NTNU – Trondheim
Norwegian University of
Science and Technology

On the Well-Posedness of a Class of Whitham-like Nonlocal Equations with Weak Dispersion

Henrik Fiskerstrand Gjørtz

Master of Science in Physics and Mathematics
Submission date: June 2015
Supervisor: Mats Ehrnstrøm, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

We consider the Cauchy problems for a class of Whitham-like nonlocal equations with weak dispersion. Specifically, based on classical theory by Kato, local well-posedness in Sobolev spaces of order $s > 3/2$ for this class of equations is proven, both on the real line and on the torus. The possibility of extending to global well-posedness is also discussed, and in one specific case a global ill-posedness result is given. Additionally, the text includes a largely self-contained treatment of the theory of Sobolev spaces of real order, both on $\mathbb{R}^d$ and on the one-dimensional torus.

Sammendrag

Vi studerer initialverdiproblemene knyttet til en klasse Whitham-lignende ikkelokale differensialligninger med svak dispersjon. Ved hjelp av klassisk teori utviklet av Kato, bevises velstiltheten til initialverdiproblemene i Sobolevrom av orden $s > 3/2$, både på tallinjen og på torusen. Muligheten for å utvide til global velstilthet diskuteres også, og i ett tilfelle viser vi at initialverdiproblemet ikke er globalt velstilt. Teksten inneholder i tillegg en utledning av teorien for Sobolevrom av reell orden, både på $\mathbb{R}^d$ og på den endimensjonale torusen.

Notation

The following notation is used throughout the text.

- The **non-negative integers**, $\mathbb{N}_0 = \{0, 1, 2, ...\}$.
- A d-dimensional **multi-index** α is an an ordered d-tuple $\alpha = (\alpha_1, \alpha_2, ..., \alpha_d)$, where each $\alpha_i \in \mathbb{N}_0$ and $|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_d$. We write $\alpha \in \mathbb{N}_0^d$.
- The **set of compactly supported smooth functions from $\mathbb{R}^d$ to $\mathbb{R}$** is denoted by $C_c^\infty(\mathbb{R}^d)$. At certain points the notation $\mathcal{D}(\mathbb{R}^d)$ is also used, as the set of compactly supported smooth functions is the predual of the distribution space $\mathcal{D}'(\mathbb{R}^d)$.
- $BC^k(\mathbb{R}^d)$ denotes **the space of k times differentiable functions from $\mathbb{R}^d$ to $\mathbb{C}$ whose derivatives are continuous and bounded**, i.e. with norm
$$\|f\|_{BC^k(\mathbb{R}^d)} := \sum_{|\alpha| \le k} \sup_{x \in \mathbb{R}^d} |D^\alpha f(x)| < \infty.$$
$BC^k(\mathbb{R}^d)$, $k \in \mathbb{N}_0$, is a Banach space with this norm [14, p. 17].
- We will sometimes write $\langle \cdot \rangle^s$ for $(1 + |\cdot|^2)^{\frac{s}{2}}$ (Japanese bracket).
- By Θ_k we mean the function $x \mapsto \exp(ikx)$.
- If $(X, \|\cdot\|_X)$ is a normed space and $S \subseteq \mathbb{R}$ is any subset, $C^0(S, X)$ denotes the space of functions from S to X that are continuous with respect to the Euclidean metric and the metric induced by $\|\cdot\|_X$. The space can be equipped with the norm
$$\|f\|_{C^0(S,X)} = \sup_{t \in S} \|f(t)\|_X.$$
- If S is a closed and bounded subset of $\mathbb{R}$ and $(X, \|\cdot\|_X)$ is a Banach space, then $C^0(S, X)$ denotes the space of functions from S to X that are continuous with respect to the Euclidean metric and the metric induced by $\|\cdot\|_X$. When equipped with the metric
$$d_{C^0(S,X)}(f, g) := \|f - g\|_{C^0(S,X)},$$
$C^0(S, X)$ is a complete metric space [30, p. 65].
- The **positive real numbers**, $\mathbb{R}^+ = (0, \infty)$.

Notation believed to be well-known and unambiguous is not included in the above. Some additional notation will be defined in the text itself.

Preface

This thesis was written during the spring semester of 2015, and its submission marks the author's completion of the five-year Sivilingeniør/M.Sc. programme Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU), within the specialisation Industrial Mathematics.

After a brief introduction to relevant aspects of mathematical water wave theory, the thesis begins with a treatment of the theory of Sobolev spaces of real order, both on $\mathbb{R}^d$ and on the one-dimensional torus (periodic Sobolev spaces). We then present two examples of well-posedness analysis for linear equations, in order to familiarise ourselves with the concept of well-posedness in function or distribution spaces, before moving on to the main part of the thesis, where we establish the results described in the abstract. In the writing of this text we have assumed from the reader only a basic knowledge of measure theory and partial differential equations. We have intentionally outsourced the proofs of some results concerning the basic function spaces of analysis such as the Schwartz space and the space of smooth functions of compact support, other than that our treatment is largely self-contained.

More precisely, the text progresses as follows:

Section 1 is a general introduction to some important equations from water wave theory, in particular the Korteweg-de Vries equation and the Whitham equation. Sobolev spaces and the concept of well-posedness are also introduced.

Section 2 begins with a derivation of fundamental Fourier theory on L^2, which is then applied in the development of the basic theory of Sobolev spaces of non-negative order on $\mathbb{R}^d$. Specifically, besides proving fundamental properties of these spaces such as completeness, we answer the questions of when one can embed a Sobolev space into the space of bounded and continuous functions, and when a Sobolev space is closed under multiplication. The text [13] was used as the primary source on the theory of Sobolev spaces, although this text only deals with spaces on $\mathbb{R}$.

Section 3 deals with extending the class of Sobolev spaces from spaces of non-negative order to spaces of real order. It starts with a short presentation of the theory of distributions where the canonical distribution space $\mathcal{D}'(\mathbb{R}^d)$ and the tempered distributions are introduced. Next it is shown that the Fourier transform is well-defined and in fact is an automorphism on the space of tempered distributions. Finally the Sobolev spaces on $\mathbb{R}^d$ of real order are defined as subspaces of the tempered distributions with finite Sobolev norm.

Section 4 is devoted to an analogous treatment of the theory of periodic Sobolev spaces of real order. Our approach to the periodic Sobolev spaces follows closely that of the book [21], although with a couple of exceptions our proofs are all original.

Section 5 includes some practical examples of well-posedness analyses for simple linear equations, and it is further specified exactly what we mean by a Cauchy problem for a PDE being well-posed in a function or distribution space.

Section 6 contains the main part of this thesis, where we investigate the well-posedness of the Cauchy problems for a class of nonlocal and nonlinear Whitham-like dispersive equations in Sobolev spaces. Adapting a method previously used in [10] to prove local well-posedness for the periodic Camassa-Holm equation, and more recently in [15] to prove local well-posedness for the Whitham equation, the Cauchy problems for these equations are shown to be locally well-posed in Sobolev spaces of order $s > 3/2$, both in the periodic case and on the real line. The method we adapt is based on a classical theorem of Kato from [23]. Furthermore, the possibility of extending to a global well-posedness result is discussed, and in one special case it is shown that results from the article [7] imply a global ill-posedness result in Sobolev spaces of order $s > 3/2$.

I am grateful to my adviser Professor Mats Ehrnström for guiding me safely through the whole thesis writing process, both by plotting the course of my thesis and by helping me pass the many mathematical hurdles I encountered. I would also like to thank Long Pei for helping to formulate the topic of this thesis, and lastly the many excellent professors and students at NTNU whom I have learned much from during my time in Trondheim.

Contents

1 Introduction

For the Vikings settled on the western coast of Norway, understanding the ocean was the key to prosperity and adventure. Eventually, their supreme knowledge of the waves would take them to America five hundred years before Columbus. In Norse mythology, the sea is ruled by the giant Ægir along with his wife, the goddess Rán. Their nine daughters became the waves of the ocean. In no particular order their names and translations thereof are: Himinglæva, the wave that reflects the sky; Dúfa, the pitching wave; Hefring, the rising wave; Udr, the frothing wave; Hrönn, the grasping wave; Bylgja, the billowing wave; Dröfn, the wave; Kólga, the chilling wave; and Blódughadda, the blood red wave after a naval battle.

Although the modern mathematician is unlikely to find himself bloodying the sea during a naval battle, the inquiring spirit of the Viking explorers is alive in the mathematical field of nonlinear water wave theory. As the many daughters of Ægir and Rán attest, the motion of water waves can be very complex, and a mathematical model perfectly describing every facet of water wave motion would have to be impossibly intricate. Still, we would like to ensure that our simple models are as nuanced as possible. In particular, we want to ensure that under the right conditions, they admit solutions which reflect certain observed behaviours of physical water waves.

Among the more miraculous species of water waves are what we today call *solitons*, localised solitary waves that maintain their shape while propagating at a constant velocity, and can cross each other and emerge from the collision unchanged. In [37], John Scott Russell describes an encounter he had with what he termed the *Wave of Translation* in 1834 on the Union Canal near Edinburgh, Scotland:

> "I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped - not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon

which I have called the Wave of Translation."

Russell's discovery could not be explained by existing water wave theory at the time, and therefore was met with some skepticism. It wasn't until the 1870s that Lord Rayleigh and Joseph Boussinesq developed theory that supported Russell's observations, and in 1877 in [5, p. 360] Boussinesq introduced (in a footnote) the shallow-water equation

$$u_t + uu_x + u_{xxx} = 0,$$

which we today know as the Korteweg-de Vries (KdV) equation, after Diederik Korteweg and Gustav de Vries who re-derived the equation in 1895 [26].

The KdV equation turns out to hit the perfect balance between so-called dispersive effects and nonlinear effects, and admits soliton solutions [13]. In order to explain how these two effects work and cancel each other out, let us consider the simplest wave equation, the transport equation

$$u_t + u_x = 0$$

with $(t, x) \in \mathbb{R} \times \mathbb{R}$. This equation admits only solutions that are translation of the initial profile $u(0, \cdot) = u_0$, i.e. the solution takes the values $u(t, x) = u_0(t - x)$. These waves display one of the properties we require of solitons, namely that they travel at constant velocity and do not change shape. If we add the nonlinear term uu_x to the transport equation, we get $u_t + uu_x + u_x = 0$. The change of variables $\tilde{u}(t, x) = u(t, x + t)$ gives us the familiar Burgers' equation, so we may equivalently consider

$$\begin{cases} u_t + uu_x = 0, & (t, x) \in \mathbb{R} \times \mathbb{R}, \\ u(0, x) = u_0(x). \end{cases}$$

A solution of this Cauchy problem is determined by the initial profile u_0 by the implicit formula

$$u(t, x) = u_0(x - ut),$$

from which we gather that the level curves of u in the x-t plane are lines, with slopes equal to the corresponding values of $1/u$. A point on the initial profile at a height h over the x-axis will therefore appear to move with the speed h. Hence the nonlinear term uu_x has the effect of making the speed of propagation of a wave dependent on its amplitude. The solution wave will therefore distort over time, yet each point on the wave remains at its initial amplitude. Given an initial profile which is decreasing and positive over some region, Burgers' equation displays another important facet of wave motion, namely it allows for solutions which break, meaning that the solution wave's profile gradually steepens yet remains bounded, until the gradient at some point is vertical [9].

If we exchange the nonlinear term uu_x in Burgers' equation for the linear term u_{xxx}, we get the linearised KdV equation

$$\begin{cases} u_t + u_{xxx} = 0, & (t,x) \in \mathbb{R} \times \mathbb{R}, \\ u(0,x) = u_0(x). \end{cases}$$

In Section 5.3, we solve this equation using the Fourier transform, and find that the solution is a superposition of waves of the form $\hat{u}_0(\xi)e^{i\xi^3 t}e^{i\xi x}$, where $\xi \in \mathbb{R}$ is the frequency and $\hat{u}_0$ is the Fourier transform of the initial profile, i.e.

$$u(t,x) = \int_{\mathbb{R}} \hat{u}_0(\xi) e^{i\xi[\xi^2 t + x]}\, d\xi.$$

We see that the value $\hat{u}_0(\xi)$ represents the amplitude of the wave of frequency ξ, and that these Fourier components of the solution remain undiminished in amplitude but travel with different frequency-dependent velocities given by $-\xi^2$. This effect, where the velocity of a component of a wave depends on the component's frequency, is called *dispersion*. The asymptotic effect of dispersion is not as easy to analyse as that of nonlinearity, but in [1, Section 3.4.3], it is demonstrated how in the limit $x/t \to 0$, the solution of the linearised KdV equation tends towards the *self-similar* solution,

$$\lim_{x/t \to 0} u(t,x) \approx \frac{\hat{u}_0(0)}{(3t)^{1/3}} \operatorname{Ai}\left(\frac{x}{(3t)^{1/3}} \right),$$

where $\operatorname{Ai}(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{isy + s^3/3}\, ds$ is the Airy function and is independent of u. This expresses how the effect of the dispersive term u_{xxx} is here to reduce the initial profile to a universal shape, which is slowly diminishing in amplitude while spreading out and flattening.

The KdV equation combines the profile-steepening effects of Burgers' equation and the flattening effects of dispersion, which struggle against each other in what turns out to be a perfectly balanced fight, leading to the equation's admittance of soliton solutions.

Certain other outcomes of the nonlinearity-dispersion balancing in the KdV equation are possible, however, the solution may not break [42, Section 13.14]. The fact that the KdV equation - which is a model of waves on shallow water surfaces - does not admit solution waves that display this very real physical trait, lead to the introduction of the alternative and more general model known as the Whitham equation. The equation is named after by Gerald Whitham, who introduced it in 1967 as a model to study breaking of non-linear dispersive water waves [43]. Whitham's model is given by

$$u_t + uu_x + Lu_x = 0, \tag{1.1}$$

where L is an operator defined via the Fourier transform by

$$\mathcal{F}(Lf)(\xi) = \left(\frac{\tanh \xi}{\xi}\right)^{1/2} \hat{f}(\xi).$$

Note that in the limit $\xi \to 0$, that is for long wave-lengths (or from another perspective, in shallow water), the KdV equation is recovered as an approximation (cf. Section 6). In [32], it is shown that a solution of the Whitham equation will indeed break if the slope of its initial profile is sufficiently large and negative at some point.

If instead of letting $\xi \to 0$ and getting the KdV equation, we let $\xi \to \infty$, we get an equation given by (1.1) with the operator L defined by

$$\mathcal{F}(Lf)(\xi) = |\xi|^{-1/2} f(\xi).$$

This equation is part of the class of Whitham-like equations that we study in the final section of this text. Our main concern is establishing the *well-posedness* of the Cauchy problems for these equations, both on the line and on the torus (the periodic Cauchy problem). The concept of well-posedness was first suggested by Hadamard in the early 1900s [16, p. 49] (or see [31, p. 451]). That a Cauchy problem for a partial differential equation is well-posed (in the sense of Hadamard) means that it has a unique solution which depends continuously on the initial data. This classification extracts the essential properties a model of a (nonchaotic) physical system should have: The existence of a unique solution reflects the definiteness of the physical situation, while the solution's continuous dependence on initial data reflects the stability of the system - changing the initial conditions only slightly should affect the outcome only slightly.

We distinguish between local and global (in time) well-posedness, with the latter being stronger than the former. That a Cauchy problem for a PDE is locally well-posed means that we can only guarantee that it has a unique solution which depends continuously on the initial data for a finite amount time T. If instead $T = \infty$, the problem is said to be globally well-posed.

It is very possible for local well-posedness to hold and global well-posedness not to hold, for instance for Burgers' equation we know that the gradient of the solution may blow up in finite time. When the solution wave breaks, clearly it no longer satisfies the differential equation since it is discontinuous. Thus, if we're only interested in classical solutions, we would say that the Cauchy problem is *ill-posed*. However, it is useful to expand our concept of solution, because like many other PDEs that model physical systems, Burgers' equation has physically correct solutions that satisfy a certain weaker formulation of the equation. The weak formulation of an equation is obtained when exchanging the classical derivatives for so-called weak derivatives, which we define in the very beginning of this report. The weak derivative coincides with the classical derivative when the latter exists, however functions that are not differentiable may be weakly differentiable. When we later state Cauchy problems for PDEs in this thesis, the spatial derivatives appearing in the equations will in general be weak derivatives. We consider solutions of this weak formulation (sometimes called weak solutions) worthy solutions, and refer to them simply as solutions. We will however point out if a solution is actually classical, or what can be done to make it classical.

The notion of weak derivatives lead to the introduction of Sobolev spaces, or more precisely the classical Sobolev spaces, which are spaces of functions with a certain amount of well-behaved weak derivatives. Specifically, a function in a Sobolev space of order k has weak derivatives of order up to k that are all square integrable. The order of a Sobolev space is therefore an expression of the regularity of the elements (functions or distributions) in that space, and we shall generally look for solutions in these spaces instead of in the classical C^k-spaces. It turns out that the Sobolev spaces have some very nice properties that make them easy to work with. We devote the next three sections to developing Sobolev theory, both for general functions on $\mathbb{R}^d$ and for periodic functions on $\mathbb{R}$.

2 Sobolev spaces on $\mathbb{R}^d$ of non-negative order

In this section we define the weak derivative and introduce the classical Sobolev spaces, and we show that these spaces are complete. Next we establish some Fourier theory, which is applied in extending the class of Sobolev spaces to include the fractional Sobolev spaces of real non-negative order, and we prove some important properties of these spaces. This section is dense with results, but we will in fact find use for each of them in later sections.

Remark 2.1. The writing on the classical Sobolev spaces is original, while the results and proofs concerning the fractional spaces are adapted from or inspired by similar results in [13], which treats case $d = 1$ (except for the proofs of Theorems 2.13 and 2.14, and also the entire Section 2.5, which are original).

2.1 The weak derivative and classical Sobolev spaces

We introduce the spaces of locally p-integrable functions:

Definition 2.1 ($L^p_{loc}(\mathbb{R}^d)$-spaces)**.** Let $1 \leq p < \infty$. We say a function $f : \mathbb{R}^d \to \mathbb{C}$ is in $L^p_{loc}(\mathbb{R}^d)$ if for every compact subset $K \subseteq \mathbb{R}^d$,

$$\int_K |f|^p \, dx < \infty.$$

A function f is in $L^\infty_{loc}(\mathbb{R}^d)$ if its essential supremum over any compact subset K is bounded, $\operatorname{ess\,sup}_K |f| < \infty$.

Remark 2.2. Clearly $L^p(\mathbb{R}^d) \subseteq L^p_{loc}(\mathbb{R}^d)$. It is also a fact that $L^p_{loc}(\mathbb{R}^d) \subseteq L^1_{loc}(\mathbb{R}^d)$ for $1 \leq p \leq \infty$. This follows from Hölder's inequality: Let K be any compact subset of $\mathbb{R}^d$, then

$$\int_K |f| \, dx \leq \left(\int_K |f|^p \, dx \right)^{1/p} \left(\int_K dx \right)^{1/q} < \infty,$$

where $1/p + 1/q = 1$. Locally integrable functions play an important role in distribution theory (which we give a brief introduction to in Section 3).

For functions in these spaces we can define the weak derivative:

Definition 2.2 (Weak derivative of order $|\alpha|$)**.** Let $u \in L^1_{loc}(\mathbb{R}^d)$. A function $v \in L^1_{loc}(\mathbb{R}^d)$ is called a weak derivative of u of order $|\alpha|$, written $v = D^\alpha u$, if

$$\int_{\mathbb{R}^d} \varphi v \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} u \, \partial^\alpha \varphi \, dx \quad \forall \varphi \in C^\infty_c(\mathbb{R}^d). \tag{2.1}$$

Here α is a d-dimensional multi-index and $\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$. We say that u is n times weakly differentiable if $D^\alpha u$ exists for all α s.t. $|\alpha| \leq n$.

When the classical derivative exists it coincides with the weak derivative, and in that case (2.1) is just the formula for partial integration over $\mathbb{R}^d$ (any boundary terms vanish due to φ having compact support).

Proposition 2.1. *The weak derivative is unique up to a set of measure zero.*

Proof. Consider two functions v and w both satisfying (2.1), i.e.

$$\int_{\mathbb{R}^d} \varphi(v - w)\, dx = 0$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$. Then the difference vanishes almost everywhere by the du Bois-Reymond lemma. □

We differentiate between the classical Sobolev spaces and the fractional Sobolev spaces. The latter class of spaces is an extension of former.

Definition 2.3 (The classical Sobolev spaces)**.** Let $k \in \mathbb{N}_0$, then

$$W_p^k(\mathbb{R}^d) = \{f \in L^p(\mathbb{R}^d) \mid D^\alpha f \in L^p(\mathbb{R}^d)\ \forall\, \alpha \in \mathbb{N}_0^d,\ |a| \leq k\},$$

with $1 \leq p \leq \infty$ an integer. These are the classical Sobolev spaces.

The classical Sobolev spaces can be equipped with a norm to make them Banach spaces:

Theorem 2.2 (Completeness of $W_p^k(\mathbb{R}^d)$)**.** *The classical Sobolev space $W_p^k(\mathbb{R}^d)$ equipped with the norm*

$$\|f\|_{W_p^k(\mathbb{R}^d)} = \left(\sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^p(\mathbb{R}^d)}^p \right)^{1/p}$$

is a Banach space.

Proof. Let us first verify that $\|\cdot\|_{W_p^k(\mathbb{R}^d)}$ is a norm: The properties $\|u\|_{W_p^k(\mathbb{R}^d)} = 0$ if and only if $f = 0$ almost everywhere and $\|\lambda f\|_{W_p^k(\mathbb{R}^d)} = |\lambda| \|f\|_{W_p^k(\mathbb{R}^d)}$ follow directly from the corresponding properties of $\|\cdot\|_{L^p(\mathbb{R}^d)}$. We also have

$$\begin{aligned} \|f + g\|_{W_p^k(\mathbb{R}^d)} &= \left(\sum_{|\alpha| \leq k} \|D^\alpha f + D^\alpha g\|_{L^p(\mathbb{R}^d)}^p \right)^{1/p} \\ &\leq \left(\sum_{|\alpha| \leq k} \left(\|D^\alpha f\|_{L^p(\mathbb{R}^d)} + \|D^\alpha g\|_{L^p(\mathbb{R}^d)} \right)^p \right)^{1/p} \\ &\leq \left(\sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^p(\mathbb{R}^d)}^p \right)^{1/p} + \left(\sum_{|\alpha| \leq k} \|D^\alpha g\|_{L^p(\mathbb{R}^d)}^p \right)^{1/p}, \end{aligned}$$

where we have used the regular Minkowski inequality with Lebesgue measure in going from the first to second line, then Minkowski's inequality with counting measure in the final step. Thus the property $\|f+g\|_{W_p^k(\mathbb{R}^d)} \le \|f\|_{W_p^k(\mathbb{R}^d)} + \|g\|_{W_p^k(\mathbb{R}^d)}$ holds, meaning that $\|\cdot\|_{W_p^k(\mathbb{R}^d)}$ is indeed a norm.

We now prove completeness: Let $\{f_n\}_n$ be a Cauchy sequence in $W_p^k(\mathbb{R}^d)$. Then $\{f_n\}_n$ is also Cauchy in $L^p(\mathbb{R}^d)$ and by the completeness of L^p-spaces, $f_n \to f$ in $L^p(\mathbb{R}^d)$ for some $f \in L^p(\mathbb{R}^d)$. For any $|\alpha| \le k$, the sequence $\{D^\alpha f_n\}_n$ will also be Cauchy in $L^p(\mathbb{R}^d)$, so $D^\alpha f_n \to f^\alpha$ in $L^p(\mathbb{R}^d)$ for some $f^\alpha \in L^p(\mathbb{R}^d)$. We want to prove $f \in W_p^k(\mathbb{R})$, i.e. we need $D^\alpha f \in L^p(\mathbb{R}^d)$ for $|\alpha| \le k$, so it will be sufficient if we can prove $D^\alpha f = f^\alpha$. To this end, let $\varphi \in C_c^\infty(\mathbb{R}^d)$ and q be such that $1/p + 1/q = 1$. Then by Hölder's inequality we have

$$\begin{aligned}\left|\int_{\mathbb{R}^d} (f_n - f) D^\alpha \varphi \, dx\right| &\le \left(\int_{\mathbb{R}^d} |f_n - f|^p \, dx\right)^{1/p} \left(\int_{\mathbb{R}^d} |D^\alpha \varphi|^q \, dx\right)^{1/q} \\ &= \|f_n - f\|_{L^p(\mathbb{R}^d)} \|D^\alpha \varphi\|_{L^q(\mathbb{R}^d)} \to 0,\end{aligned}$$

since $f_n \to f$ in $L^p(\mathbb{R}^d)$ and $\|D^\alpha \varphi\|_{L^q(\mathbb{R}^d)}$ is bounded for $\varphi \in C_c^\infty(\mathbb{R}^d)$. Similarly

$$\left|\int_{\mathbb{R}^d} (D^\alpha f_n - f^\alpha) \varphi \, dx\right| \le \|D^\alpha f_n - f^\alpha\|_{L^p(\mathbb{R}^d)} \|\varphi\|_{L^q(\mathbb{R}^d)} \to 0.$$

Thus $\int_{\mathbb{R}^d} f_n D^\alpha \varphi \, dx \to \int_{\mathbb{R}^d} f D^\alpha \varphi \, dx$ and $\int_{\mathbb{R}^d} (D^\alpha f_n) \varphi \, dx \to \int_{\mathbb{R}^d} f^\alpha \varphi \, dx$. This is exactly what we need, as recalling the definition of the weak derivative we can write

$$\begin{aligned}\int_{\mathbb{R}^d} f D^\alpha \varphi \, dx &= \lim_{n\to\infty} \int_{\mathbb{R}^d} f_n D^\alpha \varphi \, dx \\ &= \lim_{n\to\infty} (-1)^{|\alpha|} \int_{\mathbb{R}^d} (D^\alpha f_n) \varphi \, dx \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^d} f^\alpha \varphi \, dx,\end{aligned}$$

which implies $D^\alpha f = f^\alpha$. □

The case $p = 2$ is of particular interest, as the space $W_2^k(\mathbb{R}^d)$ inherits the Hilbert space property of $L^2(\mathbb{R}^d)$.

Definition 2.4 ($W_2^k(\mathbb{R}^d)$-inner product)**.** The classical Sobolev space on $\mathbb{R}^d$ of order $k \in \mathbb{N}_0$ for $p = 2$, which we usually denote by $H^k(\mathbb{R}^d)$, is the Hilbert space with inner product

$$\langle f, g\rangle_{H^k(\mathbb{R}^d)} = \sum_{|\alpha|\le k} \langle D^\alpha f, D^\alpha g\rangle_{L^2(\mathbb{R}^d)} = \sum_{|\alpha|\le k} \int_{\mathbb{R}^d} (D^\alpha f)\, D^\alpha g \, dx.$$

Remark 2.3. Each of the properties that $\langle\cdot,\cdot\rangle_{H^k(\mathbb{R}^d)}$ needs to satisfy in order to be an inner product follow from the corresponding properties of $\langle\cdot,\cdot\rangle_{L^2(\mathbb{R}^d)}$.

Clearly the classical definition of Sobolev spaces makes sense only for $k \in \mathbb{N}_0$. In order to introduce the fractional Sobolev spaces of non-negative order, that is the spaces $H^s(\mathbb{R}^d)$ for real $s \geq 0$, we need an alternative definition of Sobolev spaces that makes sense for all non-negative s, and which is equivalent in the cases $s = k \in \mathbb{N}_0$.

2.2 The Fourier transform on the Schwartz space

Our first step on the way to the fractional Sobolev spaces is establishing some Fourier theory. We will start by defining the Fourier transform on the so-called Schwartz space of rapidly decreasing smooth functions, named after the French mathematician Laurent Schwartz.

Definition 2.5 (The Schwartz space)**.** The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is the space of rapidly decreasing smooth functions, more precisely

$$\mathcal{S}(\mathbb{R}^d) = \{\varphi \in C^\infty(\mathbb{R}^d) \mid \sup_{x\in\mathbb{R}^d} |x^\alpha D^\beta \varphi(x)| < \infty \ \forall \alpha, \beta \in \mathbb{N}_0^d\}.$$

Here $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d}$ and $D^\beta \varphi$ is the classical partial derivative since φ is smooth. We can define a family of seminorms on this space,

$$\|\varphi\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}^d} |x^\alpha D^\beta \varphi(x)| \quad \alpha, \beta \in \mathbb{N}_0^d. \tag{2.2}$$

The properties $\|\lambda\varphi\|_{\alpha,\beta} = |\lambda| \|\varphi\|_{\alpha,\beta}$ and $\|\varphi + \psi\|_{\alpha,\beta} \leq \|\varphi\|_{\alpha,\beta} + \|\psi\|_{\alpha,\beta}$ clearly hold. One actually also has $\|\varphi\|_{\alpha,\beta} = 0$ if and only if $\varphi \equiv 0$ for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. This is because $\|\varphi\|_{\alpha,\beta} = 0$ implies $D^\beta\varphi(x) = 0$ for all $x \neq 0$, and by continuity $D^\beta\varphi(x) = 0$ for all $x \in \mathbb{R}^d$. Then φ must be a polynomial (in d variables), and since it vanishes at infinity, $\varphi \equiv 0$. Thus the seminorms are in fact norms on the Schwartz space. We refer to them as seminorms because $\mathcal{S}(\mathbb{R}^d)$ is not complete with respect to just finitely many of them.

The Schwartz space is a linear space with the family of seminorms (2.2) inducing a topology on the space, specifically we have the following notion of convergence:

Definition 2.6 (Convergence in $\mathcal{S}(\mathbb{R}^d)$)**.** We say a sequence $\{\varphi_n\}_n \subseteq \mathcal{S}(\mathbb{R}^d)$ converges to φ in $\mathcal{S}(\mathbb{R}^d)$ when $\|\varphi_n - \varphi\|_{\alpha,\beta} \to 0$ as $n \to \infty$ for all α, β. We will use the notation $\varphi_n \underset{\mathcal{S}}{\to} \varphi$.

We also introduce a metric on $\mathcal{S}(\mathbb{R}^d)$,

$$d_{\mathcal{S}}(\varphi,\psi) = \sum_{\alpha,\beta} \frac{1}{2^{|\alpha|+|\beta|}} \frac{\|\varphi-\psi\|_{\alpha,\beta}}{1+\|\varphi-\psi\|_{\alpha,\beta}}.$$

Note that $d_{\mathcal{S}}(\varphi_n, \varphi) \to 0$ if and only if $\varphi_n \underset{\mathcal{S}}{\to} \varphi$. The Schwartz space equipped with the metric $d_{\mathcal{S}}$ is a complete metric space, in fact it is a *Fréchet space* [18], a metric space complete with respect to a metric induced not by a norm, but by a countable family of seminorms. We also introduce an equivalent seminorm to (2.2),

$$P_{\alpha,\beta}(\varphi) = \sup_{x\in\mathbb{R}^d}(1+|x|)^{|\alpha|}|D^\beta \varphi(x)|. \tag{2.3}$$

From the finiteness of the seminorms it follows that $\mathcal{S}(\mathbb{R}^d) \subseteq L^p(\mathbb{R}^d)$, $1 \le p \le \infty$.

From the definition we see that the Schwartz space is invariant under differentiation and multiplication by polynomials. This property makes it an ideal set to define the Fourier transform $\mathcal{F}$ on, as a corresponding symmetry in $\mathcal{F}$ leads to the Schwartz space being invariant under Fourier transform.

Definition 2.7 (The Fourier transform)**.** The Fourier transform of a function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ is defined by

$$\mathcal{F}(\varphi)(\xi) := \hat{\varphi}(\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x) e^{-ix\cdot\xi}\, dx,$$

where $x \cdot \xi$ is the standard inner product on $\mathbb{R}^d$.

Definition 2.8 (The inverse Fourier transform)**.** The inverse Fourier transform of a function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ is defined by

$$\mathcal{F}^{-1}(\varphi)(\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x) e^{ix\cdot\xi}\, dx.$$

Notice that $\mathcal{F}^{-1}(\varphi)(\xi) = \mathcal{F}(\varphi)(-\xi)$.

Remark 2.4. The above notation and naming is justified by the fact that $\mathcal{F}^{-1}\mathcal{F}\varphi = \mathcal{F}\mathcal{F}^{-1}\varphi = \varphi$ for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. This is the content of Theorem 2.9.

Immediately from the definition we extract the following basic properties of the Fourier transform:

Theorem 2.3 (Properties of the Fourier transform)**.** *For $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and $\xi \in \mathbb{R}^d$,*

(i) $\mathcal{F}(D^\alpha\varphi)(\xi) = i^{|\alpha|}\xi^\alpha\hat{\varphi}(\xi)$

(ii) $\mathcal{F}(x^\alpha\varphi)(\xi) = i^{|\alpha|}D^\alpha\hat{\varphi}(\xi)$

(iii) $\mathcal{F}(e^{ia\cdot x}\varphi)(\xi) = \hat{\varphi}(\xi - a)$

(iv) $\mathcal{F}(\varphi(\lambda x))(\xi) = \frac{1}{|\lambda|^d}\hat{\varphi}(\frac{\xi}{\lambda})$ *for* $\lambda \in \mathbb{R}$.

Proof. The first result follows from partial integration since the Schwartz functions are smooth. The second follows from $\varphi(x)e^{-ix\cdot\xi}$ being smooth and integrable for $\varphi \in \mathcal{S}(\mathbb{R}^d)$, allowing us to differentiate under the integral sign in the expression for $\hat{\varphi}(\xi)$. The two final results follow from the definition of the Fourier transform and a change of variables. □

The convolution of two functions is an important concept in Fourier theory, since as we shall see in Theorem 2.5, the Fourier transform maps the convolution of two functions to the pointwise product of their Fourier transforms (modulo a constant).

Definition 2.9 (Convolution)**.** Given two functions f and g, the convolution $f * g$ is defined as

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy = \int_{\mathbb{R}^d} f(x-y)g(y)\,dy = (g * f)(x),$$

if it exists. Commutativity follows from a change of variables, given that either of the convolution integrals converge.

Remark 2.5. From the commutativity of the convolution follows the useful property that when taking the derivative of $f * g$, we may choose to let the derivative fall on either function. Therefore the convolution is at least as smooth as the smoothest of the functions involved. If we consider the convolution as an average of f about a point using weights from g (or vice versa), we see how convolution with certain smooth functions (known as approximations to the identity, or mollifiers) can be used to create a smooth approximation to a rough function.

In the following lemma we prove the very useful fact that the Schwartz space is closed under convolution. The convolution also has some other important properties as a map on certain function spaces:

Lemma 2.4. *The convolution is a bilinear continuous mapping*

(i) $L^1(\mathbb{R}^d) \times L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d), \quad 1 \le p \le \infty,$

(ii) $BC^k(\mathbb{R}^d) \times L^1(\mathbb{R}^d) \to BC^k(\mathbb{R}^d), \quad k \in \mathbb{N}_0,$

(iii) $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d).$

Proof. It is clear from the definition that the map is bilinear (linear in both arguments).

(i) This result is sometimes known as Young's inequality for convolutions. See for instance [3, p. 205] for a proof.

(ii) Assume $f \in BC^k(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$. Lebesgue's dominated convergence theorem allows us to differentiate under the integral sign of the convolution, since the derivatives of f are bounded and g is integrable, thus we have for $0 \le |\alpha| \le k$

$$\begin{aligned}|D_x^\alpha (f * g)(x)| &\le \int_{\mathbb{R}^d} |D_x^\alpha f(x-y) g(y)|\, dy \\ &\le \int_{\mathbb{R}^d} \|f\|_{BC^k(\mathbb{R}^d)} |g(y)|\, dy = \|f\|_{BC^k(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Hence $\|f * g\|_{BC^k(\mathbb{R}^d)} = \sum_{|\alpha|\le k} \sup_{x\in\mathbb{R}^d} |D^\alpha (f * g)(x)|$ is finite.

(iii) Assume $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$. Clearly we have $\varphi * \psi \in C^\infty(\mathbb{R}^d)$. By the triangle inequality and the binomial theorem

$$(1+|x|)^k \le (1+|x-y|+1+|y|)^k = \sum_{j=0}^{k} \binom{k}{j} (1+|x-y|)^j (1+|y|)^{k-j}.$$

Using this and $D_x^\beta(\varphi * \psi) = (D_x^\beta \varphi) * \psi$, we get

$$\begin{aligned}(1+|x|)^{|\alpha|} \left|D_x^\beta(\varphi * \psi)(x)\right| &\le \int_{\mathbb{R}^d} (1+|x|)^{|\alpha|} |D_x^\beta \varphi(x-y)||\psi(y)|\, dy| \\ \le \int_{\mathbb{R}^d} \sum_{j=0}^{|\alpha|} \binom{|\alpha|}{j} &(1+|x-y|)^j (1+|y|)^{|\alpha|-j} |D_x^\beta \varphi(x-y)||\psi(y)|\, dy \\ = \sum_{j=0}^{|\alpha|} \binom{|\alpha|}{j} &\left((1+|\cdot|)^j |D^\beta \varphi| * (1+|\cdot|)^{\alpha-j} |\psi|\right)(x) < \infty.\end{aligned} \tag{2.4}$$

That the sum is finite follows from φ and ψ and their derivatives decreasing faster than any polynomial. This implies that $P_{\alpha,\beta}(\varphi * \psi)$ is finite and thus $\psi * \varphi \in \mathcal{S}(\mathbb{R}^d)$. Since $P_{\alpha,\beta}(\varphi * \psi)$ can be estimated by a linear combination of seminorms of φ and ψ as per (2.4), the convolution is also continuous $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$.

□

Now that we know the Schwartz space is closed under convolution, we can prove the following very useful result:

Theorem 2.5 (Convolution theorem). *If $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$, then*

$$\mathcal{F}(\varphi * \psi) = (2\pi)^{\frac{d}{2}} \hat{\varphi}\, \hat{\psi}, \tag{2.5}$$

and

$$\hat{\varphi} * \hat{\psi} = (2\pi)^{\frac{d}{2}} \mathcal{F}(\varphi\psi). \tag{2.6}$$

Proof. Since $\mathcal{S}(\mathbb{R}^d)$ is closed under convolution, we may take the Fourier transform of $\varphi * \psi$, so by applying Fubini's theorem we get

$$\begin{aligned}\mathcal{F}(\varphi * \psi)(\xi) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(y)\psi(x-y)\,dy \right) e^{-ix\cdot\xi}\,dx \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y) \left(\int_{\mathbb{R}^d} \psi(x-y)e^{-ix\cdot\xi}\,dx \right) dy \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y)e^{-iy\cdot\xi} \left(\int_{\mathbb{R}^d} \psi(x-y)e^{-i(x-y)\cdot\xi}\,dx \right) dy \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y)e^{-iy\cdot\xi} \left(\int_{\mathbb{R}^d} \psi(z)e^{-iz\cdot\xi}\,dz \right) dy \\ &= (2\pi)^{\frac{d}{2}}\hat{\varphi}(\xi)\hat{\psi}(\xi).\end{aligned}$$

The proof of the second identity is similar. □

Remark 2.6. From the proof of Theorem 2.5 and Lemma 2.4 we see that (2.5) also holds for $f, g \in L^1(\mathbb{R}^d)$ (clearly the Fourier transform of an integrable function is well-defined, even though so far we have only talked about the Fourier transform of Schwartz functions).

The Schwartz space is also closed under Fourier transform:

Theorem 2.6. *The Fourier transform is a continuous linear map from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$. The same is true for the inverse Fourier transform.*

Proof. The linearity of $\mathcal{F}$ is clear from the definition. We shall prove that $\mathcal{F}$ maps Schwartz functions to Schwartz functions: In other words,

$$\sup_{\xi\in\mathbb{R}^d} |\xi^\alpha D^\beta \hat{\varphi}(\xi)| < \infty \qquad \forall \alpha, \beta \in \mathbb{N}_0^d,$$

for $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

We investigate $|\xi^\alpha D_\xi^\beta \hat{\varphi}(\xi)|$, using the fact that the product of a Schwartz function and a complex exponential is again a Schwartz function. This means the product and its derivatives are smooth and integrable, and we can differentiate under the integral sign:

$$\begin{aligned}
|\xi^\alpha D_\xi^\beta \hat{\varphi}(\xi)| &= \left| \frac{\xi^\alpha D_\xi^\beta}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x) e^{-ix\cdot\xi}\, dx \right| \\
&= \left| \frac{(-i)^{|\beta|}\xi^\alpha}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} x^\beta \varphi(x) e^{-ix\cdot\xi}\, dx \right| \\
&= \left| \frac{(-i)^{|\beta|-|\alpha|}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} D_x^\alpha (x^\beta \varphi(x)) e^{-ix\cdot\xi}\, dx \right| \\
&\le \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |D_x^\alpha (x^\beta \varphi(x))|\, dx \\
&\le \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \sum_{|\gamma|\le|\alpha|} \left| \binom{\alpha}{\gamma} (D_x^\gamma x^\beta) D_x^{\alpha-\gamma} \varphi(x) \right| dx < \infty
\end{aligned} \tag{2.7}$$

In the third equality we have used $D_x^\alpha e^{-ix\cdot\xi} = (-i)^{|\alpha|}\xi^\alpha e^{-ix\cdot\xi}$ and integrated by parts. The integral converges due to φ being a Schwartz function. This means $\hat{\varphi}(\xi) \in \mathcal{S}(\mathbb{R}^d)$ and so $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$. Since $\mathcal{F}^{-1}(\varphi)(\xi) = \mathcal{F}(\varphi)(-\xi)$, the same is true for the inverse Fourier transform.

Since both $\mathcal{F}$ and $\mathcal{F}^{-1}$ are linear, continuity now follows from (2.7). □

When we later show that $\mathcal{F}^{-1}$ is indeed the inverse of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^d)$, it will be clear that these functions are in fact bijections on the Schwartz space.

The following very practical equality will be used frequently:

Lemma 2.7. *For φ, $\psi \in \mathcal{S}(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} \varphi(x)\hat{\psi}(x)\, dx = \int_{\mathbb{R}^d} \hat{\varphi}(x)\psi(x)\, dx.$$

Proof. We may apply Fubini's theorem as φ and ψ are integrable:

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x)\hat{\psi}(x)\, dx &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x) \left(\int_{\mathbb{R}^d} \psi(\xi) e^{-ix\cdot\xi}\, d\xi \right) dx \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x)\psi(\xi) e^{-ix\cdot\xi}\, d\xi\, dx \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(x) e^{-ix\cdot\xi}\, dx \right) \psi(\xi)\, d\xi \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \hat{\varphi}(\xi)\psi(\xi)\, d\xi
\end{aligned}$$

□

We will need the next lemma to prove the important Fourier inversion theorem.

Lemma 2.8 (Fourier transform of a Gaussian). *The Gaussian function* $x \mapsto e^{-\frac{\lambda^2}{2}|x|^2} \in \mathcal{S}(\mathbb{R}^d)$, *with* $\lambda > 0$, *has Fourier transform*

$$\mathcal{F}\left(e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) = \frac{1}{\lambda^d} e^{-\frac{|\xi|^2}{2\lambda^2}}.$$

Proof. Firstly,

$$\sum_{i=1}^{d} \frac{\partial}{\partial x_i} e^{-\frac{\lambda^2}{2}|x|^2} = -\lambda^2 \sum_{i=1}^{d} x_i e^{-\frac{\lambda^2}{2}|x|^2},$$

meaning that

$$\sum_{i=1}^{d} (\frac{\partial}{\partial x_i} + \lambda^2 x_i) e^{-\frac{\lambda^2}{2}|x|^2} = 0. \tag{2.8}$$

We may take the Fourier transform of the left hand side of (2.8) as the Gaussian is a Schwartz function (it has exponential decay), and of course $\mathcal{F}(0) = 0$, so

$$\begin{aligned}
0 &= \mathcal{F}\left(\sum_{i=1}^{d} (\frac{\partial}{\partial x_i} + \lambda^2 x_i) e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) \\
&= \mathcal{F}\left(\sum_{i=1}^{d} \frac{\partial}{\partial x_i} e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) + \mathcal{F}\left(\sum_{i=1}^{d} \lambda^2 x_i e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) \\
&= i\left(\sum_{i=1}^{d} \xi_i\right) \mathcal{F}\left(e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) + \lambda^2 i \left(\sum_{i=1}^{d} \frac{\partial}{\partial \xi_i}\right) \mathcal{F}\left(e^{-\frac{\lambda^2}{2}|x|^2}\right)(\xi) \\
&= i\left(\sum_{i=1}^{d} (\lambda^2 \frac{\partial}{\partial \xi_i} + \xi_i) \mathcal{F}\left(e^{-\frac{\lambda^2}{2}|x|^2}\right)\right),
\end{aligned}$$

where we have employed properties (i) and (ii) of Theorem 2.3 in going from the second to the third line. In addition,

$$\mathcal{F}\left(e^{-\frac{\lambda^2}{2}|x|^2}\right)(0) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{\lambda^2}{2}|x|^2}\, dx = \frac{1}{\lambda^d}$$

so we have the initial value problem

$$\begin{cases} \sum_{i=1}^{d} (\lambda^2 \frac{\partial}{\partial \xi_i} + \xi_i) f = 0 & \text{for } \xi \in \mathbb{R}^d, \\ f(0) = \frac{1}{\lambda^d}, \end{cases}$$

with unique solution

$$f(\xi) = \frac{1}{\lambda^d} e^{-\frac{|\xi|^2}{2\lambda^2}},$$

according to the Picard-Lindelöf theorem [13, p. 96-100]. □

Knowing this we are able to prove the following:

Theorem 2.9 (Fourier inversion theorem)**.** *The Fourier transform is a bijection on* $\mathcal{S}(\mathbb{R}^d)$*: For* $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$\mathcal{F}^{-1}\mathcal{F}\varphi = \mathcal{F}\mathcal{F}^{-1}\varphi = \varphi.$$

Proof. By assumption the integral

$$\begin{aligned}\mathcal{F}^{-1}(\hat{\varphi})(x) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{\varphi}(\xi) e^{ix\cdot\xi}\, d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(y) e^{-iy\cdot\xi}\, dy \right) e^{ix\cdot\xi}\, d\xi\end{aligned}$$

exists for all $x \in \mathbb{R}^d$ since $\mathcal{F}$ and $\mathcal{F}^{-1}$ map Schwartz functions to Schwartz functions. Inserting a Gaussian to make the integral absolutely convergent we get

$$\begin{aligned}\varphi_\lambda(x) &:= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{\varphi}(\xi) e^{ix\cdot\xi} e^{-\frac{\lambda^2}{2}|\xi|^2}\, d\xi \\ &\to \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{\varphi}(\xi) e^{ix\cdot\xi}\, d\xi \quad \text{as } \lambda \searrow 0,\end{aligned}$$

by Lebesgue's dominated convergence theorem (the integrand is dominated by the absolute value of its own limit, which is integrable since $\hat{\varphi}$ is Schwartz). Changing the order of integration and applying the previous lemma we get

$$\begin{aligned}\varphi_\lambda(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(y) e^{i(x-y)\cdot\xi} e^{-\frac{\lambda^2}{2}|\xi|^2}\, dy\, d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi(y) \left(\int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-\frac{\lambda^2}{2}|\xi|^2} e^{-iy\cdot\xi}\, d\xi \right) dy \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y) \mathcal{F}_\xi (e^{ix\cdot\xi} e^{-\frac{\lambda^2}{2}|\xi|^2})(y)\, dy \\ &= \frac{1}{\lambda^d (2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y) e^{-\frac{|y-x|^2}{2\lambda^2}}\, dy,\end{aligned}$$

by property (iii) of Theorem 2.3. Making the change of variables $s := \frac{y-x}{\sqrt{2}\lambda}$ gives

$$\begin{aligned}\frac{1}{\lambda^d (2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(y) e^{-\frac{|y-x|^2}{2\lambda^2}}\, dy &= \frac{(\sqrt{2}\lambda)^d}{\lambda^d (2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x + \sqrt{2}\lambda s) e^{-|s|^2}\, ds \\ &\to \frac{1}{\pi^{d/2}} \int_{\mathbb{R}^d} \varphi(x) e^{-|s|^2}\, ds = \varphi(x)\end{aligned}$$

as $\lambda \searrow 0$, by Lebesgue's dominated convergence theorem (since φ is bounded and continuous). Hence $(2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\varphi}(\xi) e^{ix\cdot\xi}\, d\xi = \varphi(x)$.

That $\mathcal{F}^{-1}\mathcal{F}\varphi = \mathcal{F}\mathcal{F}^{-1}\varphi = \varphi$ then follows from $(\mathcal{F}^{-1}\varphi)(y) = (\mathcal{F}\varphi)(-y)$. □

This concludes our study of the Fourier transform on the Schwartz space. It is natural to ask whether there exists any larger space which the Fourier transform bijectively maps onto itself. In the following section we show that one can define the Fourier transform of a square integrable function in a natural way such that $\mathcal{F}$ is in fact a bijection on $L^2(\mathbb{R}^d)$.

2.3 Extension of $\mathcal{F}$ to $L^2(\mathbb{R}^d)$

An immediate difficulty when talking about the Fourier transform of a general $L^2(\mathbb{R}^d)$-function, is that the integral appearing in the Fourier transform may not converge. However, in the next lemma we show that the Schwartz space is dense in $L^2(\mathbb{R}^d)$. Using this we can define the Fourier transform on $L^2(\mathbb{R}^d)$ by extension from $\mathcal{S}(\mathbb{R}^d)$.

Lemma 2.10. $\mathcal{S}(\mathbb{R}^d)$ *is densely embedded in* $L^2(\mathbb{R}^d)$,

$$\mathcal{S}(\mathbb{R}^d) \overset{d}{\hookrightarrow} L^2(\mathbb{R}^d).$$

Proof. The inclusion map between $\mathcal{S}(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$ is continuous since for $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}
\|\varphi\|_{L^2(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d} |\varphi(x)|^2\, dx\right)^{\frac{1}{2}} = \left(\int_{\mathbb{R}^d} \left|(1+|x|)^{-(d+1)}(1+|x|)^{d+1}\varphi(x)\right|^2\, dx\right)^{\frac{1}{2}} \\
&\leq \left(\sup_{x\in\mathbb{R}^d}\left((1+|x|)^{2(d+1)}|\varphi(x)|^2\right)\int_{\mathbb{R}^d}\frac{1}{(1+|x|)^{2(d+1)}}\,dx\right)^{\frac{1}{2}} \\
&=: C\left(\sup_{x\in\mathbb{R}^d}\left((1+|x|)^{2(d+1)}|\varphi(x)|^2\right)\right)^{\frac{1}{2}} \\
&\leq C\sup_{x\in\mathbb{R}^d}\left((1+|x|)^{d+1})|\varphi(x)|\right) \\
&= CP_{\alpha,0},
\end{aligned}$$

where α is s.t. $|\alpha| = d+1$. The inclusion $\mathcal{S}(\mathbb{R}^d) \subseteq L^2(\mathbb{R}^d)$ also follows from this estimate. Since $C_c^\infty(\mathbb{R}^d) \subseteq \mathcal{S}(\mathbb{R}^d)$, and $C_c^\infty(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$ (cf. [39, p. 373]), $\mathcal{S}(\mathbb{R}^d)$ is also dense in $L^2(\mathbb{R}^d)$. □

Using this result we can define the Fourier transform of $L^2(\mathbb{R}^d)$ in a natural limiting sense (details in the following proof) to get the following result for $\mathcal{F}$ and $\mathcal{F}^{-1}$ as bijective operators on $L^2(\mathbb{R}^d)$:

Theorem 2.11 (Plancherel's theorem)**.** *The Fourier transform extends to a unitary operator on* $L^2(\mathbb{R}^d)$. *In particular,*

$$\langle \hat{f}, \hat{g}\rangle_{L^2(\mathbb{R}^d)} = \langle f, g\rangle_{L^2(\mathbb{R}^d)} \tag{2.9}$$

for $f, g \in L^2(\mathbb{R}^d)$*.*

Proof. For $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\|\hat{\varphi}\|_{L^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \varphi(x)\mathcal{F}\left(\overline{\mathcal{F}\varphi}\right)(x)\,dx$$

by Lemma 2.7. From the Fourier inversion theorem we get

$$\begin{aligned}\overline{\mathcal{F}(\varphi)(\xi)} &= \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} \overline{\varphi(x)e^{-ix\cdot\xi}}\,dx \\ &= \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} \overline{\varphi(x)}e^{ix\cdot\xi}\,dx \\ &= \mathcal{F}^{-1}(\overline{\varphi})(\xi),\end{aligned}$$

which means $\mathcal{F}(\overline{\mathcal{F}\varphi}) = \overline{\varphi}$ and so

$$\|\hat{\varphi}\|^2_{L^2(\mathbb{R}^d)} = \|\varphi\|^2_{L^2(\mathbb{R}^d)} \quad \text{for } \varphi \in \mathcal{S}(\mathbb{R}^d). \tag{2.10}$$

We now wish to extend this result to general $f \in L^2(\mathbb{R}^d)$. Because the Schwartz space is dense in $L^2(\mathbb{R}^d)$ by Lemma 2.10, we can define the Fourier transform on $L^2(\mathbb{R}^d)$ by extension from $\mathcal{S}(\mathbb{R}^d)$ in the following way: Given an $f \in L^2(\mathbb{R}^d)$, consider a sequence $\{\varphi_n\}_n \subseteq \mathcal{S}(\mathbb{R}^d)$ with $\varphi_n \to f$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$. Then $\{\varphi_n\}_n$ is Cauchy in $L^2(\mathbb{R}^d)$, and from the identity (2.10)

$$\|\hat{\varphi}_n - \hat{\varphi}_m\|^2_{L^2(\mathbb{R}^d)} = \|\varphi_n - \varphi_m\|^2_{L^2(\mathbb{R}^d)} \to 0 \quad \text{as } n, m \to \infty,$$

i.e. $\{\hat{\varphi}_n\}_n$ is also Cauchy in $L^2(\mathbb{R}^d)$. From the completeness of $L^2(\mathbb{R}^d)$ we then know there exists an $\hat{f} \in L^2(\mathbb{R}^d)$ such that $\mathcal{F}(\varphi_n) \to \hat{f}$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$. We define this function $\hat{f}$ to be the Fourier transform of $f \in L^2(\mathbb{R}^d)$, and we have

$$\|f\|_{L^2(\mathbb{R}^d)} = \lim_{n\to\infty} \|\varphi_n\|_{L^2(\mathbb{R}^d)} = \lim_{n\to\infty} \|\hat{\varphi}_n\|_{L^2(\mathbb{R}d)} = \|\hat{f}\|_{L^2(\mathbb{R}^d)}.$$

That $\langle\varphi, \psi\rangle_{L^2(\mathbb{R}^d)} = \langle\hat{\varphi}, \hat{\psi}\rangle_{L^2(\mathbb{R}^d)}$ for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ follows by a similar argument to the one leading to (2.10). Then another limiting argument shows the identity of the theorem also holds for $f, g \in L^2(\mathbb{R}^d)$. □

Remark 2.7. Clearly this implies that $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous on $L^2(\mathbb{R}^d)$, since they are bounded linear operators on $L^2(\mathbb{R}^d)$. Moreover, by definition and Theorem 2.9, they are bijections on $L^2(\mathbb{R}^d)$.

We are now are ready to introduce the fractional Sobolev spaces of non-negative order.

2.4 Fractional Sobolev spaces $H^s(\mathbb{R}^d)$ for $s \geq 0$

Definition 2.10 (Fractional Sobolev spaces). We define for $s \geq 0$ the fractional Sobolev space

$$H^s(\mathbb{R}^d) = \{f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d} (1+|\xi|^2)^s |\mathcal{F}(f)(\xi)|^2 \, d\xi < \infty\}, \tag{2.11}$$

equipped with the Sobolev norm

$$\|f\|_{H^s(\mathbb{R}^d)} = \|(1+|\cdot|^2)^{\frac{s}{2}} \mathcal{F}(f)\|_{L^2(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} (1+|\xi|^2)^s |\hat{f}(\xi)|^2 \, d\xi \right)^{\frac{1}{2}},$$

and the Sobolev inner product

$$\begin{aligned} \langle f, g \rangle_{H^s(\mathbb{R}^d)} &= \langle (1+|\cdot|^2)^{\frac{s}{2}} \hat{f}, (1+|\cdot|^2)^{\frac{s}{2}} \hat{g} \rangle_{L^2(\mathbb{R}^d)} \\ &= \int_{\mathbb{R}^d} (1+|\xi|^2)^s \hat{f}(\xi) \overline{\hat{g}(\xi)} \, d\xi \end{aligned}$$

Remark 2.8. Each of the properties that $\|\cdot\|_{H^s(\mathbb{R}^d)}$ should satisfy in order to be a norm follow from the corresponding properties of $\|\cdot\|_{L^2(\mathbb{R}^d)}$. The same is true for the inner product. In Theorem 2.14 we show that $H^s(\mathbb{R}^d)$ is a Hilbert space.

Remark 2.9. Clearly $H^{s_1}(\mathbb{R}^d) \subseteq H^{s_2}(\mathbb{R}^d)$ for $s_1 \geq s_2$.

The motivation behind this definition comes from a connection between the convergence of the integral in the Sobolev norm and the existence of weak derivatives of order s in $L^2(\mathbb{R}^d)$ in the case where s is a natural number:

Lemma 2.12 (Fourier version of weak differentiability). *Let $n \in \mathbb{N}$. If*

$$(1+|\cdot|^2)^{\frac{n}{2}} \hat{f} \in L^2(\mathbb{R}^d),$$

then the function $f \in L^2(\mathbb{R}^d)$ is n times weakly differentiable with weak derivatives $D^\alpha f \in L^2(\mathbb{R}^d)$ for $|\alpha| \leq n$ such that

$$\mathcal{F}(D^\alpha f) = i^{|\alpha|} \xi^\alpha \hat{f}(\xi) \in L^2(\mathbb{R}^d).$$

Proof. Firstly, since $\hat{f} \in L^2(\mathbb{R}^d)$, by Theorem 2.11 there exists an $f \in L^2(\mathbb{R}^d)$ and a sequence $\{\varphi_j\}_j \subseteq \mathcal{S}(\mathbb{R}^d)$ such that

$$\lim_{j\to\infty} \|f - \varphi_j\|^2_{L^2(\mathbb{R}^d)} = \lim_{j\to\infty} \|\hat{f} - \hat{\varphi}_j\|^2_{L^2(\mathbb{R}^d)} = 0.$$

Moreover, we have $|\xi^\alpha| \hat{f}(\xi) \in L^2(\mathbb{R}^d)$ for α s.t. $|\alpha| \leq n$ since $|\xi^\alpha| \sim |\xi|^{|\alpha|}$ (meaning $\exists\, m, M \in \mathbb{R}$ s.t. $m|\xi^\alpha| \leq |\xi|^{|\alpha|} \leq M|\xi^\alpha|$). We claim that

$$\lim_{j\to\infty} \int_{\mathbb{R}^d} |\xi^\alpha|^2 |\hat{f}(\xi) - \hat{\varphi}_j(\xi)|^2 \, d\xi = 0$$

for $|\alpha| \le n$: Suppose $\lim_{j\to\infty} \int_{\mathbb{R}^d} |\xi^\alpha|^2 |\hat{f}(\xi) - \hat{\varphi}_j(\xi)|^2 \, d\xi = I$ for some $\alpha \in \mathbb{N}_0^d$ s.t. $|\alpha| \le n$. We know $I < \infty$ since $|\xi^\alpha|\hat{f}(\xi) \in L^2(\mathbb{R}^d)$ and $\hat{\varphi}_j$ is Schwartz. On a bounded subset Ω of $\mathbb{R}^d$ s.t. $\sup_{\xi\in\Omega} |\xi^\alpha| = L^{\frac{1}{2}}$, we have

$$\lim_{j\to\infty} \int_\Omega |\xi^\alpha|^2 |\hat{f}(\xi) - \hat{\varphi}_j(\xi)|^2 \, d\xi \le L \lim_{j\to\infty} \int_\Omega |\hat{f}(\xi) - \hat{\varphi}_j(\xi)|^2 \, d\xi = 0$$

by $\hat{\varphi}_j \to \hat{f}$ in $L^2(\Omega)$ for any $\Omega \subseteq \mathbb{R}^d$. Thus for all bounded subsets Ω of $\mathbb{R}^d$,

$$\lim_{j\to\infty} \int_{\mathbb{R}^d \setminus \Omega} |\xi^\alpha|^2 |\hat{f}(\xi) - \hat{\varphi}_j(\xi)|^2 \, d\xi = I.$$

For the integral over $\mathbb{R}^d$ to be finite, the integral over the tails must go to zero, hence $I = 0$. This means $i^{|\alpha|}\xi^\alpha \hat{\varphi}_j \to i^{|\alpha|}\xi^\alpha \hat{f}$ in $L^2(\mathbb{R}^d)$.

Since $\varphi_j \in \mathcal{S}(\mathbb{R}^d)$, we have for all test functions $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi_j(x) D^\alpha \psi(x) \, dx &= (-1)^{|\alpha|} \int_{\mathbb{R}^d} \psi(x) D^\alpha \varphi_j(x) \, dx \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^d} \psi(x) \mathcal{F}^{-1} \left(i^{|\alpha|} \xi^\alpha \hat{\varphi}_j(\xi) \right)(x) \, dx\end{aligned}$$

by property (i) of Theorem 2.3. Then since $\mathcal{F}^{-1}$ is continuous on $L^2(\mathbb{R}^d)$, taking the limit $j \to \infty$ yields that

$$\int_{\mathbb{R}^d} f(x) D^\alpha \psi(x) \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} \psi(x) \mathcal{F}^{-1} \left(i^{|\alpha|} \xi^\alpha \hat{f} \right)(x) \, dx.$$

□

This lemma implies the following:

Theorem 2.13. *If $s = k \in \mathbb{N}_0$, the fractional Sobolev space $H^s(\mathbb{R}^d)$ coincides with the classical Sobolev space $W_2^k(\mathbb{R}^d)$.*

Proof. Recall the definition of $W_2^k(\mathbb{R}^d)$,

$$W_2^k(\mathbb{R}^d) = \{ f \in L^2(\mathbb{R}^d) \mid D^\alpha f \in L^2(\mathbb{R}^d) \, \forall \alpha \in \mathbb{N}_0^d, |a| \le k \}.$$

By Definition 2.10, given a function $f \in H^k(\mathbb{R}^d)$ we have $(1 + |\cdot|^2)^{k/2} \hat{f} \in L^2(\mathbb{R}^d)$. Then by the previous lemma we have $D^\alpha f \in L^2(\mathbb{R}^d)$ for all $|\alpha| \le k$, thus $f \in W_2^k(\mathbb{R}^d)$.

Given instead $g \in W_2^k(\mathbb{R}^d)$, we have for all $|\alpha| \le k$

$$\int_{\mathbb{R}^d} |D^\alpha g(x)|^2 \, dx = \int_{\mathbb{R}^d} |\mathcal{F}(D^\alpha g)(\xi)|^2 \, d\xi = \int_{\mathbb{R}^d} |\xi^\alpha|^2 |\hat{g}(\xi)|^2 \, d\xi < \infty.$$

Thus

$$\int_{\mathbb{R}^d} (1 + |\xi|^2)^{|\alpha|} |\hat{g}(\xi)|^2 \, d\xi \le 2^{|\alpha|} \int_{|\xi|\le 1} |\hat{g}(\xi)|^2 \, d\xi + C_\alpha \int_{|\xi|>1} |\xi^\alpha|^2 |\hat{g}(\xi)|^2 \, d\xi < \infty,$$

for some constant C_α since $(1 + |\xi|^2)^{|\alpha|} \sim |\xi^\alpha|^2$ for large ξ, so indeed $(1 + |\cdot|^2)^{|\alpha|/2} \hat{g} \in L^2(\mathbb{R}^d)$. □

Remark 2.10. We see that the fractional Sobolev spaces $H^s(\mathbb{R}^d)$ for non-negative real s in some sense interpolate between the classical Sobolev spaces $W_2^k(\mathbb{R}^d)$, and thus could be useful in defining the notion of a fractional weak derivative.

Like the classical Sobolev spaces, the fractional Sobolev spaces are complete:

Theorem 2.14. *The fractional Sobolev spaces $H^s(\mathbb{R}^d)$ for $s \geq 0$ are Hilbert spaces.*

Proof. The map $f \mapsto (1+|\cdot|^2)^{\frac{s}{2}}\mathcal{F}f$ is isometric $H^s(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. It is also onto, and therefore bijective, since for every $g \in L^2(\mathbb{R}^d)$, $\mathcal{F}^{-1}((1+|\cdot|^2)^{-\frac{s}{2}}g) \in H^s(\mathbb{R}^d)$. Thus $H^s(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$ are isometrically isomorphic, and thus it follows that $H^s(\mathbb{R}^d)$ is also a Hilbert space. One may see this in the following way: Consider a Cauchy sequence $\{f_n\}_n$ in $H^s(\mathbb{R}^d)$. Then $\{(1+|\cdot|^2)^{\frac{s}{2}}\mathcal{F}(f_n)\}_n$ is Cauchy in $L^2(\mathbb{R}^d)$, and so there exists a $g \in L^2(\mathbb{R}^d)$ such that $(1+|\cdot|^2)^{\frac{s}{2}}\mathcal{F}(f_n) \to g$ in $L^2(\mathbb{R}^d)$, which by the above implies $f_n \to \mathcal{F}^{-1}((1+|\cdot|^2)^{-\frac{s}{2}}g)$ in $H^s(\mathbb{R}^d)$. □

Our goal for the rest of this section is to derive two fundamental properties of the fractional Sobolev spaces that will come in handy when we later prove well-posedness results for PDEs in Sobolev spaces. Specifically, we will establish when a function $f \in H^s(\mathbb{R}^d)$ is also in $BC^k(\mathbb{R}^d)$, and when the pointwise product of two functions $f, g \in H^s(\mathbb{R}^d)$ is also in $H^s(\mathbb{R}^d)$.

2.5 Two important properties of Sobolev spaces

While a space such as $BC^k(\mathbb{R}^d)$ is truly a function space, $H^s(\mathbb{R}^d)$ is strictly speaking a space of equivalence classes of functions, where each equivalence class consists of functions that differ only on a set of measure zero. We are interested in under which conditions one can find a representative function $f \in BC^k(\mathbb{R}^d)$ of the equivalence class $[f] \in H^s(\mathbb{R}^d)$. We will need the following lemma:

Lemma 2.15 (Approximation by smooth functions)**.** *The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$ for $s \geq 0$.*

Proof. Suppose $f \in H^s(\mathbb{R}^d)$. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, there is a sequence $\{\psi_n\}_n \subseteq C_c^\infty(\mathbb{R}^d)$ such that $\psi_n \to (1+|\cdot|^2)^{s/2}\hat{f}$ in $L^2(\mathbb{R}^d)$. Let $\varphi_n := (1+|\cdot|^2)^{-s/2}\psi_n$, then $\varphi_n \in C_c^\infty(\mathbb{R}^d)$, and so $\mathcal{F}^{-1}(\varphi_n) \in \mathcal{S}(\mathbb{R}^d)$. We then have

$$F^{-1}(\varphi_n) \to f \quad \text{in } H^s(\mathbb{R}^d).$$

□

Remark 2.11. Although we haven't yet defined the Sobolev spaces of negative order, we point out that the above proof holds also for $s < 0$, hence $\mathcal{S}(\mathbb{R}^d)$ is in fact densely embedded in $H^s(\mathbb{R}^d)$ for $s \in \mathbb{R}$. For $s \geq 0$, one can actually show that $C_c^\infty(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$, see for example [17, p. 66].

Now that we have established that the Schwartz space is dense in $H^s(\mathbb{R}^d)$, we can prove the first main result of this section:

Theorem 2.16 (Fractional Sobolev embedding theorem). *Given $k \in \mathbb{N}_0$, one has*

$$H^s(\mathbb{R}^d) \hookrightarrow BC^k(\mathbb{R}^d),$$

for $s > k + \frac{d}{2}$, meaning that for such an s we can in each equivalence class $[f] \in H^s(\mathbb{R}^d)$ find a representative function $f \in BC^k(\mathbb{R}^d)$. In fact, one has

$$\|f\|_{BC^k(\mathbb{R}^d)} \leq C\|[f]\|_{H^s(\mathbb{R}^d)}.$$

for some constant C depending only on s and k.

Proof. We give a proof similar to the one found in [17, p. 71]: Since $\mathcal{S}(\mathbb{R}^d)$ is densely and continuously embedded in $H^s(\mathbb{R}^d)$, it will be sufficient to prove that there is some constant $C > 0$ such that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$|D^\alpha \varphi(x)| \leq C\|\varphi\|_{H^s(\mathbb{R}^d)}, \quad |\alpha| \leq k,\ x \in \mathbb{R}^d,$$

and the result will follow from continuity.

The Cauchy-Schwarz inequality, together with the fact that $|\xi^\alpha| \sim (1+|\xi|^2)^{|\alpha|/2}$ for large $|\xi|$, imply

$$\begin{aligned}
|D^\alpha \varphi(x)| &= |D^\alpha(\mathcal{F}^{-1}\mathcal{F}\varphi)(x)| = |\mathcal{F}^{-1}(\xi^\alpha \mathcal{F}\varphi(\xi))(x)| \\
&= \frac{1}{\sqrt{2\pi}} \Big| \int_{\mathbb{R}^d} e^{ix\cdot\xi} \xi^\alpha \hat{\varphi}(\xi)\, d\xi \Big| \\
&\leq c' \int_{\mathbb{R}^d} (1+|\xi|^2)^{\frac{s}{2}} |\hat{\varphi}(\xi)| (1+|\xi|^2)^{\frac{k-s}{2}}\, d\xi \\
&\leq c' \left(\int_{\mathbb{R}^d} (1+|\xi|^2)^s |\hat{\varphi}(\xi)|^2\, d\xi \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} \frac{1}{(1+|\xi|^2)^{s-k}}\, d\xi \right)^{\frac{1}{2}} \\
&= c' \|\varphi\|_{H^s(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} \frac{1}{(1+|\xi|^2)^{s-k}}\, d\xi \right)^{\frac{1}{2}} = C\|\varphi\|_{H^s(\mathbb{R}^d)}
\end{aligned}$$

The last integral converges since $s - k > d/2$. □

We now set out to prove the second main result of this section, which will answer the question of when the product of two functions $f, g \in H^s(\mathbb{R}^d)$ is also in $H^s(\mathbb{R}^d)$. We will need the following lemma concerning the integrability of the Fourier transform of a function in $H^s(\mathbb{R}^d)$:

Lemma 2.17. *If $f \in H^s(\mathbb{R}^d)$ for $s > d/2$, then $\hat{f} \in L^1(\mathbb{R}^d)$. In fact,*

$$\|\hat{f}\|_{L^1(\mathbb{R}^d)} \leq C_s \|f\|_{H^s(\mathbb{R}^d)}$$

where $C_s = \left(\int_{\mathbb{R}^d} \langle\xi\rangle^{-2s}\, d\xi \right)^{1/2}$.

Proof. By Cauchy-Schwarz,

$$\|\hat{f}\|_{L^1(\mathbb{R}^d)} = \| \langle\cdot\rangle^{-s} \langle\cdot\rangle^{s} \hat{f}\|_{L^1(\mathbb{R}^d)} \leq \| \langle\cdot\rangle^{-s} \|_{L^2(\mathbb{R}^d)} \|f\|_{H^s(\mathbb{R}^d)}.$$

As in the proof of Theorem 2.16, the integral appearing in $\| \langle\cdot\rangle^{-s} \|_{L^2(\mathbb{R})}$ converges for $s > d/2$. □

With this we can prove the very useful fact that $H^s(\mathbb{R}^d)$ for $s > d/2$ is a *Banach algebra*:

Definition 2.11 (Banach algebra)**.** A Banach algebra is a Banach space X paired with a product $(x, y) \in X \times Y \mapsto xy \in X$ such that, for all $x, y, z \in X$ and for all $s, r \in \mathbb{C}$,

(i) $(xy)z = x(yz)$,

(ii) $r(xy) = (rx)y = x(ry)$,

(iii) $(x + y)z = xz + yz$ and $x(y + z) = xy + xz$,

(iv) $\|xy\| \leq \|x\|\|y\|$.

The final property is the only nontrivial one for $H^s(\mathbb{R}^d)$.

Theorem 2.18. *The space $H^s(\mathbb{R}^d)$ for $s > d/2$ is closed under multiplication, and for $f, g \in H^s(\mathbb{R}^d)$ we have*

$$\|fg\|_{H^s(\mathbb{R}^d)} \leq c_s\|f\|_{H^s(\mathbb{R}^d)}\|g\|_{H^s(\mathbb{R}^d)}, \tag{2.12}$$

where c_s depends only on s. In other words, $H^s(\mathbb{R}^d)$ is a Banach algebra for $s > d/2$.

Proof. We take an approach similar to the one in [29, p. 49], and make use of the following elementary inequality:

$$\left(1 + |x|^2\right)^t \leq 2^{2t}\left(1 + |x - y|^2\right)^t + 2^{2t}\left(1 + |y|^2\right)^t \tag{2.13}$$

for any $t \in [0, \infty)$, $x, y \in \mathbb{R}^d$. We want to prove this inequality using a variational approach, therefore we fix $x \in \mathbb{R}^d$ and minimise the right hand side with respect to y: Define $y =: cx + z$, where $z \in \mathbb{R}^d$ is such that $x \cdot z = 0$ and $c \in \mathbb{R}$. Then the right hand side in (2.13) is equal to

$$\begin{aligned} 2^{2t}\left(1 + (1 - c)^2|x|^2 + |z|^2\right)^t + 2^{2t}\left(1 + |z|^2 + c^2|x|^2\right)^t \\ \geq 2^{2t}\left(1 + (1 - c)^2|x|^2\right)^t + 2^{2t}\left(1 + c^2|x|^2\right)^t, \end{aligned} \tag{2.14}$$

where we have set $z = 0$ to get the inequality. This expression is clearly greater than the left hand side of (2.13) for $c \geq 1$ and $c \leq 0$, thus we assume $c \in (0,1)$. In fact, since as a function of c the expression is symmetric about $c = 1/2$, we assume $c \in (0, 1/2]$. For such c, we have $(1-c)^2 \geq 1/4$. Thus, by dropping the final term in (2.14), we finally get the desired result:

$$\begin{aligned}2^{2t}\left(1+|x-y|^2\right)^t + 2^{2t}\left(1+|y|^2\right)^t &\geq 2^{2t}\left(1+\frac{1}{4}|x|^2\right)^t \\ &\geq \left(4+|x|^2\right)^t > \left(1+|x|^2\right)^t.\end{aligned}$$

By the previous lemma, if $s > d/2$ then $f \in H^s(\mathbb{R}^d)$ implies $\hat{f} \in L^1(\mathbb{R}^d)$. Therefore the convolution $\hat{f} * \hat{g}$ is well-defined given two functions $f, g \in H^s(\mathbb{R}^d)$, $s > d/2$. Assume we have two such functions f and g, then by Theorem 2.5 and the above inequality,

$$\begin{aligned}\langle\xi\rangle^s \left|\mathcal{F}(fg)(\xi)\right| &= \frac{1}{\sqrt{2\pi}}\langle\xi\rangle^s \left|\left(\hat{f} * \hat{g}\right)(\xi)\right| \\ &\leq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}^d}\left|\langle\xi\rangle^s \hat{f}(\xi-\eta)\hat{g}(\eta)\right| d\eta \\ &\leq \frac{2^s}{\sqrt{2\pi}}\int_{\mathbb{R}^d}\left|\left(\langle\xi-\eta\rangle^s + \langle\eta\rangle^s\right)\hat{f}(\xi-\eta)\hat{g}(\eta)\right| d\eta \\ &= \frac{2^s}{\sqrt{2\pi}}\left(\left|\langle\cdot\rangle^s \hat{f}\right| * |\hat{g}|(\xi) + \left|\hat{f}\right| * \left|\langle\cdot\rangle^s \hat{g}\right|(\xi)\right).\end{aligned}$$

Thus by Minkowski's inequality and Young's inequality,

$$\begin{aligned}\|fg\|_{H^s(\mathbb{R}^d)} &\leq \frac{2^s}{\sqrt{2\pi}}\|\left|\langle\cdot\rangle^s \hat{f}\right| * |\hat{g}|\,\|_{L^2(\mathbb{R}^d)} + \frac{2^s}{\sqrt{2\pi}}\|\left|\hat{f}\right| * \left|\langle\cdot\rangle^s \hat{g}(\cdot)\right|\,\|_{L^2(\mathbb{R}^d)} \\ &\leq \frac{2^s}{\sqrt{2\pi}}\|f\|_{H^s(\mathbb{R}^d)}\|\hat{g}\|_{L^1(\mathbb{R}^d)} + \frac{2^s}{\sqrt{2\pi}}\|g\|_{H^s(\mathbb{R}^d)}\|\hat{f}\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Finally, Lemma 2.17 then implies

$$\|fg\|_{H^s(\mathbb{R}^d)} \leq c_s\|f\|_{H^s(\mathbb{R}^d)}\|g\|_{H^s(\mathbb{R}^d)}.$$

□

3 Distribution theory and Sobolev spaces on $\mathbb{R}^d$ of real order

Since any $f \in L^2(\mathbb{R}^d)$ clearly satisfies $\|f\|_{H^s(\mathbb{R}^d)} < \infty$ if $s < 0$, we must in order to define $H^s(\mathbb{R}^d)$ for all $s \in \mathbb{R}$ expand the "ambient space" we are working in from $L^2(\mathbb{R}^d)$ to some bigger space. At the same time we want the new definition to be equivalent in the case of non-negative s. The correct definition turns out to be the subset of the continuous dual of the Schwartz space in which the elements have finite Sobolev norm $\|\cdot\|_{H^s(\mathbb{R}^d)}$. For non-negative s we know that these elements should be (equivalence classes of) functions in $L^2(\mathbb{R}^d)$. For negative s they turn out to be distributions, a type of generalised function, and the continuous dual $\mathcal{S}'(\mathbb{R}^d)$ of the Schwartz space is a special class of distributions called the tempered distributions.

Remark 3.1. The proofs in this section are all original, except in one case where we make it clear that we are following an argument from another work. However, the approach to distribution theory is standard.

3.1 Distributions in $\mathcal{D}'(\mathbb{R}^d)$

When considering functions as elements of Lebesgue spaces, we can no longer talk about pointwise values, as two p-integrable functions are inseparable by the $L^p(\mathbb{R}^d)$-metric if they differ at only countably many points. We say they belong the the same equivalence class of functions in $L^p(\mathbb{R}^d)$. In this way elements in $L^p(\mathbb{R}^d)$ are determined only by their global behaviour. With distributions we take this a step further: A distribution is determined by how it acts against a set of *test functions*. More precisely, distributions are continuous linear functionals on this space of test functions. For instance, the test functions may be the compactly supported smooth functions $C_c^\infty(\mathbb{R}^d)$, also written $\mathcal{D}(\mathbb{R}^d)$. The corresponding distribution space is denoted by $\mathcal{D}'(\mathbb{R}^d)$. Later we will consider distributions with Schwartz functions as test functions, the *tempered distributions* $\mathcal{S}'(\mathbb{R}^d)$.

Before we can give a formal definition of $\mathcal{D}'(\mathbb{R}^d)$, we need a notion of convergence on $\mathcal{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$:

Definition 3.1 (Convergence in $\mathcal{D}(\mathbb{R}^d)$)**.** We say a sequence $\{\varphi_n\}_n \subseteq \mathcal{D}(\mathbb{R}^d)$ converges to $\varphi \in \mathcal{D}(\mathbb{R}^d)$ if there exists a compact set $K \subseteq \mathbb{R}^d$ such that

$$\operatorname{supp} \varphi_n \subseteq K$$

for every $n \in \mathbb{N}$ and

$$\sup_{x\in\mathbb{R}^d} D^\alpha(\varphi_n(x) - \varphi(x)) \to 0$$

for every multi-index $\alpha \in \mathbb{N}_0^d$. We say $\varphi_n \to \varphi$ in $\mathcal{D}(\mathbb{R}^d)$.

Definition 3.2 (Distribution in $\mathcal{D}'(\mathbb{R}^d)$)**.** A distribution $T \in \mathcal{D}'(\mathbb{R}^d)$ is a continuous linear functional

$$T : \mathcal{D}(\mathbb{R}^d) \to \mathbb{C}, \quad \varphi \mapsto T\varphi.$$

Here continuity means that $T\varphi_n \to T\varphi$ if $\varphi_n \to \varphi$ in $\mathcal{D}(\mathbb{R}^d)$.

From now on we will usually write $\langle T, \varphi \rangle$ for $T\varphi$.

Example 3.1. Any locally integrable function $f \in L^1_{loc}(\mathbb{R}^d)$ determines a distribution T_f in $\mathcal{D}'(\mathbb{R}^d)$, defined by

$$\langle T_f, \varphi \rangle := \int_{\mathbb{R}^d} f\varphi \, dx$$

for all $\varphi \in \mathcal{D}(\mathbb{R}^d)$. Clearly T_f is well-defined as a linear functional on $\mathcal{D}(\mathbb{R}^d)$. To see that it is continuous, consider a sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{D}(\mathbb{R}^d)$ s.t. $\varphi_n \to 0$ in $\mathcal{D}(\mathbb{R}^d)$ as $n \to \infty$ (by linearity it is enough to consider continuity at the origin). Then, if the compact set K is s.t. $\operatorname{supp} \varphi_n \subseteq K$ for every $n \in \mathbb{N}$,

$$|\langle T_f, \varphi_n \rangle| \le \int_K |f\varphi_n| \, dx \le \sup_{x\in K} |\varphi_n(x)| \int_K |f| \, dx \to 0$$

by $\varphi_n \to 0$ in $\mathcal{D}(\mathbb{R}^d)$.

The distribution T_f is determined by f up to pointwise almost everywhere equivalence [19, Theorem 1.2.5], therefore we often simply write f for T_f and in that way identify f with its distribution. Such a distribution is called a regular distribution. By Remark 2.2, any $f \in L^p(\mathbb{R}^d)$ for $1 \le p \le \infty$, determines a regular distribution.

There are also distributions which cannot be constructed in such a way, thus distributions do indeed generalise the concept of a function. Distributions that are not regular are called singular distributions. A familiar example is the following:

Example 3.2. The δ distribution, defined by

$$\langle \delta, \varphi \rangle = \varphi(0)$$

for all $\varphi \in \mathcal{D}(\mathbb{R}^d)$, is a singular distribution. To see that there can exist no function $f \in L^1_{loc}(\mathbb{R}^d)$ such that $\langle \delta, \varphi \rangle = \int_{\mathbb{R}^d} f\varphi \, dx$, notice that the restriction of δ to $\mathbb{R}^d \setminus \{0\}$ is the zero distribution. Thus $f \equiv 0$ almost everywhere, yet by the definition of δ one would have $\int_{-\lambda}^{\lambda} f \, dx = 1$ for any $\lambda > 0$ (think of φ as a smooth function of compact support with $\varphi = 1$ in $(-\lambda, \lambda)$).

By linearity one can add distributions together and multiply them by scalars to get new distributions, thus $\mathcal{D}'(\mathbb{R}^d)$ forms a vector space. One can also define the product of a smooth function and a distribution. One cannot, however, define a general product of distributions.

Definition 3.3 (Multiplication by a smooth function)**.** Let $g \in C^\infty(\mathbb{R}^d)$ and $T \in \mathcal{D}'(\mathbb{R}^d)$. Then $gT \in \mathcal{D}'(\mathbb{R}^d)$ is the distribution with action

$$\langle gT, \varphi \rangle := \langle T, g\varphi \rangle \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d). \tag{3.1}$$

Remark 3.2. That gT is well-defined as a functional on $\mathcal{D}(\mathbb{R}^d)$ follows from $g\varphi \in \mathcal{D}(\mathbb{R}^d)$ for $\varphi \in \mathcal{D}(\mathbb{R}^d)$ and g smooth. Since gT is linear, it is enough to check continuity at the origin: Let $\{\varphi_n\}_n \subseteq \mathcal{D}(\mathbb{R}^d)$ be a sequence s.t. $\varphi_n \to 0$ in $\mathcal{D}(\mathbb{R}^d)$. Then one can easily verify from the definition of convergence in $\mathcal{D}(\mathbb{R}^d)$ that the sequence $\{g\varphi_n\}_n \subseteq \mathcal{D}(\mathbb{R}^d)$ also goes to 0 in $\mathcal{D}(\mathbb{R}^d)$, thus continuity of gT follows from the continuity of T.

The formula (3.1) is the natural way to define the product between g and T in the case where $T = T_f$ is a regular distribution. One then uses this natural concept to extend the definition of gT to any $T \in \mathcal{D}'(\mathbb{R}^d)$. This is typical of how one defines various operations on distributions.

One can define derivatives of elements in $\mathcal{D}'(\mathbb{R}^d)$, and in fact it turns out that all distributions are infinitely differentiable. This reflects the fact that the test functions we are using are smooth.

Definition 3.4 (Distributional derivative)**.** Let $T \in \mathcal{D}'(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}_0^d$. We define the distributional derivative $D^\alpha T \in \mathcal{D}'(\mathbb{R}^d)$ of T, as the distribution with action

$$\langle D^\alpha T, \varphi \rangle := (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d).$$

Continuity of the linear functional $D^\alpha T$ again follows from the continuity of T.

Remark 3.3. Note that for any weakly differentiable function f, the distributional derivative of f is the same as the weak derivative.

Before moving on to the tempered distributions, we define the following notion of convergence in $\mathcal{D}'(\mathbb{R}^d)$:

Definition 3.5 (Convergence in $\mathcal{D}'(\mathbb{R}^d)$)**.** We say a sequence $\{T_n\}_n \subseteq \mathcal{D}'(\mathbb{R}^d)$ converges to $T \in \mathcal{D}'(\mathbb{R}^d)$ if

$$\langle T_n, \varphi \rangle \to \langle T, \varphi \rangle$$

as $n \to \infty$ for every $\varphi \in \mathcal{D}(\mathbb{R}^d)$. We write $T_n \to T$ in $\mathcal{D}'(\mathbb{R}^d)$.

3.2 Tempered distributions

By using the larger space $\mathcal{S}(\mathbb{R}^d)$ as test functions, one gets the tempered distributions $\mathcal{S}'(\mathbb{R}^d)$, a subspace of $\mathcal{D}'(\mathbb{R}^d)$:

Definition 3.6 (Distribution in $\mathcal{S}'(\mathbb{R}^d)$)**.** A distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ is a continuous linear functional

$$T : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}, \quad \varphi \mapsto \langle T, \varphi \rangle.$$

We call such a distribution T a tempered distribution. Here continuity means that $\langle T, \varphi_n \rangle \to \langle T, \varphi \rangle$ if $\varphi_n \to \varphi$ in $\mathcal{S}(\mathbb{R}^d)$.

That $\mathcal{S}'(\mathbb{R}^d) \subseteq \mathcal{D}'(\mathbb{R}^d)$ follows from the fact that convergence of a sequence in $\mathcal{D}(\mathbb{R}^d)$ implies convergence of the same sequence in $\mathcal{S}(\mathbb{R}^d)$ (cf. Definitions 2.6 and 3.1), and that $\mathcal{D}(\mathbb{R}^d)$ is a dense subset of $\mathcal{S}(\mathbb{R}^d)$. Let us quickly prove the last statement:

Proposition 3.1. *We have the dense embedding*

$$\mathcal{D}(\mathbb{R}^d) \overset{d}{\hookrightarrow} \mathcal{S}(\mathbb{R}^d).$$

Proof. Consider a cutoff function $\eta \in \mathcal{D}(\mathbb{R}^d)$ such that $\eta(x) = 1$ for $|x| \le 1$, and $|\eta(x)| \le 1$ elsewhere. Define $\eta_n(x) \coloneqq \eta(\frac{x}{n})$. Then $\eta_n(x) = 1$ for $|x| \le n$ and $\eta_n \to 1$ pointwise as $n \to \infty$. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and define $\varphi_n = \eta_n \varphi$. Each $\varphi_n \in \mathcal{D}(\mathbb{R}^d)$, since φ_n is a product of smooth functions and has the compact support of η_n. We want to show $\varphi_n \underset{\mathcal{S}}{\to} \varphi$. We have

$$\begin{aligned}
\sup_{x \in \mathbb{R}^d} \left| x^\alpha D^\beta (\varphi_n - \varphi) \right| &= \sup_{x \in \mathbb{R}^d} \left| x^\alpha \left(\sum_{|\gamma| \le |\beta|} \binom{\beta}{\gamma} (D^\gamma \eta_n) \left(D^{\beta-\gamma} \varphi \right) - D^\beta \varphi \right) \right| \\
&= \sup_{x \in \mathbb{R}^d} \left| \sum_{1 \le |\gamma| \le |\beta|} x^\alpha \binom{\beta}{\gamma} (D^\gamma \eta_n) \left(D^{\beta-\gamma} \varphi \right) + x^\alpha \left(D^\beta \varphi \right) (\eta_n - 1) \right| \\
&\le \sup_{x \in \mathbb{R}^d} \left| \sum_{1 \le |\gamma| \le |\beta|} x^\alpha \binom{\beta}{\gamma} (D^\gamma \eta_n) \left(D^{\beta-\gamma} \varphi \right) \right| + \sup_{x \in \mathbb{R}^d} \left| x^\alpha \left(D^\beta \varphi \right) (\eta_n - 1) \right| \\
&\le \sum_{1 \le |\gamma| \le |\beta|} \binom{\beta}{\gamma} \|\varphi\|_{\alpha, \beta-\gamma} \sup_{x \in \operatorname{supp} \eta_n} |D^\gamma \eta_n(x)| \\
&\quad + \|\varphi\|_{\alpha,\beta} \sup_{x \in \operatorname{supp} \eta_n} |\eta_n(x) - 1| \to 0 \quad \text{as } n \to \infty,
\end{aligned}$$

since $\eta_n \to 1$ pointwise and has compact support. □

We cannot in general multiply a smooth function and a tempered distribution. We can, however, define the product between a tempered distribution and a *slowly growing*, or *tempered*, function:

Definition 3.7 (Tempered function)**.** A tempered function is a smooth function f for which there exists a $C > 0$ and an $N \in \mathbb{N}_0$ such that $|f(x)| \le C(1 + |x|)^N$ for all $x \in \mathbb{R}^d$.

Definition 3.8 (Multiplication by a tempered function)**.** Let g be a tempered function and T a tempered distribution. Then gT is the tempered distribution with action

$$\langle gT, \varphi \rangle := \langle T, g\varphi \rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$

Remark 3.4. By the definition of the Schwartz space, $g\varphi$ is a Schwartz function for g tempered and $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Thus gT is well-defined as a functional on $\mathcal{S}(\mathbb{R}^d)$. As gT is linear and T is continuous, showing continuity of gT is simply a matter of verifying from the definition of convergence in $\mathcal{S}(\mathbb{R}^d)$ that $g\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$ given a sequence $\{\varphi_n\}_n \subseteq \mathcal{S}(\mathbb{R}^d)$ such that $\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$.

The definitions of the distributional derivative of a tempered distribution and of convergence in $\mathcal{S}'(\mathbb{R}^d)$ are analogous to the definitions for $\mathcal{D}'(\mathbb{R}^d)$ (see Definitions 3.4 and 3.5), with the test functions φ now being in $\mathcal{S}(\mathbb{R}^d)$. From the definition of the distributional derivative and by the Schwartz space being invariant under differentiation, it is clear that the derivative of a tempered distribution is again a tempered distribution.

The criteria for a function to determine a regular distribution is stricter in $\mathcal{S}'(\mathbb{R}^d)$: Each function in f such that $f\varphi \in L^1(\mathbb{R}^d)$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ defines a regular tempered distribution if the functional $\int_{\mathbb{R}^d} f\varphi$ is continuous. It is no longer enough that $f \in L^1_{loc}(\mathbb{R}^d)$. In the next proposition we show that tempered functions as well as functions in $L^p(\mathbb{R}^d)$ for $1 \leq p \leq \infty$ define regular tempered distributions:

Lemma 3.2 (Regular tempered distributions)**.** *Let* $1 \leq p \leq \infty$*. Every function* $f \in L^p(\mathbb{R}^d)$ *defines a regular distribution in* $\mathcal{S}'(\mathbb{R}^d)$*. In addition, every tempered function defines a regular tempered distribution.*

Proof. Let $f \in L^p(\mathbb{R}^d)$. For $\varphi \in \mathcal{S}(\mathbb{R}^d)$, define the linear functional $\langle f, \varphi \rangle = \int_{\mathbb{R}^d} f\varphi$. That $\varphi f \in L^1(\mathbb{R}^d)$ follows from Hölder's inequality,

$$\|f\varphi\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|\varphi\|_{L^q(\mathbb{R}^d)} < \infty,$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Let $\{\varphi_n\}_n \subseteq \mathcal{S}(\mathbb{R}^d)$ be a sequence such that $\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. Then

$$|\langle f, \varphi_n \rangle| \leq \int_{\mathbb{R}^d} |f| |\varphi_n| \, dx \leq \|f\|_{L^p(\mathbb{R}^d)} \|\varphi_n\|_{L^q(\mathbb{R}^d)}.$$

Either we have $q = \infty$, in which case the term $\|\varphi_n\|_{L^q(\mathbb{R}^d)}$ tends to 0 trivially due to $\varphi_n \underset{\mathcal{S}}{\to} 0$, or $q < \infty$ and we have

$$\begin{aligned}\int_{\mathbb{R}^d} |\varphi_n(x)|^q \, dx &\leq \|\varphi_n\|_\infty^{q-1} \int_{\mathbb{R}^d} |\varphi_n(x)| \, dx \\ &= \|\varphi_n\|_\infty^{q-1} \int_{\mathbb{R}^d} \frac{\prod_{i=1}^d (1 + x_i^2) |\varphi_n(x)|}{\prod_{i=1}^d (1 + x_i^2)} \, dx \\ &\leq \|\varphi_n\|_\infty^{q-1} \| \prod_{i=1}^d (1 + x_i^2) \varphi_n \|_\infty \int_{\mathbb{R}^d} \frac{dx}{\prod_{i=1}^d (1 + x_i^2)},\end{aligned}$$

where the final integral equals π^d, thus $|\langle f, \varphi_n \rangle| \to 0$ as $n \to \infty$ by $\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$.

Concerning the second part of the lemma, note that for a tempered function f and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we have

$$|f(x)\varphi(x)| = \frac{|f(x)|}{(1 + |x|)^N} (1 + |x|)^N |\varphi(x)| \leq C(1 + |x|)^N |\varphi(x)|$$

for some $C > 0$ and $N \in \mathbb{N}_0$, thus $f\varphi$ is integrable. To show continuity, we let $\{\varphi_n\}_n \subseteq \mathcal{S}(\mathbb{R}^d)$ be a sequence s.t. $\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. Then

$$\int_{\mathbb{R}^d} f(x)\varphi_n(x) \, dx \leq C \int_{\mathbb{R}^d} (1 + |x|)^N |\varphi_n(x)| \, dx \to 0$$

by Lebesgue's dominated convergence theorem, since the integrand goes (uniformly) to 0 by $\varphi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. □

Remark 3.5. Note that this is in no way a complete classification of regular tempered distributions. For instance, it can be shown that a locally integrable function f defines a regular tempered distribution if

$$\int_{|x| \leq A} |f(x)| \, dx \leq CA^N \quad \text{as } A \to \infty$$

for some constants C and N, and that this condition is necessary if f is positive [41, p. 47].

By Lemma 3.2 we have the inclusions $L^p(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$ for $1 \leq p \leq \infty$. Since also $\mathcal{S}(\mathbb{R}^d) \subseteq L^p(\mathbb{R}^d)$ for $1 \leq p \leq \infty$, we have the relation

$$\mathcal{S}(\mathbb{R}^d) \subseteq L^p(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d).$$

3.3 The Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$

Due to the Schwartz space being invariant under Fourier transform, we may define the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ by duality:

Definition 3.9 (Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$)**.** For a tempered distribution T, the Fourier transform $\mathcal{F}(T)$ is the tempered distribution defined by

$$\langle \mathcal{F}(T), \varphi \rangle := \langle T, \mathcal{F}\varphi \rangle$$

for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. The inverse Fourier transform is defined similarly.

Remark 3.6. Note that $\mathcal{F}(T)$ is well-defined as a tempered distribution due to the invariance of $\mathcal{S}(\mathbb{R}^d)$ under Fourier transform and the continuity of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^d)$ (cf. Theorem 2.6).

Theorem 3.3 (Fourier inversion theorem on $\mathcal{S}'(\mathbb{R}^d)$)**.** *The Fourier transform is a continuous and invertible map $\mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$: For $T \in \mathcal{S}'(\mathbb{R}^d)$, we have $\mathcal{F}\mathcal{F}^{-1}(T) = \mathcal{F}^{-1}\mathcal{F}(T) = T$.*

Proof. That $\mathcal{F}\mathcal{F}^{-1}(T) = \mathcal{F}^{-1}\mathcal{F}(T) = T$ follows from the definition of $\mathcal{F}$ on $\mathcal{S}'(\mathbb{R}^d)$ and Theorem 2.9:

$$\langle \mathcal{F}\mathcal{F}^{-1}(T), \varphi \rangle = \langle \mathcal{F}^{-1}(T), \mathcal{F}\varphi \rangle = \langle T, \mathcal{F}^{-1}\mathcal{F}\varphi \rangle = \langle T, \varphi \rangle,$$

and similarly for $\mathcal{F}^{-1}\mathcal{F}(T)$.

It is enough to check continuity at the origin since $\mathcal{F}$ is linear, therefore let $T_n \to 0$ in $\mathcal{S}'(\mathbb{R}^d)$. Then

$$\langle \mathcal{F}(T_n), \varphi \rangle = \langle T_n, \mathcal{F}(\varphi) \rangle \to 0,$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, i.e. $\mathcal{F}(T_n) \to 0$ in $\mathcal{S}'(\mathbb{R}^d)$. □

Theorem 3.4 (Properties of $\mathcal{F}$ on $\mathcal{S}'(\mathbb{R}^d)$)**.** *For $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}_0^d$, we have*

$$\mathcal{F}(D^\alpha T) = i^{|\alpha|} x^\alpha \mathcal{F}(T)$$

and

$$\mathcal{F}(x^\alpha T) = i^{|\alpha|} D^\alpha \mathcal{F}(T).$$

Proof. Firstly, note that since x^α is a tempered function, $x^\alpha T$ is well-defined as a tempered distribution. For $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we have by property (ii) of Theorem 2.3 and the definition of the distributional derivative,

$$\begin{aligned}\langle \mathcal{F}(D^\alpha T), \varphi \rangle &= \langle D^\alpha T, \mathcal{F}\varphi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha(\mathcal{F}\varphi) \rangle \\ &= i^{|\alpha|} \langle T, \mathcal{F}(x^\alpha \varphi) \rangle = i^{|\alpha|} \langle \mathcal{F}(T), x^\alpha \varphi \rangle = \langle i^{|\alpha|} x^\alpha \mathcal{F}(T), \varphi \rangle.\end{aligned}$$

Similarly, by property (i) of Theorem 2.3,

$$\begin{aligned}\langle \mathcal{F}(x^\alpha T), \varphi \rangle &= \langle x^\alpha T, \mathcal{F}\varphi \rangle = \langle T, x^\alpha \mathcal{F}\varphi \rangle = (-1)^{|\alpha|} i^{|\alpha|} \langle T, \mathcal{F}(D^\alpha \varphi) \rangle \\ &= (-1)^{|\alpha|} i^{|\alpha|} \langle \mathcal{F}(T), D^\alpha \varphi \rangle = \langle i^{|\alpha|} D^\alpha \mathcal{F}(T), \varphi \rangle.\end{aligned}$$

□

Next we would like to prove a result analogous to Theorem 2.5 for the Fourier transform of the convolution of a tempered distribution and a Schwartz function. In order to figure out how to define the convolution $T * \varphi$ for $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we suppose that f is an integrable function which thus defines a regular tempered distribution. Then if $\sim$ denotes the reflection operator, that is $\tilde{\varphi}(x) = \varphi(-x)$, and τ_a the translation operator, that is $\tau_a\varphi(x) = \varphi(x-a)$, we have

$$(f * \varphi)(x) = \int_{\mathbb{R}^d} f(y)\varphi(x-y)\,dx = \langle f, \tau_x\tilde{\varphi}\rangle.$$

This motivates the following definition:

Definition 3.10 (Convolution between $\mathcal{S}'(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$)**.** For $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we define the convolution $T * \varphi$ to be the function defined by

$$(T * \varphi)(x) = \langle T, \tau_x\tilde{\varphi}\rangle, \quad x \in \mathbb{R}^d. \tag{3.2}$$

Another way to define the convolution $T * \varphi$ as a *distribution* is motivated by considering the tempered distribution generated by the convolution between a function $f \in L^1(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then $f * \varphi \in L^1(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$ by Lemma 2.4, and for any $\psi \in \mathcal{S}(\mathbb{R}^d)$ we have, by Fubini's theorem,

$$\begin{aligned}\langle f * \varphi, \psi\rangle &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(y)\varphi(x-y)\,dy\right)\psi(x)\,dx\\ &= \int_{\mathbb{R}^d} f(y)\left(\int_{\mathbb{R}^d}\varphi(x-y)\psi(x)\,dx\right)dy = \langle f, \tilde{\varphi} * \psi\rangle.\end{aligned}$$

Definition 3.11 (Convolution between $\mathcal{S}'(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$ II)**.** For $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we define the convolution $T *' \varphi$ to be the tempered distribution defined by

$$\langle T *' \varphi, \psi\rangle = \langle T, \tilde{\varphi} * \psi\rangle, \quad \forall\, \psi \in \mathcal{S}(\mathbb{R}^d).$$

Remark 3.7. That $T *' \varphi$ is a well-defined tempered distribution follows from $T \in \mathcal{S}'(\mathbb{R}^d)$ and the fact that $*$ is continuous map $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ as per Lemma 2.4.

Of course we ought to prove that these two definitions are equivalent:

Proposition 3.5. *For $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}'(\mathbb{R}^d)$, we have $T * \varphi \in \mathcal{S}'(\mathbb{R}^d)$ and*

$$\langle T *' \varphi, \psi\rangle = \langle T * \varphi, \psi\rangle$$

for all $\psi \in \mathcal{S}(\mathbb{R}^d)$.

Proof. First of all, the function $T * \varphi$ defined by (3.2) generates a tempered distribution because it is a tempered function: By definition, $T * \varphi$ is tempered if it is smooth and if there exists a $C > 0$ and an $N \in \mathbb{N}_0$ such that

$$|(T * \varphi)(x)| \leq C(1 + |x|)^N$$

for all $x \in \mathbb{R}^d$, i.e $|T * \varphi|$ should be bounded by a polynomial. We estimate

$$|(T * \varphi)(x)| = |\langle T, \tau_x \tilde{\varphi} \rangle| \leq \|T\|_{op}\, d_{\mathcal{S}}(\tau_x \tilde{\varphi}, 0) = \|T\|_{op} \sum_{\alpha,\beta} \frac{1}{2^{|\alpha|+|\beta|}} \frac{P_{\alpha,\beta}(\tau_x \tilde{\varphi})}{1 + P_{\alpha,\beta}(\tau_x \tilde{\varphi})},$$

by $T \in \mathcal{S}'(\mathbb{R}^d)$. For the term $P_{\alpha,\beta}(\tau_x \tilde{\varphi})$ we have

$$\begin{aligned}
\sup_{y \in \mathbb{R}^d} (1 + |y|)^{|\alpha|} \left|(D_y^\beta \tilde{\varphi})(y - x)\right| &= \sup_{y \in \mathbb{R}^d} (1 + |y + x|)^{|\alpha|} |(D_y^\beta \tilde{\varphi})(y)| \\
&\leq \sup_{y \in \mathbb{R}^d} (1 + |y| + |x|)^{|\alpha|} |(D_y^\beta \tilde{\varphi})(y)| \\
&\leq \sup_{y \in \mathbb{R}^d} \left((1 + |y|)^{|\alpha|} + |x|^{|\alpha|}\right) |(D_y^\beta \tilde{\varphi})(y)| \\
&\leq P_{\alpha,\beta}(\tilde{\varphi}) + \sup_{y \in \mathbb{R}^d} (1 + |x|)^{|\alpha|} |(D_y^\beta \tilde{\varphi})(y)| \\
&\leq 2 P_{\alpha,\beta}(\tilde{\varphi}),
\end{aligned}$$

thus

$$\begin{aligned}
|(T * \varphi)(x)| &\leq \|T\|_{op} \sum_{\alpha,\beta} \frac{1}{2^{|\alpha|+|\beta|}} \frac{2 P_{\alpha,\beta}(\tilde{\varphi})}{1 + 2 P_{\alpha,\beta}(\tilde{\varphi})} \\
&\leq 2 \|T\|_{op} \sum_{\alpha,\beta} \frac{1}{2^{|\alpha|+|\beta|}} \frac{P_{\alpha,\beta}(\tilde{\varphi})}{1 + P_{\alpha,\beta}(\tilde{\varphi})} \\
&= 2 \|T\|_{op}\, d_{\mathcal{S}}(\varphi, 0) =: C.
\end{aligned}$$

By the continuity of T, we have for $a \in \mathbb{R}^d$

$$\lim_{x \to a} (T * \varphi)(x) = \lim_{x \to a} \langle T, \tau_x \tilde{\varphi} \rangle = (T * \varphi)(a)$$

since the sequence $\tau_{a+\frac{1}{n}} \tilde{\varphi} \to \tau_a \tilde{\varphi}$ as $n \to \infty$ in $\mathcal{S}(\mathbb{R}^d)$. Similarly we get, if $h = (0, \cdots, h_j, \cdots,)$ for $j \in \{1, \cdots, d\}$,

$$\frac{\partial}{\partial x_j} (T * \varphi)(a) = \lim_{h_j \to 0} \langle T, \frac{\tau_{a+h} \tilde{\varphi} - \tau_a \tilde{\varphi}}{h_j} \rangle = (T * (\frac{\partial \varphi}{\partial x_j}))(a) < \infty.$$

Repeating the argument we get that $T * \varphi$ is smooth, and thus it is tempered. By Lemma 3.2 this implies that $T * \varphi$ generates a regular tempered distribution.

Finally then we get by direct computation

$$\begin{aligned}\langle T * \varphi, \psi\rangle &= \langle\langle T, \tau_x\tilde{\varphi}\rangle, \psi\rangle = \int_{\mathbb{R}^d}\langle T, \tau_x\tilde{\varphi}\rangle\psi(x)\,dx \\ &= \langle T, \int_{\mathbb{R}^d}(\tau_x\tilde{\varphi})\psi(x)\,dx\rangle = \langle T, \tilde{\varphi} * \psi\rangle.\end{aligned}$$

□

From now on we denote the convolution between $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ using either definition simply by $*$. For such convolutions we have the following result:

Theorem 3.6. *Let $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then we have*

$$\mathcal{F}(T * \varphi) = (2\pi)^{\frac{d}{2}}\hat{\varphi}\hat{T} \quad \text{and} \quad \hat{T} * \hat{\varphi} = (2\pi)^{\frac{d}{2}}\mathcal{F}(\varphi T)$$

in the sense of tempered distributions.

Proof. Let $\psi \in \mathcal{S}(\mathbb{R}^d)$. We compute

$$\begin{aligned}\langle\mathcal{F}(T * \varphi), \psi\rangle &= \langle T * \varphi, \hat{\psi}\rangle = \langle T, \tilde{\varphi} * \hat{\psi}\rangle \\ &= (2\pi)^{\frac{d}{2}}\langle T, \mathcal{F}(\mathcal{F}^{-1}(\tilde{\varphi})\psi)\rangle \\ &= (2\pi)^{\frac{d}{2}}\langle T, \mathcal{F}(\hat{\varphi}\psi)\rangle = (2\pi)^{\frac{d}{2}}\langle\hat{\varphi}\hat{T}, \psi\rangle,\end{aligned}$$

by Theorem 2.5, and

$$\begin{aligned}\langle\hat{T} * \hat{\varphi}, \psi\rangle &= \langle T, \mathcal{F}\left(\widetilde{\hat{\varphi}} * \psi\right)\rangle = \langle T, \mathcal{F}(\mathcal{F}^{-1}\varphi * \psi)\rangle \\ &= (2\pi)^{\frac{d}{2}}\langle T, \varphi\hat{\psi}\rangle = (2\pi)^{\frac{d}{2}}\langle\mathcal{F}(\varphi T), \psi\rangle.\end{aligned}$$

□

We are now ready to introduce the fractional Sobolev spaces of negative order. We will at the same time give an equivalent definition of $H^s(\mathbb{R}^d)$ for $s \geq 0$.

3.4 Fractional Sobolev spaces $H^s(\mathbb{R}^d)$ for $s \in \mathbb{R}$

Definition 3.12 (Fractional Sobolev spaces of real order)**.** For $s \in \mathbb{R}$ we define the fractional Sobolev space of order s as

$$H^s(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) \ : \ \|f\|_{H^s(\mathbb{R}^d)} < \infty\}$$

where $\|f\|^2_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d}(1 + |\xi|^2)^s \left|\hat{f}(\xi)\right|^2 d\xi$.

Remark 3.8. Note that in this definition it is implicit that the distributions in $H^s(\mathbb{R}^d)$ have measurable functions as Fourier transforms.

In Definition 2.10 we defined the spaces $H^s(\mathbb{R}^d)$ for $s \geq 0$ to be the set of $L^2(\mathbb{R}^d)$-functions with finite Sobolev norm. This is equivalent to the above definition, since $\|f\|_{H^s(\mathbb{R}^d)} < \infty$ for $s \geq 0$ clearly implies $f \in L^2(\mathbb{R}^d)$, and we know from Lemma 3.2 that $L^2(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$.

For $s < 0$, on the other hand, not all elements of $H^s(\mathbb{R}^d)$ are $L^2(\mathbb{R}^d)$-functions. In fact, even singular tempered distributions such as δ may be in $H^s(\mathbb{R}^d)$ for small enough s: By writing out the definition, one can find the Fourier transform of δ to be the regular tempered distribution induced by the constant function $(2\pi)^{-d/2}$. Therefore we have $\delta \in H^s(\mathbb{R}^d)$ for $s < -d/2$.

We would now like to establish that $H^s(\mathbb{R}^d)$ is a Hilbert space also for negative s. This actually follows from the fact that $H^{-s}(\mathbb{R}^d)$ for $s \geq 0$ can be considered the continuous dual of the Hilbert space $H^s(\mathbb{R}^d)$ (and vice versa):

Theorem 3.7. *Let $s \in \mathbb{R} \setminus \{0\}$. $H^{-s}(\mathbb{R}^d)$ is the continuous dual of $H^s(\mathbb{R}^d)$. $H^0(\mathbb{R}^d) = L^2(\mathbb{R}^d)$ is self-dual.*

Proof. We follow the proof of Proposition 4.10 in [11]. Suppose $f \in H^{-s}(\mathbb{R}^d)$, and identify f with the linear functional

$$I_f(g_n) := \int_{\mathbb{R}^d} \hat{f}(\xi)\hat{g}(\xi)\, d\xi$$

for $g \in H^s(\mathbb{R}^d)$. The functional is well-defined and continuous, as for a sequence $\{g_n\}_n \subseteq H^s(\mathbb{R}^d)$ such that $g_n \to 0$ in $H^s(\mathbb{R}^d)$, we have by Hölder's inequality

$$|I_f(g)| \leq \int_{\mathbb{R}^d} (1+|\xi|^2)^{-s/2} \left|\hat{f}(\xi)\right| (1+|\xi|^2)^{s/2} |g_n(\xi)|\, d\xi \leq \|f\|_{H^{-s}(\mathbb{R}^d)} \|g_n\|_{H^s(\mathbb{R}^d)} \to 0.$$

Consequently we have the embedding $H^{-s}(\mathbb{R}^d) \hookrightarrow H^s(\mathbb{R}^d)'$.

Now let $T \in H^s(\mathbb{R}^d)'$. By the dense embedding $\mathcal{S}(\mathbb{R}^d) \hookrightarrow H^s(\mathbb{R}^d)$ (see Lemma 2.15), we have that $H^s(\mathbb{R}^d)' \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$, hence $T \in \mathcal{S}'(\mathbb{R}^d)$. Suppose $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Since $(1+|\xi|^2)^{-s/2}$ is a tempered function we have

$$\begin{aligned}
\left|\langle (1+|\xi|^2)^{-\frac{s}{2}} \mathcal{F}(T), \varphi\rangle\right| &= \left|\langle T, \mathcal{F}((1+|\xi|^2)^{-\frac{s}{2}}\varphi)\rangle\right| \\
&\leq \|T\|_{H^s(\mathbb{R}^d)'} \|\mathcal{F}((1+|\xi|^2)^{-\frac{s}{2}}\varphi)\|_{H^s(\mathbb{R}^d)} \\
&= \|T\|_{H^s(\mathbb{R}^d)'} \|\varphi\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

Thus, by the continuous dense embedding of $\mathcal{S}(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d)$, we have that $(1+|\xi|^2)^{-s/2}\mathcal{F}(T) \in L^2(\mathbb{R}^d)$ since it defines a continuous linear functional on $L^2(\mathbb{R}^d)$ and we know $L^2(\mathbb{R}^d)$ is self-dual. This means we also have the other embedding, $H^s(\mathbb{R}^d)' \hookrightarrow H^{-s}(\mathbb{R}^d)$. □

This implies that $H^s(\mathbb{R}^d)$ for $s < 0$ is also a Hilbert space, as it is the dual of a Hilbert space:

Corollary 3.7.1. *$H^s(\mathbb{R}^d)$ for $s < 0$ is a Hilbert space.*

For $s \geq 0$, a function $f \in H^s(\mathbb{R}^d)$ can be characterised as an $L^2(\mathbb{R}^d)$-function having s (possibly fractional) weak derivatives in $L^2(\mathbb{R}^d)$. A natural question is if we can similarly characterise elements in Sobolev spaces of negative order. If we define Λ^s as the operator that corresponds to $\langle\xi\rangle^s = (1 + |\xi|^2)^{s/2}$ on the Fourier side, i.e. $\mathcal{F}(\Lambda^s u)(\xi) = \langle\xi\rangle^s \hat{u}(\xi)$, then one can also write $H^s(\mathbb{R}^d)$ for $s \in \mathbb{R}$ as

$$H^s(\mathbb{R}^d) = \Lambda^s L^2(\mathbb{R}^d). \tag{3.3}$$

This is because, as remarked in Theorem 2.14, Λ^s is an isometric isomorphism $H^s(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. In Theorem 2.12 we saw how square integrability of $\Lambda^k f$ for $f \in L^2(\mathbb{R}^d)$ corresponded to f having k weak derivatives in $L^2(\mathbb{R}^d)$. Using (3.3) one can characterise functions in $H^s(\mathbb{R}^d)$ for $s < 0$ as functions of *negative regularity*, in the sense that each such function is in some sense the (possibly fractional) derivative of an $L^2(\mathbb{R}^d)$-function.

Finally, in view of (3.3) and Theorem 3.4, we note that for all $s \in \mathbb{R}$, the distributional derivative operator D^α for $\alpha \in \mathbb{N}_0^d$ maps $H^s(\mathbb{R}^d)$ continuously into $H^{s-|\alpha|}(\mathbb{R}^d)$.

4 Periodic Sobolev spaces on the line

The purpose of this section is to develop theory for the periodic Sobolev spaces of real order, analogous to the $H^s(\mathbb{R}^d)$-theory we have already established in the previous sections. We start by defining the concept of Fourier series of periodic functions, which is the periodic analogue of the Fourier transform of functions on the line. Then we define and establish some theory for the periodic test functions and periodic distributions, and finally we develop the basic theory of periodic Sobolev spaces. We restrict ourselves to developing theory for the periodic Sobolev spaces on the real line because we don't need the multidimensional theory in this thesis, however the presentation is easily generalisable to the d-dimensional case, where periodicity is taken to mean periodicity in each variable.

Remark 4.1. The presentation in this section owes a lot to [21]. Our proofs are mostly original, in a few cases where we have taken inspiration from other sources we make this clear.

4.1 Fourier series

It is clear what it means for a function to be periodic: If we define the translation operator τ_a to be the operator with action $\tau_a f(x) = f(x-a)$ on any function f, then f has period $p > 0$ if $\tau_p f = f$. To simplify the presentation we work exclusively with 2π-periodic functions. To such a function f we associate a *Fourier series*

$$\sum_{k\in\mathbb{Z}} c_k \Theta_k, \tag{4.1}$$

where Θ_k is the function $x \mapsto \exp(ikx)$ and the *Fourier coefficients* c_k are given by

$$c_k = \hat{f}(k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = \frac{1}{2\pi}\langle f, \Theta_k\rangle_{L^2(-\pi,\pi)}, \tag{4.2}$$

provided that these coefficients are well-defined. Our interest is in what sense the infinite sum (4.1) converges to f, and in which conditions we must put on f for convergence to hold. The function f actually doesn't have to be periodic to have a Fourier series associated to it, in fact it doesn't even need to have pointwise values: For $f \in L^2(-\pi,\pi)$, the Fourier series converges to f in the $L^2(-\pi,\pi)$-sense, that is

$$\lim_{n\to\infty} \|f - \sum_{k=-n}^{n} c_k\Theta_k\|_{L^2(-\pi,\pi)} = 0,$$

a result known as the Riesz-Fischer theorem[1]. Many other and stronger results about the convergence of Fourier series are known, in fact Lennart Carleson established that the Fourier series of a square integrable function converges almost

[1]Although it is only one of several closely related results about L^2 with this name.

everywhere [6]. The elegant theory of Fourier series is complicated somewhat by the different notions of convergence. Our primary goal will be to establish convergence at least in the (rather weak) sense of *periodic distributions* for elements in periodic Sobolev spaces.

It is not as obvious how one should define periodicity for distributions, since they do not in general have pointwise values. To help make sense of the idea of a distribution being periodic, let us apply the translation operator to a regular distribution T_f determined by some continuous function f. Since we typically identify T_f with f, the natural way to define the action of τ_a on a regular distribution is $\tau_a T_f = T_{\tau_a f}$. Through a change of variables we then get

$$\langle \tau_a f, \varphi\rangle = \int f(x-a)\varphi(x)\,dx = \int f(y)\varphi(y+a)\,dx = \langle f, \tau_{-a}\varphi\rangle$$

for all test functions φ. We use this formula to define the action of the translation operator on any distribution T, and then we can say that a distribution T is periodic with period $p \in \mathbb{R}$ if $\tau_p T = T$.

Instead of defining the periodic Sobolev spaces as subspaces of 2π-periodic distributions in $\mathcal{S}'(\mathbb{R})$, we will work in the set of distributions $\mathcal{P}'$ over the test functions which we call $\mathcal{P}$, which denotes the set of smooth 2π-periodic functions. All distributions in $\mathcal{P}'$ are automatically 2π-periodic. One can prove that the space of 2π-periodic tempered distributions is isomorphic to $\mathcal{P}'$ [40], and it is also a fact that every periodic tempered distribution has a Fourier expansion which converges in the sense of tempered distributions [35, Theorem 5.10]. However, we decide to work in the space $\mathcal{P}'$ of periodic distributions because of its simpler mathematical structure.

Before we formally define the periodic test functions, we include for convenience the following fundamental result on the summability of the Fourier series of an $L^2(-\pi,\pi)$-function:

Theorem 4.1 (Parseval's identity)**.** *For a function $f \in L^2(-\pi,\pi)$, we have*

$$\|\hat{f}\|^2_{l^2(\mathbb{Z})} = \sum_{k\in\mathbb{Z}} |\hat{f}(k)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx = \frac{1}{2\pi}\|f\|^2_{L^2(-\pi,\pi)},$$

where the Fourier coefficients $\hat{f}(k)$ are defined as in (4.2)*. Equivalently, we have*

$$\langle \hat{f}, \hat{g}\rangle_{l^2(\mathbb{Z})} = \sum_{k\in\mathbb{Z}} \hat{f}(k)\overline{\hat{g}(k)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx = \frac{1}{2\pi}\langle f, g\rangle_{L^2(-\pi,\pi)}$$

for $f, g \in L^2(-\pi,\pi)$.

Proof. See for instance [44, p. 54]. □

Finally, we point out that for every complex sequence $\{\alpha_k\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$, there is a unique function $f \in L^2(-\pi,\pi)$ such that $\hat{f}(k) = \alpha_k$, that is $f := \sum_{k\in\mathbb{Z}} \alpha_k \Theta_k$ in $L^2(-\pi,\pi)$ [24, p. 32].

4.2 Periodic test functions

We now formally introduce the set of test functions which we later shall define the periodic distributions as continuous linear functionals over:

Definition 4.1 (Periodic test functions)**.** Let $\mathcal{P}$ be the set of complex valued smooth functions on the real line of period 2π. Elements in $\mathcal{P}$ are called *periodic test functions.*

Clearly $\mathcal{P}$ is a vector space. In order to define the set of periodic distributions as the continuous dual of $\mathcal{P}$, we need a notion of convergence in $\mathcal{P}$.

Definition 4.2 (Convergence in $\mathcal{P}$)**.** We say a sequence $\{\varphi_n\}_n \subseteq \mathcal{P}$ converges to φ in $\mathcal{P}$ if, for all $k \in \mathbb{N}_0$,

$$\|\varphi_n^{(k)} - \varphi^{(k)}\|_\infty = \sup_{x\in(-\pi,\pi)} |\varphi_n^{(k)}(x) - \varphi^{(k)}(x)| \to 0$$

as $n \to \infty$.

The space $\mathcal{P}$ is a metric space with the distance

$$d_{\mathcal{P}}(\varphi,\psi) := \sum_{j=0}^{\infty} \frac{1}{2^j} \frac{\|\varphi^{(j)} - \psi^{(j)}\|_\infty}{1 + \|\varphi^{(j)} - \psi^{(j)}\|_\infty},$$

for $\varphi, \psi \in \mathcal{P}$. Moreover, $(\mathcal{P}, d_{\mathcal{P}})$ is complete [21, p. 133]. Note that this metric induces same notion of convergence as the one in Definition 4.2.

We would like to study Fourier series of periodic test functions. Clearly periodic test functions are locally square integrable, thus their Fourier series are $l^2(\mathbb{Z})$-sequences by Theorem 4.1, and their Fourier series converge in $L^2(-\pi,\pi)$. Much more is true, however. We begin our study by defining the sequential analogue of the Schwartz space of rapidly decreasing functions:

Definition 4.3 (Rapidly decreasing sequences)**.** A sequence $\{\alpha_k\}_{k\in\mathbb{Z}} \subseteq \mathbb{C}$ is said to be *rapidly decreasing* if

$$\|\alpha\|_{\infty,j} := \sup_{k\in\mathbb{Z}} |k|^j |\alpha_k| < \infty$$

for all $j \in \mathbb{N}_0$. The set of rapidly decreasing sequences is denoted by $\mathcal{S}(\mathbb{Z})$.

We will regard $\mathcal{S}(\mathbb{Z})$ as a metric space with the distance

$$d_{\mathcal{S}(\mathbb{Z})}(\alpha,\beta) = \sum_{j=0}^{\infty} \frac{1}{2^j} \frac{\|\alpha - \beta\|_{\infty,j}}{1 + \|\alpha - \beta\|_{\infty,j}},$$

and $\mathcal{S}(\mathbb{Z})$ is in fact complete under this metric [21].

In the rest of this section, we will by the *Fourier transform*[2] of a test function refer to the following map:

[2]Not to be confused with the Fourier transform of, say, a Schwartz function on the line.

Definition 4.4 (Fourier transform of a periodic test function)**.** We denote by $\mathcal{F}$ the map which takes a periodic test function to its sequence of Fourier coefficients, that is

$$\mathcal{F} : \varphi \mapsto \{\hat{\varphi}(k)\}_{k\in\mathbb{Z}},$$

for $\varphi \in \mathcal{P}$.

Clearly the Fourier series of φ converges in the $L^2(-\pi,\pi)$-sense. However, as we will show in the next theorem, the far stronger property of convergence in $\mathcal{P}$ also holds. First we define the map which we call the *inverse Fourier transform* on the sequence space $\mathcal{S}(\mathbb{Z})$:

Definition 4.5 (Inverse Fourier transform on $\mathcal{S}(\mathbb{Z})$)**.** Let $\alpha = \{\alpha_k\}_{k\in\mathbb{Z}} \in \mathcal{S}(\mathbb{Z})$. The inverse Fourier transform is the map

$$\mathcal{F}^{-1} : \alpha \mapsto \sum_{k\in\mathbb{Z}} \alpha_k \Theta_k.$$

Theorem 4.2 (Fourier series of periodic test functions)**.** *The Fourier transform is a linear homeomorphism from $\mathcal{P}$ to $\mathcal{S}(\mathbb{Z})$. That is, $\mathcal{F}$ is a continuous function $\mathcal{P} \to \mathcal{S}(\mathbb{Z})$ with a continuous inverse $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{Z}) \to \mathcal{P}$ (with respect to the metrics $d_{\mathcal{P}}$ and $d_{\mathcal{S}(\mathbb{Z})}$). Any $\varphi \in \mathcal{P}$ can be written as a Fourier series*

$$\varphi = \sum_{k\in\mathbb{Z}} \hat{\varphi}(k) e^{ikx}$$

with convergence in $\mathcal{P}$.

Proof. The linearity of the Fourier transform $\mathcal{F}$ and the inverse Fourier transform $\mathcal{F}^{-1}$ are easily verified from the definitions.

Suppose $\varphi \in \mathcal{P}$. Then the sequence $\{\hat{\varphi}(k)\}_{k\in\mathbb{Z}}$ of Fourier coefficients of φ is in $\mathcal{S}(\mathbb{Z})$ by the following equality, where $j \in \mathbb{N}_0$ and $k \in \mathbb{Z}$:

$$\begin{aligned} |k|^j|\hat{\varphi}(k)| &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} \varphi(x) k^j e^{-ikx}\,dx\right| = \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} \varphi(x) D_x^j e^{-ikx}\,dx\right| \\ &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} (D_x^j \varphi(x)) e^{-ikx}\,dx\right| = |\widehat{\varphi^{(j)}}(k)| \end{aligned} \tag{4.3}$$

Hence $\sup_{k\in\mathbb{Z}} |k|^j|\hat{\varphi}(x)|$ is finite for all $j \in \mathbb{N}_0$ since $D^j\varphi \in \mathcal{P}$ for all $j \in \mathbb{N}_0$.

Now suppose $\{\alpha_k\}_{k\in\mathbb{Z}}$ is in $\mathcal{S}(\mathbb{Z})$, and consider the series

$$\psi := \sum_{k\in\mathbb{Z}} \alpha_k \Theta_k. \tag{4.4}$$

This series converges absolutely and uniformly by the Weierstrass M-test, since for $k \neq 0$,

$$|\alpha_k e^{ikx}| = |\alpha_k| \leq \frac{\|\alpha_k\|_{\infty,2}}{|k|^2} =: M_k,$$

and $\sum_{k\in\mathbb{Z}} M_k < \infty$ where we define $M_0 := |\alpha_0|$. Similarly, the term-wise derivative of any order $j \in \mathbb{N}$ of the series also converges absolutely and uniformly, since $|k^j \alpha_k| \le \frac{\|\alpha_k\|_{\infty,j+2}}{|k|^2}$ (for $k \ne 0$). Thus the infinite sum (4.4) converges in $\mathcal{P}$ to ψ (which is clearly smooth and 2π-periodic) by the definition of convergence in $\mathcal{P}$. This shows that $\mathcal{F}$ is an isomorphism $\mathcal{P}$ to $\mathcal{S}(\mathbb{Z})$.

In order to show the continuity of $\mathcal{F}$ as a map from $\mathcal{P}$ to $\mathcal{S}(\mathbb{Z})$, we let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}$ such that $\varphi_n \to 0$ in $\mathcal{P}$, i.e. for all $j \in \mathbb{N}_0$,

$$\|\varphi_n^{(j)}\|_\infty = \sup_{x\in\mathbb{R}} |\varphi_n^{(j)}(x)| \to 0 \text{ as } n \to \infty.$$

Since $\mathcal{F}$ maps $\varphi_n^{(j)}$ to $(ik)^j \hat{\varphi}(k)$ (see (4.3) above), we have for all $j \in \mathbb{N}_0$

$$\|\hat{\varphi}_n\|_{\infty,j} = \sup_{k\in\mathbb{Z}} |k|^j |\hat{\varphi}_n(k)| = \sup_{k\in\mathbb{Z}} |\widehat{\varphi_n^{(j)}}(k)| \le \|\varphi_n^{(j)}\|_\infty,$$

where the final inequality follows from the definition of the Fourier coefficients in (4.2). This implies $d_{\mathbb{S}(\mathbb{Z})}(\hat{\varphi}_n, 0) \to 0$ as $n \to \infty$, which proves the continuity of $\mathcal{F}$ as a map from $\mathcal{P}$ to $\mathcal{S}(\mathbb{Z})$.

Finally, let us prove continuity of $\mathcal{F}^{-1}$ as a map from $\mathcal{S}(\mathbb{Z})$ to $\mathcal{P}$. Let $\{\alpha^n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{S}(\mathbb{Z})$ such that $\alpha^n \to 0$ in $\mathcal{S}(\mathbb{Z})$, that is

$$d_{\mathcal{S}(\mathbb{Z})}(\alpha^n, 0) \to 0 \text{ as } n \to \infty.$$

The Fourier series given by

$$\mathcal{F}^{-1}(\alpha^n)(x) = \sum_{k\in\mathbb{Z}} \alpha_k^n e^{ikx}$$

determines a function in $\mathcal{P}$ since $\alpha^n \in \mathcal{S}(\mathbb{Z})$, and may be differentiated term-wise, which yields

$$\partial^j \mathcal{F}^{-1}(\alpha^n)(x) = \sum_{k\in\mathbb{Z}} \alpha_k^n (ik)^j e^{ikx}.$$

Thus for any $j \in \mathbb{N}_0$,

$$\begin{aligned}\|\partial^j \mathcal{F}^{-1}(\alpha^n)\|_\infty &\le \sum_{k\in\mathbb{Z}} |k|^j |\alpha_k^n| \\ &\le |\alpha_0^n| + \sum_{k\ne 0} |k|^{j+2} |\alpha_k^n| |k|^{-2} \\ &\le \|\alpha^n\|_{\infty,0} + \|\alpha^n\|_{\infty,j+2} \sum_{k\ne 0} |k|^{-2} \to 0,\end{aligned}$$

as $n \to \infty$ by the convergence of α^n to 0 in $\mathcal{S}(\mathbb{Z})$. □

From the proof we extract the following corollary, since the derivative of a periodic test function is again a periodic test function:

Corollary 4.2.1. *For $\varphi \in \mathcal{P}$, the Fourier series of $\varphi^{(j)}$ for any $j \in \mathbb{N}$ converges in $\mathcal{P}$ and the Fourier coefficients are given by*

$$\widehat{\varphi^{(j)}}(k) = (ik)^j \hat{\varphi}(k).$$

The final thing we study before moving on to the periodic distributions, is the Fourier series of products and convolutions of elements in $\mathcal{P}$. For periodic functions we define the convolution as follows:

Definition 4.6 (Convolution of periodic functions)**.** Given two 2π-periodic functions f and g on $(-\pi, \pi)$, we define the convolution $f * g$ to be the 2π-periodic function given by

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x-y)\, dy,$$

provided the integral converges almost everywhere.

For two functions $\varphi, \psi \in \mathcal{P}$, the convolution is clearly well-defined and in fact we have $\varphi * \psi \in \mathcal{P}$:

Lemma 4.3. *Let $\varphi, \psi \in \mathcal{P}$. Then $\varphi * \psi \in \mathcal{P}$ and the convolution is a continuous bilinear map.*

Proof. That $\varphi * \psi$ is 2π-periodic follows directly from the definition. The convolution $\psi * \psi$ is continuous by

$$\lim_{x \to a} (\varphi * \psi)(x) = \lim_{x \to a} \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(y)\psi(x-y)\, dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(y)\psi(a-y)\, dy = (\varphi * \psi)(a),$$

by the dominated convergence theorem (the sequence $\varphi(\cdot)\psi(a+\frac{1}{n}-\cdot)$ is dominated by $\sup_x |\varphi(x)| \sup_y |\psi(y)|$, which is of course locally integrable). That $\varphi * \psi$ is differentiable is seen by letting $a \in (-\pi, \pi)$ and considering the limit

$$\begin{aligned}
(\varphi * \psi)'(a) &= \lim_{h \to 0} \frac{(\varphi * \psi)(a+h) - (\varphi * \psi)(a)}{h} \\
&= \frac{1}{2\pi} \lim_{h \to 0} \int_{-\pi}^{\pi} \varphi(y) \frac{(\psi(a+h-y) - \psi(a-y))}{h}\, dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(y)\psi'(a-y)\, dy = (\varphi * \psi')(a),
\end{aligned}$$

again by Lebesgue's dominated convergence theorem. Note that by Fubini's theorem we have $\varphi * \psi = \psi * \varphi$ whenever the convolution is well-defined, thus we may also write $(\varphi * \psi)' = \varphi' * \psi$. Repeating the argument we get that $\varphi * \psi$ is smooth.

That the convolution is bilinear is clear from the definition. To have continuity, we would like to be able to estimate a seminorm of $\varphi * \psi$ by a seminorm of φ times a seminorm of ψ. By above we have $(\varphi * \psi)^{(j)} = \varphi^{(j)} * \psi$, thus

$$\begin{aligned}\|(\varphi * \psi)^{(j)}\|_\infty &\leq \frac{1}{2\pi} \sup_{x \in (-\pi,\pi)} \int_{-\pi}^{\pi} |\varphi^{(j)}(y)\psi(x-y)|\,dy \\ &\leq \frac{1}{2\pi}\|\varphi^{(j)}\|_\infty \int_{-\pi}^{\pi} \sup_{x\in(-\pi,\pi)} |\psi(x-y)|\,dy \\ &\leq \|\varphi^{(j)}\|_\infty \|\psi\|_\infty.\end{aligned}$$

□

The discrete convolution of two sequences will be an important concept when we consider the Fourier coefficients of a convolution.

Definition 4.7 (Discrete convolution)**.** For two complex sequences $\alpha = \{\alpha_k\}_{k\in\mathbb{Z}}$ and $\beta = \{\beta\}_{k\in\mathbb{Z}}$, we define the convolution $\alpha * \beta$ to be the sequence given by

$$(\alpha * \beta)_k = \sum_{j\in\mathbb{Z}} \alpha_j \beta_{k-j} \quad k \in \mathbb{Z},$$

whenever the sum on the right hand side converges.

The convolution of two rapidly decreasing sequences is clearly well-defined (in fact, as a consequence of the second result in the next theorem, it turns out to be a rapidly decreasing sequence). We have the following important result on the Fourier series of convolutions and products of elements in $\mathcal{P}$:

Theorem 4.4. *Let $\varphi, \psi \in \mathcal{P}$. Then we have*

$$\mathcal{F}(\varphi * \psi)(k) = \hat{\varphi}(k)\hat{\psi}(k) \quad \textit{and} \quad \mathcal{F}(\varphi\psi)(k) = (\hat{\varphi} * \hat{\psi})(k).$$

Proof. Applying Fubini's theorem we get the first result:

$$\begin{aligned}\mathcal{F}(\varphi * \psi)(k) &= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \varphi(y)\psi(x-y)e^{-ikx}\,dy\,dx \\ &= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi} \varphi(y)e^{-iky}\left(\int_{-\pi}^{\pi}\psi(x-y)e^{-ik(x-y)}\,dx\right)dy \\ &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi(y)e^{-iky}\,dy\right)\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\psi(z)e^{-ikz}\,dz\right)\end{aligned}$$

As for the second result, we have by Parseval's identity

$$\mathcal{F}(\varphi\psi)(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi(x)\psi(x)e^{-ikx}\,dx = \sum_{j\in\mathbb{Z}} \hat{\varphi}(j)\overline{\mathcal{F}(\overline{\psi}\Theta_k)(j)}.$$

By the identity $\widehat{\overline{\psi}}(k) = \overline{\hat{\psi}(-k)}$ we have

$$\mathcal{F}(\overline{\psi}\Theta_k)(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{\psi}(x)e^{-i(j-k)x}\,dx = \widehat{\overline{\psi}}(j-k) = \overline{\hat{\psi}(k-j)},$$

thus

$$\mathcal{F}(\varphi\psi)(k) = \sum_{j\in\mathbb{Z}}\hat{\varphi}(j)\hat{\psi}(k-j) = (\hat{\varphi} * \hat{\psi})(k).$$

□

4.3 Periodic distributions

Having established the necessary theory for the space of periodic test functions, we are ready to introduce the space of periodic distributions.

Definition 4.8 (Periodic distributions)**.** Let $\mathcal{P}'$ denote the continuous dual of $\mathcal{P}$. Here continuity means that $\langle T, \varphi_n\rangle \to \langle T, \varphi\rangle$ as $n \to \infty$ if $\varphi_n \to \varphi$ in $\mathcal{P}$. We call this the space of periodic distributions.

We define the following notion of convergence in $\mathcal{P}'$:

Definition 4.9 (Convergence in $\mathcal{P}'$)**.** We say a sequence $\{T_n\}_{n\in\mathbb{N}}$ in $\mathcal{P}'$ converges to $T \in \mathcal{P}'$ if

$$\lim_{n\to\infty}\langle T_n, \varphi\rangle = \langle T, \varphi\rangle$$

for all $\varphi \in \mathcal{P}$. We write $T_n \to T$ in $\mathcal{P}'$.

Example 4.1. The space of periodic distributions includes all $f \in L^1(-\pi,\pi)$, in the sense that any such f induces a linear functional T_f defined by

$$\langle T_f, \varphi\rangle = \int_{-\pi}^{\pi} f\varphi\,dx,$$

for $\varphi \in \mathcal{P}$. Clearly this is a well-defined linear functional on $\mathcal{P}$. To see that it is continuous, consider a sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}$ s.t. $\varphi_n \to 0$ in $\mathcal{P}$. Then we have

$$|\langle T_f, \varphi_n\rangle| \le \int_{-\pi}^{\pi}|f\varphi_n|\,dx \le \sup_{x\in(-\pi,\pi)}|\varphi_n(x)|\int_{-\pi}^{\pi}|f|\,dx \to 0$$

by $\varphi_n \to 0$ in $\mathcal{P}$.

Note that this examples implies $\mathcal{P} \subseteq \mathcal{P}'$, and also $L^p(-\pi,\pi) \subseteq \mathcal{P}'$ for all $p \ge 1$.

Example 4.2. The δ distribution, defined by

$$\langle \delta, \varphi\rangle = \varphi(0) \tag{4.5}$$

for $\varphi \in \mathcal{P}$, is a periodic distribution.

Like for distributions in $\mathcal{D}'(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$, we may define the distributional derivative of periodic distributions, and by the smoothness of functions in $\mathcal{P}$, all periodic distributions are infinitely differentiable:

Definition 4.10 (Distributional derivative)**.** By the distributional derivative in the sense of $\mathcal{P}'$, we mean the map D^k for $k \in \mathbb{N}$ such that for $T \in \mathcal{P}'$,

$$\langle D^k T, \varphi \rangle = (-1)^k \langle T, D^k \varphi \rangle$$

for all $\varphi \in \mathcal{P}$.

Remark 4.2. That the map D^k is well defined on $\mathcal{P}'$ follows from the fact that given $\varphi \in \mathcal{P}$, $D^k \varphi$ is also 2π-periodic and smooth for all $k \in \mathbb{N}_0$.

We can define the Fourier series of a periodic distribution by duality. Note that if $\varphi \in \mathcal{P}$, its Fourier coefficients are given by

$$\hat{\varphi}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(x) e^{-ikx} \, dx$$

for each $k \in \mathbb{Z}$. Thus if T_φ is the distribution corresponding to φ, it follows that

$$\hat{\varphi}(k) = \frac{1}{2\pi} \langle T_\varphi, \Theta_{-k} \rangle$$

for all $k \in \mathbb{Z}$, since $\Theta_{-k} \in \mathcal{P}$. This motivates the following definition:

Definition 4.11 (Fourier coefficients of periodic distributions)**.** If T is a periodic distribution, its Fourier coefficients are given by

$$\hat{T}(k) := \frac{1}{2\pi} \langle T, \Theta_{-k} \rangle$$

for each $k \in \mathbb{Z}$.

The Fourier transform on $\mathcal{P}'$ is defined in the following way:

Definition 4.12 (Fourier transform on $\mathcal{P}'$)**.** We denote by $\mathcal{F}$ the map that takes an element $T \in \mathcal{P}'$ to its sequence of Fourier coefficients, that is

$$\mathcal{F} : T \to \{\hat{T}(k)\}_{k \in \mathbb{Z}}.$$

We refer to $\mathcal{F}$ as the Fourier transform.

We have the following useful result on the Fourier transform of the derivative of an element in $\mathcal{P}'$:

Lemma 4.5 (Fourier transform of derivative)**.** *For $T \in \mathcal{P}'$ we have*

$$\widehat{T^{(n)}}(k) = (ik)^n \hat{T}(k)$$

for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$.

Proof. The result follows from a simple application of Definitions 4.10 and 4.12:

$$\widehat{T^{(n)}}(k) = \frac{1}{2\pi}\langle T^{(n)}, \Theta_{-k}\rangle = \frac{(-1)^n}{2\pi}\langle T, (-ik)^n\Theta_{-k}\rangle = \frac{(ik)^n}{2\pi}\langle T, \Theta_{-k}\rangle = (ik)^n\hat{T}(k)$$

□

Next we would like to prove convergence of the Fourier series of elements in $\mathcal{P}'$.

Theorem 4.6 (Fourier series of periodic distributions)**.** *Any $T \in \mathcal{P}'$ can be written as a Fourier series,*

$$T = \sum_{k\in\mathbb{Z}} \hat{T}(k)e^{ikx},$$

where the infinite sum converges in the sense of periodic distributions.

Proof. Given $T \in \mathcal{P}'$, define

$$T_n = \sum_{k=-n}^{n} \hat{T}(k)\Theta_k,$$

i.e. $\hat{T}_n(k) = \hat{T}(k)$ for $|k| \le n$, and $\hat{T}_n(k) = 0$ for $|k| > n$. Clearly $T_n \in \mathcal{P}$ for any $n \in \mathbb{N}_0$, and thus also $T_n \in L^2(-\pi,\pi)$. By applying Parseval's identity and the formula $\hat{\overline{\varphi}}(k) = \overline{\hat{\varphi}(-k)}$ we then get, for any $\varphi \in \mathcal{P}$,

$$\begin{aligned}\langle T_n, \varphi\rangle &= \int_{-\pi}^{\pi} T_n(x)\varphi(x)\,dx = 2\pi\sum_{k=-n}^{n}\hat{T}_n(k)\overline{\hat{\overline{\varphi}}(k)} = 2\pi\sum_{k=-n}^{n}\hat{T}(k)\hat{\varphi}(-k)\\ &= \sum_{k=-n}^{n}\langle T, \Theta_{-k}\rangle\hat{\varphi}(-k) = \langle T, \sum_{k=-n}^{n}\hat{\varphi}(k)\Theta_k\rangle \to \langle T, \varphi\rangle \quad \text{as } n \to \infty,\end{aligned}$$

by the continuity of T and convergence of the Fourier series of the periodic test function φ (cf. Theorem 4.2). This proves the convergence of the Fourier series of T in the sense of periodic distributions. □

We have now established that periodic distributions have well-defined Fourier transforms, and from the proof above we immediately extract the following useful characterisation of $T \in \mathcal{P}'$ in terms of its Fourier coefficients:

Corollary 4.6.1. *Let $T \in \mathcal{P}'$. For every $\varphi \in \mathcal{P}$,*

$$\langle T, \varphi \rangle = 2\pi \sum_{k \in \mathbb{Z}} \hat{T}(k)\hat{\varphi}(-k).$$

One may wonder about the range of the Fourier transform as a map on $\mathcal{P}'$. It turns out that $\mathcal{F}$ is a bijective map from $\mathcal{P}'$ to the space of sequences of *slow growth* (see the Theorem 4.8 below):

Definition 4.13 (Sequences of slow growth)**.** A complex sequence $\{\alpha_k\}_{k \in \mathbb{Z}}$ is said to be of slow growth if there exists $N > 0$ and $C > 0$ such that

$$|\alpha_k| \le C|k|^N \quad \forall k \in \mathbb{Z} \setminus \{0\}.$$

The set of all such sequences is denoted by $\mathcal{S}'(\mathbb{Z})$.

The behaviour of the sequences in $\mathcal{S}'(\mathbb{Z})$ is dual to the behaviour of the rapidly decaying sequences. The notation $\mathcal{S}'(\mathbb{Z})$ also indicates that the set of sequences of slow growth is the dual of $\mathcal{S}(\mathbb{Z})$. This can in fact be shown to be the case [21, p. 190], but we only include the proof that $\mathcal{S}'(\mathbb{Z}) \subseteq \mathcal{S}(\mathbb{Z})'$, since establishing the other inclusion is not vital to our presentation.

Lemma 4.7. *Any sequence of slow growth α determines a continuous linear functional on $\mathcal{S}(\mathbb{Z})$ by the formula*

$$\langle \alpha, \beta \rangle = \sum_{k \in \mathbb{Z}} \alpha_k \beta_k \quad \forall \beta \in \mathcal{S}(\mathbb{Z}). \tag{4.6}$$

Proof. The linear functional given by (4.6) is well-defined since

$$|\langle \alpha, \beta \rangle| \le \sum_{k \in \mathbb{Z}} |\alpha_k \beta_k| \le |\alpha_0 \beta_0| + C \sum_{k \ne 0} |\beta_k||k|^N \le |\alpha_0 \beta_0| + C \sum_{k \ne 0} \frac{\|\beta\|_{\infty, N+2}}{|k|^2} < \infty.$$

Let $\beta^n \to 0$ in $\mathcal{S}(\mathbb{Z})$. Then by the estimate above we get $\langle \alpha, \beta^n \rangle \to 0$ by the definition of convergence in $\mathcal{S}(\mathbb{Z})$. This proves the continuity of the linear functional given by α.

□

We define the inverse Fourier transform on $\mathcal{S}'(\mathbb{Z})$:

Definition 4.14 (Inverse Fourier transform on $\mathcal{S}'(\mathbb{Z})$)**.** For a sequence $\alpha \in \mathcal{S}'(\mathbb{Z})$, we define the inverse Fourier transform as the map

$$\mathcal{F}^{-1} : \alpha \mapsto T_\alpha$$

where T_α is the linear functional given by $\langle T_\alpha, \varphi \rangle = 2\pi \sum_{k \in \mathbb{Z}} \alpha_k \hat{\varphi}(-k)$ for $\varphi \in \mathcal{P}$.

The linear functional T_α is indeed well-defined, moreover it is continuous, as the next theorem shows:

Theorem 4.8. *The Fourier transform is a homeomorphism from $\mathcal{P}'$ to $\mathcal{S}'(\mathbb{Z})$.*

Proof. We need the intermediate result that for any $T \in \mathcal{P}'$, there exists $N \in \mathbb{N}_0$ and $C > 0$ such that

$$|\langle T, \varphi\rangle| \le C \sum_{j=0}^{N} \|D^j \varphi\|_\infty \tag{4.7}$$

for all $\varphi \in \mathcal{P}$. We say that T has *finite order*. Following the argument in the proof of Lemma 2.4.3 in [38], we set out to prove the result by contradiction: Assume that for any $N \in \mathbb{N}_0$ there is a $\varphi_N \in \mathcal{P}$ such that

$$|\langle T, \varphi_N\rangle| \ge N \sum_{j=0}^{N} \|D^j \varphi_N\|_\infty. \tag{4.8}$$

Define

$$\psi_N := \frac{1}{N}\left(\sum_{j=0}^{N} \|D^j \varphi_N\|_\infty\right)^{-1} \varphi_N.$$

For any fixed $n \in \mathbb{N}_0$, we have

$$\|D^n \psi_N\|_\infty \le \frac{1}{N}$$

for N sufficiently large. Thus for each k, $D^k\psi_N \to 0$ uniformly as $N \to \infty$, which shows that $\psi_N \to 0$ in $\mathcal{P}$. By the continuity of T we then have $\langle T, \psi_N\rangle \to 0$ as $N \to \infty$. But $|\langle T, \psi_N\rangle| \ge 1$ for all N by (4.8), which gives the desired contradiction (that is, for some choice of N there will not exist any $\varphi_N \in \mathcal{P}$ such that (4.8) holds, thus (4.7) must hold with $C = N$).

Hence we know that for the periodic distribution T, there exists $N > 0$ and $C > 0$ such that $|\langle T, \varphi\rangle| \le C \sum_{j=0}^{N} \|D^j\varphi\|_\infty$ for all $\varphi \in \mathcal{P}$. Taking $\varphi = \Theta_{-k}$ we get

$$|\hat{T}(k)| = \frac{1}{2\pi}|\langle T, \Theta_{-k}\rangle| \le \frac{1}{2\pi} C \sum_{j=0}^{N} |k|^j \le c|k|^N,$$

thus the Fourier coefficents $\hat{T}(k)$ form a sequence of slow growth.

In order to prove the converse, we assume $\alpha = \{\alpha_k\}_{k\in\mathbb{Z}} \in \mathcal{S}'(\mathbb{Z})$. The inverse Fourier transform sends α to the linear functional T_α given by the formula

$$\langle T_\alpha, \varphi\rangle = 2\pi \sum_{k\in\mathbb{Z}} \alpha_k \hat{\varphi}(-k), \quad \varphi \in \mathcal{P}.$$

This is a well-defined functional $\mathcal{P} \to \mathbb{C}$ since

$$\sum_{k\in\mathbb{Z}} |\alpha_k \hat{\varphi}(-k)| = |\alpha_0 \hat{\varphi}(0)| + \sum_{k\neq 0} |\alpha_k \hat{\varphi}(-k)| \leq |\alpha_0 \hat{\varphi}(0)| + C \sum_{k\neq 0} |k|^N |\hat{\varphi}(k)|$$
$$\leq |\alpha_0 \hat{\varphi}(0)| + C \sum_{k\neq 0} \frac{\|\hat{\varphi}\|_{\infty,N+2}}{|k|^2} < \infty.$$

We claim that $T_\alpha \in \mathcal{P}'$. Indeed

$$\langle T_\alpha, \varphi \rangle = 2\pi \sum_{k\in\mathbb{Z}} \alpha_k \hat{\varphi}(-k) = \sum_{k\in\mathbb{Z}} \alpha_k \int_{-\pi}^{\pi} \varphi(x) e^{ikx}\, dx$$
$$= \int_{-\pi}^{\pi} \varphi(x) \left(\sum_{k\in\mathbb{Z}} \alpha_k e^{ikx} \right) dx =: \int_{-\pi}^{\pi} \varphi(x) \psi(x)\, dx$$

where $\psi \in \mathcal{P}$ by Theorem 4.2. Thus since $\mathcal{P} \subseteq \mathcal{P}'$ (see Example 4.1), the linear functional T_α is also continuous, hence $T_\alpha \in \mathcal{P}'$. Moreover, we have

$$\hat{T}_\alpha(j) = \frac{1}{2\pi} \langle T_\alpha, \Theta_{-j} \rangle = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \alpha_k \int_{-\pi}^{\pi} e^{-i(j-k)x}\, dx = \alpha_j.$$

This proves that the Fourier transform is a bijection from $\mathcal{P}'$ to $\mathcal{S}'(\mathbb{Z})$. In order to prove continuity of $\mathcal{F}$, we let $T_n \to T$ in $\mathcal{P}'$. Since all elements in $\mathcal{S}'(\mathbb{Z})$ are continuous linear functionals on $\mathcal{S}(\mathbb{Z})$, we say $\hat{T}_n \to \hat{T}$ in $\mathcal{S}'(\mathbb{Z})$ if

$$\langle \hat{T}_n, \beta \rangle \to \langle \hat{T}, \beta \rangle \quad \forall \beta \in \mathcal{S}(\mathbb{Z}).$$

We have

$$\langle \hat{T}_n, \beta \rangle = \sum_{k\in\mathbb{Z}} \hat{T}_n(k) \beta_k = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \langle T_n, \Theta_{-k} \rangle \beta_k = \frac{1}{2\pi} \langle T_n, \sum_{k\in\mathbb{Z}} \beta_k \Theta_{-k} \rangle$$
$$\to \frac{1}{2\pi} \langle T, \sum_{k\in\mathbb{Z}} \beta_k \Theta_{-k} \rangle = \sum_{k\in\mathbb{Z}} \frac{1}{2\pi} \langle T, \Theta_{-k} \rangle \beta_k = \sum_{k\in\mathbb{Z}} \hat{T}(k) \beta_k$$

by $\sum_{k\in\mathbb{Z}} \beta_k \Theta_{-k} \in \mathcal{P}$ since $\beta \in \mathcal{S}(\mathbb{Z})$ and the definition of convergence in $\mathcal{P}'$.

Next suppose $\alpha^n \to \alpha$ in $\mathcal{S}'(\mathbb{Z})$. For any $\varphi \in \mathcal{P}$ we have $\hat{\varphi} \in \mathcal{S}(\mathbb{Z})$, hence by the definition of convergence in $\mathcal{S}'(\mathbb{Z})$

$$\langle T_{\alpha^n}, \varphi \rangle = 2\pi \sum_{k\in\mathbb{Z}} \alpha_k^n \hat{\varphi}(-k) = 2\pi \langle \alpha^n, \widetilde{\hat{\varphi}} \rangle \to 2\pi \langle \alpha, \widetilde{\hat{\varphi}} \rangle = 2\pi \sum_{k\in\mathbb{Z}} \alpha_k \hat{\varphi}(-k) = \langle T_\alpha, \varphi \rangle,$$

in other words $T_{\alpha^n} \to T_\alpha$ in $\mathcal{P}'$.

□

To wrap up our study of the periodic distributions, we would like to prove an important result on the convolution of elements in $\mathcal{P}'$. First let us figure out the right definition for the convolution of a distribution $T \in \mathcal{P}'$ and a test function $\varphi \in \mathcal{P}$. The motivation will come from the convolution of a continuous 2π-periodic function f and $\varphi \in \mathcal{P}$. If $\sim$ denotes the reflection operator, that is $\tilde{\varphi}(x) = \varphi(-x)$, then

$$(f * \varphi)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)\varphi(x-y)\,dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)(\tau_x \tilde{\varphi}(y))\,dy = \frac{1}{2\pi} \langle f, \tau_x \tilde{\varphi} \rangle.$$

Definition 4.15 (Convolution of $\mathcal{P}'$ and $\mathcal{P}$)**.** Let $T \in \mathcal{P}'$ and $\varphi \in \mathcal{P}$. The convolution $f * \varphi$ of f and φ is the function

$$(T * \varphi)(x) = \frac{1}{2\pi} \langle T, \tau_x \tilde{\varphi} \rangle, \quad x \in \mathbb{R}.$$

For such convolutions we have the following result:

Lemma 4.9. *For $T \in \mathcal{P}'$ and $\varphi \in \mathcal{P}$ we have $T * \varphi \in \mathcal{P}$ and*

$$\mathcal{F}(T * \varphi)(k) = \hat{f}(k)\hat{\varphi}(k)$$

for all $k \in \mathbb{Z}$.

Proof. The convolution $T * \varphi$ is easily verified to be 2π-periodic by direct computation, and it is also continuous, since

$$\lim_{x \to a} (T * \varphi)(x) = \frac{1}{2\pi} \lim_{x \to a} \langle T, \tau_x \tilde{\varphi} \rangle = \frac{1}{2\pi} \langle T, \lim_{x \to a} \tau_x \tilde{\varphi} \rangle = (T * \varphi)(a),$$

where we have used to continuity of T (one can check that the sequence $\{\tau_{a+\frac{1}{n}} \tilde{\varphi}\}_n$ converges to $\tau_a \tilde{\varphi}$ in $\mathcal{P}$ using Definition 4.2). The convolution is also differentiable, as for $a \in (-\pi, \pi)$ we similarly have

$$\lim_{h \to 0} \frac{(T * \varphi)(a+h) - (T * \varphi)(a)}{h} = \frac{1}{2\pi} \lim_{h \to 0} \langle T, \frac{(\tau_{a+h} - \tau_a)\tilde{\varphi}}{h} \rangle = T * \varphi'.$$

Repeating the argument shows that the convolution is infinitely differentiable, thus $T * \varphi \in \mathcal{P}$.

We get the identity of the lemma from the following computation:

$$\begin{aligned}
\mathcal{F}(T * \varphi)(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\langle T, \tau_x\tilde{\varphi}\rangle e^{-ikx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\langle T, \tau_x\tilde{\varphi}e^{-ikx}\rangle\,dx \\
&= \frac{1}{2\pi}\langle T, \int_{-\pi}^{\pi}\frac{1}{2\pi}\tau_x\tilde{\varphi}e^{-ikx}\,dx\rangle \\
&= \frac{1}{2\pi}\langle T, \Theta_{-k}\frac{1}{2\pi}\int_{-\pi}^{\pi}\tau_x\left(\widetilde{\varphi\Theta_{-k}}\right)\,dx\rangle \\
&= \frac{1}{2\pi}\langle T, \Theta_{-k}\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi(y)e^{-iky}\,dy\rangle \\
&= \frac{1}{2\pi}\langle T, \Theta_{-k}\rangle\hat{\varphi}(k) \\
&= \hat{T}(k)\hat{\varphi}(k)
\end{aligned}$$

□

Our goal now is to prove a similar result for convolution of two elements in $\mathcal{P}'$. We start by introducing an alternative definition of the convolution of $T \in \mathcal{P}'$ and $\varphi \in \mathcal{P}$. The definition is now motivated by the *distribution* produced by convolving a continuous 2π-periodic function f with $\varphi \in \mathcal{P}$: We know that $f * \varphi$ is continuous (even more, it is infinitely differentiable) thus it defines a regular periodic distribution, and for $\psi \in \mathcal{P}$ we have by Fubini's theorem

$$\begin{aligned}
\langle f * \varphi, \psi\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{\pi} f(y)\varphi(x-y)\,dy\right)\psi(x)\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{\pi}\psi(x)\varphi(x-y)\,dx\right)f(y)\,dy \\
&= \int_{-\pi}^{\pi}(\psi * \tilde{\varphi})(y)f(y)\,dy = \langle f, \tilde{\varphi} * \psi\rangle.
\end{aligned}$$

This computation lead us to the following alternative definition:

Definition 4.16 (Convolution of $\mathcal{P}'$ and $\mathcal{P}$ II)**.** For $T \in \mathcal{P}'$ and $\varphi \in \mathcal{P}$, we define the convolution $T *' \varphi$ to be the periodic distribution with action

$$\langle T *' \varphi, \psi\rangle = \langle T, \tilde{\varphi} * \psi\rangle \tag{4.9}$$

for all $\psi \in \mathcal{P}$.

That (4.9) defines a periodic distribution follows from the fact that $\tilde{\varphi} * \psi \in \mathcal{P}$ and that the convolution is continuous on $\mathcal{P}$ (see Lemma 4.3). Of course we ought to prove that Definitions 4.15 and 4.16 are equivalent:

Proposition 4.10. *For $T \in \mathcal{P}'$ and $\varphi \in \mathcal{P}$, we have*

$$\langle T * \varphi, \psi \rangle = \langle T *' \varphi, \psi \rangle$$

for all $\psi \in \mathcal{P}$.

Proof. By direct computation we get

$$\begin{aligned}\langle T * \varphi, \psi \rangle &= \langle \frac{1}{2\pi} \langle T, \tau . \tilde{\varphi} \rangle, \psi \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle T, \tau_x \tilde{\varphi} \rangle (x) \psi(x)\, dx \\ &= \langle T, \frac{1}{2\pi} \int_{-\pi}^{\pi} (\tau_x \tilde{\varphi}) \psi(x)\, dx \rangle = \langle T, \tilde{\varphi} * \psi \rangle.\end{aligned}$$

□

From now on we use the symbol $*$ to denote the convolution using either definition. Using Definition 4.16 we may introduce the convolution of two elements in $\mathcal{P}'$:

Definition 4.17 (Convolution of periodic distributions)**.** Let $T, S \in \mathcal{P}'$. The convolution $T * S$ is the periodic distribution defined by the formula

$$\langle T * S, \varphi \rangle = \langle T, \tilde{S} * \varphi \rangle$$

for all $\varphi \in \mathcal{P}$.

Remark 4.3. The reflection operator $\sim$ is defined as one would expect on $\mathcal{P}'$,

$$\langle \tilde{S}, \varphi \rangle = \langle S, \tilde{\varphi} \rangle \quad \text{for } \varphi \in \mathcal{P}.$$

Remark 4.4. The convolution $T * S$ is well-defined as a periodic distribution: We have $\tilde{S} * \varphi \in \mathcal{P}$ by Lemma 4.9, thus $T * S$ is a well-defined linear functional on $\mathcal{P}$ by $T \in \mathcal{P}'$. Moreover, that the linear functional $T * S$ is continuous can be verified from Definitions 4.15 and 4.17 by letting $\varphi_n \to 0$ in $\mathcal{P}$ and using the fact that T and S are continuous.

We are now ready to prove the following important theorem on the Fourier series of the convolution of periodic distributions:

Theorem 4.11. *For $T, S \in \mathcal{P}'$ we have*

$$\mathcal{F}(T * S)(k) = \hat{T}(k) \hat{S}(k).$$

Proof. By definition we have $\hat{T}(k) = \frac{1}{2\pi}\langle T, \Theta_{-k}\rangle$ for $T \in \mathcal{P}'$ and

$$\langle T * S, \varphi\rangle = \langle T, \tilde{S} * \varphi\rangle = \langle T, \frac{1}{2\pi}\langle \tilde{S}, \tau_x\tilde{\varphi}\rangle\rangle,$$

for all $\varphi \in \mathcal{P}$. Thus

$$\begin{aligned}\mathcal{F}(T * S)(k) &= \frac{1}{2\pi}\langle T, \tilde{S} * \Theta_{-k}\rangle = \frac{1}{2\pi}\langle T, \frac{1}{2\pi}\langle \tilde{S}, \tau_x\Theta_k\rangle\rangle \\ &= \frac{1}{2\pi}\langle T, \frac{1}{2\pi}\langle S, \widetilde{\tau_x\Theta_k}\rangle\rangle = \frac{1}{2\pi}\langle T, \frac{1}{2\pi}\langle S, \Theta_{-k}\rangle e^{-ikx}\rangle \\ &= \frac{1}{2\pi}\langle T, \hat{S}(k)e^{-ikx}\rangle = \hat{T}(k)\hat{S}(k).\end{aligned}$$

□

As we cannot in general define the product of distributions, we don't have a general result for $\mathcal{P}'$ analogue to the second equality in Theorem 4.4. However, we do have such a result for periodic distributions with high enough Sobolev regularity, as we shall see in Lemma 4.18.

4.4 Periodic Sobolev spaces $H^s(\mathbb{T})$ for $s \in \mathbb{R}$

We are now just about ready to introduce the periodic Sobolev spaces. First, however, we make a remark on notation. Any 2π-periodic function f on the line may be identified, in a natural way, with a function g on the unit circle $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ by $f(t) = g(e^{it})$ for $t \in (-\pi, \pi)$. In higher dimensions, one can identify periodic functions with functions on the n-torus $\mathbb{T}^n$, defined as the product of n circles, $\mathbb{T}^n = S^1 \times \cdots \times S^1$. In one dimension we have $\mathbb{T} = S^1$. This explains the notation $H^s(\mathbb{T})$.

Definition 4.18 (Periodic Sobolev spaces)**.** For $s \in \mathbb{R}$, we define the periodic Sobolev spaces by

$$H^s(\mathbb{T}) = \{f \in \mathcal{P}' \;:\; \{\langle k\rangle^s \hat{f}(k)\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})\},$$

along with the periodic Sobolev norm

$$\|f\|_{H^s(\mathbb{T})} = (2\pi)^{\frac{1}{2}}\|\langle\cdot\rangle^s \hat{f}\|_{l^2(\mathbb{Z})} = (2\pi)^{\frac{1}{2}}\left(\sum_{k\in\mathbb{Z}}\langle k\rangle^{2s}\left|\hat{f}(k)\right|^2\right)^{1/2},$$

and the periodic Sobolev inner product

$$\langle f, g\rangle_{H^s(\mathbb{T})} = 2\pi\sum_{k\in\mathbb{Z}}\langle k\rangle^{2s}\hat{f}(k)\overline{\hat{g}(k)} = 2\pi\langle\langle\cdot\rangle^s\hat{f}, \langle\cdot\rangle^s\hat{g}\rangle_{l^2(\mathbb{Z})},$$

for $f, g \in H^s(\mathbb{T})$.

Theorem 4.12. *$H^s(\mathbb{T})$ is a Hilbert space for $s \in \mathbb{R}$.*

Proof. We will here only show that $H^s(\mathbb{T})$ for $s \geq 0$ is Hilbert. That $H^s(\mathbb{T})$ with s negative is complete will follow from Theorem 4.13, where we show that for $s \in \mathbb{R}$, $H^{-s}(\mathbb{T})$ can be identified with the dual space of $H^s(\mathbb{T})$. In the rest of this proof we therefore set $s \geq 0$.

That $H^s(\mathbb{T})$ is a Hilbert space follows from the fact that the sequence space $l^2(\mathbb{Z})$ is Hilbert. First of all, it is easily verified that $H^s(\mathbb{T})$ satisfies the vector space axioms. Secondly, all the properties that $\langle \cdot, \cdot \rangle_{H^s(\mathbb{T})}$ needs to satisfy in order to qualify as an inner product follow from the corresponding properties of $\langle \cdot, \cdot \rangle_{l^2(\mathbb{Z})}$.

Finally, that $H^s(\mathbb{T})$ is complete follows from the completeness of $l^2(\mathbb{Z})$ by the following argument: Suppose the sequence $\{f_n\}_{n \in \mathbb{N}}$ is Cauchy in $H^s(\mathbb{T})$. Then the sequence $\{\left(\langle k \rangle^s \hat{f}_n(k)\right)_{k \in \mathbb{Z}}\}_{n \in \mathbb{N}}$ is Cauchy in $l^2(\mathbb{Z})$. By the completeness of $l^2(\mathbb{Z})$, there exists a sequence $(g_k)_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ such that $\left(\langle k \rangle^s \hat{f}_n(k)\right)_{k \in \mathbb{Z}} \to (g_k)_{k \in \mathbb{Z}}$ in $l^2(\mathbb{Z})$ as $n \to \infty$. If $s = 0$, the proof is done. Assume $s > 0$. That $(g_k)_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ implies that $\left(\hat{f}(k)\right)_{k \in \mathbb{Z}} := \left(\langle k \rangle^{-s} g_k\right)_{k \in \mathbb{Z}}$ is in $l^2(\mathbb{Z})$. Then we have by the convergence of $\left(\langle k \rangle^s \hat{f}_n(k)\right)_{k \in \mathbb{Z}}$ to $(g_k)_{k \in \mathbb{Z}}$ in $l^2(\mathbb{Z})$ that

$$f_n \to f \text{ in } H^s(\mathbb{T}) \text{ as } n \to \infty,$$

where we have defined $f := \sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ikx}$ (f is in $\mathcal{P}'$ by $(\hat{f}(k))_{k \in \mathbb{Z}} \in S'(\mathbb{Z})$). □

Like for the Sobolev spaces over $\mathbb{R}^d$, we can characterise the Sobolev space $H^{-s}(\mathbb{T})$ as the dual of $H^s(\mathbb{T})$:

Theorem 4.13. *For $s \in \mathbb{R} \setminus \{0\}$, $H^{-s}(\mathbb{T})$ is the dual of $H^s(\mathbb{T})$. $H^0(\mathbb{T})$ is self-dual.*

Proof. That $H^0(\mathbb{T})$ is self-dual follows from the fact that is is really the space $L^2(-\pi, \pi)$ in disguise (thus we sometimes denote $H^0(\mathbb{T})$ by $L^2(\mathbb{T})$). More precisely, the two spaces are isometrically isomorphic: Given a distribution $f \in H^0(\mathbb{T})$, we have by definition $\{\hat{f}(k)\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$. Then the function $f = \sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ikx}$ is in $L^2(-\pi, \pi)$. Conversely, given a function $g \in L^2(-\pi, \pi)$, then $g \in \mathcal{P}'$ as we saw in Example 4.1, and we have (by Parseval's identity) that $g \in H^0(\mathbb{T})$.

Let $s \neq 0$ and $f \in H^{-s}(\mathbb{T})$. We want to show that the linear functional

$$L_f(g) := \sum_{k \in \mathbb{Z}} \hat{f}(k) \overline{\hat{g}(k)}$$

for $g \in H^s(\mathbb{T})$ is continuous, in the sense that if $g_n \to 0$ in $H^s(\mathbb{T})$ then $L_f(g_n) \to 0$.

By the Cauchy-Schwarz inequality we have

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \hat{f}(k)\overline{\hat{g}_n(k)} &= \sum_{k\in\mathbb{Z}} \langle k\rangle^{-s} \hat{f}(k)\overline{\langle k\rangle^{s} \hat{g}_n(k)} \\ &\leq \left(\sum_{k\in\mathbb{Z}} \langle k\rangle^{-2s} |\hat{f}(k)|^2\right)^{\frac{1}{2}} \left(\sum_{k\in\mathbb{Z}} \langle k\rangle^{2s} |\hat{g}_n(k)|^2\right)^{\frac{1}{2}} \\ &= \frac{1}{2\pi}\|f\|_{H^{-s}(\mathbb{T})}\|g\|_{H^s(\mathbb{T})} \to 0\end{aligned}$$

as $n \to \infty$. Thus we have the embedding $H^{-s}(\mathbb{T}) \hookrightarrow H^s(\mathbb{T})'$.

Now assume $L \in H^s(\mathbb{T})'$. By the Riesz representation theorem we may identify L with an $f_L \in H^s(\mathbb{T})$, in the sense that for all $g \in H^s(\mathbb{T})$,

$$\langle L, g\rangle = \langle g, f_L\rangle_{H^s(\mathbb{T})} = 2\pi \sum_{k\in\mathbb{Z}} \hat{g}(k)\overline{\langle k\rangle^{2s} \hat{f}_L(k)}.$$

Define $\hat{f}(k) = \overline{\langle k\rangle^{2s} \hat{f}_L(k)}$, and $f := \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikx}$. Then $f \in H^{-s}(\mathbb{T})$ and

$$\langle L, g\rangle = 2\pi \sum_{k\in\mathbb{Z}} \hat{g}(k)\hat{f}(k)$$

for all $g \in H^s(\mathbb{T})$, and thus we have the other embedding, $H^s(\mathbb{T})' \hookrightarrow H^{-s}(\mathbb{T})$. □

We saw in the proof above that $H^0(\mathbb{T}) = L^2(\mathbb{T})$ can be identified with the space $L^2(-\pi, \pi)$. Similar to the regular Sobolev spaces over $\mathbb{R}^d$, we may classify $H^s(\mathbb{T})$ for $s \geq 0$ as subspaces of $L^2(-\pi, \pi)$ with functions of finite Sobolev norm:

Theorem 4.14. *For $s \geq r$, we have the dense embedding*

$$H^s(\mathbb{T}) \hookrightarrow H^r(\mathbb{T}),$$

which is continuous in the sense that $\|f\|_{H^r(\mathbb{T})} \leq \|f\|_{H^s(\mathbb{T})}$ for all $f \in H^s(\mathbb{T})$. In particular, for $s \geq 0$ we have $H^s(\mathbb{T}) \subseteq L^2(-\pi, \pi)$

Proof. Clearly $\{\langle k\rangle^s \hat{f}(k)\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$ implies that $\{\langle k\rangle^r \hat{f}(k)\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$ for $r \leq s$. Thus we have the embedding $H^s(\mathbb{T}) \hookrightarrow H^r(\mathbb{T})$ which is continuous in the sense that $\|f\|_{H^r(\mathbb{T})} \leq \|f\|_{H^s(\mathbb{T})}$ for all $f \in H^s(\mathbb{T})$.

We will show that this embedding is necessarily dense by proving that $\mathcal{P}$ is a dense subset of $H^s(\mathbb{T})$ for all $s \in \mathbb{R}$. That $\mathcal{P} \subseteq H^s(\mathbb{T})$ for $s \leq 0$ is clear by $\mathcal{P} \subseteq L^2(-\pi, \pi) \subseteq \mathcal{P}'$. That $\mathcal{P} \subseteq H^s(\mathbb{T})$ for $s > 0$ can be seen for instance from Lemma 4.5 (or see Theorem 4.15 below). To show that $\mathcal{P}$ is a dense subset of $H^s(\mathbb{T})$ for all $s \in \mathbb{R}$, suppose $f \in H^s(\mathbb{T})$ and define the function $f_n := \sum_{k=-n}^{n} \hat{f}(k)e^{ikx}$. Then $f_n \in \mathcal{P}$ for all $n \in \mathbb{N}_0$ and we have

$$\|f - f_n\|^2_{H^s(\mathbb{T})} = \sum_{|k|>n} \langle k\rangle^{2s} |\hat{f}(k)|^2 \to 0$$

as $n \to \infty$ since $f \in H^s(\mathbb{T})$. □

As with the Sobolev spaces on $\mathbb{R}^d$, we may for $s = m \in \mathbb{N}_0$ describe the periodic Sobolev spaces as spaces of functions with m square integrable (distributional) derivatives:

Theorem 4.15. *Let $m \in \mathbb{N}_0$. Then $f \in H^m(\mathbb{T})$ if and only if $D^j f \in L^2(-\pi,\pi)$ for $j \in \{0,1,2,\cdots,m\}$, where the derivatives are taken in the sense of $\mathcal{P}'$. Moreover, the norms $\|f\|_{H^s(\mathbb{T})}$ and*

$$\|f\|_{W_2^m(\mathbb{T})} := \left(\sum_{j=0}^m \|D^j f\|^2_{L^2(-\pi,\pi)}\right)^{\frac{1}{2}} = \sqrt{2\pi}\left(\sum_{j=0}^m \|\widehat{f^{(j)}}\|^2_{l^2(\mathbb{Z})}\right)^{\frac{1}{2}}$$

are equivalent.

Proof. By Theorem 4.5 we have for $T \in \mathcal{P}'$

$$|\widehat{T^{(j)}}(k)| = |(ik)^j \hat{T}(k)| \le \langle k\rangle^m |\hat{T}(k)|$$

for all $j \in \{0,1,\cdots,m\}$. If $f \in H^m(\mathbb{T})$, we have $\{\langle k\rangle^m \hat{f}(k)\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$, thus $f^{(j)} \in H^0(\mathbb{T}) = L^2(-\pi,\pi)$ for all $j \in \{0,1,\cdots,m\}$. Furthermore,

$$\|f\|_{W_2^m(\mathbb{T})} = \sqrt{2\pi}\left(\sum_{j=0}^m \|\widehat{f^{(j)}}\|^2_{l^2(\mathbb{Z})}\right)^{\frac{1}{2}} \le \sqrt{2\pi(m+1)}\|f\|_{H^m(\mathbb{T})}.$$

Now assume $f \in P'$ is such that $f^{(j)} \in L^2(-\pi,\pi)$ for all $j \in \{0,1,\cdots,m\}$. Then by Parseval's identity and Theorem 4.5 we have for all such j

$$\|\widehat{f^{(j)}}(k)\|^2_{l^2(\mathbb{Z})} = \|(ik)^j \hat{f}(k)\|^2_{l^2(\mathbb{Z})} = \frac{1}{2\pi}\|f^{(j)}\|^2_{L^2(-\pi,\pi)} < \infty,$$

which by $\langle k\rangle^j = (1+|k|^2)^{j/2} \le C_j|k|^j \ \forall k \in \mathbb{Z}$ for some constant C_j implies that $\{\langle k\rangle^j \hat{f}(k)\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$ for all $j \in \{0,1,\cdots,m\}$. Thus $f \in H^m(\mathbb{T})$. Also, we have

$$\begin{aligned}\|f\|_{H^m(\mathbb{T})} &= (2\pi)^{\frac{1}{2}}\left(\sum_{k\in\mathbb{Z}}(1+|k|^2)^m|\hat{f}(k)|^2\right)^{\frac{1}{2}} = (2\pi)^{\frac{1}{2}}\left(\sum_{k\in\mathbb{Z}}\left(\sum_{j=0}^m\binom{m}{j}|k|^{2j}\right)|\hat{f}(k)|^2\right)^{\frac{1}{2}}\\ &\le (2\pi)^{\frac{1}{2}}\left(\max_{j\in\{0,\cdots,m\}}\binom{m}{j}\right)^{\frac{1}{2}}\left(\sum_{j=0}^m\sum_{k\in\mathbb{Z}}|k|^{2j}|\hat{f}(k)|^2\right)^{\frac{1}{2}} = c_m\left(\sum_{j=0}^m\sum_{k\in\mathbb{Z}}|\widehat{f^{(j)}}(k)|^2\right)^{\frac{1}{2}}\\ &= c_m\left(\sum_{j=0}^m\|\widehat{f^{(j)}}(k)\|^2_{l^2(\mathbb{Z})}\right)^{\frac{1}{2}} = \frac{c_m}{\sqrt{2\pi}}\|f\|_{W_2^m(\mathbb{T})}.\end{aligned}$$

□

Next we prove the important Sobolev embedding theorem for periodic Sobolev spaces:

Theorem 4.16 (Periodic Sobolev embedding theorem). *Let $k \in \mathbb{N}_0$. For $s > k+1/2$ we have the embedding*

$$H^s(\mathbb{T}) \hookrightarrow BC^k(\mathbb{T}), \tag{4.10}$$

which is continuous in the sense that

$$\|f\|_{BC^k(\mathbb{T})} := \sum_{j=0}^{k} \|f^{(j)}\|_{L^\infty} \le C_{sk}\|[f]\|_{H^s(\mathbb{T})}.$$

Remark 4.5. This embedding should be interpreted to mean that in each equivalence class of functions $[f] \in H^s(\mathbb{T})$ for $s > k+1/2$, there is a representative function $f \in BC^k(\mathbb{T})$.

Proof. Let $k \in \mathbb{N}_0$ and suppose $f \in H^s(\mathbb{T})$ for $s > k + 1/2$. Then

$$\sum_{k\in\mathbb{Z}} \left|\hat{f}(k)\right| = \sum_{k\in\mathbb{Z}} \langle k\rangle^{-s} \langle k\rangle^{s} |\hat{f}(k)| \le \left(\sum_{k\in\mathbb{Z}} \langle k\rangle^{2s} |\hat{f}(k)|^2\right)^{\frac{1}{2}} \left(\sum_{k\in\mathbb{Z}} \langle k\rangle^{-2s}\right)^{\frac{1}{2}} < \infty,$$

by the Cauchy-Schwarz inequality. The first sum converges due to $f \in H^s(\mathbb{T})$, while the second sum converges due to $s > 1/2$. Then

$$|f(x)| = \left|\sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikx}\right| < \infty \quad \forall x \in (-\pi, \pi),$$

and since the terms in the series defining f are continuous periodic functions, this Fourier series converges uniformly to a continuous periodic function by the Weierstrass M-test (set $M_k := \hat{f}(k)$). Thus we have $f \in BC(\mathbb{T})$, and the embedding (4.10) follows from making similar estimates for $f^{(j)}$ for $j \in \{1, 2, \cdots, k\}$ by applying Theorem 4.5. Using this argument the inequality of the theorem also follows. □

The final result we would like to prove for the periodic Sobolev spaces, is that $H^s(\mathbb{T})$ is closed under multiplication if $s > 1/2$. We will need the following lemma:

Lemma 4.17. *Let $s > 1/2$. For $f \in H^s(\mathbb{T})$ we have $\hat{f} \in l^1(\mathbb{Z})$, and*

$$\|\hat{f}\|_{l^1(\mathbb{Z})} \le C_s\|f\|_{H^s(\mathbb{T})}$$

where C_s is a constant depending only on s.

Proof. The result again follows from a simple application of the Cauchy-Schwarz inequality: For $f \in H^s(\mathbb{T})$ with $s > 1/2$, we have

$$\|\hat{f}\|_{l^1(\mathbb{Z})} = \| \langle k\rangle^{s} \langle k\rangle^{-s} \hat{f}(k)\|_{l^1(\mathbb{Z})} \le (2\pi)^{-\frac{1}{2}} \| \langle k\rangle^{-s} \|_{l^2(\mathbb{Z})}\|f\|_{H^s(\mathbb{T})} < \infty.$$

□

We also need the next lemma concerning the convolution of the product of distributions in periodic Sobolev spaces. We didn't state a result analogue to the second equality in Theorem 4.4 for $\mathcal{P}'$ because we cannot in general define the product between distributions. However, we have seen that for $s > 1/2$, elements in $H^s(\mathbb{T})$ are bounded and continuous functions, thus we may define their product and we may compute the Fourier series of this product:

Lemma 4.18. *For $f, g \in H^s(\mathbb{T})$ with $s > 1/2$ we have*

$$\mathcal{F}(fg)(k) = \hat{f} * \hat{g}.$$

Proof. For $s > 1/2$ we know that $f, g \in H^s(\mathbb{T})$ are bounded and continuous functions, thus the product fg is in $BC(\mathbb{T})$. Then the argument in the proof of Theorem 4.4 yields the desired result. □

Finally then, we are able to prove the following:

Theorem 4.19. *Let $s > 1/2$. Given $f, g \in H^s(\mathbb{T})$ we have $fg \in H^s(\mathbb{T})$ and*

$$\|fg\|_{H^s(\mathbb{T})} \le c_s \|f\|_{H^s(\mathbb{T})} \|g\|_{H^s(\mathbb{T})}$$

where c_s is a constant that depends only on s. In other words, $H^s(\mathbb{T})$ is a Banach algebra for $s > 1/2$.

Proof. From Theorem 2.18, we have for any $t \in [0, \infty)$ the inequality

$$(1 + |x|^2)^t \le 2^{2t}(1 + |x - y|^2)^t + 2^{2t}(1 + |y|^2)^t \quad \forall x, y \in \mathbb{R}.$$

Thus, assuming $s > 1/2$ and $f, g \in H^s(\mathbb{T})$, we have by Lemma 4.18,

$$\begin{aligned} \langle k \rangle^s |\widehat{fg}(k)| = \langle k \rangle^s |(\hat{f} * \hat{g})(k)| &\le \sum_{j \in \mathbb{Z}} \langle k \rangle^s |\hat{f}(j)||\hat{g}(k - j)| \\ &\le 2^s \sum_{j \in \mathbb{Z}} \left(\langle k - j \rangle^s + \langle j \rangle^s\right) |\hat{f}(j)||\hat{g}(k - j)| \\ &= 2^s \left(|\langle \cdot \rangle^s \hat{g}| * |\hat{f}|\right)(k) + 2^s \left(|\langle \cdot \rangle^s \hat{f}| * |\hat{g}|\right)(k). \end{aligned}$$

Then, by the triangle inequality for the $l^2(\mathbb{Z})$ norm and Young's inequality (see for instance [21, Proposition 3.199]),

$$\begin{aligned} \|fg\|_{H^s(\mathbb{T})} &= \| \langle k \rangle^s \widehat{fg}(k) \|_{l^2(\mathbb{Z})} \\ &\le 2^s \| |\langle \cdot \rangle^s \hat{g}| * |\hat{f}| \|_{l^2(\mathbb{Z})} + 2^s \| |\langle \cdot \rangle^s \hat{f}| * |\hat{g}| \|_{l^2(\mathbb{Z})} \\ &\le \frac{2^s}{\sqrt{2\pi}} \|\hat{f}\|_{l^1(\mathbb{Z})} \|g\|_{H^s(\mathbb{T})} + \frac{2^s}{\sqrt{2\pi}} \|\hat{g}\|_{l^1(\mathbb{Z})} \|f\|_{H^s(\mathbb{T})}. \end{aligned}$$

Finally then by Lemma 4.17 we have $\|fg\|_{H^s(\mathbb{T})} \le c_s \|f\|_{H^s(\mathbb{T})} \|g\|_{H^s(\mathbb{T})}$. □

5 Well-posedness of Cauchy problems for linear PDEs

The purpose of this section is to give some practical examples of how one can prove well-posedness for simple linear PDEs. We start by performing an analysis of the Cauchy problem for the linearised KdV equation, where the goal is to establish properties of the problem that constitute its well-posedness in the Sobolev space $H^3(\mathbb{R})$. We then elaborate on what exactly we mean by a Cauchy problem for a PDE being well-posed in a function or distribution space. Next we introduce the concept of a Fourier multiplier, and analyse the well-posedness in Sobolev spaces for a class of linear Fourier multiplier equations. Finally we discuss some of the difficulties involved in establishing well-posedness for nonlinear equations.

Remark 5.1. In this section all the results and analyses are original.

5.1 The linearised Korteweg-de Vries equation

As mentioned in the introduction, the linearised Korteweg-de Vries equation

$$u_t + u_{xxx} = 0 \tag{5.1}$$

describes the linear evolution of the KdV equation

$$u_t + (u^2 + u_{xx})_x = 0 \tag{5.2}$$

in one-dimensional space and time, $(t,x) \in \mathbb{R} \times \mathbb{R}$. We consider the spatial derivatives in the equation to be weak derivatives in general. By applying the Fourier and Sobolev space theory from the previous sections we can solve the Cauchy problem

$$\begin{cases} u_t + u_{xxx} = 0 & \text{for } (t,x) \in \mathbb{R} \times \mathbb{R}, \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases} \tag{5.3}$$

and derive some important properties of the solution.

Let us assume that we can apply the Fourier transform with respect to x to each of the terms appearing in (5.3). Using property (i) of Theorem 2.3 (alternatively Theorem 3.4),

$$0 = \mathcal{F}(u_t) + \mathcal{F}(u_{xxx}) = \hat{u}_t(t,\xi) - i\xi^3 \hat{u}(t,\xi),$$

so we get the ODE

$$\hat{u}_t(t,\xi) = i\xi^3 \hat{u}(t,\xi), \tag{5.4}$$

with solution $\hat{u}(t,\xi) = C(\xi)e^{i\xi^3 t} = \hat{u}_0(\xi)e^{i\xi^3 t}$ by the initial condition. Applying the inverse Fourier transform we get the solution of (5.3),

$$u(t,x) = \mathcal{F}^{-1}\left(\hat{u}_0(\xi)e^{i\xi^3 t}\right)(x) = \frac{1}{\sqrt{2\pi}}\left(\mathcal{F}^{-1}\left(e^{i\xi^3 t}\right) * u_0\right)(x), \tag{5.5}$$

by Theorem 2.5, assuming the convolution is well-defined.

The solution given by (5.5) has some important properties which we elaborate on in the next four lemmas.

Lemma 5.1. *For initial data $u_0 \in H^3(\mathbb{R})$, the solution map $t \mapsto u(t,\cdot)$ is in $C^0(\mathbb{R}, H^3(\mathbb{R})) \cap C^1(\mathbb{R}, L^2(\mathbb{R}))$.*

Proof. We start by proving that u takes $\mathbb{R}$ to $H^3(\mathbb{R})$. The $H^3(\mathbb{R})$-norm of u is finite for all $t \in \mathbb{R}$, and in fact equal to the $H^3(\mathbb{R})$-norm of u_0:

$$\|u(t,\cdot)\|_{H^3(\mathbb{R})} = \left(\int_{\mathbb{R}} (1+|\xi|^2)^3 |\hat{u}_0(\xi)e^{i\xi^3 t}|^2 \, d\xi\right)^{\frac{1}{2}} = \|u_0\|_{H^3(\mathbb{R})} < \infty$$

To show that the solution map is continuous in t from $\mathbb{R}$ to $H^3(\mathbb{R})$, we fix an arbitrary t_0, and let $\{t_n\}_n \subseteq \mathbb{R}$ be a sequence such that $t_n \to t_0$. We have

$$\begin{aligned} \lim_{n\to\infty} \|u(t_n,\cdot) - u(t_0,\cdot)\|^2_{H^3(\mathbb{R})} &= \lim_{n\to\infty} \int_{\mathbb{R}} (1+|\xi|^2)^3 |\hat{u}(\xi,t) - \hat{u}(\xi,t_0)|^2 \, d\xi \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} (1+|\xi|^2)^3 |\hat{u}_0(\xi)|^2 |e^{i\xi^3 t_n} - e^{i\xi^3 t_0}|^2 \, d\xi \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} (1+|\xi|^2)^3 |\hat{u}_0(\xi)|^2 |1 - e^{i\xi^3(t_0 - t_n)}|^2 \, d\xi \to 0 \end{aligned}$$

by Lebesgue's dominated convergence theorem, as $e^{i\xi^3(t_0-t_n)} \to 1$ pointwise and the integrand is dominated by $(1+|\xi|^2)^3|\hat{u}_0(\xi)|^2$ which is integrable.

Next, since $u_t = -u_{xxx}$, we have $u_t \in L^2(\mathbb{R})$ for $u \in H^3(\mathbb{R})$. The functions u and u_t are continuous $\mathbb{R} \to L^2(\mathbb{R})$ by a similar argument to the one above. Thus u is also in $C^1(\mathbb{R}, L^2(\mathbb{R}))$. □

The proof of Lemma 5.1 demonstrates an important feature of the Cauchy problem (5.3), namely that the $H^3(\mathbb{R})$-regularity of the initial data is preserved in the solution.

Lemma 5.2. *For fixed $t \in \mathbb{R}$ and initial data $u_0 \in H^3(\mathbb{R})$, the solution $u(t,x) = u(t,x;u_0)$ given by (5.5) defines a data-to-solution map $u_0 \mapsto u(t,\cdot;u_0)$ that is Lipschitz continuous $H^3(\mathbb{R}) \to H^3(\mathbb{R})$.*

Proof. The data-to-solution map $u_0 \mapsto u(t,\cdot;u_0)$ is linear by the linearity of the Fourier transform, and we know it is bounded on $H^3(\mathbb{R})$ since $\|u(t,\cdot;u_0)\|_{H^3(\mathbb{R})} = \|u_0\|_{H^3(\mathbb{R})}$. Thus it is a Lipschitz continuous linear map from $H^3(\mathbb{R})$ to $H^3(\mathbb{R})$. □

This property of the data-to-solution map demonstrates the *stability* of the Cauchy problem (5.3): A small perturbation in the initial data leads to only a small change in the solution. Since t only appears inside a complex exponential in the solution function, continuity of the data-to-solution map holds also when we no longer consider t fixed:

Lemma 5.3. *The data-to-solution map $u_0 \mapsto u(\cdot,\cdot;u_0)$ is Lipschitz continuous from $H^3(\mathbb{R})$ to $C^0(\mathbb{R}, H^3(\mathbb{R})) \cap C^1(\mathbb{R}, L^2(\mathbb{R}))$.*

Proof. First we show Lipschitz continuity $H^3(\mathbb{R}) \to C^0(\mathbb{R}, H^3(\mathbb{R}))$. Since the map of the lemma is linear in u_0, we only need to show it is bounded, i.e. that

$$\sup_{t\in\mathbb{R}} \|u(t,\cdot;u_0)\|_{H^3(\mathbb{R})} \leq C\|u_0\|_{H^3(\mathbb{R})}$$

for some constant C. Again $\|u(t,\cdot;u_0)\|_{H^3(\mathbb{R})} = \|u_0\|_{H^3(\mathbb{R})}$ independently of t, hence the map is Lipschitz continuous from $H^3(\mathbb{R})$ to $C^0(\mathbb{R}, H^3(\mathbb{R}))$.

In order to prove that the data-to-solution map is Lipschitz continuous from $H^3(\mathbb{R})$ to $C^1(\mathbb{R}, L^2(\mathbb{R}))$, we need to show that the inequality

$$\sup_{t\in\mathbb{R}} \|u(t,\cdot;u_0)\|_{L^2(\mathbb{R})} + \sup_{t\in\mathbb{R}} \|u_t(t,\cdot;u_0)\|_{L^2(\mathbb{R})} \leq K\|u_0\|_{H^3(\mathbb{R})}$$

holds for some constant K. We have $\|u(t,\cdot;u_0)\|_{L^2(\mathbb{R})} = \|u_0\|_{L^2(\mathbb{R})} \leq \|u_0\|_{H^3(\mathbb{R})}$ for all $t \in \mathbb{R}$. We also have $\|u_t(t,\cdot;u_0)\|_{L^2(\mathbb{R})} = \|u_{xxx}(t,\cdot;u_0)\|_{L^2(\mathbb{R})} \leq \|u_0\|_{H^3(\mathbb{R})}$ for all t, therefore $K = 2$ will do. □

In finding the solution (5.5), we applied the Fourier transform with respect to the space variable to u_t and u_{xxx}. For this to make sense we need the Fourier transforms of these terms to be well-defined, which they are if we for instance require them to be $L^2(\mathbb{R})$-functions. Since functions in $W_2^k(\mathbb{R}) = H^k(\mathbb{R})$ have k weak derivatives in $L^2(\mathbb{R})$, we may require $u(t,\cdot)$ to be in $H^3(\mathbb{R})$, which we have seen is the case for $u_0 \in H^3(\mathbb{R})$. Since taking one temporal derivative of u corresponds to taking three spatial derivatives, we also have $u_t \in L^2(\mathbb{R})$ if $u \in H^3(\mathbb{R})$.

We have uniqueness of solution in $H^3(\mathbb{R})$:

Lemma 5.4. *The solution $u(t,\cdot)$ of* (5.3) *given by* (5.5) *is unique in $H^3(\mathbb{R})$.*

Proof. Suppose we have two $H^3(\mathbb{R})$-solutions for initial data $u_0 \in H^3(\mathbb{R})$, $u^1(t,\cdot)$ and $u^2(t,\cdot)$. By the linearity of the equation, $v := u^1 - u^2$ will also be a solution, with initial data $v_0 = v(0,\cdot) = 0$ a.e.. Then by (5.5), $v = 0$ a.e. and so $u^1 = u^2$ a.e, or in other words, u_1 and u_2 come from the same equivalence class of functions in $H^3(\mathbb{R})$. □

Remark 5.2. Note that since we consider the derivatives in (5.3) to be weak derivatives in general, the $H^3(\mathbb{R})$-solution that we have considered above is a weak solution. If we had more regular initial data, we could also guarantee classical solutions: From Theorem 2.16 we have $H^s(\mathbb{R}) \hookrightarrow BC^3(\mathbb{R})$ for $s > 3 + \frac{1}{2}$, so for such an s, initial data $u_0 \in H^s(\mathbb{R})$ permits a unique classical $H^s(\mathbb{R})$-solution to the linearised KdV equation given by formula (5.5).

The analysis we have just performed on (5.3) establishes the *well-posedness* of the Cauchy problem for the linearised Korteweg-de Vries equation equation in the Sobolev space $H^3(\mathbb{R})$. We discussed the concept of well-posedness of a Cauchy problem in the introduction, and in the beginning of the next subsection we give a more precise definition of what a Cauchy problem for a PDE being well-posed in a function space means. We sum up the results of our analysis of the linearised KdV equation in Lemmata 5.1 through 5.4 as follows:

Theorem 5.5. *The Cauchy problem* (5.3) *is well-posed in* $H^3(\mathbb{R})$. *Specifically, given* $u_0 \in H^3(\mathbb{R}^d)$ *we have that:*

(i) A solution $u(t,\cdot)$ *given by* (5.5) *exists in* $H^3(\mathbb{R})$ *and stays in* $H^3(\mathbb{R})$ *for all* $t \in \mathbb{R}$.

(ii) This solution is unique in $H^3(\mathbb{R})$.

(iii) For fixed $t \in \mathbb{R}$, *the data-to-solution map* $u_0 \mapsto u(t,\cdot;u_0)$ *is Lipschitz continuous* $H^3(\mathbb{R}) \to H^3(\mathbb{R})$. *More generally, the data-to-solution map* $u_0 \mapsto u(\cdot,\cdot;u_0)$ *is Lipschitz continuous* $H^3(\mathbb{R}) \to C^0(\mathbb{R}, H^3(\mathbb{R})) \cap C^1(\mathbb{R}, L^2(\mathbb{R}))$.

5.2 Well-posedness in function spaces

We now make precise the concept of well-posedness of a Cauchy problem for a PDE in a function space:

Definition 5.1 (Well-posedness)**.** We say that a Cauchy problem for a PDE is well-posed in a space X if, given initial data in X,

(i) there exists a solution in X which stays in X for some time $T > 0$,

(ii) the solution is unique is X,

(iii) the solution depends continuously on the initial data.

There are some further subtleties involved in determining the well-posedness of a Cauchy problem. Specifically, the solution might depend continuously on the initial data in several different ways, and certain forms of existence of solution are stronger than others. Concerning this last point, we will generally be satisfied with establishing the existence of weak solutions, as discussed above and in the introduction. The exact meaning of uniqueness of solution in X may also depend on the nature of the space X, and the time of existence T may vary. In our analysis of the linearised KdV equation, we saw that the solution stayed in $H^3(\mathbb{R})$ for all $t \in \mathbb{R}$, thus $T = \infty$. In such cases we say that the Cauchy problem is *globally* well-posed. This is typically the case for linear equations, while for nonlinear equations we may only be able to establish that the solutions exist for a finite time T (we say the problem is *locally* well-posed), and global well-posedness may remain an open question with both positive and negative answers possible.

5.3 Global well-posedness for a class of linear Fourier multiplier equations

One of the reasons why the Fourier transform is such a powerful tool for solving Cauchy problems for PDEs, is that is maps differential operators to polynomials. It is natural to consider equations where the operators map to general functions through Fourier transform. This is the idea of Fourier multiplier operators.

Let us consider a general linear problem of the form

$$\begin{cases} u_t + p(D)u = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases} \tag{5.6}$$

where the operator $p(D)$ is defined via the Fourier transform by

$$\widehat{p(D)f}(\xi) = p(\xi)\hat{f}(\xi).$$

The function p is called the Fourier multiplier, or just multiplier [12, p. 58], and we shall refer to the corresponding operator $p(D)$ as the (Fourier) multiplier operator. A Fourier multiplier equation such as (5.6) is a special case of what we term pseudodifferential equations. Properties (i) and (ii) of Theorem 2.3 imply that when p is a polynomial, $p(D)$ will be a differential operator. In this case $p(D)$ is a *local operator*, in the sense that the value of $p(D)f$ at a certain point depends only on the values of f in a neighbourhood of the point. When this is not the case, we say that $p(D)$ is a *nonlocal operator*, and to determine the value of $p(D)f$ at a point, information about the values of f far from that point may be needed. Integral operators are typical nonlocal operators, an example being the convolution with a certain function, which we know from Theorems 2.5, 3.6 and 4.11 can be a Fourier multiplier operator.

As an example, let us consider the well-posedness of the Cauchy problem (5.6) with $p(\xi) = (1 + |\xi|^{2s})^{-1}$, $s \geq 0$. Assuming we can apply the Fourier transform to u_t

and $p(D)u$, the solution is

$$u(t,x) = \mathcal{F}^{-1}\left(\hat{u}_0(\xi)e^{-t(1+|\xi|^{2s})^{-1}}\right) = \frac{1}{\sqrt{2\pi}}\left(\mathcal{F}^{-1}\left(e^{-t(1+|\xi|^{2s})^{-1}}\right) * u_0\right)(x). \tag{5.7}$$

In order for the Fourier transform of $p(D)u$ to be well-defined, we may require $p(D)u$ to be in $L^2(\mathbb{R})$. It turns out that $p(D)u$ is in $L^2(\mathbb{R})$ for $u \in L^2(\mathbb{R})$, since

$$\|p(\xi)\hat{u}(\xi)\|_{L^2(\mathbb{R})} = \left(\int_{\mathbb{R}}\left|\frac{1}{1+|\xi|^{2s}}\hat{u}(\xi)\right|^2 d\xi\right)^{\frac{1}{2}} \leq \|\hat{u}\|_{L^2(\mathbb{R})}.$$

This is of course true for any bounded p, since we then have $\|p\hat{u}\|_{L^2(\mathbb{R})} \leq M\|\hat{u}\|_{L^2(\mathbb{R})}$ for some constant M. Also, from (5.7) we see that ∂_t applied to u corresponds to multiplication by $-(1+|\xi|^{2s})^{-1}$ on the Fourier side. Thus $u_t \in L^2(\mathbb{R})$ for $u \in L^2(\mathbb{R})$

We have the following global well-posedness result in $L^2(\mathbb{R})$:

Theorem 5.6. *The Cauchy problem* (5.6) *with* $p(\xi) = (1+|\xi|^{2s})^{-1}$, $s \geq 0$, *is globally well-posed in* $L^2(\mathbb{R})$. *Specifically, given initial data* $u_0 \in L^2(\mathbb{R})$, *we have that*

(i) A solution $u(t,\cdot)$ *given by* (5.7) *exists in* $L^2(\mathbb{R})$ *and stays in* $L^2(\mathbb{R})$ *for all* $t \in \mathbb{R}^+$.

(ii) This solution is unique in $L^2(\mathbb{R})$.

(iii) For fixed $t \in \mathbb{R}^+$, *the data-to-solution map* $u_0 \mapsto u(t,\cdot;u_0)$ *is Lipschitz continuous* $L^2(\mathbb{R}) \to L^2(\mathbb{R})$. *More generally, the data-to-solution map* $u_0 \mapsto u(\cdot,\cdot;u_0)$ *is Lipschitz continuous* $L^2(\mathbb{R}) \to C^k(\mathbb{R}^+, L^2(\mathbb{R}))$ *for all* $k \in \mathbb{N}_0$.

Proof.

(i) Given initial data $u_0 \in L^2(\mathbb{R})$, a solution $u(t,\cdot)$ given by (5.7) exists in $L^2(\mathbb{R})$ since all the steps taken to acquire the solution (5.7) make sense for $L^2(\mathbb{R})$-functions. Also, the solution stays in $L^2(\mathbb{R})$ for all $t \in \mathbb{R}^+$, as its norm is uniformly bounded by the norm of the initial data:

$$\|\hat{u}(t,\cdot)\|_{L^2(\mathbb{R})} = \left(\int_{\mathbb{R}}\left|\hat{u}_0(\xi)e^{-\frac{t}{1+|\xi|^{2s}}}\right|^2 d\xi\right)^{\frac{1}{2}} \leq \|\hat{u}_0\|_{L^2(\mathbb{R})},$$

for any $t \in \mathbb{R}^+$, which implies $\|u\|_{L^2(\mathbb{R})} \leq \|u_0\|_{L^2(\mathbb{R})}$ by Theorem 2.11. We note here that - unlike for the linearised KdV equation - it is important that we are considering only the forward time evolution of the system, since the above inequality clearly does not hold for $t < 0$.

(ii) The solution is unique in $L^2(\mathbb{R})$ by an argument similar to the one in the proof of Theorem 5.4.

(iii) For fixed $t \in \mathbb{R}^+$, the data-to-solution map is in Lipschitz continuous from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$ by the final inequality in (i).

More generally, the data-to-solution map is Lipschitz continuous $L^2(\mathbb{R}) \to C^k(\mathbb{R}^+, L^2(\mathbb{R}))$ for all $k \in \mathbb{N}_0$: We have

$$\partial_t^k u = \frac{1}{\sqrt{2\pi}} \left(\mathcal{F}^{-1} \left(\frac{(-1)^k}{(1+|\xi|^{2s})^k} e^{-\frac{t}{1+|\xi|^{2s}}} \right) * u_0 \right)(x).$$

Thus $t \mapsto \partial_t^k u(t, \cdot; u_0)$ takes $\mathbb{R}^+$ to $L^2(\mathbb{R})$ for all $k \in \mathbb{N}_0$, since

$$\|\widehat{\partial_t^k u}(t,\cdot)\|_{L^2(\mathbb{R})} = \left(\int_{\mathbb{R}} \left| \hat{u}_0(\xi) \frac{1}{(1+|\xi|^{2s})^k} e^{-\frac{t}{1+|\xi|^{2s}}} \right|^2 d\xi \right)^{\frac{1}{2}} \leq \|\hat{u}_0\|_{L^2(\mathbb{R})}, \tag{5.8}$$

which by Theorem 2.11 implies $\|\partial_t^k u(t,\cdot)\|_{L^2(\mathbb{R})} \leq \|u_0\|_{L^2(\mathbb{R})} < \infty$. Thus we have $\partial_t^k u(t,\cdot) \in L^2(\mathbb{R})$.

In fact, the solution map $t \mapsto u(t, \cdot; u_0)$ is in $C^k(\mathbb{R}^+, L^2(\mathbb{R}))$ for any $k \in \mathbb{N}_0$: For arbitrary $t_0 \in \mathbb{R}^+$ fixed and given any $\varepsilon > 0$, we can pick $\delta > 0$ such that for any $\xi \in \mathbb{R}$

$$\left| e^{-\frac{t}{1+|\xi|^{2s}}} - e^{-\frac{t_0}{1+|\xi|^{2s}}} \right| < \varepsilon / \|u_0\|_{L^2(\mathbb{R})}$$

whenever $|t - t_0| < \delta$ by the smoothness of the exponential function. This choice of δ gives

$$\begin{aligned}
\|\partial_t^k u(t,\cdot;u_0) - \partial_t^k u(t_0,\cdot;u_0)\|_{L^2(\mathbb{R})} &= \|\widehat{\partial_t^k u}(t,\cdot;u_0) - \widehat{\partial_t^k u}(t_0,\cdot;u_0)\|_{L^2(\mathbb{R})} \\
&= \left(\int_{\mathbb{R}} \left| \hat{u}_0(\xi) \frac{1}{(1+|\xi|^{2s})^k} \left(e^{-\frac{t}{1+|\xi|^{2s}}} - e^{-\frac{t_0}{1+|\xi|^{2s}}} \right) \right|^2 d\xi \right)^{\frac{1}{2}} \\
&< \frac{\varepsilon}{\|u_0\|_{L^2(\mathbb{R})}} \left(\int_{\mathbb{R}} \left| \hat{u}_0(\xi) \frac{1}{(1+|\xi|^{2s})^k} \right|^2 d\xi \right)^{\frac{1}{2}} \leq \varepsilon
\end{aligned}$$

whenever $|t - t_0| < \delta$ and for any $k \in \mathbb{N}_0$.

Finally, the data-to-solution map $u_0 \mapsto u(\cdot, \cdot; u_0)$ is Lipschitz continuous from $L^2(\mathbb{R})$ to $C^k(\mathbb{R}^+, L^2(\mathbb{R}))$, since

$$\sup_{t \in \mathbb{R}^+} \|\widehat{\partial_t^k u}(t,\cdot;u_0)\|^2_{L^2(\mathbb{R})} = \sup_{t \in \mathbb{R}^+} \int_{\mathbb{R}} \left| \hat{u}_0(\xi) \frac{1}{(1+|\xi|^{2s})^k} e^{-\frac{t}{1+|\xi|^{2s}}} \right|^2 d\xi \leq \|\hat{u}_0\|^2_{L^2(\mathbb{R})}$$

for all $k \in \mathbb{N}_0$.

□

Remark 5.3. The first time derivative of the solution actually has Sobolev regularity of order $2s$ for $u_0 \in L^2(\mathbb{R})$, since

$$\|u_t(t,\cdot)\|^2_{H^{2s}(\mathbb{R})} = \int_{\mathbb{R}} \frac{(1+|\xi|^2)^{2s}}{(1+|\xi|^{2s})^2} \left|\hat{u}_0(\xi) e^{-\frac{t}{1+|\xi|^{2s}}}\right|^2 d\xi \leq C \int_{\mathbb{R}} |\hat{u}_0(\xi)|^2 \, d\xi = C\|u_0\|^2_{L^2(\mathbb{R})}$$

for some constant C and any $t \in \mathbb{R}^+$. In general the kth time derivative of the solution is in $H^{2sk}(\mathbb{R})$ for initial data $u_0 \in L^2(\mathbb{R})$. We say $p(D)$ is a smoothing operator of order $-2s$.

5.4 Proving well-posedness for nonlinear equations

For the linear equations we studied in Sections 5.1 and 5.3, we saw how the entire package of global well-posedness in certain Sobolev spaces followed neatly from the equations' solutions by Fourier transform. Showing well-posedness for a nonlinear equation can be more intricate. From our experiences with ordinary differential equations, we know that solutions of nonlinear equations often display certain behaviours that solutions of linear equations typically do not. Consider for instance the initial value problem

$$\begin{cases} \frac{du}{dt} = u^2 & \text{for } t \in \mathbb{R}, \\ u(0) = u_0. \end{cases}$$

This problem has solution

$$u(t) = \left(\frac{1}{u_0} - t\right)^{-1}.$$

We notice that the solution blows up as $t \to 1/u_0$. So far all our solutions and well-posedness results have been global in time. We cannot expect this to be the case for nonlinear PDEs. In fact, in the next section we focus mainly on proving local well-posedness, and our general results only guarantee that a unique solution exists and is stable for a finite time that depends on the size of the initial data.

6 Well-posedness for a class of nonlocal Whitham-like equations

In this section we shall prove local well-posedness in the Sobolev spaces $H^s(\mathbb{R})$ for $s > 3/2$ of Cauchy problems of the form

$$\begin{cases} u_t + (u^p)_x + \mathcal{L}_\alpha u_x = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases} \tag{6.1}$$

with $p \geq 2$ a positive integer and the operator $\mathcal{L}_\alpha$ (which is assumed to map real data to real data) defined on the Fourier side by

$$\mathcal{F}(\mathcal{L}_\alpha f)(\xi) = \frac{b(\xi)}{|\xi|^\alpha} \hat{f}(\xi), \quad 0 < \alpha \leq 1, \tag{6.2}$$

for all objects f for which the definition makes sense. Here b is a real-valued function of *slow* or *polynomial growth* of at most order

$$N_b \leq s - 1 + \alpha, \tag{6.3}$$

in other words, there exists constants $C_b > 0$ and $N_b \geq 0$ satisfying (6.3) s.t.

$$|b(\xi)| \leq C_b (1 + |\xi|)^{N_b} \tag{6.4}$$

for all $\xi \in \mathbb{R}$ (see (6.16) below for where this restriction comes from).

Let us immediately state our first main well-posedness result. Note that throughout this section we will often write simply H^s for $H^s(\mathbb{R})$, and the same for L^2, C_c^∞ etc.:

Theorem 6.1 (Well-posedness in $H^s(\mathbb{R})$)**.** *Let $s > 3/2$, $p \in \{2, 3, \cdots\}$ and $\alpha \in (0,1]$, and let b be a real-valued function satisfying* (6.4) *for some $C_b > 0$ and $N_b \geq 0$ satisfying* (6.3)*. The Cauchy problem* (6.1) *with the operator $\mathcal{L}_\alpha$ defined as in* (6.2) *is locally well-posed in H^s. Specifically, for any given $u_0 \in H^s$, there exists a maximal time $T > 0$ depending only on $\|u_0\|_{H^s}$ such that*

(i) There is a unique solution $u \in C^0([0,T), H^s) \cap C^1([0,T), H^q)$, where

$$q := \min\{s - 1, s - 1 + \alpha - N_b\} \geq 0.$$

(ii) The data-to-solution map $u_0 \mapsto u(\cdot; u_0)$ is continuous from H^s to $C^0([0,T), H^s) \cap C^1([0,T), H^q)$.

Remark 6.1. Note that by the embedding $H^s \hookrightarrow BC^1$ for $s > 3/2$, our solution is in fact continuously differentiable with respect to x.

Similarly we shall prove well-posedness for the corresponding periodic Cauchy problems,

$$\begin{cases} u_t + (u^p)_x + \mathcal{L}_\alpha u_x = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \text{ with } u \text{ a } 2\pi\text{-periodic distribution,} \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}. \end{cases} \tag{6.5}$$

Recall from Section 4 that the Fourier transform of a periodic function or distribution is defined as the map from the function or distribution to its sequence of Fourier coefficients. The operator $\mathcal{L}_\alpha$ (assumed to map real data to real data) is defined in terms of its Fourier coefficients by

$$\mathcal{F}(\mathcal{L}_\alpha f)(k) = \frac{b(k)}{|k|^\alpha} \hat{f}(k), \quad k \in \mathbb{Z}, \tag{6.6}$$

with $0 < \alpha \leq 1$, where b is a real sequence of slow growth of at most order

$$N_b \leq s - 1 + \alpha, \tag{6.7}$$

i.e. there exists $C_b > 0$ and $N_b \geq 0$ satisfying (6.7) such that

$$|b(k)| \leq C_b |k|^{N_b} \quad \forall\, k \neq 0. \tag{6.8}$$

Our periodic well-posedness result is analogous to Theorem (6.1):

Theorem 6.2 (Well-posedness in $H^s(\mathbb{T})$)**.** *Let $s > 3/2$, $p \in \{2,3,\cdots\}$ and $\alpha \in (0,1]$, and let b be a real sequence satisfying (6.8) for some $C_b > 0$ and $N_b \geq 0$ satisfying (6.7). The periodic Cauchy problem (6.5) with the operator $\mathcal{L}_\alpha$ defined as in (6.6) is locally well-posed in $H^s(\mathbb{T})$. Specifically, for any given $u_0 \in H^s(\mathbb{T})$, there exists a maximal time $T > 0$ depending only on $\|u_0\|_{H^s(\mathbb{T})}$ such that*

(i) There is a unique solution $u \in C^0([0,T), H^s(\mathbb{T})) \cap C^1([0,T), H^q(\mathbb{T}))$, where

$$q = \min\{s-1, s-1+\alpha-N_b\} \geq 0.$$

(ii) The data-to-solution map $u_0 \mapsto u(\cdot; u_0)$ is continuous from $H^s(\mathbb{T})$ to $C^0([0,T), H^s(\mathbb{T})) \cap C^1([0,T), H^q(\mathbb{T}))$.

Remark 6.2. Again we have the embedding $H^s(\mathbb{T}) \hookrightarrow BC^1(\mathbb{T})$ for $s > 3/2$ (cf. Theorem 4.16). Thus our solution is in fact continuously differentiable with respect to x.

As for why we characterise these equations as Whitham-like, recall that the Whitham equation is given by

$$u_t + uu_x + \mathcal{L}u_x = 0$$

with the Fourier multiplier $\mathcal{L}$ defined by

$$\mathcal{F}(\mathcal{L}f)(\xi) = \left(\frac{\tanh\xi}{\xi}\right)^{\frac{1}{2}} \hat{f}(\xi). \tag{6.9}$$

When assuming long wave lengths, one recovers the KdV equation from the Whitham equation as an approximation: Long wavelengths means we should let the frequency $\xi \to 0$. By retaining the first two terms of the Taylor series we get the approximation $\left(\frac{\tanh\xi}{\xi}\right)^{\frac{1}{2}} \approx 1 - \frac{\xi^2}{6}$, which gives us the equation

$$u_t + 2uu_x + u_x + \frac{1}{6}u_{xxx} = 0.$$

Performing the change of variables

$$\tilde{u}(t,x) = -2\sqrt[3]{6}\,u(t, \frac{x}{\sqrt[3]{6}} + t)$$

then gives us the KdV equation. In the opposite limit $\xi \to \infty$, we have $\tanh\xi \to 1$, so that $\left(\frac{\tanh\xi}{\xi}\right)^{1/2} \to |\xi|^{-1/2}$. This equation corresponds to (6.1) with $p = 2$, $\alpha = 1/2$ and $b \equiv 1$.

While in our analyses we focus on the cases where $0 < \alpha \le 1$, we note that if we were to set $\alpha = 0$ with $p = 2$ and $b \equiv 1$ in (6.1), we would have the familiar Burgers' equation (after the change of variables $\tilde{u}(t,x) = u(t, 2x + t)$). With $-1 < \alpha < 0$, we would have the *fractal* Burgers' equation, for which the Cauchy problem is studied in for instance [28] (we remark on this further in Section 6.5). Similarly, for $\alpha = -1$, we have the Benjamin-Ono equation, for which global well-posedness has been proven in L^2 in both the periodic case and on the real line [20]. Finally, setting $\alpha = 2$ gives us the KdV equation again, which is very well-studied, see for instance [8].

Remark 6.3. The work in this section is all original, but we adapt a method previously used by Ehrnström, Escher and Pei in [15] to prove local well-posedness for the Whitham equation. This is explained further in the text below.

6.1 Local well-posedness by Kato's method. The case $\alpha = 1$ and b bounded

Our analysis follows a paper by Ehrnström, Escher and Pei [15], which establishes local well-posedness for the Whitham equation. This paper applies a method used by Constantin and Escher in [10] to prove local well-posedness for the periodic Camassa-Holm equation, which in turn is based on a result by Kato [23]. We will state Kato's result below in Theorem 6.3 for convenience.

The main difference between our analysis and that in the paper [15], is that the operator $\mathcal{L}$ in the Whitham equation given by (6.9) is a bounded linear operator on L^2. This nice behaviour of $\mathcal{L}$ on square integrable functions simplifies the analysis of the Whitham equation somewhat compared to the equations that we consider, for which the operator $\mathcal{L}_\alpha$ is in general only a bounded linear operator from the *homogeneous Sobolev space* $\dot{H}^\alpha$ to L^2. The homogeneous Sobolev spaces $\dot{H}^s(\mathbb{R}^d)$ for $s \in \mathbb{R}$ are defined by

$$\dot{H}^s(\mathbb{R}^d) = \{f \in \mathcal{S}'(\mathbb{R}^d) \;:\; \hat{f} \in L^1_{loc}(\mathbb{R}^d) \text{ and } \int_{\mathbb{R}^d} |\xi|^{2s} |\hat{f}(\xi)|^2 \, d\xi < \infty\}.$$

The nice behaviour of $\mathcal{L}_\alpha$ on $\dot{H}^\alpha$ makes it natural to consider proving well-posedness in homogeneous Sobolev spaces instead of regular Sobolev spaces.

However, the homogeneous spaces lack many of the nice properties of the regular Sobolev spaces, which makes working with them difficult.

In particular, unlike the regular Sobolev spaces which are all Hilbert spaces, $\dot{H}^s(\mathbb{R}^d)$ is complete if and only if $s < d/2$ [2], and furthermore there is no analogous relation to the inclusions

$$H^{s_1} \subseteq H^{s_2} \quad \text{for } s_1 \geq s_2$$

for the homogeneous spaces. The lack of such a relation implies in particular that homogeneous Sobolev spaces of positive order cannot be classified as subspaces of L^2. Considering the hypotheses of Kato's theorem (cf. Theorem 6.3 below), it is therefore not trivial to exchange the regular Sobolev spaces for the homogeneous spaces when proving local well-posedness via Kato's theorem.

Our approach is instead to combine the operator $\mathcal{L}_\alpha$ with ∂_x in our analysis. The derivative operator cancels out the Fourier side singularity that comes from $\mathcal{L}_\alpha$, leading to the nice behaviour of the composition $\mathcal{L}_\alpha \partial_x$ as a bounded linear operator from Sobolev spaces of a certain order to L^2. This explains why we only consider $\alpha \in (0, 1]$, since for $\alpha > 1$ the derivative operator would not fully cancel out the singularity on the Fourier side. That we always have to consider the composition $\mathcal{L}_\alpha \partial_x$ as a single operator leaves us with somewhat less freedom throughout the analysis, and is the main thing that makes our proof at times more intricate than the proofs in [10] and [15].

In the case where $\alpha = 1$ and b is a bounded real-valued function, the operators $\mathcal{L}_\alpha$ and ∂_x cancel each other exactly, which simplifies the analysis and makes it easier to follow. Therefore we consider this special case first. That is, we will prove the local well-posedness of the Cauchy problem

$$\begin{cases} u_t + pu^{p-1}u_x + Lu_x = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases} \tag{6.10}$$

where $L := \mathcal{L}_1$ with b bounded and real-valued, and $p \geq 2$ is an integer (the technical differences in the proof in the general case $\alpha \in (0,1]$ and b of slow growth are discussed in the next section). In the analysis to come we actually assume $p = 2$ for simplicity, however, as will be apparent, the analysis is independent of the exact order of the nonlinearity and may thus be performed for p equal to any integer greater than or equal to 2. For similar reasons we also assume $b \equiv 1$.

Following [15], we rewrite (6.10) as

$$\begin{cases} u_t + A(u)u = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases} \tag{6.11}$$

where we have defined

$$A(y) : = (2y + L)\partial_x$$
$$\text{dom(A(y))} : = \{v \in L^2 \mid (2yv + Lv)_x \in L^2\},$$

for some $y \in H^s$, $s > 3/2$.

Remark 6.4. The ghost of the homogeneous Sobolev spaces actually show up in the definition of the domain of $A(y)$. In the articles [10] and [15], the domain of $A(y)$ is defined as $\{v \in L^2 \mid 2yv + Lv \in H^1\}$. Since we don't like to let L act on v by itself without pairing it with ∂_x, we only require of functions v in the domain of $A(y)$ that the derivative of $2yv + Lv$ is in L^2, and put no particular restrictions on $2yv + Lv$ itself. We don't define $\text{dom}(A(y))$ as the set $\{v \in L^2 \mid 2yv + Lv \in \dot{H}^1\}$ simply because we don't want to require that $\mathcal{F}(2yv + Lv) \in L^1_{loc}$ either.

Remark 6.5. In the general case $p \geq 2$ one should define the operator $A(y)$ as

$$A(y) := (py^{p-1} + L)\partial_x$$

for a fixed $y \in H^s$, $s > 3/2$, with $\text{dom}(A(y)) := \{v \in L^2 \mid (py^{p-1}v + Lv)_x \in L^2\}$.

The domain of $A(y)$ is dense in L^2. One can see this by for instance verifying that $H^1 \subseteq \text{dom}(A(y))$. Recall that, by Theorem 2.18, H^s is closed under pointwise multiplication when $s > 1/2$ (in other words, it is a Banach algebra), and notice that for $u \in H^s$,

$$\|L\partial_x u\|_{H^s} = \|u\|_{H^s},$$

since the Fourier multiplier operator $L\partial_x$ corresponds to $i\,\text{sgn}$ on the Fourier side. For this reason we also have for $u \in L^2$ that

$$\|L\partial_x u\|_{L^2} = \|u\|_{L^2},$$

in other words the operator $L\partial_x$ is an isometry on L^2.

Remark 6.6. This last point is the main reason the proof of well-posedness is simplified somewhat in the case $\alpha = 1$, since if we have $\alpha \in (0,1]$ instead, we may only say that $\mathcal{L}_\alpha \partial_x$ is a bounded operator from $H^{1-\alpha+N_b}$ to L^2.

In order to state Kato's theorem, we should first establish what it means for an operator to be accretive. In the following, $\mathcal{B}(X,Y)$ denotes the space of bounded linear operators from the Banach space X to the Banach space Y, and $\mathcal{B}(X) = \mathcal{B}(X,X)$.

Definition 6.1 (Accretive operator)**.** Let T be an operator on a Hilbert space H. We then say that

- T is accretive if $\operatorname{Re}\langle Tv, v\rangle_H \geq 0$ for all $v \in \operatorname{dom}(T)$;
- T is quasi-accretive if $T + \alpha$ is accretive for some scalar $\alpha > 0$;
- T is m-accretive if $(T+\lambda)^{-1} \in \mathcal{B}(H)$ with $\operatorname{Re}(\lambda)\|(T+\lambda)^{-1}\| \leq 1$ for $\operatorname{Re}(\lambda) > 0$;
- T is quasi-m-accretive if $T + \alpha$ is m-accretive for some scalar $\alpha > 0$.

We now state Kato's theorem for a general quasilinear evolution equation of the form (6.11). In our case we have $Y := H^s$ with $s > 3/2$ and $X := L^2$.

Theorem 6.3 (Kato's theorem)**.** *Consider the abstract quasilinear evolution equation* (6.11). *Let X and Y be Hilbert spaces such that Y is continuously and densely injected into X. Let $Q : Y \to X$ be an isomorphism. Assume that*

(i) *$A(y) \in \mathcal{B}(Y,X)$ for $y \in Y$, with*

$$\|(A(y) - A(z))w\|_X \leq \mu_A \|y - z\|_X \|w\|_Y \quad y, z, w \in Y,$$

and $A(y)$ is quasi-m-accretive, uniformly on bounded sets in Y,

(ii) *$QA(y)Q^{-1} = A(y) + B(y)$, where $B(y) \in \mathcal{B}(X)$ is bounded, uniformly on bounded sets in Y. Moreover*

$$\|(B(y) - B(z))w\|_X \leq \mu_B \|y - z\|_Y \|w\|_X, \quad y, z \in Y, w \in X.$$

Here the constants μ_A and μ_B depend only on $\max\{\|y\|_Y, \|z\|_Y\}$.

Then, for any given $v_0 \in Y$, there is a maximal $T > 0$ depending only on $\|v_0\|_Y$ and a unique solution v to (6.11) *such that $v = v(\cdot; v_0) \in C^0([0,T), Y) \cap C^1([0,T), X)$, and furthermore, the map $v_0 \mapsto v(\cdot; v_0)$ is continuous from Y to $C^0([0,T), Y) \cap C^1([0,T), X)$.*

We shall study the operator $A(y)$ for a fixed $y \in Y = H^s$, $s > 3/2$. Later y will be taken to be in a bounded subset of Y, however, note that all estimates to come are uniform with respect to any such bounded subset. Still following [15], we define for $y \in Y$,

$$\begin{aligned} Dv &:= (2yv + Lv)_x - 2y_x v \\ \operatorname{dom}(D) &:= \{v \in L^2 \mid (2yv + Lv)_x \in L^2\} \end{aligned} \tag{6.12}$$

and

$$\begin{aligned} D_0 v &:= -(2yv + Lv)_x, \\ \operatorname{dom}(D_0) &:= \{v \in L^2 \mid (2yv + Lv)_x \in L^2\}. \end{aligned} \tag{6.13}$$

Note that the choice of these domains make both D and D_0 *closed operators* in $X = L^2$ (we prove this in Lemma 6.7). Closed operators are an important class of unbounded linear operators defined as follows:

Definition 6.2 (Closed operator)**.** Let X and Y be Banach spaces, and let

$$A : X \supset \operatorname{dom}(A) \to Y$$

be a linear operator. Then A is called *closed* if for all sequences $\{x_n\}_n \subseteq \operatorname{dom}(A)$ such that there exists $x = \lim_{n\to\infty} x_n$ in X and $y = \lim_{n\to\infty} Ax_n$ in Y, it holds that $x \in \operatorname{dom}(A)$ and $Ax = y$.

We shall prove that D satisfies condition (i) of Theorem 6.3 by help of a few consecutive lemmas. This will in turn mean that $A(y)$ satisfies condition (i), since $(2yv+Lv)_x - 2y_x v = 2yv_x + Lv_x$ in the space H^{-1}, in other words the L^2 distributions Dv and $A(y)v$ for $v \in \operatorname{dom}(A(y))$ are equal a.e.

Remark 6.7. Note that $(Lv)_x = Lv_x$, i.e. L and ∂_x commute. This is because they are both Fourier multiplier operators.

Remark 6.8. Also note that $2y_x v + 2yv_x + Lv_x \in L^2$ (which is true for $v \in \operatorname{dom(D)}$ by definition), actually means that each term is in L^2 individually: Clearly Lv_x is in L^2, and so is $2y_x v$ by the embedding $H^s \hookrightarrow BC$ for $s > 1/2$. Then $2yv_x$ must be in L^2 as well, since

$$\|2yv_x\|_{L^2} = \|(2yv + Lv)_x - 2y_x v - Lv_x\|_{L^2} \le \|(2yv + Lv)_x\|_{L^2} + \|2y_x v + Lv_x\|_{L^2} < \infty.$$

We start by proving an approximation result which we shall use frequently:

Lemma 6.4. *Given $v \in \operatorname{dom(D)}$, there exists a sequence $\{v_n\}_n \subseteq C^\infty$ such that*

$$v_n \to v \quad \text{and} \quad (2yv_n + Lv_n)_x \to (2yv + Lv)_x$$

in L^2 as $n \to \infty$.

Proof. We follow the argument in [15] and pick $\rho \in C_c^\infty$ with $\rho \geq 0$ and $\int_{\mathbb{R}} \rho \, dx = 1$. For $n \geq 1$, let $\rho_x(x) \coloneqq n\rho(nx)$ be a mollifier on $\mathbb{R}$. Defining $v_n \coloneqq v * \rho_n$, we have $v_n \in C^\infty \cap L^2$ by Young's inequality and indeed $v_n \to v$ in L^2 as $n \to \infty$ (see for instance [39, Lemma 7.1 c]). This proves the first part of the lemma.

As what concerns the second part, we have

$$\begin{aligned}
&(2yv_n + Lv_n)_x - (2yv + Lv)_x \\
&= (2y(v_n)_x + L(v_n)_x - (2yv_x + Lv_x)) + (2y_x v_n - 2y_x v) \\
&= (2y(v_n)_x + L(v_n)_x - (2yv_x + Lv_x) * \rho_n) \\
&\quad + ((2yv_x + Lv_x) * \rho_n - (2yv_x + Lv_x)) \\
&\quad + (2y_x v_n - 2y_x v) =: \mathrm{I}_n(v) + \mathrm{II}_n(v) + \mathrm{III}_n(v).
\end{aligned}$$

By $2yv_x + Lv_x \in L^2$ for $v \in \operatorname{dom}(D)$, one has

$$\mathrm{II}_n(v) = (2yv_x + Lv_x) * \rho_n - (2yv_x + Lv_x) \to 0$$

in L^2 as $n \to \infty$. Also, by the embedding $H^s \hookrightarrow BC$ for $s > 1/2$, one has

$$\mathrm{III}_n(v) = 2y_x(v_n - v) \to 0$$

in L^2 as $n \to \infty$. It remains to show that $\mathrm{I}_n(v) \to 0$ in L^2 as $n \to \infty$. We claim that this holds for $v \in C_c^\infty$. For such v we have by the triangle inequality

$$\begin{aligned}
&\|2y(v_x * \rho_n) - (2yv_x) * \rho_n + L(v_x * \rho_n) - (Lv_x) * \rho_n\|_{L^2} \\
&\qquad \leq \|2y(v_x * \rho_n) - 2yv_x\|_{L^2} + \|2yv_x - (2yv_x) * \rho_n\|_{L^2} \\
&\qquad\quad + \|L(v_x * \rho_n) - Lv_x\|_{L^2} + \|Lv_x - (Lv_x) * \rho_n\|_{L^2}.
\end{aligned}$$

For the first term we may factor out y since it is bounded, leaving us with the L^2-norm of $v_x * \rho_n - v_x$, which goes to 0 as $n \to \infty$ by $v_x \in L^2$. The second term also goes to 0 since clearly $2yv_x \in L^2$ when $v \in C_c^\infty$, and similarly for the fourth term. For the third term we have by the linearity of L

$$\|L(v_x * \rho_n) - Lv_x\|_{L^2} = \|L\partial_x(v * \rho_n - v)\|_{L^2} = \|v * \rho_n - v\|_{L^2} \to 0, \tag{6.14}$$

since $L\partial_x$ is an isometry on L^2. Therefore, by C_c^∞ being densely and continuously embedded in L^2, we only need to prove that $\|\mathrm{I}_n(v)\|_{L^2} \leq C\|v\|_{L^2}$ for $v \in \operatorname{dom}(D)$ and some constant C, and the result will follow from continuity. To prove this, note that for any $v \in \operatorname{dom}(D)$,

$$\begin{aligned}
\mathcal{F}[\mathrm{I}_n(v)] &= \mathcal{F}[2y(v_n)_x + L(v_n)_x - ((2yv + Lv)_x - 2y_x v) * \rho_n] \\
&= \mathcal{F}[2y(v_n)_x] + \mathcal{F}[L(v_n)_x] - \mathcal{F}[(2yv + Lv)_x * \rho_n] + \mathcal{F}[(2y_x v) * \rho_n] \\
&= \mathcal{F}[2y(v_n)_x] + i\sqrt{2\pi}\frac{\xi}{|\xi|}\mathcal{F}[v]\mathcal{F}[\rho_n] - \mathcal{F}[(2yv) * (\rho_n)_x] \\
&\quad - i\sqrt{2\pi}\frac{\xi}{|\xi|}\mathcal{F}[v]\mathcal{F}[\rho_n] + \mathcal{F}[(2y_x v) * \rho_n] \\
&= \mathcal{F}[2y(v_n)_x - (2yv) * (\rho_n)_x + (2y_x v) * \rho_n],
\end{aligned}$$

where we have used $(2yv)_x * \rho_n = 2yv * (\rho_n)_x$ and the convolution theorem for the Fourier transform (cf. Theorems 2.5 and 3.6). Then

$$\begin{aligned}
\mathrm{I}_n(v) &= 2y(v * (\rho_n)_x) - (2yv) * (\rho_n)_x + (2y_x v) * \rho_n \\
&= 2\int_{\mathbb{R}} (y(x) - y(x-s))v(x-s)(\rho_n)_x(s)\, ds + (2y_x v) * \rho_n \\
&= 2n^2 \int_{\mathbb{R}} (y(x) - y(x-s))v(x-s)\rho_x(ns)\, ds + (2y_x v) * \rho_n \\
&=: \hat{\mathrm{I}}_n(v) + (2y_x v) * \rho_n.
\end{aligned}$$

Note that by the embedding $H^s \hookrightarrow BC$ for $s > 1/2$, we have $2y_x v \in L^2$ and thus $(2y_x v) * \rho_n \to 2y_x v$ in L^2 as $n \to \infty$. In fact, by Hölder's inequality we have

$$\begin{aligned}
((2y_x v) * \rho_n)(x) &= 2\int_{\mathbb{R}} \rho_n(x-s) y_x(s) v(s)\, ds \\
&= 2\int_{\mathbb{R}} \rho_n(x-s)^{\frac{1}{2}} \rho_n(x-s)^{\frac{1}{2}} y_x(s) v(s)\, ds \\
&\le 2\left(\int_{\mathbb{R}} |\rho(x-s)||y_x(s)v(s)|^2\, ds\right)^{1/2},
\end{aligned}$$

since $\int_{\mathbb{R}} \rho\, dx = 1$, and so by Fubini's theorem

$$\begin{aligned}
\|(2y_x v) * \rho_n\|_{L^2} &\le 2\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |\rho(x-s)||y_x(s)v(s)|^2\, ds\, dx\right)^{1/2} \\
&= 2\left(\int_{\mathbb{R}} |y_x(s)v(s)|^2\, ds\right)^{1/2} = \|2y_x v\|_{L^2}.
\end{aligned}$$

Furthermore, we have $\|y_x v\|_{L^2} \le \|y_x\|_{L^\infty}\|v\|_{L^2}$, and thus in total we have the estimate

$$\|(2y_x v) * \rho_n\|_{L^2} \le 2\|y_x\|_{L^\infty}\|v\|_{L^2}$$

for the second term of $\mathrm{I}_n(v)$. Now we need only to prove that $\|\hat{\mathrm{I}}_n(v)\|_{L^2} \le C'\|v\|_{L^2}$. To this end, suppose $\mathrm{supp}(\rho) \subseteq [-\lambda, \lambda]$. Then we have

$$\begin{aligned}
\|\hat{\mathrm{I}}_n(v)\|_{L^2} &= \|2n\int_{-\lambda}^{\lambda} (y(x) - y(x-s/n))v(x-s/n)\rho_x(s)\, ds\|_{L^2} \\
&\le \|2\sup_{s\in\mathbb{R}} |y_x(s)| \int_{-\lambda}^{\lambda} |sv(x-s/n)\rho_x(s)|\, ds\|_{L^2} \\
&\le 2\|y_x\|_{L^\infty}\|\left(\int_{-\lambda}^{\lambda} |s\rho_x(s)|^2\, ds\right)^{1/2}\left(\int_{-\lambda}^{\lambda} |v(x-s/n)|^2\, ds\right)^{1/2}\|_{L^2},
\end{aligned}$$

by Hölder's inequality. Let $M = 2\|y_x\|_{L^\infty}\left(\int_{-\lambda}^{\lambda} |s\rho_x(s)|^2\, ds\right)^{1/2} < \infty$. By Fubini's theorem we then have

$$\begin{aligned}\|\hat{\mathrm{I}}_n(v)\|_{L^2} &\le M\left(\int_{\mathbb{R}}\int_{-\lambda}^{\lambda}|v(x-s/n)|^2\,ds\,dx\right)^{1/2}\\ &= M\left(\int_{-\lambda}^{\lambda}\int_{\mathbb{R}}|v(x-s/n)|^2\,dx\,ds\right)^{1/2}\\ &= (2\lambda)^{1/2}M\|v\|_{L^2},\end{aligned}$$

which completes the proof. □

We continue by establishing the accretiveness of D and D_0:

Lemma 6.5. *The operators D and D_0 are both quasi-accretive in L^2.*

Proof. By definition, D is quasi-accretive in L^2 if and only if

$$\mathrm{Re}\langle (D+\alpha I)v, v\rangle_{L^2} \ge 0$$

for all $v \in \mathrm{dom}(D)$ and some scalar $\alpha > 0$. In view of Lemma 6.4, we can find a sequence $\{v_n\}_n \subseteq C^\infty$ such that

$$\langle Dv, v\rangle_{L^2} = \lim_{n\to\infty}\langle Dv_n, v_n\rangle_{L^2} = \lim_{n\to\infty}\int_{\mathbb{R}}(2y(v_n)_x + L(v_n)_x)v_n\,dx. \tag{6.15}$$

Note that the operator $L\partial_x$ is skew-symmetric on L^2, as for any $f, g \in L^2$ we have, by Plancherel's theorem,

$$\langle L\partial_x f, g\rangle_{L^2} = \langle \widehat{L\partial_x f}, \hat{g}\rangle_{L^2} = \langle i\,\mathrm{sgn}(\cdot)\hat{f}, \hat{g}\rangle_{L^2} = -\langle f, L\partial_x g\rangle_{L^2}.$$

This implies $\langle L\partial_x v_n, v_n\rangle_{L^2} = -\langle v_n, L\partial_x v_n\rangle_{L^2}$, and thus since L and v_n are real-valued, the term $\int_{\mathbb{R}} L(v_n)_x v_n\,dx$ in (6.15) vanishes for all n. In addition we have

$$\int_{\mathbb{R}} 2y(v_n)_x v_n\,dx = \int_{\mathbb{R}} y(v_n^2)_x\,dx = -\int_{\mathbb{R}} y_x v_n^2\,dx,$$

and so $\langle Dv_n, v_n\rangle_{L^2} = -\int_{\mathbb{R}} y_x v_n^2\,dx$, which gives

$$\mathrm{Re}\langle (D+\alpha I)v_n, v_n\rangle_{L^2} = \int_{\mathbb{R}}(\alpha - y_x)v_n^2\,dx.$$

By the embedding $H^s \hookrightarrow BC$ for $s > 1/2$, we may pick $\alpha \ge \|y_x\|_{L^\infty}$. Then

$$\mathrm{Re}\langle (D+\alpha I)v, v\rangle_{L^2} = \lim_{n\to\infty}\int_{\mathbb{R}}(\alpha - y_x)v_n^2\,dx \ge 0,$$

hence D is quasi-accretive.

Now let us consider the operator D_0. Note that $D_0 v = -Dv - 2y_x v$, thus we have

$$\langle D_0 v, v\rangle_{L^2} = \langle -Dv - 2y_x v, v\rangle_{L^2} = -\langle Dv, v\rangle_{L^2} - \langle 2y_x v, v\rangle_{L^2}.$$

Thus for v_n as above,

$$\langle D_0 v_n, v_n\rangle_{L^2} = \int_{\mathbb{R}} y_x v_n^2 \, dx - 2\int_{\mathbb{R}} y_x v_n^2 \, dx = -\int_{\mathbb{R}} y_x v_n^2 \, dx,$$

and the quasi-accretiveness of D_0 thus follows from the analysis we performed for D. □

Next we would like to prove that D_0 is the adjoint of D. For a densely defined linear operator on a Hilbert space the definition of the adjoint is as follows:

Definition 6.3 (Adjoint of densely defined operator)**.** The adjoint A^* of a densely defined linear operator A on a Hilbert space H is a linear operator with domain

$$\operatorname{dom}(A^*) = \{y \in H \mid \exists\, z \in H \text{ s.t. } \langle Ax, y\rangle_H = \langle x, z\rangle_H \; \forall x \in \operatorname{dom}(A)\}$$

and $A^* y$ equals the z defined thus.

Note that any densely defined linear operator on a Hilbert space (in our case L^2) will have a unique adjoint associated to it [36, p. 348].

Lemma 6.6. *The adjoint of D in L^2 is D_0.*

Proof. For $v \in \mathcal{S} \subseteq \operatorname{dom}(D)$ and any $\omega \in \operatorname{dom}(D^*)$, we have by definition

$$\begin{aligned}
\langle v, D^*\omega\rangle_{L^2} &= \langle Dv, \omega\rangle_{L^2} \\
&= \langle (2yv + Lv)_x - 2y_x v, \omega\rangle_{L^2} \\
&= \langle 2yv_x + Lv_x, \omega\rangle_{L^2} \\
&= \langle 2yv_x, \omega\rangle_{L^2} + \langle Lv_x, \omega\rangle_{L^2}.
\end{aligned}$$

Note that $2yv_x$ and Lv_x are both in L^2, so we are free to split up the inner product this way. For the first term we have $\langle 2yv_x, \omega\rangle_{L^2} = \langle v_x, 2y\omega\rangle_{L^2} = 2\int_{\mathbb{R}} y v_x \omega \, dx$ by y and ω real valued. In the sense of distributions we have

$$\langle 2y\omega, v_x\rangle = -\langle (2y\omega)_x, v\rangle = 2\int_{\mathbb{R}} y\omega v_x \, dx$$

where $(2y\omega)_x \in H^{-1}$ denotes the distributional derivative of the regular tempered distribution $2y\omega \in L^2$. By $L\partial_x$ being skew-symmetric on L^2, we have for the second term

$$\langle Lv_x, \omega\rangle_{L^2} = -\langle v, L\omega_x\rangle_{L^2}.$$

We have $D^*\omega \in L^2$ by definition, thus $D^*\omega$ determines a regular tempered distribution, and we note that

$$\langle D^*\omega, v\rangle = \int_{\mathbb{R}} (D^*\omega)(x)v(x)\,dx = \langle v, D^*\omega\rangle_{L^2},$$

since $D^*\omega$ is real-valued, and similarly

$$\langle L\omega_x, v\rangle = \int_{\mathbb{R}} (L\omega_x)(x)v(x)\,dx = \langle v, L\omega_x\rangle_{L^2}.$$

In total we therefore have

$$\langle D^*\omega, v\rangle = -\langle (2y\omega + L\omega)_x, v\rangle = \langle D_0\omega, v\rangle,$$

for all $v \in \mathcal{S}$. That $D^*\omega = D_0\omega$ as tempered distributions, together with the fact that $D^*\omega \in L^2$ by definition, means that $D_0\omega = D^*\omega$ as elements of L^2 since regular distributions are determined by their generating functions up to pointwise almost everywhere equivalence (see for instance [19, Theorem 1.2.5]). Thus we have $\omega \in \operatorname{dom}(D_0)$ and so $D^* \subseteq D_0$ (by this it is meant precisely that $\operatorname{dom}(D^*) \subseteq \operatorname{dom}(D_0)$ and $D^*\omega = D_0\omega$ for all $\omega \in \operatorname{dom}(D^*)$).

Assume now that $v \in \operatorname{dom}(D_0) = \operatorname{dom}(D)$. Note that for any $u \in \operatorname{dom}(D)$, we can always find a sequence $\{u_n\}_n \subseteq C^\infty$ such that Lemma 6.4 holds. Therefore we have that

$$\langle Du, v\rangle_{L^2} = \lim_{n\to\infty} \langle Du_n, v\rangle_{L^2} = \lim_{n\to\infty} \langle 2y(u_n)_x, v\rangle_{L^2} + \lim_{n\to\infty} \langle L(u_n)_x, v\rangle_{L^2}.$$

For the first term we have

$$\langle 2y(u_n)_x, v\rangle_{L^2} = \langle (u_n)_x, 2yv\rangle_{L^2} = -\langle u_n, (2yv)_x\rangle_{L^2},$$

where $(2yv)_x \in L^2$ since $v \in \operatorname{dom}(D)$. For the second term we have as before

$$\langle L(u_n)_x, v\rangle_{L^2} = -\langle u_n, Lv_x\rangle_{L^2},$$

by $L\partial_x$ being skew-symmetric on L^2. In total then

$$\langle Du, v\rangle_{L^2} = \lim_{n\to\infty} \langle u_n, -(2yv + Lv)_x\rangle_{L^2} = \langle u, D_0 v\rangle_{L^2},$$

and it follows from Definition 6.3 that $v \in \operatorname{dom}(D^*)$, hence $D_0 \subseteq D^*$.

□

Now that we know $D_0 = D^*$, we would like to prove that D and D_0 are closed:

Lemma 6.7. *The operators D and D_0 are closed.*

Proof. The fact that $D_0 = D^*$ means D_0 is closed, as it is the adjoint of a densely defined operator [36, Theorem 13.9]. We can write the operator D as $D = -D_0 + 2y_x v$. From the definition one can verify that $-D_0$ is closed, and the multiplication operator S defined by

$$Sv := 2y_x v$$

is in $\mathcal{B}(L^2)$ for $y \in H^s$ with $s > 3/2$ by the embedding $H^s \hookrightarrow BC^1$ for $s > 3/2$. Let $T := -D_0$. We would like to show that $D = T + S$ is closed. To this end, let $\{x_n\}_n \subseteq \mathrm{dom}(D)$ be a sequence s.t. $x_n \to x$ in L^2 and $Dx_n = Tx_n + Sx_n \to y$ in L^2. We need to verify that $Dx = Tx + Sx = y$. By the continuity of S we clearly have

$$Sx_n \to Sx \text{ in } L^2,$$

thus

$$Tx_n \to y - Sx$$

since $\|Tx_n - (y - Sx)\|_{L^2} \leq \|Tx_n + Sx_n - y\|_{L^2} + \|Sx - Sx_n\|_{L^2} \to 0$. Then because T is closed it holds that $Tx_n \to Tx$ in L^2 by definition, or in other words $y = Tx + Sx$. □

By Lemma 6.5 and Lemma 6.6, both D and D^* are quasi-accretive. A classical argument (cf. [34, Corollary 4.4 p. 15]) then gives the following:

Lemma 6.8. *For the closed linear operator D, densely defined on the Banach space X, with both D and its adjoint D^* quasi-accretive, there exists a scalar $\alpha \in \mathbb{R}$ such that the operator $-(D + \alpha)$ is the infinitesimal generator of a C_0-semigroup of contractions on X, i.e. D is a quasi-m-accretive operator.*

This means that $A(y)$ satisfies condition (i) of Theorem 6.3, since for $y, z, w \in H^s$, $s > 3/2$, we have

$$\begin{aligned} \|(A(y) - A(z))w\|_{L^2} &= 2\|(y - z)w_x\|_{L^2} \\ &\leq 2\|y - z\|_{L^2}\|w_x\|_{L^\infty} \\ &\leq 2C_s\|y - z\|_{L^2}\|w_x\|_{H^{s-1}} \\ &\leq 2C_s\|y - z\|_{L^2}\|w\|_{H^s}, \end{aligned}$$

where the second inequality is due to the Sobolev embedding theorem (see Theorem 2.16). Also, $A(y)$ is in $\mathcal{B}(H^s, L^2)$ with $s > 3/2$, as is clear by the rough estimate

$$\|A(y)v\|_{L^2} = \|2yv_x + Lv_x\|_{L^2} \leq (2\|y\|_{L^\infty} + 1)\|v\|_{H^s}. \tag{6.16}$$

Remark 6.9. This is where the growth condition (6.3) on b comes from, since we in general have $\|\mathcal{L}_\alpha v_x\|_{L^2} \leq C_b\|v\|_{H^{1-\alpha+N_b}}$, which we can only bound by $C_b\|v\|_{H^s}$ if $1 - \alpha + N_b \leq s$.

Now we turn our attention to condition (ii). We will need the concept of a commutator of two operators:

Definition 6.4 (Commutator)**.** Denote by

$$[T, S] = TS - ST$$

the commutator of the two operators T and S.

Remark 6.10. Note that both L and ∂_x are Fourier multiplier operators, so that $[\partial_x, L] = 0$ on H^s for all $s \in \mathbb{R}$.

Let

$$B(y) := QA(y)Q^{-1} - A(y) = [Q, A(y)]Q^{-1},$$

where $Q := \Lambda^s = (1 - \partial_x^2)^{s/2}$ is an isomorphism from H^s to L^2. We then have the following lemma:

Lemma 6.9. *For $y \in Y$, the operator $B(y)$ satisfies condition (ii) of Theorem 6.3.*

Proof. Note that

$$\begin{aligned} [Q, A(y)] &= [\Lambda^s, A(y)] \\ &= 2[\Lambda^s, y]\partial_x + [\Lambda^s, L]\partial_x \\ &= 2[\Lambda^s, y]\partial_x, \end{aligned} \tag{6.17}$$

where we have used the commutation properties $[\Lambda^s, \partial_x] = 0$ and $[\Lambda^s, L] = 0$.

In order to prove uniform boundedness of $B(y)$ for y in a bounded subset of H^s, we assume without loss of generality that $y \in \mathcal{W} \subseteq H^s$, where $\mathcal{W}$ is an open ball in H^s with radius $R > 0$. By classical estimates for (6.17) (cf. [23, Lemma A.2]), we get (for $s > 3/2$)

$$\|[\Lambda^s, y]\Lambda^{1-s}\| \leq C_0 \|\partial_x y\|_{H^{s-1}} \leq C_0 \|y\|_{H^s} \leq C_0 R =: \alpha_0(R),$$

where C_0 only depends on s and $\|\cdot\|$ on the left denotes the operator norm on L^2 (in the general case $p \in \{2, 3, \cdots\}$ one should here apply Theorem 2.18 in order to set $\|\partial_x y^{p-1}\|_{H^{s-1}} = (p-1)\|y^{p-2} y_x\|_{H^{s-1}} \leq c_s \|y^{p-2}\|_{H^{s-1}} \|y_x\|_{H^{s-1}}$ etc.). Then for any $z \in L^2$, we have

$$\begin{aligned} \|B(y)z\|_{L^2} &= 2\|[\Lambda^s, y]\Lambda^{1-s}\Lambda^{s-1}\partial_x \Lambda^{-s} z\|_{L^2} \\ &\leq 2\|[\Lambda^s, y]\Lambda^{1-s}\| \|\Lambda^{s-1}\partial_x \Lambda^{-s} z\|_{L^2} \\ &\leq 2\alpha_0(R)\|\partial_x \Lambda^{-1} z\|_{L^2} \\ &\leq 2\alpha_0(R)\|z\|_{L^2}, \end{aligned}$$

where the last step is due to the fact that $\|\partial_x \Lambda^{-1} z\|_{L^2} \le \|\Lambda^{-1} z\|_{H^1} = \|z\|_{L^2}$. Hence $B(y)$ is a bounded linear operator on L^2 for $y \in Y$. Furthermore, for any $y, z \in \mathcal{W}$ and $w \in X$, we have

$$\begin{aligned}\|(B(y) - B(z))w\|_{L^2} &= 2\|[\Lambda^s, y - z]\partial_x \Lambda^{-s} w\|_{L^2} \\ &\le 2\|[\Lambda^s, y - z]\Lambda^{1-s}\| \|\Lambda^{s-1}\partial_x \Lambda^{-s} w\|_{L^2} \\ &\le 2C_0 \|y - z\|_{H^s} \|w\|_{L^2}.\end{aligned}$$

Thus $B(y)$ satisfies condition (ii) of Theorem 6.3. □

We are now ready to prove the main theorem in its full generality for initial data $u_0 \in H^s(\mathbb{R})$, $s > 3/2$ (refer to the next section for the technical differences in the analysis above in the general case of $0 < \alpha \le 1$ and b of slow growth):

Proof of Theorem 6.1. By Lemmata 6.4 through 6.9 we can apply Theorem 6.3 to find a solution u as described in Theorem 6.1, although in the solution class $C^0([0,T), H^s(\mathbb{R})) \cap C^1([0,T), L^2(\mathbb{R}))$. However, in view of that $H^s(\mathbb{R})$ is an algebra with respect to pointwise multiplication for $s > 1/2$, and that $\mathcal{L}_\alpha \partial_x$ maps $H^s(\mathbb{R})$ continuously into $H^{s-1+\alpha-N_b}(\mathbb{R})$ (where $s - 1 + \alpha - N_b \ge 0$ by (6.3)) since

$$\|\mathcal{L}_\alpha u_x\|^2_{H^{s-1+\alpha-N_b}(\mathbb{R})} = C_b^2 \int_{\mathbb{R}} \langle \xi \rangle^{2(s-1+\alpha-N_b)} |\xi|^{2(1-\alpha+N_b)} |\hat{u}(\xi)|^2 \, d\xi \le C_b^2 \|u\|^2_{H^s(\mathbb{R})},$$

one sees that for the equation (6.1) we have

$$u_t = -(u^p)_x - \mathcal{L}_\alpha u_x \in H^q(\mathbb{R}),$$

where $q = \min\{s-1, s-1+\alpha-N_b\} \ge 0$. Hence $u \in C^1([0,T), H^q(\mathbb{R}))$

Also, since the data-to-solution map $u_0 \mapsto u$ is continuous from $H^s(\mathbb{R})$ to $C^0([0,T), H^s(\mathbb{R}))$, a similar argument can be used to conclude that the data-to-solution map is continuous from $H^s(\mathbb{R})$ to $C^1([0,T), H^q(\mathbb{R}))$. □

6.2 The general case $\alpha \in (0,1]$ and b of slow growth

In our analysis of (6.10), it was crucial that we considered the composition $\mathcal{L}_1 \partial_x$ of the two operators $\mathcal{L}_1$ and ∂_x as a single operator, since this operator corresponds simply to multiplication by $i \operatorname{sgn}$ on the Fourier side, thus it acts as an isometry on L^2. This allowed us to adapt the analysis performed in [15], with a few modifications.

In the general case where $\alpha \in (0,1]$ and b is not bounded but of slow growth, the operator $\mathcal{L}_\alpha \partial_x$ corresponds to multiplication by $ib(\xi)\xi|\xi|^{-\alpha}$ on the Fourier side. Therefore, we have for some $C_b > 0$ and $N_b \ge 0$ satisfying (6.3)

$$|\mathcal{F}(\mathcal{L}_\alpha \partial_x u)(\xi)| \le C_b (1 + |\xi|)^{1-\alpha+N_b} |\hat{u}(\xi)|,$$

and $\mathcal{L}_\alpha \partial_x$ is therefore an isometry $H^{1-\alpha+N_b} \to L^2$. We would like to use this to prove local well-posedness in H^s for $s > 3/2$ as before. Again we set $p = 2$ for simplicity, but note that the analysis easily generalises to $p \geq 2$.

The proof proceeds in the same steps, but we will need to modify some of the arguments which we used to prove Lemma 6.4-6.6. The rest of the analysis remains unchanged, since it is not as sensitive to changes in the exact form of the linear term.

First of all, note that D and D_0, defined as in (6.12) and (6.13) respectively, are still densely defined operators on L^2, since H^{1+N_b} is easily verified to be a subset of $\mathrm{dom}(D)$. Also, both operators are closed with the same domains as before, which again will follow from the fact that $D^* = D_0$ and $D = -D_0 + 2y_x v$.

We begin as before with the following approximation result:

Lemma 6.10. *Given $v \in \mathrm{dom}(\mathrm{D})$, there exists a sequence $\{v_n\}_n \subseteq C^\infty$ such that*

$$v_n \to v \quad \textit{and} \quad (2yv_n + Lv_n)_x \to (2yv + Lv)_x$$

in L^2 as $n \to \infty$.

Proof. In view of the proof of Lemma 6.4, the lemma holds by the following observations that justify the steps in that proof now that $L := \mathcal{L}_\alpha$ with $\alpha \in (0,1]$ and $|b(\xi)| \leq C_b(1+|\xi|)^{N_b}$:

- The proof of the claim that $\mathrm{I}_n(v) \to 0$ in L^2 for $v \in C_c^\infty$ is the same up to the equation (6.14). From there we instead use the following estimate to get the desired result, using the fact that $v_n = v * \rho_n \in C_c^\infty$ for $v, \rho_n \in C_c^\infty$:

$$\begin{aligned} \|L(v_x * \rho_n) - Lv_x\|_{L^2} &= \|L\partial_x(v_n - v)\|_{L^2} \\ &\leq C_b \|v_n - v\|_{H^{1-\alpha+N_b}} \\ &= C_b \|\Lambda^{1-\alpha+N_b}(v_n - v)\|_{L^2} \\ &\leq C_b \|\Lambda^{1-\alpha+N_b}\|_{C_c^\infty \subseteq L^2} \|v_n - v\|_{L^2} \to 0 \quad \text{as } n \to \infty, \end{aligned}$$

  where $\|\Lambda^{1-\alpha+N_b}\|_{C_c^\infty \subseteq L^2}$ denotes the operator norm of $\Lambda^{1-\alpha+N_b}$ over the subset of L^2 containing the compactly supported smooth functions, that is

$$\|\Lambda^{1-\alpha+N_b}\|_{C_c^\infty \subseteq L^2} = \sup_{\{v \in C_c^\infty : \|v\|_{L^2} = 1\}} \|\Lambda^{1-\alpha+N_b} v\|_{L^2},$$

  which is clearly finite since the fractional derivative operator $\Lambda^{1-\alpha+N_b}$ maps $v \in C_c^\infty$ to another compactly supported smooth function.

- Note that we still have the sum $2yv_x + Lv_x \in L^2$ for $v \in \mathrm{dom}(D)$. This is true by a similar argument to the one in Remark 6.8 since clearly $2y_x v \in L^2$. Thus we can say $\mathrm{II}_n(v) \to 0$ in L^2 as $n \to \infty$. The same is clearly true for $\mathrm{III}_n(v)$.

- It is no longer true that each term in the expression $2y_x v + 2yv_x + Lv_x$ is in L^2 individually, as we pointed out in Remark 6.8 in order to justify taking the Fourier transform of the terms $2y(v_n)_x$ and $(Lv_n)_x$ separately. However, the derivative of $v_n \in C^\infty \cap L^2$ is a bounded function since it is smooth and goes to 0 at infinity, therefore we have $2y(v_n)_x \in L^2$. Hence we also have $L(v_n)_x \in L^2$, see the argument leading up to (6.18) below. This all means that the Fourier transform of these terms is well-defined.

  Furthermore, we claim the Fourier transform of the terms $(2yv)_x * \rho_n$ and $Lv_x * \rho_n$ is also well-defined. The claim holds for the first term since the convolution between $(2yv)_x \in H^{-1} \subseteq \mathcal{S}'$ and $\rho_n \in \mathcal{S}$ is again a tempered distribution (see Theorem 3.6). As for the second term, we show in the argument leading up to (6.19) below that Lv_x is also a well-defined tempered distribution.

$\square$

The argument in the proof of Lemma 6.5 also needs some modification:

Lemma 6.11. *The operators D and D_0 are both quasi-accretive in L^2.*

Proof. In view of the proof of Lemma 6.5, we only need to show that

$$\langle L\partial_x v_n, v_n\rangle_{L^2} = -\langle v_n, L\partial_x v_n\rangle_{L^2} \quad \text{for } v_n \in C^\infty \cap \operatorname{dom}(D)$$

still holds now that $L = \mathcal{L}_\alpha$ with $\alpha \in (0, 1]$ and b is of slow growth. The claim that $L\partial_x$ is skew-symmetric on L^2 is no longer valid, since $L\partial_x$ does not map L^2 into L^2, however for $v_n \in C^\infty \cap \operatorname{dom}(D)$, we do indeed have $L\partial_x v_n \in L^2$, which allows us to apply Plancherel's theorem to the inner product $\langle L\partial_x v_n, v_n\rangle_{L^2}$ to produce the desired equality. This follows from the fact that for $v_n \in C^\infty \cap \operatorname{dom}(D)$, we have $2y(v_n)_x \in L^2$, since $(v_n)_x$ is bounded because it is smooth and goes to 0 at infinity. Then by an argument similar to the one in Remark 6.8, the definition of $\operatorname{dom}(D)$ implies that

$$L\partial_x v_n \in L^2. \tag{6.18}$$

$\square$

Next we prove that $D^* = D_0$, using a slightly different and more technical argument than in the proof of Lemma 6.6:

Lemma 6.12. *The adjoint of D in L^2 is D_0.*

Proof. We prove the equality $D^* = D_0$ by proving the two inclusions, starting with $D^* \subseteq D_0$. By definition, we have for $v \in \mathcal{S} \subseteq \mathrm{dom}(D)$ and any $\omega \in \mathrm{dom}(D^*)$

$$\begin{aligned}\langle v, D^*\omega\rangle_{L^2} &= \langle Dv, \omega\rangle_{L^2}\\ &= \langle (2yv + Lv)_x - 2y_x v, \omega\rangle_{L^2}\\ &= \langle 2yv_x + Lv_x, \omega\rangle_{L^2}\\ &= \langle 2yv_x, \omega\rangle_{L^2} + \langle Lv_x, \omega\rangle_{L^2},\end{aligned}$$

since we have $2yv_x \in L^2$ and $Lv_x \in L^2$ for $v \in \mathcal{S}$ (this is clearly true for the first term due to the invariance of $\mathcal{S}$ under differentiation, and $2yv_x \in L^2$ in fact implies $Lv_x \in L^2$ since by definition $2yv_x + Lv_x \in L^2$ for $v \in \mathrm{dom}(D)$).

By the same argument as in the proof of Lemma 6.6, we have

$$\langle 2yv_x, \omega\rangle_{L^2(\mathbb{T})} = -\langle (2y\omega)_x, v\rangle.$$

As for the term $\langle L\partial_x v, \omega\rangle_{L^2}$, we have by Plancherel's theorem

$$\begin{aligned}\langle L\partial_x v, \omega\rangle_{L^2} = \langle \widehat{L\partial_x v}, \hat{\omega}\rangle_{L^2} &= \int_{\mathbb{R}} \frac{i\xi}{|\xi|^\alpha} b(\xi)\hat{v}(\xi)\overline{\hat{\omega}(\xi)}\, d\xi\\ &= \overline{-\int_{\mathbb{R}} \frac{i\xi}{|\xi|^\alpha} b(\xi)\overline{\hat{v}(\xi)}\hat{\omega}(\xi)\, d\xi}.\end{aligned}$$

We claim that the function $\widehat{L\partial_x\omega} : \xi \mapsto i\xi\, b(\xi)|\xi|^{-\alpha}\hat{\omega}(\xi)$ generates a regular tempered distribution. Indeed, $\widehat{L\partial_x\omega}$ is a well-defined linear functional $\mathcal{S} \to \mathbb{C}$ since $\langle \widehat{L\partial_x\omega}, \varphi\rangle$ for any $\varphi \in \mathcal{S}$ can be expressed as an L^2 inner product,

$$\langle \widehat{L\partial_x\omega}, \varphi\rangle = \int_{\mathbb{R}} i\xi|\xi|^{-\alpha} b(\xi)\hat{\omega}(\xi)\varphi(\xi)\, d\xi = \langle \hat{\omega}, -i\xi|\xi|^{-\alpha} b(\xi)\varphi\rangle_{L^2} < \infty$$

by $\omega \in L^2$ and $\varphi \in \mathcal{S}$. To show continuity, we let $\varphi_n \to 0$ in $\mathcal{S}$, and define the quantities I_n and II_n by

$$\langle \widehat{L\partial_x\omega}, \varphi_n\rangle = \underbrace{\int_{|\xi|\le 1} \frac{i\xi}{|\xi|^\alpha} b(\xi)\hat{\omega}(\xi)\varphi_n(\xi)\, d\xi}_{\mathrm{I}_n} + \underbrace{\int_{|\xi|>1} \frac{i\xi}{|\xi|^\alpha} b(\xi)\hat{\omega}(\xi)\varphi_n(\xi)\, d\xi}_{\mathrm{II}_n}.$$

The integrand of I_n is dominated by $M_b|\hat{\omega}|$ for some constant $M_b > 0$, since

$$C_b(1 + |\xi|)^{1-\alpha+N_b} \le 2^{1-\alpha+N_b} C_b$$

and $|\varphi_n| \le \|\varphi_n\|_{0,0} \to 0$ as $n \to \infty$, thus $|\varphi_n| \le M$ for some M since a convergent sequence is bounded. The function $M_b|\hat{\omega}|$ is locally integrable since $\omega \in L^2$, thus $\mathrm{I}_n \to 0$ as $n \to \infty$ by Lebesgue's dominated convergence theorem.

As for II_n, note that for any $k \in \mathbb{N}_0$ we have $|\varphi_n(\xi)| = \frac{(1+|\xi|)^k|\varphi_n(\xi)|}{(1+|\xi|)^k} \le \frac{P_{k,0}(\varphi_n)}{(1+|\xi|)^k}$, where $P_{k,0}(\varphi_n)$ is the seminorm of φ_n defined as per (2.3). Thus

$$(1+|\xi|)^{1-\alpha+N_b}|\varphi_n(\xi)| \le (1+|\xi|)^{1-\alpha+N_b}\frac{P_{2-\alpha+N_b+q,0}(\varphi_n)}{(1+|\xi|)^{1+(1-\alpha+N_b)+q}} = \frac{P_{2-\alpha+N_b+q,0}(\varphi_n)}{(1+|\xi|)^{1+q}}$$

where $q := \lceil 2-\alpha+N_b \rceil - (2-\alpha+N_b)$. We have $P_{2-\alpha+N_b+q,0}(\varphi_n) \to 0$ as $n \to \infty$, thus the integrand of II_n is dominated by $m|\hat{\omega}|(1+|\xi|)^{-1-q}$ for some $m > 0$, which is integrable over $\mathbb{R} \setminus [-1,1]$. This can be seen for instance by expressing the integral of the dominating function as an L^2 inner product,

$$\begin{aligned}\mathrm{II}_n &\le m\int_{|\xi|>1}\frac{|\hat{\omega}(\xi)|}{(1+|\xi|)^{1+q}}\,d\xi \\ &= m\langle|\hat{\omega}|, \frac{1}{(1+|\xi|)^{1+q}}\rangle_{L^2(-\infty,-1)} + m\langle|\hat{\omega}|, \frac{1}{(1+|\xi|)^{1+q}}\rangle_{L^2(1,\infty)} < \infty.\end{aligned}$$

Hence we also have $\mathrm{II}_n \to 0$ as $n \to \infty$ and therefore $\lim_{n\to\infty}\langle\widehat{L\partial_x\omega}, \varphi_n\rangle = 0$. This proves the continuity of $\widehat{L\partial_x\omega}$ as a linear functional on $\mathcal{S}$.

Applying the identity $\hat{\hat{\psi}} = \tilde{\hat{\hat{\psi}}}$ (where $\sim$ is the reflection), we then have by the definition of the Fourier transform on $\mathcal{S}'$,

$$\langle\widehat{L\partial_x\omega}, \bar{\hat{v}}\rangle = \langle L\partial_x\omega, \hat{\bar{\hat{v}}}\rangle = \langle L\partial_x\omega, \bar{v}\rangle.$$

Thus

$$\langle L\partial_x v, \omega\rangle_{L^2} = -\overline{\int_{\mathbb{R}}\frac{i\xi}{|\xi|^\alpha}b(\xi)\overline{\hat{v}(\xi)}\hat{\omega}(\xi)\,d\xi} = -\langle\overline{L\partial_x\omega}, v\rangle = -\langle L\partial_x\omega, v\rangle,$$

where the final equality holds since $L\partial_x\omega$ is real-valued. Hence for all $v \in \mathcal{S}$ we have like before

$$\langle D^*\omega, v\rangle = \langle -(2y\omega + L\omega)_x, v\rangle = \langle D_0\omega, v\rangle, \tag{6.19}$$

i.e. $D^*\omega = D_0\omega$ as tempered distributions. Since $D^*\omega \in L^2$ by definition, this implies $D_0\omega \in L^2$ and so $\omega \in \mathrm{dom}(D_0)$ and $D^* \subseteq D_0$.

Now we set out to prove the other inclusion. Suppose $v \in \mathrm{dom}(D_0) = \mathrm{dom}(D)$ and $u \in \mathrm{dom}(D)$. For such u and v, we can find $\{u_n\}_n \subseteq C^\infty \cap \mathrm{dom}(D)$ and $\{v_n\}_n \subseteq C^\infty \cap \mathrm{dom}(D)$ such that Lemma 6.10 holds. Then we have that

$$\langle Du, v\rangle_{L^2} = \lim_{n\to\infty}\langle Du_n, v_n\rangle_{L^2} = \lim_{n\to\infty}\langle 2y(u_n)_x, v_n\rangle_{L^2} + \lim_{n\to\infty}\langle L(u_n)_x, v_n\rangle_{L^2}.$$

We may again split up the inner product in this way because $2y(u_n)_x \in L^2$ and $L(u_n)_x \in L^2$ for $u_n \in C^\infty \cap \mathrm{dom}(D)$ by the argument leading up to (6.18) above.

For the first term we have

$$\langle 2y(u_n)_x, v_n\rangle_{L^2} = \langle (u_n)_x, 2yv_n\rangle_{L^2} = -\langle u_n, (2yv_n)_x\rangle_{L^2},$$

where $(2yv_n)_x \in L^2$ since $v_n \in C^\infty \cap \operatorname{dom}(D)$.

For the second term we have as before

$$\langle L(u_n)_x, v_n\rangle_{L^2} = -\langle u_n, L(v_n)_x\rangle_{L^2}.$$

Thus in total

$$\langle Du, v\rangle_{L^2} = \lim_{n\to\infty} \langle u_n, -(2yv_n + Lv_n)_x\rangle_{L^2} = \langle u, D_0 v\rangle_{L^2},$$

and it then follows from Definition 6.3 that $v \in \operatorname{dom}(D^*)$ and so $D_0 \subseteq D^*$.

□

The rest of the proof continues exactly like in the special case $\alpha = 1$ and b bounded.

6.3 The periodic case for $\alpha = 1$ and b bounded

We are now interested in analysing the periodic Cauchy problem (6.5). By applying Fourier theory for periodic functions and distributions, and the theory of periodic Sobolev spaces from Section 4, we can follow the steps in the analysis of the Cauchy problem on the real line to prove local well-posedness in $H^s(\mathbb{T})$, $s > 3/2$.

A difficulty we immediately encounter in our analysis is that the sequence of Fourier coefficients $\{\mathcal{F}(\mathcal{L}_\alpha f)(k)\}_{k\in\mathbb{Z}}$ is clearly not in general summable. However, when considering the composition of operators $\mathcal{L}_\alpha \partial_x$, the differential operator cancels out the singularity on the Fourier side. In the case $\alpha = 1$ and b bounded, the operators cancel out such that the composition $\mathcal{L}_1\partial_x$, which we shall treat as a single operator, is an isometry on $L^2(\mathbb{T})$. This simplifies the analysis and makes it easier to follow. We therefore treat this case in detail first, and in the next subsection we expand on the differences in the analysis in the more intricate and general case of $\alpha \in (0,1]$ and b of slow growth.

Remark 6.11. In the case where $b \equiv 1$, by Lemma 4.5 the operator $\mathcal{L}_1\partial_x$ corresponds to $i\,\mathrm{sgn}$ on the Fourier side. One can compute the Fourier coefficients of the 2π-periodic distribution $f(x) = -\cot(\frac{x}{2})$ to be $\hat{f}(k) = -i\,\mathrm{sgn}(k)$ (cf. [21, p. 195]). Thus, by Theorem 4.11, for $f \in \mathcal{P}'$ we have

$$\mathcal{L}_1\partial_x f = -\cot(\frac{x}{2}) * f.$$

That is, the operator $\mathcal{L}_1\partial_x$ actually behaves as convolution with $-\cot(\frac{x}{2})$.

We follow the proof of well-posedness of the Cauchy problem on real line via Kato's theorem, now working in the Hilbert spaces $Y := H^s(\mathbb{T})$ with $s > 3/2$ and $X := H^0(\mathbb{T}) = L^2(\mathbb{T})$. As before, we assume for clarity that $p = 2$ and $b \equiv 1$, and we define $L := \mathcal{L}_1$. We rewrite (6.5) as

$$\begin{cases} u_t + A(u)u_x = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \text{ with } u \text{ a } 2\pi\text{-periodic distribution,} \\ u(0,x) = u_0(x) & \text{for } x \in \mathbb{R}, \end{cases}$$

where

$$\begin{aligned} A(y) &:= (2y + L)\partial_x \\ \operatorname{dom}(A(y)) &:= \{v \in L^2(\mathbb{T}) \mid (2yv + Lv)_x \in L^2(\mathbb{T})\}, \end{aligned}$$

for some $y \in H^s(\mathbb{T})$, $s > 3/2$. One can easily verify that $H^1(\mathbb{T}) \subseteq \operatorname{dom}(A(y))$, thus by Theorem 4.14 the operator $A(y)$ is densely defined in $L^2(\mathbb{T})$.

For $y \in Y$ we also define

$$\begin{aligned} Dv &:= (2yv + Lv)_x - 2y_x v, \\ \operatorname{dom}(D) &:= \{v \in L^2(\mathbb{T}) \mid (2yv + Lv)_x \in L^2(\mathbb{T})\}, \end{aligned} \tag{6.20}$$

and

$$\begin{aligned} D_0 v &:= -(2yv + Lv)_x, \\ \operatorname{dom}(D_0) &:= \{v \in L^2(\mathbb{T}) \mid (2yv + Lv)_x \in L^2(\mathbb{T})\}. \end{aligned} \tag{6.21}$$

The choice of these domains make D and D_0 closed operators in $L^2(\mathbb{T})$, see Lemma 6.16.

We now set out to prove that D, and thus $A(y)$, satisfies condition (i) of Theorem 6.3 by help of a few lemmas.

Lemma 6.13. *Given $v \in \operatorname{dom}(D)$, there exists a sequence $\{v_n\}_n \subseteq C^\infty(-\pi,\pi)$ such that*

$$v_n \to v \quad \textit{and} \quad (2yv_n + Lv_n)_x \to (2yv + Lv)_x$$

in $L^2(\mathbb{T}) = L^2(-\pi,\pi)$ as $n \to \infty$.

Proof. Pick $\rho \in C_c^\infty$ with $\rho \geq 0$ and $\int_{\mathbb{R}} \rho \, dx = 1$. For $n \geq 1$, let $\rho_x(x) := n\rho(nx)$ be a mollifier. Defining v_n by

$$v_n(x) := (v * \rho_n)(x) = \int_{-\pi}^{\pi} v(y)\rho(x-y)\,dy,$$

we have $v_n \in C^\infty(-\pi,\pi) \cap L^2(\mathbb{T})$ by Young's inequality and $v_n \to v$ in $L^2(-\pi,\pi) = L^2(\mathbb{T})$ (cf. [39, Lemma 7.1 c]). This proves the first part of the lemma.

As what concerns the second part, we have as before

$$\begin{aligned}
&(2yv_n + Lv_n)_x - (2yv + Lv)_x \\
&= (2y(v_n)_x + L(v_n)_x - (2yv_x + Lv_x)) + (2y_x v_n - 2y_x v) \\
&= (2y(v_n)_x + L(v_n)_x - (2yv_x + Lv_x) * \rho_n) \\
&\quad + ((2yv_x + Lv_x) * \rho_n - (2yv_x + Lv_x)) \\
&\quad + (2y_x v_n - 2y_x v) =: \mathrm{I}_n(v) + \mathrm{II}_n(v) + \mathrm{III}_n(v).
\end{aligned}$$

By $2yv_x + Lv_x \in L^2(\mathbb{T})$ for $v \in \mathrm{dom}(D)$, one has

$$\mathrm{II}_n(v) = (2yv_x + Lv_x) * \rho_n - (2yv_x + Lv_x) \to 0$$

in $L^2(\mathbb{T})$ as $n \to \infty$. Also, by the embedding $H^s(\mathbb{T}) \hookrightarrow BC(\mathbb{T})$ for $s > 1/2$, one has

$$\mathrm{III}_n(v) = 2y_x(v_n - v) \to 0$$

in $L^2(\mathbb{T})$ as $n \to \infty$.

It then remains to show that $\mathrm{I}_n(v) \to 0$ in $L^2(\mathbb{T})$ as $n \to \infty$. Again we claim that this holds for $v \in C_c^\infty(-\pi,\pi)$, the argument is the same as for the real line case. Then by $C_c^\infty(-\pi,\pi)$ being densely and continuously embedded in $L^2(\mathbb{T})$ [39, Theorem 7.1], we only need to prove that $\|\mathrm{I}_n(v)\|_{L^2(\mathbb{T})} \le C\|v\|_{L^2(\mathbb{T})}$ for $v \in \mathrm{dom}(D)$ and some constant $C > 0$ and the second result of the lemma will follow from continuity. We set out to prove this inequality, and first note that for any $v \in \mathrm{dom}(D)$,

$$\begin{aligned}
\mathcal{F}(\mathrm{I}_n(v)) &= \mathcal{F}(2y(v_n)_x + L(v_n)_x - ((2yv + Lv)_x - 2y_x v) * \rho_n) \\
&= \mathcal{F}(2y(v_n)_x) + \mathcal{F}(L(v_n)_x) - \mathcal{F}((2yv + Lv)_x * \rho_n) + \mathcal{F}((2y_x v) * \rho_n) \\
&= \mathcal{F}(2y(v_n)_x) + i\frac{k}{|k|}\mathcal{F}(v)\mathcal{F}(\rho_n) - \mathcal{F}((2yv) * (\rho_n)_x) \\
&\quad - i\frac{k}{|k|}\mathcal{F}(v)\mathcal{F}(\rho_n) + \mathcal{F}((2y_x v) * \rho_n) \\
&= \mathcal{F}(2y(v_n)_x - (2yv) * (\rho_n)_x + (2y_x v) * \rho_n),
\end{aligned}$$

where we have used $(2yv)_x * \rho_n = 2yv * (\rho_n)_x$ and the convolution theorem for the Fourier transform (cf. Theorem 4.11). Note that the Fourier transform of the terms $2y(v_n)_x$ and $L(v_n)_x$ is well-defined since $L(v_n)_x \in L^2(\mathbb{T})$ and $2y_x v_n \in L^2(\mathbb{T})$ implies $2y(v_n)_x \in L^2(\mathbb{T})$ for $v \in \mathrm{dom(D)}$ by the definition of dom(D). Then

$$\begin{aligned}
\mathrm{I}_n(v) &= 2y(v * (\rho_n)_x) - (2yv) * (\rho_n)_x + (2y_x v) * \rho_n \\
&= 2\int_{-\pi}^{\pi} (y(x) - y(x-s))v(x-s)(\rho_n)_x(s)\,ds + (2y_x v) * \rho_n \\
&= 2n^2 \int_{-\pi}^{\pi} (y(x) - y(x-s))v(x-s)\rho_x(ns)\,ds + (2y_x v) * \rho_n \\
&=: \hat{\mathrm{I}}_n(v) + (2y_x v) * \rho_n.
\end{aligned}$$

For the second term we have $\|2y_x v\|_{L^2(\mathbb{T})} \le 2\|y_x\|_\infty \|v\|_{L^2(\mathbb{T})}$, see the proof of Lemma 6.4. It only remains to prove that $\|\hat{\mathrm{I}}_n(v)\|_{L^2(\mathbb{T})} \le C'\|v\|_{L^2(\mathbb{T})}$. We have

$$\begin{aligned}
\|\hat{\mathrm{I}}_n(v)\|_{L^2} &= 2n\| \int_{-\pi}^{\pi} (y(x) - y(x - s/n))v(x - s/n)\rho_x(s)\, ds\|_{L^2(\mathbb{T})} \\
&\le 2 \sup_{s\in(-\pi,\pi)} |y_x(s)| \| \int_{-\pi}^{\pi} |sv(x - s/n)\rho_x(s)|\, ds\|_{L^2(\mathbb{T})} \\
&\le 2\|y_x\|_\infty \left(\int_{-\pi}^{\pi} |s\rho_x(s)|^2\, ds \right)^{1/2} \| \left(\int_{-\pi}^{\pi} |v(x - s/n)|^2\, ds \right)^{1/2} \|_{L^2(\mathbb{T})},
\end{aligned}$$

by Hölder's inequality. Let $M = 2\|y_x\|_\infty \left(\int_{-\pi}^{\pi} |s\rho_x(s)|^2\, ds\right)^{1/2} < \infty$. By Fubini's theorem then

$$\begin{aligned}
\|\hat{\mathrm{I}}_n(v)\|_{L^2} &\le M \left(\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |v(x - s/n)|^2\, ds\, dx \right)^{1/2} \\
&= M \left(\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |v(x - s/n)|^2\, dx\, ds \right)^{1/2} \\
&= (2\pi)^{1/2} M \|v\|_{L^2(\mathbb{T})},
\end{aligned}$$

which completes the proof. □

Lemma 6.14. *The operators D and D_0 are both quasi-accretive in $L^2(\mathbb{T})$.*

Proof. By definition, D is quasi-accretive in $L^2(\mathbb{T})$ if and only if

$$\mathrm{Re}\langle (D + \alpha I)v, v\rangle_{L^2(\mathbb{T})} \ge 0$$

for all $v \in \mathrm{dom}(D)$ and some scalar $\alpha > 0$. In view of Lemma 6.13, we can find a sequence $\{v_n\}_n \subseteq C^\infty(-\pi, \pi)$ such that

$$\langle Dv, v\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \langle Dv_n, v_n\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \int_{-\pi}^{\pi} (2y(v_n)_x + L(v_n)_x)v_n\, dx. \tag{6.22}$$

Note that the operator $L\partial_x$ is skew-symmetric on $L^2(\mathbb{T})$, as for any $f, g \in L^2(\mathbb{T})$ we have,

$$\langle L\partial_x f, g\rangle_{L^2(\mathbb{T})} = 2\pi \langle \widehat{L\partial_x f}, \hat{g}\rangle_{l^2(\mathbb{Z})} = 2\pi \langle i\,\mathrm{sgn}(\cdot)\hat{f}, \hat{g}\rangle_{l^2(\mathbb{Z})} = -\langle f, L\partial_x g\rangle_{L^2(\mathbb{T})}.$$

This implies

$$\langle L\partial_x v_n, v_n\rangle_{L^2(\mathbb{T})} = -\langle v_n, L\partial_x v_n\rangle_{L^2(\mathbb{T})},$$

and thus since L and v_n are real-valued, the term $\int_{-\pi}^{\pi} L(v_n)_x v_n\, dx$ in (6.22) vanishes for all n.

In addition, we have by Parseval's identity,

$$\begin{aligned}\langle 2y(v_n)_x, v_n\rangle_{L^2(\mathbb{T})} &= \langle y, 2(v_n)_x v_n\rangle_{L^2(\mathbb{T})} = \langle y, (v_n^2)_x\rangle_{L^2(\mathbb{T})} \\ &= 2\pi\langle \hat{y}(k), ik\,\widehat{v_n^2}(k)\rangle_{l^2(\mathbb{Z})} = -2\pi\langle ik\,\hat{y}(k),\, \widehat{v_n^2}(k)\rangle_{l^2(\mathbb{Z})} \\ &= -\langle y_x, v_n^2\rangle_{L^2(\mathbb{T})},\end{aligned}$$

hence $\langle Dv_n, v_n\rangle_{L^2(\mathbb{T})} = -\int_{-\pi}^{\pi} y_x v_n^2\, dx$. This gives

$$\operatorname{Re}\langle (D+\alpha I)v_n, v_n\rangle_{L^2(\mathbb{T})} = \int_{-\pi}^{\pi} (\alpha - y_x) v_n^2\, dx.$$

By the embedding $H^s(\mathbb{T}) \hookrightarrow BC(\mathbb{T})$ for $s > 1/2$, we may pick $\alpha \geq \|y_x\|_\infty$. Then

$$\operatorname{Re}\langle (D+\alpha I)v, v\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \int_{-\pi}^{\pi} (\alpha - y_x) v_n^2\, dx \geq 0,$$

hence D is quasi-accretive.

Now let us consider the operator D_0. Note that $D_0 v = -Dv - 2y_x v$, thus we have

$$\langle D_0 v, v\rangle_{L^2(\mathbb{T})} = \langle -Dv - 2y_x v, v\rangle_{L^2(\mathbb{T})} = -\langle Dv, v\rangle_{L^2(\mathbb{T})} - \langle 2y_x v, v\rangle_{L^2(\mathbb{T})}.$$

Hence for v_n as above

$$\langle D_0 v_n, v_n\rangle_{L^2(\mathbb{T})} = \int_{\mathbb{R}} y_x v_n^2\, dx - 2\int_{\mathbb{R}} y_x v_n^2\, dx = -\int_{\mathbb{R}} y_x v_n^2\, dx,$$

and the quasi-accretiveness of D_0 thus follows from the analysis we performed for D. □

Lemma 6.15. *The adjoint of D in $L^2(\mathbb{T})$ is D_0.*

Proof. For $v \in \mathcal{P} \subseteq \operatorname{dom}(D)$ and any $\omega \in \operatorname{dom}(D^*)$, we have by definition

$$\begin{aligned}\langle v, D^*\omega\rangle_{L^2(\mathbb{T})} &= \langle Dv, \omega\rangle_{L^2(\mathbb{T})} \\ &= \langle (2yv + Lv)_x - 2y_x v, \omega\rangle_{L^2(\mathbb{T})} \\ &= \langle 2yv_x + Lv_x, \omega\rangle_{L^2(\mathbb{T})} \\ &= \langle 2yv_x, \omega\rangle_{L^2(\mathbb{T})} + \langle Lv_x, \omega\rangle_{L^2(\mathbb{T})}.\end{aligned}$$

We are free to split up the inner product this way since both $2yv_x$ and Lv_x are in $L^2(\mathbb{T})$. For the first term we have $\langle 2yv_x, \omega\rangle_{L^2(\mathbb{T})} = \langle v_x, 2y\omega\rangle_{L^2(\mathbb{T})} = 2\int_{-\pi}^{\pi} yv_x\omega\, dx$ by Parseval's identity and y and ω real valued. In the sense of distributions we have

$$\langle 2y\omega, v_x\rangle = -\langle (2y\omega)_x, v\rangle = 2\int_{-\pi}^{\pi} y\omega v_x\, dx$$

where $(2y\omega)_x \in H^{-1}(\mathbb{T})$ denotes the distributional derivative of the regular periodic distribution $2y\omega \in L^2(\mathbb{T})$. By $L\partial_x$ being skew-symmetric on $L^2(\mathbb{T})$ we have for the second term

$$\langle Lv_x, \omega\rangle_{L^2(\mathbb{T})} = -\langle v, L\omega_x\rangle_{L^2(\mathbb{T})}.$$

Note that $D^*\omega$ is a regular periodic distribution since $D^*\omega \in L^2(\mathbb{T})$ by definition, and

$$\langle D^*\omega, v\rangle = \int_{-\pi}^{\pi} (D^*\omega)(x)v(x)\, dx = \langle v, D^*\omega\rangle_{L^2(\mathbb{T})}$$

by $D^*\omega$ real-valued. Similarly

$$\langle L\omega_x, v\rangle = \int_{-\pi}^{\pi} (L\omega_x)(x)v(x)\, dx = \langle Lv_x, \omega\rangle_{L^2(\mathbb{T})}.$$

In total we then have

$$\langle D^*\omega, v\rangle = -\langle (2y\omega + L\omega)_x, v\rangle = \langle D_0\omega, v\rangle,$$

i.e. $D^*\omega = D_0\omega$ as periodic distributions. Together with the fact that $D^*\omega \in L^2(\mathbb{T})$ by definition, this means that $D_0\omega = D^*\omega$ as elements of $L^2(\mathbb{T})$ since regular distributions are determined by their generating functions up to pointwise almost everywhere equivalence. Therefore $\omega \in \operatorname{dom}(D_0)$ and so $D^* \subseteq D_0$.

Assume now that $v \in \operatorname{dom}(D_0) = \operatorname{dom}(D)$. Note that for any $u \in \operatorname{dom}(D)$, we can always find a sequence $\{u_n\}_n \in C^\infty(-\pi,\pi)$ such that Lemma 6.13 holds. Therefore we have that

$$\langle Du, v\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \langle Du_n, v\rangle_{L^2} = \lim_{n\to\infty} \langle 2y(u_n)_x, v\rangle_{L^2(\mathbb{T})} + \lim_{n\to\infty} \langle L(u_n)_x, v\rangle_{L^2(\mathbb{T})}.$$

For the first term we have

$$\langle 2y(u_n)_x, v\rangle_{L^2(\mathbb{T})} = \langle (u_n)_x, 2yv\rangle_{L^2(\mathbb{T})} = -\langle u_n, (2yv)_x\rangle_{L^2(\mathbb{T})},$$

where $(2yv)_x \in L^2(\mathbb{T})$ since $v \in \operatorname{dom}(D)$. For the second term we have as before

$$\langle L(u_n)_x, v\rangle_{L^2(\mathbb{T})} = -\langle u_n, Lv_x\rangle_{L^2(\mathbb{T})}$$

by $L\partial_x$ being skew-symmetric on $L^2(\mathbb{T})$. Thus in total

$$\langle Du, v\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \langle u_n, -(2yv + Lv)_x\rangle_{L^2(\mathbb{T})} = \langle u, D_0 v\rangle_{L^2(\mathbb{T})},$$

and it then follows from Definition 6.3 that $v \in \operatorname{dom}(D^*)$ and so $D_0 \subseteq D^*$.

□

Now that we know $D_0 = D^*$, it follows that D and D_0 are closed:

Lemma 6.16. *The operators D and D_0 are closed.*

Proof. The proof is exactly the same as the proof of Lemma 6.7. □

By Lemma 6.14 and Lemma 6.15, both D and D^* are quasi-accretive. A classical argument (cf. [34, Corollary 4.4 p. 15]) then gives the following:

Lemma 6.17. *For the closed linear operator D, densely defined on the Banach space X, with both D and its adjoint D^* quasi-accretive, there exists a scalar $\alpha \in \mathbb{R}$ such that the operator $-(D+\alpha)$ is the infinitesimal generator of a C_0-semigroup of contractions on X, i.e. D is a quasi-m-accretive operator.*

This means that $A(y)$ satisfies condition (i) of Theorem 6.3, since for $y, z, w \in H^s(\mathbb{T})$, $s > 3/2$, we have

$$\begin{aligned}\|(A(y) - A(z))w\|_{L^2(\mathbb{T})} &= 2\|(y-z)w_x\|_{L^2(\mathbb{T})} \\ &\le 2\|y-z\|_{L^2(\mathbb{T})}\|w_x\|_\infty \\ &\le 2C_s\|y-z\|_{L^2(\mathbb{T})}\|w_x\|_{H^{s-1}(\mathbb{T})} \\ &\le 2C_s\|y-z\|_{L^2(\mathbb{T})}\|w\|_{H^s(\mathbb{T})},\end{aligned}$$

where the second inequality is due to the periodic Sobolev embedding theorem (see Theorem 4.16). Also, $A(y)$ is in $\mathcal{B}(H^s(\mathbb{T}), L^2(\mathbb{T}))$ with $s > 3/2$, as is clear by the rough estimate

$$\|A(y)v\|_{L^2(\mathbb{T})} = \|2yv_x + Lv_x\|_{L^2(\mathbb{T})} \le (2\|y\|_\infty + 1)\|v\|_{H^s(\mathbb{T})}.$$

We now we turn our attention to condition (ii). Let

$$B(y) := QA(y)Q^{-1} - A(y) = [Q, A(y)]Q^{-1},$$

where $Q := \Lambda^s = (1-\partial_x^2)^{s/2}$ is an isomorphism from $H^s(\mathbb{T})$ to $L^2(\mathbb{T})$. We then have the following lemma.

Lemma 6.18. *For $y \in Y$, the operator $B(y)$ satisfies condition (ii) of Theorem 6.3.*

Proof. Note that

$$\begin{aligned}[Q, A(y)] &= [\Lambda^s, A(y)] \\ &= 2[\Lambda^s, y]\partial_x + [\Lambda^s, L]\partial_x \\ &= 2[\Lambda^s, y]\partial_x,\end{aligned} \tag{6.23}$$

where we have used the commutation properties $[\Lambda^s, \partial_x] = 0$ and $[\Lambda^s, L] = 0$.

In order to prove uniform boundedness of $B(y)$ for y in a bounded subset of $H^s(\mathbb{T})$, we assume without loss of generality that $y \in \mathcal{W} \subseteq H^s(\mathbb{T})$, where $\mathcal{W}$ is an open ball in $H^s(\mathbb{T})$ with radius $R > 0$. By classical estimates for (6.23) (cf. [23, Lemma A.2], the proof may be adapted to the periodic case using appropriate theory from Section 4), we get (for $s > 3/2$)

$$\|[\Lambda^s, y]\Lambda^{1-s}\| \leq C_0 \|\partial_x y\|_{H^{s-1}(\mathbb{T})} \leq C_0 \|y\|_{H^s(\mathbb{T})} \leq C_0 R =: \alpha_0(R),$$

where C_0 only depends on s and $\|\cdot\|$ denotes the operator norm on $L^2(\mathbb{T})$. Then for any $z \in L^2(\mathbb{T})$, we have

$$\begin{aligned}\|B(y)z\|_{L^2(\mathbb{T})} &= 2\|[\Lambda^s, y]\Lambda^{1-s}\Lambda^{s-1}\partial_x\Lambda^{-s} z\|_{L^2(\mathbb{T})}\\ &\leq 2\|[\Lambda^s, y]\Lambda^{1-s}\|\|\Lambda^{s-1}\partial_x\Lambda^{-s} z\|_{L^2(\mathbb{T})}\\ &\leq 2\alpha_0(R)\|\partial_x\Lambda^{-1} z\|_{L^2(\mathbb{T})}\\ &\leq 2\alpha_0(R)\|z\|_{L^2(\mathbb{T})},\end{aligned}$$

where the last step is due to the fact that $\|\partial_x\Lambda^{-1}z\|_{L^2(\mathbb{T})} \leq \|\Lambda^{-1}z\|_{H^1(\mathbb{T})} = \|z\|_{L^2(\mathbb{T})}$. Hence $B(y)$ is a bounded linear operator on $L^2(\mathbb{T})$ for $y \in Y$. Furthermore, for any $y, z \in \mathcal{W}$ and $w \in X$, we have

$$\begin{aligned}\|(B(y) - B(z))w\|_{L^2(\mathbb{T})} &= 2\|[\Lambda^s, y - z]\partial_x\Lambda^{-s} w\|_{L^2(\mathbb{T})}\\ &\leq 2\|[\Lambda^s, y - z]\Lambda^{1-s}\|\|\Lambda^{s-1}\partial_x\Lambda^{-s} w\|_{L^2(\mathbb{T})}\\ &\leq 2C_0\|y - z\|_{H^s(\mathbb{T})}\|w\|_{L^2(\mathbb{T})}.\end{aligned}$$

Thus $B(y)$ satisfies condition (ii) of Theorem 6.3. □

We are now ready to prove Theorem 6.2 for initial data $u_0 \in H^s(\mathbb{T})$, $s > 3/2$ (refer to the next section for the proof of the previous lemmas in the general case $\alpha \in (0, 1]$ and b of slow growth):

Proof of Theorem 6.2. By Lemmata 6.13 through 6.18 we can apply Theorem 6.3 to find a solution u as described in Theorem 6.2, although in the solution class $C^0([0, T), H^s(\mathbb{T})) \cap C^1([0, T), L^2(\mathbb{T}))$. However, in view of that $H^{s-1}(\mathbb{T})$ is an algebra with respect to pointwise multiplication for $s > 3/2$, and that $\mathcal{L}_\alpha\partial_x$ maps $H^s(\mathbb{T})$ continuously into $H^{s-1+\alpha+N_b}(\mathbb{T})$, one sees that for the equation (6.5),

$$u_t = -2uu_x - \mathcal{L}_\alpha u_x \in H^q(\mathbb{T}),$$

where $q := \min\{s - 1, s - 1 + \alpha - N_b\} \geq 0$. Hence $u \in C^1([0, T), H^{s-1}(\mathbb{T}))$

Also, since the data-to-solution map $u_0 \mapsto u$ is continuous from $H^s(\mathbb{T})$ to $C^0([0, T), H^s(\mathbb{T}))$, a similar argument can be used to conclude that the data-to-solution map is continuous from $H^s(\mathbb{T})$ to $C^1([0, T), H^q(\mathbb{T}))$. □

6.4 The periodic case for $\alpha \in (0,1]$ and b of slow growth

We can perform a similar analysis when $\alpha \in (0,1]$ and b is of slow growth, in fact the proof of well-posedness proceeds in the same steps as in the case $\alpha = 1$ and b bounded, but some of the arguments in Lemmata 6.13 through 6.15 must be modified. The main difference is that the map $\mathcal{L}_\alpha \partial_x$ is not in general an isometry on $L^2(\mathbb{T})$ for $0 < \alpha \leq 1$. However, it is a continuous map from $H^{1-\alpha+N_b}(\mathbb{T})$ to $L^2(\mathbb{T})$. This implies that D and D_0, defined as per (6.20) and (6.21) respectively, are still densely defined operators on $L^2(\mathbb{T})$, since $H^{1+N_b}(\mathbb{T})$ is a subset of $\mathrm{dom}(D)$. Also, both operators are closed with the same domains as before, which again will follow from the fact that $D^* = D_0$ and $D = -D_0 + 2y_x v$.

As usual we assume $p = 2$ for simplicity, and set $L = \mathcal{L}_\alpha$. We proceed stepwise, starting with an approximation lemma:

Lemma 6.19. *Given $v \in \mathrm{dom}(D)$, there exists a sequence $\{v_n\}_n \subseteq C^\infty(-\pi,\pi)$ such that*

$$v_n \to v \quad \textit{and} \quad (2yv_n + Lv_n)_x \to (2yv + Lv)_x$$

in $L^2(\mathbb{T})$ as $n \to \infty$.

Proof. In view of the proof of Lemma 6.13, we only need to show that the Fourier transform of the terms $2y(v_n)_x$, $L(v_n)_x$, Lv_x and $(2yv)_x$ is well-defined, and the result will follow by the argument in that proof.

This claim holds if we can establish that each term is at least in $\mathcal{P}'$. Recall that the range of the Fourier transform on $\mathcal{P}'$ is $\mathcal{S}'(\mathbb{Z})$, the space of slowly growing sequences, and in fact the Fourier transform is a bijective map $\mathcal{P}' \to \mathcal{S}'(\mathbb{Z})$ (see Theorem 4.8). Also recall that by definition a complex sequence $\{\alpha_k\}_{k\in\mathbb{Z}}$ is of slow growth if there exists $N > 0$ and $C > 0$ such that

$$|\alpha_k| \leq C|k|^N \quad \forall k \in \mathbb{Z} \setminus \{0\}.$$

The term $2y(v_n)_x$ is the product of the smooth function $(v_n)_x$ and $y \in BC^1(\mathbb{T})$, thus it is locally integrable and generates a periodic distribution (cf. Example 4.1). The term $L(v_n)_x$ is defined in terms of its Fourier coefficients by $\mathcal{F}(L(v_n)_x)(k) = ik\,b(k)|k|^{-\alpha}\hat{v}_n(k)$. Since $v_n \in L^2(\mathbb{T}) \subseteq \mathcal{P}'$, we know that there is an $M > 0$ and a $C > 0$ such that $|\hat{v}_n(k)| \leq C|k|^M$ for all $k \in \mathbb{Z} \setminus \{0\}$. Thus

$$|\mathcal{F}(L(v_n)_x)(k)| \leq C|k|^{1-\alpha+N_b+M},$$

in other words the Fourier coefficients of $L(v_n)_x$ form a sequence in $\mathcal{S}'(\mathbb{Z})$, which implies $L(v_n)_x \in \mathcal{P}'$ and thus $L(v_n)_x$ has a well-defined Fourier transform. Similarly we deduce that $Lv_x \in \mathcal{P}'$. Finally we have $(2yv)_x \in H^{-1}(\mathbb{T})$.

□

The proof of the accretiveness of D and D_0 is also slightly different:

Lemma 6.20. *The operators D and D_0 are both quasi-accretive in L^2.*

Proof. In view of the proof of Lemma 6.14, we only need to show that

$$\langle L\partial_x v_n, v_n\rangle_{L^2(\mathbb{T})} = -\langle v_n, L\partial_x v_n\rangle_{L^2(\mathbb{T})}$$

holds for $v_n \in C^\infty(-\pi,\pi) \cap \operatorname{dom}(D)$. Since $v_n \in C^\infty(-\pi,\pi)$, we have $v_n^{(k)} \in L^2(\mathbb{T})$ for all $k \in \mathbb{N}_0$, therefore since $L\partial_x$ maps $H^{1-\alpha+N_b}(\mathbb{T})$ to $L^2(\mathbb{T})$, we have

$$L(v_n)_x \in L^2(\mathbb{T}). \tag{6.24}$$

□

Just as for the Cauchy problem on the real line, the proof that $D^* = D_0$ is more technical in the general case $\alpha \in (0,1]$ and b of slow growth:

Lemma 6.21. *The adjoint of D in $L^2(\mathbb{T})$ is D_0.*

Proof. We first prove the inclusion $D^* \subseteq D_0$. Suppose $v \in \mathcal{P} \subseteq \operatorname{dom}(D)$ and $\omega \in \operatorname{dom}(D^*)$. By definition we have

$$\langle v, D^*\omega\rangle_{L^2(\mathbb{T})} = \langle Dv, \omega\rangle_{L^2(\mathbb{T})} = \langle 2yv_x, \omega\rangle_{L^2(\mathbb{T})} + \langle Lv_x, \omega\rangle_{L^2(\mathbb{T})}.$$

Note that we are free to split up the inner product this way since both $2yv_x$ and Lv_x are in $L^2(\mathbb{T})$ for $v \in \mathcal{P}$.

Using the same argument as in the proof of Lemma 6.15, we get that

$$\langle 2yv_x, \omega\rangle_{L^2(\mathbb{T})} = -\langle (2y\omega)_x, v\rangle.$$

We need to employ different reasoning in treating the term Lv_x, since $L\partial_x$ is no longer a bounded operator on $L^2(\mathbb{T})$. Ideally we would like to write simply $\langle Lv_x, \omega\rangle_{L^2(\mathbb{T})} = -\langle v, L\omega_x\rangle_{L^2(\mathbb{T})}$ using Parseval's identity, however $L\omega_x$ is not necessarily an element of $L^2(\mathbb{T})$. Nonetheless, we have by definition

$$\begin{aligned}
\langle Lv_x, \omega\rangle_{L^2(\mathbb{T})} &= 2\pi\langle \widehat{Lv_x}, \hat{\omega}\rangle_{l^2(\mathbb{Z})} = 2\pi\sum_{k\in\mathbb{Z}} ik\, b(k)|k|^{-\alpha}\hat{v}(k)\overline{\hat{\omega}(k)} \\
&= -2\pi\sum_{k\in\mathbb{Z}} \hat{v}(k)\overline{ik\, b(k)|k|^{-\alpha}\hat{\omega}(k)} = -2\pi\overline{\sum_{k\in\mathbb{Z}} \overline{\hat{v}(k)}ik\, b(k)|k|^{-\alpha}\hat{\omega}(k)} \\
&= -2\pi\overline{\sum_{k\in\mathbb{Z}} \hat{v}(-k)ik\, b(k)|k|^{-\alpha}\hat{\omega}(k)}
\end{aligned}$$

where $\overline{\hat{v}(k)} = \hat{v}(-k)$ holds since v is assumed to be real valued.

We claim $L\omega_x \in \mathcal{P}'$. By Theorem 4.8 this claim holds if we can show $\widehat{L\omega_x} \in \mathcal{S}'(\mathbb{Z})$. By definition $\widehat{L\omega_x}(k) = ik\,b(k)|k|^{-\alpha}\hat{\omega}(k)$. Since $\omega \in L^2(\mathbb{T})$, ω induces a periodic distribution and thus its Fourier coefficients form a slowly increasing sequence, i.e. there is a $C > 0$ and an $M > 0$ such that $|\hat{\omega}(k)| \le C|k|^M$ for all $k \in \mathbb{Z} \setminus \{0\}$. Hence $|\widehat{L\omega_x}(k)| \le C'|k|^{M+1-\alpha+N_b}$, so we indeed have $L\omega_x \in \mathcal{P}'$. Knowing that $L\omega_x$ is a well-defined periodic distribution, we may by Corollary 4.6.1 write

$$\langle L\omega_x, \varphi\rangle = 2\pi \sum_{k\in\mathbb{Z}} \widehat{L\omega_x}(k)\hat{\varphi}(-k) = 2\pi \sum_{k\in\mathbb{Z}} ik\,b(k)|k|^{-\alpha}\hat{\omega}(k)\hat{\varphi}(-k),$$

for $\varphi \in \mathcal{P}$. Hence by Lv_x and ω real-valued,

$$\langle Lv_x, \omega\rangle_{L^2(\mathbb{T})} = \overline{\langle Lv_x, \omega\rangle}_{L^2(\mathbb{T})} = -\langle L\omega_x, v\rangle = -2\pi \sum_{k\in\mathbb{Z}} ik\,b(k)|k|^{-\alpha}\hat{\omega}(k)\hat{v}(-k)$$

Thus we have

$$\begin{aligned}\langle D^*\omega, v\rangle &= \int_{-\pi}^{\pi} (D^*\omega)(x)v(x)\,dx = \langle v, D^*\omega\rangle_{L^2(\mathbb{T})}\\ &= \langle Dv, \omega\rangle_{L^2(\mathbb{T})} = \langle 2yv_x + Lv_x, \omega\rangle_{L^2(\mathbb{T})} = \langle D_0\omega, v\rangle,\end{aligned}$$

in other words $D^*\omega \in L^2(\mathbb{T})$ and $D_0\omega$ are equal in the sense of periodic distributions, hence we must have $D_0\omega \in L^2(\mathbb{T})$ as well. This shows that $\omega \in \operatorname{dom}(D_0)$ and therefore $D^* \subseteq D_0$.

Next we prove the inclusion $D_0 \subseteq D^*$. Suppose $v \in \operatorname{dom}(D_0) = \operatorname{dom}(D)$ and $u \in \operatorname{dom}(D)$. For such u and v we can find u_n and v_n in $C^\infty(-\pi,\pi)$ such that Lemma 6.19 holds. Thus

$$\begin{aligned}\langle Du, v\rangle_{L^2(\mathbb{T})} &= \lim_{n\to\infty}\langle Du_n, v_n\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty}\langle 2y(u_n)_x + L(u_n)_x, v_n\rangle_{L^2(\mathbb{T})}\\ &= \lim_{n\to\infty}\langle 2y(u_n)_x, v_n\rangle_{L^2(\mathbb{T})} + \lim_{n\to\infty}\langle L(u_n)_x, v_n\rangle_{L^2(\mathbb{T})}.\end{aligned}$$

We are allowed to split up the inner product this way since we have $L(u_n)_x \in L^2(\mathbb{T})$ by the argument leading up to (6.24) above, and $2y(u_n)_x$ is clearly locally square integrable since it is continuous.

For the first term we have

$$\begin{aligned}\langle 2y(u_n)_x, v_n\rangle_{L^2(\mathbb{T})} &= \langle (u_n)_x, 2yv_n\rangle_{L^2(\mathbb{T})} = 2\pi\langle ik\,b(k)\hat{u}_n(k), \widehat{2yv_n}(k)\rangle_{l^2(\mathbb{Z})}\\ &= -2\pi\langle \hat{u}_n(k), ik\,b(k)\widehat{2yv_n}(k)\rangle_{l^2(\mathbb{Z})} = -\langle u_n, (2yv_n)_x\rangle_{L^2(\mathbb{T})},\end{aligned}$$

by Parseval's identity, where $(2yv_n)_x \in L^2(\mathbb{T})$ because $2yv_n$ is continuously differentiable.

For the second term we have by Parseval's identity

$$\langle L(u_n)_x, v_x\rangle_{L^2(\mathbb{T})} = -\langle u_n, L(v_n)_x\rangle_{L^2(\mathbb{T})}.$$

In total then

$$\langle Du, v\rangle_{L^2(\mathbb{T})} = \lim_{n\to\infty} \langle u_n, -(2yv_n + Lv_n)_x\rangle_{L^2(\mathbb{T})} = \langle u, D_0 v\rangle_{L^2(\mathbb{T})},$$

from which it follows by Definition 6.3 that $v \in \mathrm{dom}(D^*)$. We therefore get the other inclusion, $D_0 \subseteq D^*$. □

The rest of the well-posedness proof proceeds exactly as in the case $\alpha = 1$ and b bounded.

6.5 Some notes on global well-posedness

Having established local well-posedness in H^s for $s > 3/2$, the natural next step is to try and extend to a global result. The idea is simple: let the given initial data evolve according to the equation until time a stopping time $t_s < T$ is reached, then feed the solution $u(t_s, \cdot)$ into the equation again as initial data. A new solution will then exist for a new finite time depending on the size of $u(t_s, \cdot)$. However, it is possible that the norm of the solution blows up in finite time and that the procedure therefore fails. In that case we have a negative result for global well-posedness, and we say that the global Cauchy problem is *ill-posed*. One way to guarantee that this doesn't happen is to prove that the equation satisfies some growth condition on the norm of the solution. For instance, in [22] Kato proves a global well-posedness result in H^s with $s > 3/2$ for the KdV equation using "condition (G)". Alternatively, one may settle the question of global well-posedness by proving finite time blow-up of the solution.

We shall gather previous results about some equations of the general form (6.1), and see if these results point towards global well-posedness or ill-posedness. We will restrict us to the case $p = 2$, since the quadratic nonlinearity is perhaps the most physically relevant (it appears in equations from hydrodynamics, for instance) and thus seems to be more well-studied. In our analysis we recall that p could be any integer greater than 2 and we still had the same local result. For global well-posedness the order of the nonlinearity probably has more of an impact, since the exact balancing of the nonlinear and dispersive effects in an equation can be key to the equation admitting global well-posedness results (we discussed this in the introduction in reference to the KdV equation).

In the case where $b \equiv 1$, we saw that composition of operators $\mathcal{L}_1\partial_x$ corresponded to multiplication by $i\,\mathrm{sgn}$ on the Fourier side. If one defines the Fourier transform of the function $1/x$ using the Cauchy principal value, then its Fourier transform is $-i\sqrt{\pi/2}\,\mathrm{sgn}$, that is

$$\mathcal{F}(\frac{1}{x})(\xi) := \frac{1}{\sqrt{2\pi}}\,p.v.\int_{\mathbb{R}} \frac{e^{-ix\xi}}{x}\,dx = \frac{1}{\sqrt{2\pi}}\,p.v.\int_{\mathbb{R}} \frac{\cos(x\xi) - i\sin(x\xi)}{x}\,dx$$
$$= -\frac{i}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{\sin(x\xi)}{x}\,dx = -i\sqrt{\frac{\pi}{2}}\,\mathrm{sgn}(\xi).$$

In the case $b \equiv 1$ the operator $\mathcal{L}_1\partial_x$ therefore corresponds in some sense to convolution with $1/x$. Precisely, it is the convolution with the tempered distribution $p.v.\,\frac{1}{x}$ [33, Chapter 3]. Hence the problem (6.1) with $p = 2$, $b \equiv 1$ and $\alpha = 1$ corresponds to the Cauchy problem for the *Burgers-Hilbert equation*, typically defined via the *Hilbert transform* as

$$u_t + uu_x - \mathcal{H}u = 0, \tag{6.25}$$

where $\mathcal{H}u(x) := \frac{1}{\pi}\,p.v.\int_{\mathbb{R}} \frac{u(x-y)}{y}\,dy$ (by a change of variables similar to the ones we performed in the beginning of the section, we can scale in or out any constants in front of the terms in our equation). The Burgers-Hilbert equation is studied in [4], where global existence is established in $L^2(\mathbb{R})$. On the other hand, in [7], singularity formation in finite time in the $C^{1,\delta}$-norm for a large class of initial data is proven. Specifically, they proved that for initial data $u_0 \in L^2(\mathbb{R}) \cap C^{1,\delta}(\mathbb{R})$, $0 < \delta < 1$, for which there is a point $\beta_0 \in \mathbb{R}$ with

$$\mathcal{H}u_0(\beta_0) > 0 \quad \text{and} \quad u_0(\beta_0) \geq (32\pi\|u_0\|^2_{L^2(\mathbb{R})})^{1/3}, \tag{6.26}$$

there exists a finite time T such that

$$\lim_{t\to T} \|u(t,\cdot)\|_{C^{1,\delta}(\mathbb{R})} = \infty.$$

Here $C^{1,\delta}(\mathbb{R})$ is the *Hölder space* of continuously differentiable functions f with *Hölder continuous* derivative with exponent δ, that is

$$|f'(x) - f'(y)| \leq c|x-y|^\delta \quad \forall x, y \in \mathbb{R}$$

for some constant $c > 0$, and the $C^{k,\delta}(\mathbb{R})$-norm for $k \in \mathbb{N}_0$ is

$$\|f\|_{C^{k,\delta}(\mathbb{R})} = \|f\|_{C^k(\mathbb{R})} + \sup_{x\neq y\in\mathbb{R}} \frac{|f^{(k)}(x) - f^{(k)}(y)|}{|x-y|^\delta}.$$

This result actually implies global ill-posedness in $H^s(\mathbb{R})$, $s > 3/2$ (see Theorem 6.22 below), since one can easily find initial data in these spaces which satisfy the conditions (6.26). For instance one can set $\beta_0 = 0$ and consider smooth initial data of compact support which takes the form $A\cos(bx)$ around $x = 0$. The data could keep this form for half a period on the left side of $x = 0$ and one period on the right side before being smoothed out to 0, in order to ensure that $\mathcal{H}u_0(0) > 0$. In addition, the constants A and b should be scaled appropriately so that $u_0(0) \geq (32\pi\|u_0\|^2_{L^2(\mathbb{R})})^{1/3}$ holds. In practice this means that b should be large so that the period of the cosine-function is small, and the amplitude A should be large so that the value of $u_0(0)$ is large. In this way one sees that the condition $u_0(\beta_0) \geq (32\pi\|u_0\|^2_{L^2(\mathbb{R})})^{1/3}$ represents a steepness requirement on the graph of u_0. If one considers (6.25) as Burgers' equation with the source term $\mathcal{H}u$, the first condition in (6.26) represents positivity of the forcing term $\mathcal{H}u_0$ at β_0.

Theorem 6.22. *Let $s > 3/2$. The Cauchy problem* (6.1) *with $p = 2$, $b \equiv 1$ and $\alpha = 1$ is globally ill-posed in $H^s(\mathbb{R})$. Specifically, for initial data $u_0 \in H^s(\mathbb{R}) \cap C^{1,\delta}(\mathbb{R})$, $0 < \delta < 1$, satisfying condition* (6.26)*, there exists a finite time T such that*

$$\lim_{t\to T} \|u(t,\cdot)\|_{H^s(\mathbb{R})} = \infty.$$

Proof. The Hölder spaces $C^{k,\delta}(\mathbb{R}^d)$ for $0 < \delta < 1$ interpolate between the spaces $C^k(\mathbb{R}^d)$ and $C^{k+1}(\mathbb{R}^d)$, and $H^s(\mathbb{R}^d)$ embeds continuously in the Hölder space $C^{k,\delta}(\mathbb{R}^d)$ if $s \geq d/2 + k + \delta$ for some $k \in \mathbb{N}$ and $\delta \in (0,1)$ [2, Theorem 1.66]. In other words, we have

$$\|f\|_{C^{1,\delta}(\mathbb{R})} \leq C\|f\|_{H^s(\mathbb{R})}$$

for $s \geq 3/2 + \delta$ with $0 < \delta < 1$. The blow-up in the $C^{1,\delta}(\mathbb{R})$-norm for $0 < \delta < 1$ therefore implies the ill-posedness result in $H^s(\mathbb{R})$ for $s > 3/2$. □

Remark 6.12. Since the ill-posedness here follows from blow-up of the $C^{1,\delta}(\mathbb{R})$-norm, this result doesn't necessarily imply that the solution wave breaks (in which case the supremum of the gradient of the solution should blow up, while the supremum of the solution itself should remain finite).

The article [7] continues with a proof (by a different method) of the existence of finite time singularities in the $C^{1,\delta}$-norm for some (unspecified) $L^2(\mathbb{R})\cap C^{1,\delta}(\mathbb{R})$-initial data for the equation

$$u_t + uu_x - D^\beta \mathcal{H}u = 0 \tag{6.27}$$

with $0 < \beta < 1$, where D^β is the fractional Laplacian $(-\Delta)^{\frac{\beta}{2}}$, defined as a Fourier multiplier operator by

$$\mathcal{F}(D^\beta f)(\xi) = |\xi|^\beta \hat{f}(\xi).$$

On the Fourier side the term $D^\beta\mathcal{H}u$ corresponds to $-i\,\mathrm{sgn}(\xi)|\xi|^\beta\hat{u}(\xi)$, or

$$\mathcal{F}(-D^\beta\mathcal{H}u)(\xi) = i\,\mathrm{sgn}(\xi)|\xi|^\beta\hat{u}(\xi) = i\xi|\xi|^{\beta-1}\hat{u}(\xi).$$

From this we see that the equation (6.27) with $0 < \beta < 1$ corresponds to (6.1) with $\alpha = 1 - \beta$ and $0 < \alpha < 1$ (in the case $p = 2$ and $b \equiv 1$). Again this hints at a possible global ill-posedness result in $H^s(\mathbb{R})$, $s > 3/2$, for some of the Cauchy problems given by (6.1). However, in this general case no exact conditions on the initial data which guarantee blow-up are specified, thus we don't know whether or not there exists $H^s(\mathbb{R})$-initial data, $s > 3/2$, for which the blow-up in the $C^{1,\delta}$-norm is guaranteed to occur.

In [27] it is observed that the proof in [7] extends to when the operator $\mathcal{L}_\alpha$ on the Fourier side is not a pure power of $1/|\xi|$, thus an eventual ill-posed result may be extended to equations with more general dispersive operators (e.g. the Whitham equation). In [25], it is conjectured based on numerical experiments that given smooth L^2-initial data,

- for $0 < \alpha < 1$ the solutions to (6.1) with $p = 2$ and $b \equiv 1$ and to the Whitham equation stay smooth for all time and will be radiated away, if the initial data is of sufficiently small mass;
- for $\alpha = -1/2$ the solutions to (6.1) with $p = 2$ and $b \equiv 1$ and to the Whitham equation will develop a cusp of the form $|x - x^*|^{\frac{1}{2}}$ within a finite amount of time, if the initial data is positive and of sufficiently large norm mass.

The numerical solution's tendency to remain smooth only for initial data of sufficiently small mass perhaps reflects the first condition in (6.26).

The paper [28] collects some results about equivalent problems to (6.1) in the case $p = 2$ and $b \equiv 1$ for various values of α. In the paper it is claimed that the equations (6.1) with $p = 2$, $b \equiv 1$ and any α, that is

$$u_t + uu_x + D^{-\alpha}u_x = 0$$

where $D^{-\alpha}$ is again the fractional Laplacian $(-\Delta)^{-\frac{\alpha}{2}}$ of order $-\alpha$ (in our case $0 < \alpha \leq 1$), allows for two conserved quantities, namely the L^2-norm of the solution, and the quantity (the Hamiltonian)

$$H(u) = \int_{\mathbb{R}} \frac{1}{3}u^3(x,t) + \frac{1}{2}|D^{-\frac{\alpha}{2}}u(x,t)|^2\,dx.$$

If one could press local well-posedness down to L^2, then one would have global well-posedness due to the the first conserved quantity. However, this is clearly not easily done using the method in this thesis. The quantity $H(u)$ is not well-defined for general $u \in H^s$, $s > 3/2$ (when $0 < \alpha \leq 1$). If we instead of setting $b \equiv 1$, let b be such that $\frac{b(\xi)}{|\xi|^\alpha} = |\xi|^\beta$, where possible values for β are $-1 \leq \beta \leq s - 1$, s being the order of the Sobolev space we have our local well-posedness result in (i.e. $s > 3/2$), then the problem (6.1) with $p = 2$ corresponds again to the Cauchy problem for the equation

$$u_t + uu_x + D^\beta u_x = 0, \tag{6.28}$$

now with $-1 \leq \beta \leq s - 1$. Again we have conservation of the L^2-norm of the solution and of the Hamiltonian $H(u) = \int_{\mathbb{R}} \frac{1}{3}u^3(x,t) + \frac{1}{2}|D^{\frac{\beta}{2}}u(x,t)|^2\,dx$. In this case the Hamiltonian is in fact well-defined as long as $\beta \geq 1/3$, by the embedding $H^{\frac{1}{6}}(\mathbb{R}) \hookrightarrow L^3(\mathbb{R})$ [28]. A possible step on the way to proving global well-posedness for the equation (6.28) with $\beta \geq 1/3$ might therefore be to show that also the L^3-norm of the solution is also conserved, which would at the same time provide a conservation law for the $H^{\frac{\beta}{2}}$-norm. We do not know if this is possible, however.

When $0 < \beta < 1$, the equation (6.28) corresponds to the fractal Burgers' equation that we mentioned in the introduction of this section. In [28] some well-posedness results in various Sobolev spaces are given for the fractal Burgers' equation.

References

[1] M.J. Ablowitz, *Nonlinear Dispersive Waves: Asymptotic Analysis and Solitons*, Cambridge University Press (2011)

[2] H. Bahouri, J.-Y. Chemin & R. Danchin, *Fourier Analysis and Nonlinear Partial Differential Equations*, Springer (2011)

[3] V.I. Bogachev, *Measure Theory, Volume I*, Springer (2007)

[4] A. Bressan & K.T. Nguyen, *Global existence of weak solutions for the Burgers-Hilbert equation*, SIAM Journal on Mathematical Analysis Vol. 46, No. 4, pp 2884-2904 (2014)

[5] J.V. Boussinesq, *Essai sur la théorie des eaux courantes*, Mémoires présentés par divers Savants à l'Académie Royale des Sciences de l'Institut de France, vol. XXIII (1877)

[6] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Mathematica, Vol. 116, Issue 1, pp 135-157 (1966)

[7] A. Castro, D. Córdoba & F. Gancedo, *Singularity Formation in a Surface Wave Model*, Nonlinearity 23 (2010)

[8] J. Colliander, et al., *Sharp global well-posedness for KdV and modified KdV on* $\mathbb{R}$ *and* $\mathbb{T}$, Journal of the American Mathematical Society, Vol. 16, No. 3, pp 705-749 (2003)

[9] A. Constantin & J. Escher, *Wave breaking for nonlinear nonlocal shallow water equations*, Acta Mathematica, Volume 181, Issue 2, pp 229-243 (1998)

[10] A. Constantin & J. Escher, *Well-posedness, global existence, and blowup phenomena for a periodic quasi-linear hyperbolic equation*, Commun. Pure Appl. Math., 51 (1998)

[11] F. Demengel & G. Demengel, *Functional Spaces for the Theory of Elliptic Partial Differential Equations*, Springer (2012)

[12] J. Duoandikoetxea, *Fourier Analysis*, American Mathematical Society (2001)

[13] M. Ehrnström, *Waves and Solitons*, Lecture notes (2010)

[14] M. Ehrnström, *Linear Methods*, Lecture notes (2014)

[15] M. Ehrnström, J. Escher & L. Pei, *A note on the local well-posedness for the Whitham equation*, Springer Proceedings in Mathematics & Statistics: Elliptic and Parabolic Equations, Vol. 119, pp 63-75 (2015)

[16] J. Hadamard, *Sur les problèmes aux dérivées partielles et leur signification physique*, Princeton University Bulletin (1902)

[17] D. Haroske & H. Triebel, *Distributions, Sobolev Spaces, Elliptic Equations*, European Mathematical Society (2008)

[18] M. Hazewinkel, ed., *Fréchet space*, Encyclopedia of Mathematics, from `http://www.encyclopediaofmath.org/index.php/Fr%C3%A9chet_space`

[19] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Springer (1998)

[20] A.D. Ionescu & C.E. Kenig, *Global well-posedness of the Benjamin-Ono equation in low regularity spaces*, Journal of the American Mathematical Society, Vol. 20, No. 3, pp 753-798 (2007)

[21] R. Iorio & V.M. Iorio, *Fourier Analysis and Partial Differential Equations*, Cambridge University Press (2001)

[22] T. Kato, *On the Korteweg-de Vries equation*, Manuscripta Mathematica, Volume 28, Issue 1-3, pp 89-99 (1979)

[23] T. Kato, *Quasi-linear equations of evolution, with applications to partial differential equations*, in Spectral Theory and Differential Equations (Proc. Sympos., Dundee, 1974), Springer (1975)

[24] Y. Katznelson, *An Introduction to Harmonic Analysis*, Cambridge University Press (3rd ed., 2004)

[25] C. Klein & J.-C. Saut, *A numerical approach to blow-up issues for dispersive perturbations of Burgers' equation*, Physica D: Nonlinear Phenomena, Vol. 295-296, pp 46-65 (2015)

[26] D. Korteweg & G.d. Vries, *On the Change of Form of Long Waves Advancing in a Rectangular Canal, and on a New Type of Long Stationary Waves*, Philosophical Magazine 39 (1895)

[27] D. Lannes & J.-C. Saut, *Remarks on the full dispersion Kadomtsev-Petviashvli equaton*, Kinetic and Related Models, American Institute of Mathematical Sciences, Vol. 6, No. 4 (2013)

[28] F. Linares, D. Pilod & J.-C. Saut, *Dispersive Perturbations of Burgers and Hyperbolic Equations I: Local Theory*, SIAM Journal on Mathematical Analysis Vol. 46, No. 2, pp 1505-1537 (2014)

[29] F. Linares & G. Ponce, *Introduction to Nonlinear Dispersive Equations*, Springer (2009)

[30] T. Lindstrøm, *Mathematical Analysis*, Lecture notes (2014), from `http://uio.no/studier/emner/matnat/math/MAT2400/v14/mathanalbook.pdf`

[31] V. Mazya & T. O. Shaposhnikova, *Jacques Hadamard: A Universal Mathematician*, American Mathematical Society (1999)

[32] P.I. Naumkin & I.A. Shishmarev, *Nonlinear Nonlocal Equations in the Theory of Waves*, American Mathematical Society (1994)

[33] J.N. Pandey, *The Hilbert transform of Schwartz distributions and applications*, Wiley-Interscience (1996)

[34] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer (1983)

[35] J.I. Richards & H.K. Youn, *The Theory of Distributions: A Nontechnical Introduction*, Cambridge University Press (1990)

[36] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company (1991)

[37] J.S. Russell, *Report on Waves*, Fourteenth meeting of the British Association for the Advancement of Science (1844)

[38] M. Salo, *Fourier analysis and distribution theory*, Lecture notes (2013), from `http://www.rni.helsinki.fi/~msa/lecturenotes/FA_distributions.pdf`

[39] S. Salsa, *Partial Differential Equations in Action: From Modelling to Theory*, Springer (2009)

[40] D.C. Smith, *An introduction to distribution theory for signals analysis. Part III. Periodic tempered distributions*, Digital Signal Processing (2015)

[41] R.S. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, World Scientific (2003)

[42] G.B. Whitham, *Linear and Nonlinear Waves*, Wiley-Interscience (1999)

[43] G.B. Whitham, *Variational methods and applications to water waves*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, Vol. 299, No. 1456 (1967)

[44] N. Young, *An introduction to Hilbert space*, Cambridge University Press (1988)