Flat Fronts and Stability for the Porous Medium Equation

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Clemens Kienzler

aus

Blaubeuren

Bonn, August 2013

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

genehmigte redaktionell korrigierte Fassung

1. Gutachter: Prof. Dr. Herbert Koch

2. Gutachter: Prof. Dr. Juan José López Velázquez

Tag der Promotion: 16. September 2013

Erscheinungsjahr: 2014

Acknowledgements

There have been many a day when I was not sure if my venture was only daring or simply mad. However, seemingly against all odds, the end of my first great expedition into the fascinating, albeit sometimes very foreign world of mathematics has arrived, and with it the time for giving thanks.

First and foremost, I would like to express my deep gratitude towards my advisor Professor Herbert Koch. Not only did he introduce me to the different theories that gear into each other so beautifully in this field, but he also showed an extraordinary degree of patience, kindness and intuition in guiding me. Needless to say that this thesis would not have been possible without him.

If it were not for Maria Athanassenas, Georg Menz and Professor Felix Otto, though, I would not even have started this thesis. They smoothed my way to Bonn decisively, probably without themselves even knowing.

On the other hand, I would not have finished this thesis if it were not for Dominik John. Apart from his valuable mathematical input, his companionship has carried me through hard times, and he has become a dear personal friend to me.

Many fruitful discussions, both long and short, have helped me improve this work. Among the mathematicians that contributed are - in order of their appearance - Axel Grünrock, Sebastian Herr, Gereon Knott, Alexander Raisch, Antoine Choffrut, Stefan Steinerberger, Tobias Schottdorf, Angkana Rüland, Christian Zillinger and Christian Weiß.

I am also very much indebted to my family in Köln, Laichingen, München/Reykjavik/Stockholm and New York/Sankt Gallen, constantly spreading out and continuing to grow, as well as my friends, be they mathematicians or not. There is a life beyond mathematics - music, literature, football, philosophical discussions - and they had to remind me of that from time to time. In days and nights of joy as well as despair, they hoped and trembled, suffered and rejoiced with me, and they believed in me, sometimes more than I did myself. A hearty thanks to all of them, for so many things.

My studies have been supported by the Bonn International Graduate School at the Hausdorff Centre of Mathematics, and the Cusanuswerk. Besides the financial benefit, especially the company of the latter provided me with treasured encounters and memories that encouraged and inspired me. For all this I am very thankful.

Contents

1 Introduction 1

2 Preliminaries 11

3 Solutions of the Linear Perturbation Equation 33

4 Energy Estimates 43

5 The Intrinsic Metric 55

6 Local Estimates 69

7 Estimates Against Initial Values 85

8 Gaussian Estimate and Consequences 91

9 Estimates Against the Inhomogeneity 101

10 The Non-Linear Equation 119

A Singular Integrals in Spaces of Homogeneous Type 127

Bibliography 137

1 Introduction

The Porous Medium Equation

The density of gas in a porous medium can be modelled by the so-called porous medium equation (PME). The following definition gives a meaning to this notion in mathematical terms.

1.1 Definition *Consider an open set $\omega \subset \mathbb{R} \times \mathbb{R}^n$.*
A function u is said to be a solution of the PME on ω if and only if for $m > 1$ we have $u \in L^m_{loc}(\omega)$, $u \geq 0$ and

$$\partial_t u - \Delta_x u^m = 0 \ \textit{in} \ \mathcal{D}'(\omega).$$

From the theoretical point of view we deal with a quasilinear and strongly parabolic equation. However, the parabolicity is not necessarily uniform on all of ω, but degenerates on subsets where u is not bounded away from 0. This becomes evident by considering the equation in divergence form, namely $\partial_t u - \nabla_x \cdot (m\, u^{m-1} \nabla_x u) = 0$, where the diffusion coefficient u^{m-1} – the pressure of the gas, in physical terms – vanishes as the density u approaches zero. This is why the PME is also called slow diffusion equation. In contrast to that, the case $0 < m < 1$ – with a suitable adjustment of the concept of solution – is referred to as the fast diffusion equation. For $m = 1$, the same definition as above yields a well-known linear and uniformly strongly parabolic equation on ω, termed heat equation. The PME is therefore also sometimes labelled non-linear heat equation. More general non-linearities can of course be considered with the appropriate adjustments in the definition, and equations associated with them are also known as diffusion or filtration equations. Other generalisations include the treatment of different coefficients and additional drift terms. Also higher order equations in divergence form can have the same appearance. An example is provided by the thin film equation. For any of these equations one could additionally consider solutions that do not have to be positive, also called signed solutions.
In the literature, often a combination of several of those features is discussed and a vast amount of research papers on any of those topics is available. One can gain a good overview from [Fri82], [Váz07] and [DK07]. There is also an interpretation as a gradient flow in terms of the Wasserstein metric due to [Ott01]. We will concentrate on the PME as presented in Definition 1.1. A derivation of the equation from physical grounds as well as an overview of some of its mathematical properties and their development can be found in the surveys [Pel81], [Aro86], [Váz92] and the references therein. In the following we give a brief recapitulation of results that are important from our perspective.
In Definition 1.1 we have adopted a purely local point of view. This is already enough to obtain at least weak regularity of solutions.

1.2 Theorem *Let $\omega \subset \mathbb{R} \times \mathbb{R}^n$ be open.*
If u is a solution of the PME on ω, then u is Hölder continuous on ω.

That continuous solutions are Hölder continuous was proven by [DF85]. Almost a decade later [DK93] succeeded in removing the continuity assumption by means of potential theory in conjunction with a-priori estimates for the PME from [DK84] and [Sac83] that only require general regularity theory.

The Cauchy Problem

Henceforth we concentrate on domains without spatial boundary and consider cylinders $\omega = I \times \mathbb{R}^n$ for an open time interval $I = (t_1, t_2)$ with finite left end point $t_1 > -\infty$. There is a self-similar solution to the PME on $I \times \mathbb{R}^n$ that displays some characteristic features of the PME and often plays the role of a benchmark. It was first described by the Soviet researchers Zeldovich, Kompaneets and Barenblatt in the 1950s, thus often bearing their name as ZKB-solution. The first reference not written in Russian seems to be [Pat59], hence also Barenblatt-Pattle-solution is a commonly used denomination. We call it source-type-solution, given by

$$u_{st}(t,x) := (t-t_1)^{-\frac{n}{2+n(m-1)}} \left(C_1 - C_2 \, |x|^2 \, (t-t_1)^{-\frac{2}{2+n(m-1)}} \right)_+^{\frac{1}{m-1}}$$

with specific constants C_1, C_2 that depend only on n and m. The nomenclature can be explained by the behaviour of u_{st} for small times: It tends to a multiple of the Dirac delta in the sense of distributions as t approaches t_1 from above, therefore representing a solution that evolves from a point source.

Also in general it is a natural question to ask what happens to a solution when approaching the initial time. It was proven in [Pie82] that not only u_{st}, but any solution satisfying some additional conditions has a Borel measure as initial value. That this is indeed typical without further restrictions was shown in [AC83].

1.3 Theorem *Let $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval.*
If u is a solution of the PME on $I \times \mathbb{R}^n$, then there exists exactly one Borel measure g on $\mathbb{R}^n$ with

$$u(t,\cdot) \to g \ (t \to t_1) \ \text{ in } \mathcal{D}'(\mathbb{R}^n).$$

The proof has several important ingredients that are interesting results in their own right. One of them is a weak Harnack-type inequality that bounds the mass of gas contained in a ball by its radius and the density of the gas at its centre at a later time. Another one is a comparison principle for solutions of the PME that follows from the approximation of the original problem by suitable boundary-value-problems on bounded domains. Their theory, in turn, is developed in one spatial dimension by [ACP82], and contains an a-priori comparison principle that is proven in multiple space dimensions by [DK84]. Also a geometrical assertion taken from [CF80] is generalised and applied here.

Actually, the paper provides even more information: Any Borel measure that is an initial value possesses a property that limits the amount of mass it can place far outside in space in the sense of the following definition.

1.4 Definition *A Borel measure μ on $\mathbb{R}^n$ is said to be essentially mass bounded if and only if*

$$\sup_{r>1} r^{-n-\frac{2}{m-1}} \mu(B_r(0)) < \infty.$$

It turns out that this growth condition is not only necessary, but also sufficient for solving the initial value problem. This is the main result of [BCP84]. They make use of the existence of solutions for integrable and bounded initial data that is known by [BC81] and show a pointwise estimate with respect to the spatial variable for such solutions, using the additional properties that solutions to such initial data enjoy. The theory of non-linear semigroups and an approximation process then deliver the following existence result.

1.5 Theorem *Let $t_1 > -\infty$ and g be an essentially mass bounded Borel measure on $\mathbb{R}^n$. Then there exists $t_2 = t_2(g)$ with $t_1 < t_2 \leq \infty$, and a solution u of the PME on $(t_1, t_2) \times \mathbb{R}^n$ with*

$$u(t,\cdot) \to g \ (t \to t_1) \ \text{in } \mathcal{D}'(\mathbb{R}^n).$$

A lower bound for the maximal existence time t_2 can be calculated explicitly. If $g \in L^1(\mathbb{R}^n)$ and $g \geq 0$, one gets $t_2(g) = \infty$ and therefore global time solvability. This was already known from previous works as for example [BC79].
There remains the rather subtle question of uniqueness. In [Sab61] it was shown that bounded and square integrable initial data determine solutions uniquely. Then [BC79] produced the same statement for integrable initial data. The first result for possibly unbounded initial data followed in [BCP84], but excluded the case of Borel measures. In [Pie82], in turn, Borel measures are considered, imposing the constrained of them being finite. Moreover, all these uniqueness results were obtained within different classes of solutions and made use of the additional properties given in these classes. It was finally [DK84] who settled the issue in full generality. Their work relies heavily on all the previous papers mentioned above. Notably, the approximation by boundary-value-problems as in the proof of Theorem 1.3 is used again to show a similar pointwise bound as the one used in the proof of Theorem 1.5, but for general solutions.

1.6 Theorem *Let $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. If u is a solution of the PME on $I \times \mathbb{R}^n$, then it is determined uniquely by its initial value.*

Theorems 1.3, 1.5 and 1.6 provide us with a beautifully closed theory that characterises solutions of the Cauchy problem for the PME in analogy to the so called Widder theory for the heat equation ([Wid75]).

1.7 Remark *Thanks to the very general uniqueness result* 1.6, *any additional properties that were gained in existence proofs for several subclasses of initial data carry over to the general situation. For example, for a solution u to an integrable initial datum $g \in L^1(\mathbb{R}^n)$, $g \geq 0$, we know that*

$$u \in C([t_1,\infty); L^1(\mathbb{R}^n)) \cap L^\infty((t_1 + a, \infty) \times \mathbb{R}^n) \text{ for any } a > 0$$

according to [BC79], while for integrable and bounded $g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, $g \geq 0$, we even have

$$u \in C([t_1,\infty); L^1(\mathbb{R}^n)) \cap L^\infty((t_1, \infty) \times \mathbb{R}^n)$$

by [BC81]. In any of the two cases the initial datum is taken in the continuous sense and we can write $u(t_1, \cdot) = g$ almost everywhere on $\mathbb{R}^n$.

There is one time-local theorem that plays a major role in the proofs of any of the theorems just presented. It asserts the existence of a lower bound for the Laplacian of the pressure and is originally due to [AB79], there proven for solutions with additional regularity assumptions. In any step of the process of generalising the solutions, hereby weakening the conditions they satisfy, its validity was proven again, until finally once more [DK84] verified that it holds for any solution without constraints.

1.8 Theorem *Let $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. If u is a solution of the PME on $I \times \mathbb{R}^n$, then*

$$\Delta_x(u^{m-1}) \geq -\frac{n(m-1)}{nm(m-1) + 2m}\, t^{-1} \ \text{in } \mathcal{D}'(I \times \mathbb{R}^n).$$

Equality is attained by the source type solution, implying that the constant is optimal.
An immediate consequence of 1.8 is a lower bound also for the temporal derivative of a solution u, namely

$$\partial_t u \geq -\frac{n}{2+(m-1)n}\, t^{-1}\, u \text{ in } \mathcal{D}'(I \times \mathbb{R}^n).$$

A direct proof of this statement with a slightly weaker constant is contained in [CF79].
Theorem 1.8 highlights the spatial non-local character of the PME, since it only holds on the whole space $\mathbb{R}^n$. A weaker local version of this estimate was established rather recently in [LNVV09].

Finite Propagation, Interface and Regularity

As opposed to the situation for uniformely strongly parabolic equations, the degeneracy of the PME implies that any solution whose initial positivity set does not cover the whole space retains this property for any finite time. The source-type solution provides an explicit example for this feature, and drawing on the Harnack inequality from [AC83] on the one hand and results of [Ali85] based on [BCP84] on the other hand, it can be seen to hold in general ([CVW87]). In physical terms this means that the diffusing gas does not get to every point of space instantaneously, but that disturbances are rather propagated with finite speed. Not only does this paint a more realistic picture of the real world in terms of modelling diffusion processes, but it does also give rise to an interesting mathematical phenomenon: The time-space positivity set $\mathcal{P}(u)$ of solutions on $I \times \mathbb{R}^n$, open because of their continuity, has a non-empty boundary $\partial\mathcal{P}(u)$ that separates it from the time-space-region where u vanishes, thus constituting a sharply defined interface $\mathcal{G}(u) := \partial\mathcal{P}(u) \cap (I \times \mathbb{R}^n)$. We denote the spatial part of the positivity set by $\mathcal{P}(u(t)) := \{x \in \mathbb{R}^n \mid u(t,x) > 0\}$ and the spatial interface by $\mathcal{G}(u(t)) := \partial\mathcal{P}(u(t))$ for any fixed time $t \in I$, as well as the set of immediate positivity by

$$\mathcal{P}(u(t_1)) := \bigcap_{t \in I} \mathcal{P}(u(t)).$$

Note that for a solution u with initial value g, the initial positivity set $\mathcal{P}(g)$ does in general not coincide with the set of immediate positivity $\mathcal{P}(u(t_1))$, but is merely contained in it.
A direct consequence of the lower bound for the temporal derivative above is that the spatial positivity set does not shrink with time, that is we have $\mathcal{P}(u(t)) \subset \mathcal{P}(u(\bar{t}))$ for any $t \leq \bar{t} \in \bar{I}$. At this point it is not clear if the monotonicity is strict, since there may be a time span in the beginning, called waiting time, during which the spatial positivity set does not change. However, once it has actually started to spread out near a point in space it does not stop anymore, as was shown by [CF80]. Furthermore, symmetrisation techniques allowed [Váz82] to deduce bounds from below for the rate of growth of the interface, so it will eventually reach any point in space and can not remain motionless all the time.
The regularity of the interface and the regularity of solutions are closely connected. This becomes plausible by the physical interpretation of the equation that suggests to view the derivative of the pressure as the velocity of the extension of gas and hence the speed of the interface. Note that parabolic regularity theory ([LUS75]) immediately implies the smoothness of solutions on $\mathcal{P}(u)$, so regularity is only an issue near $\mathcal{G}(u)$. In one space dimension the picture is fairly complete. There the interface is always Lipschitz regular as was shown in [Aro70]. This is optimal since according to [ACV85], when starting to move after a waiting time the interface may have a corner. After the waiting time, however, it is in any case not only smooth ([AV87]), but even real analytic ([Ang88]). Furthermore, the pressure is Lipschitz continuous everywhere in time and space as was shown in [Bén83] and [Aro69].

In the general case of arbitrary dimension some irregularities can appear. In [CVW87] it was shown that u^{m-1} is a Lipschitz function in time and space for sufficiently large times and on all of $\mathbb{R}^n$, that is especially across the interface. As seen above, the temporal constraint solely applies for dimensions $n > 1$ and is important in case the positivity set of the initial datum contains one or more holes. Any hole is filled in finite time, but advancing interfaces may hit each other and the velocity of the interface can become unbounded at the focussing time, see [AG93]. So in general, for regularity of the pressure one has to wait until the focussing time has passed, any possible holes are filled and the positivity set of $u(t,\cdot)$ overflows the smallest ball in which the initial positivity set was contained ([CVW87]). That the positivity set of u gets round fast and that $\mathcal{G}(u)$ can be described as a Lipschitz continuous surface for sufficiently large times then follows from the Lipschitz property of u^{m-1}. The latter also implies that the Hölder continuity of solutions u on all of $\mathbb{R}^n$ that is given by Theorem 1.2 can be recovered at least for large times. Note that the example of the source-type solution again shows that in general one cannot expect more regularity than that, even though the interface is a smooth surface.
We now introduce a non-degeneracy condition on initial data that particularly ensures that the spatial interface starts to move at all points right in the beginning, thus generating solutions without waiting time ([Ves89]).

1.9 Definition *A function $g : \mathbb{R}^n \to \mathbb{R}$, $g \geq 0$, with bounded positivity set $\mathcal{P}(g)$ is said to be a non-degenerate initial datum if and only if*

$$g^{m-1} \in C^1(\overline{\mathcal{P}}(g)),$$

$$g^{m-1} + |\nabla_x(g^{m-1})| \geq c > 0 \text{ on } \mathcal{P}(g)$$

and

$$\Delta_x(g^{m-1}) \geq -C \ \text{ in } \mathcal{D}'(\mathcal{P}(g))$$

for constants c, $C > 0$.

Note that by definition $\mathcal{G}(u)$ contains the graph of the function

$$t = S(x) := \inf \ \{\tau \in I \mid u(\tau, x) > 0\}$$

that sends points $x \notin \overline{\mathcal{P}}(g)$ to the time when they are first reached by the gas. It was proved in [CF80] that the interface of solutions without waiting time, ensured by a slightly different concept of non-degeneracy, is in fact given as above with Hölder continuous S. For initial data that satisfy Definition 1.9, it is proven in [CVW87] that S is Lipschitz and the velocity is bounded from below for large times. Given this situation, [CW90] showed that the interface is indeed $C^{1,\alpha}$ for large times, although it was not clear from their work that the pressure enjoys the same regularity. This was improved by [Koc99] to reach large time smoothness of both the pressure and the interface.

1.10 Theorem *Let $t_1 > -\infty$ and g be a non-degenerate initial datum.*
If u is a solution of the PME on $(t_1, \infty) \times \mathbb{R}^n$ with $u(t_1) = g$, then both S and u^{m-1} are smooth for any sufficiently large t.

In the course of the argument for the proof of Theorem 1.10, a transformation of the equation on $\mathcal{P}(u)$ onto a fixed domain via a local coordinate change and a transfer of the problem into a perturbational setting is accomplished.

Short time smoothness of solutions before a possible blow-up time has been established in [DH98] for non-degenerate initial data, with a slightly different understanding of non-degeneracy that substitutes the lower bound for the Laplacian of the initial pressure for a Hölder condition for the second order derivatives with respect to an intrinsically arising singular distance function. That it is enough to impose this Hölder condition onto the first order derivatives was shown in [Koc99]. The need for the use of the special metric is a manifestation of the degeneracy of the equation. It also plays a role in the derivation of large time smoothness.
Finally, [DHL01] found that non-degenerate initial data in the sense of Definition 1.9 which in addition possess a weakly concave square root function of the pressure generate solutions u with convex positivity set for all times and hence smoothness of the pressure on the whole space for any time follows. As a consequence, $\mathcal{G}(u)$ is also smooth.

Transforms and Perturbations

Assume now that u is a solution of the PME on $I \times \mathbb{R}^n$ with positivity set $\mathcal{P}(u) \subsetneq I \times \mathbb{R}^n$. Abbreviating the rescaled pressure by $v := \frac{m}{m-1}\, u^{m-1}$, a direct calculation reveals that v satisfies the pressure equation

$$\partial_t v - (m-1)\, v\, \Delta_x v - |\nabla_x v|^2 = 0 \text{ on } \mathcal{P}(u)$$

pointwise in the classical sense. An interchange of dependent and independent variables on the positivity set then transforms this problem onto a fixed domain. Fixing a point (t_0, x_0, z_0) in the graph of $v|_{\mathcal{P}(u)}$, we assume that $|\partial_{x_n} v(t_0, x_0)| \neq 0$. The implicit function theorem then allows us to solve the defining equation of the graph, that is $v(t, x', x_n) = z$, locally near (t_0, x_0, z_0) for x_n. We therefore obtain a function $w(t, x', z) = x_n$ such that $v(t, x', w(t, x', z)) = z$ with $\nabla_{t,x'} w = -(\partial_{x_n} v)^{-1} \nabla_{t,x'} v$, $\partial_z w = (\partial_{x_n} v)^{-1}$. Thus the graph of the pressure v can be described locally near (t_0, x_0, z_0) by

$$\big\{(t, x', x_n, z) \mid w(t, x', z) = x_n\big\} = \operatorname{graph} w.$$

This suggests a change of coordinates $(t, x) \mapsto ((m-1)t, x', v(t, x)) =: (s, y)$, adding also a rescaling of time. Under this transformation, any subset of the spatial positivity set becomes a subset of the open upper half plane $H := \{x \in \mathbb{R}^n \mid x_n > 0\}$. Furthermore, w satisfies the transformed pressure equation

$$\partial_s w - y_n\, \Delta'_y w + y_n\, \partial_{y_n} \frac{1 + |\nabla'_y w|^2}{\partial_{y_n} w} + (1 + \sigma)\, \frac{1 + |\nabla'_y w|^2}{\partial_{y_n} w} = 0$$

on $\omega \subset (s_1, s_2) \times H$ with a suitable interval (s_1, s_2), where $\sigma := -\frac{m-2}{m-1} > -1$. The choice of the n-th coordinate as the one being interchanged is of course not essential.
In this work, we mainly consider a special class of solutions of the transformed pressure equation, namely perturbed travelling wave solutions. In general, a travelling wave is a function with a profile that propagates along a fixed direction with constant speed without changing its shape. To make our setting definite we will assume the speed to be normed and the direction to be the n-th coordinate direction in accordance with the choice in the above transformation. Henceforth, for the transformed pressure equation on an arbitrary $\omega \subset (s_1, s_2) \times H$ we thus call

$$w_{tw}(s, y) := y_n - (1 + \sigma)\,(s - s_1)$$

the travelling wave solution. We are now interested in the perturbed travelling wave $w_{tw} + \widetilde{u}$, where $\widetilde{u}$ denotes the perturbation. It is itself a solution of the transformed pressure equation on ω again if and only if $\widetilde{u}$ satisfies the perturbation equation

$$\partial_s \widetilde{u} - L_\sigma \widetilde{u} = f[\widetilde{u}]$$

on ω with linear spatial part

$$y_n\, \Delta_y \widetilde{u} + (1+\sigma)\, \partial_{y_n} \widetilde{u} =: L_\sigma \widetilde{u}$$

and non-linearity

$$-(1+\sigma)\, \frac{|\nabla_y \widetilde{u}|^2}{\partial_{y_n} \widetilde{u} + 1} - y_n\, \partial_{y_n} \frac{|\nabla_y \widetilde{u}|^2}{\partial_{y_n} \widetilde{u} + 1} =: f[\widetilde{u}].$$

This is the equation we are mainly dealing with in the present work.
Note that we can express both the spatial part of the operator and the non-linearity in divergence form as

$$L_\sigma \widetilde{u} = y_n^{-\sigma}\, \nabla_y \cdot (y_n^{1+\sigma}\, \nabla_y \widetilde{u})$$

and

$$f[\widetilde{u}] = -y_n^{-\sigma} \partial_{y_n} (y_n^{1+\sigma} \frac{|\nabla_y \widetilde{u}|^2}{1 + \partial_{y_n} \widetilde{u}}),$$

since $\omega \subset \{(s,y) \mid y_n > 0\}$. Furthermore, $\widetilde{u}$ satisfies the linearised perturbation equation

$$\partial_s \widetilde{u} - L_\sigma \widetilde{u} = f \text{ on } \omega$$

in the sense of distributions if and only if it satisfies

$$y_n^\sigma\, \partial_s \widetilde{u} - y_n^\sigma\, L_\sigma \widetilde{u} = y_n^\sigma\, f \text{ on } \omega.$$

A regular distributional solution $\widetilde{u}$ on ω is thus characterised by the integral identity

$$-\int\limits_\omega \widetilde{u}\, \partial_s \varphi\, y_n^\sigma\, d\mathcal{L}^{n+1} + \int\limits_\omega \nabla_y \widetilde{u} \cdot \nabla_y \varphi\, y_n^{1+\sigma}\, d\mathcal{L}^{n+1} = \int\limits_\omega f\, \varphi\, y_n^\sigma\, d\mathcal{L}$$

for any test function $\varphi \in C_c^\infty(\omega)$ and a suitable inhomogeneity f.

Flat Fronts and Stability

The last identity will serve as a model for our definition of solutions of the linear perturbation equation. In view of its appearance, the use of weighted measures is natural. Moreover, our particular interest in the behaviour of solutions towards $\{y_n = 0\}$ motivates to carry our considerations to the boundary of H by applying a wider class of test functions that can attain non-zero values there. The mainly technical results necessary for doing so are presented in Chapter 2.
In Chapter 3 we define a suitable and rather weak notion of energy solution for the linearised perturbation equation with and without initial data on time-space cylinders in $I \times \overline{H}$. Existence follows with a Galerkin approximation, and on the whole space $\overline{H}$ uniqueness of solutions is

shown by means of the energy identity 3.5.
Energy methods are used in Chapter 4 to obtain additional regularity of solutions on $I \times \overline{H}$. Propositions 4.1 and 4.3 are energy estimates that hold globally in space. The statement alongside with a formal proof is already contained in [Koc99].
Chapter 5 deals with the intrinsic metric of the problem that turns our weighted measure space into a space of homogeneous type. An equivalent characterisation in terms of an explicit expression is given in Theorem 5.6.
A localisation of the global energy estimates with respect to this metric is carried out in Chapter 6, culminating in the pointwise derivative estimate from Theorem 6.10 that also shows the smoothness of local solutions on a subset of their definition set. In [Koc99] a slightly weaker theorem of the same kind was proven in a less direct way.
The local pointwise estimate opens the way for the treatment of initial value problems on the whole space in Chapter 7. Here Proposition 7.5 is crucial, estimating the derivatives at a certain time pointwise by an exponentially weighted L^2_σ-norm at an earlier time. On the one hand this provides the pointwise estimates against rough norms of initial data in Theorem 7.6.
On the other hand, it is possible to consider the Green function of our problem as it is done in Chapter 8. Theorem 8.3 contains a pointwise exponential decay estimate of any derivative of the Green function with an upper bound that resembles the Gaussian function in terms of the intrinsically given metric and measure. Such an estimate is called Gaussian estimate or Aronson-type estimate after one of the first authors exploring this type of inequalities ([Aro67]). For general uniformely strongly parabolic equations their proof was originally given by means of the Harnack inequality contained in [Mos64] and [Mos67]. This order was reversed by [FS86] and Gaussian estimates were shown directly. This idea was extended by [Koc99] to cover the degenerate parabolic case with measurable coefficients. Our proof simplifies this approach in a special case of constant coefficients and at the same time adds control over the derivatives of the Green function. Compare also Remark 8.4.
The Gaussian estimate enables us to consider initial value problems with rough initial data and more general inhomogeneities as well as to gain both on-diagonal and off-diagonal kernel estimates. Their consequences are studied in Chapter 9, where the global pointwise estimate in Theorem 9.9 is derived. Furthermore, we apply the theory of singular integrals and Calderón-Zygmund operators in spaces of homogeneous type to find localised L^p-estimates against the inhomogeneity for the linear equation as in Theorem 9.10.
We can then finally turn to the non-linear equation in Chapter 10 and use the linear estimates we obtained to construct function spaces consisting of the intersection of local L^p-spaces and the global homogeneous Lipschitz space in time and space. The special shape of our non-linearity helps us to operate an analytic fixed point argument in the spirit of [KL12] in this function spaces to gain Theorem 10.3, providing existence as well as temporal and tangential analyticity of perturbations with small initial Lipschitz norm. Thanks to [Koc99], the smoothness of the perturbation follows from the resulting bound on the global Lipschitz norm that is implied by our special choice of spaces. This perturbation result is new, and a reformulation provides stability of solutions of the transformed pressure equation that are initially close to the travelling wave w_{tw} in the sense of homogeneous Lipschitz spaces. Moreover, there are precise estimates of derivatives. Note that w_{tw} is continuous down to the initial time s_1 and we have $w_{tw}(s_1, y) = y_n$ and thus $\nabla_y w_{tw}(s_1, \cdot) = \vec{e}_n$.

1.11 Theorem *Let $\sigma > -1$, $s_1 > -\infty$ and $I = (s_1, s_2) \subset \mathbb{R}$ be an open interval.*
Then there exists an $\varepsilon > 0$ such that for any $g : \overline{H} \to \mathbb{R}$ satisfying $\|\nabla_y g - \vec{e}_n\|_{L^\infty(H)} < \varepsilon$ we can find a solution w_* to the transformed pressure equation on $I \times \overline{H}$ with initial value g for which we have

$w_* \in C^\infty(I \times \overline{H})$ and

$$\sup_{(s,y)\in I\times H} |\nabla_y w_*(s,y) - \vec{e}_n| \le c\,\varepsilon$$

for a constant $c = c(n,\sigma) > 0$. Furthermore, w_* is analytic in the temporal and tangential directions on $I \times \overline{H}$ with an $R > 0$ and a $C = C(n) > 0$ such that

$$\sup_{(s,y)\in I\times H} (s-s_1)^{k+|\alpha'|}\, |\partial_s^k \partial_{y'}^{\alpha'} \nabla_y w_*(s,y)| \le C\, R^{-k-|\alpha'|}\, k!\, \alpha'!\, \varepsilon$$

for any $k \in \mathbb{N}_0$ and $\alpha' \in \mathbb{N}_0^{n-1}$ with $k + |\alpha'| > 0$.

We now set

$$T : I \times \overline{H} \ni (s,y) \mapsto (s, y', w_*(s,y)) =: (t,x),$$

interchanging dependent and independent variables. Choosing ε even smaller so that $c\,\varepsilon < 1$, from the global bound on the gradient of w_* it follows immediately that

$$(1-c\,\varepsilon)|(s,y)-(\bar{s},\bar{y})| < |T(s,y) - T(\bar{s},\bar{y})| < (1-c\,\varepsilon)^{-1}|(s,y)-(\bar{s},\bar{y})|$$

for any (s,y), $(\bar{s},\bar{y}) \in (I \times \overline{H})$. This means that T is injective and a quasi-isometry, thus allowing us to reparametrise the graph of w_* globally via T, reversing the local process that was applied above to motivate the consideration of the transformed pressure equation in the first place. The smooth function whose graph is given in terms of t and x is called v_*, and we get $T(\bar{I} \times \overline{H}) = \overline{\mathcal{P}}(v_*)$. Next we perform the rescaling of time that inverts the one given above without renaming the time variable t, doublebinding the notation here also with respect to the transformed and rescaled time interval that we name $(t_1, t_2) =: I$ again. The same calculations as above then show that v_* is a classical solution of the pressure equation pointwise on $\overline{\mathcal{P}}(v_*)$ up to and including the boundary. Since the level set of v_* at height z is given by

$$\left\{(s,y) \mid y_n = w_*(s,y',z)\right\},$$

the temporal and tangential analyticity translates into analyticity of the level sets of v^*. Note that for these transformations to hold it is necessary that the perturbation has small homogeneous Lipschitz norm. Thus in this sense the smallness condition on the initial perturbation is optimal. The repetition of this process with w_{tw} instead of w_* generates the travelling wave solution

$$v_{tw}(t,x) = (x_n + (t - t_1))$$

of the pressure equation on $\overline{\mathcal{P}}(v_{tw})$.

Finally, consider $u_*(t,x) := \left(\frac{m-1}{m} v_*(t,x)\right)^{\frac{1}{m-1}}$ for any $(t,x) \in \overline{\mathcal{P}}(v_*)$ with $u_*(t,x) := 0$ whenever $(t,x) \in \overline{\mathcal{P}}(v_*)^c$ to generate a classical solution of the PME on $\mathcal{P}(u_*) = \mathcal{P}(v_*)$ that is a weak solution of the PME on $I \times \mathbb{R}^n$ in the sense of Definition 1.1. To see the latter we simply compute

$$\int_{I\times\mathbb{R}^n} u_* \partial_t \varphi \, d\mathcal{L}^{n+1} + \int_{I\times\mathbb{R}^n} u_*^m \Delta_x \varphi \, d\mathcal{L}^{n+1} = - \int_{\mathcal{P}(u_*)} \partial_t u_* \varphi \, d\mathcal{L}^{n+1} + \int_{\mathcal{P}(u_*)} \Delta_x u_*^m \varphi \, d\mathcal{L}^{n+1} = 0$$

for any $\varphi \in C_c^\infty(I \times \mathbb{R}^n)$, where the boundary terms of the integrations by parts vanish since u_* vanishes at the boundary of its positivity set and we have

$$\nabla_x u_*^m = \left(\frac{m-1}{m}\right)^{\frac{m}{m-1}} \nabla_x v_*^{\frac{m}{m-1}} = u_* \nabla_x v_*.$$

In conjunction with the existence and uniqueness results 1.3, 1.5 and 1.6, this shows the stability of solutions of the PME whose pressure is initially close to a flat front as well as regularity for the solution and its interface, thus establishing the main result of the present work.

1.12 Theorem *Let $I = (t_1, t_2) \subset \mathbb{R}$ be open with $-\infty < t_1 < t_2 \leq \infty$ and $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ be such that $\left\| \frac{m}{m-1} \nabla_x g^{m-1} - \vec{e}_n \right\|_{L^\infty(\mathcal{P}(g))} < \varepsilon$ for an $\varepsilon > 0$ that is sufficiently small.*
If u is the solution of the PME on $I \times \mathbb{R}^n$ with initial value g, then we have $u^{m-1} \in C^\infty(\overline{\mathcal{P}}(u))$,

$$\sup_{(t,x) \in \mathcal{P}(u)} \left| \frac{m}{m-1} \nabla_x u^{m-1}(t,x) - \vec{e}_n \right| < c\,\varepsilon$$

with a constant $c = c(n,m)$, and all the level sets of u are analytic.

The level set of level 0 is nothing but $\mathcal{G}(u)$. The proof of the conjecture formulated in Remark 10.4, drawing also on the analyticity result from [Koc99], would ensure that not only the interface of u, but indeed its pressure u^{m-1} is analytic.

There are at least two questions that are interesting to pursue from this point onwards. On the one hand it should be possible to gain a local existence result as in [Koc99, Theorem 5.5.1.] with less restrictive assumptions on the initial data. This would improve short time regularity by weakening the prerequisits for the initial datum $g \geq 0$ with bounded positivity set to be only $g^{m-1} \in C^1(\overline{\mathcal{P}}(g))$ with $g^{m-1} + |\nabla_x(g^{m-1})| \geq c > 0$ on $\mathcal{P}(g)$.
On the other hand, it is known from other equations that flatness implies regularity. The most prominent example is possibly the so-called Stefan problem ([Mei92]) that has had a development somewhat parallel to the PME. There are several possibilities to make the intuitive notion of flatness mathematically concrete. In the literature there exist measure theoretic approaches as in [Caf88] or differently in [Caf77]. With the definition from [Caf89], it was shown in [ACS98] that flat weak solutions of the Stefan problem are classical and hence smooth ([Koc98]). Inspired by this result, for the PME we would propose to define flatness of a solution by trapping its pressure between two travelling wave solutions of the pressure equation. The Lipschitz framework of our Theorem 1.12 should then make it possible to show that flatness also implies smoothness of the solutions of the PME at least locally. This requires to generalise the proof of the pointwise estimate in Theorem 6.10 to the case of measurable coefficients and therefrom derive both an upper and a lower Gaussian estimate in this context. These Gaussian statements are already contained in [Koc99]. However, for the PME it is known that any solution to initial data with bounded positivity set will indeed always get caught between two flat fronts eventually, therefore becoming flat. This would then prove large time regularity for any solution of the PME that evolves from an initial datum with bounded positivity set.

2 Preliminaries

In this chapter we fix the notation and present some results needed later in places scattered all over this work. Some general background information can be found in the appendix.

Basic Notations and Conventions

Constants are denoted by c and C, and we write for example $c = c(a, b)$ if and only if the constant c depends only on the parameters a and b. In chains of inequalities, the value of c can differ in every step without a change in the notation. To avoid confusion, we therefore mostly use the notation $x \lesssim_{a,b} y$ instead, meaning that there is a constant depending only on a and b – for now called $c(a, b)$, but usually not naming it explicitely – such that the inequality $x \leq c(a, b)\, y$ holds. The mere symbol $\lesssim$ itself ensures that the constant does not contain any parameters. In every step of a chain of inequalities, exactly the parameters entering in each step are indicated by this notation, while the conclusion contains all dependencies. In an example once more we could perhaps have that $x \lesssim_a y \lesssim_b z$, implying $x \lesssim_{a,b} z$. The reverted symbol is $\gtrsim_{a,b}$, and $x \lesssim_a y \lesssim_b x$ is abbreviated to $x \eqsim_{a,b} y$.
We say that a quantity x is positive if and only if $x \geq 0$. For $x > 0$ the terminus strictly positive is reserved. The same remark applies to negative and strictly negative terms as well as to monotone functions: A function f is monotonically increasing (decreasing) if and only if $f(x_1) \leq f(x_2)$ $(f(x_1) \geq f(x_2))$ for any $x_1 < x_2$, and strictly monotonically increasing (decreasing) if and only if this inequality is strict.
The Kronecker symbol is always written as δ_{ij}.
Throughout this work, $n \in \mathbb{N}$ denotes the (spatial) dimension and we set $\mathbb{R}^1 = \mathbb{R}$. We consider subsets ω of time-space $\mathbb{R} \times \mathbb{R}^n$. Mostly ω is of a product structure itself, that is we have $\omega = I \times \Omega$ for a time interval $I \subset \mathbb{R}$ and a connected subset Ω of space $\mathbb{R}^n$. The underlying topology is always taken to be the usual one, alongside with the standard notations $\overline{I}$ and $\overline{\Omega}$ for the closures of I and Ω in $\mathbb{R}$ or in $\mathbb{R}^n$, and $\mathring{I}$ and $\mathring{\Omega}$ for the interiors of I and Ω. On subsets, the induced topology is considered and a subset of a subspace that is open or closed with respect to the induced topology will also be called relatively open or relatively closed, respectively. The compact subsets of $\Omega \subset \mathbb{R}^n$ with respect to the induced topology on Ω are given exactly by the compact subsets of $\mathbb{R}^n$ that are contained in Ω. Thus compact subsets of an open set Ω have a positive distance to $\partial\Omega$, whereas compact subsets of $\overline{\Omega}$ can touch the boundary.
Intervals are given by their end points $t_1 \leq t_2$, where $\pm\infty$ is admissible for both end points. We denote $(t_1, t_2) := \{t \in \mathbb{R} \mid t_1 < t < t_2\}$ and $[t_1, t_2] := \{t \in \mathbb{R} \mid t_1 \leq t \leq t_2\}$, with the obvious alterations in the definitions of $[t_1, t_2)$ and $(t_1, t_2]$. It is clear that (t_1, t_2) is an open set and $[t_1, t_2]$ a closed one with $\overline{(t_1, t_2)} = [t_1, t_2]$. Moreover, we introduce the notation ${}^{\ulcorner}I$ and $I^{\urcorner}$ to mean the closure of an interval I only at its left or right end point, respectively. We always exclude $\pm\infty$ from being an element of the (time) interval and therefore set by convention $[-\infty, t_2] := (-\infty, t_2]$, and likewise for the other possible cases. Therefore $t \in \overline{I}$ is always a finite point of time, regardless of the interval I being bounded or not. On the other hand, the interval $I = \mathbb{R} = \overline{\mathbb{R}}$ is included in and consistent with this notation, since it does not contain an infinite time and is both closed and open.
Another conventional setting we use is $\infty \pm h = \infty$ and $-\infty \pm h = -\infty$ for any $h \in \mathbb{R}$. For an arbitrary open interval $I = (t_1, t_2)$ and $h > 0$ we denote $I^h := (t_1, t_2 - h)$ and $I^{-h} := (t_1 + h, t_2)$.

By the convention we then have $I^{-h} = I$ if $t_1 = -\infty$ and $I^h = I$ if $t_2 = \infty$. For arbitrary $h \in \mathbb{R}$ it is obvious that $I^h + h = I^{-h}$.
We define the euclidean open ball with radius $r > 0$ centred at $x_0 \in \mathbb{R}^n$ as

$$B_r^{eu}(x_0) := \{x \in \mathbb{R}^n \mid d^{eu}(x, x_0) < r\},$$

where d^{eu} indicates the euclidean metric on $\mathbb{R}^n$. The closure of the ball is then given by

$$\overline{B}_r^{eu}(x_0) := \{x \in \mathbb{R}^n \mid d^{eu}(x, x_0) \leq r\}.$$

We denote the cube centred at $x_0 \in \mathbb{R}^n$ with side length $2\,r$ by

$$C_r(x_0) := \left\{ x \in \mathbb{R}^n \mid \max_{j=1,\ldots,n} |x_{0,j} - x_j| < r \right\},$$

which can also be written as the cartesian product of the intervals $(x_{0,j} - r, x_{0,j} + r)$ ranging over $j = 1, \ldots, n$. Obviously we have

$$C_{\frac{r}{\sqrt{n}}}(x_0) \subset B_r^{eu}(x_0) \subset C_r(x_0).$$

Other special subsets of $\mathbb{R}^n$ we consider include the open x_n-strip

$$H_a^b := \{x \in \mathbb{R}^n \mid a < x_n < b\}$$

and its closure

$$\overline{H}_a^b = \{a \leq x_n \leq b\}$$

for some numbers $a \leq b$. Here the same conventions as for time intervals prevent the inclusion of infinity into the set. For $a = 0$ and $b = \infty$ we drop the sub- and superscripts and obtain the upper half plane H, an open and unbounded subset of $\mathbb{R}^n$. Its boundary is the same as that of its closure $\overline{H}$, namely

$$\partial H = \partial \overline{H} = \{x \in \mathbb{R}^n \mid x_n = 0\} =: \{x_n = 0\}.$$

As a finite dimensional vector spaces we always equip $\mathbb{R}^n$ with the canonical basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$. Any $x \in \mathbb{R}^n$ is then given as $x = (x_1, \ldots, x_n)$. When considering a particular direction of $\mathbb{R}^n$ seperately, say the j-th direction for a $j \in \{1, \ldots, n\}$, we sometimes abbreviate the remaining $n-1$ coordinates as $x' := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ and write $x = (x', x_j)$ regardless of the exact position of x_j. The inner product in $\mathbb{R}^n$ of two elements x, $y \in \mathbb{R}^n$ is denoted by

$$x \cdot y := \sum_{j=1}^{n} x_j y_j.$$

We also introduce the function $(\,\cdot\,)_n^l : \overline{H} \ni x \mapsto x_n^l \in \mathbb{R}$ for $l \geq 0$.

The canonical measure on $\mathbb{R}^n$ is the Lebesgue measure $\mathcal{L}^n$, and we drop the exponent 1 to write $\mathcal{L}$ on $\mathbb{R}$. We also drop the Lebesgue measure itself in the notation of the Lebesgue spaces and set $L^p(\Omega, \mathcal{L}^n) =: L^p(\Omega)$ and $L^p_{loc}(\Omega, \mathcal{L}^n) =: L^p_{loc}(\Omega)$ for any $1 \leq p \leq \infty$.

For a function $u \in L^1_{loc}(\Omega)$, we regard the vanishing set $\mathcal{V}(u)$ as the union of all relatively open subsets of Ω on which $u = 0$ holds $\mathcal{L}^n$-almost everywhere. The support of u is then defined as $\operatorname{supp} u := \Omega \setminus \mathcal{V}(u)$, a relatively closed subset of Ω. If u is continuous, the support can be characterised as

$$\operatorname{supp} u = \overline{\{x \in \Omega \mid u(x) \neq 0\}} \cap \Omega.$$

This is of course again a relatively closed subset of Ω.
If $u \geq 0$ we also consider the positivity set $\mathcal{P}(u) := \{x \in \Omega \mid u(x) > 0\}$. For continuous u it is an open set, and we have $\operatorname{supp} u = \overline{\mathcal{P}(u)} \cap \Omega$.
Functions u on time-space sets that are evaluated at a time t remain functions on the spatial part of the underlying set, but we often supress this in the notation and write only $u(t, \cdot) = u(t)$.
The partial derivative of a function u in the x_j-direction is denoted by $\partial_{x_j} u$, and in general we use the notation

$$\partial_x^\alpha u = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \dots \partial_{x_n}^{\alpha_n} u$$

with a multi-index $\alpha \in \mathbb{N}_0^n$. The length of such α is given by $|\alpha| = \sum_{j=1}^{n} \alpha_j$. The collection of all derivatives of order m is $D_x^m u := \{\partial_x^\alpha u \mid |\alpha| = m\}$. Especially for $m = 1$, we often view this set as an ordered n-tuple or vector and write $\nabla_x u = (\partial_{x_1} u, \dots, \partial_{x_n} u)$. By an abuse of notation, we often ignore the set character of $D_x^m u$ and use this symbol to express a fact for every single element of it, as for example in $D_x^m u \in L^p(\Omega)$, meaning that $\partial_x^\alpha u \in L^p(\Omega)$ for any multi-index α of length m.
For the converse case of singleing out arbitrary elements of derivatives of a certain order withouth specifying excatly which, we introduce the symbolic notation $D_x^{m_1} u \star D_x^{m_2} u$ to denote any linear combination of products of derivatives of orders m_1 and m_2. For example, we have

$$\partial_{x_n} u \, \Delta_x u = \nabla_x u \star D_x^2 u$$

as well as

$$\partial_{x_n}^2 u \sum_{j=1}^{n} \partial_{x_j} u = \nabla_x u \star D_x^2 u.$$

We use $1 \star D_x^m u$ to mean a linear combination of derivatives of order m only, that is for instance

$$\partial_{x_n} u = 1 \star \nabla_x u$$

and

$$\sum_{j=1}^{n} \partial_{x_j} u = 1 \star \nabla_x u,$$

and in the same spirit allow terms like $v \star D_x^m u$ for another function v. The iterated application of $\star$ onto the same order of derivatives, as in $D_x^m u \star \dots \star D_x^m u$ – j times – is abbreviated by $(D_x^m u)^{j\star}$ with the usual conventions $(D_x^m u)^{1\star} = 1 \star D_x^m u$ and $(D_x^m u)^{0\star} = 1$.
Furthermore, we set

$$|D_x^m u| := \sqrt{\sum_{|\alpha|=m} |\partial_x^\alpha u|^2} \eqsim_n \sum_{|\alpha|=m} |\partial_x^\alpha u|.$$

For $m = 1$ the notation of the inner product in $\mathbb{R}^n$ can be applied to give $|D_x^1 u|^2 = |\nabla_x u|^2 = \nabla_x u \cdot \nabla_x u$.
We also use the common abbreviation

$$\nabla_x \cdot \vec{u} := \sum_{j=1}^{n} \partial_{x_j} u_j$$

for an $\mathbb{R}^n$-valued function $\vec{u} = (u_1, \ldots, u_n)$, and especially

$$\Delta_x u := \nabla_x \cdot \nabla_x u = \sum_{j=1}^{n} \partial_{x_j}^2 u.$$

Weighted Measures and Lebesgue Spaces

The weight function $x \mapsto |x_n|^\sigma$ is in $L^1_{loc}(\mathbb{R}^n)$ if and only if $\sigma > -1$. For any such σ we can therefore define the weighted Lebesgue measure $\mu_\sigma(x) := |x_n|^\sigma \, d\mathcal{L}^n(x)$ on $\mathbb{R}^n$, that is μ_σ is given by

$$\mu_\sigma(\Omega) = \int_\Omega |x_n|^\sigma \, d\mathcal{L}^n(x) \text{ for any } \mathcal{L}^n - \text{measurable } \Omega \subset \mathbb{R}^n.$$

By definition the μ_σ-measurable sets contain the $\mathcal{L}^n$-measurable ones, and especially all Borel sets. Of course the special case $\sigma = 0$ gives us back the Lebesgue measure $\mathcal{L}^n = \mu_0$. Instead of $\mu_\sigma(\Omega)$ we mostly write $|\Omega|_\sigma$, dropping the index for $\sigma = 0$. Note that any $\mathcal{L}^n$-nullset is also a μ_σ-nullset. The converse is also true, as one can see by a short computation. As a consequence we do not have to distinguish between $\mathcal{L}^n$-nullsets and μ_σ-nullsets, or the phrases $\mathcal{L}^n$-always everywhere and μ_σ-always everywhere, and the likes. Thus we merely write nullset, always everywhere, and so on.
It is well-known that $|B_r^{eu}(x_0)| \eqsim_n r^n$. As a first proposition we calculate the μ_σ-measure of a cube and a euclidean ball.

2.1 Proposition *Let $\sigma > -1$, $x_0 \in \mathbb{R}^n$ and $r > 0$.*

(i) If $|x_{0,n}| \geq r$, then $|C_r(x_0)|_\sigma \eqsim_{n,\sigma} r^{n-1} \left((|x_{0,n}| + r)^{1+\sigma} - (|x_{0,n}| - r)^{1+\sigma} \right)$.

(ii) If $|x_{0,n}| < r$, then $|C_r(x_0)|_\sigma \eqsim_{n,\sigma} r^{n-1} (|x_{0,n}| + r)^{1+\sigma}$.

(iii) $|B_r^{eu}(x_0)|_\sigma \eqsim_{n,\sigma} r^n (|x_{0,n}| + r)^\sigma$.

Proof: By definition we have

$$|C_r(x_0)|_\sigma = 2^{n-1} r^{n-1} \int_{(x_{0,n}-r, x_{0,n}+r)} |x_n|^\sigma \, d\mathcal{L}(x_n).$$

For $x_{0,n} \geq r > 0$ the fundamental theorem of calculus then shows that

$$\begin{aligned} |C_r(x_0)|_\sigma &= \frac{2^{n-1}}{1+\sigma} r^{n-1} \left((x_{0,n} + r)^{1+\sigma} - (x_{0,n} - r)^{1+\sigma} \right) \\ &\eqsim_{n,\sigma} r^{n-1} \left((|x_{0,n}| + r)^{1+\sigma} - (|x_{0,n}| - r)^{1+\sigma} \right). \end{aligned}$$

This calculation is possible since $\sigma > -1$, and likewise we get for $x_{0,n} \leq -r < 0$ that

$$|C_r(x_0)|_\sigma = -\frac{2^{n-1}}{1+\sigma}\, r^{n-1} \left((-x_{0,n} - r)^{1+\sigma} - (-x_{0,n} + r)^{1+\sigma} \right)$$
$$\eqsim_{n,\sigma} r^{n-1} \left((|x_{0,n}| + r)^{1+\sigma} - (|x_{0,n}| - r)^{1+\sigma} \right).$$

We now consider $-r < x_{0,n} < r$. The same direct calculation as above leads to

$$|C_r(x_0)|_\sigma \eqsim_{n,\sigma} r^{n-1} \left((-x_{0,n} + r)^{1+\sigma} + (x_{0,n} + r)^{1+\sigma} \right).$$

If $x_{0,n} \geq 0$, the second term is bigger and we get

$$|C_r(x_0)|_\sigma \lesssim_{n,\sigma} r^{n-1}(x_{0,n} + r)^{1+\sigma} = r^{n-1}(|x_{0,n}| + r)^{1+\sigma}.$$

A similar calculation is valid for $x_{0,n} < 0$, where the first term is bigger, and the last inequality therefore holds in the case $-r < x_{0,n} < r$.
For the opposite inequality in this case we note that for $x_{0,n} \geq 0$ we also have

$$|C_r(x_0)|_\sigma \gtrsim_n r^{n-1} \int\limits_{(0,x_{0,n}+r)} x_n^\sigma \, d\mathcal{L}(x_n) \eqsim_\sigma r^{n-1}(x_{0,n} + r)^{1+\sigma}$$

and the same inequality holds for $x_{0,n} < 0$ with $-x_{0,n}$ instead of $x_{0,n}$ on the right hand side. This finishes the proof for the measure of the cube.
We turn to the ball and first consider the case $r < \frac{1}{2}\,|x_{0,n}|$. For any $x \in B_r^{eu}(x_0)$ we then see that $\frac{1}{2}\,|x_{0,n}| < |x_n| < \frac{3}{2}\,|x_{0,n}|$ and can conclude that

$$|B_r^{eu}(x_0)|_\sigma \eqsim_\sigma |x_{0,n}|^\sigma \,|B_r^{eu}(x_0)| \eqsim_n |x_{0,n}|^\sigma \, r^n.$$

For $\sigma \geq 0$ we have $|x_{0,n}|^\sigma < (|x_{0,n}| + r)^\sigma$ and, now using that we are in the case $r < \frac{1}{2}|x_{0,n}|$, also

$$|x_{0,n}|^\sigma = (\frac{1}{2}\,|x_{0,n}| + \frac{1}{2}\,|x_{0,n}|)^\sigma \gtrsim_\sigma (|x_{0,n}| + r)^\sigma.$$

Reiterating the computation for $-1 < \sigma < 0$ also generates $|x_{0,n}|^\sigma \eqsim_\sigma (|x_{0,n}| + r)^\sigma$ and thus $|B_r^{eu}(x_0)|_\sigma \eqsim_{n,\sigma} r^n\,(|x_{0,n}| + r)^\sigma$ if $r < \frac{1}{2}\,|x_{0,n}|$.
On the other hand it is immediate that $r \geq \frac{1}{2}\,|x_{0,n}|$ implies

$$|B_r^{eu}(x_0)|_\sigma \leq |C_r(x_0)|_\sigma \lesssim_{n,\sigma} r^n\,(|x_{0,n}| + r)^\sigma$$

by an application of the results for the cube. To show the corresponding estimate from below we divide the case into two subcases. If $r > \sqrt{n}\,|x_{0,n}|$ we can use the second formula for the measure of the cube with radius $\frac{r}{\sqrt{n}}$ and obtain

$$|B_r^{eu}(x_0)|_\sigma \geq |C_{\frac{r}{\sqrt{n}}}(x_0)|_\sigma \eqsim_{n,\sigma} r^{n-1}\,(|x_{0,n}| + r)^{1+\sigma}.$$

The trivial inequality $(|x_{0,n}| + r) > r$ finishes this case. Finally, if $\frac{1}{2}\,|x_{0,n}| \leq r \leq \sqrt{n}\,|x_{0,n}|$, the first formula for the measure of the cube with radius $\frac{r}{\sqrt{n}}$ is applicable and reveals that

$$\begin{aligned}|B_r^{eu}(x_0)|_\sigma \geq |C_{\frac{r}{\sqrt{n}}}(x_0)|_\sigma &\eqsim_{n,\sigma} r^{n-1}\left((|x_{0,n}| + \frac{r}{\sqrt{n}})^{1+\sigma} - (|x_{0,n}| - \frac{r}{\sqrt{n}})^{1+\sigma}\right) \\ &\geq r^{n-1}\,|x_{0,n}|^{1+\sigma}\left((1+\frac{1}{2\sqrt{n}})^{1+\sigma} - (1-\frac{1}{2\sqrt{n}})^{1+\sigma}\right) \eqsim_{n,\sigma} r^{n-1}\,|x_{0,n}|^{1+\sigma},\end{aligned}$$

where we used the left hand side bound for r that is available in this special subcase in both summands. The right hand side inequality for r then yields the lower bound $r^n\,|x_{0,n}|^\sigma$ with an additional dependence on n in the constant. If $\sigma \geq 0$ we use once more that

$$|x_{0,n}| = \frac{1}{2}\,|x_{0,n}| + \frac{1}{2}\,|x_{0,n}| \geq \frac{1}{2}\,|x_{0,n}| + \frac{1}{2\sqrt{n}}\,r,$$

while for $-1 < \sigma < 0$ the trivial inequality allows us to add a positive term to $|x_{0,n}|^\sigma$ without making it bigger. ■

From a theoretical point of view, Proposition 2.1 also shows that μ_σ is not only a Borel measure, but also a Radon measure (see definition A.10). Furthermore, μ_σ has the very useful property of being countably finite on $\mathbb{R}^n$ as defined in Definition A.2.

As an abbreviation for the Lebesgue spaces with respect to μ_σ we set $L^p(\Omega,\mu_\sigma) =: L^p_\sigma(\Omega)$ for $1 \leq p \leq \infty$. Since μ_σ is a Borel measure, also the local Lebesgue spaces $L^p_{loc}(\Omega,\mu_\sigma)$ are defined for any $1 \leq p \leq \infty$. By definition we have $L^p_\sigma(\Omega) \subset L^p_{loc}(\Omega,\mu_\sigma)$ for any $1 \leq p \leq \infty$. Moreover, the fact that μ_σ is also Radon ensures that $L^p_{loc}(\Omega,\mu_\sigma) \subset L^1_{loc}(\Omega,\mu_\sigma)$ for any $1 \leq p \leq \infty$. Note also that for arbitrary Ω we have $\mathring{\Omega} \subset \overline{\Omega}$ and therefore the collection of compact subsets of $\overline{\Omega}$ contains at least all sets that are compact subsets of $\mathring{\Omega}$. Consequently we get $L^p_{loc}(\overline{\Omega},\mu_\sigma) \subset L^p_{loc}(\mathring{\Omega},\mu_\sigma)$ for any $1 \leq p \leq \infty$.
For a set Ω that is open or closed we know that the boundary is a nullset. It is therefore clear that $L^p_\sigma(\mathring{\Omega}) = L^p_\sigma(\overline{\Omega})$ for such Ω.
Now consider an arbitrary set $\Omega \subset H$. Since any compact set contained in the open upper half plain stays away from $\{x_n = 0\}$ it follows that on any compact $M \subset \Omega$ we know that both $\sup\limits_{x\in M} x_n^{-\sigma}$ and $\sup\limits_{x\in M} x_n^{\sigma}$ are finite constants. This shows that for Ω contained in the upper half plane we have $L^p_{loc}(\Omega,\mu_\sigma) = L^p_{loc}(\Omega)$ for any $1 \leq p \leq \infty$. In conjunction with the inclusions above this means that on an open set Ω that is contained in H we have

$$L^p_\sigma(\Omega) \subset L^p_{loc}(\overline{\Omega},\mu_\sigma) \subset L^p_{loc}(\Omega,\mu_\sigma) \subset L^1_{loc}(\Omega,\mu_\sigma) \subset L^1_{loc}(\Omega)$$

for any $1 \leq p \leq \infty$.
The countable finiteness of μ_σ ensures that $L^\infty(\Omega)$ is the dual space to $L^1_\sigma(\Omega)$ on any $\Omega \subset \mathbb{R}^n$. As always in the context of dual Lebesgue spaces we use the conventions $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$ when these cases arise.
On time-space-sets $I \times \Omega$ we denote the product measure by $\mathcal{L} \times \mu_\sigma$. With $L^q(I; L^p_\sigma(\Omega))$ we mean the space of $L^p_\sigma(\Omega)$-valued functions whose q-th power is $\mathcal{L}$-integrable on I in the sense of Bochner ([Yos68]). It is obvious that $L^p(I; L^p_\sigma(\Omega)) = L^p(I \times \Omega, \mathcal{L} \times \mu_\sigma)$ with a suitable identification of the objects involved.

Spaces of Continuous Functions

As usual we denote

$$C(\Omega) := \{u : \Omega \longrightarrow \mathbb{R} \mid u \text{ continuous}\}$$

and

$$C_c(\Omega) := \{u \in C(\Omega) \mid \operatorname{supp} u \text{ is compact subset of } \Omega\}.$$

Instead of adopting the intrinsic point of view one could also stress the subset nature of Ω by considering

$$\check{C}(\Omega) := \{u : \Omega \longrightarrow \mathbb{R} \mid u \text{ has extension } \check{u} \in C(\mathbb{R}^n)\} \subset C(\Omega).$$

For compactly supported continuous functions on a subset this approach yields the characterisation

$$C_c(\Omega) = \left\{u \in \check{C}(\Omega) \mid \operatorname{supp} u \text{ is compact subset of } \Omega\right\}$$

proven for example by means of Tietze's extension theorem ([Mun00]). The extension $\check{u}$ of $u \in C_c(\Omega)$ can indeed be chosen to have compact support again, so we have

$$C_c(\Omega) = \{u : \Omega \longrightarrow \mathbb{R} \mid \operatorname{supp} u \text{ is compact subset of } \Omega \text{ and } u \text{ has extension } \check{u} \in C_c(\mathbb{R}^n)\}.$$

By considering the extensions we may regard $C_c(\Omega)$ as a subspace of $C_c(\mathbb{R}^n)$ that increases with the underlying set.
Reversing the order of these operations on the function space $C(\Omega)$, we could also consider the space

$$\check{C}_c(\Omega) := \{u : \Omega \longrightarrow \mathbb{R} \mid u \text{ has extension } \check{u} \in C_c(\mathbb{R}^n)\} \subset C(\Omega).$$

This space differs from $C_c(\Omega)$ by lacking the condition on the support of the restricted function. Thus the support of functions in $\check{C}_c(\Omega)$ is always bounded, but does not have to be compact. We have $C_c(\Omega) \subset \check{C}_c(\Omega)$ with equality only for closed Ω.
Note that by considering functions in $C_c(\mathbb{R}^n)$ whose support is contained in Ω, where Ω is arbitrary again, we would get yet another space on Ω that in general does not coincide with any of the other two. Its elements are characterised by their compact support as functions on Ω and the fact that they can be extended by zero regardless of the topological properties of Ω. In contrast to that, only if Ω is open with $\partial\Omega \neq \emptyset$, it is clear that the elements of $C_c(\Omega)$ have a trivial extension to $\mathbb{R}^n$, and even vanish already at a distance from $\partial\Omega$. For general Ω, however, $u \in C_c(\Omega)$ can as well have non-zero values near and even at $\partial\Omega$. On the other hand, for any set Ω with boundary, even if it is open, the functions in $\check{C}_c(\Omega)$ do not have to be zero at $\partial\Omega$.
As a reminder we annotate that by means of Lusin's theorem ([Els07]) one sees that $C_c(\Omega)$ is a dense subset of $L^p_\sigma(\Omega)$ for any $\Omega \subset \mathbb{R}^n$ and any $1 \leq p < \infty$.
The collection of functions on an arbitrary Ω that are continuous and bounded is written as $C_b(\Omega)$. For vector space valued functions $u : I \longrightarrow L^p_\sigma(\Omega)$ that are continuous with respect to the topology induced by the norm on $L^p_\sigma(\Omega)$, the usual notation $C(I; L^p_\sigma(\Omega))$ is used, and $C_c(I; L^p_\sigma(\Omega))$, $C_b(I; L^p_\sigma(\Omega))$, and so on, have the obvious meaning.

Spaces of Differentiable Functions

In contrast to continuity, differentiability of a function can only be defined on open subsets Ω of $\mathbb{R}^n$. For such an open set and $m \in \mathbb{N}$ we denote

$$C^m(\Omega) := \{u : \Omega \to \mathbb{R} \mid u \text{ is } m \text{ times differentiable with } \partial_x^\alpha u \in C(\Omega) \text{ for all } |\alpha| \leq m\}$$

and

$$C_c^m(\Omega) := \{u \in C^m(\Omega) \mid \operatorname{supp} u \text{ is compact subset of } \Omega\}\,.$$

For general, not necessarily open subsets Ω we consistently define

$$C^m(\Omega) := \left\{u : \Omega \to \mathbb{R} \mid u \in C^m(\mathring{\Omega}),\, \partial_x^\alpha u \text{ has continuous extension to } \Omega \text{ for any } |\alpha| \leq m\right\}.$$

This is not the same as considering the functions restricted from the whole space as we did above. In fact, the space

$$\check{C}^m(\Omega) := \{u : \Omega \longrightarrow \mathbb{R} \mid u \text{ has extension } \check{u} \in C^m(\mathbb{R}^n)\}$$

equals $C^m(\Omega)$ if and only of Ω is closed. In view of this and the above characterisation of continuous functions with compact support we set

$$C_c^m(\Omega) := \{u : \Omega \to \mathbb{R} \mid \operatorname{supp} u \text{ is compact subset of } \Omega,\, u \text{ has extension } \check{u} \in C_c^m(\mathbb{R}^n)\}\,.$$

For open Ω this definition coincides with the one given above.
Additionally we consider the space

$$\check{C}_c^m(\Omega) := \{u : \Omega \to \mathbb{R} \mid u \text{ has extension } \check{u} \in C_c^m(\mathbb{R}^n)\}$$

for arbitrary $\Omega \subset \mathbb{R}^n$. As before we have $C_c^m(\Omega) = \check{C}_c^m(\Omega)$ for closed Ω and $C_c^m(\Omega) \subsetneq \check{C}_c^m(\Omega)$ for open $\Omega \subsetneq \mathbb{R}^n$. Setting $C_c^0(\Omega) := C_c(\Omega)$, with the same definition for the other spaces presented, is consistent. We have for example $C_c^m(\Omega) \subset C_c(\Omega)$ and all comments for functions in $C_c(\Omega)$ carry over to functions in $C_c^m(\Omega)$.
Finally, denote

$$C_c^\infty(\Omega) := \bigcap_{m \in \mathbb{N}} C_c^m(\Omega),$$

and likewise $\check{C}^\infty(\Omega)$, $\check{C}_c^\infty(\Omega)$, and $C^\infty(\Omega)$.
Let now $\Omega \subset \mathbb{R}^n$ be open. The space $C_c^\infty(\Omega)$ can then be equipped with seminorms to become the locally convex space $\mathcal{D}(\Omega)$. The elements of its dual are called distributions, accordingly denoted by $\mathcal{D}'(\Omega)$. Any distribution u has a derivative $\partial_x^\alpha u$ that is a distribution itself again. Although it is not clear that the product of two distributions is also a distribution, the product of a distribution with a smooth function yields a distribution again. For such products, the usual product rule of derivatives applies.
The Riesz representation theorem ([Wer05]) makes it possible to identify any function $u \in L_{loc}^1(\Omega)$ uniquely with a distribution. Conversely, the distributions on Ω that are given by locally Lebesgue integrable functions are called regular.

Weighted Sobolev Spaces

Concentrating on open subsets Ω of the upper halfplane H, the existence of a distributional derivative is guaranteed for any $u \in L^p_{loc}(\overline{\Omega}, \mu_\sigma)$ and hence for any $u \in L^p_\sigma(\Omega)$. This derivative is uniquely determined in Ω, albeit not on $\partial\Omega$.
It is not clear at all that a distribution that is given by the derivative of a regular distribution is regular itself. The notion of Sobolev spaces is designed for those that are. Here we allow different weights for every order of derivatives and define the weighted Sobolev spaces on an open set $\Omega \subset H$ as

$$W^{m,p}_{\vec{\sigma}}(\Omega) := \left\{ u \in L^p_{\sigma_0}(\Omega) \mid \partial^\alpha_x u \in L^p_{\sigma_{|\alpha|}}(\Omega) \text{ for all } |\alpha| \leq m \right\}$$

for $m \in \mathbb{N}$ and $1 \leq p < \infty$, understanding the weight exponents $\sigma_0, \ldots, \sigma_m > -1$ as the vector $(\sigma_0, \ldots, \sigma_m) = \vec{\sigma}$. If $\sigma_i = \sigma + i$ for a $\sigma > -1$ and $i = 0, \ldots, m$, we will also use the abbreviation $W^{m,p}_\sigma(\Omega)$. If on the other hand $\sigma_0 = \ldots = \sigma_m = 0$, that is all measures considered are the Lebesgue measure, we drop the $\vec{\sigma}$ entirely from the notation and merely write $W^{m,p}(\Omega)$. We also set $W^{0,p}_{\vec{\sigma}}(\Omega) := L^p_\sigma(\Omega)$. By an abuse of notation we include functions on time-space sets $I \times \Omega$ for an open interval $I \subset \mathbb{R}$ and set

$$W^{m,p}_{\vec{\sigma}}(I \times \Omega) := \left\{ u \in L^p(I \times \Omega, \mathcal{L} \times \mu_{\sigma_0}) \mid \partial^k_t \partial^\alpha_x u \in L^p(I \times \Omega, \mathcal{L} \times \mu_{\sigma_{|\alpha|}}) \text{ for all } k + |\alpha| \leq m \right\}.$$

Equipped with the norm

$$\|u\|_{W^{m,p}_{\vec{\sigma}}(\Omega)} := \left(\sum_{|\alpha| \leq m} \|\partial^\alpha_x u\|^p_{L^p_{\sigma_{|\alpha|}}(\Omega)} \right)^{\frac{1}{p}},$$

for $1 \leq p < \infty$ the weighted Sobolev spaces on the upper half plane are complete for any $m \in \mathbb{N}_0$, as is shown in the following proposition.

2.2 Proposition *Let $m \in \mathbb{N}_0$, $\sigma_0, \ldots, \sigma_m > -1$, $\Omega \subset H$ be open and $1 \leq p < \infty$.*
Then $W^{m,p}_{\vec{\sigma}}(\Omega)$ is a Banach space.

Proof: Let $(u_j)_{j \in \mathbb{N}}$ be a Cauchy sequence in $W^{m,p}_{\vec{\sigma}}(\Omega)$. Then $(\partial^\alpha_x u_j)_{j \in \mathbb{N}}$ is a Cauchy sequence in $L^p_{\sigma_{|\alpha|}}(\Omega)$ for any $0 \leq |\alpha| \leq m$. Since the Lebesgue spaces are complete for any $\sigma > -1$ and any $1 \leq p < \infty$, we know of the existence of $u_\alpha \in L^p_{\sigma_{|\alpha|}}(\Omega)$ with $\partial^\alpha_x u_j \to u_\alpha$ in $L^p_{\sigma_{|\alpha|}}(\Omega)$ for $j \to \infty$.
But $\partial^\alpha_x u_j$ as well as u_α can be seen as regular distributions for any $j \in \mathbb{N}$ and $0 \leq |\alpha| \leq m$. So for a $\varphi \in C^\infty_c(\Omega)$ we have

$$\begin{aligned} |\partial^\alpha_x u_j(\varphi) - u_\alpha(\varphi)| &\leq \int_\Omega |\partial^\alpha_x u_j(x) - u_\alpha(x)|\, x_n^{-\frac{\sigma_{|\alpha|}}{p}}\, |\varphi(x)|\, x_n^{\frac{\sigma_{|\alpha|}}{p}}\, d\mathcal{L}^n(x) \\ &\leq \left\|\partial^\alpha_x u_j - u_\alpha\right\|_{L^p_\sigma(\Omega)} \left\| (\,\cdot\,)_n^{-\frac{\sigma_{|\alpha|}}{p}} \varphi \right\|_{L^{\frac{p}{p-1}}(\Omega)} \end{aligned}$$

with Hölder's inequality. Since supp φ is compact and stays away from $\{x_n = 0\}$, the last term is bounded, while the first converges to 0. Thus $\partial^\alpha_x u_j \to u_\alpha$ in $\mathcal{D}'(\Omega)$ for any $0 \leq |\alpha| \leq m$. But then we see, again for $\varphi \in C^\infty_c(\Omega)$ and with the notation of distributions, that

$$u_\alpha(\varphi) = \lim_{j \to \infty} (-1)^{|\alpha|}\, u_j(\partial^\alpha_x \varphi) = (-1)^{|\alpha|}\, u_0(\partial^\alpha_x \varphi) = \partial^\alpha_x u_0(\varphi)$$

for any $0 \leq |\alpha| \leq m$. Hence we have $u_\alpha = \partial_x^\alpha u_0 \in L^p_{\sigma_{|\alpha|}}(\Omega)$ for any $0 \leq |\alpha| \leq m$ and thus $u_j \to u_0$ in $W^{m,p}_{\vec{\sigma}}(\Omega)$. ■

For $p = 2$ and $m \in \mathbb{N}_0$ the space $W^{m,2}_{\vec{\sigma}}(\Omega)$ is a Hilbert space with the obvious inner product

$$(u \mid v)_{W^{m,p}_{\sigma}(\Omega)} := \sum_{|\alpha| \leq m} \int_\Omega \partial_x^\alpha u \, \partial_x^\alpha v \, d\mu_{\sigma_{|\alpha|}}.$$

As an abbreviation we also write

$$(\nabla_x u \mid \nabla_x v)_{W^{m,p}_{\sigma}(\Omega)} := \sum_{j=1}^{n} (\partial_{x_j} u \mid \partial_{x_j} v)_{W^{m,p}_{\sigma}(\Omega)}.$$

A lot of the questions that arise naturally in weighted Sobolev spaces are also treated in [Kuf85]. We now prove that on the closed upper half space the norm of the highest derivative controls the Sobolev norm at least on a compact set if the weights are given adequately.

2.3 Proposition (Hardy Inequality) *Let $\sigma > -1$, $m \in \mathbb{N}$ and $1 \leq p < \infty$. If $\partial_{x_n}^m u \in L^p_{mp+\sigma}(H)$ and u has compact support contained in $\overline{H}$, then*

$$\|u\|_{L^p_\sigma(H)} \lesssim_{\sigma,m,p} \left\|\partial_{x_n}^m u\right\|_{L^p_{mp+\sigma}(H)}.$$

Proof: First assume $m = 1$ and $p = 1$. The assumptions then assure that for almost any $x' \in \mathbb{R}^{n-1}$ we have $\partial_{x_n} u(x', \cdot) \in L^1_{loc}((0,\infty))$, thus $u(x', \cdot) \in L^1_{loc}((0,\infty))$ and, moreover, the fundamental theorem of calculus in x_n-direction is applicable. This yields

$$u(x) = - \int_{(x_n,\infty)} \partial_{z_n} u(x', z_n) \, d\mathcal{L}(z_n)$$

thanks to the compact support of u. With Fubini's theorem we get

$$\|u\|_{L^1_\sigma(H)} \leq \int_{(0,\infty)} \left(\|\partial_{z_n} u(\cdot, z_n)\|_{L^1(\mathbb{R}^{n-1})} \int_{(0,z_n)} x_n^\sigma \, d\mathcal{L}(x_n) \right) d\mathcal{L}(z_n) \lesssim_\sigma \|\partial_{x_n} u\|_{L^1_{1+\sigma}(H)}.$$

Now consider $1 \leq p < \infty$. Here the prerequisits show that $\partial_{x_n} u(x', \cdot) \in L^p_{loc}((0,\infty))$ and consequently $u(x', \cdot) \in L^p_{loc}((0,\infty))$. Hölder's inequality then guarantees that

$$\partial_{x_n}(u^p)(x', \cdot) = p \, u^{p-1}(x', \cdot) \, \partial_{x_n} u(x', \cdot) \in L^1_{loc}((0,\infty))$$

and we can therefore apply the above result onto u^p to get

$$\begin{aligned} \|u\|^p_{L^p_\sigma(H)} &= \|u^p\|_{L^1_\sigma(H)} \lesssim_\sigma \|\partial_{x_n}(u^p)\|_{L^1_{1+\sigma}(H)} \eqsim_p \left\|u^{p-1} \partial_{x_n} u\right\|_{L^1_{1+\sigma}(H)} \\ &= \int_H x_n^{\sigma\frac{p-1}{p}} |u|^{p-1} x_n^{\frac{\sigma}{p}+1} |\partial_{x_n} u| \, d\mathcal{L}^n(x) \leq \|u\|^{p-1}_{L^p_\sigma(H)} \|\partial_{x_n} u\|_{L^p_{p+\sigma}(H)} \end{aligned}$$

using Hölder's inequality again.
Finally let $m \geq 1$. The inequality for $m = 1$ then implies that $\partial_{x_n}^{m-1} u \in L^p_{(m-1)p+\sigma}(H)$. An iteration of this argument finishes the proof. ■

Difference Quotient and Weak Regularisation

For an open (time) interval I let $u \in L^1_{loc}(I)$ and $h \in \mathbb{R}$, and define the (temporal) difference quotient by

$$D^h u(t) := \begin{cases} h^{-1}\left(u(t+h) - u(t)\right) & \text{for all } t \in I^h \\ 0 & \text{for all } t \in I \setminus I^h. \end{cases}$$

Obviously we, have $D^h u \in L^1_{loc}(I)$. Direct calculations show that

$$\begin{aligned} D^h(u_1\, u_2)(t) &= D^h u_1(t)\, u_2(t) + D^h u_2(t)\, u_1(t+h) \\ &= D^h u_1(t)\, u_2(t) + D^h u_2(t)\, u_1(t) + h\, D^h u_1(t) D^h u_2(t) \end{aligned}$$

holds, as well as

$$\int_I u\, D^h \varphi\, d\mathcal{L} = -\int_I D^{-h} u\, \varphi\, d\mathcal{L} \text{ for any } \varphi \in L^1_{loc}(I) \text{ with } \{\varphi \neq 0\} \subset I^h \cap I^{-h}.$$

Furthermore, we consider

$$u^h(t) := \begin{cases} h^{-1} \int\limits_{(0,h)} u(t+\tau)\, d\mathcal{L}(\tau) & \text{for all } t \in I^h \\ 0 & \text{for all } t \in I \setminus I^h. \end{cases}$$

2.4 Remark *A regularising effect of u^h is immediate, since clearly $u^h \in C(I^h)$ and also $u^h \in L^1_{loc}(I)$. But moreover, if $\partial_t u \in L^1_{loc}(I)$ it is also clear that $(\partial_t u)^h = D^h u$. Note that Fubini's theorem shows*

$$\int_I u \varphi^h\, d\mathcal{L} = \int_I u^{-h} \varphi\, d\mathcal{L} \textit{ for any } \varphi \in L^1_{loc}(I) \textit{ with } \{\varphi \neq 0\} \subset I^h \cap I^{-h}.$$

By a computation it can now be seen that for any $u \in L^1_{loc}(I)$ and h small enough we also have $\partial_t(u^h) = D^h u \in L^1_{loc}(I)$.

Convolution and Strong Regularisation

We fix $J : \mathbb{R}^n \to \mathbb{R} \in C_c^\infty(\mathbb{R}^n)$ with $\operatorname{supp} J \subset \overline{B}_1^{eu}(0)$ and $\int\limits_{\mathbb{R}^n} J\, d\mathcal{L}^n = 1$. Such a J exists, as we can choose for example

$$J(x) := \begin{cases} C\, e^{-\frac{1}{1-|x|^2}} & \text{for } x \in B_1^{eu}(0) \\ 0 & \text{else} \end{cases}$$

with a suitable normalisation constant C. For $\varepsilon > 0$, the function $J_\varepsilon(x) := \varepsilon^{-n} J(\varepsilon^{-1} x)$ is called a mollifier. It is clear that

$$J_\varepsilon \in C_c^\infty(\mathbb{R}^n),\ \operatorname{supp} J_\varepsilon = \overline{B}_\varepsilon^{eu}(0) \text{ and } \int_{\mathbb{R}^n} J_\varepsilon\, d\mathcal{L}^n = 1.$$

We now adjust the usual mollification on H in a way that is modelled to suit our case of weighted measures. Thus for $\Omega \subset H$ and $u \in L^1_{loc}(\Omega)$, in an abuse of notation we define

$$(J_{\varepsilon x_n} * u)(x) := \int_{\mathbb{R}^n} J_{\varepsilon x_n}(x-y) u(y)\, d\mathcal{L}^n(y),$$

where we always extend u by 0 outside of Ω. This is well-defined for any $x \in H$. Note that it is sufficient to consider the integral on the ball $B^{eu}_{\varepsilon x_n}(x)$ because of the special shape of supp J_ε. It is clear that $J_{\varepsilon x_n} * u \in C^\infty(\Omega)$ for any $\varepsilon > 0$, that is the adjusted mollification is indeed a strong regularisation.
A transformation of the integral easily shows that we have the characterisation

$$(J_{\varepsilon x_n} * u)(x) = \int_{\mathbb{R}^n} J_\varepsilon(y)\, u(x - x_n y)\, d\mathcal{L}^n(y) = \int_{B^{eu}_\varepsilon(0)} J_\varepsilon(y)\, u(x - x_n y)\, d\mathcal{L}^n(y).$$

If $\Omega \subset H$ is open and the derivatives of $u \in L^1_{loc}(\Omega)$ are regular distributions on Ω again, a calculation shows that for any open and bounded set $M \subset \Omega$ and any $x \in M$ we have

$$\begin{aligned}\partial^\alpha_x (J_{\varepsilon x_n} * u)(x) =& (J_{\varepsilon x_n} * \partial^\alpha_z u)(x) \\ &+ \sum_{j=1}^{\alpha_n} (-1)^j \binom{\alpha_n}{j} \sum_{|\beta|=l} \int_{B^{eu}_\varepsilon(0)} J_\varepsilon(y)\, \partial^\beta_z \partial^{\alpha'}_{z'} \partial^{\alpha_n - j}_{z_n} u(x - x_n y)\, y^\beta\, d\mathcal{L}^n(y)\end{aligned}$$

if εx_n is so small that $x - x_n y \in \overline{\Omega}$ for any $y \in B^{eu}_\varepsilon(0)$ and $x \in M$. The latter is satisfied for ε small enough if $\overline{M} \subset \Omega \cup \partial H$, that is if M keeps a distance to the boundary of Ω except at those parts of $\partial\Omega$ that are contained in $\{x_n = 0\}$. Compared to the usual mollification this is a slight gain in the expansion of M.
If $u \in C(\Omega)$, the normalised L^1-norm of J_ε enables us to calculate

$$|(J_{\varepsilon x_n} * u)(x) - u(x)| \leq \sup_{y \in B^{eu}_\varepsilon(0)} |u(x - x_n y) - u(x)|$$

which converges to 0 for $\varepsilon > 0$ pointwise and even uniformely on any set that is bounded and on which u is uniformely continuous.
That this mollification is well-adopted to our weighted measures is established in the following lemma.

2.5 Lemma *Let $\sigma > -1$, $\Omega \subset H$ and $1 \leq p < \infty$.*
If $u \in L^p_\sigma(\Omega)$, then $\|J_{\varepsilon x_n} * u\|_{L^p_\sigma(\Omega)} \lesssim_{\sigma,p} \|u\|_{L^p_\sigma(\Omega)}$ for any ε small enough.

Proof: For a fixed $i \in \mathbb{Z}$ consider the strip $H_i := H^{2^{i+1}}_{2^i}$. We also set $u_j := u\, \chi_{H_j}$ for any $j \in \mathbb{Z}$. We then have

$$\|J_{\varepsilon x_n} * u\|^p_{L^p_\sigma(H_i)} \leq \int_{H_i} \left(\int_{B^{eu}_{\varepsilon x_n}(x)} J_{\varepsilon x_n}(x-y) \sum_{j \in \mathbb{Z}} |u_j(y)|\, d\mathcal{L}^n(y) \right)^p d\mu_\sigma(x).$$

For any $x \in H_i$ and any $y \in B^{eu}_{\varepsilon x_n}(x)$ it is then clear that $u_j(y) = 0$ if $|j - i| > 2$ and $\varepsilon < \frac{1}{2}$. As a consequence we get

$$\|J_{\varepsilon x_n} * u\|^p_{L^p_\sigma(H_i)} \lesssim_p \sum_{j=i-1}^{i+1} \|J_{\varepsilon x_n} * |u_j|\|^p_{L^p_\sigma(H_i)}.$$

Hölder's inequality now shows that

$$\|J_{\varepsilon x_n} * |u_j|\|^p_{L^p_\sigma(H_i)} \leq \int\limits_{H_i} \left(\int\limits_{\mathbb{R}^n} J_{\varepsilon x_n}(x-y)|u_j(y)|^p \, d\mathcal{L}^n(y) \right) \left(\int\limits_{\mathbb{R}^n} J_{\varepsilon x_n}(x-y) \, d\mathcal{L}^n(y) \right)^{p-1} d\mu_\sigma(x).$$

The latter inner integral is 1 for any $x \in \mathbb{R}^n$ by the normalisation of J_ε. But on any strip H_i we have $x_n^\sigma \eqsim_\sigma 2^{i\sigma}$ and therefore with Fubini's theorem get

$$\begin{aligned}\|J_{\varepsilon x_n} * |u_j|\|^p_{L^p_\sigma(H_i)} &\lesssim_\sigma 2^{i\sigma} \int\limits_{H_j} |u_j(y)|^p \int\limits_{H_i} J_{\varepsilon x_n}(x-y) \, d\mathcal{L}^n(x) \, d\mathcal{L}^n(y) \\ &= 2^{i\sigma} \int\limits_{H_j} |u_j(y)|^p \int\limits_{H_i} \frac{x_n}{y_n} \, x_n^{-n-1} \, y_n \, J_\varepsilon(x_n^{-1}(x-y)) \, d\mathcal{L}^n(x) \, d\mathcal{L}^n(y).\end{aligned}$$

For $y \in H_j$ and $x \in H_i$ the quotient $\frac{x_n}{y_n}$ is bounded by 4. On $\mathbb{R}^n$ we can then transform the inner integral with $T : x \mapsto x_n^{-1}(x-y) =: z$. The Jacobi determinant of T is exactly $x_n^{-n-1} y_n$. The normalisation of J_ε therefore also implies that the inner integral is bounded by 1. This adds up to

$$\|J_{\varepsilon x_n} * |u_j|\|^p_{L^p_\sigma(H_i)} \lesssim_\sigma 2^{i\sigma} \|u_j\|^p_{L^p(H_j)}$$

and thus

$$\sum_{j=i-1}^{i+1} \|J_{\varepsilon x_n} * |u_j|\|^p_{L^p_\sigma(H_i)} \lesssim_\sigma \|u\|^p_{L^p_\sigma(H_i)}.$$

Then this is also an upper bound for $\|J_{\varepsilon x_n} * u\|^p_{L^p_\sigma(H_i)}$ if ε is small enough. The statement follows by summing over $i \in \mathbb{Z}$. ∎

2.6 Remark *With the usual mollification $J_\varepsilon * u$, the ε for which the overlap of H_i and $B^{eu}_\varepsilon(x)$ is finite depends on i: the crucial step in the argument then only works for $\varepsilon < 2^{i-1}$ and the statement of Lemma* 2.5 *thus only holds for u with support that stays away from $\{x_n = 0\}$.*
On the other hand, $J_\varepsilon * u$ is pointwise bounded by the euclidean Hardy-Littlewood maximal function. For a measure μ that is p-Muckenhoupt on the euclidean space, $1 < p < \infty$, Lemma 2.5 *then follows immediately on $L^p(\Omega, \mu)$ for any $\Omega \subset \mathbb{R}^n$. Details are discussed for example in [Kil94].*

The lemma is the key to prove L^p_σ-convergence of $J_{\varepsilon x_n} * u$ for general u.

2.7 Proposition *Let $\Omega \subset H$ and $1 \leq p < \infty$.*

(i) Let in addition $\sigma > -1$.
If $u \in L^p_\sigma(\Omega)$, then

$$\|J_{\varepsilon x_n} * u - u\|_{L^p_\sigma(\Omega)} \to 0 \quad (\varepsilon \to 0).$$

(ii) *Let in addition $m \in \mathbb{N}$ and $\sigma_0, \dots, \sigma_m > -1$.*
If Ω is open and $u \in W^{m,p}_{\vec{\sigma}}(\Omega)$, then

$$\|J_{\varepsilon x_n} * u - u\|_{W^{m,p}_{\vec{\sigma}}(M)} \to 0 \quad (\varepsilon \to 0)$$

for any open and bounded $M \subset \Omega$ with $\overline{M} \subset \Omega \cup \partial H$.

Proof: First observe that for $u \in C_c(\Omega) \subset L^p_\sigma(\Omega)$ the first part follows immediately from the uniform convergence of $J_{\varepsilon x_n} * u$, since μ_σ is a Radon measure.
Now let $u \in L^p_\sigma(\Omega)$, fix a $\delta > 0$ and let $\varphi \in C_c(\Omega)$ with $\|u - \varphi\|_{L^p_\sigma(\Omega)} < \delta$. Then

$$\begin{aligned}\|u - J_{\varepsilon x_n} * u\|_{L^p_\sigma(\Omega)} &\le \|u - \varphi\|_{L^p_\sigma(\Omega)} + \|\varphi - J_{\varepsilon x_n} * \varphi\|_{L^p_\sigma(\Omega)} + \|J_{\varepsilon x_n} * \varphi - J_{\varepsilon x_n} * u\|_{L^p_\sigma(\Omega)} \\ &\lesssim_{\sigma,p} \|u - \varphi\|_{L^p_\sigma(\Omega)} + \|\varphi - J_{\varepsilon x_n} * \varphi\|_{L^p_\sigma(\Omega)}\end{aligned}$$

thanks to Lemma 2.5. The choice of φ and the first observation in this proof then proves the first statement in general.
For the second part note that on M as in the assumptions and for ε small enough we have

$$|\partial^\alpha_x (J_{\varepsilon x_n} * u)(x) - (J_{\varepsilon x_n} * \partial^\alpha_z u)(x)| \lesssim_\alpha \varepsilon \sum_{|\gamma|=|\alpha|} (J_{\varepsilon x_n} * |\partial^\gamma_z u|)(x)$$

for any $1 \le |\alpha| \le m$. Another application of Lemma 2.5, this time on $J_{\varepsilon x_n} * |\partial^\gamma_z u|$, thus shows that

$$\|\partial^\alpha_x (J_{\varepsilon x_n} * u) - \partial^\alpha_x u\|_{L^p_{\sigma_{|\alpha|}}(\Omega)} \lesssim_p \|J_{\varepsilon x_n} * \partial^\alpha_z u - \partial^\alpha_x u\|_{L^p_{\sigma_{|\alpha|}}(\Omega)} + \varepsilon \sum_{|\gamma|=|\alpha|} \|\partial^\gamma_z u\|_{L^p_\sigma(\Omega)} .$$

Since $\partial^\gamma_x u \in L^p_{\sigma_{|\alpha|}}(\Omega) \cap L^1_{loc}(\Omega)$ by assumption for any $|\gamma| = |\alpha|$, including $\gamma = \alpha$, the first part finishes the proof of the proposition. ∎

2.8 Remark *As observed in the proof of Proposition 2.7, for $u \in C_c(\Omega)$ the first part does neither require the use of Lemma 2.5 nor any additional assumption on the measure other than it being Radon.*
Note also that compared to the statement for the usual mollification we slightly relaxed the condition on the set M for which the second part holds.

Density Propositions

We first use mollification to sharpen the density result for continuous functions with compact support in L^p_σ.

2.9 Proposition *Let $\sigma > -1$, $\Omega \subset H$ and $1 \le p < \infty$.*
Then $C^\infty_c(\Omega)$ is a dense subset of $L^p_\sigma(\Omega)$.

Proof: We know that for any $u \in L^p_\sigma(\Omega)$ and any $\delta > 0$ we can find a $\varphi \in C_c(\Omega)$ such that

$$\|u - \varphi\|_{L^p_\sigma(\Omega)} < \delta.$$

Then $J_{\varepsilon x_n} * \varphi \in C^\infty_c(\Omega)$ for small ε and we have

$$\|u - J_{\varepsilon x_n} * \varphi\|_{L^p_\sigma(\Omega)} \le \|u - \varphi\|_{L^p_\sigma(\Omega)} + \|\varphi - J_{\varepsilon x_n} * \varphi\|_{L^p_\sigma(\Omega)} .$$

The first part of Proposition 2.7 applied to $\varphi \in C_c(\Omega)$ then implies the proposition. ∎

Remark 2.8 assures that the last proposition is actually true for any Radon measure μ. Its proof could be simplified as to using merely $J_\varepsilon * \varphi$ instead of $J_{\varepsilon x_n} * \varphi$, and the restriction onto the upper half plane is superfluous.

In contrast to the situation for Lebesgue spaces, density statements in general Sobolev spaces require additional assumptions on the measure which are given by its weight character in our case.

2.10 Proposition *Let $m \in \mathbb{N}_0$, $\sigma_0, \ldots, \sigma_m > -1$, $\Omega \subset H$ and $1 \leq p < \infty$. Then $C^\infty(\Omega) \cap W^{m,p}_{\vec{\sigma}}(\Omega)$ is a dense subset of $W^{m,p}_{\vec{\sigma}}(\Omega)$.*

Proof: Let $u \in W^{m,p}_{\vec{\sigma}}(\Omega)$ and $\Psi \in C^\infty_c(\Omega)$. For any $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq m$ we then have

$$\|\partial_x^\alpha(\Psi\, u)\|^p_{L^p_{\sigma_{|\alpha|}}(\Omega)} \lesssim_{\alpha,p} \sum_{\beta \leq \alpha} \left(\sup_{x \in \operatorname{supp}\Psi} x_n^{\sigma_{|\alpha|}-\sigma_{|\beta|}} |\partial_x^{\alpha-\beta}\Psi|^p \right) \left\| \partial_x^\beta u \right\|^p_{L^p_{\sigma_{|\beta|}}(\Omega)}.$$

Since $\operatorname{supp}\Psi$ is compact, contained in Ω and hence stays away from $\{x_n = 0\}$, this shows that $\Psi\, u \in W^{m,p}_{\vec{\sigma}}(\Omega)$. Moreover, $\operatorname{supp}(\Psi\, u)$ is a compact subset of Ω.
Now the prerequisits of the second part of Proposition 2.7 are satisfied for $\Psi\, u$, so we see

$$\|J_{\varepsilon x_n} * (\Psi\, u) - \Psi\, u\|_{W^{m,p}_{\vec{\sigma}}(M)} \to 0 \quad (\varepsilon \to 0)$$

for any open and bounded set $M \subset \Omega$ that stays away from $\partial\Omega \cap (\mathbb{R}^n \setminus \partial H)$.
Next consider open and bounded $\Omega_N \subset \mathbb{R}^n$ with $\overline{\Omega}_N \subset \Omega_{N+1}$ and $\overline{\Omega}_N \subset \Omega$ for any $N \in \mathbb{N}$ as well as $\bigcup_{N \in \mathbb{N}} \Omega_N = \Omega$. Such sets exist, as the example

$$\Omega_N := \Omega \cap B_N^{eu}(0) \cap \left\{ x \in \mathbb{R}^n \mid d^{eu}(x, \partial\Omega) > \frac{1}{N} \right\}$$

shows. We also set $\Omega_0 := \emptyset$, $\Omega_{-1} := \emptyset$ and $U_j := \Omega_{j+1} \setminus \overline{\Omega}_{j-1}$, $M_j := \Omega_{j+2} \setminus \overline{\Omega}_{j-2}$ for any $j \in \mathbb{N}$. Obviously, all these sets are open subsets of Ω and we have $\overline{U}_j \subset M_j$ for any $j \in \mathbb{N}$ and $\overline{M}_j \subset (\mathbb{R}^n \setminus \Omega_N) \cap \Omega$ for any $j > N+2$ as well as $\bigcup_{j \in \mathbb{N}} U_j = \Omega = \bigcup_{j \in \mathbb{N}} M_j$.
We can then consider a partition of unity subordinate to $\{U_j \mid j \in \mathbb{N}\}$ and get the existence of $\Psi_j \in C^\infty(\Omega)$ with $\sum_{j \in \mathbb{N}} \Psi_j = 1$ on Ω and $\operatorname{supp}\Psi_j \subset U_j$ for any $j \in \mathbb{N}$. Thus

$$u(x) = \sum_{j=1}^{N+2} \Psi_j(x)\, u(x) \text{ for any } x \in \Omega_N.$$

Additionally, we set $\varphi_j(x) := (J_{\varepsilon_j x_n} * (\Psi_j\, u))$ for $\varepsilon_j > 0$ and $\varphi(x) := \sum_{j \in \mathbb{N}} \varphi_j$. If ε_j is small enough we also have $\operatorname{supp}\varphi_j \subset M_j$ and therefore

$$\varphi(x) = \sum_{j=1}^{N+2} \varphi_j(x) \text{ for any } x \in \Omega_N.$$

We now fix $\delta > 0$. Then

$$\|u - \varphi\|_{W^{m,p}_{\vec{\sigma}}(\Omega_N)} \leq \sum_{j=1}^{N+2} \left\| \Psi_j\, u - \varphi_j \right\|_{W^{m,p}_{\vec{\sigma}}(M_j)} < \delta(1 - 2^{-N-2})$$

if all ε_j are small enough by the first paragraph of this proof. The dominated convergence theorem then implies

$$\|u - \varphi\|_{W^{m,p}_{\vec{\sigma}}(\Omega)} \leq \lim_{N\to\infty} \delta\,(1 - 2^{-N-1}) = \delta$$

if all ε_j are small enough. But on any compact subset of Ω, the function φ is a finite sum of smooth functions. Hence $\varphi \in C^\infty(\Omega)$ and the proof is finished. ■

2.11 Remark *Since we have shown in Proposition* 2.2 *that our weighted Sobolev spaces is a Banach space, the last proposition in effect establishes that* $W^{m,p}_{\vec{\sigma}}(\Omega)$ *equals the space given by the completion of*

$$\left\{\varphi \in C^\infty(\Omega) \mid \|\varphi\|_{W^{m,p}_{\vec{\sigma}}(\Omega)} < \infty\right\}$$

in this norm. For unweighted spaces this result of Meyers and Serrin was first shown in their paper [MS64] with the incomparably short title "$H = W$*", reflecting the historical naming of the - presumably - two spaces involved. The new insight here was that the underlying set* Ω *does not have to satisfy any smoothness conditions at all. The definition of the sets* Ω_N *and* U_j *in our proof was taken from the original paper in the unweighted setting.*
For circumstances that do not imply the existence of distributional derivatives for the elements of weighted Lebesgue spaces, this approach still provides the means to define a Sobolev function and a notion of derivative. However, additional conditions have to be imposed to ensure that the so defined Sobolev derivative is well defined and coincides with the classical derivative for elements of $C^1(\Omega)$*, see for example [FKS82]. Even if the distributional derivative exists, it does not necessarily equal the Sobolev derivative of the same function. Defining the weighted Sobolev space as we did via distributional derivatives in such a case can result in an incomplete space. All this is discussed in greater detail in [HKM93].*
Finally, note also that in our proof the application of Proposition 2.7 *only takes place on a function whose support stays away from* $\{x_n = 0\}$*. In view of Remark* 2.6 *this means we could have used the usual mollification here instead. Likewise, Proposition* 2.10 *also holds for euclidean Muckenhoupt weights.*

In order to extend the approximation result 2.10 further and obtain the density of the smaller set of functions that are smooth and have an extension to a smooth function with compact support on $\mathbb{R}^n$, already in the unweighted case one has to impose additional geometric conditions on the underlying set Ω ([Maz85]). In addition to that, one needs the continuity of the translation operator with respect to the norm on the Lebesgue space over the set in question. In general this is not true for $L^p_\sigma(\Omega)$ unless $\sigma = 0$. Nonetheless, we can prove that a special translation on $\Omega = H$ satisfies the according inequality at least for $\sigma \geq 0$.

2.12 Lemma *Let* $\sigma \geq 0$ *and* $1 \leq p < \infty$.
If $u \in L^p_\sigma(H)$*, then*

$$\|u(\cdot + h\vec{e}_n) - u\|_{L^p_\sigma(H)} \to 0 \quad (h \searrow 0).$$

Proof: Fix $\delta > 0$. Then there exists a $\varphi \in C_c(H)$ with

$$\|\varphi - u\|_{L^p_\sigma(H)} < \delta$$

and we have

$$\|\varphi(\cdot + h\vec{e}_n) - \varphi\|_{L^p_\sigma(H)} < \delta$$

for $h > 0$ small enough thanks to to uniform continuity of φ on the compact set $\operatorname{supp}\varphi$. But moreover

$$\|u(\cdot + h\vec{e}_n) - \varphi(\cdot + h\vec{e}_n)\|^p_{L^p_\sigma(H)} = \int_{H+h\vec{e}_n} |u(x) - \varphi(x)|^p (x_n - h)^\sigma \, d\mathcal{L}^n(x) \le \|u - \varphi\|^p_{L^p_\sigma(H)}$$

holds, since $x_n - h > 0$ on $H + h\vec{e}_n \subset H$ and $\sigma \ge 0$. ∎

It is in the following proof that we truely need the special properties of the adjusted mollification.

2.13 Proposition *Let $m \in \mathbb{N}_0$, $\sigma_m \ge \ldots \ge \sigma_0 > -1$ with $\sigma_{i+1} - \sigma_i \le 1$ for $i = 0, \ldots, m-1$ as well as $\sigma_m \ge 0$, and $1 \le p < \infty$.*
Then $\breve{C}^\infty_c(H)$ is a dense subset of $W^{m,p}_{\vec{\sigma}}(H)$.

Proof: Let $\eta \in C^\infty_c(\mathbb{R}^n)$ be a cut-off function with $\eta = 1$ on $\overline{B}^{eu}_1(0)$, $\operatorname{supp}\eta \subset B^{eu}_2(0)$ and $|\partial^\alpha_x \eta| \le C$ on $\mathbb{R}^n$ for any $0 \le |\alpha| \le m$ with $C > 0$. For $\delta > 0$ define $\eta_\delta(x) := \eta(\delta x)$ for any $x \in \mathbb{R}^n$. Then we have $\eta_\delta = 1$ on $\overline{B}^{eu}_{\delta^{-1}}(0)$, $\operatorname{supp}\eta_\delta \subset \overline{B}^{eu}_{2\delta^{-1}}(0)$ and $|\partial^\alpha_x \eta_\delta| \le \delta^{|\alpha|} C$ on $\mathbb{R}^n$.
Fix $1 \le p < \infty$ and $u \in W^{m,p}_{\vec{\sigma}}(H)$, and consider $\eta_\delta\, u$, extending u by 0 outside of H. Obviously, we have that $\operatorname{supp}(\eta_\delta\, u) \subset \overline{H} \cap \overline{B}^{eu}_{2\delta^{-1}}(0)$ is a compact subset of $\mathbb{R}^n$. Furthermore, for $|\alpha| \le m$ we compute

$$\|\partial^\alpha_x(\eta_\delta\, u)\|^p_{L^p_{\sigma_{|\alpha|}}(H)} \lesssim_{\alpha,p} \sum_{\beta \le \alpha} \left(\sup_{x \in \operatorname{supp}\eta_\delta \cap H} x_n^{-\sigma_{|\alpha|} + \sigma_{|\beta|}} \right) \left\|\partial^\beta_x u\right\|_{L^p_{\sigma_{|\beta|}}(H)} \delta^{p(|\alpha| - |\beta|)}.$$

By assumption our weights are increasing, with an increase that is at most 1. This ensures that $\eta_\delta\, u \in W^{m,p}_{\vec{\sigma}}(H)$ for any $\delta \le 1$, or more precisely

$$\|\eta_\delta\, u\|_{W^{m,p}_{\vec{\sigma}}(H)} \lesssim_{m,p} \|u\|_{W^{m,p}_{\vec{\sigma}}(H)}$$

for any $\delta \le 1$. By construction we then get

$$\begin{aligned} \|\eta_\delta\, u - u\|_{W^{m,p}_{\vec{\sigma}}(H)} &= \|\eta_\delta\, u - u\|_{W^{m,p}_{\vec{\sigma}}(H \cap (\mathbb{R}^n \setminus B^{eu}_{\delta^{-1}}(0)))} \\ &\lesssim_{m,p} \|u\|_{W^{m,p}_{\vec{\sigma}}(H \cap (\mathbb{R}^n \setminus B^{eu}_{\delta^{-1}}(0)))} \to 0 \quad (\delta \to 0). \end{aligned}$$

For any point $z \in \partial H$ let now $U(z)$ and $M(z)$ be open neighbourhoods of z with $\overline{U}(z) \subset M(z)$. The set

$$F := \operatorname{supp}(\eta_\delta\, u) \setminus \Big(\bigcup_{z \in \partial H} M(z) \Big)$$

is then a compact subset of H. Consequently, we can find open sets U_0 and M_0 with $F \subset U_0$, $\overline{U}_0 \subset M_0$ and $\overline{M}_0 \subset H$. Then $\{U(z) \mid z \in \partial H\} \cup \{U_0\}$ is an open cover of $\operatorname{supp}(\eta_\delta\, u)$ and we can hence extract a finite subcover renamed to $\{U_0, \ldots, U_N\}$ with the corresponding bigger cover $\{M_0, \ldots, M_N\}$. We now consider a partition of unity subordinate to $\{U_j \mid j = 0, \ldots, N\}$, that is we have $\Psi_j \in C^\infty_c(U_j)$ with $\sum_{j=0}^N \Psi_j = 1$ on $\bigcup_{j=1}^N U_j$. Set $u_j := \Psi_j\, \eta_\delta\, u$ on $\mathbb{R}^n$. Then $\operatorname{supp} u_j \subset \overline{H} \cap U_j$, and $u_j \in W^{m,p}_{\vec{\sigma}}(H \cap M_j)$ can be seen as before, since the weights are increasing and $\operatorname{supp}\Psi_j$ is compact.
For $j = 0$ this means that $\overline{\operatorname{supp} u_0}$ is contained in H. We then define $\varphi_{0,\varepsilon_0} := J_{\varepsilon_0 x_n} * u_0$ for $\varepsilon_0 > 0$

and get

$$\|\varphi_{0,\varepsilon_0} - u_0\|_{W^{m,p}_{\vec{\sigma}}(H)} = \|\varphi_{0,\varepsilon_0} - u_0\|_{W^{m,p}_{\vec{\sigma}}(U_0)} \to 0 \quad (\varepsilon_0 \to 0)$$

by the second part of Proposition 2.7.
For $j = 1, \dots, N$, on the other hand, the support of u_j contains a portion of ∂H. For $x \in H$ and $\varepsilon_j > 0$ we consider $\varphi_{j,\varepsilon_j}(x) := (J_{\varepsilon_j x_n} * u_j)(x)$, and set it to 0 outside of H. Since we are close to the boundary of H we can choose ε_j so small that $\operatorname{supp} \varphi_{j,\varepsilon_j} \subset \overline{H} \cap M_j$ for ε_j small. Furthermore we have $\varphi_{j,\varepsilon_j} \in C^\infty(H \cap M_j)$. Close to the boundary we now take advantage of the special properties of the adopted mollification and use the second part of Proposition 2.7 on $H \cap U_j$ to get

$$\left\| u_j - \varphi_{j,\varepsilon_j} \right\|_{W^{m,p}_{\vec{\sigma}}(H \cap U_j)} \to 0 \, (\varepsilon_j \to 0).$$

We now push the regularisation over the boundary of H. To this end, for $h_j > 0$ define $\varphi_{j,\varepsilon_j,h_j}(x) := \varphi_{j,\varepsilon_j}(x + h_j \vec{e}_n)$ for any $x \in \mathbb{R}^n$. The corresponding support is then partially translated into the lower half plane. However, for h_j small enough it stays in M_j, and so we have $\operatorname{supp} \varphi_{j,\varepsilon_j,h_j} \subset (\overline{H} - h_j \vec{e}_n) \cap M_j$ and $\varphi_{j,\varepsilon_j,h_j}$ is smooth on $(H - h_j \vec{e}_n) \cap M_j$. Then $\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j} \in C^\infty(H)$ with $\operatorname{supp}(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \subset M_j \cap \overline{H}$ and for any $i = 0, \dots, m-1$ we can apply Hardy's inequality 2.3 to get

$$\left\| D^i(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \right\|_{L^p_{\sigma_i}(H)} \lesssim_{\sigma,m,p} \left\| D^m(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \right\|_{L^p_{(m-i)p+\sigma_i}(H)}.$$

On H it is enough to consider the bounded set $M_j \cap H$ since the support of the function is contained in it. But there we get

$$\left\| D^m(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \right\|^p_{L^p_{(m-i)p+\sigma_i}(H)} \le \left(\sup_{x \in M_j \cap H} x_n^{(m-i)p+\sigma_i-\sigma_m} \right) \left\| D^m(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \right\|^p_{L^p_{\sigma_m}(H)}$$

and by the moderate increase of the weight exponents we have that $(m-i)p + \sigma_i - \sigma_m$ is positive for any i. But $\sigma_m \ge 0$ by assumption and thus Lemma 2.12 can be applied to see that for any $i = 0, \dots m$ the upper bound just given approximates 0 if h_j vanishes. Hence

$$\left\| D^m(\varphi_{j,\varepsilon_j,h_j} - \varphi_{j,\varepsilon_j}) \right\|_{W^{m,p}_{\vec{\sigma}}(H)} \to 0 \quad (h_j \to 0).$$

Let us now set $\varphi_{0,\varepsilon_0,h_0} := \varphi_{0,\varepsilon_0}$. Then $\sum_{j=0}^{N} \varphi_{j,\varepsilon_j,h_j}$ is smooth on a neighbourhood of H including the boundary and has compact support. We can then find a function $\varphi \in C_c^\infty(\mathbb{R}^n)$ with $\varphi|_{\overline{H}} = (\sum_{j=0}^{N} \varphi_{j,\varepsilon_j,h_j})|_{\overline{H}}$ and the property

$$\begin{aligned} \|u - \varphi\|_{W^{m,p}_{\vec{\sigma}}(H)} \le \|u - \eta_\delta u\|_{W^{m,p}_{\vec{\sigma}}(H)} + \|u_0 - \varphi_{0,\varepsilon_0}\|_{W^{m,p}_{\vec{\sigma}}(H)} \\ + \sum_{j=1}^{N} \left(\left\| u_j - \varphi_{j,\varepsilon_j} \right\|_{W^{m,p}_{\vec{\sigma}}(H)} + \left\| \varphi_{j,\varepsilon_j} - \varphi_{j,\varepsilon_j,h_j} \right\|_{W^{m,p}_{\vec{\sigma}}(H)} \right). \end{aligned}$$

This gets arbitrarily small for $\delta, \varepsilon_0, \dots, \varepsilon_N, h_1, \dots, h_N$ small enough. ■

2.14 Remark *We would like to point out that we cannot reverse the order of the last two steps in the proof of Proposition* 2.13*: If the translation is done first we drop out of H where the adopted mollification is* *not available any more. Conversely, first mollifying and then translating would not work for the usual mollification. It is precisely in this step that we need not only the convergences of the mollified function in the weighted Lebesgue spaces and close to ∂H, but also the slight gain in the condition under which the derivative of the adopted mollification can be computed.*

2.15 Remark *An immediate consequence of Proposition 2.13 is that also $\check{C}_c^\infty(\overline{H})$ is dense in $W_{\vec{\sigma}}^{m,p}(H)$ if $\vec{\sigma}$ complies with the hypotheses of Proposition 2.13. This in turn implies that $C_c^\infty(\overline{H})$ is dense in the Sobolev spaces. Of course it is not true even in the unweighted case that $C_c^\infty(H)$ is dense in $W^{m,p}(H)$ for an open Ω unless Ω equals $\mathbb{R}^n$ or is in a suitable sense very close to it and very regular. More on this is contained in [AF03].*

An Interpolation Inequality

We finally would like to derive a weighted version of the Gagliardo-Nirenberg interpolation inequality. Both Sobolev's and Morrey's inequality could be unified into one result that in turn could be considered to be special cases of a generalised interpolation inequality, as is demonstrated in [Fri08]. We follow this text also in the proofs we present for the unweighted case and start with a one-dimensional lemma.

2.16 Lemma *Let $I \subset \mathbb{R}$ be a bounded and open interval and $u \in C^2(I)$.*
Then

$$\|\partial_x u\|_{L^{p_1}(I)} \lesssim_{p_1} |I|^{-1-\frac{1}{p_0}+\frac{1}{p_1}} \|u\|_{L^{p_0}(I)} + |I|^{1-\frac{1}{p_2}+\frac{1}{p_1}} \left\|\partial_x^2 u\right\|_{L^{p_2}(I)}$$

for any $1 \leq p_0, p_1, p_2 \leq \infty$.

Proof: Let $I := (a,b)$ and consider $y \in (a, a+\frac{1}{4}|I|)$, $z \in (a+\frac{3}{4}|I|, b)$. By the intermediate value theorem we then get $\xi \in (y,z)$ with

$$\partial_x u(\xi) = \frac{u(z)-u(y)}{z-y}$$

and hence

$$|\partial_x u(x)| \leq \frac{|u(z)|+|u(y)|}{|z-y|} + \int\limits_{(\xi,x)} |\partial_x^2 u| \, d\mathcal{L}$$

for any $x \in I$ by the fundamental theorem of calculus. Because of the above choice of y and z we know that $|z-y| \geq \frac{1}{2}|I|$. An integration with respect to $(y,z) \in (a, a+\frac{1}{4}|I|) \times (a+\frac{3}{4}|I|, b)$ then results in

$$\frac{1}{16}|I|^2 |\partial_x u(x)| \leq \frac{1}{2}\int\limits_I |u| \, d\mathcal{L} + \frac{1}{16}|I|^2 \int\limits_I |\partial_x^2 u| \, d\mathcal{L}.$$

Taking the p_1-th power on both sides for $1 \leq p_1 < \infty$ as well as using Hölder's inequality for an arbitrary $1 \leq p_0 \leq \infty$ in the first integral, and for an arbitrary $1 \leq p_2 \leq \infty$ in the second, after another integration we obtain

$$\|\partial_x u\|_{L^{p_1}(I)}^{p_1} \lesssim_{p_1} |I|^{-p_1-\frac{p_1}{p_0}+1} \|u\|_{L^{p_0}(I)}^{p_1} + |I|^{p_1-\frac{p_1}{p_2}+1} \left\|\partial_x^2 u\right\|_{L^{p_2}(I)}^{p_1}.$$

This implies the statement for $1 \leq p_1 < \infty$, where we can also let p_1 tend to ∞ to finish the proof.■

By means of Lemma 2.16 we can prove the unweighted Gagliardo-Nirenberg inequality.

2.17 Proposition *Let $m \in \mathbb{N}$, $i \in \mathbb{N}$ with $i < m$ and $1 \leq p_0, p_1, p_2 \leq \infty$ with $\frac{1}{p_1} = \frac{m-i}{mp_0} + \frac{i}{mp_2}$. Further let $a \in \mathbb{R} \cup \{-\infty\}$.*
If $u \in \check{C}_c^m(H_a)$, then

$$\left\| D^i u \right\|_{L^{p_1}(H_a)} \lesssim_{n,m,p_1} \|u\|_{L^{p_0}(H_a)}^{\frac{m-i}{m}} \|D^m u\|_{L^{p_2}(H_a)}^{\frac{i}{m}}.$$

Proof: Let first $m = 2$, $i = 1$. The dimension be $n = 1$ and we fix $a \in \mathbb{R}$. Of course $u \in \check{C}_c^2((a,\infty))$ implies $u|_I \in C^2(I)$ and we can apply Lemma 2.16, raised to the power $p_1 < \infty$, on every subinterval of (a,∞). We consider a bounded and open interval $I = (a,b) \subsetneq (a,\infty)$. On $\widetilde{I}_1 := (a, a + j^{-1}|I|) \subset I$, $j \in \mathbb{N}$ fixed, we now compare the terms of the right hand side of the inequality resulting from the application of Lemma 2.16 there. If

$$|\widetilde{I}_1|^{-p_1 - \frac{p_1}{p_0} + 1} \|u\|_{L^{p_0}(\widetilde{I}_1)}^{p_1} \leq |\widetilde{I}_1|^{p_1 - \frac{p_1}{p_2} + 1} \left\| \partial_x^2 u \right\|_{L^{p_2}(\widetilde{I}_1)}^{p_1}$$

we set $I_1 := \widetilde{I}_1 := (b_0, b_1)$ with $b_0 := a$ and $b_1 := b_0 + j^{-1} |I|$. Hence on $I_1 \subset I$ we get

$$\|\partial_x u\|_{L^{p_1}(I_1)}^{p_1} \lesssim_{p_1} j^{-p_1 + \frac{p_1}{p_2} - 1} |I|^{p_1 - \frac{p_1}{p_2} + 1} \left\| \partial_x^2 u \right\|_{L^{p_2}(I)}.$$

In the opposite case we define $b_0 := a$ and $b_1 > b_0 + j^{-1} |I|$ in such a way, that on $I_1 := (b_0, b_1)$ equality of the two terms holds. A b_1 like this exists, since due to the exponents of $|I|$, excluding $p_0 = \infty$, the term involving the norm of u decreases with increasing length of the interval, while the other term containing $\partial_x^2 u$ increases. Moreover, assuming that $u \in \check{C}_c^2((a,\infty))$ assures that $\partial_x^2 u$ is not the constant zero function on all of (a,∞) unless $u = 0$. So by construction we get

$$\|\partial_x u\|_{L^{p_1}(I_1)}^{p_1} \lesssim_{p_1} |I_1|^{1 - \frac{p_1}{2p_0} - \frac{p_1}{2p_2}} \|u\|_{L^{p_0}(I_1)}^{\frac{p_1}{2}} \left\| \partial_x^2 u \right\|_{L^{p_2}(I_1)}^{\frac{p_1}{2}}$$

in this case.
We continue this procedure by comparing the terms on $\widetilde{I}_2 := (b_1, \min\left\{b, b_1 + j^{-1}|I|\right\})$ and define I_2 accordingly, and so forth, until we stop after after N steps as soon as we have $b_N \geq b$. Then $N \leq j$ and $b_{N-1} < b$, and we can take the union of the disjoint intervals $I_1, \ldots, I_N$ to get

$$\|\partial_x u\|_{L^{p_1}(I)}^{p_1} \lesssim_{p_1} j^{-p_1 + \frac{p_1}{p_2}} |I|^{p_1 - \frac{p_1}{p_2} + 1} \left\| \partial_x^2 u \right\|_{L^{p_2}(I)}^{p_1} + \sum_{i=1}^{N} |I_i|^{1 - \frac{p_1}{2p_0} - \frac{p_1}{2p_2}} \|u\|_{L^{p_0}(I_i)}^{\frac{p_1}{2}} \left\| \partial_x^2 u \right\|_{L^{p_2}(I_i)}^{\frac{p_1}{2}}.$$

Hölder's inequality for sums in the second summand then reveals the estimate

$$\|\partial_x u\|_{L^{p_1}(I)}^{p_1} \lesssim_{p_1} j^{-p_1 + \frac{p_1}{p_2}} |I|^{p_1 - \frac{p_1}{p_2} + 1} \left\| \partial_x^2 u \right\|_{L^{p_2}(I)}^{p_1} + (b_N - a)^{1 - \frac{p_1}{2p_0} - \frac{p_1}{2p_2}} \|u\|_{L^{p_0}((a,b_N))}^{\frac{p_1}{2}} \left\| \partial_x^2 u \right\|_{L^{p_2}((a,b_N))}^{\frac{p_1}{2}}.$$

For $j \to \infty$, the first summand goes to 0 if $p_2 > 1$. However, in the limit we can not control b_N and thus the second term could be very big. We therefore cause it to vanish by demanding

$$1 - \frac{p_1}{2p_0} - \frac{p_1}{2p_2} = 0.$$

The interval (a, b_N) is of course always contained in (a, ∞). The result holds for any $b > a$, so also the norm on the left hand side can be taken to be on (a, ∞). But that again is true for any $a \in \mathbb{R}$, making it possible to allow $a = -\infty$. Taking the $\frac{1}{p_1}$-th power on both sides yields

$$\|\partial_x u\|_{L^{p_1}((a,\infty))} \lesssim_{p_1} \|u\|^{\frac{1}{2}}_{L^{p_0}((a,\infty))} \left\|\partial_x^2 u\right\|^{\frac{1}{2}}_{L^{p_2}((a,\infty))}$$

for $1 \leq p_0, p_1 < \infty$, $1 < p_2 \leq \infty$ with $\frac{1}{p_1} - \frac{1}{2p_0} - \frac{1}{2p_2} = 0$, if $u \in \check{C}^2_c((a, \infty))$. By taking the limit we also get the cases $p_0 = \infty$, $p_1 = \infty$, $p_2 = 1$.
For higher dimensions we use the one-dimensional version just proven on $x_j \mapsto u(x', x_j) \in \check{C}^2_c((a_j, \infty))$ with $a_1 = \ldots = a_{n-1} = -\infty$, $a_n = a$, and the notation

$$H'_a := (a_1, \infty) \times \ldots \times (a_{j-1}, \infty) \times (a_{j+1}, \infty) \times \ldots \times (a_n, \infty).$$

In conjunction with Hölder's inequality and the condition on p_0, p_1, p_2 we then calculate

$$\begin{aligned}
\left\|\partial_{x_j} u\right\|^{p_1}_{L^{p_1}(H_a)} &\leq \int\limits_{H'_a} \left\|\partial_{x_j} u(x', \cdot)\right\|^{p_1}_{L^{p_1}((a_j,\infty))} \, d\mathcal{L}^{n-1}(x') \\
&\lesssim_{p_1} \int\limits_{H'_a} \|u(x', \cdot)\|^{\frac{p_1}{2}}_{L^{p_0}((a_j,\infty))} \left\|\partial^2_{x_j} u(x', \cdot)\right\|^{\frac{p_1}{2}}_{L^{p_2}((a_j,\infty))} \, d\mathcal{L}^{n-1}(x') \\
&\leq \|u\|^{\frac{p_1}{2}}_{L^{p_0}(H_a)} \left\|\partial^2_{x_j} u\right\|^{\frac{p_1}{2}}_{L^{p_2}(H_a)}.
\end{aligned}$$

The dependency of the constant on n enters when writing $|\nabla_x u|$ instead of $\partial_{x_j} u$ on the left hand side.
Finally, an induction on m finishes the proof. ∎

The weighted Gagliardo-Nirenberg inequality follows from the unweighted one. For weights similar to ours, the result for differentiable functions can also be found in [Lin86]. For different and rather general weight functions, [NS96] contains a Gagliardo-Nirenberg inequality that generalises the assumption to Sobolev spaces, as it is also done in the following proposition.

2.18 Proposition *Let $\sigma \geq 0$, $m \in \mathbb{N}$ and $i \in \mathbb{N}$ with $i < m$. Further let $1 \leq p_0 \leq \infty$ and $1 \leq p_1, p_2 < \infty$ with $\frac{1}{p_1} = \frac{m-i}{mp_0} + \frac{i}{mp_2}$.*
If $u \in L^{p_0}_\sigma(H) \cap W^{m,p_2}_{\vec{\sigma}}(H)$ with $\sigma_0 = \ldots = \sigma_m = \sigma$, then

$$\left\|D^i u\right\|_{L^{p_1}_\sigma(H)} \lesssim_{n,m,p_1} \|u\|^{\frac{m-i}{m}}_{L^{p_0}_\sigma(H)} \|D^m u\|^{\frac{i}{m}}_{L^{p_2}_\sigma(H)}.$$

Proof: Let $\sigma > 0$ and first $u \in \check{C}^m_c(H)$. We calculate

$$\int\limits_H |D^i u(x)|^{p_1} \, d\mathcal{L}^n(x) \leq \sigma \int\limits_{(0,\infty)} z_n^{\sigma-1} \left\|D^i u\right\|^{p_1}_{L^{p_1}(H_{z_n})} \, d\mathcal{L}(z_n)$$

with the fundamental theorem of calculus and Fubini's theorem. The unweighted Gagliardo-

Nirenberg inequality 2.17 on H_{z_n} then generates the bound

$$\sigma \int\limits_{(0,\infty)} z_n^{\sigma-1} \left(\int\limits_{H_{z_n}} |u|^{p_0} \, d\mathcal{L}^n \right)^{\frac{m-1}{m}\frac{p_1}{p_0}} \left(\int\limits_{H_{z_n}} |D_x^m u|^{p_2} \, d\mathcal{L}^n \right)^{\frac{i}{m}\frac{p_1}{p_2}} d\mathcal{L}(z_n)$$
$$\leq \left(\int\limits_{(0,\infty)} z_n^{\sigma-1} \int\limits_{H_{z_n}} |u|^{p_0} \, d\mathcal{L}^n \, d\mathcal{L}(z_n) \right)^{\frac{m-1}{m}\frac{p_1}{p_0}} \left(\int\limits_{(0,\infty)} z_n^{\sigma-1} \int\limits_{H_{z_n}} |D_x^m u|^{p_2} \, d\mathcal{L}^n \, d\mathcal{L}(z_n) \right)^{\frac{i}{m}\frac{p_1}{p_2}}$$

since $\frac{m-i}{m}\frac{p_1}{p_0} + \frac{i}{m}\frac{p_1}{p_2} = 1$. The constant here depends on n, m and p_1.
Reversing the order of integration and integrating out the z_n-integral cancels the factor σ again and delivers the statement for $u \in \check{C}_c^m(H)$.
But the latter is dense in $L_\sigma^{p_0}(H) \cap W_{\vec{\sigma}}^{m,p_2}(H)$ for the conditions on σ_i stated in the prerequisits, so we can find $(\varphi_j)_{j\in\mathbb{N}} \subset \check{C}_c^m(H)$ that converges to u in $L_\sigma^{p_0}(H) \cap W_{\vec{\sigma}}^{m,p_2}(H)$. The convergence of $D_x^i \varphi_j$ to $D^i u$ in $L_\sigma^{p_1}(H)$ follows. This concludes the proof. ∎

3 Solutions of the Linear Perturbation Equation

We first define a suitable notion of solution of the linear perturbation equation $\partial_t u - L_\sigma u = f$, where

$$L_\sigma u = (\,\cdot\,)_n \, \Delta_x u + (1+\sigma)\, \partial_{x_n} u.$$

Here we consider both the initial value problem and a setting that is local in time. Since we are interested in what happens near ∂H, we use test functions that do not always vanish near the spatial boundary of the underlying set. For a reasonable theory we define solutions in the energy sense, equipping them with some additional regularity properties that allow for energy techniques. In this context, however, the requirements we set are the weakest possible.

3.1 Definition *Consider $\sigma > -1$, an open interval $I = (t_1, t_2) \subset \mathbb{R}$ and a relatively open subset Ω of $\overline{H}$.*

- *Given $f \in L^1_{loc}\big(I;\, L^2_\sigma(\Omega)\big)$, we say that u is a σ-solution to f on $I \times \Omega$, if and only if $u \in L^2_{loc}\big(I;\, L^2_\sigma(\Omega)\big)$, $\nabla_x u \in L^2\big(I;\, L^2_{1+\sigma}(\Omega)\big)$ and*

$$-\int_I (\, u \mid \partial_t \varphi \,)_{L^2_\sigma(\Omega)} \, d\mathcal{L} + \int_I (\, \nabla_x u \mid \nabla_x \varphi \,)_{L^2_{1+\sigma}(\Omega)} \, d\mathcal{L} = \int_I (\, f \mid \varphi \,)_{L^2_\sigma(\Omega)} \, d\mathcal{L}$$

  for all $\varphi \in C^\infty_c(I \times \Omega)$.

- *Given $t_1 > -\infty$, $f \in L^1_{loc}\big(\lceil I;\, L^2_\sigma(\Omega)\big)$ and $g \in L^2_\sigma(\Omega)$, we say that u is a σ-solution to f on $\lceil I \times \Omega$ with initial value g, if and only if $u \in L^2_{loc}\big(\lceil I;\, L^2_\sigma(\Omega)\big)$, $\nabla_x u \in L^2\big(I;\, L^2_{1+\sigma}(\Omega)\big)$ and*

$$-\int_I (\, u \mid \partial_t \varphi \,)_{L^2_\sigma(\Omega)} \, d\mathcal{L} + \int_I (\, \nabla_x u \mid \nabla_x \varphi \,)_{L^2_{1+\sigma}(\Omega)} \, d\mathcal{L} = \int_I (\, f \mid \varphi \,)_{L^2_\sigma(\Omega)} \, d\mathcal{L} + (\, g \mid \varphi(t_1) \,)_{L^2_\sigma(\Omega)}$$

  for all $\varphi \in C^\infty_c(\lceil I \times \Omega)$.

Of course any σ-solution to an initial value problem on $\lceil I \times \Omega$ is also a σ-solution to the time-local problem on $I \times \Omega$. Moreover, a short calculation shows that a time-local solution on $I \times \Omega$ gives rise to a solution to the initial value problem on $[\widetilde{t}_1, t_2) \times \Omega$ for any point $\widetilde{t}_1 \in I$. This means that statements on $\lceil I$, proven for an initial value solution on $\lceil I \times \Omega$, immediately imply statements on $[\widetilde{t}_1, t_2)$ for the time-local solution on $I \times \Omega$ for almost any $\widetilde{t}_1 \in I$. We will make use of this remark in the sequel without further comment.

3.2 Remark *It is worth pointing out at this point that the equation possesses an invariant scaling. Considering the coordinate transformation*

$$T_\lambda : (\hat{t}, \hat{x}) \mapsto (\lambda\, \hat{t}, \lambda\, \hat{x}) =: (t, x),$$

a calculation shows that if u is a σ-solution to f and g on $\lceil I \times \Omega$ with respect to (t, x), then $u \circ T_\lambda$ is a σ-solution to $\lambda\, (f \circ T_\lambda)$ and $\hat{g}$ on $T^{-1}(\lceil I \times \Omega)$ with respect to $(\hat{t}, \hat{x})$, where $\hat{g}(\hat{x}) := g(\lambda\, \hat{x})$ for any

$\hat{x} \in \hat{\Omega} := \lambda^{-1}\,\Omega$.
Moreover, translations in any temporal and spatial direction save the x_n-direction commute with the differential operator.

A Galerkin approximation now shows that the above choice of conditions for the solution is reasonable: under a weak assumption on the inhomogeneity it provides us with a σ-solution to the initial value problem.

3.3 Proposition *Let $t_1 > -\infty$, $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $\Omega \subset \overline{H}$ be relatively open. Then for any $f \in L^1\big(I;\, L^2_\sigma(\Omega)\big)$ and any $g \in L^2_\sigma(\Omega)$ there exists a σ-solution to f on $\overline{I} \times \Omega$ with initial value g.*

Proof: In the following we drop I and Ω from the notation whenever spaces with respect to these sets are involved and the meaning is clear from the context, and moreover write for example $L^1L^2_\sigma$ instead of $L^1\big(I;\, L^2_\sigma(\Omega)\big)$.
The Hilbert space $W^{1,2}_\sigma$ has a Schauder basis $\{w_i \,:\, i \in \mathbb{N}\}$. We define

$$A_m := \Big(\big(\, w_j \mid w_i \,\big)_{L^2_\sigma}\Big)_{i,j=1}^m,\, B_m := \Big(\big(\, \nabla_x w_j \mid \nabla_x w_i \,\big)_{L^2_{1+\sigma}}\Big)_{i,j=1}^m \text{ and } \vec{f}_m(t) := \Big(\big(\, f(t) \mid w_i \,\big)_{L^2_\sigma}\Big)_{i=1}^m$$

for any positive integer m. Using the linear independence of the Schauder basis, one sees that A_m is positive definite and thus invertible for any $m \in \mathbb{N}$.
Obviously, $A_m^{-1}B_m,\, A_m^{-1}\vec{f}_m \in L^1_{loc}(\overline{I})$, and thus by standard theory of ordinary differential equations, for any given $t_0 \in \overline{I}$, especially for $t_0 = t_1$, and for any $\vec{c} \in \mathbb{R}^m$ we can find a unique $\vec{d}_m \in W^{1,1}_{loc}(\overline{I})$ such that $\vec{d}_m(t_1) = \vec{c}$ exists and

$$\partial_t \vec{d}_m + A_m^{-1} B_m\, \vec{d}_m = A_m^{-1} \vec{f}_m \text{ almost everywhere on } I.$$

Extending these expressions explicitely shows that we can be sure of the existence of a function

$$u_m := \sum_{j=1}^m d_j^{(m)}\, w_j \in L^1_{loc}(\overline{I};\, W^{1,2}_\sigma(\Omega))$$

with

$$\partial_t u_m = \sum_{j=1}^m \partial_t d_j^{(m)}\, w_j \in L^1_{loc}(\overline{I};\, W^{1,2}_\sigma(\Omega)),$$

$$u_m(t_1) = \sum_{j=1}^m c_j w_j$$

and

$$(*) \qquad \big(\, \partial_t u_m(t) \mid w_i \,\big)_{L^2_\sigma} + \big(\, \nabla_x u_m(t) \mid \nabla_x w_i \,\big)_{L^2_{1+\sigma}} = \big(\, f(t) \mid w_i \,\big)_{L^2_\sigma}$$

for almost all $t \in I$ and for all $i = 1, \ldots, m$.
We multiply these equations by $d_i^{(m)}(t)$ and sum over i. The product rule for bilinear forms implies that we have $\big(\, \partial_t u_m(t) \mid u_m \,\big)_{L^2_\sigma} = \frac{1}{2}\,\partial_t\, \|u_m(t)\|^2_{L^2_\sigma}$. This means that for any $\vec{c} \in \mathbb{R}^m$ there is a u_m as above with

$$\frac{1}{2}\,\partial_t\, \|u_m(t)\|^2_{L^2_\sigma} + \|\nabla_x u_m(t)\|^2_{L^2_{1+\sigma}} = \big(\, f(t) \mid u_m(t) \,\big)_{L^2_\sigma}$$

for almost all $t \in I$.
For the purpose of integrating in time consider $\widetilde{t} \in \overline{I}$ and the bounded interval $(t_1, \widetilde{t}) =: \widetilde{I} \subset I$ as the region of integration. In combination with the fundamental theorem of calculus this shows

$$\frac{1}{2}\, \|u_m(\widetilde{t})\|^2_{L^2_\sigma} - \frac{1}{2}\, \|u_m(t_1)\|^2_{L^2_\sigma} + \int\limits_{\widetilde{I}} \|\nabla_x u_m\|^2_{L^2_{1+\sigma}}\, d\mathcal{L} = \int\limits_{\widetilde{I}} (\, f(t) \mid u_m(t)\,)_{L^2_\sigma}\, d\mathcal{L}$$

for any $\widetilde{t} \in \overline{I}$.
On the right hand side we can apply Hölder's inequality first in the spatial integral and then in the temporal integral to get

$$\begin{aligned}
\frac{1}{2}\, \|u_m(\widetilde{t})\|^2_{L^2_\sigma} - \frac{1}{2}\, \|u_m(t_1)\|^2_{L^2_\sigma} + \int\limits_{\widetilde{I}} \|\nabla_x u_m\|^2_{L^2_{1+\sigma}}\, d\mathcal{L} \leq \\
\leq \sup_{t \in \widetilde{I}}\, \|u_m(t)\|_{L^2_\sigma} \int\limits_{\widetilde{I}} \|f(t)\|_{L^2_\sigma}\, d\mathcal{L} \leq \\
\leq \frac{1}{4} \sup_{t \in I}\, \|u_m(t)\|^2_{L^2_\sigma} + \|f\|^2_{L^1 L^2_\sigma}
\end{aligned}$$

where we used Young's inequality and a substitution of $\widetilde{I}$ for the at most bigger interval I in the last step.
At this point it is possible to take the supremum over $\widetilde{t} \in I$ on the left hand side. Rearranging terms then produces

$$\frac{1}{4} \sup_{t \in I}\, \|u_m(t)\|^2_{L^2_\sigma} + \int\limits_{I} \|\nabla_x u_m\|^2_{L^2_{1+\sigma}}\, d\mathcal{L} \leq \|f\|^2_{L^1 L^2_\sigma} + \frac{1}{2}\, \|u_m(t_1)\|^2_{L^2_\sigma}\,.$$

Now set $\vec{c} := ((\, g \mid w_j\,)_{L^2_\sigma})_{j=1}^m$. Then

$$\left\| \sum_{j=1}^{m} (\, g \mid w_j\,)_{L^2_\sigma}\, w_j \right\|_{L^2_\sigma} \leq \|g\|_{L^2_\sigma}$$

and we thus have a u_m with

$$u_m(t_1) = \sum_{j=1}^{m} (\, g \mid w_j\,)_{L^2_\sigma}\, w_j$$

and

$$\frac{1}{4} \sup_{t \in I}\, \|u_m(t)\|^2_{L^2_\sigma} + \int\limits_{I} \|\nabla_x u_m\|^2_{L^2_{1+\sigma}}\, d\mathcal{L} \leq \|f\|^2_{L^1 L^2_\sigma} + \frac{1}{2}\, \|g\|^2_{L^2_\sigma}\,.$$

This implies that $(u_m)_{m \in \mathbb{N}}$ is a bounded sequence in $L^\infty L^2_\sigma$, and $(\nabla_x u_m)_{m \in \mathbb{N}}$ is a bounded sequence in $L^2 L^2_{1+\sigma}$. Hence we can extract subsequences - not relabelled in the following - such that for $m \to \infty$ we have

$$u_m \rightharpoonup^* u \in L^\infty L^2_\sigma$$

and

$$\nabla_x u_m \rightharpoonup \vec{v} \in L^2 L^2_{1+\sigma}.$$

Explicitely, this means that there exist $u \in L^\infty L^2_\sigma$ and $\vec{v} \in L^2 L^2_{1+\sigma}$ with

$$\int_I (\, u_m(t) \mid \xi(t) \,)_{L^2_\sigma} \, d\mathcal{L} \to \int_I (\, u(t) \mid \xi(t) \,)_{L^2_\sigma} \, d\mathcal{L} \text{ for all } \xi \in L^1 L^2_\sigma$$

and

$$\int_I (\, \nabla_x u_m(t) \mid \vec{\zeta}(t) \,)_{L^2_{1+\sigma}} \, d\mathcal{L} \to \int_I (\, \vec{v}(t) \mid \vec{\zeta}(t) \,)_{L^2_{1+\sigma}} \, d\mathcal{L} \text{ for all } \vec{\zeta} \in L^2 L^2_{1+\sigma}$$

as $m \to \infty$. It follows that indeed $\vec{v} = \nabla_x u$.
Let us now turn back to the equations $(*)$. We choose an arbitrary $N \in \mathbb{N}$ and let m be big, precisely $m > N$. Then for $i = 1, \ldots, N$ we multiply the equations by $\widetilde{d}_i \in C_c^\infty(\overline{I})$. Summing up and denoting $\widetilde{\varphi} := \sum_{i=1}^{N} \widetilde{d}_i \, w_i$ then gives

$$(\, \partial_t u_m(t) \mid \widetilde{\varphi}(t) \,)_{L^2_\sigma} + (\, \nabla_x u_m(t) \mid \nabla_x \widetilde{\varphi}(t) \,)_{L^2_{1+\sigma}} = (\, f(t) \mid \widetilde{\varphi}(t) \,)_{L^2_\sigma}$$

for almost all $t \in I$. We can integrate this equation over I and execute a temporal integration by parts on the first summand. Then u_m satisfies

$$-\int_I (\, u_m \mid \partial_t \widetilde{\varphi} \,)_{L^2_\sigma} \, d\mathcal{L} + \int_I (\, \nabla_x u_m \mid \nabla_x \widetilde{\varphi} \,)_{L^2_{1+\sigma}} \, d\mathcal{L} = \int_I (\, f \mid \widetilde{\varphi} \,)_{L^2_\sigma} \, d\mathcal{L} + (\, u_m(t_1) \mid \widetilde{\varphi}(t_1) \,)_{L^2_\sigma}.$$

Now $\partial_t \widetilde{\varphi}$ can be used as test function ξ in the convergence above, and likewise $\nabla_x \widetilde{\varphi}$ as test function $\vec{\zeta}$. Furthermore we have

$$\Big(\, \sum_{j=1}^{m} (\, g \mid w_j \,)_{L^2_\sigma} \, w_j \mid \widetilde{\varphi}(t_1) \,\Big)_{L^2_\sigma} \to (\, g \mid \widetilde{\varphi}(t_1) \,)_{L^2_\sigma}$$

for $m \to \infty$. Thus for any $\widetilde{\varphi}$ of the above form we constructed a $u \in L^\infty L^2$ with $\nabla_x u \in L^2 L^2_{1+\sigma}$ such that

$$-\int_I (\, u \mid \partial_t \widetilde{\varphi} \,)_{L^2_\sigma} \, d\mathcal{L} + \int_I (\, \nabla_x u \mid \nabla_x \widetilde{\varphi} \,)_{L^2_{1+\sigma}} \, d\mathcal{L} = \int_I (\, f \mid \widetilde{\varphi} \,)_{L^2_\sigma} \, d\mathcal{L} + (\, g \mid \widetilde{\varphi}(t_1) \,)_{L^2_\sigma}.$$

But since $\operatorname{span}\{w_i;\, i \in \mathbb{N}\}$ is dense in $W^{1,2}_\sigma$ and we let m tend to ∞, for any fixed t we know that functions of the form of $\widetilde{\varphi}(t)$ are dense in $W^{1,2}_\sigma$. The equality then certainly holds for all test functions and so we obtained a σ-solution to f and g on $\overline{I} \times \Omega$, seeing that $L^\infty(I) \subset L^2_{loc}(\overline{I})$. ∎

3.4 Remark *The construction especially yields the existence of a σ-solution to the initial value problem with the additional property $u \in L^\infty(I;\, L^2_\sigma(\Omega))$. Since we are interested in the weakest possible definition of solution, we did not add this to the defining properties. However, the next result will show that boundedness of any σ-solution to an initial value problem for $\Omega = \overline{H}$ follows from the definition. See also Remark* 3.6 *later.*

Let us now consider $\Omega = \overline{H}$, the only relatively open subset of $\overline{H}$ that is closed as a subset of $\mathbb{R}^n$. Since then $C_c^\infty(\overline{H})$ is dense in $W^{1,2}_\sigma(H)$ by Remark 2.15, any $\varphi \in L^2\big(I;\, L^2_\sigma(H)\big)$ with

$\nabla_x \varphi \in L^2\big(I;\, L^2_{1+\sigma}(H)\big)$ and $\partial_t \varphi \in L^2\big(I;\, L^2_\sigma(H)\big)$ that has compact support in time contained in I (in $\overline{I}$) serves as an admissible test function for the time-local equation (the initial value problem). The compact temporal support together with the temporal square-integrability of both φ and $\partial_t \varphi$ implies $\varphi \in L^\infty\big(I;\, L^2_\sigma(H)\big)$, so that the defining equation remains reasonable with such test functions for the class of inhomogeneities considered in the definition.
With this remark, almost all properties of test functions are met by a time-localised σ-solution on $I \times \overline{H}$ itself. Only regularity of the temporal derivative is missing. In the proof of the energy identity that follows next, we therefore apply the weak regularisation from Remark 2.4 in the temporal variable to be able to use a σ-solution as a test function.

3.5 Proposition *Let $t_1 > -\infty$, $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $f \in L^1\big(I;\, L^2_\sigma(H)\big)$ and $g \in L^2_\sigma(H)$.*
If u is a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value g, then $u \in C\big(\overline{I};\, L^2_\sigma(H)\big)$ with $u(t_1) = g$ and for any $\widetilde{I} = (\widetilde{t}_1, \widetilde{t}_2) \subset I$ we have the energy identity

$$\frac{1}{2}\, \big\|u(\widetilde{t}_2)\big\|^2_{L^2_\sigma(H)} + \int_{\widetilde{I}} \|\nabla_x u\|^2_{L^2_{1+\sigma}(H)}\ d\mathcal{L} = \frac{1}{2}\, \big\|u(\widetilde{t}_1)\big\|^2_{L^2_\sigma(H)} + \int_{\widetilde{I}} (\, f \mid u\,)_{L^2_\sigma(H)}\ d\mathcal{L}.$$

Proof: First fix $\widetilde{t}_1 \le \widetilde{t}_2 \in I$, so that $\widetilde{I} \subsetneq I$. In a formal calculation with $\chi_{\widetilde{I}}\, u$ as a testing function the energy identity follows immediatley. However, both $\chi_{\widetilde{I}}$ and u are missing the L^2-regularity of the time derivative needed for a justification.
Therefore we regularise in time as in Remark 2.4 to make this approach rigorous. The regularity of the given σ-solution u carries over to its regularisation u^h and we thus have $u^h \in L^2_{loc} L^2_\sigma$ and $\nabla_x u^h \in L^2 L^2_{1+\sigma}$. But moreover, for any $h > 0$ the temporal derivatives satisfies $\partial_t u^h = D^h u \in L^2_{loc} L^2_\sigma$ and hence u^h has all the regularity we need to use it as a test function. As a consequence, for any cut-off function $\eta \in W^{1,2}_{loc}(I)$ with $\operatorname{supp} \eta \subset I^h \cap I^{-h}$ we have that $(\eta\, u^h)^{-h}$ is an admissible test function in the equation for u on $I \times \overline{H}$. Note that this is equivalent to the observation that u^h is a σ-solution to f^h on $(I^h \cap I^{-h}) \times \overline{H}$ and testing the equation for u^h with $\eta\, u^h$.
Taking into account that with a temporal integration by parts we get

$$\int_I u^h\, \partial_t(\eta\, u^h)\ d\mathcal{L} = \frac{1}{2} \int_I (u^h)^2\, \partial_t \eta\ d\mathcal{L},$$

this shows that

$$-\frac{1}{2} \int_I \partial_t \eta\ \left\|u^h\right\|^2_{L^2_\sigma}\ d\mathcal{L} + \int_I \eta\ \left\|\nabla_x u^h\right\|^2_{L^2_{1+\sigma}}\ d\mathcal{L} = \int_I \eta\, (\, f^h \mid u^h\,)_{L^2_\sigma}\ d\mathcal{L}.$$

For suitably small ε_1, $\varepsilon_2 > 0$ now define $\eta_\varepsilon := \eta_{\varepsilon_1}\, \eta_{\varepsilon_2}$ with

$$\eta_{\varepsilon_1}(t) := \begin{cases} 0 & \text{for all } t \in (t_1, \widetilde{t}_1) \\ \frac{t - \widetilde{t}_1}{\varepsilon_1} & \text{for all } t \in (\widetilde{t}_1, \widetilde{t}_1 + \varepsilon_1) \\ 1 & \text{for all } t \in (\widetilde{t}_1 + \varepsilon_1, \widetilde{t}_2) \end{cases}$$

and

$$\eta_{\varepsilon_2}(t) := \begin{cases} 1 & \text{for all } t \in (t_1, \widetilde{t}_2 - \varepsilon_2) \\ -\frac{t - \widetilde{t}_2}{\varepsilon_2} & \text{for all } t \in (\widetilde{t}_2 - \varepsilon_2, \widetilde{t}_2) \\ 0 & \text{for all } t \in (\widetilde{t}_2, t_2). \end{cases}$$

Since $\widetilde{I} \subset I^h \cap I^{-h}$ for h small enough it is clear that $\operatorname{supp} \eta_\varepsilon \subset I^h \cap I^{-h}$ for $\varepsilon_1, \varepsilon_2$ small enough and we have $\eta_\varepsilon \to \chi_{\widetilde{I}}$ in $L^2(I)$ as well as

$$\partial_t \eta_\varepsilon = \begin{cases} \frac{1}{\varepsilon_1} & \text{on } (\widetilde{t}_1, \widetilde{t}_1 + \varepsilon_1) \\ -\frac{1}{\varepsilon_2} & \text{on } (\widetilde{t}_2 - \varepsilon_2, \widetilde{t}_2) \\ 0 & \text{elsewhere.} \end{cases}$$

This explicit computation enables us to see that

$$\begin{aligned} \int_I \partial_t \eta_\varepsilon \left\| u^h \right\|_{L^2_\sigma}^2 \, d\mathcal{L} &= \frac{1}{\varepsilon_1} \int_{(\widetilde{t}_1, \widetilde{t}_1 + \varepsilon_1)} \left\| u^h \right\|_{L^2_\sigma}^2 \, d\mathcal{L} - \frac{1}{\varepsilon_2} \int_{(\widetilde{t}_2 - \varepsilon_2, \widetilde{t}_2)} \left\| u^h \right\|_{L^2_\sigma}^2 \, d\mathcal{L} \\ &= (\left\| u^h(\widetilde{t}_1) \right\|_{L^2_\sigma}^2)^{\varepsilon_1} - (\left\| u^h(\widetilde{t}_2) \right\|_{L^2_\sigma}^2)^{-\varepsilon_2}. \end{aligned}$$

Putting this together, we arrive at

$$\begin{aligned} &-\frac{1}{2} (\left\| u^h(\widetilde{t}_1) \right\|_{L^2_\sigma}^2)^{\varepsilon_1} + \frac{1}{2} (\left\| u^h(\widetilde{t}_2) \right\|_{L^2_\sigma}^2)^{-\varepsilon_2} \\ &\quad = -\int_I \eta_\varepsilon \left\| \nabla_x u^h \right\|_{L^2_{1+\sigma}}^2 \, d\mathcal{L} + \int_I \eta_\varepsilon \, (\, f^h \mid u^h \,)_{L^2_\sigma} \, d\mathcal{L}. \end{aligned}$$

This is a regularised version of the energy identity. Here, in all terms the limit $h \to 0$ poses no difficulties.
In the next step note that $\widetilde{t}_2 \mapsto (\|u(\widetilde{t}_2)\|_{L^2_\sigma}^2)^{-\varepsilon_2}$ is a sequence in $C((\widetilde{t}_1, t_2))$. Moreover, for arbitrary $\varepsilon_2, \varepsilon_2'$ we have

$$\begin{aligned} &\sup_{\widetilde{t}_2 \in (\widetilde{t}_1, t_2)} \left| (\|u(\widetilde{t}_2)\|_{L^2_\sigma}^2)^{-\varepsilon_2} - (\|u(\widetilde{t}_2)\|_{L^2_\sigma}^2)^{-\varepsilon_2'} \right| \\ &\qquad = 2 \sup_{\widetilde{t}_2 \in (\widetilde{t}_1, t_2)} \left| \int_I \eta_{\varepsilon_1} \, (\eta_{\varepsilon_2} - \eta_{\varepsilon_2'}) \, ((\, f \mid u \,)_{L^2_\sigma} - \|\nabla_x u\|_{L^2_{1+\sigma}}^2) \, d\mathcal{L} \right| \\ &\qquad \lesssim \sup_{\widetilde{t}_2 \in (\widetilde{t}_1, t_2)} \int_{(\widetilde{t}_2 - \max\{\varepsilon_2, \varepsilon_2'\}, \widetilde{t}_2)} |(\, f \mid u \,)_{L^2_\sigma}| + \|\nabla_x u\|_{L^2_{1+\sigma}}^2 \, d\mathcal{L}. \end{aligned}$$

This implies that $(\|u(\widetilde{t}_2)\|_{L^2_\sigma}^2)^{-\varepsilon_2}$ is a Cauchy sequence in $C((\widetilde{t}_1, t_2))$ and thus has a continuous limit. But in any Lebesgue point we have that $(\|u(\widetilde{t}_2)\|_{L^2_\sigma}^2)^{-\varepsilon_2} \to \|u(\widetilde{t}_2)\|_{L^2_\sigma}^2$ for $\varepsilon_2 \to 0$, since $t \mapsto \|u(t)^2\|_{L^2_\sigma} \in L^2(I)$. After the same considerations for $\widetilde{t}_1$ the continuity on I follows and the energy identity is immediate.
If we repeate the whole process with the test function $\eta_\varepsilon \, \widetilde{\varphi}$ for $\widetilde{\varphi} \in C_c^\infty(\overline{H})$, it becomes clear that $t \mapsto (\, u \mid \widetilde{\varphi} \,)_{L^2_\sigma}$ is bounded and continuous on $\overline{[\widetilde{t}_1, t_2)}$ for any $\widetilde{t}_1 \in I$. By density, this is nothing but weak continuity of $t \mapsto u(t) \in L^2_\sigma$. Since L^2_σ is a Hilbert space, this amounts to the continuity of $t \mapsto u(t) \in L^2_\sigma$ on $\overline{[\widetilde{t}_1, t_2)}$ for any $\widetilde{t}_1 \in I$ and thus $u \in C\big(I; L^2_\sigma(H)\big)$. For the inclusion of the initial value fix a $t_0 \in I$ and define

$$\eta_\varepsilon := \begin{cases} 1 & \text{for all } t \in (t_1, t_0) \\ 1 - \frac{t - t_0}{\varepsilon} & \text{for all } t \in (t_0, t_0 + \varepsilon) \\ 0 & \text{for all } t \in (t_0 + \varepsilon, t_2) \end{cases} \quad \text{for small } \varepsilon > 0.$$

Then $\widetilde{\varphi}\,\eta_\varepsilon$ is an admissible test function, and the same calculations as above lead to

$$((\, u(t_0) \mid \widetilde{\varphi}\,)_{L^2_\sigma})^\varepsilon = -\int_I \eta_\varepsilon\,(\, \nabla_x u \mid \nabla_x \widetilde{\varphi}\,)_{L^2_{1+\sigma}}\; d\mathcal{L} + \int_I \eta_\varepsilon\,(\, f \mid \widetilde{\varphi}\,)_{L^2_\sigma}\; d\mathcal{L} + (\, g \mid \widetilde{\varphi}\,)_{L^2_\sigma}.$$

In view of the result from the first part we can now let $\varepsilon \to 0$ and then $t_0 \to t_1$ to get

$$\lim_{t_0 \to t_1} (\, u(t_0) \mid \widetilde{\varphi}\,)_{L^2_\sigma} = (\, g \mid \widetilde{\varphi}\,)_{L^2_\sigma}.$$

This proves that weak continuity can be extended into t_1. But on (t_1, t_2) we know by the first part that $\|u(t)\|_{L^2_\sigma}$ is uniformely bounded. Full continuity in L^2_σ down to t_1 is therefore proven. ∎

3.6 Remark *The energy identity* 3.5 *also shows that for a σ-solution u to the initial value problem with $f \in L^1\big(I;\, L^2_\sigma(H)\big)$ and $g \in L^2_\sigma(H)$ we have*

$$\sup_{t \in (\widetilde{t}_1, t_2)} \|u(t)\|^2_{L^2_\sigma} + \int_{(\widetilde{t}_1, t_2)} \|\nabla_x u\|^2_{L^2_{1+\sigma}(H)}\; d\mathcal{L} \le 2\, \|g\|^2_{L^2_\sigma(H)} + 4\, \|f\|^2_{L^1\left((\widetilde{t}_1, t_2);\, L^2_\sigma(H)\right)}$$

for any $\widetilde{t}_1 \in \overline{I}$. This implies $u \in C_b\big(I;\, L^2_\sigma(H)\big) \subset L^\infty\big(I;\, L^2_\sigma(H)\big)$ as "predicted" by the Galerkin approximation (compare Remark 3.4*).*
Note also that for time-local solutions we do not necessarily get boundedness near the initial time t_1, but an evaluation at the terminal time t_2 is always possible as long as t_2 is finite.

An easy consequence of the energy identity 3.5 is the norm-decrease of σ-solutions of the homogeneous problem.

3.7 Corollary *Let $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let $g \in L^2_\sigma(H)$.*
If u is a σ-solution to $f = 0$ on $\ulcorner\!I \times \overline{H}$ with initial value g, then the function $t \mapsto \|u(t)\|_{L^2_\sigma(H)}$ is monotonically decreasing on $\overline{I}$.

Uniqueness of initial value solutions now follows directly.

3.8 Proposition *Let $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let $f \in L^1\big(I;\, L^2_\sigma(H)\big)$ and $g \in L^2_\sigma(H)$.*
Then there is exactly one σ-solutions to f on $\ulcorner\!I \times \overline{H}$ with initial value g.

Proof: Given two σ-solutions u_1 and u_2 to f on $\ulcorner\!I \times \overline{H}$ with initial value g, the difference $u_1 - u_2$ is a σ-solution to 0 with initial value 0. Since the right hand side vanishes, we know with corollary 3.7 that the norm of u is bounded by the the norm of the initial value. But this is also zero, hence $\|u_1 - u_2\|_{L^2_\sigma} = 0$ and we must have $u_1 = u_2$. ∎

We can finally use the continuity from Proposition 3.5 to show existence and uniqueness of solutions in the only case not covered yet, that is for $t_1 = -\infty$.

3.9 Proposition *Let $I := (-\infty, t_2) \subset \mathbb{R}$ and $f \in L^1\big(I;\, L^2_\sigma(H)\big)$.*
Then there exists exactly one σ-solution to f on $I \times \overline{H}$ with $\lim\limits_{t \to -\infty} \|u(t)\|_{L^2_\sigma(H)} = 0$ and the energy identity 3.5 *holds for any $\widetilde{I} = (\widetilde{t}_1, \widetilde{t}_2) \subset I$.*

Proof: Consider a sequence $t^j \in I$ with $t^j \to -\infty$ as $j \to \infty$. Then by the existence theorem 3.3 and the energy identity 3.5 there are σ-solutions $u^j \in C(\overline{[t^j, t_2)}; L^2_\sigma(H))$ to f with $u^j(t^j) = 0$ for any of these t^j. Extending u^j by zero onto I therefore delivers a sequence $(u^j)_{j\in\mathbb{N}} \subset C(I; L^2_\sigma(H))$. But for indices j, j', where we assume $j < j'$ without loss of generality, the difference $u^j - u^{j'}$ is a σ-solution to $\chi_{(t^{j'},t^j)} f$ on $[t^{j'}, t_2)$ with $(u^j - u^{j'})(t^{j'}) = 0$. Hence by the energy identity we get

$$\sup_{t\in I} \left\| u^j(t) - u^{j'}(t) \right\|_{L^2_\sigma(H)} \leq \int\limits_{(t^{j'},t^j)} \|f\|_{L^2_\sigma(H)} \, d\mathcal{L},$$

and by virtue of $t \mapsto \|f(t)\|_{L^2_\sigma(H)} \in L^1(I)$ this is a Cauchy sequence. It follows that $\lim\limits_{j\to\infty} u^j =: u$ exists in $C(I; L^2_\sigma(H))$ and is a σ-solution to f on I with

$$\left\| u(t^j) \right\|_{L^2_\sigma(H)} = \left\| u(t^j) - u^j(t^j) \right\|_{L^2_\sigma(H)} \leq \sup_{t\in I} \left\| u(t) - u^j(t) \right\|_{L^2_\sigma(H)} \to 0$$

for $j \to \infty$. The energy identity is then obvious and uniqueness follows by the decrease of the norm as before in corollary 3.7. ∎

Henceforth, we will include the case $t_1 = -\infty$ into the notation for the initial value problem: Whenever we speak of a σ-solutions to f on $\overline{I} \times \overline{H}$ with intial value g for $t_1 = -\infty$, we implicitely set $g = 0$ and mean the σ-solution to f on $(-\infty, t_2) \times \overline{H}$ that vanishes at $-\infty$. Uniqueness is then given for the initial value problem in any case, so we can always refer to "the" σ-solution to f with initial value g. Note also that for $t_1 = -\infty$ and $f = 0$ we deal with the trivial solution $u = 0$.

For the homogeneous equation we can use the same method of proof as in the energy identity with a second σ-solution as test function to get a duality result that will be used later.

3.10 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and g_1, $g_2 \in L^2_\sigma(H)$.*
If u_1 is a σ-solutions to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g_1, and if u_2 is a σ-solutions to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g_2, then for any $\widetilde{t}_1 \leq \widetilde{t}_2 \in \overline{I}$ we have the duality equality

$$(\, u_1(\widetilde{t}_1) \mid u_2(\widetilde{t}_2) \,)_{L^2_\sigma(H)} = (\, u_1(\widetilde{t}_2) \mid u_2(\widetilde{t}_1) \,)_{L^2_\sigma(H)}.$$

Proof: Fix $\widetilde{t}_1 \leq \widetilde{t}_2 \in \overline{I}$ and define

$$T : I \to \mathbb{R},\ t \mapsto \widetilde{t}_1 + \widetilde{t}_2 - t =: s.$$

Note that if $\widetilde{I} := (\widetilde{t}_1, \widetilde{t}_2)$ then $T|_{\widetilde{I}}$ maps $\widetilde{I}$ bijectively onto $\widetilde{I}$ and we have $(u_2 \circ T)(\widetilde{t}_1) = u_2(\widetilde{t}_2)$ as well as $(u_2 \circ T)(\widetilde{t}_2) = u_2(\widetilde{t}_1)$.
Let us now test the equation for u_1 formally with $\chi_{\widetilde{I}}(u_2 \circ T)$. We obtain

$$\begin{aligned} 0 = -\int\limits_I \partial_t \chi_{\widetilde{I}} \int\limits_H u_1\,(u_2 \circ T)\, d\mu_\sigma\, d\mathcal{L} &- \int\limits_I \chi_{\widetilde{I}} \int\limits_H u_1\, \partial_t(u_2 \circ T)\, d\mu_\sigma\, d\mathcal{L} \\ &+ \int\limits_I \chi_{\widetilde{I}} \int\limits_H \nabla_x u_1 \cdot \nabla_x(u_2 \circ T)\, d\mu_{1+\sigma}\, d\mathcal{L}. \end{aligned}$$

An integration by parts followed by a transformation of the integral, using that $\chi_{\widetilde{I}}$ vanishes outside $\widetilde{I}$ and that T and T^{-1} map $\widetilde{I}$ onto itself, yield

$$\begin{aligned} -\int_I \chi_{\widetilde{I}}\, u_1\, \partial_t(u_2 \circ T)\, d\mathcal{L} &= -\int_I \left(\partial_t(u_1\, \chi_{\widetilde{I}}) \circ T^{-1}\right) u_2\, d\mathcal{L} \\ &= \int_I \partial_s\left((u_1\, \chi_{\widetilde{I}}) \circ T^{-1}\right) u_2\, d\mathcal{L}. \end{aligned}$$

Consequently, considering $(u_1\, \chi_{\widetilde{I}}) \circ T^{-1}$ as a test function for u_2 we see that

$$-\int_I \chi_{\widetilde{I}} \int_H u_1\, \partial_t(u_2 \circ T)\, d\mu_\sigma\, d\mathcal{L} = -\int_I \int_H \nabla_x\left((\chi_{\widetilde{I}}\, u_1) \circ T^{-1}\right) \cdot \nabla_x u_2\, d\mu_\sigma\, d\mathcal{L}$$

and by reversing the transformation and inserting the result above we finally get

$$0 = -\big(\, u_1(\widetilde{t}_1) \,|\, (u_2 \circ T)(\widetilde{t}_1) \,\big)_{L^2_\sigma(H)} + \big(\, u_1(\widetilde{t}_2) \,|\, (u_2 \circ T)(\widetilde{t}_2) \,\big)_{L^2_\sigma(H)}.$$

This can be made rigorous in the same way as the proof of Proposition 3.5. ∎

4 Energy Estimates

We now use similar regularisation techniques as before to show weighted L^2-boundedness of derivatives of σ-solutions. In order to do so, we have to impose extra conditions onto the right hand side. Furthermore, for reaching the initial point in time we also need additional regularity conditions for the initial value. Therefore, in favor of a clearer presentation we first consider only the time-local case.

It turns out that the weak choice of prerequisits with respect to time in the definition of a σ-solution does not constitute a real restriction if $\Omega = H$.

4.1 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2\big(I;\, L^2_\sigma(H)\big)$.*
If u is a σ-solution to f on $I \times \overline{H}$, then

$$t \mapsto \|\nabla_x u(t)\|_{L^2_{1+\sigma}(H)} \in C([\widetilde{t}_1, t_2))$$

and we have

$$\int\limits_{(\widetilde{t}_1, t_2)} \|\partial_t u\|^2_{L^2_\sigma(H)}\; d\mathcal{L} \le \big\|\nabla_x u(\widetilde{t}_1)\big\|^2_{L^2_{1+\sigma}(H)} + \int\limits_{(\widetilde{t}_1, t_2)} \|f\|^2_{L^2_\sigma(H)}\; d\mathcal{L}$$

for any $\widetilde{t}_1 \in I$.

Proof: Fix $\widetilde{t}_1 \le \widetilde{t}_2 \in I$ so that $\widetilde{I} \subsetneq I$. In the formal proof we test the equation with $\chi_{\widetilde{I}}\, \partial_t u$.
To make this rigorous note that for any σ-solution u with the additional property $\partial_t u \in L^2_{loc} L^2_\sigma$, by an integration by parts in time and a density argument we equivalently know that

$$\int\limits_I (\, \partial_t u \mid \varphi\,)_{L^2_\sigma}\; d\mathcal{L} + \int\limits_I (\, \nabla_x u \mid \nabla_x \varphi\,)_{L^2_{1+\sigma}}\; d\mathcal{L} = \int\limits_I (\, f \mid \varphi\,)_{L^2_\sigma}\; d\mathcal{L}$$

for all $\varphi \in L^2 L^2_\sigma$ with $\nabla_x \varphi \in L^2 L^2_{1+\sigma}$ and compact temporal support. Since u^h is such a σ-solution with regular time derivative - see the proof of Proposition 3.5 - we can use this formulation with the test function $\eta_\varepsilon\, \partial_t u^h$, where $\eta_\varepsilon = \eta_{\varepsilon_1} \eta_{\varepsilon_2}$ also as before. Taking into account that

$$\int\limits_I \eta_\varepsilon\, \nabla_x u^h \cdot \nabla_x \partial_t u^h\; d\mathcal{L} = -\frac{1}{2} \int\limits_I \partial_t \eta_\varepsilon\, |\nabla_x u^h|^2\; d\mathcal{L}$$

with an integration by parts in t, it follows that

$$\int\limits_I \eta_\varepsilon\, \left\|\partial_t u^h\right\|^2_{L^2_\sigma}\; d\mathcal{L} - \frac{1}{2}\, (\left\|\nabla_x u^h(\widetilde{t}_1)\right\|^2_{L^2_{1+\sigma}})^{\varepsilon_1} + \frac{1}{2}\, (\left\|\nabla_x u^h(\widetilde{t}_2)\right\|^2_{L^2_{1+\sigma}})^{-\varepsilon_2} = \int\limits_I \eta_\varepsilon\, (\, f^h \mid \partial_t u^h\,)_{L^2_\sigma}\; d\mathcal{L}.$$

Hölder's inequality combined with Young's inqeuality then reveals that

$$\frac{1}{2} \int\limits_I \eta_\varepsilon\, \left\|\partial_t u^h\right\|^2_{L^2_\sigma}\; d\mathcal{L} \le \frac{1}{2} \int\limits_I \eta_\varepsilon\, \left\|f^h\right\|^2_{L^2_\sigma}\; d\mathcal{L} + \frac{1}{2}\, (\left\|\nabla_x u^h(\widetilde{t}_1)\right\|^2_{L^2_{1+\sigma}})^{\varepsilon_1}.$$

At least if $\widetilde{t}_1$ is a Lebesgue point, the limits can now be taken. Going back to the regularised equation above, with a Cauchy argument as in the proof of Proposition 3.5 it is then possible to show the continuity of $t \mapsto \|\nabla_x u(t)\|_{L^2(H,\mu_{1+\sigma})}$ on $[\widetilde{t}_1, t_2)$ as well as the convergence of the equation itself. ∎

4.2 Remark *The weak regularity gain in the temporal derivative enables us to restrict ourselves to elliptic equations if $\Omega = H$: If u is a σ-solution to f on $I \times \overline{H}$, then for almost all $t \in I$ we have that $u(t)$ satisfies*

$$-L_\sigma u(t) = f(t) - \partial_t u(t) =: \widetilde{f}(t) \text{ on } \overline{H}.$$

In a slight abuse of notation we will thus supress the time dependence and consider $u \in W^{1,2}_\sigma(H)$ with

$$\int_H \nabla_x u \cdot \nabla_x \varphi \, d\mu_{1+\sigma} = \int_H \widetilde{f} \, \varphi \, d\mu_\sigma,$$

$\varphi \in C_c^\infty(\overline{H})$, as the energy formulation of the elliptic equation. As before, the density statement 2.15 *implies that any function in $W^{1,2}_\sigma(H)$ is an admissible test function.*

4.3 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2\big(I; L^2_\sigma(H)\big)$. If u is a σ-solution to f on $I \times \overline{H}$, then we have*

$$\int_{(\widetilde{t}_1,t_2)} \|\nabla_x u\|^2_{L^2_\sigma(H)} \, d\mathcal{L} + \int_{(\widetilde{t}_1,t_2)} \left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(H)} \, d\mathcal{L} \lesssim_{n,\sigma} \int_{(\widetilde{t}_1,t_2)} \|f\|^2_{L^2_\sigma(H)} \, d\mathcal{L} + \left\|\nabla_x u(\widetilde{t}_1)\right\|^2_{L^2_{1+\sigma}(H)}$$

for any $\widetilde{t}_1 \in I$.

Proof: Formally, the estimate for both the tangential and vertical derivatives follows by testing the elliptic equation with $\partial_{x_n} u$, while for the second order derivative we consider $x_n \Delta_x u$ as a test function. Making this rigorous in the same way as above requires an analysis of certain commutators. We use a transformation onto an ordinary differential equation instead.
So starting with the energy formulation of the elliptic equation

$$(\,\cdot\,)_n \, \Delta_x u + (1+\sigma)\, \partial_{x_n} u = -\widetilde{f}$$

on H as in Remark 4.2, we perform a Fourier transformation in the tangential directions with Fourier variable $\xi' \in \mathbb{R}^{n-1}$ and without renaming the functions u and $\widetilde{f}$ to get

$$-(\,\cdot\,)_n \, |\xi'|^2 \, u + (\,\cdot\,)_n \, \partial^2_{x_n} u + (1+\sigma)\, \partial_{x_n} u = -\widetilde{f}.$$

The transformation $z := |\xi'| \, x_n$ then yields

$$(\,\cdot\,) \, |\xi'| \, \partial_z^2 u + (1+\sigma) \, |\xi'| \, \partial_z u - (\,\cdot\,) \, |\xi'| \, u = -\widetilde{f}$$

with $u = u(\xi', z)$ and $\widetilde{f} = \widetilde{f}(\xi', z)$, where $(\,\cdot\,) : z \mapsto z$ represents the identity operator on $\mathbb{R}$. We consider ξ' as parameters and rename $|\xi'| \, u(\xi', z)$ to $u(z)$, reaching the equation

$$(*) \qquad z \, \partial_z^2 u + (1+\sigma) \, \partial_z u - z \, u = -\widetilde{f}$$

on $(0, \infty)$.
A solution u of the homogeneous version of this equation defines a solution $v = z^{\frac{\sigma}{2}} \, u$ of the

modified Bessel equation with parameter $\frac{\sigma}{2}$, that is

$$z^2\,\partial_z^2 v + z\,\partial_z v - z^2\,v - \frac{\sigma^2}{4}\,v = 0,$$

and vice versa. The modified Bessel functions $I_{\frac{\sigma}{2}}$ and $K_{\frac{\sigma}{2}}$, described in detail in [OM10], form a fundamental system of this ordinary differential equation, hence a fundamental system for the homogeneous equation $(*)$ is given by

$$\Psi_1(z) := z^{-\frac{\sigma}{2}}\,I_{\frac{\sigma}{2}}$$

and

$$\Psi_2(z) := z^{-\frac{\sigma}{2}}\,K_{\frac{\sigma}{2}}.$$

The asymptotics of the modified Bessel functions are known, and up to constants depending on σ, for $\sigma > 0$ we get

$$\begin{aligned} \Psi_1(z) &\sim 1 \quad (z \to 0), & \Psi_1(z) &\sim z^{-\frac{1+\sigma}{2}} e^{z} \quad (z \to \infty) \\ \Psi_2(z) &\sim z^{-\sigma} \quad (z \to 0), & \Psi_2(z) &\sim z^{-\frac{1+\sigma}{2}} e^{-z} \quad (z \to \infty). \end{aligned}$$

For $\sigma < 0$ we have $\Psi_2(z) \sim 1 \quad (z \to 0)$, and $\Psi_2(z) \sim \ln z \quad (z \to 0)$ for $\sigma = 0$, while the other three relations remain as before. The Wronskian of Ψ_1 and Ψ_2 can be computed to be $z^{-1-\sigma}$.
All this leads to the fundamental solution

$$k(z,y) := \begin{cases} y^{\sigma}\,\Psi_1(z)\,\Psi_2(y), & z < y \\ y^{\sigma}\,\Psi_1(y)\,\Psi_2(z), & z > y, \end{cases}$$

with first order derivative having a jump discontinuity of the type y^{-1} at $z = y$. Therefore, solutions u to $(*)$ are characterised by the representation

$$z^l\,u(z) = -\int\limits_{(0,\infty)} z^l\,k(z,y)\,\widetilde{f}(y)\,d\mathcal{L}(y)$$

for any $l \in \mathbb{R}$.
We rewrite this to get the operator

$$\widetilde{f} \mapsto z^l\,u = -\int\limits_{(0,\infty)} z^l\,k(z,y)\,y^{-\sigma}\,\widetilde{f}(y)\,d\mu_\sigma(y).$$

If both

$$\sup_{z\in(0,\infty)} \int\limits_{(0,\infty)} z^l\,|k(z,y)|\,y^{-\sigma}\,d\mu_\sigma(y) < \infty$$

and

$$\sup_{y\in(0,\infty)} \int\limits_{(0,\infty)} z^l\,|k(z,y)|\,y^{-\sigma}\,d\mu_\sigma(z) < \infty,$$

then Schur's lemma A.20 ensures that

$$\|u\|_{L^2_{2l+\sigma}((0,\infty))} \lesssim \left\|\widetilde{f}\right\|_{L^2_\sigma((0,\infty))}.$$

To see this, fix a small $\varepsilon > 0$ and a large $R > 0$. We start with the first condition and divide the range of each supremum into three disjoint pieces. The definition of the fundamental solution and the asymptotic expansions of Ψ_1 and Ψ_2 for $\sigma > 0$ then show that

$$\begin{aligned}
&\sup_{z\in(0,\varepsilon)} \int\limits_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(y) \\
&\quad= \sup_{z\in(0,\varepsilon)} z^l \Bigg(|\Psi_2(z)| \int\limits_{(0,z)} y^\sigma \, |\Psi_1(y)| \, d\mathcal{L}(y) + |\Psi_1(z)| \int\limits_{(z,\varepsilon)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) \\
&\qquad + |\Psi_1(z)| \int\limits_{(\varepsilon,R)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) + |\Psi_1(z)| \int\limits_{(R,\infty)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) \Bigg) \\
&\quad\lesssim \sup_{z\in(0,\varepsilon)} \Bigg(z^{l-\sigma} \int\limits_{(0,z)} y^\sigma \, d\mathcal{L}(y) + z^l \int\limits_{(z,\varepsilon)} 1 \, d\mathcal{L}(y) \\
&\qquad + z^l \int\limits_{(\varepsilon,R)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) + z^l \int\limits_{(R,\infty)} y^{\frac{\sigma-1}{2}} \, e^{-y} \, d\mathcal{L}(y) \Bigg).
\end{aligned}$$

The first integral can be computed since $\sigma > -1$. We also evaluate the second integral explicitely, while the third one is surely bounded depending only on ε, R and σ, and the same is true for the last one with a constant depending on R and σ. We find that

$$\sup_{z\in(0,\varepsilon)} \int\limits_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(y) \lesssim_{\sigma,\varepsilon,R} \sup_{z\in(0,\varepsilon)} \left(z^{l+1} + z^l \right)$$

which is bounded for $l \geq 0$. The cases $\sigma \leq 0$ follow in the same fashion with an additional additive term $z^{l+\sigma+1}$ for $\sigma < 0$, and $z^{l+1} \ln z$ for $\sigma = 0$. Also this is bounded near 0 for $l \geq 0$ in both cases.

The interval (ε, R) is treated straightforwardly. Here we get

$$\begin{aligned}
&\sup_{z\in(\varepsilon,R)} z^l \int\limits_{(0,\infty)} |k(z,y)| \, d\mathcal{L}(y) \\
&\quad= \sup_{z\in(\varepsilon,R)} z^l \Bigg(|\Psi_2(z)| \int\limits_{(0,\varepsilon)} y^\sigma \, |\Psi_1(y)| \, d\mathcal{L}(y) + |\Psi_2(z)| \int\limits_{(\varepsilon,z)} y^\sigma \, |\Psi_1(y)| \, d\mathcal{L}(y) \\
&\qquad + |\Psi_1(z)| \int\limits_{(z,R)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) + |\Psi_1(z)| \int\limits_{(R,\infty)} y^\sigma | \, \Psi_2(y)| \, d\mathcal{L}(y) \Bigg) \\
&\quad\lesssim \sup_{z\in(\varepsilon,R)} z^l \Bigg(|\Psi_2(z)| \int\limits_{(0,\varepsilon)} y^\sigma \, d\mathcal{L}(y) + |\Psi_2(z)| \int\limits_{(\varepsilon,z)} y^\sigma \, |\Psi_1(y)| \, d\mathcal{L}(y) \\
&\qquad + |\Psi_1(z)| \int\limits_{(z,R)} y^\sigma \, |\Psi_2(y)| \, d\mathcal{L}(y) + |\Psi_1(z)| \int\limits_{(R,\infty)} y^{\frac{\sigma-1}{2}} \, e^{-y} \, d\mathcal{L}(y) \Bigg).
\end{aligned}$$

As before, all the integrals are bounded by a constant depending only on σ, ε and R.
Finally, on (R,∞) we have

$$\begin{aligned}
&\sup_{z\in(R,\infty)} \int_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(y) \\
&\quad \lesssim \sup_{z\in(R,\infty)} \Bigg(z^{l-\frac{1+\sigma}{2}} \, e^{-z} \int_{(0,\varepsilon)} y^\sigma \, d\mathcal{L}(y) + z^{l-\frac{1+\sigma}{2}} \, e^{-z} \int_{(\varepsilon,R)} y^\sigma \, |\Psi_1(y)| \, d\mathcal{L}(y) \\
&\qquad + z^{l-\frac{1+\sigma}{2}} \, e^{-z} \int_{(R,z)} y^{\frac{\sigma-1}{2}} \, e^y \, d\mathcal{L}(y) + z^{l-\frac{1+\sigma}{2}} \, e^{z} \int_{(z,\infty)} y^{\frac{\sigma-1}{2}} \, e^{-y} \, d\mathcal{L}(y) \Bigg).
\end{aligned}$$

The first two integrals are obviously bounded. But we can also compute

$$\sup_{z\in(R,\infty)} z^{l-\frac{1+\sigma}{2}} \, e^{-z} \int_{(R,z)} y^{\frac{\sigma-1}{2}} \, e^y \, d\mathcal{L}(y) = \sup_{z\in(R,\infty)} z^{l-1-\frac{\sigma-1}{2}} \, e^{-z} \int_{(R,z)} y^{\frac{\sigma-1}{2}} \, e^y \, d\mathcal{L}(y) \lesssim_{\sigma,R} \sup_{z\in(R,\infty)} z^{l-1}$$

and likewise

$$\sup_{z\in(R,\infty)} z^{l-\frac{1+\sigma}{2}} \, e^{z} \int_{(z,\infty)} y^{\frac{\sigma-1}{2}} \, e^{-y} \, d\mathcal{L}(y) = \sup_{z\in(R,\infty)} z^{l-1-\frac{\sigma-1}{2}} \, e^{z} \int_{(z,\infty)} y^{\frac{\sigma-1}{2}} \, e^{-y} \, d\mathcal{L}(y) \lesssim \sup_{z\in(R,\infty)} z^{l-1}.$$

This means that we have

$$\sup_{z\in(R,\infty)} \int_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(y) \lesssim_{\sigma,\varepsilon,R} \sup_{z\in(R,\infty)} \left(z^{l-\frac{1+\sigma}{2}} \, e^{-z} + z^{l-1} \right)$$

which is bounded whenever $l \leq 1$.
For the verification of the second condition we proceed along the same lines and get

$$\begin{aligned}
&\sup_{y\in(0,\varepsilon)} \int_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(z) \\
&\quad = \sup_{y\in(0,\varepsilon)} \Bigg(|\Psi_2(y)| \int_{(0,y)} z^{l+\sigma} \, |\Psi_1(z)| \, d\mathcal{L}(z) + |\Psi_1(y)| \int_{(y,\varepsilon)} z^{l+\sigma} \, |\Psi_2(z)| \, d\mathcal{L}(z) \\
&\qquad + |\Psi_1(y)| \int_{(\varepsilon,R)} z^{l+\sigma} \, |\Psi_2(z)| \, d\mathcal{L}(z) + |\Psi_1(y)| \int_{(R,\infty)} z^{l+\sigma} \, |\Psi_2(z)| \, d\mathcal{L}(z) \Bigg) \\
&\quad \lesssim \sup_{y\in(0,\varepsilon)} \Bigg(y^{-\sigma} \int_{(0,y)} z^{l+\sigma} \, d\mathcal{L}(z) + \int_{(y,\varepsilon)} z^l \, d\mathcal{L}(z) \\
&\qquad + \int_{(\varepsilon,R)} z^{l+\sigma} \, |\Psi_2(z)| \, d\mathcal{L}(z) + \int_{(R,\infty)} z^{l+\frac{\sigma-1}{2}} \, e^{-z} \, d\mathcal{L}(z) \Bigg) \\
&\quad \lesssim_{\sigma,\varepsilon,R} \sup_{y\in(0,\varepsilon)} \left(y^{l+1} + 1 \right)
\end{aligned}$$

for $\sigma \geq 0$. Here, for $\sigma < 0$ we get the additional term $y^{l+1+\sigma}$, while for $\sigma = 0$ and $l \in \mathbb{N}_0$ we have an extra $y^{l+1} \ln y$. Any of these summands is bounded on $(0,\varepsilon)$ if $l \geq 0$.

Furthermore we have

$$
\begin{aligned}
&\sup_{y\in(R,\infty)} \int\limits_{(0,\infty)} z^l \, |k(z,y)| \, y^{-\sigma} \, d\mu_\sigma(z) \\
&\quad = \sup_{y\in(R,\infty)} \Bigg(|\Psi_2(y)| \int\limits_{(0,\varepsilon)} z^{l+\sigma} \, |\Psi_1(z)| \, d\mathcal{L}(z) + |\Psi_2(y)| \int\limits_{(\varepsilon,R)} z^{l+\sigma} \, |\Psi_1(z)| \, d\mathcal{L}(z) \\
&\qquad + |\Psi_2(y)| \int\limits_{(R,y)} z^{l+\sigma} \, |\Psi_1(z)| \, d\mathcal{L}(z) + |\Psi_1(y)| \int\limits_{(y,\infty)} z^{l+\sigma} \, |\Psi_2(z)| \, d\mathcal{L}(z) \Bigg) \\
&\quad \lesssim \sup_{y\in(R,\infty)} \Bigg(y^{-\frac{\sigma+1}{2}} \, e^{-y} \int\limits_{(0,\varepsilon)} z^{l+\sigma} \, d\mathcal{L}(z) + y^{-\frac{\sigma+1}{2}} \, e^{-y} \int\limits_{(\varepsilon,R)} z^{l+\sigma} \, |\Psi_1(z)| \, d\mathcal{L}(z) \\
&\qquad + y^{-\frac{\sigma+1}{2}} \, e^{-y} \int\limits_{(R,y)} z^{l+\frac{\sigma-1}{2}} \, e^{z} \, d\mathcal{L}(z) + y^{-\frac{\sigma+1}{2}} \, e^{y} \int\limits_{(y,\infty)} z^{l+\frac{\sigma-1}{2}} \, e^{-z} \, d\mathcal{L}(z) \Bigg) \\
&\quad \lesssim_{\sigma,\varepsilon,R} \sup_{y\in(R,\infty)} \left(y^{-\frac{\sigma+1}{2}} \, e^{-y} + y^{l-1} \right)
\end{aligned}
$$

and this is bounded again if $l \leq 1$.
It follows that $l = 0$ and $l = 1$ are both admissible, so for solutions u of $(*)$ we get

$$\|u\|_{L^2_\sigma((0,\infty))} + \|(\,\cdot\,)\, u\|_{L^2_\sigma((0,\infty))} \lesssim_\sigma \left\|\widetilde{f}\right\|_{L^2_\sigma((0,\infty))}.$$

But then it makes sense to incorporate $(\,\cdot\,)\, u$ onto the right hand side of $(*)$. This results in a first order ordinary differential equation for $\partial_z u =: v$, namely

$$(**) \qquad (\,\cdot\,)\, \partial_z v + (1+\sigma)\, v = -\widetilde{f} + (\,\cdot\,)\, u =: \overline{f}.$$

A solution to the homogeneous equation is clearly given by $z \mapsto z^{-1-\sigma}$, and so for this equation we get the fundamental solution

$$\overline{k}(z,y) := y^\sigma \, z^{-1-\sigma} \text{ if } z > y$$

and 0 elsewhere. Now we consider the operator

$$z^\delta \, \overline{f}(z) \mapsto z^\delta \, v(z) = \int\limits_{(0,\infty)} \left(\frac{z}{y}\right)^\delta \overline{k}(z,y) \, y^\delta \, \overline{f}(y) \, d\mathcal{L}(y).$$

For $\delta = \delta_1$ we have

$$\sup_{z\in(0,\infty)} \int\limits_{(0,\infty)} \left(\frac{z}{y}\right)^{\delta_1} |\overline{k}(z,y)| \, d\mathcal{L}(y) = \sup_{z\in(0,\infty)} z^{\delta_1-1-\sigma} \int\limits_{(0,z)} y^{\sigma-\delta_1} \, d\mathcal{L}(y) = \frac{1}{1+\sigma-\delta_1}$$

if $\delta_1 < 1+\sigma$. This shows that

$$\left\|(\,\cdot\,)^{\delta_1} \, v\right\|_{L^\infty((0,\infty))} \lesssim_{\sigma,\delta_1} \left\|(\,\cdot\,)^{\delta_1} \, \overline{f}\right\|_{L^\infty((0,\infty))}$$

in this case.
Similarly, for a $\delta = \delta_2$ we see

$$\sup_{y\in(0,\infty)} \int\limits_{(0,\infty)} \left(\frac{z}{y}\right)^{\delta_2} |\bar{k}(z,y)| \, d\mathcal{L}(z) = \sup_{y\in(0,\infty)} y^{\sigma+\delta_2} \int\limits_{(y,\infty)} z^{\delta_2-1-\sigma} \, d\mathcal{L}(z) = -\frac{1}{\delta_2-\sigma}$$

if $\delta_2 < \sigma$, thus

$$\left\| (\,\cdot\,)^{\delta_2} v \right\|_{L^1((0,\infty))} \lesssim_{\sigma,\delta_2} \left\| (\,\cdot\,)^{\delta_2} \bar{f} \right\|_{L^1((0,\infty))}.$$

An interpolation yields

$$\left\| (\,\cdot\,)^{\frac{\delta_1+\delta_2}{2}} v \right\|_{L^2((0,\infty))} \lesssim_{\sigma,\delta_1,\delta_2} \left\| (\,\cdot\,)^{\frac{\delta_1+\delta_2}{2}} \bar{f} \right\|_{L^2((0,\infty))}$$

for $\delta_1 < 1+\sigma$ and $\delta_2 < \sigma$. But we can choose δ_1 and δ_2 with $\delta_1 + \delta_2 = \sigma$, if only $\sigma < 1 + 2\sigma$. This condition, however, is equivalent to $\sigma > -1$ and hence we get

$$\|\partial_z u\|_{L^2_\sigma((0,\infty))} = \|v\|_{L^2_\sigma((0,\infty))} \lesssim_\sigma \left\|\bar{f}\right\|_{L^2_\sigma((0,\infty))} \lesssim_\sigma \left\|\tilde{f}\right\|_{L^2_\sigma((0,\infty))}$$

with the first result for $(\,\cdot\,)\, u$.
An immediate consequence of $(**)$ is then

$$\left\|\partial_z^2 u\right\|_{L^2_{2+\sigma}((0,\infty))} = \left\|(\,\cdot\,)\,\partial_z^2 u\right\|_{L^2_\sigma((0,\infty))} \eqsim_\sigma \left\|\tilde{f} + \partial_z u + (\,\cdot\,)\, u\right\|_{L^2_\sigma((0,\infty))} \lesssim \left\|\tilde{f}\right\|_{L^2_\sigma((0,\infty))}.$$

Summing up, after reverting the notation back to the starting point we have shown that

$$\||\xi'|\, u\|_{L^2_\sigma((0,\infty))} + \||\xi'|\, u\|_{L^2_{2+\sigma}((0,\infty))} + \||\xi'|\,\partial_z u\|_{L^2_\sigma((0,\infty))} + \left\||\xi'|\,\partial_z^2 u\right\|_{L^2_{2+\sigma}((0,\infty))} \lesssim_\sigma \left\|\tilde{f}\right\|_{L^2_\sigma((0,\infty))}.$$

The retransformation from z to x_n and an additional integration in the ξ' direction reveals that

$$\||\xi'|\, u\|_{L^2_\sigma(H)} + \left\||\xi'|^2\, u\right\|_{L^2_{2+\sigma}(H)} + \|\partial_{x_n} u\|_{L^2_\sigma(H)} + \left\|\partial_{x_n}^2 u\right\|_{L^2_{2+\sigma}(H)} \lesssim_{n,\sigma} \left\|\tilde{f}\right\|_{L^2_\sigma(H)}.$$

With Plancherel's theorem in the reverse Fourier transformation we arrive at

$$\left\|\nabla_x' u\right\|_{L^2_\sigma(H)} + \left\|\Delta_x' u\right\|_{L^2_{2+\sigma}(H)} + \|\partial_{x_n} u\|_{L^2_\sigma(H)} + \left\|\partial_{x_n}^2 u\right\|_{L^2_{2+\sigma}(H)} \lesssim_{n,\sigma} \left\|\tilde{f}\right\|_{L^2_\sigma(H)}.$$

Finally, the mixed second order derivatives can be gained thanks to the formula

$$\left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(H)} \eqsim_\sigma \|\Delta_x u\|^2_{L^2_{2+\sigma}(H)} + \left\|\nabla_x' u\right\|^2_{L^2_\sigma(H)}$$

by means of the density of $C_c^\infty(\overline{H})$ from Remark 2.15. ∎

We will now briefly return to the initial value problem and collect the results for this case. To this end, we consider only the initial datum $g = 0$. This could be generalised to other initial data under some regularity conditions on g.

4.4 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2\big(I;\, L^2_\sigma(H)\big)$.*
If u is the σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$, then

$$t \mapsto \|\nabla_x u(t)\|_{L^2_{1+\sigma}(H)} \in C(\overline{I}) \text{ with } \nabla_x u(t_1) = 0$$

and

$$\int\limits_{(\widetilde{t}_1, t_2)} \|\nabla_{t,x} u\|^2_{L^2_\sigma(H)} \; d\mathcal{L} + \int\limits_{(\widetilde{t}_1, t_2)} \Big\|D^2_x u\Big\|^2_{L^2_{2+\sigma}(H)} \; d\mathcal{L} \lesssim_{n,\sigma} \int\limits_{(\widetilde{t}_1, t_2)} \|f\|^2_{L^2_\sigma(H)} \; d\mathcal{L}$$

for any $\widetilde{t}_1 \in \overline{I}$.

Proof: For $g = 0$, the σ-solution u can be extended to a bigger interval by zero. We can then use the time-local results 4.1, 4.3 and obtain the statement. ∎

4.5 Remark *From the last proposition it follows that for $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with*

$$(l, k, |\alpha|) \in \{(0,1,0),\ (0,0,1),\ (1,0,2)\}\,,$$

the mappings

$$f \mapsto (\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha u$$

that send the inhomogeneity to certain derivatives of the σ-solution of the zero initial value problem on $\overline{I} \times \overline{H}$ are bounded operators from $L^2(I;\, L^2_\sigma(H))$ to $L^2(I;\, L^2_\sigma(H))$.
Elements of the above set of exponents will be referred to as Calderón-Zygmund-exponents in the following. Note that they are exactly given by those $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ that satisfy the conditions $l - k - |\alpha| = -1$ and $2\,l - |\alpha| \leq 0$.

We now prove recursively that also derivatives of σ-solutions solve an adjusted equation in the energy sense. To this end we need two additional estimates for mixed second order derivatives that can be gained by similar methods as the energy estimates above. Note that unlike there, the estimates for the second spatial derivatives do not carry extra weights here. However, we have to place additional assumptions on the inhomogeneity.

4.6 Lemma *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2\big(I;\, L^2_\sigma(H)\big)$.*

(i) If in addition $\partial_t f \in L^2\big(I;\, L^2_\sigma(H)\big)$ and u is a σ-solution to f on $I \times \overline{H}$, then

$$t \mapsto \|\partial_t u(t)\|_{L^2_\sigma(H)} \in C\big([\widetilde{t}_1, t_2)\big)$$

and we have

$$\int\limits_{(\widetilde{t}_1, t_2)} \|\nabla_x \partial_t u\|^2_{L^2_{1+\sigma}(H)} \; d\mathcal{L} \lesssim \big\|\partial_t u(\widetilde{t}_1)\big\|^2_{L^2_\sigma(H)} + \big\|\nabla_x u(\widetilde{t}_1)\big\|^2_{L^2_{1+\sigma}(H)} + \int\limits_{(\widetilde{t}_1, t_2)} \|\partial_t f\|^2_{L^2_\sigma(H)} \; d\mathcal{L} + \int\limits_{(\widetilde{t}_1, t_2)} \|f\|^2_{L^2_\sigma(H)} \; d\mathcal{L}$$

for any $\widetilde{t}_1 \in I$.

(ii) If in addition $\partial_{x_j} f \in L^2\big(I;\, L^2_\sigma(H)\big)$ for a $j \in \{1, \ldots, n-1\}$ and u is a σ-solution to f on $I \times \overline{H}$, then

$$t \mapsto \left\|\partial_{x_j} u(t)\right\|_{L^2_\sigma(H)} \in C([\widetilde{t}_1, t_2))$$

and we have

$$\int\limits_{(\widetilde{t}_1,t_2)} \left\|\nabla_x \partial_{x_j} u\right\|^2_{L^2_{1+\sigma}(H)} \, d\mathcal{L} \lesssim_{n,\sigma} \left\|\partial_{x_j} u(\widetilde{t}_1)\right\|^2_{L^2_\sigma(H)} + \left\|\nabla_x u(\widetilde{t}_1)\right\|^2_{L^2_{1+\sigma}(H)}$$
$$+ \int\limits_{(\widetilde{t}_1,t_2)} \left\|\partial_{x_j} f\right\|^2_{L^2_\sigma(H)} \, d\mathcal{L} + \int\limits_{(\widetilde{t}_1,t_2)} \|f\|^2_{L^2_\sigma(H)} \, d\mathcal{L}$$

for any $\widetilde{t}_1 \in I$.

Proof: Consider the same temporal cut-off function η_ε as in the proof of the energy identity 3.5 and note that the temporal difference quotient $D^h u$ solves the equation for $D^h f$ and has a well-behaved time derivative on $(\widetilde{t}_1, t_2)$ by Proposition 4.1. Testing the equation for $D^h u$ with $\eta_\varepsilon\, D^h u$ is thus possible and leads to

$$\frac{1}{2}\,(\left\|D^h u(\widetilde{t}_2)\right\|^2_{L^2_\sigma})^{-\varepsilon} + \int\limits_I \left\|\nabla_x D^h u\right\|^2_{L^2_{1+\sigma}} \eta_\varepsilon \, d\mathcal{L} = \frac{1}{2}\,(\left\|D^h u(\widetilde{t}_1)\right\|^2_{L^2_\sigma})^{\varepsilon} + \int\limits_I D^h f\, D^h u\, \eta_\varepsilon \, d\mathcal{L}.$$

Given that $\partial_t f \in L^2 L^2_\sigma$, this shows that $t \mapsto \|\partial_t u(t)\|_{L^2_\sigma}$ is continuous on $[\widetilde{t}_1, t_2)$ as well as

$$2 \int\limits_{(\widetilde{t}_1,t_2)} \|\nabla_x \partial_t u\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} \le \left\|\partial_t u(\widetilde{t}_1)\right\|^2_{L^2_\sigma} + \int\limits_{(\widetilde{t}_1,t_2)} \|\partial_t f\|^2_{L^2_\sigma} \, d\mathcal{L} + \int\limits_{(\widetilde{t}_1,t_2)} \|\partial_t u\|^2_{L^2_\sigma} \, d\mathcal{L}.$$

We can use Proposition 4.1 again to finish the proof for the temporal derivative.
For the second part consider $D^h u$ as the difference quotient in x_j-direction for a $j \in \{1, \ldots, n-1\}$. Using Proposition 4.3, the same reasoning as above leads to the continuity of $t \mapsto \left\|\partial_{x_j} u\right\|_{L^2_\sigma}$ and the estimate

$$2 \int\limits_{(\widetilde{t}_1,t_2)} \left\|\nabla_x \partial_{x_j} u\right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} \le \left\|\partial_{x_j} u(\widetilde{t}_1)\right\|^2_{L^2_\sigma} + \int\limits_{(\widetilde{t}_1,t_2)} \left\|\partial_{x_j} f\right\|^2_{L^2_\sigma} \, d\mathcal{L} + \int\limits_{(\widetilde{t}_1,t_2)} \left\|\partial_{x_j} u\right\|^2_{L^2_\sigma} \, d\mathcal{L}.$$

■

For future reference we annotate the following very easy consequence.

4.7 Corollary *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2\big(I;\, L^2_\sigma(H)\big)$ as well as $\nabla'_x f \in L^2\big(I;\, L^2_\sigma(H)\big)$.*
If u is a σ-solution to f on $I \times \overline{H}$, then we have

$$\int\limits_{(\widetilde{t}_1,t_2)} \|\Delta'_x u\|^2_{L^2_{1+\sigma}(H)} \, d\mathcal{L} \lesssim_{n,\sigma} \left\|\nabla'_x u(\widetilde{t}_1)\right\|^2_{L^2_\sigma(H)} + \left\|\nabla_x u(\widetilde{t}_1)\right\|^2_{L^2_{1+\sigma}(H)}$$
$$+ \int\limits_{(\widetilde{t}_1,t_2)} \|\nabla'_x f\|^2_{L^2_\sigma(H)} \, d\mathcal{L} + \int\limits_{(\widetilde{t}_1,t_2)} \|f\|^2_{L^2_\sigma(H)} \, d\mathcal{L}.$$

for any $\widetilde{t}_1 \in I$.

4.8 Remark *As in Proposition* 4.4*, the range of points in time for which estimates* 4.6 *and* 4.7 *hold can be extended to include the initial time* t_1 *for solution to the initial value problem with* $g = 0$.

We can now iterate the notion of solution in terms of their derivatives as announced above. We set $\partial_{x_n}^{\alpha_n - 1} u := 0$ if $\alpha_n = 0$.

4.9 Proposition *Let* $I = (t_1, t_2) \subset \mathbb{R}$ *be an open interval,* $f \in L^1\big(I; L^2_\sigma(H)\big)$ *and* u *be a* σ*-solution to* f *on* $I \times \overline{H}$*. Further let* $\tilde{t}_1 \in I$*,* $k \in \mathbb{N}_0$ *and* $\alpha \in \mathbb{N}_0^n$*.*
If $\partial_t^i \partial_{x'}^{\beta' + \gamma'} \partial_{x_n}^{\gamma_n} f \in L^2(I; L^2_{\gamma_n + \sigma}(H))$ *for any* $0 \le i \le k$*,* $0 \le \beta' \le \alpha'$ *and* $0 \le |\gamma| \le \alpha_n$*, then* $\partial_t^k \partial_x^\alpha u$ *is an* $(\alpha_n + \sigma)$*-solution to* $\partial_t^k \partial_x^\alpha f + \alpha_n \Delta'_x \partial_t^k \partial_{x'}^{\alpha'} \partial_{x_n}^{\alpha_n - 1} u$ *on* $[\tilde{t}_1, t_2) \times \overline{H}$*.*

Proof: The temporal and tangential derivatives commute with the differential expression by Remark 3.2 and hence formally one sees that $\partial_t^k u$ is a σ-solution to $\partial_t^k f$ and $\partial_{x_j}^{\alpha_j} u$ is a σ-solution to $\partial_{x_j}^{\alpha_j} f$. The formal treatment of the normal derivative is not as straight forward as that, but a calculation shows that $\partial_t \partial_{x_n} u - L_{1+\sigma} u = \partial_{x_n} f + \Delta'_x u$ and thus recursively $\partial_t \partial_{x_n}^{\alpha_n} u - L_{\alpha_n + \sigma} \partial_{x_n}^{\alpha_n} u = \partial_{x_n}^{\alpha_n} f + \alpha_n \Delta'_x \partial_{x_n}^{\alpha_n - 1} u$. We will therefore turn to the verification of the regularity properties needed.
If $\partial_t f \in L^2 L^2_\sigma$, Proposition 4.1 and the temporal part of Lemma 4.6 imply the right regularity for $\partial_t u$ on $(\tilde{t}_1, t_2) \times \overline{H}$. Since we can always fit yet another open interval in between (t_1, t_2) and $(\tilde{t}_1, t_2)$, $\partial_t u$ really becomes a σ-solution on $[\tilde{t}_1, t_2) \times \overline{H}$. An iteration of this argument in the time direction is obviously possible if we adjust the interval in any step and assume $\partial_t^i f \in L^2 L^2_\sigma$ for any $i \le k$.
The proof in tangential directions works exactly the same way, this time using Proposition 4.3 and the spatial part of 4.6. As before we can iterate this to see that $\partial_{x_j}^{\alpha_j} u$ is a σ-solution on $[\tilde{t}_1, t_2) \times \overline{H}$ whenever $\partial_{x_j}^{\beta_j} f \in L^2 L^2_\sigma$ for any $\beta_j \le \alpha_j$.
Let us now consider the x_n-direction. By the second derivative estimate from 4.3 we directly see that $\partial_{x_n} u$ already has the right regularity itself. Moreover, with $f \in L^2 L^2_\sigma$ and $\nabla'_x f \in L^2 L^2_\sigma$ we have by Corollary 4.7 that

$$
(*) \qquad \begin{aligned}
\int_{(\tilde{t}_1, t_2)} \left\| \partial_{x_n} f + \Delta'_x u \right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} &\le \int_{(\tilde{t}_1, t_2)} \left\| \partial_{x_n} f \right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} + \int_{(\tilde{t}_1, t_2)} \left\| \Delta'_x u \right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} \\
&\lesssim \left\| \nabla'_x u(\tilde{t}_1) \right\|^2_{L^2_\sigma} + \left\| \nabla_x u(\tilde{t}_1) \right\|^2_{L^2_{1+\sigma}} + \sum_{0 \le |\gamma| \le 1} \int_{(\tilde{t}_1, t_2)} \left\| \partial_x^\gamma f \right\|^2_{L^2_{\gamma_n + \sigma}} \, d\mathcal{L}
\end{aligned}
$$

This means that if in addition $\partial_{x_n} f \in L^2 L^2_{1+\sigma}$, then $\partial_{x_n} f + \Delta'_x u \in L^2 L^2_{1+\sigma}$ with respect to the smaller time interval, and hence $\partial_{x_n} u$ is a $(1+\sigma)$-solution to this right hand side on $[\tilde{t}_1, t_2) \times \overline{H}$ that satisfies the inequalities 4.1 and 4.3 for the temporal and the tangential derivatives, meaning that

$$
\begin{aligned}
&\int_{(\tilde{t}_1, t_2)} \left\| \nabla_{t,x} \partial_{x_n} u \right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} + \int_{(\tilde{t}_1, t_2)} \left\| D_x^2 \partial_{x_n} u \right\|^2_{L^2_{3+\sigma}} \, d\mathcal{L} \\
&\quad \lesssim \left\| \nabla_x \partial_{x_n} u(\tilde{t}_1) \right\|^2_{L^2_{2+\sigma}} + \int_{(\tilde{t}_1, t_2)} \left\| \partial_{x_n} f + \Delta'_x u \right\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} \\
&\quad \lesssim \left\| \nabla'_x u(\tilde{t}_1) \right\|^2_{L^2_\sigma} + \left\| \nabla_x u(\tilde{t}_1) \right\|^2_{L^2_{1+\sigma}} + \left\| \nabla_x \partial_{x_n} u(\tilde{t}_1) \right\|^2_{L^2_{2+\sigma}} + \sum_{0 \le |\gamma| \le 1} \int_{(\tilde{t}_1, t_2)} \left\| \partial_x^\gamma f \right\|^2_{L^2_{\gamma_n + \sigma}} \, d\mathcal{L}
\end{aligned}
$$

where in the second inequality we used $(*)$ again.
But under the assumptions in place, also $\nabla'_x u$ is a σ-solution to $\nabla'_x f$ by the first part of this proof,

so if in addition the condition $\nabla_x'\nabla_x' f \in L^2 L^2_\sigma$ holds we can apply Corollary 4.7 onto $\nabla_x' u$ and $\nabla_x' f$ and consequently $(*)$ onto $\partial_{x_n}\nabla_x' f + \Delta_x' \nabla_x' u$ and learn that

$$\int_{(\tilde{t}_1,t_2)} \|\nabla_x'(\partial_{x_n} f + \Delta_x' u)\|^2_{L^2_{1+\sigma}} \, d\mathcal{L}$$
$$\lesssim \|\nabla_x'\nabla_x' u(\tilde{t}_1)\|^2_{L^2_\sigma} + \|\nabla_x \nabla_x' u(\tilde{t}_1)\|^2_{L^2_{1+\sigma}} + \sum_{0\le|\gamma|\le 1} \int_{(\tilde{t}_1,t_2)} \|\partial_x^\gamma \nabla_x' f\|^2_{L^2_{\gamma_n+\sigma}} \, d\mathcal{L}.$$

Imposing the requirement $\nabla_x'\partial_{x_n} f \in L^2 L^2_{1+\sigma}$ on top, this shows $\nabla_x'(\partial_{x_n} f + \Delta_x' u) \in L^2 L^2_{1+\sigma}$. It is thus possible to apply Corollary 4.7 on the $(1+\sigma)$-solution $\partial_{x_n} u$ to $\partial_{x_n} f + \Delta_x' u$ itself. We gain

$$\begin{aligned}
\int_{(\tilde{t}_1,t_2)} \|\Delta_x'\partial_{x_n} u\|^2_{L^2_{2+\sigma}} \, d\mathcal{L} &\lesssim \|\nabla_x'\partial_{x_n} u(\tilde{t}_1)\|^2_{L^2_{1+\sigma}} + \|\nabla_x \partial_{x_n} u(\tilde{t}_1)\|^2_{L^2_{2+\sigma}} \\
&\quad + \int_{(\tilde{t}_1,t_2)} \|\partial_{x_n} f + \Delta_x' u\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} + \int_{(\tilde{t}_1,t_2)} \|\nabla_x'(\partial_{x_n} f + \Delta_x' u)\|^2_{L^2_{1+\sigma}} \, d\mathcal{L} \\
&\lesssim \sum_{0\le|\gamma|\le 1} \|\nabla_x'\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{\gamma_n+\sigma}} + \sum_{0\le|\gamma|\le 1} \|\nabla_x\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{\gamma_n+1+\sigma}} \\
&\quad + \sum_{0\le|\gamma|\le 1} \int_{(\tilde{t}_1,t_2)} \|\partial_x^\gamma f\|^2_{L^2_{\gamma_n+\sigma}} \, d\mathcal{L} + \sum_{0\le|\gamma|\le 1} \int_{(\tilde{t}_1,t_2)} \|\partial_x^\gamma \nabla_x' f\|^2_{L^2_{\gamma_n+\sigma}} \, d\mathcal{L}
\end{aligned}$$

and hence

$$\begin{aligned}
\int_{(\tilde{t}_1,t_2)} \left\|\partial_{x_n}^2 f + \Delta_x'\partial_{x_n} u\right\|^2_{L^2_{2+\sigma}} \, d\mathcal{L} &\lesssim \sum_{0\le|\gamma|\le 1} \|\nabla_x'\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{\gamma_n+\sigma}} + \sum_{0\le|\gamma|\le 1} \|\nabla_x\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{\gamma_n+1+\sigma}} \\
&\quad + \sum_{0\le|\gamma|\le 2} \int_{(\tilde{t}_1,t_2)} \|\partial_x^\gamma f\|^2_{L^2_{\gamma_n+\sigma}} \, d\mathcal{L}.
\end{aligned}$$

This allows us the performance of an induction on $\alpha_n \ge 1$ to prove that $\partial_{x_n}^{\alpha_n} u$ is an $(\alpha_n+\sigma)$-solution to $\partial_{x_n}^{\alpha_n} f + \alpha_n \Delta_x' \partial_{x_n}^{\alpha_n - 1} u$ if $\partial_x^\gamma f \in L^2 L^2_{\gamma_n+\sigma}$ for any $0 \le |\gamma| \le \alpha_n$. Consequently, the estimates 4.1 and 4.3 hold accordingly. By the inductive result of the calculations above these estimates get the form

$$\begin{aligned}
&\int_{(\tilde{t}_1,t_2)} \|\nabla_{t,x}\partial_{x_n}^{\alpha_n} u\|^2_{L^2_{\alpha_n+\sigma}(H)} \, d\mathcal{L} + \int_{(\tilde{t}_1,t_2)} \left\|D_x^2 \partial_{x_n}^{\alpha_n} u\right\|^2_{L^2_{2+\alpha_n+\sigma}(H)} \, d\mathcal{L} \\
&\lesssim \sum_{0\le|\gamma|\le\alpha_n} \int_{(\tilde{t}_1,t_2)} \|\partial_x^\gamma f\|^2_{L^2_{\gamma_n+\sigma}(H)} \, d\mathcal{L} + \|\nabla_x \partial_{x_n}^{\alpha_n} u(\tilde{t}_1)\|^2_{L^2_{1+\alpha_n+\sigma}(H)} \\
&\quad + \sum_{0\le|\gamma|\le\alpha_n-1} \|\nabla_x\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{1+\gamma_n+\sigma}(H)} + \sum_{0\le|\gamma|\le\alpha_n-1} \|\nabla_x'\partial_x^\gamma u(\tilde{t}_1)\|^2_{L^2_{\gamma_n+\sigma}(H)}.
\end{aligned}$$

Applying this vertical result on the σ-solution $\partial_t^k \partial_{x'}^{\alpha'} u$ then yields the complete statement. ∎

5 The Intrinsic Metric

Distance functions that are naturally attached to linear differential expressions have proved to be extremely useful and powerful. For a self-adjoint second order equation, [FP83] used the principal symbol to generate balls that govern the behaviour of the solutions. The same definition was brought forth by [SC84], and applied in the special case of a sum of squares satisfying the Hörmander condition, which means that the equation is subelliptic, see [Hö67] and [Ego75]. In this same situation, a suitable metric was constructed by [NSW85] in a different way, but resulting in the same notion.
And indeed are we given a spatial differential expression

$$L_\sigma = (\,\cdot\,)_n \, \Delta_x + (1+\sigma)\, \partial_{x_n}$$

that is not only self-adjoint, but has also a connection to subelliptic operators: By a coordinate transformation

$$x_n \mapsto \sqrt{x_n} =: \widetilde{x}_n$$

it is transformed into

$$(\tilde{\cdot})_n^2 \, \Delta'_{\widetilde{x}} + \frac{1}{4}\, \partial^2_{\widetilde{x}_n} + \left(\frac{1+\sigma}{2} - \frac{1}{4} \right) (\tilde{\cdot})_n^{-1} \, \partial_{\widetilde{x}_n},$$

a subelliptic sum of squares of $(\tilde{\cdot})_n \partial_{\widetilde{x}_j}$, $j = 1, \dots, n-1$ and $\partial_{\widetilde{x}_n}$.
Following [Koc99] from here on, for any $x \in H$ we consider the principal symbol $p(x,\xi) = x_n \, |\xi|^2$ that induces the bilinear form

$$\langle\, \xi \mid \zeta \,\rangle_x = x_n^{-1} \, (\xi \cdot \zeta)$$

on the tangent space at x. Its scaling behaviour is described by $\langle\, \lambda\,\xi \mid \lambda\,\zeta \,\rangle_{\lambda x} = \lambda \,\langle\, \xi \mid \zeta \,\rangle_x$. Obviously, this is not only an inner product on the whole tangent space, but also a Riemannian metric, giving us the ability to measure the length of curves $\Gamma : \mathbb{R} \supset [a,b] \to H$ by

$$\mathcal{L}(\Gamma) := \int\limits_{(a,b)} \sqrt{\langle\, \partial_\varsigma \Gamma(\varsigma) \mid \partial_\varsigma \Gamma(\varsigma) \,\rangle_{\Gamma(\varsigma)}} \; d\mathcal{L}(\varsigma).$$

Therefore, the Carnot-Caratheodory-metric

$$d(x,y) = \inf \{ \mathcal{L}(\Gamma) \mid \Gamma \text{ joins } x \text{ and } y \}$$

turns H into a metric space with a finite intrinsic distance function induced by the spatial part of our differential expression.
In order to get to know the metric d better, we consider the minimisation problem for the energy

$$\mathcal{E}(\Gamma) := \int\limits_{(a,b)} \langle\, \partial_\varsigma \Gamma(\varsigma) \mid \partial_\varsigma \Gamma(\varsigma) \,\rangle_{\Gamma(\varsigma)} \; d\mathcal{L}(\varsigma)$$

on an at first arbitrary, but fixed interval (a, b). By the Cauchy-Schwarz inequality it is immediately clear that $\mathcal{L}(\Gamma)^2 \leq 2\,(b-a)\,\mathcal{E}(\Gamma)$ with equality if and only if $\langle\, \partial_\varsigma\Gamma(\varsigma) \mid \partial_\varsigma\Gamma(\varsigma)\,\rangle_{\Gamma(\varsigma)} = 1$ on (a, b). In this equality case of curves that are parametrised by arc length, or traversed with constant speed, a curve that minimises $\mathcal{E}$ over (a, b) is a curve that minimises $\mathcal{L}$ over (a, b) and vice versa. For fixed $x,\, y \in H$ assume now that $\Gamma : [a, b] \to H$ is a curve that indeed realises the infimum over the length of all curves joining x and y – that is a distance minimising curve, also called geodesic – and is parametrised by arclength. Then $\Gamma(a) = x$, $\Gamma(b) = y$ and over (a, b) the curve satisfies

$$1 = \Gamma_n^{-1}\,(\partial_\varsigma\Gamma \cdot \partial_\varsigma\Gamma) = \Gamma_n^{-1}\,|\partial_\varsigma\Gamma|^2$$

as well as the Euler-Lagrange equations for the energy:

$$\begin{aligned} 0 &= 2\,\partial_\varsigma(\Gamma_n^{-1}\,\partial_\varsigma\Gamma_j) + \Gamma_n^{-2}\,(\partial_\varsigma\Gamma \cdot \partial_\varsigma\Gamma)\,\delta_{jn} \\ &= 2\,\partial_\varsigma(\Gamma_n^{-1}\,\partial_\varsigma\Gamma_j) + \Gamma_n^{-1}\,\delta_{jn} \qquad\qquad \text{for any } j = 1, \ldots, n. \end{aligned}$$

It follows that $\Gamma_n^{-1}\,\partial_\varsigma\Gamma_j$ is constant for $j = 1, \ldots, n-1$, a fact that can be used to reduce the unit speed equation to an equation for Γ_n only, which in turn makes a simplification of the the n-th Euler-Lagrange equation possible. After integrating the first $n-1$ Euler-Lagrange equations in time it becomes evident that for all $\varsigma \in [a, b]$ the curve Γ satisfies the equations

$$\begin{aligned} (\partial_\varsigma\Gamma_n(\varsigma))^2 &= \Gamma_n(\varsigma) - c^2\,\Gamma_n^2(\varsigma) \\ \Gamma_j(\varsigma) &= x_j + c_j \int\limits_{(a,\varsigma)} \Gamma_n \, d\mathcal{L} \qquad \text{for } j = 1, \ldots, n-1 \\ \partial_\varsigma^2\Gamma_n(\varsigma) &= \frac{1}{2} - c^2\,\Gamma_n(\varsigma) \end{aligned}$$

for constants $c_j \in \mathbb{R}$ and with $c := \left(\sum\limits_{j=1}^{n-1} c_j^2\right)^{\frac{1}{2}}$.
If x and y are vertically aligned, that is for $y = (x', y_n)$, then $c_j = 0$ for any $j = 1, \ldots n-1$ and hence $c = 0$. In this case $\Gamma_j = x_j$ are constant on $[a, b]$ for $j = 1, \ldots, n-1$, meaning that the geodesic is indeed a vertical line. Because of the ordinary differential equations for Γ_n it is clear that for any $\varsigma \in [a, b]$ we have

$$\Gamma_n(\varsigma) = \frac{1}{4}(\varsigma - \vartheta)^2$$

for a $\vartheta \in \mathbb{R}$.
On the other hand, if x and y are not vertically aligned, then at least one c_j does not vanish and thus $c > 0$. Again solving the ordinary differential equations for Γ_n and using the result in the equations for Γ_j we obtain for any $\varsigma \in [a, b]$ that

$$\begin{aligned} \Gamma_j(\varsigma) &= x_j + \frac{c_j}{2c^2}\left(\varsigma - a + \frac{1}{c}\,(\sin(c\,(a-\vartheta)) - \sin(c\,(\varsigma-\vartheta)))\right) \\ \Gamma_n(\varsigma) &= \frac{1}{2c^2}\,(1 - \cos(c\,(\varsigma - \vartheta))) \end{aligned}$$

for a $\vartheta \in \mathbb{R}$. Thus arc-length parametrised geodesics are cycloid curves. Compare also [DH98] for a treatment of the issue in two dimensions.

Conversely, it is clear that a curve Γ in H given by one of the two mapping prescriptions above is parametrised by arc length and satisfies the Euler-Lagrange equations for the energy regardless of the interval (a,b). Therefore, for fixed a and b it realises the miminal length in going from $\Gamma(a)$ to $\Gamma(b)$ among all curves over (a,b). By a suitable choice of a and b and the parameters ϑ and (if neccessary) c_j, any points $x, y \in H$ can be joined by a curve of this kind. Among these choices, the one that belongs to the setting that minimises the length $b-a$ is then a geodesic between x and y. But this allows us to expand our definition of the metric onto $\overline{H}$. For $\vartheta = a$ we have $\Gamma_n(a) = 0$ both if x and y are vertically aligned and if not. On the other hand, if ϑ is chosen so that $\Gamma(a) = x$, it is always possible to find b and c_j such that $\Gamma(b) = 0$, so y with $y_n = 0$ can also be reached. If x_n and y_n vanish, that is if x and y are situated on the boundary of H, then the points are not vertically aligned unless they coincide. In any case we can then define $d(x,y)$ as the minimal value of $b-a$ such that $\Gamma : [a,b] \to \overline{H}$ is of the type presented above and joins x and y.

A weighted bound of the gradient of a continuously differentiable function now ensures that this function is Lipschitz with respect to d.

5.1 Proposition *Let $\Psi : \overline{H} \to \mathbb{R}$ be continuous with $\Psi \in C^1(H)$ and $C > 0$.*
If $\sqrt{x_n}\,|\nabla_x \Psi(x)| \leq C$ for any $x \in H$, then $|\Psi(x) - \Psi(y)| \leq C\, d(x,y)$ for all , $x, y \in \overline{H}$.

Proof: Fix $x, y \in H$ and let $\Gamma : [a,b] \to H$ be the arc-length parametrised geodesic connecting them. Then the chain rule and the fundamental theorem of calculus show that

$$\begin{aligned}|\Psi(x) - \Psi(y)| &= \Big| \int_{(a,b)} \nabla_x \Psi(\Gamma(\varsigma)) \cdot \partial_\varsigma \Gamma(\varsigma)\, d\mathcal{L}(\varsigma) \Big| \\ &\leq (b-a) \sup_{\varsigma \in (a,b)} \left\{ \Gamma_n^{\frac{1}{2}}(\varsigma)\, |\nabla_x \Psi(\Gamma(\varsigma))| \right\} \sup_{\varsigma \in (a,b)} \left\{ |\partial_\varsigma \Gamma(\varsigma)| \Gamma_n^{-\frac{1}{2}}(\varsigma) \right\}.\end{aligned}$$

With the assumption in the first supremum and the unit speed property of Γ in the second one, we get

$$|\Psi(x) - \Psi(y)| \leq C\,(b-a) = C\, d(x,y)$$

and therefore the Lipschitz property on H. By a limiting procedure this can be extended onto $\overline{H}$ thanks to the continuity of $d(\cdot, y)$ and Ψ on $\overline{H}$. ∎

We now turn to to a precise description of the intrinsic metric. If x and y are vertically aligned, we can calculate $b-a$ explicitely. In case the points are horizontally aligned, at least an inequality can be proven, and in the special subcase that both x and y are on the boundary, a precise expression for d can be obtained again. Here we use the notation d^{eu} for the euclidean metric regardless of the dimension.

5.2 Proposition *Let $x, y \in \overline{H}$.*

(i) If $y' = x'$ then

$$d(x,y) = 2\,|\sqrt{x_n} - \sqrt{y_n}|.$$

(ii) If $y_n = x_n$, then

$$d(x,y) \leq x_n^{-\frac{1}{2}}\, d^{eu}(x,y).$$

(iii) If $x_n = y_n = 0$ *then*

$$d(x,y) = 2\sqrt{\pi\, d^{eu}(x,y)}.$$

Proof: First assume $y = (x', y_n)$ with $y_n \le x_n$. We then consider $\Gamma : [a,b] \to \overline{H}$ given by $\Gamma_j = x_j$ for $j = 1, \dots, n-1$ and $\Gamma_n(\varsigma) = \frac{1}{4}(\varsigma - \vartheta)^2$, where $\vartheta = a + 2\sqrt{x_n}$. Then $\Gamma(a) = x$ is obviously true, and for $\Gamma(b) = y$ we only need to have

$$y_n = \Gamma_n(b) = \frac{1}{4}(b - a - 2\sqrt{x_n})^2.$$

But this is the case if

$$b = \vartheta - 2\sqrt{y_n} = a + 2\sqrt{x_n} - 2\sqrt{y_n}$$

which is bigger than a since $x_n \ge y_n$, but at the same time the smallest possible b that we can get within this setting. Thus Γ – together with the interval $[a,b]$ – is a length minimising curve joining x and y and we have

$$d(x,y) = b - a = 2\sqrt{x_n} - 2\sqrt{y_n}.$$

For $y_n \ge x_n$ we choose $\vartheta = a - 2\sqrt{x_n}$ and then get

$$b = \vartheta + 2\sqrt{y_n} = a + 2\sqrt{y_n} - 2\sqrt{x_n} \ge a$$

and hence

$$d(x,y) = b - a = 2\sqrt{y_n} - 2\sqrt{x_n}$$

as desired.

We turn to the horizontally aligned cases and denote the curve that traverses the line between x and y parametrised over the interval $(0,1)$ by $\overline{\Gamma}$. Its length is given by

$$\mathcal{L}(\overline{\Gamma}) = x_n^{-\frac{1}{2}} \int_{(0,1)} \sqrt{\partial_\varsigma \overline{\Gamma} \cdot \partial_\varsigma \overline{\Gamma}}\, d\mathcal{L} = x_n^{-\frac{1}{2}} d^{eu}(x,y).$$

By definition, this is an upper bound of $d(x,y)$.

Let now $x_n = y_n = 0$. This time the curve is of the form $\Gamma : [a,b] \to \overline{H}$,

$$\Gamma_j(\varsigma) = x_j + \frac{c_j}{2c^2}\left(\varsigma - a + \frac{1}{c}\left(\sin\left(c\,(a - \vartheta)\right) - \sin\left(c\,(\varsigma - \vartheta)\right)\right)\right)$$
$$\Gamma_n(\varsigma) = \frac{1}{2c^2}\left(1 - \cos\left(c\,(\varsigma - \vartheta)\right)\right)$$

with $\vartheta = a$. Then again we have $\Gamma(a) = x$. To obtain $\Gamma_n(b) = 0$ we therefore need the smallest $b > a$ in a way that it satisfies

$$0 = \frac{1}{2c^2}\left(1 - \cos\left(c(b-a)\right)\right).$$

Then obviously $b = a + \frac{2\pi}{c}$ is the point we are looking for, so

$$d(x,y) = b - a = \frac{2\pi}{c}.$$

To determine the value of c we consider the tangential components

$$\begin{aligned} y_j = \Gamma_j(b) &= x_j + \frac{c_j}{2c^2}\left(b - a + \frac{1}{c}\left(\sin\left(c\left(a-a\right)\right) - \sin\left(c\left(b-a\right)\right)\right)\right) \\ &= x_j + \frac{c_j}{2c^3}\left(2\,\pi + \left(\sin\left(0\right)\right) - \sin\left(2\,\pi\right)\right)\right)) = x_j + \frac{\pi\, c_j}{c^3}. \end{aligned}$$

Thus for the $n-1$-dimensional euclidean distance between x' and y' we get

$$d^{eu}(x',y')^2 = \sum_{j=1}^{n-1}(x_j - y_j)^2 = \frac{\pi^2}{c^6}\,c^2.$$

But then it follows that $c = \sqrt{\frac{\pi}{d^{eu}(x',y')}}$ and the statement is proven. ∎

We can make use of the formulas for the special cases from Proposition 5.2 to deduce estimates that hold on all of $\overline{H}$, starting with estimates from above.

5.3 Lemma *Let $x,\ y \in \overline{H}$. We then have:*

(i) $d(x,y) \le 2\left(\sqrt{x_n} + \sqrt{y_n} + \sqrt{\pi\, d^{eu}(x'-y')}\right)$.

(ii) $d(x,y) \le 2|\sqrt{x_n} - \sqrt{y_n}| + \frac{d^{eu}(x',y')}{\max\{\sqrt{x_n},\sqrt{y_n}\}}$.

Proof: By the triangle inequality we have

$$d(x,y) \le d(x,(x',0)) + d((x',0),(y',0)) + d((y',0),y)$$

as well as

$$d(x,y) \le d(x,(x',y_n)) + d((x',y_n),y)$$

and

$$d(x,y) \le d(x,(y',x_n)) + d((y',x_n),y).$$

The application of the appropriate parts of Proposition 5.2 yields the desired estimates. ∎

For similar estimates from below we have to take the geodesic into consideration again.

5.4 Lemma *Let $x,\ y \in \overline{H}$, $\Gamma : [a,b] \to \overline{H}$ be the arc-length parametrised geodesic that joins them and $\Gamma_n^{max} := \max_{\varsigma\in[a,b]} \Gamma_n(\varsigma)$. We then have:*

(i) $d(x,y) \ge \frac{d^{eu}(x,y)}{\sqrt{\Gamma_n^{max}}}$.

(ii) $d(x,y) \ge 2\left(\sqrt{\Gamma_n^{max}} - \min\left\{\sqrt{x_n},\sqrt{y_n}\right\}\right)$.

Proof: By definition it is clear that

$$d(x,y) \geq \frac{1}{\sqrt{\Gamma_n^{max}}} \int\limits_{(a,b)} \sqrt{\partial_\varsigma \Gamma \cdot \partial_\varsigma \Gamma}\, d\mathcal{L}.$$

The integral equals the euclidean length of Γ and is therefore not smaller than the euclidean distance between the points x and y joined by Γ.
Let now $z = (z', \Gamma_n^{max})$ be the point in the image of Γ where Γ_n^{max} is realised for the first time. We denote the part of Γ that joins x with z by Γ_1 and let $\overline{\Gamma}$ be the line connecting x and (x', Γ_n^{max}). We then have

$$d(x,y) = \mathcal{L}(\Gamma) \geq \mathcal{L}(\Gamma_1) \geq \mathcal{L}(\overline{\Gamma}) = d(x, (x', \Gamma_n^{max}))$$

since the line is indeed the geodesic for vertically aligned points. The same calculation can be done with (y', Γ_n^{max}). Using Proposition 5.2 we thus get

$$d(x,y) \geq \max\left\{2\left(\sqrt{\Gamma_n^{max}} - \sqrt{x_n}\right), 2\left(\sqrt{\Gamma_n^{max}} - \sqrt{y_n}\right)\right\}$$

and the claimed estimate follows. ∎

From the second estimate of Lemma 5.4 we can conclude that the distance of a point from a given horizontal level set is realised by the vertical projection of this point onto the level set $\{x_n = C\}$.

5.5 Proposition *Let $C \geq 0$ and $x_0 \in \overline{H}$.*
We have

$$d(x_0, \{x_n = C\}) = d(x_0, (x_0', C)) = 2|\sqrt{x_{0,n}} - C|.$$

Proof: For $x_{0,n} = C$ the claim is obvious, so we consider only $x_{0,n} \neq C$. Because of Proposition 5.2 we know that

$$d(x_0, \{x_n = C\}) \leq d(x_0, (x_0', C)) = 2|\sqrt{x_{0,n}} - \sqrt{C}|.$$

Now let $x_{0,n} > C$. For any $y \in \{x_n = C\}$ such that the arc-length parametrised geodesic $\Gamma : [a,b] \longrightarrow \overline{H}$ connecting y with x_0 has maximal x_n-value $\Gamma_n^{max} > x_{0,n}$, Lemma 5.4 shows that

$$d(x_0, y) \geq 2\left(\sqrt{\Gamma_n^{max}} - \sqrt{\min\{x_{0,n}, y_n\}}\right) > 2\left(\sqrt{x_{0,n}} - \sqrt{C}\right).$$

Therefore, the distance realising point y_* can not be among those y, and the arc-length parametrised geodesic $\Gamma_* : [a_*, b_*] \to \overline{H}$ between x_0 and y_* does not go above $x_{0,n}$ in the x_n-direction. Since $\Gamma_{*,n}(a_*) = x_{0,n}$, it follows $\Gamma_{*,n}^{max} = x_{0,n}$. By Lemma 5.4 again we then see that

$$d(x_0, \{x_n = C\}) \geq 2\left(\sqrt{\Gamma_{*,n}^{max}} - \sqrt{\min\{x_{0,n}, y_{*,n}\}}\right) = 2\left(\sqrt{x_{0,n}} - \sqrt{C}\right),$$

which is exactly the distance of the vertically projected point in $\{x_n = C\}$ from x_0.
If we have $x_{0,n} < C$ for a $C > 0$, the same argument shows that the minimising point y_* in this case can not be element of $\{y \in \{x_n = C\} \mid \Gamma_n^{max} < C\}$. The equality follows as before. ∎

By means of Lemmas 5.3 and 5.4 we obtain an equivalent quasi-metric for d on all of $\overline{H}$.

5.6 Theorem *For all $x,\, y \in \overline{H}$ we have*

$$c_d^{-1}\, d(x,y) \leq \frac{d^{eu}(x,y)}{\sqrt{x_n} + \sqrt{y_n} + \sqrt{d^{eu}(x,y)}} \leq d(x,y)$$

with $c_d := 12$.

Proof: Fix arbitrary $x,\, y \in \overline{H}$. Because of Lemma 5.3 we have

$$d(x,y) \leq 2\left(\sqrt{x_n} + \sqrt{y_n} + \sqrt{\pi}\,\sqrt{d^{eu}(x-y)}\right)$$

and

$$d(x,y) \leq 2\,\frac{|x_n - y_n| + d^{eu}(x',y')}{\sqrt{x_n} + \sqrt{y_n}},$$

where we used $d^{eu}(x',y') \leq d^{eu}(x,y)$ in the first inequality, and $\max\{a,b\} \geq \frac{1}{2}(a+b)$ in the second. Still working in the second estimate, we can apply the inequality $\sqrt{a} + \sqrt{b} \leq \sqrt{2}\,\sqrt{a+b}$ for any positive a and b to get

$$d(x,y) \leq 2\sqrt{2}\,\frac{d^{eu}(x,y)}{\sqrt{x_n} + \sqrt{y_n}}.$$

Now in case of $\sqrt{d^{eu}(x,y)} \geq \sqrt{x_n} + \sqrt{y_n}$, the first estimate implies

$$d(x,y) \leq 2\,(1+\sqrt{\pi})\,\sqrt{d^{eu}(x,y)} = 2\,(1+\sqrt{\pi})\,\frac{d^{eu}(x,y)}{\sqrt{d^{eu}(x,y)}}.$$

In view of the second estimate we adjust the denominator and get

$$d(x,y) \leq 4\,(1+\sqrt{\pi})\,\frac{d^{eu}(x,y)}{\sqrt{d^{eu}(x,y)} + \sqrt{x_n} + \sqrt{y_n}}$$

in this case, while in the opposite case of $\sqrt{d^{eu}(x,y)} < \sqrt{x_n} + \sqrt{y_n}$ the denominator of the other estimate can be brought into the same shape to yield

$$d(x,y) \leq 4\sqrt{2}\,\frac{d^{eu}(x,y)}{\sqrt{x_n} + \sqrt{y_n} + \sqrt{d^{eu}(x,y)}}.$$

Since

$$\sqrt{2} \leq 1 + \sqrt{2} \leq 1 + \sqrt{\pi} \leq 3$$

the definition of c_d is justified and the left hand part of the statement follows.
Naturally, for the proof of the right hand side of the statement we use the lower estimates of the metric from Lemma 5.4. In view of these estimates and the result of the first part, we distinguish the cases $\sqrt{\Gamma_n^{max}} \leq \sqrt{x_n} + \sqrt{y_n} + \sqrt{d^{eu}(x,y)}$ and $\sqrt{\Gamma_n^{max}} > \sqrt{x_n} + \sqrt{y_n} + \sqrt{d^{eu}(x,y)}$ where Γ_n^{max} is again the highest x_n-value of the arc-length parametrised geodesic joining x and y. In the first case, the first part of Lemma 5.4 immediately delivers the statement, whereas in the second case

we use the second part of Lemma 5.4 and the fact that $\min\{a,b\} \le a+b$ to get

$$\begin{aligned} d(x,y) &\ge 2\sqrt{\Gamma_n^{max}} - (\sqrt{x_n} + \sqrt{y_n}) \ge \sqrt{d^{eu}(x,y)} \\ &= \frac{d^{eu}(x,y)}{\sqrt{d^{eu}(x,y)}} \ge \frac{d^{eu}(x,y)}{\sqrt{d^{eu}(x,y)} + \sqrt{x_n} + \sqrt{y_n}} \end{aligned}$$

as desired. ∎

5.7 Remark *It is sometimes convenient to consider the quasi-metric*

$$\tilde{d}(x,y) := \frac{d^{eu}(x,y)}{(x_n^2 + y_n^2 + d^{eu}(x,y)^2)^{\frac{1}{4}}}$$

instead. Then we have

$$c_d^{-1}\, d(x,y) \le \tilde{d}(x,y) \le 4\, d(x,y)$$

for any $x,\ y \in \overline{H}$.

In the following, all balls $B_r(x_0)$ are understood with respect to the intrinsic metric d. The relation between intrinsic and euclidean geometry, represented by the metric balls, is described in the next result.

5.8 Proposition *Let $x_0 \in \overline{H}$ and $r > 0$.*
Then we have

$$\overline{H} \cap B^{eu}_{c_d^{-2} r(r+\sqrt{x_{0,n}})}(x_0) \subset B_r(x_0) \subset B^{eu}_{2r(r+2\sqrt{x_{0,n}})}(x_0).$$

Proof: Let $y \in B_r(x_0)$. Theorem 5.6 implies that

$$d^{eu}(x_0,y) \le d(x_0,y)\left(\sqrt{x_{0,n}} + \sqrt{y_n} + \sqrt{d^{eu}(x_0,y)}\right) \le r\,(\sqrt{x_{0,n}} + \sqrt{y_n}) + \frac{1}{2}\,(r^2 + d^{eu}(x,y))$$

where we also used Young's inequality. It follows that

$$\frac{1}{2}\, d^{eu}(x,y) \le r\,(\sqrt{x_{0,n}} + \sqrt{y_n}) + \frac{1}{2}\, r^2.$$

But for $y \in B_r(x_0)$, thanks to Proposition 5.5 and Proposition 5.2 we furthermore have that

$$r > d(x_0,y) \ge d(x_0,(x_0',y_n)) = 2\,|\sqrt{x_{0,n}} - \sqrt{y_n}|.$$

If $y_n \ge x_{0,n}$, this means that

$$\sqrt{y_n} < \sqrt{x_{0,n}} + \frac{1}{2}\, r$$

and this estimate is also trivially true for $y_n < x_{0,n}$. So we have

$$\frac{1}{2}\, d^{eu}(x,y) \le r\,(2\sqrt{x_{0,n}} + \frac{1}{2}\, r) + \frac{1}{2}\, r^2 = r\,(2\sqrt{x_{0,n}} + r)$$

which proves the right inclusion.
For the left hand side we fix an $x \in B^{eu}_R(x_0)$ with an arbitrary $R > 0$. Using Theorem 5.6 again we

get

$$d(x_0, y) \le c_d \, \frac{d^{eu}(x_0, y)}{\sqrt{x_{0,n}} + \sqrt{d^{eu}(x_0, y)}} \le c_d \, \frac{R}{\sqrt{x_{0,n}} + \sqrt{R}}$$

since the function $a \mapsto \frac{a}{\sqrt{a+c}}$ is monotonically increasing.
Now set $R := c_d^{-2}\, r\,(r + \sqrt{x_{0,n}})$. Then

$$c_d \, \frac{R}{\sqrt{x_{0,n}} + \sqrt{R}} < c_d \, \frac{c_d^{-2}\, r\,(r + \sqrt{x_{0,n}})}{\sqrt{x_{0,n}} + c_d^{-1}\sqrt{r^2}} < r$$

concluding the proof. ∎

Thus the topologies defined by the Carnot-Caratheodory-metric d and the euclidean metric d^{eu} are equivalent. Intrinsic balls with a small ratio $\frac{\sqrt{x_{0,n}}}{r}$, indicating that they are situated rather close to ∂H, behave like euclidean balls with squared radius, while they are similar to euclidean balls of the same radius if $\frac{\sqrt{x_{0,n}}}{r}$ is large and the balls therefore further outside in H. In balls that stay away far enough from ∂H, the vertical component of points is comparable independent of the radius or any parameters.

5.9 Corollary *Let $x_0 \in \overline{H}$ and $r > 0$.*
If $5\, r \le \sqrt{x_{0,n}}$, then $x_n \approx x_{0,n}$ for any $x \in B_r(x_0)$.

Proof: By the ball inclusion from Proposition 5.8 it is clear that

$$B_r(x_0) \subset B^{eu}_{\frac{22}{25} x_{0,n}}(x_0).$$

Therefore, for any $x \in B_r(x_0)$ we get that

$$|x_{0,n} - x_n| \le d^{eu}(x_0, x) \le \frac{22}{25}\, x_{0,n}.$$

If $x_{0,n} \ge x_n$ this means that we also have $x_n > \frac{3}{25}\, x_{0,n}$, whereas for $x_{0,n} < x_n$ we additionally get $x_n < \frac{47}{25} x_{0,n}$. ∎

We now turn to the interplay between the intrinsic metric d and the naturally arising measure μ_σ.

5.10 Proposition *Let $\sigma > -1$, $x_0 \in \overline{H}$ and $r > 0$.*
Then we have

$$|B_r(x_0)|_\sigma \approx_{n,\sigma} r^n (r + \sqrt{x_{0,n}})^{n+2\sigma}.$$

Proof: Note first that by the inclusions from Proposition 5.8 we can easily compute an approximate value for the case $\sigma = 0$, since of course the Lebesgue measure of euclidean balls is known. Therefore we get

$$|B_r(x_0)| \approx_n r^n (r + \sqrt{x_{0,n}})^n.$$

In case the ball in consideration stays far away from the boundary, that is for $r < \frac{1}{5}\sqrt{x_{0,n}}$, this comment together with Corollary 5.9 implies

$$|B_r(x_0)|_\sigma = \int\limits_{B_r(x_0)} y_n^\sigma \, d\mathcal{L}^n(y) \approx x_{0,n}^\sigma \, |B_r(x_0)| \approx_n \sqrt{x_{0,n}}^{2\sigma}\, r^n \,(r + \sqrt{x_{0,n}})^n.$$

If $\sigma \geq 0$, then it is always true that $\sqrt{x_{0,n}}^{2\sigma} < (\sqrt{x_{0,n}} + r)^{2\sigma}$, and in the case away from the boundary we also have $\sqrt{x_{0,n}}^{2\sigma} > 2^{-2\sigma}\,(\sqrt{x_{0,n}} + r)^{2\sigma}$.
Conversely, if $-1 < \sigma < 0$, then $\sqrt{x_{0,n}}^{2\sigma} < 2^{-2\sigma}\,(\sqrt{x_{0,n}} + r)^{2\sigma}$ since we are away from the boundary, and $\sqrt{x_{0,n}}^{2\sigma} > (\sqrt{x_{0,n}} + r)^{2\sigma}$ regardless of the position of the ball. For balls far away from the boundary we therefore get the statement with constants depending on n and σ.
In an abuse of notation denote now the part of the cube centred at x_0 with side length $2R$ that lies within the upper half plane by

$$C_R(x_0) := \left\{ y \in \overline{H} \mid \max_{j=1,\dots,n} |x_{0,j} - y_j| < R \right\}.$$

Then we have

$$C_{\frac{R}{\sqrt{n}}}(x_0) \subset B_R^{eu}(x_0) \subset C_R(x_0),$$

and thus by the ball inclusions from Proposition 5.8 we get

$$|C_{c_d^{-2} n^{-\frac{1}{2}} r(r+\sqrt{x_{0,n}})}(x_0)|_\sigma \leq |B_r(x_0)|_\sigma \leq |C_{2r(r+2\sqrt{x_{0,n}})}(x_0)|_\sigma.$$

Similar as in Proposition 2.1 we compute $|C_R(x_0)|_\sigma$. If $x_{0,n} > R$, then

$$|C_R(x_0)|_\sigma = (2R)^{n-1} \int\limits_{(x_{0,n}-R, x_{0,n}+R)} x_n^\sigma \, d\mathcal{L}(x_n) = \frac{(2R)^{n-1}}{1+\sigma} \left((x_{0,n} + R)^{1+\sigma} - (x_{0,n} - R)^{1+\sigma} \right),$$

and if $x_{0,n} \leq R$, then

$$|C_R(x_0)|_\sigma = (2R)^{n-1} \int\limits_{(0, x_{0,n}+R)} x_n^\sigma \, d\mathcal{L}(x_n) = \frac{(2R)^{n-1}}{1+\sigma} \, (x_{0,n} + R)^{1+\sigma}.$$

Back to the main line of argument, we consider the case where $\sqrt{x_{0,n}} \leq 5\,r$. In this case we have $B_r(x_0) \subset C_{25r^2}(x_0)$ with $25\,r^2 \geq x_{0,n}$. Hence the second formula for the measure of the cube can be applied with $R = 25\,r^2$ to arrive at

$$|B_r(x_0)|_\sigma \lesssim_{n,\sigma} r^{2n-2}\,(x_{0,n} + r^2)^{1+\sigma} \leq r^{2n-2}\,(\sqrt{x_{0,n}} + r)^{2+2\sigma}.$$

But it is obvious that $(r + \sqrt{x_{0,n}})^2 \leq 6^2\,r^2$ in the case we are in, and $r^n \leq (r + \sqrt{x_{0,n}})^n$ always holds. For $\sqrt{x_{0,n}} \leq 5\,r$ we have thus seen that

$$|B_r(x_0)|_\sigma \lesssim_{n,\sigma} r^n\,(\sqrt{x_{0,n}} + r)^{n+2\sigma}.$$

For the estimate from below we divide the remaining case $\sqrt{x_{0,n}} \leq 5\,r$ into two subcases. Consider first the situation very close to the boundary, that is suppose $\sqrt{x_{0,n}} \leq n^{-\frac{1}{4}}\,c_d^{-1}\,r$. It is always true that

$$B_r(x_0) \supset C_{c_d^{-2} n^{-\frac{1}{2}} r^2}(x_0).$$

In our case, for half of the side length of this cube we have $c_d^{-2}\,n^{-\frac{1}{2}}\,r^2 > x_{0,n}$ and we can use the second formula for the measure of the cube again, this time with $R = c_d^{-2}\,n^{-\frac{1}{2}}\,r^2$. Therefore we get

$$|B_r(x_0)|_\sigma \gtrsim_{n,\sigma} r^{2n-2}\,(x_{0,n}+r^2)^{1+\sigma} \gtrsim_\sigma r^{2n-2}\,(\sqrt{x_{0,n}}+r)^{2+2\sigma} \geq r^{2n}\,(\sqrt{x_{0,n}}+r)^{2\sigma}.$$

Using that in the case in consideration we have

$$r^n \geq 2^{-n}\,(r+n^{\frac{1}{4}}\,c_d\,\sqrt{x_{0,n}})^n \geq 2^{-n}\,(r+\sqrt{x_{0,n}})^n$$

finishes this part.
It remains to prove the estimate from below in case we have $n^{-\frac{1}{4}}\,c_d^{-1}\,r < \sqrt{x_{0,n}} \leq 5\,r$. Independent of the location of the ball we always have

$$B_r(x_0) \supset C_{c_d^{-2}n^{-\frac{1}{2}}r\sqrt{x_{0,n}}}(x_0),$$

and in our case in addition $c_d^{-2}\,n^{-\frac{1}{2}}\,r\,\sqrt{x_{0,n}} \leq c_d^{-1}\,n^{-\frac{1}{4}}\,x_{0,n}$ by the left hand side bound on $\sqrt{x_{0,n}}$. Since $n^{-\frac{1}{4}}\,c_d^{-1} < 1$ this means that we can use the first formula for the measure of the cube to see that

$$|B_r(x_0)|_\sigma \gtrsim_{n,\sigma} r^{n-1}\,\sqrt{x_{0,n}}^{\,n-1}\left((x_{0,n}+c_d^{-2}\,n^{-\frac{1}{2}}\,r\,\sqrt{x_{0,n}})^{1+\sigma}-(x_{0,n}-c_d^{-2}\,n^{-\frac{1}{2}}\,r\,\sqrt{x_{0,n}})^{1+\sigma}\right).$$

By the right hand side bound on $\sqrt{x_{0,n}}$ given in this case we find $r\,\sqrt{x_{0,n}} \geq \frac{x_{0,n}}{5}$, and since $1+\sigma > 0$ we can use this in both summands to get

$$|B_r(x_0)|_\sigma \gtrsim_{n,\sigma} r^{n-1}\,\sqrt{x_{0,n}}^{\,n-1}\,x_{0,n}^{1+\sigma}\left((1+\frac{1}{5c_d^2\sqrt{n}})^{1+\sigma}-(1-\frac{1}{5c_d^2\sqrt{n}})^{1+\sigma}\right).$$

The term in parenthesis, however, is nothing but a positive constant depending on n and σ and we are left with

$$r^{n-1}\,\sqrt{x_{0,n}}^{\,n+1+2\sigma} \gtrsim_n r^n\,\sqrt{x_{0,n}}^{\,n+2\sigma} \gtrsim_{n,\sigma} r^n\,(r+\sqrt{x_{0,n}})^{n+2\sigma}$$

after using the left hand side bound on $\sqrt{x_{0,n}}$ if $n+2\sigma \geq 0$ and the right hand side bound on $\sqrt{x_{0,n}}$ if $n+2\sigma < 0$. ∎

5.11 Remark *By means of the formula from Proposition* 5.10 *we can not only expand the region where the μ_σ-measure of the ball is approximately given by the Lebesgue measure multiplied by the vertical component of the centre point – although paying for the improvement by additional dependencies on n and σ – but also use the same argument to show that far away from ∂H, where $x_{0,n} \eqsim x_n$ for any $x \in B_r(x_0)$, the measure of balls is independent of the centre point:*

(i) Let $\delta > 0$ (and possibly small). If $\delta\,r < \sqrt{x_{0,n}}$, then

$$|B_r(x_0)|_\sigma \eqsim_{n,\sigma,\delta} x_{0,n}^\sigma\,|B_r(x_0)|.$$

(ii) If $5\,r < \sqrt{x_{0,n}}$, then

$$|B_r(x_0)|_\sigma \eqsim_{n,\sigma} |B_r(x)|_\sigma$$

for any $x \in B_r(x_0)$.

As a consequence we can quantify the amount the ball measure grows when dilating or shrinking the radius.

5.12 Proposition *Let $\sigma > -1$, $x_0 \in \overline{H}$ and $r > 0$.*

(i) If $n + 2\sigma \geq 0$, then

$$|B_{\kappa r}(x_0)|_\sigma \lesssim_{n,\sigma} \begin{cases} \kappa^{2n+2\sigma}\,|B_r(x_0)|_\sigma \text{ for any } \kappa \geq 1 \\ \kappa^n\,|B_r(x_0)|_\sigma \text{ for any } 0 < \kappa < 1. \end{cases}$$

(ii) If $n + 2\sigma < 0$, then

$$|B_{\kappa r}(x_0)|_\sigma \lesssim_{n,\sigma} \begin{cases} \kappa^n\,|B_r(x_0)|_\sigma \text{ for any } \kappa \geq 1 \\ \kappa^{2n+2\sigma}\,|B_r(x_0)|_\sigma \text{ for any } 0 < \kappa < 1. \end{cases}$$

Proof: Let first $n + 2\sigma \geq 0$. If $\kappa \geq 1$, Proposition 5.10 extends to

$$|B_{\kappa r}(x_0)|_\sigma \lesssim_{n,\sigma} (\kappa\, r)^n\, \kappa^{n+2\sigma}\, (\sqrt{x_{0,n}} + r)^{n+2\sigma} \eqsim_{n,\sigma} \kappa^{2n+2\sigma}\, |B_r(x_0)|_\sigma.$$

If $\kappa < 1$, on the other hand, we get

$$|B_{\kappa r}(x_0)|_\sigma \lesssim_{n,\sigma} (\kappa\, r)^n\, (\sqrt{x_{0,n}} + r)^{n+2\sigma} \eqsim_{n,\sigma} \kappa^n\, |B_r(x_0)|_\sigma$$

as stated.
For $n + 2\sigma < 0$ the same arguments are interchanged, leading to the switch of cases in the statement. ∎

5.13 Remark *Note that the case $n + 2\sigma < 0$ can only occur for $n = 1$.*

In terms of the general theory, in any of the two possible cases the implication of Proposition 5.12 is a reformulation of the doubling condition A.13. It therefore holds whenever $\sigma > -1$.
Conversely, for $\sigma \leq -1$ the measure of balls containing 0 is not finite any more, so μ_σ can not be locally finite. For reference we formulate these facts in the following corollary.

5.14 Corollary *μ_σ is a locally finite doubling measure with respect to d on $\overline{H}$ if and only if $\sigma > -1$.*

Since μ_σ is even a Radon measure (compare Chapter 2), this means that $(\overline{H}, d, \mu_\sigma)$ is a space of homogeneous type as defined in A.14.
In this general context it also makes sense to ask which weight functions are in the p-Muckenhoupt class $A_p(\overline{H}, d, \mu_\sigma)$ given by Definition A.23.

5.15 Proposition *Let $\sigma > -1$ and $1 < p < \infty$. We then have:*

$$(\,\cdot\,)_n^{\rho-\sigma} \in A_p(\overline{H}, d, \mu_\sigma) \text{ if and only if } -1 < \rho < p\,(\sigma+1) - 1.$$

Proof: Let $(\,\cdot\,)_n^{\rho-\sigma} \in A_p(\overline{H}, d, \mu_\sigma)$. The general theory then asserts that

$$(\,\cdot\,)_n^{\rho-\sigma}\, d\mu_\sigma = (\,\cdot\,)_n^{\rho}\, d\mathcal{L}^n$$

is doubling on $\overline{H}$ with respect to d, see Proposition A.25. Corollary 5.14 shows that weighted measures are locally finite and doubling with respect to d if and only if their exponent is bigger than -1, so $\rho > -1$ follows.
Now by Proposition A.26 we know that $(\,\cdot\,)_n^{\rho-\sigma} \in A_p(\overline{H}, d, \mu_\sigma)$ is equivalent to

$$(\,\cdot\,)_n^{-\frac{\rho-\sigma}{p-1}} \in A_{\frac{p}{p-1}}(\overline{H}, d, \mu_\sigma).$$

The same arguments as before then imply that

$$(\,\cdot\,)_n^{-\frac{\rho-\sigma}{p-1}}\, d\mu_\sigma = (\,\cdot\,)_n^{-\frac{\rho-\sigma}{p-1}+\sigma}\, d\mathcal{L}^n$$

is doubling and locally finite, hence equivalently

$$-\frac{\rho-\sigma}{p-1} + \sigma > -1$$

which in turn is equivalent to

$$\rho < p\,(1+\sigma) - 1.$$

Conversely, if the condition on ρ holds both $(\,\cdot\,)_n^{\rho}\, d\mathcal{L}^n$ and $(\,\cdot\,)_n^{-\frac{\rho-\sigma}{p-1}+\sigma}\, d\mathcal{L}^n$ are locally finite and doubling, and especially the measure of any ball is finite. More precisely, with Proposition 5.10 we get for an arbitrary ball $B_r(x_0)$ that

$$\begin{aligned} &|B_r(x_0)|_\sigma^{-1} \int\limits_{B_r(x_0)} x_n^{\rho-\sigma}\, d\mu_\sigma(x) \left(|B_r(x_0)|_\sigma^{-1} \int\limits_{B_r(x_0)} x_n^{-\frac{\rho-\sigma}{p-1}}\, d\mu_\sigma(x) \right)^{p-1} \\ &= |B_r(x_0)|_\sigma^{-1}\, |B_r(x_0)|_\rho \left(|B_r(x_0)|_\sigma^{-1} |B_r(x_0)|_{\sigma-\frac{\rho-\sigma}{p-1}} \right)^{p-1} \eqsim_{n,\sigma} 1 \end{aligned}$$

regardless of x_0 and r. Therefore the Muckenhoupt condition for $(\,\cdot\,)_n^{\rho-\sigma}$ is satisfied. ∎

5.16 Remark *Up to this point we only dealt with the geometry of $-L_\sigma$, the elliptic, second order spatial part of our linear differential expression. To incorporate the temporal part ∂_t on an interval $I \subset \mathbb{R}$ we proceed as in [FSC86] and define a metric*

$$D((t,x),(s,y)) := \sqrt{|t-s| + d(x,y)^2}$$

on $I \times \overline{H}$. We get that $(I \times \overline{H}, D, \mathcal{L} \times \mu_\sigma)$ is a space of homogeneous type and that

$$(\,\cdot\,)_n^{\rho-\sigma} \in A_p(I \times \overline{H}, D, \mathcal{L} \times \mu_\sigma) \text{ if and only if } -1 < \rho < p(\sigma+1) - 1.$$

For later use we now investigate the behaviour of the measure of a ball when we replace the centre point. As a direct consequence of Proposition 5.12 it turns out that the exchange causes a loss that we can describe quite precisely.

5.17 Proposition *Let $\sigma > -1$, x_0, $y_0 \in \overline{H}$ and $r > 0$.*

(i) If $n + 2\sigma \geq 0$, then

$$|B_r(x_0)|_\sigma^\theta \lesssim_{n,\sigma,\theta} \left(1 + \frac{d(x_0, y_0)}{r}\right)^{(2n+2\sigma)|\theta|} |B_r(y_0)|_\sigma^\theta$$

for any $\theta \in \mathbb{R}$.

(ii) If $n + 2\sigma < 0$, then

$$|B_r(x_0)|_\sigma^\theta \lesssim_{n,\sigma,\theta} \left(1 + \frac{d(x_0, y_0)}{r}\right)^{n|\theta|} |B_r(y_0)|_\sigma^\theta$$

for any $\theta \in \mathbb{R}$.

Proof: We use the triangle inequality to see that $B_r(x_0) \subset B_{r+d(x_0,y_0)}(y_0)$. With Proposition 5.12 we then get

$$|B_r(x_0)|_\sigma \lesssim_{n,\sigma} \left(1 + \frac{d(x_0, y_0)}{r}\right)^{2n+2\sigma} |B_r(y_0)|_\sigma$$

if $n + 2\sigma \geq 0$, and the same expression with exponent n instead of $2n + 2\sigma$ if $n + 2\sigma < 0$. This proves the statement for $\theta \geq 0$.
However, note that the expression is symmetric with respect to x_0 and y_0. That means that for $\theta \geq 0$ we also get

$$|B_r(y_0)|_\sigma^\theta \lesssim_{n,\sigma,\theta} \left(1 + \frac{d(x_0, y_0)}{r}\right)^{(2n+2\sigma)\theta} |B_r(x_0)|_\sigma^\theta$$

if $n + 2\sigma \geq 0$, which in turns can be regrouped to

$$|B_r(x_0)|_\sigma^{-\theta} \lesssim_{n,\sigma,\theta} \left(1 + \frac{d(x_0, y_0)}{r}\right)^{(2n+2\sigma)\theta} |B_r(y_0)|_\sigma^{-\theta},$$

again with the obvious adjustments for the case $n + 2\sigma < 0$. ∎

5.18 Remark *The special form of the formula for the measure of balls from Proposition* 5.10 *allows us to use Proposition* 5.17 *to exchange x_n by y_n in any expression $(r + \sqrt{x_n})^\theta$, $\theta \in \mathbb{R}$ at the expense of the same factor as before. To see this write*

$$(r + \sqrt{x_n})^\theta = ((r + \sqrt{x_n})^n\, r^n)^{\frac{\theta}{n}}\; r^{-\theta} \eqsim_n |B_r(x)|^{\frac{\theta}{n}}\, r^{-\theta}$$

and use Proposition 5.17 *for $\sigma = 0$ with $\widetilde{\theta} = \frac{\theta}{n}$. We then obtain*

$$(r + \sqrt{x_n})^\theta \lesssim_{n,\theta} \left(1 + \frac{d(x, y)}{r}\right)^{2|\theta|} (r + \sqrt{y_n})^\theta$$

for any $\theta \in \mathbb{R}$.

6 Local Estimates

We first aim at a spatial localisation of the energy estimates in terms of the intrinsic metric. At the same time we also localise in time. Clearly, the metric depends on the relative position of points with respect to ∂H, and thus there will be different treatments near and far away from the boundary of H. However, we start with a proposition that does not depend on the point of the localisation or the metric at all. For a function u given on a subset of $\mathbb{R} \times \overline{H}$, the extension by zero onto $\mathbb{R} \times \overline{H}$ is denoted by the same symbol u.

6.1 Proposition *Let $\sigma > -1$, $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $\Omega \subset \overline{H}$ relatively open and $f \in L^2\big(I; L^2_\sigma(\Omega)\big)$. Further let $\eta \in C_c^\infty(\,\overline{I}\!\!\urcorner \times \Omega)$.*
If u is a σ-solution to f on $I \times \Omega$, then ηu is a σ-solution to $\eta\, f + \partial_t \eta\, u - L_\sigma \eta\, u - 2\,(\,\cdot\,)_n\, \nabla_x \eta \cdot \nabla_x u$ on $\overline{I}\!\!\urcorner \times \overline{H}$ with initial value $g = 0$, and on $\overline{I}$ we have

$$\big(\,\eta(t)\, f(t) - L_\sigma \eta(t)\, u(t) - 2\,(\,\cdot\,)_n\, \nabla_x \eta(t) \cdot \nabla_x u(t) \mid \eta(t)\, u(t)\,\big)_{L^2_\sigma(H)}$$
$$\lesssim r^2\, \|\eta(t)\, f(t)\|^2_{L^2_\sigma(H)} + r^{-2}\, \|\eta(t)\, u(t)\|^2_{L^2_\sigma(H)} + \|\nabla_x \eta(t)\, u(t)\|^2_{L^2_{1+\sigma}(H)}$$

for any $r > 0$.

Proof: For an arbitrary test function $\varphi \in C_c^\infty(\,\overline{I}\!\!\urcorner \times \overline{H})$ we compute

$$\begin{aligned}
&- \int_{I\times H} \eta\, u\, \partial_t \varphi\, d(\mathcal{L} \times \mu_\sigma) + \int_{I\times H} \nabla_x(\eta\, u) \cdot \nabla_x \varphi\, d(\mathcal{L} \times \mu_{1+\sigma}) \\
&= - \int_{I\times H} u\, \partial_t(\eta\, \varphi)\, d(\mathcal{L} \times \mu_\sigma) + \int_{I\times H} u\, \nabla_x \eta \cdot \nabla_x \varphi\, d(\mathcal{L} \times \mu_{1+\sigma}) + \int_{I\times H} u\, \varphi\, \partial_t \eta\, d(\mathcal{L} \times \mu_\sigma) \\
&\quad - \int_{I\times H} \varphi\, \nabla_x \eta \cdot \nabla_x u\, d(\mathcal{L} \times \mu_{1+\sigma}) + \int_{I\times H} \nabla_x u \cdot \nabla_x(\eta\, \varphi)\, d(\mathcal{L} \times \mu_{1+\sigma}).
\end{aligned}$$

Using that u is a σ-solution to f with test function $\eta\, \varphi \in C_c^\infty(I \times \Omega)$ then shows

$$\begin{aligned}
&- \int_{I\times H} \eta\, u\, \partial_t \varphi\, d(\mathcal{L} \times \mu_\sigma) + \int_{I\times H} \nabla_x(\eta\, u) \cdot \nabla_x \varphi\, d(\mathcal{L} \times \mu_{1+\sigma}) = \\
&= \int_{I\times H} u\, \varphi\, \partial_t \eta\, d(\mathcal{L} \times \mu_\sigma) + \int_{I\times H} u\, \nabla_x \eta \cdot \nabla_x \varphi\, d(\mathcal{L} \times \mu_{1+\sigma}) - \int_{I\times H} \varphi\, \nabla_x \eta \cdot \nabla_x u\, d(\mathcal{L} \times \mu_{1+\sigma}) \\
&\quad + \int_{I\times H} \varphi\, \eta\, f\, d(\mathcal{L} \times \mu_\sigma).
\end{aligned}$$

In a spatial integration by parts in the second term all the boundary terms vanish: although at $x_n = 0$ both φ and η can have values, $x_n^{1+\sigma} = 0$ holds there. We hence have that $\eta\, u$ is a σ-solution on $\overline{I}\!\!\urcorner \times \overline{H}$ to the inhomogeneity stated and with vanishing initial values.

Now we use another spatial integration by parts to see that

$$\begin{aligned}
-2\int_H (\,\cdot\,)_n^{1+\sigma}\, \nabla_x u(t)\cdot \nabla_x \eta(t)\, \eta(t)\, u(t)\, d\mathcal{L}^n \\
= 2\int_H L_\sigma \eta(t)\, \eta(t)\, u(t)^2\, d\mu_\sigma + 2\int_H (\,\cdot\,)_n^{1+\sigma}\, |\nabla_x \eta(t)|^2\, u(t)^2\, d\mathcal{L}^n \\
+ 2\int_H (\,\cdot\,)_n^{1+\sigma}\, \nabla_x u(t)\cdot \nabla_x \eta(t)\, \eta(t)\, u(t)\, d\mathcal{L}^n
\end{aligned}$$

and therefore

$$\begin{aligned}
-2(\, (\,\cdot\,)_n\, \nabla_x \eta(t)\cdot \nabla_x u(t) \mid \eta(t)\, u(t)\,)_{L^2_\sigma(H)} \\
= (\, L_\sigma \eta(t)\, u(t) \mid \eta(t)\, u(t)\,)_{L^2_\sigma(H)} + \|\nabla_x \eta(t)\, u(t)\|^2_{L^2_{1+\sigma}(H)}\,.
\end{aligned}$$

For the estimate stated we use Hölder's inequality on $(\, \eta(t)\, f(t) \mid \eta(t)\, u(t)\,)_{L^2_\sigma(H)}$ and multiply the result by $1 = r^2\, r^{-2}$ before using Young's inequality. ∎

We now specify the way in which we localise precisely. Reflecting the behaviour of the metric, the localisation will be chosen differently in a small spatial ball centred at $(0', 1) =: \vec{1}$ than in one centred at 0.
We also shift our paradigm concerning the time interval. Instead of defining it by means of its two end points, we will rather characterise it by its length r^2 and the situation of a special point s_0. We use the abbreviation $I_{r,\varepsilon}(s_0) := (s_0 + \varepsilon\, r^2, s_0 + r^2)$ for $r > 0$ and $\varepsilon \in [0, 1)$, dropping the second index if $\varepsilon = 0$.

6.2 Lemma *Let $s_0 \in \mathbb{R}$, $\varepsilon_1 \in [0, 1)$ and $\delta_1 \in (0, 1]$.*

(i) If $r \le 1$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \frac{1}{6c_d^2}\, \delta_1)$ there exists

$$\eta \in C_c^\infty(\, \bar{I}_{r,\varepsilon_1}(s_0) \times B_{\delta_1 r}(\vec{1}))$$

with

$$\eta = 1 \text{ on } \overline{I}_{r,\varepsilon_2}(s_0) \times \overline{B}_{\delta_2 r}(\vec{1})$$

and

$$|\partial_t^k \partial_x^\alpha \eta| \lesssim_{n,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-2k-|\alpha|} \text{ for any } k \in \mathbb{N}_0 \text{ and } \alpha \in \mathbb{N}_0^n.$$

(ii) For any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \frac{1}{\sqrt{2}c_d}\, \delta_1)$ there exists

$$\eta \in C_c^\infty(\, \bar{I}_{1,\varepsilon_1}(s_0) \times B_{\delta_1}(0))$$

with

$$\eta = 1 \text{ on } \overline{I}_{1,\varepsilon_2}(s_0) \times \overline{B}_{\delta_2}(0)$$

and

$$|\partial_t^k \partial_x^\alpha \eta| \lesssim_{n,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} 1 \text{ for any } k \in \mathbb{N}_0 \text{ and } \alpha \in \mathbb{N}_0^n.$$

Proof: For $\varepsilon_1 \in [0,1)$ consider the interval $(s_0 + \varepsilon_1 r^2, s_0 + (1+\varepsilon_2) r^2)$ with $\varepsilon_2 \in (\varepsilon_1, 1)$. It is then clear that for any such ε_2 the set $[s_0 + \varepsilon_2 r^2, s_0 + r^2]$ is contained in this first interval and hence there exists a cut-off function

$$\eta_t \in C_c^\infty((s_0 + \varepsilon_1 r^2, s_0 + (1+\varepsilon_1) r^2))$$

with

$$\eta_t = 1 \text{ on } \overline{I}_{r,\varepsilon_2}(s_0)$$

and

$$|\partial_t^k \eta_t| \lesssim_{k,\varepsilon_1,\varepsilon_2} r^{-2k} \text{ for any } k \in \mathbb{N}_0.$$

This holds for any $r > 0$.
In the spatial direction we first look at the case near $\vec{1}$ for $r \leq 1$. Given a $\delta_1 \in (0,1]$, the ball inclusions from Proposition 5.8 indicate that in this case we have

$$B_{\delta_1 r}(\vec{1}) \supset B^{eu}_{c_d^{-2}\delta_1 r}(\vec{1}) \text{ and } B_{\delta_2 r}(\vec{1}) \subset B^{eu}_{6\delta_2 r}(\vec{1}) \text{ for any } \delta_2 > 0.$$

Therefore, if $\delta_2 < \frac{1}{6c_d^2}\,\delta_1$, we have that $\overline{B}^{eu}_{6\delta_2 r}(\vec{1}) \subset B^{eu}_{c_d^{-2}\delta_1 r}(\vec{1})$ and hence get a cut-off function

$$\eta_x \in C_c^\infty(B^{eu}_{c_d^{-2}\delta_1 r}(\vec{1}))$$

with

$$\eta_x = 1 \text{ on } \overline{B}^{eu}_{6\delta_2 r}(\vec{1})$$

and

$$|\partial_x^\alpha \eta_x| \lesssim_{\alpha,\delta_1,\delta_2} r^{-|\alpha|} \text{ for any } \alpha \in \mathbb{N}_0^n.$$

Close to 0 the same reasoning applies with $r = 1$. Here Proposition 5.8 shows that

$$B_{\delta_1}(0) \supset B^{eu}_{c_d^{-2}\delta_1^2}(0) \cap \overline{H} \text{ and } B_{\delta_2}(0) \subset B^{eu}_{2\delta_2^2}(0) \text{ for any } \delta_2 > 0,$$

and we get an η_x with the same properties as above for $\delta_2 < \frac{1}{\sqrt{2}c_d}\,\delta_1$.
Both statements now follow if we take

$$\eta := \eta_t|_{\bar{I}_{r,\varepsilon_1}(s_0)}\,\eta_x|_{\overline{H}}$$

with the appropriate η_x, where the restriction onto $\overline{H}$ is redundant near $\vec{1}$ and we set $r = 1$ near 0. ∎

For a σ-solution u that is localised according to Lemma 6.2 with an inhomogeneity given by Proposition 6.1, we can now apply the global energy estimates to arrive at their local version. They are only valid on spatial sets that are small enough compared to the set $B_{\delta_1 r}(x)$ we start from, where $\delta_1 \in (0,1]$. More precisely, we assert that there exists a small $\delta_0 \in (0,\delta_1)$, such that for any $\delta_2 \in (0,\delta_0)$ the estimate holds on $B_{\delta_2 r}(x)$. We express this by simply saying that it holds for any $\delta_2 \in (0, \delta_1)$ small enough.

6.3 Proposition *Let $s_0 \in \mathbb{R}$, $\varepsilon_1 \in [0,1)$ and $\delta_1 \in (0,1]$.*

(i) If $r \leq 1$, $f \in L^2\big(I_{r,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1 r}(\vec{1}))\big)$ and u is a σ-solution to f on $I_{r,\varepsilon_1}(s_0) \times B_{\delta_1 r}(\vec{1})$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \delta_1)$ small enough we have

$$\int\limits_{I_{r,\varepsilon_2}(s_0)} \|\partial_t u\|^2_{L^2_\sigma(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\varepsilon_2}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\varepsilon_2}(s_0)} \left\|D^2_x u\right\|^2_{L^2_{2+\sigma}(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L}$$
$$\lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}.$$

(ii) If $f \in L^2\big(I_{1,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1}(0))\big)$ and u is a σ-solution to f on $I_{1,\varepsilon_1}(s_0) \times B_{\delta_1}(0)$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \delta_1)$ small enough we have

$$\int\limits_{I_{1,\varepsilon_2}(s_0)} \|\partial_t u\|^2_{L^2_\sigma(B_{\delta_2}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_2}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\delta_2}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_2}(s_0)} \left\|D^2_x u\right\|^2_{L^2_{2+\sigma}(B_{\delta_2}(0))}\, d\mathcal{L}$$
$$\lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} \int\limits_{I_{1,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L}.$$

Proof: We first concentrate on the case near $\vec{1}$ and consider $r \leq 1$. Let η be a cut-off function as in Lemma 6.2. Proposition 6.1 then tells us that $\eta\, u$ is a σ-solution to $\eta\, f + \partial_t \eta\, u - L_\sigma \eta\, u - 2x_n \nabla_x \eta \cdot \nabla_x u$ on $\overline{I}_{\varepsilon_1, r}(s_0)$ with initial value 0, where the prerequisits on f and the properties of η ensure that the inhomogeneity is in

$$L^2\big(I_{r,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1 r}(\vec{1}))\big) \cap L^1\big(I_{r,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1 r}(\vec{1}))\big).$$

We can hence apply the estimate from Remark 3.6, following from the global energy identity, onto $\eta\, u$. Together with the estimate from Proposition 6.1 this yields

$$\int\limits_{I_{r,\varepsilon_1}(s_0)} \|\nabla_x(\eta\, u)\|^2_{L^2_{1+\sigma}(H)}\, d\mathcal{L} \lesssim r^2 \int\limits_{I_{r,\varepsilon_1}(s_0)} \|\eta\, f\|^2_{L^2_\sigma(H)}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|\eta\, u\|^2_{L^2_\sigma(H)}\, d\mathcal{L}$$
$$+ \int\limits_{I_{r,\varepsilon_1}(s_0)} \left\|\sqrt{\partial_t \eta\, \eta}\, u\right\|^2_{L^2_\sigma(H)}\, d\mathcal{L} + \int\limits_{I_{r,\varepsilon_1}(s_0)} \|\nabla_x \eta\, u\|^2_{L^2_{1+\sigma}(H)}\, d\mathcal{L}.$$

By construction, η vanishes outside $B_{\delta_1 r}(\vec{1})$, while we have $\eta = 1$ on the smaller set

$$I_{r,\varepsilon_2}(s_0) \times B_{\delta_2 r}(\vec{1}) \subset I_{r,\varepsilon_1}(s_0) \times H$$

for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and $\delta_2 \in (0, \frac{1}{6c_d^2}\, \delta_1)$. With the bounds for the derivatives of η from Lemma 6.2 we then deduce

$$\int\limits_{I_{r,\varepsilon_2}(s_0)} \|\nabla_x u\|^2_{L^2_{1+\sigma}(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L} \lesssim_{n,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^2 \int\limits_{I_{r,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}$$
$$+ r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_{1+\sigma}(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}.$$

But locally near $\vec{1}$, lowering the weight exponent of $x_n^{1+\sigma}$ to x_n^{σ} is possible at the expense of an upper bound not depending on any parameters, causing the last term to blend in with the second to last.

Now consider $\widetilde{\varepsilon}_1 > \varepsilon_1$ and $\widetilde{\delta}_1 < \delta_1$. Then u is also a σ-solution to f on the smaller set $I_{r,\widetilde{\varepsilon}_1}(s_0) \times B_{\widetilde{\delta}_1 r}(\vec{1})$ and we choose η as in Lemma 6.2 with respect to this other set to get a global solution in space with zero initial values. This time we apply the global energy estimates from Proposition 4.4 and with the same reasoning as above get

$$\begin{aligned}
&\int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \|\partial_t u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} \\
&\lesssim_{n,\sigma,\widetilde{\varepsilon}_1,\widetilde{\varepsilon}_2,\widetilde{\delta}_1,\widetilde{\delta}_2} \int\limits_{I_{r,\widetilde{\varepsilon}_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\widetilde{\delta}_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-4} \int\limits_{I_{r,\widetilde{\varepsilon}_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_1 r}(\vec{1}))}\, d\mathcal{L} \\
&+ r^{-4} \int\limits_{I_{r,\widetilde{\varepsilon}_1}(s_0)} \|u\|^2_{L^2_{2+\sigma}(B_{\widetilde{\delta}_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\widetilde{\varepsilon}_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\widetilde{\varepsilon}_1}(s_0)} \|\nabla_x u\|^2_{L^2_{2+\sigma}(B_{\widetilde{\delta}_1 r}(\vec{1}))}\, d\mathcal{L}
\end{aligned}$$

for any $\widetilde{\varepsilon}_2 \in (\widetilde{\varepsilon}_1, 1)$ and any $\widetilde{\delta}_2$ that is small enough with respect to $\widetilde{\delta}_1$. Now in the fourth summand of the right hand side we can use that $r^{-2} \leq r^{-4}$ for $r \leq 1$, whereas in the third and in the last summand we lower the weight exponents as before, this time from $x_n^{2+\sigma}$ to x_n^{σ} and $x_n^{1+\sigma}$, respectively. This version of the last summand can then be estimated with the result of the first considerations in this proof if only $\widetilde{\delta}_1$ is small enough in relation to δ_1. Adjusting the sets appropriately we then gain the estimate

$$\begin{aligned}
&\int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \|\partial_t u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\widetilde{\varepsilon}_2}(s_0)} \left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(B_{\widetilde{\delta}_2 r}(\vec{1}))}\, d\mathcal{L} \\
&\qquad \lesssim_{n,\sigma,\varepsilon_1,\widetilde{\varepsilon}_2,\delta_1,\widetilde{\delta}_2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-4} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}
\end{aligned}$$

for any $\widetilde{\varepsilon}_2 \in (\varepsilon_1, 1)$ and $\widetilde{\delta}_2 \in (0, \delta_1)$ small enough, that is precisely for any $\widetilde{\delta}_2 \in (0, \frac{1}{36c_d^4}\,\delta_1)$.

But near $\vec{1}$, according to Corollary 5.9 we can not only decrease weights, but also increase them without loosing more than a constant factor in the estimates. Thus the first result above also delivers a bound for the time-space L^2_σ-norm of $\nabla_x u$ on $I_{r,\widetilde{\varepsilon}_2}(s_0) \times B_{\widetilde{\delta}_2 r}(\vec{1})$ that is better in terms of the exponent of r than the one just proven. This shows the full statement near $\vec{1}$.

We now turn to the situation near 0. The very same arguments as before, with the cut-off function η chosen appropriately according to Lemma 6.2 and with $r = 1$, result in the estimate

$$\begin{aligned}
&\int\limits_{I_{1,\widetilde{\varepsilon}_2}(s_0)} \|\nabla_x u\|^2_{L^2_{1+\sigma}(B_{\widetilde{\delta}_2}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\widetilde{\varepsilon}_2}(s_0)} \|\partial_t u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2}(0))}\, d\mathcal{L} \\
&\qquad + \int\limits_{I_{1,\widetilde{\varepsilon}_2}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\widetilde{\delta}_2}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\widetilde{\varepsilon}_2}(s_0)} \left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(B_{\widetilde{\delta}_2}(0))}\, d\mathcal{L} \\
&\qquad \lesssim_{n,\sigma,\varepsilon_1,\widetilde{\varepsilon}_2,\delta_1,\widetilde{\delta}_2} \int\limits_{I_{1,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L}
\end{aligned}$$

for any $\widetilde{\varepsilon}_2 \in (\varepsilon_1, 1)$ and $\widetilde{\delta}_2 \in (0, \delta_1)$ small enough, where the exact condition on the smallness of $\widetilde{\delta}_2$ is slightly different from the one above. Note carefully that apart from Lemma 6.2 and Proposition

6.1 only the lowering of the weight exponent plays a role in the argument, and this can be done also close to the origin. The absence of the different factors involving r in this case then allows us to use the third summand on the left hand side directly as an upper bound for the first one, giving the estimate the shape stated. ∎

We also need a local version of an auxiliary energy estimate to obtain a starting point for the iteration of the local energy estimates in terms of the derivatives. In fact, close to $\vec{1}$ the statement is redundant, since for any spatial second derivative, Proposition 6.3 delivers a result there that is much stronger than the following in various respects. However, near 0 the auxiliary result gives us control on the L^2-norm with weight $x_n^{1+\sigma}$ which otherwise we would not obtain. For a uniform treatment we include the superfluous case nonetheless.

6.4 Lemma *Let $s_0 \in \mathbb{R}$, $\varepsilon_1 \in [0,1)$ and $\delta_1 \in (0,1]$.*

(i) If $r \le 1$, $f \in L^2\big(I_{r,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1 r}(\vec{1}))\big)$ with $\nabla_x' f \in L^2\big(I_{r,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1 r}(\vec{1}))\big)$, and if u is a σ-solution to f on $I_{r,\varepsilon_1}(s_0) \times B_{\delta_1 r}(\vec{1})$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \delta_1)$ small enough we have

$$\int\limits_{I_{r,\varepsilon_2}(s_0)} \left\|\nabla_x \nabla_x' u\right\|^2_{L^2_{1+\sigma}(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L} \lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-6} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\varepsilon_1}(s_0)} \left\|\nabla_x' f\right\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}.$$

(ii) If $f \in L^2\big(I_{1,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1}(0))\big)$ with $\nabla_x' f \in L^2\big(I_{1,\varepsilon_1}(s_0);\, L^2_\sigma(B_{\delta_1}(0))\big)$, and if u is a σ-solution to f on $I_{1,\varepsilon_1}(s_0) \times B_{\delta_1}(0)$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0,\delta_1)$ small enough we have

$$\int\limits_{I_{1,\varepsilon_2}(s_0)} \left\|\nabla_x \nabla_x' u\right\|^2_{L^2_{1+\sigma}(B_{\delta_2}(0))}\, d\mathcal{L} \lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} \int\limits_{I_{1,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L} + \int\limits_{I_{1,\varepsilon_1}(s_0)} \left\|\nabla_x' f\right\|^2_{L^2_\sigma(B_{\delta_1}(0))}\, d\mathcal{L}.$$

Proof: In the same setting as in the proof of Proposition 6.3 we apply the spatial part of Lemma 4.6 onto the global solution $\eta\, u$. This requires one tangential derivative of the inhomogeneity given in Proposition 6.1. As a result, near $\vec{1}$ we get

$$\begin{aligned}\int\limits_{I_{r,\varepsilon_2}(s_0)} \left\|\nabla_x \nabla_x' u\right\|^2_{L^2_{1+\sigma}(B_{\delta_2 r}(\vec{1}))}\, d\mathcal{L} &\lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-6} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} \\ &+ r^{-4} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|\nabla_x u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \left\|D_x^2 u\right\|^2_{L^2_{2+\sigma}(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} \\ &+ r^{-2} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|f\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L} + \int\limits_{I_{r,\varepsilon_1}(s_0)} \left\|\nabla_x' f\right\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}\end{aligned}$$

where we also lowered the weight exponents again. Thanks to the local energy estimate 6.3 the result follows after possibly choosing δ_2 even smaller.
Close to 0 there is no difference in the argumentation. ∎

We turn to the iterated local energy estimates. In favor of a clearer presentation we now set $f = 0$. Remember also the convention $\partial_{x_n}^{-1} u := 0$ already used in the global setting.

6.5 Proposition *Let $s_0 \in \mathbb{R}$, $\varepsilon_1 \in [0,1)$ and $\delta_1 \in (0,1]$. Further let $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$,*

(i) If $r \leq 1$ and u is a σ-solution to $f = 0$ on $I_{r,\varepsilon_1}(s_0) \times B_{\delta_1 r}(\vec{1})$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ as well as any $\delta_2 \in (0, \delta_1)$ small enough we have that $\partial_t^k \partial_x^\alpha u$ is an $(\alpha_n + \sigma)$-solution to $\Delta'_x \partial_{x_n}^{\alpha_n - 1} u$ on $I_{r,\varepsilon_2}(s_0) \times B_{\delta_2 r}(\vec{1})$ and

$$\int\limits_{I_{r,\varepsilon_2}(s_0)} \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B_{\delta_2 r}(\vec{1}))} \, d\mathcal{L} \lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4k-2|\alpha|} \int\limits_{I_{r,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))} \, d\mathcal{L}.$$

(ii) If u is a σ-solution to $f = 0$ on $I_{1,\varepsilon_1}(s_0) \times B_{\delta_1}(0)$, then for any $\varepsilon_2 \in (\varepsilon_1, 1)$ and any $\delta_2 \in (0, \delta_1)$ small enough we have that $\partial_t^k \partial_x^\alpha u$ is an $(\alpha_n + \sigma)$-solution to $\Delta'_x \partial_{x_n}^{\alpha_n - 1} u$ on $I_{1,\varepsilon_2}(s_0) \times B_{\delta_2}(0)$ and

$$\int\limits_{I_{1,\varepsilon_2}(s_0)} \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B_{\delta_2}(0))} \, d\mathcal{L} \lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} \int\limits_{I_{1,\varepsilon_1}(s_0)} \|u\|^2_{L^2_\sigma(B_{\delta_1}(0))} \, d\mathcal{L}.$$

Proof: As in the global Proposition 4.9 we only have to check the regularity of the derivatives to ensure that they are solutions to the inhomogeneity stated. The reasoning here is the same at 0 and at $\vec{1}$, so we will not distinguish between the cases and merely write I and B in the notation. Naturally, the balls and intervals on the left hand side of the estimates are smaller than the ones on the right hand side in the same fashion as before, and the balls can be centred at $\vec{1}$ or at 0 in the same way.

We start with the tangential directions. For first order derivatives it is immediately clear by the local energy estimate 6.3 and the auxiliary estimate 6.4 that the right regularity is given on the smaller sets for which those estimates hold. Therefore $\nabla'_x u$ is a σ-solution to $f = 0$ there, and both estimates can be applied to $\nabla'_x u$. In an induction we make the sets smaller in every step by choosing δ_2 smaller every time. For any given $\alpha' \in \mathbb{N}_0^{n-1}$ we then get that $\partial_{x'}^{\alpha'} u$ is a σ-solution to 0 and consequently the energy estimate holds, that is we have

$$\begin{aligned} \int\limits_I \left\| \partial_t \partial_{x'}^{\alpha'} u \right\|^2_{L^2_\sigma(B)} \, d\mathcal{L} + r^{-2} \int\limits_I \left\| \nabla_x \partial_{x'}^{\alpha'} u \right\|^2_{L^2_\sigma(B)} \, d\mathcal{L} + \int\limits_I \left\| D_x^2 \partial_{x'}^{\alpha'} u \right\|^2_{L^2_{2+\sigma}(B)} \, d\mathcal{L} \\ \lesssim_{n,\sigma,\alpha',\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4-2|\alpha'|} \int\limits_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L}. \end{aligned}$$

The vertical direction is trickier. Here for a given $\alpha_n \in \mathbb{N}$ as an induction hypothesis we assume that $\partial_{x_n}^{\alpha_n} u$ is an $(\alpha_n + \sigma)$-solution to $\nabla'_x \partial_{x_n}^{\alpha_n} u$ and we have both the estimates

$$(*) \quad \begin{aligned} \int\limits_I \|\partial_t \partial_{x_n}^{\alpha_n} u\|^2_{L^2_{\alpha_n+\sigma}(B)} \, d\mathcal{L} + r^{-2} \int\limits_I \|\nabla_x \partial_{x_n}^{\alpha_n} u\|^2_{L^2_{\alpha_n+\sigma}(B)} \, d\mathcal{L} + \int\limits_I \left\| D_x^2 \partial_{x_n}^{\alpha_n} u \right\|^2_{L^2_{2+\alpha_n+\sigma}(B)} \, d\mathcal{L} \\ \lesssim_{n,\sigma,\alpha_n,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4-2\alpha_n} \int\limits_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L} \end{aligned}$$

and

$$(**) \quad \int\limits_I \|\nabla_x \nabla'_x \partial_{x_n}^{\alpha_n} u\|^2_{L^2_{1+\alpha_n+\sigma}(B)} \, d\mathcal{L} \lesssim_{n,\sigma,\alpha_n,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-6-2\alpha_n} \int\limits_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L}.$$

In case of the base clause $\alpha_n = 1$, by the local energy estimate and a decrease of weight exponents it is clear that $\partial_{x_n} u$ has the regularity required for a $(1+\sigma)$-solution. The auxiliary energy estimate 6.4 ensure that also the inhomogeneity $\Delta'_x u$ is in the right space and satisfies

$$\int_I \|\Delta'_x u\|^2_{L^2_{1+\sigma}(B)} \; d\mathcal{L} \lesssim_{n,\sigma,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-6} \int_I \|u\|^2_{L^2_\sigma(B)} \; d\mathcal{L}.$$

Therefore $\partial_{x_n} u$ is a $(1+\sigma)$-solution to $\Delta'_x u$. We can thus apply the energy estimate 6.3 onto this solution, decrease the weight exponents once and use the local energy estimate as well as the estimate for $\Delta'_x u$ above to get $(*)$ for $\alpha_n = 1$. Similarly, for $(**)$ we use the auxiliary energy estimate 6.4 on the newly found solution, lower the weight exponents and use the energy estimate as well as the knowledge on the Laplacian of a solution, albeit this time for $\nabla'_x u$ instead of u.
Likewise, in the inductive step $\alpha_n + 1$ the right regularity follows from the energy estimate and the auxiliary estimate for α_n. More precisely, the latter implies that

$$\int_I \|\Delta'_x \partial^{\alpha_n}_{x_n} u\|^2_{L^2_{1+\alpha_n+\sigma}(B)} \; d\mathcal{L} \lesssim_{n,\sigma,\alpha_n,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-6-2\alpha_n} \int_I \|u\|^2_{L^2_\sigma(B)} \; d\mathcal{L}.$$

Then $\partial^{\alpha_n+1}_{x_n} u$ is an $(\alpha_n+1+\sigma)$-solution to $\Delta'_x \partial^{\alpha_n}_{x_n} u$. Applying the local energy estimate 6.3 on this solution shows that

$$\begin{aligned}
\int_I \left\|\partial_t \partial^{\alpha_n+1}_{x_n} u\right\|^2_{L^2_{\alpha_n+1+\sigma}(B)} \; d\mathcal{L} + r^{-2} \int_I \left\|\nabla_x \partial^{\alpha_n+1}_{x_n} u\right\|^2_{L^2_{\alpha_n+1+\sigma}(B)} \; d\mathcal{L} + \int_I \left\|D^2_x \partial^{\alpha_n+1}_{x_n} u\right\|^2_{L^2_{2+\alpha_n+1+\sigma}(B)} \; d\mathcal{L} \\
\lesssim_{n,\sigma,\alpha_n,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4} \int_I \left\|\partial^{\alpha_n+1}_{x_n} u\right\|^2_{L^2_{\alpha_n+1+\sigma}(B)} \; d\mathcal{L} + \int_I \left\|\Delta'_x \partial^{\alpha_n+1}_{x_n} u\right\|^2_{L^2_{\alpha_n+1+\sigma}(B)} \; d\mathcal{L} \\
\lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4-2(\alpha_n+1)} \int_I \|u\|^2_{L^2_\sigma(B)} \; d\mathcal{L}
\end{aligned}$$

where the weight exponents were decreased in the first summand of the right hand side before we used the energy estimate $(*)$ from the α_n-th step on it, while the bound for the second summand was already computed above.
The iterated auxiliary estimate also follows as before: first apply the non-iterated auxiliar estimate 6.4 onto the solution, then lower the weight exponents and apply the iterated energy estimate $(*)$ from the α_n-th step on one summand and the inequality for the Laplacian of the α_n-th derivative once for u itself and once for $\nabla'_x u$. This finishes the induction and thus the proof in the vertical direction.
Finally we can deal with the time direction now. Given both the tangential and the vertical derivatives, a simple induction as in the tangential case shows that $\partial^k_t u$ is a σ-solution to 0 and we have

$$\begin{aligned}
\int_I \left\|\partial_t \partial^k_t u\right\|^2_{L^2_\sigma(B)} \; d\mathcal{L} + r^{-2} \int_I \left\|\nabla_x \partial^k_t u\right\|^2_{L^2_\sigma(B)} \; d\mathcal{L} + \int_I \left\|D^2_x \partial^k_t u\right\|^2_{L^2_{2+\sigma}(B)} \; d\mathcal{L} \\
\lesssim_{n,\sigma,k,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4-4k} \int_I \|u\|^2_{L^2_\sigma(B)} \; d\mathcal{L}.
\end{aligned}$$

We can then apply first the x_n-directional result, and then the other two estimates succintly on

$\partial_t^k \partial_x^\alpha u$ and obtain

$$\int_I \left\| \partial_t \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_{\alpha_n+\sigma}(B)} \, d\mathcal{L} + r^{-2} \int_I \left\| \nabla_x \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_{\alpha_n+\sigma}(B)} \, d\mathcal{L} + \int_I \left\| D_x^2 \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_{2+\alpha_n+\sigma}(B)} \, d\mathcal{L}$$
$$\lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4-4k-2|\alpha|} \int_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L}$$

Knowing now that $\partial_t^k \partial_x^\alpha u$ is a solution for any k and α we do not need the gain in derivatives any more and can thus state the estimate in the more compact form

$$\int_I \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_{\alpha_n+\sigma}(B)} \, d\mathcal{L} \lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} r^{-4k-2|\alpha|} \int_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L}.$$

Then it only remains to treat the weights on the left hand side. It is here that the different cases enter again. Close to $\vec{1}$ an increase of the weights is possible by Corollary 5.9. Near 0 on the other hand, we use Hardy's inequality 2.3 with $m = \alpha_n$ for $p = 2$ onto $\eta\, \partial_t^k \partial_x^\alpha u$, where η is a cut-off function near 0 in space only as constructed in Lemma 6.2. After possibly decreasing the domain of integration once more, this leads to

$$\begin{aligned}
\int_I \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B)} \, d\mathcal{L} &= \int_I \left\| \eta\, \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B)} \, d\mathcal{L} \\
&\lesssim_{n,\sigma,\alpha_n} \sum_{|\beta|=\alpha_n} \int_I \left\| \partial_x^\beta (\eta\, \partial_t^k \partial_x^\alpha u) \right\|^2_{L^2_{2\alpha_n+\sigma}(B)} \, d\mathcal{L} \\
&\lesssim_{n,\sigma,\alpha_n} \sum_{|\beta|=\alpha_n} \sum_{\gamma \le \beta} \int_I \left\| \partial_x^{\beta-\gamma} \eta\, \partial_t^k \partial_x^{\alpha+\gamma} u \right\|^2_{L^2_{2\alpha_n+\sigma}(B)} \, d\mathcal{L}.
\end{aligned}$$

According to Lemma 6.2, the derivatives of η are bounded by a constant depending on α and δ_1, δ_2 only. But we also have $\alpha_n = |\beta| \ge \beta_n \ge \gamma_n$ and can consequently lower the weight exponent $x_n^{2\alpha_n+\sigma}$ to $x_n^{\alpha_n+\gamma_n+\sigma}$. An application of the result with additional weights in balls centred at 0 shows that

$$\int_I \left\| \partial_t^k \partial_x^{\alpha+\gamma} u \right\|^2_{L^2_{\alpha_n+\gamma_n+\sigma}(B)} \, d\mathcal{L} \lesssim_{n,\sigma,k,\alpha,\varepsilon_1,\varepsilon_2,\delta_1,\delta_2} \int_I \|u\|^2_{L^2_\sigma(B)} \, d\mathcal{L}$$

holds there and the proposition is proven. ∎

We now use the invariance of the equation under the scaling T_λ from Remark 3.2 to gain a local estimate in any ball $B_r(x_0)$. The interplay between this coordinate transformation and the intrinsic metric is described by the following simple lemma that will also be useful later on.

6.6 Lemma *Let $\lambda > 0$ and $T_\lambda : (\hat{t}, \hat{x}) \mapsto (\lambda\, \hat{t}, \lambda\, \hat{x}) := (t, x)$. Further let s_0, $\hat{t}_0 \in \mathbb{R}$, $x_0, \hat{x}_0 \in \overline{H}$ and $r > 0$, $\varepsilon \in [0, 1)$.*
Then we have

$$T_\lambda(I_{r,\varepsilon}(\hat{t}_0) \times B_r(\hat{x}_0)) \subset I_{\sqrt{\lambda} r,\varepsilon}(\lambda \hat{t}_0) \times B_{4c_d^2 \sqrt{\lambda} r}(\lambda \hat{x}_0),$$

$$T_\lambda(I_{r,\varepsilon}(\hat{t}_0) \times B_r(\hat{x}_0)) \supset I_{\sqrt{\lambda} r,\varepsilon}(\lambda \hat{t}_0) \times B_{\frac{1}{4c_d^2} \sqrt{\lambda} r}(\lambda \hat{x}_0),$$

and

$$T_\lambda^{-1}(I_{r,\varepsilon}(s_0) \times B_r(x_0)) \subset I_{\frac{r}{\sqrt{\lambda}},\varepsilon}(\frac{s_0}{\lambda}) \times B_{4c_d^2 \frac{r}{\sqrt{\lambda}}}(\frac{1}{\lambda}x_0),$$

$$T_\lambda^{-1}(I_{r,\varepsilon}(s_0) \times B_r(x_0)) \supset I_{\frac{r}{\sqrt{\lambda}},\varepsilon}(\frac{s_0}{\lambda}) \times B_{\frac{1}{4c_d^2} \frac{r}{\sqrt{\lambda}}}(\frac{1}{\lambda}x_0).$$

Proof: Due to Proposition 5.8 we know that

$$\frac{1}{\lambda} B_r(x_0) \subset B^{eu}_{2\frac{r}{\sqrt{\lambda}}(\frac{r}{\sqrt{\lambda}}+2\sqrt{\frac{x_{0,n}}{\lambda}})}(\frac{1}{\lambda}x_0) \subset B_{4c_d^2 \frac{r}{\sqrt{\lambda}}}(\frac{1}{\lambda}x_0).$$

Conversely it is clear that

$$\frac{1}{\lambda} B_r(x_0) \supset B^{eu}_{\frac{1}{c_d^2}\frac{r}{\sqrt{\lambda}}(\frac{r}{\sqrt{\lambda}}+\sqrt{\frac{x_{0,n}}{\lambda}})}(\frac{1}{\lambda}x_0) \supset B_{\frac{1}{4c_d^2} \frac{r}{\sqrt{\lambda}}}(\frac{1}{\lambda}x_0).$$

With λ^{-1} replaced by λ the computation remains the same. ■

According to the position of a ball $B_r(x_0)$ relative to ∂H, expressed by the boundedness or unboundedness of the ratio of r and $\sqrt{x_{0,n}}$, we use different choices of λ for the scaling to get a globally valid local L^2_σ-estimate for arbitrary derivatives of a solution. For the sake of simplicity we set $f = 0$. This time also the temporal set has to shrink substantially, governed by a parameter ε. Similar as before we write that the estimate holds for any $\varepsilon \in (0,1)$ close enough to 1, meaning that there exists a $\varepsilon_0 \in (0,1)$ close to 1 such that it holds for any $\varepsilon \in (\varepsilon_0, 1)$.

6.7 Proposition *Let $s_0 \in \mathbb{R}$, $x_0 \in \overline{H}$ and $r > 0$. Further let $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$. If u is a σ-solution to $f = 0$ on $I_r(s_0) \times B_r(x_0)$, then for any $\varepsilon \in (0,1)$ close enough to 1 and any $\delta \in (0,1)$ small enough we have that $\partial_t^k \partial_x^\alpha u$ is an $(\alpha_n + \sigma)$-solution to $\Delta_x' \partial_{x_n}^{\alpha_n - 1} u$ on $I_{r,\varepsilon}(s_0) \times B_{\delta r}(x_0)$ and*

$$\int_{I_{r,\varepsilon}(s_0)} \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B_{\delta r}(x_0))} \, d\mathcal{L} \lesssim_{n,\sigma,k,\alpha,\varepsilon,\delta} r^{-4k-2|\alpha|}(r + \sqrt{x_{0,n}})^{-2|\alpha|} \int_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))} \, d\mathcal{L}.$$

Proof: Let $x_0 = (0, \dots, 0, x_{0,n})$ with $x_{0,n} \geq 0$ and fix $\kappa > 0$. We first consider the case $\sqrt{x_{0,n}} < \kappa r$ and set $\lambda := r^2$ for the invariant scaling T_λ from Remark 3.2. For $\delta \in (0,1)$, the triangle inequality for the metric d and Proposition 5.2 imply that

$$B_{\delta r}(x_0) \subset B_{(\delta + 2\kappa) r}(0)$$

and hence with $\varepsilon \in (0,1)$ also

$$\int_{I_{r,\varepsilon}(s_0)} \left\| \partial_t^k \partial_x^\alpha u \right\|^2_{L^2_\sigma(B_{\delta r}(x_0))} \, d\mathcal{L}(t) \leq \lambda^{-2k-2|\alpha|+n+1+\sigma} \int_{I_{1,\varepsilon}(\frac{s_0}{\lambda})} \left\| \partial_{\hat{t}}^k \partial_{\hat{x}}^\alpha (u \circ T_\lambda) \right\|^2_{L^2_\sigma(B_{4c_d^2(\delta+2\kappa)}(0))} \, d\mathcal{L}(\hat{t})$$

due to an application of the integral transformation formula in conjunction with Lemma 6.6 and our special choice of λ. On the other hand, Lemma 6.6 also shows that under this transformation, a σ-solution u on $I_r(s_0) \times B_r(x_0)$ becomes a σ-solution on

$$I_1(\frac{s_0}{r^2}) \times B_{\frac{1}{4c_d^2}}(\frac{1}{r^2}x_0).$$

The case we are considering now ensures that this ball is close enough to ∂H as to contain a ball centred at zero with radius $\frac{1}{4c_d^2} - 3\,\kappa$, if κ was given to be smaller than $\frac{1}{12c_d^2}$. In the calculation we use Proposition 5.2 again. The last ball is a superset of $B_{4c_d^2(\delta+2\kappa)}(0)$ if we choose $\delta \in (0,1)$ so small that

$$\delta < \frac{1}{16c_d^4} - \frac{8c_d^2+3}{4c_d^2}\kappa.$$

Such a δ can be found if from the start we had $\kappa < \frac{1}{32c_d^4+12c_d^2}$, so we could have set for example $\kappa := \frac{1}{33c_d^4}$, remembering that $c_d = 12$. All this allows us to use Proposition 6.5 near the origin after possibly choosing δ even smaller, and gain the upper bound

$$\lambda^{-2k-2|\alpha|+n+1+\sigma} \int\limits_{I_1(\frac{s_0}{\lambda})} \|u \circ T_\lambda\|^2_{L^2_\sigma(B_{\frac{1}{4c_d^2}-3\kappa}(0))} \; d\mathcal{L}(\hat{t}) \le \lambda^{-2k-2|\alpha|} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_{(1-12c_d^2\kappa)r}(0))} \; d\mathcal{L}(t)$$

with a constant depending on n, σ, k, α, ε and δ. The last inequality follows from the re-transformation, annulating the volume factor of the transformation above, and once again Lemma 6.6 and our special choice of λ.
The ball centred at zero on the right hand side of this inequality is contained in the ball $B_{r+2\kappa r-12c_d^2\kappa r}(x_0)$ by Proposition 5.8, and the latter is in turn a subset of $B_r(x_0)$ whenever $\kappa < \frac{1}{12c_d^2-2}$ which is obviously satisfied by $\kappa = \frac{1}{33c_d^4}$. Putting all this together we have shown that if $\sqrt{x_{0,n}} < \frac{1}{33c_d^4}\,r$ and u is a σ-solution to $f = 0$ on $I_r(s_0) \times B_r(x_0)$, then for all $k \in \mathbb{N}_0$, $\alpha \in \mathbb{N}_0^n$, for all $\varepsilon \in (0,1)$ and all $\delta \in (0,1)$ small enough we have

$$\int\limits_{I_{r,\varepsilon}(s_0)} \left\|\partial_t^k \partial_x^\alpha u\right\|^2_{L^2_\sigma(B_{\delta r}(x_0))} \; d\mathcal{L}(t) \lesssim_{n,\sigma,k,\alpha,\varepsilon,\delta} r^{-4k-4|\alpha|} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))} \; d\mathcal{L}(t).$$

Finally note that the factor in front of the integral can be estimated to become

$$r^{-4k-4|\alpha|} \le r^{-4k-2|\alpha|}\, 2^{2|\alpha|}\, (r+r)^{-2|\alpha|} \lesssim_\alpha r^{-4k-2|\alpha|}\, (\sqrt{x_{0,n}} + r)^{-2|\alpha|}$$

in the case we are considering.
We now turn to the opposite case $\sqrt{x_{0,n}} \ge \frac{1}{33c_d^4}\,r$. Here we set $\lambda := x_{0,n}$ and use κ as an abbreviation for the same value as above. For $\varepsilon \in (0,1)$ and $\delta \in (0,1)$, the transformation formula, Lemma 6.6 and the choice of λ in this case imply

$$\int\limits_{I_{r,\varepsilon}(s_0)} \left\|\partial_t^k \partial_x^\alpha u\right\|^2_{L^2_\sigma(B_{\delta r}(x_0))} \; d\mathcal{L}(t) \le \lambda^{-2k-2|\alpha|+n+1+\sigma} \int\limits_{I_{\frac{r}{\sqrt{\lambda}},\varepsilon}(\frac{s_0}{\lambda})} \left\|\partial_{\hat{t}}^k \partial_{\hat{x}}^\alpha (u \circ T_\lambda)\right\|^2_{L^2_\sigma(B_{4c_d^2\delta\frac{r}{\sqrt{\lambda}}}(\vec{1}))} \; d\mathcal{L}(\hat{t}).$$

Also by Lemma 6.6 we know that $u \circ T_\lambda$ is a σ-solution on

$$I_{\frac{r}{\sqrt{\lambda}}}(\frac{s_0}{\lambda}) \times B_{\frac{1}{4c_d^2}\frac{r}{\sqrt{\lambda}}}(\vec{1})$$

and thus on

$$I_{\frac{r}{\sqrt{\lambda}}}(\frac{s_0}{\lambda}) \times B_{\frac{\sqrt{\delta}r}{\sqrt{\lambda}}}(\vec{1})$$

if $\sqrt{\delta} \leq \frac{1}{4c_d^2}$. In order to get the same radius in time and in space we also make the interval smaller, but without changing the right end point. A short computation shows that we are dealing with a solution on

$$I_{\frac{\sqrt{\delta}r}{\sqrt{\lambda}}}(\frac{s_0}{\lambda} + (1-\delta)\frac{r^2}{\lambda}) \times B_{\frac{\sqrt{\delta}r}{\sqrt{\lambda}}}(\vec{1}).$$

The interval in the integral above equals

$$I_{\frac{\sqrt{\delta}r}{\sqrt{\lambda}},\frac{\delta-1+\varepsilon}{\delta}}(\frac{s_0}{\lambda} + (1-\delta)\frac{r^2}{\lambda})$$

with the new parameter $\frac{\delta-1+\varepsilon}{\delta} \in (0,1)$ if we started with an ε so close to 1 that $\varepsilon \in (1-\delta, 1)$. We can now apply Proposition 6.5 near $\vec{1}$ if only δ is so small that $4\,c_d^2\,\sqrt{\delta}$ is small enough compared to 1 and that the radius $\frac{\sqrt{\delta}r}{\sqrt{\lambda}}$ is bounded by 1. But the first condition for δ can always be fulfilled, and since $\lambda = x_{0,n}$ and $\sqrt{x_{0,n}} \geq \kappa\, r$, the second one is always satisfied for $\delta \leq \kappa^2$. For δ small enough and ε close enough to 1, Proposition 6.5 and the re-transformation with Lemma 6.6 again then show

$$\int\limits_{I_{r,\varepsilon}(s_0)} \left\|\partial_t^k \partial_x^\alpha u\right\|^2_{L^2_\sigma(B_{\delta r}(x_0))} \, d\mathcal{L}(t) \lesssim_{n,\sigma,k,\alpha,\varepsilon,\delta} r^{-4k-2|\alpha|}\,\sqrt{x_{0,n}}^{-2|\alpha|} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))}\, d\mathcal{L}(t).$$

Similar as above we can use the case assumption to see that

$$r^{-4k-2|\alpha|}\,\sqrt{x_{0,n}}^{-2|\alpha|} \lesssim_\alpha r^{-4k-2|\alpha|}\,(\sqrt{x_{0,n}}+r)^{-2|\alpha|}.$$

Since the equation is invariant under translation in any tangential direction, this finishes the proof. ∎

6.8 Remark *The factor involving r and $x_{0,n}$ that appears in estimate 6.7 is important. Henceforth we will abbreviate it by $c_{k,\alpha}(r,x_0) := r^{-2k-|\alpha|}\,(r+\sqrt{x_{0,n}})^{-|\alpha|}$.*

We now adapt a time-space Morrey-type inequality to our metric and measure. Once more we consider the cases near $\vec{1}$ and near 0 seperately first.

6.9 Lemma *Let $s_0 \in \mathbb{R}$, $\varepsilon \in [0,1)$ and $\delta_1 \in (0,1]$. Further let $m_1 \in \mathbb{N}$ with $m_1 > \frac{n+1}{2}$, $\sigma > -1$ and $m_2 \in \mathbb{N}_0$ with $m_2 \geq \frac{\sigma}{2}$.*

(i) If $r \leq 1$ and $u \in W^{m_1,2}_{\vec{\sigma}}(I_{r,\varepsilon}(s_0) \times B_{\delta_1 r}(\vec{1}))$ with $\sigma_i = \sigma$ for $i = 0,\ldots,m_1$, then for any $\delta_2 \in (0,\delta_1)$ small enough we have

$$|u(t,x)|^2 \lesssim_{n,\sigma,m_1,\varepsilon,\delta_1,\delta_2} \sum_{j+|\beta|\leq m_1} r^{4j+2|\beta|-n-1} \int\limits_{I_{r,\varepsilon}(s_0)} \left\|\partial_t^j \partial_x^\beta u\right\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))}\, d\mathcal{L}$$

for almost any $(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta_2 r}(\vec{1})$.

(ii) If $u \in W^{m_1+m_2,2}_{\vec{\sigma}}(I_{1,\varepsilon}(s_0) \times B_{\delta_1}(0))$ with $\sigma_i = \sigma$ for $i = 0, \dots, m_1 + m_2$, then for any $\delta_2 \in (0, \delta_1)$ small enough we have

$$|u(t,x)|^2 \lesssim_{n,\sigma,m_1,m_2,\varepsilon,\delta_1,\delta_2} \sum_{j+|\beta|\le m_1} \sum_{|\gamma|\le m_2} \int_{I_{1,\varepsilon}(s_0)} \left\| \partial_t^j \partial_x^{\beta+\gamma} u \right\|^2_{L^2_\sigma(B_{\delta_1}(0))} \, d\mathcal{L}$$

for almost any $(t,x) \in \overline{I}_{1,\varepsilon}(s_0) \times B_{\delta_2}(0)$.

Proof: Since near $\vec{1}$ we can both increase and decrease the weight exponent on expense of an upper bound depending on the difference between old and new weight exponents, it is clear that

$$\int_{I_{r,\varepsilon}(s_0)} \left\| \partial_t^j \partial_x^\beta u \right\|^2_{L^2(B_{\delta_1 r}(\vec{1}))} \, d\mathcal{L} \lesssim_\sigma \int_{I_{r,\varepsilon}(s_0)} \left\| \partial_t^j \partial_x^\beta u \right\|^2_{L^2_\sigma(B_{\delta_1 r}(\vec{1}))} \, d\mathcal{L}.$$

The right hand side is bounded by the prerequisits for any $j + |\beta| \le m_1$, and the ball $B_{\delta_1 r}(\vec{1})$ contains a euclidean ball centred at $\vec{1}$ with radius $c_d^{-2}\, \delta_1\, r$ by Proposition 5.8. We can then apply the usual time-space Morrey-type inequality for $m_1 > \frac{n+1}{2}$ on a euclidean time-space cylinder to obtain the statement. Here δ_2 has to be so small that

$$B_{\delta_2 r}(\vec{1}) \subset B^{eu}_{c_d^{-2}\delta_1 r}(\vec{1}).$$

Close to 0, the same argumentation applies if $\sigma \le 0$, since the weight can be lowered again. We then have

$$\int_{I_{1,\varepsilon}(s_0)} \left\| \partial_t^j \partial_x^\beta u \right\|^2_{L^2(B_{\delta_1}(0))} \, d\mathcal{L} \lesssim_\sigma \int_{I_{1,\varepsilon}(s_0)} \left\| \partial_t^j \partial_x^\beta u \right\|^2_{L^2_\sigma(B_{\delta_1}(0))} \, d\mathcal{L},$$

which is bounded for $j + |\beta| \le m_1$ again.
However, for $\sigma > 0$ we have to reason differently since near 0 the weight exponents can not be increased. We therefore consider $\delta_2 \in (0, \delta_1)$ small enough and a cut-off function η as in Lemma 6.2, but purely spatial, with $\eta = 1$ on $B_{\delta_2}(0)$, $\operatorname{supp} \eta \subset B_{\delta_1}(0)$ and any derivative bounded by a constant depending on δ_1, δ_2 only. An application of Hardy's inequality 2.3 with m_2 on $\eta\, \partial_t^j \partial_x^\beta u$ is possible for any $j + |\beta| \le m_1$ by the assumptions, and so we get

$$\begin{aligned}
\left\| \partial_t^j \partial_x^\beta u \right\|_{L^2(B_{\delta_2}(0))} &\le \left\| \eta\, \partial_t^j \partial_x^\beta u \right\|_{L^2(H)} \\
&\lesssim_{n,m_2} \sum_{|\widetilde{\gamma}|=m_2} \sum_{\gamma \le \widetilde{\gamma}} \left\| \partial_x^{\widetilde{\gamma}-\gamma} \eta\, \partial_t^j \partial_x^{\beta+\gamma} u \right\|_{L^2_{2m_2}(H)} \\
&\lesssim_{n,m_2,\delta_1,\delta_2} \sum_{|\gamma|\le m_2} \left\| \partial_t^j \partial_x^{\beta+\gamma} u \right\|_{L^2_{2m_2}(B_{\delta_1}(0))}.
\end{aligned}$$

But the prerequisits on m_2 assure that $2\, m_2 \ge \sigma$, so we can lower the weight exponent again and get an upper bound for the L^2-norm of $\partial_t^j \partial_x^\beta u$ that is finite as before for $j + |\beta| \le m_1$ thanks to the assumption on u. The time-space Morrey-type inequality in the euclidean setting for $m_1 > \frac{n+1}{2}$ then finishes the proof as above after possibly choosing δ_2 smaller. ■

This enables us to get a pointwise estimate for any derivatives on a smaller time-space set, showing that local σ-solutions to the homogeneous problem are indeed smooth.

6.10 Theorem *Let $s_0 \in \mathbb{R}$, $x_0 \in \overline{H}$ and $r > 0$. Further let $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$.*
If u is a σ-solution to $f = 0$ on $I_r(s_0) \times B_r(x_0)$, then for any $\varepsilon \in (0,1)$ close enough to 1 and any $\delta \in (0,1)$ small enough we have that

$$|\partial_t^k \partial_x^\alpha u(t,x)|^2 \lesssim_{n,\sigma,k,\alpha,\varepsilon,\delta} c_{k,\alpha}(r,x_0)^2\, r^{-2}\, |B_r(x_0)|_\sigma^{-1} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))}\, d\mathcal{L}$$

for any $(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta r}(x_0)$.

Proof: Thanks to translation invariance in any direction save the vertical one, we can again assume without loss of generality that $x_0 = (0,\dots,0,x_{0,n})$ with $x_{0,n} \geq 0$. We define $\kappa := \frac{1}{33c_d^4}$, denote the invariant scaling from Remark 3.2 by T_λ and also elsewhere follow closely the proof of Proposition 6.7. Furthermore, m_1 will always denote the first positive integer that is bigger than $\frac{n+1}{2}$, and m_2 the first positive integer that is bigger than $\frac{\sigma}{2}$.
Assume first that $\sqrt{x_{0,n}} < \kappa\, r$ and set $\lambda := r^2$. For any $\delta_2 \in (0,1)$ small enough we have

$$B_{\delta_2 r}(x_0) \subset B_{(\delta_2 + 2\kappa) r}(0)$$

due to Proposition 5.2. Given also $\varepsilon \in (0,1)$ this means that

$$(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta_2 r}(x_0)$$

implies

$$T^{-1}(t,x) =: (\hat{t},\hat{x}) \in \overline{I}_{1,\varepsilon}(\hat{t}_0) \times B_{4c_d^2(\delta_2+2\kappa)}(0)$$

with $\hat{t}_0 = \frac{s_0}{r^2}$. If δ_2 is so small compared to a $\delta_1 \in (0,1)$ that $4\,c_d^2\,(\delta_2 + 2\kappa)$ is small enough compared to $4\,c_d^2\,(\delta_1 + 2\kappa)$ for Lemma 6.9 to be applied, then for almost any $(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta_2 r}(x_0)$ we see that

$$\begin{aligned} |\partial_t^k \partial_x^\alpha u(t,x)|^2 &= \lambda^{-2k-2|\alpha|}\, |\partial_{\hat{t}}^k \partial_{\hat{x}}^\alpha (u \circ T_\lambda)(\hat{t},\hat{x})|^2 \\ &\lesssim_{n,\sigma,\varepsilon,\delta_2} \lambda^{-2k-2|\alpha|} \sum_{j+|\beta| \leq m_1} \sum_{|\gamma| \leq m_2} \int\limits_{I_{1,\varepsilon}(\hat{t}_0)} \left\| \partial_{\hat{t}}^{j+k} \partial_{\hat{x}}^{\beta+\gamma+\alpha} (u \circ T_\lambda) \right\|^2_{L^2_\sigma(B_{4c_d^2(\delta_1+2\kappa)}(0))} d\mathcal{L}. \end{aligned}$$

As before we can now use Proposition 6.5 near the origin if ε is close enough to 1, and δ_1 and all the more δ_2 are small enough. With a constant depending on n, σ, k, α, δ_1 and ε, this yields the upper bound

$$\lambda^{-2k-2|\alpha|} \int\limits_{I_1(\hat{t}_0)} \|u \circ T_\lambda\|^2_{L^2_\sigma(B_{\frac{1}{4c_d^2}-3\kappa}(0))}\, d\mathcal{L} \leq \lambda^{-2k-2|\alpha|-1-n-\sigma} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))}\, d\mathcal{L}$$

by the integral transformation formula, adding the volume factor in front of the integral. Since $\lambda = r^2$, the whole factor can be estimated to be

$$\begin{aligned} r^{-4k-4|\alpha|-2-2n-2\sigma} &\lesssim_\alpha c_{k,\alpha}(r,x_0)^2\, r^{-2}\, r^{-n}\, r^{-n-2\sigma} \lesssim_{n,\sigma} c_{k,\alpha}(r,x_0)^2\, r^{-2}\, r^{-n}\, (r + \sqrt{x_{0,n}})^{-n-2\sigma} \\ &\eqsim_{n,\sigma} c_{k,\alpha}(r,x_0)^2\, r^{-2}\, |B_r(x_0)|_\sigma^{-1} \end{aligned}$$

by Proposition 5.10. Note that in the step second to last we distinguish between the case where $n + 2\sigma \geq 0$, then using the assumption $\sqrt{x_{0,n}} < \kappa\, r$, and the case $n + 2\sigma < 0$.

Assume now $\sqrt{x_{0,n}} \geq \kappa\, r$. As before we then set $\lambda := x_{0,n}$, and moreover use the abbreviations $\hat{r} := \frac{r}{\sqrt{x_{0,n}}}$ and $\hat{t}_0 := \frac{s_0}{x_{0,n}}$. Let δ_2 and ε be given. If $(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta_2 r}(x_0)$ we have

$$T^{-1}(t,x) =: (\hat{t}, \hat{x}) \in \overline{I}_{\hat{r},\varepsilon}(\hat{t}_0) \times B_{4c_d^2 \delta_2 \hat{r}}(\vec{1})$$

and Lemma 6.9 can be used if δ_2 is small enough compared to a $\delta_1 < \frac{1}{4c_d^2}$. This means that for almost any $(t,x) \in \overline{I}_{r,\varepsilon}(s_0) \times B_{\delta_2 r}(x_0)$ we get

$$|\partial_t^k \partial_x^\alpha u(t,x)|^2 \lesssim_{n,\sigma,\varepsilon,\delta_2} \lambda^{-2k-2|\alpha|} \sum_{j+|\beta| \leq m_1} \hat{r}^{4j+2|\beta|-1-n} \int\limits_{I_{\hat{r},\varepsilon}(\hat{t}_0)} \left\| \partial_{\hat{t}}^{j+k} \partial_{\hat{x}}^{\beta+\alpha} (u \circ T_\lambda) \right\|^2_{L^2_\sigma(B_{4c_d^2 \delta_1 \hat{r}}(\vec{1}))} \, d\mathcal{L}.$$

An application of Proposition 6.5 as before annulates the factor $\hat{r}^{4j+2|\beta|}$ and yields in addition $\hat{r}^{-4k-2|\alpha|}$. Transformation of the integral furthermore results in a volume factor. Putting this together then delivers

$$|\partial_t^k \partial_x^\alpha u(t,x)|^2 \lesssim_{n,\sigma,\varepsilon,\delta_2} \lambda^{-2k-2|\alpha|-1-n-\sigma}\, \hat{r}^{-4k-2|\alpha|-n-1} \int\limits_{I_r(s_0)} \|u\|^2_{L^2_\sigma(B_r(x_0))} \, d\mathcal{L}.$$

But we had $\lambda = x_{0,n}$ and $\hat{r} = \frac{r}{x_{0,n}}$ and thus get the factor

$$\begin{aligned} r^{-4k-2|\alpha|-1-n} \sqrt{x_{0,n}}^{-2|\alpha|-1-n-2\sigma} &\lesssim_{\alpha,n,\sigma} c_{k,\alpha}(r,x_0)^2\, r^{-1}\, r^{-n} \left(\sqrt{x_{0,n}} + r\right)^{-1-n-2\sigma} \\ &\lesssim_{n,\sigma} c_{k,\alpha}(r,x_0)^2\, r^{-2}\, |B_r(x_0)|_\sigma^{-1} \end{aligned}$$

as stated. ∎

7 Estimates Against Initial Values

We now consider initial value problems on $\overline{H}$ again, bearing in mind that any solution to the time-local problem on (t_1, t_2) is also a solution to the initial value problem on $[\widetilde{t}_1, t_2)$ for any $\widetilde{t}_1 \in I$.
The first result is a very easy consequence of the pointwise estimate 6.10, stating that the smoothness of σ-solutions to $f = 0$ carries over from the local situation to the Cauchy problem.

7.1 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $g \in L^2_\sigma(H)$. Further let $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$.*
If u is a σ-solution to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g, then

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} \frac{c_{k,\alpha}(\sqrt{t_0 - s_0}, x_0)}{|B_{\sqrt{t_0-s_0}}(x_0)|_\sigma^{\frac{1}{2}}} \, \|u(s_0)\|_{L^2_\sigma(H)}$$

for any $x_0 \in \overline{H}$ and any $s_0 < t_0 \in \overline{I}$.

Proof: A σ-solution u to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g is also one to $f = 0$ on $[s_0, t_0) \times B_r(x_0)$ with initial value $u(s_0)$ for any pair of points $s_0 < t_0 \in \overline{I}$, any point $x_0 \in \overline{H}$ and any radius r. For $r := \sqrt{t_0 - s_0}$ it is therefore possible to use the pointwise estimate 6.10 in the temporal end point t_0 and the spatial centre point x_0 and get

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r, x_0)\, r^{-1}\, |B_r(x_0)|_\sigma^{-\frac{1}{2}}\, r \sup_{t \in (s_0, t_0)} \|u(t)\|_{L^2_\sigma(B_r(x_0))}$$

for any $s_0 < t_0 \in \overline{I}$ and $x_0 \in \overline{H}$. Passing from $B_r(x_0)$ to the whole space H in the L^2_σ-norm, we know from Corollary 3.7 that the latter decreases in time not only on (s_0, t_0), but also on $[s_0, t_0)$ thanks to there being a solution to the inital-value-problem. ∎

The decrease of the L^2_σ-norm on whole space was crucial in this proof. We can now extend the norm-decrease 3.7 of solutions by an exponential factor depending on the equivalent quasi-metric $\tilde{d}$ defined in Remark 5.7. To this end we consider the function

$$\Psi(x; c_\Psi, \varepsilon_\Psi, y_0) := c_\Psi \frac{\tilde{d}(x, y_0)^2}{\sqrt{\varepsilon_\Psi + \tilde{d}(x, y_0)^2}} \text{ for any } x \in \overline{H},$$

where $c_\Psi \in \mathbb{R}$, $\varepsilon_\Psi > 0$ are constants and $y_0 \in \overline{H}$ is arbitrary, but fixed.

7.2 Lemma $\Psi(\cdot; c_\Psi, \varepsilon_\Psi, y_0) \in C^1(H)$ *and*

$$\sqrt{x_n}\, |\nabla_x \Psi(x; c_\Psi, \varepsilon_\Psi, y_0)| \le c_L\, |c_\Psi| \textit{ for any } x \in \overline{H} \textit{ with } c_L := 2^6.$$

Proof: Direct calculations show that for any $x \in \overline{H}$ we have $x_n\, |\nabla_x \tilde{d}(x, y_0)|^2 \le 2^7$ and thus

$$|\nabla_x \Psi(x; c_\Psi, \varepsilon_\Psi, y_0)|^2 \le c_\Psi^2\, 2^5 |\nabla_x \tilde{d}(x, y_0)| \le 2^{12}\, c_\Psi^2\, x_n^{-1}$$

as stated. ∎

7.3 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $g \in L^2_\sigma(H)$ and $y_0 \in \overline{H}$. If u is a σ-solution to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g, then we have*

$$\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t) \right\|_{L^2_\sigma(H)} \leq e^{c_L^2 c_\Psi^2 (t-t_1)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, g \right\|_{L^2_\sigma(H)}$$

for any $t \in \overline{I}$.

Proof: For $t \in I$ we define

$$F(t) := e^{-2c_L^2 c_\Psi^2 (t-t_1)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t) \right\|^2_{L^2_\sigma} + 2 \int\limits_{(t_1,t)} e^{-2c_L^2 c_\Psi^2 (\tau-t_1)} \left\| \nabla_x (e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(\tau)) \right\|^2_{L^2_{1+\sigma}} d\mathcal{L}(\tau).$$

Then obviously we have

$$\begin{aligned}\partial_t F(t) = e^{-2c_L^2 c_\Psi^2 (t-t_1)}\, \partial_t \Big(\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t) \right\|^2_{L^2_\sigma} \Big) - 2\, c_L^2\, c_\Psi^2\, e^{-2c_L^2 c_\Psi^2 (t-t_1)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t) \right\|^2_{L^2_\sigma} \\ + 2\, e^{-2c_L^2 c_\Psi^2 (t-t_1)} \left\| \nabla_x (e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t)) \right\|^2_{L^2_{1+\sigma}}\end{aligned}$$

for all $t \in I$.
So consider

$$\partial_t \Big(\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t) \right\|^2_{L^2_\sigma} \Big) = 2 \int\limits_H e^{2\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, u(t)\, \partial_t u(t)\, d\mu_\sigma.$$

The differentiation under the integral is justified by the product rule for bilinear forms. Abbreviating $\Psi(\cdot; c_\Psi, \varepsilon_\Psi, y_0) =: \Psi$ henceforth, we use $e^{2\Psi}\, u(t)$ formally as a test function for the σ-solution u in the equivalent formulation for solutions with regular temporal derivative – compare the proof and the statement of Proposition 4.1 and Remark 4.2 – and find that for almost any $t \in I$ we have

$$\partial_t \Big(\left\| e^{\Psi} u(t) \right\|^2_{L^2_\sigma} \Big) = -2 \int\limits_H (\nabla_x (e^{\Psi}\, u(t)) + e^{\Psi}\, u(t) \nabla_x \Psi) \cdot \nabla_x u(t)\, e^{\Psi}\, d\mu_{1+\sigma}$$

by virtue of the chain rule. With $e^{\Psi}\, \nabla_x u(t) = \nabla_x (e^{\Psi}\, u(t)) - e^{\Psi}\, u(t)\, \nabla_x \Psi$ this yields

$$\begin{aligned}\partial_t \Big(\left\| e^{\Psi}\, u(t) \right\|^2_{L^2_\sigma} \Big) &= -2 \left\| \nabla_x (e^{\Psi}\, u(t)) \right\|^2_{L^2_{1+\sigma}} + 2 \left\| e^{\Psi}\, u(t)\, \nabla_x \Psi \right\|^2_{L^2_{1+\sigma}} \\ &\leq -2 \left\| \nabla_x (e^{\Psi}\, u(t)) \right\|^2_{L^2_{1+\sigma}} + 2\, c_L^2\, c_\Psi^2 \left\| e^{\Psi}\, u(t) \right\|^2_{L^2_\sigma},\end{aligned}$$

where the use of Lemma 7.2 in the last step is crucial to reduce the $(1+\sigma)$-norm in the last summand to a σ-norm.
But this shows that $\partial_t F(t) \leq 0$ on I. Thus F is monotonically decreasing and we have

$$F(t) \leq F(t_1) = \left\| e^{\Psi} u(t_1) \right\|^2_{L^2_\sigma} \text{ for all } t \in I.$$

To justify the formal calculation we have to use a bounded approximation of Ψ that does not change the estimate for $|\nabla_x \Psi|$ in the limit. ■

7.4 Remark *Although we will not use it in the sequel, we would like to note that the proof showed more than we stated, namely also the estimate*

$$2 \int_{(t_1,t)} e^{-2c_L^2 c_\Psi^2 (\tau - t_1)} \left\| \nabla_x (e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(\tau)) \right\|_{L^2_{1+\sigma}(H)} d\mathcal{L}(\tau) \leq \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} g \right\|^2_{L^2_\sigma(H)} \text{ for all } t \in \overline{I}.$$

The decrease of the exponential L^2_σ-norm enables us now to get an exponential version of 7.1.

7.5 Proposition *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $g \in L^2_\sigma(H)$ and $y_0 \in \overline{H}$. If u is a σ-solution to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g, then*

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} \frac{c_{k,\alpha}(\sqrt{t_0 - s_0}, x_0)}{|B_{\sqrt{t_0-s_0}}(x_0)|_\sigma^{\frac{1}{2}}} e^{2c_L^2 c_\Psi^2 (t_0 - s_0) - \Psi(x_0;c_\Psi,\varepsilon_\Psi,y_0)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(s_0) \right\|_{L^2_\sigma(H)}$$

for any $x_0 \in \overline{H}$ and any $s_0 < t_0 \in \overline{I}$.

Proof: Let again $r := \sqrt{t_0 - s_0}$. As in the proof of Proposition 7.1, but with an additional factor $1 = e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}$ in the norm on the right hand side, it is clear that

$$\begin{aligned} &|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \\ &\quad \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r, x_0)\, r^{-1}\, |B_r(x_0)|_\sigma^{-\frac{1}{2}} \sup_{x \in B_r(x_0)} e^{-\Psi(x;c_\Psi,\varepsilon_\Psi,y_0)}\, r \sup_{t \in (s_0,t_0)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(t) \right\|_{L^2_\sigma(B_r(x_0))}. \end{aligned}$$

On the whole space, however, Proposition 7.3 ensures that the L^2_σ-norm of $e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(t)$ on (s_0, t_0) is bounded by $e^{c_L^2 c_\Psi^2 r^2} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(s_0) \right\|_{L^2_\sigma(H)}$, since we consider an initial value problem on $[s_0, t_0)$. If we multiply the resulting inequality by $e^{\Psi(x_0;c_\Psi,\varepsilon_\Psi,y_0)}$ on both sides and use that by Lemma 7.2 and Proposition 5.1 we have $\Psi(x_0; c_\Psi, \varepsilon_\Psi, y_0) - \Psi(x; c_\Psi, \varepsilon_\Psi, y_0) \leq c_L\, |c_\Psi|\, d(x, x_0)$, then we arrive at

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r, x_0)\, |B_r(x_0)|_\sigma^{-\frac{1}{2}}\, e^{c_L^2 c_\Psi^2 r^2 + c_L |c_\Psi| r - \Psi(x_0;c_\Psi,\varepsilon_\Psi,y_0)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} u(s_0) \right\|_{L^2_\sigma(H)}.$$

We can compute $e^{c_L^2 c_\Psi^2 r^2 + c_L |c_\Psi| r} \leq e^{2c_L^2 c_\Psi^2 r^2}$ if $c_L |c_\Psi| r > 1$ and $e^{c_L^2 c_\Psi^2 r^2 + c_L |c_\Psi| r} \leq e\, e^{c_L^2 c_\Psi^2 r^2} \leq e\, e^{2c_L^2 c_\Psi^2 r^2}$ if $c_L |c_\Psi| r < 1$. This finishes the proof. ∎

Of course Proposition 7.1 is contained in Proposition 7.5 with $c_\Psi = 0$. The exponential factor, however, allows us to obtain an estimate for rather rough norms of the initial value.

7.6 Theorem *Let $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $g \in L^2_\sigma(H)$ and u be a σ-solution to $f = 0$ on $\overline{I} \times \overline{H}$ with initial value g.*

(i) If $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$, then

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(\sqrt{t_0 - s_0}, x_0)\ \|u(s_0)\|_{L^\infty(H)}$$

for any $x_0 \in \overline{H}$ and any $s_0 < t_0 \in \overline{I}$.

(ii) If $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with $k + |\alpha| > 0$, then

$$|\partial_t^k \partial_x^\alpha u(t_0, x_0)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(\sqrt{t_0 - s_0}, x_0)\, \sqrt{t_0 - s_0}\, (\sqrt{t_0 - s_0} + \sqrt{x_{0,n}})\ \|\nabla_x u(s_0)\|_{L^\infty(H)}$$

for any $x_0 \in \overline{H}$ and any $s_0 < t_0 \in \overline{I}$.

Proof: Fix $s_0 < t_0 \in \overline{I}$ and $x_0 \in \overline{H}$. For a radius $r > 0$ we can view H as a disjoint union of annular rings $B_{jr}(x_0) \smallsetminus \overline{B}_{(j-1)r}(x_0) =: R_{r_j}(x_0)$ for $j \in \mathbb{N}$. We now specify $\Psi(\cdot; c_\Psi, \varepsilon_\Psi, y_0)$ by setting $c_\Psi := -\frac{1}{r}$, choosing ε_Ψ such that $\varepsilon_\Psi < r^2$ and letting x_0 play the role of the parameter point. For $x \in R_{r_j}(x_0)$ we then have

$$\Psi(x; c_\Psi, \varepsilon_\Psi, x_0) \leq -\frac{1}{r} \frac{\tilde{d}(x,x_0)^2}{\sqrt{r^2 + \tilde{d}(x,x_0)^2}} \leq \frac{-j+3}{4c_d^2}$$

thanks to an application of the equivalence estimate 5.7. This shows that

$$\int\limits_{R_{r_j}(x_0)} e^{2\Psi(\cdot;c_\Psi,\varepsilon_\Psi,x_0)}\, u(s_0)^2\, d\mu_\sigma \lesssim e^{-\frac{1}{2c_d^2}j}\, |B_{jr}(x_0)|_\sigma\, \|u(s_0)\|^2_{L^\infty(H)}$$

$$\lesssim_{n,\sigma} j^{2n+2\sigma}\, e^{-\frac{1}{2c_d^2}j}\, |B_r(x_0)|_\sigma\, \|u(s_0)\|^2_{L^\infty(H)}$$

by Proposition 5.12 in case $n + 2\sigma \geq 0$, and with j^n instead of $j^{2n+2\sigma}$ if $n + 2\sigma < 0$. Summing this over j gives the bound

$$\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,x_0)}\, u(s_0) \right\|^2_{L^2_\sigma(H)} \lesssim_{n,\sigma} |B_r(x_0)|_\sigma\, \|u(s_0)\|^2_{L^\infty(H)},$$

where the convergent series

$$\sum_{j\in\mathbb{N}} j^{2n+2\sigma}\, e^{-\frac{1}{2c_d^2}j}$$

is subsumed into the constant.
Proposition 7.5 thus shows that

$$|\partial_t^k \partial_x^\alpha u(t_0,x_0)| \lesssim_{n,\sigma,k,\alpha} \frac{c_{k,\alpha}(\sqrt{t_0-s_0}, x_0)}{|B_{\sqrt{t_0-s_0}}(x_0)|_\sigma^{\frac{1}{2}}}\, e^{2c_L^2 \frac{t_0-s_0}{r^2}}\, |B_r(x_0)|_\sigma^{\frac{1}{2}}\, \|u(s_0)\|_{L^\infty(H)}$$

since $\Psi(x_0; c_\Psi, \varepsilon_\Psi, x_0) = 0$. Setting $r = \sqrt{t_0 - s_0}$ then finishes the proof of the first part.
For the treatment of the second part we consider the same specification of Ψ as above and let $C \in \mathbb{R}$ be an arbitrary constant. Then $u - C$ is also a σ-solution on $\overline{I} \times \overline{H}$, and Proposition 7.5 amounts to

$$|\partial_t^k \partial_x^\alpha (u(t,x) - C)|_{t=t_0, x=x_0}|$$
$$\lesssim_{n,\sigma,k,\alpha} \frac{c_{k,\alpha}(\sqrt{t_0-s_0}, x_0)}{|B_{\sqrt{t_0-s_0}}(x_0)|_\sigma^{\frac{1}{2}}}\, e^{c_L^2 c_\Psi^2 (t_0-s_0)} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,x_0)}\, (u(s_0) - C) \right\|_{L^2_\sigma(H)}.$$

For $C := u(s_0, x_0)$ we have

$$|u(s_0,x) - C| \leq |x - x_0|\, \|\nabla_x u(s_0)\|_{L^\infty(H)},$$

so we get

$$\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,x_0)}\, (u(s_0) - C) \right\|^2_{L^2_\sigma(H)} \leq \|\nabla_x u(s_0)\|^2_{L^\infty(H)} \int\limits_H e^{2\Psi(x;c_\Psi,\varepsilon_\Psi,x_0)} |x - x_0|^2\, d\mu_\sigma(x).$$

The integral is treated as above, where for the additional factor we have

$$|x - x_0| \lesssim j^2\, r\, (r + \sqrt{x_{0,n}}) \text{ on } B_{jr}(x_0)$$

by Proposition 5.8. The series in consideration still converges, and in the end we now get

$$\begin{aligned}|\partial_t^k \partial_x^\alpha (u(t,x) - u(s_0,x_0))|_{t=t_0,x=x_0}| \\ \lesssim_{n,\sigma,k,\alpha} \frac{c_{k,\alpha}(\sqrt{t_0 - s_0}, x_0)}{|B_{\sqrt{t_0-s_0}}(x_0)|^{\frac{1}{\sigma}}_{\frac{1}{2}}}\, e^{c_L^2 \frac{t_0-s_0}{r^2}}\, \|\nabla_x u(s_0)\|_{L^\infty(H)}\, r\,(r + \sqrt{x_{0,n}})\, |B_r(x_0)|^{\frac{1}{\sigma}}_{\frac{1}{2}}\end{aligned}$$

Since for $k + |\alpha| > 0$ a derivative falls onto the constant on the left hand side and causes it to vanish, we get the statement with choosing $r = \sqrt{t_0 - s_0}$. ∎

7.7 Remark *We also get an estimate involving $\|\nabla u(s_0)\|_{L^\infty(H)}$ for $k = |\alpha| = 0$: Using the triangle inequality from below in the last step of the proof above immediately delivers*

$$|u(t_0,x_0)| \lesssim_{n,\sigma,k,\alpha} \sqrt{t_0 - s_0}\,(\sqrt{t_0 - s_0} + \sqrt{x_{0,n}})\; \|\nabla_x u(s_0)\|_{L^\infty(H)} + \|u(s_0)\|_{L^\infty(H)}\,.$$

7.8 Remark *We would like to point out that we can bound both u and the gradient of u in all of $I \times H$, so that for $s_0 = t_1$ the map $g \mapsto u$ becomes a continuous curve in the solution space: The special cases $\alpha = k = 0$ and $|\alpha| = 1$, $k = 0$ in the last proposition read*

$$\|u\|_{L^\infty(I\times H)} \lesssim_{n,\sigma,k,\alpha} \|g\|_{L^\infty(H)}$$

and

$$\|\nabla_x u\|_{L^\infty(I\times H)} \lesssim_{n,\sigma,k,\alpha} \|\nabla_x g\|_{L^\infty(H)}\,.$$

Local L^p-estimates in time and space are an immediate consequence of this, as long as we stay away from the time s_0. As a notational shortcut we define cylinders

$$Q_r(s_0,x_0) := I_{r,\frac{1}{2}}(s_0) \times B_r(x_0)$$

for $s_0 \in [t_1,t_2)$, $x_0 \in \overline{H}$ and $0 < r \leq \sqrt{t_2 - s_0}$, where we always – especially for $t_2 = \infty$ – assume that $r < \infty$. Note that we still consider balls with respect to the metric d, and will stick to that, whereas the L^p-spaces will be understood with respect to the usual and unweighted Lebesgue measure in the sequel.

7.9 Corollary *Let $I = (t_1,t_2) \subset \mathbb{R}$ be an open interval, $g \in L^2_\sigma(H)$ and u be a σ-solution to $f = 0$ and g on $\overline{I} \times \overline{H}$.*

(i) If $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$, then

$$\begin{aligned}r^{2k+|\alpha|}\,(r + \sqrt{x_{0,n}})^{-2l+|\alpha|}\,|Q_r(s_0,x_0)|^{-\frac{1}{p}}\, \left\|(\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha u\right\|_{L^p(Q_r(s_0,x_0))} \\ \lesssim_{n,\sigma,k,\alpha} \|u(s_0)\|_{L^\infty(H)}\end{aligned}$$

for any $1 \leq p < \infty$ as well as any $s_0 \in \overline{I}$, $x_0 \in \overline{H}$ and $0 < r \leq \sqrt{t_2 - s_0}$.

(ii) If $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with $k + |\alpha| > 0$, then

$$r^{2k+|\alpha|-1} \, (r + \sqrt{x_{0,n}})^{-2l+|\alpha|-1} \, |Q_r(s_0, x_0)|^{-\frac{1}{p}} \, \left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha u \right\|_{L^p(Q_r(s_0,x_0))} \lesssim_{n,\sigma,k,\alpha} \|\nabla_x u(s_0)\|_{L^\infty(H)}$$

for any $1 \leq p < \infty$ as well as any $s_0 \in \overline{I}$, $x_0 \in \overline{H}$ and $0 < r \leq \sqrt{t_2 - s_0}$.

Proof: Fix $p \geq 1$, $s_0 \in I$, $x_0 \in \overline{H}$ and $0 < r \leq \sqrt{t_2 - s_0}$. Obviously we have

$$\left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha u \right\|_{L^p(Q_r(s_0,x_0))} \leq |Q_r(s_0, x_0)|^{\frac{1}{p}} \sup_{(t,x) \in Q_r(s_0,x_0)} |x_n^l \, \partial_t^k \partial_x^\alpha u(t, x)|$$

so we can apply Theorem 7.6 after multiplying both sides by x_n^l. It is clear that

$$x_n^l \leq (\sqrt{x_n} + \sqrt{t - s_0})^{2l}$$

since both l and $\sqrt{t - s_0}$ for $t > s_0$ are positive. We then get

$$x_n^l \, |\partial_t^k \partial_x^\alpha u(t, x)| \lesssim_{n,\sigma,k,\alpha} \sqrt{t - s_0}^{-2k-|\alpha|} \, (\sqrt{t - s_0} + \sqrt{x_n})^{-|\alpha|+2l} \, \|u(s_0)\|_{L^\infty(H)}$$

and

$$x_n^l \, |\partial_t^k \partial_x^\alpha u(t, x)| \lesssim_{n,\sigma,k,\alpha} \sqrt{t - s_0}^{-2k-|\alpha|+1} \, (\sqrt{t - s_0} + \sqrt{x_n})^{-|\alpha|+1+2l} \, \|\nabla_x u(s_0)\|_{L^\infty(H)} \, .$$

For $(t, x) \in Q_r(s_0, x_0)$ we have $\sqrt{t - s_0} \approx r$ since the cylinders are bounded away from time s_0 and we only have to deal with an upper bound for $\sqrt{x_n}$ here. Thanks to Proposition 5.8 we see that for any $x \in B_r(x_0)$ we have $r + \sqrt{x_n} \approx r + \sqrt{x_{0,n}}$. Collecting the constants then proves the corollary. ∎

8 Gaussian Estimate and Consequences

A Green function is understood to be a function that represents solutions to the initial value problem for the homogeneous equation as an integral with respect to the Lebesgue measure. If existence and uniqueness of solutions is known, this equivalently means that any function given as such an integral is a solution. For arbitrary equations it is not clear a priori that a Green function exists. However, the bounds we obtained so far guarantee that our linear perturbation equation does indeed possess one.

8.1 Proposition *Let $\sigma > -1$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval.*
Then the Green function on $I \times \overline{H}$ exists, that is $G_\sigma : I \times \overline{H} \times \lceil I \times \overline{H} \longrightarrow \mathbb{R}$ with $G_\sigma(t,x,s,y) = 0$ whenever $t < s \in I$, and

$$\partial_t^k \partial_x^\alpha u(t,x) = \int_H \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)\, u(s,y)\, d\mathcal{L}(y)$$

for any $t > s \in \lceil I$, $x \in \overline{H}$, any $k \in \mathbb{N}_0$, $\alpha \in \mathbb{N}_0^n$, and any σ-solution u to $f = 0$ on $\lceil I \times \overline{H}$ with initial value $u(t_1) = g \in L^2_\sigma(H)$.

Proof: Let u be the σ-solution to $f = 0$ on $\lceil I \times \overline{H}$ with initial value $g \in L^2_\sigma(H)$. Existence and uniqueness is ensured by Propositions 3.3, 3.8 and 3.9. Now fix $t > s \in \lceil I$ as well as $x \in \overline{H}$ and $k \in \mathbb{N}_0$, $\alpha \in \mathbb{N}_0^n$.
Proposition 7.1 shows that the linear functional

$$L^2_\sigma(H) \ni u(s) \mapsto \partial_t^k \partial_x^\alpha u(t,x) \in \mathbb{R}$$

is in fact continuous. By the Riesz representation theorem we can therefore find a $K^{k,\alpha}_{t,x,s} \in L^2_\sigma(H)$ with

$$\partial_t^k \partial_x^\alpha u(t,x) = \int_H K^{k,\alpha}_{t,x,s}(y)\, u(s,y)\, d\mu_\sigma(y).$$

This proves the existence of $G_\sigma(t,x,s,y) := y_n^\sigma\, K^{0,0}_{t,x,s}(y)$ for any $t > s \in \lceil I$ and almost all x, $y \in \overline{H}$ that represents a solution as desired. Moreover, since

$$\partial_t^k \partial_x^\alpha u(t,x) = \partial_t^k \partial_x^\alpha \int_H G_\sigma(t,x,s,y)\, u(s,y)\, d\mathcal{L}^n(y) = \int_H \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)\, u(s,y)\, d\mathcal{L}^n(y)$$

by Lebesgue's theorem we have $\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y) = y_n^\sigma\, K^{k,\alpha}_{t,x,s}(y)$. ∎

For a given interval I we henceforth always denote the Green function on $I \times \overline{H}$ by G_σ without mentioning the interval in the notation.

8.2 Remark *A calculation shows that $y_n^{-\sigma}\, G_\sigma(\cdot,\cdot,s,y)$ is a time-local σ-solution to $f = 0$ on $I \times \overline{H}$ for almost any $(s,y) \in \lceil I \times \overline{H}$.*

But because of the duality identity 3.10 *we also have that*

$$G_\sigma(t,x,s,y) = \left(\frac{y_n}{x_n}\right)^\sigma G_\sigma(s,y,t,x)$$

for almost all $(t,x,s,y) \in I \times \overline{H} \times \bar{I} \times \overline{H}$. *Therefore,* $(\,\cdot\,)_n^{-\sigma}\, G_\sigma(t,x,\cdot,\cdot)$ *and consequently* $(\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,\cdot,\cdot)$ *is a* σ*-solution to* $f = 0$ *on* $I \times \overline{H}$ *for almost any* $(t,x) \in I \times \overline{H}$.

Now the exponential estimates we proved before put us into the position to show that G_σ is not only in L^∞, but decays exponentially with a bound in the shape of the Gaussian function with respect to d and μ_σ.

8.3 Theorem (Gaussian Estimate) *Let* $\sigma > -1$ *and* $I = (t_1, t_2) \subset \mathbb{R}$ *be an open interval. Further let* $k \in \mathbb{N}_0$ *and* $\alpha \in \mathbb{N}_0^n$.
Then we have

$$|\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(\sqrt{t-s}, x)\, |B_{\sqrt{t-s}}(x)|_\sigma^{-\frac{1}{2}}\, |B_{\sqrt{t-s}}(y)|_\sigma^{-\frac{1}{2}}\, y_n^\sigma\, e^{-\frac{d(x,y)^2}{32 c_d^2 c_L^2 (t-s)}}$$

for all $x \neq y \in \overline{H}$ *and all* $t > s \in \bar{I}$.

Proof: Fix $t > s \in \bar{I}$ as well as $y_0 \in \overline{H}$, and denote $r := \sqrt{t-s}$. Since L^∞ is the dual space to L^1_σ we have for any $x \in \overline{H}$ that

$$\begin{aligned}
&\left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, |B_r(\cdot)|_\sigma^{\frac{1}{2}}\, (\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,\cdot) \right\|_{L^\infty(H)} \\
&= \sup_{\|\zeta\|_{L^1_\sigma(H)} \le 1} \left| \int_H e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, |B_r(\cdot)|_\sigma^{\frac{1}{2}}\, (\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,\cdot)\, \zeta\, d\mu_\sigma \right| \\
&= \sup \left\{ |\partial_t^k \partial_x^\alpha v(t,x)| \mid v \text{ is } \sigma\text{-solution, } v(s) = e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} |B_r(\cdot)|_\sigma^{\frac{1}{2}}\, \zeta \text{ and } \|\zeta\|_{L^1_\sigma(H)} \le 1 \right\}
\end{aligned}$$

thanks to the representing property of the Green function.
Applying estimate 7.5 onto such v in the points $s < \frac{t+s}{2}$ for $\alpha = 0$ and $k = 0$ shows that

$$e^{\Psi(x;c_\Psi,\varepsilon_\Psi,y_0)}\, |B_r(x)|_\sigma^{\frac{1}{2}}\, |v(\frac{t+s}{2}, x)| \lesssim_{n,\sigma} e^{c_L^2 c_\Psi^2 r^2} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(s) \right\|^2_{L^2_\sigma(H)}$$

for any $x \in \overline{H}$. We now denote the multiplication operator on $L^2_\sigma(H)$ with respect to the function $x \mapsto |B_r(x)|_\sigma^{\frac{1}{2}}$ by M. For any σ-solutions to the linear perturbation equation we furthermore define a modified solution operator $S_s(\frac{t+s}{2})$ on $L^2_\sigma(H)$, incorporating the exponential function by assigning $e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(s)$ to later time evaluations $e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(\frac{t+s}{2})$. In terms of the operators the last estimate means that

$$\left\| (M S_s(\frac{t+s}{2})) e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} v(s) \right\|_{L^\infty(H)} \lesssim_{n,\sigma} e^{c_L^2 c_\Psi^2 r^2} \left\| e^{\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)} v(s) \right\|_{L^2_\sigma(H)}$$

and therefore $M S_s(\frac{t+s}{2})$ is an operator not only from $L^2_\sigma(H)$ to $L^2_\sigma(H)$, but in fact maps into $L^\infty(H)$ and has operator norm bounded by $C\, e^{c_L^2 c_\Psi^2 r^2}$ for a constant $C = C(n,\sigma) > 0$. Consequently, since the dual operator has the same norm, for any $\xi \in L^1_\sigma(H)$ we have

$$\left\| \left(M S_s(\frac{t+s}{2}) \right)^* \xi \right\|_{L^2_\sigma(H)} \lesssim_{n,\sigma} e^{c_L^2 c_\Psi^2 r^2} \, \|\xi\|_{L^1_\sigma(H)} .$$

Now the multiplication operator is self-adjoint, and so is the solution operator by the duality equation 3.10. We then find that for the dual of S_s we have

$$S_s(\frac{t+s}{2})^* : e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(s) \mapsto e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(\frac{t+s}{2}).$$

Choosing

$$\xi = |B_r(\cdot)|_\sigma^{-\frac{1}{2}} \, e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(s) = \zeta$$

we see that

$$\left\| e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(\frac{t+s}{2}) \right\|_{L^2_\sigma(H)} \lesssim_{n,\sigma} e^{c_L^2 c_\Psi^2 r^2} \, \|\zeta\|_{L^1_\sigma(H)} \leq e^{c_L^2 c_\Psi^2 r^2}.$$

But estimate 7.5 can as well be applied to v in the points $\frac{t+s}{2} < t$ and with the function $\Psi(\cdot;-c_\Psi,\varepsilon_\Psi,y_0) = -\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)$ instead of $\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)$ to give

$$|\partial_t^k \partial_x^\alpha v(t,x)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r,x)\, |B_r(x)|_\sigma^{-\frac{1}{2}} \, e^{c_L^2 c_\Psi^2 r^2 + \Psi(x;c_\Psi,\varepsilon_\Psi,y_0)} \left\| e^{-\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)}\, v(\frac{t+s}{2}) \right\|_{L^2_\sigma(H)}$$

for any $x \in \overline{H}$. The combination of the two estimates for v then yields

$$|\partial_t^k \partial_x^\alpha v(t,x)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r,x)\, |B_r(x)|_\sigma^{-\frac{1}{2}} \, e^{2c_L^2 c_\Psi^2 r^2 + \Psi(x;c_\Psi,\varepsilon_\Psi,y_0)}.$$

This means that we have shown

$$\left\| e^{\Psi}\, |B_r(\cdot)|_\sigma^{\frac{1}{2}} \, (\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,\cdot) \right\|_{L^\infty(H)} \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r,x)\, |B_r(x)|_\sigma^{-\frac{1}{2}} \, e^{2c_L^2 c_\Psi^2 r^2 + \Psi(x;c_\Psi,\varepsilon_\Psi,y_0)}$$

for any $x \in \overline{H}$.
Consequently we get

$$|y_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r,x)\, |B_r(y)|_\sigma^{-\frac{1}{2}} \, |B_r(x)|_\sigma^{-\frac{1}{2}} \, e^{-\Psi(y;c_\Psi,\varepsilon_\Psi,y_0)} \, e^{2c_L^2 c_\Psi^2 r + \Psi(x;c_\Psi,\varepsilon_\Psi,y_0)}$$

for almost all $y \in \overline{H}$. Now fix $y = y_0$ as well as $x \neq y_0$, and specify $\Psi(\cdot;c_\Psi,\varepsilon_\Psi,y_0)$ by setting $\varepsilon_\Psi := 3\,\tilde{d}(x,y) > 0$ and $c_\Psi := -c$ for a positive, but apart from that arbitrary $c > 0$. In conjunction with Theorem 5.6 we have then shown that

$$|y_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(r,x)\, |B_r(y)|_\sigma^{-\frac{1}{2}} \, |B_r(x)|_\sigma^{-\frac{1}{2}} \, e^{2c_L^2 c^2 r^2 - \frac{c}{2c_d} d(x,y)}.$$

Finally, we optimise over the constant c. Straightforward calculation shows that

$$c \mapsto 2\, c_L^2\, c^2\, r^2 - \frac{c}{2c_d}\, d(x,y)$$

has a minimum for

$$c_* = \frac{d(x,y)}{8c_L^2 c_d r^2}.$$

Inserting this value into the inequality then gives the exponential decay estimate stated. ■

8.4 Remark *We want to emphasise that Proposition* 7.5*, derived from Proposition* 7.3*, was crucial in the proof of the Gaussian estimate. For $\alpha = k = 0$, we could have proven the same estimate also using Proposition 7.3 alongside with Remark 7.4, but by means of a Moser iteration instead of the pointwise estimate. This also works for the more general spatial part $L_\sigma u = (\,\cdot\,)_n^{-\sigma} \nabla_x \cdot ((\,\cdot\,)_n^{1+\sigma} A \nabla_x u)$ with a time-space coefficient matrix A that is uniformely strongly parabolic, measurable and bounded. See [Koc99] for a proof. Apart from loosing the direct control on the derivatives, another drawback of the Moser approach is that it cannot be generalised to higher order equations.*

8.5 Remark *The Gaussian estimate makes it possible to solve the initial value problem also for more general data not contained in $L^2_\sigma(H)$. Truncating those initial data to reach the L^2_σ-setting and gain solutions through the representation by the Green function, the exponential decay ensures in many cases - as for example for initial data in L^1_σ, L^∞ and the homogeneous Lipschitz space - that the truncated solutions converge. It thus makes sense to speak of a solution with such initial values. See also Theorem* 7.6*.*

We can replace the measure of the ball centred at y or the one centred at x by the mutually other one with Proposition 5.17, paying for it by an additional factor $(1 + \frac{d(x,y)}{\sqrt{t-s}})^{n+\sigma}$. The exponential decay in $d(x,y)$ now makes the loss in this exchange controllable. In fact, for any $a \geq 0$ we have that $e^{-a^2} \lesssim_m (1+a)^{-m}$ for any $m \geq 0$. We state the application of this trivial fact in our situation as a lemma.

8.6 Lemma *Let $c > 0$ be an arbitrary constant and $t > s$.*
We have both

$$e^{-c\frac{d(x,y)^2}{t-s}} \lesssim_{m,c} \left(1 + \frac{d(x,y)}{\sqrt{t-s}}\right)^{-m} e^{-\frac{c}{2}\frac{d(x,y)^2}{t-s}} \text{ for any } m \geq 0$$

and

$$e^{-c\frac{1}{t-s}} \lesssim_{m,c} \sqrt{t-s}^{\,m} e^{-\frac{c}{2}\frac{1}{t-s}} \text{ for any } m \geq 0.$$

8.7 Remark *The first part of Lemma* 8.6 *in conjunction with Proposition* 5.17 *now yields*

$$|B_{\sqrt{t-s}}(y)|_\sigma^{-\frac{1}{2}} e^{-\frac{d(x,y)^2}{32c_d^2c_L^2(t-s)}} \lesssim_{n,\sigma,m} \left(1 + \frac{d(x,y)}{\sqrt{t-s}}\right)^{n+\sigma} |B_{\sqrt{t-s}}(x)|_\sigma^{-\frac{1}{2}} \left(1 + \frac{d(x,y)}{\sqrt{t-s}}\right)^{-m} e^{-\frac{d(x,y)^2}{64c_d^2c_L^2(t-s)}}$$

for $n + 2\sigma \geq 0$, and similarly for $n + 2\sigma < 0$ with exponent $\frac{n}{2}$ instead of $n+\sigma$. Since $\sigma > -1$ and hence $n + \sigma > 0$, we can set $m = n+\sigma$, or $m = \frac{n}{2} > 0$ in the opposite case, and the additional factor disappears. We could of course have exchanged balls the other way around, that is we could have substituted the ball centred at x by the one centred at y. Furthermore, the same argument works for $c_{k,\alpha}(\sqrt{t-s}, x)$ by Remark 5.18*. Here we get*

$$c_{k,\alpha}(\sqrt{t-s}, x)\, e^{-\frac{d(x,y)^2}{32c_d^2c_L^2(t-s)}} \lesssim_{n,\alpha,m} \left(1 + \frac{d(x,y)}{\sqrt{t-s}}\right)^{2|\alpha|} c_{k,\alpha}(\sqrt{t-s}, y) \left(1 + \frac{d(x,y)}{\sqrt{t-s}}\right)^{-m} e^{-\frac{d(x,y)^2}{64c_d^2c_L^2(t-s)}}$$

and can chose $m = 2\,|\alpha|$.
Calculations in this fashion will be made frequently in the following. From now on we do not specify the constant in the exponent any more, but merely write C even if it changes in the course of the argument.

8.8 Remark *Note also that thanks to proposition* 5.10 *we know*

$$|B_{\sqrt{t-s}}(y)|_\sigma^{-1}\, y_n^\sigma \eqsim_{n,\sigma} |B_{\sqrt{t-s}}(y)|^{-1} \left(\frac{\sqrt{y_n}}{\sqrt{y_n}+\sqrt{t-s}}\right)^{2\sigma}$$

and can thus simplify the appereance of the Gaussian estimate if $\sigma \geq 0$: then we have

$$\left(\frac{\sqrt{y_n}}{\sqrt{y_n}+\sqrt{t-s}}\right)^{2\sigma} \leq 1.$$

We can now make use of the pointwise estimate 6.10 once more. In conjunction with Remark 8.2 this allows us to gain the following generalised version of the Gaussian estimate.

8.9 Proposition *Let $\sigma > -1$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let $k \in \mathbb{N}_0$ and $j \in \mathbb{N}_0$ as well as $\alpha \in \mathbb{N}_0^n$ and $\beta \in \mathbb{N}_0^n$.*
Then we have

$$\left|\partial_s^j \partial_y^\beta \left(y_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)\right)\right| \lesssim_{n,\sigma,k,\alpha,j,\beta} c_{j+k,\beta+\alpha}(\sqrt{t-s},x)\, |B_{\sqrt{t-s}}(y)|_\sigma^{-1}\, e^{-\frac{d(x,y)^2}{C(t-s)}}$$

for all $x \neq y \in \overline{H}$ and all $s < t \in I$, and with any possible combination of the points x and y in the factors $c_{j+k,\beta+\alpha}(\sqrt{t-s},x)\,|B_{\sqrt{t-s}}(y)|_\sigma^{-1}$.

Proof: Fix $s < t \in I$. Then there exists a constant $c > 0$ with $c^2 < \frac{s-t_1}{t-s}$. We set $r := c\,\sqrt{t-s}$ and $s_0 := (1+c^2)\,s - c^2\,t > t_1$, thus obtaining $I_r(s_0) = (s_0, s) \subset I$. Note that for any $\tau \in I_r(s_0)$ we have $r < t - \tau < (1+c^2)\,r$ and hence $t - \tau \eqsim r$.
Now fix also x, $y \in \overline{H}$. Because of Remark 8.2 and the considerations on the time interval we know that $(\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,\cdot,\cdot)$ is a σ-solution on $I_r(s_0) \times B_r(y)$. We can thus apply the local pointwise estimate 6.10 in the temporal end point s and the spatial centre point y and get

$$\begin{aligned}
&\left|\partial_s^j \partial_y^\beta \left(y_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)\right)\right| \\
&\quad\lesssim_{n,\sigma,j,\beta} c_{j,\beta}(r,y)\, r^{-1}\, |B_r(y)|_\sigma^{-\frac{1}{2}} \left(\int\limits_{I_r(s_0)} \left\|(\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,\tau,\cdot)\right\|_{L^2_\sigma(B_r(y))}^2 \, d\mathcal{L}(\tau)\right)^{\frac{1}{2}} \\
&\quad\lesssim c_{j,\beta}(\sqrt{t-s},y) \sup_{\tau \in I_r(s_0)} \left\|(\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,\tau,\cdot)\right\|_{L^\infty(B_r(y))}.
\end{aligned}$$

But by the Gaussian estimate 8.3, taking Remark 8.7 into account, we find that

$$\begin{aligned}
&\sup_{\tau \in I_r(s_0)} \left\|(\,\cdot\,)_n^{-\sigma}\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,\tau,\cdot)\right\|_{L^\infty(B_r(y))} \\
&\quad\lesssim_{n,\sigma,k,\alpha} \sup_{\tau \in I_r(s_0)} c_{k,\alpha}(\sqrt{t-\tau},x)\, |B_{\sqrt{t-\tau}}(x)|_\sigma^{-1} \left\|e^{-\frac{d^2(x,\cdot)}{C(t-\tau)}}\right\|_{L^\infty(B_r(y))} \\
&\quad\lesssim c_{k,\alpha}(\sqrt{t-s},x)\, |B_{\sqrt{t-s}}(x)|_\sigma^{-1}\, e^{-\frac{d^2(x,y)}{C(t-s)}} \left\|e^{\frac{d^2(y,\cdot)}{C(t-s)}}\right\|_{L^\infty(B_r(y))}
\end{aligned}$$

since on $I_r(s_0)$ we have $t-\tau \approx r \approx \sqrt{t-s}$, and because of the triangle inequality for d coupled with Young's inequality. It is clear that

$$\left\| e^{\frac{d^2(y,\cdot)}{C(t-s)}} \right\|_{L^\infty(B_r(y))} \leq e^{\frac{c^2}{C}}$$

and the statement follows after possibly exchanging x by y and vice versa. ∎

The Gaussian estimate has numerous useful consequences. We start with an estimate of the L^q-norm with respect to x as well as y.

8.10 Proposition *Let $\sigma > -1$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$.*

(i) We have

$$\begin{aligned} &\left\| x_n^l\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,\cdot) \right\|_{L^q(H)} \\ &\lesssim_{n,\sigma,k,\alpha,q} (\sqrt{t-s} + \sqrt{x_n})^{2l}\, c_{k,\alpha}(\sqrt{t-s}, x)\, |B_{\sqrt{t-s}}(x)|^{\frac{1}{q}-1} \end{aligned}$$

for all $t > s \in \overline{I}$ and almost all $x \in \overline{H}$, and for any q with $\sigma q > -1$.

(ii) We have

$$\begin{aligned} &\left\| (\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha G_\sigma(t,\cdot,s,y) \right\|_{L^q(H)} \\ &\lesssim_{n,\sigma,l,k,\alpha,q} (\sqrt{t-s} + \sqrt{y_n})^{2l}\, c_{k,\alpha}(\sqrt{t-s}, y)\, |B_{\sqrt{t-s}}(y)|^{\frac{1}{q}-1} \left(\frac{\sqrt{y_n}}{\sqrt{t-s} + \sqrt{y_n}} \right)^{2\sigma} \end{aligned}$$

for all $t > s \in \overline{I}$ and almost all $y \in \overline{H}$, and for any $q \geq 1$.

Proof: Fix $(t,x) \in I \times \overline{H}$ and $q \geq 1$. For any $t_1 \leq s < t$, the Gaussian estimate 8.3 with Remark 8.7, where both the derivative constant and the ball are taken with respect to x, then implies that

$$\int_H |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)|^q \, d\mathcal{L}^n(y) \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(\sqrt{t-s}, x)^q\, |B_{\sqrt{t-s}}(x)|_\sigma^{-q} \int_H y_n^{\sigma q} e^{-q\frac{d(x,y)^2}{C(t-s)}} \, d\mathcal{L}^n(y).$$

We consider the right hand side integral seperately, fix $r > 0$ and cover H with annular rings

$$B_{jr}(x) \setminus B_{(j-1)r}(x)$$

for $j \in \mathbb{N}$. Using the obvious estimate

$$e^{-q\frac{d(x,y)^2}{C(t-s)}} \leq e^{-q\frac{(j-1)^2 r^2}{C(t-s)}}$$

on any of these rings we get

$$\int_H y_n^{\sigma q}\, e^{-q\frac{d(x,y)^2}{C(t-s)}} \, d\mathcal{L}^n(y) \leq \sum_{j\in\mathbb{N}} e^{-q\frac{(j-1)^2 r^2}{C(t-s)}} \int_{B_{jr}(x)} y_n^{\sigma q} \, d\mathcal{L}(y).$$

At this point we have to assume $\sigma q > -1$ to be sure that the integral exists. Then

$$\sum_{j\in\mathbb{N}} e^{-q\frac{(j-1)^2r^2}{C(t-s)}} \, |B_{jr}(x)|_{\sigma q} \lesssim_{n,\sigma} |B_r(x)|_{\sigma q} \sum_{j\in\mathbb{N}} e^{-q\frac{(j-1)^2r^2}{C(t-s)}} \, j^{2n+2\sigma q}$$

where we used Proposition 5.12 for $n+2\,\sigma\,q \geq 0$. In case we have $n+2\,\sigma\,q < 0$ the exponent of j becomes merely n. Now choose $r=\sqrt{t-s}$ to get a convergent series in any of the two cases. We have then shown that

$$\int_H |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)|^q \, d\mathcal{L}^n(y) \lesssim_{n,\sigma,k,\alpha,q} c_{k,\alpha}(\sqrt{t-s},x)^q \, |B_{\sqrt{t-s}}(x)|_\sigma^{-q} \, |B_{\sqrt{t-s}}(x)|_{\sigma q}.$$

But Proposition 5.10 implies that

$$|B_{\sqrt{t-s}}(x)|_\sigma^{-q} \, |B_{\sqrt{t-s}}(x)|_{\sigma q} \eqsim_{n,\sigma,q} |B_{\sqrt{t-s}}(x)|^{1-q}.$$

Multiplying the equation with weights $x_n^l \leq (\sqrt{t-s}+\sqrt{x_n})^{2l}$ proves the first statement.
For the other integration we fix $(s,y) \in \overline{I} \times \overline{H}$ and first incorporate the weight x_n^l. This time the Gaussian estimate is used with y instead of x wherever it is possible, then yielding

$$\begin{aligned}
&\int_H (\,\cdot\,)_n^l \, |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)|^q \, d\mathcal{L}^n(x) \\
&\qquad \lesssim_{n,\sigma,k,\alpha} c_{k,\alpha}(\sqrt{t-s},y)^q \, |B_{\sqrt{t-s}}(y)|_\sigma^{-q} \, y_n^{\sigma q} \int_H (\sqrt{t-s}+\sqrt{x_n})^{2l} \, e^{-q\frac{d(x,y)^2}{C(t-s)}} \, d\mathcal{L}^n(x).
\end{aligned}$$

Thanks to Remark 5.18 and the first part of Lemma 8.6, under the integral we can make the x_n into a y_n on expense of a portion of the exponential decay and the dependency of the constant on l. Proceeding as before with annular rings we see that

$$\int_H e^{-q\frac{d(x,y)^2}{C(t-s)}} \, d\mathcal{L}^n(x) \lesssim_n |B_{\sqrt{t-s}}(y)|$$

without any restriction on q, where C is double the size as before. The stated estimate follows. ∎

The exponential decay has another technical consequence contained in the following lemma.

8.11 Lemma *Let $0 < \delta \leq \frac{\rho}{2} < \infty$.*
If $\theta \geq 0$ and $0 \leq c \leq 1$, then

$$\sqrt{t-s}^{\,-\theta} e^{-c\frac{d(x,y)^2}{t-s}} \lesssim_{c,\theta,\delta,\rho} e^{-\frac{c}{2}d(x,y)}$$

for all $t \in [2\,\delta,\rho]$, almost all $x \in \overline{H}$ and almost all $(s,y) \in \big([0,t)\times\overline{H}\big) \setminus \big((\delta,t)\times B_\rho(x)\big)$.

Proof: We fix δ and ρ as in the prerequisit. For $t > \delta$ we can consider the set

$$M := \big([0,t)\times\overline{H}\big) \setminus \big((\delta,t)\times B_\rho(x)\big).$$

If $(s,y) \in M$, then

$$(t-s,y) \in \big([t-\delta,t]\times B_\rho(x)\big) \cup \big([t-\delta,t]\times B_\rho(x)^c\big) \cup \big((0,t-\delta)\times B_\rho(x)^c\big),$$

where the unions are disjoint, and we will distinguish these three cases in the following.
Given $(t-s,y) \in [t-\delta,t] \times B_\rho(x)$, we get

$$\sqrt{t-s}^{-\theta}\, e^{-c\frac{d(x,y)^2}{t-s}} \leq \sqrt{t-\delta}^{-\theta} \leq \sqrt{\delta}^{-\theta}\, e^{\rho}\, e^{-d(x,y)}$$

if $t \geq 2\,\delta$, since we also know $\rho - d(x,y) \geq 0$.
In the second case we have $(t-s,y) \in [t-\delta,t] \times B_\rho(x)^c$ and therefore $\frac{d(x,y)}{t-s} \geq 1$, given that $t \leq \rho$. It follows that

$$\sqrt{t-s}^{-\theta}\, e^{-c\frac{d(x,y)^2}{(t-s)}} \leq \sqrt{\delta}^{-\theta}\, e^{-cd(x,y)},$$

where we need $t \geq 2\,\delta$ once more.
Finally, for $(t-s) \in (0,t-\delta) \times B_\rho(x)^c$, we have again that $\frac{d(x,y)}{t-s} \geq 1$ if $t \leq \rho$. This leads to

$$\sqrt{t-s}^{-\theta}\, e^{-c\frac{d(x,y)^2}{t-s}} \leq \sqrt{t-s}^{-\theta}\, e^{-\frac{c\rho^2}{2}\frac{1}{t-s}}\, e^{-\frac{c}{2}d(x,y)}$$

The second part of Lemma 8.6 gives the upper bound $C\,e^{-\frac{c}{2}d(x,y)}$ with a constant $C = C(c,\theta,\rho)$.■

If we restrict the time interval to $(0,1)$, the appearance of the Gaussian estimate can be considerably simplified. We first state this as a lemma before using it in different contexts together with the results already obtained above. From here on we often distinguish cases $z \gtrsim 1$ and $z \lesssim 1$, meaning that we fix a constant $c > 0$ and then have $z \geq c$ bounded away from zero and the complementary case $z \leq c$, without specifying the exact size of c.

8.12 Lemma *Let $I := (0,1)$, $t > s \in \bar{I}$ and x, $y \in \overline{H}$.*

(i) If $\theta \geq 0$, then

$$\left(\frac{\sqrt{t-s}}{\sqrt{t-s}+\sqrt{x_n}}\right)^{\theta} \leq (1+\sqrt{x_n})^{-\theta}$$

and the same holds true for $\sqrt{y_n}$ instead of $\sqrt{x_n}$.

(ii) If $\sigma > -1$, then

$$\left(\frac{\sqrt{y_n}}{\sqrt{t-s}+\sqrt{y_n}}\right)^{2\sigma} \lesssim (y_n^\sigma)^{\chi_{\{\sigma<0\}}\chi_{\{\sqrt{y_n}\lesssim 1\}}}.$$

(iii) If $x \neq y$ as well as $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with $2\,l - |\alpha| \leq 0$, then

$$\begin{aligned} &x_n^l\, |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \\ &\quad\lesssim_{n,\sigma,k,\alpha} \sqrt{t-s}^{2(l-k-|\alpha|-n)}\, (1+\sqrt{x_n})^{2l-|\alpha|}\, |B_1(y)|^{-1}\, (y_n^\sigma)^{\chi_{\{\sigma<0\}}\chi_{\{y_n\lesssim 1\}}}\, e^{-\frac{d(x,y)^2}{C(t-s)}} \end{aligned}$$

and the same statement holds with any possible combination of the points x and y in the factors $(1+\sqrt{x_n})^{2l-|\alpha|}\,|B_1(y)|^{-1}$.

Proof: For any $t > s \in \bar{I} = [0,1)$ we also have $\sqrt{t-s} \leq 1$. In conjunction with the monotonicity of the function $a \mapsto \frac{a}{a+c}$ for any positive value c this proves the first part.
For the second part remember that we have already seen in Remark 8.8 that the term in consideration disappears regardless of the interval I if only $\sigma \geq 0$. On $I = (0,1)$, we can also treat the

opposite case of $-1 < \sigma < 0$, since then we get

$$\left(\frac{\sqrt{y_n}}{\sqrt{y_n} + \sqrt{t-s}}\right)^{2\sigma} \leq \left(\frac{\sqrt{y_n}}{\sqrt{y_n} + 1}\right)^{2\sigma}.$$

If in addition we have that $\sqrt{y_n} \gtrsim 1$, this term is bounded by a constant again. If on the other hand we have $\sqrt{y_n} \lesssim 1$, we obtain y_n^{σ} as an upper bound.
Now, reformulating the Gaussian estimate 8.3 with the derivative constant taken with respect to x and the ball centred at y according to Remark 8.7, with additional weights on the left hand side and with the abbreviation $r = \sqrt{t-s}$, we get

$$\begin{aligned} x_n^l\, |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \\ \lesssim_{n,\sigma,k,\alpha} r^{2(l-k-|\alpha|-n)}\, e^{-\frac{d(x,y)^2}{Cr^2}} \left(\frac{r}{r+\sqrt{x_n}}\right)^{-2l+|\alpha|} \left(\frac{r}{r+\sqrt{y_n}}\right)^{n} \left(\frac{\sqrt{y_n}}{r+\sqrt{y_n}}\right)^{2\sigma}. \end{aligned}$$

By the first parts of this lemma, using also

$$\left(\frac{1}{1+\sqrt{y_n}}\right)^{n} \eqsim_n |B_1(y)|^{-1}$$

from Proposition 5.10, the statement follows with the better exponential decay.
In order to obtain the other combinations of x and y we have to apply Lemma 8.6 and Proposition 5.17 with Remark 8.7 at most twice (possibly in a different choice of x and y in the Gaussian estimate), each time giving away a portion of the decay. ∎

The last two lemmas now immediately amount to a pointwise estimate of Green's function on $(0,1) \times \overline{H}$ on a rather complicated range of values.

8.13 Proposition *Let $\sigma > -1$ and $I := (0,1)$. Futher let $0 < \delta \leq \frac{1}{2}$ and $1 \leq \rho < \infty$. If $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with $2\,l - |\alpha| \leq 0$, then*

$$x_n^l\, |\partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)| \lesssim_{n,\sigma,l,k,\alpha,\delta,\rho} (1+\sqrt{x_n})^{2l-|\alpha|}\, |B_1(y)|^{-1}\, (y_n^{\sigma})^{\chi_{\{\sigma<0\}}\chi_{\{y_n \lesssim 1\}}}\, e^{-\frac{d(x,y)}{C}}$$

for all $t \in [2\,\delta, 1]$, almost all $x \in \overline{H}$ and almost all $(s,y) \in \big([0,t) \times \overline{H}\big) \smallsetminus \big((\delta,t) \times B_\rho(x)\big)$, and the same statement holds with any possible combination of x and y in the factors $(1+\sqrt{x_n})^{2l-|\alpha|}\, |B_1(y)|^{-1}$.

Proof: Use Lemma 8.11 with $\theta = -2\,(l-k-|\alpha|-n)$ on the right hand side of the third part of Lemma 8.12, where the condition $2\,l - |\alpha| \leq 0$ coming from the latter ensures that the exponent is positiv. For $\delta \leq \frac{1}{2}$ and $1 \leq \rho < \infty$ we furthermore have both the condition $0 < \delta \leq \frac{\rho}{2} < \infty$ from Lemma 8.11 satisfied and know that $[2\,\delta, \rho] \supset [2\,\delta, 1]$, so the statement is also valid on the smaller interval that is usable in the situation $I = (0,1)$. ∎

On $I = (0,1)$ we can also derive a time-space-version of the L^q-boundedness from Proposition 8.10.

8.14 Proposition *Let $\sigma > -1$ and $I := (0,1)$. Further let $l \geq 0$ and $\alpha \in \mathbb{N}_0^n$ such that $2\,l - |\alpha| \leq 0$ and $l - |\alpha| > -1$.*

(i) We have

$$\left\| x_n^l\, \partial_x^\alpha G_\sigma(t,x,\cdot,\cdot) \right\|_{L^q((0,t)\times H)} \lesssim_{n,\sigma,k,\alpha,q} (1+\sqrt{x_n})^{2l-|\alpha|}\, |B_1(x)|^{\frac{1}{q}-1}$$

for all $t \in I$ and almost all $x \in \overline{H}$, and for any q with $1 \le q < \frac{n+1}{n+|\alpha|-l}$ and $\sigma q > -1$.

(ii) *We have*

$$\left\|(\,\cdot\,)_n^l\,\partial_x^\alpha G_\sigma(\cdot,\cdot,s,y)\right\|_{L^q((s,1)\times H)} \lesssim_{n,\sigma,l,k,\alpha,q} (1+\sqrt{y_n})^{2l-|\alpha|}\,|B_1(y)|^{\frac{1}{q}-1}\,(y_n^\sigma)^{\chi_{\{\sigma<0\}}\chi_{\{y_n\lesssim 1\}}}$$

for all $s \in \mathring{I}$ and almost all $y \in \overline{H}$, and for any q with $1 \le q < \frac{n+1}{n+|\alpha|-l}$.

Proof: By Proposition 8.10 it is immediately clear that

$$\begin{aligned}&\left\|x_n^l\,\partial_t^k\partial_x^\alpha G_\sigma(t,x,\cdot,\cdot)\right\|_{L^q((0,t)\times H)}^q\\ &\lesssim_{n,\sigma,k,\alpha,q} \int\limits_{(0,t)} (\sqrt{t-s}+\sqrt{x_n})^{2ql}\,c_{k,\alpha}(\sqrt{t-s},x)^q\,|B_{\sqrt{t-s}}(x)|^{1-q}\,d\mathcal{L}(s)\end{aligned}$$

for all $t \in (0,1]$ and almost all $x \in \overline{H}$, and for any q with $\sigma q > -1$.
Explicitely spelling out the factors using the formula 5.10 for the measure of intrinsic balls, we therefore look at the integrand

$$\sqrt{t-s}^{2q(l-k-|\alpha|-n)+2n}\left(\frac{\sqrt{t-s}}{\sqrt{t-s}+\sqrt{x_n}}\right)^{-q(2l-|\alpha|)}\left(\frac{\sqrt{t-s}}{\sqrt{t-s}+\sqrt{x_n}}\right)^{-n(1-q)}.$$

Since $I=(0,1)$ and $q \ge 1$, we can use the first part of Lemma 8.12 if $2l-|\alpha| \le 0$. In conjunction with Proposition 5.10 and a substitution in the integration variable we thus get

$$\begin{aligned}&\left\|x_n^l\,\partial_t^k\partial_x^\alpha G_\sigma(t,x,\cdot,\cdot)\right\|_{L^q((0,t)\times H)}^q\\ &\lesssim_{n,\sigma,k,\alpha,q} (1+\sqrt{x_n})^{(2l-|\alpha|)q}\,|B_1(x)|^{1-q}\int\limits_{(0,t)} s^{(l-k-|\alpha|-n)q+n}\,d\mathcal{L}(s).\end{aligned}$$

If the exponent of the integrand is greater than -1, that is for any $q < \frac{1+n}{-l+k+|\alpha|}$, this integral converges to

$$c(n,\alpha,k,l,q)\,t^{(l-k-|\alpha|-n)q+n+1}$$

which in turns is bounded for $t \in (0,1]$. It is obvious that the convergence condition can only be satisfied for a $q \ge 1$ if $-l+k+|\alpha| < 1$. In conjunction with the condition $2l-|\alpha| \le 0$ that arouse from the calculations above, it becomes clear that only $k=0$ is admissible and we therefore need to have $-l+|\alpha| < 1$.
The second part of the statment is proven identically, with the additional factor coming from Proposition 8.10 estimated by the second part of Lemma 8.12 and without the additional constraint on q, but an additional dependency of the constant on l as before. ∎

8.15 Remark *We already noted in the proof that no temporal derivative can be treated in that way and thus $k=0$ is the only possible case that can appear. It is worth pointing out that we can not achieve a bound for the gradient of G_σ in L^q without extra weights to compensate for the spatial derivative either. In fact, the condition on the exponents l and α are satisfied if and only if either $l=0$ and $\alpha=0$ or $l \in (0,\frac{1}{2}]$ and $|\alpha|=1$. We will call such exponents (l,k,α) Green-exponents.*

9 Estimates Against the Inhomogeneity

We now abandon the situation where $f = 0$ and gain estimates against the inhomogeneity.

9.1 Remark *To see that solutions with inhomogeneity can be expressed in terms of G_σ, we invoke Duhamel's principle:*
Let $\sigma > -1$, $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval and $f \in L^2(I; L^2_\sigma(H))$.
Any σ-solution u to f and $g = 0$ on $\overline{I} \times \overline{H}$ is given by

$$u(t,x) = \int\limits_{(t_1,t)\times H} G_\sigma(t,x,s,y) f(s,y)\, d\mathcal{L}^{n+1}(s,y)$$

for all $t \in I$ and almost all $x \in \overline{H}$.
Similar as for initial values in Remark 8.5, we can make sense of the term σ-solution for inhomogeneities in other spaces than $L^2(I;\, L^2_\sigma(H))$ by means of this formula as long as the integral converges. Thus, we will not specify conditions on f any more in the sequel.

Duhamel's principle also enables us to view the operator that maps inhomogeneities to solutions as integral kernel operators. Now Remark 5.16 ensures that $(I \times \overline{H}, \mathcal{L} \times \mu_\sigma, D)$ is a space of homogenous type, and thus the theory of Calderón-Zygmund can be applied in this non-euclidean setting. Therefore, the solution operators belonging to Calderón-Zygmund-exponents as defined in Remark 4.5 are of Calderón-Zygmund-type.

9.2 Proposition *Let $\sigma > -1$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ are Calderón-Zygmund-exponents, then

$$\int\limits_I \left\| (\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha u \right\|^p_{L^p_\sigma(H)} \, d\mathcal{L} \lesssim_{n,\sigma,l,k,\alpha,p} \int\limits_I \|f\|^p_{L^p_\sigma(H)} \, d\mathcal{L}$$

for any $1 < p < \infty$.

Proof: For fixed $l \geq 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ consider the operator

$$L^2(I \times H, \mathcal{L} \times \mu_\sigma) \ni f \mapsto (\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha u \in L^2(I \times H, \mathcal{L} \times \mu_\sigma).$$

Duhamel's principle 9.1 allows us to view them as kernel operators with kernels

$$K(t,x,s,y) = y_n^{-\sigma}\, x_n^l\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)$$

that are continuous where $(t,x) \neq (s,y)$. Here we suppress the parameters l, k, α in the notation of the kernels.
We want to show that the Calderón-Zygmund cancellation conditions A.34 hold. To this end define

$$V := |B^D_{D((t,x),(s,y))}(t,x)|_{0\times\sigma} + |B^D_{D((t,x),(s,y))}(s,y)|_{0\times\sigma}$$

where the balls B^D are understood with respect to the time-space-metric D and

$$| \cdot |_{0\times\sigma} := (\mathcal{L} \times \mu_\sigma)(\,\cdot\,).$$

Note that then

$$D((t,x),(s,y))^{-2} \left(|B_{D((t,x),(s,y))}(x)|_\sigma + |B_{D((t,x),(s,y))}(y)|_\sigma \right)^{-1} \lesssim V^{-1}.$$

Therefore, for the first cancellation condition to hold it is enough to show that

$$|K(t,x,s,y)| \lesssim D((t,x),(s,y))^{-2} \left(|B_{D((t,x),(s,y))}(x)|_\sigma + |B_{D((t,x),(s,y))}(y)|_\sigma \right)^{-1}$$

for any $s < t \in I$ and almost any $x \neq y \in \overline{H}$. The Gaussian estimate 8.9 reveals that

$$|K(t,x,s,y)| \lesssim_{n,\sigma,k,\alpha} x_n^l \, c_{k,\alpha}(\sqrt{t-s}, x) \, |B_{\sqrt{t-s}}(x)|_\sigma^{-1} \, e^{-\frac{d(x,y)^2}{C(t-s)}}.$$

By an exchange of the ball centre as before we see that

$$|B_{\sqrt{t-s}}(x)|_\sigma^{-1} \, e^{-\frac{d(x,y)^2}{C(t-s)}} \lesssim_{n,\sigma} \left(|B_{\sqrt{t-s}}(x)|_\sigma + |B_{\sqrt{t-s}}(y)|_\sigma \right)^{-1} e^{-\frac{d(x,y)^2}{C(t-s)}}$$

where we needed that $n + \sigma > 0$ and doubled the constant C.
The doubling condition in Proposition 5.12 together with the exponential decay then enables us to change also the radius of the balls, that is we get

$$\begin{aligned} &\left(|B_{\sqrt{t-s}}(x)|_\sigma + |B_{\sqrt{t-s}}(y)|_\sigma \right)^{-1} e^{-\frac{d(x,y)^2}{C(t-s)}} \\ &\qquad \lesssim_{n,\sigma} \left(|B_{\sqrt{(t-s)+d(x,y)^2}}(x)|_\sigma + |B_{\sqrt{(t-s)+d(x,y)^2}}(y)|_\sigma \right)^{-1} e^{-\frac{d(x,y)^2}{C(t-s)}} \end{aligned}$$

where again the C on the right hand side is twice the one on the left hand side.
On the other hand we have that

$$x_n^l \, c_{k,\alpha}(\sqrt{t-s}, x) \lesssim \sqrt{t-s}^{\,2l-2k-2|\alpha|}$$

if $2\,l - |\alpha| \leq 0$. Thanks to the exponential decay we can modify this to obtain

$$\sqrt{t-s}^{\,2l-2k-2|\alpha|} \, e^{-\frac{d(x,y)^2}{C(t-s)}} \lesssim_{n,\sigma,l,k,\alpha} \sqrt{(t-s)+d(x,y)^2}^{\,2(l-k-|\alpha|)}.$$

For $l - k - |\alpha| = -1$, these estimates exactly imply the condition we wanted to show.
For the second cancellation condition we consider (t,x), (s,y), $(\bar{t},\bar{x})$, $(\bar{s},\bar{y})$ that satisfy the condition

$$(*) \qquad \frac{D((t,x),(\bar{t},\bar{x})) + D((s,y),(\bar{s},\bar{y}))}{D((t,x),(s,y)) + D((\bar{t},\bar{x}),(\bar{s},\bar{y}))} \leq \varepsilon$$

for an $\varepsilon \in (0,1)$. For those points, the triangle inequality then ensures that

$$\frac{1-\varepsilon}{1+\varepsilon} D((t,x),(s,y)) \leq D((\bar{t},\bar{x}),(\bar{s},\bar{y})) \leq \frac{1+\varepsilon}{1-\varepsilon} D((t,x),(s,y))$$

and hence both

$$D((t,x),(\bar{t},\bar{x})) + D((s,y),(\bar{s},\bar{y})) \leq \frac{2\varepsilon}{1-\varepsilon} D((t,x),(s,y))$$

and

$$D((t,x),(\bar{t},\bar{x})) + D((s,y),(\bar{s},\bar{y})) \leq \frac{2\varepsilon}{1-\varepsilon} D((\bar{t},\bar{x}),(\bar{s},\bar{y}))$$

if $\varepsilon < \frac{1}{3}$. From these inequalities it follows among others that

$$\begin{aligned}
\frac{1-3\varepsilon}{1-\varepsilon} D((t,x),(s,y)) &\leq D((\bar{t},\bar{x}),(s,y)) \leq \frac{1+\varepsilon}{1-\varepsilon} D((t,x),(s,y)) \\
\frac{1-3\varepsilon}{1-\varepsilon} D((t,x),(s,y)) &\leq D((\bar{t},x),(s,y)) \leq \frac{1+\varepsilon}{1-\varepsilon} D((t,x),(s,y)) \\
\frac{1-3\varepsilon}{1-\varepsilon} D((\bar{t},\bar{x}),(\bar{s},\bar{y})) &\leq D((\bar{t},\bar{x}),(\bar{s},y)) \leq \frac{1+\varepsilon}{1-\varepsilon} D((\bar{t},\bar{x}),(\bar{s},\bar{y})) \\
\frac{1-3\varepsilon}{1-\varepsilon} D((t,x),(s,y)) &\leq D((\bar{t},x),(\bar{s},y)) \leq \frac{1+3\varepsilon}{1-\varepsilon} D((t,x),(s,y)).
\end{aligned}$$

We now fix $\varepsilon = \frac{1}{6}$ and estimate

$$\begin{aligned}
|K(t,x,s,y) - K(\bar{t},\bar{x},\bar{s},\bar{y})| \\
&\leq |K(t,x,s,y) - K(\bar{t},x,s,y)| + |K(\bar{t},x,s,y) - K(\bar{t},\bar{x},s,y)| \\
&+ |K(\bar{t},\bar{x},s,y) - K(\bar{t},\bar{x},\bar{s},y)| + |K(\bar{t},\bar{x},\bar{s},y) - K(\bar{t},\bar{x},\bar{s},\bar{y})| \\
&=: (I) + (II) + (III) + (IV).
\end{aligned}$$

For $t < s$ and $\bar{t} < \bar{s}$, the left hand side is zero and we have nothing to show. Similarly, the cases $t < s$ and $\bar{t} < s$ can be excluded in the consideration of (I). There it remains to find a cancellation bound if either s is between t and $\bar{t}$ or smaller than both t and $\bar{t}$. The fundamental theorem of calculus infers that

$$(I) \leq |t - \bar{t}| \sup_{\tau} |\partial_\tau K(\tau, x, s, y)|$$

where the supremum is taken over all τ between t and $\bar{t}$. If s is between t and $\bar{t}$, it can happen that $\tau < s$. In this case $K(\tau, x, s, y) = 0$. So in any case the Gaussian estimate 8.9 shows that

$$(I) \lesssim_{n,\sigma,k,\alpha} |t - \bar{t}| \sup_{\tau} x_n^l \, c_{k+1,\alpha}(\sqrt{|\tau - s|}, x) \, |B_{\sqrt{|\tau-s|}}(x)|_\sigma^{-1} e^{-\frac{d(x,y)^2}{C|\tau-s|}}.$$

As above the exponential decay enables us to add a ball centred at y, increase the radius of the balls and manipulate the appearance of the preceding factors if $2l - |\alpha| \leq 1$ to get

$$(I) \lesssim_{n,\sigma,l,k,\alpha} |t - \bar{t}| \sup_{\tau} D((\tau,x),(s,y))^{2l-2k-2-2|\alpha|} \left(|B_{D((\tau,x),(s,y))}(x)|_\sigma + |B_{D((\tau,x),(s,y))}(y)|_\sigma\right)^{-1}$$

with the supremum again ranging over any τ between t and $\bar{t}$. Depending on the relation of t and $\bar{t}$, an upper bound of this is given either by replacing any τ by t or $\bar{t}$. In the latter case the condition $(*)$ makes it possible to exchange $\bar{t}$ by t again. All together we then find

$$\begin{aligned}(I) &\lesssim_{n,\sigma,l,k,\alpha} |t-\bar{t}|\, D((t,x),(s,y))^{2l-2k-2-2|\alpha|} \left(|B_{D((t,x),(s,y))}(x)|_\sigma + |B_{D((t,x),(s,y))}(y)|_\sigma\right)^{-1} \\ &\lesssim \frac{\sqrt{|t-\bar{t}|}}{D((t,x),(s,y)) + D((t,x),(s,y))} \frac{\sqrt{|t-\bar{t}|}}{D((t,x),(s,y))} D((t,x),(s,y))^{2l-2k+2-2|\alpha|}\, V^{-1} \\ &\lesssim \frac{\sqrt{|t-\bar{t}|}}{D((t,x),(s,y)) + D((\bar{t},\bar{x}),(\bar{s},\bar{y}))} D((t,x),(s,y))^{2l-2k+2-2|\alpha|}\, V^{-1}\end{aligned}$$

by virtue of the above consequences of $(*)$.
For (II) let $\Gamma : [a,b] \to H$ be the arc-length parametrised geodesic that connects x and $\bar{x}$. Then the chain rule and the fundamental theorem of calculus show that

$$(II) = \Big| \int\limits_{(a,b)} \nabla_{\Gamma(\tau)} K(\bar{t},\Gamma(\tau),s,y) \cdot \partial_\tau \Gamma(\tau)\, d\mathcal{L}(\tau) \Big| \leq d(x,\bar{x}) \sup_{z \in B_{d(x,\bar{x})}(x)} z_n^{\frac{1}{2}}\, |\nabla_z K(\bar{t},z,s,y)|$$

with $z = \Gamma(\tau)$ thanks to the unit speed property of Γ and the geodesic character of d, compare Chapter 5. The Gaussian estimate 8.9 then implies that

$$(II) \lesssim_{n,\sigma,k,\alpha} d(x,\bar{x}) \sup_{z \in B_{d(x,\bar{x})}(x)} \left(z_n^{l+\frac{1}{2}}\, c_{k,\alpha+1}(\sqrt{\bar{t}-s},z) + z_n^{l-\frac{1}{2}}\, c_{k,\alpha}(\sqrt{\bar{t}-s},z)\right) |B_{\sqrt{\bar{t}-s}}(y)|_\sigma^{-1}\, e^{-\frac{d(z,y)^2}{C(\bar{t}-s)}}$$

with the obvious meaning of the symbolic notation $c_{k,\alpha+1}$ and for $l > 0$ as well as $\bar{t} > s$, the left hand side being 0 for $\bar{t} < s$. If we have $l = 0$ the second summand on the right hand side does not turn up and the following calculations simplify accordingly. Given $l > 0$, we need indeed $l \geq \frac{1}{2}$ and $2\,l - |\alpha| \leq 0$ to reach

$$(II) \lesssim_{n,\sigma,l,k,\alpha} d(x,\bar{x}) \sup_{z \in B_{d(x,\bar{x})}(x)} \sqrt{\bar{t}-s}^{\,2l-2k-2|\alpha|-1} |B_{\sqrt{\bar{t}-s}}(y)|_\sigma^{-1} e^{-\frac{d(z,y)^2}{C(\bar{t}-s)}}.$$

Now for any $z \in B_{d(x,\bar{x})}(x)$ we have

$$\begin{aligned}d(x,y) &\leq d(x,\bar{x}) + d(z,y) \leq D((t,x),(\bar{t},\bar{x})) + d(z,y) \\ &\leq \frac{2\varepsilon}{1-\varepsilon} D((t,x),(s,y)) + d(z,y) \leq \frac{2\varepsilon}{1-3\varepsilon} D((\bar{t},x),(s,y)) + d(z,y) \\ &\leq \frac{2\varepsilon}{1-3\varepsilon} (\sqrt{\bar{t}-s} + d(x,y)) + d(z,y)\end{aligned}$$

by condition $(*)$ and hence

$$-d(z,y)^2 \lesssim -d(x,y)^2 + (\bar{t}-s)$$

for any $z \in B_{d(x,\bar{x})}(x)$ if $(*)$ is satisfied for an $\varepsilon < \frac{1}{5}$. This is obviously the case for our choice of $\varepsilon = \frac{1}{6}$. As a consequence we can gain an exponential factor $e^{-\frac{d(x,y)^2}{C(\bar{t}-s)}}$ in the above estimate and perform the same operation as above to reach

$$\begin{aligned}(II) &\lesssim_{n,\sigma,l,k,\alpha} d(x,\bar{x})\, D(\bar{t},x,s,y)^{2l-2k-2|\alpha|-1} \left(|B_{D((\bar{t},x),(s,y))}(x)|_\sigma + |B_{D((\bar{t},x),(s,y))}(y)|_\sigma\right)^{-1} \\ &\lesssim \frac{d(x,\bar{x})}{D(t,x,s,y) + D(\bar{t},\bar{x},\bar{s},\bar{y})} D(t,x,s,y)^{2l-2k-2|\alpha|+2}\, V^{-1}\end{aligned}$$

after using the consequences of $(*)$ again.
As for (I), for the next term we get

$$(III) \lesssim |s-\bar{s}| \sup_{\tau} \sqrt{|\bar{t}-\tau|}^{2l-2k-2-2|\alpha|} |B_{\sqrt{|\bar{t}-\tau|}}(y)|_{\sigma}^{-1} e^{-\frac{d(\bar{x},y)^2}{C|\bar{t}-\tau|}}$$

where the supremum is taken over all τ between s and $\bar{s}$. Condition $(*)$ implies both

$$d(x,y) \leq \frac{1-\varepsilon}{1-3\varepsilon} D((\bar{t},\bar{x}),(s,y))$$

and

$$d(x,y) \leq \frac{1+\varepsilon}{1-3\varepsilon} D((\bar{t},\bar{x}),(\bar{s},y)).$$

As a consequence, for any τ between s and $\bar{s}$ we obtain

$$-d(\bar{x},y)^2 \lesssim |\bar{t}-\tau| - d(x,y)^2.$$

It follows that

$$(III) \lesssim_{n,\sigma,l,k,\alpha} \frac{\sqrt{|s-\bar{s}|}}{D(t,x,s,y)+D(\bar{t},\bar{x},\bar{s},\bar{y})} D(t,x,s,y)^{2l-2k-2|\alpha|+2} V^{-1}.$$

Finally, with $M(y,\bar{y}) := B_{d(y,\bar{y})}(y) \cap B_{d(y,\bar{y})}(\bar{y})$ we find

$$\begin{aligned}(IV) &\lesssim_{n,\sigma,l,k,\alpha} d(y,\bar{y}) \sup_{z\in M(y,\bar{y})} \bar{x}_n^l\, c_{k,\alpha}(\sqrt{\bar{t}-\bar{s}},\bar{x})\, z_n^{\frac{1}{2}}\, c_{0,1}(\sqrt{\bar{t}-\bar{s}},z)\, |B_{\sqrt{\bar{t}-\bar{s}}}(z)|_{\sigma}^{-1} e^{-\frac{d(\bar{x},z)^2}{C(\bar{t}-\bar{s})}} \\ &\leq d(y,\bar{y}) \sup_{z\in M(y,\bar{y})} \sqrt{\bar{t}-\bar{s}}^{2l-2k-2|\alpha|-1} |B_{\sqrt{\bar{t}-\bar{s}}}(z)|_{\sigma}^{-1} e^{-\frac{d(\bar{x},z)^2}{C(\bar{t}-\bar{s})}}.\end{aligned}$$

By condition $(*)$, for all $z \in B_{d(y,\bar{y})}(\bar{y})$ we have

$$-d(\bar{x},z)^2 \lesssim (\bar{t}-\bar{s}) - d(x,z)^2$$

which makes it possible to get a ball centred at x in the inequality. Once again $(*)$ also ensures that

$$-d(x,z)^2 \lesssim (\bar{t}-\bar{s}) - d(x,y)^2$$

for any $z \in B_{d(y,\bar{y})}(y)$. This leads to

$$\begin{aligned}(IV) &\lesssim_{n,\sigma,l,k,\alpha} d(y,\bar{y})\, D((\bar{t},x),(\bar{s},y))^{2l-2k-2|\alpha|-1} \left(|B_{D((\bar{t},x),(\bar{s},y))}(x)|_{\sigma} + |B_{D((\bar{t},x),(\bar{s},y))}(y)|_{\sigma}\right)^{-1} \\ &\lesssim \frac{d(y,\bar{y})}{D((t,x),(s,y))+D((\bar{t},\bar{x}),(\bar{s},\bar{y}))} D((t,x),(s,y)^{2l-2k-2|\alpha|+2} V^{-1}\end{aligned}$$

by virtue of yet another application of the consequences from $(*)$.
The kernels in consideration have therefore been proven to satisfy the Calderón-Zygmund cancellation conditions on $(I\times\overline{H}, D, \mathcal{L}\times\mu_{\sigma})$ if and only if $l-k-|\alpha| = -1$ and $2l-|\alpha| \leq 0$, that is if (l,k,α) are Calderón-Zygmund-exponents in the sense of Remark 4.5. As noted there, these are exactly the exponents for which the operators are bounded on $L^2(I\times H, \mathcal{L}\times\mu_{\sigma})$. It follows that those operators are of Calderón-Zygmund type according to Definition A.34 and

hence extend to bounded operators on $L^p(I \times H, \mathcal{L} \times \mu_\sigma)$ for any $1 < p < \infty$ by Proposition A.36. ∎

The theory of Muckenhoupt weights allows us to formulate the L^p-result in unweighted spaces. Once we get rid of the weights, we can also take derivatives into consideration.

9.3 Corollary *Let $\sigma > -1$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$, $k \in \mathbb{N}_0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l, k and $\alpha - \gamma$ are Calderón-Zygmund-exponents, then

$$\left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p} \left\| D_x^{|\gamma|} f \right\|_{L^p(I \times H)}$$

for any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$.

Proof: We know that $x_n^{\rho - \sigma}$ is a p-Muckenhoupt weight with respect to μ_σ if and only if

$$-1 < \rho < p\,(\sigma + 1) - 1$$

by Proposition 5.15. By Proposition A.37, any Calderón-Zygmund-operator with respect to μ_σ also extends to all spaces $L^p(x_n^{\rho-\sigma} d\mu_\sigma) = L^p(\mu_\rho)$, and the same is true for the corresponding time-space setting by Remark 5.16. The statement for $\gamma = 0$ follows with $\rho = 0$, which is always admissible if $\sigma \geq 0$, whereas for negative σ we need the extra condition $p > \frac{1}{1+\sigma}$. Note that this does not constitute a restriction for p in case $\sigma \geq 0$.
Given this and using that $\nabla_x' u$ is a σ-solution to $\nabla_x' f$, we have also shown that

$$\int_I \left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha \nabla_x' u \right\|_{L^p(H)}^p \, d\mathcal{L} \lesssim_{n,\sigma,l,k,\alpha} \int_I \left\| \nabla_x' f \right\|_{L^p(H)}^p \, d\mathcal{L}$$

for any exponents l, k and α that are of Calderón-Zygmund-type, that is especially for $l = 0$, $k = 0$ and $|\alpha| = 1$. This yields the estimate

$$\int_I \left\| \nabla_x \nabla_x' u \right\|_{L^p(H)}^p \, d\mathcal{L} \lesssim_{n,\sigma,l,k,\alpha} \int_I \left\| \nabla_x' f \right\|_{L^p(H)}^p \, d\mathcal{L}.$$

For the vertical direction, however, we can use that $\partial_{x_n} u$ is a $(1+\sigma)$-solution to

$$\partial_{x_n} f + (1+\sigma)\, \Delta_x' u$$

by Proposition 4.9. The first part of the proof then shows that

$$\int_I \left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha \partial_{x_n} u \right\|_{L^p(H)}^p \, d\mathcal{L} \lesssim_{n,\sigma,l,k,\alpha} \int_I \left\| \partial_{x_n} f \right\|_{L^p(H)}^p \, d\mathcal{L} + \int_I \left\| \Delta_x' u \right\|_{L^p(H)}^p \, d\mathcal{L}$$

which is bounded by $\int_I \left\| \nabla_x f \right\|_{L^p(H)}^p \, d\mathcal{L}$ because of the previous considerations. An induction yields the statement. ∎

At least on the time interval $(0,1)$, thanks to Proposition 8.14 we can gain similar estimates for a different set of exponents, namely Green-exponents as defined in Remark 8.15. Note that we need $\sigma \geq 0$ in the proof of the following result and also use the unweighted Calderón-Zygmund-estimate 9.3.

9.4 Proposition *Let $\sigma \geq 0$ and $I := (0,1)$. Further let u be a σ-solution to f on $\bar{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l and $\alpha - \gamma$ are Green-exponents, then

$$\left\| (\,\cdot\,)_n^l \, \partial_x^\alpha u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p} \left\| D_x^{|\gamma|} f \right\|_{L^p(I \times H)}$$

for any $1 \leq p \leq \infty$.

Proof: By Duhamel's principle 9.1 we can consider the kernels $K(t,x,s,y) = x_n^l \, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)$ for the unweighted L^2-operators that send f to derivatives of σ-solutions $x_n^l \, \partial_t^k \partial_x^\alpha u$ to f. Since $\sigma \geq 0$, the L^1-norms of these kernels taken with respect to (t,x) as well as the ones taken with respect to (s,y) are bounded whenever $k = 0$ and l, α are Green-exponents in the sense of Remark 8.15, as an application of Proposition 8.14 for $q = 1$ shows. Thus Schur's lemma A.20 can be used to conclude that

$$\left\| (\,\cdot\,)_n^l \, \partial_x^\alpha u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha} \|f\|_{L^p(I \times H)}$$

for any $1 \leq p \leq \infty$ and any Green-exponents l and α.
Now for $j = 1, \dots, n-1$ also $\partial_{x_j} u$ is a σ-solution to $\partial_{x_j} f$ and the result just obtained can be applied to get

$$\left\| (\,\cdot\,)_n^l \, \partial_x^\alpha \partial_{x_j} u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha} \left\| \partial_{x_j} f \right\|_{L^p(I \times H)}$$

for any Green-exponents l and α.
As before, for the vertical direction we use that $\partial_{x_n} u$ is a $(1+\sigma)$-solution to $\partial_{x_n} f + (1+\sigma)\Delta'_x u$. Thus

$$\left\| (\,\cdot\,)_n^l \, \partial_x^\alpha \partial_{x_n} u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha} \|\partial_{x_n} f\|_{L^p(I \times H)} + \left\| \Delta'_x u \right\|_{L^p(I \times H)},$$

and the last summand is bounded by $\|\nabla_x f\|_{L^p(I \times H)}$ thanks to the unweighted Calderón-Zygmund-estimate from Corollary 9.3. We have therefore shown for Green-exponents l and α and for any $1 \leq p \leq \infty$ that

$$\left\| (\,\cdot\,)_n^l \, \partial_x^\alpha \partial_x^\gamma u \right\|_{L^p(I \times H)} \lesssim_{n,\sigma,l,k,\alpha,p} \left\| D_x^{|\gamma|} f \right\|_{L^p(I \times H)}$$

with $|\gamma| = 1$. The statement now follows by induction. ∎

In passing we note that Proposition 8.14 also implies the following pointwise estimate, this time with the restriction of positive σ removed again.

9.5 Proposition *Let $\sigma > -1$ and $I := (0,1)$. Further let u be a σ-solution to f on $\bar{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l and $\alpha - \gamma$ are Green-exponents, then for almost all $x \in \overline{H}$ we have

$$x_n^l \, |\partial_x^\alpha u(1,x)| \lesssim_{n,\sigma,l,k,\alpha,\gamma,p} (1 + \sqrt{x_n})^{2l - |\alpha|} \, |B_1(x)|^{-\frac{1}{p}} \left\| D_x^{|\gamma|} f \right\|_{L^p(I \times H)}$$

for any $\max\left\{ \frac{n+1}{l - |\alpha| + |\gamma| + 1},\ 1,\ \frac{1}{1+\sigma} \right\} < p < \infty$.

Proof: We use Duhamel's principle to get

$$x_n^l\,|\partial_t^k\partial_x^\alpha u(1,x)| \leq \int\limits_{(0,1)\times H} x_n^l\,|\partial_t\partial_x^\alpha G_\sigma(1,x,s,y)|\,|f(s,y)|\,d\mathcal{L}^{n+1}(s,y)$$
$$\leq \left\|x_n^l\partial_t^k\partial_x^\alpha G(1,x,\cdot,\cdot)\right\|_{L^{\frac{p}{p-1}}((0,1)\times H)}\ \|f\|_{L^p((0,1)\times H)}$$

with Hölder's inequality for any $1\leq p\leq\infty$. But for $k=0$ and Green-exponents l and α, the first factor can be treated with Proposition 8.14. This immediately yields the statement for $\gamma=0$ if both $1\leq\frac{p}{p-1}<\frac{n+1}{n-l+|\alpha|}$ and $\sigma\frac{p}{p-1}>-1$, and these conditions on p are equivalent to

$$\max\left\{\frac{n+1}{l-|\alpha|+1},\,1,\,\frac{1}{1+\sigma}\right\}<p<\infty.$$

As before we now use that $\partial_{x_j}u$ is a σ-solution to $\partial_{x_j}f$ for $j=1,\ldots,n-1$ to get

$$x_n^l\,|\partial_x^\alpha\partial_{x_j}u(1,x)| \lesssim_{n,\sigma,l,k,\alpha,p} (1+\sqrt{x_n})^{2l-|\alpha|}\,|B_1(x)|^{-\frac{1}{p}}\,\left\|\partial_{x_j}f\right\|_{L^p(I\times H)},$$

and that $\partial_{x_n}u$ is a $(1+\sigma)$-solution to $\partial_{x_n}f+(1+\sigma)\,\Delta'_x u$ to get

$$x_n^l\,|\partial_x^\alpha\partial_{x_n}u(1,x)| \lesssim_{n,\sigma,l,k,\alpha,p} (1+\sqrt{x_n})^{2l-|\alpha|}\,|B_1(x)|^{-\frac{1}{p}}\left(\|\partial_{x_n}f\|_{L^p(I\times H)}+\left\|\Delta'_x u\right\|_{L^p(I\times H)}\right).$$

Again we can apply the unweighted Calderón-Zygmund-estimate from Corollary 9.3 to bound the norm of the Laplacian. This proves that for Green-exponents l and α and for

$$\max\left\{\frac{n+1}{l-|\alpha|+|\gamma|+1},\,1,\,\frac{1}{1+\sigma}\right\}<p<\infty$$

we have

$$x_n^l\,|\partial_x^\alpha\partial_x^\gamma u(1,x)| \lesssim_{n,\sigma,l,k,\alpha,p} |B_1(x)|^{-\frac{1}{p}}\,\left\|D_x^{|\gamma|}f\right\|_{L^p(I\times H)}$$

if $|\gamma|=1$. The statement for all γ follows by induction. ■

In the next step we would like to localise the Calderón-Zygmund-estimate onto time-space-cylinders bounded away from the initial time as before in the context of L^p-estimates against initial data, compare Chapter 7. To this end we first consider only inhomogeneities f that are supported on cylinders and therefore in fact look at points on the diagonal of $(I\times\overline{H})\times(I\times\overline{H})$. There we get a local version of the L^p-estimate of Corollary 9.3 that can be improved by virtue of Proposition 9.4.
We use the abbreviation

$$Q_r(x_0):=Q_r(0,x_0)$$

for time-space-cylinders bounded away from $t=0$ and denote increased cylinders by

$$\widetilde{Q}_r(x_0):=(\frac{r^2}{4},r^2)\times B_{2r}(x_0).$$

9.6 Proposition *Let $\sigma > -1$, $I := (0,1)$, $x_0 \in \overline{H}$ and $f : I \times \overline{H} \longrightarrow \mathbb{R}$ with* $\operatorname{supp} f \subset \overline{\widetilde{Q}_1(x_0)}$. *Further let u be a σ-solution to f on $\bar{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$, $k \in \mathbb{N}_0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l, k and $\alpha - \gamma$ are Calderón-Zygmund-exponents, then

$$(1+\sqrt{x_{0,n}})^{-2l+|\alpha|-|\gamma|}\,|Q_1(x_0)|^{-\frac{1}{p}}\,\left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u\right\|_{L^p(Q_1(x_0))}$$
$$\lesssim_{n,\sigma,l,k,\alpha,\gamma,p,\theta}\sup_{\substack{0<R\leq 1\\ z\in H}} R^\theta\,|Q_R(z)|^{-\frac{1}{p}}\,\left\|D_x^{|\gamma|}f\right\|_{L^p(Q_R(z))}$$

for any $\theta \in \mathbb{R}$ and any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$.

Proof: The Calderón-Zygmund-estimate 9.3 immediately yields

$$\left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u\right\|_{L^p(Q_1(x_0))} \leq \left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u\right\|_{L^p(I\times H)} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p} \left\|D_x^{|\gamma|}f\right\|_{L^p(I\times H)} = \left\|D_x^{|\gamma|}f\right\|_{L^p(\widetilde{Q}_1(x_0))}$$

for Calderón-Zygmund-exponents l, k and $\alpha-\gamma$ and any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$.
Consider now radii $R_1 := \frac{1}{\sqrt{2}}$, $R_2 := 1$ and $R_3 := \frac{\sqrt{3}}{2}$. Independent of the location of x_0 in $\overline{H}$ we can find N_1 points $\{z_{j_1}\}$, N_2 points $\{z_{j_2}\}$ and N_3 points $\{z_{j_3}\}$, all of them contained in $\widetilde{Q}_1(x_0)$ and with N_i only depending on n for $i = 1, 2, 3$, such that the collection $\{Q_{R_i}(z_{j_i}) \mid j_i = 1,\ldots,N_i,\ i = 1, 2, 3\}$ covers $\widetilde{Q}_1(x_0)$. It follows that for $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$ we have

$$\left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u\right\|_{L^p(Q_1(x_0))} \lesssim_{n,\sigma,l,k,\alpha,p} \sum_{i=1}^{3}\sum_{j_i=1}^{N_i}\left\|D_x^{|\gamma|}f\right\|_{L^p(Q_{R_i}(z_{j_i}))}$$
$$\lesssim_\theta \sum_{i=1}^{3}\sum_{j_i=1}^{N_i}|Q_{R_i}(z_{j_i})|^{\frac{1}{p}}\,R_i^\theta\,|Q_{R_i}(z_{j_i})|^{-\frac{1}{p}}\,\left\|D_x^{|\gamma|}f\right\|_{L^p(Q_{R_i}(z_{j_i}))}$$
$$\lesssim_{n,p} |Q_1(x_0)|^{\frac{1}{p}}\sup_{\substack{0<R\leq 1\\ z\in H}} R^\theta\,|Q_R(z)|^{-\frac{1}{p}}\,\left\|D_x^{|\gamma|}f\right\|_{L^p(Q_R(z))},$$

where we first used the doubling property from Proposition 5.10 to shrink balls into $Q_1(x_0)$ and then took the supremum in any summand in the last step. This is the localisation of the Calderón-Zygmund-estimate.
Now recall that for any Calderón-Zygmund-exponents $(l, k, \alpha-\gamma)$ we have that $-2\,l + |\alpha| - |\gamma| \geq 0$. If we consider an x_0 with $\sqrt{x_{0,n}} \lesssim 1$, this means that

$$1 \leq (1+\sqrt{x_{0,n}})^{-2l+|\alpha|-|\gamma|} \lesssim_{l,\alpha,\gamma} 1$$

and the inequlity we showed is equivalent to the statement of this proposition. Note that the only Calderón-Zygmund-exponents for which $-2\,l + |\alpha| - |\gamma|$ is non-zero are the ones that satisfy $l = 0$ and $|\alpha| - |\gamma| = 1$.
For the converse case of $\sqrt{x_{0,n}} \gtrsim 1$ we first consider $\sigma \geq 0$ only. If l and $\alpha-\gamma$ are Green-exponents, then the global estimate from Proposition 9.4 can be localised with the same arguments as above to show for the same range of p that

$$\left\|(\,\cdot\,)_n^l\,\partial_x^\alpha u\right\|_{L^p(Q_1(x_0))} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p,\theta} |Q_1(x_0)|^{\frac{1}{p}}\sup_{\substack{0<R\leq 1\\ z\in H}} R^\theta\,|Q_R(z)|^{-\frac{1}{p}}\,\left\|D_x^{|\gamma|}f\right\|_{L^p(Q_R(z))}.$$

However, for any $x \in B_1(x_0)$ we have that $x_n \eqsim x_{0,n}$ if $\sqrt{x_{0,n}} \gtrsim 1$ as in our present case. Together with $l = \frac{1}{2}$, any α and γ such that $|\alpha| - |\gamma| = 1$ are admissible Green-exponents, and so we have

$$\sqrt{x_{0,n}}\ \|\partial_x^\alpha u\|_{L^p(Q_1(x_0))} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p,\theta} |Q_1(x_0)|^{\frac{1}{p}} \sup_{\substack{0<R\leq 1\\ z\in H}} R^\theta\, |Q_R(z)|^{-\frac{1}{p}}\ \left\| D_x^{|\gamma|} f \right\|_{L^p(Q_R(z))}$$

if $|\alpha| - |\gamma| = 1$ and $\sqrt{x_{0,n}} \gtrsim 1$ for $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$.
On the other hand, $l = 0$, $k = 0$ and α, γ with $|\alpha| - |\gamma| = 1$ are Calderón-Zygmund-exponents and we have therefore shown above – regardless of the position of x_0 – that for $|\alpha| - |\gamma| = 1$ we also have

$$\|\partial_x^\alpha u\|_{L^p(Q_1(x_0))} \lesssim_{n,\sigma,l,k,\alpha,\gamma,p,\theta} |Q_1(x_0)|^{\frac{1}{p}} \sup_{\substack{0<R^2\leq 1\\ z\in H}} R^\theta\, |Q_R(z)|^{-\frac{1}{p}}\ \left\| D_x^{|\gamma|} f \right\|_{L^p(Q_R(z))}.$$

This implies

$$\begin{aligned}(1+\sqrt{x_{0,n}})^{-2l+|\alpha|-|\gamma|}\ &\left\| (\,\cdot\,)_n^l\, \partial_t^k \partial_x^\alpha u \right\|_{L^p(Q_1(x_0))}\\ &\lesssim_{n,\sigma,l,k,\alpha,\gamma,p,\theta} |Q_1(x_0)|^{\frac{1}{p}} \sup_{\substack{0<R^2\leq 1\\ z\in H}} R^\theta\, |Q_R(z)|^{-\frac{1}{p}}\ \left\| D_x^{|\gamma|} f \right\|_{L^p(Q_R(z))}\end{aligned}$$

for $l = k = 0$ and $|\alpha| - |\gamma| = 1$ also in case of $\sqrt{x_{0,n}} \gtrsim 1$. Since for any other possible Calderón-Zygmund-exponent (l, k, α) we have that $-2\,l + |\alpha| - |\gamma| = 0$ this proves the statement for $\sigma \geq 0$.
We now consider $-1 < \sigma < 0$ and $\sqrt{x_{0,n}} \gtrsim 1$, the last remaining case. Let first $\gamma = 0$. As in the first step of the proof of Proposition 9.4 we consider the L^2-operators that send f to derivatives of σ-solutions $x_n^l\, \partial_t^k \partial_x^\alpha u$ to f, this time with kernels $K(t,x,s,y) = x_n^l\, \partial_t^k \partial_x^\alpha G_\sigma(t,x,s,y)\, \chi_{\operatorname{supp} f}(s,y)$. Since $\sqrt{x_{0,n}} \gtrsim 1$ we have $\operatorname{supp} f \subset \{y \mid y_n \gtrsim 1\}$ and thus Proposition 8.14 shows also in this case where σ is not in the good range that the L^1-norms of the kernels are bounded and therefore Schur's lemma A.20 is applicable for any set of Green-exponents l and α. Consequently, we get

$$\left\| (\,\cdot\,)_n^l\, \partial_x^\alpha u \right\|_{L^p(I\times H)} \lesssim_{n,\sigma,l,k,\alpha} \|f\|_{L^p(I\times H)}$$

for $-1 < \sigma < 0$, if l and α are Green-exponents and $\operatorname{supp} f \subset \widetilde{Q}_1(x_0)$ for x_0 with $\sqrt{x_{0,n}} \gtrsim 1$. For any tangential derivative these prerequisits are satisfied and we can iterate the argument as before. For a similar treatment of the vertical derivative, however, we lack the guarantee that the corresponding inhomogeneity $\partial_{x_n} f + (1+\sigma)\,\Delta'_x u$ has support in the right region. But $\partial_{x_n} u$ is a $(1+\sigma)$-solution to this right hand side, and since $(1+\sigma) \geq 0$ we apply Proposition 9.4 on this solution as above. We can therefore use an induction to show that

$$\left\| (\,\cdot\,)_n^l\, \partial_x^\alpha u \right\|_{L^p(I\times H)} \lesssim_{n,\sigma,l,k,\alpha,\gamma} \left\| D_x^{|\gamma|} f \right\|_{L^p(I\times H)}$$

for $-1 < \sigma < 0$, if l and α are Green-exponents and $\operatorname{supp} f \subset \widetilde{Q}_1(x_0)$ for x_0 with $\sqrt{x_{0,n}} \gtrsim 1$. We can procede by localising and applying this result in the same fashion as before to finish the proof. ∎

For inhomogeneities that are supported away from the diagonal we can achieve a pointwise estimate against a similar expression as before, albeit with the additional condition that the range of the parameter that was called θ above is not entirely arbitrary any more.

9.7 Proposition *Let $\sigma > -1$ and $I := (0,1)$. Fix $x \in \overline{H}$ as well as $0 < \delta \le \frac{1}{2}$ and let $f : I \times \overline{H} \longrightarrow \mathbb{R}$ with* $\operatorname{supp} f \subset \big([0,1] \times \overline{H}\big) \setminus ((\delta,1) \times B_1(x))$. *Further let u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \ge 0$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ with $2\,l - |\alpha| \le 0$, then for all $t \in [2\,\delta, 1]$ we have

$$(1+\sqrt{x_n})^{-2l+|\alpha|-\varepsilon_1}\, x_n^l\, |\partial_t^k \partial_x^\alpha u(t,x)| \lesssim_{n,\sigma,l,k,\alpha,p,\delta,\varepsilon_2} \sup_{\substack{0<R\le 1\\ z\in \overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1}\, R^{2-\varepsilon_2}\, |Q_R(z)|^{-\frac{1}{p}}\, \|f\|_{L^p(Q_R(z))}$$

for any $\varepsilon_1 \ge 0$, $\varepsilon_2 > 0$ and any $\frac{1}{1+\sigma} < p < \infty$.

Proof: We first consider the case $\sigma \ge 0$.
Let $M := \big((0,t) \times \overline{H}\big) \setminus ((\delta,t) \times B_1(x))$. By Duhamel's principle 9.1 and the prerequisit on the support of f we get

$$x_n^l\, |\partial_t^k \partial_x^\alpha u(t,x)| \le \int_M x_n^l\, |\partial_t \partial_x^\alpha G_\sigma(t,x,s,y)|\, |f(s,y)|\, d\mathcal{L}^{n+1}(s,y)$$

where G_σ is the Green function on $(0,1) \times \overline{H}$. This is only possible for $\delta \le t \le 1$. For $2\,\delta \le t \le 1$, the application of Proposition 8.13 under the integral is then justified if $2\,l - |\alpha| \le 0$, and gives us the upper bound

$$(1+\sqrt{x_n})^{2l-|\alpha|+\varepsilon_1} \int_{(0,1)\times H} (1+\sqrt{y_n})^{-\varepsilon_1}\, |B_1(y)|^{-1}\, e^{-\frac{d(x,y)}{C}}\, |f(s,y)|\, d\mathcal{L}^{n+1}(s,y)$$

for an arbitrary $\varepsilon_1 \ge 0$ and with a constant that depends on n, σ, l, k, α and δ. Note that we have here exchanged x_n by y_n in a portion of the factor $(1+\sqrt{x_n})^{2l-\alpha}$ as we can do thanks to the exponential decay.
For the computation of the integral we first cover $\overline{H}$ with a countable number of balls $B_1(y_0)$. Together with the triangle inequality we then get

$$\begin{aligned}
&\int_{(0,1)\times H} (1+\sqrt{y_n})^{-\varepsilon_1}\, |B_1(y)|^{-1}\, e^{-\frac{d(x,y)}{C}}\, |f(s,y)|\, d\mathcal{L}^{n+1}(s,y) \\
&\quad \le \sum_{y_0} \int_{(0,1)\times B_1(y_0)} (1+\sqrt{y_n})^{-\varepsilon_1}\, |B_1(y)|^{-1}\, e^{-\frac{d(x,y)}{C}}\, |f(s,y)|\, d\mathcal{L}^{n+1}(s,y) \\
&\quad \le \sup_{y_0} \int_{(0,1)\times B_1(y_0)} (1+\sqrt{y_n})^{-\varepsilon_1}\, |B_1(y)|^{-1}\, |f(s,y)|\, d\mathcal{L}^{n+1}(s,y) \sum_{y_0} e^{-\frac{d(x,y_0)}{C}}
\end{aligned}$$

and the series in the back converges uniformely in x.
We now cover the temporal interval $(0,1)$ by $(\frac{1}{2}\,R_m^2, R_m^2)$, $m \in N_0$, where $R_m := 2^{-\frac{m}{2}}$. Furthermore, for any $m \in \mathbb{N}_0$ and any y_0 we cover $B_1(y_0)$ with $N(m)$ balls $B_{R_m}(z_i)$, $z_i \in B_1(y_0)$. This is possible by Vitali's covering lemma ([Koc04]), which also ensures that

$$\sum_{i=1}^{N(m)} |B_{R_m}(z_i)|_\sigma \lesssim_{n,\sigma} |B_1(y_0)|_\sigma$$

independent of m. Note that $(2^{-m-1}, 2^{-m}) \times B_{R_m}(z_i) = Q_{R_m}(z_i)$ and we therefore now look at the expression

$$\sup_{y_0} \sum_{m\in\mathbb{N}_0} \sum_{i=1}^{N(m)} \int_{Q_{R_m}(z_i)} (1+\sqrt{y_n})^{-\varepsilon_1} \, |B_1(y)|^{-1} \, |f(s,y)| \, d\mathcal{L}^{n+1}(s,y).$$

Consider the cases $y_{0,n} \lesssim 1$, that is the situation close to the boundary of H, and $y_{0,n} \gtrsim 1$ separately, starting with the latter one away from the boundary.
If the distance of y_0 to the boundary is sufficiently big, we know that $|B_1(y)| \eqsim_n |B_1(y_0)|$ for any $y \in B_1(y_0)$ from Remark 5.11 and thus also $(1+\sqrt{y_n})^{-\varepsilon_1} \eqsim_n (R_m + \sqrt{z_{i,n}})^{-\varepsilon_1}$ for any $y \in B_{R_m}(z_i)$ with $z_i \in B_1(y_0)$. An application of Hölder's inequality then shows that

$$\begin{aligned}
&\sup_{y_{0,n}\gtrsim 1} |B_1(y_0)|^{-1} \sum_{m\in\mathbb{N}_0} \sum_{i=1}^{N(m)} (R_m+\sqrt{z_{i,n}})^{-\varepsilon_1} \int_{Q_{R_m}(z_i)} |f(s,y)| \, d\mathcal{L}^{n+1}(s,y) \\
&\leq \sup_{y_{0,n}\gtrsim 1} |B_1(y_0)|^{-1} \sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2} \sum_{i=1}^{N(m)} (R_m+\sqrt{z_{i,n}})^{-\varepsilon_1} \, |B_{R_m}(z_i)| \, R_m^{2-\varepsilon_2} \, |Q_{R_m}(z_i)|^{-\frac{1}{p}} \, \|f\|_{L^p(Q_{R_m}(z_i))} .
\end{aligned}$$

Taking the supremum over $0 < R \leq 1$ and $z \in \overline{H}$ then leads to the upper bound

$$\begin{aligned}
&\sup_{y_{0,n}\gtrsim 1} |B_1(y_0)|^{-1} \sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2} \sum_{i=1}^{N(m)} |B_{R_m}(z_i)| \sup_{\substack{0<R\leq 1\\ z\in\overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1} \, R^{2-\varepsilon_2} \, |Q_R(z)|^{-\frac{1}{p}} \, \|f\|_{L^p(Q_R(z))} \\
&\qquad \leq \sup_{\substack{0<R\leq 1\\ z\in\overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1} \, R^{2-\varepsilon_2} \, |Q_R(z)|^{-\frac{1}{p}} \, \|f\|_{L^p(Q_R(z))} \sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2}
\end{aligned}$$

by virtue of the Vitali-property discussed above. For any $\varepsilon_2 > 0$ also this series converges and can therefore be subsumed into the constant.
In the case that y_0 is close to the boundary we vary the arguments slightly. First we use the rather rough estimate $(1+\sqrt{y_n})^{-\varepsilon_1} \, |B_1(y)|^{-1} \leq 1$. The same reasoning as above then asserts that

$$\begin{aligned}
&\sup_{y_{0,n}\lesssim 1} \sum_{m\in\mathbb{N}_0} \sum_{i=1}^{N(m)} \int_{Q_{R_m}(z_i)} |f(s,y)| \, d\mathcal{L}^{n+1}(s,y) \\
&\qquad \leq \sup_{y_{0,n}\lesssim 1} |B_1(y_0)| \sup_{\substack{0<R\leq 1\\ z\in\overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1} \, R^{2-\varepsilon_2} \, |Q_R(z)|^{-\frac{1}{p}} \, \|f\|_{L^p(Q_R(z))} \sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2}
\end{aligned}$$

where we used that $\varepsilon_1 \geq 0$ and

$$R_m + \sqrt{z_{i,n}} \leq 1 + \sqrt{z_{i,n}} \lesssim_n (1+d(z_i,y_0))^2 \, (1+\sqrt{y_{0,n}}) \lesssim_n 1$$

thanks to Remark 5.18 and the fact that $z_i \in B_1(y_0)$ for $\sqrt{y_{0,n}} \lesssim 1$.
But for any $y_{0,n} \lesssim 1$ it is also clear that $|B_1(y_0)| \lesssim_n 1$ and thus the same upper bound as above follows for any $\varepsilon_2 > 0$.
It remains to consider the case $-1 < \sigma < 0$ that generates an additional factor $(y_n^\sigma)^{\chi_{\{y_n\lesssim 1\}}}$. Whenever $y_n \lesssim 1$ – a possibility that only realises itself if $y_{0,n} \lesssim 1$ – we therefore have to deal with an additional y_n^σ under the integral. The application of Hölder's inequality then results in

$$\sup_{y_{0,n}\lesssim 1} \sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2} \sum_{i=1}^{N(m)} (R_m+\sqrt{z_{i,n}})^{-\varepsilon_1} \, R_m^{2(1-\frac{1}{p})-\varepsilon_2} \, |B_{R_m}(z_i)|_{\sigma\frac{p}{p-1}}^{1-\frac{1}{p}} \, \|f\|_{L^p(Q_{R_m}(z_i))}$$

if $\sigma\,(1-\frac{1}{p}) > -1$. The last condition is equivalent to $p > \frac{1}{1+\sigma}$ for $-1 < \sigma < 0$, and does not pose a restriction on p if $\sigma \geq 0$. Now Proposition 5.10 asserts that

$$|B_{R_m}(z_i)|^{1-\frac{1}{p}}_{\sigma\frac{p}{p-1}} \eqsim_{n,\sigma} |B_{R_m}(z_i)|^{1-\frac{1}{p}}\,(R_m+\sqrt{z_{i,n}})^{2\sigma} \eqsim_{n,\sigma} |B_{R_m}(z_i)|_\sigma\,|B_{R_m}(z_i)|^{-\frac{1}{p}}$$

and a reiteration of the steps above, this time using Vitali's lemma for the measure μ_σ with σ not neccessarily 0, implies the upper bound

$$\sup_{y_{0,n}\lesssim 1} |B_1(y_0)|_\sigma \sup_{\substack{0<R\leq 1\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon_1}\,R^{2-\varepsilon_2}\,|Q_R(z)|^{-\frac{1}{p}}\;\|f\|_{L^p(Q_R(z))}\sum_{m\in\mathbb{N}_0} R_m^{\varepsilon_2}.$$

Similar as before we can conclude that $|B_1(y_0)|_\sigma \lesssim_{n,\sigma} 1$ if $y_{0,n} \lesssim 1$. ∎

Together with Proposition 9.5 this yields a pointwise estimate for any inhomogeneity on $I=(0,1)$ that can be rescaled onto arbitrary intervals $I=(t_1,t_2)$ by the invariant scaling T_λ of the equation, see Remark 3.2. We first formulate the scaling of the right hand side in a separate lemma.

9.8 Lemma *Let $\lambda > 0$, $s_1 \in \mathbb{R}$ and $T_\lambda : (\hat{s},\hat{y}) \mapsto (s_1+\lambda\,\hat{s}, \lambda\,\hat{y}) := (s,y)$. Further let $\gamma \in \mathbb{N}_0^n$. We then have*

$$\sup_{\substack{0<R\leq 1\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon}\,R^\theta\,|Q_R(z)|^{-\frac{1}{p}}\;\left\|D_{\hat{y}}^{|\gamma|}(f\circ T_\lambda)\right\|_{L^p(Q_R(z))} \lesssim_{n,p}$$

$$\lambda^{|\gamma|+\frac{\varepsilon}{2}-\frac{\theta}{2}} \sup_{\substack{0<R^2\leq\lambda\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon}\,R^\theta\,|Q_R(s_1,z)|^{-\frac{1}{p}}\;\left\|D_y^{|\gamma|}f\right\|_{L^p(Q_R(s_1,z))}$$

for any $\varepsilon \geq 0$, $\theta \in \mathbb{R}$ and any $1 \leq p \leq \infty$.

Proof: The transformation rule for integrals and the chain rule assert that

$$\left\|D_{\hat{y}}^{|\gamma|}(f\circ T_\lambda)\right\|_{L^p(Q_R(z))} = \lambda^{-\frac{n+1}{p}+|\gamma|}\;\left\|D_y^{|\gamma|}f\right\|_{L^p(T_\lambda(Q_R(z)))}.$$

According to Lemma 6.6 we have

$$T_\lambda(Q_R(z)) \subset \left(s_1+\frac{(\sqrt{\lambda}R)^2}{2}, s_1+(\sqrt{\lambda}R)^2\right) \times B_{4c_d^2\sqrt{\lambda}R}(\lambda z).$$

We can now find N points $\lambda\,z_i \in B_{4c_d^2\sqrt{\lambda}R}(\lambda z)$ such that $\bigcup_{i=1}^N B_{\sqrt{\lambda}R}(\lambda z_i)$ cover $B_{4c_d^2\sqrt{\lambda}R}(\lambda z)$. Note that N only dependes on the dimension n. Hence we get

$$T_\lambda(Q_R(z)) \subset \bigcup_{i=1}^N Q_{\sqrt{\lambda}R}(s_1,\lambda z_i)$$

and therefore

$$\left\|D_{\hat{y}}^{|\gamma|}(f\circ T_\lambda)\right\|_{L^p(Q_R(z))} \lesssim_p \lambda^{-\frac{n+1}{p}+|\gamma|}\sum_{i=1}^N \left\|D_y^{|\gamma|}f\right\|_{L^p(Q_{\sqrt{\lambda}R}(s_1,\lambda z_i))}$$

for $\lambda\,z_i \in B_{4c_d^2\sqrt{\lambda}R}(\lambda z)$.

Considering that

$$|Q_{\sqrt{\lambda}R}(s_1, \lambda z)| \eqsim_n \lambda^{1+n} |Q_R(z)|$$

by Proposition 5.10, the last line leads to

$$\sup_{\substack{0<R\leq 1\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon} R^\theta |Q_R(z)|^{-\frac{1}{p}} \left\| D_{\tilde{y}}^{|\gamma|}(f\circ T_\lambda)\right\|_{L^p(Q_R(z))} \lesssim_{n,p}$$

$$\lambda^{-\frac{n+1}{p}+|\gamma|+\frac{\varepsilon}{2}-\frac{\theta}{2}+\frac{n+1}{p}} \sum_{i=1}^N \sup_{\substack{0<R\leq 1\\ z\in H}} (\sqrt{\lambda}R+\sqrt{\lambda z_n})^{-\varepsilon} (\sqrt{\lambda}R)^\theta |Q_{\sqrt{\lambda}R}(s_1,\lambda z)|^{-\frac{1}{p}} \left\| D_y^{|\gamma|} f\right\|_{L^p(Q_{\sqrt{\lambda}R}(s_1,\lambda z_i))}.$$

However, due to the triangle inequality for d and the doubling property of this distance function we get

$$|B_{\sqrt{\lambda}R}(\lambda z)|^{-\frac{1}{p}} \lesssim_{n,p} |B_{\sqrt{\lambda}R}(\lambda z_i)|^{-\frac{1}{p}}$$

for any $\lambda z_i \in B_{4c_d^2\sqrt{\lambda}R}(\lambda z)$. But since we are taking the supremum over all $z \in H$, each summand can be considered independently and the statement follows with an additional dependency on N or rather n. ∎

In the following we still allow $t_2 = \infty$ and use the convention that $r^2 \leq t_2 - t_1$ means $r^2 < \infty$ in this case.

9.9 Theorem *Let $\sigma > -1$, $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l and $\alpha - \gamma$ are Green-exponents, then

$$\sup_{(t,x)\in I\times H} \sqrt{t-t_1}^{|\alpha|+\varepsilon_1-2|\gamma|-2+\theta} (\sqrt{t-t_1}+\sqrt{x_n})^{-2l+|\alpha|-\varepsilon_1} x_n^l |\partial_x^\alpha u(t,x)|$$

$$\lesssim_{n,\sigma,l,k,\alpha,p,p_0,\theta,\gamma,\varepsilon_1} \sup_{\substack{0<R^2\leq t_2-t_1\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon_1} R^{\theta-2|\gamma|+\varepsilon_1} |Q_R(t_1,z)|^{-\frac{1}{p}} \|f\|_{L^p(Q_R(t_1,z))}$$

$$+ \sup_{\substack{0<R^2\leq t_2-t_1\\ z\in H}} R^\theta |Q_R(t_1,z)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f\right\|_{L^{p_0}(Q_R(t_1,z))}$$

for any $\varepsilon_1 \geq 0$, $\theta < 2+2|\gamma|-\varepsilon_1$ and $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$, $\max\left\{\frac{n+1}{l-|\alpha|+|\gamma|+1}, 1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$.

Proof: Let first $t_1 = 0$ and $t_2 = 1$ so that u is a σ-solution to f on $[0,1) \times \overline{H}$ with initial value 0. Fix an arbitrary $x \in \overline{H}$ and set $f_1 := \chi_{Q_1(x)} f$, $f_2 := (1-\chi_{Q_1(x)}) f$. Then $\operatorname{supp} f_1 \subset \overline{Q_1(x)}$, $\operatorname{supp} f_2 \subset \left([0,1]\times\overline{H}\right) \setminus Q_1(x)$, and denoting the σ-solutions to f_1 and f_2 on $[0,1)\times\overline{H}$ with vanishing initial value by u_1 and u_2, respectively, we also get $u = u_1 + u_2$. With Proposition 9.5 and Proposition 9.7, the latter for $\delta = \frac{1}{2}$, we then find

$$x_n^l |\partial_x^\alpha u(1,x)| \leq x_n^l |\partial_x^\alpha u_1(1,x)| + x_n^l |\partial_x^\alpha u_2(1,x)|$$

$$\lesssim_{n,\sigma,l,k,\alpha,\gamma,p_0,p,\varepsilon_2} (1+\sqrt{x_n})^{2l-|\alpha|} |B_1(x)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f_1\right\|_{L^{p_0}((0,1)\times H)}$$

$$+ (1+\sqrt{x_n})^{2l-|\alpha|+\varepsilon_1} \sup_{\substack{0<R\leq 1\\ z\in H}} (R+\sqrt{z_n})^{-\varepsilon_1} R^{2-\varepsilon_2} |Q_R(z)|^{-\frac{1}{p}} \|f_2\|_{L^p(Q_R(z))}$$

for any $\varepsilon_1 \geq 0$, $\varepsilon_2 > 0$ and any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$, $\max\left\{\frac{n+1}{l-|\alpha|+|\gamma|+1}, 1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$, if l and $\alpha - \gamma$ are Green-exponents. We use that $\operatorname{supp} f_1 \subset \overline{Q_1(x)}$ and multiply the first summand with $1 = 1^\theta$ before taking the supremum over $0 < R \leq 1$ and $z \in \overline{H}$ to see that

$$
\begin{aligned}
&(1+\sqrt{x_n})^{-2l+|\alpha|-\varepsilon_1}\, x_n^l\, |\partial_x^\alpha u(1,x)| \\
&\qquad \lesssim_{n,\sigma,l,k,\alpha,p_0,p,\varepsilon_2} \sup_{\substack{0<R\leq 1\\ z\in \overline{H}}} R^\theta\, |Q_R(z)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f \right\|_{L^{p_0}(Q_R(z))} \\
&\qquad\qquad + \sup_{\substack{0<R\leq 1\\ z\in \overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1}\, R^{2-\varepsilon_2}\, |Q_R(z)|^{-\frac{1}{p}}\, \|f\|_{L^p(Q_R(z))}
\end{aligned}
$$

for any $\varepsilon_1 \geq 0$, $\varepsilon_2 > 0$, $\theta \in \mathbb{R}$ and $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$, $\max\left\{\frac{n+1}{l-|\alpha|+|\gamma|+1}, 1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$. Turning to the rescalation of this estimate, we let now u be a σ-solution to f on $(t_1, t_2) \times \overline{H}$ for arbitrary $t_1 < t_2$ with initial value 0. Then u is also a σ-solution to f on $(t_1, t) \times \overline{H}$ for any $t \in (t_1, t_2]$ with vanishing initial value. Denoting the time-shifted scaling function by $T_\lambda : (\hat{t}, \hat{x}) \mapsto (t_1 + \lambda\, \hat{t}, \lambda\, \hat{x}) = (t, x)$ as above, for any $(t, x) \in (t_1, t_2) \times \overline{H}$ we get

$$
\begin{aligned}
&\left(\sqrt{t-t_1}+\sqrt{x_n}\right)^{-2l+|\alpha|-\varepsilon_1} x_n^l\, |\partial_x^\alpha u(t,x)| \\
&\qquad = \lambda^{-\frac{|\alpha|}{2}-\frac{\varepsilon_1}{2}} \left(\sqrt{\frac{t-t_1}{\lambda}}+\sqrt{\frac{x_n}{\lambda}}\right)^{-2l+|\alpha|-\varepsilon_1} \left(\frac{x_n}{\lambda}\right)^l \left|\partial_{\hat{x}}^\alpha (u\circ T_\lambda)\left(\frac{t-t_1}{\lambda}, \frac{1}{\lambda}x\right)\right|.
\end{aligned}
$$

Now fix a $t \in (t_1, t_2)$ and choose $\lambda = t - t_1$. The time-shifted scaling in consideration is invariant for our equation, that is we have that $u \circ T_{t-t_1}$ is a σ-solution to $(t - t_1)\,(f \circ T_{t-t_1})$ on $(0, 1) \times \overline{H}$ with initial value 0 (see Remark 3.2), and we can then apply the above estimate on it to get the inequalities

$$
\begin{aligned}
&\left(\sqrt{t-t_1}+\sqrt{x_n}\right)^{-2l+|\alpha|-\varepsilon_1} x_n^l\, |\partial_x^\alpha u(t,x)| \\
&\lesssim_{n,\sigma,l,k,\alpha,\gamma,p_0,p,\varepsilon_2} (t-t_1)^{-\frac{|\alpha|}{2}-\frac{\varepsilon_1}{2}+1} \Bigg(\sup_{\substack{0<R\leq 1\\ z\in \overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1}\, R^{2-\varepsilon_2}\, |Q_R(z)|^{-\frac{1}{p}}\, \|f\circ T_{t-t_1}\|_{L^p(Q_R(z))} \\
&\qquad + \sup_{\substack{0<R\leq 1\\ z\in \overline{H}}} R^\theta\, |Q_R(z)|^{-\frac{1}{p_0}} \left\| D_{\hat{x}}^{|\gamma|} (f\circ T_{t-t_1}) \right\|_{L^{p_0}(Q_R(z))} \Bigg) \\
&\lesssim_{n,p,p_0} (t-t_1)^{-\frac{|\alpha|}{2}-\frac{\varepsilon_1}{2}+1} \Bigg((t-t_1)^{\frac{\varepsilon_1}{2}-1+\frac{\varepsilon_2}{2}} \sup_{\substack{0<R^2\leq t-t_1\\ z\in \overline{H}}} (R+\sqrt{z_n})^{-\varepsilon_1}\, R^{2-\varepsilon_2}\, |Q_R(t_1,z)|^{-\frac{1}{p}}\, \|f\|_{L^p(Q_R(t_1,z))} \\
&\qquad + (t-t_1)^{|\gamma|-\frac{\theta}{2}} \sup_{\substack{0<R^2\leq t-t_1\\ z\in \overline{H}}} R^\theta\, |Q_R(t_1,z)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f \right\|_{L^{p_0}(Q_R(t_1,z))} \Bigg)
\end{aligned}
$$

with Lemma 9.8. We then set $\varepsilon_2 := -\varepsilon_1 - \theta + 2 + 2\,|\gamma|$, hereby restricting ourselves to

$$\theta < 2 + 2\,|\gamma| - \varepsilon_1.$$

■

We have now finally reached the localised L^p-estimate we started out for earlier.

9.10 Theorem *Let $\sigma > -1$, $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value $g = 0$.*
If $l \geq 0$, $k \in N_0$ and $\gamma \leq \alpha \in \mathbb{N}_0^n$ are such that l, k and $\alpha - \gamma$ are Calderón-Zygmund-exponents, then

$$\sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in \overline{H}}} R^{2k+|\alpha|-|\gamma|-2+\theta} \left(R+\sqrt{z_n}\right)^{-2l+|\alpha|-|\gamma|} |Q_R(t_1,z)|^{-\frac{1}{p_0}} \left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha u \right\|_{L^{p_0}(Q_R(t_1,z))}$$
$$\lesssim_{n,\sigma,l,k,\alpha,\gamma,p,p_0,\theta} \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in \overline{H}}} \left(R+\sqrt{z_n}\right)^{-|\gamma|} R^{\theta-|\gamma|} |Q_R(t_1,z)|^{-\frac{1}{p}} \|f\|_{L^p(Q_R(t_1,z))}$$
$$+ \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in \overline{H}}} R^{\theta} |Q_R(t_1,z)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f \right\|_{L^{p_0}(Q_R(t_1,z))}$$

for any $\theta < 2 + |\gamma|$ and any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p, p_0 < \infty$.

Proof: Let first be $I = (0,1)$ and fix an arbitrary $x_0 \in \overline{H}$. We consider

$$f = f\,\chi_{\widetilde{Q}_1(x_0)} + f\,(1-\chi_{\widetilde{Q}_1(x_0)}) =: f_1 + f_2.$$

Note that then

$$\operatorname{supp} f_1 \subset \overline{\widetilde{Q}_1(x_0)} \text{ and } \operatorname{supp} f_2 \subset \left([0,1]\times \overline{H}\right) \setminus \left((\frac{1}{4},1) \times B_2(x_0) \right).$$

This splits u into a sum of u_1 and u_2, σ-solution to f_1 and f_2, respectively.
For the on-diagonal part we use Proposition 9.6 and immediately get for almost any $x_0 \in \overline{H}$ that

$$\left(1+\sqrt{x_{0,n}}\right)^{-2l+|\alpha|-|\gamma|} |Q_1(x_0)|^{-\frac{1}{p_0}} \left\| (\,\cdot\,)_n^l \, \partial_t^k \partial_x^\alpha u_1 \right\|_{L^{p_0}(Q_1(x_0))}$$
$$\lesssim_{n,\sigma,l,k,\alpha,\gamma,p_0,\theta} \sup_{\substack{0<R\leq 1 \\ z\in \overline{H}}} R^{\theta} |Q_R(z)|^{-\frac{1}{p_0}} \left\| D_x^{|\gamma|} f_1 \right\|_{L^{p_0}(Q_R(z))}$$

for any $\theta \in \mathbb{R}$ and any $\max\left\{1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$, if $(l, k, \alpha-\gamma)$ are Calderón-Zygmund-exponents.
Off-diagonal, on the other hand, we first note that by the triangle inequality we have $B_1(x) \subset B_2(x_0)$ for any $x \in \overline{B}_1(x_0)$. For any such x and any $t \in [\frac{1}{2}, 1]$, that is for $(t,x) \in \overline{Q_1(x_0)}$ it is therefore clear that

$$\operatorname{supp} f_2 \subset \left([0,1]\times \overline{H}\right) \setminus \left((\frac{1}{4},1) \times B_1(x) \right)$$

and thus for $2l - |\alpha| \leq 0$ the requirements needed for 9.7 are met with $\delta = \frac{1}{4}$. Adding the calculation

$$\left(1+\sqrt{x_n}\right)^{2l-|\alpha|+\varepsilon_1} \lesssim_{l,\alpha,\varepsilon_1} \left(1+\sqrt{x_{0,n}}\right)^{2l-|\alpha|+\varepsilon_1}$$

that is made possible by Remark 5.18 in these circumstances, for almost any $x_0 \in \overline{H}$ an integration yields

$$(1+\sqrt{x_{0,n}})^{-2l+|\alpha|-\varepsilon_1}\,|Q_1(x_0)|^{-\frac{1}{p_0}}\,\left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u_2\right\|_{L^{p_0}(Q_1(x_0))}$$
$$\lesssim_{n,\sigma,l,k,\alpha,p,\varepsilon_1,\varepsilon_2}\sup_{\substack{0<R\leq 1\\ z\in\overline{H}}}(R+\sqrt{z_n})^{-\varepsilon_1}\,R^{2-\varepsilon_2}\,|Q_R(z)|^{-\frac{1}{p}}\,\|f_2\|_{L^p(Q_R(z))}$$

for any $\varepsilon_1\geq 0$, $\varepsilon_2>0$ and any $\max\left\{1,\frac{1}{1+\sigma}\right\}<p<\infty$, $1\leq p_0<\infty$, if $2l-|\alpha|\leq 0$. Together with the on-diagonal estimates this gives the statement on $(0,1)\times\overline{H}$ if we set $\varepsilon_1=|\gamma|$.
For the rescaled version we let u now be a σ-solution to f on $[t_1,t_2)\times\overline{H}$ with initial value $g=0$. Then u is also a σ-solution to f on $[t_1,t_1+r^2)\times\overline{H}$ for any $0<r^2\leq t_2-t_1$ with initial value $g=0$. For the scaling

$$T_\lambda:(\hat{t},\hat{x})\mapsto(t_1+\lambda\,\hat{t},\lambda\,\hat{x})=:(t,x)$$

we then have

$$\left\|(\,\cdot\,)_n^l\,\partial_t^k\partial_x^\alpha u\right\|_{L^{p_0}(Q_r(t_1,x_0))}=\lambda^{\frac{n+1}{p_0}+l-k-|\alpha|}\,\left\|(\,\hat{\cdot}\,)\,\partial_{\hat{t}}^k\partial_{\hat{x}}^\alpha(u\circ T_\lambda)\right\|_{L^{p_0}(T_\lambda^{-1}(Q_r(t_1,x_0)))}.$$

As in the proof of Lemma 9.8 we can see that

$$T_\lambda^{-1}(Q_r(t_1,x_0))\subset\bigcup_{i=1}^N Q_{\frac{r}{\sqrt{\lambda}}}(\lambda^{-1}x_i)\text{ for }\lambda^{-1}x_i\in B_{4c_d^2\frac{r}{\sqrt{\lambda}}}(\lambda^{-1}x_0)$$

and a number N only depending on the dimension n. With $\lambda=r^2$ and the invariance of the time-shifted scaling that makes $u\circ T_{r^2}$ a σ-solution to $r^2\,(f\circ T_{r^2})$ on $(0,1)\times\overline{H}$, we can apply the statement proven above in every summand to get

$$\left\|(\,\cdot\,)_n^l\partial_t^k\partial_x^\alpha u\right\|_{L^{p_0}(Q_r(t_1,x_0))}$$
$$\lesssim_{n,\sigma,l,k,\alpha,\gamma,p,p_0,\varepsilon_2,\theta} r^{\frac{2n+2}{p_0}+2l-2k-2|\alpha|+2}\sum_{i=1}^N\left(1+\sqrt{\frac{x_{i,n}}{r^2}}\right)^{2l-|\alpha|+|\gamma|}|Q_1(r^{-2}x_i)|^{\frac{1}{p_0}}$$
$$\sup_{\substack{0<R\leq 1\\ z\in\overline{H}}}\Big(R^\theta\,|Q_R(z)|^{-\frac{1}{p_0}}\,\left\|D_{\hat{x}}^{|\gamma|}(f\circ T_{r^2})\right\|_{L^{p_0}(Q_R(z))}$$
$$+(R+\sqrt{z_n})^{-|\gamma|}\,R^{2-\varepsilon_2}\,|Q_R(z)|^{-\frac{1}{p}}\,\|f\circ T_{r^2}\|_{L^p(Q_R(z))}\Big).$$

Using the same calculations as in the proof of Theorem 9.9 for the measure of the cylinder and the rescalation of the right hand side from Lemma 9.8 as well as $\varepsilon_2:=-\theta+|\gamma|+2$ for $\theta<2+|\gamma|$, we arrive at the upper bound

$$\sup_{\substack{0<R\leq r\\ z\in\overline{H}}}\left(R^\theta\,|Q_R(t_1,z)|^{-\frac{1}{p_0}}\,\left\|D_x^{|\gamma|}f\right\|_{L^{p_0}(Q_R(t_1,z))}+(R+\sqrt{z_n})^{-|\gamma|}\,R^{\theta-|\gamma|}\,|Q_R(t_1,z)|^{-\frac{1}{p}}\,\|f\|_{L^p(Q_R(t_1,z))}\right)$$
$$r^{-2k-|\alpha|+|\gamma|+2-\theta}\sum_{i=1}^N(r+\sqrt{x_{i,n}})^{2l-|\alpha|+|\gamma|}\,|Q_r(x_i)|^{\frac{1}{p_0}}$$

for $x_i \in B_{16c_d^4 r}(x_0)$. As before we use doubling and triangle inequality to see that

$$|B_r(x_i)|^{\frac{1}{p_0}} \lesssim_n |B_r(x_0)|^{\frac{1}{p_0}}$$

and do the same trick on

$$(r + \sqrt{x_{i,n}})^{2l-|\alpha|-|\gamma|} \eqsim_n |B_r(x_i)|^{\frac{2l-|\alpha|+|\gamma|}{n}} r^{-2l+|\alpha|-|\gamma|}$$

in reversed order, since here the exponents on the ball measure are negative. This finishes the proof. ∎

10 The Non-Linear Equation

We finally turn to the non-linear perturbation equation we first started out with. Following [KL12] we would like to use a fixed point argument in special function spaces to get existence and uniqueness for the non-linear problem. In [KT01], the choice of these function spaces was motivated by the square function characterisation of BMO.

As noted before, the best initial values we can hope for in the transformation are contained in a homogeneous Lipschitz space. Now we need to choose our function spaces such that the inhomogeneity bounds the same left hand side as the Lipschitz initial values do. In light of the estimate from Remark 7.8 we therefore have to use the pointwise bound 9.9 for $l = 0$ and $|\alpha| = 1$. Since there we need l and $\alpha - \gamma$ to be Green-exponents, we have to have $|\gamma| = 1$. Accordingly, to achieve the same factors as in the bound for the initial values, we set $\varepsilon_1 = 1$ and $\theta = 2$. Note that for these values of l, $|\alpha|$ and $|\gamma|$ the condition on p_0 reduces to $\max\left\{n+1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$.

Now, fortunately, this very choice of the parameters $|\gamma|$ and θ in the localised L^p-bound for Calderón-Zygmund exponents by the inhomogeneity from Theorem 9.10 leads to the same combination of factors as the ones that can be bounded by the initial values by Corollary 7.9. Spelling out the possible combination of Calderón-Zygmund-exponents naturally suggests the following definitions:

For an interval $I = (t_1, t_2) \subset \mathbb{R}$ and $1 \leq q, p, p_0 \leq \infty$ let the function spaces $X := X(I, q)$ and $Y := Y(I, p, p_0)$ be given by

$$\begin{aligned}
\|u\|_{X(I,q)} = &\|\nabla_x u\|_{L^\infty(I\times H)} \\
&+ \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in H}} R^2 \left|Q_R(t_1,z)\right|^{-\frac{1}{q}} \|\nabla_x \partial_t u\|_{L^q(Q_R(t_1,z))} \\
&+ \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in H}} R\left(R+\sqrt{z_n}\right) \left|Q_R(t_1,z)\right|^{-\frac{1}{q}} \left\|D_x^2 u\right\|_{L^q(Q_R(t_1,z))} \\
&+ \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in H}} R^2 \left|Q_R(t_1,z)\right|^{-\frac{1}{q}} \left\|(\,\cdot\,)_n\, D_x^3 u\right\|_{L^q(Q_R(t_1,z))}
\end{aligned}$$

and

$$\begin{aligned}
\|f\|_{Y(I,p,p_0)} := &\sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in H}} R\left(R+\sqrt{z_n}\right)^{-1} \left|Q_R(t_1,z)\right|^{-\frac{1}{p}} \|f\|_{L^p(Q_R(t_1,z))} \\
&+ \sup_{\substack{0<R^2\leq t_2-t_1 \\ z\in H}} R^2 \left|Q_R(t_1,z)\right|^{-\frac{1}{p_0}} \|\nabla_x f\|_{L^{p_0}(Q_R(t_1,z))}\,.
\end{aligned}$$

This defines norms modulo constants. As an intersection of complete spaces, the spaces $X(I,q)$ and $Y(I,p,p_0)$ are also complete.

We furthermore denote the homogeneous Lipschitz space by

$$\dot{C}^{0,1}(H) := \left\{ g : \overline{H} \to \mathbb{R} : \|\nabla_x g\|_{L^\infty(H)} < \infty \right\}.$$

For functions depending on time and space whose spatial gradient is bounded we will alter this notation to $\dot{C}_x^{0,1}(I \times H)$. The space $X(I, q)$ is obviously contained in this space for any $1 \le q \le \infty$.

10.1 Remark *Rephrasing Remark 7.8, Corollary 7.9 and Theorems 9.9 and 9.10 in terms of these definitions yields the following statement:*
Let $\sigma > -1$, $t_1 > -\infty$, $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval, $\max\left\{n+1, \frac{1}{1+\sigma}\right\} < p_0 < \infty$ *and* $\max\left\{1, \frac{1}{1+\sigma}\right\} < p < \infty$*. Further let $f \in Y(I, p)$, $g \in \dot{C}^{0,1}(H)$ and u be a σ-solution to f on $\overline{I} \times \overline{H}$ with initial value g.*
Then we have

$$\|u\|_{X(I,p_0)} \lesssim_{n,\sigma,p,p_0} \|g\|_{\dot{C}^{0,1}(H)} + \|f\|_{Y(I,p,p_0)} .$$

Now consider the non-linearity more closely and remember that we have

$$\begin{aligned} f[u] &= -(1+\sigma)\,\frac{|\nabla_x u|^2}{\partial_{x_n} u + 1} - (\,\cdot\,)_n\,\partial_{x_n}\frac{|\nabla_x u|^2}{\partial_{x_n} u + 1} \\ &= -(1+\sigma)\frac{|\nabla_x u|^2}{\partial_{x_n} u + 1} - 2\,(\,\cdot\,)_n\,\frac{\nabla_x \partial_{x_n} u \cdot \nabla_x u}{\partial_{x_n} u + 1} + (\,\cdot\,)_n\,\frac{|\nabla_x u|^2\,\partial_{x_n}^2 u}{(\partial_{x_n} u + 1)^2}. \end{aligned}$$

Thanks to the special structure of this non-linearity we can establish mapping properties of $u \mapsto f[u]$ in the spaces defined above.
For a function space Z we denote closed balls centred at zero by $\overline{B}_\rho^Z := \{u \in Z \mid \|u\|_Z \le \rho\}$.

10.2 Proposition *Let $t_1 > -\infty$, $I = (t_1, t_2) \subset \mathbb{R}$, $1 \le p \le \infty$ and $0 < \rho < 1$.*
Then the mapping $\overline{B}_\rho^{X(I,p)} \ni u \mapsto f[u] \in Y(I, p, p)$ is analytic and we have

$$\|f[u]\|_{Y(I,p,p)} \lesssim_{n,p} \frac{1}{(1-\rho)^3}\ \|u\|^2_{X(I,p)} \text{ for all } u \in \overline{B}_\rho^{X(I,p)}$$

as well as

$$\|f[u_1] - f[u_2]\|_{Y(I,p,p)} \lesssim_{n,p} \frac{\rho}{(1-\rho)^6}\ \|u_1 - u_2\|_{X(I,p)} \text{ for all } u_1,\, u_2 \in \overline{B}_\rho^{X(I,p)}.$$

Proof: A direct calculation shows that

$$f[u] = f_1[u] \star \nabla_x u \star \nabla_x u + f_2[u] \star \nabla_x u \star (\,\cdot\,)_n D_x^2 u$$

and, for $j = 1, \dots, n$,

$$\partial_{x_j} f[u] = f_2[u] \star \nabla_x u \star D_x^2 u + f_2[u] \star \nabla_x u \star (\,\cdot\,)_n D_x^3 u + f_3[u] \star D_x^2 u \star (\,\cdot\,)_n D_x^2 u$$

with

$$f_i[u] := \sum_{k=1}^{i} (\partial_{x_n} u + 1)^{-k} \star (\nabla_x u)^{(k-1)\star},\ i = 1, 2, 3.$$

Fix arbitrary $1 \le p, p_0 \le \infty$. The analyticity is a consequence of the fact that polynomials are analytic and the $f_i[u]$, as functions of $\nabla_x u$, can be extended into power series. In any neighbourhood in $\overline{B}_\rho^{\dot{C}_x^{0,1}(I\times H)}$ this yields a power series expansion that converges in $Y(I, p, p_0)$.

We now consider the parts of $\|f[u]\|_{Y(I,p,p_0)}$ term by term. We will drop the initial time t_1 from the notation of the cylinders as well as the set in the notation of the L^∞-norm on the whole space $I \times H$.

$$\begin{aligned} R\,(R+\sqrt{z_n})^{-1}\,|Q_R(z)|^{-\frac{1}{p_0}}\ &\|f_1[u] \star \nabla_x u \star \nabla_x u\|_{L^{p_0}(Q_R(z))} \\ &\leq \|f_1[u]\|_{L^\infty}\ \|\nabla_x u\|_{L^\infty}^2\,, \end{aligned}$$

since $\|\nabla_x u\|_{L^{p_0}(Q_R(z))} \leq \|\nabla_x u\|_{L^\infty}\ |Q_R(z)|^{\frac{1}{p_0}}$ and $(R+\sqrt{z_n})^{-1} \leq R$.

$$\begin{aligned} R\,(R+\sqrt{z_n})^{-1}\,|Q_R(z)|^{-\frac{1}{p_0}}\ &\left\|f_2[u] \star \nabla_x u \star (\,\cdot\,)_n D_x^2 u\right\|_{L^{p_0}(Q_R(z))} \\ &\lesssim \|f_2[u]\|_{L^\infty}\ \|\nabla_x u\|_{L^\infty}\ R\,(R+\sqrt{z_n})\,|Q_R(z)|^{-\frac{1}{p_0}}\ \left\|D_x^2 u\right\|_{L^{p_0}(Q_R(z_0))}, \end{aligned}$$

this time since $\sqrt{x_n} \lesssim R+\sqrt{z_n}$ for any $x \in B_R(z)$.

$$\begin{aligned} R^2\,|Q_R(z)|^{-\frac{1}{p}}\ &\left\|f_2[u] \star \nabla_x u \star D_x^2 u\right\|_{L^p(Q_R(z))} \\ &\leq \|f_2[u]\|_{L^\infty}\ \|\nabla_x u\|_{L^\infty}\ R\,(R+\sqrt{z_n})\,|Q_R(z)|^{-\frac{1}{p}}\ \left\|D_x^2 u\right\|_{L^p(Q_R(z_0))}, \end{aligned}$$

again with $R \leq R+\sqrt{z_n}$.

$$\begin{aligned} R^2\,|Q_R(z)|^{-\frac{1}{p}}\ &\left\|f_2[u] \star \nabla_x u \star (\,\cdot\,)_n D_x^3 u\right\|_{L^p(Q_R(z))} \\ &\leq \|f_2[u]\|_{L^\infty}\ \|\nabla_x u\|_{L^\infty}\ R^2\,|Q_R(z)|^{-\frac{1}{p}}\ \left\|D_x^2 u\right\|_{L^p(Q_R(z_0))} \end{aligned}$$

straightforward.

$$\begin{aligned} R^2\,|Q_R(z)|^{-\frac{1}{p}}\ &\left\|f_3[u] \star D_x^2 u \star (\,\cdot\,)_n D_x^2 u\right\|_{L^p(Q_R(z))} \\ &\leq \|f_3[u]\|_{L^\infty}\ R^2\,|Q_R(z)|^{-\frac{1}{p}}\ \left\|(\,\cdot\,)_n\,|D_x^2 u|^2\right\|_{L^p(Q_R(z))} \\ &\lesssim_{n,p} \|f_3[u]\|_{L^\infty}\ \|\nabla_x u\|_{L^\infty}\ R^2\,|Q_R(z)|^{-\frac{1}{p}}\ \left\|(\,\cdot\,)_n D_x^3 u\right\|_{L^p(Q_R(z))} \end{aligned}$$

thanks to the identity

$$\left\|(\,\cdot\,)_n\,|D_x^2 u|^2\right\|_{L^p(Q_R(z))} = \left\|(\,\cdot\,)_n^{\frac{1}{2}}\,D_x^2 u\right\|_{L^{2p}(Q_R(z))}^2$$

and the weighted Gagliardo-Nirenberg interpolation inequality from Proposition 2.18 in its local form.
Setting $p = q = p_0$ we thus see that by the definition of $X(I,q)$ we have

$$\|f[u]\|_{Y(I,p,p)} \lesssim_{n,p} \left(\|f_1[u]\|_{L^\infty} + \|f_2[u]\|_{L^\infty} + \|f_3[u]\|_{L^\infty}\right) \|u\|_{X(I,p)}^2\,.$$

We therefore only need to consider $\|f_i\|_{L^\infty}$ for $i = 1, 2, 3$. By assumption we have $u \in \overline{B}_\rho^{\dot{C}^{0,1}(H)}$ with $\rho < 1$ and hence both

$$\|\nabla_x u\|_{L^\infty} \leq \rho \text{ and } \|\partial_{x_n} u + 1\|_{L^\infty}^{-k} \leq (1-\rho)^{-k}.$$

It follows that

$$\|f_i[u]\|_{L^\infty} \leq \sum_{k=1}^{i} \frac{\rho^{k-1}}{(1-\rho)^k}$$

which implies

$$\|f[u]\|_{Y(I,p,p)} \lesssim_{n,p} \sum_{k=1}^{3} \frac{\rho^{k-1}}{(1-\rho)^k} \, \|u\|_{X(I,p)}^2$$

and thus the first inequality.
We turn to the second inequality and note that for the first term we have

$$\begin{aligned} f_1[u_1] \star \nabla_x u_1 \star \nabla_x u_1 - f_1[u_2] \star \nabla_x u_2 \star \nabla_x u_2 =& (f_1[u_1] - f_1[u_2]) \star \nabla_x u_1 \star \nabla_x u_1 \\ &+ f_1[u_2] \star \nabla_x (u_1 - u_2) \star \nabla_x u_1 \\ &+ f_1[u_2] \star \nabla_x u_2 \star \nabla_x (u_1 - u_2) \end{aligned}$$

and likewise for the other terms. In the resulting summands we can proceed exactly as before and get an estimate

$$\begin{aligned} &\|f[u_1] - f[u_2]\|_{Y(I,p,p)} \\ &\lesssim_{n,p} \sum_{i=1}^{3} \Big(\|f_i[u_1] - f_i[u_2]\|_{L^\infty} \|u_1\|_{X(I,p)}^2 + \|f_i[u_2]\|_{L^\infty} \left(\|u_1\|_{X(I,p)} + \|u_2\|_{X(I,p)}\right) \|u_1 - u_2\|_{X(I,p)} \Big) . \end{aligned}$$

Note that in the part involving $f_3[u]$ we use Hölder's inequality to obtain

$$\begin{aligned} &\left\| (\,\cdot\,)_n \, D_x^2 u_1 \, D_x^2 (u_1 - u_2) \right\|_{L^p(Q_R(z))} \\ &\leq \left\| (\,\cdot\,)_n^{\frac{1}{2}} \, D_x^2 u_1 \right\|_{L^{2p}(Q_R(z_0))} \left\| (\,\cdot\,)_n^{\frac{1}{2}} \, D_x^2 (u_1 - u_2) \right\|_{L^{2p}(Q_R(z_0))} \\ &\lesssim_{n,p} \|u_1\|_{L^\infty}^{\frac{1}{2}} \left\| (\,\cdot\,)_n \, D_x^3 u_1 \right\|_{L^p(Q_R(z))}^{\frac{1}{2}} \|u_1 - u_2\|_{L^\infty}^{\frac{1}{2}} \left\| (\,\cdot\,)_n \, D_x^3 (u_1 - u_2) \right\|_{L^p(Q_R(z))}^{\frac{1}{2}} \end{aligned}$$

with the interpolation inequality applied onto both factors.
In view of the first part it remains to estimate the terms

$$\begin{aligned} f_i[u_1] - f_i[u_2] = \sum_{k=1}^{i} \Bigg(&\left((\partial_{x_n} u_1 + 1)^{-k} - (\partial_{x_n} u_2 + 1)^{-k} \right) \star (\nabla_x u_1)^{(k-1)\star} \\ &+ (\partial_{x_n} u_2 + 1)^{-k} \star \nabla_x (u_1 - u_2) \star \sum_{l=1}^{k-1} (\nabla_x u_1)^{(l-1)\star} (\nabla_x u_2)^{(k-1-l)\star} \Bigg) \end{aligned}$$

for $i = 1, 2, 3$. But it is clear that

$$
\begin{aligned}
&(\partial_{x_n} u_1 + 1)^{-k} - (\partial_{x_n} u_2 + 1)^{-k} \\
&= (\partial_{x_n} u_2 - \partial_{x_n} u_1)\,(\partial_{x_n} u_1 + 1)^{-k}\,(\partial_{x_n} u_2 + 1)^{-k} \sum_{l=1}^{k} \sum_{m=0}^{l-1} \binom{k}{l} (-1)^{k-l}\, \partial_{x_n} u_2^m\, \partial_{x_n} u_1^{l-1-m}
\end{aligned}
$$

and thus, because of $u_1, u_2 \in \overline{B}_\rho^X$ and $\rho < 1$, we get

$$
\left\| (\partial_{x_n} u_1 + 1)^{-k} - (\partial_{x_n} u_2 + 1)^{-k} \right\|_{L^\infty} \lesssim \|\partial_{x_n}(u_1 - u_2)\|_{L^\infty}\ (1-\rho)^{-2k} \sum_{l=1}^{k} \rho^{l-1}
$$

and consequently

$$
\|f_i[u_1] - f_i[u_2]\|_{L^\infty} \lesssim \|u_1 - u_2\|_{X(I,p)} \sum_{k=1}^{i} \left(\frac{\rho^{k-1}}{(1-\rho)^{2k}} + \frac{\rho^{k-2}}{(1-\rho)^{k}} \right).
$$

Using

$$
\|u_1\|_{X(I,p)}^2 \leq \rho\,(\|u_1\|_{X(I,p)} + \|u_2\|_{X(I,p)}),
$$

all this amounts to

$$
\|f[u_1] - f[u_2]\|_{Y(I,p,p)} \lesssim_{n,p} \sum_{k=1}^{3} \frac{\rho^{k-1}}{(1-\rho)^{2k}} \left(\|u_1\|_{X(I,p)} + \|u_2\|_{X(I,p)}\right) \|u_1 - u_2\|_{X(I,p)}\,,
$$

finishing the proof. ∎

This allows us to close the argument for the non-linear equation by means of a fixed point theorem in a way that closely follows [KL12]. With an idea developped in [Ang90] that was pushed further also by [KL12], we additionally get analyticity in the temporal and tangential directions of the solution we will construct as a consequence.

10.3 Theorem *Let $\sigma > -1$, $t_1 > -\infty$ and $I = (t_1, t_2) \subset \mathbb{R}$ be an open interval. Further let $\max\left\{n+1, \frac{1}{1+\sigma}\right\} < p < \infty$.*
Then there exists an $\varepsilon > 0$ and a $C(n,\sigma) > 1$ such that for any $g \in B_\varepsilon^{\dot{C}^{0,1}(H)}$ we can find a σ-solution $u_* \in X(I,p)$ for the non-linear perturbation equation on $\overline{I} \times \overline{H}$ with initial value g and satisfying $\|u_*\|_{X(I,p)} \lesssim_{n,\sigma,p} \|g\|_{\dot{C}^{0,1}(H)}$ that is unique within $B_{C\varepsilon}^{X(I,p)}$.
Moreover, u_* depends analytically on the data g, we have $u_* \in C^\infty(I \times \overline{H})$, and u_* is analytic in the temporal and tangential directions on $I \times \overline{H}$ with an $r > 0$ such that

$$
\sup_{(t,x) \in I \times H} (t - t_1)^{k + |\alpha'|}\, |\partial_t^k \partial_{x'}^{\alpha'} \nabla_x u_*(t,x)| \lesssim_n r^{-k-|\alpha'|}\, k!\, \alpha'!\ \|g\|_{\dot{C}^{0,1}(H)}
$$

for any $k \in \mathbb{N}_0$ and $\alpha' \in \mathbb{N}_0^{n-1}$ with $k + |\alpha'| > 0$.

Proof: In this proof the same arguments as in [KL12, Theorem 3.1.] apply. So define the function

$$
F : \dot{C}^{0,1}(H) \times X(I,p) \to X(I,p)
$$

by assigning $(g,u) \in \dot{C}^{0,1}(H) \times X(I,p)$ to the σ-solution to $f[u]$ on $\overline{I} \times \overline{H}$ with initial value g. By

Remark 10.1 and Proposition 10.2 we can find a $\rho_1 > 0$ and an $\varepsilon_1 > 0$ such that for any $g \in \overline{B}_{\varepsilon_1}^{\dot{C}^{0,1}(H)}$ the map $F(g,\cdot)$ is a contraction within $\overline{B}_{\rho_1}^{X(I,p)}$. Hence the contraction mapping principle provides us with a unique fixed point $u_* \in \overline{B}_{\rho_1}^{X(I,p)}$ that depends on g in a Lipschitz continuous way. Thus u_* is the unique global σ-solution for the non-linear equation with initial value g. The bound on $\|u_*\|_{X(I,p)}$ especially implies that u_* is Lipschitz in time and space. By [Koc99, Theorem 5.6.1.] the smoothness of u_* follows.
But moreover, F and therefore also

$$G : \dot{C}^{0,1}(H) \times X(I,p) \to X(I,p),\ G(g,u) := u - F(g,u)$$

are analytic on $\dot{C}^{0,1}(H) \times \overline{B}_{\rho_1}^{X(I,p)}$. Since $G(0,0) = 0$ and $D_u G(0,0) = \mathrm{id}$, the analytic implicit function theorem on Banach spaces is applicable ([Dei85]) and yields the existence of balls

$$\overline{B}_{\varepsilon_2}^{\dot{C}^{0,1}(H)} \subset \dot{C}^{0,1}(H) \text{ and } \overline{B}_{\rho_2}^{X(I,p)} \subset \overline{B}_{\rho_1}^{X(I,p)}$$

alongside with an analytic function

$$A : B_{\varepsilon_2}^{\dot{C}^{0,1}(H)} \to B_{\rho_2}^{X(I,p)}$$

that satisfies $A(0) = 0$ and $F(g,u) = u$ for any $g \in B_{\varepsilon_2}^{\dot{C}^{0,1}(H)}$ and $u \in B_{\rho_2}^{X(I,p)}$ if and only if $u = A(g)$. By the uniqueness of the fixed point in $\overline{B}_{\rho_1}^{X(I,p)} \supset B_{\rho_2}^{X(I,p)}$ above we conclude that A sends $g \in B_{\varepsilon_3}^{\dot{C}^{0,1}(I)}$ to u_* analytically, where $\varepsilon_3 := \min\ \{\varepsilon_1, \varepsilon_2\}$.
We now consider $(\tau, \xi') \in B_{\kappa_1}^{\mathbb{R}}(1) \times B_{\delta_1}^{\mathbb{R}^{n-1}}$ and define

$$f_{\tau,\xi'}[u] := \tau\, f[u] - (1-\tau)\, L_\sigma u - \xi' \cdot \nabla_x' u.$$

Note that then $f_{1,0}[u] = f[u]$. Similar as above we define

$$\widetilde{F} : B_{\kappa_1}^{\mathbb{R}}(1) \times B_{\delta_1}^{\mathbb{R}^{n-1}} \times \dot{C}^{0,1}(H) \times B_{\rho_2}^{X(I,p)} \to X(I,p)$$

that maps (τ, ξ', g, u) to the σ-solution to g and $f_{\tau,\xi'}[u]$ on $\overline{I} \times \overline{H}$. A straightforward calculation as in the proof of Proposition 10.2 shows that we have

$$\left\| \widetilde{F}(\tau, \xi', g, u) \right\|_{X(I,p)} \lesssim_{n,\sigma,p} \|g\|_{\dot{C}^{0,1}} + \tau\, \|u\|_{X(I,p)}^2 + (|1-\tau| + |\xi'|)\, \|u\|_{X(I,p)}\,.$$

Hence we can conclude that $D_u \widetilde{F}|_{(1,0,0,0)}$ vanishes. Applying the analytic implicit function theorem once more, this time on

$$\widetilde{G} := \mathrm{id}_{X(I,p)} - \widetilde{F}$$

at the point $(\tau, \xi', g, u) = (1,0,0,0)$ we then obtain the existence of $\kappa_2 < \kappa_1$, $\delta_2 < \delta_1$, $\rho_3 < \rho_2$ and ε_4 as well as a unique analytic function

$$\widetilde{A} : B_{\kappa_2}^{\mathbb{R}}(1) \times B_{\delta_2}^{\mathbb{R}^{n-1}} \times B_{\varepsilon_4}^{\dot{C}^{0,1}(H)} \to B_{\rho_3}^{X(I,p)}$$

such that

$$\widetilde{G}(\tau, \xi', g, \widetilde{A}(\tau, \xi', g)) = 0.$$

For $(\tau, \xi') \in B^{\mathbb{R}}_{\kappa_2}(1) \times B^{\mathbb{R}^{n-1}}_{\delta_2}$ and $(t, x) \in I \times \overline{H}$ let us now consider the transformation

$$T : (\tau, \xi', t, x) \mapsto (\tau(t - t_1) + t_1, x' - \xi'(t - t_1), x_n).$$

A simple calculation shows that $u \circ T(\tau, \xi', \cdot, \cdot)$ is a σ-solution to $f_{\tau,\xi'}[u]$ on $[t_1, \tau(t_2 - t_1) + t_1) \times \overline{H}$ with initial value g if u is a σ-solution to $f[u]$ on $I \times \overline{H}$ with initial value g. Therefore we have that

$$\widetilde{G}(\tau, \xi', A(g) \circ T(\tau, \xi', \cdot, \cdot)) = 0.$$

It is also clear that

$$A(g) \circ T(\tau, \xi', t_1, \cdot) = g = \widetilde{A}(\tau, \xi', g)(t_1, \cdot)$$

and thus the above uniqueness results for $\varepsilon := \min\ \{\varepsilon_3, \varepsilon_4\}$ and ρ_3 imply that $A(\cdot) \circ T = \widetilde{A}$. Especially, $u_* \circ T(\cdot, \cdot, t, x)$ is analytic as a function of τ and ξ' into $X(I, p)$ near $\tau = 1$ and $\xi' = 0$ for any $(t, x) \in I \times \overline{H}$. For finite t we have that

$$\begin{aligned}
\left|\nabla_x \partial_\tau (u_* \circ T(\tau, \xi', t, x))|_{(\tau,\xi')=(1,0)}\right| &= \left|\nabla_x (t - t_1)\, \partial_t u_*(t, x)\right|, \\
\left|\nabla_x \partial_{\xi_j} (u_* \circ T(\tau, \xi', t, x))|_{(\tau,\xi')=(1,0)}\right| &= \left|-\nabla_x (t - t_1)\, \partial_{x_j} u_*(t, x)\right| \text{ for } j = 1, \ldots, n-1
\end{aligned}$$

with similar formulas for higher order and mixed derivatives. This implies the analyticity of u_* on $I \times \mathbb{R}^n$ in t and x' as well as the formula given in the statement. ∎

10.4 Remark *It is shown in [Koc99, Proposition 6.3.1.] that our solution u_* and all its temporal derivatives are also analytic on $I \times \overline{H}$ in any spatial direction including the vertical one. We conjecture that u_* is indeed analytic on $I \times \overline{H}$ in time and space.*

A Singular Integrals in Spaces of Homogeneous Type

Measure Theory

We introduce some notions and notations of measure theory on arbitrary sets as well as topological spaces. For a deeper treatment consult [Els07], [EG92] or [Fed69]. The power set of a set X is denoted by $\mathcal{P}(X)$.

A.1 Definition *Consider an arbitrary set X.*
We say that $\mu : \mathcal{P}(X) \longrightarrow [0,\infty) \cup \{\infty\}$ is a measure on X if and only if

- $\mu(\emptyset) = 0$,
- $\mu(A_1) \leq \mu(A_2)$ *for all* $A_1 \subset A_2 \subset X$ *(monotonicity),*
- $\mu(\bigcup_{j\in\mathbb{N}} A_j) \leq \sum_{j\in\mathbb{N}} \mu(A_j)$ *for all* $(A_j)_{j\in\mathbb{N}} \subset \mathcal{P}(X)$ *(countable sub-additivity).*

What we call measure here is often referred to as outer measure in the literature. A set endowed with a measure is also called measure space.
Remark that $\mu \geq 0$ actually follows from the conditions in the definition and thus would not have to be assumed. On the other hand, it is not excluded by definiton that a measure takes on infinite values. A measure that decomposes its underlying set into subsets with finite measure is easier to handle.

A.2 Definition *Consider an arbitrary set X and a measure μ on X.*
We say that μ is countably finite if and only if there exist $(A_j)_{j\in\mathbb{N}} \subset \mathcal{P}(X)$ such that $\mu(A_j) < \infty$ for any $j \in \mathbb{N}$ and

$$X = \bigcup_{j\in\mathbb{N}} A_j.$$

By our definition, a measure assigns a number to any subset of X, thus measuring it. Nonetheless, for a subset to be called measurable we require an additional property.

A.3 Definition *Consider an arbitrary set X and a measure μ on X.*
We say that $A \subset X$ is measurable if and only if

$$\mu(C \cap A) + \mu(C \cap A^c) \leq \mu(C) \text{ for all } C \subset X.$$

The collection of all measurable subsets of X is denoted by $\mathcal{A}_\mu(X)$.

Because of the monotonicity this means that a subset is measurable if and only if it divides all other subsets in a measure-theoretically correct way. It is easy to see that any nullset is measurable, as are X itself and $\emptyset$, and that $\mathcal{A}_\mu(X)$ is closed with respect to complement and countable union. Moreover, μ is countably additive on its measurable sets.

A.4 Proposition *Let X be an arbitrary set and μ be a measure on X. Then*

$$\mu\Big(\bigcup_{j\in\mathbb{N}} A_j\Big) = \sum_{j\in\mathbb{N}} \mu(A_j) \text{ for all disjoint } A_j \in \mathcal{A}_\mu(X).$$

On a topological space X, the family of sets that is generated by the open subsets through complement and countable union are denoted by $\mathcal{B}(X)$. The elements of $\mathcal{B}(X)$ are called Borel sets.

A.5 Definition *Consider a topological space X and a measure μ on X. We say that μ is a Borel measure if and only if* $\mathcal{B}(X) \subset \mathcal{A}_\mu(X)$.

A Borel measure is therefore a measure for which at least the Borel sets are measurable, including all open and all closed subsets of X. Caratheodory proved that on metrisable topological spaces with metric d_X there is another characterisation of Borel measures.

A.6 Proposition *Let X be a metrisable topological space and μ be a measure on X. We then have: μ is a Borel measure if and only if*

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) \text{ for all } A_1,\, A_2 \subset X \text{ with } d_X(A_1, A_2) > 0.$$

A proof can for example be found in [Fal85].
Still, of course, not all subsets are necessarily measurable. Measures that can compensate this deficit are called regular.

A.7 Definition *Consider an arbitrary set X and a measure μ on X. We say that μ is regular if and only if for all* $A \subset X$ *there exists a measurable set* $C \in \mathcal{A}_\mu(X)$ *such that* $C \supset A$ *and* $\mu(C) = \mu(A)$.

The following definition sharpens the notion of a regular Borel measure a little, but decisive bit further.

A.8 Definition *Consider a topological space X and a Borel measure μ on X. We say that μ is Borel regular if and only if for all* $A \subset X$ *there exists a Borel set* $B \in \mathcal{B}(X)$ *such that* $B \supset A$ *and* $\mu(B) = \mu(A)$.

Instead of a Borel regular Borel measure we will merely speak of a Borel regular measure.
On certain topological spaces we can make sense of a local form of finiteness. Remember that in a Hausdorff space any compact set is closed.

A.9 Definition *Consider a locally compact Hausdorff space X and a measure μ on X. We say that μ is locally finite if and only if we have* $\mu(K) < \infty$ *for any compact* $K \subset X$.

Local compactness is often included in the definition of local finiteness, demanding that for any $x \in X$ there exists a compact neighbourhood with finite measure. However, a clearer separation of topological and measure theoretical properties is achieved by means of our definitions.
Locally finite Borel regular measures get a name in their own right.

A.10 Definition *Consider a locally compact Hausdorff space X and a measure μ on X. We say that μ is a Radon measure if and only if μ is a locally finite and Borel regular measure.*

Radon measures possess another regularity property that is important.

A.11 Definition *Consider a Hausdorff space X and a measure μ on X.*
We say that μ is tight if and only if

$$\mu(B) = \sup\ \{\mu(K) \mid K \text{ compact and } K \subset B\}\,.$$

A.12 Proposition *Let X be a locally compact Hausdorff space and μ be a measure on X.*
If μ is a Radon measure, then μ is tight.

This is proven for example in [EG92].
We recall now that metrisable topological spaces are Hausdorff. Balls with radius r centred at $x \in X$ taken with respect to the metric d_X on X are denoted by $B_r^X(x)$.

A.13 Definition *Consider a metrisable topological space X and a measure μ on X.*
μ is called doubling if and only if

$$\mu(B_{2r}^X(x)) \lesssim \mu(B_r^X(x)) \text{ for any } r > 0 \text{ and } x \in X.$$

We do not include the doubling constant explicitely into the definition, but would like to point out that it is not allowed to depend on any parameter at all. By monotonicity it is also obvious that the converse inequality holds without a constant.
Metric spaces that are endowed with a measure that is compatible with the metric as well as regular are important. They are close enough to the euclidean space, equipped with the Lebesgue measure, as to gain a lot of the results that hold in this model case.

A.14 Definition *We say that a metrisable locally compact topological space with a non-trivial doubling Radon measure is a space of homogeneous type.*

It is also possible to give a general definition that requires a quasi-metric on the measure space. However, we consent ourselves to the framework presented here.

Lebesgue Spaces

We study real-valued functions on measure spaces.

A.15 Definition *Consider an arbitrary set X, a measure μ on X and a function $u : X \longrightarrow \mathbb{R}$.*
We say that u is measurable if and only if $u^{-1}(B) \in \mathcal{A}_\mu(X)$ for any $B \in \mathcal{B}(\mathbb{R})$.

For measurable functions u on X, the construction of the integral over a measurable set $A \subset \mathcal{A}_\mu(X)$ with respect to μ, denoted by

$$\int_A u\, d\mu,$$

can be done via positive simple functions and indicator functions in the usual way. Note that the measurability of u implies that $|u|$ is also measurable.

A.16 Definition *Consider an arbitrary set X, a measure μ on X and a measurable set $A \in \mathcal{A}_\mu(X)$, as well as a measurable function $u : X \longrightarrow \mathbb{R}$ and $1 \leq p < \infty$.*
We say that u is p-integrable on A if and only if

$$\int_A |u|^p \, d\mu < \infty.$$

For $p = 1$ we simply say integrable instead of 1-integrable.
There is a similar notion for $p = \infty$. By the essential supremum of a function u on a measurable set A we mean

$$\operatorname*{ess\,sup}_A u := \inf \{ c \in \mathbb{R} : \mu(x \in A : f(x) > c) = 0 \} .$$

A.17 Definition *Consider an arbitrary set X, a measure μ on X, and a measurable set $A \in \mathcal{A}_\mu(X)$, as well as a measurable function $u : X \longrightarrow \mathbb{R}$ and $1 \leq p < \infty$.*
We say that u is essentially bounded or ∞-integrable on A if and only if

$$\operatorname*{ess\,sup}_A |u| < \infty.$$

We can now define the Lebesgue spaces.

A.18 Definition *Consider an arbitrary set X, a measure μ on X and a measurable set $A \in \mathcal{A}_\mu(X)$, as well as $1 \leq p \leq \infty$.*
The p-Lebesgue space on A is defined as

$$L^p(A, \mu) := \{ u : X \longrightarrow \mathbb{R} \mid u \text{ is } p\text{-integrable on } A \} .$$

After identifying functions that coincide almost everywhere, these spaces become Banach spaces with the norms

$$\|u\|_{L^p(A,\mu)} := \left(\int_A |u|^p \, d\mu \right)^{\frac{1}{p}} , \quad 1 \leq p < \infty$$

and

$$\|u\|_{L^\infty(A,\mu)} := \operatorname*{ess\,sup}_A |u|.$$

Regarding the dual spaces of the Lebesgue spaces up to isometric isomorphisms, the following is known.

A.19 Proposition *Let X be an arbitrary set, μ be a measure on X and $A \in \mathcal{A}_\mu(X)$.*

(i) If $1 < p < \infty$, then $L^p(A, \mu)' = L^{\frac{p}{p-1}}(A, \mu)$.

(ii) If μ is countably finite, then $L^1(A, \mu)' = L^\infty(A, \mu)$.

Briefly, we now consider operators between Lebesgue spaces that are given by a kernel. In this context the following result, called Schur's lemma, is important. A proof is contained in [Fol84].

A.20 Proposition *Let X and Y be arbitrary sets, μ be a measure on X and ν be a measure on Y, $A \in \mathcal{A}_\mu(X)$ and $C \in \mathcal{A}_\nu(Y)$. Furthermore, let $k : A \times C \longrightarrow \mathbb{R}$ be $\mu \times \nu$-measurable.*
If

$$\int_A |k(x,y)| \, d\mu(x) < \infty \textit{ for } \nu - \textit{almost every } y \in C$$

and

$$\int_C |k(x,y)| \, d\nu(y) < \infty \textit{ for } \mu - \textit{almost every } x \in A,$$

then

$$T : f \mapsto \int_C k(\cdot, y) f(y) \, d\nu(y)$$

maps $L^p(A,\mu)$ into $L^p(C,\nu)$ continuously and we have

$$\|Tf\|_{L^p(C,\nu)} \lesssim \|f\|_{L^p(A,\mu)}$$

for any $1 \leq p \leq \infty$.

On topological spaces with Borel measures, there is a local version of the Lebesgue spaces as well.

A.21 Definition *Consider a topological space X, a Borel measure μ on X and a measurable set $A \in \mathcal{A}_\mu(X)$, as well as $1 \leq p \leq \infty$.*
The local p-Lebesgue space on A is defined as

$$L^p_{loc}(A,\mu) := \{u : X \longrightarrow \mathbb{R} \mid u \textit{ is } p\textit{-integrable on } K \textit{ for any compact } K \subset A\}.$$

Any Lebesgue space is clearly contained in its corresponding local Lebesgue space. Local finiteness also guarantees a relation among local Lebesgue spaces.

A.22 Proposition *Let X be locally compact Hausdorff space, μ be a measure on X and $A \in \mathcal{A}_\mu(X)$.*
If μ is a locally finite Borel measure, then

$$L^p_{loc}(A,\mu) \subset L^{\overline{p}}_{loc}(A,\mu) \textit{ for any } 1 \leq \overline{p} \leq p \leq \infty.$$

Muckenhoupt Weights

We are now concerned with weighted measures in spaces of homogeneous type. The proofs of the following statements can be found in [Koc04].

A.23 Definition *Consider a space of homogeneous type (X,d,μ) as well as a μ-measurable function $\nu : X \longrightarrow [0,\infty)$ and $1 < p < \infty$.*
We say that ν is a p-Muckenhoupt weight on (X,d,μ) if and only if

$$\sup_{B \textit{ ball}} \mu(B)^{-1} \int_B \nu \, d\mu \left[\mu(B)^{-1} \int_B \nu^{-\frac{1}{p-1}} \, d\mu \right]^{p-1} \lesssim 1.$$

The class of p-Muckenhoupt weights on (X, d, μ) is denoted by $A_p(X, d, \mu)$.

In the sequel we often identify the Muckenhoupt weight ν with the measure $\nu\, d\mu$.

There is a characterisation of the p-Muckenhoupt weights in terms of the locally integrable functions on (X, d, μ).

A.24 Proposition *Let (X, d, μ) be a space of homogeneous type, $1 < p < \infty$ and $\nu \geq 0$ a μ-measurable function on X. We then have:*
ν is a p-Muckenhoupt weight if and only if

$$\left| \mu(B)^{-1} \int_B u\, d\mu \right|^p \lesssim \nu(B)^{-1} \int_B |u|^p\, d\nu$$

for any ball B and any $u \in L^1_{loc}(X, \mu)$.

The doubling property is inherited by Muckenhoupt weights.

A.25 Proposition *Let (X, d, μ) be a space of homogeneous type and $1 < p < \infty$.*
If $\nu \in A_p(X, d, \mu)$, then ν is a doubling measure on X with respect to d.

Finally, we state a duality property of Muckenhoupt classes.

A.26 Proposition *Let (X, d, μ) be a space of homogeneous type and $1 < p < \infty$. We then have:*
$\nu \in A_p(X, d, \mu)$ if and only if $\nu^{-\frac{1}{p-1}} \in A_{\frac{p}{p-1}}(X, d, \mu)$.

Maximal Functions

We discuss three different maximal functions in the general setting of a space of homogeneous type (X, d, μ) and study their relationship. The basic reference for this material is [Ste93]. It was noted by [CW77] that the arguments do not depend on the euclidean structure of the underlying space. We orientate ourselves at [Koc04] and [Koc08], where also the proofs we omit can be found.

For the definition of the maximal function, for any fixed ball $B_r(x_0)$ we introduce a class of functions

$$\mathcal{L}(B_r(x_0)) := \Bigg\{ \varphi \in C(X) \mid |\varphi(x)| \leq \mu(B_r(x_0))^{-1} \frac{\max\, \{r - d(x, x_0), 0\}}{r} \text{ for any } x \in B_r(x_0),$$
$$|\varphi(x) - \varphi(y)| \leq \mu(B_r(x_0))^{-1} \frac{d(x,y)}{r} \text{ for any } x,\, y \in B_r(x_0) \Bigg\}.$$

Obviously, we have $\|\varphi\|_{L^\infty(X,\mu)} \leq |B_r(x_0)|^{-1}$ and $\operatorname{supp} \varphi \subset B_r(x_0)$ as well as $-\varphi \in \mathcal{L}(B_r(x_0))$ for any $\varphi \in \mathcal{L}(B_r(x_0))$.

A.27 Definition *Consider $f \in L^1_{loc}(X, \mu)$.*

- *The maximal function of f is given by*

$$Mf(x) := \sup_{\substack{B \text{ ball} \\ B \ni x}} \sup_{\varphi \in \mathcal{L}(B)} \int_X f\, \varphi\, d\mu \text{ for any } x \in X.$$

- *The Hardy-Littlewood maximal function of f is given by*

$$M_{HL}f(x) := \sup_{\substack{B \text{ ball} \\ B \ni x}} \mu(B)^{-1} \int_B |f| \, d\mu \text{ for any } x \in X.$$

- *The maximal oscillation function of f is given by*

$$f^{\#}(x) := \sup_{\substack{B \text{ ball} \\ B \ni x}} \mu(B)^{-1} \int_B \left| f - \mu(B)^{-1} \int_B f \, d\mu \right| d\mu.$$

Mf, $M_{HL}f$ and $f^{\#}$ are all μ-measurable by their definitions. Furthermore, it is immediately clear that

$$0 \leq Mf(x) \leq M_{HL}f(x) \text{ for any } x \in X$$

holds as well as

$$f^{\#}(x) \lesssim M_{HL}f(x) \text{ for any } x \in X.$$

The first result is a weak-type L^1-estimate for M_{HL} whose proof requires the use of the Vitali covering theorem ([Koc04]).

A.28 Proposition *Let $f \in L^1(X, \mu)$.*
Then

$$\mu(\{x \in X \mid M_{HL}f(x) > \lambda\}) \lesssim \frac{1}{\lambda} \|f\|_{L^1(X,\mu)} .$$

From there we get L^p-boundedness of the Hardy-Littlewood maximal operator if $1 < p < \infty$.

A.29 Proposition *Let $1 < p < \infty$ and $f \in L^p(X, \mu)$.*
Then

$$\|M_{HL}f\|_{L^p(X,\mu)} \lesssim_p \|f\|_{L^p(X,\mu)} .$$

Immediate consequences are the bounds

$$(*) \qquad \|Mf\|_{L^p(X,\mu)} \lesssim_p \|f\|_{L^p(X,\mu)}$$

and

$$(**) \qquad \left\|f^{\#}\right\|_{L^p(X,\mu)} \lesssim_p \|f\|_{L^p(X,\mu)}$$

for any $1 < p < \infty$.

It turns out that any weight function for which M_{HL} is bounded on $L^p(X, \nu)$ is a p-Muckenhoupt weight. Moreover, the class $A_p(x, d, \mu)$ is indeed characterised by the behaviour of the Hardy-Littlewood maximal function. The proof of this requires a reverse Hölder inequality which in turn depends on an adjusted Whitney-type theorem ([Koc04]).

A.30 Proposition *Let $1 < p < \infty$ and $\nu \in L^1_{loc}(X, \mu)$ with $\nu \geq 0$. We then have: M_{HL} is a bounded operator on $L^p(X, \nu)$ if and only if $\nu \in A_p(X, d, \mu)$.*

The maximal function and the maximal oscillation function are dual objects in a certain sense. This is a famous result due to Fefferman in the euclidean setting. For the proof in spaces of homogeneous type we need a decomposition of $L^1(X, \mu)$-functions into a sum of so-called atoms, characterised by their bounded support, mean zero and controlled size in terms of their maximal function, and a bounded function. The duality statement then reads as follows.

A.31 Proposition *Let $1 < p < \infty$, $f_1 \in L^p(X, \mu)$ and $f_2 \in L^{\frac{p}{p-1}}(X, \mu)$. Then*

$$\int_X f_1 \, f_2 \, d\mu \lesssim \int_X M f_1 \, f_2^{\#} \, d\mu.$$

This implies that also the converse of the above inequality $(**)$ involving the maximal oscillation function holds. The proof also uses $(*)$.

A.32 Corollary *Let $1 < p < \infty$ and $f \in L^p(X, \mu)$. Then*

$$\|f\|_{L^p(X,\mu)} \lesssim_p \left\| f^{\#} \right\|_{L^p(X,\mu)}.$$

Actually, the same inequality holds for any p-Muckenhoupt weight ν.

A.33 Remark *With the maximal function and the maximal oscillation function it is possible to define the Hardy space $\mathcal{H}^1(X, d, \mu)$ and the space of functions of bounded mean oscillation $BMO(X, d, \mu)$, respectively. Proposition* A.31 *then ensures the duality of these spaces.*

Singular Integrals of Calderón-Zygmund-Type

We return to integral kernel operators now, this time explicitely considering singular kernels that satisfy certain cancellation conditions.

A.34 Definition *Consider $1 < q < \infty$ and a bounded linear operator T from $L^q(X, \mu)$ into itself that is given by a $\mu \times \mu$ measurable kernel $k : X \times X \longrightarrow \mathbb{R}$ such that*

$$y \mapsto k(x, y) \in L^1_{loc}(X \setminus \{x\}, \mu),$$

$$x \mapsto k(x, y) \in L^1_{loc}(X \setminus \{y\}, \mu),$$

and for any compactly supported $f \in L^\infty(X, \mu) \cap L^q(X, \mu)$ we have

$$Tf(x) = \int_X k(x, y) f(y) \, d\mu(y) \text{ for all } x \in (\operatorname{supp} f)^c.$$

We say that T is a Calderón-Zygmund operator on $L^q(X, \mu)$ if in addition the kernel k satisfies the Calderón-Zygmund cancellation conditions

$$|k(x, y)| \lesssim \Big(\mu(B_{d(x,y)}(x)) + \mu(B_{d(x,y)}(y)) \Big)^{-1} \text{ for any } x \neq y \in X,$$

and

$$|k(x,y) - k(\overline{x},\overline{y})| \lesssim \left(\mu(B_{d(x,y)}(x)) + \mu(B_{d(x,y)}(y))\right)^{-1} \left(\frac{d(x,\overline{x}) + d(y,\overline{y})}{d(x,y) + d(\overline{x},\overline{y})}\right)^{\delta}$$

for any $x \neq y \in X$ and $\overline{x} \neq \overline{y} \in X$ with

$$\frac{d(x,\overline{x}) + d(y,\overline{y})}{d(x,y) + d(\overline{x},\overline{y})} \leq \varepsilon$$

for some $\varepsilon \in (0,1)$ and some $\delta \in (0,1]$.

Decomposing the function f at a given point into one function supported close to x and one supported away from x, the cancellation conditions ensure the following pointwise estimate for Calderón-Zygmund operators.

A.35 Proposition *Let $1 < q < \infty$, T be a Calderón-Zygmund operator on $L^q(X,\mu)$ and $f \in L^q(X,\mu)$. Then*

$$(Tf)^{\#}(x) \lesssim_q (M(|f|^q)(x))^{\frac{1}{q}} \text{ for almost all } x \in X.$$

Together with corollary A.32 and $(*)$, the last proposition enables us to show the crucial Calderón-Zygmund bound in spaces of homogeneous type.

A.36 Proposition *Let $1 < q < \infty$.*
If T is a Calderón-Zygmund operator on $L^q(X,\mu)$, then for any $1 < p < \infty$ we have that T is an operator on $L^p(X,\mu)$ with

$$\|Tf\|_{L^p(X,\mu)} \lesssim_p \|f\|_{L^p(X,\mu)}.$$

Finally, this continues to hold for Muckenhoupt weights thanks to the extension of Corollary A.32 for this case.

A.37 Proposition *Let $1 < q < \infty$.*
If T is a Calderón-Zygmund operator on $L^q(X,\mu)$, then for $1 < p < \infty$ and $\nu \in A_p(X,d,\mu)$ we have that T is an operator on $L^p(X,\nu)$ with

$$\|Tf\|_{L^p(X,\nu)} \lesssim_p \|f\|_{L^p(X,\nu)}.$$

Bibliography

[AB79] Donald Gary Aronson and Philippe Bénilan. *Regularité des Solutions de l'Équation des Milieux Poreux dans* $\mathbb{R}^N$. Comptes Rendus Hebdomadaires des Séances de l' Académie des Sciences Paris, Série A. Sciences Mathemátiques **288** 103 – 105, 1979.

[AC83] Donald Gary Aronson and Luis Ángel Caffarelli. *The Initial Trace of a Solution of the Porous Medium Equation*. Transactions of the American Mathematical Society **280**(1) 351 – 366, 1983.

[ACP82] Donald Gary Aronson, Michael Grain Crandall and Lambertus Adrianus Peletier. *Stabilization of Solutions of a Degenerate Nonlinear Diffusion Problem*. Nonlinear Analysis, Theory, Methods & Applications **6**(10) 1001 – 1022, 1982.

[ACS98] Ioannis Athanasopoulos, Luis Ángel Caffarelli and Sandro Salsa. *Phase Transition Problems of Parabolic Type: Flat Free Boundaries Are Smooth*. Communications on Pure and Applied Mathematics **51**(1) 77 – 112, 1998.

[ACV85] Donald Gary Aronson, Luis Ángel Caffarelli and Juan Luis Vázquez. *Interfaces with a Corner Point in One-Dimensional Porous Medium Flow*. Communications on Pure and Applied Mathematics **38** 375 – 404, 1985.

[AF03] Robert Alexander Adams and John J. F. Fournier. *Sobolev Spaces*. Academic Press, Amsterdam, 2003.

[AG93] Donald Gary Aronson and J. L. Graveleau. *A selfsimilar solution to the focusing problem for the porous medium equation*. European Journal of Applied Mathematics **4** 65 – 81, 1993.

[Ali85] Nicholas Dimitrios Alikakos. *On the Pointwise Behavior of the Solutions of the Porous Medium Equation as t Approaches Zero or Infinity*. Nonlinear Analysis, Theory, Methods & Applications **9** 1095 – 1113, 1985.

[Ang88] Sigurd B. Angenent. *Analyticity of the Interface of the Porous Media Equation after the Waiting Time*. Proceedings of the American Mathematical Society **102** 329 – 336, 1988.

[Ang90] Sigurd B. Angenent. *Nonlinear analytic semiflows*. Proceedings of the Royal Society of Edinburgh. Section A. Mathematics **115** 91 – 107, 1990.

[Aro67] Donald Gary Aronson. *Bounds for the Fundamental Solution of a Parabolic Equation*. Bulletin of the American Mathematical Society **73** 890 – 896, 1967.

[Aro69] Donald Gary Aronson. *Regularity Properties of Flows through Porous Media*. SIAM Journal on Applied Mathematics **17**(2) 461 – 467, 1969.

[Aro70] Donald Gary Aronson. *Regularity Properties of Flows through Porous Media: The Interface*. Archive for Rational Mechanics and Analysis **37** 1 – 10, 1970.

[Aro86] Donald Gary Aronson. *The Porous Medium Equation*. In: *Nonlinear Diffusion Problems* (Antonio Fasano and Mario Primicerio, editors), pages 1 – 46. Springer-Verlag, Berlin, 1986.

[AV87] Donald Gary Aronson and Juan Luis Vázquez. *Eventual C^∞-Regularity and Concavity for Flows in One-Dimensional Porous Media*. Archive for Rational Mechanics and Analysis **99**(4) 329 – 348, 1987.

[BC79] Haïm Brézis and Michael Grain Crandall. *Uniqueness of Solutions of the Initial-Value Problem for $u_t - \Delta\phi(u) = 0$*. Journal de Mathématiques Pures et Appliquées **58**(2) 153 – 163, 1979.

[BC81] Philippe Bénilan and Michael Grain Crandall. *The Continuous Dependence on Φ of Solutions of $u_t - \Delta\Phi(u) = 0$*. Indiana University Mathematics Journal **30**(2) 161 – 177, 1981.

[BCP84] Philippe Bénilan, Michael Grain Crandall and Michel Pierre. *Solutions of the Porous Medium Equation in $\mathbb{R}^N$ under Optimal Conditions on Initial Values*. Indiana University Mathematics Journal **33**(1) 51 – 87, 1984.

[Bén83] Philippe Bénilan. *A strong regularity L^p for solutions of the porous media equation*. In: *Contributions to nonlinear partial differential equations* (Claude Williams Bardos, Alain Damlamian, Jesús Ildefonso Díaz and Jesús Hernandéz, editors), pages 39 – 58. Pitman Publishing, Marshfield, 1983.

[Caf77] Luis Ángel Caffarelli. *The Regularity of Free Boundaries in Higher Dimensions*. Acta Mathematica **139** 155 – 184, 1977.

[Caf88] Luis Ángel Caffarelli. *A Harnack Inequality Approach to the Regularity of Free Boundaries. Part III: Existence Theory, Compactness, and Dependence on X*. Annali della Scuola Normale Superiore di Pisa **15**(4) 583 – 602, 1988.

[Caf89] Luis Ángel Caffarelli. *A Harnack Inequality Approach to the Regularity of Free Boundaries. Part II: Flat Free Boundaries are Lipschitz*. Communications on Pure and Applied Mathematics **42** 55 – 78, 1989.

[CF79] Luis Ángel Caffarelli and Avner Friedman. *Continuity of the Density of a Gas Flow in a Porous Medium*. Transactions of the American Mathematical Society **252** 99 – 113, 1979.

[CF80] Luis Ángel Caffarelli and Avner Friedman. *Regularity of the Free Boundary of a Gas Flow in an n-Dimensional Porous Medium*. Indiana University Mathematics Journal **29**(3) 361 – 391, 1980.

[CVW87] Luis Ángel Caffarelli, Juan Luis Vázquez and Noemí Irene Wolanski. *Lipschitz Continuity of Solutions and Interfaces of the N-Dimensional Porous Medium Equation*. Indiana University Mathematics Journal **36**(2) 373 – 401, 1987.

[CW77] Ronald R. Coifman and Guido Weiss. *Extensions of Hardy Spaces and their Use in Analysis*. Bulletin of the American Mathematical Society **83**(4) 569 – 645, 1977.

[CW90] Luis Ángel Caffarelli and Noemí Irene Wolanski. *$C^{1,\alpha}$ Regularity of the Free Boundary for the N-Dimensional Porous Media Equation*. Communications on Pure and Applied Mathematics **43**(7) 885 – 902, 1990.

[Dei85] Klaus Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.

[DF85] Emmanuele DiBenedetto and Avner Friedman. *Hölder Estimates for Nonlinear Degenerate Parabolic Systems*. Journal für die Reine und Angewandte Mathematik **357** 1 – 22, 1985.

[DH98] Panagiota Daskalopoulos and Richard Hamilton. *Regularity of the Free Boundary for the Porous Medium Equation*. Journal of the American Mathematical Society **11**(4) 899 – 965, 1998.

[DHL01] Panagiota Daskalopoulos, Richard Hamilton and Ki-Ahm Lee. *All Time C^∞-Regularity of the Interface in Degenerate Diffusion: A Geometric Approach*. Duke Mathematical Journal **108**(2) 295 – 327, 2001.

[DK84] Björn E. J. Dahlberg and Carlos Eduardo Kenig. *Non-Negative Solutions of the Porous Medium Equation*. Communications in Partial Differential Equations **9**(5) 409 – 437, 1984.

[DK93] Björn E. J. Dahlberg and Carlos Eduardo Kenig. *Weak Solutions of the Porous Medium Equation*. Transactions of the American Mathematical Society **336**(2) 711 – 725, 1993.

[DK07] Panagiota Daskalopoulos and Carlos Eduardo Kenig. *Degenerate Diffusions*. European Mathematical Society, Zürich, 2007.

[EG92] Lawrence Craig Evans and Ronald Francis Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, Boca Raton, 1992.

[Ego75] Yuri Vladimirovich Egorov. *Subelliptic Operators*. Russian Mathematical Surveys **30**(2) 59 – 118, 1975.

[Els07] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer-Verlag, Berlin, 2007.

[Fal85] Kenneth Falconer. *The geometry of fractal sets*. Cambridge University Press, Cambridge, 1985.

[Fed69] Herbert Federer. *Geometric Measure Theory*. Springer-Verlag, Berlin, 1969.

[FKS82] Eugene Barry Fabes, Carlos Eduardo Kenig and Raul Paolo Serapioni. *The Local Regularity of Solutions of Degenerate Elliptic Equations*. Communications in Partial Differential Equations **7** 77 – 116, 1982.

[Fol84] Gerald Budge Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & sons, New York, 1984.

[FP83] Charles Louis Fefferman and Duong Hong Phong. *Subelliptic eigenvalue problems*. In: *Conference on Harmonic Analysis in Honor of Antoni Zygmund* (William Beckner, Alberto Pedro Calderón, Robert Allen Fefferman and Peter Wilcox Jones, editors), volume 2, pages 590 – 606. Wadsworth, Belmont, 1983.

[Fri82] Avner Friedman. *Variational Principles and Free-Boundary Problems*. John Wiley & sons, Chichester, 1982.

[Fri08] Avner Friedman. *Partial Differential Equations*. Dover Publications, Mineola, 2008.

[FS86] Eugene Barry Fabes and Daniel Wyler Stroock. *A New Proof of Moser's Parabolic Harnack Inequality Using the Old Ideas of Nash*. Archive for Rational Mechanics and Analysis **96**(4) 327 – 338, 1986.

[FSC86] Charles Louis Fefferman and Antonio Sánchez-Calle. *Fundamental solutions for second order subelliptic operators*. Annals of Mathematics **124**(2) 247 – 272, 1986.

[HKM93] Juha Heinonen, Tero Kilpeläinen and Olli Martio. *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Oxford University Press, Oxford, 1993.

[Hö67] Lars Hörmander. *Hypoelliptic second order differential equations*. Acta Mathematica **119**(1) 147 – 171, 1967.

[Kil94] Tero Kilpeläinen. *Weighted Sobolev Spaces and Capacity*. Annales Academiae Scientiarum Fennicae. Series A I. Mathematica **19** 95 – 113, 1994.

[KL12] Herbert Koch and Tobias Lamm. *Geometric Flows with Rough Initial Data*. Asian Journal of Mathematics **16**(2) 209 – 235, 2012.

[Koc98] Herbert Koch. *Classical Solutions to Phase Transition Problems are Smooth*. Communications in Partial Differential Equations **23**(3) 389 – 437, 1998.

[Koc99] Herbert Koch. *Non-Euclidean Singular Integrals and the Porous Medium Equation*. Habilitation, Ruprecht-Karls-Universität Heidelberg, 1999.

[Koc04] Herbert Koch. *Partial differential equations and singular integrals*. In: *Dispersive Nonlinear Problems in Mathematical Physics* (Piero D'Ancona and Vladimir Georgev, editors), pages 59 – 122. Dipartimenta di Matematica della Seconda Università di Napoli, Napoli, 2004.

[Koc08] Herbert Koch. *Partial Differential Equations with Non-Euclidean Geometries*. Discrete and Continuous Dynamical Systems. Series S **1**(3) 481 – 504, 2008.

[KT01] Herbert Koch and Daniel Tataru. *Well-posedness for the Navier-Stokes equations*. Advances in Mathematics **157**(1) 22 – 35, 2001.

[Kuf85] Alois Kufner. *Weighted Sobolev Spaces*. John Wiley & sons, Chichester, 1985.

[Lin86] Chang-Shou Lin. *Interpolation Inequalities with Weights*. Communications in Partial Differential Equations **11**(14) 1515 – 1538, 1986.

[LNVV09] Peng Lu, Lei Ni, Juan Luis Vázquez and Cédric Villani. *Local Aronson-Bénilan estimates and entropy formulae for porous medium and fast diffusion equations on manifolds*. Journal de Mathématiques Pures et Appliquées **91** 1 – 19, 2009.

[LUS75] Olga Aleksandrovna Ladyženskaya, Nina Nikolaevna Ural'ceva and Vsevolod Alekseevich Solonnikov. *Linear and Quasi-Linear Equations of Parabolic Type*. American Mathematical Society, Providence, 1975.

[Maz85] Vladimir G. Maz'ya. *Sobolev Spaces*. Springer-Verlag, Berlin, 1985.

[Mei92] Anvarbek M. Meirmanov. *The Stefan Problem*. De Gruyter, Berlin, 1992.

[Mos64] Jürgen Moser. *A Harnack Inequality for Parabolic Differential Equations*. Communications on Pure and Applied Mathematics **17** 101 – 134, 1964.

[Mos67] Jürgen Moser. *Correction to "A Harnack Inequality for Parabolic Differential Equations"*. Communications on Pure and Applied Mathematics **20** 231 – 236, 1967.

[MS64] Norman George Meyers and James B. Serrin, Jr. $H = W$. Proceedings of the National Academy of Sciences of the United States of America **51** 1055 – 1056, 1964.

[Mun00] James Raymond Munkres. *Topology*. Prentice Hall, Upper Saddle River, 2000.

[NS96] Francesco Nicolosi and Igor V. Skrypnik. *Nirenberg-Gagliardo Interpolation Inequality and Regularity of Solutions of Nonlinear Higher Order Equations*. Topological Methods in Nonlinear Analysis **7** 327 – 347, 1996.

[NSW85] Alexander Nagel, Elias Menachem Stein and Stephen Wainger. *Balls and metrics defined by vector fields I: Basic properties*. Acta Mathematica **155**(1) 103 – 147, 1985.

[OM10] Frank William John Olver and Leonard C. Maximon. *Bessel Functions*. In: *NIST Handbook of Mathematical Functions* (Frank William John Olver, Daniel William Lozier, Ronald F. Boisvert and Charles Winthrop Clark, editors), pages 215 – 286. Cambridge University Press, Cambridge, 2010.

[Ott01] Felix Otto. *The geometry of dissipative evolution equations: the porous medium equation* **26**(1&2) 101 – 174, 2001.

[Pat59] Richard Eric Pattle. *Diffusion from an Instantaneous Point Source with a Concentration-Dependent Coefficient*. Quaterly Journal of Mechanics and Applied Mathematics **12**(4) 407 – 409, 1959.

[Pel81] Lambertus Adrianus Peletier. *The Porous Media Equation*. In: *Applications of Nonlinear Analysis in the Physical Sciences* (Herbert Amann, Norman William Bazley and Klaus Kirchgässner, editors), pages 229 – 241. Pitman Publishing, London, 1981.

[Pie82] Michel Pierre. *Uniqueness of the Solutions of $u_t - \Delta\phi(u) = 0$ with Initial Datum a Measure*. Nonlinear Analysis, Theory, Methods & Applications **6**(2) 175 – 187, 1982.

[Sab61] Evgeniya Sergeevna Sabinina. *On the Cauchy Problem for the Equation of Nonstationary Gas Filtration in Several Space Variables*. Soviet Mathematics Doklady **2** 166 – 169, 1961.

[Sac83] Paul E. Sacks. *Continuity of Solutions of a Singular Parabolic Equation*. Nonlinear Analysis, Theory, Methods & Applications **7**(4) 387 – 409, 1983.

[SC84] Antonio Sánchez-Calle. *Fundamental solutions and geometry of the sum of squares of vector fields*. Inventiones mathematicae **78**(1) 143 – 160, 1984.

[Ste93] Elias Menachem Stein. *Harmonic Analysis*. Princeton University Press, Princeton, 1993.

[Váz82] Juan Luis Vázquez. *Symétrisation pour $u_t = \Delta\varphi(u)$ et Applications*. Comptes Rendus Hebdomadaires des Séances de l' Académie des Sciences Paris, Série A. Sciences Mathemátiques **295** 71 – 74, 1982.

[Váz92] Juan Luis Vázquez. *An Introduction to the Mathematical Theory of the Porous Medium Equation*. In: *Shape Optimization and Free Boundaries* (Michel Claude Delfour and Gert Sabidussi, editors), pages 347 – 389. Kluwer Academic Publishers, Dordrecht, 1992.

[Váz07] Juan Luis Vázquez. *The Porous Medium Equation*. Clarendon Press, Oxford, 2007.

[Ves89] Vincenzo Vespri. *Analytic Semigroups, Degenerate Elliptic Operators and Applications to Nonlinear Cauchy Problems*. Annali di Matematica Pura ed Applicata **155** 353 – 388, 1989.

[Wer05] Dirk Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, 2005.

[Wid75] David Vernon Widder. *The Heat Equation*. Academic Press, London, 1975.

[Yos68] Kôsaku Yosida. *Functional Analysis*. Springer-Verlag, Berlin, 1968.

Summary

We are concerned with the Cauchy problem for the porous medium equation and prove stability and regularity of the pressure of solutions close to flat fronts under conditions on the initial data that are optimal for the methods we use.
To this end we consider a perturbed travelling wave solution of the transformed porous medium pressure equation on the closed upper half space and obtain stability in the homogeneous Lipschitz sense, temporal and tangential analyticity and precise estimates of derivatives. This type of stability is the least possible one enabling us to perform a change of dependent and independent variables that allows to transfer the results to the positivity set of the porous medium equation.
The major part of the argument concentrates on the linear perturbation equation on the closed upper half space. In a first step we define a weak notion of energy solution and show that such solutions satisfy an energy identity and energy estimates. A localisation of these results in terms of an intrinsic metric induced by the spatial part of the linear differential operator is then extended to a pointwise estimate, and thus smoothness for local solutions follows. For solutions to the initial value problem on the closed half space, an exponential factor can be included in the arguments to give estimates against rough norms of initial data. But moreover, we also get a pointwise exponential decay estimate of any derivative of the Green function with an upper bound resembling the Gaussian function with respect to the intrinsic metric and a naturally arising weighted measure. Consequences of the Gaussian estimate are on-diagonal and off-diagonal kernel estimates that imply a global pointwise estimate against the inhomogeneity. Furthermore, the theory of singular integrals and Calderón-Zygmund operators in spaces of homogeneous type can be applied to find a globally valid localised L^p-estimates against the inhomogeneity.
From there we turn to the nonlinear equation. The linear results suggest to construct certain function spaces in which the special shape of the nonlinearity leads to the possibility of operating a fixed point argument to get existence and uniqueness of solutions as well as Lipschitz continuity and thus smoothness. As a byproduct, varying the argument slightly provides us with the means to get analyticity in any direction save the vertical one, alongside with a precise estimate for the perturbation.