, Ch. 2.1], cf. also [152, Ch. 7].

Definition 1.2.34 (Hausdorff measure)**.** Let $A \subset \mathbb{R}^d$ and $0 \leq s < \infty$. We define the *s-dimensional Hausdorff measure* to be

$$\mathcal{H}^s(A) \\ := \lim_{\delta \searrow 0} \inf\left\{ \sum_{j=1}^{\infty} \alpha(s) \left(\frac{\mathrm{d}(A_j)}{2}\right)^s : A_j \subset \mathbb{R}^d,\, A \subset \bigcup_{j=1}^{\infty} A_j,\, \mathrm{d}(A_j) \leq \delta \right\}$$

with the normalization parameter

$$\alpha(s) := \frac{\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2}+1)},$$

where Γ is the usual gamma function, and $\mathrm{d}(A_j) := \operatorname{diam} A_j$.

Remark 1.2.35.

(i) The normalization parameter α is present for more or less cosmetic purposes and there are various further possibilities to define the Hausdorff measure. The stated version will allow to establish exact coincidence of surface measures on (parts Ξ of) the boundary $\partial\Lambda$ of domains $\Lambda \subset \mathbb{R}^d$ with $\mathcal{H}^{d-1}$ restricted to Ξ or $\partial\Lambda$, respectively. Note that $\lambda^d(\mathbb{B}(r)) = \alpha(d)r^d$ for all d-dimensional balls around the origin (and thus for all).

(ii) The Hausdorff measures are universal in the sense that the family of measures $\mathcal{H}^s$ for $0 \leq s < \infty$ is indeed defined on *every* Euclidean space $\mathbb{R}^d$ for $d \in \mathbb{N}$ (as opposed to being defined on one fixed dimension). This makes them particularly useful to measure "dimensions" and lower-dimensional contents of sets $A \subset \mathbb{R}^d$. See also [152, Ch. 7, §27, Rem. 27.8].

The next definition is a characterization by Jonsson and Wallin [94] for sufficiently nicely behaved sets in $\mathbb{R}^d$ in the sense that these sets are exactly those onto which we will be able to define a suitable trace.

Definition 1.2.36 (N-set, [94, Ch. II.1])**.** Let $\Upsilon \subset \mathbb{R}^d$ be a closed set and $N \in (0, d]$. We say that Υ is an *N-set*, if there exist constants $c_0, c_1 > 0$ such that the inequality

$$c_0 r^N \leq \mathcal{H}^N\big(\Upsilon \cap \mathbb{B}(\mathrm{x}, r)\big) \leq c_1 r^N$$

is true for all $0 < r \leq r_0$ for some $r_0 > 0$ and all $\mathrm{x} \in \Upsilon$.

Remark 1.2.37.

(i) Every finite union of N-sets is again a N-set. Moreover, the upper bound for the radius is arbitrary (in the original definition, it is $r_0 = 1$). See [56, Lem. 1.2.23/1.2.24].

(ii) A common alternative name for N-sets is Ahlfors-David *regular of dimension N*, or short N-Ahlfors regular.

It was already mentioned in Remark 1.2.32 that a $\mathrm{W}^{k,p}$-extension domain $\Lambda \subset \mathbb{R}^d$ is already a d-set which we can now finally state properly. The result is due to Hajłasz, Koskela and Tuominen.

Proposition 1.2.38 ([76, Thm. 2])**.** *Let $\Lambda \subset \mathbb{R}^d$ be a domain and let $k \in \mathbb{N}$ and $1 < p < \infty$. If Λ is a $\mathrm{W}^{k,p}$-extension domain, then it is already a d-set.*

Last but not least, let us introduce the strictly defined representative of an $\mathrm{W}^{1,p}(\mathbb{R}^d)$-function, motivated by the Lebesgue differentiation theorem.

Definition 1.2.39 (Strictly defined representative)**.** Let $k \in \mathbb{N}$ and $1 < p < \infty$. Then, given $g \in \mathrm{W}^{1,p}(\mathbb{R}^d)$, we define its *strictly defined representative* $\mathcal{R}g$ by

$$(\mathcal{R}g)(\mathrm{x}) \coloneqq \lim_{r \searrow 0} \frac{1}{\lambda^d(\mathbb{B}(\mathrm{x},r))} \int_{\mathbb{B}(\mathrm{x},r)} g(\mathrm{y})\,\mathrm{dy} \quad \text{for } \mathrm{x} \in \mathbb{R}^d.$$

For $f \in \mathrm{W}^{k,p}(\mathbb{R}^d)$, we set $\mathcal{R}^k f \coloneqq \big(\mathcal{R}(D^s f)\big)_{|s| \le k-1}$, i.e., the collection of all strictly defined representatives of all partial derivatives up to order $k-1$ of f. Given a set $\Upsilon \subseteq \mathbb{R}^d$, we write $\mathcal{R}_\Upsilon f \coloneqq \mathcal{R}f \restriction \Upsilon$ and analogously for $\mathcal{R}^k_\Upsilon$ for the restrictions of the strictly defined representatives to Υ, and call $\mathcal{R}_\Upsilon f$ and $\mathcal{R}^k_\Upsilon f$, respectively, the (Υ-)*trace* of f.

The previous definitions now allow to precisely state in which sense the restriction of $D^s f$ to Υ is to be understood. But first, we have to make sure that the limit in Definition 1.2.39 actually *exists*. Clearly, if $p > d$, each function $D^s f$ for $|s| \le k-1$ is continuous by virtue of the Sobolev embeddings (1.14) and thus the strictly defined representatives $\mathcal{R}_\Upsilon(D^s f)$, $|s| \le k-1$, exist and coincide with the pointwise restriction $D^s f \restriction \Upsilon$ for *all* $\mathrm{x} \in \Upsilon$ for any nonempty set $\Upsilon \subseteq \mathbb{R}^d$. The case $p \le d$ is subject of the following lemma:

Lemma 1.2.40. *Let $k \in \mathbb{N}$, $1 < p \le d$ and $f \in \mathrm{W}^{k,p}(\mathbb{R}^d)$. Then $(\mathcal{R}^k f)(\mathrm{x})$ exists for $\mathcal{H}^\ell$-almost every $\mathrm{x} \in \mathbb{R}^d$, where $\ell > d - p$. To be precise: For* *every multiindex s with $|s| \le k-1$, there is a $\mathcal{H}^\ell$-nullset $\mathcal{N}(s)$ such that $\mathcal{R}(D^s f)(\mathrm{x})$ exists for every $\mathrm{x} \in \mathbb{R}^d \setminus \mathcal{N}(s)$. In particular, $\mathcal{R}^k_\Upsilon$ is well-defined if $\mathcal{H}^\ell(\Upsilon) > 0$.*

Proof. This result is a combination of results from the book of D.R. Adams and Hedberg [1]. We refer to [57, Sect. 4.2]. ☐

Remark 1.2.41. We have not told the whole truth in order to simplify the presentation: First, the strictly defined representatives may very well also be defined on the Bessel scale $\mathrm{H}^{s,p}(\mathbb{R}^d)$ for $s \in \mathbb{R}^+$ and, somewhat related, the exceptional set for which $\mathcal{R}_\Upsilon g$ does not exist is in fact a nullset in the sense of suitable Bessel capacities, see [1, Thm. 6.2.1], a stronger result than the one given in Lemma 1.2.40.

We now put the three definitions from above together to obtain a characterization of $\mathrm{W}^{k,p}_\Xi(\Lambda)$-functions in terms of their trace on Ξ. The driving force behind the result is the amazing theory of Jonsson and Wallin [94, Ch. VII].

Theorem 1.2.42. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $\Xi \subseteq \partial\Lambda$ be a closed $(d-1)$-set. Suppose that Λ is a $\mathrm{W}^{k,p}_\Xi$-extension domain for some $k \in \mathbb{N}$ and $1 < p < \infty$. Then*

$$\mathrm{W}^{k,p}_\Xi(\Lambda) = \Big\{\mathfrak{R}_\Lambda f \colon f \in \mathrm{W}^{k,p}(\mathbb{R}^d) \textit{ such that } \mathcal{R}^k_\Xi f = 0,\ \mathcal{H}^{d-1}\textit{-a.e. in } \Xi\Big\}.$$

Proof. The main work is done in [28, Thm. 4.4], where it is shown that

$$\mathrm{W}^{k,p}_\Xi(\mathbb{R}^d) = \Big\{f \in \mathrm{W}^{k,p}(\mathbb{R}^d) \colon \mathcal{R}^k_\Xi f = 0,\ \mathcal{H}^{d-1}\text{-a.e. in } \Xi\Big\}, \tag{1.19}$$

see also [77, Thm. 3.7]. Since we have assumed that Λ is a $\mathrm{W}^{k,p}_\Xi$-extension domain, we know that in fact $\mathrm{W}^{k,p}_\Xi(\Lambda) = \mathfrak{R}_\Lambda \mathrm{W}^{k,p}_\Xi(\mathbb{R}^d)$, and this is exactly what is stated in the theorem. ☐

Remark 1.2.43. A characterization of $\mathrm{W}^{1,p}_\Xi(\Lambda)$ functions from "within", under the condition that $\Xi \subseteq \partial\Lambda$ is a closed $(d-1)$-set, has been obtained recently by EGERT and TOLKSDORF in [59, Thm. 2.1] (see also [28, Cor. 5.3]). It states that, for $1 < p < \infty$, $f \in \mathrm{W}^{1,p}(\Lambda)$ is a member of $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ if and only if

$$\lim_{r \searrow 0} \frac{1}{\lambda^d(\mathbb{B}(\mathrm{x}, r))} \int_{\mathbb{B}(\mathrm{x},r)\cap\Lambda} |f(\mathrm{y})| \,\mathrm{dy} = 0 \quad \text{for } \mathcal{H}^{d-1}\text{-almost all } \mathrm{x} \in \Xi.$$

After this brief interlude, we return to some more properties of $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ spaces and finally give a name to the (anti-) dual space of $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$.

Definition 1.2.44. Let $\Lambda \subset \mathbb{R}^d$ be a domain, let $\Xi \subseteq \Lambda$ be closed and let $k \in \mathbb{N}$ and $1 \leq p < \infty$. Then we define $\mathrm{W}^{-k,p'}_{\Xi}(\Lambda)$ to be the space of *antilinear* continuous functionals on $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$. We keep the notation $\mathrm{W}^{-k,p'}_{\emptyset}(\Lambda)$ even if Λ is a $\mathrm{W}^{k,p}$-extension domain.

This notation is in agreement with the already noted $\left(\mathrm{H}^{s,p}(\mathbb{R}^d)\right)' = \mathrm{H}^{-s,p'}(\mathbb{R}^d)$, cf. (1.13). It is, however, not obvious how to identify dual spaces of function spaces on Λ with each other (see [146, Ch. 4.8.2]). We will return to an identification of elements of the dual space of $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ later, cf. Lemma 2.1.15.

Remark 1.2.45. By considering functions from $\mathrm{C}^{\infty}_{\Xi}(\mathbb{R}^d)$ it is immediate that the spaces $\mathrm{W}^{k,p}_{\Xi}(\mathbb{R}^d)$ enjoy the same embeddings as the $\mathrm{W}^{k,p}(\mathbb{R}^d)$ spaces do, cf. Theorem 1.2.15, and that these embedding transfer to the $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ spaces if Λ is a $\mathrm{W}^{k,p}_{\Xi}$-extension domain, including compactness, cf. Theorem 1.2.27. We will also use established (dense) embeddings for $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ spaces to obtain such for their dual spaces via the adjoint of the embedding mapping as in Proposition 1.0.2.

As a logical continuation of the thoughts in Remark 1.2.45, we consider interpolation identities for the $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ scale. These are, in the generality of this chapter, a much more difficult topic. We state the following recent theorem by HALLER-DINTELMANN, JONSSON, KNEES and REHBERG suitable for our means:

Theorem 1.2.46 ([77, Thm. 3.3/Cor. 3.4])**.** *Let $\Lambda \subset \mathbb{R}^d$ be a domain and let $\Xi \subseteq \partial\Lambda$ be closed. Assume that Λ is a $\mathrm{W}^{1,q}_{\Xi}$-extension domain uniformly for $1 < q < \infty$ and that Ξ is a $(d-1)$-set. Then we have the interpolation identities*

$$\big[\mathrm{W}^{1,p_0}_{\Xi}(\Lambda), \mathrm{W}^{1,p_1}_{\Xi}(\Lambda)\big]_\theta \doteq \big(\mathrm{W}^{1,p_0}_{\Xi}(\Lambda), \mathrm{W}^{1,p_1}_{\Xi}(\Lambda)\big)_{\theta,p} \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$$

and

$$\big[\mathrm{W}^{-1,p_0}_{\Xi}(\Lambda), \mathrm{W}^{-1,p_1}_{\Xi}(\Lambda)\big]_\theta \doteq \mathrm{W}^{-1,p}_{\Xi}(\Lambda)$$

for $0 < \theta < 1$, $1 < p_0, p_1 < \infty$ and $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

The second assertion follows from the first by duality properties, cf. Lemma 1.1.14. Since the quoted duality properties do not account for the *antilinear* dual spaces, we need to take a slight detour via the retraction/coretraction theorem in the form of Corollary 1.1.6 applied to the coretraction $E\colon f \mapsto [\psi \mapsto \langle f, \overline{\psi}\rangle]$, mapping an antilinear continuous functional to a linear one, and the corresponding retraction R with the same action, this time interpreted as mapping a linear continuous functional to an antilinear one (see also [80, Cor. 3.5]).

We do not know, at present, any interpolation identities for varying order of smoothness apart from the Hilbert scale $p = 2$, for which we refer to [56, Thm. 2.5.17] and [58, Sec. 7]. Note that there is a more fleshed out theory under less general assumptions on Λ and Ξ such as the existence of suitable boundary charts [69, 119]. We will also introduce such concepts below, cf. Definition 1.3.9.

There is one notable setting for the interpolation between spaces of higher smoothness and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$-spaces which in some sort circumvents the difficulties of getting a “handle” on the spaces by reducing interpolation with $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ to interpolation with $\mathrm{L}^p(\Lambda)$, but is still applicable in general contexts by using the reiteration theorem for domains of fractional powers of operators on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$. It builds upon the knowledge about one particular domain of a fractional power of a positive operator on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$.

Lemma 1.2.47. *Let $\Lambda \subset \mathbb{R}^d$ be a domain with $\Xi \subseteq \partial\Lambda$ closed, let $1 < p < \infty$ and assume that A is a positive operator on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$.*

1. *Let $0 < \theta < 1$. Then*

$$\big(W^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A\big)_{\theta,1} \hookrightarrow \operatorname{dom} A^{\theta} \hookrightarrow \big(W^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A\big)_{\theta,\infty}.$$

2. *If $\operatorname{dom} A^{1/2} \hookrightarrow \mathrm{L}^p(\Lambda)$, then*

$$\big(W^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A\big)_{\theta,1} \hookrightarrow \big(\mathrm{L}^p(\Lambda), \operatorname{dom} A\big)_{2\theta-1,1}$$

for all $\frac{1}{2} < \theta < 1$.

3. *If even $\operatorname{dom} A^{1/2} \doteq \mathrm{L}^p(\Lambda)$, then*

$$\big(W^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A\big)_{\theta,r} \doteq \big(\mathrm{L}^p(\Lambda), \operatorname{dom} A\big)_{2\theta-1,r}$$

for all $1 \le r \le \infty$ and $\frac{1}{2} < \theta < 1$.

Proof. The first assertion is a general principle for positive operators and proven in [146, Ch. 1.15.2]. For the second, we use the reiteration theorem for domains of fractional powers, Theorem 1.1.13, and Corollary 1.1.10:

$$\begin{aligned}\big(W^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A\big)_{\theta,1} &\doteq \big(\operatorname{dom} A^{1/2}, \operatorname{dom} A\big)_{2\theta-1,1} \\ &\hookrightarrow \big(\mathrm{L}^p(\Lambda), \operatorname{dom} A\big)_{2\theta-1,1}.\end{aligned}$$

The last assertion follows analogously, but there is no need to use Corollary 1.1.10. ☐

Remark 1.2.48.

(i) We note that even in the case $\operatorname{dom} A \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$, the previous lemma does *not* give a description of the interpolation spaces between $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$, since we merely have shifted the problem to the L^p scale and still do not know what happens to the vanishing trace property during interpolation. The usefulness of Lemma 1.2.47 manifests in now being able to use interpolation results for $\operatorname{dom} A$ and $\mathrm{L}^p(\Lambda)$ to obtain further information about $(\mathrm{W}^{-1,p}_{\Xi}(\Lambda), \operatorname{dom} A)_{\theta,1}$, which is particularly effective when $\operatorname{dom} A \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$.

(ii) The exponent $1/2$ for the requirement $\operatorname{dom} A^{1/2} = \mathrm{L}^p(\Lambda)$ is in a sense arbitrary since, as we see in the proof, we only need any domain of a fractional power for which we have more precise information. We have chosen the exponent $1/2$ since the information $\operatorname{dom} A^{1/2} \doteq \mathrm{L}^p(\Lambda)$ is exactly the assertion in the famous *Kato square root problem* and thus highly investigated, even in the context of Sobolev spaces with partially vanishing traces, see the recent works [20, 56, 58] and the references in there. Note that the condition can only be true for $p \geq 2$, cf. [20, Rem. 5.2]. See also Proposition 1.5.5 below.

(iii) While there are no explicit conditions on the domain Λ in Lemma 1.2.47, there are certainly implicit assumptions on Λ hidden within the assumptions on the operator A. We think it is of interest to obtain a general statement in any way. Of course, the divergence-gradient operators with which we work later on are exactly of the kind as required in Lemma 1.2.47, cf. Section 1.5.

1.2.4 Function spaces on (subsets of) the boundary

In this section we consider function spaces defined on the boundary $\partial\Lambda$ of a domain $\Lambda \subset \mathbb{R}^d$ or on subsets $\Xi \subseteq \partial\Lambda$ thereof. We are particularly interested in the connection to the function spaces on Λ, i.e., so-called trace mappings, where we are however content with traces of up to first order Sobolev spaces.

It turns out that an appropriate regularity assumption for (parts F of) the boundary $\partial\Lambda$ can be formulated by requiring the existence of a suitable measure μ on F, where we define "suitable" as follows:

Definition 1.2.49 (Upper $(d-1)$-Ahlfors measure)**.** Let $F \subset \mathbb{R}^d$ be compact and let μ be a Borel measure on F. If

$$\sup_{\mathrm{x}\in F}\sup_{0<r<r_0} r^{1-d}\mu(\mathbb{B}(\mathrm{x},r)) < \infty, \tag{1.20}$$

then we say that μ is an *upper* $(d-1)$*-Ahlfors measure* on F.

Remark 1.2.50. In Definition 1.2.49, μ being "a measure on F" means that the support of μ is within F and not necessarily that μ is defined only on a σ-algebra of subsets of F. Of course, there is a one-to-one correspondence between measures on $\mathbb{R}^d$—or any set containing F—with support in F and measures on F, but generally, the condition should be read as $\mu(A) = 0$ for all μ-measurable sets $A \subset \mathbb{R}^d \setminus F$.

It is clear that if F is a (compact) $(d-1)$-set, then $\mathcal{H}^{d-1} \restriction F$ is exactly an upper $(d-1)$-Ahlfors measure on F, cf. Definition 1.2.36. Interestingly, the restriction of the Hausdorff measure to F is already an upper $(d-1)$-Ahlfors measure—and even slightly more, note that the supremum in (1.21) below is taken over all $\mathrm{x} \in \mathbb{R}^d$— whenever $\mathcal{H}^{d-1}(F) < \infty$:

Lemma 1.2.51. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let F be a closed subset of $\partial\Lambda$ with $\mathcal{H}^{d-1}(F) < \infty$. Then the restriction of the $(d-1)$-dimensional Hausdorff measure to F, that is, $\mu := \mathcal{H}^{d-1} \restriction F$ given by $\mu(A) := \mathcal{H}^{d-1}(A \cap F)$ for all Borel sets $A \subset \mathbb{R}^d$, satisfies the measure condition*

$$\sup_{\mathrm{x}\in \mathbb{R}^d}\sup_{0<r<r_0} r^{1-d}\mu(\mathbb{B}(\mathrm{x},r)) < \infty. \tag{1.21}$$

In particular, μ is an upper $(d-1)$-Ahlfors measure on F.

Proof. Note that F is a closed subset of the compact set $\overline{\Lambda}$ and thus compact itself. In [61, Ch. 2.3], it is shown that if $\mathcal{H}^{d-1}(F) < \infty$, then we have the measure densities

$$\lim_{r\searrow 0} \frac{\mathcal{H}^{d-1}(\mathbb{B}(\mathrm{x},r)\cap F)}{\alpha(d-1)r^{d-1}} = 0 \quad \text{for } \mathrm{x} \in \mathbb{R}^d \setminus F,$$

and
$$\frac{1}{2^{d-1}} \leq \limsup_{r\searrow 0} \frac{\mathcal{H}^{d-1}(\mathbb{B}(\mathrm{x},r)\cap F)}{\alpha(d-1)r^{d-1}} \leq 1 \quad \text{for } \mathrm{x} \in F,$$

where $\alpha(d-1)$ it the volume of the $(d-1)$-dimensional unit ball, cf. Remark 1.2.35. We infer that there exists $r_0 > 0$ such that

$$\frac{\mathcal{H}^{d-1}(\mathbb{B}(\mathrm{x},r)\cap F)}{\alpha(d-1)r^{d-1}} \leq 2 \quad \text{for all } \mathrm{x} \in \mathbb{R}^d, \quad 0 < r < r_0. \tag{1.22}$$

For the remaining range of radii r, we use

$$\frac{\mathcal{H}^{d-1}(\mathbb{B}(\mathrm{x},r)\cap F)}{\alpha(d-1)r^{d-1}} \leq \frac{\mathcal{H}^{d-1}(F)}{\alpha(d-1)r_0^{d-1}} < \infty \quad \text{for all } \mathrm{x} \in \mathbb{R}^d, \quad r \geq r_0.$$

Together with (1.22), this gives the assertion. □

Remark 1.2.52. The typical situation where the preceding lemma is applicable is when there is a finite number of sets $A_1, \ldots, A_n$ with $\mathcal{H}^{d-1}(A_i) < \infty$ for $i = 1, \ldots, n$ such that $F \subseteq \cup_{i=1}^n A_i$. This is in particular the case when there are local descriptions of F or $\partial\Lambda$ available, cf. Chapter 1.3.

Lemma 1.2.51 moreover shows that for the special measure $\mathcal{H}^{d-1} \restriction F$ on a closed subset F of $\partial\Lambda$ with $\mathcal{H}^{d-1}(F) < \infty$, the two measure conditions (1.20) and (1.21) are already equivalent, and we even obtain that these conditions are themselves equivalent to $\mathcal{H}^{d-1}(F) < \infty$:

Corollary 1.2.53. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $F \subseteq \partial\Lambda$ be closed. Then the following are equivalent for $\mu := \mathcal{H}^{d-1} \restriction F$:*

(i) μ satisfies the measure condition (1.21)*.*

(ii) μ is an upper $(d-1)$-Ahlfors measure on F.

(iii) $\mathcal{H}^{d-1}(F) < \infty$.

Proof. It is obvious that (i) implies (ii).

Assume (ii), so that $\mu = \mathcal{H}^{d-1} \restriction F$ is an upper $(d-1)$-Ahlfors measure. Since F is a closed subset of the compact set $\overline{\Omega}$ and thus itself compact, covering F by finitely many balls $\mathbb{B}_r(\mathrm{x}_i)$ with $1 \leq i \leq n$ for some fixed $0 < r < r_0$ and using that μ is an upper $(d-1)$-Ahlfors measure on F, we find that

$$\mathcal{H}^{d-1}(F) \leq \sum_{i=1}^{n} \mathcal{H}^{d-1}(\mathbb{B}_r(\mathrm{x}_i) \cap F) \leq nMr^{d-1} < \infty,$$

which was (iii).
Finally, for $\mathcal{H}^{d-1}(F) < \infty$, Lemma 1.2.51 strikes to show that $\mu = \mathcal{H}^{d-1} \restriction F$ satisfies (1.21), so (i) is true. □

The fundamental property of upper $(d-1)$-Ahlfors measures on $F \subseteq \partial\Lambda$ for this chapter is that they allow to obtain a well-defined *trace operator* for Sobolev functions:

Theorem 1.2.54 ([25, Thm. 1.1])**.** *Let $\Lambda \subset \mathbb{R}^d$ be a bounded $\mathrm{W}^{1,p}$-extension domain for $1 < p < \infty$, let $F \subseteq \partial\Lambda$ be closed and let μ be an upper $(d-1)$-Ahlfors measure on F. Suppose that $1 \leq q \leq \infty$ satisfies $\frac{d-p}{p} \leq \frac{d-1}{q}$ with $q < \infty$ if $p = d$. Then there exists a* continuous linear *operator*

$$\mathrm{tr} \colon \mathrm{W}^{1,p}(\Omega) \to \mathrm{L}^q(F;\mu)$$

such that

$$\mathrm{tr}\, f = f \restriction F \quad \textit{for every } f \in \mathrm{W}^{1,p}(\Omega) \cap \mathrm{C}(\overline{\Omega}),$$

the so-called trace operator*. The operator is even* compact *if $\frac{d-p}{p} < \frac{d-1}{q}$.*

Note that F in the original version of foregoing theorem might, again, in general be a subset of $\overline{\Lambda}$ instead of only $\partial\Lambda$. We have omitted this possibility since we will not need it in the following. For the sake of completeness, we next cite a result by Maz'ya which yields the same assertion as in the compact case in Theorem 1.2.54 and a little more, but requires the stronger measure condition (1.21):

Theorem 1.2.55 ([112, Ch. 1.4.7, Cor. 2]). *Let $\Lambda \subset \mathbb{R}^d$ be a bounded $\mathrm{W}^{k,p}$-extension domain for $k = 0, 1$ and $1 < p < \infty$, and let μ be a measure on $\overline{\Lambda}$ satisfying the measure condition* (1.21)*. Suppose that $\frac{d-p}{p} < \frac{d-1}{q}$ and $p \leq q$. Then*

$$\| \operatorname{tr} f\|_{\mathrm{L}^q(\overline{\Lambda};\mu)} \leq C\|f\|_{\mathrm{L}^p(\Lambda)}^{1-\theta}\|f\|_{\mathrm{W}^{1,p}(\Lambda)}^{\theta} \quad \textit{for} \quad \theta = \frac{d}{p} - \frac{d-1}{q} \tag{1.23}$$

for all $f \in \mathrm{C}^1(\Lambda)$.

Remark 1.2.56.

(i) It is of interest to note that the requirement (1.20) in Theorem 1.2.55 is "minimal" in the sense that if (1.23) is true for all restrictions of $\mathrm{C}^1(\mathbb{R}^d)$-functions to Λ, then the measure μ *must* already satisfy (1.21). See [112, Ch. 1.4.7, Cor. 2].

(ii) The inequality (1.23) strongly resembles the interpolation inequality for a Banach space E, so

$$\|f\|_E \leq C\|f\|_{A_0}^{1-\theta}\|f\|_{A_1}^{\theta},$$

which we had used to show that $(A_0, A_1)_{\theta,1} \hookrightarrow E$ in Lemma 1.1.9. Indeed, it seems tempting to conjecture a statement of the following form: Suppose that $A \in \mathscr{L}(A_0 \cap A_1; Z)$ and that

$$\|Af\|_Z \leq C\|f\|_{A_0}^{1-\theta}\|f\|_{A_1}^{\theta} \quad \text{for all } f \in A_0 \cap A_1.$$

Then A can be continuously extended to a continuous linear operator from $(A_0, A_1)_{\theta,1}$ to Z. In the present case of (1.23), such a result would allow to continuously extend the trace operator tr from $\mathrm{W}^{1,p}$ to the $\mathrm{H}^{s,p}$-scale in an optimal way: For $q = p$, it would imply that the trace operator tr maps $\mathrm{H}^{\frac{1}{p}+\varepsilon,p}(\Omega)$ continuously into $\mathrm{L}^p(\partial\Lambda)$, which is the classical result that one "looses" $\frac{1}{p}$ of smoothness for the trace. However, the author is unaware how such a claim in this generality should be proved, and all successful attempts using additional assumptions on the linear operator A were rendered unusable by the trace operator resisting to verify these assumptions.

We have already noted in the previous Chapter 1.2.3 that the requirement of the $\mathrm{W}^{k,p}$-extension property for Λ may be significantly weakened if we are content with $\mathrm{W}^{k,p}_\Xi$-extension with a closed set $\Xi \subseteq \partial\Lambda$ (or if that is the goal in the first place!). We transfer Theorems 1.2.54 and 1.2.55 to this setting, more precisely, to that of Theorem 1.2.31. Thereby, we only consider the maximal case $F = \overline{\Gamma} = \overline{(\partial\Lambda \setminus \Xi)}$, from which the case of general closed $F \subseteq \overline{\Gamma}$ follows by further restriction.

Lemma 1.2.57. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $\Xi \subseteq \partial\Lambda$ be closed. Assume that for every $\mathrm{x} \in \overline{\Gamma} = \overline{(\partial\Lambda \setminus \Xi)}$ there is an open neighborhood U_x of x such that $\Lambda \cap U_\mathrm{x}$ is a $\mathrm{W}^{k,p}$-extension domain for $k = 0, 1$ and $1 \leq p < \infty$, and that μ is an upper $(d-1)$-Ahlfors measure on $\overline{\Gamma}$. Let moreover $1 \leq q \leq \infty$ satisfy $\frac{d-p}{p} \leq \frac{d-1}{q}$ with $q < \infty$ if $p = d$. Then there is a continuous linear trace operator*

$$\operatorname{tr}\colon \mathrm{W}^{1,p}_\Xi(\Omega) \to \mathrm{L}^q(\overline{\Gamma}; \mu)$$

such that

$$\operatorname{tr} f = f \restriction \overline{\Gamma} \quad \textit{for every } f \in \mathrm{W}^{1,p}_\Xi(\Omega) \cap \mathrm{C}(\overline{\Omega}),$$

which is even compact if $\frac{d-p}{p} < \frac{d-1}{q}$. If in this case also $p \leq q$ holds true and μ satisfies the measure condition (1.21)*, then we also have*

$$\|\operatorname{tr} f\|_{\mathrm{L}^q(\overline{\Gamma};\mu)} \leq C \|f\|^{1-\theta}_{\mathrm{L}^p(\Lambda)} \|f\|^{\theta}_{\mathrm{W}^{1,p}(\Lambda)} \quad \textit{for} \quad \theta = \frac{d}{p} - \frac{d-1}{q} \tag{1.24}$$

for all $f \in \mathrm{C}^1(\Lambda)$.

The proof of Lemma 1.2.57 consists of a simple localization argument to make use of Theorems 1.2.54 and 1.2.55, to which the lemma collapses for the special case $\Xi = \emptyset$.

Proof. Let U_x be neighborhoods of $\mathrm{x} \in \overline{\Gamma}$ as in the assumption. Then the system of sets $\{U_\mathrm{x} \colon \mathrm{x} \in \overline{\Gamma}\}$ forms an open covering of the compact set $\overline{\Gamma}$, hence there exists a finite collection of points $\mathrm{x}_1, \ldots, \mathrm{x}_n \in \overline{\Gamma}$ such that $\cup_{i=1}^n U_{\mathrm{x}_i}$ still covers $\overline{\Gamma}$. We moreover know that each $U_{\mathrm{x}_i} \cap \Lambda$ is a $\mathrm{W}^{k,p}$-extension domain. Since for $f \in \mathrm{C}_\Xi^\infty(\Omega)$, we have $f \restriction (U_{\mathrm{x}_i} \cap \Lambda) \in \mathrm{C}^\infty(U_{\mathrm{x}_i} \cap \Lambda) \subset \mathrm{W}^{1,p}(U_{\mathrm{x}_i} \cap \Lambda)$, Theorem 1.2.54 thus shows

$$\|\operatorname{tr} f\|_{\mathrm{L}^q(\overline{\Gamma};\mu)} \leq \sum_{i=1}^n \|\operatorname{tr} f\|_{\mathrm{L}^q(U_{\mathrm{x}_i} \cap \overline{\Gamma})} \leq C \sum_{i=1}^n \|f\|_{\mathrm{W}^{1,p}(U_{\mathrm{x}_i} \cap \Lambda)} \leq C_n \|f\|_{\mathrm{W}^{1,p}(\Lambda)},$$

which extends to $f \in \mathrm{W}_\Xi^{1,p}(\Lambda)$ by density. The restriction property for $\operatorname{tr} f$ for continuous f is clear from the construction. Now let in addition the measure condition (1.21) be satisfied and $p \leq q$. Then, Theorem 1.2.55 yields for each $i = 1, \ldots, n$ and for all $f \in \mathrm{C}^1(U_{\mathrm{x}_i} \cap \Lambda)$

$$\|f\|_{\mathrm{L}^q(U_{\mathrm{x}_i} \cap \overline{\Gamma};\mu)} \leq C \|f\|_{\mathrm{L}^p(U_{\mathrm{x}_i} \cap \Lambda)}^{1-\theta} \|f\|_{\mathrm{W}^{1,p}(U_{\mathrm{x}_i} \cap \Lambda)}^\theta$$

for the values of p, q and θ as in (1.24). Now let us consider $f \in \mathrm{C}^1(\Lambda)$. Then clearly $f \restriction (U_{\mathrm{x}_i} \cap \Lambda) \in \mathrm{C}^1(U_{\mathrm{x}_i} \cap \Lambda)$ and we have

$$\begin{aligned} \|\operatorname{tr} f\|_{\mathrm{L}^q(\overline{\Gamma};\mu)} &\leq \sum_{i=1}^n \|\operatorname{tr} f\|_{\mathrm{L}^q(U_{\mathrm{x}_i} \cap \overline{\Gamma})} \\ &\leq C \sum_{i=1}^n \|f\|_{\mathrm{L}^p(U_{\mathrm{x}_i} \cap \Lambda)}^{1-\theta} \|f\|_{\mathrm{W}^{1,p}(U_{\mathrm{x}_i} \cap \Lambda)}^\theta \leq C_n \|f\|_{\mathrm{L}^p(\Lambda)}^{1-\theta} \|f\|_{\mathrm{W}^{1,p}(\Lambda)}^\theta. \end{aligned}$$

This was the claim. □

In Remark 1.2.56, we have already mentioned the standard and folklore result for the trace operator, namely that it is acting as a continuous linear operator $\mathrm{W}^{1,p}(\Lambda) \to \mathrm{W}^{1-\frac{1}{p},p}(\partial\Lambda)$, so between Sobolev-type spaces of integrability $1 < p < \infty$ with a differentiability gap of $\frac{1}{p}$. In addition, classically, one is also able to actually *characterize* the trace of $\mathrm{W}^{1,p}(\Lambda)$ functions because the mapping into $\mathrm{W}^{1-\frac{1}{p}}(\partial\Lambda)$ is usually shown to be *surjective*, so *onto*. See also Remark 1.2.63 for the case of fractional differentiability order.

We will obtain an analogous result in Theorem 1.2.62 below, further building upon the work by Jonsson and Wallin in [94]. First we introduce one more function space defined in [94, Ch. V] – not surprisingly, it is a space of fractional smoothness on a lower-dimensional set in the form of a $(d-1)$-set.

Definition 1.2.58. Let $F \subset \mathbb{R}^d$ be a $(d-1)$-set, let $0 < s < 1$ and $1 \leq p \leq \infty$. Set $\mu = \mathcal{H}^{d-1} \upharpoonright F$ and define

$$[f]_{s,p} \coloneqq \left(\iint_{\mathbb{F}} \frac{|f(\mathrm{x}) - f(\mathrm{y})|^p}{|\mathrm{x} - \mathrm{y}|^{d-1+sp}} \,\mathrm{d}\mu(\mathrm{x}) \,\mathrm{d}\mu(\mathrm{y}) \right)^{\frac{1}{p}}$$

where $\mathbb{F} = \{(\mathrm{x}, \mathrm{y}) \in F \times F \colon |\mathrm{x} - \mathrm{y}| < 1\}$. Then the *Besov space* on F is given by

$$\mathrm{B}^s_{p,p}(F) \coloneqq \left\{ f \in \mathrm{L}^p(F;\mu) \colon \|f\|_{\mathrm{B}^s_{p,p}(F)} \coloneqq \|f\|_{\mathrm{L}^p(F;\mu)} + [f]_{s,p} < \infty \right\}.$$

The space $\mathrm{B}^s_{p,p}(F)$ is a Banach space, and it is clear from the definition of the norm on $\mathrm{B}^s_{p,p}(F)$ that $\mathrm{B}^s_{p,p}(F) \hookrightarrow \mathrm{L}^p(F;\mu)$. In fact, the embeddings between $\mathrm{B}^s_{p,p}(F)$ and $\mathrm{L}^q(F;\mu)$ work as one would expect, cf. [25, Thm. 6.8].

Remark 1.2.59. For the reader familiar with function spaces incorporating non-integer orders of smoothness, it might seem that the seminorm $[\cdot]_{s,p}$ in Definition 1.2.58 resembles the usual Slobodeckij-seminorm more so than the norms “expected” of Besov spaces (cf. also [56, Rem. 1.1.2]). Indeed, it would also be justified to use the notation $\mathrm{W}^{s,p}(F)$, which would also line up nicely with the folklore result mentioned before Definition 1.2.58. The terminology in this case seems like a matter of taste, as the Slobodeckij-norm on $\mathbb{R}^d$ is merely an equivalent norm on the Besov space on $\mathbb{R}^d$ [146, Ch. 2.5.1] and as such it is hard to argue which one “came first”. We thus decide to not pick a side and simply stick with the notation used by the inventors in [94].

The Besov spaces introduced in Definition 1.2.58 are of fundamental importance in the following results, since they will prove to exactly constitute the space of traces of $\mathrm{W}^{1,p}$-functions. First we recall the result for $\mathrm{W}^{1,p}(\mathbb{R}^d)$.

Proposition 1.2.60 ([94, Ch. VII])**.** *Let $F \subset \mathbb{R}^d$ be a closed $(d-1)$-set and $1 < p < \infty$.*

(i) The trace operator $\mathcal{R}_F$ from Definition 1.2.39 gives rise to a continuous linear operator from $\mathrm{W}^{1,p}(\mathbb{R}^d)$ to $\mathrm{B}^{1-1/p}_{p,p}(F)$.

(ii) There exists a continuous linear extension operator $\mathcal{E}_F$ from $\mathrm{B}^{1-1/p}_{p,p}(F)$ to $\mathrm{W}^{1,p}(\mathbb{R}^d)$ which serves as a right-inverse for $\mathcal{R}_F$.

Remark 1.2.61. Again, Proposition 1.2.60 is merely a special case of the results presented in [94], where also traces of the Bessel potential scale $\mathrm{H}^{s,p}(\mathbb{R}^d)$ for $s \in \mathbb{R}^+$, and of Besov spaces $\mathrm{B}^s_{p,q}(\mathbb{R}^d)$ are characterized, cf. also Remark 1.2.41.

Next, we use Proposition 1.2.60 to define a suitable trace of $\mathrm{W}^{1,p}_\Xi(\Lambda)$ functions. The way to proceed is generally clear: Given a $(d-1)$-set $F \subset \partial\Lambda$ and a function $f \in \mathrm{W}^{1,p}_\Xi(\Lambda)$, we extend f to a $\mathrm{W}^{1,p}(\mathbb{R}^d)$-function $\mathfrak{E}_\Lambda f$ and then define the F-trace of f by the F-trace of $\mathfrak{E}_\Lambda f$ as provided by Proposition 1.2.60:

$$f \in \mathrm{W}^{1,p}_\Xi(\Lambda) \xrightarrow{\mathfrak{E}_\Lambda} g \in \mathrm{W}^{1,p}(\mathbb{R}^d) \xrightarrow{\mathcal{R}_F} \text{“}f \restriction F\text{”} \in \mathrm{B}^{1-1/p}_{p,p}(F).$$

However, in doing so, we need to make sure that this procedure is unambiguous w.r.t. the extension operator $\mathfrak{E}_\Lambda$, i.e., that the such-defined trace of $f \in \mathrm{W}^{1,p}_\Xi(\Lambda)$ is independent of the function $g \in \mathrm{W}^{1,p}(\mathbb{R}^d)$, satisfying $\mathfrak{R}_\Lambda g = f$, used to define the trace. Luckily, the conditions to make sure that this is indeed the case are already known [57, Lem. 6.7]. The author is indebted to Moritz Egert for pointing out the deeper meaning of the lemma.

Theorem 1.2.62. *Let $\Lambda \subset \mathbb{R}^d$ be a domain, let $\Xi \subseteq \partial\Lambda$ be a closed set, $1 < p < \infty$, and assume that Λ is a $\mathrm{W}^{1,p}_\Xi$-extension domain by virtue of the extension operator $\mathfrak{E}_\Lambda$. Let $F \subseteq \overline{\Gamma}$ be a closed $(d-1)$-set. Let at least* *one of these conditions be satisfied:*

(i) The asymptotically nonvanishing relative volume *condition*

$$\liminf_{r\searrow 0} \frac{\lambda^d(\mathbb{B}(\mathrm{x},r)\cap\Lambda)}{r^d} > 0$$

is satisfied for $\mathcal{H}^{d-1}$-almost all $\mathrm{x}\in F$, or

(ii) there exists an extension operator $\mathfrak{E}_\Lambda^\bullet$ (not necessarily $\mathfrak{E}_\Lambda$ itself) which maps $\mathrm{C}_\Xi^\infty(\Lambda)$ into $\mathrm{W}^{1,q}(\mathbb{R}^d)$ for some $q>d$.

Then the trace operator $\mathcal{R}_F^\Lambda\colon \mathrm{W}_\Xi^{1,p}(\Lambda)\to \mathrm{B}_{p,p}^{1-1/p}(F)$ given by

$$\mathcal{R}_F^\Lambda f := \mathcal{R}_F g \quad \text{for any } g\in W^{1,p}(\mathbb{R}^d) \text{ such that } \mathfrak{R}_\Lambda g = f$$

is a well-defined, continuous, linear operator which is unambiguously defined w.r.t. the choice of the function g. Moreover, the continuous, linear extension operator

$$\mathcal{E}_F^\Lambda := \mathfrak{R}_\Lambda\mathcal{E}_F\colon \mathrm{B}_{p,p}^{1-1/p}(F)\to \mathrm{W}_\Xi^{1,p}(\Lambda)$$

serves as a right-inverse for $\mathcal{R}_F^\Lambda$.

Proof. We only need to make sure that $\mathcal{R}_F^\Lambda$ is well-defined; the other claims follow directly from Proposition 1.2.60. In particular, we may realize $\mathcal{R}_F^\Lambda$ via $\mathcal{R}_F\mathfrak{E}_\Lambda$ once we have shown that it is unambiguously defined. Let $f\in \mathrm{W}_\Xi^{1,p}(\Lambda)$ be given. Since Λ is a $\mathrm{W}_\Xi^{1,p}$-extension domain, there exists a function $g\in \mathrm{W}^{1,p}(\mathbb{R}^d)$ such that $\mathfrak{R}_\Lambda g = f$ in the first place.[3]

Now let $g_1,g_2\in \mathrm{W}^{1,p}(\mathbb{R}^d)$ with $\mathfrak{R}_\Lambda g_1 = \mathfrak{R}_\Lambda g_2 = f$ be given and set $\mathfrak{g} := g_1 - g_2$. Clearly, $\mathfrak{R}_\Lambda \mathfrak{g} = 0$. We need to show that $\mathcal{R}_F(\mathfrak{g}) = 0$ is true $\mathcal{H}^{d-1}$-almost everywhere on F. According to (1.19), this is the case if and only if $\mathfrak{g}\in \mathrm{W}_F^{1,p}(\mathbb{R}^d)$.

(i) Assume that condition (i) is satisfied and let $\mathcal{N}\subset F$ be the exceptional $\mathcal{H}^{d-1}$-nullset where the asymptotically nonvanishing relative

[3]Even $g\in \mathrm{W}_\Xi^{1,p}(\mathbb{R}^d)$, but it will turn out that this information is in fact not needed for our proof; it may however come in handy if one wants to allow $\Xi\cap F\neq\emptyset$, see also Remark 1.2.63.

volume condition does not hold or $\mathcal{R}_F$ is not defined. Then it is shown in the proof of [57, Lem. 6.7] that a function $\mathfrak{g} \in \mathrm{W}^{1,p}(\mathbb{R}^d)$ satisfying $\mathfrak{R}_\Lambda \mathfrak{g} = 0$ on the intersection of an open neighborhood of $\mathrm{x} \in F \setminus \mathcal{N}$ with Λ for each such x is indeed from $\mathrm{W}^{1,p}_F(\mathbb{R}^d)$ – in their proof, the authors have $\mathfrak{R}_\Lambda \mathfrak{g} \in \mathrm{C}^\infty_F(\Lambda)$, but our case with $\mathfrak{R}_\Lambda \mathfrak{g} = 0$ on Λ is clearly sufficient.

(ii) Assuming condition (ii), the assertions are exactly proven in the first part of the proof of [28, Thm. 5.1], where nothing but a particular extension operator satisfying condition (ii) is used. □

Remark 1.2.63.

(i) We have omitted a variant of condition (i) in Theorem 1.2.62 which also allows the asymptotically nonvanishing relative volume condition to be satisfied only $(1,p)$-quasieverywhere, i.e., in the sense of capacities (which we do not want to introduce). The proof remains exactly the same in this case.

(ii) As already noted in the proof, there is a similar result for so-called *locally (ε,δ)-domains* in [28]. The authors there also consider the case where F is a general subset of $\partial\Lambda$, using modifications of the Besov spaces $\mathrm{B}^s_{p,p}(F)$, and show that the two traces on F and Ξ are compatible. They also consider $\mathrm{W}^{k,p}_\Xi(\Lambda)$-spaces with $k \in \mathbb{N}$. It should be possible to straightforwardly modify Theorem 1.2.62 appropriately as done in [28, Thm. 5.1] if one is interested in an analogous result when Ξ is also a closed $(d-1)$-set, since we already have Theorem 1.2.42 at hand.

(iii) The asymptotically nonvanishing relative volume condition is satisfied *a fortiori* for *all* $\mathrm{x} \in \partial\Lambda$ if Λ is a d-set. Indeed, under this assumption, Theorem 1.2.62 is shown already in [94, Ch. VIII] for $\Xi = \emptyset$ and $F = \partial\Lambda$. Moreover, in [82], it is established that the existence of a continuous trace operator onto the Besov space is *nearly* sufficient for the existence of a $\mathrm{W}^{1,p}$-extension operator.

(iv) In relation to the previous remark, let $\Lambda \subset \mathbb{R}^d$ satisfy the assumptions of Theorem 1.2.31, i.e., $\Xi \subseteq \partial\Lambda$ is closed and for every $\mathrm{x} \in \overline{\Gamma}$

there exists an open neighborhood U_{x} of x such that $\Lambda \cup U_{\mathrm{x}}$ is a $\mathrm{W}^{1,p}$-extension domain. We have already noted in Proposition 1.2.38 that a $\mathrm{W}^{1,p}$-extension domain must already be a d-set. Hence a domain $\Lambda \subset \mathbb{R}^d$ satisfying the assumptions of Theorem 1.2.31 already satisfies the assumptions of Theorem 1.2.62 for all closed $(d-1)$-sets $F \subseteq \overline{\Gamma}$.

(v) An analogous result to Theorem 1.2.62 holds true for the fractional Besov scale $\mathrm{B}^s_{p,q}$ in the setting of Theorem 1.2.31; in particular one re-obtains that (formally speaking)

$$\mathcal{R}_F^\Lambda \colon \mathrm{W}^{s,p}_\Xi(\Lambda) \to \mathrm{B}^{s-\frac{1}{p}}_{p,p}(F) \quad \text{for} \quad \frac{1}{p} < s < 1$$

is a continuous, linear and *surjective* operator. This can be inferred from [28, Thm. 8.7]: There, the authors prove this under the assumption that the underlying domain is *d-thick*, which in the context of Theorem 1.2.31 follows for each boundary neighborhood as a $\mathrm{W}^{1,p}$-extension domain again by the d-set property as per Proposition 1.2.38. Then one can use a localization procedure as in Lemma 1.2.57 to construct the total boundary restriction operator.

From the existence of an extension operator $\mathcal{E}_F^\Lambda \colon \mathrm{B}^{1-1/p}_{p,p}(F) \to \mathrm{W}^{1,p}_\Xi(\Lambda)$ it is already clear that $\mathcal{R}_F^\Lambda$ in Theorem 1.2.62 is surjective. Thus, the theorem completely characterizes the traces of first order Sobolev spaces on extension domains (note that $\Xi = \emptyset$ is allowed). Since the strictly defined representative $\mathcal{R}_F f$ for a continuous function f coincides with $f \restriction F$, we indeed find $\mathcal{R}_F^\Lambda = \mathrm{tr}$, where tr is the trace operator introduced in Theorem 1.2.54 or Lemma 1.2.57: the operators coincide on $\mathrm{C}^\infty_\Xi(\Lambda)$, which is dense in $\mathrm{W}^{1,p}_\Xi(\Lambda)$ by construction. In this sense, together with the embedding $\mathrm{B}^{1-1/p}_{p,p}(F) \hookrightarrow \mathrm{L}^q(F)$ mentioned in Definition 1.2.58, Theorem 1.2.62 is a straight "upgrade" to the foregoing results; however, there is no assertion about *compactness* in Theorem 1.2.62, which makes Theorems 1.2.54 and 1.2.55 or Lemma 1.2.57 earn their rightful places. In fact, we would expect such a compactness property to arise precisely from the embedding $\mathrm{B}^{1-1/p}_{p,p}(F) \hookrightarrow \mathrm{L}^q(F)$—and this would indeed make the corresponding assertion in Theorem 1.2.54 a straightforward consequence of Theorem 1.2.62—, but the author was unable to locate a corresponding result in [46] or [25].

1.3 Geometry for the underlying spatial sets

We now give assumptions on the domains $\Lambda \subset \mathbb{R}^d$ that allow to use the abstract results from the previous chapters. Fix a space dimension $d \in \mathbb{N}$. The fundamental idea is to provide boundary charts for neighborhoods of points on $\partial\Lambda$, and we will use the bi-Lipschitz class of mappings for these boundary charts. We will see that we may relax these conditions for many results for the spaces $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ due to functions from this space being "fixed" as zero on the part Ξ of $\partial\Lambda$.

Definition 1.3.1. Let $\Upsilon \subseteq \mathbb{R}^d$. A mapping $\phi\colon \Upsilon \to \mathbb{R}^n$ is called *bi-Lipschitz*, if it is injective and Lipschitz-continuous and $\phi^{-1}\colon \phi(\Upsilon) \to \Upsilon$ is also Lipschitz continuous.

It turns out that for Lipschitz-mappings, the Hausdorff dimension of the image of a set Υ is not larger than the Hausdorff dimension of Υ itself:[4]

Lemma 1.3.2 ([152, Ch. 7, §28, Thm. 28.4])**.** *Let $\Upsilon \subset \mathbb{R}^d$ and assume that $f\colon \Upsilon \to \mathbb{R}^n$ is Lipschitz-continuous. Then*

$$\mathcal{H}^s(f(\Upsilon)) \leq L_f \mathcal{H}^s(\Upsilon)$$

for $0 \leq s < \infty$, where L_f denotes the Lipschitz-constant of f on Υ.

Now let $\phi\colon \Upsilon \to \mathbb{R}^n$ be a bi-Lipschitz mapping. Then we have the property

$$\mathbb{B}_{\mathbb{R}^d}\big(\phi^{-1}(\mathrm{y}), L_\phi^{-1} r\big) \subset \phi^{-1}\big(\mathbb{B}_{\mathbb{R}^n}(\mathrm{y}, r)\big) \subset \mathbb{B}_{\mathbb{R}^d}\big(\phi^{-1}(\mathrm{y}), L_{\phi^{-1}} r\big) \qquad (1.25)$$

[4]This is in contrast to Hölder-mappings, see the cited theorem.

for all $y \in \phi(\Upsilon)$, where L_ϕ and $L_{\phi^{-1}}$ again denote the Lipschitz-constants of ϕ and ϕ^{-1}. Together with Lemma 1.3.2, (1.25) implies that the class of N-sets is invariant under bi-Lipschitz mappings, cf. also [94, Ch. II.1, Ex. 1].

Corollary 1.3.3. *Let $\Upsilon \subset \mathbb{R}^d$, assume that $\phi\colon \Upsilon \to \mathbb{R}^n$ is a bi-Lipschitz mapping and let $0 < N \le d$. Then Υ is an N-set if and only if $\phi(\Upsilon)$ is an N-set.*

We aim to characterize the local boundary regularity of a domain $\Lambda \subset \mathbb{R}^d$ by bi-Lipschitz mappings between neighborhoods of boundary points and suitably regular model sets. The model sets we use are as follows, cf. also Figure 1.1:

$$
\begin{aligned}
K &:= (-1,1)^d, && \text{the open unit cube in } \mathbb{R}^d \text{ centered around } 0,\\
K^- &:= \{x \in K : x_d < 0\}, && \text{the open “lower” half of } K,\\
\Sigma &:= \{x \in K : x_d = 0\}, && \text{the midplate of } K \text{ or “upper” plate of } K^-,\\
\Sigma^- &:= \{x \in \Sigma : x_{d-1} \le 0\}, && \text{the (relatively) closed “left” half of } \Sigma.
\end{aligned}
$$

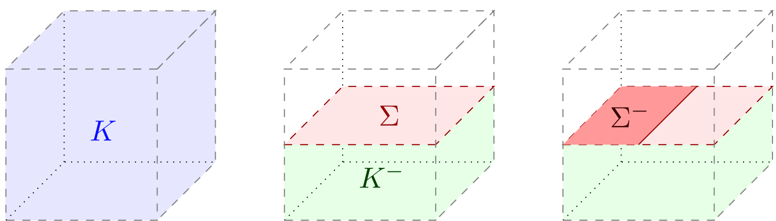


Figure 1.1. Model sets for $d = 3$

The idea of using bi-Lipschitz mappings onto model sets to describe the local geometry around the boundary is exactly the foundation of the fundamental geometric framework of a Lipschitz domain ([112, Ch. 1.1.9], [72, Ch. 1.2]):

Definition 1.3.4 (Lipschitz domain)**.** Let $\Lambda \subset \mathbb{R}^d$ be a (bounded) domain. Then we say that Λ is a (bounded) *Lipschitz domain* if for every $\mathrm{x} \in \partial\Lambda$ there exists an open neighborhood U_x of x, a bi-Lipschitz mapping ϕ_x from an open neighborhood of $\overline{U_\mathrm{x}}$ into $\mathbb{R}^d$, and a number $\tau_\mathrm{x} > 0$ such that $\phi_\mathrm{x}(\mathrm{x}) = 0$ with $\phi_\mathrm{x}(U_\mathrm{x}) = \tau_\mathrm{x} K$ and $\phi_\mathrm{x}(U_\mathrm{x} \cap \Lambda) = \tau_\mathrm{x} K^-$.

From the requirements in Definition 1.3.4 it follows that for each boundary point $\mathrm{x} \in \partial\Lambda$, the set $U_\mathrm{x} \cap \Lambda$ has only one connected component, that is, Λ always lies on only *one* side of its boundary; in other words, Λ does not touch itself. Moreover, it also follows that $\phi_\mathrm{x}(U_\mathrm{x} \cap \partial\Lambda) = \tau_\mathrm{x}\Sigma$, which is indeed the "illustrative" information one looks for. The actual choice of the model sets K etc. is not necessarily restricted to the ones presented above; it is only necessary that the model sets (or, more precisely, the designated image of $U_\mathrm{x} \cap \partial\Lambda$) are regular enough and force the domain to lie on one side of its boundary. From this it is clear that the condition $\phi_\mathrm{x}(\mathrm{x}) = 0$ in Definition 1.3.4 is only present for convenience at this point. By the above definition, a Lipschitz domain in $\mathbb{R}^d$ is exactly the analogue of a d-dimensional C^1-manifold in $\mathbb{R}^d$ for the Lipschitz setting, justifying the name *boundary charts* for the mappings ϕ. The above notion of Lipschitz domain is more general than the one of a *strong Lipschitz-* or *Lipschitz graph domain*, which is in fact equal to a *uniform cone domain*, cf. [72, Thm. 1.2.2.2]. The most striking example of a domain that is a Lipschitz domain, but *not* a Lipschitz graph domain is that of crossing beams (see also [112, Ch. 1.1.9] for a more abstract example). Interestingly, in more smooth situations, requiring Λ to be, say, a d-dimensional C^k-manifold in $\mathbb{R}^d$ for $k \geq 1$ is equivalent to requiring Λ to be a C^k-graph domain. The failure of this equivalency for the Lipschitz situation is ultimately attributed to the failure of the implicit function theorem for mere Lipschitz functions, see [72, Thm. 1.2.1.5].

Remark 1.3.5. Further deepening the likeness between Lipschitz domain and Lipschitz manifold, there is a surface measure ω for Lipschitz domains Λ which on the one hand coincides with the standard surface measure on more smooth domains and, on the other hand, is exactly the restriction of the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$ to $\partial\Lambda$. The measure moreover satisfies the condition (1.21) and is thus an upper $(d-1)$-Ahlfors measure on $\partial\Lambda$. See [81, Ch. 3.1] and [61, Ch. 3.3.4C] and the remarks there for details.

Next we incorporate a setting for mixed boundary conditions while staying in the class of Lipschitz domains. Let $\Xi \subseteq \partial\Lambda$ be closed, allowed to be empty. If we assume Λ to be a Lipschitz domain in the first place, then there is already enough information for the extension property to hold without any further assumptions on Ξ.

Theorem 1.3.6. *Let $\Lambda \subset \mathbb{R}^d$ be a Lipschitz domain. Then it is a universal extension domain. Moreover, if $\Xi \subseteq \partial\Lambda$ is closed, allowed to be empty, then Λ is a universal $\mathrm{W}^{k,p}_{\Xi}$-extension domain.*

Proof. The universal extension property of Lipschitz domains is a classical result, see e.g. [67, Thm. 7.25]. Corollary 1.2.33 then readily yields the remaining claim. □

It is not surprising that the boundary of a Lipschitz domain $\Lambda \subset \mathbb{R}^d$ is well-behaved in the sense of being a $(d-1)$ set, which allows to obtain a continuous (even compact) boundary trace when combined with Theorem 1.3.6.

Lemma 1.3.7. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Then $\partial\Lambda$ is a $(d-1)$ set with $\mathcal{H}^{d-1}(\partial\Lambda) < \infty$.*

Proof. The collection of sets U_{x} from Definition 1.3.1 for $\mathrm{x} \in \partial\Lambda$ forms an open covering of $\partial\Lambda$. But since Λ is bounded, $\partial\Lambda$ is a compact subset of $\mathbb{R}^d$ and may thus be covered by finitely many of the neighborhoods $U_{\mathrm{x}_1}, \dots, U_{\mathrm{x}_n}$. For each of these neighborhoods, $U_{\mathrm{x}_i} \cap \partial\Lambda$ is mapped onto $\tau_j\Sigma$ by the bi-Lipschitz mapping ϕ_j. Using the property of the model sets $\tau_i\Sigma$ to be $(d-1)$-sets, see [94, Ch. II.1, Ex. 1], Corollary 1.3.3 shows that $\partial\Lambda$ is in fact the union of finitely many $(d-1)$-sets $U_{\mathrm{x}_i} \cap \partial\Lambda$ and thus a $(d-1)$-set itself, cf. Remark 1.2.37. Since the model surfaces $\tau_j\Sigma$

have finite measure w.r.t. $\mathcal{H}^{d-1}$, so have $U_{\mathrm{x}_i} \cap \partial\Lambda$ by Lemma 1.3.2 and $\mathcal{H}^{d-1}(\partial\Lambda)$ follows as in Remark 1.2.52. □

Corollary 1.3.8. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded Lipschitz domain and set $\omega = \mathcal{H}^{d-1} \restriction \partial\Lambda$. Then ω satisfies the measure condition* (1.21)*. In particular, the assertions of Lemma 1.2.57 hold true.*

Proof. The claim follows from $\partial\Lambda$ having finite $\mathcal{H}^{d-1}$-measure by Lemma 1.3.7 and then using Corollary 1.2.53. This also gives an alternative proof to the claim that the surface measure ω as noted in Remark 1.3.5 satisfies the measure condition (1.21). Theorem 1.3.6 now allows to use Lemma 1.2.57. □

While Theorem 1.3.6 and Corollary 1.3.8 answer the questions of the extension property and (compact) boundary trace for Lipschitz domains, they are not the end of the story. For instance, we have already noted in Remark 1.2.45 that the question of interpolation of (at least) $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ spaces is not yet (fully) answered in the generality assumed so far. To make use of such and more results, we need more decisive information about Ξ. Intuitively, it is clear that a crucial point in this situation is the regularity of the boundary of Ξ within $\partial\Lambda$. Note that a condition on this regularity is implicitly also contained in the assumptions of Theorem 1.2.31 by including the *closure* $\overline{\Gamma}$ of $\Gamma = \partial\Lambda \setminus \Xi$ in the conditions for the local extension property. The precise assumptions are given in the following definition which was introduced by Gröger in [73], calling the set $\Lambda \cup \Xi$ *regular*. The concept of regular sets has turned out to be very suitable for a functional-analytic treatment, cf. for instance [52, 71, 74, 75, 79–81, 87, 88, 98].

Definition 1.3.9 (Regular in the sense of Gröger)**.** Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $\Xi \subseteq \partial\Lambda$ be a closed subset of its boundary, allowed to be empty.

(i) We say that $\Lambda \cup \Xi$ is *regular in the sense of Gröger* if for every $\mathrm{x} \in \partial\Lambda$ there exists an open neighborhood U_{x} of x, a number $\tau_{\mathrm{x}} > 0$, and a bi-Lipschitz mapping ϕ_{x} from an open neighborhood of $\overline{U_{\mathrm{x}}}$ into $\mathbb{R}^d$

such that $\phi_{\mathrm{x}}(\mathrm{x}) = 0$ and $\phi_{\mathrm{x}}(U_{\mathrm{x}} \cap (\Lambda \cup \Xi))$ is either one of the model sets $\tau_{\mathrm{x}} K^-$ or $\tau_{\mathrm{x}}(K^- \cup \Sigma)$ or $\tau_x(K^- \cup \Sigma^-)$.

(ii) We further say that $\Lambda \cup \Xi$ is *volume-preserving* regular in the sense of Gröger, if each of the mappings ϕ_{x} are volume-preserving, satisfying $\det \phi'_{\mathrm{x}} \equiv 1$ almost everywhere.

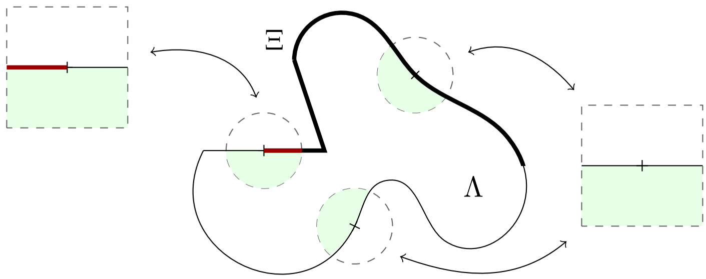


Figure 1.2. Regular in the sense of Gröger for $d = 2$

Remark 1.3.10. Recall $\Gamma = \partial\Lambda \setminus \Xi$. The following properties of regular sets are shown in [79, Ch. 5]:

(i) If $\Lambda \cup \Xi$ is regular in the sense of Gröger, then Λ is a Lipschitz domain.

(ii) There are the following alternative characterizations of regular sets for space dimensions $d = 2$ and $d = 3$: $\Lambda \cup \Xi$ is regular in the sense of Gröger, iff Λ is a Lipschitz domain and the following two conditions are satisfied:

- $d = 2$: the boundary $\Xi \cap \overline{\Gamma}$ of Ξ within $\partial\Lambda$ is a finite set and no connected component of Ξ consists of a single point,
- $d = 3$: Ξ is the closure of its relative interior in $\partial\Lambda$ and its boundary $\Xi \cap \overline{\Gamma}$ within $\partial\Lambda$ is locally bi-Lipschitz diffeomorphic to the open unit interval $(-1, 1)$.

Note that the authors of [79] call $\Lambda \cup \Gamma$ regular instead of $\Lambda \cup \Xi$, which gives an equivalent definition.

Clearly, the notions of regular sets and Lipschitz domain are closely intertwined. Indeed, if Λ is a Lipschitz domain, then we can already find neighborhoods U_{x}, bi-Lipschitz mappings ϕ_{x} and numbers τ_{x} such that $\phi_{\mathrm{x}}(U_{\mathrm{x}} \cap (\Lambda \cup \Xi)) = \phi_{\mathrm{x}}(U_{\mathrm{x}} \cap \Lambda) = \tau_{\mathrm{x}} K^-$ for $\mathrm{x} \notin \Xi$ and $\phi_{\mathrm{x}}(U_{\mathrm{x}} \cap (\Lambda \cup \Xi)) = \tau_{\mathrm{x}}\bigl(K^- \cup \Sigma\bigr)$ for $\mathrm{x} \in \Xi \setminus \overline{\Gamma}$. The only "new" requirement in Definition 1.3.9 compared to Definition 1.3.4 is thus the regularity around $\Xi \cap \overline{\Gamma}$. This is of course also reflected by the alternative characterizations in Remark 1.3.10. We also note that, in contrast to Definition 1.3.4, the condition $\phi_{\mathrm{x}}(\mathrm{x}) = 0$ in Definition 1.3.9 is *not* posed for convenience, but is critical in its interplay with the set Σ^-, as 0 is a boundary point of Σ^-.

Remark 1.3.11. The additional property of *volume-preserving* bi-Lipschitz mappings ϕ_{x} has been required in several different contexts, e.g. [74] by Gröger himself or [69], the latter being the paper which establishes interpolation properties for $W^{\pm 1,q}_{\Xi}(\Lambda)$ spaces for $\Lambda \cup \Xi$ regular in the sense of Gröger. Albeit seeming restrictive, the class of volume-preserving bi-Lipschitz mappings has turned out to be quite capable of mapping "non-smooth" objects onto smooth ones, such as the ball onto the cylinder, the cube or another half ball, cf. [70] or [64]. In particular, it is easily seen that strong Lipschitz domains always admit bi-Lipschitz boundary charts which are volume preserving. In this sense, the requirement of volume-preserving maps ϕ_{x} should be seen as not too restrictive. Moreover, it should be possible to get rid of this requirement in many contexts by introducing function spaces with weights, which has not been worked out yet, however.

For $\Lambda = \emptyset$ or $\Lambda = \partial\Lambda$, we recover exactly Definition 1.3.4 from Definition 1.3.9, i.e., the concept of *regular in the sense of Gröger* is a true generalization of that of a Lipschitz domain. In this sense, we accept the concept of Gröger's regular sets as the adequate construct for mixed boundary conditions within the class of Lipschitz domains. We introduce another generalization with which we leave the Lipschitz domain class, allowing Λ to lie on both sides of Ξ in the following sense, see also Figure 1.3. Recall again $\Gamma = \partial\Lambda \setminus \Xi$.

Definition 1.3.12 (Generalized regular in the sense of Gröger)**.** Let Λ be a bounded domain in $\mathbb{R}^d$ and let $\Xi \subseteq \partial\Lambda$ be a closed subset of its boundary, allowed to be empty. We say that $\Lambda \cup \Xi$ is *generalized regular in the sense of Gröger*, if the following conditions are satisfied:

(i) For every $\mathrm{x} \in \partial\Lambda \setminus \overline{\Gamma}$ there exists a domain $\mathcal{U}_\mathrm{x}$ with $\mathrm{x} \in \mathcal{U}_\mathrm{x}$ satisfying $\mathcal{U}_\mathrm{x} \cap \overline{\Gamma} = \emptyset$ and such that $\mathcal{U}_\mathrm{x} \cap \Lambda$ consists of k_x connected components $V_1, \dots, V_{k_\mathrm{x}}$. For each $j \in 1, \dots, k_\mathrm{x}$, there moreover exists an open neighborhood U_j of x such that $V_j \subset U_j \subseteq \mathcal{U}_\mathrm{x}$, a bi-Lipschitz mapping ϕ_j which maps an open neighborhood of $\overline{U_j}$ into $\mathbb{R}^d$, and a number $\tau_j > 0$ such that $\phi_j(\mathrm{x}) = 0$ with $\phi_j(U_j) = \tau_j K$ and $\phi_j\big(V_j \cup (\partial V_j \cap U_j)\big) = \tau_j(K^- \cup \Sigma)$.

(ii) For every $\mathrm{x} \in \overline{\Gamma}$ there exists an open neighborhood U_x of x, a bi-Lipschitz mapping ϕ_x which maps an open neighborhood of $\overline{U_\mathrm{x}}$ into $\mathbb{R}^d$, and a number $\tau_\mathrm{x} > 0$ such that $\phi_\mathrm{x}(\mathrm{x}) = 0$ with $\phi_\mathrm{x}(U_\mathrm{x}) = \tau_\mathrm{x} K$. Moreover, the following mapping properties for the boundary are true:

   (a) If $\mathrm{x} \in \partial\Lambda \setminus \Xi$, then $U_\mathrm{x} \cap \Xi = \emptyset$ or equivalently $\phi_\mathrm{x}\big(U_\mathrm{x} \cap (\Lambda \cup \Xi)\big) = \tau_\mathrm{x} K^-$,

   (b) if $\mathrm{x} \in \Xi \cap \overline{\Gamma}$, then $\phi_\mathrm{x}\big(U_\mathrm{x} \cap (\Lambda \cup \Xi)\big) = \tau_\mathrm{x}(K^- \cup \Sigma^-)$.

We further say that $\Lambda \cup \Xi$ is *volume-preserving* generalized regular in the sense of Gröger, if each of the occurring bi-Lipschitz mappings ϕ in both previous cases is *volume-preserving*, i.e., $\det |D\phi| \equiv 1$ almost everywhere.

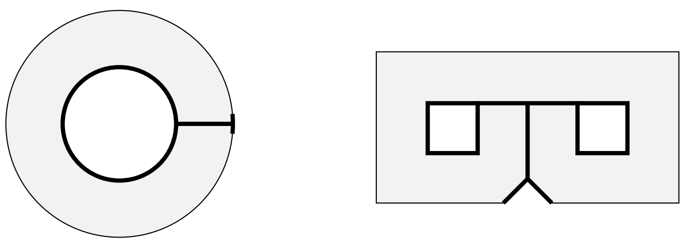


Figure 1.3. Exemplary sets Λ which are generalized regular in the sense of Gröger, with Ξ marked with thick lines.

Remark 1.3.13. Of course, the reason for freely talking about derivatives Df of Lipschitz-functions f is the theorem of RADEMACHER, which says that a (locally) Lipschitz-continuous function $f\colon \mathbb{R}^d \to \mathbb{R}^m$ for some $m \in \mathbb{N}$ is differentiable λ^d-almost everywhere [61, Ch. 3.1.2, Thm. 2].

Definition 1.3.12 relaxes the conditions of Definition 1.3.9 on the relative interior $\partial\Lambda \setminus \overline{\Gamma}$ of Ξ: Clearly, condition (ii) exactly reproduces the condition on the regular sets as in Definition 1.3.9 with a slightly more precise distinction of the two cases, so for $\overline{\Gamma}$ the situation has generally not changed from Definition 1.3.9. For points x from the relative interior $\partial\Lambda \setminus \overline{\Gamma}$ of Ξ however, we now allow a finite number of connected components $V_1, \dots, V_{k_\mathrm{x}}$ of $\mathcal{U}_\mathrm{x} \cap \Lambda$, which in its essence allows the domain to lie on multiple sides of its boundary, i.e., to touch itself, see Figure 1.3.

The regularity requirement for *each* of these connected components is essentially the one for Ξ in Definition 1.3.9 with Λ and Ξ locally replaced by V_j and $\partial V_j \cap \Xi$ and the neighborhood U_j corresponding to U_x. To verify this, let us quickly substitute $\Lambda^* = V_j$ and $\Xi^* = \partial V_j \cap \Xi$. Then we have $U_j \cap \Lambda^* = U_j \cap V_j = V_j$ and $U_j \cap \Xi^* = U_j \cap \partial V_j \cap \Xi = U_j \cap \partial V_j$, thus

$$U_j \cap (\Lambda^* \cup \Xi^*) = U_j \cap (V_j \cup (\partial V_j \cap U_j)) = V_j \cup (\partial V_j \cap U_j).$$

Hence, the condition $\phi_j\big(V_j \cup (\partial V_j \cap U_j)\big) = \tau_j(K^- \cup \Sigma)$ is indeed the analogue to the one in Definition 1.3.9 with the stated identifications, which shows that each connected component V_j satisfies a condition equivalent to the one in Definition 1.3.9.

Remark 1.3.14. Let $\mathrm{x} \in \Xi \setminus \overline{\Gamma}$. For *only one* connected component V_1 of $\mathcal{U}_\mathrm{x} \cap \Lambda$, assuming $U_1 = \mathcal{U}_\mathrm{x}$ for simplicity, we even have $V_1 = U_1 \cap \Lambda$ and $\partial V_1 \cap U_1 = \Xi \cap U_1$. This implies $U_1 \cap (\Lambda \cup \Xi) = V_1 \cup (\partial V_1 \cap U_1)$ and

$$\phi_1\big(U_1 \cap (\Lambda \cup \Xi)\big) = \phi_1\big(V_1 \cup (\partial V_1 \cap U_1)\big) = \tau_1\big(K^- \cup \Sigma\big).$$

Hence, if $\Lambda \cup \Xi$ is (volume-preserving) generalized regular in the sense of Gröger with $k_\mathrm{x} = 1$ for every $\mathrm{x} \in \partial\Lambda \setminus \overline{\Gamma}$, then it is exactly (volume-preserving) regular in the sense of Gröger and vice versa, so we obtain again a true generalization as promised by nomenclature. In particular, Λ is a Lipschitz domain in this case, as it is for $\Xi = \emptyset$.

Considering the "severed" annulus in Figure 1.3, which is generalized regular in the sense of Gröger, it is clear that a set satisfying the requirements of Definition 1.3.12 need not be a $\mathrm{W}^{k,p}$-extension domain.[5] However, thanks to the Lipschitz-conditions on points from $\overline{\Gamma}$, we may apply Theorem 1.2.31:

Theorem 1.3.15. *Let $\Lambda \cup \Xi$ be generalized regular in the sense of Gröger. Then Λ is a universal $\mathrm{W}^{k,p}_{\Xi}$-extension domain.*

Proof. By assumption, $\Xi \subseteq \partial\Lambda$ is closed. We further need that $\Lambda \cap U_{\mathrm{x}}$ for each $\mathrm{x} \in \overline{\Gamma}$, with the open neighborhoods U_{x} provided by Definition 1.3.12, are universal extension domains in order to apply Theorem 1.2.31. But exactly this is proven in [56, Lem. 2.2.20, Thm. 2.2.21], hence the assertion follows. □

Theorem 1.3.16. *Let $\Lambda \cup \Xi$ be generalized regular in the sense of Gröger. Then Ξ is a $(d-1)$-set.*

Proof. For each $\mathrm{x} \in \Xi \setminus \overline{\Gamma}$, consider the domain $\mathcal{U}_{\mathrm{x}}$, the neighborhoods $U_{\mathrm{x},j}$ of the connected components $V_{\mathrm{x},j}$, the numbers $\tau_{\mathrm{x},j}$, the bi-Lipschitz mappings $\phi_{\mathrm{x},j}$ from Definition 1.3.12 (i) and define the sets

$$W_{\mathrm{x},j} := \phi_{\mathrm{x},j}^{-1}(\tau_{\mathrm{x},j}\Sigma) = \partial V_{\mathrm{x},j} \cap U_{\mathrm{x},j},$$

all for $j = 1, \dots, k_{\mathrm{x}}$. For $\mathrm{y} \in \Xi \cap \overline{\Gamma}$ we collect the bi-Lipschitz mappings ϕ_{y} and the neighborhoods U_{y} of y from case (ii) of Definition 1.3.12. Then the systems $\{U_{\mathrm{y}} \cap \Xi \colon \mathrm{y} \in \Xi \cap \overline{\Gamma}\}$ and $\{W_{\mathrm{x},j} : \mathrm{x} \in \Xi \setminus \overline{\Gamma}, 1 \le j \le k_{\mathrm{x}}\}$ form a relatively open covering of the compact set Ξ from which we choose a finite subcovering $\mathscr{C}$, corresponding to the finite number of points $\mathrm{y}_1, \dots, \mathrm{y}_m$

[5] We have already mentioned the standard counter-example for extension domains, the sliced disk, cf. the considerations below Definition 1.2.21; one argues analogously.

and $x_1, \dots, x_n$. For $x \in \Xi \setminus \overline{\Gamma}$, define $J(x) = \{j \in \{1, \dots, k_x\} : W_{x,j} \in \mathscr{C}\}$. Then we have

$$\mathscr{C} = \{U_{y_1} \cap \Xi, \dots, U_{y_m} \cap \Xi\} \cup \bigcup_{\ell=1}^{n} \{W_{x_\ell,j} : j \in J(x_\ell)\}$$

which means we can write Ξ in the form

$$\Xi = \bigcup_{\ell=1}^{m} (U_{y_l} \cap \Xi) \cup \bigcup_{\ell=1}^{n} \bigcup_{j \in J(x_\ell)} W_{x_l,j}.$$

The next step is now to note that every finite union of $(d-1)$-sets is again a $(d-1)$-set, cf. [56, Lem. 1.2.24]. Hence, one has to show only that each of the sets forming $\mathscr{C}$ is a $(d-1)$-set. For the sets $\Xi \cap U_{y_\ell}$ this is immediate by Corollary 1.3.3 and the supposition on the mappings ϕ_{y_ℓ}, whereas each $W_{x_\ell,j}$ is exactly constructed as the bi-Lipschitz preimage of the $(d-1)$-set $\tau_{x_\ell,j}\Sigma$, such that again Corollary 1.3.3 applies and shows that $W_{x_\ell,j}$ is also a $(d-1)$-set. This altogether makes Ξ a $(d-1)$-set. □

The next aim is to establish the existence of a surface measure for Λ by giving an explicit construction if $\Lambda \cup \Xi$ is a generalized regular set, which extends the considerations in [81, Ch. 3.1] to sets which are generalized regular in the sense of Gröger, compare also [61, Ch. 3.3.4C] and Remark 1.3.5. We are interested in having a surface measure for the boundary of a domain Λ which agrees with restriction of the $(d-1)$-dimension Hausdorff measure because the coincidence with the Hausdorff measure is the reassurance that the constructed measure behaves reasonably in measuring $(d-1)$-dimensional content and that the Gauss-Green theorem in its usual form is applicable [61, Ch. 5.8], which is a link between strong and weak formulations of PDEs. In order to establish such a result, we quote, with suitable modifications, the following change of variables formula:

Theorem 1.3.17 (Change of variables, [61, Ch. 3.3.3, Thm. 2])**.** *Let $f : \mathbb{R}^{d-1} \to \mathbb{R}^d$ be Lipschitz-continuous. Then, for all $g \in L^1(\mathbb{R}^{d-1})$,*

$$\int_{\mathbb{R}^{d-1}} g(\mathrm{x})\sqrt{\det\big(Df(\mathrm{x})^\top Df(\mathrm{x})\big)}\,\mathrm{dx} \\ = \int_{\mathbb{R}^d} \Bigg[\sum_{\mathrm{x}\in f^{-1}(\mathrm{y})} g(\mathrm{x})\Bigg]\,\mathrm{d}\mathcal{H}^{d-1}(\mathrm{y}). \quad (1.26)$$

Remark 1.3.18. The assumption that the functions in Theorem 1.3.17 are defined on the entire Euclidean space is no loss of generality. For a function $g \in \mathrm{L}^1(\Upsilon)$, we know that the extension $\mathfrak{E}^0 g$ by zero is again in $\mathrm{L}^1(\mathbb{R}^d)$, while a Lipschitz-continuous function $f\colon \Upsilon \to \mathbb{R}^n$ may be extended to a Lipschitz-function $\bar{f}$ on the whole space by [61, Ch. 3.1.1, Thm. 1]. Due to the extension of g by zero, we may thus replace the integration over $\mathbb{R}^{d-1}$ in (1.26) by an integration over Υ and still use the assertion.

Clearly, if f in Theorem 1.3.17 is bijective, the integrand on the right hand side simplifies to $g\big(f^{-1}(\mathrm{y})\big)$, more resembling the usual change of variables formula. The interesting thing about the previous theorem is that it is also a transformation formula for *measures* – from the $(d-1)$-dimensional Lebesgue measure in $\mathbb{R}^{d-1}$, which is the same as the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$ on $\mathbb{R}^{d-1}$, to $\mathcal{H}^{d-1}$ in $\mathbb{R}^d$. We use (1.26) in the following theorem to establish the coincidence with the Hausdorff measure.

Theorem 1.3.19. *Let $\Lambda \cup \Xi$ be generalized regular in the sense of Gröger. Then there exists a surface measure ω on $\partial\Lambda$ which coincides with $\mathcal{H}^{d-1} \upharpoonright \partial\Lambda$. Moreover, $\omega(\partial\Lambda) < \infty$, and thus* (1.21) *holds true for $\mu = \omega$.*

Proof. We define a general integration formula for measurable functions on $\partial\Lambda$, where "measurable" is to be understood with respect to the restriction of the Borel-σ-algebra on $\mathbb{R}^d$ to $\partial\Lambda$. The measure of such a measurable set $\Upsilon \subset \partial\Lambda$ is then given by integrating the indicator function χ_Υ, and the boundary $\partial\Lambda$ itself is measurable since it is a Borel set in $\mathbb{R}^d$ in the first place. The idea is, of course, again to use the bi-Lipschitz charts to describe the surface $\partial\Lambda$ or subsets thereof via the model set Σ.

From Definition 1.3.12, we choose an open covering of $\partial\Lambda$ as follows: For each point $\mathrm{y} \in \overline{\Gamma}$ we choose the open neighborhood U_y, and for each $\mathrm{x} \in \Xi\backslash\overline{\Gamma}$ we choose $W_\mathrm{x} := \cup_{j=1}^{k_\mathrm{x}} U_{j,\mathrm{x}}$, which is again an open neighborhood of x. Clearly, the union of all these sets forms an open covering of the compact set $\partial\Lambda$, hence there exists a finite subcovering which we denote by $W_{\mathrm{x}_1} \cap \partial\Lambda, \ldots, \mathrm{W}_{\mathrm{x}_n} \cap \partial\Lambda, U_{\mathrm{y}_1} \cap \partial\Lambda, \ldots, U_{\mathrm{y}_m} \cap \partial\Lambda$. Let $\eta_1, \ldots, \eta_{n+m}$ be a continuous partition of unity on $\partial\Lambda$ subordinated to this finite open covering. Now, for each y_ℓ, there exists a bi-Lipschitz mapping ϕ_{y_ℓ} between $U_{\mathrm{y}_\ell} \cap \partial\Lambda$ and $\tau_{\mathrm{y}_\ell}\Sigma$ and we define $\zeta_\ell(\mathrm{z}_1, \ldots, \mathrm{z}_{d-1}) := \phi_{\mathrm{y}_\ell}^{-1}(\tau_{\mathrm{z}_\ell}\mathrm{z}_1, \ldots, \tau_{\mathrm{z}_\ell}\mathrm{z}_{d-1}, 0)$ for $\mathrm{z} \in (-1,1)^{d-1}$ or, equivalently, $(\tau_{\mathrm{y}_\ell}\mathrm{z}, 0) \in \tau_{\mathrm{y}_\ell}\Sigma$. Analogously, for each x_i, there exist k_i bi-Lipschitz mappings $\phi_{\mathrm{x}_i,j}$, each between $\partial V_{\mathrm{x}_i,j} \cap \partial\Lambda$ and $\tau_{\mathrm{x}_i,j}\Sigma$, and we define $\zeta_{i,j}(\mathrm{z}_1, \ldots, \mathrm{z}_{d-1}) := \phi_{\mathrm{x}_i,j}^{-1}(\tau_{\mathrm{x}_i,j}\mathrm{z}_1, \ldots, \tau_{\mathrm{x}_i,j}\mathrm{z}_{d-1}, 0)$ on $(-1,1)^{d-1}$. Clearly, the mappings ζ are still bi-Lipschitz ones. For $i \in \{1, \ldots, n\}$, let $\eta_1^i, \ldots, \eta_{k_i}^i$ also be continuous partitions of unity of $W_{\mathrm{x}_i} \cap \partial\Lambda$, subordinated to the open coverings $W_{\mathrm{x}_i} \cap \partial\Lambda = \cup_{j=1}^{k_i}(\partial V_{\mathrm{x}_i,j} \cap \partial\Lambda)$. Lastly, we define the Jacobian determinants as functions on $(-1,1)^{d-1}$ by

$$\mathcal{J}_\ell = \sqrt{\det\big((D\zeta_\ell)^\top D\zeta_\ell\big)} = \left(\sum_{k=1}^{d-1} \det\left(\frac{\partial(\zeta_\ell^1, \ldots, \zeta_\ell^{k-1}, \zeta_\ell^{k+1}, \ldots \zeta_\ell^d)}{\partial(\mathrm{z}_1, \ldots, \mathrm{z}_{d-1})}\right)^2\right)^{\frac{1}{2}},$$

with the analogous definition for $\mathcal{J}_{i,j}$ on $(-1,1)^{d-1}$ for $j \in k_i$, cf. also [61, Ch. 3.2.1]. Finally, let $f \colon \partial\Lambda \to \mathbb{R}$ be measurable in the above sense. Then we set

$$\begin{aligned}\int_{\partial\Lambda} f \,\mathrm{d}\omega := \int_{(-1,1)^{d-1}} \sum_{\ell=1}^{m} \big[(\eta_{n+\ell} \cdot f) \circ \zeta_\ell\big] \cdot \mathcal{J}_\ell \\ + \sum_{i=1}^{n} \sum_{j=1}^{k_i} \Big[(\eta_i \cdot \eta_j^i \cdot f) \circ \zeta_{i.j}\Big] \cdot \mathcal{J}_{i,j} \,\mathrm{d}\mathrm{z}. \quad (1.27)\end{aligned}$$

Inserting $f = \chi_\Upsilon$ and using (1.26) from Theorem 1.3.17 together with the properties of the partitions of unity, we indeed obtain

$$\omega(\Upsilon) = \int_\Upsilon 1 \,\mathrm{d}\omega = \mathcal{H}^{d-1}(\Upsilon)$$

for every measurable set $\Upsilon \subseteq \partial\Lambda$.

It only remains to show that $\mathcal{H}^{d-1}(\partial\Lambda) < \infty$. This follows from boundedness of Λ which allows to work with a finite covering as introduced above. Then, the integral on the right-hand side in (1.27) is bounded by at most $(m+k_1+\cdots+k_n)$ times the maximal Jacobian determinant over $(-1,1)^{d-1}$ which is finite, cf. [61, Ch. 4.2.3]. Hence, $\mathcal{H}^{d-1}(\partial\Lambda) < \infty$ and we can use Corollary 1.2.53 to infer that (1.21) holds true for $\mu = \omega$. □

Remark 1.3.20.

(i) From the construction of ω and (1.25), we infer that if $\Lambda \cup \Xi$ is (generalized) regular in the sense of Gröger and $\Xi \neq \emptyset$, then $\omega(\Xi) > 0$. In particular, there is a Poincaré inequality available in this case by [155, Thm. 4.8.1], cf. also [81, Thm. 3.5] and [20, Rem. 3.4], i.e., $f \mapsto \|\nabla f\|_{\mathrm{L}^p(\Lambda)}$ is an equivalent norm on $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ for $1 < p < \infty$.

(ii) Theorem 1.3.19 allows to use Lemma 1.2.57 for a boundary trace operator onto $\Gamma = \overline{\partial\Lambda \setminus \Xi}$ and subsets thereof, if $\Lambda \cup \Xi$ is generalized regular in the sense of Gröger (see the proof of Theorem 1.3.15 for the local extension property requirement in Lemma 1.2.57).

We introduce one last class of domains which includes the foregoing ones. Here, we go one step further compared to Definition 1.3.12 and also dispose of the Lipschitz charts for Ξ.

Definition 1.3.21 (Lipschitz around $\partial\Lambda \setminus \Xi$)**.** Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain and let $\Xi \subset \partial\Lambda$ be a closed $(d-1)$-set. Then we say that $\Lambda \cup \Xi$ is *Lipschitz around* $\partial\Lambda \setminus \Xi = \Gamma$ if for every $\mathrm{x} \in \overline{\Gamma}$ there exists an open neighborhood U_{x} of x, a number $\tau_{\mathrm{x}} > 0$, and a bi-Lipschitz mapping ϕ_{x} from an open neighborhood of $\overline{U_{\mathrm{x}}}$ into $\mathbb{R}^d$ such that $\phi_{\mathrm{x}}(\mathrm{x}) = 0$ and $\phi_{\mathrm{x}}(U_{\mathrm{x}} \cap \Lambda) = \tau_{\mathrm{x}} K^-$ as well as $\phi_{\mathrm{x}}(U_{\mathrm{x}} \cap \partial\Lambda) = \tau_{\mathrm{x}}(K^- \cup \Sigma)$

It is clear from the definition, Lemma 1.3.7 and Theorem 1.3.16 that both $\Lambda \cup \Xi$ being regular or generalized regular in the sense of Gröger implies being Lipschitz around $\partial\Lambda \setminus \Xi$. Observing the proof of Theorem 1.3.15, we note that the result there, Λ being a universal extension $\mathrm{W}^{k,p}_{\Xi}$-extension domain if $\Lambda \cup \Xi$ was regular in the sense of Gröger, did not rely on the Lipschitz charts around Ξ at all. Hence, we obtain the same result for the new class of domains.

Theorem 1.3.22. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$. Then Λ is a universal $\mathrm{W}^{k,p}_{\Xi}$-extension domain.*

There is also a surface measure available on $\overline{\Gamma}$ for $\Lambda \cup \Xi$ Lipschitz around $\partial\Lambda \setminus \Xi$, and this measure coincides with the restriction of the $(d-1)$-dimensional Hausdorff-measure to $\overline{\Gamma}$. In particular, it satisfies $\mathcal{H}^{d-1}(\overline{\Gamma}) < \infty$ and thus the measure condition (1.21) by Corollary 1.2.53. Since we have seen the details of the proofs of corresponding results for the other classes of domains already, cf. Corollary 1.3.8 and Theorem 1.3.19, we just collect the result.

Theorem 1.3.23. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$. Then there exists a surface measure ω on $\overline{\Gamma}$ which coincides with $\mathcal{H}^{d-1} \restriction \overline{\Gamma}$. Moreover, $\omega(\overline{\Gamma}) < \infty$ and thus (1.21) holds true for $\mu = \omega$.*

At this point, let us comment on where to place the regularity assumptions from Definition 1.3.12 or Definition 1.3.21 within the possible general settings. The developments at the "frontier" of function space theory on highly irregular domains and the associated questions have taken rather (positively) astonishing forms in recent years, leaving the already general class of Lipschitz domains behind by far. It seems safe to say that the concept of N-set itself, in the particular form of $(d-1)$-set for boundary parts Ξ or d-set for Λ, has turned out to be very powerful and allows to obtain a large array of results for mixed boundary problems and the corresponding spaces $\mathrm{W}^{k,p}_{\Xi}(\Lambda)$ previously known only for much smoother cases. We have moreover already listed some results above which do not need any kind of boundary charts or even an actual description of the boundary regularity around $\partial\Lambda \setminus \Xi$, cf. Theorem 1.2.31 (where Ξ only needs to be closed) or Theorem 1.2.55. However, it seems that for more in-depth results, at least local descriptions within the Lipschitz context are still needed, cf. [20, 56, 58, 69], see also [28] for the analogous idea with (ε, δ)-domains. We will see later that our own little contribution in Chapter 2.1 for generalized regular in the sense of Gröger domains relies critically on such descriptions even on Ξ. This concept anyway seems to sit at a crossroads: on the one hand, it is generally outside of the Lipschitz domain class (cf. Figure 1.3) and Ξ is indeed a $(d-1)$-set by Theorem 1.3.16, just as in Definition 1.3.21, but on the other hand we have constructed it in such a way that there are still locally tractable properties of the boundary which allow for localization procedures.

1.4 Maximal parabolic regularity

We are ultimately concerned with abstract quasilinear parabolic partial differential equations in divergence form. A particular ansatz to deal with this kind of equations relies on *maximal parabolic regularity*, which we introduce in an abstract setting in this section. The usefulness of maximal parabolic regularity is however by far not limited to quasilinear equations, but generally allows to treat discontinuous inhomogeneities in evolution equations and many more applications. We refer to the monographs [3, 100] for more general information and historical remarks.

Let us fix a few often needed objects for the following. Set $J = (T_0, T_1) \subset \mathbb{R}_0^+$ for $T_0 < T_1 < \infty$ to be a given interval, and let X, Y be two Banach spaces with $Y \hookrightarrow_{\mathrm{d}} X$.

Definition 1.4.1 (Maximal regularity spaces)**.** Let $1 < r, s < \infty$.

(i) We set

$$\mathrm{W}^{1,r}(J;X) \coloneqq \Big\{ f \in \mathrm{L}^r(J;X) \colon f' \in \mathrm{L}^r(J;X) \Big\},$$

with

$$\|f\|_{\mathrm{W}^{1,r}(J;X)} \coloneqq \Big(\|f'\|^r_{\mathrm{L}^r(J;X)} + \|f\|^r_{\mathrm{L}^r(J;X)} \Big)^{\frac{1}{r}}.$$

Here, f' denotes the distributional derivative in the sense of vector-valued distributions, i.e., f' in the above definition satisfies

$$\int_J f(t)\phi'(t)\,\mathrm{d}t = -\int_J f'(t)\phi(t)\,\mathrm{d}t \quad \text{in } X \quad \text{for all } \phi \in \mathrm{C}_c^\infty(J)$$

and is an element of $\mathrm{L}^r(J;X)$.

(ii) We further define the *maximal regularity spaces*

$$\mathbb{W}_s^{1,r}(J;X,Y) := \mathrm{W}^{1,r}(J;X) \cap \mathrm{L}^s(J;Y),$$

equipped with the norm

$$\|f\|_{\mathbb{W}_s^{1,r}(J;X,Y)} := \|f'\|_{\mathrm{L}^r(J;X)} + \|f\|_{\mathrm{L}^s(J;Y)}$$

and set $\mathbb{W}^{1,r}(J;X,Y) := \mathbb{W}_r^{1,r}(J;X,Y)$.

Of course, $\mathrm{W}^{1,r}(J;X)$ and thereby also $\mathbb{W}_s^{1,r}(J;X,Y)$ is a Banach space (cf. [146, Ch. 1.8.1]).

Remark 1.4.2. If $Y \hookrightarrow X$, then the chosen norm for $\mathbb{W}_s^{1,r}(J;X,Y)$ is equivalent to the usual sum norm on an intersection space, that is,

$$\|f\|_{\mathbb{W}_s^{1,r}(J;X,Y)} \cong \|f\|_{\mathrm{W}^{1,r}(J;X)} + \|f\|_{\mathrm{L}^s(J;Y)} \quad \text{for all } f \in \mathbb{W}_s^{1,r}(J;X,Y).$$

The upper bound "$\leq$" is obvious. For the lower bound, we get rid of the root in the $\mathrm{W}^{1,r}(J;X)$ norm via $a^r + b^r \leq (a+b)^r$ for all $1 < r < \infty$ and $a, b \geq 0$. If $s \geq r$, then the embedding $\mathrm{L}^s(J;Y) \hookrightarrow \mathrm{L}^r(J;X)$ directly implies the lower bound "$\geq$". If $s < r$, then from $f \in \mathrm{W}^{1,r}(J;X)$ we infer that $f \in \mathrm{W}^{1,s}(J;X)$. This space is continuously embedded into $\mathrm{L}^\infty(J;X)$ (see e.g. (1.30) below for $X = Y$) and thus

$$\begin{aligned}\|f\|_{\mathrm{L}^r(J;X)} \leq C\|f\|_{\mathrm{W}^{1,s}(J;X)} &= C\big(\|f'\|_{\mathrm{L}^s(J;X)}^s + \|f\|_{\mathrm{L}^s(J;X)}^s\big)^{\frac{1}{s}} \\ &\leq C\big(\|f'\|_{\mathrm{L}^r(J;X)}^s + \|f\|_{\mathrm{L}^s(J;Y)}^s\big)^{\frac{1}{s}}.\end{aligned}$$

From this the lower bound follows again with the elementary inequality as above.

The maximal regularity spaces have the following crucial properties, which relates them to (Hölder-) continuous functions with values in interpolation spaces:

Proposition 1.4.3 ([53, Lem. 3.4])**.** *Let X, Y be two Banach spaces with $Y \hookrightarrow X$ and let $1 < r < \infty$. Then the following embeddings are true:*

$$\mathbb{W}^{1,r}(J; X, Y) \hookrightarrow \mathrm{C}\big(\overline{J}; (X, Y)_{1/r',r}\big) \tag{1.28}$$

and

$$\mathbb{W}^{1,r}(J; X, Y) \hookrightarrow \mathrm{C}^{\alpha}\big(J; (X, Y)_{\theta,1}\big) \tag{1.29}$$

for $0 < \theta < 1/r'$ and $0 < \alpha < 1/r' - \theta$.

One may generalize these embeddings to the spaces $\mathbb{W}^{1,r}_s(J; X, Y)$ "for free", since the embedding (1.28) in fact follows from an equivalent definition of the real interpolation spaces (see [146, Ch. 1.8.3]) and the proof of (1.29) starting from (1.28) does not rely on $s = r$ in any way. Hence, we obtain the following embeddings for spaces with mixed integrability:

Lemma 1.4.4. *Let X, Y be two Banach spaces with $Y \hookrightarrow X$ and let $1 < r, s < \infty$. Set $\xi \coloneqq s\,(1 + \frac{1}{s} - \frac{1}{r})$. Then the following embeddings are true:*

$$\mathbb{W}^{1,r}_s(J; X, Y) \hookrightarrow \mathrm{C}\big(\overline{J}; (X, Y)_{1/\xi',\xi}\big) \tag{1.30}$$

and

$$\mathbb{W}^{1,r}_s(J; X, Y) \hookrightarrow \mathrm{C}^{\alpha}\big(J; (X, Y)_{\theta,1}\big) \tag{1.31}$$

for $0 < \theta < 1/\xi' = \frac{1}{r'}\big(1 + \frac{1}{s} - \frac{1}{r}\big)^{-1}$ and $0 < \alpha < 1/r' - \theta\big(1 + \frac{1}{s} - \frac{1}{r}\big)$.

Proposition 1.4.3 will prove to be very valuable in the following considerations, in particular the combination of embedding (1.29) and the Arzelà-Ascoli Theorem 1.2.5, cf. Corollary 1.2.6. We will return to Lemma 1.4.4

later in Chapter 3.

A first immediate consequence of (1.30) not involving compactness is that the *point evaluation* $\delta_\tau \colon f \mapsto f(\tau)$ for τ from the domain of definition of f satisfies

$$\delta_\tau \in \mathscr{L}\big(\mathbb{W}^{1,r}_s(J;X,Y);(X,Y)_{1/\xi',\xi}\big) \quad \text{for all } \tau \in \overline{J} \tag{1.32}$$

for $\xi = s\,(1+\frac{1}{s}-\frac{1}{r})$. This shows that

$$\mathbb{W}^{1,r}_0(J;X,Y) := \Big\{ f \in \mathbb{W}^{1,r}(J;X,Y) \colon \delta_{T_0} f = 0 \Big\}$$

is a closed subspace of $\mathbb{W}^{1,r}(J;X,Y)$ and thus also a Banach space. Moreover, it allows to formulate the following integration by parts formula:

Theorem 1.4.5. *Let X, Y be Banach spaces with $Y \hookrightarrow_{\mathrm{d}} X$, let $1 < r, s < \infty$, and assume $u \in \mathbb{W}^{1,r}_s(J;X,Y)$ and $v \in \mathbb{W}^{1,s'}_{r'}(J;Y',X')$. Then for every $t \in \overline{J}$ we have*

$$\int_{T_0}^t \big\langle u'(s), v(s)\big\rangle_{X,X'} + \big\langle u(s), v'(s)\big\rangle_{Y,Y'} \,\mathrm{d}s = \big\langle u(t), v(t)\big\rangle_\xi - \big\langle u(T_0), v(T_0)\big\rangle_\xi, \tag{1.33}$$

where $\langle\cdot,\cdot\rangle_\xi$ denotes the duality pairing between $(X,Y)_{1/\xi',\xi}$ and its dual space $(X,Y)'_{1/\xi',\xi}$ with $\xi = s(1+\frac{1}{s}-\frac{1}{r})$.

Proof. We show (1.33) by employing the fundamental theorem of calculus. From Lemma 1.4.4 we know that

$$\mathbb{W}^{1,r}_s(J;X,Y) \hookrightarrow \mathrm{C}\big(\overline{J};(X,Y)_{1/\xi',\xi}\big)$$

and

$$\mathbb{W}^{1,s'}_{r'}(Y',X') \hookrightarrow \mathrm{C}\big(\overline{J};(Y',X')_{1/\xi',\xi'}\big),$$

as well as

$$(Y',X')_{1/\xi',\xi'} = (X',Y')_{1/\xi,\xi'} \doteq (X,Y)'_{1/\xi',\xi}$$

due to Lemmata 1.1.8 and 1.1.14. Thus,

$$(u,v) \mapsto \Big[t \mapsto \big\langle u(t), v(t)\big\rangle_\xi - \big\langle u(T_0), v(T_0)\big\rangle_\xi\Big] \tag{1.34}$$

is continuous as a mapping from $\mathbb{W}^{1,r}_s(J;X,Y) \times \mathbb{W}^{1,s'}_{r'}(J;Y',X')$ to $\mathrm{C}(\overline{J})$. Clearly,

$$(u,v) \mapsto \langle u'(t), v(t)\rangle_{X,X'} + \langle u(t), v'(t)\rangle_{Y,Y'}$$

maps $\mathbb{W}^{1,r}_s(J;X,Y) \times \mathbb{W}^{1,s'}_{r'}(J;Y',X')$ continuously into $L^1(J)$, hence

$$(u,v) \mapsto \Big[t \mapsto \int_{T_0}^t \big\langle u'(s), v(s)\big\rangle_{X,X'} + \big\langle u(s), v'(s)\big\rangle_{Y,Y'} \,\mathrm{d}s\Big] \tag{1.35}$$

is also a continuous mapping from $\mathbb{W}^{1,r}_s(J;X,Y) \times \mathbb{W}^{1,s'}_{r'}(J;Y',X')$ to $\mathrm{C}(\overline{J})$.

Due to the *dense* embeddings

$$Y \overset{\mathrm{d}}{\hookrightarrow} (X,Y)_{\zeta,1} \overset{\mathrm{d}}{\hookrightarrow} X \quad \text{and} \quad Y' \overset{\mathrm{d}}{\hookrightarrow} (Y',X')_{\zeta,1} \overset{\mathrm{d}}{\hookrightarrow} X' \tag{1.36}$$

for all $0 < \zeta < 1$, the dual pairing $\langle \mathfrak{u}(t), \mathfrak{v}(t)\rangle_\xi$ coincides with $\langle \mathfrak{u}(t), \mathfrak{v}(t)\rangle_{X,X'}$ and $\langle \mathfrak{u}(t), \mathfrak{v}(t)\rangle_{Y,Y'}$ if $\mathfrak{u}(t) \in Y$ and $\mathfrak{v}(t) \in X'$, cf. Proposition 1.0.2. Thus, we calculate for $\mathfrak{u} \in \mathrm{C}^1(\overline{J}) \otimes Y$ and $\mathfrak{v} \in \mathrm{C}^1(\overline{J}) \otimes X'$ that

$$\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \mathfrak{u}(t), \mathfrak{v}(t)\big\rangle_\xi = \big\langle \mathfrak{u}'(t), \mathfrak{v}(t)\big\rangle_{X,X'} + \big\langle \mathfrak{u}(t), \mathfrak{v}'(t)\big\rangle_{Y,Y'}$$

for all $t \in \overline{J}$, and hence, by the fundamental theorem of calculus,

$$\int_{T_0}^t \big\langle \mathfrak{u}'(s), \mathfrak{v}(s)\big\rangle_{X,X'} + \big\langle \mathfrak{u}(s), \mathfrak{v}'(s)\big\rangle_{Y,Y'} \,\mathrm{d}s = \big\langle \mathfrak{u}(t), \mathfrak{v}(t)\big\rangle_\xi - \big\langle \mathfrak{u}(T_0), \mathfrak{v}(T_0)\big\rangle_\xi$$

for all $\mathfrak{u} \in \mathrm{C}^1(\overline{J}) \otimes Y$ and $\mathfrak{v} \in \mathrm{C}^1(\overline{J}) \otimes X'$. But $\mathrm{C}^1(\overline{J}) \otimes Y$ and $\mathrm{C}^1(\overline{J}) \otimes X'$ are dense in $\mathrm{C}^1(\overline{J};Y)$ and $\mathrm{C}^1(\overline{J};X')$, which in turn are dense in $\mathbb{W}^{1,r}_s(J;X,Y)$ and $\mathbb{W}^{1,s'}_{r'}(J;Y',X')$, respectively (cf. Lemma 2.1.17 below and [3, Ch. V, Thm. 2.4.6]). Moreover, we already had noted that both sides of the preceding equation, seen as continuous functions on $\overline{J}$, also give rise to continuous functions in $(u,v) \in \mathbb{W}^{1,r}_{s}(J;X,Y) \times \mathbb{W}^{1,s'}_{r'}(J;Y',X')$ with values in $\mathrm{C}(\overline{J})$, cf. (1.34) and (1.35). This implies (1.33) by density. $\square$

Remark 1.4.6. Let $Y \hookrightarrow_{\mathrm{d}} X$. Due to the *density* of this embedding and thus also of the interpolation embeddings as seen in (1.36), the notation of the dual pairings with varying subscripts indicating the spaces in the foregoing theorem is in fact more or less arbitrary, since all dual pairings must agree, as already seen in Proposition 1.0.2 (see also the considerations in [3, Ch. V.1]). We use it mostly to indicate the spaces to which the involved functions belong.

After these preparations, let us introduce the fundamental notion of maximal parabolic regularity. Note that, given an operator $B\colon J \to \mathscr{L}(Y;X)$ and a function $u\colon J \to Y$, we freely use the identification $(Bu)(\cdot) := B(\cdot)u(\cdot)\colon J \to X$ in the following. In this sense, we also identify such an operator B with a linear mapping from a function space on J with values in Y to a function space on J with values in X.

Definition 1.4.7 (Maximal parabolic regularity)**.** Let $J = (T_0, T_1)$ and $1 < r < \infty$. Let A be a closed operator on X with dense domain Y. We say that A satisfies *maximal parabolic* L^r*-regularity on* X *over* J if for every $f \in \mathrm{L}^r(J;X)$ the problem

$$u'(t) + Au(t) = f(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = 0 \tag{1.37}$$

has a unique solution $u \in \mathbb{W}^{1,r}_0(J;X,Y)$.

Remark 1.4.8. Staring at the definition for a minute, it is clear that the name “maximal regularity” stems from the property that both terms u' and Au on the left-hand side in (1.37) exhibit the same regularity as f does, i.e., the solution u has exactly as much regularity as it *can* have from the assumption on f.

Let us first collect some well-known properties of operators satisfying maximal parabolic regularity in the following lemma. For the proofs we refer to the survey article of Dore [55], cf. also [127, 140].

Lemma 1.4.9. *Let $1 < r < \infty$ and let A be a closed operator on X with dense domain satisfying maximal parabolic L^r regularity over J. Then the following assertions are true:*

(i) $-A$ generates an analytic semigroup on X.

(ii) A satisfies maximal parabolic L^s-regularity for any $1 < s < \infty$.

(iii) A satisfies maximal parabolic L^r-regularity over (T_0, T) for any $T_0 < T < \infty$.

In view of Lemma 1.4.9, we usually only say or assume that an operator A *satisfies maximal parabolic regularity* (on X, if not clear from the context) and drop the explicit references to the interval J and the integrability order r.

Remark 1.4.10. By Lemma 1.4.9, it is a necessary condition for A to satisfy maximal parabolic regularity on X that $-A$ generates an analytic semigroup on X. de Simon has shown that this is also sufficient if X is a Hilbert space [49]. Even more, it is known from the work of Kalton and Lancien that a Banach space X on which every negative generator of an analytic semigroup satisfies maximal parabolic regularity must essentially already be a Hilbert space [95]. We refer to [3, 111] for a comprehensive treatment of analytic semigroups.

From the open mapping theorem, we immediately obtain the following, very useful reformulation of maximal parabolic regularity in terms of continuous invertibility of the "total" differential operator $\partial + A$.

Lemma 1.4.11. *Let $1 < r < \infty$ and let A be a closed operator on X with dense domain Y. Then A satisfies maximal parabolic L^r-regularity if and only if*

$$\partial + A \in \mathscr{L}_{\mathrm{iso}}\big(\mathbb{W}_0^{1,r}(J; X, Y), \mathrm{L}^r(J; X)\big),$$

where ∂ denotes the distributional (time) derivative.

Now let us consider the equation

$$u'(t) + Au(t) = f(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0 \tag{1.38}$$

for $f \in \mathrm{L}^r(J; X)$ with nonzero initial value u_0 from some function space. We already know that $u(T_0) = \delta_{T_0} u \in (X, Y)_{1/r',r}$ if $u \in \mathbb{W}^{1,r}(J; X, Y)$, cf. (1.32). The following lemma shows that δ_{T_0} is in fact surjective from $\mathbb{W}^{1,r}(J; X, Y)$ *onto* $(X, Y)_{1/r',r}$.

Lemma 1.4.12 ([3, Prop. III.4.10.3])**.** *Let B be a closed operator on X with dense domain Y and assume that B is the generator of an analytic semigroup T on X. Then $E\colon x \mapsto T(\cdot - T_0)x$ is a coretraction from $(X, Y)_{1/r',r}$ to $\mathbb{W}^{1,r}(J; X, Y)$ with δ_{T_0} being the corresponding retraction.*

Indeed, maximal parabolic regularity of A already implies unique solvability of (1.38) within the $\mathbb{W}^{1,r}(J; X, Y)$-class for $u_0 \in (X, Y)_{1/r',r}$, as we see in the next lemma. The proof uses Lemma 1.4.12 for $B = -A$ which is a valid choice by Lemma 1.4.9. In view of Lemma 1.4.12, this is again the optimal regularity for the given data, cf. Remark 1.4.8.

Lemma 1.4.13 ([9, Prop. 2.1])**.** *Let $1 < r < \infty$ and let A be a closed operator on X with dense domain Y. Then the following assertions are equivalent:*

(i) A satisfies maximal parabolic L^r-regularity.

(ii) Eq. (1.38) *has a unique solution $u \in \mathbb{W}^{1,r}(J; X, Y)$ for every $u_0 \in (X, Y)_{1/r',r}$ and every $f \in \mathrm{L}^r(J; X)$.*

(iii) $(\partial + A, \delta_{T_0}) \in \mathscr{L}_{\mathrm{iso}}\big(\mathbb{W}^{1,r}(J; X, Y), \mathrm{L}^r(J; X) \times (X, Y)_{1/r',r}\big)$.

It is exactly the reformulation in terms of continuous invertibility of the total differential operator $\partial + A$ that makes maximal parabolic regularity particularly resistant against perturbations. We will return to this topic later.

Let us now turn to *nonautonomous* maximal parabolic regularity. As already mentioned in the introduction, this is a most natural topic to investigate when working with quasilinear evolution equations or optimal control of such equations. So, let $(A(t))_{t\in J}$ be a family of closed operators on a Banach space X. There are two rather strictly divided cases which need to be distinguished: Do the domains $\operatorname{dom} A(t)$ vary with t or not? We concentrate on the case where they are constant and define as follows:

Definition 1.4.14 (Nonautonomous maximal parabolic regularity)**.** Let $1 < r < \infty$, let X, Y be two Banach spaces such that $Y \hookrightarrow_{\mathrm{d}} X$ and assume that

$$A \in \mathrm{L}^1(J; \mathscr{L}(Y;X)) \cap \mathscr{L}(\mathbb{W}^{1,r}(J;X,Y); \mathrm{L}^r(J;X)).$$

We say that A satisfies *nonautonomous maximal parabolic* L^r*-regularity on* X *over* J if for every $f \in \mathrm{L}^r(J;X)$ the problem

$$u'(t) + A(t)u(t) = f(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0 \tag{1.39}$$

has a unique solution $u \in \mathbb{W}^{1,r}_0(J;X,Y)$ whenever $u_0 = 0$.

Definition 1.4.14 is nearly analogous to the one for the autonomous case, Definition 1.4.7. Indeed, nonautonomous maximal parabolic regularity for constant domains shares many properties with the autonomous case, in particular the reformulation via continuous invertibility of the total differential $\partial + A$ as in Lemma 1.4.11:

Lemma 1.4.15. *Let* $1 < r < \infty$ *and let* A *be as in Definition 1.4.14. Then* A *satisfies nonautonomous maximal parabolic* L^r*-regularity on* X *over* J *if and only if*

$$\partial + A \in \mathscr{L}_{\mathrm{iso}}(\mathbb{W}^{1,r}_0(J;X,Y), \mathrm{L}^r(J;X)),$$

where ∂ *denotes the distributional (time) derivative.*

The properties of operators satisfying autonomous maximal parabolic regularity as in Lemma 1.4.9 are not necessarily satisfied in the case of nonautonomous maximal parabolic regularity any more. Both independence of the integrability index r and the property of being the negative generator of an analytic semigroup for each $t \in J$ do *not* follow automatically from the maximal regularity property. This in particular means that we cannot use Lemma 1.4.12 directly to obtain nonautonomous maximal parabolic regularity for nonzero initial values as done in Lemma 1.4.13 without assuming that there *exists* a generator of an analytic semigroup on X with domain Y at this stage. We will see that both are available under the sufficient conditions we use for nonautonomous maximal parabolic regularity. Indeed, the first result is as follows:

Lemma 1.4.16. *Let $1 < r < \infty$ and let A be as in Definition 1.4.14. Assume further that $A \in \mathrm{C}(\overline{J};\mathscr{L}(Y;X))$. Then the following assertions are equivalent:*

(i) A satisfies nonautonomous maximal parabolic L^r-regularity on X over J.

(ii) Eq. (1.39) *has a unique solution $u \in \mathbb{W}^{1,r}(J;X,Y)$ for every $u_0 \in (X,Y)_{1/r',r}$ and every $f \in \mathrm{L}^r(J;X)$.*

(iii) $(\partial + A, \delta_{T_0}) \in \mathscr{L}_{\mathrm{iso}}(\mathbb{W}^{1,r}(J;X,Y), \mathrm{L}^r(J;X) \times (X,Y)_{1/r',r})$.

If one and thus all of the equivalent assertions are true, then $A(\tau)$ already satisfies autonomous *maximal parabolic regularity on X for every $\tau \in \overline{J}$.*

Proof. The conditions (ii) and (iii) are clearly equivalent, cf. also Lemma 1.4.13, and of course imply (i), even without the continuity assumption. The other way around, it is known from [128, Rem. 2.6], [7, Prop. 7.1] that if A is continuous as a mapping into $\mathscr{L}(Y;X)$ and satisfies maximal nonautonomous parabolic L^r-regularity on X over J, then $A(\tau)$ already satisfies *autonomous* maximal parabolic regularity for all $\tau \in \overline{J}$. But then each operator $-A(\tau)$ generates an analytic semigroup on X with domain Y (see Lemma 1.4.9), and the proof of Lemma 1.4.13 applies to the nonautonomous setting word for word, thanks to Lemma 1.4.12. □

It even turns out that $A(\tau)$ satisfying maximal parabolic regularity for every time $\tau \in \overline{J}$ is not only necessary for $A \in \mathrm{C}(\overline{J};\mathscr{L}(Y;X))$ to satisfy nonautonomous maximal parabolic regularity, but even sufficient as proven by Prüss and Schnaubelt [128] and Amann [7]:

Theorem 1.4.17 ([128, Thm. 2.5], [7, Thm. 7.1])**.** *Let $1 < r < \infty$ and let A be as in Definition 1.4.14. Assume further that $A \in \mathrm{C}(\overline{J};\mathscr{L}(Y;X))$. Then the following conditions are equivalent:*

(i) $A(\tau)$ satisfies maximal parabolic regularity on X for every $\tau \in \overline{J}$.

(ii) The problem

$$u'(t) + A(t)u(t) = f(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0$$

admits a unique solution $u \in \mathbb{W}^{1,r}(J;X,Y)$ for every $f \in \mathrm{L}^r(J;X)$ and every $u_0 \in (X,Y)_{1/r',r}$ for $1 < r < \infty$. In particular, A satisfies nonautonomous maximal parabolic regularity on X over J.

Remark 1.4.18.

(i) Since the assumption that each $A(\tau)$ satisfies maximal parabolic regularity for each $\tau \in \overline{J}$ in Theorem 1.4.17 is independent of $1 < r < \infty$ as noted in Lemma 1.4.9, we also obtain nonautonomous maximal parabolic L^r-regularity for all such r in Theorem 1.4.17.

(ii) We have already used that each operator $-A(\tau)$ in the setting of Theorem 1.4.17 is the generator of an analytic semigroup on X with domain Y, per Lemma 1.4.9.

(iii) Under the assumptions of Theorem 1.4.17, A clearly also satisfies nonautonomous parabolic L^r-regularity on X over every subinterval $J^* \subseteq J$. It is thus acceptable to skip the reference to J in this context.

We finally collect perturbation results for nonautonomous maximal parabolic regularity from [7, Cor. 5.1] and [127, Cor. 3.4]. The first, a surprisingly elementary application of perturbation of invertibility, is as follows:

Lemma 1.4.19 ([7, Cor. 5.1]). *Let $1 < r < \infty$, let A be as in Definition 1.4.14, and let numbers $0 < \theta < 1$ and $1 < \varrho \leq \infty$ be given such that $0 \leq 1/\varrho < \min(1-\theta, 1/r)$. Assume that A satisfies nonautonomous maximal parabolic L^r-regularity on X over J and that $B \in \mathrm{L}^\varrho(J; \mathscr{L}((X,Y)_{\theta,\infty}; X))$. Then $A + B$ satisfies nonautonomous maximal parabolic L^r-regularity on X over J.*

The second perturbation result is in fact one of the cornerstones when considering the linearization of quasilinear parabolic equations in divergence form. It is stated in [127] without an explicit proof as a special case of a much more potent theorem whose proof is challenging (see Theorem 2.2.7). We thus give a proof for the perturbation result with some more details. Let us remark that this result is not present in either [7, Thm. 7.1] or [128, Thm. 3.1], although the authors there consider a similar setting. The stated case is a limit point of the admissible combination of integrability in time and spatial regularity as in Lemma 1.4.19, cf. Remark 1.4.22 below.

Lemma 1.4.20. *Let $1 < r < \infty$ and let A be as in Definition 1.4.14. Assume further that $A \in \mathrm{C}(\overline{J}; \mathscr{L}(Y;X))$ such that $A(\tau)$ satisfies maximal parabolic regularity for all $\tau \in \overline{J}$. Let moreover $B \in \mathrm{L}^r(J; \mathscr{L}((X,Y)_{1/r',r}; X))$. Then the problem*

$$u'(t) + A(t)u(t) + B(t)u(t) = f(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0$$

admits a unique solution $u \in \mathbb{W}^{1,r}(J; X, Y)$ for every $f \in \mathrm{L}^r(J; X)$ and every $u_0 \in (X,Y)_{1/r',r}$. In particular, $A + B$ satisfies nonautonomous maximal parabolic L^r-regularity on X.

Proof. We construct an appropriate setting for a fixed point argument. This will require to perform a continuation procedure for local solutions, for which we first collect some uniform bounds.

By the assumption that all $A(\tau)$, $\tau \in \overline{J}$, satisfy maximal parabolic regularity, Lemma 1.4.11 implies that for every $\tau \in \overline{J}$ there exists a constant

$C_J(\tau) > 0$ such that the unique solutions $w \in \mathbb{W}_0^{1,r}(J;X,Y)$ of

$$w'(t) + A(\tau)w(t) = g(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad w(T_0) = 0$$

satisfy $\|w\|_{\mathbb{W}_0^{1,r}(J;X,Y)} \leq C_J(\tau)\|g\|_{\mathrm{L}^r(J;X)}$ for all $g \in \mathrm{L}^r(J;X)$, where $C_J(\tau)$ is the norm of the "solution operator" $(\partial + A(\tau))^{-1}$. Moreover, from the continuity of A and the inversion mapping we infer that

$$\left[\tau \mapsto (\partial + A(\tau))^{-1}\right] \in \mathrm{C}(\overline{J}; \mathscr{L}(\mathrm{L}^r(J;X); \mathbb{W}_0^{1,r}(J;X,Y))).$$

This implies that

$$C_J := \max_{\tau\in\overline{J}} C_J(\tau) = \max_{\tau\in\overline{J}} \|(\partial + A(\tau))^{-1}\|_{\mathscr{L}\left(\mathrm{L}^r(J;X);\mathbb{W}_0^{1,r}(J;X,Y)\right)} < \infty.$$

In the following, let C_T be the embedding constant of the embedding $\mathbb{W}_0^{1,r}(J;X,Y) \hookrightarrow \mathrm{C}(\overline{J};(X,Y)_{1/r',r})$, cf. Proposition 1.4.3.

Next, we choose "stepsizes" for the continuation procedure. First, we use the uniform continuity of A on $\overline{J}$ to choose a number $\delta > 0$ such that

$$\sup_{\substack{t,s\in\overline{J} \\ |t-s|<\delta}} \|A(t) - A(s)\|_{\mathscr{L}(Y;X)} \leq \frac{1}{4C_J}.$$

Second, let us abbreviate $\phi(t) := \|B(t)\|_{\mathscr{L}((X,Y)_{1/r',r};X)}$. Then $\phi \in \mathrm{L}^r(J)$, and we may pick a number $\varepsilon \leq \delta$ such that

$$\|\phi\|_{\mathrm{L}^r(t_i,t_{i+1};X)} \leq \frac{1}{2C_JC_T},$$

where

$$t_i := (T_0 + i\varepsilon) \wedge T_1, \quad i = 0,\dots,n := \frac{\lceil T_1 - T_0\rceil}{\varepsilon},$$

cf. [152, Thm. 8.20].

Now consider the intervals $J^i = (t_i, t_{i+1}) \subset J$. Let $v \in \mathbb{W}^{1,r}(J^i;X,Y)$ be given and consider the problem family

$$w'(t) + A(t_i)w(t) = f(t) - \big(B(t) + A(t) - A(t_i)\big)v(t) \qquad \text{in } X \quad \text{for a.a. } t \in J^i, \quad w(t_i) = u_i \tag{1.40}$$

for $u_i \in (X,Y)_{1/r',r}$ and $i = 0, \dots, n$. Due to $\mathbb{W}^{1,r}(J^i; X, Y) \hookrightarrow \mathrm{C}(\overline{J}^i; (X,Y)_{1/r',r})$ as in Proposition 1.4.3, we have $Bv \in \mathrm{L}^r(J^i; X)$. Hence, by Theorem 1.4.17, there exists a unique solution $u \in \mathbb{W}^{1,r}(J^i; X, Y)$ of (1.40). Define the (affine-linear) mapping $\mathcal{T}_i \colon v \mapsto u$ such that u is the solution to (1.40). We show that $\mathcal{T}_i$ admits a unique fixed point in $\mathbb{W}^{1,r}(J^i; X, Y)$. Observe that $\mathcal{T}_i v - \mathcal{T}_i \bar{v}$ is the $\mathbb{W}^{1,r}_0(J^i; X, Y)$-solution of

$$w'(t) + A(t_i)w(t) = \big(B(t) + A(t) - A(t_i)\big)\big(\bar{v}(t) - v(t)\big) \qquad \text{in } X \quad \text{for a.a. } t \in J^i, \quad w(t_i) = 0.$$

Again Lemma 1.4.11 yields a constant $C_i > 0$ such that

$$\begin{aligned} \|\mathcal{T} v - \mathcal{T} \bar{v}\|_{\mathbb{W}^{1,r}_0(J^i;X,Y)} &\le C_i \big\|\big(B(\cdot) + A(\cdot) - A(t_i)\big)\big(\bar{v}(\cdot) - v(\cdot)\big)\big\|_{\mathrm{L}^r(J^i;X)} \\ &\le C_i \Big(\|\phi\|_{\mathrm{L}^r(J^i)} \|\bar{v} - v\|_{\mathrm{C}(\overline{J}^i;(X,Y)_{1/r',r})} \\ &\qquad + \|A(\cdot) - A(t_i)\|_{\mathrm{C}(J^i;\mathscr{L}(Y;X))} \|\bar{v} - v\|_{\mathbb{W}^{1,r}(J^i;X,Y)} \Big). \end{aligned}$$

Now it remains to observe that C_i may not be greater than $C_J(t_i)$, hence in particular $C_i \le C_J$, and that the embedding constant of $\mathbb{W}^{1,r}_0(J^i; X, Y) \hookrightarrow \mathrm{C}(\overline{J}^i; (X,Y)_{1/r',r})$ is not larger than C_T, both by applying translations and extensions by zero (note that we work with *autonomous* operators $A(t_i)$ here!). But then by inserting the properties of ε we obtain

$$\begin{aligned} \|\mathcal{T} v - \mathcal{T} \bar{v}\|_{\mathbb{W}^{1,r}_0(J^i;X,Y)} &\le C_J \left(C_T \frac{1}{4 C_J C_T} + \frac{1}{4 C_J} \right) \|v - \bar{v}\|_{\mathbb{W}^{1,r}(J^i;X,Y)} \\ &= \frac{1}{2} \|v - \bar{v}\|_{\mathbb{W}^{1,r}(J^i;X,Y)}. \end{aligned}$$

Now Banach's fixed point theorem shows that there exists a unique fixed point of $\mathcal{T}$ on each interval J^i which solves

$$u'(t) + A(t)u(t) + B(t)u(t) = f(t) \qquad \text{in } X \quad \text{for a.a. } t \in J^i, \quad u(t_i) = u_i. \tag{1.41}$$

It remains to construct the solution on the whole interval J. We start at $i = 0$, obtaining a solution u^0 of (1.41) on $J^0 = (T_0, t^1)$ such that $u^0 \in \mathbb{W}^{1,r}(J^0; X, Y) \hookrightarrow \mathrm{C}(\overline{J}^0; (X,Y)_{1/r',r})$. Now we set $u_1 \coloneqq u^0(t_1) \in (X,Y)_{1/r',r}$ and proceed iteratively in i in the same fashion. "Gluing" together the such produced functions $u^0, \ldots, u^n$ yields a function $u \in \mathbb{W}^{1,r}(J; X, Y)$ (see [10, Lem. 7.1]) which is a solution to the equation under consideration, and it is even unique because all the u^i were unique. □

We combine the two preceding lemmata to a final statement.

Corollary 1.4.21. *Let $1 < r < \infty$, let A be as in Definition 1.4.14. Assume that $A \in \mathrm{C}(\overline{J}; \mathscr{L}(Y; X))$ satisfies nonautonomous maximal parabolic regularity over J. Let moreover $B \in \mathrm{L}^r(J; \mathscr{L}((X,Y)_{1/r',r}; X))$ and $C \in \mathrm{L}^\varrho(J; \mathscr{L}((X,Y)_{\theta,\infty}; X))$ for $1/r' < \theta < 1$ and $r < \varrho \leq \infty$. Then $A+B+C$ satisfies nonautonomous maximal parabolic L^r-regularity on X over J.*

Proof. The result follows form first applying Lemma 1.4.20 for the perturbation by B to the operator A and afterwards Lemma 1.4.19 for the perturbation by C to the operator $A+B$. Comparing with Lemma 1.4.19, we have omitted the case where $0 < \theta \leq 1/r'$. For these θ, we have $(X,Y)_{1/r',r} \hookrightarrow (X,Y)_{\theta,\infty}$ by Lemma 1.1.8 and $r < \varrho \leq \infty$ due to $\min(1-\theta, 1/r) = 1/r$. But then

$$\mathrm{L}^\varrho(J; \mathscr{L}((X,Y)_{\theta,\infty}; X)) \hookrightarrow \mathrm{L}^r(J; \mathscr{L}((X,Y)_{1/r',r}; X))$$

and this case is already covered by Lemma 1.4.20. Hence Lemma 1.4.19 only brings the case $1/r' < \theta < 1$ to the table. □

Remark 1.4.22.

(i) Note that we only need the continuity property of A in the foregoing Corollary 1.4.21 for Lemma 1.4.20, whereas Lemma 1.4.19 works for every operator A satisfying nonautonomous maximal parabolic L^r-regularity on X over J. Thus, for the proof of the combined result, we first apply Lemma 1.4.20 to A and afterwards Lemma 1.4.19 to the operator $A+B$ which need not be continuous in time any more.

(ii) It is imperative to take a closer look at the relation between time integrability and spatial regularity in Corollary 1.4.21 or Lemma 1.4.19, respectively. Assume that $1/r' < \theta < 1$. Then $\min(1-\theta, 1/r) = 1-\theta$ such that $(X,Y)_{\theta,\infty} \hookrightarrow (X,Y)_{1/r',r}$ and the condition on ϱ becomes $1/(1-\theta) < \varrho \leq \infty$. This means that the operators $C(t)$ for $t \in J$ need only be continuous on a *stronger* space compared to $(X,Y)_{1/r',r}$. The informally calculated limit points are

$$\theta \nearrow 1: \qquad \mathrm{L}^\varrho(J;(X,Y)_{\theta,\infty};X) \text{ “}\searrow\text{” } \mathrm{L}^\infty(J;\mathscr{L}(Y;X))$$

where “$\searrow$” is meant in a descending set inclusion way, and

$$\theta \searrow 1/r': \quad \mathrm{L}^\varrho(J;\mathscr{L}((X,Y)_{\theta,\infty};X)) \text{ “}\nearrow\text{” } \mathrm{L}^r(J;\mathscr{L}((X,Y)_{1/r',\infty},X))$$

with “$\nearrow$” analogously in an ascending set inclusion sense. The second limit is essentially the setting of Lemma 1.4.20, whereas the first one is also obtained as a limiting case in a rigorous setting, cf. [7, Thm. 7.1] and [18, Thm. 2.11].

1.5 The divergence-gradient operator

The main protagonist in the class of parabolic evolution equations under consideration is the general divergence-gradient operator $-\nabla\cdot\rho\nabla$ for some coefficient function ρ as an operator on spaces of type $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for a domain $\Lambda \subset \mathbb{R}^d$ and a closed subset of its boundary $\Xi \subseteq \partial\Lambda$. It seems thus worthwhile to establish its properties in a rather general context assembled in one place. We will give more specific assumptions on the regularity of the domain Λ in this chapter in order to make sense of the assertions and operators, such as boundary forms. Thereby, we generally give the results for $\Lambda \cup \Xi$ Lipschitz around $\partial\Lambda \setminus \Xi$ and indicate when they are available for even more general geometries.

To have a concise notion for the classes of coefficient functions ρ at hand, we first make the following definition.

Definition 1.5.1. By $\mathbb{M}_d$ we denote the set of *real* $(d \times d)$-matrices, equipped with the operator norm $\|\cdot\|_{\mathbb{M}_d}$ generated by the Euclidean norm on $\mathbb{R}^d$. Moreover, we call $\mathbb{S}_d$ the subset of $\mathbb{M}_d$ consisting of *symmetric* matrices. Given $\rho_\bullet, \rho^\bullet \in \mathbb{R}^+$ with $\rho_\bullet < \rho^\bullet$, we further set

$$\mathbb{M}_d(\rho_\bullet) := \Big\{\rho \in \mathbb{M}_d \colon \rho_\bullet \|v\|_2^2 \leq v^\top \rho v \text{ for all } v \in \mathbb{R}^d\Big\}.$$

and

$$\mathbb{M}_d(\rho_\bullet, \rho^\bullet) := \Big\{\rho \in \mathbb{M}_d \colon \rho_\bullet \|v\|_2^2 \leq v^\top \rho v \leq \rho^\bullet \|v\|_2^2 \text{ for all } v \in \mathbb{R}^d\Big\},$$

as well as $\mathbb{S}_d(\rho_\bullet) := \mathbb{M}_d(\rho_\bullet) \cap \mathbb{S}_d$ and $\mathbb{S}_d(\rho_\bullet, \rho^\bullet) := \mathbb{M}_d(\rho_\bullet, \rho^\bullet) \cap \mathbb{S}_d$. Note that $\|\rho\|_{\mathbb{M}_d} \leq \rho^\bullet$ if $\rho \in \mathbb{S}_d(\rho_\bullet, \rho^\bullet)$ due to the spectral theorem.

We agree that, from now on, using the expressions $\mathbb{M}_d(\rho_\bullet)$ or $\mathbb{M}_d(\rho_\bullet, \rho^\bullet)$ implies the relations $0 < \rho_\bullet < \rho^\bullet$ without further explicitly noting so.

Remark 1.5.2. Since there are different ways to set the conditions for the coefficient functions ρ in the context of complex functions, we briefly mention that if $\rho \in \mathbb{M}_d(\kappa_\bullet)$, then ρ also satisfies

$$2\kappa_\bullet \|v\|_2^2 \leq \operatorname{Re}\left(v^{\mathrm{H}} \rho v\right) \quad \text{for all } v \in \mathbb{C}^d.$$

Moreover, if $\rho \in \mathbb{S}_d(\kappa_\bullet)$, we obtain $\kappa_\bullet \|v\|_2^2 \leq v^{\mathrm{H}} \rho v \in \mathbb{R}$ for all $v \in \mathbb{C}^d$.

Hence it is no limitation to restrict ourselves to real vectors v in Definition 1.5.1.

We now define the divergence-gradient operator and explore various properties.

Definition 1.5.3 (Divergence-gradient operator)**.** Let $\Lambda \subseteq \mathbb{R}^d$ be a domain, let $\Xi \subseteq \partial\Lambda$ be a closed subset of its boundary, allowed to be empty, and assume that $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d)$. We define the operator $-\nabla\cdot\rho\nabla \colon \mathrm{W}^{1,2}_\Xi(\Lambda) \to \mathrm{W}^{-1,2}_\Xi(\Lambda)$ by

$$\langle -\nabla\cdot\rho\nabla\psi, \xi\rangle \coloneqq \int_\Lambda \rho\nabla\psi\cdot\nabla\overline{\xi}\,\mathrm{d}x \quad \text{for } \psi,\xi \in \mathrm{W}^{1,2}_\Xi(\Lambda),$$

where $\langle\cdot,\cdot\rangle$ stands for the dual pairing between $\mathrm{W}^{-1,2}_\Xi(\Lambda)$ and $\mathrm{W}^{1,2}_\Xi(\Lambda)$, extending the $\mathrm{L}^2(\Lambda)$ scalar product, cf. [23, Ch. 1.§1]. We keep the notation $-\nabla\cdot\rho\nabla$ for the operator obtained by the maximal corestriction of $-\nabla\cdot\rho\nabla$ to $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ for $p > 2$ and denote its domain by $\mathcal{D}_p(\rho)$ which we equip with the graph norm.

It is often of interest to consider $-\nabla\cdot\rho\nabla$ as a closed operator on $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ with domain $\mathcal{D}_p(\rho)$ for $p \geq 2$. We will do so freely.

Remark 1.5.4. The estimate

$$\begin{aligned}\|-\nabla\cdot\rho\nabla\psi\|_{\mathrm{W}^{-1,p}_\Xi(\Lambda)} &= \sup_{\|\xi\|_{\mathrm{W}^{1,p'}_\Xi(\Lambda)}=1} \left|\int_\Lambda \rho\nabla\psi\cdot\nabla\overline{\xi}\,\mathrm{d}x\right| \\ &\leq \|\rho\|_{\mathrm{L}^\infty(\Lambda;\mathbb{M}_d)}\|\psi\|_{\mathrm{W}^{1,p}_\Xi(\Lambda)}\end{aligned}$$

shows that $\mathrm{W}^{1,p}_\Xi(\Lambda) \hookrightarrow \mathcal{D}_p(\rho)$ for all $p \geq 2$ and all $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d)$ and that

$$\mathrm{L}^\infty(\Lambda;\mathbb{M}_d) \ni \rho \mapsto -\nabla\cdot\rho\nabla \in \mathscr{L}(\mathrm{W}^{1,p}_\Xi(\Lambda);\mathrm{W}^{-1,p}_\Xi(\Lambda))$$

is a linear bounded operator.

The first fundamental property of the divergence-gradient operators is the following, the so called *Kato square root property*:

Proposition 1.5.5. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$ and let $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ as well as $p \geq 2$. Assume that*

$$\left(-\nabla \cdot \rho\nabla + 1 \restriction \mathrm{L}^2(\Lambda)\right)^{-1/2} \in \mathscr{L}_{\mathrm{iso}}\left(\mathrm{L}^2(\Lambda); \mathrm{W}^{1,2}_\Xi(\Lambda)\right). \tag{1.42}$$

Then the following assertions are true:

(i) The operator $-\nabla \cdot \rho\nabla + 1$ is a positive *one on any space $W^{-1,p}_\Xi(\Lambda)$, cf. Definition 1.1.12.*

(ii) The square root satisfies

$$\left(-\nabla \cdot \rho\nabla + 1\right)^{-1/2} \in \mathscr{L}_{\mathrm{iso}}\left(\mathrm{W}^{-1,p}_\Xi(\Lambda); \mathrm{L}^p(\Lambda)\right),$$

or in other words,

$$\mathrm{dom}\left((-\nabla \cdot \rho\nabla + 1)^{1/2}\right) = \mathrm{L}^p(\Lambda).$$

Proof. The second assertion is exactly the main theorem in [20], from which the first one is derived in [20, Thm. 11.5]. The geometric assumptions are the same as ours, and Assumption 4.2 in [20] is exactly (1.42). □

Remark 1.5.6. Sufficient conditions for Assumption (1.42) in Proposition 1.5.5 to be satisfied are, for our special cases:

- $\Lambda \cup \Xi$ is regular in the sense of Gröger, or
- $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{S}_d(\rho_\bullet))$.

This can be seen as follows: If $\Lambda \cup \Xi$ is regular in the sense of Gröger, then Λ is a Lipschitz domain and hence a universal extension domain as seen in Theorem 1.3.6 and a d-set by Proposition 1.2.38. But then, together with the other assumptions on $\Lambda \cup \Xi$, the isomorphism property in (1.42) is the main result in [58]. On the other hand, if $\Lambda \cup \Xi$ is only generalized regular in the sense of Gröger or only Lipschitz around $\partial\Lambda \setminus \Xi$, we do no longer know that Λ is a d-set, hence we suppose a symmetric coefficient function in this case, which also enforces (1.42) (see [20, Rem. 4.3]).

Proposition 1.5.5 opens the door to use Lemma 1.2.47 to interpolate $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ and $\mathcal{D}_p(\rho)$, however this is by (very) far not the only useful consequence of the square root property, see e.g. [20, 53, 54, 89, 135], see also Theorem 1.5.16 below.

Corollary 1.5.7. *Let the assumptions of Proposition 1.5.5 be satisfied. Then*

$$\big(\mathrm{W}^{-1,p}_{\Xi}(\Lambda), \mathcal{D}_p(\rho)\big)_{\theta,r} \doteq \big(\mathrm{L}^p(\Lambda), \mathcal{D}_p(\rho)\big)_{2\theta-1,r}$$

for all $1 \le r \le \infty$ *and* $\frac{1}{2} < \theta < 1$.

Remark 1.5.8. The attentive reader will have noticed that we have used the operators $-\nabla\cdot\rho\nabla+1$ in conjunction with a coercive coefficient function $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$ in Proposition 1.5.5. Of course, the reason behind the addend "+1" is to make sure that we still have an *elliptic* differential operator at hand, even if $\Xi = \emptyset$. If $\Xi \neq \emptyset$ and $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$, then $-\nabla\cdot\rho\nabla$ is already elliptic on $\mathrm{W}^{1,2}_{\Xi}(\Lambda)$ by itself, since

$$\operatorname{Re}\langle -\nabla\cdot\rho\nabla\psi, \psi\rangle \ge 2\rho_\bullet \|\nabla\psi\|^2_{\mathrm{L}^2(\Lambda)} \cong 2\rho_\bullet \|\psi\|^2_{\mathrm{W}^{1,2}_{\Xi}(\Lambda)} \quad \text{for } \psi \in \mathrm{W}^{1,2}_{\Xi}(\Lambda)$$

by the Poincaré inequality (see Remarks 1.3.20 and 1.5.2). Since we neither want to introduce even more terminology nor do a case distinction for every occurrence of the operator, we leave the possibility of $\Xi = \emptyset$ open and thrust it upon the reader to add a virtual "+1" in the following results in that case, if there is no other source of coercivity available. For a positive example of the latter case, the operator $-\nabla\cdot\rho\nabla + B_\gamma$ is still elliptic on $\mathrm{W}^{1,2}_{\Xi}(\Lambda)$ if $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega,\mathbb{R})$ as in Definition 1.5.11 below satisfies $\int_\Gamma \gamma \,\mathrm{d}\omega > 0$ (recall that $\overline{\Gamma} = \partial\Lambda$ if $\Xi = \emptyset$).

Even if we have in general no further information about $\mathcal{D}_p(\rho)$, a result which by all means deserves to be mentioned is that for $p > d$, $\mathcal{D}_p(\rho)$ indeed embeds into a Hölder space if $\Lambda \cup \Xi$ is regular in the sense of Gröger and $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$, and that this embedding is stable under interpolation. Note that this is essentially a Hölder regularity result for the solution u of the equation $-\nabla\cdot\rho\nabla u = f$ with $f \in \mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for $p > d$, complemented with mixed boundary conditions.

Theorem 1.5.9 ([144, Thm. 1.1]/[79, Thm. 3.3]/[53, Lem. 4.8]). *Let* $\Lambda \cup \Xi$ *be regular in the sense of Gröger and let* $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ *and* $p > d$. *Then there exists* $0 < \alpha < 1$ *such that*

$$\mathcal{D}_p(\rho) \hookrightarrow \mathrm{C}^\alpha(\Lambda)$$

holds true. Even more, there exist $\theta > 0$ *and* $0 < \beta < \alpha$ *such that even*

$$\big[\mathrm{W}^{-1,p}_\Xi(\Lambda), \mathcal{D}_p(\rho)\big]_{1-\theta} \hookrightarrow \mathrm{C}^\beta(\Lambda)$$

holds true.

Hölder estimates as in the previous theorem are of great interest with respect to fixed point theorems and continuity properties of solution operators and we will in fact obtain a somewhat analogous theorem for the parabolic case as one of the main results later (see Theorem 2.1.4).

The admissible geometry in [53, 144] is much broader than just regular in the sense of Gröger, but does not quite allow *generalized* regular in the sense of Gröger, unfortunately. The result in [79] is the earlier and original one, but it is restricted to space dimensions $2 \le d \le 4$.

Remark 1.5.10. In fact, we can show that an interpolation space between $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ and $\mathcal{D}_p(\rho)$ always embeds into a Hölder space under the assumptions of Proposition 1.5.5 as long as $\mathcal{D}_p(\rho)$ does so: Thanks to Corollaries 1.5.7 and 1.1.10, we have

$$\big(\mathrm{W}^{-1,p}_\Xi(\Lambda), \mathcal{D}_p(\rho)\big)_{\theta,1} \doteq \big(\mathrm{L}^p(\Lambda), \mathcal{D}_p(\rho)\big)_{2\theta-1,1} \hookrightarrow \big[\mathrm{L}^p(\Lambda), \mathrm{C}^\alpha(\Lambda)\big]_{2\theta-1}.$$

By [147], the latter space may be identified with a *Triebel-Lizorkin* space which in turn can be shown to embed into a Hölder space $\mathrm{C}^\beta(\Lambda)$ whenever $\beta = \alpha - 2\theta(\alpha + \frac{d}{p}) > 0$, cf. [146, Ch. 2.8.1], under our assumptions.

Let us next introduce an extensional operator which allows to incorporate

Robin boundary conditions into the divergence-gradient operator.

Definition 1.5.11. Let $\Lambda \cup \Xi \subset \mathbb{R}^d$ with Ξ possibly empty be Lipschitz around $\partial\Lambda \setminus \Xi$ and set $\Gamma = \partial\Lambda \setminus \Xi$. Assume that $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega)$, let $p \geq 2$ and choose r such that $\frac{d-p}{d-1}\frac{1}{p} < \frac{1}{r} < \frac{d}{d-1}\frac{1}{p}$. Then we define $B_\gamma\colon \mathrm{L}^r(\overline{\Gamma}) \to \mathrm{W}^{-1,p}_\Xi(\Lambda)$ by

$$\langle B_\gamma\psi, \xi\rangle := \int_\Gamma \gamma\psi \,\mathrm{tr}\,\overline{\xi}\,\mathrm{d}\omega \quad \text{for } \psi \in \mathrm{L}^r(\overline{\Gamma}) \text{ and } \xi \in \mathrm{W}^{1,p'}_\Xi(\Lambda).$$

and $\mathsf{B}_\gamma := B_\gamma \circ \mathrm{tr}\colon \mathrm{W}^{1,p}_\Xi(\Lambda) \to \mathrm{W}^{-1,p}_\Xi(\Lambda)$. We further set $\mathsf{B} := \mathsf{B}_1$ for $\gamma \equiv 1$, the constant function, and $B := B_1$ accordingly.

In the setting of Definition 1.5.11, the trace operator is available from Lemma 1.2.57 via Theorem 1.3.23. The choice of r is done exactly such that the trace operator tr maps both $\mathrm{W}^{1,p'}_\Xi(\Lambda)$ to $\mathrm{L}^{r'}(\overline{\Gamma})$ and $\mathrm{W}^{1,p}_\Xi(\Lambda)$ to $\mathrm{L}^r(\overline{\Gamma})$, hence $B_\gamma\psi$ is indeed an element of $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ and $\mathsf{B}_\gamma = B_\gamma \circ \mathrm{tr}$ is well defined.

Remark 1.5.12. It is clear from definition that $B_\gamma\psi = \mathrm{tr}^*(\gamma\psi) \in \mathrm{W}^{-1,p}_\Xi(\Lambda)$ for tr considered as an operator in $\mathscr{L}(\mathrm{W}^{1,p'}_\Xi(\Lambda);\mathrm{L}^{r'}(\overline{\Gamma}))$, so in particular $B = \mathrm{tr}^*$ and $\mathsf{B} = \mathrm{tr}^*\,\mathrm{tr} \in \mathscr{L}(\mathrm{W}^{1,p}_\Xi(\Lambda);\mathrm{W}^{-1,p}_\Xi(\Lambda))$.

Lemma 1.5.13. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$ and let $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$.*

(i) Let $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega,\mathbb{R}_0^+)$ and let possibly $\int_\Gamma \gamma\,\mathrm{d}\omega > 0$ if $\Xi = \emptyset$ (see Remark 1.5.8). Then the operator $-\nabla\cdot\rho\nabla + \mathsf{B}_\gamma$ is continuously invertible if considered on $\mathrm{W}^{1,2}_\Xi(\Lambda)$, i.e.,

$$-\nabla\cdot\rho\nabla + \mathsf{B}_\gamma \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,2}_\Xi(\Lambda);\mathrm{W}^{-1,2}_\Xi(\Lambda)\big), \tag{1.43}$$

and the norm of the inverse is bounded by $1/\rho_\bullet$.

(ii) Moreover, we have

$$-\nabla\cdot\rho\nabla \in \mathscr{L}_{\mathrm{iso}}(\mathcal{D}_p(\rho);\mathrm{W}^{-1,p}_{\Xi}(\Lambda))$$

for $p \geq 2$.

Proof. The first assertion follows from the famous Lax-Milgram lemma [47, Ch. VII §1.1, Thm. 1], cf. also Remark 1.5.8. The second one is a consequence of the definition of $-\nabla\cdot\rho\nabla$ on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ and an application of the open mapping theorem: By the first assertion, for every $f \in \mathrm{W}^{-1,p}_{\Xi}(\Lambda) \hookrightarrow \mathrm{W}^{-1,2}_{\Xi}(\Lambda)$ there exists a unique $u = u_f \in \mathrm{W}^{1,2}_{\Xi}(\Lambda)$ such that $-\nabla\cdot\rho\nabla u = f$. But this implies that $u \in \mathcal{D}_p(\rho)$ and thus the open mapping theorem gives a continuous inverse of the continuous linear bijective operator $-\nabla\cdot\rho\nabla\colon \mathcal{D}_p(\rho) \to \mathrm{W}^{-1,p}_{\Xi}(\Lambda)$. $\square$

We next give conditions for the domain of $-\nabla\cdot\rho\nabla + \mathsf{B}_\gamma$ on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ to be still $\mathcal{D}_p(\rho)$.

Lemma 1.5.14. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$. Assume that $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$ and $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega)$ and let $p \geq 2$. Suppose further that there is $r > p\frac{d-1}{d}$ such that the trace operator* tr *extends to a compact operator from $\mathcal{D}_p(\rho)$ to $\mathrm{L}^r(\overline{\Gamma};\omega)$.*
Then B_γ is bounded on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ relative to $-\nabla\cdot\rho\nabla$ with arbitrarily small relative bound, i.e., for every $\varepsilon > 0$ there exists $\eta \geq 0$ such that

$$\|\mathsf{B}_\gamma\psi\|_{\mathrm{W}^{-1,p}_{\Xi}(\Lambda)} \leq \varepsilon\|-\nabla\cdot\rho\nabla\psi\|_{\mathrm{W}^{-1,p}_{\Xi}(\Lambda)} + \eta\|\psi\|_{\mathrm{W}^{-1,p}_{\Xi}(\Lambda)} \quad \textit{for all } \psi \in \mathcal{D}_p(\rho).$$

Moreover, the domain of $-\nabla\cdot\rho\nabla + \mathsf{B}_\gamma$ in $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ is still $\mathcal{D}_p(\rho)$.

Proof. The operator $\mathsf{B}_\gamma = B_\gamma \circ \mathrm{tr}$ is a *compact* one if considered on $\mathcal{D}_p(\rho)$ due to the assumption on the trace operator, and hence compact relative to $-\nabla\cdot\rho\nabla$ on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$, cf. [96, Ch. IV.1.3]. But this implies the relative-boundedness assertion (see [26]), and the equality of domains follows from either by a classic perturbation theorem [96, Ch. IV.1.1]. $\square$

Remark 1.5.15.

(i) We point out that the assumption on the trace operator for *some* $r > p\frac{d-1}{d}$ is satisfied automatically as soon as $\mathcal{D}_p(\rho) \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$ due to Lemma 1.2.57.

(ii) In their proof of the same result [80, Lem. 5.15] for $\Lambda \cup \Xi$ (volume-preserving) regular in the sense of Gröger, the authors of [80] use that $\mathrm{tr}\colon \mathcal{D}_p(\rho) \to \mathrm{L}^\infty(\overline{\Gamma};\omega)$ is compact for $p > d$ (see Theorem 1.5.9), that the assertion holds true for $p = 2$, and interpolation techniques for the remaining range $2 \leq p \leq d$. The latter is the point where we get stuck for more general geometries, cf. Remark 1.2.45. If this interpolation problem was resolved, one could of course imitate the proof in [80] also for more general geometries.

(iii) One could use an analogous technique to augment first order differential operators with suitably bounded coefficient functions to the divergence-gradient operators. We omit the details here.

We next formulate the maximal parabolic regularity result for divergence-gradient operators under the assumptions that the Kato square root property (1.42) is satisfied.

Theorem 1.5.16. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$, let $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ and let $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega)$. Assume that* (1.42) *is satisfied. Then the following assertions are true:*

(i) The operators $-\nabla \cdot \rho\nabla$ satisfy maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for every $p \geq 2$.

(ii) If the assumptions of Lemma 1.5.14 are satisfied for some $p \geq 2$, then $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ also satisfies maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$.

Proof. It is a consequence of the Kato square root property that $-\nabla \cdot \rho\nabla$ satisfies maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for every $p \geq 2$ as proven in [20, Thm. 11.5] (see also [54, 80]). We have verified the assumptions in [20] already in Proposition 1.5.5. Maximal parabolic regularity including the boundary form B_γ follows, if B_γ is bounded in $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ relative to $-\nabla\cdot\rho\nabla$ with arbitrarily small relative bound by [99, Cor. 2], which is exactly the assertion of Lemma 1.5.14. We refer to [80, Thm. 5.16] for more details. □

1.5.1 Maximal Sobolev regularity

While Proposition 1.5.5 shows that the *square root* of a divergence-gradient operator admits maximal Sobolev regularity in the sense of "spending" exactly one order of differentiability in a continuously invertible fashion, the lack of more precise knowledge about $\mathcal{D}_p(\rho)$ and maximal Sobolev regularity of the divergence-gradient operator itself, i.e., the question whether

$$-\nabla\cdot\rho\nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_{\Xi}(\Lambda);\mathrm{W}^{-1,p}_{\Xi}(\Lambda)\big) \quad \text{for } p>2,$$

is still a problem to overcome when dealing with these operators on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$. This is, however, not a "special feature", or lack thereof, of the divergence-gradient operators on negative Sobolev spaces but is shared by their siblings, the corresponding operators on $\mathrm{L}^p(\Lambda)$ spaces if Λ is not regular enough and ρ is generally discontinuous.

Due to Lemma 1.5.13, maximal Sobolev regularity for $-\nabla\cdot\rho\nabla$ is in fact equivalent to $\mathcal{D}_p(\rho) \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$. We will use both formulations interchangeably, depending on the context. A particular point where we get stuck by dealing with just $\mathcal{D}_p(\rho)$ is that we will have to assume that $\mathcal{D}_p(\rho) \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$ in order to also obtain maximal parabolic regularity for $-\nabla\cdot\rho\nabla+\mathsf{B}_\gamma$ via Theorem 1.5.16 if $\Lambda\cup\Xi$ is only Lipschitz around $\partial\Lambda\setminus\Xi$. It was already mentioned also in Remark 1.5.15 that the assertions of Lemma 1.5.14 are valid in this case.

Corollary 1.5.17. *Let $\Lambda\cup\Xi$ be Lipschitz around $\partial\Lambda\setminus\Xi$, let $\rho\in\mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$ and let $\gamma\in\mathrm{L}^\infty(\overline{\Gamma};\omega)$. Assume that the assumptions of Proposition 1.5.5 are satisfied and that $\mathcal{D}_p(\rho) \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$ for some*

$p > 2$. *Then* $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ *satisfies maximal parabolic regularity on* $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ *with domain* $\mathrm{W}^{1,p}_\Xi(\Lambda)$.

However, we will need maximal Sobolev regularity, so, the isomorphism property of $-\nabla \cdot \rho\nabla$ between $\mathrm{W}^{1,p}_\Xi(\Lambda)$ and $\mathrm{W}^{-1,p}_\Xi(\Lambda)$, primarily to obtain certain permanence principles for the treatment of quasilinear parabolic evolution equations involving the divergence-gradient operators ("*in divergence form*"), also on domains which are Lipschitz around $\partial\Lambda \setminus \Xi$, and less for establishing properties of the divergence-gradient operators on domains with rough boundaries. To illustrate this, let us briefly consider the exemplary quasilinear problem

$$u'(t) - \nabla \cdot \sigma(u)(t)\rho\nabla u(t) = F(u)(t) \quad \text{in } \mathrm{W}^{-1,p}_\Xi(\Lambda) \quad \text{on } J, \quad u(T_0) = u_0, \tag{1.44}$$

known already from the introduction, for $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ and for simplicity $\sigma \in \mathrm{C}^{1-}_{\mathrm{loc}}(\mathbb{R})$ with $\sigma \geq \sigma_\bullet > 0$. Quite similar to the situation regarding nonautonomous maximal parabolic regularity in Section 1.4, one has to deal with potentially varying domains $\mathcal{D}_p(\sigma(u(t))\rho)$ here, which vary not only in $t \in J$ for a fixed function u, but also for every possible function u. While there *are* concepts to treat nonconstant domains (see e.g. [3, Ch. IV]), we will be concerned with only constant ones and will take appropriate measures to ensure this property. The case $p > d$ is of particular interest, both with respect to optimal control problems built around quasilinear parabolic equations in divergence form, as well as inherently in the quasilinear problem themselves (see Chapter 2.2).

A particular and very noteworthy case where $\mathcal{D}_p(\rho)$ is exactly known and coincides with $\mathrm{W}^{1,p}_\Xi(\Lambda)$, i.e., $-\nabla \cdot \rho\nabla$ admits maximal Sobolev regularity, is when p is close to 2, and then we even have maximal Sobolev regularity for $-\nabla\cdot\rho\nabla+\mathsf{B}_\gamma$, so an extension of the Hilbert space case as in (1.43). The corresponding results were obtained by Gröger [73] and Gröger and Rehberg [75] already in 1989 for $\Lambda\cup\Xi$ regular in the sense of Gröger and extended only very recently by Haller-Dintelmann, Jonsson, Knees and Rehberg [77].

Theorem 1.5.18. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$, let $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ and let $\gamma \in \mathrm{L}^\infty(\overline{\Gamma}; \omega, \mathbb{R}_0^+)$. Then the following assertions are true:*

(i) There exists $p_0 > 2$ such that for all $2 \le p < p_0$ we have

$$-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_\Xi(\Lambda); \mathrm{W}^{-1,p}_\Xi(\Lambda)\big),$$

i.e., $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ admits maximal Sobolev regularity.

(ii) Both the number p_0 and the norm of $(-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma)^{-1}$ can be estimated uniformly *with respect to $\rho_\bullet$ and $\|\rho\|_{\mathrm{L}^\infty(\Lambda;\mathbb{M}_d)}$.*

Proof. The existence of p_0 such that

$$-\nabla \cdot \rho\nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_\Xi(\Lambda); \mathrm{W}^{-1,p}_\Xi(\Lambda)\big) \tag{1.45}$$

for all $2 \le p < p_0$ is proven in [77, Thm. 5.6]. Uniformity of p_0 and the norm of the inverse is proven in [77, Rem. 5.7] or [73, 75]. From there, we can augment the differential operator with B_γ without losing the isomorphism property and the uniform bounds by applying the argument as in [75, Thm. 4], i.e., boot-strapping the regularity of

$$u = (-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma)^{-1} f \in \mathrm{W}^{1,2}_\Xi(\Lambda)$$

for $f \in \mathrm{W}^{-1,p}_\Xi(\Lambda)$ using (1.45). This works because of the isomorphism property for $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ for $p = 2$ and the uniformity of constants w.r.t. $\rho_\bullet$ and $\|\rho\|_{\mathrm{L}^\infty(\Lambda;\mathbb{M}_d)}$ (see Lemma 1.5.13); it only remains to observe that B_γ maps $\mathrm{W}^{1,2}_\Xi(\Lambda)$ continuously into $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ if p is close enough to 2. This in turn is the case because if r in Definition 1.5.11 is chosen by $\frac{d-2}{d-1}\frac{1}{2} < \frac{1}{r} < \frac{d}{d-1}\frac{1}{2}$, i.e., B_γ is well-defined on $\mathrm{W}^{1,2}_\Xi(\Lambda)$, then we can find $2 < p < p_0$ such that also $\frac{d-p}{d-1}\frac{1}{p} < \frac{1}{r} < \frac{d}{d-1}\frac{1}{p}$, hence B_γ maps $\mathrm{W}^{1,2}_\Xi(\Lambda)$ into $\mathrm{W}^{-1,p}_\Xi(\Lambda)$ and we can use the isomorphism assumption (1.45). $\square$

We note that in [77] it is only assumed that Ξ is a $(d-1)$-set and that Λ is a $\mathrm{W}^{1,p}_\Xi$-extension domain uniformly for all $1 \le p < \infty$. To augment the divergence-gradient operator with a boundary form w.r.t. $\mathcal{H}^{d-1} \restriction \overline{\Gamma}$ as we did, one only needs the assumptions of Lemma 1.2.57 (or Lemma 1.2.51 for the measure, respectively), which already nearly imply the extension property by Theorem 1.2.31, and of course the assumptions on γ as in Lemma 1.5.13.

Remark 1.5.19.

(i) If in the situation of Theorem 1.5.18, the coefficient function ρ is symmetric, i.e., $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{S}_d(\rho_\bullet))$, then the adjoint operator of $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ acts in the same way on $\mathrm{W}^{1,p'}_\Xi(\Lambda)$ and is still a topological isomorphism. In this sense, we obtain an interval of the form $(\frac{p_0}{p_0-1}, p_0)$ around 2 as the "isomorphism range" for the integrability orders and the uniformity assertion is still true.

(ii) The phenomenon of an open interval around a given "isomorphy anchor", in this case $\mathrm{W}^{1,2}_\Xi(\Lambda)$, is most natural in an interpolation scale (see Theorem 1.2.46 and [52, Lem. 5.5] or [77, Ch. 5]) due to the result of Šneĭberg [139], cf. also [56, Ch. 1.3.5].

Now, it seems like maximal Sobolev regularity for the operators $-\nabla \cdot \rho\nabla + \mathsf{B}_\gamma$ for $p > d$ is "done" for $d = 2$ and rather general sets by the cited Theorem 1.5.18 above. The bad news is that this is the only case which is "done" and will probably stay so, even for the operators without the boundary form B_γ (for these, it becomes apparent in the proof of Theorem 1.5.18 that the proximity of p_0 to 2 is crucial):

First, there is no hope *at all* in obtaining maximal Sobolev regularity for divergence-gradient operators involving mixed boundary conditions and $p \geq 4$ due to the famous counter-example by Shamir [136, Introduction]: There exists a harmonic function u on a smooth domain, satisfying homogeneous mixed boundary conditions, for which $\|\nabla u\|_2^4$ is not integrable.

This leaves us with dimension $d = 3$ for the case of mixed boundary conditions in general. As the next negative result, there is no hope to obtain a result analogous to Theorem 1.5.18 for $p > d \geq 3$ and discontinuous coefficient functions ρ, since the general integrability of the gradient ∇u of a solution to $-\nabla \cdot \rho\nabla u = 0$ may be arbitrarily close to 2 in that case (see [60, Ch. 4], [117]) even if there are no mixed boundary conditions present. This means one has to pose further assumptions on Λ, Ξ and ρ. Even then, one already *a priori* knows that the "easy" cases of the Laplacian, so $\rho \equiv 1$, on a *strong* Lipschitz domain, complemented with pure Dirichlet ($\Xi = \partial\Lambda$, [92]) or pure Neumann ($\Xi = \emptyset$, [154]) boundary conditions still only admit maximal Sobolev regularity for $p > d = 3$ arbitrarily close to 3, depending on the properties of the domain. Still, there is a large variety of admissible constellations between Λ, Ξ and ρ with $\Lambda \cup \Xi$ regular in the sense of Gröger as exhibited in the work of Disser, Kaiser and Rehberg [52], assuring that $-\nabla \cdot \rho\nabla$ is a topological isomorphism between $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for some $p > 3 = d$. We will use this isomorphism—or maximal Sobolev regularity—property as a "black box" and refer to said paper for practical model constellations.

In view of the asserted form of the quasilinear operators as in (1.44), it also seems natural to start with $-\nabla \cdot \rho\nabla$ and to establish permanence principles for maximal Sobolev regularity from there. Indeed, maximal Sobolev regularity is stable under multiplication with strictly positive uniformly continuous functions:

Proposition 1.5.20 ([52, Lem. 6.2])**.** *Let $\Lambda \subset \mathbb{R}^d$ be Lipschitz around $\partial\Lambda \setminus \Xi$. Assume that $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet))$ as well as $2 \leq p \leq 2^\star = \frac{2d}{d-2}$, and let $\phi \in \mathrm{C}(\overline{\Lambda}; \mathbb{R})$ satisfy $\phi > 0$ on $\overline{\Lambda}$. If $-\nabla \cdot \rho\nabla$ is a topological isomorphism between $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$, i.e.,*

$$-\nabla \cdot \rho\nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_{\Xi}(\Lambda); \mathrm{W}^{-1,p}_{\Xi}(\Lambda)\big)$$

then

$$-\nabla \cdot \phi\rho\nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_{\Xi}(\Lambda); \mathrm{W}^{-1,p}_{\Xi}(\Lambda)\big),$$

so $-\nabla \cdot \phi\rho\nabla$ is also a topological isomorphism between $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$.

The limitation of integrability to $2 \leq p \leq 2^\star = \frac{2d}{d-2}$ is a rather unfortunate technical obstruction coming from the localization argument used in the proof (see also [60, Lem. 3.9]). Due to $2^\star = 4$ for $d = 4$, the previous proposition is useless for the consideration of $p > d = 4$ (and of course for all higher dimensions), which we, however, had ruled out in the presence of mixed boundary conditions any way. For $d = 3$, we have $2^\star = 6$, which seems acceptable in view of the explanations regarding the case $p \gg 3$. In this sense, we will limit ourselves to the case $d \leq 3$ when using Proposition 1.5.20. Note that we only need Λ to be a $\mathrm{W}^{k,p}_{\Xi}$-extension domain for $k = 0, 1$ for the proposition to hold.

Extending the result of Proposition 1.5.20, the dependence on ϕ of $-\nabla\cdot\rho\nabla$ in $\mathscr{L}(\mathrm{W}^{1,p}_{\Xi}(\Lambda); \mathrm{W}^{-1,p}_{\Xi}(\Lambda))$, i.e., the mapping

$$\Big\{\phi \in \mathrm{C}(\overline{\Lambda}; \mathbb{R})\colon \phi > 0\Big\} \ni \phi \to -\nabla\cdot\phi\rho\nabla \in \mathscr{L}\big(\mathrm{W}^{1,p}_{\Xi}(\Lambda); \mathrm{W}^{-1,p}_{\Xi}(\Lambda)\big) \quad (1.46)$$

is a (Lipschitz-)continuous one. This follows immediately from the observation in Remark 1.5.4 and allows for the following extension:

Proposition 1.5.21 ([52, Thm. 6.3/Cor. 6.4])**.** *Let the assumptions of Proposition 1.5.20 be satisfied, let $\mathfrak{C}$ be a compact subset of $C(\overline{\Lambda}; \mathbb{R})$ such that $\phi > 0$ on $\overline{\Lambda}$ for all $\phi \in \mathfrak{C}$, and assume that*

$$-\nabla \cdot \rho\nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,p}_{\Xi}(\Lambda); \mathrm{W}^{-1,p}_{\Xi}(\Lambda)\big).$$

Then the mapping

$$\mathfrak{C} \ni \phi \to \big(-\nabla \cdot \phi\rho\nabla\big)^{-1} \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{-1,p}_{\Xi}(\Lambda); \mathrm{W}^{1,p}_{\Xi}(\Lambda)\big)$$

is uniformly bounded and Lipschitz-continuous.

Remark 1.5.22. In the proof of Proposition 1.5.21, the following general principle is used, which will become handy in various places: Let $A, B \in \mathscr{L}(X;Y)$ such that $A^{-1}, B^{-1} \in \mathscr{L}(Y;X)$. Then we re-arrange

$$\begin{aligned}\left\|A^{-1} - B^{-1}\right\|_{\mathscr{L}(Y;X)} &= \left\|B^{-1}(A-B)A^{-1}\right\|_{\mathscr{L}(Y;X)} \\ &\leq \|B^{-1}\|_{\mathscr{L}(Y;X)}\|A^{-1}\|_{\mathscr{L}(Y;X)}\|A-B\|_{\mathscr{L}(X;Y)}.\end{aligned}$$

Hence, the inversion mapping $\mathscr{L}(X;Y) \ni A \mapsto A^{-1} \in \mathscr{L}(Y;X)$ is indeed Lipschitz-continuous on every bounded subset of the (open) set of continuously invertible operators in $\mathscr{L}(X;Y)$.

The propositions above may be put to good use immediately by establishing nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ for a family of divergence-gradient operators:

Lemma 1.5.23. *Let $\Lambda \cup \Xi$ be Lipschitz around $\partial\Lambda \setminus \Xi$ and let $\rho \in \mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$ and $\varphi \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Lambda};\mathbb{R}))$ with $\varphi(t) > 0$ on $\overline{\Lambda}$ for every $t \in \overline{J}$. Let moreover $\gamma \in \mathrm{C}(\overline{J};\mathrm{L}^\infty(\overline{\Gamma};\omega))$ and assume that (1.42) is satisfied and $\mathcal{D}_p(\rho) = \mathrm{W}^{1,p}_{\Xi}(\Lambda)$ for some $2 \le p \le 2^\star = \frac{2d}{d-2}$. Then the following assertions hold true:*

(i) We have

$$-\nabla \cdot \varphi(\cdot)\rho\nabla \in \mathrm{C}\Big(\overline{J};\mathscr{L}_{\mathrm{iso}}(\mathrm{W}^{1,p}_{\Xi}(\Lambda);\mathrm{W}^{-1,p}_{\Xi}(\Lambda))\Big)$$

and the operator satisfies nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ over J and every subinterval $J_\bullet \subseteq J$.

(ii) The operator $t \mapsto -\nabla \cdot \varphi(t)\rho\nabla + \mathsf{B}_{\gamma(t,\cdot)}$ also satisfies nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ over J and every subinterval $J_\bullet \subseteq J$.

Proof. By Proposition 1.5.20, we know that

$$-\nabla \cdot \varphi(t)\rho\nabla \in \mathscr{L}_{\mathrm{iso}}(\mathrm{W}^{1,p}_{\Xi}(\Lambda);\mathrm{W}^{-1,p}_{\Xi}(\Lambda)) \quad \text{for every } t \in \overline{J}.$$

For fixed t, each of these operators satisfies autonomous maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ by Theorem 1.5.16. Moreover, $t \mapsto -\nabla\cdot\varphi(t)\rho\nabla$ is a continuous mapping on $\overline{J}$ (see also (1.46)), hence Theorem 1.4.17 shows that $t \mapsto -\nabla\cdot\varphi(t)\rho\nabla$ satisfies nonautonomous maximal parabolic regularity over $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$.
To also include the boundary form, we need a perturbation result for which we appeal—as in Theorem 1.5.16—to [99, Cor. 2]. Together with

Lemma 1.5.14, this shows that $-\nabla\cdot\varphi(t)\rho\nabla+\mathsf{B}_{\gamma(t,\cdot)}$ satisfies (autonomous) maximal parabolic regularity on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$. Since we have assumed γ to be continuous in time, we can again refer to Theorem 1.4.17 to obtain nonautonomous maximal parabolic regularity for $t\mapsto -\nabla\cdot\varphi(t)\rho\nabla+\mathsf{B}_{\gamma(t,\cdot)}$ on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ with domain $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$.

The proof also shows that the result holds true for every subinterval $J_\bullet\subseteq J$ just by restriction. ☐

We will extend this result in Chapter 2.2 to continuous dependency of the inverses of the total differential operators on φ, cf. Lemma 2.2.13.

Let us finally observe two more consequences of maximal Sobolev regularity. Firstly, we know that if there is maximal Sobolev regularity for $-\nabla\cdot\rho\nabla$ for some $p_0>2$, then we know the same already for all $2\le p\le p_0$ by interpolation:

Lemma 1.5.24. *Let $\Lambda\cup\Xi$ be Lipschitz around $\partial\Lambda\setminus\Xi$ and let $\rho\in\mathrm{L}^\infty(\Lambda;\mathbb{M}_d(\rho_\bullet))$. Assume that $\mathcal{D}_{p_0}(\rho)\doteq\mathrm{W}^{1,p_0}_{\Xi}(\Lambda)$ for some $p_0>2$. Then it follows that $\mathcal{D}_p(\rho)\doteq\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ for all $2<p<p_0$.*

Proof. By Theorem 1.2.46, we know that

$$\mathrm{W}^{-1,p}_{\Xi}(\Lambda)=\big[\mathrm{W}^{-1,2}_{\Xi}(\Lambda),\mathrm{W}^{-1,p_0}_{\Xi}(\Lambda)\big]_\theta$$

and

$$\mathrm{W}^{1,p}_{\Xi}(\Lambda)=\big[\mathrm{W}^{1,2}_{\Xi}(\Lambda),\mathrm{W}^{1,p_0}_{\Xi}(\Lambda)\big]_\theta,$$

each for $\frac{1}{p}=\frac{1-\theta}{2}+\frac{\theta}{p_0}$ and $0<\theta<1$, cf. also Theorem 1.3.15. But then the isomorphism property for the operators on $\mathrm{W}^{1,2}_{\Xi}(\Lambda)$ and $\mathrm{W}^{1,p_0}_{\Xi}(\Lambda)$ follows by interpolating both $-\nabla\cdot\rho\nabla$ and its inverse on these spaces and observing that this gives rise to continuous operators between $\mathrm{W}^{1,p}_{\Xi}(\Lambda)$ and $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ which are inverse to each other and agree with $-\nabla\cdot\rho\nabla$ and its inverse on $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ on the dense subset $\mathrm{W}^{1,p_0}_{\Xi}(\Lambda)$. ☐

Again, we in fact only need that Λ is a $\mathrm{W}^{1,p}_{\Xi}$-extension domain uniformly for $1 \le p < \infty$ and that Ξ is a closed $(d-1)$-set for the result of the previous lemma.

Secondly, we have already observed in Theorem 1.5.9 and Remark 1.5.10 that certain interpolation spaces between $\mathrm{W}^{-1,p}_{\Xi}(\Lambda)$ and $\mathcal{D}_p(\rho)$ embed into Hölder spaces for $p > d$ if we have the square root property at our disposal. This of course does not change when $\mathcal{D}_p(\rho) \doteq \mathrm{W}^{1,p}_{\Xi}(\Lambda)$. In view of the maximal regularity property of the divergence-gradient operators and the maximal regularity embeddings as in Proposition 1.4.3, we thus give conditions for both the interpolation spaces and the maximal regularity spaces to embed into the space of (Hölder-) continuous functions. For further use, we formulate it directly in a quite general fashion.

Lemma 1.5.25. *Let the assumptions of Proposition 1.5.5 hold and let $J \subset \mathbb{R}_0^+$ be a finite open interval. If $\Omega \cup D$ is* volume-preserving *regular in the sense of Gröger, let $1 < s \le p$. Otherwise, let $2 \le s \le p$.*

(i) Let $\left(\frac{1}{2} + \frac{d}{s} - \frac{d}{2p}\right)\left(1 + \frac{d}{s} - \frac{d}{p}\right)^{-1} < \zeta < \theta < 1$. Then

$$\left(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right)_{\theta,p} \hookrightarrow \mathrm{C}^{\beta}(\Lambda)$$

for $\beta = 2\zeta\left(1 + \frac{d}{s} - \frac{d}{p}\right) - 1 + \frac{d}{p} - \frac{2d}{s}$.

(ii) Let $r > 2\left(1-\frac{d}{p}\right)^{-1}\left(1+\frac{d}{s}-\frac{d}{p}\right)$ and $\left(\frac{1}{2}+\frac{d}{s}-\frac{d}{2p}\right)\left(1+\frac{d}{s}-\frac{d}{p}\right)^{-1} < \xi < 1-\frac{1}{r}$. Then

$$\mathbb{W}^{1,r}\left(J; \mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right) \hookrightarrow \mathrm{C}^{\alpha}\left(J; \mathrm{C}^{\beta}(\Lambda)\right)$$

for $0 < \alpha < 1 - \frac{1}{r} - \xi$ and $\beta = 2\xi\left(1 + \frac{d}{s} - \frac{d}{p}\right) - 1 + \frac{d}{p} - \frac{2d}{s}$.

Proof. Using (1.5) and the Reiteration Theorem 1.1.11, we find

$$\begin{aligned}
&\left(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right)_{\theta,p} \hookrightarrow \left(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right)_{\zeta,1} \\
&\doteq \Big(\left(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right)_{\frac{1}{2},1}, \mathrm{W}^{1,p}_{\Xi}(\Lambda)\Big)_{2\zeta-1,1} \hookrightarrow \left(\mathrm{L}^{s}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\right)_{2\zeta-1,1}.
\end{aligned}$$

For $2 \le s \le p$, the latter embedding follows from $\mathrm{W}^{1,p}_{\Xi}(\Lambda) \hookrightarrow \mathrm{W}^{1,s}_{\Xi}(\Lambda) \hookrightarrow$

$\mathcal{D}_s(\rho)$, cf. Remark 1.5.4, the interpolation embedding principle from Corollary 1.1.10, and Lemma 1.2.47 together with Proposition 1.5.5. If only $1 < s < 2$, we have to rely on [69, Lem. 3.4] and (1.10) for the last step, because we do not know that the divergence-gradient operators are suitable positive operators on the spaces $\mathrm{W}^{-1,s}_\Xi(\Lambda)$. From here, we can argue for both cases at once.

For $\Xi = \emptyset$, we could now employ the embedding of $(\mathrm{L}^s(\Lambda), \mathrm{W}^{1,p}_\Xi(\Lambda))_{2\zeta-1,1}$ into the corresponding complex interpolation space and use Theorem 1.2.26 (interpolation identities for the Bessel scale on Λ) and Corollary 1.2.28 (embeddings for the Bessel scale on Λ) to obtain the wished-for embedding. For the general case, we will have to circumvent the obstacle of having no interpolation identities at hand.

Let $f \in \mathrm{W}^{1,p}_\Xi(\Lambda)$. Then on the one hand $f \in \mathrm{C}^{1-\frac{d}{p}}(\Lambda)$ by Theorem 1.2.27, which is in turn embedded into $\mathrm{C}^\beta(\Lambda)$. On the other hand, setting

$$\frac{1}{q} := \frac{2-2\zeta}{s} + \frac{2\zeta-1}{p},$$

we have by the definition of β together with Theorem 1.2.26 and Corollary 1.2.28:

$$\mathrm{W}^{1,p}_\Xi(\Lambda) \hookrightarrow \mathrm{W}^{1,p}(\Lambda) \hookrightarrow \big[\mathrm{L}^s(\Lambda), \mathrm{W}^{1,p}(\Lambda)\big]_{\frac{d}{q}+\beta} = \mathrm{H}^{\frac{d}{q}+\beta,q} \hookrightarrow \mathrm{C}^\beta(\Lambda).$$

Hence, we are allowed to estimate for all $f \in \mathrm{W}^{1,p}_\Xi(\Lambda) = \mathrm{W}^{1,p}_\Xi(\Lambda) \cap \mathrm{L}^s(\Lambda)$:

$$\|f\|_{\mathrm{C}^\beta(\Lambda)} \le C\|f\|_{\mathrm{H}^{\frac{d}{q}+\beta,q}(\Lambda)} \le C\|f\|_{\mathrm{L}^s(\Lambda)}^{1-\frac{d}{q}-\beta}\|f\|_{\mathrm{W}^{1,p}(\Lambda)}^{\frac{d}{q}+\beta}$$

Due to Lemma 1.1.9, this shows that

$$\big((\mathrm{L}^s(\Lambda), \mathrm{W}^{1,p}_\Xi(\Lambda)\big)_{\frac{d}{q}+\beta,1} \hookrightarrow \mathrm{C}^\beta(\Lambda),$$

from which the claim follows.

Now let us turn to the embedding for the maximal regularity space. Due to Proposition 1.4.3, we have

$$\mathbb{W}^{1,r}\big(J; \mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\big) \hookrightarrow \mathrm{C}^{\alpha}\big(J; \big(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\big)_{\xi,1}\big)$$

for all $1 < r < \infty$, where $0 < \xi < 1-\frac{1}{r}$ and $0 < \alpha < 1-\frac{1}{r}-\xi$. If r is as in the assumption, then the interval $\big(\big(\frac{1}{2}+\frac{d}{s}-\frac{d}{2p}\big)\big(1+\frac{d}{s}-\frac{d}{p}\big)^{-1}, 1-\frac{1}{r}\big)$ is nonempty. Hence, we can choose ξ such that $\big(\frac{1}{2}+\frac{d}{s}-\frac{d}{2p}\big)\big(1+\frac{d}{s}-\frac{d}{p}\big)^{-1} < \xi < 1-\frac{1}{r}$, for which we have

$$\big(\mathrm{W}^{-1,s}_{\Xi}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\big)_{\xi,1} \hookrightarrow \big(\mathrm{L}^{s}(\Lambda), \mathrm{W}^{1,p}_{\Xi}(\Lambda)\big)_{2\xi-1,1} \hookrightarrow \mathrm{C}^{\beta}(\Lambda)$$

for $\beta = 2\xi\big(1+\frac{d}{s}-\frac{d}{p}\big)-1+\frac{d}{p}-\frac{2d}{s}$, as in the proof of the foregoing assertion. Putting everything together proves the lemma. □

Remark 1.5.26. An alternative proof for Lemma 1.5.25 with $s = p$ can be obtained via the technique explained in Remark 1.5.10, which would in fact allow to show that, for different values of r, α, β than before,

$$\mathbb{W}^{1,r}\big(J; \mathrm{W}^{-1,p}_{\Xi}(\Lambda), \mathcal{D}_p(\rho)\big) \hookrightarrow \mathrm{C}^{\alpha}\big(J; \mathrm{C}^{\beta}(\Lambda)\big)$$

whenever $\mathcal{D}_p(\rho) \hookrightarrow \mathrm{C}^{\varrho}(\Lambda)$. The drawback of this ansatz is that α, β and in particular r would depend on ϱ which is in general not exactly known. Hence, one could not, for instance, give a precise minimum requirement for r depending only on p, which might be of interest.

1.6 Real spaces

Lastly, we introduce and do "the twist": Having worked with complex vector spaces of complex-valued functions so far, we will switch to real vector spaces of real-valued functions in the following. Many definitions and results introduced above are essentially complex in their nature, such as the complex interpolation functor or the Bessel potential spaces $\mathrm{H}^{s,p}(\mathbb{R}^d)$, but also the considerations regarding analytical semigroups in Section 1.4 or the square root property as in Proposition 1.5.5, which makes it preferable to introduce these concepts in their natural habitat. However, as we consider evolution equations connected to real-world applications in Chapter 2.2 and following, a generally real setup also seems adequate.

On a more practical level, we will have to require differentiability properties of functions inducing superposition– or Nemytskii-operators such as the function σ in (1.44), both for analytical purposes and for reasons related to the optimal control of such equations. But assuming a function to be *complex* differentiable makes it already analytical, which is a very strong property and generally not satisfied in many applications. A problem with differentiability also manifests in $x \mapsto \|x\|_X^2$ not being Fréchet-differentiable if X is a *complex* Hilbert space since the derivative cannot be $\mathbb{C}$-linear.

So, whenever we agree to consider *real* spaces, we talk about the real subspaces of real-valued functions of the function spaces introduced so far. It is clear that embeddings established for the complex versions are still valid for the real spaces. For the complex interpolation functor, we first complexify the real spaces, then apply complex interpolation and take the (correct) real subspace afterwards. This comes at the expense of equality up to equivalence of norms. We refer to [3, Ch. I.2.4] for details. Further, we have defined $\mathrm{W}_\Xi^{-k,p}(\Lambda)$ to be the *anti*-dual space of $\mathrm{W}_\Xi^{k,p}(\Lambda)$. The elements of the dual space of the *real* version of $\mathrm{W}_\Xi^{k,p}(\Lambda)$ are to be identified exactly with the antilinear forms of the complex $\mathrm{W}_\Xi^{-k,p}(\Lambda)$ which take real values when applied to real functions in the complex space $\mathrm{W}_\Xi^{k,p}(\Lambda)$. We still call the so-obtained real dual space $\mathrm{W}_\Xi^{-k,p}(\Lambda)$. Lastly, we have worked with *real* coefficient functions μ for the divergence-gradient operators $-\nabla\cdot\mu\nabla$ in Section 1.5, which makes the operator commute with complex conjugation (i.e., the real subspace of real functions is mapped to the same one in the image space). In this sense, the results of Section 1.5 are still true for the real versions of the spaces. If necessary, the boundary function $\gamma \in \mathrm{L}^\infty(\overline{\Gamma};\omega)$ also needs to be chosen as real-valued, of course.

CHAPTER 2

Analysis of quasilinear parabolic equations in divergence form

This chapter is devoted to the analysis of quasilinear parabolic evolution equations in divergence form

$$u'(t) - \nabla \cdot \sigma(u)\rho\nabla u(t) + u(t) = F(u)(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0 \tag{2.1}$$

for a finite time interval J. Here, $\Omega \subset \mathbb{R}^d$ is a domain with $\rho \in \mathrm{L}^\infty(\Omega; \mathbb{M}_d(\rho_\bullet))$ and $D \subseteq \partial\Omega$ closed, whereas σ, F and u_0 are chosen appropriately such that this problem is well-posed in a suitable sense. The reader may for now imagine $\sigma \in \mathrm{C}^{1-}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^+)$. We will also allow for a boundary operator as in Definition 1.5.11 in the formulation, accounting for Robin boundary conditions. Note that we have added the "+1" in the differential operator to make sure that it is coercive, even if $D = \emptyset$, cf. Remark 1.5.8. Moreover, we use the letter "u" for the solution as a tribute to the Russian notation as in the works of LADYZHENSKAYA ET AL [101, 102] upon which the result in Chapter 2.1 is based.

We will show in Theorem 2.2.10 below that (2.1) admits local-in-time solutions in maximal regularity spaces $\mathbb{W}^{1,r}(J_\bullet; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ for suitable subintervals $J_\bullet$ of J, under a Lipschitz-assumption of F on the maximal regularity space and very weak assumptions on the domain, namely $\Omega \cup D$ being Lipschitz around $\partial\Omega \setminus D$. This is done by showing that the divergence-gradient operator (including a boundary form) satisfies the assumptions on the differential operator in the abstract quasilinear "solution theorems" of Amann or Prüss, Theorems 2.2.4 and 2.2.7. Remarkably, the former even allows for a nonlocal-in-time dependence of σ and F on u. Unfortunately, these theorems and their proofs do not allow to read off direct conditions under which we can guarantee to obtain global-in-time solutions, even if F does not depend on u at all.

This is the point where the self-imposed additional structure of a divergence-gradient operator comes into play. Since the nonlinearity in (2.1) is "trapped" inside the coefficient function σ, the family of differential operators $-\nabla \cdot \sigma(u)(t)\rho\nabla + 1$ is *uniform* in their upper bounds and coercivity constants, depending on σ and ρ, of course. We will exploit this additional structure by showing that solutions u to the nonautonomous equation

$$u'(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) + u(t) = f(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = 0 \tag{2.2}$$

are Hölder-continuous on $Q := J \times \Omega$ for $f \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$ with $q > d$ and $s > 2(1 - \frac{d}{q})^{-1}$, even though μ is merely measurable, coercive and bounded, and that the set of solutions corresponding to f from bounded sets in $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$ is bounded in the Hölder space uniformly with respect to the coercivity- and upper bound. The proof of this result in the general context of $\Omega \cup D$ generalized regular in the sense of Gröger is contained in Chapter 2.1 with an extension to nonzero initial value in combination with inhomogeneous Dirichlet trace in Chapter 2.1.5.

We will then apply the new result to the model equation with frozen coefficients, that is,

$$u'(t) - \nabla \cdot \sigma(w)(t)\rho\nabla u(t) + u(t) = F(w)(t) \qquad \text{in } \mathrm{W}_D^{-1,q}(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0. \tag{2.3}$$

In conjunction with Schauder's fixed point theorem, this will allow to prove existence and uniqueness of *global* solutions in the maximal regularity space to the quasilinear equation (2.1) in Chapter 2.2, albeit under much stronger Lipschitz-continuity assumptions on F than necessary for the theorems of Amann and Prüss, but still allowing nonlocal-in-time operators. The continuity of the designated fixed point mapping $w \mapsto u_w$ such that u_w solves (2.3) depends heavily on maximal Sobolev regularity of the operators $-\nabla \cdot \sigma(u)(t)\rho\nabla + 1$ for which we assume that $-\nabla \cdot \rho\nabla + 1$ admits maximal Sobolev regularity, cf. Proposition 1.5.20, and the maximal parabolic regularity of the divergence-gradient operators (see Lemma 1.5.23) via the reformulation in terms of continuous operator invertibility as seen in Lemmata 1.4.11 and 1.4.15.

This way, it will become clear that the interplay between the well-behaved dependence of the divergence-gradient operator on $u \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and the Hölder-regularity of solutions u gives rise to a rather satisfying theory for global solutions of (2.1). The results in Chapter 2.1 and the global existence result have been published together with Joachim Rehberg in the article "Hölder-estimates for non-autonomous parabolic problems with rough data" [116].

2.1 Uniform Hölder-estimates for nonautonomous equations

We establish Hölder regularity for solutions of the nonautonomous equation

$$u'(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) + u(t) = f(t) \\ \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = 0 \quad (2.2)$$

for $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$ and $f \in \mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))$ for $s > 2(1-\frac{d}{q})^{-1}$ with $q > d$. The critical result for quasilinear equations will be that the solutions to (2.2) are even uniformly bounded in the Hölder space with respect to $\mu_\bullet, \mu^\bullet$ and f from bounded sets in $\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))$. Even for merely measurable, coercive and bounded coefficient functions μ, Hölder regularity under the given assumptions on q and s is classical ever since the monograph [101] of Ladyzhenskaya, Solonnikov and Ural'tseva, at least as long as there are no mixed boundary conditions involved.

Unfortunately, the investigations in [101] contain—in their generality—some peculiarities which make it not easy to apply them to problems originating from modern applications: First, the Hölder spaces under consideration, see [101, P. 7], are *not* the classical ones: the oscillation of the function is only measured over the connected components of the intersection of the domain with suitable balls (what is indeed adequate in case of *general* Dirichlet boundary data). Secondly, the estimates affect distributional right hand sides f which are represented as the (spatial) divergence of vector-valued L^q-functions $\mathfrak{f}$, i.e., $f = \mathfrak{f}_0 + \nabla\cdot\mathfrak{f}$ (see also Lemma 2.1.15 below). As is well-known, such representations are highly non-unique; in particular the zero-functional may be represented as the divergence of a non-zero vector valued function. Lastly, it is not quite clear how broad the admissible geometric setting *really* is: on one hand "piecewise C^1" is demanded, on the other the crucial "Condition A" ([101, P. 9], compare also [97, Ch. II.B, Definition B.3])—well-known from elliptic theory—comes into play.

Our intention is to deliver a treatment which

- uses a clearly defined underlying geometric concept for the domain Ω, thereby avoiding "Condition A",
- incorporates mixed boundary conditions within an appropriately defined framework in a setting with weak boundary regularity requirements,

- allows for right hand sides from $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$, and
- gives a result in the formulation of *classical* Hölder spaces.

The chapter is organized as follows: As a starting point, we quote the classical result on the existence and uniqueness of solutions for nonautonomous parabolic equations in a Hilbert spaces setting which serves as the origin of the solutions for which we later show Hölder regularity. Afterwards the main result is announced, cf. Theorem 2.1.4, and the rest of the chapter is mostly preoccupied with the proof of this theorem, starting with the quotation of several classical results from [101]. The idea is to take these in a setting—the half cube and the ball and *homogeneous* Dirichlet boundary conditions—where the inherent technical difficulties are still not appearing: here it is clear that the general suppositions on the domain posed in [101] are fulfilled and the results fall back to the classical Hölder spaces. After establishing some preliminaries in Ch. 2.1.2, we establish permanence principles such as localization, bi-Lipschitz transformations and reflection in Ch. 2.1.3 which allow us to treat suitable sub-problems in the geometric setting established above. Now employing the classical results of Ladyzhenskaya et al., we are able to deduce the required Hölder results for the solution of (2.2) in space and time in Ch. 2.1.4. Here the *homogeneous* Dirichlet condition allows us to establish *global* Hölder continuity from the Hölder continuity on the connected components of the intersection of Ω with suitable domains.

Up to this point, the considerations are restricted to initial value zero and, as mentioned before, homogeneous Dirichlet conditions. In Ch. 2.1.5 we deviate from this and admit nonzero initial values together with inhomogeneous Dirichlet data as an add-on to Theorem 2.1.4. Since, for $\Omega \cup D$ generalized regular in the sense of Gröger, D is a $(d-1)$-set (see Theorem 1.3.16), one may apply the restriction– and extension results for Sobolev spaces with partially vanishing trace of Jonsson and Wallin, cf. Proposition 1.2.60. This allows again to prove Hölder regularity for the solution in space and time.

The following assumptions hold true for the rest of this chapter:

(i) The set $\Omega \subset \mathbb{R}^d$ is a bounded domain and D (like *Dirichlet*) is a closed subset of $\partial\Omega$. The cases $D = \emptyset$ and $D = \partial\Omega$ are not excluded. We suppose that $\Omega \cup D$ is *volume-preserving* generalized regular in the sense of Gröger. In all what follows, $\partial\Omega \setminus D$ will be denoted by N (like *Neumann*).

(ii) We consider a finite interval $J = (T_0, T_1) \subset \mathbb{R}_0^+$.

(iii) All Banach spaces and all occurring functions are supposed to be real ones, i.e., we are working in a *real* setting in the sense of Chapter 1.6.

Note that we assume the additional property of *volume-preserving* generalized regular in the sense of Gröger for Chapter 2.1. This is for mainly technical reasons. It should, in principle, be able to get rid of it by working in spaces with spatial weights, but this has not been worked out yet.

In order to establish the frame in which our main result can be formulated, we quote the following classical result from Dautray and J.-L. Lions, cf. [48, Ch. XVIII §3 and Ch. XVIII §4.2]. Essentially, it is a maximal nonautonomous parabolic regularity result for Hilbert spaces.

Proposition 2.1.1. *Suppose that $V \hookrightarrow H \hookrightarrow V'$ is a Gelfand triplet of* real *Hilbert spaces with dense embeddings. Let $\{\mathfrak{a}_t\}_{t\in J}$ be a family of bilinear forms on V the norms of which are uniformly bounded and such that each $\mathfrak{a}_t$ is coercive with a coercivity constant $\varkappa$, also uniformly in $t \in J$. Suppose that the mapping $J \ni t \mapsto \mathfrak{a}_t(\psi, \varphi)$ is measurable for all $\psi, \varphi \in V$. Then, for any $f \in \mathrm{L}^2(J; V')$, there is a unique $u = u_f \in \mathbb{W}_0^{1,2}(J; V', V)$ such that*

$$\langle u'(t), \psi\rangle_V + \mathfrak{a}_t(u(t), \psi) = \langle f(t), \psi\rangle_V \quad \text{for all } \psi \in V \text{ for a.a. } t \in J. \tag{2.4}$$

Moreover, u admits the following estimates:

$$\|u\|_{\mathrm{L}^2(J;V)} \le \frac{1}{\varkappa}\|f\|_{\mathrm{L}^2(J;V')}, \quad \|u\|_{\mathrm{C}(\overline{J};H)} \le \sqrt{\frac{1}{\varkappa}}\|f\|_{\mathrm{L}^2(J;V')}.$$

Thus, the mapping which assigns to the right hand side $f \in \mathrm{L}^2(J;V')$ *the solution* u *of* (2.4) *with initial value* $u(T_0) = 0$ *is well-defined and continuous from* $\mathrm{L}^2(J;V')$ *into* $\mathrm{L}^2(J;V) \cap \mathrm{C}(\overline{J};H)$*, and its norm is not larger than* $\frac{1}{\varkappa} + \sqrt{\frac{1}{\varkappa}}$.

Remark 2.1.2. Defining, for $t \in J$, the operator $A(t)\colon V \to V'$ by

$$\langle A(t)w, \psi\rangle_V := \mathfrak{a}_t(w,\psi), \quad \text{for all } w,\psi \in V,$$

equation (2.4) reads as

$$\langle u'(t), \psi\rangle_V + \langle A(t)u(t), \psi\rangle_V = \langle f(t), \psi\rangle_V \qquad \text{for all } \psi \in V \text{ for a.a. } t \in J \tag{2.5}$$

or

$$u'(t) + A(t)u(t) = f(t) \quad \text{in } V' \quad \text{for a.a. } t \in J.$$

In the latter, we have re-obtained the familiar form for nonautonomous evolution equations as in Chapter 1.4 and Proposition 2.1.1 exactly states that the operator A satisfies nonautonomous maximal L^2 regularity on V over J (recall Lemma 1.4.15). Albeit we are talking about "only" a Hilbert space V, the result is astonishing in that it requires *no regularity* in addition to measurability of the mapping $t \mapsto A(t)$.

Of course, aiming at a result for (2.2), the form $\mathfrak{a}_t$ will be of type

$$\mathrm{W}^{1,2}_D(\Omega)\times\mathrm{W}^{1,2}_D(\Omega) \ni (\psi,\varphi) \mapsto \mathfrak{a}_t(\psi,\varphi) := \int_\Lambda \mu(t,\cdot)\nabla\psi\cdot\nabla\varphi + \psi\varphi \,\mathrm{d}x \tag{2.6}$$

Clearly, the resulting operator $A(t)$ as in Remark 2.1.2 is then the corresponding divergence operator $-\nabla\cdot\mu(t,\cdot)\nabla + 1$ on $\mathrm{W}^{1,2}_D(\Omega)$ and we call these operators $\mathcal{A}_\mu$ in the following, with the usual understanding of the polarity between a time-dependent "spatial" operator and an operator acting on a function which itself is time-dependent, i.e.,

$$J \ni t \mapsto \mathcal{A}_\mu(t) = -\nabla \cdot \mu(t,\cdot)\nabla + 1 \in \mathscr{L}(\mathrm{W}^{1,2}_D(\Omega); \mathrm{W}^{-1,2}_D(\Omega)) \tag{2.7}$$

on the one hand, and

$$\mathcal{A}_\mu \colon \mathrm{L}^2(J; \mathrm{W}^{1,2}_D(\Omega)) \to \mathrm{L}^2(J; \mathrm{W}^{-1,2}_D(\Omega)), \quad (\mathcal{A}_\mu u)(t) = \mathcal{A}_\mu(t)u(t) \tag{2.8}$$

on the other.

Remark 2.1.3. Let us once more point out that the following considerations may also be carried out for the operators of the form $-\nabla \cdot \mu(t,\cdot)\nabla$ alone, i.e., without "+1", if $D \neq \emptyset$, cf. Remark 1.5.8.

The subsequent theorem contains the main result of this chapter.

Theorem 2.1.4. *Let $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$ for some $0 < \mu_\bullet < \mu^\bullet$ and let $q > d$ and $s > 2(1-\frac{d}{q})^{-1}$ be fixed and suppose that $f \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$. Then the solution $u = u_f \in \mathbb{W}^{1,2}_0(J; \mathrm{W}^{-1,2}_D(\Omega), \mathrm{W}^{1,2}_D(\Omega))$ of the equation*

$$u'(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) + u(t) = f(t) \quad \text{in } \mathrm{W}^{-1,2}_D(\Omega) \quad \text{on } J \tag{2.9}$$

in the sense of Proposition 2.1.1 or Remark 2.1.2 exists and is unique. Moreover, let $\mathbb{B}_{s,q}(0)$ denote the unit ball in $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$. Then the following holds true:

(i) The supremum $\sup_{f \in \mathbb{B}_{s,q}(0)} \|u_f\|_{\mathrm{L}^\infty(Q)}$ is finite and uniform in the parameters $\mu_\bullet, \mu^\bullet, q$ and s.

(ii) There is an $\alpha > 0$ such that even $\sup_{f \in \mathbb{B}_{s,q}(0)} \|u_f\|_{\mathrm{C}^\alpha(Q)}$ is finite and uniform in $\mu_\bullet, \mu^\bullet, q$ and s. In other words: Let $(\partial + \mathcal{A}_\mu)^{-1}$ denote the linear operator which assigns to the right-hand side of the parabolic equation in (2.9) the solution $u = u_f$ with initial value $u(T_0) = 0$.

Then the mapping

$$(\partial + \mathcal{A}_\mu)^{-1} : \mathrm{L}^s(J; \mathrm{W}_D^{-1,q}(\Omega)) \to \mathrm{C}^\alpha(J \times \Omega) \tag{2.10}$$

is well-defined and continuous for some α. For fixed $\mu_\bullet, \mu^\bullet$, the mappings (2.10) *are* equicontinuous *for all coefficient functions $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$.*

Remark 2.1.5.

(i) It is straight-forward to check that for $q, s \geq 2$,

$$\mathrm{L}^s(J; \mathrm{W}_D^{-1,q}(\Omega)) \hookrightarrow \mathrm{L}^2(J; \mathrm{W}_D^{-1,2}(\Omega))$$

with embedding constant $\lambda^d(\Omega)^{\frac{q-2}{2q}} \lambda(J)^{\frac{s-2}{2s}}$. In this sense, right-hand sides f from $\mathrm{L}^s(J; \mathrm{W}_D^{-1,q}(\Omega))$ in equation (2.9) are implicitly always to be understood as right-hand sides from $\mathrm{L}^2(J; \mathrm{W}_D^{-1,2}(\Omega))$ without further comment in the sequel.

(ii) It is worthwhile to compare the choice of q and s in Theorem 2.1.4 with the requirement on the corresponding quantities for the maximal regularity embedding into Hölder spaces as in Lemma 1.5.25.

(iii) From continuity of the mapping (2.10) it follows immediately that the domain of $\partial + \mathcal{A}_\mu$ in $\mathrm{L}^s(J; \mathrm{W}_D^{-1,q}(\Omega))$ embeds continuously into $\mathrm{C}^\alpha(Q)$.

The main result in Theorem 2.1.4 is formulated for homogeneous Dirichlet- and possibly inhomogeneous Neumann-data. We however obtain the following extension to Robin boundary conditions simply by "hijacking" the previous result and observing that the boundary form is subordinated to the total differential operator in the same way as it was to the divergence-gradient operator, cf. Lemma 1.5.14 and Remark 1.5.15. Analogously to the definition of $\mathcal{A}_\mu$, we consider the boundary operator $\mathcal{B}_\gamma$ as the "time-extension" of the operators B_γ introduced in Definition 1.5.11, that is,

$$J \ni t \mapsto \mathcal{B}_\gamma(t) := \mathsf{B}_{\gamma(t,\cdot)} \in \mathscr{L}(\mathrm{W}_D^{1,2}(\Omega); \mathrm{W}_D^{-1,2}(\Omega)) \tag{2.11}$$

and

$$\mathcal{B}_\gamma \colon \mathrm{L}^2(J;\mathrm{W}^{1,2}_D(\Omega)) \to \mathrm{L}^2(J;\mathrm{W}^{-1,2}_D(\Omega)), \quad (\mathcal{B}_\gamma u)(t) = \mathcal{B}_\gamma(t)u(t) \tag{2.12}$$

for a function $\gamma \in \mathrm{L}^\infty(J \times \overline{N}; \lambda \otimes \omega)$. We again use $\mathcal{B} := \mathcal{B}_1$ for $\gamma \equiv 1$, the constant function in time and space.

Corollary 2.1.6. *Adopt the assumptions of Theorem 2.1.4 and suppose that $\gamma \in \mathrm{L}^\infty(J \times \overline{N}; \lambda \otimes \omega, \mathbb{R}_0^+)$. Then the results of Theorem 2.1.4 still hold true for the equation*

$$u'(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) + u(t) + \mathsf{B}_{\gamma(t,\cdot)}u(t) = f(t) \quad \text{in } \mathrm{W}^{-1,2}_D(\Omega) \quad \text{on } J \tag{2.13}$$

and the operators $\partial + \mathcal{A}_\mu + \mathcal{B}_\gamma$. That is, there exists a unique solution $u = u_f$ from $\mathbb{W}^{1,2}_0(J;\mathrm{W}^{-1,2}_D(\Omega),\mathrm{W}^{1,2}_D(\Omega))$ which satisfies (2.13) for almost every $t \in J$ and the mapping

$$(\partial + \mathcal{A}_\mu + \mathcal{B}_\gamma)^{-1} \colon \mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega)) \to \mathrm{C}^\alpha(Q),$$

assigning $f \mapsto u_f$, is again equicontinuous with respect to all coefficient functions $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$.

Proof. It is clear that the form

$$\begin{aligned} \mathrm{W}^{1,2}_D(\Omega) \times \mathrm{W}^{1,2}_D(\Omega) \ni (\psi,\varphi) \mapsto \mathfrak{b}_t(\psi,\varphi) \\ := \mathfrak{a}_t(\psi,\varphi) + \int_N \gamma(t) \operatorname{tr}\psi \operatorname{tr}\varphi \,\mathrm{d}\omega \end{aligned}$$

with $\mathfrak{a}$ as in (2.6) still satisfies the assumptions of Proposition 2.1.1, such that (2.13) indeed admits a unique solution $u = u_f$ in the maximal regularity space for every $f \in \mathrm{L}^2(J;\mathrm{W}^{-1,2}_D(\Omega))$ (recall the condition $\int_N \gamma(t)\,\mathrm{d}\omega > 0$ if $D = \emptyset$, cf. Remark 1.5.8). To show the remaining assertions, we prove that the domains of $\partial + \mathcal{A}_\mu$ and $\partial + \mathcal{A}_\mu + \mathcal{B}_\gamma$ in $\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))$ agree, from which the assertions follow. This we do by showing $\mathcal{B}_\gamma$ is compact relative to $\partial + \mathcal{A}_\mu$, analogously to the reasoning in Lemma 1.5.14.

The mapping $\mathcal{B}_\gamma$ certainly maps $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})) \doteq \mathrm{C}(\overline{Q})$ continuously into $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$. Due to the Arzelà-Ascoli Theorem 1.2.5, we have $\mathrm{C}^\alpha(Q) \hookrightarrow \mathrm{C}(\overline{Q})$, cf. Corollary 1.2.6. But then the continuous embedding of the domain of $\partial + \mathcal{A}_\mu$ in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$ into $\mathrm{C}^\alpha(J \times \Omega)$ implies that $\mathcal{B}_\gamma$ is a compact operator from that domain into $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$, and from this we infer that $\mathcal{B}_\gamma$ is compact relative to $\partial + \mathcal{A}_\mu$. Now [96, Ch. IV.1.3] shows that the domains of $\partial + \mathcal{A}_\mu$ and $\partial + \mathcal{A}_\mu + \mathcal{B}_\gamma$ in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$ agree. □

Let us give the proof of Theorem 2.1.4. We first collect some classical results of Ladyzhenskaya, Solonnikov and Ural'tseva [101] adopted for our cause. The basis of our considerations will be Corollaries 2.1.11 and 2.1.13 which are based on space-time local estimates for so-called generalized solutions of corresponding equations in [101, Ch. III]. However, in order to use those, we invest quite some work and introduce a non-trivial localization-procedure for (2.9) which allows to transform the localized equation onto a very regular object, namely the lower half-cubes τK^- and (via reflection) the full cubes τK in such a way that the resulting equation still provides a generalized equation in the sense of Ladyzhenskaya.

2.1.1 Classical results

We begin by introducing the notion of a generalized equation, which is essentially a very weak formulation for the underlying evolution equation. The crucial link to the concept of Lions is the space $\mathrm{V}_2^{1,0}(J\times\Lambda)$ introduced in the next definition, which corresponds to the spaces $\mathrm{L}^2(J;V)\cap\mathrm{C}(\overline{J};H)$ in Proposition 2.1.1, there choosing $V = \mathrm{W}^{1,2}(\Lambda)$ and $H = \mathrm{L}^2(\Lambda)$.

Definition 2.1.7 (Generalized solution)**.** Let $\Lambda \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $\sigma \in \mathrm{L}^\infty(J \times \Lambda;\mathbb{M}_d(\sigma_\bullet))$ for some $\sigma_\bullet > 0$.

(i) Let $\mathrm{V}_2^{1,0}(J\times\Lambda)$ be the space $\mathrm{L}^2(J;\mathrm{W}^{1,2}(\Lambda))\cap\mathrm{C}(\overline{J};\mathrm{L}^2(\Lambda))$, equipped with the norm

$$v\mapsto\sup_{t\in\overline{J}}\|v(t,\cdot)\|_{\mathrm{L}^2(\Lambda)}+\Bigl(\int_J\int_\Lambda v(t,\mathrm{x})^2+\|\nabla v(t,\mathrm{x})\|_2^2\,\mathrm{d}\mathrm{x}\,\mathrm{d}t\Bigr)^{1/2}.$$

(ii) Suppose that $\mathfrak{f}=(\mathfrak{f}_0,\mathfrak{f}_1,\dots,\mathfrak{f}_n)\in\mathrm{L}^2(J;\mathrm{L}^2(\Lambda;\mathbb{R}^{d+1})\simeq\mathrm{L}^2(J\times\Lambda;\mathbb{R}^{d+1})$. We say that a function $u\in\mathrm{V}_2^{1,0}(J\times\Lambda)$ is a *generalized solution* of the equation

$$u'-\sum_{i,j=1}^d\frac{\partial}{\partial\mathrm{x}_i}\Bigl(\sigma_{ij}\frac{\partial u}{\partial\mathrm{x}_j}\Bigr)+u=\sum_{k=1}^d\frac{\partial\mathfrak{f}_k}{\partial\mathrm{x}_k}+\mathfrak{f}_0,\tag{2.14}$$

if for every $\vartheta\in\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J\times\Lambda)$ the integral identity

$$\begin{aligned}0=\int_\Lambda u(\tau,\mathrm{x})\vartheta(\tau,\mathrm{x})\,\mathrm{d}\mathrm{x}-\int_{T_0}^\tau\int_\Lambda u\frac{\partial\vartheta}{\partial t}\,\mathrm{d}\mathrm{x}\,\mathrm{d}t\\+\int_{T_0}^\tau\int_\Lambda\sum_{i,j=1}^d\sigma_{ij}\frac{\partial u}{\partial\mathrm{x}_j}\frac{\partial\vartheta}{\partial\mathrm{x}_i}\,\mathrm{d}\mathrm{x}\,\mathrm{d}t+\int_{T_0}^\tau\int_\Lambda u\vartheta\,\mathrm{d}\mathrm{x}\,\mathrm{d}t\\-\int_{T_0}^\tau\int_\Lambda\mathfrak{f}_0\vartheta+\sum_{k=1}^d\mathfrak{f}_k\frac{\partial\vartheta}{\partial\mathrm{x}_k}\,\mathrm{d}\mathrm{x}\,\mathrm{d}t\end{aligned}\tag{2.15}$$

holds true for all $\tau\in\overline{J}$.

Remark 2.1.8.

(i) We refer to [101, Ch. III §1] for the validation that, given $\mathfrak{f}\in\mathrm{L}^2(J;\mathrm{L}^2(\Lambda;\mathbb{R}^{d+1}))$, the (very) weak formulation (2.15) is well-posed for $\vartheta\in\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J\times\Lambda)$ and $u\in\mathrm{V}_2^{1,0}(J\times\Lambda)$ for every $\tau\in\overline{J}$. Moreover, it is shown there that, given such u and τ, the right-hand side gives rise to a continuous linear form on $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J\times\Lambda)$.

(ii) Assume we have a generalized solution u of (2.14) at hand. Then, *formally* integrating by parts, we find

$$\int_{T_0}^{\tau}\int_\Lambda u\frac{\partial\vartheta}{\partial t}\,\mathrm{dx}\,\mathrm{d}t = -\int_{T_0}^{\tau}\int_\Lambda \frac{\partial u}{\partial t}\vartheta\,\mathrm{dx}\,\mathrm{d}t \\ + \int_\Lambda u(\tau,\mathrm{x})\vartheta(\tau,\mathrm{x})\,\mathrm{dx} - \int_\Lambda u(T_0,\mathrm{x})\vartheta(T_0,\mathrm{x})\,\mathrm{dx}.$$

If we further assume that u is in fact a “usual” weak solution to (2.14), i.e., as exemplary in Proposition 2.1.1, and substitute the corresponding Λ-integrated identities for the Λ-integral with the time derivative of u and finally plug everything into (2.15), then all terms cancel out (as it should be!), except for the integral involving $u(T_0,\cdot)$. Thus, if test functions ϑ are admitted which are nonzero on $\{T_0\}\times\Lambda$, such as those from $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J\times\Lambda)$, this enforces $u(T_0,\cdot)$ to be the zero function, i.e., a generalized solution as in Definition 2.1.7 is always to be considered as having initial value zero – on a *formal* level, but more is not to expected at this point without further knowledge.

The next results are in their essence space-time local estimates for generalized solutions if the right-hand side in (2.14) is regular enough. However, for initial value 0 we may re-obtain the estimates for the *whole* time interval J, see Corollaries 2.1.11 and 2.1.13. We begin with L^∞ bounds:

Proposition 2.1.9 ([101, Ch. III, Thm. 8.1])**.** *Let $\Lambda\subset\mathbb{R}^d$ be a bounded Lipschitz domain and let $\sigma\in\mathrm{L}^\infty(J\times\Lambda;\mathbb{M}_d(\sigma_\bullet,\sigma^\bullet))$ for some $0<\sigma_\bullet<\sigma^\bullet$. Fix $q>d$ and $s>2(1-\frac{d}{q})^{-1}$. Let the set $\mathcal{F}$ be given such that*

$$\mathcal{F}\subseteq\Big\{\mathfrak{f}\in\mathrm{L}^s(J;\mathrm{L}^q(\Lambda;\mathbb{R}^{d+1})):\|\mathfrak{f}\|_{\mathrm{L}^s(J;\mathrm{L}^q(\Lambda;\mathbb{R}^{d+1}))}\le C\Big\}, \tag{2.16}$$

for some $C\ge 0$. Moreover, assume that for every $\mathfrak{f}\in\mathcal{F}$ a generalized solution $u=u_\mathfrak{f}$ of (2.14) exists and that $\{u_\mathfrak{f}:\mathfrak{f}\in\mathcal{F}\}$ is contained in a ball around 0 in $\mathrm{V}_2^{1,0}(J\times\Lambda)$ with radius r_V.

(i) Let $\Lambda_0\subset\Lambda$ be a subdomain which has a positive distance $\mathfrak{d}<T_1-T_0$ to $\partial\Lambda$. Then $\sup_{\mathfrak{f}\in\mathcal{F}}\|u_\mathfrak{f}\|_{\mathrm{L}^\infty((T_0+\mathfrak{d},T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_V, $\mathfrak{d}$, q, s and C.

(ii) *Let F be a closed part of $\partial\Lambda$ and let all $u_{\mathfrak{f}}$ belong to the space* $\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda))$. *If a subdomain Λ_0 of Λ has a positive distance $\mathfrak{d} < T_1 - T_0$ to $\partial\Lambda \setminus F$, then also $\sup_{\mathfrak{f}\in\mathcal{F}} \|u_{\mathfrak{f}}\|_{\mathrm{L}^\infty((T_0+\mathfrak{d},T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_{V}, $\mathfrak{d}$, q, s and C.*

Since the assumptions in [101, Ch. III, Thm. 8.1] rely on boundary values of functions in $\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda))$, a comment is in order.

Remark 2.1.10. As Λ in Proposition 2.1.9 is supposed to be a Lipschitz domain, we already know that we have trace operator as in Lemma 1.2.57 at hand, cf. Corollary 1.3.8. In particular, tr gives rise to a continuous operator from $\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda)))$ to $\mathrm{L}^2(J;\mathrm{L}^2(F;\omega)) \doteq \mathrm{L}^2(J\times F;\lambda\otimes\omega)$ and for almost all $t\in J$, a function $f\in\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda))$ satisfies $\operatorname{tr} f(t,\cdot) = 0$ in the ω-almost everywhere sense on F (recall that ω coincides with $\mathcal{H}^{d-1} \upharpoonright F$). The zero values on F follow from denseness of $\mathrm{C}^\infty_F(\Lambda)$ in $\mathrm{W}^{1,2}_F(\Lambda)$ and continuity of the trace operator. In fact, we will employ the results of Ladyzhenskaya et al. only for F being a $(d-1)$-set, and for these we even know that the strictly defined representative of $f \upharpoonright F$ satisfies $\mathcal{R}^F_\Lambda f \equiv 0$, again ω-a.e., for all $f\in\mathrm{W}^{1,2}_F(\Lambda)$ by Theorem 1.2.42.

As announced above, if the generalized solution u in fact has initial value $u(T_0) = 0$, then we are able to upgrade to a *global* estimate in time, i.e., for the whole given time interval J.

Corollary 2.1.11. *Suppose the general conditions of Proposition 2.1.9 and assume that the generalized solution u satisfies $u(T_0) = 0$.*

(i) *Let $\Lambda_0 \subset \Lambda$ be a subdomain which has a positive distance $\mathfrak{d}$ to $\partial\Lambda$. Then for the generalized solutions $u_{\mathfrak{f}}$, the supremum $\sup_{\mathfrak{f}\in\mathcal{F}} \|u_{\mathfrak{f}}\|_{\mathrm{L}^\infty((T_0,T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_{V}, $\mathfrak{d}$, q, s and C.*

(ii) *Let F be a closed part of $\partial\Lambda$ and let all $u_{\mathfrak{f}}$ belong to the space* $\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda))$. *If a subdomain Λ_0 has a positive distance $\mathfrak{d}$ to $\partial\Lambda\setminus F$ then $\sup_{\mathfrak{f}\in\mathcal{F}} \|u_{\mathfrak{f}}\|_{\mathrm{L}^\infty((T_0,T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_{V}, $\mathfrak{d}$, q, s and C.*

Proof. One associates to the problem (2.14) another one on the prolonged interval $J_0 := (T_0 - \mathfrak{d} - 1, T_1)$ in the following manner: one defines a coefficient function $\check{\sigma}$ on $J_0 \times \Lambda$ by

$$\check{\sigma}(t, \mathrm{x}) = \begin{cases} \frac{\sigma_\bullet + \sigma^\bullet}{2} \, \mathrm{id}_d & \text{if} \quad t \in J_0 \setminus J, \\ \sigma(t, \mathrm{x}) & \text{else.} \end{cases}$$

Moreover, one defines a new right-hand side $\check{\mathfrak{f}}$ as 0 on $J_0 \setminus J$ and as $\mathfrak{f}$ on J and finds the solution $\check{u}$ on $J_0 \times \Lambda$ with $u(T_0 - \mathfrak{d} - 1) = 0$. This solution $\check{u}$ is zero on $(J_0 \setminus J) \times \Lambda$ and coincides with u on $J \times \Lambda$. Applying Proposition 2.1.9 (i) to the function $\check{u}$ one gets (i). Point (ii) is deduced analogously from (ii) of the foregoing Proposition 2.1.9. $\square$

If one already has the information that the generalized solutions are essentially bounded, then one every obtains Hölder continuity and even uniform boundedness in the Hölder space:

Proposition 2.1.12 ([101, Ch. III, Thm. 10.1])**.** *Let $\Lambda \subset \mathbb{R}^d$ be a bounded,* convex *domain (and, hence, a Lipschitz domain), and suppose $\sigma \in \mathrm{L}^\infty(J \times \Lambda; \mathbb{M}_d(\sigma_\bullet, \sigma^\bullet))$ for some $0 < \sigma_\bullet < \sigma^\bullet$. Fix $q > d$ and $s > 2(1 - \frac{d}{q})^{-1}$. Assume that $\mathcal{F}$ is again a subset of the set in* (2.16)*, such that for every $\mathfrak{f} \in \mathcal{F}$ a generalized solution $u = u_\mathfrak{f}$ of* (2.14) *exists and that this set of generalized solutions is contained in a ball around* 0 *in $\mathrm{L}^\infty(J \times \Lambda)$ with radius r_∞. Then there is an $\alpha > 0$ such that the following is true:*

- *(i) For every subdomain $\Lambda_0 \subset \Lambda$ having a positive distance $0 < \mathfrak{d} < T_1 - T_0$ to the boundary $\partial\Lambda$, $\sup_{\mathfrak{f} \in \mathcal{F}} \|u_\mathfrak{f}\|_{\mathrm{C}^\alpha((T_0 + \mathfrak{d}, T_1) \times \Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_∞, $\mathfrak{d}$, q, s and C.*
- *(ii) Let F be a closed part of $\partial\Lambda$ and suppose that all $u_\mathfrak{f}$ belong to the space $\mathrm{L}^2(J; \mathrm{W}^{1,2}_F(\Lambda))$. If a subdomain Λ_0 of Λ has a positive distance $0 < \mathfrak{d} < T_1 - T_0$ to $\partial\Lambda \setminus F$, then the supremum $\sup_{\mathfrak{f} \in \mathcal{F}} \|u_\mathfrak{f}\|_{\mathrm{C}^\alpha((T_0 + \mathfrak{d}, T_1) \times \Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_∞, $\mathfrak{d}$, q, s and C.*

Again, initial value 0 allows to get rid of the restriction to a local time interval:

Corollary 2.1.13. *Suppose the assumptions of Proposition 2.1.12 to hold and assume, additionally, that the initial value $u(T_0)$ of the solution is zero. Then there is an $\alpha > 0$ such that the following is true:*

(i) *For every subdomain $\Lambda_0 \subset \Lambda$ having a positive distance $\mathfrak{d}$ to the boundary $\partial\Lambda$, $\sup_{\mathfrak{f}\in\mathcal{F}} \|u_\mathfrak{f}\|_{\mathrm{C}^\alpha((T_0,T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_∞, $\mathfrak{d}$, q, s and C.*

(ii) *Let F be a closed part of $\partial\Lambda$ and suppose that each $u_\mathfrak{f}$ belongs to the space $\mathrm{L}^2(J;\mathrm{W}^{1,2}_F(\Lambda))$. Then, for any subdomain Λ_0 with a positive distance $\mathfrak{d}$ to $\partial\Lambda \setminus F$, $\sup_{\mathfrak{f}\in\mathcal{F}} \|u_\mathfrak{f}\|_{\mathrm{C}^\alpha((T_0,T_1)\times\Lambda_0)}$ is finite and uniform in d, $\sigma_\bullet$, $\sigma^\bullet$, r_∞, $\mathfrak{d}$, q, s and C.*

The proof works analogously to the one of Corollary 2.1.11.

Remark 2.1.14. In fact, the quoted result in Proposition 2.1.12 holds for much more general domains as convex ones. However, we have good reasons to restrict ourselves to this case:

- If Λ is convex and $B \subset \mathbb{R}^d$ is a ball, then $\Lambda \cap B$ is still convex and therefore always consists of only one component. Thus, one may deal with the classical notion of Hölder continuity – and not of the much more sophisticated one in [101, Ch. I]
- Secondly, if Λ is convex, then every point $\mathrm{x} \in \partial\Lambda$ admits a supporting hyperplane such that Λ lies on one side of this hyperplane. Thus, for any ball $B \subset \mathbb{R}^d$ with center x, the intersection $\Lambda \cap B$ has at most half the measure of B, what makes the crucial "Condition A" ([101, Ch. 1, P.9]) obviously fulfilled in our context, with the constant $\theta_0 = \frac{1}{2}$ – universal for all convex domains and all balls.
- We will need the result only in case of balls, cubes and half cubes, serving as our local model sets.

Next, we establish the link between generalized solutions and solutions in the sense of Proposition 2.1.1. For doing so, we restrict ourselves to the case of right hand sides from $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$ which are *step functions* in time only (these being dense in the whole space). The reason is as follows: By a classical theorem, the elements f from $\mathrm{W}^{-1,q}(\Omega)$ may be represented as the sum of the weak divergence of a $\mathbb{R}^d$-valued function $\mathfrak{f} \in \mathrm{L}^q$ and $\mathfrak{f}$ itself (see for instance [112, Ch. 1.1.14]). From there, we obtain the same representation for elements of $\mathrm{W}_D^{-1,q}(\Omega)$ since every continuous functional on $\mathrm{W}_D^{1,q'}(\Omega)$ may be continuously extended to $\mathrm{W}^{1,q'}(\Omega)$ by the Hahn-Banach theorem. The problem is now that this representation is highly non-unique and, the worse, not even obviously linear, which makes it a rather bad choice to use for (almost) every $t \in J$. So we prefer to restrict ourselves to step functions and to use the corresponding representation theorem separately on any of the "constancy intervals" only. Since the step functions with values in $\mathrm{W}_D^{-1,q}(\Omega)$ are dense in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$, this will pose no problem later on.

The representation result, whose proof we just have sketched, is as follows, where the norm bounds are also to be found in [112, Ch. 1.1.14].

Lemma 2.1.15. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $F \subset \partial\Lambda$ be closed and $q, s \geq 2$. Assume that we are given a* step function *$f \in \mathrm{L}^s(J;\mathrm{W}_F^{-1,q}(\Lambda))$, i.e., there exists a partition (J_k) of J and functions $f_k \in \mathrm{W}_F^{-1,q}(\Lambda)$ such that*

$$f(t) = \sum_k \chi_k(t) f_k \quad \text{for almost all } t \in J,$$

where χ_k is the indicator function of the interval J_k. Then, for every k, there is a collection of functions $\mathfrak{f}_k = (\mathfrak{f}_{k,0}, \dots, \mathfrak{f}_{k,d}) \in \mathrm{L}^q(\Lambda;\mathbb{R}^{d+1})$ such that f_k is represented by

$$\langle f_k, \varphi \rangle_{\mathrm{W}_F^{1,q'}(\Lambda)} = \int_\Lambda \mathfrak{f}_{k,0}\varphi - \sum_{j=1}^d \mathfrak{f}_{k,j} \frac{\partial \varphi}{\partial x_j} \,\mathrm{dx} \quad \text{for all } \varphi \in \mathrm{W}_F^{1,q'}(\Lambda).$$

Moreover, the norms of $\mathfrak{f}_k$ and $\mathfrak{f} = \sum_k \chi_{J_k} \mathfrak{f}_k$ are bounded by

$$\|\mathfrak{f}_k\|_{\mathrm{L}^q(\Lambda;\mathbb{R}^{d+1})} \leq 2\|f_k\|_{\mathrm{W}_F^{-1,q}(\Lambda)}$$

and

$$\|\mathfrak{f}\|_{\mathrm{L}^s(J;\mathrm{L}^q(\Lambda;\mathbb{R}^{d+1}))} \le 2\|f\|_{\mathrm{L}^s(J;\mathrm{W}_F^{-1,q}(\Lambda))}.$$

for all (step functions) $f \in \mathrm{L}^s(J;\mathrm{W}_F^{-1,q}(\Lambda))$.

Using the so-obtained representation of $f \in \mathrm{L}^s(J;\mathrm{W}_F^{-1,q}(\Lambda))$ compatible with the notion in Definition 2.1.7, we are able to show that a solution of a "usual" weak evolution equation in the spirit of Proposition 2.1.1, involving f, is also a generalized solution in the sense of Ladyzhenskaya et al. when $q, s \ge 2$. Due to the test functions in Definition 2.1.7, we only need the equation to hold in $\mathrm{W}_0^{-1,2}(\Lambda)$ (see also Lemma 2.1.17 below). This will play a crucial role later in Chapter 2.1.4.

Proposition 2.1.16. *Let* $\Lambda \subset \mathbb{R}^d$ *be a bounded Lipschitz domain, let* $F \subset \partial\Lambda$ *be closed and let* $\sigma \in \mathrm{L}^\infty(J \times \Lambda; \mathbb{M}_d(\sigma_\bullet))$ *for some* $\sigma_\bullet > 0$. *Let moreover* $q, s \ge 2$ *and suppose that* $f \in \mathrm{L}^s(J;\mathrm{W}_F^{-1,q}(\Lambda))$ *is a step function. Assume that* $u \in \mathrm{L}^2(J;\mathrm{W}_F^{1,2}(\Lambda)) \cap \mathrm{C}(\overline{J};\mathrm{L}^2(\Lambda))$ *with* $u' \in \mathrm{L}^2(J;\mathrm{W}_0^{-1,2}(\Lambda))$ *and* $u(T_0) = 0$ *is a solution of the equation*

$$u'(t) - \nabla\cdot\sigma(t,\cdot)\nabla u(t) + u(t) = f(t) \quad \text{in } \mathrm{W}_0^{-1,2}(\Lambda) \quad \text{for a.a. } t \in J.$$

Then u *is a generalized solution of the equation* (2.14), *there choosing* $\mathfrak{f} = \sum_k \chi_k \mathfrak{f}_k$ *as the* $\mathrm{L}^s(J;\mathrm{L}^q(\Lambda;\mathbb{R}^{d+1}))$*-representation of* f *as in Lemma 2.1.15.*

For the proof of Proposition 2.1.16, we first collect some auxiliary results which will be of use in the proof. We have ordered them in such a way to subtly hint at the strategy, cf. also Remark 2.1.10.

Lemma 2.1.17. *Let* $\Lambda \subset \mathbb{R}^d$ *be a bounded Lipschitz domain.*

(i) For every Banach space X, *the set*

$$\mathrm{C}_c^\infty(\overline{J}) \otimes X = \left\{\sum_{j=1}^k \eta_j \otimes v_j \colon \eta_j \in \mathrm{C}_c^\infty(\overline{J}),\ v_j \in X,\ k \in \mathbb{N}\right\}$$

is sequentially dense in $\mathrm{C}_c^\infty(\overline{J}; X)$.

(ii) The set $\mathrm{C}_c^\infty(\overline{J}; \mathrm{W}_0^{1,2}(\Lambda))$ *is dense in* $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J \times \Lambda)$.

Proof.

(i) This is proven in the book of Amann [3, Ch. V, Prop. 2.4.1].

(ii) Firstly, it is known that $\mathrm{C}_c^\infty(\overline{J}; \mathrm{W}_0^{1,2}(\Lambda))$ is dense in $\mathrm{L}^2(J; \mathrm{W}_0^{1,2}(\Lambda)) \cap \mathrm{W}^{1,2}(J; \mathrm{L}^2(\Lambda))$ (see [48, Ch. XVIII §2, Lem. 1]). On the other hand, we have

$$\mathrm{W}^{1,2}(J \times \Lambda) \doteq \mathrm{L}^2(J; \mathrm{W}^{1,2}(\Lambda)) \cap \mathrm{W}^{1,2}(J; \mathrm{L}^2(\Lambda)),$$

both algebraically and topologically, cf. e.g. [48, Ch. XVIII §1.3] for a particular case, otherwise argue via density of $\mathrm{C}^\infty(J\times\Lambda)$. Restricting this isomorphism to the set $\mathrm{C}^\infty(J) \otimes \mathrm{C}^\infty_{\partial\Lambda}(\Lambda)$, which is dense in $\mathrm{C}^\infty_{\overline{J}\times\partial\Lambda}(J \times \Lambda)$, one obtains that $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J \times \partial\Lambda)$ is isomorphic to $\mathrm{L}^2(J; \mathrm{W}_0^{1,2}(\Lambda)) \cap \mathrm{W}^{1,2}(J; \mathrm{L}^2(\Lambda))$. □

It follows the proof of Proposition 2.1.16.

Proof of Proposition 2.1.16. By the suppositions, equation (2.4) can be written as

$$\begin{aligned}
&\langle u'(t), v\rangle_{\mathrm{W}_0^{1,2}(\Lambda)} - \langle \nabla\cdot\sigma(t,\cdot)\nabla u(t) + u(t), v\rangle_{\mathrm{W}_0^{1,2}(\Lambda)} \\
&\qquad = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Lambda u(t)v\,\mathrm{dx} + \int_\Lambda \sigma(t,\cdot)\nabla u(t)\cdot\nabla v + u(t)v\,\mathrm{dx} \\
&\qquad = \langle f(t), v\rangle_{\mathrm{W}_0^{1,2}(\Lambda)} = \int_\Lambda \mathfrak{g}_{k,0}(t)v - \sum_{j=1}^d \mathfrak{g}_{k,j}(t)\frac{\partial v}{\partial x_j}\,\mathrm{dx}, \qquad (2.17)
\end{aligned}$$

for all $v \in \mathrm{W}_0^{1,2}(\Lambda) \hookrightarrow \mathrm{W}_0^{1,q'}(\Lambda)$ and then for almost all $t \in J_k$ (see Remarks 2.1.2 and 2.1.5). Note that we may interpret $-\nabla\cdot\sigma(t,\cdot)\nabla u(t)$ as a functional on $\mathrm{W}_0^{1,2}(\Lambda)$ by restricting its actual domain $\mathrm{W}_F^{1,2}(\Lambda)$ to $\mathrm{W}_0^{1,2}(\Lambda) \subset \mathrm{W}_F^{1,2}(\Lambda)$ (cf. Definition 1.5.3), and the same for f. For the reformulation of the distributional time derivative, see [48, Ch. XVIII,

§1.2 Prop. 7] – here it is crucial that $u'(t)$ gives rise to an element of $\mathrm{W}_0^{-1,2}(\Lambda)$.
Take now any function $\eta \in \mathrm{C}_c^\infty(\overline{J})$, multiply (2.17) with η and integrate from T_0 to $\tau \in J$. Integrating by parts and considering $u(T_0) = 0$, one then obtains

$$
\begin{aligned}
\int_{T_0}^{\tau} \Big(\frac{\mathrm{d}}{\mathrm{d}t} \int_\Lambda uv \,\mathrm{dx}\Big) \eta \,\mathrm{d}t &+ \int_{T_0}^{\tau} \Big(\int_\Lambda \sigma \nabla u \cdot \nabla v + uv \,\mathrm{dx}\Big) \eta \,\mathrm{d}t \\
&= \int_\Lambda u(\tau,\mathrm{x})\, (\eta \otimes v)(\tau,\mathrm{x}) \,\mathrm{dx} - \int_{T_0}^{\tau} \int_\Lambda u \,\frac{\partial(\eta\otimes v)}{\partial t} \,\mathrm{dx}\,\mathrm{d}t \\
&\qquad + \int_{T_0}^{\tau} \int_\Lambda \sigma \nabla u \cdot \nabla(\eta \otimes v) + u(\eta\otimes v) \,\mathrm{dx}\,\mathrm{d}t \\
= \int_{T_0}^{\tau} \int_\Lambda \Big(\sum_k \chi_{J_k} \mathfrak{g}_{k,0}\Big)(\eta \otimes v) &- \int_{T_0}^{\tau} \sum_{j=1}^{d} \Big(\sum_k \chi_{J_k} \mathfrak{g}_{k,j}\Big) \frac{\partial(\eta\otimes v)}{\partial \mathrm{x}_j} \,\mathrm{dx}\ \mathrm{d}t
\end{aligned}
$$

for all $\eta \in \mathrm{C}_c^\infty(\overline{J})$ and all $v \in \mathrm{W}_0^{1,2}(\Lambda)$ and hence for all functions in $\mathrm{C}_c^\infty(\overline{J}) \otimes \mathrm{W}_0^{1,2}(\Lambda)$ by linearity. But we have already seen that $\mathrm{C}_c^\infty(\overline{J}) \otimes \mathrm{W}_0^{1,2}(\Lambda)$ is dense in $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J \times \Lambda)$ in Lemma 2.1.17, and that the very weak formulation used in the definition of a generalized solution depends continuously on the test functions in $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J \times \Lambda)$. Hence, the above identity, which is exactly this very weak formulation for the test function $\eta \otimes v$, extends continuously to the whole $\mathrm{W}^{1,2}_{\overline{J}\times\partial\Lambda}(J \times \Lambda)$. $\square$

Now we have the necessary tools at hand to make use of *both* the existence result of Lions as in Proposition 2.1.1 and the toolbox by Ladyzhenskaya, Solonnikov and Ural'tseva. To use the latter appropriately, we have to introduce some more details.

2.1.2 Preliminaries

One of the main technical ingredients of our proof is a certain localization procedure of the equation (2.9). In contrast to [73] and many following papers it is *not* carried out by multiplying the solution with suitable cutoff functions and afterwards deriving a corresponding equation for the product. We only restrict the function to open subsets of the domain and deduce a corresponding equation for this restriction – in an adequate weak formulation. In fact, this idea was developed in [144] for elliptic problems. The following lemmata allow us in the sequel to perform this procedure in an appropriate manner. The first lemma covers the cases of neighborhoods of interior points of Ω and from the closure of the Neumann boundary (i.e., satisfying case (ii) of Definition 1.3.12).

Lemma 2.1.18. *Let $U \subset \mathbb{R}^d$ be an open neighborhood of a point $\mathrm{x} \in \Omega \cup \overline{N}$. Set $\Lambda := U \cap \Omega$, $S := N \cap U$ and $E := \partial\Lambda \setminus S$ and let $1 \leq p < \infty$.*

(i) The set S is open in $\partial\Lambda$ and E is closed.

(ii) There exists a unique isometric map $\mathfrak{E}_U^0 \colon \mathrm{W}_E^{1,p}(\Lambda) \to \mathrm{W}_D^{1,p}(\Omega)$ such that $\mathfrak{E}_U^0 w$ is the extension of w to Ω by 0 for all $w \in \mathrm{C}_E^\infty(\Lambda)$.

(iii) Set $R := \overline{D \cap U}$. Then $R \subset \partial\Lambda$ and $R \subseteq D \cap E$. Moreover, $u \in \mathrm{W}_D^{1,p}(\Omega)$ implies that $u \restriction \Lambda \in \mathrm{W}_R^{1,p}(\Lambda)$. Thus, the restriction operator from $\mathrm{W}_D^{1,p}(\Omega)$ is a continuous one into $\mathrm{W}_R^{1,p}(\Lambda)$ with norm not larger than 1.

Proof. For (i) and (ii), see [144, Lem. 6.13].
For assertion (iii), observe that $D \cap U \subseteq \partial\Omega \cap U \subseteq \partial\Lambda$. Since $\partial\Lambda$ is closed, this gives $R \subseteq \partial\Lambda$. On the other hand, $R = \overline{D \cap U} \subseteq \overline{D} = D$, since D is closed. From the relations $D \cap U \subseteq \partial\Lambda$ and $U \cap D \cap S = U \cap \big(D \cap (\partial\Omega \setminus D)\big) = \emptyset$ we obtain $D \cap U \subseteq E$ what implies $R \subseteq E$, thanks to the closedness of E. Hence, if $u \in \mathrm{C}_D^\infty(\Omega)$, then the restriction $u \restriction \Lambda$ belongs to $\mathrm{C}_R^\infty(\Lambda)$ with the obvious estimate

$$\|u \restriction \Lambda\|_{\mathrm{W}^{1,p}(\Lambda)} = \|u \restriction \Lambda\|_{\mathrm{W}_R^{1,p}(\Lambda)} \leq \|u\|_{\mathrm{W}_D^{1,p}(\Omega)} = \|u\|_{\mathrm{W}^{1,p}(\Omega)} \qquad \square$$

In case (i) in Definition 1.3.12 the local model set is allowed to be disconnected. Nevertheless, one can also in this case find an adequate localization procedure. In the spirit of the comments right after Definition 1.3.12, this relies on the localization procedure for each of the connected components.

Lemma 2.1.19. *Let $1 \leq p < \infty$. In the terminology of Definition 1.3.12 (i) the following holds true for each $j \in \{1, \dots, k\}$:*

(i) There is an isometric operator $\mathfrak{E}_j^0$ which extends any function from $\mathrm{W}_0^{1,p}(V_j)$ by 0 to a function from $\mathrm{W}_0^{1,p}(\Omega) \subseteq \mathrm{W}_D^{1,p}(\Omega)$.

(ii) We have $\partial V_j \subseteq \partial(U_j \cap \Omega)$.

(iii) Let $R_j = \overline{\partial V_j \cap U_j}$. Then $R_j \subset \partial V_j$ and $u \in \mathrm{W}_D^{1,p}(\Omega)$ implies $u \upharpoonright V_j \in \mathrm{W}_{R_j}^{1,p}(V_j)$.

Proof. (i): The support of every function from $\mathrm{C}_0^\infty(V_j)$ has a positive distance to $\partial\Omega$; thus the extension by zero leads to a function from $\mathrm{C}_0^\infty(\Omega)$ in this case. The general claim follows by density.

(ii) By the definition of V_j it is clear that ∂V_j is contained in $\overline{U_j \cap \Omega}$. Now suppose that a point $\mathrm{y} \in \partial V_j$ lies in $U_j \cap \Omega$ (i.e., not on $\partial(U_j \cap \Omega)$). Since $U_j \cap \Omega$ is open, we find an open ball B containing y which is still a subset of $U_j \cap \Omega$. By supposition, y is a boundary point of V_j, hence $V_j \cap B \neq \emptyset$. Thus, the connectedness of both V_j and B implies that $V_j \cup B \supset V_j$ is also open and connected – and, hence, identical with V_j. But then $B \subset V_j$ which is a contradiction to y being a boundary point of V_j. So indeed $\partial V_j \subseteq \partial(U_j \cap \Omega)$.

(iii) The inclusion $R_j \subset \partial V_j$ is obvious. Let $u \in \mathrm{W}_D^{1,p}(\Omega)$. Repeating the arguments in Lemma 2.1.18 with Λ replaced by V_j and the choice $U = U_j$ shows that indeed $u \upharpoonright V_j \in \mathrm{W}_{R_j}^{1,p}(V_j)$: By (ii), we have $\partial V_j \subset \partial(U_j \cap \Omega)$. But then we find

$$R_j = \overline{D \cap U_j} = \overline{D \cap U_j} \cap \partial V_j = \overline{U_j \cap \partial V_j}$$

due to $U_j \cap \partial V_j \subseteq \mathcal{U} \cap \partial V_j \subset D$. □

We aim lastly at equations on τK^- and τK for localized equations in neighborhoods of boundary points of Ω, to be achieved via the bi-Lipschitzian transformations occurring in Definition 1.3.12. Hence it is, of course, of interest onto which sets the different boundary parts are mapped by these transformations:

Lemma 2.1.20. *Let* $\mathrm{x} \in \partial\Omega$.

(i) If $\mathrm{x} \in D \setminus \overline{N}$, *then for each* $j \in \{1, \dots, k\}$ *one has* $\phi_j(\partial V_j) = \partial(\tau_j K^-)$ *and, in the terminology of Lemma 2.1.19,* $\phi_j(R_j) = \phi_j(\overline{\partial V_j \cap U_j}) = \tau_j \overline{\Sigma}$.

(ii) If $\mathrm{x} \in \overline{N}$, *one has in the terminology of Lemma 2.1.18 for the corresponding cases in Definition 1.3.12 (ii):*

(a) $\phi_{\mathrm{x}}(E) = \partial(\tau K^-) \setminus \tau\Sigma$ *and* $\phi_{\mathrm{x}}(R) = \emptyset$, *or*

(b) $\phi_{\mathrm{x}}(E) = \partial(\tau K^-) \setminus (\tau\Sigma \setminus \tau\Sigma^-)$ *and* $\phi_{\mathrm{x}}(R) = \tau\overline{\Sigma^-}$.

Proof. This is straight-forward from the mapping properties of the transformations ϕ_{x} and ϕ_j. □

It turns out that the model constellation in Definition 1.3.12 (ii) (b) is indeed suggestive, but not optimal for further analytical purpose. We show in the next lemma that it can be replaced by another one which is much more controllable later, cf. Chapter 2.1.4. See. [79, Sect. 4.2] for the general transformation ansatz.

Lemma 2.1.21. *For every* $\tau > 0$, *there exists a volume-preserving, bi-Lipschitzian mapping* $\varsigma_d \colon \mathbb{R}^d \to \mathbb{R}^d$ *that maps* τK^- *onto* τK^-, $\partial(\tau K^-) \setminus (\tau\Sigma \setminus \tau\Sigma^-)$ *onto* $\partial(\tau K^-) \setminus \tau\Sigma$ *and* $\tau\overline{\Sigma^-}$ *onto the set* $(-\tau, \tau)^{n-2} \times \{-\tau\} \times [-\tau, 0]$. *Finally,* $\varsigma_n(\frac{\tau}{2}K^-) = \tau\underline{K}^-$ *with* $\underline{K}^- := (-\frac{1}{2}, \frac{1}{2})^{n-2} \times (-1, 0) \times (-\frac{1}{2}, 0)$.

Proof. Let us start with the case $d = 2$, thereby focusing first on the case $\tau = 1$. We define on the lower halfspace $\{(x, y) \in \mathbb{R}^2 \colon y \leq 0\}$

$$\xi_1(x, y) := \begin{cases} (x - y/2, y/2) & \text{if } x \leq 0,\ y \geq x, \\ (x/2, -x/2 + y) & \text{if } x \leq 0,\ y < x, \\ (x/2, x/2 + y) & \text{if } x > 0,\ y < -x, \\ (x + y/2, y/2) & \text{if } x > 0,\ y \geq -x. \end{cases}$$

Observing that ξ_1 acts as the identity on the x-axis, we may define ξ_1 on the upper half space $\{(x, y) \in \mathbb{R}^2 \colon y > 0\}$ by $\xi_1(x, y) = (x, y/2)$. In this way we obtain a globally bi-Lipschitz transformation ξ_1 from $\mathbb{R}^2$ onto itself that transforms $K \cup \Sigma^-$ onto the triangle shown in Figure 2.1.

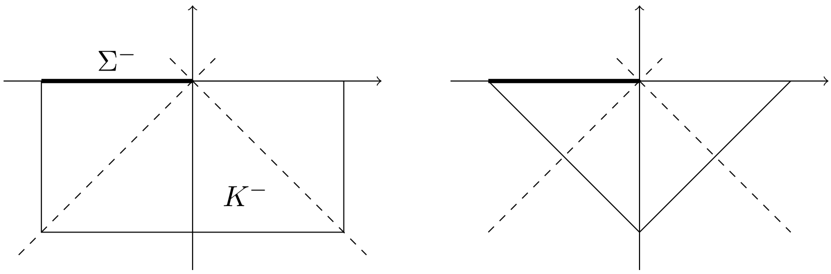


Figure 2.1. $K^- \cup \overline{\Sigma}^-$ and $\xi_1(K^- \cup \overline{\Sigma}^-)$

Next we define the bi-Lipschitz mapping $\xi_2 \colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\xi_2(x,y) := \begin{cases} (x, x+2y+1) & \text{if } x \leq 0, \\ (x, -x+2y+1) & \text{if } x > 0, \end{cases}$$

in order to get the geometric constellation in Figure 2.2. If ξ_3 is the

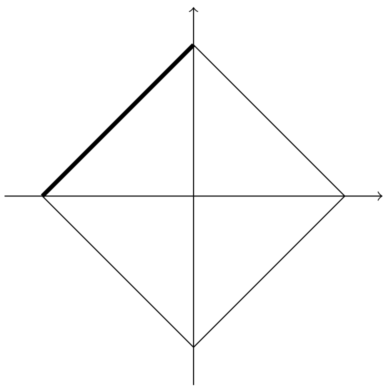


Figure 2.2. $\xi_2(\xi_1(K^- \cup \overline{\Sigma}^-))$

(counter-clockwise) rotation of $\pi/4$ around $0 \in \mathbb{R}^2$, we thus have achieved that $\xi := \xi_3\xi_2\xi_1 : \mathbb{R}^2 \to \mathbb{R}^2$ is bi-Lipschitzian and satisfies

$$\xi(K^-) = \frac{1}{\sqrt{2}}K, \quad \text{and} \quad \xi(\overline{\Sigma}^-) = \left\{\frac{-1}{\sqrt{2}}\right\} \times \left[\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right].$$

Let $\xi_4 \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the affine mapping $(x, y) \mapsto (\sqrt{2}x, \frac{1}{\sqrt{2}}y - \frac{1}{2})$. Then $\varsigma_2 = \xi_4\xi$ is bi-Lipschitzian and maps K^- onto itself, $\partial K^- \setminus (\Sigma \setminus \Sigma^-)$ onto $\partial K^- \setminus \Sigma$, and $\overline{\Sigma^-}$ bi-Lipschitzian onto $\{-1\} \times [-1, 0]$. The assertion for $\underline{K}^-$ is verified by a straight forward calculation. As is easy to check, the determinant of the Jacobian is identically one almost everywhere. Hence, ς_2 is volume-preserving.
If $\tau \neq 1$, then one first applies the homothety $\mathrm{y} \mapsto \frac{1}{\tau}\mathrm{y}$, then the mapping ς_2 just constructed for the case $\tau = 1$ and afterwards the inverse homothety $\mathrm{y} \mapsto \tau\mathrm{y}$.
For $d \geq 3$, one simply puts $\varsigma_d(x_1, \ldots, x_d) := (x_1, \ldots, x_{d-2}, \varsigma_2(x_{d-1}, x_d))$. □

Corollary 2.1.22. *For every point* x *from $\partial D = D \cap \overline{N}$, i.e., in the situation of Definition 1.3.12 (ii) (b), there is a an open neighborhood U_{x}, a positive number $\tau = \tau_{\mathrm{x}}$ and a bi-Lipschitzian, volume-preserving mapping from a neighborhood of $\overline{U_{\mathrm{x}}}$ into $\mathbb{R}^d$, which maps $U_{\mathrm{x}} \cap \Omega$ onto τK^-, E onto $\partial(\tau K^-) \setminus \tau\Sigma$, and R onto the set $[-\tau, \tau]^{n-2} \times \{-\tau\} \times [-\tau, 0]$, where E, R are defined as in Lemma 2.1.18.*

Proof. If one defines the asserted mapping as the composition $\varsigma_n \circ \phi_{\mathrm{x}}$, then the application of Lemma 2.1.20 and Lemma 2.1.21 gives the assertion. □

Having the bi-Lipschitz mappings ϕ and ς defined above at hand, we collect properties of bi-Lipschitzian transformations if applied to the typical data of parabolic equations as (2.9). It turns out that (volume-preserving) bi-Lipschitz mappings essentially preserve the structure of the underlying problem.

Proposition 2.1.23. *Let Λ be a bounded Lipschitz domain, and let F be a closed portion of its boundary. Assume that ζ is a bi-Lipschitzian mapping* *from a neighborhood of* $\overline{\Lambda}$ *into* $\mathbb{R}^d$*. Define for any function* $\varphi\colon \zeta(\Lambda) \to \mathbb{R}$ *the function* $\Phi\varphi\colon \Lambda \to \mathbb{R}$ *by*

$$(\Phi\varphi)(\mathrm{x}) := \varphi(\zeta(\mathrm{x})) = (\varphi \circ \zeta)(\mathrm{x}), \quad \mathrm{x} \in \Lambda.$$

(i) For every $\varphi \in \mathrm{W}^{1,1}(\zeta(\Lambda))$*, the (generalized) gradient of the function* $\varphi \circ \zeta$ *is calculated for almost all* $\mathrm{x} \in \Lambda$ *as follows:*

$$\begin{aligned}\nabla(\varphi \circ \zeta)(\mathrm{x}) &= \begin{pmatrix} \frac{\partial \zeta_1}{\partial \mathrm{x}_1}(\mathrm{x}) & \dots & \frac{\partial \zeta_n}{\partial x_1}(\mathrm{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial \zeta_1}{\partial \mathrm{x}_n}(\mathrm{x}) & \dots & \frac{\partial \zeta_n}{\partial x_n}(\mathrm{x}) \end{pmatrix} \begin{pmatrix} \frac{\partial \varphi}{\partial \mathrm{x}_1}(\zeta(\mathrm{x})) \\ \vdots \\ \frac{\partial \varphi}{\partial \mathrm{x}_n}(\zeta(\mathrm{x})) \end{pmatrix} \\ &= (D\zeta)^\top(\mathrm{x}) \nabla\varphi(\zeta(\mathrm{x})).\end{aligned}$$

(ii) For every $1 < p < \infty$*, the mapping* Φ *induces linear, topological isomorphisms*

$$\Phi_{1,p}\colon \mathrm{W}^{1,p}_{\zeta(F)}(\zeta(\Lambda)) \to \mathrm{W}^{1,p}_F(\Lambda)$$

and

$$\Phi^*_{1,p'}\colon \mathrm{W}^{-1,p}_F(\Lambda) \to \mathrm{W}^{-1,p}_{\zeta(F)}(\zeta(\Lambda))$$

as well as $\Phi_p\colon \mathrm{L}^p(\zeta(\Lambda)) \to \mathrm{L}^p(\Lambda)$*. These are consistent for different values of* p*.*

(iii) If $0 < \alpha < 1$*, then* Φ *induces a topological isomorphism* $\Phi_{0,\alpha}$ *between* $\mathrm{C}^\alpha(\zeta(\Lambda))$ *and* $\mathrm{C}^\alpha(\Lambda)$*. The norms of* $\Phi_{0,\alpha}$ *and* $\Phi_{0,\alpha}^{-1}$ *only depend on the Lipschitz-constants of* ζ *and* ζ^{-1}*.*

(iv) Let $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d)$*. Then one has for every* $1 < p < \infty$ *and every pair* $(\psi, \varphi) \in \mathrm{W}^{1,p}(\zeta(\Lambda)) \times \mathrm{W}^{1,p'}(\zeta(\Lambda))$ *the identity*

$$\int_\Lambda \rho \nabla(\psi \circ \zeta) \cdot \nabla(\varphi \circ \zeta)\,\mathrm{dx} = \int_{\zeta(\Lambda)} \rho_\zeta \nabla\psi \cdot \nabla\varphi\,\mathrm{dy}.$$

with

$$\rho_\zeta(\mathrm{y}) = C_\zeta(\mathrm{y})(D\zeta)(\zeta^{-1}(\mathrm{y}))\rho(\zeta^{-1}(\mathrm{y}))\,(D\zeta)^\top\,(\zeta^{-1}(\mathrm{y})) \qquad (2.18)$$

for almost all $\mathrm{y} \in \zeta(\Lambda)$ *with*

$$C_\zeta(\mathrm{y}) := \frac{1}{\left|\det((D\zeta)(\zeta^{-1}(\mathrm{y}))\right|}.$$

Here, $D\zeta$ *denotes the Jacobian of* ζ*.*

(v) *Let* $\mathfrak{l}_\zeta$ *and* $\mathfrak{l}_{\zeta^{-1}}$ *denote the Lipschitz-constants of* ζ *and* ζ^{-1}*, respectively. If* ζ *is volume preserving and* $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\rho_\bullet, \rho^\bullet))$*, then* $\rho_\zeta \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d(\bar{\rho}_\bullet, \bar{\rho}^\bullet))$*, where* $\bar{\rho}_\bullet := \rho_\bullet \mathfrak{l}_{\zeta^{-1}}^{-2}$ *and* $\bar{\rho}^\bullet := \rho^\bullet \mathfrak{l}_\zeta^2$*.*

Proof. For (i) see [112, Ch. 1.1.7]. The proof of (ii) is contained in [69, Thm. 2.7/2.10]. (iii) is obvious. Assertion (iv) can be deduced from (i), for a complete proof see [78, Prop. 16].

It remains to prove (v). Firstly, one observes that for a volume-preserving mapping ζ the function $|\det(D\zeta)(\cdot)|$ is identically 1, [61, Ch. 3]. Secondly, Rademacher's theorem shows that $\|D\zeta\|_{\mathrm{L}^\infty(\Lambda;\mathbb{M}_d)} \le \mathfrak{l}_\zeta$ and $\|D(\zeta^{-1})\|_{\mathrm{L}^\infty(\zeta(\Lambda);\mathbb{M}_d)} \le \mathfrak{l}_{\zeta^{-1}}$. With all this in mind, one easily calculates for almost all $\mathrm{y} \in \zeta(\Lambda)$ and all $\mathrm{z} \in \mathbb{R}^d$ as follows:

$$\begin{aligned}\|\rho_\zeta(\mathrm{y})\mathrm{z}\cdot\mathrm{z}\|_2 &= \left\|\rho(\zeta^{-1}\mathrm{y})(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\cdot(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\right\|_2 \\ &\le \rho^\bullet\left\|(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\right\|_2^2 \le \rho^\bullet\mathfrak{l}_\zeta^2\,\|\mathrm{z}\|_2^2.\end{aligned}$$

In order to deduce the lower bound, one first recalls the equality

$$(D\zeta)(\zeta^{-1}\mathrm{y}) = \big(D\zeta^{-1})(\mathrm{y})\big)^{-1},$$

which holds for almost all $\mathrm{y} \in \zeta(\Lambda)$, see [61, Ch. 3.1.2, Cor. 1]. Having this at hand, one estimates for almost all $\mathrm{y} \in \zeta(\Lambda)$ and all $\mathrm{z} \in \mathbb{R}^d$

$$\begin{aligned}\|\rho_\zeta(\mathrm{y})\mathrm{z}\cdot\mathrm{z}\|_2 &= \left\|\rho(\zeta^{-1}\mathrm{y})(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\cdot(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\right\|_2 \\ &\ge \rho_\bullet\left\|(D\zeta)^\top(\zeta^{-1}(\mathrm{y}))\mathrm{z}\right\|_2^2 = \rho_\bullet\left\|\left((D(\zeta^{-1}))^\top(\mathrm{y})\right)^{-1}\mathrm{z}\right\|_2^2 \\ &\ge \frac{\rho_\bullet}{\operatorname{ess\,sup}_{\mathrm{y}\in\zeta(\Lambda)}\|D(\zeta^{-1})(\mathrm{y})\|^2}\|\mathrm{z}\|_2^2 \ge \frac{\rho_\bullet}{\mathfrak{l}_{\zeta^{-1}}^2}\|\mathrm{z}\|_2^2.\end{aligned}$$ ☐

Remark 2.1.24. Assume the setting of Proposition 2.1.23.

(i) Under the assumption that ζ is *volume-preserving*, the operator $\Phi_2 \colon \mathrm{L}^2(\zeta(\Lambda)) \to \mathrm{L}^2(\Lambda)$ is in fact *unitary*, i.e., we have

$$\left(\varphi,\psi\right)_{\mathrm{L}^2(\zeta(\Lambda))} = \left(\Phi_2\varphi, \Phi_2\psi\right)_{\mathrm{L}^2(\Lambda)} \quad \text{for all } \varphi,\psi \in \mathrm{L}^2(\zeta(\Lambda))$$

by the usual transformation of variables. This is indeed the functional-analytic manifestation of the usefulness of the *volume-preserving* property.

(ii) If μ is a coefficient function on $J \times \Lambda$, then we denote by μ_ζ the coefficient function $t \mapsto \mu_\zeta(t,\cdot)$ on $J \times \zeta(\Lambda)$ given as in (2.18).

2.1.3 Localization, transformation, reflection

Now we have the principle ideas at hand and will first *localize* the parabolic equation suitably in order to consider it on smaller sets. The resulting equations are then *transformed* by bi-Lipschitzian mappings, corresponding of course to Definition 1.3.12, to equations on the (scaled) half cube τK^-. In the case of points from the Neumann boundary part, one finally needs a reflection argument, which will be established in the last part of this subsection.

We start with the localization procedure. Therefore, for the rest of this section we assume that given $f \in \mathrm{L}^2(J;\mathrm{W}^{-1,2}_D(\Omega))$, we have a unique solution $u \in \mathbb{W}^{1,2}_0(J;\mathrm{W}^{-1,2}_D(\Omega),\mathrm{W}^{1,2}_D(\Omega))$ of the equation

$$\begin{aligned}\langle u'(t),\varphi\rangle_{\mathrm{W}^{1,2}_D(\Omega)} - \langle \nabla\cdot\mu(t,\cdot)\nabla u(t),\varphi\rangle_{\mathrm{W}^{1,2}_D(\Omega)} + \int_\Omega u(t)\varphi\,\mathrm{dx}& \\ = \langle f(t),\varphi\rangle_{\mathrm{W}^{1,2}_D(\Omega)} \quad \text{for all } \varphi\in\mathrm{W}^{1,2}_D(\Omega) \quad \text{for a.a. } t\in J& \qquad (2.19)\end{aligned}$$

at hand. Note that such a unique solution indeed exists and lies additionally in the space $\mathrm{C}(\overline{J};\mathrm{L}^2(\Omega))$, cf. Propositions 2.1.1 and 1.4.3.

Let us fix an arbitrary point $\mathrm{x} \in \overline{\Omega}$ and consider an open neighborhood U of x. If $\mathrm{x} \in \Omega$, we assume $\overline{U} \subset \Omega$. We will now localize the equation around x according to the constructions from Lemmata 2.1.18 (for the first two cases) and 2.1.19 (the last case), respectively:

- If $\mathrm{x} \in \Omega$, set $\Lambda = U$, $E = \partial\Lambda$ and $R = \emptyset$.
- For $\mathrm{x} \in \overline{N}$, we choose $\Lambda = \Omega \cap U$ and E, R as in Lemma 2.1.18, i.e., $E = \partial\Lambda \setminus (N \cap U)$ and $R = \overline{D \cap U}$.
- In case of $\mathrm{x} \in D \setminus \overline{N}$, $\Omega \cap U$ may be disconnected with, say, k connected components V_j. We thus set $\Lambda_j = V_j$, $E_j = \partial V_j$ and $R_j = \overline{\partial V_j \cap U_j}$, where U_j is an open set with $V_j \subset U_j \subset U$, for each $j \in \{1, \ldots, k\}$. The following localization procedure then has to be done for every $j \in \{1, \ldots, k\}$. We will, however, omit the index j to simplify the notation.

We will need to work with restrictions to Λ extensively in the following. To reduce clutter, we agree, for a function $\mathfrak{u}$ defined on Ω, that $\mathfrak{u}_{[\Lambda]} := \mathfrak{u} \restriction \Lambda$. In this terminology, one calculates for $w \in \mathrm{W}^{1,2}_D(\Omega)$, a coefficient function $\rho \in \mathrm{L}^\infty(\Lambda; \mathbb{M}_d)$, and every $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$:

$$\begin{aligned}\langle -\nabla \cdot \rho_{[\Lambda]} \nabla w_{[\Lambda]}, \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} &= \int_\Lambda \rho_{[\Lambda]} \nabla w_{[\Lambda]} \cdot \nabla \varphi \,\mathrm{dx} \\ &= \int_\Omega \rho \nabla w \cdot \nabla (\mathfrak{E}^0_U \varphi) \,\mathrm{dx} = \langle -\nabla \cdot \rho \nabla w, (\mathfrak{E}^0_U \varphi) \rangle_{\mathrm{W}^{1,2}_D(\Omega)}. \end{aligned} \tag{2.20}$$

Remark 2.1.25. The first term in (2.20) does *not* contain abuse of the above introduced notation in the following sense: for $w \in \mathrm{W}^{1,2}_D(\Omega)$ the restriction $w_{[\Lambda]}$ belongs to the space $\mathrm{W}^{1,2}_R(\Lambda)$, cf. Lemmata 2.1.18 (iii) and 2.1.19 (iii). The operators $-\nabla \cdot \rho_{[\Lambda]} \nabla$ are well-defined from $\mathrm{W}^{1,2}_R(\Lambda)$ to $\mathrm{W}^{-1,2}_R(\Lambda)$ since R is closed, giving $-\nabla \cdot \rho_{[\Lambda]} \nabla w_{[\Lambda]} \in \mathrm{W}^{-1,2}_R(\Lambda)$. But R is contained in E, which yields $\mathrm{W}^{1,2}_E(\Lambda) \hookrightarrow \mathrm{W}^{1,2}_R(\Lambda)$ with isometric injection. Thus,

$$\mathrm{W}^{1,2}_E(\Lambda) \ni \varphi \mapsto \langle -\nabla \cdot \rho_{[\Lambda]} \nabla w_{[\Lambda]}, \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)}$$

is to be understood as the restriction of the linear form $-\nabla \cdot \rho_{[\Lambda]} \nabla w_{[\Lambda]} \in \mathrm{W}^{-1,2}_R(\Lambda)$ to the subspace $\mathrm{W}^{1,2}_E(\Lambda) \subseteq \mathrm{W}^{1,2}_R(\Lambda)$.

Let us now return to the unique function $u \in \mathbb{W}^{1,2}_0(J; \mathrm{W}^{-1,2}_D(\Omega), \mathrm{W}^{1,2}_D(\Omega))$

which satisfies (2.19) with $f \in \mathrm{L}^2(J; \mathrm{W}^{-1,2}_D(\Omega))$. From the identity

$$\int_\Lambda u(t)\varphi \,\mathrm{dx} = \int_\Omega u(t)\mathfrak{E}^0_U \varphi \,\mathrm{dx} \quad \text{for all } \varphi \in \mathrm{W}^{1,2}_E(\Lambda),$$

one deduces (see [48, Ch. XVIII §1.2 Prop. 7]) that the distributional time derivative of $u_{[\Lambda]}$ is given by $(u_{[\Lambda]})'(t) = (\mathfrak{E}^0_U)^* u'(t) \in \mathrm{W}^{-1,2}_E(\Lambda)$ for each $t \in J$, since we have for all $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$ the identities

$$\begin{aligned} \langle (u_{[\Lambda]})'(t), \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} &= \frac{\mathrm{d}}{\mathrm{d}t} \int_\Lambda u_{[\Lambda]}(t)\varphi \,\mathrm{dx} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega u(t)\mathfrak{E}^0_U \varphi \,\mathrm{dx} = \langle u'(t), \mathfrak{E}^0_U \varphi \rangle_{\mathrm{W}^{1,2}_D(\Omega)}. \end{aligned} \tag{2.21}$$

Here, the time derivative on the left hand side is taken in the sense of $\mathrm{W}^{-1,2}_E(\Lambda)$-valued distributions and in the sense of $\mathrm{W}^{-1,2}_D(\Omega)$-valued distributions on the right-hand side.

Note carefully that everything is indeed in order since $\mathfrak{E}^0_U \colon \mathrm{W}^{1,2}_E(\Lambda) \to \mathrm{W}^{1,2}_D(\Omega)$ is well-defined and continuous, thanks to Lemma 2.1.18 (ii) and Lemma 2.1.19 (i). Here one may notice the crucial part in the construction of the sets E and R in dependence of Λ: E is built exactly such that it includes the "interior" border of Λ to the rest of Ω. This way, we are able to extend a function $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$ by zero to the whole Ω which allows to fall back to integration over Ω.

One step further, using (2.21) and (2.20) in case of $w = u(t)$ and $\rho = \mu(t, \cdot)$, one obtains for every $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$ and for almost all $t \in J$:

$$\begin{aligned} \langle (u_{[\Lambda]})'(t), \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} &- \langle \nabla \cdot \mu_{[\Lambda]}(t,\cdot)\nabla u_{[\Lambda]}(t), \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} \\ &+ \int_\Lambda u_{[\Lambda]}(t)\varphi \,\mathrm{dx} = \langle f(t), \mathfrak{E}_U \varphi \rangle_{\mathrm{W}^{1,2}_D(\Omega)}. \end{aligned} \tag{2.22}$$

For $1 < q < \infty$ and $g \in \mathrm{W}^{-1,q}_D(\Omega)$, we denote the linear form $\mathrm{W}^{1,q'}_E(\Lambda) \ni$

$\varphi \mapsto \langle g, \mathfrak{E}_U^0 \varphi \rangle$ by $g_U = (\mathfrak{E}_U^0)^* g$. One easily estimates

$$\|g_U\|_{\mathrm{W}_E^{-1,q}(\Lambda)} \le \|(\mathfrak{E}_U^0)^*\|_{\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_E^{-1,q}(\Lambda))} \|g\|_{\mathrm{W}_D^{-1,q}(\Omega)} \le \|g\|_{\mathrm{W}_D^{-1,q}(\Omega)} \tag{2.23}$$

since $\mathfrak{E}_U^0$ is an isometry. This shows the following: the function $J \ni t \mapsto f_U(t)$, defining the right-hand side in (2.22), belongs to $\mathrm{L}^2(J;\mathrm{W}_E^{-1,2}(\Lambda))$, and its norm does not exceed $\|f\|_{\mathrm{L}^2(J;\mathrm{W}_D^{-1,2}(\Omega))}$. Analogously, if $2 < q, s < \infty$, and $f \in \mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$, then $f_U \in \mathrm{L}^s(J;\mathrm{W}_E^{-1,q}(\Lambda))$ with a similar estimate.

Remark 2.1.26. In any of the localization cases, the property $u \in \mathrm{L}^2(J;\mathrm{W}_D^{1,2}(\Omega)) \cap \mathrm{C}(\overline{J};\mathrm{L}^2(\Omega))$ implies that we have $u_{[\Lambda]} \in \mathrm{L}^2(J;\mathrm{W}_R^{1,2}(\Lambda)) \cap \mathrm{C}(\overline{J};\mathrm{L}^2(\Lambda))$, and the corresponding $\mathrm{V}_2^{1,0}(J \times \Lambda)$-norm of $u_{[\Lambda]}$ is not larger than the $\mathrm{V}_2^{1,0}(Q)$-norm of u, cf. Lemmata 2.1.18 and 2.1.19. Moreover, $u'_{[\Lambda]} \in \mathrm{L}^2(J;\mathrm{W}_E^{-1,2}(\Lambda))$ due to $(u_{[\Lambda]})'(t) = (\mathfrak{E}_U^0)^* u'(t)$ for almost all $t \in J$.

This means we end up with the following equation on each of the local sets Λ, satisfied for the restriction $u_{[\Lambda]}$ of the unique solution u of (2.19):

Localization: The function $u_{[\Lambda]}$ is from $\mathrm{L}^2(J;\mathrm{W}_R^{1,2}(\Lambda)) \cap \mathrm{C}(\overline{J};\mathrm{L}^2(\Lambda))$, with $u'_{[\Lambda]} \in \mathrm{L}^2(J;\mathrm{W}_E^{-1,2}(\Lambda))$ given by $(u_{[\Lambda]})'(t) = (\mathfrak{E}_U^0)^* u'(t)$ for almost all $t \in J$, and satisfies

$$\begin{aligned}\big\langle (u_{[\Lambda]})'(t), \varphi \big\rangle_{\mathrm{W}_E^{1,2}(\Lambda)} + \big\langle -\nabla \cdot \mu_{[\Lambda]}(t,\cdot)\nabla u_{[\Lambda]}(t), \varphi \big\rangle_{\mathrm{W}_E^{1,2}(\Lambda)} \\ + \int_\Lambda u_{[\Lambda]}(t)\varphi \,\mathrm{dx} = \big\langle f_U(t), \varphi \big\rangle_{\mathrm{W}_E^{1,2}(\Lambda)}\end{aligned} \tag{2.24}$$

for all $\varphi \in \mathrm{W}_E^{1,2}(\Lambda)$ and almost all $t \in J$. The functional f_U is from $\mathrm{L}^s(J;W_E^{-1,q}(\Lambda))$ and given by $f_U(t) = (\mathfrak{E}_U^0)^* f(t)$ for almost all $t \in J$.

This completes the *localization* procedure so far: For every possible constellation in and around a point $\mathrm{x} \in \overline{\Omega}$, we have constructed a suitable local equation in $\mathrm{W}_E^{-1,2}(\Lambda)$ which is satisfied by the global solution u. Next, we *transform* these local equations according to Definition 1.3.12 using the properties of the transformations established in Proposition 2.1.23. Suppose from now on that for every point $\mathrm{x} \in \partial\Omega$, a neighborhood U of x is given as declared in the fitting case in Definition 1.3.12 and that Λ, E and R are chosen accordingly as in the localization procedure above (with the obvious adjustments).

We now exploit the *volume-preserving* part of Definition 1.3.12, that is, for each case of boundary points x, there is a volume-preserving, bi-Lipschitzian mapping ζ from a neighborhood of $\overline{\Lambda}$ onto a neighborhood of the cube τK^-. Let us assume that E is mapped onto $E_\bullet \subset \partial(\tau K^-)$, and that R is mapped onto $R_\bullet \subset \partial(\tau K^-)$ – where ζ and $E_\bullet, R_\bullet$ will be specified later and, of course, in correspondence with Definition 1.3.12, Lemma 2.1.20 and Corollary 2.1.22.

For almost all $t \in J$, we know that $u_{[\Lambda]}(t) \in \mathrm{W}_R^{1,2}(\Lambda)$ is of the form

$$u_{[\Lambda]}(t) = \Phi v(t) = v(t) \circ \zeta$$

for a $v(t) \in \mathrm{W}_{R_\bullet}^{1,2}(\tau K^-)$, just as every $\varphi \in \mathrm{W}_E^{1,2}(\Lambda)$ is of the form $\varphi = \Phi\psi = \psi \circ \zeta$ for some $\psi \in \mathrm{W}_{E_\bullet}^{1,2}(\tau K^-)$, both thanks to Proposition 2.1.23 (ii) (see also there for the definition of Φ). Taking this into account and using that ζ is volume-preserving, i.e., $|\det(D\zeta)| = |\det(D\zeta^{-1})| \equiv 1$ almost everywhere on τK^-, one obtains

$$\begin{aligned}
\langle (u_{[\Lambda]})'(t), \varphi \rangle_{\mathrm{W}_E^{1,2}(\Lambda)} &= \frac{\mathrm{d}}{\mathrm{d}t} \langle u_{[\Lambda]}(t), \varphi \rangle_{\mathrm{W}_E^{1,2}(\Lambda)} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int_\Lambda u_{[\Lambda]}(t) \varphi \,\mathrm{dx} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int_\Lambda (v(t) \circ \zeta)(\psi \circ \zeta) \,\mathrm{dy} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\tau K^-} v(t)\, \psi \,\mathrm{dx} = \langle v'(t), \psi \rangle_{\mathrm{W}_{E_\bullet}^{1,2}(\tau K^-)},
\end{aligned}$$

for all $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$ and thus for all $\psi \in \mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)$, and almost all $t \in J$ (see also Remark 2.1.24). This shows that v' is given by

$$v'(t) = \Phi^*_{1,2}(u_{[\Lambda]})'(t) = \Phi^*_{1,2}(\mathfrak{E}^0_U)^* u'(t) \in \mathrm{W}^{-1,2}_{E_\bullet}(\tau K^-)$$

for almost all $t \in J$. On the other hand, one gets for every $\varphi \in \mathrm{W}^{1,2}_E(\Lambda)$ for almost every $t \in J$

$$\begin{aligned}
\langle \nabla \cdot \mu(t,\cdot)_{[\Lambda]} \nabla u_{[\Lambda]}(t), \varphi \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} &= \langle \nabla \cdot \mu(t,\cdot)_{[\Lambda]} \nabla \big(v(t) \circ \zeta\big), \psi \circ \zeta \rangle_{\mathrm{W}^{1,2}_E(\Lambda)} \\
&= -\int_\Lambda \mu(t,\cdot)_{[\Lambda]} \nabla\big(v(t) \circ \zeta\big) \cdot \nabla\big(\psi \circ \zeta\big) \,\mathrm{dy} \\
&= -\int_{\tau K^-} \mu_\zeta(t,\cdot) \nabla v(t) \cdot \nabla \psi \,\mathrm{dx} \\
&= \langle \nabla \cdot \mu_\zeta(t,\cdot) \nabla v(t), \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)},
\end{aligned}$$

cf. Proposition 2.1.23 (iv). Finally, for almost all $t \in J$,

$$\int_\Lambda u(t)_{[\Lambda]} \varphi \,\mathrm{dx} = \int_{\tau K^-} v(t) \psi \,\mathrm{dx},$$

since ζ is volume-preserving cf. again Remark 2.1.24. Hence, (2.24) leads to the following equation for the transformed function v:

$$\begin{aligned}
\langle v', \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} - \langle \nabla \cdot \mu_\zeta(t,\cdot) \nabla v, \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} + \int_{\tau K^-} v(t) \psi \,\mathrm{dx}& \\
= \langle f_U, \psi \circ \zeta \rangle_{\mathrm{W}^{1,2}_E(\Lambda)}&
\end{aligned}$$

for all $\psi \in \mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)$. In view of (2.23), for every $\psi \in \mathrm{W}^{1,q'}_{E_\bullet}(\tau K^-)$ and almost all $t \in J$ one obtains

$$\begin{aligned}
\Big| \langle f_U(t), \psi \circ \zeta \rangle_{\mathrm{W}^{1,q'}_E(\Lambda)} \Big| &\le \| f_U(t) \|_{\mathrm{W}^{-1,q}_E(\Lambda)} \| \psi \circ \zeta \|_{\mathrm{W}^{1,q'}_E(\Lambda)} \\
&\le C_\zeta \| f(t) \|_{\mathrm{W}^{-1,q}_D(\Lambda)} \| \psi \|_{\mathrm{W}^{1,q'}_{E_\bullet}(\tau K^-)}, \qquad (2.25)
\end{aligned}$$

the constant C_ζ only depending on ζ, see Proposition 2.1.23 (i). Thus, for almost every $t \in J$, the linear form

$$\mathrm{W}^{1,q'}_{E_\bullet}(\tau K^-) \ni \psi \mapsto \langle f_U(t), \psi \circ \zeta \rangle$$

belongs to $\mathrm{W}^{-1,q}_{E_\bullet}(K)$ and is in fact given by $\Phi^*_{1,q'} f_U(t)$ (from which the estimate (2.25) would have followed as well, of course). If one denotes this linear form by $g(t) = \Phi^*_{1,q'} f_U(t)$, then (2.25) shows the following: if f in (2.5), cf. also (2.9), even belongs to $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$, then g is from $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_{E_\bullet}(\tau K^-))$ and, additionally, fulfills the estimate

$$\|g\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_{E_\bullet}(\tau K^-))} \le C_\zeta \|f\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))}, \tag{2.26}$$

the constant C_ζ only depending on the mapping ζ.

Remark 2.1.27. Again, the property $u_{[\Lambda]} \in \mathrm{L}^2(J; \mathrm{W}^{1,2}_R(\Lambda) \cap \mathrm{C}(\overline{J}; \mathrm{L}^2(\Lambda))$ leads to v being from $\mathrm{L}^2(J; \mathrm{W}^{1,2}_{R_\bullet}(\tau K^-)) \cap \mathrm{C}(\overline{J}; \mathrm{L}^2(\tau K^-))$ including a corresponding estimate whose constant depends *only* on the bi-Lipschitz mapping ζ, cf. Proposition 2.1.23 (ii). Moreover, we have already noted that $v'(t) = \Phi^*_{1,2}(u_{[\Lambda]})'(t)$, hence $v' \in \mathrm{L}^2(J; \mathrm{W}^{-1,2}_{E_\bullet}(\tau K^-))$ together with estimates for the corresponding norms.

This altogether gives the final product of the transformation stage starting from $u_{[\Lambda]}$, where v was given by $v(t) = \Phi^{-1} u_{[\Lambda]}(t)$.

Transformation: The function $v = \Phi^{-1} u_{[\Lambda]}$ is from $\mathrm{L}^2(J; \mathrm{W}^{1,2}_{R_\bullet}(\tau K^-)) \cap \mathrm{C}(\overline{J}; \mathrm{L}^2(\tau K^-))$, with $v' \in \mathrm{L}^2(J; \mathrm{W}^{-1,2}_{E_\bullet}(\tau K^-))$ given by $v'(t) = \Phi^*_{1,2}(\mathfrak{E}^0_U)^* u'(t)$ for almost all $t \in J$, and satisfies

$$\begin{aligned} \langle v'(t), \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} &- \langle \nabla \cdot \mu_\zeta(t,\cdot) \nabla v(t), \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} \\ &+ \int_{\tau K^-} v(t)\psi \,\mathrm{dx} = \langle g(t), \psi \rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} \end{aligned} \tag{2.27}$$

for all $\psi \in \mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)$ and almost all $t \in J$. The functional g is from

$\mathrm{L}^s(J; \mathrm{W}^{-1,q}_{E_\bullet}(\tau K^-))$ and given by $f_U(t) = \Phi^*_{1,q'}(\mathfrak{E}^0_U)^* f(t)$ for almost all $t \in J$.

Let us now specify the mapping ζ in dependence of the different cases in Definition 1.3.12 and the conventions from the beginning of the localization procedure, defining the sets $E_\bullet = \zeta(E)$ and $R_\bullet = \zeta(R)$ correspondingly:

- In case (i) one puts $\zeta_j \coloneqq \phi_j$, thus obtaining

$$E_{j,\bullet} = \zeta_j(E_j) = \partial(\tau_j K^-) \quad \text{and} \quad R_{j,\bullet} = \zeta_j(R_j) = \tau_j \overline{\Sigma}, \tag{2.28}$$

  for each $j \in \{1, \ldots, k\}$, see Lemma 2.1.20.
- In case (ii) (a), we set $\zeta = \phi_{\mathrm{x}}$, such that

$$E_\bullet = \zeta(E) = \partial(\tau K^-) \setminus \tau\Sigma \quad \text{and} \quad R_\bullet = \zeta(R) = \emptyset, \tag{2.29}$$

  cf. Lemma 2.1.20.
- In case (ii) (b) we choose $\zeta \coloneqq \varsigma_n \circ \phi_{\mathrm{x}}$ and obtain, in view of Corollary 2.1.22,

$$\begin{aligned} &E_\bullet = \zeta(E) = \partial(\tau K^-) \setminus \tau\Sigma \quad \text{and} \\ &R_\bullet = \zeta(R) = [-\tau, \tau]^{n-2} \times \{-\tau\} \times [-\tau, 0]. \end{aligned} \tag{2.30}$$

  Observe that in this last case $\zeta(\mathrm{x}) = (0, \ldots, 0, -\tau, 0)$.

Having the transformed equations on the half cubes with transformed boundary conditions at hand, we lastly introduce *reflection* for case (ii) from Definition 1.3.12.

Inspection of Corollaries 2.1.11 and 2.1.13 reveals why this is necessary: Both corollaries require a subdomain Λ_0 which has a positive distance to the whole boundary $\partial\Lambda$ or to the complement of the Dirichlet boundary part F. But in case (ii) of Definition 1.3.12, after the localization and transformation procedure we end up with $\zeta(\mathrm{x})$ being a boundary point on the half square without prescribed Dirichlet boundary part (remember

$\zeta(R) = \emptyset$ in case (a)) and $\zeta(\mathrm{x})$ being at the boundary of the Dirichlet boundary part itself, respectively. Both cases do not admit a suitable neighborhood of $\zeta(\mathrm{x})$ which would satisfy the assumptions of Corollaries 2.1.11 and 2.1.13. By reflecting the equation across the "upper" plate of the half cubes, we obtain $\zeta(\mathrm{x})$ being inner points of the whole cube and the (combined) Dirichlet boundary part, respectively, allowing to use the aforementioned corollaries.

We follow [80, Sect. 4.2] and first define for $\mathrm{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ the symbol

$$\mathrm{x}_- := (x_1, \dots, x_{d-1}, -x_d),$$

as well as for $\varrho \in \mathbb{M}_d$ the matrix ϱ^- by

$$\varrho^-_{i,j} := \begin{cases} \varrho_{i,j} & \text{if } i, j < d, \\ -\varrho_{i,j} & \text{if } i = d \text{ and } j \neq d \text{ or } j = d \text{ and } i \neq d, \\ \varrho_{i,j} & \text{if } i = j = d. \end{cases}$$

Corresponding to a coefficient function $\rho \in \mathrm{L}^\infty(\tau K^-; \mathbb{M}_d)$, we then define the coefficient function $\hat{\rho} \in \mathrm{L}^\infty(\tau K; \mathbb{M}_d)$ by

$$\hat{\rho}(\mathrm{x}) := \begin{cases} \rho(\mathrm{x}) & \text{if } \mathrm{x} \in \tau K^-, \\ \big(\rho(\mathrm{x}_-)\big)^- & \text{if } \mathrm{x}_- \in \tau K^-, \\ 1 & \text{if } \mathrm{x} \in \Sigma. \end{cases}$$

Finally, we define for $w \in \mathrm{L}^1(\tau K)$ the function w_- by $w_-(\mathrm{x}) = w(\mathrm{x}_-)$, and for $w \in \mathrm{L}^1(\tau K^-)$ the (symmetrically) reflected function by

$$\mathfrak{E}\colon \mathrm{L}^1(\tau K^-) \to \mathrm{L}^1(\tau K), \quad (\mathfrak{E} w)(\mathrm{x}) = \begin{cases} w(\mathrm{x}) & \text{if } \mathrm{x} \in \tau K^-, \\ w(\mathrm{x}_-) & \text{if } \mathrm{x}_- \in \tau K^-, \end{cases}$$

The character $\mathfrak{E}$ used above is a (vertically) reflected "$\mathfrak{E}$" as a symbol for the symmetrically reflected extension over $\tau\Sigma$. Indeed and most importantly, the such-defined reflection is compatible with the vanishing traces property. For concise notation, we set

$$F^{\mathfrak{r}} := F \cup \{\mathrm{x} : \mathrm{x}_{-} \in F\}$$

for a closed subset F of $\partial(\tau K^{-}) \setminus \tau\Sigma$.

Lemma 2.1.28. *Let F be a closed subset of $\partial(\tau K^{-})\setminus\tau\Sigma$, put and assume $1 \le p < \infty$. Then $w \in \mathrm{W}^{1,p}_{F}(\tau K^{-})$ if and only if $\mathfrak{I}w \in \mathrm{W}^{1,p}_{F^{\mathfrak{r}}}(\tau K)$.*

Proof. First, $\mathfrak{I}w \in \mathrm{W}^{1,p}_{F^{\mathfrak{r}}}(\tau K)$ trivially implies $w \in \mathrm{W}^{1,p}_{F}(\tau K^{-})$. In view of the converse assertion, it is known that $w \in \mathrm{W}^{1,p}_{F}(\tau K^{-}) \subseteq \mathrm{W}^{1,p}(\tau K^{-})$ implies $\mathfrak{I}w \in \mathrm{W}^{1,p}(\tau K)$, see [68, Lemma 3.4]. Lastly, standard arguments show that $\mathfrak{I}w$ may be approximated in the $\mathrm{W}^{1,p}$-norm by restrictions of $\mathrm{C}^{\infty}_{c}(\mathbb{R}^{d})$-functions the support of which avoids $F^{\mathfrak{r}}$. ☐

Let us next introduce an extension operator for distribution-type objects: For $1 < p < \infty$, define the extension operator $\mathfrak{E}\colon \mathrm{W}^{-1,p}_{F}(\tau K^{-}) \to \mathrm{W}^{-1,p}_{F^{\mathfrak{r}}}(\tau K)$ acting on $f \in \mathrm{W}^{-1,p}_{F}(\tau K^{-})$ as the adjoint operator of the symmetric projection from τK onto τK^{-}, i.e.,

$$\langle \mathfrak{E}f, \varphi\rangle_{\mathrm{W}^{1,p'}_{F^{\mathfrak{r}}}(\tau K)} := \langle f, \varphi_{[\tau K^{-}]} + (\varphi_{-})_{[\tau K^{-}]}\rangle_{\mathrm{W}^{1,p'}_{F}(\tau K^{-})} \quad \text{for } \varphi \in \mathrm{W}^{1,p'}_{F^{\mathfrak{r}}}(\tau K).$$

Here, we have used the subscript notation with brackets again for the restriction, this time of functions on τK. We immediately obtain the following properties:

Lemma 2.1.29. *Assume $1 < p < \infty$.*

(i) If $\psi \in \mathrm{L}^{1}(\tau K^{-}) \cap \mathrm{W}^{-1,p}_{F}(\tau K^{-})$, then $\mathfrak{E}\psi$ is given by $\mathfrak{I}\psi$ in the $\mathrm{L}^{2}(\tau K)$ scalar product.

(ii) For any closed subset $F \subseteq \partial(\tau K^{-})\setminus\tau\Sigma$, the operator $\mathfrak{E}$ is a continuous linear one between $\mathrm{W}^{-1,p}_{F}(\tau K^{-})$ and $\mathrm{W}^{-1,p}_{F^{\mathfrak{r}}}(\tau K)$ whose operator norm is bounded by 2.

Proof. One has for all $\varphi \in \mathrm{W}^{1,p'}_{F^{\mathfrak{r}}}(\tau K)$ the identity

$$\langle \mathfrak{E}\psi, \varphi\rangle_{\mathrm{W}^{1,p'}_{F^{\mathfrak{r}}}(\tau K)} = \int_{\tau K^-} \psi\big(\varphi_{[\tau K^-]} + (\varphi_-)_{[\tau K^-]}\big)\,\mathrm{dx} = \int_{\tau K} \mathfrak{I}\psi\,\varphi\,\mathrm{dx},$$

which proves the first point. Moreover, as noted above, the operator $\mathfrak{E}$ under consideration is the adjoint of the continuous operator with norm bounded by 2 realizing the symmetric projection to τK^-, given by

$$\mathrm{W}^{1,p'}_{F^{\mathfrak{r}}}(\tau K) \ni \varphi \mapsto \varphi_{[\tau K^-]} + (\varphi_-)_{[\tau K^-]} \in \mathrm{W}^{1,p'}_{F}(\tau K^-),$$

which implies both assertions from the second point. ☐

The next lemma ensures that a function satisfying a suitable differential equation on the lower half cube does so also in a reflected sense on the whole cube.

Lemma 2.1.30. *Let $E_\bullet, R_\bullet$ with $R_\bullet \subseteq E_\bullet$ be two closed subsets of $\partial(\tau K^-) \setminus \tau\Sigma$ and let $\rho \in \mathrm{L}^\infty(\tau K^-;\mathbb{M}_d)$ be a coefficient function. Assume that $w \in \mathrm{W}^{1,2}_{R_\bullet}(\tau K^-)$ satisfies*

$$\big\langle -\nabla\cdot\rho\nabla w + w, \psi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} = \langle h, \psi\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} \tag{2.31}$$

for all $\psi \in \mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)$ and some $h \in \mathrm{W}^{-1,2}_{E_\bullet}(\tau K^-)$. Then

$$\big\langle -\nabla\cdot\hat\rho\nabla(\mathfrak{I}w) + \mathfrak{I}w, \varphi\big\rangle_{\mathrm{W}^{1,2}_{E^{\mathfrak{r}}_\bullet}(\tau K)} = \langle \mathfrak{E}h, \varphi\rangle_{\mathrm{W}^{1,2}_{E^{\mathfrak{r}}_\bullet}(\tau K)}$$

is satisfied for all $\varphi \in \mathrm{W}^{1,2}_{E^{\mathfrak{r}}_\bullet}(\tau K)$.

Proof. The assertion is obtained by the definitions of $\mathfrak{I}w$, $\mathfrak{E}h$, the connection between $-\nabla\cdot\rho\nabla$ and $-\nabla\cdot\hat\rho\nabla$, and straightforward calculations based on Proposition 2.1.23 applied to the transformation $\mathrm{x} \mapsto \mathrm{x}_-$. ☐

Now let us return to the transformed equation (2.27) and suppose that the preceding localization procedure was done with respect to a point $\mathrm{x} \in \overline{N}$, i.e., in case (ii) of Definition 1.3.12. Then $E_\bullet$ and $R_\bullet$ are closed subsets of $\partial(\tau K^-) \setminus \tau\Sigma$.

Rearranging (2.27), we obtain

$$\begin{aligned}\big\langle -\nabla \cdot \mu(t,\cdot)\nabla v(t) + v(t), \psi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(K)} \\ = \big\langle g(t), \psi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(K)} - \big\langle v'(t), \psi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(K)}\end{aligned}$$

for all $\psi \in \mathrm{W}^{1,2}_{E_\bullet}(K)$, which for almost all $t \in J$ is an equation of type (2.31). The foregoing Lemma 2.1.30 tells us that this leads to the equation

$$\begin{aligned}\big\langle -\nabla \cdot \widehat{\mu(t,\cdot)}\nabla\big(\mathfrak{J}v(t)\big) + \mathfrak{J}v(t), \varphi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)} \\ = \big\langle \mathfrak{E}g(t), \varphi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)} - \big\langle \mathfrak{E}v'(t), \varphi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)},\end{aligned}$$

true for all $\varphi \in \mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)$ and almost all $t \in J$. For the time derivative, one calculates for each $\varphi \in \mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)$ using the definition of v' and $\mathfrak{J}$ as follows:

$$\begin{aligned}\big\langle \mathfrak{E}v'(t), \varphi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)} &= \big\langle v'(t), \varphi_{[\tau K^-]} + (\varphi_-)_{[\tau K^-]}\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet}(\tau K^-)} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\tau K^-} v(t)\big(\varphi + \varphi_-\big)\,\mathrm{dx} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\tau K} \mathfrak{J}v(t)\,\varphi\,\mathrm{dx} = \big\langle (\mathfrak{J}v)'(t), \varphi\big\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)},\end{aligned}$$

which shows that $(\mathfrak{J}v)'(t)$ is given by $\mathfrak{E}v'(t) \in \mathrm{W}^{-1,2}_{E_\bullet^{\mathfrak{r}}(\tau K)}$ for almost all $t \in J$. Clearly, we then have $\mathfrak{J}v \in \mathrm{L}^2(J; \mathrm{W}^{1,2}_{R_\bullet^{\mathfrak{r}}}(\tau K)) \cap \mathrm{C}(\overline{J}; \mathrm{L}^2(\tau K))$ with $(\mathfrak{J}v)' \in \mathrm{L}^2(J; \mathrm{W}^{-1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K))$. Finally, if $g \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_{E_\bullet}(\tau K^-))$, then $\mathfrak{E}g \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_{E_\bullet^{\mathfrak{r}}}(\tau K))$ and

$$\|\mathfrak{E}g\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_{E_\bullet^{\mathfrak{r}}}(\tau K))} \le 2\|g\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_{E_\bullet}(\tau K^-))}$$

due to Lemma 2.1.29. We sum these considerations up, in relation to the function v obtained at the end of the transformation stage for case (ii) of Definition 1.3.12.

Reflection: The function $\mathfrak{J}v$ is from $\mathrm{L}^2(J;\mathrm{W}^{1,2}_{R_\bullet^{\mathfrak{r}}}(\tau K))\cap\mathrm{C}(\overline{J};\mathrm{L}^2(\tau K))$ with $(\mathfrak{J}v)'\in\mathrm{L}^2(J;\mathrm{W}^{-1,2}_{E_\bullet^{\mathfrak{r}}(\tau K)})$ given by $(\mathfrak{J}v)'(t)=\mathfrak{T}v'(t)$ for almost all $t\in J$, and satisfies

$$\begin{aligned}\langle\mathfrak{J}v'(t),\varphi\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)}&+\langle-\nabla\cdot\widehat{\mu(t,\cdot)}\nabla\mathfrak{J}v(t)\\&+\mathfrak{J}v(t),\varphi\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)}=\langle\mathfrak{T}g(t),\varphi\rangle_{\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)}\end{aligned}\tag{2.32}$$

for all $\varphi\in\mathrm{W}^{1,2}_{E_\bullet^{\mathfrak{r}}}(\tau K)$ and almost all $t\in J$. The functional $\mathfrak{T}g$ is from $\mathrm{L}^s(J;\mathrm{W}^{-1,q}_{E_\bullet^{\mathfrak{r}}}(\tau K))$.

2.1.4 The core of the proof

Now we have all preparations at hand and will prove our main result, Theorem 2.1.4. For this, we take the assumptions of that theorem as given from now on.

The following lemma is the starting point for the usage of the foregoing results because it opens the door to use the classical results presented in Chapter 2.1.1 by establishing uniform bounds in the space $\mathrm{V}_2^{1,0}(Q)$.

Lemma 2.1.31. *Let $\mathbb{B}_{s,q}(0)$ be the unit ball in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$. For every $f\in\mathbb{B}_{s,q}(0)$, the equation*

$$u'(t)-\nabla\cdot\mu(t,\cdot)\nabla u(t)+u(t)=f(t)\quad\textit{in }\mathrm{W}_D^{-1,2}(\Omega)\quad\textit{for a.a. }t\in J$$

admits a unique solution $u = u_f \in \mathbb{W}_0^{1,2}(J; \mathrm{W}_D^{-1,2}(\Omega); \mathrm{W}_D^{1,2}(\Omega))$ *which is contained in a ball* $\mathbb{B}(0, r_\mathrm{V})$ *in* $\mathrm{V}_2^{1,0}(Q)$ *with radius*

$$r_\mathrm{V} := \left(\frac{1}{\varkappa} + \sqrt{\frac{1}{\varkappa}}\right) \lambda^d(\Omega)^{\frac{q-2}{2q}} \lambda(J)^{\frac{s-2}{2s}},$$

where $\varkappa = \min(\mu_\bullet, 1)$*. Hence, for all coefficient functions* μ *admitting the same ellipticity constant* $\mu_\bullet$*, in particular all those from* $\mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$*, the radii* r_V *may be taken uniformly.*

Proof. The forms

$$\mathrm{W}_D^{1,2}(\Omega) \times \mathrm{W}_D^{1,2}(\Omega) \ni (\psi, \varphi) \mapsto \mathfrak{a}_t(\psi, \varphi) := \int_\Omega \mu(t, \cdot)\nabla\psi \cdot \nabla\varphi + \psi\varphi \, \mathrm{dx}$$

for $t \in J$ satisfy the assumptions of Proposition 2.1.1 using the triple $\mathrm{W}_D^{-1,2}(\Omega) \hookrightarrow \mathrm{L}^2(\Omega) \hookrightarrow \mathrm{W}_D^{1,2}(\Omega)$ and admit the uniform coercivity constants $\varkappa$. Moreover, the unit ball $\mathbb{B}_{s,q}(0)$ is contained in a corresponding ball in $\mathrm{L}^2(J; \mathrm{W}_D^{-1,2}(\Omega))$ with radius $\lambda^d(\Omega)^{\frac{q-2}{2q}} \lambda(J)^{\frac{s-2}{2s}}$, cf. Remark 2.1.5. Hence, it remains to observe that the norm bounds in Proposition 2.1.1 imply exactly the $\mathrm{V}_2^{1,0}(Q)$ norm estimate. □

We now proceed to construct a finite open covering of $\overline{\Omega}$ and to show uniform L^∞- and Hölder-bounds on the intersection of each of the covering sets with Ω. To this end, we localize the parabolic equation (2.9) with respect to a suitable neighborhood of each point, transform the localized equations to such on the half cubes and reflect the problem to the whole cube, if necessary. This allows to use Corollaries 2.1.11 and 2.1.13, respectively, to deduce the wished-for estimates.

Choose for any point $\mathrm{x} \in \Omega$ a ball $B_\mathrm{x}^\bullet$ around x which satisfies $B_\mathrm{x}^\bullet \subset \Omega$ and which has a positive distance to $\partial\Omega$. Define B_x as the ball with half the radius of $B_\mathrm{x}^\bullet$. Further, for every $\mathrm{y} \in \partial\Omega$, let U_y be an open neighborhood of y which satisfies the conditions in Definition 1.3.12. In case (i) of that definition, we put $W_\mathrm{y} = \cap_j \phi_j^{-1}(\frac{\tau_j}{2}K)$. If y fulfills case (ii) of the definition, then we put $W_\mathrm{y} = \phi_\mathrm{y}^{-1}(\frac{\tau_\mathrm{y}}{2}K)$, which implies $W_\mathrm{y} \cap \Omega =$

$\phi_{\mathrm{y}}^{-1}(\frac{\tau_{\mathrm{y}}}{2}K^-)$. Obviously, the collection of the sets $\{B_{\mathrm{x}}\}_{\mathrm{x}\in\Omega}$ and $\{W_{\mathrm{y}}\}_{\mathrm{y}\in\partial\Omega}$ forms an open covering of $\overline{\Omega}$. Let $B_{\mathrm{x}_1},\dots,B_{\mathrm{x}_{m_0}},W_{\mathrm{y}_1},\dots,W_{\mathrm{y}_{m_1}}$ be a finite subcovering.

Before we continue, we need the following property of the sets W_{y} in case of Definition 1.3.12 (i):

Lemma 2.1.32. *In the situation of Definition 1.3.12 (i), with* $W := \bigcap_{j=1}^k \phi_j^{-1}(\frac{\tau_j}{2}K)$ *one has*

$$W\cap\Omega \subseteq \bigcup_{j=1}^k \phi_j^{-1}(\frac{\tau_j}{2}K^-), \tag{2.33}$$

the right hand side being a disjoint *union.*

Proof. Since $W \subset \bigcap_{j=1}^k U_j \subseteq \mathcal{U}$, we find

$$\begin{aligned} W\cap\Omega = W\cap\Omega\cap\mathcal{U} = W\cap\bigcup_{j=1}^k V_j = \bigcup_{j=1}^k (V_j\cap W) \\ \subseteq \bigcup_{j=1}^k \Big(V_j\cap\phi_j^{-1}(\frac{\tau_j}{2}K)\Big) = \bigcup_{j=1}^k \Big(\phi_j^{-1}(\tau_j K^-)\cap\phi_j^{-1}(\frac{\tau_j}{2}K)\Big) \\ = \bigcup_{j=1}^k \phi_j^{-1}\Big(\tau_j K^- \cap \frac{\tau_j}{2}K\Big) = \bigcup_{j=1}^k \phi_j^{-1}(\frac{\tau_j}{2}K^-). \quad \square \end{aligned}$$

Let $\mathbb{B}_{s,q}(0)$ be again the unit ball in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$, and let $\mathbb{B}_{s,q}^{\mathrm{step}}$ denote the set of step functions in $\mathbb{B}_{s,q}(0)$. Suppose the assumptions of Theorem 2.1.4.

Step 1: For every $f\in \mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega)) \hookrightarrow \mathrm{L}^2(J;\mathrm{W}_D^{-1,2}(\Omega))$ a (unique) solution $u = u_f$ of (2.9) exists, cf. Proposition 2.1.1 or Remark 2.1.2. The set of solutions $\{u_f\colon f\in\mathbb{B}_{s,q}(0)\}$ is bounded in $\mathrm{V}_2^{1,0}(Q)$, and the bound in this space can be taken uniformly with respect to all coefficient functions $\mu\in\mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$, cf. Lemma 2.1.31.

Step 2: We consider the restricted problem on each of the balls $B^\bullet_{\mathrm{x}_\ell}$ according to Ch. 2.1.3 (there setting $U = B^\bullet_{\mathrm{x}_\ell}$), ending up with (2.24). There, the right-hand side f_U in the restricted problem is still bounded by 1 for $f \in \mathbb{B}_{s,q}(0)$, and the norm of $u \restriction J \times B^\bullet_{\mathrm{x}_\ell}$ in $\mathrm{V}^{1,0}_2(J \times B^\bullet_{\mathrm{x}_\ell})$ is bounded by the $\mathrm{V}^{1,0}_2(Q)$-norm of u itself. For $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$, however, the solution u_f is a generalized solution of a corresponding generalized problem on $B^\bullet_{\mathrm{x}_\ell}$ with right-hand side $\mathfrak{f}_U$, cf. Proposition 2.1.16, $\mathfrak{f}_U$ still being a step function in time and contained in the ball with radius 2 in $\mathrm{L}^s(J; \mathrm{L}^q(B^\bullet_{\mathrm{x}_\ell}; \mathbb{R}^{d+1}))$. Thanks to Corollary 2.1.11, the functions $u_f \restriction J \times B_{\mathrm{x}_\ell}$ are essentially bounded, and the norms $\|u \restriction J \times B_{\mathrm{x}_\ell}\|_{\mathrm{L}^\infty(J \times B_{\mathrm{x}_\ell})}$ are bounded uniformly in $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$ and in $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$. This of course implies uniform boundedness for all $\ell \in \{1, \dots, m_0\}$.

Step 3: Let us now consider the boundary points, thereby temporarily fixing such a point $\mathrm{y} = \mathrm{y}_\ell \in \partial\Omega$.

We start with case (i) of Definition 1.3.12: Intersecting Ω with U_y, the restriction of the function $u = u_f$ to each of the connected components V_j belongs to $\mathrm{W}^{1,2}_{R_j}(V_j)$ when taking R_j as $\overline{\partial V_j \cap U_j}$, cf. Lemma 2.1.19. One obtains a restricted problem on V_j which is of the same quality as (2.9), cf. (2.24) with $\Lambda = V_j$ and $E = \partial V_j$. Further, we transform this resulting problem to a problem for the function $v_j := (u \restriction V_j) \circ \phi_j^{-1}$ on $\tau_j K^-$. According to (2.27) and (2.28), one ends up with an equation for the transformed function v_j on $\tau_j K^-$ with new right-hand side $g_j \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_0(\tau_j K^-))$, which is still a step function in time. By Proposition 2.1.16, v_j is then a generalized solution of the transformed equation (2.27) on $\tau_j K^-$ with right-hand side $\mathfrak{g}_j \in \mathrm{L}^s(J; \mathrm{L}^q(\tau_j K^-; \mathbb{R}^{d+1}))$ and coefficient function μ_{ϕ_j}. This is the setting for all $j \in \{1, \dots, k\}$. Let us show that we are in the situation to use Corollary 2.1.11 for each problem on V_j.

- The new right-hand sides $\mathfrak{g}_j$ may be estimated suitably with respect to the original ones, cf. (2.26) and Proposition 2.1.16, giving

$$\|\mathfrak{g}_j\|_{\mathrm{L}^s(J;\mathrm{L}^q(\tau_j K^-;\mathbb{R}^{d+1}))} \leq 2\|g_j\|_{\mathrm{L}^s(J;\mathrm{W}_0^{-1,q}(\tau_j K^-))} \leq 2c_j.$$

- The resulting transformed coefficient functions μ_{ϕ_j} on $J \times \tau_j K^-$ still admit uniform upper bounds $\mu_j^\bullet$, and uniform ellipticity constants $\mu_\bullet^j$, cf. Proposition 2.1.23 (v).
- Moreover, it is clear that $\|u \restriction J \times V_j\|_{\mathrm{V}_2^{1,0}(J\times V_j)}$ is not larger than $\|u\|_{\mathrm{V}_2^{1,0}(Q)}$, which was uniformly bounded over $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$ and with respect to $f \in \mathbb{B}_{s,q}(0)$ by the constant r_{V} thanks to Lemma 2.1.31. Proposition 2.1.23 (ii) shows that $\|v_j\|_{\mathrm{V}_2^{1,0}(J\times\tau_j K^-)}$ may be estimated by $\tilde{c}_j r_{\mathrm{V}}$ for some constant $\tilde{c}_j$ depending on j via ϕ_j.
- By Remarks 2.1.26 and 2.1.27, we have $v_j \in \mathrm{L}^2(J;\mathrm{W}^{1,2}_{\tau_j\overline{\Sigma}}(\tau_j K^-))$.

Summing up, we have, for each j, coefficient functions from $\mathrm{L}^\infty(J \times \tau_j K^-;\mathbb{M}_d(\mu_\bullet^j,\mu_j^\bullet))$ and right-hand sides $\mathfrak{g}_j$ contained in the $2c_j$-ball around 0 in $\mathrm{L}^s(J;\mathrm{L}^q(\tau_j K^-;\mathbb{R}^{d+1}))$ such that the generalized solutions v_j to all those right-hand sides are in turn contained in a ball with radius $\tilde{c}_j r_{\mathrm{V}}$ in $\mathrm{V}_2^{1,0}(J \times \tau_j K^-)$ and even belong to $\mathrm{L}^2(J;\mathrm{W}^{1,2}_{\tau_j\overline{\Sigma}}(\tau_j K^-))$. Applying Corollary 2.1.11 (ii) with the subdomain $\frac{\tau_j}{2}K^-$, we obtain L^∞-bounds on $J \times \frac{\tau_j}{2}K^-$ for every v_j which are uniform in $f \in \mathbb{B}_{s,q}^{\mathrm{step}}$ and $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$, cf. Figure 2.3. Thanks to (2.33), this gives L^∞-bounds for u on $J\times(W_{\mathrm{y}}\cap\Omega)$, uniformly for $f \in \mathbb{B}_{s,q}^{\mathrm{step}}$ and $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$.

Next we will consider the case (ii) in Definition 1.3.12. We abbreviate $\tau_{\mathrm{y}} =: \tau$. Localizing around y with respect to U_{y} according to Ch. 2.1.3 results in a problem for $u_{[\Lambda]}$ in the form (2.24) with $\Lambda = U_{\mathrm{y}} \cap \Omega$ and $E = \partial\Lambda\setminus(N\cap U_{\mathrm{y}})$. By afterwards transforming the resulting problem via $\zeta = \phi_{\mathrm{y}}$ (case (a)) and $\zeta = \varsigma_n \circ \phi_{\mathrm{y}}$ (case (b)), one again ends up with a problem on τK^- as in (2.27), which we interpret as a generalized problem solved by the function $v = u_{[\Lambda]} \circ \zeta$. We obtain analogous estimates and bounds, especially uniformly in $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d(\mu_\bullet,\mu^\bullet))$ and $f \in \mathbb{B}_{s,q}^{\mathrm{step}}$, for the coefficient function $\mu_{\phi_{\mathrm{y}}}$, right-hand side $\mathfrak{g}$ and solution $v \in \mathrm{V}_2^{1,0}(J\times\tau K^-)$ as we did for each j in the previously handled case (i). The following

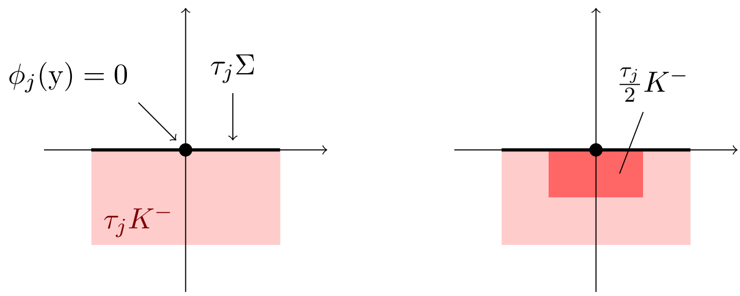


Figure 2.3. Situation (for $d = 2$) in case (i) of Definition 1.3.12, $\tau_j\Sigma$ carries Dirichlet data.

considerations require further distinguishing the assumptions.

In case (ii) (a) of Definition 1.3.12, we get $v \in \mathrm{L}^2(J; \mathrm{W}^{1,2}(\tau K^-))$ according to Remark 2.1.27 and (2.29). Here, the upper plate $\tau\Sigma$ is *disjoint* to the (transformed) Dirichlet boundary part (which in fact is even empty here, cf. (2.29)), permitting the direct application of Corollary 2.1.11 for a neighborhood of $\phi_{\mathrm{y}}(\mathrm{y}) = 0$, since the latter is obviously also a boundary point of τK^-. However, we may reflect the problem across $\tau\Sigma$ according to Lemma 2.1.30, thus obtaining the corresponding equation (2.32) on τK for the symmetrically reflected function $\mathfrak{I}v$. It is clear that the bounds for the data and the $\mathrm{V}_2^{1,0}$-estimate for v carry over to τK in a straight forward manner, cf. Lemma 2.1.29 and the definition of the reflection operator $\mathfrak{I}$, and that $\phi_{\mathrm{y}}(\mathrm{y}) = 0$ is an *interior* point in τK, cf. also Figure 2.4. Hence, we may apply Corollary 2.1.11 (i) for the subdomain $\frac{\tau}{2}K$ and obtain an L^∞-bound for $\mathfrak{I}v$ on $J \times \frac{\tau}{2}K$, again uniformly in $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$ and $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$. Obviously, this implies an L^∞-bound with the same property for u on $J \times (W_{\mathrm{y}} \cap \Omega) = J \times \phi^{-1}(\frac{\tau}{2}K^-)$.

In case (ii) (b) of Definition 1.3.12, where y sits at the boundary between Neumann- and Dirichlet boundary parts, Remark 2.1.27 and (2.30) give $v \in \mathrm{L}^2(J; \mathrm{W}^{1,2}_{R_\bullet}(\tau K^-))$ with $R_\bullet = [-\tau, \tau]^{n-2} \times \{-\tau\} \times [-\tau, 0] \subset \partial(\tau K^-)$ and $\zeta(\mathrm{y}) = (0, \ldots, 0, -\tau, 0)$. Since this point is *not* an interior one of the Dirichlet boundary part $R_\bullet$, we also reflect this problem across $\tau\Sigma$ and, again, end up with a corresponding parabolic equation for the symmet-

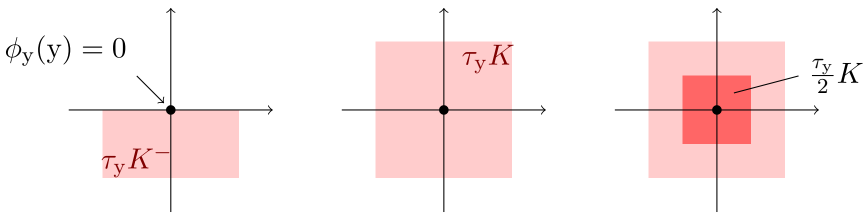


Figure 2.4. Situation (for $d = 2$) in case (ii) (a) of Definition 1.3.12, no Dirichlet data.

rically extended function $\mathfrak{I}v$ on the set τK. The Dirichlet part of the extended solution $\mathfrak{I}v$ now equals $\hat{R}_\bullet = [-\tau, \tau] \times \{-\tau\} \times [-\tau, \tau] \subset \partial(\tau K)$, cf. Lemma 2.1.28. Recalling Lemma 2.1.21, we had

$$\tau \underline{K}^- := \varsigma_n \left(\frac{\tau}{2} K^-\right) = \left]-\frac{\tau}{2}, \frac{\tau}{2}\right[^{d-2} \times (-\tau, 0) \times \left(-\frac{\tau}{2}, 0\right)$$

and one observes that the reflected set $(\tau \underline{K}^-)^{\mathfrak{r}}$ has the distance $\frac{\tau}{2}$ to the set

$$\partial(\tau K) \setminus \hat{R}_\bullet = \partial(\tau K) \setminus \left([-\tau, \tau]^{d-2} \times \{-\tau\} \times [-\tau, \tau]\right),$$

see also Figure 2.5. Another application of Corollary 2.1.11 (ii), this time for the subdomain $(\tau \underline{K}^-)^{\mathfrak{r}}$, gives an L^∞-bound for v on $J \times \tau \underline{K}^-$, and, correspondingly, on $J \times \zeta^{-1}(\tau \underline{K}^-) = J \times \phi_{\mathrm{y}}^{-1}(\frac{\tau}{2} K^-) = J \times (W_{\mathrm{y}} \cap \Omega)$, again uniformly for $f \in \mathbb{B}_{s,q}^{\mathrm{step}}$ with respect to $\mu_\bullet, \mu_\bullet$.

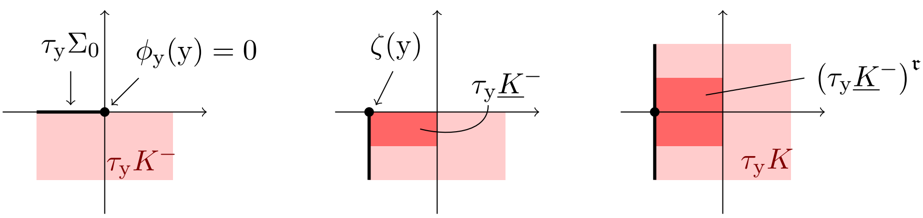


Figure 2.5. Situation (for $d = 2$) in case (ii) (b) of Definition 1.3.12, $\tau_{\mathrm{y}} \Sigma_0$ carries initial Dirichlet data.

Hence we have L^∞-bounds on $J \times W_{\mathrm{y}_\ell} \cap \Omega$ for each $\ell \in \{1, \ldots, m_1\}$ which then clearly implies L^∞-bounds uniform in l. Since the finite system

$B_{\mathrm{x}_1}, \dots, B_{\mathrm{x}_{m_0}}, W_{\mathrm{y}_1} \cap \Omega, \dots, W_{\mathrm{y}_{m_1}} \cap \Omega$ is an open covering of Ω, this altogether gives L^∞-bounds on the whole set Q, which are uniform for all $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$ for the corresponding functions u_f and which do only depend on the constants $\mu_\bullet, \mu^\bullet$. This was the first point of Theorem 2.1.4.

Step 4: Having the essential boundedness at hand, we will now establish the Hölder estimates by essentially re-iterating the considerations in the foregoing steps, this time investing the obtained uniform *global* L^∞-bounds instead of the $\mathrm{V}_2^{1,0}$-estimates and then applying Corollary 2.1.13 instead of Corollary 2.1.11.

In detail: Both Step 2, which was the case of the balls B_{x_ℓ}, and the considerations in case (ii) of Definition 1.3.12 in Step 3 work exactly as above, using Corollary 2.1.13 this time. In case (i) of Step 3, the situation is a bit more complicated and needs more care: Repeating the procedure outlined above to the point where Lemma 2.1.32 and (2.33) are used, one obtains the Hölder property for every transformed local solution v_j (including estimates uniform in $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$, depending only on $\mu_\bullet, \mu^\bullet$) on the set $J \times \frac{\tau_j}{2} K^-$ for each $j \in \{1, \dots, k\}$. Due to the disjoint union in (2.33), u can be represented as $u = \sum_{j=1}^k v_j \circ \phi_j$ on $J \times (\Omega \cap W_y)$. It is essential to observe, however, that this implies only Hölder continuity for u on each of the disjoint sets $J \times \phi_j^{-1}(\frac{\tau_j}{2} K^-) \subset J \times V_j$ on its own – it is not (yet) clear why the Hölder property should hold "across" different connected components. Let us note that this is exactly the result of Ladyzhenskaya in [101]. In the sequel we will show that our setting allows to derive from this the required *global* Hölder estimates on the sets $J \times (\Omega \cap W_{\mathrm{y}})$.

Let us in the following identify the Hölderian function v_j, defined on $J \times \frac{\tau_j}{2} K^-$, with its unique Hölderian extension on $\overline{J} \times \frac{\tau_j}{2} \overline{K^-}$. The crucial point is here that we imposed in our general ansatz a very special boundary value on the whole Dirichlet part D of the boundary – namely, 0. Indeed, the property $v_j \in \mathrm{L}^2(J; \mathrm{W}^{1,2}_{\tau_j \overline{\Sigma}}(\tau_j K^-))$ implies that $v_j(t, \cdot)$ has trace 0 on $\frac{\tau_j}{2} \overline{\Sigma}$, i.e., vanishes there almost everywhere with respect to the boundary measure $\mathcal{H}^{d-1}$ for almost all $t \in J$, see Remark 2.1.10. However, $v_j(t, \cdot)$ is also a continuous function on $\frac{\tau_j}{2} \overline{K^-}$, and $\frac{\tau_j}{2} \overline{K^-}$ has a Lipschitz-boundary around 0, hence in fact $v_j(t,\cdot) \equiv 0$ on $\frac{\tau_j}{2}\overline{\Sigma}$ for almost all $t \in J$. But then, this time due to continuity in time, v_j must be identically 0 on the *whole* $\overline{J} \times \frac{\tau_j}{2}\overline{\Sigma}$. It is straight forward to verify that the continuation $\hat{v}_j$ of v_j to $J \times \frac{\tau_j}{2}K$ by zero is also Hölder continuous – with the same Hölder-norm as v_j on $J \times \frac{\tau_j}{2}\overline{K^-}$. This means we may extend u via $\hat{u} \coloneqq \sum_j \hat{v}_j \circ \phi_j$ to the set $J \times W_{\mathrm{y}}$ (which indeed is an extension of $u = \sum_j v_j \circ \phi_j$ due to $\hat{v}_j = v_j$ and $\hat{v}_i = 0$ on $\phi_j^{-1}(\frac{\tau_j}{2}K^-)$ for $i \neq j$) and obtain a Hölder-continuous function, such that $u = \hat{u} \restriction W_{\mathrm{y}} \cap \Omega$ is also Hölderian on $W_{\mathrm{y}} \cap \Omega$ with the same estimates.

Let us inspect the corresponding Hölder bounds in some more detail: Let $t_1, t_2 \in J$ and $\mathrm{z}_1, \mathrm{z}_2$ be from two different connected components of $W_{\mathrm{y}} \cap \Omega$, that is, $\mathrm{z}_1 \in \phi_j^{-1}(\frac{\tau_j}{2}K^-) \cap W_{\mathrm{y}}$ and $\mathrm{z}_2 \in \phi_i^{-1}(\frac{\tau_i}{2}K^-) \cap W_{\mathrm{y}}$ for $j \neq i$ and let α_j, α_i be the degree of Hölder continuity of $\hat{v}_j$ and $\hat{v}_i$ on $J \times \frac{\tau_j}{2}K$ and $J \times \frac{\tau_i}{2}K$, respectively. We write

$$\begin{aligned}
|u(t_1,\mathrm{z}_1) - u(t_2,\mathrm{z}_1)| &= |v_j(t_1,\phi_j(\mathrm{z}_1)) - v_i(t_2,\phi_i(\mathrm{z}_2))| \\
&= |\hat{v}_j(t_1,\phi_j(\mathrm{z}_1)) - \hat{v}_j(t_2,\phi_j(\mathrm{z}_2)) + \hat{v}_i(t_1,\phi_i(\mathrm{z}_1)) - \hat{v}_i(t_2,\phi_i(\mathrm{z}_2))| \\
&\qquad \le (1 \vee \mathfrak{l}_{\phi_j})|\hat{v}_j|_{\alpha_j} \|(t_1,\mathrm{z}_1) - (t_2,\mathrm{z}_2)\|^{\alpha_j} \\
&\qquad\qquad + (1 \vee \mathfrak{l}_{\phi_i})|\hat{v}_i|_{\alpha_i} \|(t_1,\mathrm{z}_1) - (t_2,\mathrm{z}_2)\|^{\alpha_i}
\end{aligned}$$

since $\hat{v}_j(t,\phi_j(\mathrm{z}_2)) = \hat{v}_i(t,\phi_i(\mathrm{z}_1)) = 0$ for all $t \in J$. This shows that the Hölder seminorm of u may be estimated as follows, using $\alpha^* = \min_{j \in \{1,\dots,k\}} \alpha_j = \alpha_{j^*}$:

$$\begin{aligned}
|u|_{\alpha^*} \le \max_{j \in \{1,\dots,k\}} \Big(&(1 \vee \mathfrak{l}_{\phi_j})|\hat{v}_j|_{\alpha_j} \mathrm{diam}(J \times (W_{\mathrm{y}} \cap \Omega))^{\alpha_j - \alpha^*} \\
&+ (1 \vee \mathfrak{l}_{\phi_{j^*}})|\hat{v}_{j^*}|_{\alpha^*} \Big).
\end{aligned}$$

In particular, the Hölder seminorm estimate does not depend on all k of connected components of $W_{\mathrm{y}} \cap \Omega$ but only on two of those at once.

Now we have achieved the following: There exist constants

$\alpha(\mathrm{x}_1), \ldots, \alpha(\mathrm{x}_{m_0})$ and $\alpha(\mathrm{y}_1), \ldots, \alpha(\mathrm{y}_{m_1})$, such that

$$\sup_{f \in \mathbb{B}^{\mathrm{step}}_{s,q}} \|u_f\|_{\mathrm{C}^{\alpha(\mathrm{x}_i)}(J \times B_{\mathrm{x}_i})} < \infty \tag{2.34}$$

and

$$\sup_{f \in \mathbb{B}^{\mathrm{step}}_{s,q}} \|u_f\|_{\mathrm{C}^{\alpha(\mathrm{y}_\ell)}(J \times (W_{\mathrm{y}_l} \cap \Omega))} < \infty \tag{2.35}$$

for each $i \in \{1, \ldots, m_0\}$ and $\ell \in \{1, \ldots, m_1\}$, and these suprema are even uniform for all coefficient functions $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$. Diminishing the $\alpha(\mathrm{x}_i)$ and $\alpha(\mathrm{y}_l)$ in (2.34) and (2.35) to their common minimum, called α, (2.34) and (2.34) certainly remain true and we have Hölder-continuity of degree α on each of the sets $B_{\mathrm{x}_1}, \ldots, B_{\mathrm{x}_{m_0}}, W_{\mathrm{y}_1} \cap \Omega, \ldots, W_{\mathrm{y}_{m_1}} \cap \Omega$.

Step 5: In order to deduce *global* Hölder continuity from the previous considerations, we need the following "globalization" lemma:

Lemma 2.1.33. *There exists an $\varepsilon > 0$ such that, for every $\mathrm{x} \in \Omega$, the balls in Ω with center x and radius not larger than ε lie* completely *in at least one of the sets B_{x_i} or W_{y_l}.*

Proof. Consider the function

$$\overline{\Omega} \ni \mathrm{y} \mapsto \varepsilon(\mathrm{y}) := \frac{1}{m_0 + m_1}\Big(\sum_{i=1}^{m_0} \operatorname{dist}\big(\mathrm{y}, \mathbb{R}^d \setminus B_{\mathrm{x}_i}\big) + \sum_{\ell=1}^{m_1} \operatorname{dist}\big(\mathrm{y}, \mathbb{R}^d \setminus W(\mathrm{y}_\ell)\big)\Big).$$

This function is continuous and strictly positive, since every $\mathrm{y} \in \overline{\Omega}$ is contained in at least one of the sets B_{x_i} or W_{y_ℓ}. Therefore, it has to attain its minimum, say, $\varepsilon > 0$. Then it is straight forward to see that this ε fulfills the asserted condition, since at least one summand in the definition has to be bigger or equal to $\varepsilon(\mathrm{y})$ for each $\mathrm{y} \in \overline{\Omega}$. □

Now Lemma 2.1.33 in combination with the locality of the Hölder property, cf. the considerations before Remark 1.2.4, allows to fall back to the sets B_{x_i} and $W_{\mathrm{y}_\ell} \cap \Omega$ and thus implies *global* Hölder bounds on Q, and this uniformly in $f \in \mathbb{B}^{\mathrm{step}}_{s,q}$ and in $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$.

Step 6: The previous considerations show that, for each $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$, the linear mapping $(\partial + \mathcal{A}_\mu)^{-1}$ maps step functions from bounded sets in $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$ into bounded sets in the space $\mathrm{C}^\alpha(Q)$, the bounds being uniform in $\mu_\bullet, \mu^\bullet$. Consequently, these mappings are equicontinuous with respect to $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$. Since the step functions are dense in $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$, the mappings hence possess extensions to the whole $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$ which are still equicontinuous. This was the claim in Theorem 2.1.4.

2.1.5 Inhomogeneous Dirichlet boundary data

Up to now, the fundamental difference between the approach of Ladyzhenskaya et al. in [101] and ours consists in the fact that here only the zero Dirichlet datum is allowed, which allowed to deduce *global* Hölder continuity for the solution (it is clear that also *constant* nonzero data is admissible by obvious modifications). In this chapter we will show a way how to admit (nonconstant) nonzero Dirichlet data – without losing the classical Hölder property for the solution. We restrict ourselves to the case where the Dirichlet datum does *not* depend on time. Moreover, aiming at Hölder continuity for the solution in both time and space, it is clear that the initial value must admit the correct boundary behavior. In particular, in this context one can never expect that a solution with initial value 0 admits a nonzero Dirichlet datum.

Let us recall from Chapter 1.2.4 the Besov spaces $\mathrm{B}^{1-1/q}_{q,q}(D)$. Since D was a $(d-1)$-set by Theorem 1.3.16, Proposition 1.2.60 tells us that the trace operator $\mathcal{R}_D$ maps $\mathrm{W}^{1,q}(\mathbb{R}^d)$ linearly and continuously *onto* $\mathrm{B}^{1-1/q}_{q,q}(D)$ and there exists a corresponding linear continuous extension operator $\mathcal{E}_D$. Now consider a function $\iota \in \mathrm{B}^{1-1/q}_{q,q}(D)$. There exists a function $v = \mathcal{E}_D \iota \in \mathrm{W}^{1,q}(\mathbb{R}^d)$ which satisfies $\mathcal{R}_D v = \iota$ on D. Moreover, the restriction $\mathfrak{R}_\Lambda v$

is an element of $\mathrm{H}^{1,q}(\Lambda) \subset \mathrm{W}^{1,q}(\Lambda)$ – we do not know that the spaces are equal because Λ is in general *not* a $\mathrm{W}^{1,q}$-extension domain, cf. Figure 1.3. Let us denote $\mathfrak{R}_\Lambda v = \mathfrak{R}_\Lambda \mathcal{E}_D \iota$ by $u_0(\iota)$.

Definition 2.1.34. Let $\iota \in \mathrm{B}^{1-1/q}_{q,q}(D)$ and $g \in \mathrm{L}^2(J; \mathrm{W}^{-1,2}_D(\Omega))$, together with $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$ be given, and set $u_0(\iota) = \mathfrak{R}_\Lambda \mathcal{E}_D \iota$. Then we say that the function $w \in \mathbb{W}^{1,2}(J; \mathrm{W}^{-1,2}(\Omega), \mathrm{W}^{1,2}(\Omega))$ is a solution of the equation

$$\begin{aligned} w'(t) - \nabla \cdot \mu(t,\cdot)\nabla w(t) + w(t) = g(t) \quad \text{in } \mathrm{W}^{-1,2}_D(\Omega), \\ w(t) \restriction D = \iota, \\ u(T_0) = u_0(\iota) \end{aligned} \tag{2.36}$$

for almost all $t \in J$, if w is of the form $w = u + u_0(\iota)$, where $u \in \mathbb{W}^{1,2}_0(J; \mathrm{W}^{-1,2}_D(\Omega), \mathrm{W}^{1,2}_D(\Omega))$ is the solution of

$$\begin{aligned} u'(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) + u(t) = g(t) + \nabla \cdot \mu(t,\cdot)\nabla u_0(\iota) - u_0(\iota) \\ \text{in } \mathrm{W}^{-1,2}_D(\Omega) \quad \text{for a.a. } t \in J. \end{aligned}$$

The following theorem upgrades Theorem 2.1.4 to inhomogeneous Dirichlet data.

Theorem 2.1.35. *Suppose that $q > d$ and $s > 2(1 - \frac{d}{q})^{-1}$, and let $\iota \in \mathrm{B}^{1-1/q}_{q,q}(D)$ as well as $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$ and $g \in \mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$.*

(i) The inhomogeneous equation (2.36) *admits a unique solution $w = w_{\iota,g}$ in the sense of Definition 2.1.34 which is even Hölder-continuous on $J \times \Omega$, i.e., there is $\alpha > 0$ such that $w \in \mathrm{C}^\alpha(Q)$.*

(ii) The solution operator which maps the right-hand side g and Dirichlet data ι to $w_{g,\iota}$, i.e., (symbolically)

$$\left(\partial + \mathcal{A}_\mu, \mathcal{R}_D\right)^{-1} \colon \mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega)) \times \mathrm{B}^{1-1/q}_{q,q}(D) \to \mathrm{C}^\alpha(Q),$$

is linear and equicontinuous with respect to $\mu \in \mathrm{L}^\infty(Q; \mathbb{M}_d(\mu_\bullet, \mu^\bullet))$.

Proof. The theorem follows essentially from applying Theorem 2.1.4 to the equation

$$u'(t) - \nabla\cdot\mu(t,\cdot)\nabla u(t) + u(t) = g(t) + \nabla\cdot\mu(t,\cdot)\nabla u_0(\iota) - u_0(\iota) \qquad \text{in } \mathrm{W}^{-1,2}_D(\Omega) \quad \text{for a.a. } t\in J.$$

This is justified because $u_0(\iota)\in \mathrm{H}^{1,q}(\Omega)\subset \mathrm{W}^{1,q}(\Omega)$ is mapped to the space $\mathrm{W}^{-1,q}_\emptyset(\Omega)$ by $-\nabla\cdot\mu(t,\cdot)\nabla$ for almost every $t\in J$ and we can interpret the resulting operator as one acting on $\mathrm{W}^{1,q'}_D(\Omega)$ by restriction. Moreover, due to Remark 1.5.4 and $u_0(\iota)$ being constant in time, we know that the right hand sides involving $u_0(\iota)$ are indeed in $\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))$ for *every* $1\le s\le\infty$. Hence, the foregoing equation admits the unique solution $u\in \mathbb{W}^{1,2}_0(J;\mathrm{W}^{-1,2}_D(\Omega),\mathrm{W}^{1,2}_D(\Omega))$ which is additionally Hölder-continuous on Q. Since $u_0(\iota)=\mathfrak{R}_\Lambda\mathcal{E}_D\iota\in\mathrm{H}^{1,q}(\Omega)$ for $q>d$ (see Proposition 1.2.60 and Definition 1.2.21), we also know that $u_0(\iota)$ is Hölder-continuous on Ω and, since it is constant in time, on Q (with continuous embedding). But this makes $w=u+u_0(\iota)$ Hölder-continuous on Q as well and proves the first point.

For the second, observe that

$$u(t) = (\partial+\mathcal{A}_\mu)^{-1}\big(g(t)+\nabla\cdot\mu(t,\cdot)\nabla u_0(\iota)-u_0(\iota)\big)$$

with $(\partial+\mathcal{A}_\mu)^{-1}$ as in Theorem 2.1.4, hence we find for $\alpha>0$ being the degree of Hölder-continuity of w,

$$\begin{aligned}
\|w\|_{\mathrm{C}^\alpha(Q)} &\le \|u\|_{\mathrm{C}^\alpha(Q)} + \|u_0(\iota)\|_{\mathrm{C}^\alpha(Q)}\\
&\le C_{\mu_\bullet,\mu^\bullet}\Big(\|g+\nabla\cdot\mu\nabla u_0(\iota)-u_0(\iota)\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))}\Big)\\
&\qquad + C\|\iota\|_{\mathrm{B}^{1-1/q}_{q,q}(D)}\\
&\le C_{\mu_\bullet,\mu^\bullet}\Big(\|g\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))} + C_D(\mu^\bullet+1)\|\iota\|_{\mathrm{B}^{1-1/q}_{q,q}(D)}\Big)\\
&\qquad + C\|\iota\|_{\mathrm{B}^{1-1/q}_{q,q}(D)}\\
&= C_{\mu_\bullet,\mu^\bullet,D}\big(\|g\|_{\mathrm{L}^s(J;\mathrm{W}^{-1,q}_D(\Omega))}+\|\iota\|_{\mathrm{B}^{1-1/q}_{q,q}(D)}\big).
\end{aligned}$$

We have used that $(\partial + \mathcal{A})^{-1}$ is equicontinuous w.r.t. $\mu_\bullet, \mu^\bullet$ between $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$ and $\mathrm{C}^\alpha(Q)$, and that $\mathcal{E}_D \colon \mathrm{B}^{1-1/q}_{q,q}(D) \to \mathrm{W}^{1,q}(\mathbb{R}^d)$ as well as the obvious inequality $\|\cdot\|_{\mathrm{H}^{1,q}(\Omega)} \leq \|\cdot\|_{\mathrm{W}^{1,q}(\mathbb{R}^d)}$. □

We close this section by some remarks to the foregoing theorem.

Remark 2.1.36.

(i) Following the strategy to split off the initial value in Definition 2.1.34 or Theorem 2.1.35 requires $u_0(\iota)$ to be in the domain of $-\nabla \cdot \mu(t, \cdot)\nabla + 1$ for each $t \in J$. As we have learned in Chapter 1.5, this is *in general* only to be achieved if $u_0(\iota) \in \mathrm{W}^{1,q}(\Omega)$, cf. Remark 1.5.4. Hence, in view of Proposition 1.2.60, the space $\mathrm{B}^{1-1/q}_{q,q}(D)$ for the boundary values on D is exactly the “optimal” one.

(ii) Let us remark that, due to $q > d$, the solution w as in Theorem 2.1.35 satisfies $w \restriction D = g$ indeed in the sense of pointwise restriction.

(iii) Theorem 2.1.35 remains true, with exactly the same proof, for the differential operators $\partial + \mathcal{A}_\mu + \mathcal{B}_\gamma$ for $\gamma \in \mathrm{L}^\infty(J \times \overline{N}; \lambda \otimes \omega, \mathbb{R}^+_0)$, cf. Corollary 2.1.6.

2.2 Quasilinear parabolic equations in divergence form via maximal parabolic regularity

Let us return to the model equation for this chapter,

$$u'(t) - \nabla \cdot \sigma(u)(t)\rho\nabla u(t) + u(t) = F(u)(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0. \tag{2.1}$$

We collect and review the known existence– and uniqueness results for abstract quasilinear parabolic evolution equations, and use them to give such a result for the special case (2.1). The geometric framework in which we work will be that of Ω being Lipschitz around $\partial\Omega \setminus D$.

The following assumptions hold true for the rest of this chapter:

(i) The set $\Omega \subset \mathbb{R}^d$ is a bounded domain and D (like *Dirichlet*) is a closed subset of $\partial\Omega$. The cases $D = \emptyset$ and $D = \partial\Omega$ are not excluded. We suppose that $\Omega \cup D$ is Lipschitz around $\partial\Omega \setminus D$. In all what follows, $\partial\Omega \setminus D$ will be denoted by N (like *Neumann*).

(ii) We consider a finite interval $J = (T_0, T_1) \subset \mathbb{R}_0^+$.

(iii) All Banach spaces and all occurring functions are supposed to be real ones, i.e., we are working in a *real* setting in the sense of Chapter 1.6.

It was already noted in the introduction of this thesis that we may have to deal with non-global solutions to (2.1), hence we formalize the notion of a *maximal solution* as follows. For compatibility reasons with respect to the following theorems, we therefore switch to a general differential operator A:

Definition 2.2.1 (Maximal & global solution)**.** Let X, Y with $Y \hookrightarrow_{\mathrm{d}} X$ be Banach spaces and let $1 < r < \infty$. Set $J = (T_0, T_1)$ to be a given time interval and consider the quasilinear equation

$$u'(t) + A(u)u(t) = F(u)(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0 \tag{2.37}$$

for suitable operators A, F. Assume that for every feasible initial value $u_0 \in (X, Y)_{1/r',r}$, there is a number $T_0 < T^\bullet \leq T_1$ and a unique solution u of (2.37) on $J^\bullet = (T_0, T^\bullet)$ such that $u \in \mathbb{W}^{1,r}(T_0, T_\bullet; X, Y)$ for every $T_\bullet \in J^\bullet$.

(i) If $T^\bullet = T_1$ or $u \notin \mathbb{W}^{1,r}(J^\bullet; X, Y)$, then we say that $J^\bullet(u_0)$ is the *maximal interval of existence* and u is the *maximal solution.*

(ii) If $T^\bullet = T_1$ and $u \in \mathbb{W}^{1,r}(J; X, Y)$, then we say that u is the *global solution.*

Otherwise, we just say that u is a *local* solution.

Of course, the assumed situation in Definition 2.2.1 is exactly the one which is produced in the proofs of Theorem 2.2.4 and 2.2.7 below.

Remark 2.2.2. The somewhat convoluted formulation in Definition 2.2.1 is necessary because of the embedding $\mathbb{W}^{1,r}(I;X,Y) \hookrightarrow \mathrm{C}(\overline{I};(X,Y)_{1/r',r})$ for any interval I provided by Proposition 1.4.3: If, for some possibly maximal interval $J^\bullet(u_0)$, the solution u is already in $\mathbb{W}^{1,r}(J^\bullet(u_0);X,Y)$, then $u(T^\bullet(u_0)) \in (X,Y)_{1/r',r}$ and we may extend u beyond $T^\bullet(u_0)$ by using $u(T^\bullet(u_0))$ as a new initial value in (2.37). Hence, we must characterize maximal intervals of existence $J^\bullet(u_0) \neq J$ exactly by the property that $u \notin \mathbb{W}^{1,r}(J^\bullet(u_0);X,Y)$ – or, equivalently, $\lim_{t \nearrow T^\bullet(u_0)} u(t)$ does not exist in $(X,Y)_{1/r',r}$, cf. [127, Cor. 3.2].

Now, the “all-in-one” theorem regarding existence and uniqueness of general abstract quasilinear parabolic equations which seems the maximum to be achieved in this very general framework is the following one by Amann which can be seen as the culmination of his efforts regarding quasilinear parabolic equations via maximal parabolic regularity. For the formulation, we need the notion of a *Volterra* map:

Definition 2.2.3 (Volterra property)**.** Let $\mathcal{X},\mathcal{Y}$ be vector spaces, let $\mathcal{F}\colon \mathcal{X} \to \mathcal{Y}$ and suppose that $u,v\colon J \to \mathcal{X}$ are functions on a time interval $J = (T_0,T_1)$ mapping into $\mathcal{X}$. Then we say that $\mathcal{F}$ has the *Volterra property* if the implication

$$(u \restriction J_\bullet = v \restriction J_\bullet) \quad \Longrightarrow \quad \mathcal{F}(u) \restriction J_\bullet = \mathcal{F}(v) \restriction J_\bullet.$$

is true for every subinterval $J_\bullet = (T_0,T_\bullet) \subseteq J$.

The Volterra property is a certain “determinism” property of $\mathcal{F}$ in the sense that the function given by $\mathcal{F}(u) \restriction J_\bullet$ does not depend on the values of u on the “future” time interval $(T_\bullet,T_1)$. This means that for a Volterra map $\mathcal{F}$, the function value $\mathcal{F}(u)(t)$ for $t \in J_\bullet$ is allowed to depend on *all* function values of u up to $T_\bullet$, as opposed to only on $u(t)$ alone, and may thus be a *nonlocal* map. Now the existence– and uniqueness theorem of

Amann is as follows:

Theorem 2.2.4 ([8, Thm. 3.1]). *Let X, Y be Banach spaces such that $Y \hookrightarrow_{\mathrm{d}} X$ and let $1 < r < \infty$ and $J = (T_0, T_1)$. Let moreover the following assumptions be satisfied:*

(i) The operator A maps $\mathbb{W}^{1,r}(J; X, Y)$ into $\mathrm{L}^\infty(J; \mathscr{L}(Y; X))$, has the Volterra property, and is uniformly Lipschitz-continuous on bounded sets.

(ii) The operator $A(u) \restriction J_\bullet$ satisfies nonautonomous maximal parabolic L^r regularity on X over $J_\bullet$ with domain Y for every $u \in \mathbb{W}^{1,r}(J; X, Y)$ and every subinterval $J_\bullet := (T_0, T_\bullet) \subseteq J$.

(iii) The operator F maps $\mathbb{W}^{1,r}(J; X, Y)$ into $\mathrm{L}^r(J; X)$, has the Volterra property, and there exists a number $r < s \leq \infty$ such that $F - F(0)$ is uniformly Lipschitz-continuous on bounded sets in $\mathbb{W}^{1,r}(J; X, Y)$ with values in $\mathrm{L}^s(J; X)$.

Then, for every initial value $u_0 \in (X, Y)_{1/r', r}$, the equation

$$u'(t) + A(u)u(t) = F(u)(t) \quad \text{in } X \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0,$$

admits a unique maximal solution u on the maximal interval of existence $J^\bullet(u_0)$ in the sense of Definition 2.2.1.

Remark 2.2.5. The original reference to Theorem 2.2.4 should be the paper [10]. Unfortunately, the paper contains a slight oversight in the formulation of the theorem, as pointed out by its author in [8, Rem. 3.2], which is why we cited the version from [8] – the version in [10] lacks the requirement that the operators $A(u)$ satisfy nonautonomous maximal parabolic regularity over every shorter subinterval. On the other hand, the reader is to be advised that the notion of the Volterra property seems incorrect in [8], but is correctly done in [10].

The not-quite-as-general theorem which we state for comparison purposes and the sake of completeness is the one by Prüss, for which we introduce the notion of a Carathéodory function:

Definition 2.2.6 (Carathéodory function)**.** Let X, Y be Banach spaces. We say that a function $F\colon J \times X \to Y$ is a *Carathéodory function* if $F(\cdot, x)\colon J \to Y$ is measurable for every $x \in X$ and $F(t, \cdot)\colon X \to Y$ is continuous for every $t \in J$.

The Carathéodory property of F ensures that the joint mapping $F\colon J \times X \to Y$ is (Borel-) measurable, cf. [36, Lem. III-14]. This means that if $v\colon J \to X$ is Bochner-measurable, then $t \mapsto F(t, v(t))$ is Bochner-measurable as a function from J to Y.

Theorem 2.2.7 ([127, Thm. 3.1])**.** *Let X, Y be Banach spaces such that $Y \hookrightarrow_{\mathrm{d}} X$ and let $1 < r < \infty$ and $J = (T_0, T_1)$. Let moreover the following assumptions be satisfied:*

(i) The mapping $A\colon J \times (X, Y)_{1/r',r} \to X$ is continuous. Moreover, $v \mapsto A(\cdot, v)$ maps $(X, Y)_{1/r',r}$ into $\mathrm{L}^\infty(J; \mathscr{L}(Y; X))$ and is uniformly continuous on bounded sets.

(ii) The mapping $F\colon J \times (X, Y)_{1/r',r} \to X$ is a Carathéodory function. Moreover, $F(\cdot, 0) \in \mathrm{L}^r(J; X)$ and for every $R > 0$ there exists a function $\phi_R \in \mathrm{L}^r(J)$ such that

$$\big\|F(t, u) - F(t, v)\big\|_X \le \phi_R(t) \|u - v\|_{(X,Y)_{1/r',r}} \quad \textit{for all } u, v \in \overline{\mathbb{B}(0, R)}$$

is true for almost all $t \in J$, where $\mathbb{B}(0, R)$ denotes the ball with radius R in $(X, Y)_{1/r',r}$.

(iii) The operators $A(t, v)$ satisfy maximal parabolic regularity in X with domain Y for every $(t, v) \in J \times (X, Y)_{1/r',r}$.

Then, for every initial value $u_0 \in (X, Y)_{1/r',r}$, the equation

$$u'(t) + A(t, u(t))u(t) = F(t, u(t)) \quad \textit{in } X \quad \textit{for a.a. } t \in J, \quad u(T_0) = u_0$$

admits a unique maximal solution u on the maximal interval of existence $J^\bullet(u_0)$ in the sense of Definition 2.2.1.

Comparing the two fundamental theorems above, it is apparent that the assumptions are essentially the same, with the theorem of AMANN relying on a more "total" approach in time via the Volterra property and *nonautonomous* maximal parabolic regularity, but this is in the nature of things due to the ansatz for nonlocal equations.

Remark 2.2.8. The ability of Theorem 2.2.4 to handle *nonlocal* dependencies on the searched-for function u itself is a quite astonishing feature. This allows for instance to treat time-delayed equations (see [8,11], also for more examples), but also coupled *systems* of a quasilinear and multiple parabolic equations by solving the adjacent equations in dependence of the function u, determined by the quasilinear equation, and inserting this dependence into said equation for u. Since the other functions are subject to evolution equations themselves, they depend on the function u over the *whole* time horizon, as opposed to depending on u only pointwisely in time via $u(t)$. This technique has been successfully applied to a classical model in chemotaxis in the L^p setting by HORSTMANN, REHBERG and the author [89], and is also the backbone of the treatment of the thermistor problem in Chapter 4, albeit "only" for an additional elliptic equation. There, one may in principle solve the additional equation for each point in time in dependency of $u(t)$ alone, if u enters the equation locally-in-time. This would place the procedure within the setting of Theorem 2.2.7.

Remark 2.2.9. A difference in the assumptions of Theorems 2.2.4 and 2.2.7 is the regularity gap required by AMANN for the right-hand sides F, i.e., the requirement that $F - F(0)$ maps into $\mathrm{L}^s(J;X)$ for some $r < s \leq \infty$. Such a regularity gap is routinely needed to force the fixed point mapping in function spaces used in the proof to be a contraction on smaller time intervals $J_\bullet$ via

$$\begin{aligned}\big\|F(u) - F(v)\big\|_{\mathrm{L}^r(J_\bullet;X)} &\leq \lambda(J_\bullet)^{\frac{1}{r}-\frac{1}{s}}\big\|F(u) - F(v)\big\|_{\mathrm{L}^s(J_\bullet;X)} \\ &\leq L_F\lambda(J_\bullet)^{\frac{1}{r}-\frac{1}{s}}\|u - v\|_{\mathbb{W}^{1,r}(J;X,Y)}\end{aligned}$$

for $u, v \in \mathbb{W}^{1,r}(J;X,Y)$ from a bounded set in $\mathbb{W}^{1,r}(J;X,Y)$. However, such a construction is implicitly contained in the assumptions of

Theorem 2.2.7 by the maximal regularity embedding $\mathbb{W}^{1,r}(J;X,Y) \hookrightarrow \mathrm{C}(\overline{J};(X,Y)_{1/r',r})$, cf. Proposition 1.4.3, since the Lipschitz property of F as assumed there implies that

$$\big\|F(\cdot,u(\cdot)) - F(\cdot,v(\cdot))\big\|_{\mathrm{L}^r(J_\bullet;X)} \leq \lambda(J_\bullet)^{\frac{1}{r}}\,\|\phi_R\|_{\mathrm{L}^r(J)}\|u-v\|_{\mathrm{C}(\overline{J};(X,Y)_{1/r',r})}$$

for $u,v \in \mathbb{W}^{1,r}(J;X,Y)$ from a bounded set in $\mathbb{W}^{1,r}(J;X,Y)$ enclosed in a ball of radius R. This required gap in Theorem 2.2.4 is essentially the price one has to pay for being able to treat nonlocal equations. If we assume that each of the Lipschitz-constant functions ϕ_R in Theorem 2.2.7 are in fact from $\mathrm{L}^s(J)$ for $r < s \leq \infty$, then the assumptions in Theorem 2.2.7 imply those of Theorem 2.2.4.

We state the main existence- and uniqueness result for divergence-gradient operators as a special case of Theorem 2.2.4. Thanks to the extensive preparations in Chapter 1, the proof that the divergence-gradient setting satisfies the assumptions in Theorem 2.2.4 will be very concise. We will stay in the regime of constant domains for the divergence-gradient operators on $\mathrm{W}^{-1,q}_D(\Omega)$ and fix these domains to $\mathrm{W}^{1,q}_D(\Omega)$, using the minimal assumption $\mathcal{D}_q(\rho) = \mathrm{W}^{1,q}_D(\Omega)$. This furthermore allows to use the continuity properties as established in Chapter 1.5 (see Proposition 1.5.21). Since ρ will then be fixed, it makes sense to introduce another abbreviation of the divergence-gradient operators. Recalling the operators $\mathcal{A}_\mu$ for $\mu \in \mathrm{L}^\infty(Q;\mathbb{M}_d)$ used in Chapter 2.1, cf. (2.7) and (2.8), we set

$$\mathcal{A}_\rho(\varphi) \coloneqq \mathcal{A}_{\rho\varphi} \colon \mathrm{L}^2(J;\mathrm{W}^{1,2}_D(\Omega)) \to \mathrm{L}^2(J;\mathrm{W}^{-1,2}_D(\Omega))$$

for $\rho \in \mathrm{L}^\infty(\Omega;\mathbb{M}_d)$ and $\varphi \in \mathrm{L}^\infty(Q)$. In the same fashion, we also use the boundary form operators $\mathcal{B}_\gamma$ for $\gamma \in \mathrm{L}^\infty(J \times \overline{N};\lambda \otimes \omega)$, cf. (2.11) and (2.12). Then we have the following local existence- and uniqueness result for quasilinear parabolic evolution equations in divergence form on rough domains:

Theorem 2.2.10 (Local existence and uniqueness)**.** *Let $\Omega \cup D \subset \mathbb{R}^d$ be Lipschitz around $\partial\Omega \setminus D$ for $d \in \{2,3\}$ and let the following assumptions be satisfied for some $q > d$ and $r > 2(1-\frac{d}{q})^{-1}$:*

(i) The coefficient functions satisfy the following:

a) σ maps $\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ into $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$, satisfies $0 < \sigma_\bullet \le \sigma(u) \le \sigma^\bullet$ on $\overline{Q}$ for all $u \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$, is uniformly Lipschitz-continuous on bounded sets, and has the Volterra property.

b) $\rho \in \mathrm{L}^\infty(\Omega; \mathbb{M}_d(\rho_\bullet, \rho^\bullet))$ and

$$-\nabla \cdot \rho \nabla + 1 \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,q}_D(\Omega); \mathrm{W}^{-1,q}_D(\Omega)\big),$$

i.e., $\mathcal{D}_q(\rho) \doteq \mathrm{W}^{1,q}_D(\Omega)$. Moreover, we assume the assumptions of Proposition 1.5.5 to be satisfied for $-\nabla \cdot \rho\nabla$, and $\gamma \in \mathrm{C}(\overline{J}; \mathrm{L}^\infty(\overline{N}; \omega))$.

(ii) The operator F maps $\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ into $\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))$, has the Volterra property, and there exists a number $r < s \le \infty$ such that $F - F(0)$ is uniformly Lipschitz-continuous on bounded sets in $\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ with values in $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$.

Then, for every initial value $u_0 \in \big(\mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\big)_{1/r',r}$, the equation

$$u'(t) + (\mathcal{A}_\rho(\sigma(u)) + \mathcal{B}_\gamma)u(t) = F(u)(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0, \tag{2.38}$$

admits a unique maximal solution u on the maximal interval of existence in the sense of Definition 2.2.1.

Remark 2.2.11. Let us note that the idea to use Theorems 2.2.4 or 2.2.7 to obtain local existence- and uniqueness results for quasilinear parabolic evolution equations in divergence form is of course not new, see e.g. [80,84]. We have just tried to put it in a context which is as broad as possible to have a general result basing in recent developments at hand.

Before we prove it, let us point out a critical consequence of the assumptions of Theorem 2.2.10: Due to the assumptions of Proposition 1.5.5 being satisfied in Theorem 2.2.10, we have the Kato square root property together with maximal Sobolev regularity for $-\nabla \cdot \rho\nabla$ at our disposal. Thus

$$\bigl(\mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\bigr)_{1/r',r} \hookrightarrow \mathrm{C}^\beta(\Omega) \hookrightarrow\hookrightarrow \mathrm{C}(\overline{\Omega}) \tag{2.39}$$

for some $\beta > 0$ due to $r > 2(1-\frac{d}{q})^{-1}$, see Lemma 1.5.25, and also

$$\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)) \hookrightarrow \mathrm{C}^\alpha(Q) \hookrightarrow\hookrightarrow \mathrm{C}(\overline{Q}) \tag{2.40}$$

for some $\alpha > 0$. These embeddings will prove invaluable in the following and are the main reason why we need at least $\Omega \cup D$ Lipschitz around $\partial\Omega \setminus D$ as the regularity of the spatial domain. As noted in Remark 1.5.6, we can enforce the Kato square root property for $-\nabla \cdot \rho\nabla$ "by hand" by assuming ρ to take values in the space of *symmetric* matrices, and it is always satisfied when $\Omega \cup D$ is a $\mathrm{W}^{1,p}$-extension domain; this is true in particular if it is regular in the sense of Gröger.

Proof of Theorem 2.2.10. We prove that the assumptions of the theorem of Amann, Theorem 2.2.4, are satisfied. For F, this is immediate because the assumptions on F are exactly the ones from Theorem 2.2.4. We moreover have $\sigma(u)(t) \in \mathrm{C}(\overline{\Omega})$ for every $t \in \overline{J}$ and every $u \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ due to the assumption on σ. Hence, Proposition 1.5.21 and Lemma 1.5.14 show that the operator

$$A(u) \coloneqq \mathcal{A}_\rho(\sigma(u)) + \mathcal{B}_\gamma$$

maps $\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ into $\mathrm{L}^\infty(J; \mathscr{L}(\mathrm{W}^{-1,q}_D(\Omega); \mathrm{W}^{1,q}_D(\Omega)))$. It moreover inherits the Volterra property and Lipschitz-continuity from σ, cf. also Remark 1.5.4, and $A(u)$ satisfies nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,q}_D(\Omega)$ with domain $\mathrm{W}^{1,q}_D(\Omega)$ over J and every subinterval $J_\bullet \subseteq J$ for every $u \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ by

Lemma 1.5.23. This shows that the assumptions of Theorem 2.2.4 are satisfied. □

2.2.1 Global solutions

The existence- and uniqueness theorem above yields optimal regularity of the solution under minimal assumptions, however, at the price of maximal, but *local in time* solutions in the sense of Definition 2.2.1. While global solutions are not to be expected from the very general assumptions with no requirements regarding growth or monotonicity at all—recall (3) in the preface—, it turns out that even in the case where F is *constant* w.r.t. the searched-for function u, we do not obtain global solutions and it is not obvious from the proofs how this should be achieved. This implies that whatever assumptions one poses for F, the quasilinear structure may still "destroy" existence of global solutions.

We are able to remove this drawback using the uniform Hölder estimates established in Chapter 2.1 at the cost of a much stronger Lipschitz condition, cf. the discussion below, and on domains Ω which are volume-preserving generalized regular in the sense of Gröger, together with D. The theorem is as follows:

Theorem 2.2.12 (Global existence and uniqueness). *Let $\Omega \cup D \subset \mathbb{R}^d$ be volume-preserving generalized regular in the sense of Gröger for $d \in \{2,3\}$. Let moreover the following assumptions be satisfied for some $q > d$ and $r > 2(1-\frac{d}{q})^{-1}$:*

(i) The coefficient functions satisfy the following:

a) σ maps $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ into itself, satisfies $0 < \sigma_\bullet \leq \sigma(u) \leq \sigma^\bullet$ on $\overline{Q}$ for all $u \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$, is uniformly Lipschitz-continuous on bounded sets, and has the Volterra property.

b) $\rho \in \mathrm{L}^\infty(\Omega;\mathbb{M}_d(\rho_\bullet,\rho^\bullet))$ and

$$-\nabla\cdot\rho\nabla + 1 \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,q}_D(\Omega);\mathrm{W}^{-1,q}_D(\Omega)\big),$$

i.e., $\mathcal{D}_q(\rho) \doteq \mathrm{W}^{1,q}_D(\Omega)$. *Moreover, we assume the assumptions of Proposition 1.5.5 to be satisfied for* $-\nabla \cdot \rho\nabla$*, and* $\gamma \in \mathrm{C}(\overline{J}; \mathrm{L}^\infty(\overline{N}; \omega))$.

(ii) The mapping F *maps* $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ *into* $\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))$*, has the Volterra property, and there exists a number* $r < s \leq \infty$ *such that* $F - F(0)$ *is uniformly Lipschitz-continuous on bounded sets in* $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ *with values in* $\mathrm{L}^s(J; \mathrm{W}^{-1,q}_D(\Omega))$. *Finally,*

$$\sup_{u \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))} \|F(u)\|_{\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))} =: C_F < \infty. \tag{2.41}$$

Then, for every initial value $u_0 \in \big(\mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\big)_{1/r',r}$*, the equation*

$$u'(t) + (\mathcal{A}_\rho(\sigma(u)) + \mathcal{B}_\gamma)u(t) = F(u)(t) \\ \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0, \tag{2.38}$$

admits a unique global *solution* $u \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$.

Before we prove it, let us first compare the assumptions in Theorem 2.2.12 with the ones of Theorem 2.2.10. First and foremost, we have required $\Omega \cup D$ to be stronger than before, namely volume-preserving generalized regular in the sense of Gröger. The reason has been announced multiple times: Our proof bases fundamentally upon the uniform Hölder estimates from Chapter 2.1 in the form of Theorem 2.1.4, which we were only able to prove in that framework.

Regarding the other assumptions, it was already meantioned in the introduction that one generally has to pose a growth condition on F to obtain *global* solutions, enforced by the uniform boundedness condition (2.41). Unfortunately, the usual sublinear growth condition for the nonlinear inhomogeneity (see e.g. [126, Ch. 6, Thm. 3.3]) is not sufficient for the type of proof we do. On the other hand, a monotonicity requirement in the abstract $\mathrm{W}^{-1,q}_D(\Omega)$ setting for Banach spaces seems also not suitable.

Considering the designated way of proving the theorem via the uniform

Hölder bounds obtained by Theorem 2.1.4, it is not surprising that we work with the space of continuous functions as the fundamental function space in the proof of Theorem 2.2.12 instead of the maximal regularity space as in Theorem 2.2.10, and pose a suitable Lipschitz condition there. However, one has to be aware that the Lipschitz conditions on σ and F in Theorem 2.2.12 are *much* stronger than the corresponding ones in Theorem 2.2.4. This can be seen by the embeddings

$$\left(\mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\right)_{1/r',r} \hookrightarrow \mathrm{C}^\beta(\Omega) \hookrightarrow \mathrm{C}(\overline{\Omega}) \tag{2.39}$$

for some $\beta > 0$ and

$$\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)) \hookrightarrow \mathrm{C}^\alpha(Q) \hookrightarrow \mathrm{C}(\overline{Q}) \tag{2.40}$$

for some $\alpha > 0$, as above. If exemplarily F is well-defined on $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ in the first place, then the assumption on F in Theorem 2.2.10 implies Lipschitz-continuity of F on *compact* sets in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ due to (2.40), whereas the assumption in Theorem 2.2.12 above requires Lipschitz-continuity of F on *bounded* sets in that space. Since we deal with infinite-dimensional function spaces, this "small" detail is an *enormous* leap in strength of assumptions which may be too strong for many problems, cf. also Chapter 4.1.

Summing up, the assumptions in Theorem 2.2.12 are always at least as strong as the ones in Theorem 2.2.10. In particular, every instance of (2.38) satisfying the assumptions of Theorem 2.2.12 also satisfies those of Theorem 2.2.10.

We lastly collect a preliminary result for the divergence-gradient operators in an analogous situation to the one in Theorems 2.2.10 and 2.2.12. This sets the result up for later use and will allow to concentrate on the "essentials" in the proof of Theorem 2.2.12. The result is the promised extension of the nonautonomous maximal parabolic regularity of divergence-gradient operators established in Lemma 1.5.23. Recall the point evaluation δ_{T_0}, cf. (1.32).

Lemma 2.2.13. *Let $\Omega \cup D \subset \mathbb{R}^d$ be Lipschitz around $\partial\Omega \setminus D$ for $d \in \{2,3\}$ and let $\rho \in \mathrm{L}^\infty(\Omega; \mathbb{M}_d(\rho_\bullet))$ such that the assumptions of Proposition 1.5.5 are satisfied for $-\nabla\rho\nabla$, as well as $\gamma \in \mathrm{C}(\overline{J}; \mathrm{L}^\infty(\overline{N}; \omega))$. Assume that $\mathcal{D}_q(\rho) = \mathrm{W}^{1,q}_D(\Omega)$ for some $2 \le q < 2^\star = \frac{2d}{d-2}$. Then, for $1 < r < \infty$ and every $\varphi \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ with $\varphi(t) > 0$ on $\overline{\Omega}$ for every $t \in \overline{J}$, the linear operator*

$$\begin{aligned}(\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma, \delta_{T_0}) \colon \mathbb{W}^{1,r}\big(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\big) \\ \to \mathrm{L}^r\big(J; \mathrm{W}^{-1,q}_D(\Omega)\big) \times \big(\mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)\big)_{1/r',r}\end{aligned} \tag{2.42}$$

is continuously invertible. Let moreover $\varphi_\bullet > 0$ be given and set

$$M_{\varphi_\bullet} := \Big\{\varphi \in \mathrm{C}\big(\overline{J}; \mathrm{C}(\overline{\Omega})\big) \colon \varphi(t) \ge \varphi_\bullet \text{ on } \overline{\Omega} \text{ for every } t \in \overline{J}\Big\}.$$

Then the mapping

$$M_{\varphi_\bullet} \ni \varphi \mapsto \big(\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma, \delta_{T_0}\big)^{-1}$$

is continuous.

Proof. We have already seen in Lemma 1.5.23 that each operator $\mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma$ satisfies nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,q}_D(\Omega)$ with domain $\mathrm{W}^{1,q}_D(\Omega)$ due to the continuity of φ and is a continuous function in time, cf. Lemma 1.5.23. But then nonautonomous maximal parabolic regularity of $\mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma$ is equivalent to continuous invertibility of $(\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma, \delta_{T_0})$ between the stated spaces by Lemma 1.4.16.

Let us consider the continuity assertion: Let $(\varphi_k) \subset \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ be a sequence of functions such that $\varphi(t) \ge \varphi_\bullet$ on $\overline{\Omega}$ for every $t \in \overline{J}$, converging to some φ in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Lambda}))$. Then φ also satisfies $\varphi(t) \ge \varphi_\bullet$ on $\overline{\Omega}$ for every $t \in \overline{J}$. With Remark 1.5.4 in mind, we estimate

$$\sup_{t \in \overline{J}} \big\| -\nabla \cdot \varphi_k(t)\rho\nabla + \nabla \cdot \varphi(t)\rho\nabla \big\|_{\mathscr{L}(\mathrm{W}^{1,q}_D(\Omega); \mathrm{W}^{-1,q}_D(\Omega))} \le \|\rho\|_{\mathrm{L}^\infty} \|\varphi_k - \varphi\|_{\mathrm{C}(\overline{Q})},$$

hence $\mathcal{A}_\rho(\varphi_k) \to \mathcal{A}_\rho(\varphi)$ in $\mathrm{L}^\infty(J; \mathscr{L}(\mathrm{W}^{-1,q}_D(\Omega); \mathrm{W}^{1,q}_D(\Omega)))$. But this shows that

$$\mathcal{A}_\rho(\varphi_k) \to \mathcal{A}_\rho(\varphi) \quad \text{in} \quad \mathscr{L}(\mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega)); \mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega)))$$

and of course also $(\partial + \mathcal{A}_\rho(\varphi_k) + \mathcal{B}_\gamma, \delta_{T_0}) \to (\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma, \delta_{T_0})$ as operators as in (2.42). Since the operators $(\partial + \mathcal{A}_\rho(\varphi_k) + \mathcal{B}_\gamma, \delta_{T_0})$ themselves and the limit operator $(\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma, \delta_{T_0})$ are continuously invertible, the claim follows from continuity of the inversion mapping $A \mapsto A^{-1}$. $\square$

It is essential to observe that the assertion in the foregoing lemma relies critically on the uniformity of spatial domains of $\mathcal{A}_\rho(\varphi)$ and the continuous dependence on φ from the continuous functions.

Remark 2.2.14. The assertions of the previous Lemma 2.2.13 still hold for the operator $\partial + \mathcal{A}_\rho(\varphi) + \mathcal{B}_\gamma$ alone if considered on $\mathbb{W}^{1,r}_0(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ with image space $\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))$. This follows from observing that δ_{T_0} maps that maximal regularity space to $\{0\}$ in the interpolation space.

It follows the proof of Theorem 2.2.12. The strategy is as follows: First, we modify the equation to get rid of the initial value u_0. Then we show that the remaining equation with the nonlinear functions "frozen" admits solutions which are uniformly bounded in a Hölder space, i.e., lie in some ball there and thus in a compact convex set in the space of continuous functions. Re-inserting these solutions into the nonlinear functions shows that this compact convex set is mapped into itself by the continuous "solution mapping", which allows to use Schauder's fixed point theorem.

Proof of Theorem 2.2.12. We begin by splitting off the initial value: Choose an arbitrary function $w \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_D(\Omega), \mathrm{W}^{1,q}_D(\Omega))$ with the initial value $w(T_0) = u_0$. This is possible due to the very definition of the interpolation space, see Lemma 1.4.12. Note that, due to Lemma 1.5.25, cf. also (2.40), w is a continuous function on $\overline{Q}$. We plan to establish the searched-for function u in the form $u = w + v$ for some

$v \in \mathbb{W}_0^{1,r}(J; \mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega))$ satisfying

$$\begin{aligned} v'(t) &+ \big(\mathcal{A}_\rho(\sigma(w+v))(t) + \mathcal{B}_\gamma(t)\big)v(t) \\ &= F(w+v)(t) - w'(t) - \big(\mathcal{A}_\rho(\sigma(w+v))(t) + \mathcal{B}_\gamma(t)\big)w(t) \end{aligned} \tag{2.43}$$

in $\mathrm{W}_D^{-1,q}(\Omega)$ for almost all $t \in J$.

To this end, we consider for $\psi \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ satisfying $\psi(T_0) = 0$ the following equation with “frozen” nonlinearities:

$$\begin{aligned} v'(t) + \big(\mathcal{A}_\rho(\sigma(w+\psi))(t) + \mathcal{B}_\gamma(t)\big)v(t) &= F(w+\psi)(t) - w'(t) \\ &\quad - \big(\mathcal{A}_\rho(\sigma(w+\psi))(t) + \mathcal{B}_\gamma(t)\big)w(t) \end{aligned} \tag{2.44}$$

in $\mathrm{W}_D^{-1,q}(\Omega)$ for almost all $t \in J$. Let us denote the right-hand side in (2.44) by $\mathcal{F}(\psi)(t)$. Clearly, $\mathcal{F}(\psi) \in \mathrm{L}^r(J; \mathrm{W}_D^{-1,q}(\Omega))$ due to the regularity of w and the assumptions on F. Since the operators $\mathcal{A}_\rho(\sigma(w+\psi))(t) + \mathcal{B}_\gamma(t)$ admit the same upper– and coercivity bound for all $t \in \overline{J}$, Proposition 2.1.1 thus yields a unique solution $v_\psi \in \mathbb{W}^{1,2}(J; \mathrm{W}_D^{-1,2}(\Omega), \mathrm{W}_D^{1,2}(\Omega))$ for every $\psi \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$, which allows to define a mapping $\mathcal{T}$ which maps ψ to v_ψ. We are interested in the image of $\mathcal{T}$. Therefore, we first note that $\mathcal{F}$ is uniformly bounded in $\mathrm{L}^r(J; \mathrm{W}_D^{-1,q}(\Omega))$: uniform boundedness of F is exactly assumption (2.41), and for the divergence-gradient-term we estimate with Remark 1.5.4 in mind:

$$\|\mathcal{A}_\rho(\sigma(w+\psi))w\|_{\mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))} \le \big(\sigma^\bullet \|\rho\|_{\mathrm{L}^\infty(J\times\Omega;\mathbb{M}_d)} + 1\big)\|w\|_{\mathrm{L}^r(J;\mathrm{W}_D^{1,q}(\Omega))},$$

which is independent of ψ. As w is fixed and $\mathcal{B}_\gamma$ bounded, this means the right-hand sides $\mathcal{F}(\psi)$ in (2.44) are contained in a ball $\mathbb{B}_{s,q}(0)$ around the origin in $\mathrm{L}^r(J; \mathrm{W}_D^{-1,q}(\Omega))$ independent of ψ.

Now set

$$\begin{aligned} \mathcal{M}_{1,2} \coloneqq \Big\{ & v \in \mathbb{W}_0^{1,2}(J; \mathrm{W}_D^{-1,2}(\Omega), \mathrm{W}_D^{1,2}(\Omega)) : \\ & v = \big(\partial + \mathcal{A}_\rho(\zeta) + \mathcal{B}_\gamma\big)^{-1} g \text{ with } g \in \mathbb{B}_{s,q}(0) \end{aligned}$$

$$\text{and } \zeta \in \mathrm{C}\big(\overline{J};\mathrm{C}(\overline{\Omega})\big),\ \sigma_\bullet \le \zeta \le \sigma^\bullet \Big\}.$$

Theorem 2.1.4 shows that $\mathcal{M}_{1,2}$ is in fact contained in a ball $\mathbb{B}_\alpha(0)$ in some Hölder space $\mathrm{C}^\alpha(Q)$, which in turn is *compactly* included in some ball $\mathbb{B}_0(0)$ in the space of continuous functions $\mathrm{C}(\overline{Q}) \doteq \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$. Clearly, $\mathcal{T}$ maps $\mathbb{B}_0(0)$ to $\mathcal{M}_{1,2} \subset \mathbb{B}_\alpha(0)$ and the set $\{\mathcal{T}(\psi) \colon \psi \in \mathbb{B}_0(0)\}$ is compact in $\mathbb{B}_0(0)$. Hence, Schauder's fixed point theorem yields a fixed point $v = \mathcal{T}(v)$ in $\mathbb{B}_0(0)$, provided we are able to show continuity of the mapping $\mathcal{T}$ from $\mathbb{B}_0(0)$ to $\mathbb{B}_0(0)$. This we do as follows:

The mapping $\psi \mapsto \sigma(w+\psi)$ is continuous from $\mathbb{B}_0(0)$ into $\mathrm{C}(\overline{Q})$ by the Lipschitz assumption on σ, such that Lemma 2.2.13 implies that $\psi \mapsto \big(\partial + \mathcal{A}_\rho(\sigma(w+\psi)) + \mathcal{B}_\gamma\big)^{-1}$ is continuous from $\mathbb{B}_0(0)$ to the linear bounded operators from $\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$ to $\mathbb{W}^{1,r}_0(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$, cf. Remark 2.2.14. Using again Lipschitz-continuity of σ and additionally that of F shows that $\psi \mapsto \mathcal{F}(\psi)$ is also a continuous map from $\mathbb{B}_0(0)$ to $\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$. Now let $(\psi_k) \subset \mathbb{B}_0(0)$ such that $\psi_k \to \psi$ in $\mathbb{B}_0(0)$. Then we find, using the maximal regularity embeddings from Proposition 1.4.3,

$$\begin{aligned}\big\|\mathcal{T}(\psi) - \mathcal{T}(\psi_k)\big\|_{\mathrm{C}(\overline{Q})} \le C\Big\|&\big(\partial + \mathcal{A}_\rho(\sigma(w+\psi) + \mathcal{B}_\gamma\big)^{-1}\mathcal{F}(\psi) \\ &- \big(\partial + \mathcal{A}_\rho(\sigma(w+\psi_k)) + \mathcal{B}_\gamma\big)^{-1}\mathcal{F}(\psi_k)\Big\|_{\mathbb{W}^{1,r}_0(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))}\end{aligned}$$

and a simple triangle argument shows that this goes to 0 as k goes to infinity thanks to the continuity properties established before. This is exactly the searched-for continuity of $\mathcal{T}$.

Finally, a so-obtained fixed point v of $\mathcal{T}$ is obviously a solution (2.43) and, thanks to Lemma 2.2.13, in fact from $\mathbb{W}^{1,r}_0(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$, making $u := w + v$ a solution of (2.38) in the optimal space $\mathbb{W}^{1,r}(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$ with the correct initial value. Concerning uniqueness, we have seen from the embedding (2.40) that the assumptions of Theorem 2.2.10 are satisfied, which shows uniqueness of the solution. □

Corollary 2.2.15. *For fixed u_0, consider the set of admissible data $\{\rho, \sigma, \gamma, F\}$ for the problem (2.38) as in the assumptions of Theorem 2.2.12, where $\rho_\bullet, \rho^\bullet, \sigma_\bullet, \sigma^\bullet, C_F$ and the L^∞-norm of γ are fixed. Then there is $\alpha > 0$ and a ball $\mathbb{B}_\alpha(0)$ in $\mathrm{C}^\alpha(Q)$ such that the set of associated solutions $u_{(\rho,\sigma,\gamma,F)}$ is contained in $\mathbb{B}_\alpha(0)$.*

Proof. Inspecting the proof of Theorem 2.2.12, one observes that the set $\mathcal{M}_{1,2}$ is always the same for all data $\{\sigma, \gamma, F\}$ when $\sigma_\bullet, \sigma^\bullet$, C_F, and the L^∞-norm of γ are fixed, and that the size of $\mathbb{B}_\alpha(0)$ in the Hölder space is also uniform in $\rho_\bullet, \rho^\bullet$ by Theorem 2.1.4. Hence, the size of $\mathbb{B}_\alpha(0)$ is also uniform in $\rho_\bullet, \rho^\bullet, \sigma_\bullet, \sigma^\bullet, C_F$ and the L^∞-norm of γ. □

In the "pure quasilinear" case where F does not depend on u, we still obtain the following useful result from Theorem 2.2.12 and Corollary 2.2.15. Let us again stress that we are *not* able to obtain this result from a simple modification of the local existence– and uniqueness result in Theorem 2.2.10 even if we suppose σ to be given as in Theorem 2.2.12, i.e., well-defined and Lipschitz-continuous on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$, there.

Corollary 2.2.16. *Let the assumptions of Theorem 2.2.12 be satisfied, with the special case that $F(\cdot) \equiv f$ for some $f \in \mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$. Then, for every such f, there exists a* unique global *solution $u = u_f \in \mathbb{W}^{1,r}(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$ of*

$$u'(t) + (\mathcal{A}_\rho(\sigma(u)) + \mathcal{B}_\gamma)u(t) = f(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0.$$

In particular, there is $\alpha > 0$ such that the (nonlinear) solution operator mapping f to u_f transports bounded sets in $\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$ into bounded sets in $\mathrm{C}^\alpha(Q)$ for fixed u_0.

Remark 2.2.17.

(i) Inhomogeneous Dirichlet data may be incorporated into the results above in the same fashion as in Chapter 2.1.5.

(ii) With Theorem 2.1.4 and essentially analogous techniques as displayed in this chapter, one might show existence of global Hölder-continuous solutions to semilinear equations subject to mixed boundary conditions with uniformly bounded and “locally” Lipschitz-continuous F in the form as in Theorem 2.2.12, where the coefficient functions in the divergence-gradient operator are only *measurable* in time. We omit the details.

While we have assumed the nonlinear functions F and σ in Theorem 2.2.12 to be abstract, potentially nonlocal, mappings, the classic situation is that these mappings are realized as Nemytskii operators. We briefly state that such operators satisfy the assumptions of Theorem 2.2.4; in fact, even the assumptions of Theorem 2.2.7 are satisfied due to the interpolation embedding (2.39). In this sense, the theorem of Prüss, Theorem 2.2.7, is already appropriate for this setting.

Lemma 2.2.18. *The following conditions are sufficient for the respective functions in Theorem 2.2.10 (in case $\mathcal{X} = (\mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega))_{1/r',r}$) and Theorem 2.2.12 (in case $\mathcal{X} = \mathrm{C}(\overline{\Omega})$), respectively, to satisfy the assumptions posed there:*

(i) $\sigma \in \mathrm{C}^{1-}_{\mathrm{loc}}(\mathbb{R})$ *such that* $0 < \sigma_\bullet \leq \sigma(\mathrm{x}) \leq \sigma^\bullet$ *for all* $\mathrm{x} \in \mathbb{R}$*, and*

(ii) $F\colon J \times \mathcal{X} \to \mathrm{W}_D^{-1,q}(\Omega)$ *is a Carathéodory function. Moreover, we have* $F(\cdot, 0) \in \mathrm{L}^r(J; \mathrm{W}_D^{-1,q}(\Omega))$ *and for every* $R > 0$ *there exists a function* $\phi_R \in \mathrm{L}^s(J)$ *such that*

$$\|F(t,u) - F(t,v)\|_{\mathrm{W}_D^{-1,q}(\Omega)} \leq \phi_R(t)\|u - v\|_{\mathcal{X}} \quad \textit{for all } u, v \in \overline{\mathbb{B}(0,R)}$$

is true for almost all $t \in J$*, where* $\mathbb{B}(0,R)$ *denotes the ball with radius* R *in* $\mathcal{X}$ *and* $r < s \leq \infty$*. In case* $\mathcal{X} = \mathrm{C}(\overline{\Omega})$*, also suppose that*

$$\sup_{u \in \mathrm{C}(\overline{\Omega})} \|F(\cdot, u)\|_{\mathrm{L}^r(J; \mathrm{W}_D^{-1,q}(\Omega))} =: C_F < \infty. \tag{2.45}$$

Proof. The values $\sigma(u)(t) \in \mathrm{C}(\overline{\Omega})$ and $F(u)(t) \in \mathrm{W}_D^{-1,q}(\Omega)$ depend only on u at the time point $t \in J$, which implies the Volterra property for the induced Nemytskii operators. It is well-known that (locally) Lipschitz-continuous functions give rise to (locally) Lipschitz-continuous Nemytskii operators in the spaces of continuous functions, cf. [148, Lem. 4.11], which implies the assumptions on σ. The desired properties of F in Theorem 2.2.12 are immediate from the assumptions posed here. □

We lastly establish continuity of solutions in the sense of Theorem 2.2.12 of the quasilinear model equation (2.38) with respect to the data in the problem. Thereby, allow the coefficient function σ and the right-hand sides F to vary and show that the solutions of the corresponding equations depend sequentially continuously on this data. While interesting in itself, such results may also become very useful in optimal control theory. We will see that we do not have to pose any additional regularity assumptions on the general data in the problems to obtain the assertions because global existence is already built-in. Let us note that there are also quite sophisticated continuity results for the setting with local-in-time solutions in [10, Sect. 3].

Albeit the equation (2.38) is nonlinear in its structure, we are able to obtain a weak-strong continuity result for the dependence on the right-hand sides by using Theorem 2.1.4 twice, once in the form of Corollary 2.2.15.

Proposition 2.2.19. *Adopt the assumptions of Theorem 2.2.12. Let additionally, for $k \in \mathbb{N}$, the families of functions σ_k and F_k with the following properties be given:*

- *The functions σ_k satisfy the assumptions on σ in Theorem 2.2.12, they admit common Lipschitz-constants on every bounded set in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$, and $\sigma_k(w) \to \sigma(w)$ in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ for every $w \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})$.*
- *The functions F_k satisfy the assumptions on F in Theorem 2.2.12 with a common bound C_F, they admit common Lipschitz-constants on every bounded set in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$, and $F_k(w) \rightharpoonup F(w)$ in $\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$ for every $w \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})$.*

Then the solutions u_k of the equations

$$u'(t) + (\mathcal{A}_\rho(\sigma_k(u)) + \mathcal{B}_\gamma)u(t) = F_k(u)(t) \qquad \text{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \text{for a.a. } t \in J, \quad u(T_0) = u_0,$$

converge strongly *in* $\mathrm{C}(\overline{Q})$ *to the solution* $\bar{u}$ *of* (2.38).

Proof. Without loss of generality, we assume $u_0 = 0$ in the proof. One arrives at this situation by repeating the "split-off"-procedure done at the beginning of the proof of Theorem 2.2.12 and the obvious modifications from thereon without changing the fundamental properties of the problem, as seen there.

It is clear from Theorem 2.2.12 that for every $k \in \mathbb{N}$ there exists a unique solution $u_k \in \mathbb{W}^{1,r}(J; \mathrm{W}^{1,q}_D(\Omega), \mathrm{W}^{-1,q}_D(\Omega))$ to the equations under consideration. Due to the assumption concerning the upper bound C_F, the sequence $F_k(u_k)$ is bounded in $\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))$. Corollary 2.2.15 shows that (u_k) is thus contained in a ball in some Hölder space $\mathrm{C}^\alpha(Q)$ and admits a convergent subsequence $(u_\ell) \to \tilde{u}$ in $\mathrm{C}(\overline{Q})$, for some $\tilde{u} \in \mathrm{C}(\overline{Q})$. We want to show that $\tilde{u} = \bar{u}$. Re-inserting the newly found convergence of u_ℓ into the nonlinear functions gives $F_\ell(u_\ell) \rightharpoonup F(\tilde{u})$ in $\mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega))$ by the assumptions on F_ℓ (here we use the common Lipschitz-constants). Moreover, the analogous Lipschitz properties of σ_ℓ imply that also $\sigma_\ell(u_\ell) \to \sigma(\tilde{u})$ in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and thus

$$\left(\partial + \mathcal{A}_\rho(\sigma_\ell(u_\ell)) + \mathcal{B}_\gamma\right)^{-1} \longrightarrow \left(\partial + \mathcal{A}_\rho(\sigma(\tilde{u})) + \mathcal{B}_\gamma\right)^{-1},$$

thanks to Lemma 2.2.13. Now we have to combine weak convergence of the right-hand sides and strong convergence of the differential operators. Luckily, we still have the *compactness* of the mapping $\left(\partial + \mathcal{A}_\rho(\sigma(\tilde{u}))\right)^{-1} \colon \mathrm{L}^r(J; \mathrm{W}^{-1,q}_D(\Omega)) \to \mathrm{C}(\overline{Q})$ at hand, due to Theorem 2.1.4. This shows that indeed

$$u_\ell = \left(\partial + \mathcal{A}_\rho(\sigma_\ell(u_\ell)) + \mathcal{B}_\gamma\right)^{-1} F_\ell(u_\ell) \longrightarrow \left(\partial + \mathcal{A}_\rho(\sigma(\tilde{u})) + \mathcal{B}_\gamma\right)^{-1} F(\tilde{u}) = \tilde{u}.$$

Rewriting the limit, we find that

$$\tilde{u}'(t) + \big(\mathcal{A}_\rho(\sigma(\tilde{u}))(t) + \mathcal{B}_\gamma(t)\big)\tilde{u}(t) = F(\tilde{u})(t) \quad \text{in } \mathrm{W}^{-1,q}_D(\Omega), \quad \tilde{u}(T_0) = 0,$$

for almost all $t \in J$. But this means exactly that $\tilde{u} = \bar{u}$, the solution of (2.38). Since we can repeat this proof for every subsequence of the original (u_k) with the same limit, the whole sequence (u_k) must converge. □

Lastly, we give a result for *strong* convergence in the data, which yields even convergence of the solutions in the maximal regularity space.

Lemma 2.2.20. *Adopt the assumptions of Proposition 2.2.19, but assume that $F_k(w) \to F(w)$ in $\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))$ for every $w \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})$. Then the solutions u_k of the equations*

$$u'(t) + \big(\mathcal{A}_\rho(\sigma(u))(t) + \mathcal{B}_\gamma(t)\big)u(t) = F_k(u)(t)$$
$$\textit{in } \mathrm{W}^{-1,q}_D(\Omega) \quad \textit{for a.a. } t \in J, \quad u(T_0) = u_0,$$

converge to the solution $\bar{u}$ of (2.38) *in the maximal regularity space $\mathbb{W}^{1,r}(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$.*

Proof. The proof is the same as in Proposition 2.2.19, the difference being that we do not even need to rely on Theorem 2.1.4 a second time in this case. □

CHAPTER 3

Optimal control of quasilinear parabolic equations in divergence form

Finally, we turn to the "ultimate" goal: Optimal control of the abstract quasilinear equations considered in Chapter 2.2. As explained in the introduction of this thesis, we now assume that we are able to manipulate a control function u inside the quasilinear parabolic evolution equation. For simplicity, we consider being able to do so only via the inhomogeneity F. Let us note again that the very abstract $\mathrm{W}_D^{-1,q}(\Omega)$-setting includes most interesting control cases, in particular boundary control. The determining equation is now given by

$$y'(t) + (\mathcal{A}_\rho(\sigma(y)) + \mathcal{B}_\gamma)y(t) = F(y,u)(t) \quad \text{in } \mathrm{W}_D^{-1,q}(\Omega)$$
$$\text{for a.a. } t \in J, \quad y(T_0) = y_0. \quad (3.1)$$

The reader is advised to note the switch from u to y for the searched-for function or *state*, whereas u is now the *control*. Since we build upon the results of Chapter 2.2, we presume the reader to be familiar with its content.

Let us fix the setting in which we will work. The underlying domain Ω is, together with the designated Dirichlet-part of the boundary, supposed to be Lipschitz around $\partial\Omega \setminus D$.

The following assumptions hold true for the rest of this chapter:

(i) The set $\Omega \subset \mathbb{R}^d$ is a bounded domain for $d \in \{2,3\}$ and D (like *Dirichlet*) is a closed subset of $\partial\Omega$. The cases $D = \emptyset$ and $D = \partial\Omega$ are not excluded. We suppose that $\Omega \cup D$ is Lipschitz around $\partial\Omega \setminus D$. In all what follows, $\partial\Omega \setminus D$ will be denoted by N (like *Neumann*).

(ii) We consider a finite interval $J = (T_0, T_1) \subset \mathbb{R}_0^+$.

(iii) All Banach spaces and all occurring functions are supposed to be real ones, i.e., we are working in a *real* setting in the sense of Chapter 1.6.

For the data in the abstract partial differential equation (3.1) as above we take the assumptions of Theorem 2.2.10 to be satisfied for every possible control function u from some abstract vector-valued function space $\mathcal{X}(J;U) \subseteq \mathrm{L}^1(J;U)$. Note that we do neither assume the nonlinear right-hand side to be Lipschitz-continuous on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ nor do we require the uniform boundedness assumption (2.41) to hold, as in Theorem 2.2.12. We will come back to this choice below.

The following assumption holds true for the rest of this chapter. There exist $q > d$ and $r > 2(1-\frac{d}{q})^{-1}$ such that the following properties are true:

(i) The coefficient functions satisfy the following:

a) σ maps $\mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega))$ into $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$, satisfies $0 < \sigma_\bullet \leq \sigma(u) \leq \sigma^\bullet$ on $\overline{Q}$ for all $u \in \mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega))$, is uniformly Lipschitz-continuous on bounded sets, and has the Volterra property.

b) $\rho \in \mathrm{L}^\infty(\Omega;\mathbb{M}_d(\rho_\bullet,\rho^\bullet))$ and

$$-\nabla\cdot\rho\nabla + 1 \in \mathscr{L}_{\mathrm{iso}}(\mathrm{W}_D^{1,q}(\Omega);\mathrm{W}_D^{-1,q}(\Omega)),$$

i.e., $\mathcal{D}_q(\rho) \doteq \mathrm{W}_D^{1,q}(\Omega)$. Moreover, we assume the assumptions of Proposition 1.5.5 to be satisfied for $-\nabla\cdot\rho\nabla$, and $\gamma \in \mathrm{C}(\overline{J};\mathrm{L}^\infty(\overline{N};\omega)$.

(ii) The mapping F maps $\mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega)) \times \mathcal{X}(J;U)$ into $\mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))$. For every $u \in \mathcal{X}(J;U)$ the mapping

$$F_u \coloneqq F(\cdot,u)\colon \mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega)) \to \mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))$$

has the Volterra property, and there exists a number $r < s \leq \infty$ such that $F_u - F_u(0)$ is uniformly Lipschitz-continuous on bounded sets in $\mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega))$ with values in $\mathrm{L}^s(J;\mathrm{W}_D^{-1,q}(\Omega))$.

(iii) The initial value satisfies $y_0 \in (\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega))_{1/r',r}$.

Let us recall that from Lemma 1.5.25 and Proposition 1.5.5 we have the embeddings

$$\big(\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega)\big)_{1/r',r} \hookrightarrow \mathrm{C}^\beta(\Omega) \hookrightarrow\hookrightarrow \mathrm{C}(\overline{\Omega}) \tag{2.39}$$

for some $\beta > 0$ and

$$\mathbb{W}^{1,r}(J;\mathrm{W}_D^{-1,q}(\Omega),\mathrm{W}_D^{1,q}(\Omega)) \hookrightarrow \mathrm{C}^\alpha(Q) \hookrightarrow\hookrightarrow \mathrm{C}(\overline{Q}) \tag{2.40}$$

for some $\alpha > 0$ at our disposal under the assumptions above. We will make free use of them throughout this chapter.

Concerning the question of why we have not chosen the setting for *global* solutions from Theorem 2.2.12: We have already noted in Chapter 2.2.1 that the assumption for F to be Lipschitz-continuous on the space of uniformly continuous functions is *quite* strong. Moreover, it turns out that we are able to still do optimal control theory under reasonable assumptions even if we cannot guarantee that every control u gives rise to a global solution y in the sense of Definition 2.2.1 by restricting the optimal control problem to the set of controls whose associated solutions are indeed global ones. But this means we would not pose the uniform boundedness assumption (2.41) anyway, and then there is very little reason not to work with the much weaker assumptions on Lipschitz-continuity of F.

In the context of optimal control, global solutions for a given time interval J are of particular interest when dealing with so-called end-time tracking, i.e., the aim of the optimal control procedure is to drive the solution $y(T_1)$ at the end-time T_1 to a designated state. This of course necessitates to be able to guarantee that the solution y exists in a suitable sense up to and including T_1 in the first place, cf. also [12], where this situation is discussed. We do this by restricting the set of controls to those which admit such global solutions. If one wants to *definitely* guarantee global solutions for a given, or every, control, then the assumptions in Theorem 2.2.12 are appropriate, including having the nonlinear functions live and be locally Lipschitz-continuous on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$. We briefly summarize the intermediate results in a lemma. Since the initial value y_0 is fixed, the maximal interval of existence now depends on u alone.

Lemma 3.0.1. *For every control* $u \in \mathcal{X}(J;U)$*, there exists a unique maximal solution* $y = y_u \in \mathbb{W}^{1,r}(J^\bullet(u);\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega))$ *of* (3.1) *in the sense of Definition 2.2.1. If* $\Omega\cup D$ *is in fact volume-preserving generalized regular in the sense of Gröger,* σ *and* F_u *live and satisfy their assumptions even on* $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ *and*

$$\sup_{v\in\mathrm{C}(\overline{J};\mathrm{C}\overline{\Omega}))} \|F_u(v)\|_{\mathrm{L}^r(J;\mathrm{W}^{-1,q}_D(\Omega))} =: C_{F,u} < \infty$$

holds true for some $u \in \mathcal{X}(J;U)$*, then* y_u *is a* global *solution.*

Proof. This is just a rephrasing of the observation that the general assumptions in this chapter imply those of Theorem 2.2.10 and, under the additional assumptions as posed in this lemma, Theorem 2.2.12. □

There are (probably many) real-life examples for which the additional assumptions for global solutions in Lemma 3.0.1 are not satisfied; the thermistor problem in three spatial dimensions being such a one, see Chapter 4.

Before we define the optimal control problem, we introduce yet another control space, $\mathcal{U}$. We will use this space as the actual control space for the optimal control setup. This is to accommodate for the very common situation where one is able to achieve satisfying results for the underlying partial differential equation for u coming from a weaker space than the space in which one wants or needs to pose the optimal control problem. In our case, the weaker space is $\mathcal{X}(J;U)$, whereas we formulate the optimal control problem for the (possibly) stronger space $\mathcal{U}$. The need to use a stronger space for the optimal control problem may have many different reasons: It may be impossible to formulate certain control constraints for $u \in \mathcal{X}(J;U)$ only, or $\mathcal{U}$ may be more suitable for numerical analysis or computations, or we may need the stronger properties of $\mathcal{U}$ for further analytic purposes, just to name a few. A particular space U for which these difficulties arise would be $\mathrm{W}^{-1,q}_D(\Omega)$ or generally every abstract dual space. It is, however, perfectly valid to choose the same spaces for analysis and optimal control if there are no pressuring reasons to do otherwise.

We introduce the following abbreviations to be used for the rest of this chapter:

$$\begin{aligned}
\mathcal{Y}_{r,q} &:= \mathbb{W}^{1,r}\bigl(J;\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega)\bigr),\\
\mathcal{Z}_{r,q} &:= \mathrm{L}^r\bigl(J;\mathrm{W}^{-1,q}_D(\Omega)\bigr),\\
Y_{r,q} &:= \bigl(\mathrm{W}^{-1,q}_D(\Omega),\mathrm{W}^{1,q}_D(\Omega)\bigr)_{1/r',r}.
\end{aligned}$$

Moreover, we encode the quasilinear equation (3.1) into

$$e(y,u) := \bigl((\partial + \mathcal{A}_\rho(\sigma(y)) + \mathcal{B}_\gamma)y - F(y,u), \delta_{T_0}y - y_0\bigr) = 0$$

as a function
$$e\colon \mathcal{Y}_{r,q} \times \mathcal{X}(J;U) \to \mathcal{Z}_{r,q} \times Y_{r,q}.$$
Note that $e(y,u) = 0$ in $\mathcal{Z}_{r,q} \times Y_{r,q}$ already implies that y is a global solution to (3.1). To account for the control space $\mathcal{U}$, we moreover set

$$\mathsf{e}\colon \mathcal{Y}_{r,q} \times \mathcal{U} \to \mathcal{Z}_{r,q} \times Y_{r,q}, \quad \mathsf{e}(y,u) := e(y,\mathsf{E}u),$$

where $\mathsf{E} \in \mathscr{L}(\mathcal{U};\mathcal{X}(J;U))$ gives the connection from $\mathcal{U}$ to $\mathcal{X}(J;U)$; the generic case is that E is an embedding, but also the trace operator tr or tr^* occurs often. The general optimal control problem is then given as follows.

Definition 3.0.2 (General optimal control problem)**.** The *(general) optimal control problem* is given by

$$\min_{(y,u)\in\mathcal{Y}_{r,q}\times\mathcal{U}} \quad \mathsf{J}(y,u) \qquad \text{s.t.} \quad \begin{cases} \mathsf{e}(y,u) = 0 \\ u \in \mathcal{U}^{\mathrm{ad}} \\ \mathcal{G}(y) \in \mathcal{K} \end{cases} \tag{OC}$$

with the feasible set

$$\mathcal{M}^{\mathrm{ad}} := \Bigl\{(y,u) \in \mathcal{Y}_{r,q} \times \mathcal{U}\colon e(y,u) = 0 \text{ in } \mathcal{Z}_{r,q},\, u \in \mathcal{U}^{\mathrm{ad}},\, \mathcal{G}(y) \in \mathcal{K}\Bigr\}$$

and the coordinate-wise projections

$$\mathcal{M}^{\mathrm{ad}}_s := \Bigl\{y \in \mathcal{Y}_{r,q}\colon \exists u \in \mathcal{U}\colon (y,u) \in \mathcal{M}^{\mathrm{ad}}\Bigr\}$$

and

$$\mathcal{M}^{\mathrm{ad}}_c := \Bigl\{u \in \mathcal{U}\colon \exists y \in \mathcal{Y}_{r,q}\colon (y,u) \in \mathcal{M}^{\mathrm{ad}}\Bigr\}.$$

We say that a state $y \in \mathcal{Y}_{r,q}$ or a control $u \in \mathcal{U}$ are *feasible* if $y \in \mathcal{M}^{\mathrm{ad}}_s$

and $u \in \mathcal{M}_c^{\mathrm{ad}}$, respectively.

We fix the following assumptions on the data in (OC).

The following assumptions hold true for the rest of this chapter.

(i) $U, \mathcal{X}(J;U)$ and $\mathcal{U}$ are Banach spaces with $\mathcal{X}(J;U) \hookrightarrow \mathrm{L}^1(J;U)$. Moreover, $\mathcal{U}$ is reflexive.

(ii) J maps $\mathcal{Y}_{r,q} \times \mathcal{U}$ to $\mathbb{R}_0^+$ and is of separated form $\mathsf{J}(y,u) = \mathsf{J}_s(y) + \mathsf{J}_c(u)$.

  a) J_s is nonnegative and lower semicontinuous on $\mathcal{Y}_{r,q}$,

  b) J_c is nonnegative and weakly lower semicontinuous on $\mathcal{U}$. If $\mathcal{M}_c^{\mathrm{ad}}$ is unbounded in $\mathcal{U}$, then J_c is also coercive on $\mathcal{U}$.

(iii) $\mathcal{U}^{\mathrm{ad}} \subseteq \mathcal{U}$ is closed and convex in $\mathcal{U}$.

(iv) $\mathcal{C}$ is a Banach space, $\mathcal{G} \colon \mathrm{C}(\overline{Q}) \to \mathcal{C}$ is continuous, and $\mathcal{K} \subseteq \mathcal{C}$ is a closed convex cone in $\mathcal{C}$ with nonempty interior. Moreover, $\mathcal{C} \hookrightarrow \left[\mathrm{L}^\infty(Q)\right]^k$ for some number $k \in \mathbb{N}$.

(v) $\mathcal{M}^{\mathrm{ad}} \neq \emptyset$.

Note that $\mathcal{G}(y)$ is well-defined for $y \in \mathcal{Y}_{r,q}$ due to $\mathcal{Y}_{r,q} \hookrightarrow \mathrm{C}^\alpha(Q)$ for some $0 < \alpha < 1$, cf. (2.40). Typical examples for $\mathcal{G}, \mathcal{C}$ and $\mathcal{K}$ are

$$\mathcal{G}(y) := \begin{pmatrix} y - y^\bullet \\ y_\bullet - y \end{pmatrix}, \quad \mathcal{C} := \left[\mathrm{C}(\overline{Q})\right]^2 \quad \text{and} \quad \mathcal{K} := \left[K_-\right]^2$$

for continuous functions $y_\bullet, y^\bullet \in \mathrm{C}(\overline{Q})$ and

$$K_- := \left\{ f \in \mathrm{C}(\overline{Q}) \colon f \leq 0 \text{ on } \overline{Q} \right\}. \tag{3.2}$$

This particular choice is the abstract representation of classical *state constraints* $y_\bullet(t,\mathrm{x}) \leq y(t,\mathrm{x}) \leq y^\bullet(t,\mathrm{x})$ for all $(t,\mathrm{x}) \in \overline{Q}$.

Remark 3.0.3. Since we have assumed $\mathcal{K}$ to have nonempty interior in $\mathcal{C}$, it is not really a loss of generality to assume $\mathcal{C} \hookrightarrow \left[\mathrm{L}^\infty(Q)\right]^k$ for some number $k \in \mathbb{N}$, since the most relevant cone, the one consisting of non-negative (or nonpositive) functions in $\mathcal{C}$, has nonempty interior exactly in spaces equipped with supremum-type norms.

The go-to standard example for the control constraints set is quite similar, for instance $U = \mathrm{L}^p(\Omega)$ for some $1 \leq p \leq \infty$ and

$$\mathcal{U}^{\mathrm{ad}} := \Big\{ u \in \mathcal{U} \colon u_\bullet \leq u \leq u^\bullet \text{ a.e. in } Q \Big\}$$

for $u_\bullet, u^\bullet \in \mathcal{U}$. The reader may imagine having such examples for the constraints in (OC).

For the following considerations, we will work with (OC) in its reduced form. This is mostly for convenience and without loss of generality. As seen in Lemma 3.0.1, there is a unique maximal solution $y = y_u$ for every control $u \in \mathcal{X}(J;U)$, and we have already noted that $e(y_u, u) = 0$ already encodes that y_u is a *global* solution. To make up for this, we also introduce the set of “global controls”.

Definition 3.0.4 (Reduced optimal control problem)**.** Let $\mathcal{U}_g$ denote the set of *global controls*, that is, controls u for which the associated unique state $y_u \in \mathcal{Y}_{r,q}$ is a global solution to (3.1), or equivalently,

$$U_g := \Big\{ u \in \mathcal{X}(J;U) \colon e(y_u, u) = 0 \in \mathcal{Z}_{r,q} \Big\}, \quad \mathcal{U}_g := \mathcal{U} \cap U_g = \mathsf{E}^{-1}\big[U_g\big],$$

where E was a continuous linear mapping from $\mathcal{U}$ to $\mathcal{X}(J;U)$ and $\mathsf{E}^{-1}[\cdot]$ denotes the preimage. We moreover set the *control-to-state operators* $\mathcal{S}, \mathsf{S}$ to be

$$\mathcal{S} \colon U_g \to \mathcal{Y}_{r,q}, \quad u \mapsto y_u =: \mathcal{S}(u), \quad \text{and} \quad \mathsf{S} \colon \mathcal{U}_g \to \mathcal{Y}_{r,q}, \quad \mathsf{S} := \mathcal{S} \circ \mathsf{E},$$

and use this to define the *reduced optimal control problem* by

$$
\begin{aligned}
\min_{u \in \mathcal{U}_g} \quad & \mathsf{j}(u) \coloneqq \mathsf{J}(y_u, u) = \mathsf{J}_s(y_u) + \mathsf{J}_c(u) \\
\text{s.t.} \quad & \begin{cases} u \in \mathcal{U}^{\mathrm{ad}} \\ \mathcal{G}(y_u) \in \mathcal{K}. \end{cases}
\end{aligned}
\tag{rOC}
$$

Let us also lastly set $\mathcal{U}_g^{\mathrm{ad}} \coloneqq \mathcal{U}_g \cap \mathcal{U}^{\mathrm{ad}}$ and obtain

$$
\mathcal{M}_c^{\mathrm{ad}} = \left\{ u \in \mathcal{U}_g^{\mathrm{ad}} \colon y_u \in \mathcal{G}^{-1}[\mathcal{K}] \right\}
$$

as the feasible set for (rOC).

After this exhaustive amount of nomenclature and definitions, we finish this introduction with some remarks.

Remark 3.0.5.

(i) We will, by slight abuse of notation, keep on writing y_u for $\mathsf{S}(u)$.

(ii) Note that $U_g, \mathcal{U}_g, \mathcal{U}_g^{\mathrm{ad}}$ are all nonempty due to the assumption $\mathcal{M}^{\mathrm{ad}} \neq \emptyset$.

(iii) The reader may wonder why we have kept the different spaces $\mathcal{X}(J;U)$ and $\mathcal{U}$ instead of $\mathcal{U}$ alone. Indeed, this will become only truly relevant in Chapter 3.1.1 in which we give a framework on how to choose a suitable space $\mathcal{U}$ if $\mathcal{X}(J;U)$ alone is not sufficient for existence theory as displayed in Chapter 3.1, which is the generic case if the dependence of F_u on u is nonlinear. In general, we imagine the space $\mathcal{X}(J;U)$ to be as weak as possible for a well-rounded PDE theory.

3.1 Existence of optimal controls

Attempting the usual proof from the calculus of variations to show that minimizers to (rOC) exist (see e.g. [86, Ch. 1.5.2]), one soon encounters the problem that due to the nonlinear nature of the equation and the restriction to global controls, the sequence of solutions corresponding to a minimizing sequence of controls need not be bounded in suitable "strong" spaces to achieve compactness for weak-strong continuity of S on $\mathcal{U}$ without further assumptions. In other words, we cannot assume that S is weakly continuous without loss of generality. This means we need to use the admissible set $\mathcal{M}_c^{\mathrm{ad}}$ and the objective functional at hand, including a suitable choice of the space $\mathcal{U}$. We encode this in the following existence result.

Theorem 3.1.1. *Let $M > 0$ be a number for which $\mathcal{N} := \mathsf{J}^{-1}[[0, M)] \cap \mathcal{M}_c^{\mathrm{ad}} \neq \emptyset$. Suppose that the weak (sequential) closure $\mathcal{N}^w$ of $\mathcal{N}$ in $\mathcal{U}$ is contained in $\mathcal{U}_g$ and that S is weak-strong continuous from $\mathcal{N}^w$ to $\mathcal{Y}_{r,q}$. Then the reduced optimal control problem (rOC) admits a global solution $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$.*

Proof. Due to assumption $\mathcal{M}^{\mathrm{ad}} \neq \emptyset$, we know that also $\mathcal{M}_c^{\mathrm{ad}} \neq \emptyset$ and hence $\inf_{u \in \mathcal{M}_c^{\mathrm{ad}}} \mathsf{j}(u)$ is a finite value (recall that J takes only nonnegative values). Thus there exists a number M such that $\mathcal{N}$ is nonempty and an infimal sequence $(u_k) \subset \mathcal{M}_c^{\mathrm{ad}}$ such that $\lim_{k\to\infty} \mathsf{j}(u_k) \to \inf_{u\in\mathcal{M}_c^{\mathrm{ad}}} \mathsf{j}(u)$, from which we can without loss of generality assume that $(u_k) \subset \mathcal{N}$. From the assumptions on J_s and J_c, we obtain that (u_k) must be bounded in the reflexive space $\mathcal{U}$ and thus admits a weakly convergent subsequence $u_\ell \rightharpoonup \bar{u} \in \mathcal{U}$.

Let us show $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$. We already know that $\bar{u} \in \mathcal{U}_g^{\mathrm{ad}}$: It is in $\mathcal{U}^{\mathrm{ad}}$ since the properties "closed and convex" in $\mathcal{U}$ imply $\mathcal{U}^{\mathrm{ad}}$ being weakly closed in $\mathcal{U}$, and $\bar{u} \in \mathcal{U}_g$ follows from the assumption that $\mathcal{N}^w$ is contained in $\mathcal{U}_g$. Continuity of S on $\mathcal{N}^w$ shows that $y_\ell := y_{u_\ell} \to \bar{y} := y_{\bar{u}}$, which due to continuity of $\mathcal{G}$ then shows that $\mathcal{G}(y_\ell) \to \mathcal{G}(\bar{y})$ in $\mathcal{C}$ and thus $\mathcal{G}(\bar{y}) \in \mathcal{K}$. This altogether gives $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$.

Finally, from the lower semicontinuity properties of J_s and J_c, we obtain

$$\begin{aligned}\inf_{u \in \mathcal{M}_c^{\mathrm{ad}}} \mathsf{j}(u) = \lim_{\ell \to \infty} \mathsf{j}(u_\ell) &\geq \liminf_{\ell \to \infty} \mathsf{J}_s(y_\ell) + \liminf_{\ell \to \infty} \mathsf{J}_c(u_\ell) \\ &\geq \mathsf{J}_s(\bar{y}) + \mathsf{J}_c(\bar{u}) = \mathsf{j}(\bar{u}),\end{aligned}$$

hence indeed $\mathsf{j}(\bar{u}) = \inf_{u \in \mathcal{M}_c^{\mathrm{ad}}} \mathsf{j}(u)$ and $\bar{u}$ is the searched-for minimizer of (rOC). ☐

The assumptions in Theorem 3.1.1 have two rather different backgrounds: The first is a weak closedness assumption on $\mathcal{U}_g$ in the context of the optimal control problem. If $\mathcal{U}_g = \mathcal{U}$ or, more generally, $\mathcal{U}_g^{\mathrm{ad}}$ is weakly closed, it is clearly fulfilled, but the general case may be delicate to verify. The second assumption is a weak continuity requirement for S, also in the context of the optimal control problem. The two assumptions might very well be intertwined, that is, the weak continuity of S might indeed only hold on $\mathcal{N}^w$. Depending on the form of F_u, this may also require to choose $\mathcal{U}$ as a quite strong space to obtain such weak continuity, as the next lemma illustrates.

Lemma 3.1.2. *Let $\mathcal{N}$ be as in Theorem 3.1.1 and assume that the closure $\overline{\mathsf{E}\mathcal{N}}$ in $\mathcal{X}(J;U)$ is contained in U_g. Suppose further that $\mathcal{S}$ is continuous as a mapping from $\overline{\mathsf{E}\mathcal{N}}$ in $\mathcal{X}(J;U)$ to $\mathcal{Y}_{r,q}$. If $\mathsf{E}\,\mathcal{U}^{\mathrm{ad}}$ is a compact subset of $\mathcal{X}(J;U)$ or if E is in fact a compact linear mapping, then the assumptions in Theorem 3.1.1 are satisfied.*

Proof. Let $\bar{u}$ be an element of the weak (sequential) closure $\mathcal{N}^w$ of $\mathcal{N}$ in $\mathcal{U}$. Then there exists a sequence $(u_k) \subset \mathcal{N}$ such that $u_k \rightharpoonup \bar{u}$ in $\mathcal{U}$. From both compactness assumptions it follows that $\mathsf{E}u_k \to \mathsf{E}\bar{u} \in \mathcal{X}(J;U)$ and since $\mathsf{E}\bar{u}$ is an element of $\overline{\mathsf{E}\mathcal{N}}$ in $\mathcal{X}(J;U)$, it must be a global control, $\mathsf{E}\bar{u} \in U_g$. Whence also $\bar{u} \in \mathcal{U}_g$. This shows that $\mathcal{N}^w$ is contained in $\mathcal{U}_g$, which implies $\mathcal{S}(u_k) \to \mathcal{S}(\bar{u})$ in $\mathcal{Y}_{r,q}$. Weak continuity of S on $\mathcal{N}^w$ as in Theorem 3.1.1 follows. ☐

Will encounter a particular example of the previous setting in Chapter 4. A often-considered special case is the one where the right-hand side consists of an an affine-linear function in the control in $\mathcal{X}(J;U)$. If in addition σ satisfies its assumptions already on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ and $\Omega \cup D$ is volume-preserving generalized regular in the sense of Gröger, then we always obtain weak-strong convergence of S. This shows that the linear case is always well-defined regarding existence of globally optimal controls under reasonable assumptions.

Corollary 3.1.3. *Let $C \in \mathscr{L}(U;W_D^{-1,q}(\Omega))$, $f \in \mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))$, and let $F_u(y)$ be given by*

$$F_u(y) \coloneqq G(u) \coloneqq Cu + f \quad \text{for every } y \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})).$$

If $\mathsf{E}\mathcal{U}^{\mathrm{ad}}$ is a compact subset of $\mathcal{X}(J;U)$ or if E is in fact a compact linear mapping, then $u \mapsto G(u)$ is weak-strong continuous as a mapping from $\mathcal{U}^{\mathrm{ad}}$ or $\mathcal{U}$, respectively, to $\mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))$. If in addition σ satisfies its assumptions already on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ and $\Omega \cup D$ is volume-preserving generalized regular in the sense of Gröger, then S is weak-strong continuous on $\mathcal{U}^{\mathrm{ad}}$ or $\mathcal{U}$.

Proof. The first assertion is clear from the linear structure, whereas we can apply Lemma 2.2.20 for the continuity result for S. □

Remark 3.1.4. The above results all require weak-*strong* continuity of S in a suitable sense, which is in general to be expected since we deal with nonlinear problems. However, we have seen in Proposition 2.2.19 that the uniform Hölder estimates from Chapter 2.1 also allow to work with only weak convergence of the data, at the price of the much stronger form of the Lipschitz-conditions on F. Moreover, one obtains only convergence of the solutions in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$. Changing the assumptions on J_s and σ accordingly, one could modify the proof and assertions of Theorem 3.1.1 to this setting, too. In particular, assuming weak continuity of $u \mapsto F_u(\mathsf{S}(u))$, one could dispose of the compactness assumptions in Lemma 3.1.2 and Corollary 3.1.3. Unfortunately, the case of linear dependence on u in $F_u(\cdot)$

is in general the only one in which one can expect true weak continuity: It is known that Nemytskii operators Φ acting on $\mathrm{L}^p(J)$ in the form $\Phi(w) = [t \mapsto \varphi(w(t))]$ are weak-weak continuous if and only if φ is an affine-linear function, see e.g. [29, Exercise 4.20][6], and our assumptions on the right-hand sides include suitable Nemytskii operators. In this sense, we suggest to only fall back to the setting in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ including stronger Lipschitz-conditions on F if F_u is of the form $G(u)$ as in Corollary 3.1.3.

3.1.1 Interlude: A compact embedding for the control space

Lemma 3.1.2 shows that it is of interest to investigate compactness in vector-valued function spaces when aiming to treat optimal control problems whose defining partial differential equations are nonlinear in both control and state. In this subchapter, propose a family of spaces $\mathcal{U}$ which admit a compact embedding into $\mathcal{X}(J;U)$ for a large set of possible incarnations of $\mathcal{X}(J;U)$ while being relatively well-suited to the optimal control setting.

Compactness in vector-valued function spaces is an important topic for theory of nonlinear partial differential equations, popularized by Lions [109], and has attracted a number of researchers in the last decades, see e.g. [4, 19, 38, 137, 138]. Basically, for a set $\mathfrak{C}$ to be compact in $\mathcal{X}(J;U)$, one needs "spatial" compactness of $\mathfrak{C}$ together with certain convergence properties of the translation in time $h \mapsto f(\cdot + h) \in U$ as $h \to 0$ for functions $f \in \mathfrak{C}$, classically obtained by weak differentiability or Hölder continuity properties. For the optimal control setting, this gives us essentially two options:

(i) Use characterizations of compactness to put further constraints $u \in \mathfrak{C}$ on $\mathcal{U}^{\mathrm{ad}}$ which force controls $u \in \mathcal{U}^{\mathrm{ad}}$ to be of a specific structure, for instance uniformly piecewise Hölder-continuous with values in U. An example for this ansatz can be found in [88, Rem. 4.9], basing

[6] Note that Brezis supplies solutions to his exercises.

on [137, Ch. 6, Thm. 3] for the case $\mathcal{X}(J;U) = \mathrm{L}^\infty(J;\mathrm{W}_D^{-1,q}(\Omega))$. Note that for this technique $\mathcal{U}^{\mathrm{ad}}$ must already induce some (weak) kind of spatial compactness in U, which however is rather easily obtained. Any way, the such-obtained set $\mathcal{U}^{\mathrm{ad}}$ will be of difficult structure with respect to treating optimality conditions since one has to deal with the polar cone of $\mathfrak{C}$ there. One could of course use the set $\mathfrak{C}$ only to establish existence of globally optimal controls for the optimal control problem and drop it for optimality conditions, but this would result in "incomplete" theory and is as such slightly unfavorable.

(ii) Make the control space $\mathcal{U}$ strong enough such that $\mathsf{E}\colon \mathcal{U} \to \mathcal{X}(J;U)$ is compact. As already mentioned in the introduction of this chapter, it is indeed often the case that one obtains a analytically suitable dependence on the control u in the partial differential equation for u from weaker spaces than the "original" regularity the control function would have from modeling, which makes this ansatz seem rather natural. A particular case of this setting is boundary control where the boundary term $u \in \mathrm{L}^p(\partial\Omega)$ for suitable $1 \leq p \leq \infty$ in the partial differential equation involving the control is interpreted as $\mathrm{W}_D^{-1,q}(\Omega)$ valued, analogously to how we have done it in Definition 1.5.11. Since J_c in the objective functional corresponds directly to $\mathcal{U}$ because of the coercivity– and lower semicontinuity assumption, we need to be careful with the choice of the function space $\mathcal{U}$, in particular in view of the next step, first order necessary conditions. For these, we need J_c to be continuously differentiable on $\mathcal{U}$ and its derivative directly enters the final optimality conditions, so it is preferable to have a somewhat nice form at hand.

We will show a favorable choice of spaces for which the second technique results in a generalized semilinear differential equation in Banach space in the form of a variational inequality in the final first order necessary optimality conditions. The general requirement is that there exists a compact linear operator $\mathsf{E}\colon \mathcal{U} \to \mathcal{X}(J;U)$, and we need to find a continuously differentiable function J_c on $\mathcal{U}$ which is in addition coercive on $\mathcal{U}$ if $\mathcal{M}_c^{\mathrm{ad}}$

is unbounded there. Since it is rather difficult to find coercive functions on vector-valued function spaces which are not immediately related to the norm function on that space, we will concentrate on (coercive parts of) norms on $\mathcal{U}$ as the choices for J_c. Unfortunately, this rules out already a class of function spaces whose compactness properties we have already used extensively: the Hölder spaces. This is essentially a consequence of the characterization of smooth points of the supremum norm from [106] and reads as follows:

Lemma 3.1.5. *Let X be a Banach space and consider the Hölder space $\mathrm{C}^\alpha(J;X)$ for $\alpha \in [0,1) \cup \{1-\}$. The norm function $\|\cdot\|_{\mathrm{C}^\alpha(J;X)}$ fails to be continuously differentiable in every constant function $f\colon \overline{J} \to X$.*

Proof. Let $f\colon \overline{J} \to X$ be constant. Clearly, $f \in \mathrm{C}^\alpha(J;X)$. Now, if the norm function on $\mathrm{C}^\alpha(J;X)$ was continuously differentiable in f, then the limit of

$$\frac{\|f+th\|_{\mathrm{C}^\alpha(J;X)} - \|f\|_{\mathrm{C}^\alpha(J;X)}}{t} = \frac{\|f+th\|_{\mathrm{C}(\overline{J};X)} - \|f\|_{\mathrm{C}(\overline{J};X)}}{t} + [h]_{\alpha,\overline{J},X}$$

as $t \searrow 0$ must exist for every $h \in \mathrm{C}^\alpha(J;X)$, cf. [50, Ch. I.1], where $[h]_{\alpha,\overline{J},X}$ was the Hölder semi-norm, recall Definition 1.2.3. However, it is shown in [106, Cor. 3.2] that the limit on the right-hand side cannot exist since continuous functions cannot be smooth points of the supremum norm.[7] But then the limit on the left-hand side also cannot exist. □

Remark 3.1.6.

(i) We point out that the result in Lemma 3.1.5 holds true even if X is a Banach space whose norm is continuously Fréchet-differentiable in $X \setminus \{0\}$, in particular if $X = \mathbb{R}$.

(ii) It seems tempting to dismiss the “evil” supremum norm part of the Hölder norm and just use the Hölder semi-norm for J_c. This choice is invalid without even talking about differentiability because the

[7] This seems to be proven already by Banach in 1932 in [21]. Unfortunately, we were unable to locate the result there.

Hölder seminorm is not a coercive function on the Hölder space, as again the constant functions demonstrate.

This means that we need another class of vector-valued function spaces which admit compact embeddings. Since compactness in such spaces is a scarce property, it will still be the Hölder spaces which deliver the crucial compact embedding. The starting point is the embedding (1.31) for two Banach spaces $Y \hookrightarrow_{\mathrm{d}} X$ and $1 < \varrho, p < \infty$, that is,

$$\mathbb{W}^{1,\varrho}_p(J; X, Y) \hookrightarrow \mathrm{C}^\alpha(J; (X, Y)_{\theta,1}) \tag{1.31}$$

for $0 < \theta < 1/\xi' = \frac{1}{\varrho'}(1 + \frac{1}{p} - \frac{1}{\varrho})^{-1}$ and $0 < \alpha < 1/\varrho' - \theta(1 + \frac{1}{p} - \frac{1}{\varrho})$ with $\xi := p\,(1 + \frac{1}{p} - \frac{1}{\varrho})$. From this embedding we infer that, given $1 < \varrho, p < \infty$ and a *compact* linear operator $E\colon (X, Y)_{\theta,1} \to U$ for some $0 < \theta < 1/\xi'$, then the time-extension E of E also maps $\mathbb{W}^{1,\varrho}_p(J; X, Y)$ compactly to $\mathrm{C}^\beta(J; U)$ for $0 < \beta < \alpha$ by the Arzelà-Ascoli Theorem 1.2.5 in the form of Corollary 1.2.6. This means we could use $\mathcal{U} = \mathbb{W}^{1,\varrho}_p(J; X, Y)$ as long as $\mathrm{C}^\beta(J; U) \hookrightarrow \mathcal{X}(J; U)$ which seems like a reasonable requirement. Note that the degree of Hölder continuity and the index θ for the interpolation spaces in the embedding (1.31) are balanced against each other.
In this sense, the term J_c could be chosen suitable for $\mathcal{U} = \mathbb{W}^{1,\varrho}_p(J; X, Y)$ by

$$\mathsf{J}_c\colon \mathbb{W}^{1,\varrho}_p(J; X, Y) \to \mathbb{R}, \quad \mathsf{J}_c(u) := \|u'\|^\varrho_{\mathrm{L}^\varrho(J;X)} + \|u\|^p_{\mathrm{L}^p(J;Y)}, \tag{3.3}$$

which is clearly coercive and lower semicontinuous on $\mathbb{W}^{1,\varrho}_p(J; X, Y)$. It is even continuously differentiable, provided the norm functions n_X and n_Y on X and Y are so. This follows from the following two results:

Lemma 3.1.7 ([105, Thm. 2.5])**.** *Let $(\Upsilon, \mathfrak{A}, \mu)$ be a measure space, $1 < \varsigma < \infty$ and let E be a Banach space. Then the norm function on $\mathrm{L}^\varsigma(\Upsilon; \mu, E)$ is continuously differentiable except in 0 if and only if the norm function on E is so.*

Lemma 3.1.8. *Let E be a Banach space, $\nu > 1$, and let $f\colon E \to \mathbb{R}$ be con-*

tinuously differentiable in $E \setminus \{0\}$ *and locally Lipschitz-continuous around* 0*. Then* $g\colon E \to \mathbb{R}$*,* $g(x) \coloneqq |f(x)|^{\nu-1} f(x)$*, is continuously differentiable on* E *with* $g'(0) \equiv 0$ *in* E'*.*

Proof. Assume, without loss of generality, that $f(0) = 0$. It is clear that g is continuously differentiable in $E \setminus \{0\}$ with the derivative $g'(x) = \nu |f(x)|^{\nu-1} f'(x)$. In 0, we have

$$\big|g(h) - g(0) - 0 \cdot h\big| = |g(h)| \leq L_R^\nu \|h\|_E^\nu = o(\|h\|_E) \quad \text{for } h \in \mathbb{B}(0,R)$$

with R so small that f is Lipschitz-continuous on $\mathbb{B}(0,R)$ with Lipschitz-constant $L_R > 0$. Hence, g is Fréchet-differentiable in 0 with the derivative $g'(0) \equiv 0$. Since $\|f'(x)\|_{E'}$ is bounded by L_R for $x \in \mathbb{B}(0,R) \setminus \{0\}$, we also obtain

$$\|g'(x)\|_{E'} = \nu |f(x)|^{\nu-1} \|f'(x)\|_{E'} \leq L_R^\nu \nu \|x\|_E^{\nu-1} \quad \text{for all } x \in \mathbb{B}(0,R) \setminus \{0\},$$

i.e., $g'(x) \to 0 = g'(0)$ in E' if $x \to 0$ in E. □

Note that the norm function on a normed space *cannot* be continuously differentiable in 0 since this would imply continuous differentiability of the real absolute value function in 0. The derivative of J_c as in (3.3) in a point $u \in \mathbb{W}_p^{1,\varrho}(J; X, Y)$ in direction h from the same space is given by the expression

$$\begin{aligned} \mathsf{J}_c'(u)h = \int_J \varrho \|u'(t)\|_X^{\varrho-1} &\langle \mathsf{n}_X'(u'(t)), h'(t) \rangle_{X',X} \\ &+ p \|u(t)\|_Y^{p-1} \langle \mathsf{n}_Y'(u(t)), h(t) \rangle_{Y',Y} \,\mathrm{d}t. \end{aligned} \tag{3.4}$$

We summarize the above considerations in the following theorem.

Theorem 3.1.9. *Let* X, Y *be Banach spaces with* $Y \hookrightarrow_{\mathrm{d}} X$*, let* $1 < \varrho, p < \infty$*, and set* $\xi \coloneqq p\,(1 + \frac{1}{p} - \frac{1}{\varrho})$*.*

(i) Assume that there exists $0 < \theta < 1/\xi'$ *such that both* $(X, Y)_{\theta,1} \hookrightarrow U$ *and* $\mathrm{C}^\beta(J; U) \hookrightarrow \mathcal{X}(J; U)$ *for a* $0 < \beta < 1/\varrho' - \theta(1 + \frac{1}{p} - \frac{1}{\varrho})$*. Then* $\mathbb{W}_p^{1,\varrho}(J; X, Y) \hookrightarrow \mathcal{X}(J; U)$*.*

(ii) Assume that the norm functions n_X *and* n_Y *are each continuously differentiable away from* 0*. Then* J_c *as defined in* (3.3) *is coercive, lower semicontinuous and continuously differentiable on* $\mathbb{W}_p^{1,\varrho}(J; X, Y)$ *with its derivative given as in* (3.4)*.*

Since $\mathsf{J}_c'(u)$ will play a prominent role in the final form of first order necessary conditions for the (reduced) optimal control problem, it is of particular interest to obtain a form of $\mathsf{J}_c'(u)$ which is favorable for further considerations from there. The most critical part in $\mathsf{J}_c'(u)$ is the one involving the time derivatives.

We propose to choose $\varrho = 2$ and X to be a *Hilbert* space, as far as the application allows it. Let us investigate the consequences of this choice. First, the derivative of J_c in direction h becomes

$$\mathsf{J}_c'(u)h = \int_J 2\big(u'(t), h'(t)\big)_X + p\|u(t)\|_Y^{p-1}\langle \mathsf{n}_Y'(u(t)), h(t)\rangle_{Y',Y}\,\mathrm{d}t,$$

which already looks much nicer, but much more importantly depends *linearly* on the time derivative u', which will prove valuable. We illustrate this by a simple example: Suppose that we are dealing with the unconstrained optimization problem

$$\min_{u \in \mathbb{W}_p^{1,2}(J;X,Y)} \mathsf{J}_c(u).$$

The first order optimality conditions of this toy problem are classically given by $\mathsf{J}_c'(u)h = 0$, i.e.,

$$\int_J \big(u'(t), h'(t)\big)_X + \frac{p}{2}\|u(t)\|_Y^{p-1}\langle \mathsf{n}_Y'(u(t)), h(t)\rangle_{Y',Y}\,\mathrm{d}t = 0, \tag{3.5}$$

for all $h \in \mathbb{W}_p^{1,2}(J; X, Y)$. The *linear* dependence on u' now allows to derive the following additional information about u: Choose $h \in \mathrm{C}_c^\infty(J) \otimes$

$Y \subset \mathbb{W}_p^{1,2}(J; X, Y)$ of the form $h = \varphi \otimes v$ with $\varphi \in \mathrm{C}_c^\infty(J)$ and $v \in Y$. Then we have, cf. [48, Ch. VIII, §1],

$$\begin{aligned}\int_J (u'(t), h'(t))_X \,\mathrm{d}t &= \left(\int_J u'(t)\varphi'(t)\,\mathrm{d}t, v\right)_X \\ &= -\frac{p}{2}\left\langle \int_J \varphi(t)\|u(t)\|_Y^{p-1}\mathsf{n}'_Y(u(t))\,\mathrm{d}t, v\right\rangle_{Y',Y}\end{aligned}$$

for all $v \in Y$. Using the Riesz representative $\mathbf{u'(t)} \in X'$ of $u'(t) \in X$ for every $t \in J$, we obtain

$$\begin{aligned}&\int_J \mathbf{u'(t)}\varphi'(t)\,\mathrm{d}t \\ &\qquad = -\frac{p}{2}\int_J \varphi(t)\|u(t)\|_Y^{p-1}\mathsf{n}'_Y(u(t))\,\mathrm{d}t \quad \text{in } Y' \quad \text{for all } \varphi \in \mathrm{C}_c^\infty(J).\end{aligned}$$

But by the definition of the distributional time derivative this implies exactly that

$$\mathbf{u''(t)} = \frac{p}{2}\|u(t)\|_Y^{p-1}\mathsf{n}'_Y(u(t)) \quad \text{in } Y' \quad \text{f.a.a. } t \in J \tag{3.6}$$

and even $\mathbf{u''} \in \mathrm{L}^{p'}(J; Y')$. Going back to (3.5), we may now apply the partial integration formula from Theorem 1.4.5 for $\mathbf{u'} \in \mathbb{W}_2^{1,p'}(J; Y', X')$ and $h \in \mathbb{W}_p^{1,2}(J; X, Y)$ to obtain

$$\begin{aligned}\int_J (u'(t), h'(t))_X \,\mathrm{d}t &= \int_J \langle \mathbf{u'(t)}, h'(t)\rangle_{X',X}\,\mathrm{d}t \\ = -\int_J \langle \mathbf{u''(t)}, h(t)\rangle_{Y',Y}\,\mathrm{d}t &+ \langle \mathbf{u'(T_1)}, h(T_1)\rangle_\xi - \langle \mathbf{u'(T_0)}, h(T_0)\rangle_\xi,\end{aligned}$$

where $\langle \cdot, \cdot \rangle_\xi$ denotes the dual pairing between $(X, Y)_{1/\xi',\xi}$ and its dual space with $\xi = p(\frac{1}{2} + \frac{1}{p}) = \frac{p+2}{2}$. Inserting the derived expression for $\mathbf{u''(t)}$ shows that $\mathbf{u'(T_1)} = \mathbf{u'(T_0)} = 0$ in $(X, Y)_{1/\xi',\xi}$. Hence we have derived that the optimal solution u of the above exemplary unconstrained optimization problem satisfies $\mathbf{u''} \in \mathrm{L}^{p'}(J; Y')$ given by (3.6) with the boundary conditions $\mathbf{u'(T_1)} = \mathbf{u'(T_0)} = 0$ in $(X, Y)_{1/\xi',\xi}$.

The drawbacks of the choice $\varrho = 2$ and X a Hilbert space are that we now have to find a Hilbert space X for which $(X,Y)_{\theta,1}$ embeds compactly into U, and that the range of θ and thus also that of α is now more restricted: We have

$$\mathbb{W}^{1,2}_p(J;X,Y) \hookrightarrow \mathrm{C}^\alpha(J;(X,Y)_{\theta,1}) \tag{3.7}$$

for $0 < \theta < \frac{p}{p+2}$ and $0 < \alpha < 1/2 - \theta\,\frac{2+p}{2p}$. The limitation of the interpolation order together with the choice of a Hilbert space X, expected to be an L^2 space, may require to choose Y as a rather strong space if U itself is so.

A paper about this setting including further results involving the spaces $\mathbb{W}^{1,2}_p(J;X,Y)$ in an abstract optimal control setting including control constraints is in preparation by the author together with Joachim Rehberg and Christian Meyer [115]. For control-constrained problems, one obtains merely a variational inequality instead of an equation as in in (3.5), which poses an interesting problem. We moreover refer to Chapter 4 for an application of the above technique with $U = \mathrm{W}^{-1,q}_D(\Omega)$, $X = \mathrm{L}^2(\Gamma)$ and $Y = \mathrm{L}^p(\Gamma)$ for p large enough, for $E = \mathrm{tr}^*$, the adjoint trace operator.

3.2 First order necessary optimality conditions

Having established that there *exist* (globally) optimal solutions to (rOC), we next aim to *characterize* such optimal solutions. It is known already from classical finite-dimensional nonlinear programming theory that such characterizations will in general only be available for *locally* optimal solutions, and there is no way to identify a globally optimal solution solely from its optimality conditions.

Definition 3.2.1 (Locally optimal control/solution)**.** Let $\bar{u} \in \mathcal{M}^{\mathrm{ad}}_c$. We then say that $\bar{u}$ is *locally optimal* (for (rOC)) if there exists $\varepsilon > 0$ such that $\mathsf{j}(\bar{u}) \le \mathsf{j}(u)$ for all $u \in \mathcal{M}^{\mathrm{ad}}_c \cap \mathbb{B}(\bar{u},\varepsilon)$.

We have seen above that $\mathcal{U}_g$, the set of global controls, is implicitly contained in the optimal control formulation as we have done it. In principle, $\mathcal{U}_g$ and thus also $\mathcal{U}_g^{\mathrm{ad}}$ could be an extremely unpleasant set, in particular non-convex which would permit the applicability of classical KKT theory in function space. However, it will turn out that, when preparing to derive first order necessary optimality conditions, one obtains openness of $\mathcal{U}_g$ in $\mathcal{U}$ for free. This means that $\mathcal{U}_g$ poses no restriction locally in $\mathcal{U}^{\mathrm{ad}}$, which will allow to obtain classical first order necessary optimality conditions in which one does *not* have to refer to $\mathcal{U}_g$ in any way, except for the designated optimal control $\bar{u}$ being from $\mathcal{U}_g^{\mathrm{ad}}$ of course. This also includes the constraint qualification we use, cf. Theorem 3.2.10 below. The treatment will be fairly standard, except for the bits where the set of global controls comes into play. We refer to [148, Ch. 6], [86, Ch. 1.7.3] and [27, Ch. 2.3.4, Ch. 3] for a comprehensive treatment of first order necessary conditions.

First order optimality conditions will require a first order approximation of our problem to exist – in other words, we need differentiability assumptions on the data in (rOC). Maybe a bit surprisingly, this will necessarily require the coefficient function σ to be a *local* mapping in time. The reason behind this is that we need perturbation results for nonautonomous maximal parabolic regularity which require to be able to identify the operator under consideration in a pointwise-in-time sense, see Corollary 1.4.21. We implicitly assume the same on F.

The following assumptions hold true for the rest of this chapter.

(i) The functions $\mathsf{J}_s \colon \mathcal{Y}_{r,q} \to \mathbb{R}_0^+$, $\mathsf{J}_c \colon \mathcal{U} \to \mathbb{R}_0^+$ and $\mathcal{G} \colon \mathrm{C}(\overline{Q}) \to \mathcal{C}$ are continuously differentiable.

(ii) The coefficient function σ is continuously differentiable as a mapping from $\mathbb{W}^{1,r}(J; \mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega))$ to $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and we can identify its derivative $\sigma'(y)$ in $y \in \mathcal{Y}_{r,q}$ with a function in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ itself.

(iii) The function $(y, u) \ni \mathcal{Y}_{r,q} \times \mathcal{X}(J; U) \to F_u(y) \in \mathcal{Z}_{r,q}$ is continuously differentiable with

$$\partial_y F_u(y_u) \in \mathrm{L}^r\Big(J; \mathscr{L}(\mathrm{Y}_{r,q}; \mathrm{W}_D^{-1,q}(\Omega))\Big) + \mathrm{L}^\varrho\Big(J; \mathscr{L}((\mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega))_{\theta,\infty}; \mathrm{W}_D^{-1,q}(\Omega))\Big)$$

for some $1/r' < \theta < 1$ and $r < \varrho \leq \infty$, for all $u \in U_g$.

Remark 3.2.2. The conditions on F are quite abstract and may be sensible to verify. They ensure that $\partial_y F_{\bar{u}}(y_{\bar{u}})$ is a suitable perturbation for nonautonomous maximal parabolic regularity, see Corollary 1.4.21, which includes the necessity to be able to interpret it in a pointwise-in-time fashion. Every other characterization of suitable perturbations for nonautonomous maximal parabolic regularity could also be used, but the author is not aware of any results with a global-in-time approach.

Since we work with the reduced formulation, we will rely on the classical implicit function theorem which we borrow from the book of Lang [103, Ch. XIV, Thm. 2.1]:

Theorem 3.2.3 (Implicit function theorem)**.** *Let* Y, U *and* Z *be Banach spaces, let* $\mathscr{Y} \subseteq \mathsf{Y}$ *and* $\mathscr{U} \subseteq \mathsf{U}$ *be open subsets, and suppose that the function* $\mathsf{G}\colon \mathscr{Y} \times \mathscr{U} \to \mathsf{Z}$ *is continuously differentiable. Let* $(\bar{y}, \bar{u}) \in \mathscr{Y} \times \mathscr{U}$ *and assume that*

$$\partial_y \mathsf{G}(\bar{y}, \bar{u}) \in \mathscr{L}_{\mathrm{iso}}(\mathsf{Y}; \mathsf{Z}).$$

Then there exist open neighborhoods $\mathfrak{Y} \subseteq \mathscr{Y}$ *and* $\mathfrak{U} \subseteq \mathscr{U}$ *of* $\bar{y}$ *and* $\bar{u}$*, respectively, together with an unique* implicit function $\varphi\colon \mathfrak{U} \to \mathfrak{Y}$ *satisfying* $\varphi(\bar{u}) = \bar{y}$ *such that*

$$\mathsf{G}(\varphi(u), u) = \mathsf{G}(\bar{y}, \bar{u}) \quad \textit{for all } u \in \mathfrak{U}.$$

This implicit function φ *is also continuously differentiable and its derivative is given by*

$$\varphi'(u) = -\big[\partial_y \mathsf{G}(\varphi(u), u)\big]^{-1} \partial_u \mathsf{G}(\varphi(u), u)$$

for all $u \in \mathfrak{U}$.

Usually, the implicit function theorem is used in optimal control theory to validate continuous differentiability of the control-to-state operator $\mathcal{S}$ or S, which clearly coincides with the implicit function φ for the choice $\mathsf{G} = e$ or $\mathsf{G} = \mathsf{e}$, respectively. The assertion that there *exists* such a function in the first place is usually known in advance from PDE theory for a continuum of controls, which also lets the open neighborhoods on which the implicit function is defined often become overlooked as a technical detail. In our case however, they are exactly the wished-for information, because they tell us that if we are having a global control $u_g \in U_g$ satisfying $e(y_{u_g}, u_g) = 0$ at hand, then there exists an *open neighborhood* $\mathfrak{U}$ of u_g in $\mathcal{X}(J;U)$ such that $e(\varphi(u), u) = 0$ with $\varphi(u) \in \mathcal{Y}_{r,q}$ for all $u \in \mathfrak{U}$ – or, in other words: $\mathfrak{U} \subset U_g$ due to uniqueness of solutions. Before we prove the theorem stating the corresponding result, let us note that there are variants of the implicit function theorem with weaker assumptions regarding differentiability at the expense of assuming existence of the implicit function, which is a suitable setting for optimal control problems. We refer to [63, 150].

Theorem 3.2.4. *The set* U_g *of global controls is open in* $\mathcal{X}(J;U)$ *and the control-to-state operator* $\mathcal{S}\colon U_g \to \mathcal{Y}_{r,q}$ *is continuously differentiable with the derivative*

$$\mathcal{S}(u)' = -\big[\partial_y e(y_u, u))\big]^{-1} \partial_u e(y_u, u) \in \mathscr{L}(\mathcal{X}(J;U); \mathcal{Y}_{r,q})$$

for all $u \in U_g$.

Proof. Let $(y, u) \in \mathcal{Y}_{r,q} \times \mathcal{X}(J;U)$. Recall that $e\colon \mathcal{Y}_{r,q} \times \mathcal{X}(J;U) \to \mathcal{Z}_{r,q} \times Y_{r,q}$ was given by

$$e(y, u) \coloneqq \big((\partial + \mathcal{A}_\rho(\sigma(y)) + \mathcal{B}_\gamma)y - F(y, u), \delta_{T_0} y\big).$$

We first show that e is continuously differentiable. The distributional time derivative ∂ and the boundary form operator $\mathcal{B}_\gamma$ are continuous linear operators from $\mathcal{Y}_{r,q}$ to $\mathcal{Z}_{r,q}$ and thus continuously differentiable. The same holds for the point evaluation δ_{T_0} as a continuous linear operator from $\mathcal{Y}_{r,q}$ to $Y_{r,q}$, recall (1.32). Moreover, we have assumed that F is continuously differentiable as a mapping from $\mathcal{Y}_{r,q} \times \mathcal{X}(J;U)$ to $\mathcal{Z}_{r,q}$. This leaves us with the divergence-gradient operator

$$\mathcal{Y}_{r,q} \ni y \mapsto \mathcal{A}_\rho(\sigma(y))y = -\nabla \cdot \sigma(y)\rho\nabla y \in \mathcal{Z}_{r,q}. \tag{3.8}$$

We dissect the operator: The most critical is the nonlinear one, $y \mapsto \sigma(y)$, which we however have assumed to be continuously differentiable mapping $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ into itself. From there, both

$$\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})) \ni \sigma(y) \to \sigma(y)\rho \in \mathrm{C}(\overline{J};\mathrm{L}^\infty(\Omega))$$

and

$$\mathrm{C}(\overline{J};\mathrm{L}^\infty(\Omega)) \ni \sigma(y)\rho \to -\nabla \cdot \sigma(y)\rho\nabla \in \mathrm{C}\Big(\overline{J};\mathscr{L}(\mathrm{W}^{1,q}_D(\Omega);\mathrm{W}^{-1,q}_D(\Omega))\Big)$$

are just linear continuous mappings, which means that $y \mapsto -\nabla \cdot \sigma(y)\rho\nabla$ is continuously differentiable from $\mathcal{Y}_{r,q} \hookrightarrow \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ to $\mathrm{C}(\overline{J};\mathscr{L}(\mathrm{W}^{1,q}_D(\Omega);\mathrm{W}^{-1,q}_D(\Omega)))$. Identifying $-\nabla \cdot \sigma(y)\rho\nabla$ as an operator in $\mathscr{L}(\mathcal{Y}_{r,q};\mathcal{Z}_{r,q})$, a product rule shows that the mapping in (3.8) is indeed continuously differentiable.

Next, we calculate the derivative of the y-component of e in a point (y,u) in direction $\zeta \in \mathcal{Y}_{r,q}$ to

$$\begin{aligned}&\partial_y e(y,u)\zeta \\ &\quad = (\partial\zeta + \mathcal{A}_\rho(\sigma(y))\zeta + \mathcal{B}_\gamma\zeta + \mathcal{A}_\rho(\sigma'(y)\zeta)y - \partial_y F(y,u)\zeta, \delta_{T_0}\zeta).\end{aligned} \tag{3.9}$$

For the assumption in the implicit function theorem, we have to show that $\partial_y e(y_{\bar u},\bar u)$ is a topological isomorphism between $\mathcal{Y}_{r,q}$ and $\mathcal{Z}_{r,q} \times Y_{r,q}$ for every $\bar u \in U_g$. We have already seen in Lemma 1.5.23 that $\mathcal{A}_\rho(\sigma(y_{\bar u}))$ satisfies nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,q}_D(\Omega)$ with domain $\mathrm{W}^{1,q}_D(\Omega)$ while depending continuously on t. Hence, $\partial_y e(y_{\bar u},\bar u)$ being a topological isomorphism is equivalent to nonautonomous maximal parabolic regularity of the operator $\mathcal{A}_\rho(\sigma(y_{\bar u})) + \mathcal{B}_\gamma + \mathcal{A}_\rho(\sigma'(y_{\bar u})\cdot)y_{\bar u} - \partial_y F(y_{\bar u},\bar u)$ on $\mathrm{W}^{-1,q}_D(\Omega)$ with domain $\mathrm{W}^{1,q}_D(\Omega)$, thanks to Lemma 1.4.16. This means that it remains to verify that the remaining addends $\mathcal{B}_\gamma + \mathcal{A}_\rho(\sigma'(y_{\bar u})\cdot)y_{\bar u} - \partial_y F(y_{\bar u},\bar u)$ are suitable perturbations of nonautonomous maximal parabolic regularity in the sense of Corollary 1.4.21. For $\partial_y F(y_{\bar u},\bar u)$, this is the case by assumption, and we also have seen that $\mathcal{B}_\gamma$ is a feasible perturbation in Lemma 1.5.23. Regarding $\mathcal{A}_\rho(\sigma'(y_{\bar u})\cdot)y_{\bar u}$, we observe that

$$\mathcal{A}_\rho(\sigma'(y_{\bar u})(\cdot)\xi)y_{\bar u}(\cdot) \in \mathrm{L}^r\big(J;\mathrm{W}^{-1,q}_D(\Omega)\big) \quad \text{for all } \xi \in \mathrm{L}^\infty(\Omega),$$

which by definition of strong measurability for operator-valued functions and Remark 1.5.4 further implies that

$$\Big[t \mapsto \mathcal{A}_\rho(\sigma'(y_{\bar u})(t)\cdot)y_{\bar u}(t)\Big] \in \mathrm{L}^r\Big(J;\mathscr{L}(\mathrm{L}^\infty(\Omega);\mathrm{W}^{-1,q}_D(\Omega))\Big).$$

Here, we have needed the assumption on $\sigma'(y_{\bar u})$ to be meaningful as a function from $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ itself. Now recalling $Y_{r,q} \hookrightarrow \mathrm{C}(\overline{\Omega})$, we find the assumptions of the perturbation Corollary 1.4.21 to be satisfied by $\mathcal{B}_\gamma + \mathcal{A}_\rho(\sigma'(y_{\bar u})\cdot)y_{\bar u} - \partial_y F(y_{\bar u},\bar u)$, which shows that $\partial_y e(y_{\bar u},\bar u)$ is indeed a topological isomorphism. The assertions follow from the implicit function theorem. □

We obtain two corollaries which are mere reformulations of the assertion in Theorem 3.2.4 and (3.9).

Corollary 3.2.5. *The set $\mathcal{U}_g = \mathsf{E}^{-1}[U_g] \subseteq \mathcal{U}$ is open and the control-to-state operator S on $\mathcal{U}$ is also continuously differentiable with the derivative*

$$\mathsf{S}(u)' = \mathcal{S}(\mathsf{E}u)' \circ \mathsf{E} = -\big[\partial_y e(y_u,\mathsf{E}u))\big]^{-1}\partial_u e(y_u,\mathsf{E}u)\mathsf{E} \in \mathscr{L}(\mathcal{U};\mathcal{Y}_{r,q})$$

for all $u \in U_g$.

Corollary 3.2.6 (Solution of the abstract linearized system)**.** *The directional derivative of the derivative of the control-to-state operator*

$\zeta := \mathcal{S}'(u)h \in \mathcal{Y}_{r,q}$ corresponding to $u \in U_g$ and $h \in \mathcal{X}(J;U)$ is given by the unique solution of the abstract linearized equation

$$\begin{aligned}\zeta'(t) + (\mathcal{A}_\rho(\sigma(y_u)) + \mathcal{B}_\gamma)\zeta(t) + \mathcal{A}_\rho(\sigma'(y_u)\zeta)y_u(t) \\ = \partial_y F(y_u,u)\zeta(t) - \partial_u F(y_u,u)h(t) \quad \text{in } \mathrm{W}_D^{-1,q}(\Omega), \quad \zeta(T_0) = 0,\end{aligned}$$

satisfied for almost all $t \in J$.

Now that we know that U_g and thus also $\mathcal{U}_g$ are in fact *open* sets, we will be able to derive first order necessary conditions which do not depend on these sets in the variational formulations. Let us note that continuous differentiability of $\mathcal{S}$ and S, respectively, would have been a crucial cornerstone for the following considerations in any way. In this sense, the openness property is indeed for free.

Definition 3.2.7 (Lagrangian function)**.** The *Lagrangian function* $\mathfrak{L}\colon \mathcal{U}_g \times \mathcal{C}' \to \mathbb{R}$ associated with (rOC) is given by

$$\mathfrak{L}(u,p) := \mathsf{j}(u) + \langle p, \mathcal{G}(\mathsf{S}(u))\rangle_{\mathcal{C}',\mathcal{C}}.$$

Our assumptions imply that the Lagrangian function is continuously differentiable with respect to u:

Corollary 3.2.8. *The Lagrangian function $\mathfrak{L}$ is continuously differentiable with respect to the $\mathcal{U}_g$-variable and its derivative is given by*

$$\partial_u \mathfrak{L}(u,p) = \mathsf{J}_c'(u) + \mathsf{S}'(u)^*\big(\mathsf{J}_s'(y_u) + \mathcal{G}'(y_u)^* p\big) \quad \text{in } \mathcal{U}'$$

for all $(u,p) \in \mathcal{U}_g \times \mathcal{C}'$.

Proof. We calculate as follows, where $\mathcal{G}$ is assumed to be continuously differentiable by assumption and S is so by Theorem 3.2.4:

$$\langle \partial_u \mathfrak{L}(u,p), h\rangle_{\mathcal{U}',\mathcal{U}}$$

$$= \langle \mathsf{J}_s'(y_u), \mathsf{S}'(u)h \rangle_{\mathcal{Y}_{r,q}', \mathcal{Y}_{r,q}} + \langle \mathsf{J}_c'(u), h \rangle_{\mathcal{U}', \mathcal{U}} + \langle p, \mathcal{G}'(y_u)\mathsf{S}'(u)h \rangle_{\mathcal{C}', \mathcal{C}}$$
$$= \langle \mathsf{J}_c'(u) + \mathsf{S}'(u)^* (\mathsf{J}_s'(y_u) + \mathcal{G}'(y_u)^* p), h \rangle_{\mathcal{U}', \mathcal{U}} .$$

Since the preceding equality holds true for *all* directions $h \in \mathcal{U}$, the assertion follows. □

We define the notion of a Lagrangian multiplier, already in a suitable sense without a reference to $\mathcal{U}_g$. The formulation of first order necessary conditions for (rOC) will essentially be that there *exists* such a Lagrangian multiplier.

Definition 3.2.9 (Lagrangian multiplier and KKT conditions)**.** We say that $\bar{p} \in \mathcal{C}'$ is a *Lagrangian multiplier* associated with the state constraint in (rOC), if for a locally optimal control $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$ of (rOC) the *Karush-Kuhn-Tucker (KKT) conditions*

$$\bar{p} \in \mathcal{K}^\circ, \tag{3.10}$$
$$\langle \bar{p}, \mathcal{G}(y_{\bar{u}}) \rangle_{\mathcal{C}', \mathcal{C}} = 0, \tag{3.11}$$
$$\langle \partial_u \mathfrak{L}(\bar{u}, \bar{p}), u - \bar{u} \rangle_{\mathcal{U}', \mathcal{U}} \geq 0 \quad \text{for all } u \in \mathcal{U}^{\mathrm{ad}} \tag{3.12}$$

are satisfied, where

$$\mathcal{K}^\circ := \Big\{ p \in \mathcal{C}' \colon \langle p, \varphi \rangle_{\mathcal{C}', \mathcal{C}} \leq 0 \text{ for all } \varphi \in \mathcal{K} \Big\}$$

is the *polar cone* of $\mathcal{K}$.

Classically, one would expect at least a reference to $\mathcal{U}_g^{\mathrm{ad}}$ in the variational inequality (3.12), as $\mathcal{U}_g^{\mathrm{ad}}$ is the set which puts constraints on the control. The next theorem shows that one indeed obtains the optimality conditions as in Definition 3.2.9 if a Lagrangian multiplier exists for a locally optimal control, even though $\mathcal{U}_g^{\mathrm{ad}}$ is in general non-convex. It is well-known that one needs a so-called *regularity condition* or *constraint qualification* in order to ensure the existence of a Lagrangian multiplier. The classical constraint qualifications are those of Robinson [129], originating from perturbation theory, and Zowe and Kurcyusz [156] from the 1970ies. We will use a variant of those, the *linearized Slater condition.* Let us, in addition to the two classical papers, also point to the book of Bonnans and Shapiro [27, Ch. 2.3.4, Ch. 3] for a (very) comprehensive treatment.

Theorem 3.2.10 (Existence of a Lagrangian multiplier)**.** *Let $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$ be a locally optimal control for (rOC) and let the following so-called* linearized Slater condition *be satisfied: There exists $\bar{u} \neq u^\star \in \mathcal{U}^{\mathrm{ad}}$ such that*

$$G(\mathsf{S}(\bar{u})) + G'(\mathsf{S}(\bar{u}))\big[\mathsf{S}'(\bar{u})(u^\star - \bar{u})\big] \in \operatorname{int} \mathcal{K}. \tag{3.13}$$

Then there exists a Lagrangian multiplier $\bar{p} \in \mathcal{C}'$ associated with the state constraint in (rOC), i.e., such that (3.10)–(3.12) *is satisfied.*

Proof. Since $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$ must be a *global* control, Theorem 3.2.3 shows that there exists a number $\varepsilon > 0$ such that $\mathbb{B}(\bar{u}, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}} \subset \mathcal{U}_g^{\mathrm{ad}}$. Note that $\mathbb{B}(\bar{u}, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}}$ is a *convex* set. This means we can apply standard KKT theory to the auxiliary optimal control problem

$$\begin{aligned} \min_{u \in \mathcal{U}} \quad & \mathsf{j}(u) \\ \text{s.t.} \quad & \begin{cases} u \in \mathbb{B}(\bar{u}, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}} \\ \mathcal{G}(y_u) \in \mathcal{K}, \end{cases} \end{aligned} \tag{rOC$_\varepsilon$}$$

for which $\bar{u}$ must still be a locally optimal control. Note that the existence of $u^\star$ such that (3.13) is satisfied is not an immediate constraint qualification for (rOC)$_\varepsilon$ since we have not assumed $u^\star \in \mathbb{B}(\bar{u}, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}}$. But let $0 < \alpha \leq 1$ and set $u^\star(\alpha) := \alpha u^* + (1-\alpha)\bar{u}$ with $u^\star(1) = u^\star$. Then

$$\mathsf{S}(\bar{u}) + \mathsf{S}'(\bar{u})(u^\star(\alpha) - \bar{u}) = \underbrace{\alpha\big(\mathsf{S}(\bar{u}) + \mathsf{S}'(\bar{u})(u^\star - \bar{u})\big)}_{\in \operatorname{int} \mathcal{K}} + \underbrace{(1-\alpha)\mathsf{S}(\bar{u})}_{\in \mathcal{K}} \in \operatorname{int} \mathcal{K},$$

since $\mathcal{K}$ was a closed convex cone. This means that every point $u^\star(\alpha)$ for $0 < \alpha \leq 1$ is a suitable linearized-Slater point such that (3.13) is satisfied—recall that $\mathcal{U}^{\mathrm{ad}}$ was convex—and there exists $\bar{\alpha}$ small enough such that $u^\star(\bar\alpha) \in \mathbb{B}(\bar u, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}}$. But then $u^\star(\bar\alpha)$ serves as a linearized-Slater point for $(\mathrm{rOC})_\varepsilon$. In this sense, the original condition posed in the assumptions of this theorem is indeed a constraint qualification for $(\mathrm{rOC})_\varepsilon$ and from [34, Thm. 5.2] or [27, Thm. 3.9] or the classical [156, Thm. 3.1] we obtain the existence of a Lagrangian multiplier $\bar p$ such that (3.10) and (3.11) and

$$\left\langle \partial_u \mathfrak{L}(\bar u, \bar p), u - \bar u \right\rangle_{\mathcal{U}',\mathcal{U}} \geq 0 \quad \text{for all } u \in \mathbb{B}(\bar u, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}} \tag{3.14}$$

are satisfied. We apply the convexity trick from above again: For an arbitrary $u \in \mathcal{U}^{\mathrm{ad}}$ and $0 \leq \alpha \leq 1$, let $u(\alpha) := \alpha u + (1-\alpha)\bar u$. Then $u(\alpha) \in \mathcal{U}^{\mathrm{ad}}$ for all $0 \leq \alpha \leq 1$, and choosing $\bar\alpha$ small enough we again have $u(\bar\alpha) \in \mathbb{B}(\bar u, \varepsilon) \cap \mathcal{U}^{\mathrm{ad}}$. Inserting this $u(\bar\alpha)$ in the above variational inequality (3.14), we find

$$\left\langle \partial_u \mathfrak{L}(\bar u, \bar p), u(\bar\alpha) - \bar u \right\rangle_{\mathcal{U}',\mathcal{U}} = \bar\alpha \left\langle \partial_u \mathfrak{L}(\bar u, \bar p), u - \bar u \right\rangle_{\mathcal{U}',\mathcal{U}} \geq 0,$$

hence (3.14) in fact implies (3.12). $\square$

Remark 3.2.11. The constraint qualification in Theorem 3.2.10 is exactly the one which one would pose for the optimal control problem

$$\begin{aligned} &\min_{u \in \mathcal{U}} \quad \mathsf{j}(u) \\ &\text{s.t.} \quad \begin{cases} u \in \mathcal{U}^{\mathrm{ad}} \\ \mathcal{G}(y_u) \in \mathcal{K}, \end{cases} \end{aligned} \tag{$(\mathrm{rOC})_\varepsilon$}$$

so the classical optimal control problem in which one does not worry about global solutions y_u or global controls u. This means that the technical restriction to global solutions and global controls has no influence at all on first order optimality theory in itself. The technique displayed above is of course not only applicable to the restriction to global controls, but generally for every subproblem obtained by intersecting the admissible set $\mathcal{U}^{\mathrm{ad}}$ with an open set. Another generic case for which this is necessary is when e or e are continuously differentiable only for a certain subset of controls and states, see e.g. [63].

In principle, Theorem 3.2.10 admits a complete description of first order necessary optimality conditions to be satisfied for every locally optimal control $\bar{u}$. For the sake of completeness and also in view of the benefit when doing numerical implementations, we also derive the usual adjoint calculus for a more direct interpretation of the variational inequality (3.12). Inserting the formula for $\partial_u \mathfrak{L}(\bar{u}, \bar{p})$ from Corollary 3.2.8, we find (3.12) to be equivalent to

$$\Big\langle \mathsf{J}_c'(\bar{u}) + \mathsf{S}'(\bar{u})^*\big(\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^*\bar{p}\big), u - \bar{u}\Big\rangle_{\mathcal{U}',\mathcal{U}} \geq 0 \quad \text{for all } u \in \mathcal{U}^{\mathrm{ad}}.$$

The goal is to find a suitable interpretation of $\mathsf{S}'(\bar{u})^*\big(\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^*\bar{p}\big)$. In view of the form of $\mathsf{S}'(\bar{u}) = -[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})]^{-1}\partial_u e(y_{\bar{u}}, \mathsf{E}\bar{u})\mathsf{E}$ as seen in Theorem 3.2.4 or Corollary 3.2.8, we need to make sense of

$$\bar{\mathsf{p}} := -\mathsf{E}^*\big[\partial_u e(y_{\bar{u}}, \mathsf{E}\bar{u})\big]^*\big[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})\big]^{-*}\big(\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^*\bar{p}\big) \in \mathcal{U}', \quad (3.15)$$

of which the interesting part is of course $[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})]^{-*} \in \mathscr{L}((\mathcal{Z}_{r,q}' \times Y_{r,q}); \mathcal{Y}_{r,q}')$. Since we will operate in dual spaces a lot now, let us briefly recall that

$$\mathcal{Z}_{r,q}' = \mathrm{L}^{r'}(J; \mathrm{W}_D^{1,q'}(\Omega)) \quad \text{and} \quad Y_{r,q}' = \big(\mathrm{W}_D^{1,q'}(\Omega), \mathrm{W}_D^{-1,q'}(\Omega)\big)_{1/r,r'} = Y_{r',q'},$$

as seen in Lemma 1.1.14 and (1.12).

Definition 3.2.12 (Abstract adjoint equation)**.** Let the pair $(y,u) \in \mathcal{Y}_{r,q} \times \mathcal{X}(J;U)$ be fixed and let $f \in \mathcal{Y}_{r,q}'$ and $\xi_T \in Y_{r,q}'$ be given. Then the following equation for $(\xi, \chi) \in \mathcal{Z}_{r,q}' \times Y_{r,q}'$ in $\mathcal{Y}_{r,q}'$ is called the *abstract adjoint equation*:

$$\begin{aligned} &-\partial\xi + \mathcal{A}_{\rho^\top}(\sigma(y))\xi + \mathcal{B}_\gamma\xi \\ &\quad = -(\sigma'(y)\rho\nabla y)\cdot\nabla\xi + \big[\partial_y F(y,u)\big]^*\xi + \delta_{T_1}^*\otimes\xi_T - \delta_{T_0}^*\otimes\chi + f, \quad (3.16)\end{aligned}$$

where $\delta_{T_0}^*$ and $\delta_{T_1}^*$ are the adjoint operators to the linear continuous point evaluations on $\mathrm{C}(\overline{J}; Y_{r,q})$. Moreover, we have identified $\big[\mathcal{A}_\rho(\sigma(y))\big]^* \xi$ and $\mathcal{B}_\gamma^* \xi$ with the formal but intuitive expressions $\mathcal{A}_{\rho^\top}(\sigma(y))\xi$ and $\mathcal{B}_\gamma \xi$, and

$$\Big[\zeta \mapsto \Big\langle (\sigma'(y)\rho\nabla y) \cdot \nabla \xi, \zeta \Big\rangle \\ \coloneqq \int_J \int_\Omega (\sigma'(y)\zeta\rho\nabla y)\nabla\xi \,\mathrm{d}x\,\mathrm{d}t\Big] \in \mathrm{L}^\infty(J; \mathrm{L}^\infty(\Omega))'$$

on the right-hand side.

Note that $-\partial\xi$ in the preceding definition is indeed meant only in a distributional sense since $\xi \in \mathcal{Z}_{r,q}'$ admits no weak derivative, which is also the reason why we have to "manually" add the initial– and terminal value in the right-hand side in (3.16). In this sense, (3.16) is to be seen as a very weak formulation of a backwards-in-time evolution equation. As the attentive reader has probably already guessed, the terms involving ξ and χ in the equation are altogether exactly $[\partial_y e(y,u)]^*(\xi,\chi)$, thus the following theorem is not very surprising:

Theorem 3.2.13. *For every $(y,u) = (y_u,u) \in \mathcal{Y}_{r,q} \times U_g$ and $(f,\xi_T) \in \mathcal{Y}_{r,q}' \times Y_{r,q}'$, the abstract adjoint equation* (3.16) *admits a unique solution $(\xi,\chi) \in \mathcal{Z}_{r,q}' \times Y_{r,q}'$ which is given by*

$$(\xi,\chi) = \big[e_y(y_u,u)\big]^{-*}(f + \delta_{T_0}^* \xi_T).$$

Proof. The abstract adjoint equation is just a more elaborate way to put the equation

$$\big[\partial_y e(y_u,u)\big]^*(\xi,\chi) = f + \delta_{T_0}^*\xi_T \quad \text{in } \mathcal{Y}_{r,q}'.$$

But we have already seen in the proof of Theorem 3.2.4 that $e_y(y_u,u)$ is continuously invertible as a continuous linear operator from $\mathcal{Y}_{r,q}$ to $\mathcal{Z}_{r,q}$ when $u \in U_g$. Hence, $[e_y(y_u,u)]^*$ is also continuously invertible as a continuous linear operator from $\mathcal{Z}_{r,q}'$ to $\mathcal{Y}_{r,q}'$ and (ξ,χ) is uniquely given by $[e_y(y_u,u)]^{-*}(f + \delta_{T_0}^*\xi_T)$. ☐

If f and the adjoint operator of $\partial_y F(y,u)$ possess slightly better properties, then we obtain a weak derivative for the solution ξ of the abstract adjoint equation:

Lemma 3.2.14. *Let $f \in \mathrm{L}^1(J;\mathrm{W}_D^{-1,q'}(\Omega))$ and assume that $[\partial_y F(y_u,u)]^*$ is an element of $\mathscr{L}(\mathcal{Z}'_{r,q};\mathrm{L}^1(J;\mathrm{W}_D^{-1,q'}(\Omega))$ for some $u \in U_g$. Then the unique solution $(\xi,\chi) \in \mathcal{Z}'_{r,q} \times Y'_{r,q}$ of* (3.16) *even satisfies*

$$\xi \in \mathrm{W}^{1,1}(J;\mathrm{W}_D^{-1,q}(\Omega)) \cap \mathcal{Z}'_{r,q} = \mathbb{W}^{1,1}_{r'}(J;\mathrm{W}_D^{-1,q'}(\Omega),\mathrm{W}_D^{1,q'}(\Omega)).$$

Proof. Let $(\xi,\chi) \in \mathcal{Z}'_{r,q} \times Y'_{r,q}$ be the unique solution of (3.16) for the given data. Due to the assumption on σ', we can interpret the term $(\sigma'(y_u)\rho\nabla y_u)\cdot\nabla\xi$ as an $\mathrm{L}^1(J;\mathrm{L}^1(\Omega))$- and thus $\mathrm{L}^1(J;\mathrm{W}_D^{-1,q'}(\Omega))$-function via $\mathrm{W}_D^{1,q}(\Omega) \hookrightarrow \mathrm{L}^\infty(\Omega)$. We argue in a similar fashion as in Chapter 3.1.1. Let us choose the test function $\zeta = \varphi \otimes v$ with $\varphi \in \mathrm{C}_c^\infty(J)$ and $v \in \mathrm{W}_D^{1,q}(\Omega)$. Then $\zeta \in \mathcal{Y}_{r,q}$ and we can test (3.16) against ζ:

$$\begin{aligned}\Big\langle \int_J \xi(t)\varphi'(t) + \varphi(t)\Big(\mathcal{A}_{\rho^\top}(\sigma(y_u)(t))\xi(t) + \mathcal{B}_\gamma(t)\xi(t)\Big)\,\mathrm{d}t, v\Big\rangle \\ = \Big\langle \int_J \varphi(t)\Big(-(\sigma'(y_u)(t)\rho\nabla y_u(t))\cdot\nabla\xi(t) \\ + \big[\partial_y F(y_u,u)\big]^*\xi(t) + f(t)\Big)\,\mathrm{d}t, v\Big\rangle\end{aligned}$$

with the brackets denoting dual pairings of $\mathrm{W}_D^{-1,q'}(\Omega)$ against $\mathrm{W}_D^{1,q}(\Omega)$. The involved operators are from $\mathrm{W}_D^{-1,q'}(\Omega)$ for almost all $t \in J$ by construction or assumption. Since the preceding equality is true for all $v \in \mathrm{W}_D^{1,q}(\Omega)$, we have

$$\begin{aligned}&\int_J \xi(t)\varphi'(t) + \varphi(t)\Big(\mathcal{A}_{\rho^\top}(\sigma(y_u)(t))\xi(t) + \mathcal{B}_\gamma(t)\xi(t)\Big)\,\mathrm{d}t \\ &= \int_J \varphi(t)\Big(-(\sigma'(y_u)(t)\rho\nabla y_u(t))\cdot\nabla\xi(t) + \big[\partial_y F(y_u,u)\big]^*\xi(t) + f(t)\Big)\,\mathrm{d}t\end{aligned}$$

in $\mathrm{W}_D^{-1,q'}(\Omega)$ for all $\varphi \in \mathrm{C}_c^\infty(J)$, which means exactly that the distributional derivative of ξ is given by

$$\xi' = \mathcal{A}_{\rho^\top}(\sigma(y_u))\xi + \mathcal{B}_\gamma\xi + \big(\sigma'(y_u)\rho\nabla y_u\big)\cdot\nabla\xi - \big[\partial_y F(y_u,u)\big]^*\xi - f$$

revealing itself to be an $\mathrm{L}^1(J;\mathrm{W}^{-1,q'}_D(\Omega))$ function. In particular, ξ' is indeed a weak derivative. Finally, the inclusion $\xi \in \mathbb{W}^{1,1}_{r'}(J;\mathrm{W}^{-1,q'}_D(\Omega),\mathrm{W}^{1,q'}_D(\Omega))$ follows from the embedding $\mathcal{Z}'_{r,q} \hookrightarrow \mathrm{L}^1(J;\mathrm{W}^{-1,q'}_D(\Omega))$. □

Remark 3.2.15.

(i) The regularity obtained for ξ in Theorem 3.2.13 is quite low, which is to be expected for the extremely low-regularity data $f \in \mathcal{Y}'_{r,q}$ and the mapping properties of $[\partial_y F(y_u,u)]^*$, cf. [5] for a systematic treatment. We have seen in Lemma 3.2.14 how to improve this regularity, but even under additional assumptions, say, $f \in \mathrm{L}^{r'}(J;\mathrm{W}^{-1,q'}_D(\Omega))$ and $[\partial_y F(y_u,u)]^*$ also mapping to that space, one will *not* be able to derive more than $\mathrm{L}^1(J;\mathrm{W}^{-1,q'}_D(\Omega))$-regularity for ξ'. The culprit here is the term $\big(\sigma'(y_u)\rho\nabla y_u\big)\cdot\nabla\xi$ on the right-hand side in (3.16) which admits only L^1-integrability, since the integrabilities $\mathrm{L}^r(J;\mathrm{L}^q(\Omega;\mathbb{R}^d))$ of ∇y_u and $\mathrm{L}^{r'}(J;\mathrm{L}^{q'}(\Omega;\mathbb{R}^d))$ of $\nabla\xi$ add up exactly to 1. The origin of this problem is generic, namely that σ needs to operate on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$.

(ii) The lack of regularity as just explained also implies that $[\partial_y e(y_u,u)]^*$ *cannot* admit nonautonomous maximal parabolic $\mathrm{L}^{r'}$-regularity on $\mathrm{W}^{-1,q'}_D(\Omega)$ with domain $\mathrm{W}^{1,q'}_D(\Omega)$. Such a phenomenon does not appear in the autonomous case, where the adjoint of an operator satisfying autonomous maximal parabolic regularity also does so in the fitting adjoint spaces.

(iii) In view of the considerations laid out in Chapter 3.1.1, one might expect to also be able to derive boundary values $\xi(T_1)$ and $\xi(T_0)$ of ξ in the setting of Lemma 3.2.14. Note, however, that the partial integration formula from Theorem 1.4.5 only works for $1 < r, s < \infty$ because it builds upon the maximal regularity embedding (1.30) in

Lemma 1.4.4, cf. [146, Ch. 1.8.2].

Now that we have identified the action of $[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})]^{-*}$, we can characterize $\bar{\mathsf{p}}$ completely within the adjoint calculus:

Theorem 3.2.16 (First order necessary optimality conditions)**.** *Let $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$ be a locally optimal control of (rOC) for which the linearized Slater condition* (3.13) *is satisfied. Then there exists a Lagrangian multiplier $\bar{p} \in \mathcal{C}'$ such that the following* first order necessary optimality conditions *are satisfied:*

$$\bar{p} \in \mathcal{K}^\circ, \tag{3.10}$$

$$\langle \bar{p}, \mathcal{G}(y_{\bar{u}}) \rangle_{\mathcal{C}',\mathcal{C}} = 0, \tag{3.11}$$

$$\big\langle \mathsf{J}_c'(\bar{u}) + \bar{\mathsf{p}}, u - \bar{u} \big\rangle_{\mathcal{U}',\mathcal{U}} \geq 0 \quad \text{for all } u \in \mathcal{U}^{\mathrm{ad}}, \tag{3.17}$$

where $\bar{\mathsf{p}} \in \mathcal{U}'$ is the adjoint state *given by*

$$\bar{\mathsf{p}} = -\mathsf{E}^* \big[\partial_u e(y_{\bar{u}}, \mathsf{E}\bar{u})\big]^* (\bar{\xi}, \bar{\chi})$$

for the unique solution $(\bar{\xi}, \bar{\chi}) \in \mathcal{Z}_{r,q}' \times Y_{r,q}'$ of the abstract adjoint equation

$$\begin{aligned} &- \partial\xi + \mathcal{A}_{\rho^\top}(\sigma(y_{\bar{u}}))\xi + \mathcal{B}_\gamma \xi \\ &\qquad = -(\sigma'(y_{\bar{u}})\rho\nabla y) \cdot \nabla\xi + \big[\partial_y F(y_{\bar{u}}, \bar{u})\big]^* \xi - \delta_{T_0}^* \otimes \chi \\ &\qquad\qquad + \mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^* \bar{p}. \end{aligned}$$

Proof. The existence of the Lagrangian multiplier $\bar{p}$ under the constraint qualification (3.13) is exactly the result of Theorem 3.2.10. This includes the relations (3.10) and (3.11). From the variational inequality (3.12) combined with the expression for $\partial_u \mathfrak{L}(\bar{u}, \bar{p})$ from Corollary 3.2.8 and the formula for $\mathsf{S}'(\bar{u})$ from Corollary 3.2.5, we obtain exactly (3.17) for $\bar{\mathsf{p}}$ given by

$$\bar{\mathsf{p}} \coloneqq -\mathsf{E}^* \big[\partial_u e(y_{\bar{u}}, \mathsf{E}\bar{u})\big]^* \big[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})\big]^{-*} \big(\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^* \bar{p}\big) \in \mathcal{U}'. \tag{3.15}$$

Comparing this formula with the result of Theorem 3.2.13 shows that

$$(\bar{\xi}, \bar{\chi}) = \left[\partial_y e(y_{\bar{u}}, \mathsf{E}\bar{u})\right]^{-*} \left(\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^* \bar{p}\right).$$

Thereby, $\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^* \bar{p}$ is admissible data for the abstract adjoint equation because $\mathsf{J}_s'(y_{\bar{u}}) \in \mathcal{Y}_{r,q}'$ and $\mathcal{G}'(y_{\bar{u}})^* \in \mathscr{L}(\mathcal{C}'; \mathcal{M}(\overline{Q}))$, hence $\mathcal{G}'(y_{\bar{u}})^* \bar{p} \in \mathcal{M}(\overline{Q}) \hookrightarrow \mathcal{Y}_{r,q}'$ due to

$$\mathcal{Y}_{r,q} \hookrightarrow \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega})) \doteq \mathrm{C}(\overline{Q}),$$

and Theorem 1.2.8, cf. (2.40) for the embedding. Inserting all these identities yields the claim. □

Let us close this section with some concluding remarks.

Remark 3.2.17.

(i) In view of the variational inequality (3.17), let us once more point to a possibly favorable choice of $\mathcal{U}$ in Chapter 3.1.1 and to Chapter 4 where this technique is applied.

(ii) The term $\delta_{T_1}^* \xi_T$ corresponding to the designated terminal value present in the abstract adjoint equation in Definition 3.2.12 is not explicitly stated in the corresponding equation in Theorem 3.2.16, but may well be included in $\mathsf{J}_s'(y_{\bar{u}}) + \mathcal{G}'(y_{\bar{u}})^* \bar{p}$. We have already seen in the proof of Theorem 3.2.16 that $\nu \coloneqq \mathcal{G}'(y_{\bar{u}})^* \bar{p} \in \mathcal{M}(\overline{Q})$ is a measure. By restriction, we can decompose this measure into a sum of measures ν_0, ν_1, ν_Q in $\mathcal{M}(\overline{Q})$ with support in $\{T_0\} \times \overline{\Omega}$, in $\{T_1\} \times \overline{\Omega}$, and in $J \times \overline{\Omega}$. In turn, the first two of these measures may be written in the form $\nu_0 = \delta_{T_0}^* \otimes \nu_{T_0}$ and $\nu_1 = \delta_{T_1}^* \otimes \nu_{T_1}$, i.e., we have

$$\mathcal{G}'(y_{\bar{u}})^* \bar{p} = \delta_{T_0}^* \otimes \nu_{T_0} + \delta_{T_1}^* \otimes \nu_{T_1} + \nu_Q.$$

This corresponds to $\xi_T = \nu_{T_1}$ with $\nu_{T_1} \in \mathcal{M}(\overline{\Omega}) \hookrightarrow Y_{r,q}'$, the latter thanks to $Y_{r,q} \hookrightarrow \mathrm{C}(\overline{\Omega}) \hookrightarrow \mathrm{L}^2(\Omega)$, cf. (2.39), and Theorem 1.2.8.

(iii) Another generic case for a T_1-based measure to appear in the abstract adjoint equation is when J_s is of end-time tracking type, typically

$$\mathsf{J}_s(y) := \frac{1}{2}\int_E |y(T_1) - y_{\mathrm{ob}j}|^2 \,\mathrm{dx}$$

for some design area $E \subseteq \Omega$ and an objective state $y_{\mathrm{obj}} \in \mathrm{L}^2(E)$. Then we have

$$\mathsf{J}_s'(y)\zeta = \int_E (y(T_1) - y_{\mathrm{obj}})\zeta(T_1)\,\mathrm{dx}$$

and hence $\mathsf{J}_s'(y)$ can be written as $\delta_{T_1}^* \otimes \mathbf{1}_E(y(T_1) - y_{\mathrm{obj}})$ with the indicator function $\mathbf{1}_E$ and $\mathbf{1}_E(y(T_1) - y_{\mathrm{obj}}) \in \mathrm{L}^2(\Omega) \hookrightarrow \mathcal{M}(\overline{\Omega}) \hookrightarrow Y_{r,q}'$ as above. This is exactly the form in which ξ_T in Definition 3.2.12 enters the equation.

(iv) If the initial value y_0 in the evolution equation (3.1) is not subject to the control process, as we have implicitly assumed from the start, then the derivative $\partial_u e(y,u)$ of the state equation with respect to u will be zero in the second component, cf. Corollary 3.2.6. Accordingly, the adjoint operator of $\partial_u e(y,u)$ will then be constantly zero for the second component of its input $\mathscr{Z}_{r,q}' \times Y_{r,q}'$. In this sense, the second variable χ in the abstract adjoint equation plays no big role in our considerations, but would become important if we also aimed at controlling the initial value y_0.

CHAPTER 4

The thermistor problem

This final chapter serves as a showcase for a practical application and a proof-of-concept for the abstract theory collected in the previous chapters. We will show that a quite challenging optimal control problem fits in the developed framework. This optimal control problem is built around the so-called *thermistor problem* and looks as follows:

$$\left.\begin{aligned} \min \quad & \frac{1}{2}\|\theta(T_1)-\theta_{\mathrm{obj}}\|^2_{\mathrm{L}^2(E)} + \frac{\gamma}{\mathfrak{r}}\|\nabla\theta\|^{\mathfrak{r}}_{\mathrm{L}^{\mathfrak{r}}(J;\mathrm{L}^q(\Omega))} \\ & \qquad\qquad + \frac{\beta}{2}\int_{\Sigma_N} (u')^2 + |u|^p \,\mathrm{d}(\lambda\otimes\omega) \\ \text{s.t.} \quad & (4.1)\text{–}(4.6) \\ \text{and} \quad & \theta(t,\mathrm{x}) \le \theta_{\max}(t,\mathrm{x}) \quad \text{a.e. in } J\times\Omega, \\ & 0 \le u(t,\mathrm{x}) \le u_{\max}(t,\mathrm{x}) \quad \text{a.e. on } J\times N, \end{aligned}\right\} \tag{P}$$

where (4.1)–(4.6) refer to the following coupled PDE system consisting of the instationary nonlinear heat equation and the quasi-static potential equation, also known as the actual *thermistor problem*:

$$\partial_t\theta - \operatorname{div}(\eta(\theta)\kappa\nabla\theta) = (\sigma(\theta)\rho\nabla\varphi)\cdot\nabla\varphi \qquad \text{in } Q \tag{4.1}$$

$$\nu\cdot\eta(\theta)\kappa\nabla\theta + \alpha\theta = \alpha\theta_l \qquad \text{on } \Sigma_{\partial\Omega} := J\times\partial\Omega \tag{4.2}$$

$$\theta(T_0) = \theta_0 \qquad \text{in } \Omega \tag{4.3}$$

$$-\operatorname{div}(\sigma(\theta)\rho\nabla\varphi) = 0 \qquad \text{in } Q \tag{4.4}$$
$$\nu \cdot \sigma(\theta)\rho\nabla\varphi = u \qquad \text{on } \Sigma_N := J \times N \tag{4.5}$$
$$\varphi = 0 \qquad J \times D. \tag{4.6}$$

We fix all occurring quantities below. Since this is the first full partial differential equation which we encounter in all its strong form glory, we spend some time explaining the precise model.

The function θ models the the temperature in a conducting material covered by the domain Ω, while φ refers to the electric potential. As usual, the boundary of Ω is denoted by $\partial\Omega$, with the *unit normal* ν facing outward of Ω in almost every boundary point (with respect to the boundary measure ω). We have the boundary decomposition $D \,\dot\cup\, N = \partial\Omega$, where D is closed within $\partial\Omega$. The functions $\eta(\cdot)\kappa$ and $\sigma(\cdot)\rho$ represent heat– and electric conductivity. While κ and ρ are given, prescribed matrix functions, η and σ are allowed to depend on the temperature θ. Moreover, α is the heat transfer coefficient regulating the heat flux through the boundary $\partial\Omega$, and θ_l and θ_0 are given boundary– and initial data, respectively. The quadratic gradient term in (4.1) is known as the *Joule heat*.

Note that a realistic model of heat evolution includes a volumetric heat capacity $\varrho C_p(\theta)$, generally depending on θ, in front of the time derivative. We assume this term to be normalized to one, which can be achieved by re-scaling θ by so-called enthalpy transformation. The effects of this transformation on the remaining quantities in the equation may be absorbed into η, σ and α which does not influence the theory if C_p is reasonably smooth and strictly monotone (see e.g. [22, Sect. 3]).

Finally, u stands for a current which is induced via the boundary part N and which is to be controlled. The bounds in the optimization problem (P) as well as the desired temperature θ_{obj} are given functions and β is the usual Tikhonov regularization parameter; γ has an analogous meaning. In all what follows, the system (4.1)–(4.6) is frequently also called *state system.*

The PDE system (4.1)–(4.6) models the heating of a conducting material by means of an electric current, described by u, induced on the part N of the boundary, which is done for some time $T_1 - T_0$. At the grounding D, homogeneous Dirichlet boundary conditions are given, i.e., the potential is zero, inducing electron flow. Note that, usually, u will be zero on a subset N_0 of N, which corresponds to having insulation at this part of the boundary. We emphasize that the different boundary conditions are essential for a realistic modeling of the process. The objective of (P), realized in the objective functional by the first term, is to adjust the induced current u to minimize the L^2-distance between the desired and the resulting temperature at end time T_1 on the set $E \subseteq \Omega$, the latter representing the area of the material in which one is interested. The other terms are present to minimize thermal stresses (second term) and to ensure a certain smoothness of the controls (third term), whose influence to the objective functional, however, may be controlled by the weights γ and β. The actual form of these terms and the size of the integrability orders $\mathfrak{r}$ and p are motivated by functional-analytic considerations, see Chapter 4.2. Moreover, the optimization procedure is subject to pointwise control and state constraints. The control constraints reflect a maximum heating power, while the state constraints limit the temperature evolution to prevent possible damage, e.g. by melting of the material. Similarly to the mixed boundary conditions, the inequality constraints in (P) are essential for a realistic model as demonstrated by the numerical example which we exhibit in Chapter 4.3. Problem (P) is relevant in various applications, such as for instance the heat treatment of steel by means of an electric current.

Up to the authors' best knowledge, there are only few contributions dealing with the optimal control of the thermistor problem. We refer to [40, 90, 91, 104], where two-dimensional problems are discussed. In [104], a completely parabolic problem is discussed, while [91] considers the purely elliptic counterpart to (4.1)–(4.6). In [13, 40, 90], the authors investigate a parabolic-elliptic system similar to (4.1)–(4.6), assuming a particular structure of the controls. In contrast to [91, 104], mixed boundary conditions are considered in [40, 90]. However, all these contributions do not consider pointwise state constraints and non-smooth data or nonlocal coefficients. Thus, (P) differs significantly from the problems considered in the aforementioned papers.

The author was led to the thermistor problem by the paper [88] where problem (P) in two space dimensions is considered, and the friendly invitation of Christian Meyer and Joachim Rehberg to participate in the follow-up project for the treatment of (P) in three spatial dimensions. The paper [88] also accounts for mixed boundary conditions, non-smooth data, and pointwise state constraints. However, the analysis in [88] substantially differs from ours, mainly because of the quasilinear structure in the parabolic equation (4.1) considered here. Hence, main aspects of the present work do not appear in the two-dimensional not fully quasilinear setting. We have to use the full machinery developed in Chapters 2 and 3 to obtain satisfying results for (P) in three spatial dimensions. Thereby, we will also treat the two-dimensional quasilinear case as a byproduct. We will see that this gives a quite nice reflection on the different characters of the problem in differing space dimensions.

We proceed as follows. Since the optimal control problem is put "on top" of the analysis for the state system, we first verify the assumptions of the quasilinear existence theorems as in Chapter 2.2.1. The weapon of choice will be the theorem of Amann, Theorem 2.2.4, for which we have to reduce the system of equations to a single parabolic evolution equation, which will be the most involved part. It will turn out that we are indeed able to verify the assumptions of the global-existence Theorem 2.2.12, except for the uniform boundedness condition (2.45) – the latter is only available for space dimension $d = 2$. This puts us exactly in the setting of Chapter 3, from which we then use the abstract conditions developed in Chapter 3.1 to show existence of globally optimal controls for (P). We further give first order necessary conditions, and finish with a chapter quite different in nature from the previous ones: We validate the necessity to consider the optimal control problem (P) in the form as we do it by showing that the control process needs to be subjected to both control– and state constraints to obtain a correct solution. For this, we have employed numerical calculations whose results we put on display. The results in this chapter in slightly less general form will be published together with Christian Meyer and Joachim Rehberg [113, 114].

We fix the assumptions on the data in (P).

> The following assumptions hold true for the rest of this chapter:
>
> (i) The set $\Omega \subset \mathbb{R}^d$ is a bounded domain for $d \in \{2, 3\}$ and D (like *Dirichlet*) is a closed subset of $\partial\Omega$ with $D \neq \emptyset$. We suppose that $\Omega \cup D$ is *volume-preserving* regular in the sense of Gröger. In all what follows, $\partial\Omega \setminus D$ will be denoted by N (like *Neumann*).
>
> (ii) We consider a finite interval $J = (T_0, T_1) \subset \mathbb{R}_0^+$.
>
> (iii) All Banach spaces and all occurring functions are supposed to be real ones, i.e., we are working in a *real* setting in the sense of Chapter 1.6.

Let us point out that we are per foregoing assumptions in the regime of a Lipschitz domain. This is necessary because we consider a coupled system with different types of boundary conditions in each equation, but need a suitable extension property for *both* of them. For this, we need to fall back to a Lipschitz domain, cf. Chapter 1.3. The Lipschitz property of Ω also provides full boundary embeddings via Corollary 1.3.8 and the square root property in Proposition 1.5.5, recall Remark 1.5.6. The additional assumption of admitting *volume-preserving* boundary charts is mostly a technical one and allows us to use consequences of the global existence Theorem 2.2.12 and the full range of the interpolation embeddings in Lemma 1.5.25.

4.1 The state system

As a logical first step, we address the assumptions regarding (local) existence and uniqueness for the state equation (4.1)–(4.6). In this context, we treat u as a fixed, given inhomogeneity, whereas it is an unknown control function when considering the optimal control problem (P).

The following assumption on the quantities in the state system (4.1)–(4.6) holds true for the rest of this chapter. There exist $q > d$ and $r > 2(1 - \frac{d}{q})^{-1}$ such that the following properties are true:

(i) The functions σ and η map $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ into itself, they satisfy $0 < \sigma_\bullet \leq \sigma(u) \leq \sigma^\bullet$ and $0 < \eta_\bullet \leq \eta(u) \leq \eta^\bullet$ on $\overline{Q}$ for all $u \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$, are uniformly Lipschitz-continuous on bounded sets, and have the Volterra property.

(ii) The coefficient functions satisfy $\rho \in \mathrm{L}^\infty(\Omega; \mathbb{M}_d(\rho_\bullet))$ with

$$-\nabla \cdot \rho \nabla \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,q}_D(\Omega); \mathrm{W}^{-1,q}_D(\Omega)\big) \tag{4.7}$$

and $\kappa \in \mathrm{L}^\infty(\Omega; \mathbb{M}_d(\kappa_\bullet))$ with

$$-\nabla \cdot \kappa \nabla + 1 \in \mathscr{L}_{\mathrm{iso}}\big(\mathrm{W}^{1,q}(\Omega); \mathrm{W}^{-1,q}_\emptyset(\Omega)\big). \tag{4.8}$$

(iii) $\theta_l \in \mathrm{L}^\infty(J; \mathrm{L}^\infty(\partial\Omega))$.

(iv) $\alpha \in \mathrm{L}^\infty(\partial\Omega)$ with $\alpha \geq 0$ ω-a.e. on $\partial\Omega$ and $\int_{\partial\Omega} \alpha \,\mathrm{d}\omega > 0$.

(v) $\theta_0 \in \big(\mathrm{W}^{-1,q}_\emptyset(\Omega), \mathrm{W}^{1,q}(\Omega)\big)_{1/r',r}$.

(vi) $u \in \mathrm{L}^{2s}(J; L^{\mathfrak{p}}(N; \omega))$ for some $r < s \leq \infty$ and $\mathfrak{p} > \frac{d-1}{d} q$.

Note that it is not presumptuous to assume that both differential operators in (4.7) and (4.8) provide topological isomorphisms at the same time, since the latter property mainly depends on the behavior of the discontinuous coefficient functions (versus the geometry of D), and these correspond to the material properties in the workpiece described by the domain Ω, i.e., the coefficient functions should exhibit similar properties with regard to jumps or discontinuities in general, the main obstacles to overcome for the isomorphism property. We refer to [52] for more information.

Moreover, the reader should note that a similar condition as in (4.7) was also posed in [15, Ch. 3] in order to get smoothness of the solution; compare also [65], where exactly this regularity for the solution of Poisson's equation is needed in order to show uniqueness for the semiconductor equations.

Remark 4.1.1.

(i) The theorem of GRÖGER, cf. Theorem 1.5.18, shows that the isomorphism assumptions (4.7) and (4.8) are always satisfied for $d = 2$.

(ii) The regularity of the initial value θ_0 is not supposed to be a "hard" assumption in the sense of posing an obstruction at any point. In other words, we fix its regularity to whatever is needed in order to be able to use suitable existence- and uniqueness results. This may sound like a cheap way to not care about the initial value, and to some extent it is, but is justifiable when considering practical applications. In the application presented in Chapter 4.3 for example, which is concerned with the heat treatment of a workpiece, it seems reasonable to assume that the spatial temperature profile of workpiece at hand is homogeneously constant to e.g. the surrounding temperature.

Let us define what we understand as a *solution* to the system (4.1)–(4.6). Not surprisingly, it is essentially an adaption of Definition 2.2.1 to the system case. Recall the boundary operators from Definition 1.5.11 and their time-extension in Section 2.1.

Definition 4.1.2 (Abstract solution concept)**.** We say that the pair of functions (θ, φ) with $\theta(T_0) = \theta_0$ is a *local solution* of the general thermistor problem for $\mathfrak{u} \in \mathrm{L}^s(J; \mathrm{W}_D^{-1,q}(\Omega))$ if there is a number $T_0 < T^\bullet \leq T_1$ such that (θ, φ) satisfies the equations

$$\begin{aligned} \partial\theta + (\mathcal{A}_\kappa(\eta(\theta)) + \mathcal{B}_\alpha)\theta &= (\sigma(\theta)\rho\nabla\varphi)\cdot\nabla\varphi + \mathcal{B}_\alpha\theta_l \quad \text{in } \mathrm{W}^{-1,q}_{\emptyset}(\Omega), && (4.9)\\ -\nabla\cdot\sigma(\theta)\rho\nabla\varphi &= \mathfrak{u} \quad \text{in } \mathrm{W}^{-1,q}_{D}(\Omega) && (4.10)\end{aligned}$$

on $J^\bullet = (T_0, T^\bullet)$ and thereby admits the regularity

$$(\theta, \varphi) \in \mathbb{W}^{1,r}(T_0, T_\bullet; \mathrm{W}^{-1,q}_{\emptyset}(\Omega), \mathrm{W}^{1,q}(\Omega)) \times \mathrm{L}^{2r}(T_0, T_\bullet; \mathrm{W}^{1,q}_{D}(\Omega)) \quad (4.11)$$

for every $T_\bullet \in J^\bullet$.

(i) If $T^\bullet = T_1$ or (4.11) is not true for $T_\bullet = T^\bullet$, then we say that (θ, φ) is a *maximal* local solution and call $J^\bullet$ the *maximal interval of existence.*

(ii) If $T^\bullet = T_1$ and

$$(\theta, \varphi) \in \mathbb{W}^{1,r}(J; \mathrm{W}^{-1,q}_{\emptyset}(\Omega), \mathrm{W}^{1,q}(\Omega)) \times \mathrm{L}^{2r}(J; \mathrm{W}^{1,q}_{D}(\Omega)),$$

then we say that (θ, φ) is a *global solution.*

For $\mathfrak{u} = \mathcal{B}u$, we obtain a solution to the thermistor problem as stated in (4.1)–(4.6).

Remark 4.1.3.

(i) The reader will verify that the boundary conditions imposed on φ in (4.5) for $\mathfrak{u} = \mathcal{B}u$ and in (4.6) are incorporated in this definition in the sense of [66, Ch. II.2] or [39, Ch. 1.2]. For an adequate interpretation of the boundary conditions for θ as in (4.2), see [108, Ch. 3.3.2] and the in-book references there.

(ii) Due to the assumptions on Ω and the integrabilities r and q, a solution θ in the above sense is in fact Hölder-continuous on $[T_0, T_\bullet] \times \overline{\Omega}$ for every $T_\bullet \in J^\bullet$ if we have a local solution, and Hölder-continuous on $\overline{Q}$ for a global solution. This follows from Lemma 1.5.25 and Proposition 1.5.5, as we have the embeddings, exemplarily given for the whole time interval J,

$$\big(\mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega)\big)_{1/r',r} \hookrightarrow \mathrm{C}^\beta(\Omega) \hookrightarrow \mathrm{C}(\overline{\Omega}) \tag{2.39}$$

for some $\beta > 0$ and

$$\mathbb{W}^{1,r}(J; \mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega)) \hookrightarrow \mathrm{C}^\alpha(Q) \hookrightarrow \mathrm{C}(\overline{Q}) \tag{2.40}$$

for some $\alpha > 0$.

(iii) Let us mention that the quadratic gradient term on the right-hand side in (4.9) is frequently treated by the so-called "thermistor trick". We briefly explain the basic idea in our setting. Assuming $\mathcal{B}u(t) \in \mathrm{W}_D^{-1,2}(\Omega)$, the Lax-Milgram lemma immediately implies $\varphi(t) \in \mathrm{W}_D^{1,2}(\Omega)$, cf. Lemma 1.5.13. In addition, the classical Stampacchia argument gives $\varphi(t) \in \mathrm{L}^\infty(\Omega)$ (see [97, Thm. II.B.2]). This makes it possible to reformulate the quadratic gradient term via the product rule for the divergence such that one ends up with a right-hand side in (4.9) which is also in $\mathrm{W}_\emptyset^{-1,2}(\Omega)$, whereas the interpretation as it stands would only be in $\mathrm{L}^1(\Omega)$, so merely in $\mathrm{W}_\emptyset^{-1,d'}(\Omega)$. For space dimension $d = 2$, this is no improvement, where it is a rather interesting phenomenon for $d = 3$. While this Hilbert space setting is tempting to use, we have already learned in Chapter 2 that our treatment of the quasilinear structure is massively dependent on the uniform continuity of the designated solution θ, which seems very difficult to obtain in the classical energy setting. We have thus decided not to pursue the thermistor trick any further.

(iv) We briefly recall the considerations from the introduction regarding a volumetric heat capacity term in the form $\varrho C_p(\theta)$ in front of the time derivative of θ. As explained there, one may use the so-called enthalpy transformation to get rid of the additional dependency on θ, thereby modifying the data η, σ and α. Now, considering that we are allowing κ and ρ to be spatially discontinuous to account for heterogeneous material, one might be tempted to let ϱ also be of that form, say, $\varrho \in \mathrm{L}^\infty(\Omega)$ with a strictly positive essential lower bound. However, in order to return to a divergence-gradient structure as in (4.1) after applying the enthalpy transformation, this essentially requires ϱ to act as a multiplier on $\mathrm{W}^{-1,q}_{\emptyset}(\Omega)$ which calls for $\varrho \in \mathrm{W}^{1,q}(\Omega)$ – in particular, $\mathrm{L}^\infty(\Omega)$ is *not* enough. We refer to [22, Sect. 3] and [80, Ch. 6].

We formulate the main result for this section.

Theorem 4.1.4 (Existence and uniqueness for the state system)**.** *For every general control function* $\mathfrak{u} \in \mathrm{L}^{2s}(J;\mathrm{W}^{-1,q}_D(\Omega))$ *with* $r < s \leq \infty$ *there exists a* unique maximal *solution of the general thermistor problem* (4.9) *and* (4.10) *in the sense of Definition 4.1.2. If* $d = 2$*, this is always a* global *solution. In particular, there is a unique solution of* (4.1)–(4.6) *for every* $u \in \mathrm{L}^s(J;L^{\mathfrak{p}}(N;\omega))$ *in this sense.*

The proof of this theorem is given in the next subsection by verifying that we can fit the system (4.9)–(4.10) in the framework of Theorem 2.2.10.

Remark 4.1.5. From related work, it is known that one may have to deal with blow-up of solutions already in not fully quasilinear thermistor systems, cf. [15, Ch. 5], [16] and the references therein. Note that in [88], global existence in a quite similar setting as ours for $d = 2$ is proven, albeit only for a linear parabolic equation.

4.1.1 Existence and uniqueness

As indicated above, we will prove Theorem 4.1.4 by reducing the thermistor system to an equation in the temperature θ only and apply Theorem 2.2.10 to the resulting equation. In fact, our assumptions on κ, η and α, the latter identified with its constant time-extension, already imply that the operator $\theta \mapsto \mathcal{A}_\kappa(\eta(\theta)) + \mathcal{B}_\alpha$ satisfies the assumptions on the differential operator in Theorem 2.2.10. For the remaining data, we solve the elliptic equation (4.10) for φ uniquely for every time point t in dependence of θ and u and re-insert the such-obtained $\varphi(\theta, u)$ into the right-hand side in (4.9), which we then can write as a function in θ and u. We then show that this function, at this point for fixed u, satisfies the suppositions on F in the local existence– and uniqueness theorem.

So, let us first consider quadratic gradient term in (4.9) together with (4.10) in dependence of θ and u. A preliminary result highlighting the quadratic structure is the following.

Lemma 4.1.6. *For every* $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$,

$$\mathfrak{b}_\theta(\varphi) \coloneqq \mathfrak{b}_\theta(\varphi, \varphi) \coloneqq \big(\sigma(\theta)\rho\nabla\varphi\big) \cdot \nabla\varphi$$

defines a continuous quadratic form

$$\mathfrak{b}_\theta \colon \mathrm{L}^{2s}\big(J; \mathrm{W}^{1,q}_D(\Omega)\big) \to \mathrm{L}^s\big(J; \mathrm{L}^{q/2}(\Omega)\big).$$

Moreover, the mapping $(\theta, \varphi) \mapsto \mathfrak{b}_\theta(\varphi)$ *is uniformly Lipschitz-continuous on bounded sets in* $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega})) \times \mathrm{L}^{2s}\big(J; \mathrm{W}^{1,q}_D(\Omega)\big)$.

Proof. Symmetry, bilinearity and continuity of

$$(\varphi_1, \varphi_2) \mapsto \mathfrak{b}_\theta(\varphi_1, \varphi_2) \coloneqq \big(\sigma(\theta)\rho\nabla\varphi_1\big) \cdot \nabla\varphi_2$$

for each $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ are clear from the assumptions on ρ and σ and Hölder's inequality, hence $\mathfrak{b}_\theta$ is a continuous quadratic form. The second assertion follows from a straightforward calculation with the resulting estimate

$$\begin{aligned}
&\|\mathfrak{b}_{\theta_1}(\varphi_1) - \mathfrak{b}_{\theta_2}(\varphi_2)\|_{\mathrm{L}^{2s}(J;\mathrm{L}^{q/2}(\Omega))} \\
&\qquad \le \|\sigma(\theta_1) - \sigma(\theta_2)\|_{\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))}\|\rho\|_{\mathrm{L}^\infty(\Omega;\mathbb{S}_d)}\|\varphi_1\|^2_{\mathrm{L}^{2s}(J;\mathrm{W}^{1,q}_D(\Omega))} \\
&+ 2\|\sigma(\theta_2)\|_{\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))}\|\rho\|_{\mathrm{L}^\infty(\Omega;\mathbb{S}_d)}\|\varphi_1\|_{\mathrm{L}^{2s}(J;\mathrm{W}^{1,q}_D(\Omega))}\|\varphi_1 - \varphi_2\|_{\mathrm{L}^{2s}(J;\mathrm{W}^{1,q}_D(\Omega))},
\end{aligned}$$

the assumed Lipschitz-continuity of σ and boundedness of the underlying sets. $\square$

The next lemma shows that the mapping of θ onto the "solution operator" for the elliptic equation (4.10) is bounded and thus Lipschitz-continuous on *compact* sets in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$. We already incorporate a continuous dependence on the control u for further use.

Lemma 4.1.7. *Let $\mathfrak{C} \subset \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ be compact. Then the following assertions are true:*

(i) The mapping

$$\mathfrak{C} \ni \theta \mapsto \mathcal{A}_\rho(\sigma(\theta))^{-1} \in \mathrm{L}^\infty\big(\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_D^{1,q}(\Omega))\big)$$

is bounded and Lipschitz-continuous.

(ii) For every bounded set $\mathfrak{B} \subset \mathrm{L}^{2s}(J;\mathrm{W}_D^{-1,q}(\Omega))$,

$$\mathfrak{C} \times \mathfrak{B} \ni (\theta,\mathfrak{u}) \mapsto \varphi(\theta,\mathfrak{u}) := \mathcal{A}_\rho(\sigma(\theta))^{-1}\mathfrak{u} \in \mathrm{L}^{2s}\big(J;\mathrm{W}_D^{1,q}(\Omega)\big) \tag{4.12}$$

is bounded and Lipschitz-continuous.

Proof. We have already seen in Proposition 1.5.21 that for every $t \in \overline{J}$ we have $\mathcal{A}_\rho(\sigma(\theta))(t) \in \mathscr{L}_{\mathrm{iso}}(\mathrm{W}_D^{1,q}(\Omega);\mathrm{W}_D^{-1,q}(\Omega))$ and that the dependence on t of these operators is continuous for every $\theta \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$. Moreover, with the proof technique as displayed in Remark 1.5.22, we derive

$$\begin{aligned}
\big\|\mathcal{A}_\rho(\sigma(\theta_1))(t)^{-1} - \mathcal{A}_\rho(\sigma(\theta_2))(t)^{-1}\big\|_{\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_D^{1,q}(\Omega))} \\
\le \big\|\mathcal{A}_\rho(\sigma(\theta_1))(t)^{-1}\big\|_{\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_D^{1,q}(\Omega))} \\
\cdot \big\|\mathcal{A}_\rho(\sigma(\theta_2))(t)^{-1}\big\|_{\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_D^{1,q}(\Omega))} \\
\cdot \big\|\mathcal{A}_\rho(\sigma(\theta_1))(t) - \mathcal{A}_\rho(\sigma(\theta_2))(t)\big\|_{\mathscr{L}(\mathrm{W}_D^{1,q}(\Omega);\mathrm{W}_D^{-1,q}(\Omega))}.
\end{aligned}$$

Taking the supremum over $t \in \overline{J}$ on both sides and using the Lipschitz-continuity of $\theta \mapsto \mathcal{A}_\rho(\sigma(\theta))$ on bounded sets (cf. the assumptions on σ

and Remark 1.5.4), it remains to observe that

$$\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega})) \ni \theta \mapsto \mathcal{A}_\rho(\sigma(\theta))^{-1} \in \mathrm{L}^\infty(\mathscr{L}(\mathrm{W}_D^{-1,q}(\Omega);\mathrm{W}_D^{1,q}(\Omega))) \qquad (4.13)$$

is continuous as a concatenation of continuous functions and thus *bounded* over the compact set $\mathfrak{C}$. This proves the first assertion. From there, the second one follows immediately by a triangle argument, using again boundedness of (4.13) over $\mathfrak{C}$. □

Remark 4.1.8.

(i) The foregoing lemma rests fundamentally upon the maximal Sobolev regularity assumption for $-\nabla\cdot\rho\nabla$ and the fact that we work with uniformly continuous θ. Since we have $\nabla\varphi$ in the parabolic equation (4.9), we need maximal Sobolev regularity for $-\nabla\cdot\rho\nabla$ anyway, but it is only together with the regularity of θ that it gives the critical continuity properties as in Lemma 4.1.7.

(ii) We have needed the *compactness* of $\mathfrak{C}$ in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ only to make the mapping (4.13) bounded. In fact, Theorem 1.5.18 tells us that (4.13) is *always* bounded for $d=2$, cf. Remark 4.1.1. We will use this below to obtain global solutions in the two-dimensional case.

The next results establish the right-hand side in the parabolic equation (4.9) with the correct regularity and mapping properties. Moreover, Lipschitz-continuity with respect to the control $\mathfrak{u}$ in the elliptic equation is shown along the way, which will become useful for the optimal control procedure. Let us set

$$\Psi_{\mathfrak{u}}(\theta) := \mathfrak{b}_\theta(\varphi(\theta,\mathfrak{u}))$$

with φ as in (4.12).

The next lemma then follows from combining the foregoing Lemmata 4.1.6 and 4.1.7.

Lemma 4.1.9. *Let* $\mathfrak{C} \subset \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ *be compact and let* $\mathfrak{B}$ *be a bounded set in* $\mathrm{L}^{2s}(J;\mathrm{W}_D^{-1,q}(\Omega))$. *Then*

$$\mathfrak{C} \times \mathfrak{B} \ni (\theta,\mathfrak{u}) \mapsto \Psi_\mathfrak{u}(\theta) \in \mathrm{L}^s(J;\mathrm{L}^{q/2}(\Omega))$$

is Lipschitz-continuous and the Lipschitz-constant of $\theta \mapsto \Psi_\mathfrak{u}(\theta)$ *is bounded over* $\mathfrak{u} \in \mathfrak{B}$.

Following the strategy outline above, we will now specify the mapping F for the application of Theorem 2.2.10.

Proposition 4.1.10. *Let* $\mathfrak{u} \in \mathrm{L}^{2s}(J;\mathrm{W}_D^{-1,q}(\Omega))$ *for* $r < s \leq \infty$ *and set*

$$F(\theta,\mathfrak{u}) := \Psi_\mathfrak{u}(\theta) + \mathcal{B}_\alpha\theta_l. \tag{4.14}$$

Then $F(\cdot,\mathfrak{u})$ *satisfies the assumptions of Theorem 2.2.10 for the spaces* $X = \mathrm{W}_\emptyset^{-1,q}(\Omega)$ *and* $Y = \mathrm{W}^{1,q}(\Omega)$.

Proof. Let $\mathbb{B}_{r,q} := \overline{\mathbb{B}(0,R)} \subset \mathbb{W}^{1,r}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega);\mathrm{W}^{1,q}(\Omega))$ for some $R > 0$. Then $\mathbb{B}_{r,q}$ is a *compact* set in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ by Lemma 1.5.25, cf. also (2.40). The operator F maps into $\mathrm{L}^s(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$ since $\mathcal{B}_\alpha\theta_l \in \mathrm{L}^\infty(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$, see Definition 1.5.11, and Lemma 4.1.9 shows that $F(\theta,\mathfrak{u}) - F(0,\mathfrak{u}) = \Psi_u(\theta)$ is Lipschitz-continuous on $\mathbb{B}_{r,q}$ with values in $\mathrm{L}^s(J;\mathrm{L}^{q/2}(\Omega))$, where the latter space embeds into $\mathrm{L}^s(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$ by the Sobolev embedding Theorem 1.2.27, thanks to $q > d$. The Volterra property of σ is transferred to $\varphi(\theta,\mathfrak{u})$ and thus also to $\Psi_\mathfrak{u}(\theta)$ for every $\theta \in \mathbb{W}^{1,r}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega);\mathrm{W}^{1,q}(\Omega))$. □

The previous proposition allows to use Theorem 2.2.10, which gives the "unique maximal solution" part of Theorem 4.1.4. For $d = 2$, we obtain global solutions by observing that we can dispose of the dependence on *compact* sets in $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ due to the result of Gröger, Theorem 1.5.18. Observe that $F(\cdot,\mathfrak{u})$ is in fact well-defined on $\mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))$ by construction.

Proposition 4.1.11. *Let $d = 2$ and $\mathfrak{u} \in \mathrm{L}^{2s}(J; \mathrm{W}_D^{-1,q}(\Omega))$. Then $F(\cdot, \mathfrak{u})$ as defined in* (4.14) *is uniformly Lipschitz-continuous on bounded sets in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and the uniform boundedness condition*

$$\sup_{\theta \in \mathrm{C}(\overline{J};\mathrm{C}(\overline{\Omega}))} \|F(\theta, \mathfrak{u})\|_{\mathrm{L}^r(J;\mathrm{W}_D^{-1,q}(\Omega))} =: C_{F,\mathfrak{u}} < \infty$$

is satisfied.

Proof. The operators $\mathcal{A}_\rho(\sigma(\theta))(t)$ admit common coercivity bounds $\sigma_\bullet \rho_\bullet$ and common upper bounds $\sigma^\bullet \|\rho\|_{\mathrm{L}^\infty(\Omega;\mathbb{M}_d)}$ of their coefficient functions, for every $t \in \overline{J}$. Theorem 1.5.18 thus shows that there is $q_0 > 2$ such that $\mathcal{A}_\rho(\sigma(\theta))(t)$ is continuously invertible as an operator in $\mathscr{L}(\mathrm{W}_D^{1,q}(\Omega); \mathrm{W}_D^{-1,q}(\Omega))$ and the norm of the inverses are uniformly bounded for all $2 \leq q < q_0$ and *all* $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$. Without loss of generality, we can assume that the number q from the general assumptions of this chapter falls into that range; if it does not, we can diminish it without losing the isomorphism properties (4.7) and (4.8) (see Lemma 1.5.24). Inspecting Lemmata 4.1.7 and 4.1.9, we then find that we can replace $\mathfrak{C}$ by any *bounded* set in $\mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and still retain Lipschitz-continuity of $\theta \mapsto \Psi_\mathfrak{u}(\theta)$ on these sets, because $\varphi(\theta, \mathfrak{u})$ is uniformly bounded in $\mathrm{L}^{2s}(J; \mathrm{W}_D^{1,q}(\Omega))$ for all $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ and $\mathfrak{u}$ from bounded sets now, thanks to the uniform boundedness of $\mathcal{A}_\rho(\sigma(\theta))^{-1}$. This also implies that $\Psi_\mathfrak{u}(\theta)$ and thus $F(\theta, \mathfrak{u})$ are uniformly bounded in $\mathrm{L}^r(J; \mathrm{W}_\emptyset^{-1,q}(\Omega))$ over all $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$. □

We formally collect the overall proof.

Proof of Theorem 4.1.4. The general assumptions of this chapter together with Propositions 4.1.10 and 4.1.11 show that the operators in the equation

$$\partial\theta + (\mathcal{A}_\kappa(\eta(\theta)) + \mathcal{B}_\alpha)\theta = \Psi_\mathfrak{u}(\theta) + \mathcal{B}_\alpha \theta_l \quad \text{in } \mathrm{W}_\emptyset^{-1,q}(\Omega), \quad \theta(T_0) = \theta_0$$

satisfy the assumptions of Theorem 2.2.10, and for space dimension $d = 2$ even the assumptions of Theorem 2.2.12. Thus, the equation admits a unique maximal solution $\theta \in \mathbb{W}^{1,r}(T_0, T_\bullet; \mathrm{W}_\emptyset^{-1,q}(\Omega), \mathrm{W}^{1,q}(\Omega))$ for every $T_\bullet \in J^\bullet(\mathfrak{u})$, the maximal interval of existence; in case of a global solution, $T_\bullet = T_1$. For $d = 2$, the solution is always global. We now set $\varphi := \varphi(\theta, \mathfrak{u})$ as in (4.12). Then the pair (θ, φ) is a local solution of the general thermistor system in the sense of Definition 4.1.2. The case $\mathfrak{u} = \mathcal{B}u$ follows directly, for which we obtain a solution of (4.1)–(4.6). □

4.2 The optimal control problem

This subchapter marks the transition to the optimal control problem (P). We will show that this optimal control problem fits in the abstract framework developed in Chapter 3. Having done so, we "only" need to show that the particular right-hand side $(\theta, u) \mapsto F(\theta, \mathcal{B}u) = \Psi_{\mathcal{B}u}(\theta) + \mathcal{B}_\alpha \theta_l$ satisfies the assumptions posed in Chapter 3 to obtain a concise and well-rounded first-order optimality theory. It will turn out that we need the considerations in Chapter 3 in their full strength, at least for the three-dimensional case, so we assume that the reader is familiar with them. The assumptions on the state system in Chapter 3 are clearly satisfied by ours and the foregoing Proposition 4.1.10.

The following assumptions on the quantities in the optimal control problem (P) hold true for the rest of this chapter.

(i) On the integrability orders and on γ in the objective functional we suppose the following:

   a) For $d = 2$: $\gamma = 0$ and $p = 2$ with formally $\mathfrak{r} > r$,

   b) for $d = 3$: $\gamma > 0$ and $p > \frac{4}{3}q - 2$. Moreover, let $2 < \mathfrak{q} < q_0$ where q_0 is the number from Gröger's Theorem 1.5.18 corresponding to the operators $-\nabla \cdot \sigma(\theta)(t)\rho\nabla$ with $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$, and set $\varsigma := \frac{d\mathfrak{q}}{2d-\mathfrak{q}}$. Then we require $\mathfrak{r}$ to satisfy $\mathfrak{r} > 2(1 - \frac{d}{q})^{-1}(1 + \frac{d}{\varsigma} - \frac{d}{q})$.

(ii) E is an open, not necessarily proper, subset of Ω.

(iii) $\theta_{\mathrm{obj}} \in \mathrm{L}^2(E)$.

(iv) $\theta_{\max} \in \mathrm{C}(\overline{Q})$ with $\max(\max_{\overline{\Omega}} \theta_0, \operatorname{ess\,sup}_{\Sigma_{\partial\Omega}} \theta_l) \leq \theta_{\max}(t, \mathrm{x})$ for all $(t, \mathrm{x}) \in \overline{Q}$ and $\theta_0(\mathrm{x}) < \theta_{\max}(T_0, \mathrm{x})$ for all $\mathrm{x} \in \overline{\Omega}$.

(v) $u_{\max}$ is a given function on Σ_N with $u_{\max}(t, \mathrm{x}) \geq 0$ a.e. on Σ_N.

(vi) $\beta > 0$.

Note that we do not impose any regularity assumptions on the function $u_{\max}$. In particular, it is allowed that $u_{\max} \equiv \infty$ so that no upper bound is present.

We recall the setting from Chapter 3 and show that we are in the situation depicted there. For brevity, we also use the abbreviations introduced there:

$$\begin{aligned}
\mathcal{Y}_{r,q} &\coloneqq \mathbb{W}^{1,r}\big(J; \mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega)\big), \\
\mathcal{Z}_{r,q} &\coloneqq \mathrm{L}^r\big(J; \mathrm{W}_D^{-1,q}(\Omega)\big), \\
Y_{r,q} &\coloneqq \big(\mathrm{W}_D^{-1,q}(\Omega), \mathrm{W}_D^{1,q}(\Omega)\big)_{1/r',r}.
\end{aligned}$$

As the basic control space, we choose

$$\mathcal{X}(J; U) \coloneqq \mathrm{L}^{2\mathfrak{s}}\big(J; \mathrm{W}_D^{-1,q}(\Omega)\big)$$

with $\mathfrak{s} \geq \mathfrak{r}$, whereas the stronger space is chosen as

$$\mathcal{U} \coloneqq \mathbb{W}_p^{1,2}\big(J; \mathrm{L}^2(N; \omega), \mathrm{L}^p(N; \omega)\big).$$

This is exactly one of the proposed spaces from Chapter 3.1.1, for which we obtain the following result:

Proposition 4.2.1. *The adjoint trace operator $\mathcal{B}$ maps the space $\mathcal{U}$ compactly into $\mathcal{X}(J; U) = \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega))$.*

Proof.

(i) For $d = 2$, we have chosen $p = 2$, thus $\mathcal{U} = \mathrm{W}^{1,2}(J; \mathrm{L}^2(N))$ here, for which we already know that it embeds into $\mathrm{C}^{1/2}(J; \mathrm{L}^2(N))$. The Arzelà-Ascoli Theorem 1.2.5 together with compactness of the boundary trace and thus also that of its adjoint as in Lemma 1.2.57 then gives the assertion.

(ii) For $d = 3$, we use the special case (3.7) of Theorem 3.1.9, which tells us that

$$\mathcal{U} := \mathbb{W}^{1,2}_p(J; \mathrm{L}^2(N; \omega); \mathrm{L}^p(N; \omega)) \hookrightarrow \mathrm{C}^{\varrho}(J; (\mathrm{L}^2(N; \omega), \mathrm{L}^p(N; \omega))_{\tau,1})$$

for $0 < \tau < \frac{p}{2+p}$ and $0 < \varrho < 1/2 - \tau\,\frac{2+p}{2p}$. The compatibility embedding between real- and complex interpolation (1.10) and general interpolation principles for the Lebesgue spaces as in [146, Ch. 1.18.4] show that

$$\big(\mathrm{L}^2(N; \omega), \mathrm{L}^p(N; \omega)\big)_{\tau,1} \hookrightarrow \big[\mathrm{L}^2(N; \omega), \mathrm{L}^p(N; \omega)\big]_{\tau} = \mathrm{L}^{\mathfrak{p}}(N; \omega)$$

with $2 < \mathfrak{p} = \mathfrak{p}(\tau) = (\frac{1-\tau}{2} + \frac{\tau}{p})^{-1} < \frac{2+p}{2}$ for $0 < \tau < \frac{p}{2+p}$. This means for all $2 < \mathfrak{p} < \frac{2+p}{2}$, we can find $\varrho > 0$ depending on $\mathfrak{p}$ such that $\mathcal{U} \hookrightarrow \mathrm{C}^{\varrho}(J; \mathrm{L}^{\mathfrak{p}}(N))$.

If $\mathfrak{p} > \frac{2}{3}q$, then the adjoint trace operator $\mathrm{tr}^*\colon \mathrm{L}^{\mathfrak{p}}(N) \to \mathrm{W}^{-1,q}_D(\Omega)$ is compact, cf. Lemma 1.2.57. To be able to choose such $\mathfrak{p}$, we need $p > \frac{4}{3}q - 2$. Now again the Arzelà-Ascoli Theorem 1.2.5 yields the assertion. ☐

Let us again set

$$e(\theta, u) := \big((\partial + \mathcal{A}_\kappa(\eta(\theta)) + \mathcal{B}_\alpha)\theta - F(\theta, u), \delta_{T_0}\theta - \theta_0\big) = 0$$

as a function

$$e\colon \mathcal{Y}_{r,q} \times \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega)) \to \mathcal{Z}_{r,q} \times Y_{r,q}.$$

To account for the control space $\mathcal{U}$, we moreover again set

$$\mathsf{e}\colon \mathcal{Y}_{r,q} \times \mathcal{U} \to \mathcal{Z}_{r,q} \times Y_{r,q}, \quad \mathsf{e}(y,u) := e(y, \mathsf{E}u)$$

where $\mathsf{E} = \mathcal{B} \in \mathscr{L}(\mathcal{U}; \mathcal{X}(J;U))$, cf. Proposition 4.2.1.
Collecting the data from the objective functional, we find for the part J_s corresponding to the state θ:

$$\mathsf{J}_s(\theta) := \frac{1}{2}\|\delta_{T_1}\theta - \theta_{\mathrm{obj}}\|^2_{\mathrm{L}^2(E)} + \frac{\gamma}{\mathfrak{r}}\|\nabla\theta\|^{\mathfrak{r}}_{\mathrm{L}^{\mathfrak{r}}(J;\mathrm{L}^q(\Omega))},$$

which consists of continuous functions on $\mathcal{Y}_{r,q}$; in fact, it is even continuously differentiable with the derivative represented by

$$\mathsf{J}_s'(\theta) = \delta_{T_1}^* \otimes \mathbf{1}_E(\theta(T_1) - \theta_{\mathrm{obj}}) + \gamma\|\nabla\theta_u\|^{\mathfrak{r}-q}_{\mathrm{L}^q(\Omega)}\Delta_q\theta. \tag{4.15}$$

Here, we have used the considerations in Remark 3.2.17 and the (scaled weak) *q-Laplacian*, which we define via

$$\langle \|\nabla\theta_u\|^{\mathfrak{r}-q}_{\mathrm{L}^q(\Omega)}\Delta_q\theta, \zeta\rangle_{\mathcal{Y}_{r,q}'} := \int_{T_0}^{T_1} \|\nabla\theta_u(t)\|^{\mathfrak{r}-q}_{\mathrm{L}^q}\langle\Delta_q\theta_u(t), \zeta(t)\rangle_{\mathrm{W}^{-1,q'}_{\emptyset}(\Omega)}\,\mathrm{d}t$$

and

$$\langle\Delta_q\psi, \xi\rangle_{\mathrm{W}^{-1,q'}_{\emptyset}(\Omega)} := \int_\Omega |\nabla\psi|^{q-2}\nabla\psi\cdot\nabla\xi\,\mathrm{dx} \qquad \text{for } \psi, \xi \in \mathrm{W}^{1,q}(\Omega).$$

Quite similarly we find

$$\mathsf{J}_c(u) := \frac{1}{2}\int_J \|u'(t)\|^2_{\mathrm{L}^2(N;\omega)} + \|u(t)\|^p_{\mathrm{L}^p(N;\omega)}\,\mathrm{d}t$$

to be continuous and convex on $\mathcal{U} = \mathbb{W}^{1,2}_p\big(J; \mathrm{L}^2(N;\omega); \mathrm{L}^p(N;\omega)\big)$, which implies weak lower semicontinuity. Moreover, J_c is clearly coercive on $\mathcal{U}$ and continuous differentiable with the derivative

$$\begin{aligned}
&\mathsf{J}_c'(u)h \\
&\quad = \beta \int_J (u'(t), h(t))_{\mathrm{L}^2(N;\omega)} + \frac{p}{2}\langle |u(t)|^{p-2}u(t), h(t)\rangle_{\mathrm{L}^{p'}(N;\omega),\mathrm{L}^p(N;\omega)} \,\mathrm{d}t \\
&\qquad\qquad = \beta \int_J \int_N u'(t)h'(t) + \frac{p}{2}|u(t)|^{p-2}u(t)h(t)\,\mathrm{d}\omega\,\mathrm{d}t, \quad (4.16)
\end{aligned}$$

cf. Chapter 3.1.1.
The admissible set for the control $\mathcal{U}^{\mathrm{ad}}$ is given by

$$\mathcal{U}^{\mathrm{ad}} := \Big\{ u \in \mathcal{U} \colon 0 \le u \le u_{\max} \text{ a.e. in } J \times \Omega \Big\},$$

which is clearly closed and convex. We further have classical unilateral state constraints which we model by

$$\mathcal{G}(\theta) := \theta - \theta_{\max}, \quad \mathcal{C} = \mathrm{C}(\overline{Q}) \quad \text{and} \quad K_- = \Big\{ f \in \mathrm{C}(\overline{Q}) \colon f \le 0 \text{ on } \overline{Q} \Big\},$$

cf. (3.2).
A sensible condition to verify is that the feasible set is nonempty, i.e., $\mathcal{M}^{\mathrm{ad}} \neq \emptyset$, or equivalently, that the feasible set for the reduced problem is nonempty, that is, $\mathcal{M}_c^{\mathrm{ad}} \neq \emptyset$. We first obtain that the set of global admissible controls $\mathcal{U}_g^{\mathrm{ad}}$ is indeed nonempty, since the first candidate, the zero control, does the job:

Proposition 4.2.2. *The zero control $u_0 \equiv 0$ is a global control, that is, the associated solution θ_{u_0} is a global one, and $\mathcal{U}_g^{\mathrm{ad}} \neq \emptyset$.*

Proof. The zero control u_0 is obviously included in $\mathcal{U}^{\mathrm{ad}}$, and inserting it into the elliptic equation immediately implies $\varphi(\theta, \mathcal{B}u_0) = 0$ for all $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$ due to the isomorphism property (4.7) and the permanence principle from Proposition 1.5.20. But then $F(\theta, \mathcal{B}u_0) = \mathcal{B}_\alpha \theta_l$ which is independent of θ, and we can apply Corollary 2.2.16 to obtain the global solution θ_{u_0}. □

It remains to show that θ_{u_0} indeed satisfies the state constraints. To give a motivation for the following result, let us briefly assume for simplicity that the ambient temperature θ_l and the initial temperature θ_0 are homogeneous. Intuitively, it is clear that if we apply the zero control, then the temperature profile in Ω should slowly become (homogeneously) equal to the ambient temperature via the Robin boundary conditions. That means that the temperature profile will always be larger than the smaller value of the ambient temperature and the initial temperature profile, and always smaller than the larger value of those two. If the latter is smaller than the state constraints, which seems a very reasonable assumption to begin with, then the solution θ_{u_0} corresponding to the zero control also obeys the state constraints.

We formulate the result, obtained from a classical technique using the variational formulation, slightly more general than just explained, also including the case where we apply a nonzero control and in which we clearly also satisfy the same lower bounds. This is because the Joule heat term in quadratic gradient form on the right-hand side in (4.9) is always nonnegative.

Lemma 4.2.3. *The solutions (θ, φ) of the thermistor system in the sense of Definition 4.1.2 obey natural bounds:*

(i) For every solution (θ, φ) in the sense of Theorem 4.1.4 with maximal existence interval $J^\bullet$, it is true that $\theta(t, \mathrm{x}) \geq \min(\operatorname{ess\,inf}_{\Sigma_{\partial\Omega}} \theta_l, \min_{\overline{\Omega}} \theta_0)$ for all $(t, \mathrm{x}) \in [T_0, T_\bullet] \times \overline{\Omega}$ where $T_\bullet \in J^\bullet(u)$.

(ii) We have $\theta_{u_0} \in \mathcal{G}^{-1}[\mathcal{K}]$, that is, $\theta_{u_0} \leq \theta_{\max}$ on $\overline{Q}$.

Proof. We prove the first assertion. Set $m_{\inf} := \min(\operatorname{ess\,inf}_{\Sigma_{\partial\Omega}} \theta_l, \min_{\overline{\Omega}} \theta_0)$ and

$$\zeta(t) = \theta(t) - m_{\inf}$$

and decompose $\zeta(t)$ into its positive and negative part, that is, $\zeta(t) = \zeta^+(t) - \zeta^-(t)$ with both $\zeta^+(t)$ and $\zeta^-(t)$ being positive functions. By [47, Ch. IV, §7, Prop. 6/Rem. 12], $\zeta^-(t)$ is still an element of $\mathrm{W}^{1,q}(\Omega)$ for almost every $T_0 < t < T_\bullet$. In particular, we may test (4.9) against $-\zeta^-(t)$, insert $\theta = \zeta + m_{\text{inf}}$ and use that m_{inf} is constant:

$$\begin{aligned}
&-\int_\Omega \zeta'(t)\zeta^-(t)\,\mathrm{dx} - \int_\Omega \big(\eta(\theta)(t)\kappa\nabla\zeta(t)\big)\cdot\nabla\zeta^-(t)\,\mathrm{dx} - \int_\Gamma \alpha\zeta(t)\zeta^-(t)\,\mathrm{d}\omega \\
&= -\int_\Gamma \alpha\big(\theta_l(t) - m_{\text{inf}}\big)\zeta^-(t)\,\mathrm{d}\omega - \int_\Omega \zeta^-(t)\big(\sigma(\theta)(t)\rho\nabla\varphi(t)\big)\cdot\nabla\varphi(t)\,\mathrm{dx}.
\end{aligned}$$

Observe that the support of products of $\zeta(t)$ and $\zeta^-(t)$ is exactly the support of $\zeta^-(t)$, and $\zeta(t) = -\zeta^-(t)$ there. We thus obtain (see [149])

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\zeta^-(t)\|^2_{\mathrm{L}^2(\Omega)} + \int_\Omega \big(\eta(\theta)(t)\kappa\nabla\zeta^-(t)\big)\cdot\nabla\zeta^-(t)\,\mathrm{dx} + \int_\Gamma \alpha\zeta^-(t)^2\,\mathrm{d}\omega& \\
= -\int_\Gamma \alpha\big(\theta_l(t) - m_{\text{inf}}\big)\zeta^-(t)\,\mathrm{d}\omega& \\
- \int_\Omega \zeta^-(t)\big(\sigma(\theta)(t)\rho\nabla\varphi(t)\big)\cdot\nabla\varphi(t)\,\mathrm{dx}.& \quad (4.17)
\end{aligned}$$

Let us show that $\frac{\mathrm{d}}{\mathrm{d}t}\|\zeta^-(t)\|^2_{\mathrm{L}^2} \leq 0$. By the assumptions on the coefficient functions,

$$\big(\eta(\theta)(t)\kappa\nabla\zeta^-(t)\big)\cdot\nabla\zeta^-(t) \geq \eta_\bullet\kappa_\bullet\|\nabla\zeta^-(t)\|^2$$

and

$$-\big(\sigma(\theta)(t)\rho\nabla\varphi(t)\big)\cdot\nabla\varphi(t) \leq -\sigma_\bullet\rho_\bullet\|\nabla\varphi(t)\|^2.$$

This means that both integrals on the left-hand side in (4.17) are positive (since $\alpha \geq 0$), while the second term on the right-hand side is negative. The constant m_{inf} is constructed exactly such that $\theta_l(t) - m_{\text{inf}}$ is greater or equal than zero almost everywhere, such that $-\alpha(\theta_l(t) - m_{\text{inf}})\zeta^-(t) \leq 0$. Hence, from (4.17) it follows that $\frac{\mathrm{d}}{\mathrm{d}t}\|\zeta^-(t)\|^2_{\mathrm{L}^2(\Omega)} \leq 0$. But, due to the construction of ζ, we have $\zeta(T_0) \geq 0$, which means that $\zeta^-(T_0) \equiv 0$ and thus $\zeta^-(t) \equiv 0$ for all $T_0 < t < T_\bullet$.

The assertion for θ_{u_0} follows completely analogously without the quadratic gradient term. One uses the assumption $\max(\max_{\overline{\Omega}} \theta_0, \operatorname{ess\,sup}_{\Sigma_{\partial\Omega}} \theta_l) \leq \theta_{\max}$ on $\overline{Q}$ to show that $\theta_{u_0} \leq \theta_{\max}$. $\square$

Corollary 4.2.4. *The feasible sets $\mathcal{M}^{\text{ad}}, \mathcal{M}^{\text{ad}}_s$ and $\mathcal{M}^{\text{ad}}_c$ are nonempty.*

Proof. This follows from Proposition 4.2.2 and Lemma 4.2.3. $\square$

Finally, we have everything together to conclude that (P) indeed fits in the framework developed in Chapter 3. We lastly recall the control-to-state operators

$$\mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega)) \supset U_g \ni u \mapsto \mathcal{S}(u) = \theta_u \in \mathcal{Y}_{r,q}$$

and

$$\mathcal{U} \supset \mathcal{U}_g \ni u \mapsto \mathsf{S}(u) = \theta_u \in \mathcal{Y}_{r,q},$$

and the reduced optimal control problem:

$$\begin{aligned} \min_{u \in \mathcal{U}_g} \quad & \mathsf{j}(u) := \mathsf{J}_s(\theta_u) + \mathsf{J}_c(u) \\ \text{s.t.} \quad & \begin{cases} 0 \le u \le u_{\max} & \text{on } J \times N \\ \theta_u \le \theta_{\max} & \text{on } \overline{Q}. \end{cases} \end{aligned} \tag{T-rOC}$$

4.2.1 Existence of globally optimal controls

Let us turn to the question of existence of an optimal control of (P). We have already seen in Chapter 3.1 that passing to the limit starting from weak convergence in the controls, as customary in the standard direct method in the calculus of variations proof for existence of globally optimal controls, may prove very difficult. The objective functional J_s already includes a nonstandard term for $\nabla\theta$ with a curious integrability order which we have already adopted for the control space $\mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega))$. We will prove below that exactly this term, together with Lemma 4.2.3 and the state constraints imposed on the system, gives us the needed *a priori* estimates for the solutions associated to a minimizing sequence. This only applies to $d = 3$, whereas we can argue more directly in case of $d = 2$, as already indicated by having set $\gamma = 0$ in this case.

The idea for space dimension $d = 3$ is to use Theorem 1.5.18 to show that solutions θ_u associated to controls from bounded sets in $\mathrm{L}^{\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega))$

are uniformly bounded in $\mathrm{W}^{1,\mathfrak{s}}(J;\mathrm{W}_\emptyset^{-1,\varsigma}(\Omega))$, where in general only $\varsigma \sim \frac{3}{2} \ll d = 3$, cf. the assumptions on $\mathfrak{r}$ from the beginning of Chapter 4.2. Together with the state constraints posed in (T-rOC) and the lower pointwise bounds inherent in the problem (see Lemma 4.2.3), this gives an additional bound in $\mathrm{L}^{\mathfrak{s}}(J;\mathrm{W}^{1,q}(\Omega))$ in the setting of the proof of existence of globally optimal controls for (T-rOC). Then we can employ Lemma 1.4.4 to "lift" this boundedness result to a Hölder space, which is suitable for passing to the limit in the nonlinear state system with a minimizing sequence. This idea also explains the increment in $\mathfrak{s}$ compared to s from the foregoing Chapter 4.1.

We recall that we had set $\varsigma := \frac{d\mathfrak{q}}{2d-\mathfrak{q}}$, for some $2 < \mathfrak{q} < q_0$, where q_0 is the number from Gröger's Theorem 1.5.18 corresponding to the operators $-\nabla \cdot \sigma(\theta)(t)\rho\nabla$. Moreover, we had assumed $\mathfrak{r}$ to satisfy $\mathfrak{r} > 2(1-\frac{d}{q})^{-1}(1+\frac{d}{\varsigma}-\frac{d}{q})$, and $\mathfrak{s} \geq \mathfrak{r}$. Using these properties, we have the following fundamental result:

Proposition 4.2.5 (Closedness properties of U_g)**.**

(i) Consider a sequence $U_g \supset (u_n)$ such that $u_n \to \bar{u}$ with $\bar{u} \in \mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{-1,q}(\Omega))$. If the associated sequence of solutions (θ_{u_n}) admits a subsequence which converges to some $\bar{\theta}$ in $\mathrm{C}(\overline{Q})$, then $\bar{u} \in U_g$ and $\bar{\theta} = \theta_{\bar{u}}$.

(ii) Let $U_g^b \subseteq U_g$ be bounded in $\mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{-1,q}(\Omega))$ and suppose that the associated set of solutions $\mathcal{S}(U_g^b) := \{\theta_u : u \in U_g^b\}$ is bounded in $\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}^{1,q}(\Omega))$. Then $\mathcal{S}(U_g^b)$ is even relatively compact *in $\mathrm{C}(\overline{Q})$ and the closure of U_g^b in $\mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{-1,q}(\Omega))$ is still contained in U_g.*

Proof.

(i) For the first assertion, consider the sequence (u_n) from the assumptions with the associated states $(\theta_n) := (\theta_{u_n})$. By assumption, there exists a subsequence of (θ_n), called (θ_ℓ), which converges to some $\tilde{\theta}$ in $\mathrm{C}(\overline{Q})$. Lemma 4.1.9 shows that $\Psi_{u_\ell}(\theta_\ell) \to \Psi_{\bar{u}}(\bar{\theta})$ in $\mathrm{L}^{\mathfrak{s}}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$. Now, we apply the trick already used quite similarly in the proof of Theorem 4.1.4: Using Lemma 2.2.13, we find

$$\tilde{\theta} \longleftarrow \theta_l = \left(\partial + \mathcal{A}_\kappa(\eta(\theta_\ell)) + \mathcal{B}_\alpha, \delta_{T_0}\right)^{-1}\left(F(\theta_\ell, u_\ell), \theta_0\right)$$
$$\longrightarrow \left(\partial + \mathcal{A}_\kappa(\eta(\tilde{\theta})) + \mathcal{B}_\alpha, \delta_{T_0}\right)^{-1}\left(F(\tilde{\theta}, \bar{u}), \theta_0\right),$$

where we have interpreted $\alpha\theta_l \in \mathrm{L}^\infty(J; \mathrm{L}^\infty(\partial\Omega))$ as an element of $\mathrm{L}^{\mathfrak{r}}(J; \mathrm{W}^{-1,q}_\emptyset(\Omega))$. From this relation we infer that $\zeta = \tilde{\theta} \in \mathcal{Y}_{\mathfrak{r},q}$ is the solution to

$$(\partial + \mathcal{A}_\kappa(\eta(\tilde{\theta})) + \mathcal{B}_\alpha)\zeta = F(\tilde{\theta}, \bar{u}) \quad \text{in } \mathrm{W}^{-1,q}_\emptyset(\Omega), \quad \zeta(T_0) = \theta_0,$$

which by uniqueness of solutions as in Theorem 4.1.4 via Theorem 2.2.10 must coincide with $\theta_{\bar{u}}$. In particular, $\theta_{\bar{u}} \in \mathcal{Y}_{\mathfrak{r},q}$ and $\bar{u} \in U_g$.

(ii) We show that $\mathcal{S}(U_g^b)$ is bounded in the maximal-regularity space $\mathbb{W}^{1,\mathfrak{r}}(J; \mathrm{W}^{-1,\varsigma}_\emptyset(\Omega), \mathrm{W}^{1,q}(\Omega))$. To this end, we first investigate the right-hand side in the parabolic equation (4.9) and show that the time derivatives θ_u' are bounded in $\mathrm{W}^{1,\mathfrak{r}}(J; \mathrm{W}^{-1,\varsigma}_\emptyset(\Omega))$ over U_g^b.
Denote by (θ_u, φ_u) the solution for a given $u \in U_g^b$. Choosing a number $\mathfrak{q}$ such that the assertions in the theorem of Gröger are satisfied, we know that $\mathcal{A}_\rho(\sigma(\theta))(t)$ is a topological isomorphism between $\mathrm{W}^{1,\mathfrak{q}}_D(\Omega)$ and $\mathrm{W}^{-1,\mathfrak{q}}_D(\Omega)$ for every $t \in \overline{J}$ with even

$$\sup_{\theta \in \mathcal{S}(U_g^b)} \left\| \mathcal{A}_\rho(\sigma(\theta))^{-1} \right\|_{\mathrm{L}^\infty(J; \mathscr{L}(\mathrm{W}^{-1,\mathfrak{q}}_D(\Omega); \mathrm{W}^{1,\mathfrak{q}}_D(\Omega)))} < \infty.$$

In fact, the supremum is even finite over all $\theta \in \mathrm{C}(\overline{J}; \mathrm{C}(\overline{\Omega}))$. Anyway, for every $u \in U_g^b$, the function $\varphi_u = \varphi(\theta_u, u)$ is given by

$$\varphi_u(t) = \mathcal{A}_\rho(\sigma(\theta_u))^{-1} u(t) \quad \text{in } \mathrm{W}^{1,q}_D(\Omega)$$

for almost every $T_0 < t < T_1$. But this means that φ_u is bounded with respect to $u \in U_g^b$ when considered as $\mathrm{W}_D^{1,\mathfrak{q}}(\Omega)$-valued, that is,

$$\sup_{u \in U_g^b} \|\varphi_u\|_{\mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{1,\mathfrak{q}})} < \infty.$$

Arguing as in Proposition 4.1.11, we find that also

$$\sup_{u \in U_g^b} \big\|(\sigma(\theta_u)\rho\nabla\varphi_u)\cdot\nabla\varphi_u\big\|_{\mathrm{L}^{\mathfrak{s}}(J;\mathrm{L}^{\mathfrak{q}/2})(\Omega)} < \infty.$$

Using the boundedness assumption on $\mathcal{S}(U_g^b)$ in $\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}^{1,q}(\Omega))$, both the family of functionals $\mathcal{B}_\alpha\theta_u$ and, here also employing boundedness of η, the divergence-operators $\mathcal{A}_\kappa(\eta(\theta_u))\theta_u$ are uniformly bounded in $\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$ over U_g^b, i.e.,

$$\sup_{u \in U_g^b}\Big(\big\|\mathcal{A}_\kappa(\eta(\theta_u))\big\|_{\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))} + \|\mathcal{B}_\alpha\theta_u\|_{\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))}\Big) < \infty.$$

The Sobolev embedding Theorem 1.2.27 gives $\mathrm{L}^{\mathfrak{q}/2}(\Omega) \hookrightarrow \mathrm{W}_\emptyset^{-1,\varsigma}(\Omega)$ for $\varsigma = \frac{d\mathfrak{q}}{2d-\mathfrak{q}}$, and certainly $\mathrm{W}_\emptyset^{-1,q}(\Omega) \hookrightarrow \mathrm{W}_\emptyset^{-1,\varsigma}(\Omega)$ due to $q > \varsigma$. Hence,

$$\theta_u' = -\mathcal{A}_\kappa(\eta(\theta_u)) - \mathcal{B}_\alpha\theta_u + \big(\sigma(\theta_u)\rho\nabla\varphi_u\big)\cdot\nabla\varphi_u + \mathcal{B}_\alpha\theta_l$$

is *uniformly bounded* over U_g^b in $\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}_\emptyset^{-1,\varsigma}(\Omega))$.

Together with the boundedness assumption on $\mathcal{S}(U_g^b)$ in $\mathrm{L}^{\mathfrak{r}}(J;\mathrm{W}^{1,q}(\Omega))$, this shows that $\mathcal{S}(U_g^b)$ is bounded in the space $\mathbb{W}^{1,\mathfrak{r}}(J;\mathrm{W}_\emptyset^{-1,\varsigma}(\Omega),\mathrm{W}^{1,q}(\Omega))$. Due to the choice of $\mathfrak{r}$, the embedding result in Lemma 1.5.25 transfers this boundedness to a Hölder space $\mathrm{C}^{\varrho}(Q)$ and thus to (relative) compactness in $\mathrm{C}(\overline{Q})$. This was the first claim.

Now let $\bar{u} \in \overline{U_g^b}$ be given and consider a sequence $(u_n) \subset U_g^b$ converging to $\bar{u}$ in $\mathrm{L}^{2s}(J;\mathrm{W}_D^{-1,q}(\Omega))$. By relative compactness of $\mathcal{S}(U_g^b)$, the sequence of associated solutions (θ_{u_n}) admits a subsequence which converges in $\mathrm{C}(\overline{Q})$. But then (i) shows that $\bar{u} \in U_g$, hence $\overline{U_g^b} \subseteq U_g$.

This was the claim. $\square$

Remark 4.2.6. The proof of the above theorem explains the choice of the modified integrability order $\mathfrak{r}$, with $\mathfrak{s}$ accordingly, compared to Chapter 4.1. In order to have the maximal regularity space incorporating the much weaker space $\mathrm{W}_\emptyset^{-1,\varsigma}(\Omega)$ still embed into a Hölder space, we need to move the resulting interpolation spaces more towards $\mathrm{W}^{1,q}(\Omega)$ compared to before. This requires a higher integrability in time, cf. the embedding Lemma 1.4.4. In principle, we could have kept the "original" orders r, s from Chapter 4.1 and just have required the setting in Proposition 4.2.5 on top of that. But since the control space $\mathcal{U}$ admits arbitrary integrability in time (see Proposition 4.2.1) and the integrability of the control u more or less directly transfers to θ_u, we have decided to adjust the integrability index for the analytic control space $\mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{-1,q}(\Omega))$, too. In view of the initial value regularity, see Remark 4.1.1.

Proposition 4.2.5 now immediately allows to validate the assumptions in the abstract existence of globally optimal controls result in Theorem 3.1.1 via Lemma 3.1.2 and Proposition 4.2.1.

Theorem 4.2.7 (Existence of optimal controls)**.** *There exists a globally optimal solution $\bar{u} \in \mathcal{M}_c^{\mathrm{ad}}$ to the reduced optimal control problem (T-rOC).*

Proof. We already have validated the assumptions on the data posed in Chapter 3 in the set-up in Chapter 4.2, in particular $\mathcal{M}_c^{\mathrm{ad}}$ being nonempty. This means there exists a number $M > 0$ such that $\mathcal{N} := \mathsf{J}^{-1}[[0, M)] \cap \mathcal{M}_c^{\mathrm{ad}}$ is nonempty. Let $(u_n) \subset \mathcal{N} \subseteq \mathcal{U}$ be a sequence such that $(\mathcal{B}u_n)$ converges in $\mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}_D^{-1,q}(\Omega))$ to some $\bar{w}$ in the closure $\overline{\mathcal{B}\mathcal{N}}$ in that space and consider the associated states $\theta_n := \theta_{u_n}$. We distinguish between dimensions:

(i) For $d = 2$: We have seen from Proposition 4.1.11 that $\bar{w} \in U_g$ and that $(F(\theta_n, \mathcal{B}u_n))$ is uniformly bounded in $\mathrm{L}^r(J;\mathrm{W}_\emptyset^{-1,q}(\Omega))$, because the sequence $(\mathcal{B}u_n)$ is clearly bounded. Via Corollary 2.2.15, this boundedness also implies that θ_n admits a subsequence which converges in $C(\overline{Q})$, and Proposition 4.2.5 (i) then shows that θ_n converges to $\theta_{\bar{w}}$.

(ii) For $d = 3$: The states (θ_n) are bounded in a pointwise sense in $\overline{Q}$ due to the state constraints and the lower bounds inherent in the system, cf. Lemma 4.2.3. Moreover, the objective functional value $\mathsf{j}(u_n)$ is bounded by M, hence the gradients $(\nabla\theta_n)$ are bounded in $L^{\mathfrak{r}}(J; L^q(\Omega; \mathbb{R}^d))$. But this means that (θ_n) is uniformly bounded in $L^{\mathfrak{r}}(J; W^{1,q}(\Omega))$, such that Proposition 4.2.5 shows that $\bar{w} \in U_g$ and that $\theta_n \to \theta_{\bar{w}}$.

From $\bar{w} \in U_g$ we infer that the closure $\overline{\mathcal{B}\mathcal{N}}$ in $L^{2\mathfrak{s}}(J; W_D^{-1,q}(\Omega))$ is still contained in U_g. Moreover, Lemma 4.1.9 shows that $F(\theta_n, \mathcal{B}u_n) \to F(\theta_{\bar{u}}, \bar{w})$ in $L^{\mathfrak{s}}(J; W_\emptyset^{-1,q}(\Omega))$, which means that $\mathfrak{u} \mapsto F(\mathcal{S}(\mathfrak{u}), \mathfrak{u})$ is continuous on the closure $\overline{\mathcal{B}\mathcal{N}}$ in $L^{2\mathfrak{s}}(J; W_D^{-1,q}(\Omega))$. Together with $\mathsf{E} = \mathcal{B} \colon \mathcal{U} \to \mathcal{X}(J; U) = L^{2\mathfrak{s}}(J; W_D^{-1,q}(\Omega))$ being compact by Proposition 4.2.1, these are the assumptions in Lemma 3.1.2 which imply those of Theorem 3.1.1. □

4.2.2 First order necessary optimality conditions

We next treat first order necessary optimality conditions for (T-rOC). This of course requires differentiability of involved functions. However, we have already seen that J_s and J_c are continuously differentiable, cf. (4.15) and (4.16), such that the following assumption is rather short, cf. also the comments on the corresponding assumption in Chapter 3.2.

The following assumption holds true for the rest of this chapter:

- The coefficient functions η and σ are continuously differentiable as mappings on $C(\overline{J}; C(\overline{\Omega}))$ and we can identify their derivatives $\eta'(\theta)$ and $\sigma'(\theta)$ in a point $\theta \in \mathcal{Y}_{\mathfrak{r},q}$ with a function in $C(\overline{J}; C(\overline{\Omega}))$ itself.

Thanks to the preparations in Chapter 3.2, we only need to show that the right-hand side mapping $(\theta, u) \mapsto F(\theta, \mathcal{B}u)$ is continuously differentiable from $\mathcal{Y}_{\mathfrak{r},q} \times \mathcal{U}$ to $\mathcal{Z}_{\mathfrak{r},q}$ and that $\partial_\theta F(\theta, \mathcal{B}u)$ gives rise to a suitable perturbation of nonautonomous maximal parabolic regularity on $\mathrm{W}^{-1,q}_\emptyset(\Omega)$. We thereby also prove openness of $\mathcal{U}_g$ and continuous differentiability of the control-to-state operator.

Theorem 4.2.8. *The right-hand side mapping $F(\theta, u) = \Psi_u(\theta) + \mathcal{B}_\alpha \theta_l$ is continuously differentiable as a mapping from $\mathcal{Y}_{\mathfrak{r},q} \times \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$ to $\mathrm{L}^{\mathfrak{r}}(J; \mathrm{W}^{-1,q}_\emptyset(\Omega))$ with*

$$\partial_\theta F(\theta_u, u) \in \mathrm{L}^{\mathfrak{r}}\Big(J; \mathscr{L}\big((\mathrm{W}^{-1,q}_\emptyset(\Omega), \mathrm{W}^{1,q}(\Omega))_{1/\mathfrak{r}',\mathfrak{r}}); \mathrm{W}^{-1,q}_\emptyset(\Omega)\big)\Big)$$

for every $u \in U_g$. In particular, U_g is open, the control-to-state operator $\mathcal{S}$ is continuously differentiable on U_g, and its directional derivative $\zeta := \mathcal{S}(u)h \in \mathcal{Y}_{\mathfrak{r},q}$ for $u \in U_g$ is the unique solution of the equation

$$\begin{aligned} \zeta' + (\mathcal{A}_\kappa(\eta(\theta_u)) + \mathcal{B}_\alpha)\zeta &= (\sigma'(\theta_u)\zeta\rho\nabla\varphi_u) \cdot \nabla\varphi_u - \mathcal{A}_\kappa(\eta'(\theta_u)\zeta)\theta_u \\ &\quad - 2(\sigma(\theta_u)\rho\nabla\varphi_u) \cdot \nabla\Big[\mathcal{A}_\rho(\sigma(\theta_u))^{-1}\big(\mathcal{A}_\rho(\sigma'(\theta_u)\zeta)\varphi_u + h\big)\Big] \end{aligned}$$

with $\zeta(T_0) = 0$, for all $h \in \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$.

Proof. The term $\mathcal{B}_\alpha\theta_l$ in F depends neither on u nor on θ and is thus neglected for the rest of this proof. Let us first consider the partial derivative of F with respect to the control u. Let $u \in \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$. Recalling the quadratic form $\mathfrak{b}$ from Lemma 4.1.6, we observe that

$$F(\theta, u) = \Psi_u(\theta) = \mathfrak{b}_\theta(\varphi(\theta, u)) = \mathfrak{b}_\theta\big(\mathcal{A}_\rho(\sigma(\theta))^{-1}u\big),$$

which due to the continuous linear dependence of φ on u gives rise to a continuous quadratic form on $\mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$. Continuous quadratic forms, however, are continuously differentiable, and we obtain the directional derivative

$$\partial_u F(\theta, u)h = -2\mathfrak{b}_\theta\big(\mathcal{A}_\rho(\sigma(\theta))^{-1}u, \mathcal{A}_\rho(\sigma(\theta))^{-1}h\big)$$

for $h \in \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega))$.

Now let us turn to $\partial_\theta F(\theta, u)$, for which we need to consider

$$\mathcal{Y}_{\mathfrak{r},q} \ni \theta \mapsto \Psi_u(\theta) = \mathfrak{b}_\theta(\mathcal{A}_\rho(\sigma(\theta))^{-1}u).$$

Recalling that the derivative of the (continuously differentiable) inversion mapping $\mathscr{L}(X; Y) \ni A \mapsto A^{-1} \in \mathscr{L}(Y; X)$ in A is given by $H \mapsto -A^{-1}HA^{-1}$, we need to consider the derivative of $\theta \mapsto \mathcal{A}_\rho(\sigma(\theta))$. As seen in Theorem 3.2.4, its directional derivative is given by $h \mapsto \mathcal{A}_\rho(\sigma'(\theta)h)$ for $h \in \mathcal{Y}_{\mathfrak{r},q}$. But then the mentioned formula, the form of $\mathfrak{b}_\theta$, and the observation that $\mathfrak{b}_\theta$ is a continuous quadratic form shows that $\theta \mapsto F(\theta, u)$ is indeed continuously differentiable as a mapping $\mathcal{Y}_{\mathfrak{r},q}$ to $\mathcal{Z}_{\mathfrak{r},q}$ with the directional derivative

$$\begin{aligned}\partial_\theta F(\theta, u)h = -2\mathfrak{b}_\theta\big(\varphi(\theta, u), \mathcal{A}_\rho(\sigma(\theta))^{-1}\mathcal{A}_\rho(\sigma'(\theta)h)\varphi(\theta, u)\big)& \\ + \big(\sigma'(\theta)h\rho\nabla\varphi(\theta, u)\big) \cdot \nabla\varphi(\theta, u)&\end{aligned} \tag{4.18}$$

for every $h \in \mathcal{Y}_{\mathfrak{r},q}$.

Now let $u \in U_g$ be a global control. We can identify $\partial_\theta F(\theta_u, u)$ with a function in time by writing the time-space functions in their pointwise form, where we have to use the extra assumption on the derivative $\sigma'(\theta)$:

$$\begin{aligned}\partial_\theta F(\theta_u, u)(t)\xi = -2\mathfrak{b}_{\theta_u}\big(\varphi_u(t), \mathcal{A}_\rho(\sigma(\theta_u))(t)^{-1}\big(-\nabla\cdot\sigma'(\theta_u)(t)\xi\rho\nabla\varphi_u(t)\big)\big)& \\ + \big(\sigma'(\theta_u)(t)\xi\rho\nabla\varphi_u(t)\big) \cdot \nabla\varphi_u(t).&\end{aligned}$$

For every $t \in J$, this gives rise to a continuous linear operator mapping $\mathrm{C}(\overline{\Omega})$ and thus also $Y_{\mathfrak{r},q}$ to $\mathrm{W}_D^{-1,q}(\Omega)$, and rigorously counting integrability orders shows that the mapping is also indeed $\mathfrak{r}$-integrable on J. This completes the assertions on F.[8]

The remaining assertions now follow immediately from Theorem 3.2.4. Inserting all formulas and the actual expression for $\mathfrak{b}_\theta$, we also obtain the

[8] In fact, $\partial_\theta F(\theta, u)$ satisfies the assumptions for Theorem 3.2.4 for *every* pair $(\theta, u) \in \mathcal{Y}_{\mathfrak{r},q} \times \mathcal{X}(J; U)$.

equation stated in the theorem, cf. also the abstract linearized equation in Corollary 3.2.6. □

Remark 4.2.9. From the slightly involved considerations in the foregoing theorem, one could guess that one pays the price the reduction to one single abstract equation for θ here. However, this way we were able to rather elegantly apply perturbation theory for nonautonomous maximal parabolic regularity, and furthermore, there is no loss in generality, since one may split the equation solved by $\zeta = \mathcal{S}'(u)h \in \mathcal{Y}_{\mathfrak{r},q}$ in the previous Theorem 4.2.8 back into two equations: Introducing

$$\pi = \Phi(\zeta) := \mathcal{A}_\rho(\sigma(\theta_u))^{-1}\left(\mathcal{A}_\rho(\sigma'(\theta_u)\zeta)\varphi_u + h\right) \in \mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}^{1,q}_D(\Omega)),$$

we find that for every $h \in \mathrm{L}^{2r}(J;\mathrm{W}^{-1,q}_D(\Omega))$ the pair (ζ,π) is the *unique solution* of the system

$$\begin{aligned}\partial\zeta + \mathcal{A}_\kappa(\eta(\theta_u))\zeta &= \left(\sigma'(\theta_u)\zeta\rho\nabla\varphi_u\right)\cdot\nabla\varphi_u - \mathcal{A}_\kappa(\eta'(\theta_u)\zeta)\theta_u \\ &\qquad + 2\left(\sigma(\theta_u)\rho\nabla\varphi_u\right)\cdot\nabla\pi \\ \mathcal{A}_\rho(\sigma(\theta_u))\pi &= \mathcal{A}_\rho(\sigma'(\theta_u)\zeta)\varphi_u + h\end{aligned}$$

with $\zeta(T_0) = 0$. These equations are exactly the *linearized state system* for (4.9) and (4.10). This also shows, not unexpectedly, that from a functional-analytical point of view it makes no difference working with θ only and considering φ as a function obtained by θ instead of considering both functions at once.

We can proceed in giant steps towards first order necessary optimality conditions. To make the abstract results from Chapter 3.2 a bit more graspable, we spell out the abstract adjoint equation for the thermistor system. Therefor, we need to understand how $\left[\partial_\theta F(\theta_u,u)\right]^*$ acts on a function $\vartheta \in \mathcal{Z}_{\mathfrak{r},q}$, cf. Definition 3.2.12. We had

$$\begin{aligned}\partial_\theta F(\theta_u,u)h = -2\mathfrak{b}_{\theta_u}(\varphi_u,\mathcal{A}_\rho(\sigma(\theta_u))^{-1}&\mathcal{A}_\rho(\sigma'(\theta_u)h)\varphi_u) \\ &+ \left(\sigma'(\theta_u)h\rho\nabla\varphi_u\right)\cdot\nabla\varphi_u. \quad (4.18)\end{aligned}$$

It is easy to see that the second addend is formally self-adjoint since it is a multiplication operator for the function h, that is,

$$\big\langle (\sigma'(\theta_u) h \rho \nabla\varphi_u)\cdot\nabla\varphi_u, \vartheta\big\rangle_{\mathcal{Z}_{\mathfrak{r},q}} = \big\langle (\sigma'(\theta_u)\vartheta\rho\nabla\varphi_u)\cdot\nabla\varphi_u, h\big\rangle_{\mathcal{Y}_{r,q}},$$

where $(\sigma'(\theta_u)\vartheta\rho\nabla\varphi_u)\cdot\nabla\varphi_u$ is now merely an $\mathrm{L}^1(J;\mathrm{L}^1(\Omega))$-function, cf. also the discussion in Remark 3.2.15. Let us turn to the less pleasant term in (4.18). For the moment, we introduce

$$\phi_h \coloneqq \mathcal{A}_\rho(\sigma(\theta_u))^{-1}\mathcal{A}_\rho(\sigma'(\theta_u)h)\varphi_u \in \mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}^{1,q}_D(\Omega)),$$

similarly to π from Remark 4.2.9. Writing the actions of the operators explicitly as integrals,[9] we then find

$$\begin{aligned}\big\langle \mathfrak{b}_{\theta_u}(\varphi_u,\phi_h),\vartheta\big\rangle &\coloneqq \big\langle (\sigma(\theta_u)\rho\nabla\varphi_u)\cdot\nabla\phi_h,\vartheta\big\rangle\\ &= \big\langle \mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u,\phi_h\big\rangle.\end{aligned} \tag{4.19}$$

The last dual product is formally written and means

$$\phi_h \mapsto \int_J\int_\Omega (\sigma(\theta_u)\vartheta\rho\nabla\varphi_u)\nabla\phi_h \,\mathrm{d}x\,\mathrm{d}t, \tag{4.20}$$

from which we identify $\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u \in \mathrm{L}^{(2\mathfrak{s})'}(J;\mathrm{W}^{-1,q'}_D(\Omega))$. Expanding ϕ_h again, we obtain

$$\big\langle \mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u,\phi_h\big\rangle = \big\langle \mathcal{A}_\rho(\sigma(\theta_u))^{-*}\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u, \mathcal{A}_\rho(\sigma'(\theta_u)h)\varphi_u\big\rangle.$$

Now h is again only a multiplier in $\mathcal{A}_\rho(\sigma'(\theta_u)h)\varphi_u \in \mathrm{L}^{2\mathfrak{s}}(J;\mathrm{W}^{-1,q'}_D(\Omega))$, from which we finally find

$$\begin{aligned}&\big\langle -2\mathfrak{b}_{\theta_u}(\varphi_u,\mathcal{A}_\rho(\sigma(\theta_u))^{-1}\mathcal{A}_\rho(\sigma'(\theta_u)h)\varphi_u),\vartheta\big\rangle\\ &\qquad = \big\langle -(\sigma'(\theta_u)\rho\nabla\varphi_u)\cdot\nabla\left[\mathcal{A}_\rho(\sigma(\theta_u))^{-*}(2\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u)\right], h\big\rangle\end{aligned}$$

[9] The slightly frustrated author would be very grateful for suggestions on how to handle this computation more elegantly.

Now we can formulate the abstract adjoint equation, cf. also Definition 3.2.12. In contrast to Definition 3.2.12, we consider the system only in a pair (θ_u, u) of solution and associated control to save ourselves from writing $\varphi(\theta, u)$ instead of φ_u each time. Note however that the abstract adjoint equation here could be handled for every pair $(\theta, u) \in \mathcal{Y}_{\mathfrak{r},q} \times \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}_D^{-1,q}(\Omega))$, including Theorem 4.2.12, cf. the footnote in the proof of Theorem 4.2.8.

Definition 4.2.10 (Abstract adjoint equation)**.** Let $u \in U_g$ be fixed and let moreover the functions $f \in \mathcal{Y}'_{\mathfrak{r},q}$, $\vartheta_T \in Y'_{\mathfrak{r},q}$ and $g \in \mathrm{L}^{(2\mathfrak{s})'}(J; \mathrm{W}_D^{-1,q'}(\Omega))$ be given. Then the following equation for $(\vartheta, \chi) \in \mathcal{Z}'_{\mathfrak{r},q} \times Y'_{\mathfrak{r},q}$ in $\mathcal{Y}'_{\mathfrak{r},q}$ is called the *abstract adjoint equation* for the thermistor system:

$$\begin{aligned} &- \partial\vartheta + \mathcal{A}_{\kappa^\top}(\eta(\theta_u))\vartheta + \mathcal{B}_\alpha\vartheta \\ &= -(\eta(\theta_u)\kappa\nabla\theta) \cdot \nabla\vartheta + \delta_{T_1}^* \otimes \vartheta_T - \delta_{T_0}^* \otimes \chi + f + (\sigma'(\theta_u)\vartheta\rho\nabla\varphi_u) \cdot \nabla\varphi_u \\ &\qquad - (\sigma'(\theta_u)\rho\nabla\varphi_u) \cdot \nabla \left[\mathcal{A}_\rho(\sigma(\theta_u))^{-*}(2\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u + g)\right], \end{aligned} \tag{4.21}$$

where $\delta_{T_0}^*$ and $\delta_{T_1}^*$ are the adjoint operators to the linear continuous point evaluations on $\mathrm{C}(\overline{J}; Y_{r,q})$. Moreover, we have identified $\left[\mathcal{A}_\kappa(\eta(\theta_u))\right]^*\vartheta$ and $\mathcal{B}_\alpha^*\vartheta$ with the formal but intuitive expressions $\mathcal{A}_{\kappa^\top}(\eta(\theta_u))\vartheta$ and $\mathcal{B}_\alpha\vartheta$, and

$$\begin{aligned} &\Big[\zeta \mapsto \Big\langle (\eta'(\theta_u)\kappa\nabla y)\nabla\vartheta, \zeta \Big\rangle \\ &\qquad := \int_J \int_\Omega (\eta'(\theta_u)\zeta\kappa\nabla y)\nabla\vartheta \,\mathrm{d}x\,\mathrm{d}t\Big] \in \mathrm{L}^\infty(J; \mathrm{L}^\infty(\Omega))' \end{aligned}$$

on the right-hand side. For the interpretation of the remaining terms, see directly above.

Remark 4.2.11. Similarly to Remark 4.2.9, we introduce

$$\psi(\vartheta) := \mathcal{A}_\rho(\sigma(\theta_u))^{-*}(2\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u),$$

which allows to split (4.21) back into two equations, namely

$$\begin{aligned}-\partial\vartheta + \mathcal{A}_{\kappa^\top}(\eta(\theta_u))\vartheta + \mathcal{B}_\alpha\vartheta = -(\eta(\theta_u)\kappa\nabla\theta_u)\cdot\nabla\vartheta + f\\ + (\sigma'(\theta_u)\vartheta\rho\nabla\varphi_u)\cdot\nabla\varphi_u\\ - (\sigma'(\theta_u)\rho\nabla\varphi_u)\cdot\nabla\psi)\\ + \delta^*_{T_1}\otimes\vartheta_T - \delta^*_{T_0}\otimes\chi\\ \mathcal{A}_\rho(\sigma(\theta_u))^*\psi = 2\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u + g.\end{aligned}$$

This is exactly the adjoint system corresponding to the linearized state system encountered in Remark 4.2.9.

Theorem 4.2.12. *For every $u \in U_g$ and all*

$$(f, g, \vartheta_T) \in \mathcal{Y}'_{\mathfrak{r},q} \times \mathrm{L}^{(2\mathfrak{s})'}(J; \mathrm{W}^{-1,q'}_D(\Omega)) \times Y'_{\mathfrak{r},q},$$

the abstract adjoint equation admits a unique solution $(\vartheta, \chi) \in \mathcal{Z}'_{\mathfrak{r},q} \times Y'_{\mathfrak{r},q}$.

Proof. This follows from setting

$$\mathfrak{f} \coloneqq f - (\sigma'(\theta_u)\rho\nabla\varphi_u)\cdot\nabla\left[\mathcal{A}_\rho(\sigma(\theta_u))^{-*}g\right] + \delta^*_{T_1}\otimes\vartheta_T \in \mathcal{Y}_{\mathfrak{r},q} \qquad (4.22)$$

and applying Theorem 3.2.13 to the data $(\mathfrak{f}, \vartheta_T)$. ☐

As a last corollary, we characterize the action of $\mathsf{S}'(u)$ for $u \in \mathcal{U}_g$ via the abstract adjoint system. Recall that

$$\mathsf{S}'(u)^* = -\mathsf{E}^*\left[\partial_u e(\theta_u, \mathsf{E}u)\right]^*\left[\partial_y e(\theta_{\bar{u}}, \mathsf{E}u)\right]^{-*}, \qquad (4.23)$$

where E was the adjoint trace operator $\mathcal{B}\colon \mathcal{U} \to \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$, cf. Corollary 3.2.5 and Proposition 4.2.1.

Corollary 4.2.13. *Let $u \in \mathcal{U}_g$ and let $(\vartheta, \chi) \in \mathcal{Z}'_{\mathfrak{r},q} \times Y'_{\mathfrak{r},q}$ be the solution of* (4.21) *in the sense of Definition 4.2.10 with inhomogeneities f and g and terminal value ϑ_T. Then the operator $\mathsf{S}'(u)^*$ assigns to f, g and ϑ_T in the form $\mathfrak{f} \in \mathcal{Y}'_{\mathfrak{r},q}$ as in* (4.22) *the functional $\mathsf{E}^*\psi \in \mathcal{U}'$, where*

$\psi(\vartheta) \in \mathrm{L}^{(2\mathfrak{s})'}(J; \mathrm{W}^{1,q'}_D(\Omega))$ *is given by*

$$\psi(\vartheta) := \mathcal{A}_\rho(\sigma(\theta_u))^{-*}\big(2\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u\big),$$

with $\mathcal{A}_\rho(\sigma(\theta_u)\vartheta)^*\varphi_u$ *as defined by* (4.20).

Proof. In view of the form of $\mathsf{S}'(u)$ in (4.23) and Theorem 3.2.13, which tells us that $\big[\partial_y e(\theta_{\bar u}, \mathsf{E}u)\big]^{-*}$ is exactly the solution operator of the abstract adjoint equation, it only remains to verify that the action of $\big[\partial_u e(\theta_u, \mathsf{E}u)\big]^*$ on (ϑ, χ) is exactly as stated. This however follows from a repetition of the considerations starting from (4.19). $\square$

Having $\mathsf{S}'(u)^*$ at hand, we now proceed to establish the actual necessary optimality conditions in their final form. The basis is of course Theorem 3.2.16.

Theorem 4.2.14 (First Order Necessary Conditions)**.** *Let* $\bar u \in \mathcal{M}_c^{\mathrm{ad}}$ *be a locally optimal control for which the linearized Slater condition for (T-rOC) is satisfied: There exists* $\bar u \neq u^* \in \mathcal{U}^{\mathrm{ad}}$ *and* $\varepsilon > 0$ *such that*

$$\theta_{\bar u} + \mathsf{S}'(\bar u)(u^* - \bar u) \leq \theta_{\max} - \varepsilon \quad \textit{on } \overline{Q}.$$

Then there exists a Lagrangian multiplier $\bar p \in \mathcal{M}(\overline{Q})$ *such that the following* first order necessary optimality conditions *are satisfied: The* complementarity conditions

$$\theta_{\bar u} \leq \theta_{\max} \textit{ on } \overline{Q}, \qquad \bar p \geq 0, \qquad \textit{and} \quad \int_{\overline{Q}} (\theta_{\bar u} - \theta_{\max}) \,\mathrm{d}\bar p = 0, \tag{4.24}$$

together with the variational inequality

$$\int_{\Sigma_N} \bar u'(u - \bar u)' + \left(\frac{p}{2}|\bar u|^{p-2}\bar u + \frac{1}{\beta}\operatorname{tr}\psi\right)(u - \bar u)\,\mathrm{d}(\lambda \otimes \omega) \geq 0$$
$$\textit{for all } u \in \mathcal{U}^{\mathrm{ad}} = \Big\{u \in \mathcal{U} \colon 0 \leq u \leq u_{\max} \textit{ on } \Sigma_N\Big\}. \tag{4.25}$$

Here, $\psi \in \mathcal{U}'$ *is the* adjoint state *given by*

$$\psi = \mathcal{A}_\rho(\sigma(\theta_{\bar{u}}))^{-*}\big(2\mathcal{A}_\rho(\sigma(\theta_{\bar{u}})\vartheta)^*\varphi_{\bar{u}}\big)$$

such that

$$\bar{\mathsf{p}} = \mathsf{S}'(\bar{u})^*(\bar{\vartheta}, \bar{\chi}) = \mathsf{E}^*\mathcal{A}_\rho(\sigma(\theta_{\bar{u}}))^{-*}\big(2\mathcal{A}_\rho(\sigma(\theta_{\bar{u}})\vartheta)^*\varphi_{\bar{u}}\big) = \mathsf{E}^*\psi$$

for the unique solution $(\bar{\vartheta}, \bar{\chi}) \in \mathcal{Z}'_{r,q} \times Y'_{r,q}$ *of the abstract adjoint equation*

$$\begin{aligned}-\partial\vartheta + \mathcal{A}_{\kappa^\top}(\eta(\theta_{\bar{u}}))\vartheta + \mathcal{B}_\alpha\vartheta = \delta^*_{T_1} \otimes \mathbf{1}_E\big(\theta(T_1) - \theta_{\mathrm{obj}}\big) + \gamma\|\nabla\theta_u\|^{\mathfrak{r}-q}_{\mathrm{L}^q(\Omega)}\Delta_q\theta_{\bar{u}} + \bar{p}\\ - \big(\eta(\theta_{\bar{u}})\kappa\nabla\theta_{\bar{u}}\big)\cdot\nabla\vartheta - \delta^*_{T_0}\otimes\chi + \big(\sigma'(\theta_{\bar{u}})\vartheta\rho\nabla\varphi_{\bar{u}}\big)\cdot\nabla\varphi_{\bar{u}}\\ - \big(\sigma'(\theta_{\bar{u}})\rho\nabla\varphi_{\bar{u}}\big)\cdot\nabla\left[\mathcal{A}_\rho(\sigma(\theta_{\bar{u}}))^{-*}\big(2\mathcal{A}_\rho(\sigma(\theta_{\bar{u}})\vartheta)^*\varphi_{\bar{u}}\big)\right]\end{aligned}$$

in the sense of Definition 4.2.10.

Proof of Theorem 4.2.14. The assertions follow in general from Theorem 3.2.16, Corollary 4.2.13, the form of the objective functional and its derivatives, cf. (4.15) and (4.16), and the choice of $\mathcal{G}$. We have made some more precise statements owing to the specific setting we are in:

(i) Since the Lagrangian multiplier is a measure and $\mathcal{K} = K_-$ is the cone of nonpositive functions, the condition $\bar{p} \in \mathcal{K}^\circ$ means exactly that $\bar{p}$ is a positive measure in the sense of $\langle\varphi, \bar{p}\rangle \leq 0$ for all $\varphi \in K_-$. This is (4.24).

(ii) The variational inequality is obtained from inserting the derivative of J_c as in (4.16). It remains to make sense of $\langle\mathsf{p}, u - \bar{u}\rangle_{\mathcal{U}',\mathcal{U}}$. Recall that E was the adjoint trace operator $\mathcal{B}\colon \mathcal{U} \to \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega))$, which we can decompose into $\mathsf{E} = B\mathsf{E}_\bullet = \mathrm{tr}^*\,\mathsf{E}_\bullet$ with

$$\mathcal{U} \overset{\mathsf{E}_\bullet}{\hookrightarrow} \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{L}^{\mathfrak{p}}(N)) \xrightarrow{\mathrm{tr}^*} \mathrm{L}^{2\mathfrak{s}}(J; \mathrm{W}^{-1,q}_D(\Omega)),$$

for suitable $\mathfrak{p}$, see Proposition 4.2.1 and Definition 1.5.11. But then we have for $\mathsf{p} = \mathsf{E}^*\psi$ with $\psi = \mathcal{A}_\rho(\sigma(\theta_{\bar{u}}))^{-*}\big(\mathcal{A}_\rho(2\sigma(\theta_{\bar{u}})\vartheta)^*\varphi_{\bar{u}}\big) \in$

$\mathrm{L}^{(2\mathfrak{s})'}(J; \mathrm{W}^{1,q'}_D(\Omega))$,

$$\begin{aligned}\langle \mathsf{p}, u - \bar{u} \rangle_{\mathcal{U}'} &= \langle \mathsf{E}^*_\bullet B^* \psi, u - \bar{u} \rangle_{\mathcal{U}'} \\ &= \langle B^* \psi, \mathsf{E}_\bullet (u - \bar{u}) \rangle_{\mathrm{L}^{2\mathfrak{s}}(J; \mathrm{L}^{\mathfrak{p}}(N))} = \int_{\Sigma_N} \operatorname{tr} \psi (u - \bar{u}) \, \mathrm{d}(\lambda \otimes \omega),\end{aligned}$$

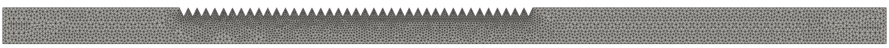
which is exactly the form in (4.25). □

Remark 4.2.15. Analogously to the technique in Remark 3.2.17, we may decompose the measure $\bar{p} \in \mathcal{M}(\overline{Q})$ by restriction into a sum of measures $\bar{p}_1, \bar{p}_2, \bar{p}_Q$ in $\mathcal{M}(\overline{Q})$ with support in $\{T_0\} \times \overline{\Omega}$, in $\{T_1\} \times \overline{\Omega}$, and in $J \times \overline{\Omega}$. Note that due to the assumption on compatibility of initial value and upper state bound, that is, $\theta_0(\mathrm{x}) < \theta_{\max}(T_0, \mathrm{x})$ for all $\mathrm{x} \in \overline{\Omega}$, and the complementarity condition (4.24), the measure $\bar{p}_1$ corresponding to T_0 cannot be active. Let us also note that this condition is, besides its very reasonable real-life meaning, also *necessary* for $\bar{p}$ to exist in the first place: If $\theta_0(\mathrm{x}) = \theta_{\max}(T_0, \mathrm{x})$ for some $\mathrm{x} \in \overline{\Omega}$, then the linearized Slater condition as in Theorem 4.2.14 *cannot* hold due to $\mathsf{S}'(\bar{u})h(T_0) = 0$ for every direction $h \in \mathcal{U}$, cf. Remark 4.2.9.

4.3 Application and numerical example

As already outlined in [88] and the introduction of this chapter, a typical example of an application for a problem in the form (P) is the optimal heating of a conducting material such as steel by means of an electric current. We present numerical calculations done in the setting of this practical application, which underline the necessity to consider the model as set up in (P). In particular, it will become clear that both control– and state constraints are necessary to obtain the correct behavior within the system.

The aim of such heating procedures is to heat up a workpiece by electric current and to cool it down rapidly with water nozzles in order to harden it. In case of steel, this treatment indeed produces a hard martensitic outer layer, see for instance [33, Ch. 9.18] for a phase diagram and [33, Chapters 10.5/10.7 about Martensite], and is thus used for instance for rack-and-pinion actuators, to be found e.g. in steering mechanisms. The part of the workpiece to be heated up corresponds to the design area E in the objective functional in (P). In order to avoid thermal stresses in the material, it is crucial to produce a homogeneous temperature distribution in the design area, which is reflected by the first term of the objective functional if we choose θ_{obj} appropriately. The gradient term in the objective functional further enforces minimal thermal stresses. Moreover, the temperatures necessary for the hardening process as described above are rather close to the melting point of the material, thus the state constraints are used to prevent the temperature exceeding the melting temperature $\theta_{\max}$. The control constraints in (P) represent a maximum electrical current which can be induced in the workpiece.



(a) Ω with underlying mesh from the side (x_1x_2-plane).

(b) Ω from above (x_1x_3-plane) with N (left) and D (right) emphasized.

Figure 4.1. The computational domain Ω used in the numerical example.

In the following we exhibit numerical examples for the optimal control of the three-dimensional thermistor problem in the form (P), underlining in particular the importance of the state-constraints. The considered computational domain Ω is a (simplified) three-dimensional gear-rack as seen in Figure 4.1 of dimensions 0.5 m $\times$ 0.02 m $\times$ 0.02 m, where the design area E consists of the saw-teeth between the 0.1 m and 0.3 m mark. The mesh consists of about 80000 nodes, inducing 400000 cells with cell diameters ranging from $8.8 \cdot 10^{-4}$ m to $7.6 \cdot 10^{-3}$ m.

The heat-equation we use in the computations is as follows:

$$\varrho C_p \partial_t \theta - \operatorname{div}(\eta(\theta)\kappa\nabla\theta) = (\sigma(\theta)\nabla\varphi) \cdot \nabla\varphi.$$

It deviates from (4.1) by the factor ϱC_p, the so-called volumetric heat capacity, where ϱ is the density of the material and C_p is its specific heat capacity. However, we assume ϱC_p to be constant for this numerical example, so it certainly has no influence on the theory presented above. We have already mentioned how to deal with non-constant heat capacity in the introduction, see also Remark 4.1.3 for a comment on a spatially inhomogeneous density ϱ. For a realistic modeling of the process, we use the data gathered in [37], i.e., the workpiece Ω is supposedly made of non-ferromagnetic stainless steel (#1.4301). The constants used can be found in Table 4.1 and the conductivity functions are given by

$$\sigma(\theta) := \frac{1}{a_\sigma + b_\sigma\theta + c_\sigma\theta^2 + d_\sigma\theta^3} \quad \text{for} \quad \theta \in [0, 10000]\ \mathrm{K},$$

with the constants $a_\sigma = 4.9659 \cdot 10^{-7}\ \Omega\mathrm{m}$, $b_\sigma = 8.4121 \cdot 10^{-10}\ \Omega\mathrm{mK}^{-1}$, $c_\sigma = -3.7246 \cdot 10^{-13}\ \Omega\mathrm{mK}^{-2}$ and $d_\sigma = 6.1960 \cdot 10^{-17}\ \Omega\mathrm{mK}^{-3}$ for the electrical conductivity (resulting in $\Omega^{-1}\mathrm{m}^{-1}$), and

$$\eta(\theta) := a_\eta + b_\eta\theta \quad \text{for} \quad \theta \in [0, 10000]\ \mathrm{K}$$

with $a_\eta = 11.215\ \mathrm{Wm}^{-1}\mathrm{K}^{-1}$ and $b_\eta = 1.4087 \cdot 10^{-4}\ \mathrm{Wm}^{-1}\mathrm{K}^{-2}$ for the thermal conductivity (resulting in $\mathrm{Wm}^{-1}\mathrm{K}^{-1}$). Both functions are extended outside of $[0, 10000]$ in a smooth and bounded way, such that the assumptions on them are satisfied. Note that ρ and κ are each chosen as the identity matrix, as we do not account for heterogeneous materials in this numerical example. To counter-act on the different scales inherent in the problem, cf. the value for $u_{\max}$ and θ_0 in Table 4.1, the model was nondimensionalized for the implementation.

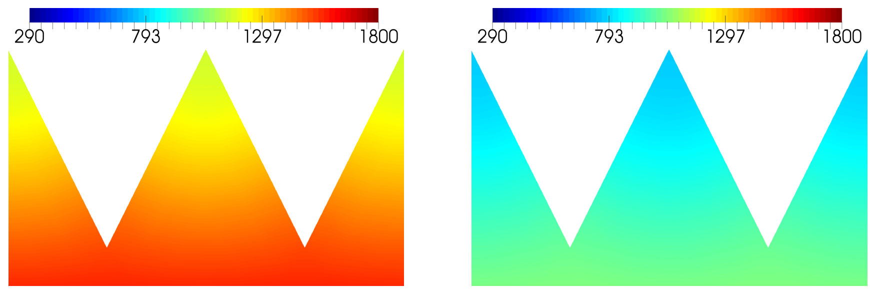
ϱ	C_p	α	θ_0	θ_l	θ_{obj}	$\theta_{\max}$	$u_{\max}$
$7900\ \frac{\mathrm{kg}}{\mathrm{m}^3}$	$455\ \frac{\mathrm{J}}{\mathrm{kg\,K}}$	$20\ \frac{\mathrm{W}}{\mathrm{m}^3\,\mathrm{K}}$	290 K	290 K	1500 K	1700 K	$10 \cdot 10^7\,\frac{\mathrm{A}}{\mathrm{m}^2}$

Table 4.1. Material parameters used in the numerical tests

The optimization problem (P) is solved by means of a Nonlinear

Conjugate-Gradients Method in the form as described in [45], modified to a projected method to account for the admissible set $\mathcal{U}_{\mathrm{ad}}$. The method needed up to 150 iterations to meet the stopping criterion, which required the relative change in the objective functional to be smaller than 10^{-5}. The state constraints in (P) are incorporated by a quadratic penalty approach—so-called Moreau-Yosida regularization—, cf. [85] and the references therein, where the penalty-parameter was increased up to a maximum of 10^{10}, stopping earlier if the violation of the state constraints was smaller than 10^{-2} K. This resulted in a violation of $9.54 \cdot 10^{-2}$ K, which is about 0.0056% of the upper bound of 1700 K. In each step of the optimization algorithm, the nonlinear state equations (4.1)-(4.6) and the adjoint equations in strong form have to be solved. We use an Implicit Euler Scheme for the time-discretization of these equations, whereas the spatial discretization is done via piecewise continuous linear finite elements. The nonlinear system of equations arising in each time-step is solved via Newton's method. Here, we do a semi-implicit pre-step to obtain a suitable initial guess for the discrete φ for Newton's method. For the control, we also choose piecewise continuous linear functions in space where the values in the first and last timestep were pre-set to 0. In the calculation of the gradient of the reduced objective functional j, the gradient representation with respect to the $\mathrm{L}^2(J;\mathrm{L}^2(N))$ scalar product of the derivative of $u \mapsto \frac{1}{2}(u')^2$ is needed, which one formally computes as u''. We used the second order central difference quotient $\frac{u_{k+1}-2u_k-u_{k-1}}{\Delta t^2}$ to approximate $u''(t_k)$ at time step k with the appropriate modifications for the first and last time step, respectively. All computations were performed within the FEnICS framework [110].

For the experiment duration, we set $T_1 - T_0 = 2.0$ s with timesteps $\Delta t = 0.02$ s and $T_0 = 0.0$ s, while we use $\gamma = 10^{-8}$ and $\beta = 10^{-5}$. This comparatively large value for β is only possible due to the nondimensionalization performed for the computations because otherwise the control function takes immense values, cf. the value of $u_{\max}$ in Table 4.1, so that the objective is dominated by its last two addends unless β is chosen very small, as done e.g. for the numerical computations in [88], where β was set to 10^{-13}. In the following, we elaborate on two settings: one in which we enforce the state constraint $\theta \leq \theta_{\max}$ and one in which we do not.



(a) Free optimization.

(b) State constrained optimization.

Figure 4.2. Detail of the sawteeth in E at end time $t = 2.0$ s with distribution of the temperature θ in K.

Figure 4.2 shows the temperature distribution at end time $T_1 = 2.0$ s in E in both cases. The desired temperature distribution close to uniformly 1500 K has been nearly achieved in the free optimization, see Figure 4.2a, at the price of very high temperature values around D and N already early in the heating process. We come back to this below, cf. also Figure 4.6. For the state-constrained optimization, we achieve a much worse result (note the same scales in both Figure 4.2a and 4.2b), which again corresponds to the rapid evolution to high temperatures at the critical areas, since these crucially limit the maximal amount of energy induced into the workpiece if one wants to prevent the temperature rising higher than the given bounds $\theta_{\max}$. This can also be seen in the development of the optimal controls in both cases over time, see below.

The potential φ and its gradient $\nabla\varphi$ associated with the optimal control to the state-constrained optimization problem, at time $t = 1.0$ s are depicted in Figures 4.3 and 4.4. Here, $\nabla\varphi$ is to be understood as the projection of the potentially discontinuous gradient of φ to the space of continuous

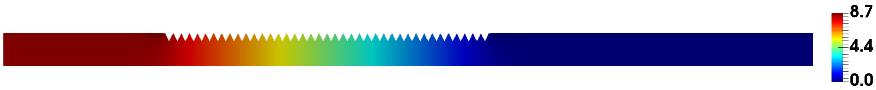


Figure 4.3. The potential φ (in V) associated with the optimal solution at time $t = 1.0$ s, view from the side (x_1x_2-plane).

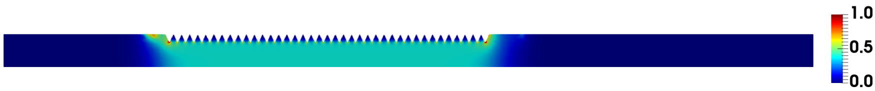
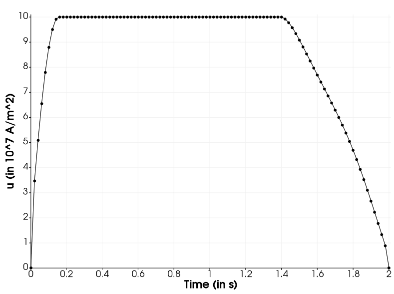
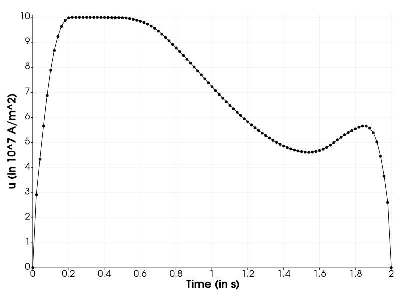


Figure 4.4. Magnitude of the gradient $\nabla\varphi$ (in V/m) associated with the optimal solution at time $t = 1.0$ s, view from the side (x_1x_2-plane).

linear finite elements. The potential φ decreases from N to the grounding with prescribed value $\varphi \equiv 0$ at D, cf. Figure 4.1b, thus inducing a current flow and acting as a heat source between D and N, since the corresponding term in the heat equation $(\sigma(\theta)\rho\nabla\varphi) \cdot \nabla\varphi$ is proportional to $|\nabla\varphi|^2$ due to the coercivity and boundedness of ρ and the bounds on σ. This is confirmed by the magnitude of $\nabla\varphi$ as seen in Figure 4.4. In particular one observes that $\nabla\varphi$ is very small or 0 in E, which means that the current flows only through the area between D and N and right *below* E, heating only this part of the workpiece.



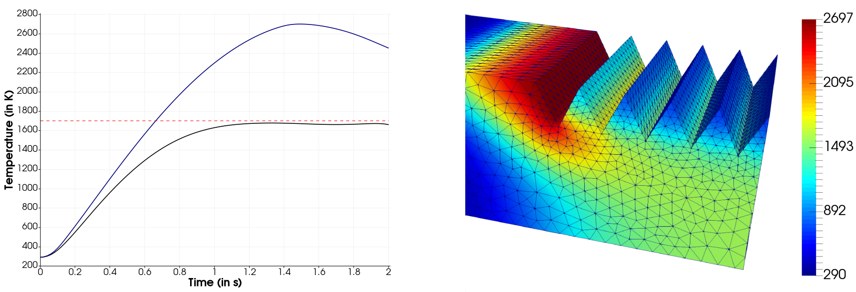
(a) For the unconstrained problem.



(b) For the state-constrained problem.

Figure 4.5. Time plot of the optimal controls, taken at an arbitrary but fixed grid point in N.

The optimal controls are shown in Figure 4.5, taken at an arbitrary but fixed grid point in N. The high values in the control at the beginning of the process seem to be the result of the inability to heat up the tooth rack in the design-area E directly as explained above, which makes heating of the teeth reliant on diffusion. This in turn requires the needed total energy to be inserted into the system as fast as possible, resulting in high control values, which also agrees with the requirement to obtain a *uniform* temperature distribution in the tooth rack. These considerations also underline the necessity of control bounds in this example. In decreasing the control values after the initial period, the optimization procedure in the free optimization is avoiding to "over-shoot", i.e., to produce a higher temperature than desired. In the case of state-constrained optimization, the presence of the state constraints forces an earlier decrease in control values in order to not violate the upper bound $\theta_{\max}$, which is then compensated by a slightly higher level of values towards the end of the simulation. This, however, is clearly not enough to make up for the earlier decrease as seen in Figure 4.2.



(a) Time plot of the temperature in a point close to N.

(b) Temperature θ in K at the critical area near N at time $t = 1.4$ s.

Figure 4.6. Influence and necessity of state constraints.

Figure 4.6 illustrates why state constraints are a necessary addition to an appropriate model of the industrial steel heating process. Figure 4.6a shows the temperature evolution in a point in one of the two critical regions, which are the points near D and N, see also Figure 4.6b and the magnitude of $\nabla\varphi$ at this region in Figure 4.4. In this case, the point lies in E close to N, but we emphasize that the state constraints hold in the whole Ω and are not limited to E. The upper line in Figure 4.6a corresponds to the temperature associated to the optimal solution of the unconstrained optimization, while the lower belongs to the state-constrained optimal solution, with the upper bound $\theta_{\max} = 1700$ K marked by the dashed line. In the free optimization case, the temperature exceeds the bounds already at about one third of the simulation time and continues to rise to almost 1000 K above $\theta_{\max}$. On the other hand, the temperature obtained from the state-constrained case stays below the threshold, as required. Note here that the evaluated point is chosen as one of those where the temperature rises highest overall, compare the temperature distribution as seen in Figure 4.6b and the maximal temperature achieved in the free optimization case in Figure 4.6a.

Concluding from the results presented above, it becomes apparent that the prescribed time of 2.0 s is too short to heat up the workpiece in the given geometry enough to reach the required temperature for Austenite to form in the workpiece (cf. [33, Ch. 9.18]) in E, if melting is to be prevented. Further computational experiments have shown that, not unexpectedly, the situation improves when increasing the duration to, for example, 4.0s s, which resulted in a significantly smaller objective value. On the other hand, the process then becomes both less interesting, both mathematically and from the manufacturer's point of view who might be interested in a higher production rate. It may thus be interesting to consider the thermistor problem also from a time-optimal control point of view.

List of symbols

General

Symbol	Description
$\mathbb{N}, \mathbb{N}_0, \mathbb{R}, \mathbb{C}$	Natural numbers, natural numbers including zero, real numbers, and complex numbers ....
$\mathbb{N}_0^s$	Multiindices of length $s \in \mathbb{N}_0$ ....
$\mathbb{R}_0^+, \mathbb{R}^+$	Real positive numbers, with and without zero ......
$\mathbb{R}^d$	d-dimensional Euclidean space ....
λ^d	d-dimensional Lebesgue measure ....
$\doteq$	Equality up to equivalent norms ....
$\cong$	Estimated from above and below ....
$\hookrightarrow$, $\hookrightarrow\hookrightarrow$, $\hookrightarrow_{\mathrm{d}}$	Embedding, compact embedding, dense embedding 1
p'	Hölder conjugate to $1 \le p \le \infty$ given by $\frac{1}{p'} = 1 - \frac{1}{p}$.
$p^\star(\ell)$	ℓ-Sobolev conjugate to $1 \le \ell p < d$, $\frac{1}{p^\star(\ell)} = \frac{1}{p} - \frac{\ell}{d}$. 22
$\mathbf{1}_E$	Indicator function for the set E ....
$f^{-1}[M]$	Preimage of the set M under f ....
$\mathbb{M}_d, \mathbb{S}_d$	Real $(d \times d)$-matrices and subspace of symmetric matrices 83

Function spaces

Geometry

Ξ	Closed subset of $\partial\Lambda$	
Γ	Complement of Ξ within $\partial\Lambda$, so $\Gamma = \partial\Lambda \setminus \Xi$	
$\mathcal{H}^s$	s-dimensional Hausdorff measure	33
K, K^-, Σ, Σ^-	Model sets of local geometries	53
ω	Surface measure	63
Q	$J \times \Omega$, time-space cylinder over Ω	

Linear operators

$\mathscr{L}(X;Y)$	Linear continuous operators between X and Y	
$\mathscr{L}_{\mathrm{iso}}(X;Y)$	Bijective operators in $\mathscr{L}(X;Y)$ whose inverses are in $\mathscr{L}(Y;X)$	
X'	Short for $\mathscr{L}(X;\mathbb{C})$, the continuous linear functionals on X	
$\mathscr{L}(X)$	Short for $\mathscr{L}(X;X)$	
$\langle\cdot,\cdot\rangle_{X,Y}$	Dual pairing between X and Y	
$\langle\cdot,\cdot\rangle_X$	Short for $\langle\cdot,\cdot\rangle_{X',X}$	
$(\cdot,\cdot)_X$	Scalar product in the Hilbert space X	
A^*	Adjoint operator of an operator A	

Particular operators

$\mathscr{F}$	$(\mathscr{F}f)(\xi) := (\frac{1}{2\pi})^{d/2}\int_{\mathbb{R}^d} e^{-\mathrm{i}\mathrm{x}\cdot\xi} f(\mathrm{x})\,\mathrm{dx}$, Fourier transform	
$\mathfrak{R}_\Upsilon$	Restriction of distributions to Υ	19
$\mathfrak{E}^0, \mathfrak{E}^0_j, \mathfrak{E}^0_U$	Extension by zero	

Optimal control

Bibliography

[1] D. R. Adams and L. I. Hedberg, *Function Spaces and Potential Theory*, Springer Berlin Heidelberg, 1996.

[2] R. A. Adams and J. J. Fournier, *Sobolev Spaces*, vol. 140 of Pure and Applied Mathematics, Elsevier Ltd, Oxford, 2003.

[3] H. Amann, *Linear and Quasilinear Parabolic Problems*, Birkhäuser Basel, 1995.

[4] ——, *Compact embeddings of vector-valued Sobolev and Besov spaces*, Glasnik Matematicki, 35 (2000), pp. 161–177.

[5] ——, *Linear parabolic equations involving measures*, Revista de la Real Academia de Ciencias Exactas, 95 (2001), pp. 85–119.

[6] ——, *Maximal regularity and quasilinear parabolic boundary value problems*, in Recent Advances in Elliptic and Parabolic Problems, Proc. Int. Conf., Hsinchu, Taiwan, C.-S. L. C.-C. Chen, M. Chipot, ed., World Scientific, 2004, pp. 1–17.

[7] ——, *Maximal regularity for nonautonomous evolution equations*, Adv. Nonlinear Stud., 4 (2004), pp. 417–430.

[8] ——, *Non-local quasi-linear parabolic equations*, Russian Math. Surveys, 60 (2005), p. 1021.

[9] ——, *Nonautonomous parabolic equations involving measures*, J Math Sci, 130 (2005), pp. 4780–4802.

[10] ——, *Quasilinear parabolic problems via maximal regularity*, Adv. Differential Equations, 10 (2005), pp. 1081–1110.

[11] ——, *Time-delayed Perona-Malik type problems*, Acta Mathematica Universitatis Comenianae. New Series, 76 (2007), pp. 15–38.

[12] H. Amann and P. Quittner, *Optimal control problems with final observation governed by explosive parabolic equations*, SIAM J. Control Optim., 44 (2005), pp. 1215–1238.

[13] M. R. S. Ammi, *Optimal control for a nonlocal parabolic problem resulting from thermistor system*, International Journal on Ecological Economics and Statistics, 9 (2007), pp. 116–122.

[14] S. B. Angenent, *Nonlinear analytic semiflows*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics, 115 (1990), pp. 91–107.

[15] S. N. Antontsev and M. Chipot, *The thermistor problem: Existence, smoothness uniqueness, blowup*, SIAM J. Math. Anal., 25 (1994), pp. 1128–1156.

[16] ——, *Analysis of blowup for the thermistor problem*, Sib Math J, 38 (1997), pp. 827–841.

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

[18] W. Arendt, R. Chill, S. Fornaro, and C. Poupaud, *L^p-maximal regularity for non-autonomous evolution equations*, J. Differential Equations, 237 (2007), pp. 1–26.

[19] J.-P. Aubin, *Un théoreme de compacité*, CR Acad. Sci. Paris, 256 (1963), pp. 5042–5044.

[20] P. Auscher, N. Badr, R. Haller-Dintelmann, and J. Rehberg, *The square root problem for second-order, divergence form operators with mixed boundary conditions on L^p*, J. Evol. Equ., 15 (2014), pp. 165–208.

[21] S. Banach, *Théorie des opérations linéaires*, Instytut Matematyczny Polskiej Akademi Nauk, 1932.

[22] S. Bartels and T. Roubíček, *Thermo-visco-elasticity with rate-independent plasticity in isotropic materials undergoing thermal expansion*, ESAIM Math. Model. Numer. Anal., 45 (2010), pp. 477–504.

[23] I. M. Berezanskii, *Selfadjoint operators in spaces of functions of infinitely many variables*, American Mathematical Society Providence, R.I, 1986.

[24] J. Bergh and J. Löfström, *Interpolation Spaces*, Springer Berlin Heidelberg, 1976.

[25] M. Biegert, *On traces of Sobolev functions on the boundary of extension domains*, Proceedings of the American Mathematical Society, 137 (2009), pp. 4169–4176.

[26] P. Binding and R. Hryniv, *Relative boundedness and relative compactness for linear operators in Banach spaces*, Proceedings of the American Mathematical Society, 128 (2000), pp. 2287–2291.

[27] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Science + Business Media, 2000.

[28] K. Brewster, D. Mitrea, I. Mitrea, and M. Mitrea, *Extending sobolev functions with partially vanishing traces from locally (ϵ, δ) domains and applications to mixed boundary problems*, J. Funct. Anal., 266 (2014), pp. 4314–4421.

[29] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer New York, 2010.

[30] V. Burenkov, *Sobolev Spaces on Domains*, Vieweg+Teubner Verlag, 1998.

[31] ——, *Extension theorems for Sobolev spaces*, in The Maz'ya Anniversary Collection, Volume 1: On Maz'ya's work in functional analysis, partial differential equations and applications, Springer Science + Business Media, 1999, pp. 187–200.

[32] A. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math., 24 (1964), pp. 113–190.

[33] W. Callister, *Materials Science And Engineering: An Introduction*, John Wiley & Sons, 2007.

[34] E. Casas, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim., 31 (1993), pp. 993–1006.

[35] E. Casas, L. A. Fernández, and J. Yong, *Optimal control of quasilinear parabolic equations*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics, 125 (1995), pp. 545–565.

[36] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Springer Berlin Heidelberg, 1977.

[37] C. Chaboudez, S. Clain, R. Glardon, J. Rappaz, M. Swierkosz, and R. Touzani, *Numerical modelling of induction heating of long workpieces*, IEEE Transactions on Magnetics, 30 (1994), pp. 5028–5037.

[38] X. Chen, A. Jüngel, and J.-G. Liu, *A note on aubin-lions-dubinskiÄη lemmas*, Acta Appl. Math., 133 (2013), pp. 33–43.

[39] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[40] G. Cimatti, *Optimal control for the thermistor problem with a current limiting device*, IMA J. Math. Control Inform., 24 (2006), pp. 339–345.

[41] P. Clément and S. Li, *Abstract parabolic quasilinear equations and application to a groundwater flow problem.*, Adv. Math. Sci. Appl., 3 (1994), pp. 17–32.

[42] P. Clément and G. Simonett, *Maximal regularity in continuous interpolation spaces and quasilinear parabolic equations*, Journal of Evolution Equations, 1 (2001), pp. 39–67.

[43] J. B. Conway, *A Course in Functional Analysis*, vol. 96 of Graduate Texts in Mathematics, Springer-Verlag New York, 2007.

[44] G. Da Prato and P. Grisvard, *Equations d'évolution abstraites non linéaires de type parabolique*, Annali di Matematica Pura ed Applicata, 120 (1979), pp. 329–396.

[45] Y. H. Dai and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., 10 (1999), pp. 177–182.

[46] D. Danielli, N. Garofalo, and D.-M. Nhieu, *Non-doubling Ahlfors measures, Perimeter measures, and the characterization of the trace spaces of Sobolev functions in Carnot-Carathéodory spaces*, American Mathematical Soc., 2006.

[47] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 2*, Springer Science + Business Media, 1988.

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

[49] L. de Simon, *Un'applicazione della teoria degli integrali singolari allo studio delle equazioni differenziali lineari astratte del primo ordine*, Rendiconti del Seminario Matematico della Università di Padova, 34 (1964), pp. 205–223.

[50] R. Deville, G. Godefroy, and V. Zizler, *Smoothness and renormings in Banach spaces*, Longman Scientific & Technical, 1993.

[51] J. Diestel and J. Uhl, *Vector measures*, American Mathematical Society, Providence, Rhode Island, 1977.

[52] K. Disser, H.-C. Kaiser, and J. Rehberg, *Optimal Sobolev regularity for linear second-order divergence elliptic operators occurring in real-world problems*, SIAM J. Math. Anal., 47 (2015), pp. 1719–1746.

[53] K. Disser, A. ter Elst, and J. Rehberg, *Hölder estimates for parabolic operators on domains with rough boundary*, Ann. Sc. Norm. Sup. Pisa, (2015).

[54] ——, *On maximal parabolic regularity for non-autonomous parabolic operators*, J. Differential Equations, 262 (2017), pp. 2039–2072.

[55] G. Dore, *L^p regularity for abstract differential equations*, in Functional Analysis and Related Topics, Springer Science + Business Media, 1993, pp. 25–38.

[56] M. Egert, *On Kato's conjecture and mixed boundary conditions*, PhD thesis, TU Darmstadt, 2015.

[57] M. Egert, R. Haller-Dintelmann, and J. Rehberg, *Hardy's inequality for functions vanishing on a part of the boundary*, Potential Anal, 43 (2015), pp. 49–78.

[58] M. Egert, R. Haller-Dintelmann, and P. Tolksdorf, *The Kato Square Root Problem for mixed boundary conditions*, J. Funct. Anal., 267 (2014), pp. 1419–1461.

[59] M. Egert and P. Tolksdorf, *Characterizations of Sobolev functions that vanish on a part of the boundary*, in PDE 2015 – Theory and Applications of Partial Differential Equations, Nov 2015, Berlin, Germany. PDE 2015: Conference Proceedings – Special Issue of Discrete and Continuous Dynamical Systems - Series S, Sept. 2016.

[60] J. Elschner, J. Rehberg, and G. Schmidt, *Optimal regularity for elliptic transmission problems including C^1 interfaces*, Interfaces Free Bound., 9 (2007), pp. 233–252.

[61] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[62] L. A. Fernández, *Integral state-constrained optimal control problems for some quasilinear parabolic equations*, Nonlinear Analysis: Theory, Methods & Applications, 39 (2000), pp. 977–996.

[63] M. Fischer, F. Lindemann, M. Ulbrich, and S. Ulbrich, *Fréchet differentiability of unsteady incompressible Navier-Stokes flow with respect to domain variations of low regularity by using a general analytical framework*, SIAM J. Control Optim., (2017). In press.

[64] I. Fonseca and G. Parry, *Equilibrium configurations of defective crystals*, Arch. Rational Mech. Anal., 120 (1992), pp. 245–283.

[65] H. Gajewski and K. Gröger, *Semiconductor equations for variable mobilities based on boltzmann statistics or fermi-dirac statistics*, Math. Nachr., 140 (1989), pp. 7–36.

[66] H. Gajewski, K. Zacharias, and K. Gröger, *Global behaviour of a reaction-diffusion system modelling chemotaxis*, Math. Nachr., 195 (1998), pp. 77–114.

[67] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer Science + Business Media, 2001.

[68] E. Giusti, *Metodi diretti nel calcolo delle variazioni*, Unione Matematica Italiana, Bologna, 1994.

[69] J. A. Griepentrog, K. Gröger, H.-C. Kaiser, and J. Rehberg, *Interpolation for function spaces related to mixed boundary value problems*, Math. Nachr., 241 (2002), pp. 110–120.

[70] J. A. Griepentrog, W. Höppner, H.-C. Kaiser, and J. Rehberg, *A bi-Lipschitz continuous, volume preserving map from the unit ball onto a cube*, Note Mat., 28 (2008), pp. 177–193.

[71] J. A. Griepentrog, H.-C. Kaiser, and J. Rehberg, *Heat kernel and resolvent properties for second order elliptic differential operators with general boundary conditions on L^p*, Adv. Math. Sci. Appl., 11 (2001), pp. 87–112.

[72] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Society for Industrial and Applied Mathematics, Jan. 2011.

[73] K. Gröger, *A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations*, Math. Ann., 283 (1989), pp. 679–687.

[74] ——, *$W^{1,p}$-estimates of solutions to evolution equations corresponding to nonsmooth second order elliptic differential operators*, Nonlinear Analysis: Theory, Methods & Applications, 18 (1992), pp. 569–577.

[75] K. Gröger and J. Rehberg, *Resolvent estimates in $W^{-1,p}$ for second order elliptic differential operators in case of mixed boundary conditions*, Math. Ann., 285 (1989), pp. 105–113.

[76] P. Hajłasz, P. Koskela, and H. Tuominen, *Sobolev embeddings, extensions and measure density condition*, Journal of Functional Analysis, 254 (2008), pp. 1217–1234.

[77] R. Haller-Dintelmann, A. Jonsson, D. Knees, and J. Rehberg, *Elliptic and parabolic regularity for second- order divergence*

operators with mixed boundary conditions, Math. Methods Appl. Sci., 39 (2015), pp. 5007–5026.

[78] R. Haller-Dintelmann, H.-C. Kaiser, and J. Rehberg, *Elliptic model problems including mixed boundary conditions and material heterogeneities*, Journal de Mathématiques Pures et Appliquées, 89 (2008), pp. 25–48.

[79] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela, *Hölder continuity and optimal control for nonsmooth elliptic problems*, Appl. Math. Optim., 60 (2009), pp. 397–428.

[80] R. Haller-Dintelmann and J. Rehberg, *Maximal parabolic regularity for divergence operators including mixed boundary conditions*, J. Differential Equations, 247 (2009), pp. 1354–1396.

[81] ——, *Coercivity for elliptic operators and positivity of solutions on lipschitz domains*, Arch. Math. (Basel), 95 (2010), pp. 457–468.

[82] P. Harjulehto, *Traces and Sobolev extension domains*, Proceedings of the American Mathematical Society, 134 (2006), pp. 2373–2383.

[83] W. Hensgen, *A simple proof of singer's representation theorem*, Proc. Amer. Math. Soc., 124 (1996), pp. 3211–3212.

[84] M. Hieber and J. Rehberg, *Quasilinear parabolic systems with mixed boundary conditions on nonsmooth domains*, SIAM Journal on Mathematical Analysis, 40 (2008), pp. 292–305.

[85] M. Hintermüller and K. Kunisch, *Feasible and noninterior path-following in constrained minimization with low multiplier regularity*, SIAM J. Control Optim., 45 (2006), pp. 1198–1221.

[86] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Springer Netherlands, 2009.

[87] D. Hömberg, K. Krumbiegel, and J. Rehberg, *Optimal control of a parabolic equation with dynamic boundary condition*, Appl. Math. Optim., 67 (2012), pp. 3–31.

[88] D. Hömberg, C. Meyer, J. Rehberg, and W. Ring, *Opti-*

mal control for the thermistor problem, SIAM J. Control Optim., 48 (2010), pp. 3449–3481.

[89] D. Horstmann, H. Meinlschmidt, and J. Rehberg, *The full Keller-Segel model is well-posed on nonsmooth domains*. Submitted, 2016.

[90] V. Hrynkiv, *Optimal boundary control for a time dependent thermistor problem.*, Electronic Journal of Differential Equations (EJDE) [electronic only], 2009 (2009), pp. Paper No. 83, 22 p., electronic only–Paper No. 83, 22 p., electronic only.

[91] V. Hrynkiv, S. Lenhart, and V. Protopopescu, *Optimal control of a convective boundary condition in a thermistor problem*, SIAM J. Control Optim., 47 (2008), pp. 20–39.

[92] D. Jerison and C. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal., 130 (1995), pp. 161–219.

[93] P. W. Jones, *Quasiconformal mappings and extendability of functions in Sobolev spaces*, Acta Mathematica, 147 (1981), pp. 71–88.

[94] A. Jonsson and H. Wallin, *Function spaces on subsets of* $\mathbb{R}^n$, vol. 2, Harwood Academic Publishers Gmbh, 1984.

[95] N. Kalton and G. Lancien, *A solution to the problem of* L^p*-maximal regularity.*, Math. Z., 235 (2000), pp. 559–568.

[96] T. Kato, *Perturbation Theory for Linear Operators*, Springer Berlin Heidelberg, 1976.

[97] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Society for Industrial and Applied Mathematics, Jan. 2000.

[98] K. Krumbiegel and J. Rehberg, *Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints*, SIAM J. Control Optim., 51 (2013), pp. 304–331.

[99] P. C. Kunstmann and L. Weis, *Perturbation theorems for maximal* L_p*-regularity*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, 30 (2001), pp. 415–435.

[100] P. C. Kunstmann and L. Weis, *Maximal L^p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus*, in Functional Analytic Methods for Evolution Equations, Springer Science + Business Media, 2004, pp. 65–311.

[101] O. Ladyzhenskaya, V. Solonnikov, and N. Ural'tseva, *Linear and quasilinear equations of parabolic type*, vol. 23 of Translations of Mathematical Monographs,, American Mathematical Society, Providence, R.I., 1968. Translated from the Russian by S. Smith.

[102] O. A. Ladyzhenskaya and N. N. Uraltseva, *Linear and quasilinear elliptic equations*, vol. 46 of Mathematics in science and engineering, New York, Academic Press, 1968.

[103] S. Lang, *Real and Functional Analysis*, Springer New York, 1993.

[104] H.-C. Lee and T. Shilkin, *Analysis of optimal control problems for the two-dimensional thermistor system*, SIAM J. Control Optim., 44 (2005), pp. 268–282.

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

[106] I. E. Leonard and K. F. Taylor, *Supremum norm differentiability*, Internat. J. Math. Math. Sci., 6 (1983), pp. 705–713.

[107] J. L. Lions, *Optimisation pour certaines classes d'équations d'évolution non linéaires*, Annali di Matematica Pura ed Applicata, Series 4, 72 (1966), pp. 275–293.

[108] ——, *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Etudes Mathématiques, Dunod, 1968.

[109] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris, 1969.

[110] A. Logg, K.-A. Mardal, and G. Wells, eds., *Automated Solution of Differential Equations by the Finite Element Method*, Springer Berlin Heidelberg, 2012.

[111] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Springer Nature, 1995.

[112] V. Maz'ya, *Sobolev Spaces*, Springer Berlin Heidelberg, 2011.

[113] H. Meinlschmidt, C. Meyer, and J. Rehberg, *Optimal control of the thermistor problem in three spatial dimensions, part 1: Existence of optimal solutions*, SIAM J. Control Optim., (2017). In press.

[114] ——, *Optimal control of the thermistor problem in three spatial dimensions, part 2: Optimality conditions*, SIAM J. Control Optim., 55 (2017), pp. 2368–2392.

[115] ——, *A space of controls for optimal control problems associated to nonlinear evolution equations*. In preparation, 2017.

[116] H. Meinlschmidt and J. Rehberg, *Hölder-estimates for non-autonomous parabolic problems with rough data*, Evol. Equ. Control Theory, 5 (2016), pp. 147–184.

[117] N. G. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, 17 (1963), pp. 189–206.

[118] N. G. Meyers and J. Serrin, $H = W$, Proc. Nat. Acad. Sci. U.S.A., 51 (1964), pp. 1055–1056.

[119] I. Mitrea and M. Mitrea, *The Poisson problem with mixed boundary conditions in Sobolev and Besov spaces in non-smooth domains*, Transactions of the American Mathematical Society, 359 (2007), pp. 4143–4183.

[120] C. B. Morrey, Jr., *Functions of several variables and absolute continuity, ii*, Duke Math. J., 6 (1940), pp. 187–215.

[121] S. Nababan and E. S. Noussair, *Existence theorems for optimal control of quasilinear parabolic partial differential equations*, The Journal of the Australian Mathematical Society. Series B. Applied Mathematics, 21 (1979), p. 90.

[122] L. Nirenberg, *On elliptic partial differential equations*, Ann. Scuola. Norm. Sup. Pisa, 13 (1959), pp. 115–162.

[123] E. S. Noussair, S. Nababan, and K. L. Teo, *On the existence of optimal controls for quasilinear parabolic partial differential equations*, Journal of Optimization Theory and Applications, 34 (1981), pp. 99–115.

[124] N. S. Papageorgiou, *On the optimal control of strongly nonlinear evolution equations*, Journal of Mathematical Analysis and Applications, 164 (1992), pp. 83–103.

[125] N. S. Papageorgiou, *Optimal control of nonlinear evolution equations with nonmonotone nonlinearities*, Journal of Optimization Theory and Applications, 77 (1993), pp. 643–660.

[126] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer New York, 1983.

[127] J. Prüss, *Maximal regularity for evolution equations in L^p-spaces*, Conf. Semin. Mat. Univ. Bari, 285 (2002), pp. 1–39.

[128] J. Prüss and R. Schnaubelt, *Solvability and maximal regularity of parabolic evolution equations with coefficients continuous in time*, J. Math. Anal. Appl., 256 (2001), pp. 405–430.

[129] S. M. Robinson, *Stability theory for systems of inequalities, part II: Differentiable nonlinear systems*, SIAM Journal on Numerical Analysis, 13 (1976), pp. 497–513.

[130] L. G. Rogers, *Degree-independent Sobolev extension on locally uniform domains*, Journal of Functional Analysis, 235 (2006), pp. 619–665.

[131] W. Rudin, *Functional Analysis*, International series in pure and applied mathematics, McGraw-Hill, 1991.

[132] L. Schwartz, *Théorie des distributions. Tome II.*, Publ. Inst. Math. Univ. Strasbourg. X. Paris: Hermann & Cie., 1951.

[133] ——, *Théorie des distributions. Tome I. Deuxième éd. revue et corrigée.*, Publ. Inst. Math. Univ. Strasbourg. X. Paris: Hermann & Cie., 1957.

[134] T. I. Seidman and H.-X. Zhou, *Existence and uniqueness of op-*

timal controls for a quasilinear parabolic equation, SIAM Journal on Control and Optimization, 20 (1982), pp. 747–762.

[135] A. M. Selitskii, *L^p-solvability of parabolic problems with an operator satisfying the kato conjecture*, Differential Equations, 51 (2015), pp. 776–782.

[136] E. Shamir, *Regularization of mixed second-order elliptic problems*, Israel J. Math., 6 (1968), pp. 150–168.

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

[138] ——, *Sobolev, Besov and Nikolskii fractional spaces: Imbeddings and comparisons for vector valued spaces on an interval*, Ann. Mat. Pura Appl. (4), 157 (1990), pp. 117–148.

[139] I. J. Šneĭberg, *Spectral properties of linear operators in interpolation families of Banach spaces*, Mat. Issled., 9 (1974), pp. 214–229, 254–255.

[140] P. E. Sobolevskii, *Coerciveness inequalities for abstract parabolic equations*, Soviet. Mat. (Doklady), 5 (1965), pp. 894–897.

[141] J. Sokolowski, *On parametric optimal control for a class of linear and quasilinear equations of parabolic type.*, Control Cybern., 4 (1975), pp. 19–38.

[142] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, Studies in advanced mathematics, World Scientific, 2003.

[143] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Springer Berlin Heidelberg, 2007.

[144] A. ter Elst and J. Rehberg, *Hölder estimates for second-order operators on domains with rough boundary*, Adv. Differential Equations, 20 (2015), pp. 299–360.

[145] D. Tiba, *Optimality conditions for distributed control problems with nonlinear state equation*, SIAM Journal on Control and Optimization, 23 (1985), pp. 85–110.

[146] H. Triebel, *Interpolation theory, function spaces, differential op-*

erators, vol. 18 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam-New York, 1978.

[147] ——, *On spaces of $B^s_{\infty,q}$ Type and $\mathfrak{C}^s$ Type*, Math. Nachr., 85 (1978), pp. 75–90.

[148] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, American Mathematical Society, apr 2010.

[149] D. Wachsmuth, *The regularity of the positive part of functions in $L^2(I;H^1(\Omega)) \cap H^1(I;H^1(\Omega)^*)$ with applications to parabolic equations*, Comment Math Univ Carolin, 57 (2016), pp. 327–332.

[150] G. Wachsmuth, *Differentiability of implicit functions: Beyond the implicit function theorem*, Journal of Mathematical Analysis and Applications, 414 (2014), pp. 259–272.

[151] S. Yang, *A Sobolev extension domain that is not uniform*, manuscripta mathematica, 120 (2006), pp. 241–251.

[152] J. Yeh, *Theory of Measure and Integration*, World Scientific Publishing Co. Pte. Ltd., 2006.

[153] K. Yosida, *Functional Analysis*, Springer Nature, 1980.

[154] D. Z. Zanger, *The inhomogeneous Neumann problem in Lipschitz domains*, Communications in Partial Differential Equations, 25 (2000), pp. 1771–1808.

[155] W. P. Ziemer, *Weakly Differentiable Functions*, Springer New York, 1989.

[156] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.

Wissenschaftlicher Werdegang

Hannes Meinlschmidt
geboren am 21. 6. 1985 in Groß-Gerau

2017	Promotion in Mathematik
seit 10/2011	Wissenschaftlicher Mitarbeiter am Fachbereich Mathematik der Technischen Universität Darmstadt
08/2011	Diplom in Mathematik
04/2005–08/2011	Studium der Mathematik an der Technischen Universität Darmstadt
06/2004	Abitur an der Prälat-Diehl Schule Groß-Gerau