NTNU
Norwegian University of
Science and Technology

On the Existence of Periodic Traveling Waves to the Whitham Equation

Fredrik Hildrum

Master of Science in Physics and Mathematics
Submission date: November 2015
Supervisor: Mats Ehrnstrøm, MATH
Co-supervisor: Mathias Nikolai Arnesen, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

We study the existence of periodic traveling waves to the Whitham equation, which is a nonlinear, nonlocal and dispersive differential equation proposed by Whitham [65, 66] as a model for surface gravity waves featuring the exact linear dispersion relation for water waves. Based on a work by Ehrnström *et al.* [21] we consider a generalized Whitham equation with power nonlinearities $n(u) = |u|^q$ or $u|u|^{q-1}$ for $q \in (1,5)$. It is found that there exist periodic traveling waves for all sufficiently large periods in any Sobolev space H^s of order $s \in \left(\frac{1}{2}, q\right)$, or $s \in \left(\frac{1}{2}, \infty\right)$ if $n(u) = u^q$ and $q = 2$, 3 or 4. The waves are shown to be of small amplitude, measured by the H^s norm, for a subset of the orders. In addition, we provide an explicit lower bound on the wave speeds.

The existence technique treats the Whitham equation as the Euler-Lagrange equation of a constrained minimization problem. As a background we perform a detailed study of Fourier series and Sobolev spaces with arbitrary periods and the calculus of variations.

Sammendrag

Vi studerer eksistensen av periodiske reisende bølger til Whitham-ligningen, som er en ikke-lineær, ikke-lokal og dispersiv differensialligning introdusert av Whitham [65, 66] som en modell for gravitasjonsdrevne overflatebølger med den eksakte lineære dispersjonsrelasjonen for vannbølger. Basert på et arbeid av Ehrnström *et al.* [21] betrakter vi en generalisert Whitham-ligning med ikke-lineæriteter på potensformene $n(u) = |u|^q$ eller $u|u|^{q-1}$ for $q \in (1,5)$. Vi finner at det eksisterer periodiske reisende bølger for alle tilstrekkelige lange perioder i Sobolevrom H^s av orden $s \in \left(\frac{1}{2}, q\right)$, eller $s \in \left(\frac{1}{2}, \infty\right)$ hvis $n(u) = u^q$ og $q = 2$, 3 eller 4. For en delmengde av ordene vises det at bølgene har liten amplitude målt med H^s-normen. I tillegg gis en eksplisitt nedre skranke for bølgefartene.

Eksistensteknikken behandler Whitham-ligningen som Euler-Lagrange-ligningen til et minimiseringsproblem med føringer, og vi gjør et detaljert bakgrunnsstudie av Fourierrekker og Sobolevrom med vilkårlige perioder samt variasjonskalkulus.

Preface

This text represents the outcome of my work in the course "TMA4900 Mathematics, Master's Thesis" and marks the end of the 5-year integrated master's programme in "Applied Physics and Mathematics" at the Norwegian University of Science and Technology.

My sincere gratitude goes to my parents for their love and support during my studies. I would also like to express thankfulness to my supervisor Prof. Mats Ehrnström for his kindness, availability and construction of an interesting problem. Moreover, I appreciate the help and proofreading of PhD. cand. Mathias Nikolai Arnesen as a co-supervisor.

The thesis is in broad terms structured as follows.

Chapter 1 introduces the Whitham equation as a model for surface water waves, with emphasis on the linear dispersion relation. We next survey research on Whitham's model and present the contribution of this thesis.

Chapter 2 refreshes some facts about operators in normed and inner product spaces, properties of C^n and L^p spaces, sequence spaces, integral calculus and convolution. As such, the audience is expected to have sufficient experience in real and functional analysis with measure theory.

Chapter 3 studies Fourier series of periodic functions, weak differentiability and periodic Sobolev spaces.

Chapter 4 briefly treats the Fourier transform on $\mathbb{R}$ and beyond and introduces Fourier multiplier operators.

Chapter 5 considers the calculus of variations, or optimization of functionals, with focus on the existence of extreme points and Lagrange multipliers in constrained problems.

Chapter 6 proves the existence of periodic traveling waves to a generalized version of the Whitham equation in periodic Sobolev spaces using the developed variational methods.

Appendix A contains an introduction to fundamental characteristics of waves.

Trondheim, November 27, 2015

Fredrik Hildrum

(Updated on March 4, 2016 with a few corrections of notation and spelling.)

CONTENTS

Custom notation

- $0 \in \mathbb{N}$, so that $\mathbb{N} = \{0\} \cup \mathbb{Z}_+$.
- $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is the scalar field and range of functions. Another option, with details left to the reader, is that $\mathbb{R}$ is the field and $\mathbb{F}$ is the range.
- $\subset$ denotes strict set inclusion, whereas $\subseteq$ includes the possibility of equality.
- If $f : X \to Y$ and $U \subseteq X$, then $f\restriction_U$ is the restriction of f to U.
- A dot $\cdot$ inside a function symbolizes the unspecified argument. An example:

$$f(\cdot + 2\pi) \qquad \text{means the function} \qquad x \mapsto f(x + 2\pi).$$

- $|\Omega|$ is the Lebesgue measure of $\Omega \subseteq \mathbb{R}$.
- $\overline{X}$ denotes the closure of a set X.
- $\operatorname{span} \mathscr{B}$ is the linear span of a set $\mathscr{B}$, that is, the set of finite linear combinations of elements of $\mathscr{B}$.
- $A \lesssim B$ is a short-hand for $A \leq cB$ whenever $c > 0$ is a constant not depending on B. This is useful in estimation. Moreover, if $A \lesssim B \lesssim A$, we write $A \asymp B$.
- Except for derivatives and indices, subscripts denote an unspecified dependence on some parameter(s). For example, $P_{q,\mu}$ means that $P = P(q, \mu)$. Similarly, we write $A \lesssim_s B$ if $A \leq c_s B$ for some $c_s > 0$, and $A \asymp_s B$ whenever $A \lesssim_s B \lesssim_s A$.

Chapter 1

Introduction

Surface water waves is a fascinating phenomenon in the exceedingly beautiful creation. Their appearance ranges from peaceful ripples and perpetual swells near a coastline to roaring ocean waves in a violent storm. Usually generated by winds and the effects of surface tension, water waves feature a rich amount of complexity.

Of particular interest is the class of traveling waves, which move progressively in one direction with fixed speed and shape; see Appendix A for an introduction to fundamental wave characteristics. The basic mechanism behind this phenomenon is a balancing of dispersion and nonlinear effects [22]. Most of the time these steady waves repeat themselves periodically, leading to periodic traveling waves. Another possible pattern is single waves which vanish at infinity. Famous for his observation in 1844 of such a solitary wave, the naval engineer Russell called it "the wave of translation" [55].

1.1 The water-waves problem and Whitham's model equation

The governing equations describing the motion of an inviscid, incompressible and irrotational fluid (water) under the influence of gravity is given by the Euler equations, supplemented by a set of boundary conditions; see Lannes [44, Chapter 1] and Socha [60]. Combined, these nonlinear equations constitute what is known as the *free surface water-waves problem*, where the surface waves arise from the *a priori* unknown free surface.

The water-waves problem is difficult to solve and many approximations and asymptotic models for surface waves have been developed. Typically one classifies the water depth into the three regimes of deep water, shallow water or arbitrary depth, and then makes assumptions on the wavelength and amplitude relative to the depth or each other. Important examples are linear (Airy) wave theory, Stokes expansions, the shallow water equations and Boussinesq approximations.

A popular model for unidirectional, small-amplitude and long gravity waves in shallow water is given by the Korteweg–de Vries (KdV) equation

$$\eta_t + c_0\,\eta_x + \frac{3c_0}{4h_0}\,(\eta^2)_x + \tfrac{1}{6}c_0 h_0^2\,\eta_{xxx} = 0, \tag{1.1}$$

where $\eta = \eta(x,t)$ describes the deflection of the water surface at point x at time t, $c_0 = \sqrt{gh_0}$ is the limiting long-wave speed, h_0 is the undisturbed water depth and g is the gravitational constant. The KdV equation is a nonlinear and dispersive partial differential equation (PDE), and by inserting a monochromatic wave, we infer that the linearized phase speed equals

$$c_{\mathrm{KdV}}(k) = c_0\left(1 - \tfrac{1}{6}(kh_0)^2\right),$$

where k is the wave number. This is a second-order approximation of the linearized phase speed

$$c_{\mathrm{Euler}}(k) = \sqrt{\frac{g}{k}\tanh(kh_0)} \tag{1.2}$$

associated with the full water-waves problem. As seen in Figure 1.1, the KdV equation does only give a faithful representation for very long waves in shallow water. The so-called BBM equation

$$\eta_t + c_0\,\eta_x + \tfrac{3c_0}{4h_0}(\eta^2)_x - \tfrac{1}{6}h_0^2\,\eta_{xxt} = 0$$

presented by Benjamin, Bona and Mahony [3] yields an improvement to (1.1), with

$$c_{\mathrm{BBM}}(k) = \frac{c_0}{1 + \frac{1}{6}(kh_0)^2}.$$

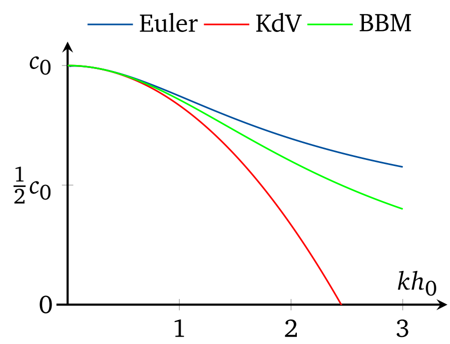

Figure 1.1: Comparison of linearized phase speeds.

In 1967 and 1974, however, Whitham [65, 66, Section 13.14] suggested to modify the dispersive term in the KdV equation so that it features (1.2) exactly. This led to the model

$$\eta_t + \tfrac{3c_0}{4h_0}(\eta^2)_x + K_{h_0} * \eta_x = 0, \tag{1.3}$$

possibly valid for shorter waves and later known as the *Whitham equation*. Here $K_{h_0} * \eta_x$ is the convolution of η_x and the integral kernel

$$K_{h_0} = \frac{1}{\sqrt{2\pi}}\mathscr{F}^{-1}(c_{\mathrm{Euler}}), \tag{1.4}$$

and $\mathscr{F}^{-1}$ denotes the (distributional) Fourier transform. Whitham's original motivation was to find a model which could capture *breaking of waves*, which means that the slope of a wave becomes unbounded in finite time while the wave itself stays bounded [22, 38]. Another interest was *wave peaking*, a phenomenon that occurs when a wave forms a cusp or sharp crest [22]. Nowadays, there is a distinction between waves with nonzero peaking angles (bounded derivative), called peakons, and cusp-like waves with zero angle (unbounded derivative). The former is seen the Stokes waves for the water-waves problem featuring a 120° peaking angle [66, Section 13.13].

The Whitham equation (1.3) is a nonlinear, nonlocal and dispersive differential equation. The *nonlocality* is of integro-differential type, described by the convolution, and there is a fundamental difference between traditional differential equations and the nonlocal variants [50]. An ordinary/partial differential equation is satisfied at a given point if the associated function values are known in an arbitrarily small neighborhood around the point. To the contrary, nonlocal equations require information about the global behavior of the function in order to validate the equation at a single point. This complicates the analysis of (1.3) compared to for example that of the KdV equation.

Due to its generic form, (1.3) is also studied for other kernels which yield nonlocal models in mathematical physics; see Naumkin and Shishmarev [49, Introduction] for illustrations.

1.2 A SURVEY OF RESEARCH ON THE WHITHAM EQUATION

The recent years have seen an increasing interest in the Whitham equation as a model for water waves and in this section we highlight many of the analytical and numerical research advancements.

Beyond introducing (1.3), Whitham [65, 66] gave formal arguments for the presence of a cusped wave, listed some properties of the kernel K_{h_0} and mentioned that the equation could be derived from a variational principle.

Two early studies of the Whitham equation include that of Gabov [29] and Zaitsev [70]. The book by Naumkin and Shishmarev [49] from 1994 is devoted to the analysis of (1.3) for a mixture of kernels, treating among others the Cauchy problem both in the periodic case and on the line, and provided an affirmative answer to the question of wave breaking in both situations. Another focus was the asymptotical behavior of solutions. As Hur [36] commented in 2015, however, the wave-breaking arguments in [49] seem to contain glitches. Nevertheless, by building on a work with Tao [38], she proved breaking for the solitary problem.

Although the breaking of waves excludes general global well-posedness of (1.3), Ehrnström *et al.* [20] in 2015 used Kato's method and established local well-posedness in Sobolev spaces of order greater than 3/2, both for solitary and periodic initial data.

In 2009 Ehrnström and Kalisch [22] proved the existence of small-amplitude periodic traveling waves via bifurcation analysis, with velocities approaching the long-wave speed c_0 as the periods tends to infinity. Whereas [22] established a local bifurcation branch, Ehrnström and Kalisch [23] demonstrated in 2013 the existence of a global branch of smooth periodic traveling waves, with comparison to (1.1). Moreover, the existence and conditional stability of solitary waves were settled in 2012 in the work by Ehrnström *et al.* [21], who considered a class of evolution equations including (1.3). It was also found that the Whitham solitary waves were approximated by scalings of the corresponding solutions to (1.1).

Until recently the singular kernel K_{h_0} was not thoroughly understood. Ehrnström and Kalisch [22, Section 2 and 4] established its integrability in certain L^p spaces and smoothness away from the origin. In a work in preparation, however, Ehrnström and Wahlén [19, 24] have obtain a closed formula both for K_{h_0} and its periodic version. This shows that the kernel is completely monotone on the interval $(0, \infty)$ and is analytic with exponential decay away from the origin. Using global bifurcation theory, they furthermore prove the existence of a highest, cusped periodic traveling wave which has exactly Hölder regularity $1/2$ at the crest, thereby resolving Whitham's conjecture.

Another work in preparation is that of Brüll *et al.* [8]. Their geometrical findings yield essentially that a solution to (1.3) is symmetric if and only if it is a periodic traveling wave. Moreover, any solitary-wave solution is exponentially decaying.

Switching to numerical investigations, note first that [22, 23] studied numerical bifurcation of periodic traveling waves and the problem of time evolution using spectral Fourier collocation methods. Both articles support the analysis of [19, 24] and indicate the presence of a cusped wave, with [22] estimating its maximum height to $0.642h_0$. Additionally, [22] found that periodic traveling waves converge to a solitary wave as the wavelength increases, which is in agreement with [21].

The studies by Borluk *et al.* [5] in 2013 and Moldabayev *et al.* [48] in 2014 then examined the validity of (1.3) as a model for surface waves by comparing it with numerical approximations to the full Euler water-waves problem. Both investigations applied spectral schemes analogously to [22, 23]. Focusing on steady waves, [5] found that for short wavelengths the Whitham solutions give a closer approximation to the Euler waves than the KdV solutions. For larger wavelengths, however, the opposite is true. Moreover, whereas (1.1) had an infinite bifurcation branch, both (1.3) and the Euler equations observed finite branches. The paper [48] studied the performance of (1.3) during time evolution for a wide parameter range of amplitudes and wavelengths. Their conclusion was that the Whitham equation gives a close representation of the Euler equations, mostly on par with or better than KdV and BBM. In addition, they "identified a scaling regime in which the Whitham equation can be derived from the Hamiltonian theory of surface water waves."

It is known that periodic wavetrains to the water-waves problem may feature a so-called Benjamin-Feir or modulational instability, leading to sidebands growth; see for instance the historical review by Zakharov and Ostrovsky [71]. Benjamin and Hasselmann [4] calculated that small-amplitude (Stokes') periodic traveling waves are unstable provided $kh_0 > 1.363$. Periodic traveling waves to the KdV and the BBM equations do not exhibit this property, but are spectrally stable [6, 7, 41]. Hur and Johnson [37], however, proved in 2015 that small-amplitude periodic traveling waves to (1.3) are modulationally unstable whenever $kh_0 > 1.146$, but spectrally stable to square integrable perturbations otherwise. Interestingly, the numerical study by Sanford *et al.* [56] from 2014 obtained modulational instability for the same range. Additionally, [56] found "that all large-amplitude solutions are unstable, while small-amplitude solutions with large enough wavelength are stable."

1.3 THE WORK AT HAND

On their path to solitary waves for a large class of nonlocal equations of Whitham type, Ehrnström *et al.* [21] first considered the periodic problem. For all sufficiently large periods P they established the existence of small-amplitude, P-periodic traveling waves in the periodic Sobolev space H^1_P (see Definition 3.32). The proof technique viewed the equations as the Euler-Lagrange equation of a constrained minimization problem using penalization and *a priori* estimates.

The original goal of this thesis was to simplify the arguments in [21] restricted to periodic traveling waves in H^1_P for the Whitham equation (1.3). This required a study of Fourier theory, Sobolev spaces and variational methods. A second task was to replace the nonlinearity $(\eta^2)_x$ in (1.3) with $(\eta^q)_x$ and determine the upper bound on $2 \leq q \in \mathbb{Z}_+$ for which the method still applies.

During the analysis, however, we found that the nonlinearity can be generalized to $(n(\eta))_x$, where $n\colon \mathbb{R} \to \mathbb{R}$ satisfies

$$n(x) = |x|^q \qquad \text{or} \qquad n(x) = x|x|^{q-1}$$

for any $q \in (1,5)$. The first case is already included in [21], but with $q \in [2,5)$. In addition, we extend the theory to the fractional-order spaces H^s_P of order $s \in \left(\frac{1}{2}, q\right)$, or even $s \in \left(\frac{1}{2}, \infty\right)$ if $n(x) = x^q$ and $q = 2$, 3 or 4. The waves are shown to be of small amplitude, measured by the H^s_P norm, for a subset of the given orders. This has been done using the natural estimate

$$\|n(u)\|_{\mathrm{H}^s_P} \lesssim_{q,s,P} \|u\|^q_{\mathrm{H}^s_P}$$

of the nonlinearity, which we prove by direct calculations. Note that the estimate likely extends to all the given orders, possibly using more advanced theory on composition operators.

Furthermore, the *a priori* estimates in [21] rely on several comparisons with the periodic and solitary-wave problems. We remove this dependence, which in particular gives a more explicit lower bound on the wave speeds.

Since we deal simultaneously with a class of nonlinearities and orders combined with arbitrary periods, the arguments are quite technical. Special concern is given to the P-dependence of estimates. To cope with this, we develop Fourier series and periodic Sobolev spaces for arbitrary periods with an emphasis on estimation/embedding constants.

With the exception of Chapter 2, Section 3.1 and Chapter 4, we clearly provide references to borrowed material and proofs throughout the text. In any case, there is a list of general references in the first part of each chapter. Consequently, each result and/or proof or example without citation has been constructed by the author.

Chapter 2

PRELIMINARIES

In this chapter we review some fundamental tools and spaces from real and functional analysis which provide a firm base for the rest of our discussion. It is assumed that the reader has basic knowledge of measure theory, Lebesgue integration and abstract algebra. We begin with rudimentary notions and results in the theory of operators between normed and inner product spaces. Next follows a survey of the classical C^n and L^p function spaces and ℓ^p sequence spaces, with some highlights of important theorems in integral calculus. At last comes an introduction to the concept of convolution.

We state most of the results without specific references and omit all proofs. The monograph by Kreyszig [43] contains details on functional analysis, while the works by Gasquet and Witomski [30, Chapter IV and Lesson 15], McDonald and Weiss [47] and Shkoller [58, Chapter 1] cover the remaining topics.

Throughout the chapter $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ denotes the scalar field and the range of functions. Another option, left to the reader, is that $\mathbb{R}$ is the field and $\mathbb{F}$ is the range.

2.1 OPERATORS ON NORMED AND INNER PRODUCT SPACES

Recall that a normed space $(X, \|\cdot\|)$ is said to be a *Banach space* if it is complete as a metric space. Likewise, an inner product space $(H, \langle\cdot,\cdot\rangle)$ over $\mathbb{F}$ is a *Hilbert space* whenever H constitutes a Banach space with the induced norm $\|x\| = \sqrt{\langle x, x\rangle}$.

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are called *equivalent* if

$$\|x\|_1 \asymp \|x\|_2 \quad \text{for all} \quad x \in X.$$

In addition, we remember that an *algebra* is an (additive) vector space with a compatible product structure. As such, we define a Banach space X to be a *Banach algebra* if it is an algebra satisfying

$$\|xy\| \lesssim \|x\|\|y\| \quad \text{for all} \quad x, y \in X.$$

Definition 2.1 (Operators and functionals). A map $T: X \to Y$ between two vector spaces X and Y over $\mathbb{F}$ is said to be an *operator*, and if $Y = \mathbb{F}$, we call T a *functional*. Moreover, T is *linear* provided

$$T(\lambda x + y) = \lambda Tx + Ty$$

for all scalars $\lambda \in \mathbb{F}$ and vectors $x, y \in X$, using the notational convention $Tx = T(x)$. □

Definition 2.2. Let $T: (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ be an operator. Then T is

i) *continuous* if it preserves convergence, that is,

$$x_n \to x \text{ in } X \qquad \text{implies} \qquad Tx_n \to Tx \text{ in } Y;$$

ii) *bounded* whenever it maps bounded sets to bounded sets, that is,

$$\|Tx\|_Y \lesssim \|x\|_X \quad \text{for all} \quad x \in X;$$

iii) *isometric* (norm-preserving) if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$; and

iv) *compact* if for every bounded sequence $\{x_n\}_{n\in\mathbb{N}}$ in X, the sequence $\{Tx_n\}_{n\in\mathbb{N}}$ has a subsequence which converges in Y.

We say that X is *continuously embedded* into Y, written $X \hookrightarrow Y$, whenever the identity map $\mathrm{id}: X \to Y$ is continuous, and if id is compact, we have a *compact embedding*. □

Continuity and boundedness are equivalent for linear operators. Moreover, the set $\mathscr{B}(X,Y)$ of all bounded linear operators $X \to Y$, with $\mathscr{B}(X) = \mathscr{B}(X,X)$, defines a normed space with the *operator norm*

$$\|T\|_{\mathscr{B}(X,Y)} = \sup_{\|x\|_X=1} \|Tx\|_Y.$$

In fact, if Y is a Banach space, so is $\mathscr{B}(X,Y)$, and if $T \in \mathscr{B}(X,Y)$ and $S \in \mathscr{B}(Y,Z)$, then the composition $S \circ T$ is in $\mathscr{B}(X,Z)$.

Definition 2.3 (Dual space). The space $\mathscr{B}(X,\mathbb{F})$ of all bounded linear functionals $X \to \mathbb{F}$ is called the *dual space* of X and is denoted by X'. □

Since $\mathbb{R}$ and $\mathbb{C}$ are complete, dual spaces always constitute Banach spaces. Furthermore, two fundamental results related to linear functionals are the Riesz' representation theorem and the Hahn-Banach theorem. We state the former only as a corollary of the full version.

Theorem 2.4 (Hahn-Banach theorem). *Let Y be a subspace of a normed space X. If $f \in Y'$, then there exists an extension $g \in X'$ with $g\restriction_Y = f$ and $\|g\|_{X'} = \|f\|_{Y'}$.* □

Theorem 2.5 (Riesz' representation theorem). *If $(H, \langle\cdot,\cdot\rangle)$ is a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, then the operator $x \mapsto \langle\cdot, x\rangle$ defines an isometric (anti-, if $\mathbb{F} = \mathbb{C}$) isomorphism $H \to H'$. In particular, every bounded linear functional on H is given by an inner product.* □

As a consequence of Theorem 2.5 it is often customary to identify H with (the complex conjugate of) H'.

Definition 2.6 (Adjoint operator). Let $T \in \mathscr{B}(X,Y)$, where X and Y are Hilbert spaces. The *(Hilbert) adjoint* or *dual* of T is the operator $T^* \in \mathscr{B}(Y,X)$ defined by

$$\langle Tx, y\rangle_Y = \langle x, T^*y\rangle_X$$

for all $x \in X$ and all $y \in Y$. □

The adjoint is well-defined and satisfies $\|T^*\|_{\mathscr{B}(Y,X)} = \|T\|_{\mathscr{B}(X,Y)}$.

Definition 2.7 (Unitary operator). Let H be a Hilbert space. Then $U \in \mathscr{B}(H)$ is said to be a *unitary operator* if

$$UU^* = U^*U = \mathrm{id}\colon H \to H.$$ □

In particular, U^{-1} exists and equals U^* for a unitary operators, and Definition 2.6 yields that

$$\langle Ux, Uy\rangle = \langle x, y\rangle \quad \text{for all} \quad x, y \in H.$$

Hence, unitary operators define automorphisms.

An important example of a unitary operator is the Fourier transform in $\mathrm{L}^2(\mathbb{R})$, which is the content of Plancherel's theorem in Chapter 4. But first we need to introduce the underlying L^p spaces.

2.2 C^n AND L^p FUNCTION SPACES AND ℓ^p SEQUENCE SPACES

If $\Omega \subseteq \mathbb{R}$ is open and $n \in \mathbb{N}$, we let $\mathrm{C}^n(\Omega)$ denote the standard space of all n times continuously differentiable maps $f : \Omega \to \mathbb{F}$, with the exponent omitted when $n = 0$. We also define $\mathrm{C}^\infty(\Omega) = \bigcap_{n\in\mathbb{N}} \mathrm{C}^n(\Omega)$ as the space of smooth or infinitely differentiable functions. In general, the label "real" if $\mathbb{F} = \mathbb{R}$ or "complex" if $\mathbb{F} = \mathbb{C}$ is added in front of a function space whenever the distinction is important.

We next express by $\mathrm{BC}^n(\Omega)$ the space of every $f \in \mathrm{C}^n(\Omega)$ for which $f^{(j)}$ is bounded on Ω for all $j = 0, \ldots, n$, and $\mathrm{BUC}^n(\Omega)$ denotes the space of $f \in \mathrm{BC}^n(\Omega)$ for which all $f^{(j)}$ are uniformly continuous. Moreover, $\mathrm{C}^n(\overline{\Omega})$ is the space of every $f \in \mathrm{C}^n(\Omega)$ for which all $f^{(j)}$ extend continuously to $\overline{\Omega}$. In each case we treat $n = 0$ and ∞ similarly as for $\mathrm{C}^n(\Omega)$.

Recall now that the *support* of $f \in \mathrm{C}(\Omega)$ is given as the closure of the set of points for which f does not vanish, that is,

$$\operatorname{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}} \subseteq \overline{\Omega}.$$

From this we introduce subspace $\mathscr{D}(\Omega)$ of all $f \in \mathrm{C}^\infty(\Omega)$ such that $\operatorname{supp} f^{(n)} \subset \Omega$ is compact for all $n \in \mathbb{N}$. Usually, elements of $\mathscr{D}(\Omega)$ are known as *test functions*.

When $n \neq \infty$, both $\mathrm{BC}^n(\Omega)$ and $\mathrm{BUC}^n(\Omega)$, and $\mathrm{C}^n(\overline{\Omega})$ if Ω is bounded, constitute Banach spaces when furnished with the norm

$$\|f\|_{\mathrm{BC}^n(\Omega)} = \sum_{j=0}^{n} \left\|f^{(j)}\right\|_\infty \asymp_n \max_{j=0,\dots,n} \left\|f^{(j)}\right\|_\infty,$$

where $\|f\|_\infty = \sup_{x\in\Omega} |f(x)|$ is the *supremum/uniform norm*. In fact, if Ω is bounded, then $\mathrm{BC}^n(\Omega)$, $\mathrm{BUC}^n(\Omega)$ and $\mathrm{C}^n(\overline{\Omega})$ coincide. This is a consequence of the extreme value theorem (Theorem 5.8), the fact that continuous functions on compact sets are uniformly continuous plus that bounded (and uniformly) continuous extensions are unique.

Definition 2.8 (L^p spaces). Let $\Omega \subseteq \mathbb{R}$ be Lebesgue measurable and $1 \le p \le \infty$. Then $\mathrm{L}^p(\Omega)$ denotes the set of all equivalence classes of Lebesgue measurable functions $f : \Omega \to \mathbb{F}$ for which

$$\|f\|_{\mathrm{L}^p(\Omega)} = \begin{cases} \left(\displaystyle\int_\Omega |f(x)|^p \,\mathrm{d}x\right)^{1/p} & \text{if } 1 \le p < \infty; \\ \operatorname*{ess\,sup}_{x\in\Omega} |f(x)| & \text{if } p = \infty, \end{cases}$$

is finite. Here

$$\operatorname*{ess\,sup}_{x\in\Omega} |f(x)| = \inf\left\{M : |f(x)| \le M \text{ for a.e. } x \in \Omega\right\}$$

is the *essential supremum* and *a.e.* means "almost every(where)." Functions in $\mathrm{L}^p(\Omega)$ belonging to the same equivalence class differ only on a set of measure 0 and are identified. Moreover, a map f is called *p-integrable* whenever $f \in \mathrm{L}^p(\Omega)$. □

The L^p spaces are sometimes referred to as *Lebesgue spaces* and $\mathbb{F}$ can even be the extended real number line $\mathbb{R} \cup \{-\infty, \infty\}$.

Remark. The a.e. precision of functions in $\mathrm{L}^p(\Omega)$ must for example be understood in the following sense. By "$f \in \mathrm{L}^p(\Omega)$ is continuous" we mean that there is a continuous representative $\tilde{f}$ in the equivalence class of f with $f(x) = \tilde{f}(x)$ for a.e. $x \in \Omega$. Similar implicit statements applies to function space inclusions associated with L^p spaces.

Moreover, the definition of support is not satisfactory for measurable functions (test with the characteristic function of $\mathbb{Q}$), but we can extend it. The *(essential) support* of a measurable function $f : \Omega \to \mathbb{F}$, still denoted by $\operatorname{supp} f$, is given as the smallest closed set $S \subseteq \mathbb{R}$ such that $f(x) = 0$ for a.e. $x \in \Omega \setminus S$. $\square$

As Definition 2.8 suggests, $L^p(\Omega)$ spaces form normed spaces, and the triangle inequality is also known as Minkowski's inequality. In fact, we have the following result.

Theorem 2.9 (Riesz-Fischer theorem). *$\big(L^p(\Omega), \|\cdot\|_{L^p(\Omega)}\big)$ constitutes a Banach space for all $1 \le p \le \infty$.* $\square$

Most important to us are the $L^2(\Omega)$ spaces since these become Hilbert spaces when endowed with the inner product

$$\langle f, g\rangle_{L^2(\Omega)} = \int_\Omega f(x)\overline{g(x)}\,\mathrm{d}x.$$

Proposition 2.10 (Hölder's inequality). *Let $1 \le p, q \le \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$ and*

$$\|fg\|_{L^1(\Omega)} \le \|f\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}. \qquad \square$$

We notice that Hölder's inequality coincides with the Cauchy-Schwarz inequality when $p = 2$.

There are in general no inclusion relationships between $L^p(\Omega)$ spaces with different exponents p if Ω has infinite measure. When $|\Omega| < \infty$, however, Hölder's inequality implies the following result.

Proposition 2.11 (L^p comparisons). *If $|\Omega| < \infty$, there exist continuous embeddings*

$$L^q(\Omega) \hookrightarrow L^p(\Omega) \quad \textit{whenever} \quad 1 \le p < q \le \infty,$$

and the embedding constant equals $|\Omega|^{\frac{1}{p}-\frac{1}{q}}$ with the convention that $\frac{1}{\infty} = 0$. $\square$

There are various ways to approximate L^p functions; for example we can use certain continuous functions.

Proposition 2.12 (Density in L^p). *Let $\Omega \subseteq \mathbb{R}$ be open and $1 \le p < \infty$. Then the space of continuous functions $f : \Omega \to \mathbb{F}$ with compact support is dense in $L^p(\Omega)$.* $\square$

L^p spaces make sense also for abstract measure spaces. In particular, if we endow $\mathbb{Z}$ with counting measure, the classical ℓ^p sequence spaces show up.

Definition 2.13 (ℓ^p spaces). Let $1 \le p \le \infty$. Then $\ell^p = \ell^p(\mathbb{Z})$ is the Banach space of all $\mathbb{F}$-valued sequences $x = \{x_k\}_{k\in\mathbb{Z}}$ with norm

$$\|x\|_{\ell^p} = \begin{cases} \Big(\sum_{k\in\mathbb{Z}} |x_k|^p\Big)^{1/p} & \text{if } 1 \le p < \infty; \\ \sup_{k\in\mathbb{Z}} |x_k| & \text{if } p = \infty. \end{cases}$$

Moreover, the inner product

$$\langle x, y\rangle_{\ell^2} = \sum_{k\in\mathbb{Z}} x_k \overline{y_k}$$

makes ℓ^2 a Hilbert space. □

Hölder's inequality becomes

$$\|xy\|_{\ell^1} \le \|x\|_{\ell^p} \|y\|_{\ell^q}$$

if $x \in \ell^p$ and $y \in \ell^q$, where $1 \le p, q \le \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ and $(xy)_k = x_k y_k$ for all $k \in \mathbb{Z}$. But contrary to Proposition 2.11, the ℓ^p spaces are increasing in p, that is,

$$\ell^p \hookrightarrow \ell^q \quad \text{whenever} \quad 1 \le p < q \le \infty, \tag{2.1}$$

with $\|x\|_{\ell^q} \le \|x\|_{\ell^p}$.

We also introduce the subspace c_0 of ℓ^∞ consisting of all convergent sequences whose limit is 0. Since c_0 is closed in ℓ^∞, it defines a Banach space with the ℓ^∞ norm.

2.3 Some important integral calculus

In this section we consider several major results which justify the interchange of limiting operations with respect to integration. Even though the uniform limit theorem for Riemann integral functions allows us to interchange limits and integration, the uniformity requirement is quite restrictive. The Lebesgue integral, however, is more adequate, and we have the following fundamental property.

Theorem 2.14 (Dominated convergence theorem). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : \Omega \to \mathbb{F}$ converging a.e. to $f : \Omega \to \mathbb{F}$. If there exists a nonnegative $g \in \mathrm{L}^1(\Omega)$ which "dominates" $\{f_n\}$, that is, for all $n \in \mathbb{N}$ we have $|f_n| \le g$ a.e., then f is integrable and*

$$\lim_{n\to\infty} \int_\Omega f_n(x)\,\mathrm{d}x = \int_\Omega f(x)\,\mathrm{d}x.$$

Moreover, $f_n \to f$ in $\mathrm{L}^1(\Omega)$. □

The sequential version of Theorem 2.14 becomes

$$\lim_{n\to\infty} \sum_{k\in\mathbb{Z}} x_{k,n} = \sum_{k\in\mathbb{Z}} x_k$$

whenever, for each $n \in \mathbb{N}$, the $\mathbb{F}$-valued sequence $\{x_{k,n}\}_{k\in\mathbb{Z}}$ satisfies $\lim_{n\to\infty} x_{k,n} = x_k$ and $|x_{k,n}| \le y_k$ for all $k \in \mathbb{Z}$, where $\{y_k\}_{k\in\mathbb{Z}} \in \ell^1$ is nonnegative.

Two other convergence results are given by the monotone convergence theorem and Fatou's lemma, but they are more or less equivalent to Theorem 2.14. Furthermore, by combining dominated convergence with the mean value theorem, we can differentiate under the integral sign, also known as Leibniz' integral rule.

Theorem 2.15 (Leibniz' integral rule). *Let $\Omega \subseteq \mathbb{R}$ be measurable and assume $I \subseteq \mathbb{R}$ is open. Suppose that $f : \Omega \times I \to \mathbb{F}$ satisfies the following properties.*

i) *$f(\cdot, t) \in \mathrm{L}^1(\Omega)$ for each $t \in I$;*

ii) *for a.e. $x \in \Omega$ the partial derivative $\frac{\partial f}{\partial t}(x,t)$ exists for all $t \in I$; and*

iii) *there exists $g \in \mathrm{L}^1(\Omega)$ such that for each $t \in I$ it is true that*

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x) \quad \textit{for a.e.} \quad x \in \Omega.$$

Then the map $\int_\Omega f(x,\cdot)\,\mathrm{d}x$ is differentiable for all $t \in I$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega f(x,t)\,\mathrm{d}x = \int_\Omega \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x. \qquad \square$$

Remark (Continuity of integrals). Similarly as Leibniz' integral rule, Theorem 2.14 also gives continuity of

$$\int_\Omega f(x,\cdot)\,\mathrm{d}x$$

on I provided ii) and iii) in Theorem 2.15 are replaced with

ii) for a.e. $x \in \Omega$ we have $f(x,\cdot) \in \mathrm{C}(I)$; and

iii) there is $g \in \mathrm{L}^1(\Omega)$ such that for each $t \in I$, we have $|f(x,t)| \leq g(x)$ for a.e.$x \in \Omega$. $\square$

We next consider the problem of interchanging the order of integration in iterated integrals. Two crucial results are given by that of Fubini and Tonelli, which we state in a suitably combined version.

Theorem 2.16 (Fubini-Tonelli theorem). *Let $\Omega \times \Gamma$ be a measurable subset of $\mathbb{R} \times \mathbb{R}$. Suppose that $f : \Omega \times \Gamma \to \mathbb{F}$ is a measurable function for which any of the three integrals*

$$\int_{\Omega\times\Gamma} |f(x,y)|\,\mathrm{d}(x,y), \qquad \int_\Omega \left(\int_\Gamma |f(x,y)|\,\mathrm{d}y\right)\mathrm{d}x \qquad \textit{or} \qquad \int_\Gamma \left(\int_\Omega |f(x,y)|\,\mathrm{d}x\right)\mathrm{d}y$$

is finite. Then

$$\int_{\Omega\times\Gamma} f(x,y)\,\mathrm{d}(x,y) = \int_\Omega \left(\int_\Gamma f(x,y)\,\mathrm{d}y\right)\mathrm{d}x = \int_\Gamma \left(\int_\Omega f(x,y)\,\mathrm{d}x\right)\mathrm{d}y.$$

Similarly, if any of the quantities

$$\sum_{k\in\mathbb{Z}}\Big(\sum_{n\in\mathbb{Z}}|x_{k,n}|\Big) \qquad \textit{or} \qquad \sum_{n\in\mathbb{Z}}\Big(\sum_{k\in\mathbb{Z}}|x_{k,n}|\Big)$$

is finite, then

$$\sum_{k\in\mathbb{Z}}\Big(\sum_{n\in\mathbb{Z}}x_{k,n}\Big)=\sum_{n\in\mathbb{Z}}\Big(\sum_{k\in\mathbb{Z}}x_{k,n}\Big).$$ □

2.4 THE CONVOLUTION OF FUNCTIONS AND SEQUENCES

Our last concern in this review is the concept of convolution.

Definition 2.17 (Convolution). The *convolution* $f*g$ of two functions $f,g\colon\mathbb{R}\to\mathbb{F}$ is a map given by

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

Moreover, if $x=\{x_k\}_{k\in\mathbb{Z}}$ and $y=\{y_k\}_{k\in\mathbb{Z}}$ are two $\mathbb{F}$-valued sequences, then

$$(x*y)_k=\sum_{n\in\mathbb{Z}}x_{k-n}\,y_n$$

defines the *(discrete) convolution* $x*y=\{(x*y)_k\}_{k\in\mathbb{Z}}$. □

When the convolution exists, its commutativity is assured by a simple change of variables, while distributivity is inherited from integration. We next display how the convolution behaves under different assumptions on the functions f and g.

Proposition 2.18 (Continuity of the convolution). *The convolution constitutes a continuous bilinear operator*

$$*\colon \mathrm{L}^p(\mathbb{R})\times\mathrm{L}^q(\mathbb{R})\to\mathrm{L}^r(\mathbb{R}) \quad \textit{for} \quad 1\le p,q,r\le\infty \quad \textit{with} \quad \tfrac{1}{p}+\tfrac{1}{q}-\tfrac{1}{r}=1; \tag{2.2a}$$

$$*\colon \mathrm{L}^p(\mathbb{R})\times\mathrm{L}^q(\mathbb{R})\to\mathrm{BC}(\mathbb{R}) \quad \textit{for} \quad 1\le p,q\le\infty \quad \textit{with} \quad \tfrac{1}{p}+\tfrac{1}{q}=1; \tag{2.2b}$$

$$*\colon \mathrm{L}^1(\mathbb{R})\times\mathrm{BC}^n(\mathbb{R})\to\mathrm{BC}^n(\mathbb{R}) \quad \textit{for} \quad n\in\mathbb{N}. \tag{2.2c}$$

Additionally, $(f*g)^{(j)}=f*g^{(j)}$ *holds for all* $j=0,\dots,n$ *in case* (2.2c). □

Property (2.2a) is known as Young's inequality for convolutions and also applies to give a continuous bilinear mapping

$$*\colon \ell^p\times\ell^q\to\ell^r \quad \text{for} \quad 1\le p,q,r\le\infty \quad \text{with} \quad \tfrac{1}{p}+\tfrac{1}{q}-\tfrac{1}{r}=1 \tag{2.3}$$

in the discrete case. Furthermore, it shows that the convolution is associative on $\mathrm{L}^1(\mathbb{R})$ and ℓ^1 and thus defines a product structure.

The situation (2.2c) is delicate. Given a highly irregular function $f \in \mathrm{L}^1(\mathbb{R})$, we can still make the convolution $f * g$ as smooth as we wish by choosing g nice enough. For particular g it is even possibly to approximate f via $f * g$, and more generally this leads to the the density of $\mathscr{D}(\Omega)$ in $\mathrm{L}^p(\Omega)$ whenever $\Omega \subseteq \mathbb{R}$ is open and $1 < p < \infty$. We give an example of this so-called *mollification* procedure in Section 3.3.

Definition 2.19 (Locally integrable functions). Let $1 \le p \le \infty$ and $\Omega \subseteq \mathbb{R}$ be open. Then

$$\mathrm{L}^p_{\mathrm{loc}}(\Omega) = \left\{ f : \Omega \to \mathbb{F} \text{ measurable} : f\restriction_K \in \mathrm{L}^p(K) \text{ for all compact subsets } K \subset \Omega \right\}.$$

denotes the space of *locally p-integrable functions*. A sequence $\{f_n\}_{n\in\mathbb{N}}$ converges to f in $\mathrm{L}^p_{\mathrm{loc}}(\Omega)$ provided $f_n \to f$ in $\mathrm{L}^p(K)$ for all compact subsets $K \subset \Omega$. □

Evidently $\mathrm{L}^p(\Omega) \subset \mathrm{L}^p_{\mathrm{loc}}(\Omega)$ for all $1 \le p \le \infty$, and analogously to Proposition 2.11 there exist continuous embeddings

$$\mathrm{L}^q_{\mathrm{loc}}(\Omega) \hookrightarrow \mathrm{L}^p_{\mathrm{loc}}(\Omega) \quad \text{whenever} \quad 1 \le p < q \le \infty.$$

We end our discussion with an important result, also known as the fundamental lemma of the calculus of variations. It establishes a correspondence between the weak (variational) and strong formulation of elliptic differential problems, see Section 5.1, and is a basic component of distribution theory.

Theorem 2.20 (Du Bois-Reymond's lemma). *Let $\Omega \subseteq \mathbb{R}$ be open and $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. If*

$$\int_\Omega f(x)\,\varphi(x)\,\mathrm{d}x = 0 \quad \textit{for all} \quad \varphi \in \mathscr{D}(\Omega),$$

then $f = 0$ a.e. in Ω. □

The proof of Theorem 2.20 follows quickly by contradiction for $f \in \mathrm{C}(\Omega)$, whereas the general statement can be established via mollification.

Chapter 3

Fourier series and periodic Sobolev spaces

In the early 19th century the French mathematician and physicist Fourier revolutionized both mathematics and physics by investigating the decomposition of a periodic function f into a countable sum of sines and cosines [67], that is,

$$f(x) \stackrel{?}{=} \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\tfrac{2\pi k x}{P}\right) + b_k \sin\left(\tfrac{2\pi k x}{P}\right) \tag{3.1a}$$

or

$$f(x) \stackrel{?}{=} \sum_{k\in\mathbb{Z}} c_k\, \mathrm{e}^{2\pi \mathrm{i} k x/P}, \tag{3.1b}$$

where P is the period of f. Utilizing orthogonality properties of sine and cosine, he found simple formulas for the coefficients a_k and b_k, or c_k, and then applied his technique in the analysis of the heat equation with periodic boundary conditions.

The underlying method of Fourier has since been rigorously established, and the right-hand side expressions in (3.1) are today known as the *Fourier series* associated with f. There are numerous applications in science and engineering, such as signal analysis, image processing and quantum mechanics [67], and as a branch of harmonic analysis, Fourier methods also share deep connections with many different fields inside mathematics.

Their relevance is especially true within the study of differential equations. Here it is often difficult to prove the existence of solutions in the classical sense, that is, of functions which are pointwise differentiable. By weakening the notion of differentiability, however, we may establish so-called *weak solutions*, and regularity theory can sometimes show that they in fact satisfy the equations classically. The weak derivatives may for instance be defined via Fourier series, which leads to the study of *periodic Sobolev spaces*.

Our discussion starts abstractly in Section 3.1 with Fourier series in general Hilbert spaces. Section 3.2 introduces C^n and L^p spaces for periodic functions, and in Section 3.3 we prove validity of the classical Fourier series representation of periodic L^2 functions. Section 3.4 next explores fundamental properties of Fourier series related to convergence, differentiability, decay and convolution. In Section 3.5 we finally examine weak differentiation and the periodic Sobolev spaces.

There is a vast literature on the topic of Fourier series. We follow mainly the works by Hunter and Nachtergaele [35, Chapters 6-7], Gasquet and Witomski [30, Chapter 2 and Lesson 16] and Iorio and Iorio [40, Chapters 2-3].

Remark. The base field and function range, if relevant, for all the considered spaces is $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ throughout the chapter. (The case with $\mathbb{R}$ as the field and $\mathbb{F}$ as the range is left to the reader.) □

3.1 FOURIER SERIES IN HILBERT SPACES

Our study begins with a refreshment of the concepts of orthogonality, projection, convergence, bases and subsequently Fourier series in abstract Hilbert spaces. The material, along with its omitted proofs, is based upon Ehrnström [18, Chapter 4], Hunter and Nachtergaele [35, Chapter 6], Gasquet and Witomski [30, Lesson 16] and Kreyszig [43, Chapter 3].

Recall that two vectors f and g in an inner product space $(H, \langle\cdot,\cdot\rangle)$ are said to be *orthogonal* if $\langle f, g\rangle = 0$, and from this we can introduce the *orthogonal complement*

$$C^{\perp} = \left\{ f \in H : \langle f, g\rangle = 0 \text{ for all } g \in C \right\}$$

of any $C \subseteq H$. It is clear that nonzero orthogonal vectors are linearly independent, and the Pythagorean identity

$$\left\| \sum_{k=1}^{n} \varphi_k \right\|^2 = \sum_{k=1}^{n} \|\varphi_k\|^2.$$

holds whenever $\varphi_1, \ldots, \varphi_n \in H$ are pairwise orthogonal.

By pairing orthogonality with completeness of the space, we obtain the following important orthogonal projection theorem, for which the success of least squares methods hinges on.

Theorem 3.1 (Orthogonal projection theorem). *Let C be a nonempty, closed and convex subset of a Hilbert space H. Then for any $f \in H$ there exists a unique $f^* \in C$, called the* best approximation of f in C *or the* closest point to f in C*, such that*

$$\|f - f^*\| = \inf_{g \in C} \|f - g\|.$$

If C in addition is a subspace of H, then H equals (isomorphically) the direct sum $C \oplus C^{\perp}$. □

The second statement $H \cong C \oplus C^{\perp}$ means that for every $f \in H$ there is a unique $f^* \in C$ such that $f - f^* \in C^{\perp}$. We refer to f^* as the *orthogonal projection of f into C*, which coincides with the best approximation of f in C.

Definition 3.2 (Orthogonal system). Let Λ be an index set. A subset $\{\varphi_k\}_{k\in\Lambda}$ of nonzero elements in an inner product space is said to be an *orthogonal system* if $\langle\varphi_k,\varphi_l\rangle = 0$ whenever $k \neq l$. If in addition $\|\varphi_k\| = 1$ for all $k \in \Lambda$, the system is called *orthonormal*. □

For simplicity we focus on orthonormal sequences with $\Lambda = \mathbb{Z}$, but remark that the upcoming results extend to arbitrary index sets.

When we project vectors into the linear span of a finite orthonormal system, the orthogonal projection theorem takes a very neat form.

Proposition 3.3. *Let $\{\varphi_1,\dots,\varphi_n\}$ be an orthonormal system in a Hilbert space H. Given any $f \in H$, the best approximation of f in $\operatorname{span}\{\varphi_1,\dots,\varphi_n\}$ equals $g = \sum_{k=1}^{n}\langle f,\varphi_k\rangle\,\varphi_k$, with*

$$\|f-g\|^2 = \|f\|^2 - \sum_{k=1}^{n}|\langle f,\varphi_k\rangle|^2.$$

In particular, if $f \in \operatorname{span}\{\varphi_1,\dots,\varphi_n\}$, then

$$f = \sum_{k=1}^{n}\langle f,\varphi_k\rangle\,\varphi_k. \tag{3.2}$$

□

The orthogonality of $\{\varphi_1,\dots,\varphi_n\}$ in Proposition 3.3 also allows us to conclude that the best approximation of f improves with n, that is, if $\{\varphi_1,\dots,\varphi_{n+1}\}$ is an orthonormal system in a Hilbert space, then

$$\left\|f - \sum_{k=1}^{n+1}\langle f,\varphi_k\rangle\,\varphi_k\right\| \le \left\|f - \sum_{k=1}^{n}\langle f,\varphi_k\rangle\,\varphi_k\right\|.$$

Proposition 3.4 (Bessel's inequality). *Orthonormal sequences $\{\varphi_k\}_{k\in\mathbb{Z}}$ in an inner product space H satisfy*

$$\sum_{k\in\mathbb{Z}}|\langle f,\varphi_k\rangle|^2 \le \|f\|^2 \quad \text{for all} \quad f\in H.$$

□

Bessel's inequality implies especially that $\langle f,\varphi_k\rangle \to 0$ as $|k|\to\infty$.

Suppose now that we are given a general orthonormal sequence $\{\varphi_k\}_{k\in\mathbb{Z}}$ in a Hilbert space H and a vector $f \in H$. Based on (3.2) it would be nice if we could write

$$f = \sum_{k\in\mathbb{Z}}\langle f,\varphi_k\rangle\,\varphi_k.$$

It is, however, neither obvious that the series on the right exists in H nor that equality with f is obtained. By the last property we mean that

$$\lim_{m,n\to\infty}\left\|f - \sum_{k=-m}^{n}\langle f,\varphi_k\rangle\,\varphi_k\right\| = 0.$$

Happily, the first attribute is true via the following result, which relates convergence in H to that of ℓ^2.

Proposition 3.5 (Convergence). *If $\{\varphi_k\}_{k\in\mathbb{Z}}$ is orthonormal in a Hilbert space H and $\{c_k\}_{k\in\mathbb{Z}}$ is a sequence of scalars, then*

$$\sum_{k\in\mathbb{Z}} c_k\varphi_k \text{ converges in } H \quad \Leftrightarrow \quad \sum_{k\in\mathbb{Z}} |c_k|^2 < \infty. \qquad \square$$

Putting $c_k = \langle f, \varphi_k\rangle$, the existence of the series $\sum_{k\in\mathbb{Z}}\langle f, \varphi_k\rangle\,\varphi_k$ is then assured from Bessel's inequality.

Definition 3.6 (Fourier series and orthonormal basis). Assume that $\{\varphi_k\}_{k\in\mathbb{Z}}$ is orthonormal in a Hilbert space H and let $f\in H$. Then we call $\sum_{k\in\mathbb{Z}}\widehat{f}_k\varphi_k$ the *Fourier series* of f with respect to $\{\varphi_k\}$, where $\widehat{f}_k = \langle f, \varphi_k\rangle$ is the *kth Fourier coefficient* of f.

If $f = \sum_{k\in\mathbb{Z}}\widehat{f}_k\varphi_k$ for all $f\in H$, then $\{\varphi_k\}$ is said to be an *orthonormal basis*. $\square$

Orthonormal bases can be characterized in several important ways.

Theorem 3.7 (Fourier series theorem). *Suppose that $\{\varphi_k\}_{k\in\mathbb{Z}}$ is orthonormal in a Hilbert space H. Then the following statements are equivalent.*

i) *$\{\varphi_k\}$ defines an orthonormal basis;*

ii) *the set of finite linear combinations from $\{\varphi_k\}$ is dense in H, that is, $\overline{\operatorname{span}\{\varphi_k\}} = H$;*

iii) *$\{\varphi_k\}$ is complete (or total): If $\widehat{f}_k = 0$ for all $k\in\mathbb{Z}$, then $f = 0$; and*

iv) *Parseval's identity holds:*

$$\langle f, g\rangle = \sum_{k\in\mathbb{Z}}\widehat{f}_k\,\overline{\widehat{g}_k} \quad \text{for all} \quad f, g\in H. \tag{3.3}$$

In particular, $\|f\|^2 = \sum_{k\in\mathbb{Z}}|\widehat{f}_k|^2$. $\square$

If $\{\varphi_k\}_{k\in\mathbb{Z}}$ is orthonormal in H, we can introduce the operator

$$\mathscr{F} : f\in H \mapsto \widehat{f} = \{\widehat{f}_k\}_{k\in\mathbb{Z}}$$

called the *Fourier transform* with respect to $\{\varphi_k\}$. This is a linear map by linearity of inner products. If $\{\varphi_k\}$ in addition forms a basis, then Proposition 3.5 and Theorem 3.7 show that

$$\mathscr{F} : H \to \ell^2$$

defines an isometric isomorphism satisfying $\langle f, g\rangle_H = \langle\widehat{f}, \widehat{g}\rangle_{\ell^2}$. In particular, the completeness of $\{\varphi_k\}$ gives injectivity of $\mathscr{F}$ and uniqueness of the Fourier coefficients.

But do orthonormal bases exist? Yes, Zorn's lemma implies that any Hilbert space contains an orthonormal basis provided we allow for uncountable index sets Λ. For separable Hilbert spaces H there is, however, a more constructive approach.

Recall that a metric space is *separable* when it contains a countable dense subset. If $\{\phi_n\}_{n\in\mathbb{Z}}$ is dense in H, we can use the *Gram–Schmidt process* to construct an orthonormal sequence $\{\varphi_k\}_{k\in\mathbb{Z}}$ satisfying $\overline{\operatorname{span}\{\varphi_k\}} = \overline{\{\phi_n\}}$. The Gram-Schmidt process builds $\{\varphi_k\}$ inductively by adding, whenever nonzero, the normalized orthogonal projection of ϕ_n into the current, finite linear span of the φ_k's; see Hunter and Nachtergaele [35, Chapter 6]. Thus $\{\varphi_k\}$ is an orthonormal basis and we have the following result.

Theorem 3.8. *Every separable Hilbert space is isomorphic to ℓ^2 via the Fourier transform $\mathscr{F}$.* □

We will prove in Section 3.3 that the periodic L^2 space to be defined in Section 3.2 is separable via the Fourier basis of complex exponentials.

3.2 PERIODIC C^n AND L^p SPACES

In this section we discuss the analogues of the C^n and L^p spaces for (P-)*periodic functions* $f : \mathbb{R} \to \mathbb{F}$, which satisfy

$$f(x+P) = f(x) \quad \text{for all} \quad x \in \mathbb{R}, \tag{3.4}$$

where $P > 0$ is called the *period*; see also Appendix A. Periodic functions are closed under pointwise addition and scalar multiplication, so that the following definition makes sense.

Definition 3.9 (Periodic L^q and C^n spaces). Let $1 \le q \le \infty$ and $n \in \mathbb{N} \cup \{\infty\}$ and assume that $P > 0$. Then L^q_P is the subspace of all P-periodic functions $f \in L^q_{\mathrm{loc}}(\mathbb{R})$ and C^n_P denotes the subspace of all P-periodic functions $f \in C^n(\mathbb{R})$ (with no superscript when $n = 0$). □

Remark. We use q instead of p for clarity. Pointwise properties of elements of L^q_P hold, as usual, up to a set of measure zero. This applies in particular to (3.4). □

All P-periodic functions f are fully determined on $[0, P)$ or any half-open interval of length P. For example,

$$\int_y^{P+y} f(x)\,\mathrm{d}x = \int_0^P f(x)\,\mathrm{d}x$$

for any $y \in \mathbb{R}$. Thus we may equally view f as a map $[0, P] \to \mathbb{F}$ satisfying $f(0) = f(P)$. Another representation is $f : \mathbb{T} \to \mathbb{F}$, where $\mathbb{T}$ is the circle of radius P defined naturally by the quotient $\mathbb{R}/P\mathbb{Z}$. This gives the more common notations $\mathrm{L}^q(\mathbb{T})$ and $\mathrm{C}^n(\mathbb{T})$ for L^q_P and C^n_P. Since we shall work with arbitrary periods in Chapter 6, however, our choice seems beneficial.

Definition 3.10 (Periodic convolution). The *P-periodic convolution* $f *_P g$ between two P-periodic functions $f, g\colon \mathbb{R} \to \mathbb{F}$ is given by

$$f *_P g(x) = \int_0^P f(x-y)\,g(y)\,\mathrm{d}y. \qquad \square$$

This notion is similar as that of Definition 2.17 and $f *_P g$ is clearly P-periodic. Moreover, commutativity follows for example by periodicity.

Theorem 3.11 (Properties of L^q_P and C^n_P). *Let $1 \le q, r \le \infty$ and $n \in \mathbb{N}$.*

i) L^q_P *is a Banach space when endowed with the norm*

$$\|f\|_{\mathrm{L}^q_P} = \left(\int_0^P |f(x)|^q \,\mathrm{d}x\right)^{1/q},$$

and the inner product

$$\langle f, g\rangle_{\mathrm{L}^2_P} = \int_0^P f(x)\overline{g(x)}\,\mathrm{d}x$$

makes L^2_P *a Hilbert space;*

ii) C^n_P *forms a Banach space with the equivalent norms*

$$\|f\|_{\mathrm{C}^n_P} = \sum_{j=0}^{n} \left\|f^{(j)}\right\|_\infty \asymp_n \max_{j=0,\dots,n} \left\|f^{(j)}\right\|_\infty, \tag{3.5}$$

where the supremum norm in practice runs over $[0,P]$*;*

iii) C_P *is dense in* L^q_P *for* $1 \le q < \infty$*;*

iv) there are continuous embeddings

$$\mathrm{C}^n_P \hookrightarrow \mathrm{BUC}^n(\mathbb{R}), \qquad \mathrm{C}^m_P \hookrightarrow \mathrm{C}^n_P \quad \textit{if} \quad m > n \qquad \textit{and} \qquad \mathrm{L}^q_P \hookrightarrow \mathrm{L}^r_P \quad \textit{if} \quad q > r,$$

where the embedding constant in the last case equals $P^{\frac{1}{r}-\frac{1}{q}}$*; and*

v) the P-periodic convolution defines continuous bilinear mappings

$$*_P\colon \mathrm{L}^p_P \times \mathrm{L}^q_P \to \mathrm{L}^r_P \quad \textit{for} \quad 1 \le p,q,r \le \infty \quad \textit{with} \quad \tfrac{1}{p} + \tfrac{1}{q} - \tfrac{1}{r} = 1; \tag{3.6a}$$

$$*_P\colon \mathrm{L}^p_P \times \mathrm{L}^q_P \to \mathrm{C}_P \quad \textit{for} \quad 1 \le p,q \le \infty \quad \textit{with} \quad \tfrac{1}{p} + \tfrac{1}{q} = 1; \tag{3.6b}$$

$$*_P\colon \mathrm{L}^1_P \times \mathrm{C}^n_P \to \mathrm{C}^n_P \quad \textit{for} \quad n \in \mathbb{N}. \tag{3.6c}$$

The formula $(f *_P g)^{(j)} = f *_P g^{(j)}$ *holds for all* $j = 0, \dots, n$ *in case* (3.6c).

Proof. This is a consequence of results, particularly for $\mathrm{L}^q([0,P])$ and $\mathrm{BUC}^n((0,P))$, from Chapter 2; we omit the details. □

There is a canonical, continuous embedding from the set of functions $f \in \mathrm{L}^q(\mathbb{R})$ with $\operatorname{supp} f \subseteq [0,P]$ (any interval of length P) to L^q_P via the *P-periodic extension map*

$$f \mapsto f_P = \sum_{j\in\mathbb{Z}} f(\cdot + jP). \tag{3.7}$$

This is valid for all $1 \leq q \leq \infty$ and f_P is called the *P-periodic extension* of f; see Ehrnström *et al.* [21].

3.3 THE CLASSICAL FOURIER BASIS OF COMPLEX EXPONENTIALS FOR L^2_P

Recall from (3.1) that Fourier's original idea was to decompose a periodic function f into a sum of sines and cosines. We shall prove that this representation is valid in L^2_P, which, based on Theorem 3.7, amounts to show that the system

$$\left\{ \frac{1}{\sqrt{P}}, \quad \frac{1}{\sqrt{P}}\cos\left(\tfrac{2\pi kx}{P}\right), \quad \frac{1}{\sqrt{P}}\sin\left(\tfrac{2\pi kx}{P}\right) \right\}_{k\in\mathbb{Z}_+}$$

forms an orthonormal basis for L^2_P. By de Moivre's formula $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$ we have

$$\frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos\left(\tfrac{2\pi kx}{P}\right) + b_k \sin\left(\tfrac{2\pi kx}{P}\right) = \sum_{|k|\leq n} c_k\, \mathrm{e}^{2\pi\mathrm{i}kx/P} \tag{3.8}$$

for some coefficients

$$c_0 = \tfrac{1}{2}a_0 \qquad \text{and} \qquad c_k = \begin{cases} \tfrac{1}{2}(a_k - \mathrm{i}b_k) & \text{if } k > 0; \\ \tfrac{1}{2}(a_k + \mathrm{i}b_k) & \text{if } k < 0. \end{cases}$$

We therefore work, for simplicity, with the system of *complex exponentials* $\{\varphi_k\}_{k\in\mathbb{Z}}$ defined by

$$\varphi_k(x) = \frac{1}{\sqrt{P}}\,\mathrm{e}^{2\pi\mathrm{i}kx/P}.$$

Linear combinations of functions in both systems are called *trigonometric polynomials*.

Remark. The complex exponentials are not included in real spaces with $\mathbb{F} = \mathbb{R}$. But since the Fourier coefficients of $f\colon \mathbb{R} \to \mathbb{R}$ with respect to $\{\varphi_k\}_{k\in\mathbb{Z}}$ satisfy

$$\overline{\widehat{f}_k} = \widehat{f}_{-k} \quad \text{for all} \quad k \in \mathbb{Z},$$

we can still treat $\{\varphi_k\}$ as a basis and represent f by $\sum_{k\in\mathbb{Z}} \widehat{f}_k\, \varphi_k$ (in some sense). □

The main challenge is to prove that $\{\varphi_k\}$ is a basis for L^2_P, because evidently

$$\langle \varphi_k, \varphi_l \rangle_{\mathrm{L}^2_P} = \begin{cases} 1 & \text{if } k = l; \\ 0 & \text{if } k \neq l. \end{cases}$$

We follow Hunter and Nachtergaele [35, Section 7.1] and establish that any $f \in \mathrm{C}_P$ can be uniformly approximated by trigonometric polynomials through mollification. Since C_P is dense in L^2_P by Theorem 3.11 iii) and uniform convergence implies L^2_P-convergence in the compact case, $\operatorname{span}\{\varphi_k\}$ must be dense in L^2_P.

The *P-periodic mollification* of a function $f \in \mathrm{C}_P$ is defined as $\rho_n *_P f$, where $\{\rho_n\}_{n\in\mathbb{N}}$ is a P-periodic mollifier according to the following definition.

Definition 3.12 (Periodic mollifier). A sequence $\{\rho_n\}_{n\in\mathbb{N}} \subset \mathrm{C}_P$ is a *P-periodic mollifier* if it satisfies

i) positivity: $\rho_n(y) \geq 0$ for all $y \in \mathbb{R}$;

ii) normalization: $\|\rho_n\|_{\mathrm{L}^1_P} = 1$; and

iii) concentration: $\displaystyle\lim_{n\to\infty} \int_{\delta \leq |y| \leq \frac{P}{2}} \rho_n(y)\,\mathrm{d}y = 0 \quad \text{for every} \quad \delta \in \left(0, \tfrac{P}{2}\right].$ □

Lemma 3.13. *Let $f \in \mathrm{C}_P$ and $\{\rho_n\}_{n\in\mathbb{N}}$ be a P-periodic mollifier. Then $\rho_n *_P f$ converges uniformly to f as $n \to \infty$.*

Proof. Inspired by Hunter and Nachtergaele [35, Theorem 7.2], let $\epsilon > 0$ and observe by normalization and positivity of $\{\rho_n\}$ that

$$\left|f(x) - \rho_n *_P f(x)\right| \leq \int_{-\frac{P}{2}}^{\frac{P}{2}} \rho_n(y) \left|f(x) - f(x-y)\right| \mathrm{d}y.$$

For every $\delta_\epsilon > 0$ we can split the integral in two and estimate

$$\sup_{|x|\leq\frac{P}{2}} \int_{|y|<\delta_\epsilon} \rho_n(y)\left|f(x)-f(x-y)\right|\mathrm{d}y \leq \sup_{|x|\leq\frac{P}{2}} \sup_{|y|<\delta_\epsilon} \left|f(x)-f(x-y)\right|$$

and

$$\sup_{|x|\leq\frac{P}{2}} \int_{\delta_\epsilon\leq|y|\leq\frac{P}{2}} \rho_n(y)\left|f(x)-f(x-y)\right|\mathrm{d}y \leq 2\|f\|_\infty \int_{\delta_\epsilon\leq|y|\leq\frac{P}{2}} \rho_n(y)\,\mathrm{d}y.$$

Since f is uniformly continuous, the first part is small for sufficiently small δ_ϵ. Given such δ_ϵ, the concentration of $\{\rho_n\}$ shows that the second part vanishes as $n \to \infty$. Thus $\|f - \rho_n *_P f\|_\infty < \epsilon$ for all sufficiently large n. □

Lemma 3.14. *The set* $\operatorname{span}\{\varphi_k\}_{k\in\mathbb{Z}}$ *of trigonometric polynomials is dense in* C_P. *In particular,* C_P^∞ *is dense in* L_P^q *for* $1 \le q < \infty$.

Proof. The argument is adapted from Hunter and Nachtergaele [35, Theorem 7.3].

If $f \in \mathrm{L}_P^1$ and $g \in \operatorname{span}\{\varphi_k\}$, then $f *_P g \in \operatorname{span}\{\varphi_k\}$ also, because

$$f *_P \varphi_k = \sqrt{P}\langle f, \varphi_k\rangle_{\mathrm{L}_P^2}\, \varphi_k,$$

combined with linearity of convolution. By Lemma 3.13 it therefore suffices to construct a P-periodic mollifier $\{\rho_n\}_{n\in\mathbb{N}} \subset \operatorname{span}\{\varphi_k\}$. To this end, define the P-periodic functions

$$\rho_n(x) = c_n\left(1+\cos\left(\tfrac{2\pi x}{P}\right)\right)^n,$$

where c_n chosen such that $\{\rho_n\}$ satisfies positivity and normalization. It is possible to write

$$\rho_n = \sum_{|k|\le n} a_{n,k}\,\varphi_k, \qquad \text{where} \qquad a_{n,k} = \frac{c_n\sqrt{P}}{2^n}\binom{2n}{n+k},$$

and thus $\{\rho_n\} \subset \operatorname{span}\{\varphi_k\}$. Moreover, on the interval $\left[-\frac{P}{2}, \frac{P}{2}\right]$, the map $1+\cos\left(\frac{2\pi\cdot}{P}\right)$ is symmetric, has a unique maximal point at the origin and is strictly decreasing on $\left[0, \frac{P}{2}\right]$. Hence, given $\delta \in \left(0, \frac{P}{2}\right]$ and $\delta \le |x| \le \frac{P}{2}$, we infer that $\rho_n(x) \searrow 0$ as $n \to \infty$. This yields concentration and so $\{\rho_n\}$ is a mollifier.

The last statement follows from Theorem 3.11 iii). □

The density of $\operatorname{span}\{\varphi_k\}_{k\in\mathbb{Z}}$ in L_P^2 is now established and we deduce the targeted result.

Theorem 3.15 (Fourier basis for L_P^2). *The complex exponentials* $\{\varphi_k\}_{k\in\mathbb{Z}}$ *form an orthonormal basis for* L_P^2. *In particular, the Fourier transform is an isometric isomorphism*

$$\mathscr{F}: \mathrm{L}_P^2 \to \ell^2, \qquad f \mapsto \widehat{f} = \{\widehat{f}_k\}_{k\in\mathbb{Z}},$$

so that every $f \in \mathrm{L}_P^2$ *has a unique Fourier series representation*

$$f = \mathscr{F}^{-1}(\widehat{f}) = \frac{1}{\sqrt{P}}\sum_{k\in\mathbb{Z}} \widehat{f}_k\, \mathrm{e}^{2\pi i k\cdot/P} \quad \text{in} \quad \mathrm{L}_P^2. \qquad □$$

3.4 Convergence, differentiability, decay and convolution

Although Theorem 3.15 is pleasing, it only shows that the Fourier series of $f \in \mathrm{L}_P^2$ converges in the norm sense

$$\lim_{m,n\to\infty} \|f - S_{m,n}(f)\|_{\mathrm{L}_P^2} = 0,$$

where

$$S_{m,n}(f)(x) = \frac{1}{\sqrt{P}} \sum_{k=-m}^{n} \widehat{f}_k \, e^{2\pi i k x/P}$$

is the sequence of partial sums. For general $f \in L^2_P$ we do not know whether $S_{\infty,\infty}(f)$ exists pointwise, or if it it exists at x, that the *Fourier inversion formula*

$$f(x) = S_{\infty,\infty}(f)(x) = \frac{1}{\sqrt{P}} \sum_{k\in\mathbb{Z}} \widehat{f}_k \, e^{2\pi i k x/P} \tag{3.9}$$

holds. As such, it is of interest to find sufficient conditions on f for the validity of its pointwise Fourier representation. Due to (3.8), the inversion problem is often also posed for the symmetric limit

$$f(x) = S_{\infty}(f)(x) = \lim_{n\to\infty} \frac{1}{\sqrt{P}} \sum_{|k|\le n} \widehat{f}_k \, e^{2\pi i k x/P},$$

where $S_n(f) = S_{n,n}(f)$ is the symmetric partial sum. Of course, if $S_{\infty,\infty}(f)(x)$ exists, then $S_\infty(f)(x) = S_{\infty,\infty}(f)(x)$.

Since

$$\widehat{f}_k = \frac{1}{\sqrt{P}} \int_0^P f(x) \, e^{-2\pi i k x/P} \, dx,$$

the Fourier transform makes sense for all $f \in L^1_P$. From Bessel's inequality (Proposition 3.4) we know that $\widehat{f} \in c_0$ whenever $f \in L^2_P$, and the result extends.

Theorem 3.16 (Riemann-Lebesgue lemma). *The Fourier transform constitutes a bounded linear operator $\mathscr{F} : L^1_P \to c_0$, so that $\widehat{f}_k \to 0$ as $|k| \to \infty$ when $f \in L^1_P$.*

Proof. As in Gasquet and Witomski [30, 5.1.1 Theorem], integration by parts and periodicity imply that

$$\widehat{f}_k = \tfrac{P}{2\pi i k}\Big(-\tfrac{1}{\sqrt{P}}\big(f(P) - f(0)\big) + \widehat{f'}_k\Big) = \tfrac{P}{2\pi i k}\,\widehat{f'}_k$$

for $k \neq 0$ if $f \in C^1_P$. Hence,

$$\big|\widehat{f}_k\big| = \tfrac{P}{2\pi|k|}\,\big|\widehat{f'}_k\big| \le \tfrac{\sqrt{P}}{2\pi|k|}\,\big\|f'\big\|_{L^1_P} \to 0 \quad \text{as} \quad |k| \to \infty,$$

so that $\mathscr{F} : C^1_P \to c_0$.

When $f \in L^1_P$, we estimate

$$\big|\widehat{f}_k\big| \le \tfrac{1}{\sqrt{P}}\big\|f - g\big\|_{L^1_P} + |\widehat{g}_k|,$$

where $g \in C^1_P$, and use the density of C^1_P in L^1_P from Lemma 3.14 to conclude. □

The property $\widehat{f} \in c_0$ is not enough to guarantee absolute convergence of $S_{m,n}(f)(x)$, but if f is sufficiently well-behaved near x, we do get the inversion formula.

Theorem 3.17 (Dirichlet-Chernoff theorem). *Let $f \in \mathrm{L}^1_P$ and $x \in \mathbb{R}$ be given.*

i) *If the four one-sided limits*

$$f(x^\pm) = \lim_{y\to x^\pm} f(y) \qquad \text{and} \qquad f'(x^\pm) = \lim_{y\to x^\pm} \frac{f(x)-f(y)}{x-y}$$

exist at x, then

$$S_n(f)(x) \to \tfrac{1}{2}\big(f(x^+)+f(x^-)\big) \quad \text{as} \quad n\to\infty; \text{ and}$$

ii) *if f is differentiable at x, then*

$$S_{m,n}(f)(x) \to f(x) \quad \text{as} \quad m,n\to\infty.$$

Proof. We refer to Chernoff [11], but note that case i) follows from the identity

$$S_n(f)(x) = D_n *_P f(x),$$

where

$$D_n(x) = \frac{1}{P}\sum_{|k|\le n} \mathrm{e}^{2\pi i k x/P}$$

is the so-called *Dirichlet kernel*. A slick argument for part ii) rewrites $S_{m,n}(f)(x)$ as a telescoping series and applies the Riemann-Lebesgue lemma. See also Gasquet and Witomski [30, 5.1.4 Theorem]. □

The symmetric partial sums $S_n(f)$ in Theorem 3.17 i) need not converge uniformly to f. In an interval around a jump discontinuity they may oscillate, and the magnitude of the oscillations is about 9 % of the jump size. This feature is captured in the Dirichlet kernel and is known as the Gibbs phenomenon; see Hunter and Nachtergaele [35, Section 7.1].

For $f \in \mathrm{C}^1_P$, however, we do get uniform convergence $S_{m,n}(f) \to f$. This is the content of Corollary 3.21, but more generally we have the following result.

Theorem 3.18 (Uniform convergence). *If $f \in \mathrm{L}^2_P$ and $\widehat{f} \in \ell^1$, then $f \in \mathrm{C}_P$ and*

$$S_{m,n}(f) \to f \quad \text{uniformly (on } \mathbb{R}) \quad \text{as} \quad m,n\to\infty,$$

or more precisely, to its continuous representative.

Proof. We adapt from Gasquet and Witomski [30, 5.3.1 Theorem and 5.3.2 Corollary] and observe that $\{S_{m,n}(f)\}_{m\in\mathbb{N}}$ is a Cauchy sequence in C_P for all $n \in \mathbb{N}$, because

$$\left\|S_{m',n}(f) - S_{m,n}(f)\right\|_\infty \leq \frac{1}{\sqrt{P}} \sum_{k=-m'}^{-m+1} \left|\widehat{f}_k\right| < \frac{1}{\sqrt{P}} \|\widehat{f}\|_{\ell^1} < \infty$$

whenever $m' > m \geq 1$. By the uniform limit theorem, $S_{m,n}(f)$ converges uniformly to $S_{\infty,n}(f) \in \mathrm{C}_P$ as $m \to \infty$. Similarly $\{S_{\infty,n}(f)\}_{n\in\mathbb{N}}$ is Cauchy, so that in total $S_{m,n}(f)$ converges uniformly to $S_{\infty,\infty}(f) \in \mathrm{C}_P$ as $m, n \to \infty$.

Since uniform convergence implies L^2_P-convergence, Theorem 3.15 yields $f = S_{\infty,\infty}(f)$ in L^2_P. Hence, $f \in \mathrm{C}_P$. □

In passing, we mention a deep, striking result by Carleson [10] (for $q = 2$) and Hunt [33] on the pointwise convergence of symmetric Fourier series.

Theorem 3.19 (Carleson-Hunt theorem). *Suppose that $f \in \mathrm{L}^q_P$ for some $1 < q < \infty$. Then*

$$S_n(f)(x) \to f(x) \quad \text{for a.e.} \quad x \in \mathbb{R}.$$ □

We now inspect how the Fourier transform relates to differentiation.

Theorem 3.20 (Differentiation and decay). *Let $f \in \mathrm{C}^n_P$ for some $n \in \mathbb{Z}_+$. Then $\widehat{f^{(j)}} \in \ell^1$ for $j = 0, \ldots, n-1$ and $\widehat{f^{(n)}} \in c_0$, with*

$$\widehat{f^{(j)}}_k = \left(\tfrac{2\pi ik}{P}\right)^j \widehat{f}_k$$

for all $j = 1, \ldots, n$ and $k \in \mathbb{Z}$.

Proof. Partial integration and periodicity give

$$\widehat{f'}_k = \tfrac{1}{\sqrt{P}}\big(f(P) - f(0)\big) + \tfrac{2\pi ik}{P} \widehat{f}_k = \tfrac{2\pi ik}{P} \widehat{f}_k$$

for $j = 1$. Next proceed by induction.

From the Riemann-Lebesgue lemma (Theorem 3.16) we know that $\widehat{f^{(n)}} \in c_0$. Thus

$$\left|\widehat{f^{(n)}}_k\right| \to 0, \qquad \text{implying that} \qquad \left|\widehat{f}_k\right| \asymp_P |k|^{-(n+\epsilon)} \quad \text{as} \quad |k| \to \infty,$$

where $\epsilon > 0$. Hence, if $n - j \geq 1$, we get

$$\left\|\widehat{f^{(j)}}\right\|_{\ell^1} \asymp_P \sum_{k\in\mathbb{Z}} |k|^j \left|\widehat{f}_k\right| \asymp \sum_{k\neq 0} |k|^{-(n-j+\epsilon)} < \infty.$$ □

The consequences of Theorem 3.20 are far-reaching. It shows that $\mathscr{F}$ maps derivatives to mononomials, which gives a powerful recipe for solving constant-coefficient linear differential equations with periodic boundary conditions. The differential equation is first mapped to an algebraic equation on the "Fourier side," which we (try to) solve. Finally, we apply $\mathscr{F}^{-1}$ and obtain a solution to the differential equation.

Another important aspect of Theorem 3.20 is that it relates differentiability with decay of the Fourier coefficients. Combined with Theorem 3.18 we readily deduce the following result; see also Gasquet and Witomski [30, 5.3.1 Theorem].

Corollary 3.21 (Uniform convergence). *Let $r \in \mathbb{Z}_+$ and $f \in \mathrm{C}_P^r$. Then for all $j = 0, \dots, r-1$,*

$$S_{m,n}\big(f^{(j)}\big) \to f^{(j)} \quad \textit{uniformly} \quad \textit{as} \quad m, n \to \infty. \qquad \square$$

Moreover, if $f \in \mathrm{C}_P^\infty$, we get that $\widehat{f^{(n)}} \in c_0$ for all $n \in \mathbb{N}$ by the Riemann-Lebesgue lemma. In particular, the differentiation formula yields $\{k^n \widehat{f}_k\}_{k \in \mathbb{Z}} \in c_0$ for every $n \in \mathbb{N}$.

Definition 3.22 (Schwartz space). The *Schwartz space* $\mathscr{S}(\mathbb{Z})$ consists of all scalar sequences $\{s_k\}_{k\in\mathbb{Z}}$ which *decay/decrease rapidly*, in the sense that $\{k^n s_k\}_{k\in\mathbb{Z}} \in c_0$ for all $n \in \mathbb{N}$. $\square$

Proposition 3.23. *Suppose that $f \in \mathrm{C}_P^\infty$. Then $\widehat{f} \in \mathscr{S}(\mathbb{Z})$. Conversely, if $f \in \mathrm{L}_P^2$ and $\widehat{f} \in \mathscr{S}(\mathbb{Z})$, then $f \in \mathrm{C}_P^\infty$.*

Proof. The first case is true by the preceding paragraph. Conversely, we follow Gasquet and Witomski [30, 5.3.4 Theorem] and notice that $\widehat{f} \in \mathscr{S}(\mathbb{Z})$ implies $\{k^n \widehat{f}_k\}_{k\in\mathbb{Z}} \in \ell^1$ for every $n \in \mathbb{N}$. Thus $\widehat{f} \in \ell^1$, so that

$$f = \frac{1}{\sqrt{P}} \sum_{k\in\mathbb{Z}} \widehat{f}_k \, \mathrm{e}^{2\pi \mathrm{i} k\cdot/P} \in \mathrm{C}_P$$

by Theorem 3.18, where the convergence is uniform. Likewise,

$$\frac{1}{\sqrt{P}} \sum_{k\in\mathbb{Z}} \frac{2\pi \mathrm{i} k}{P} \widehat{f}_k \, \mathrm{e}^{2\pi \mathrm{i} kx/P}$$

converges uniformly to some continuous function, which by

$$\frac{2\pi \mathrm{i} k}{P} \widehat{f}_k \, \mathrm{e}^{2\pi \mathrm{i} kx/P} = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\widehat{f}_k \, \mathrm{e}^{2\pi \mathrm{i} kx/P}\Big)$$

and uniform convergence equals f'. Now iterate the argument to get $f \in \mathrm{C}_P^\infty$. $\square$

Our last concern is the important convolution theorem which displays that $\mathscr{F}$ interchanges convolution and products.

Theorem 3.24 (Convolution theorem). *If $f, g \in \mathrm{L}^1_P$, then*

$$\widehat{f *_P g} = \sqrt{P}\,\widehat{f} \cdot \widehat{g} \quad \text{in} \quad c_0.$$

If $f, g \in \mathrm{L}^2_P$ and $\widehat{f}, \widehat{g} \in \ell^1$, then

$$\widehat{f \cdot g} = \frac{1}{\sqrt{P}}\,\widehat{f} * \widehat{g} \quad \text{in} \quad \ell^2.$$

Proof. Since $f *_P g \in \mathrm{L}^1_P$ by (3.6a), the Riemann-Lebesgue lemma (Theorem 3.16) shows that $\widehat{f *_P g} \in c_0$. The Fubini-Tonelli theorem (Theorem 2.16) and periodicity next yield

$$\begin{aligned}
\big(\widehat{f *_P g}\big)_k &= \frac{1}{\sqrt{P}} \int_0^P \int_0^P f(x-y)\,g(y)\,\mathrm{d}y\,\mathrm{e}^{-2\pi\mathrm{i}kx/P}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{P}} \int_0^P \int_0^P f(x-y)\,\mathrm{e}^{-2\pi\mathrm{i}k(x-y)/P}\,g(y)\,\mathrm{e}^{-2\pi\mathrm{i}ky/P}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{\sqrt{P}} \int_0^P \left(\int_0^P f(x)\,\mathrm{e}^{-2\pi\mathrm{i}kx/P}\,\mathrm{d}x\right) g(y)\,\mathrm{e}^{-2\pi\mathrm{i}ky/P}\,\mathrm{d}y \\
&= \sqrt{P}\,\widehat{f}_k\,\widehat{g}_k.
\end{aligned}$$

As regards the second case, note first that $\widehat{f * g}$ is well-defined by Hölder's inequality, while (2.1) and Young's inequality (2.3) imply that $\widehat{f} * \widehat{g} \in \ell^2$. Inspired by Iorio and Iorio [40, Theorem 3.200], we exploit Theorem 3.18 and calculate

$$\begin{aligned}
\big(\widehat{f \cdot g}\big)_k &= \frac{1}{\sqrt{P}} \int_0^P \left(\frac{1}{\sqrt{P}} \sum_{n\in\mathbb{Z}} \widehat{f}_n\,\mathrm{e}^{2\pi\mathrm{i}nx/P}\right) g(x)\,\mathrm{e}^{-2\pi\mathrm{i}kx/P}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{P}} \sum_{n\in\mathbb{Z}} \widehat{f}_n \frac{1}{\sqrt{P}} \int_0^P g(x)\,\mathrm{e}^{-2\pi\mathrm{i}(k-n)x/P}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{P}} \sum_{n\in\mathbb{Z}} \widehat{f}_n\,\widehat{g}_{k-n} = \frac{1}{\sqrt{P}}\big(\widehat{f} * \widehat{g}\big)_k,
\end{aligned}$$

justified by uniform convergence of the Fourier series. □

3.5 WEAK DERIVATIVES AND PERIODIC SOBOLEV SPACES

In this section we introduce weak differentiation based on the integration by parts formula and the Fourier transform, and also study crucial results for the associated Sobolev spaces.

If $u, \varphi \in C_P^n$, then integration by parts and periodicity show that

$$\int_0^P \varphi\, u^{(j)}\, dx = (-1)^j \int_0^P u\, \varphi^{(j)}\, dx$$

for all $j = 1, \ldots, n$. The key observation here is that the derivatives of u have moved entirely to φ. This motivates the following notion.

Definition 3.25 (Weak derivative). Let $u \in L_P^q$ for some $1 \le q \le \infty$ and suppose that there exist $v_1, \ldots, v_n \in L_P^q$ such that

$$\int_0^P u\, \varphi^{(j)}\, dx = (-1)^j \int_0^P \varphi\, v_j\, dx \quad \text{for all} \quad \varphi \in C_P^\infty \tag{3.10}$$

and $j = 1, \ldots, n$. Then u is called *n times weakly differentiable* and v_j is the *(q-integrable) jth weak derivative* of u, written as $v_j = u^{(j)}$. □

Continuously differentiable functions are by construction weakly differentiable with the weak derivatives equal to the classical ones. In applications it will be evident which type of derivative is considered and so the same notation is not confusing. As suggested, weak derivatives are unique, because if u' and $\widetilde{u'}$ are two derivatives, then

$$\int_0^P \left(u' - \widetilde{u'}\right) \varphi\, dx = 0 \quad \text{for all} \quad \varphi \in C_P^\infty.$$

From du Bois Reymond's lemma (Theorem 2.20) we conclude that $u' = \widetilde{u'}$ in L_P^q. A similar argument yields the usual commutativity $\left(u^{(j)}\right)^{(k)} = u^{(j+k)} = \left(u^{(k)}\right)^{(j)}$. Moreover, in Lemma 6.5 we present a weak version of the chain rule, and for Leibniz' (product) rule we refer to Hunter [34, Proposition 3.21].

By linearity of integrals, weak derivatives satisfy $(u + c\,v)' = u' + c\,v'$. Hence, we can consider new a class of function spaces.

Definition 3.26 (Periodic Sobolev spaces). Let $n \in \mathbb{N}$, $1 \le q \le \infty$ and $P > 0$. The *periodic Sobolev space* $W_P^{n,q}$ consists of all n times weakly differentiable functions $u \in L_P^q$, that is,

$$W_P^{n,q} = \left\{u \in L_P^q : u^{(j)} \in L_P^q \text{ for all } j = 1, \ldots, n\right\}.$$

When $q = 2$, we write H_P^n for $W_P^{n,2}$, also known as *Bessel potential spaces*. □

Clearly C_P^n is a subspace of $W_P^{n,q}$, and the inclusion is strict, as is verified in the following example.

Example 3.27. Let u be the 2-periodic extension of the hat function $\max(1-|\cdot|, 0)$ on $[-1,1]$. Then $u \in \mathrm{W}_2^{1,q}$ for all $1 \leq q \leq \infty$, with weak derivative

$$u'(x) = \begin{cases} 1 & \text{if } x \in [-1,0); \\ -1 & \text{if } x \in (0,1], \end{cases}$$

extended to be 2-periodic. The values at $x = 2j$ for $j \in \mathbb{Z}$ are unimportant. □

Analogously to the norms (3.5) on C_P^n, we can put norms on $\mathrm{W}_P^{n,q}$ using that of L_P^q. Theorem 3.11 i) extends naturally to this new setting. We omit the details, but note that other norms are also possible.

Theorem 3.28. *The periodic Sobolev space $\mathrm{W}_P^{n,q}$ is a Banach space when endowed with the norm*

$$\|u\|_{\mathrm{W}_P^{n,q}} = \sum_{j=0}^{n} \|u^{(j)}\|_{\mathrm{L}_P^q} \asymp_{n,q} \begin{cases} \left(\sum_{j=0}^{n} \|u^{(j)}\|_{\mathrm{L}_P^q}^q\right)^{1/q} & \text{if } 1 \leq q < \infty; \\ \max\limits_{j=0,\ldots,n} \|u^{(j)}\|_{\mathrm{L}_P^\infty} & \text{if } q = \infty. \end{cases}$$

In particular, the inner product

$$\langle u, v\rangle_{\mathrm{H}_P^n} = \sum_{j=0}^{n} \langle u^{(j)}, v^{(j)}\rangle_{\mathrm{L}_P^2}$$

makes H_P^n a Hilbert space. □

Proposition 3.29. *The following continuous embeddings hold.*

$$\mathrm{W}_P^{m,q} \hookrightarrow \mathrm{W}_P^{n,q} \quad \text{if} \quad m > n \qquad \text{and} \qquad \mathrm{W}_P^{n,q} \hookrightarrow \mathrm{W}_P^{n,r} \quad \text{if} \quad q > r.$$

Proof. This is direct from the definition of $\|\cdot\|_{\mathrm{W}_P^{n,q}}$ and Theorem 3.11 iv). □

We now wish to relate weak differentiation with the Fourier transform and first out is a generalization of Theorem 3.20.

Proposition 3.30. *If $u \in \mathrm{W}_P^{n,q}$, then $\widehat{u^{(j)}} \in \ell^1$ for $j = 0, \ldots, n-1$ and $\widehat{u^{(n)}} \in c_0$, with*

$$\widehat{u^{(j)}}_k = \left(\tfrac{2\pi i k}{P}\right)^j \widehat{u}_k$$

for $j = 1, \ldots, n$ and $k \in \mathbb{Z}$.

Proof. The differentiation formula is evident from (3.10) using that $\{\varphi_k\}_{k\in\mathbb{Z}} \subset \mathrm{C}_P^\infty$. Since $u^{(n)} \in \mathrm{L}_P^q \hookrightarrow \mathrm{L}_P^1$, we can estimate exactly as in Theorem 3.20 by the Riemann-Lebesgue lemma (Theorem 3.16). □

When $q = 2$, the differentiation formula even characterizes H^n_P.

Corollary 3.31. *A function $u \in \mathrm{L}^2_P$ is in H^n_P if and only if*

$$u^{(j)} = \mathscr{F}^{-1}\Big(\big(\tfrac{2\pi\mathrm{i}\cdot}{P}\big)^j\,\widehat{u}\Big) \quad \textit{for} \quad j = 1, \dots, n.$$

Proof. If $u \in \mathrm{H}^n_P$, then Proposition 3.30 yields $\big(\frac{2\pi\mathrm{i}\cdot}{P}\big)^j\,\widehat{u} = \widehat{u^{(j)}} \in \ell^1 \hookrightarrow \ell^2$, and we next apply $\mathscr{F}^{-1}\colon \ell^2 \to \mathrm{L}^2_P$ on both sides.

Conversely, Parseval's identity (3.3) and Theorem 3.20 give

$$\begin{aligned}\int_0^P u\,\varphi^{(j)}\,\mathrm{d}x = \big\langle u, \overline{\varphi^{(j)}}\big\rangle_{\mathrm{L}^2_P} &= \sum_{k\in\mathbb{Z}} \widehat{u}_k\,\big(\tfrac{-2\pi\mathrm{i}k}{P}\big)^j\,\overline{\widehat{\varphi}_k}\\ &= (-1)^j\Big\langle \mathscr{F}^{-1}\Big(\big(\tfrac{2\pi\mathrm{i}\cdot}{P}\big)^j\,\widehat{u}\Big), \overline{\varphi}\Big\rangle_{\mathrm{L}^2_P} = (-1)^j \int_0^P \mathscr{F}^{-1}\Big(\big(\tfrac{2\pi\mathrm{i}\cdot}{P}\big)^j\,\widehat{u}\Big)\,\varphi\,\mathrm{d}x,\end{aligned}$$

so that $u^{(j)} = \mathscr{F}^{-1}\Big(\big(\frac{2\pi\mathrm{i}\cdot}{P}\big)^j\,\widehat{u}\Big)$ in L^2_P by uniqueness of weak derivatives. □

According to Corollary 3.31, Theorem 3.28 and the fact that $\mathscr{F}\colon \mathrm{L}^2_P \to \ell^2$ is an isometric isomorphism, equivalent norms on H^n_P are

$$\|u\|^2_{\mathrm{H}^n_P} \eqsim \sum_{k\in\mathbb{Z}}\sum_{j=0}^n \Big|\tfrac{2\pi k}{P}\Big|^{2j}\,|\widehat{u}_k|^2 \eqsim_n \sum_{k\in\mathbb{Z}} \Big(1 + \Big|\tfrac{2\pi k}{P}\Big|^{2n}\Big)\,|\widehat{u}_k|^2, \tag{3.11}$$

where we have used that $\sum_{j=0}^n a^j \eqsim_n 1 + a^n$ for $a \geq 0$. The right-hand side in (3.11) makes sense also for $n \notin \mathbb{N}$, which enables us to generalize differentiation even further via the operator $\big|\frac{\mathrm{d}}{\mathrm{d}x}\big|^s$ defined by

$$\Big(\widehat{\Big|\tfrac{\mathrm{d}}{\mathrm{d}x}\Big|^s u}\Big)_k = \Big|\tfrac{2\pi k}{P}\Big|^s\,\widehat{u}_k \tag{3.12}$$

for any $s \geq 0$.

Definition 3.32. Let $s \geq 0$ and $P > 0$. The *periodic (fractional) Sobolev space* H^s_P consists of all functions $u \in \mathrm{L}^2_P$ for which $\big|\frac{\mathrm{d}}{\mathrm{d}x}\big|^s u \in \mathrm{L}^2_P$. □

The results of Theorem 3.28 extend naturally, so that H^s_P is a Hilbert space with the inner product

$$\langle\!\langle u, v\rangle\!\rangle_{\mathrm{H}^s_P} = \langle u, v\rangle_{\mathrm{L}^2_P} + \Big\langle \Big|\tfrac{\mathrm{d}}{\mathrm{d}x}\Big|^s u, \Big|\tfrac{\mathrm{d}}{\mathrm{d}x}\Big|^s v\Big\rangle_{\mathrm{L}^2_P} = \sum_{k\in\mathbb{Z}} \Big(1 + \Big|\tfrac{2\pi k}{P}\Big|^{2s}\Big)\,\widehat{u}_k\,\overline{\widehat{v}_k}$$

and the norm

$$\|u\|_{\mathrm{H}^s_P} = \bigg(\|u\|^2_{\mathrm{L}^2_P} + \Big\|\Big|\tfrac{\mathrm{d}}{\mathrm{d}x}\Big|^s u\Big\|^2_{\mathrm{L}^2_P}\bigg)^{1/2}.$$

We shall, however, use the equivalent inner product

$$\langle u, v\rangle_{\mathrm{H}^s_P} = \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s}\, \widehat{u}_k\, \overline{\widehat{v}_k} \qquad \text{and the norm} \qquad \|u\|_{\mathrm{H}^s_P} \asymp_s \Big(\sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s} |\widehat{u}_k|^2\Big)^{1/2},$$

where

$$\langle k\rangle_P = \Big(1 + \Big|\tfrac{2\pi k}{P}\Big|^2\Big)^{1/2}$$

is a so-called *Japanese bracket*. The equivalence relies on the fact that

$$(a+b)^s \asymp_s a^s + b^s \tag{3.13}$$

for all $a, b \geq 0$ and $s \geq 0$; see Iorio and Iorio [40, Lemma 3.197].

Since $\mathrm{H}^s_P \subseteq \mathrm{L}^2_P$ constitutes a separable Hilbert space, it is isomorphic to ℓ^2, but there is a more canonical representation.

Definition 3.33 (Sobolev sequence space). Let $s \geq 0$ and $P > 0$ be given. The *Sobolev sequence space* h^s_P is the Hilbert space of all $x = \{x_k\}_{k\in\mathbb{Z}}$ for which $\langle\cdot\rangle^s_P x = \big\{\langle k\rangle^s_P x_k\big\}_{k\in\mathbb{Z}} \in \ell^2$, equipped with the inner product

$$\langle x, y\rangle_{h^s_P} = \Big\langle \langle\cdot\rangle^s_P x, \langle\cdot\rangle^s_P y\Big\rangle_{\ell^2} \qquad \text{and the norm} \qquad \|x\|_{h^s_P} = \big\|\langle\cdot\rangle^s_P x\big\|_{\ell^2}. \qquad \square$$

Proposition 3.34. *Let $s \geq 0$ and $P > 0$. Then the complex exponentials $\{\varphi_k\}_{k\in\mathbb{Z}}$ form an orthonormal basis for H^s_P, so that*

$$\mathscr{F} \colon \mathrm{H}^s_P \to h^s_P$$

is a canonical isomorphism, with $\langle u, v\rangle_{\mathrm{H}^s_P} = \langle \widehat{u}, \widehat{v}\rangle_{h^s_P}$.

Proof. This is immediate from Theorem 3.15 and the construction of h^s_P. $\square$

We now consider some fundamental properties of H^s_P which are very useful in the analysis of differential equations. Case ii) in the following result shows for example that sufficiently smooth functions in the weak sense are classically differentiable.

Theorem 3.35 (Sobolev embedding theorem). *The following continuous embeddings hold.*

i) $\mathrm{H}^t_P \hookrightarrow \mathrm{H}^s_P$ *whenever $t > s \geq 0$; and*

ii) $\mathrm{H}^s_P \hookrightarrow \mathrm{C}_P$ *provided $s > \frac{1}{2}$. The embedding constant has an upper bound which decreases with increasing s and/or P. More generally, $\mathrm{H}^s_P \hookrightarrow \mathrm{C}^n_P$ for $n \in \mathbb{N}$ and $s > n + \frac{1}{2}$.*

Additionally, the embeddings are compact.

Proof. Part i) is immediate from Definition 3.32 and $\langle k\rangle^t_P \leq \langle k\rangle^s_P$ for $t > s$. As regards ii), let $u \in \mathrm{H}^s_P$ and observe by the Cauchy-Schwarz inequality that

$$\|\widehat{u}\|_{\ell^1} = \sum |\widehat{u}_k| \le \Big(\sum \langle k\rangle_P^{-2s}\Big)^{1/2} \Big(\sum \langle k\rangle_P^{2s}\, |\widehat{u}_k|^2\Big)^{1/2} \lesssim \|u\|_{\mathrm{H}^s_P} < \infty, \tag{3.14}$$

because $\sum \langle k\rangle_P^{-2s} < \infty$ as $2s > 1$. Thus $\widehat{u} \in \ell^1$, so that $u \in \mathrm{C}_P$ by Theorem 3.18. Moreover,

$$\|u\|_\infty \le \tfrac{1}{\sqrt{P}} \|\widehat{u}\|_{\ell^1} \le M_{s,P} \|u\|_{\mathrm{H}^s_P},$$

where

$$M_{s,P}^2 = \frac{1}{P} \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{-2s}.$$

This gives $\mathrm{H}^s_P \hookrightarrow \mathrm{C}_P$, and it is clear that $M_{s,P}$ decreases with increasing s. Using the integral test for convergence,

$$\int_0^\infty f(k)\,\mathrm{d}k \le \sum_{k=0}^\infty f(k) \le f(0) + \int_0^\infty f(k)\,\mathrm{d}k,$$

for $f \ge 0$ continuous and monotonically decreasing, we put $f = \frac{1}{P}\langle\cdot\rangle_P^{-2s}$ and find that

$$M_{s,P}^2 \in \left[M_s - \frac{1}{P}, M_s + \frac{1}{P}\right], \qquad \text{where} \qquad M_s = \frac{1}{\pi}\int_0^\infty \frac{\mathrm{d}x}{\big(1+x^2\big)^s} < \infty$$

again since $s > \frac{1}{2}$. Hence, $M_{s,P}$ has an upper bound which decreases with increasing P. Similar arguments apply for the general case.

We refer to Taylor [62, (3.12) and Proposition 3.4] for the compactness of i), which is a consequence of $[0,P]$ being compact. This yields compactness of ii), because if $r \in \big(n + \frac{1}{2}, s\big)$, then $\mathrm{H}^s_P \hookrightarrow \mathrm{H}^r_P$ is compact and $\mathrm{H}^r_P \hookrightarrow \mathrm{C}^n_P$ is continuous. □

Theorem 3.36 (Banach algebra). *If $s > \frac{1}{2}$, then H^s_P is a Banach algebra under pointwise multiplication, with*

$$\|uv\|_{\mathrm{H}^s_P} \lesssim_{s,P} \|u\|_{\mathrm{H}^s_P} \|v\|_{\mathrm{H}^s_P}.$$

The estimation constant has an upper bound which decreases with increasing P.

Proof. Based on Iorio and Iorio [40, Theorem 3.200], we first notice from (3.13) that

$$\langle k\rangle_P^s \lesssim_s \langle k-n\rangle_P^s + \langle n\rangle_P^s$$

for all $k, n \in \mathbb{Z}$. Moreover, (3.14) yields $\widehat{u}, \widehat{v} \in \ell^1$ because $s > \frac{1}{2}$, and we can apply the convolution theorem (Theorem 3.24) to get

$$\begin{aligned}
\langle k\rangle_P^s\, |\widehat{uv}_k| &= \frac{1}{\sqrt{P}} \sum_{n\in\mathbb{Z}} \langle k\rangle_P^s |\widehat{u}_{k-n}|\, |\widehat{v}_n| \\
&\lesssim_s \frac{1}{\sqrt{P}} \sum_{n\in\mathbb{Z}} \Big(\langle k-n\rangle_P^s |\widehat{u}_{k-n}|\, |\widehat{v}_n| + |\widehat{u}_{k-n}|\, \langle n\rangle_P^s |\widehat{v}_n|\Big) \\
&= \frac{1}{\sqrt{P}} \Big[\Big(\big|\langle\cdot\rangle_P^s \widehat{u}\big| * |\widehat{v}|\Big)_k + \Big(|\widehat{u}| * \big|\langle\cdot\rangle_P^s \widehat{v}\big|\Big)_k\Big].
\end{aligned}$$

By assumption, $\widehat{u}, \widehat{v} \in h_P^s$, or equivalently, $\left|\langle\cdot\rangle_P^s \widehat{u}\right|, \left|\langle\cdot\rangle_P^s \widehat{v}\right| \in \ell^2$. Hence, Young's inequality (2.3) implies that

$$\begin{aligned} \|uv\|_{\mathrm{H}_P^s}^2 &\lesssim_s \frac{2}{P}\left(\left\|\left|\langle\cdot\rangle_P^s \widehat{u}\right| * |\widehat{v}|\right\|_{\ell^2}^2 + \left\||\widehat{u}| * \left|\langle\cdot\rangle_P^s \widehat{v}\right|\right\|_{\ell^2}^2\right) \\ &\leq \frac{2}{P}\left(\|\widehat{u}\|_{h_P^s}^2 \|\widehat{v}\|_{\ell^1}^2 + \|\widehat{u}\|_{\ell^1}^2 \|\widehat{v}\|_{h_P^s}^2\right). \end{aligned}$$

We now use estimate (3.14) and $\|\widehat{\cdot}\|_{h_P^s} = \|\cdot\|_{\mathrm{H}_P^s}$ to conclude that

$$\|uv\|_{\mathrm{H}_P^s} \leq \widetilde{M}_{s,P} \|u\|_{\mathrm{H}_P^s} \|v\|_{\mathrm{H}_P^s},$$

where $\widetilde{M}_{s,P} \lesssim_s 2M_{s,P}$ and $M_{s,P}$ is as in the proof of Theorem 3.35 ii). □

Theorem 3.37 (Sobolev interpolation inequality). *Suppose that $0 \leq r < s < t < \infty$, where $s = \theta r + (1-\theta)t$ for some interpolation exponent $\theta \in (0,1)$. If $u \in \mathrm{H}_P^t$, then u satisfies the estimate*

$$\|u\|_{\mathrm{H}_P^s} \leq \|u\|_{\mathrm{H}_P^r}^{\theta} \|u\|_{\mathrm{H}_P^t}^{1-\theta}. \tag{3.15}$$

Proof. By writing $\left|\langle k\rangle_P^s \widehat{u}_k\right| = \left|\langle k\rangle_P^r \widehat{u}_k\right|^{\theta} \cdot \left|\langle k\rangle_P^t \widehat{u}_k\right|^{1-\theta}$, Hölder's inequality implies that

$$\sum \left|\langle k\rangle_P^s \widehat{u}_k\right|^2 \leq \left(\sum \left|\langle k\rangle_P^r \widehat{u}_k\right|^2\right)^{2\theta} \left(\sum \left|\langle k\rangle_P^t \widehat{u}_k\right|^2\right)^{2(1-\theta)},$$

which becomes (3.15); see also Shkoller [58, Theorem 5.31]. □

Up till now H_P^s has been characterized in terms of the Fourier transform. We end our discussion with an equivalent description using difference quotients. Due to its more local appearance, this version is sometimes easier to work with.

Theorem 3.38 (Sobolev–Slobodeckij definition of H_P^s). *Suppose $s \notin \mathbb{N}$ and write $s = m + \sigma$, where $m = \lfloor s \rfloor \in \mathbb{N}$ and $\sigma = s - \lfloor s \rfloor \in (0,1)$. If we define the Sobolev–Slobodeckij or Gagliardo semi-norm*

$$|u|_{\mathrm{H}_P^\sigma} = \left(\int_0^P \int_0^P \frac{|u(x) - u(y)|^2}{|x-y|^{1+2\sigma}} \, \mathrm{d}x \, \mathrm{d}y\right)^{1/2},$$

then an equivalent norm on H_P^s is

$$\|u\|_{\mathrm{H}_P^s} \eqsim_s \left(\|u\|_{\mathrm{H}_P^m}^2 + \left|u^{(m)}\right|_{\mathrm{H}_P^\sigma}^2\right)^{1/2}, \tag{3.16}$$

and the equivalence does not depend on P.

Proof. We refer to an argument by Rosenzweig [51] for 1-periodic functions, which extends to the P-periodic setting. □

Chapter 4

The Fourier transform on $\mathbb{R}$ and beyond

Whereas the Fourier series from Theorem 3.7 is a way to represent periodic functions, we would like to have a similar approach also in the aperiodic case. Heuristically, as was done by Emerton [26], let $f:\mathbb{R}\to\mathbb{F}$ be a function which is restricted to the interval $\left[-\frac{P}{2},\frac{P}{2}\right]$, where $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$. Then f has the representation

$$f(x)=\frac{1}{\sqrt{P}}\sum_{k\in\mathbb{Z}}\widehat{f}_k\,\mathrm{e}^{2\pi\mathrm{i}kx/P}$$

in $\left[-\frac{P}{2},\frac{P}{2}\right]$. By introducing $\xi=2\pi k/P$ and $\Delta\xi=2\pi/P$, the Fourier series of f may be written as

$$f(x)=\frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\left(\frac{1}{\sqrt{2\pi}}\int_{-\frac{P}{2}}^{\frac{P}{2}}f(x)\,\mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}x\right)\mathrm{e}^{\mathrm{i}\xi x}\,\Delta\xi.$$

If we now, formally, let $P\to\infty$, so that $\Delta\xi\to 0$, define

$$\widehat{f}(\xi)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}f(x)\,\mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}x \tag{4.1}$$

and view the above sum as a Riemann sum, then suggestively

$$f(x)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\widehat{f}(\xi)\,\mathrm{e}^{\mathrm{i}\xi x}\,\mathrm{d}\xi \tag{4.2}$$

on all of $\mathbb{R}$. Hence, if we define the *Fourier transform* $\mathscr{F}(f)=\widehat{f}$ of an aperiodic function $f:\mathbb{R}\to\mathbb{F}$ by (4.1), then (4.2) resembles the Fourier inversion formula (3.9) for periodic functions. This analysis was not rigorous, but the representation is in fact valid a.e. for $\mathrm{L}^1(\mathbb{R})$ functions; see Theorem 4.3.

In Section 4.1 we briefly study the Fourier transform on $\mathrm{L}^1(\mathbb{R})$, the Schwartz space $\mathscr{S}(\mathbb{R})$ and its extension to $\mathrm{L}^2(\mathbb{R})$. Section 4.2 next outlines generalizations of the Fourier transform to other spaces, and at last comes Section 4.3 which introduces so-called Fourier multiplier operators.

No proofs will be given, but we remark that many of the arguments from Chapter 3 extend to this new setting. The work by Gasquet and Witomski [30] is an in-depth reference on the Fourier transform of functions and distributions. Moreover, both Demengel and Demengel [15, Chapters 2 and 4], Shkoller [58] and Iorio and Iorio [40] provide details on Fourier methods and Sobolev spaces. See also McDonald and Weiss [47].

4.1 THE FOURIER TRANSFORM ON $\mathrm{L}^1(\mathbb{R})$, THE SCHWARTZ SPACE $\mathscr{S}(\mathbb{R})$ AND $\mathrm{L}^2(\mathbb{R})$

All results in this section are stated freely, but can be found in the book by Gasquet and Witomski [30].

We first note that (4.1) makes sense if and only if $f \in \mathrm{L}^1(\mathbb{R})$, and the "periodic" Riemann-Lebesgue lemma (Theorem 3.16) extends to $\mathrm{L}^1(\mathbb{R})$.

Theorem 4.1 (Riemann-Lebesgue lemma). *The Fourier transform is a bounded linear operator $\mathscr{F}: \mathrm{L}^1(\mathbb{R}) \to \mathrm{BUC}(\mathbb{R})$ satisfying $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$ when $f \in \mathrm{L}^1(\mathbb{R})$.* □

The Fubini-Tonelli theorem (Theorem 2.16) yields the relation

$$\int_{\mathbb{R}} f(x)\,\widehat{g}(x)\,\mathrm{d}x = \int_{\mathbb{R}} \widehat{f}(x)\,g(x)\,\mathrm{d}x \tag{4.3}$$

for $f, g \in \mathrm{L}^1(\mathbb{R})$, called the "change of hats formula." We next list a collection of basic properties which follow either directly or using integration by parts.

Theorem 4.2 (Basic operations with the Fourier transform). *The Fourier transform on $\mathrm{L}^1(\mathbb{R})$ obeys the following operations.*

i) Translation, multiplication by exponentials and scaling:

$$\begin{aligned} f(x-a) &\overset{\mathscr{F}}{\mapsto} \mathrm{e}^{-\mathrm{i}a\xi}\widehat{f}(\xi); \\ \mathrm{e}^{\mathrm{i}ax} f(x) &\overset{\mathscr{F}}{\mapsto} \widehat{f}(\xi - a); \\ f(ax) &\overset{\mathscr{F}}{\mapsto} \frac{1}{|a|}\widehat{f}\left(\frac{\xi}{a}\right), \quad a \neq 0. \end{aligned}$$

ii) Conjugation and parity:

$$\begin{aligned} f \text{ even (odd)} &\overset{\mathscr{F}}{\mapsto} \widehat{f} \text{ even (odd)}; \\ f \text{ real and even} &\overset{\mathscr{F}}{\mapsto} \widehat{f} \text{ real and even}; \\ f \text{ real and odd} &\overset{\mathscr{F}}{\mapsto} \widehat{f} \text{ imaginary and odd}. \end{aligned}$$

iii) *Differentiation: If* $f \in \mathrm{C}^n(\mathbb{R}) \cap \mathrm{L}^1(\mathbb{R})$ *with* $f^{(j)} \in \mathrm{L}^1(\mathbb{R})$ *for all* $j = 1, \ldots, n$, *then*

$$f^{(j)}(x) \quad \overset{\mathscr{F}}{\mapsto} \quad (\mathrm{i}\xi)^j \widehat{f}(\xi).$$

iv) *Multiplication by monomials: If* $(x \mapsto x^j f(x)) \in \mathrm{L}^1(\mathbb{R})$ *for* $j = 0, 1, \ldots, n$, *then* $\widehat{f}$ *is in* $\mathrm{BUC}^n(\mathbb{R})$ *and*

$$x^j f(x) \quad \overset{\mathscr{F}}{\mapsto} \quad \mathrm{i}^j \widehat{f}^{(j)}(\xi).$$ □

Analogously to Fourier series, Theorem 4.2 iii) and iv) display the usefulness of $\mathscr{F}$ in the analysis of differential equations in that derivatives are mapped to monomials. Moreover, we can invert the operation.

Theorem 4.3 (Fourier inversion formula). *If both f and $\widehat{f}$ are in $\mathrm{L}^1(\mathbb{R})$, then*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(\xi) \mathrm{e}^{\mathrm{i}\xi x} \, \mathrm{d}\xi \quad \textit{for a.e.} \quad x \in \mathbb{R}.$$ □

It is not always the case that $\widehat{f}$ is integrable whenever f is. But if $f \in \mathrm{C}^2(\mathbb{R})$, with f, f' and f'' in $\mathrm{L}^1(\mathbb{R})$, we do have $\widehat{f} \in \mathrm{L}^1(\mathbb{R})$.

We next consider how the Fourier transform relates to decay. If $f \in \mathrm{L}^1(\mathbb{R})$ has compact support, then $\widehat{f}$ is in $\mathrm{BUC}^\infty(\mathbb{R})$. In particular, this applies to integrable functions which are *rapidly decaying/decreasing*, in the sense that

$$\sup_{x \in \mathbb{R}} \left| x^\alpha f(x) \right| < \infty \quad \text{for all} \quad \alpha \in \mathbb{N}.$$

Conversely, if $f \in \mathrm{C}^\infty(\mathbb{R})$ satisfies $f^{(n)} \in \mathrm{L}^1(\mathbb{R})$ for all $n \in \mathbb{N}$, then $\widehat{f}$ decays rapidly. Similarly as Definition 3.22, we can then introduce the following space.

Definition 4.4 (Schwartz space). The *Schwartz space* $\mathscr{S}(\mathbb{R})$ is the a set of all $f \in \mathrm{C}^\infty(\mathbb{R})$ for which f and all of its derivatives decay rapidly. That is,

$$\mathscr{S}(\mathbb{R}) = \left\{ f \in \mathrm{C}^\infty(\mathbb{R}) : \|f\|_{\alpha,\beta} < \infty \text{ for all } \alpha, \beta \in \mathbb{N} \right\},$$

where $\|f\|_{\alpha,\beta} = \sup_{x \in \mathbb{R}} \left| x^\alpha f^{(\beta)}(x) \right|$. □

The Schwartz space constitutes a complete topological vector space via $\|\cdot\|_{\alpha,\beta}$, which define a family of seminorms. Additionally, the inclusions

$$\mathscr{D}(\mathbb{R}) \subset \mathscr{S}(\mathbb{R}) \subset \mathrm{BUC}^\infty(\mathbb{R}) \cap \mathrm{L}^p(\mathbb{R})$$

hold for all $1 \le p < \infty$, so that $\mathscr{S}(\mathbb{R})$ is dense in $\mathrm{L}^p(\mathbb{R})$ by Proposition 2.12.

It is readily seen that $\mathscr{S}(\mathbb{R})$ is closed both under multiplication by polynomials and differentiation, but more is true.

Theorem 4.5. *The Fourier transform is a continuous automorphism $\mathscr{F}: \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$.* □

In particular, $\mathscr{F}$ is a valid mapping $\mathscr{S}(\mathbb{R}) \to \mathrm{L}^2(\mathbb{R})$, so that density of $\mathscr{S}(\mathbb{R})$ and completeness of $\mathrm{L}^2(\mathbb{R})$ imply the following celebrated result.

Theorem 4.6 (Plancherel's theorem). *The Fourier transform extends uniquely to a unitary operator $\mathscr{F}: \mathrm{L}^2(\mathbb{R}) \to \mathrm{L}^2(\mathbb{R})$. Especially,*

$$\langle \widehat{f}, \widehat{g} \rangle_{\mathrm{L}^2(\mathbb{R})} = \langle f, g \rangle_{\mathrm{L}^2(\mathbb{R})} \quad \textit{for all} \quad f, g \in \mathrm{L}^2(\mathbb{R}).$$ □

Plancherel's theorem is an analog of Theorem 3.15 and Parseval's identity (3.3). Moreover, the Fourier transforms on $\mathrm{L}^1(\mathbb{R})$ and $\mathrm{L}^2(\mathbb{R})$ coincide on $\mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$.

Switching to the interplay between $\mathscr{F}$ and convolution, note first that the convolution is continuous operator

$$*: \mathscr{S}(\mathbb{R}) \times \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R}),$$

which then defines a product structure on the Schwartz space. As we saw for Fourier series, $\mathscr{F}$ interchanges convolution and products. Arguing with help of Proposition 2.18, this is still the case.

Theorem 4.7 (Convolution theorem). *For $f, g \in \mathrm{L}^1(\mathbb{R})$ we have that*

$$\widehat{f * g} = \sqrt{2\pi}\, \widehat{f} \cdot \widehat{g}. \tag{4.5}$$

If additionally $\widehat{f}, \widehat{g} \in \mathrm{L}^1(\mathbb{R})$, then

$$\widehat{f \cdot g} = \sqrt{2\pi}\, \widehat{f} * \widehat{g}. \tag{4.6}$$

Thus (4.5) and (4.6) hold for all $f, g \in \mathscr{S}(\mathbb{R})$. Moreover, if $f, g \in \mathrm{L}^2(\mathbb{R})$, then

$$f * g = \sqrt{2\pi}\, \mathscr{F}^{-1}\big(\widehat{f} \cdot \widehat{g}\big) \qquad \textit{and} \qquad \widehat{f \cdot g} = \sqrt{2\pi}\, \widehat{f} * \widehat{g}.$$

All relations are valid pointwise. □

The first convolution formula in $\mathrm{L}^2(\mathbb{R})$ is not completely satisfying. This is because $f * g$ is in $\mathrm{BC}(\mathbb{R})$ from (2.2b), but not necessarily in $\mathrm{L}^2(\mathbb{R})$. As such, we cannot apply the Fourier transform on $f * g$. There is, however, a remedy if we extend $\mathscr{F}$ further. This is the content of Section 4.2.

4.2 AN OUTLINE OF FOURIER TRANSFORMS ON OTHER SPACES

Up till now we have studied the Fourier transform of functions $f : \mathbb{R} \to \mathbb{F}$ which are sufficiently integrable and either periodic or aperiodic. There exists, however, a unified, extended theory if we switch to the so-called *tempered distributions*. These are continuous linear functionals on $\mathscr{S}(\mathbb{R})$, where the continuity is with respect to the topology on $\mathscr{S}(\mathbb{R})$ defined via the seminorms $\|\cdot\|_{\alpha,\beta}$. The Fourier transform of a tempered distribution $T \in \mathscr{S}'(\mathbb{R})$ is then constructed by duality as the functional $\mathscr{F}(T) = \widehat{T}$ given by

$$\widehat{T}(\varphi) = T(\widehat{\varphi}) \quad \text{for all} \quad \varphi \in \mathscr{S}(\mathbb{R}).$$

This resembles the change of hats formula (4.3), and from Theorem 4.5 we obtain the following result.

Theorem 4.8. *The Fourier transform extends to a continuous automorphism*

$$\mathscr{F} : \mathscr{S}'(\mathbb{R}) \to \mathscr{S}'(\mathbb{R}). \qquad \square$$

It can be shown that $L^1_{\text{loc}}(\mathbb{R}) \subset \mathscr{S}'(\mathbb{R})$, so that $\mathscr{S}'(\mathbb{R})$ covers all the previously defined function spaces. There exist, however, also more exotic elements in $\mathscr{S}'(\mathbb{R})$, of which the Dirac delta distribution is one example. Furthermore, the Fourier transform on $\mathscr{S}'(\mathbb{R})$ has many good properties. In particular, the convolution formula $\widehat{f * g} = \sqrt{2\pi}\,\widehat{f} \cdot \widehat{g}$ holds in $\mathscr{S}'(\mathbb{R})$ for $f, g \in L^2(\mathbb{R})$. We note that all the above features can be found in the work by Gasquet and Witomski [30].

Another natural generalization of the Fourier transform is to replace the domains $\mathbb{R}$ and $\mathbb{T}$ with the higher-dimensional variants $\mathbb{R}^n$ and $\mathbb{T}^n$ for $n \in \mathbb{Z}_+$. The theory is straightforward; we omit the details.

More interestingly, we can define the Fourier transform on measures or on locally compact abelian groups; see McDonald and Weiss [47, Section 15.4] and Rudin [52]. In the latter theory, the torus $\mathbb{T}^n$ is an example of the compact case, while the Euclidean space $\mathbb{R}^n$ illustrates the non-compact locally compact situation.

4.3 AN INTRODUCTION TO FOURIER MULTIPLIER OPERATORS

As we know from both Proposition 3.30 and Theorem 4.2 iii), the Fourier transform maps derivatives to monomials, so that

$$\left(\widehat{u^{(n)}}\right)_k = \left(\tfrac{2\pi i k}{P}\right)^n \widehat{u}_k \qquad \text{or} \qquad \widehat{u^{(n)}}(\xi) = (i\xi)^n\, \widehat{u}(\xi).$$

From (3.12) we can also consider the fractional-order differentiation operator $\left|\frac{\mathrm{d}}{\mathrm{d}x}\right|^s u$ defined by

$$\left(\widehat{\left|\tfrac{\mathrm{d}}{\mathrm{d}x}\right|^s u}\right)_k = \left|\tfrac{2\pi k}{P}\right|^s \widehat{u}_k \qquad \text{or} \qquad \left(\widehat{\left|\tfrac{\mathrm{d}}{\mathrm{d}x}\right|^s u}\right)(\xi) = |\xi|^s\, \widehat{u}(\xi),$$

where $s \geq 0$. Both situations are examples of a *Fourier multiplier operator* $T = T_m$, which is defined implicitly by its action on the Fourier side, that is,

$$\widehat{Tu}(\xi) = m(\xi)\widehat{u}(\xi)$$

for some function $m\colon \mathbb{R} \to \mathbb{F}$. We focus on the Euclidean case, but note that $m = \{m_k\}_{k\in\mathbb{Z}}$ is an $\mathbb{F}$-valued sequence when $u\colon \mathbb{T} \to \mathbb{F}$. The function m is called the *multiplier* or *symbol* associated with T and allows for a wide range actions. For example, the kernel K_{h_0} (1.4) in the Whitham equation (1.3) gives rise to the multiplier operator L defined by

$$\widehat{Lu} = c_{\text{Euler}}\, \widehat{u},$$

and L can be written as $Lu = K_{h_0} * u$ (formally) using the convolution theorem.

The study of Fourier multiplier operators examines in broad terms how symbols affect mapping properties of the operators. This is contained in the subject of *microlocal analysis*; see Grigis and Sjöstrand [32] and [68]. For instance, Plancherel's theorem (Theorem 4.6) yields that $m \in \mathrm{L}^\infty(\mathbb{R})$ is a sufficient condition for the validity of the multiplier operator $T_m\colon \mathrm{L}^2(\mathbb{R}) \to \mathrm{L}^2(\mathbb{R})$. Moreover, we mention that commutativity of multiplier operators is evident from

$$\widehat{STu} = m_S\, \widehat{Tu} = m_S\, m_T\, \widehat{u} = m_T\, \widehat{Su} = \widehat{TSu},$$

where m_S and m_T are the $\mathbb{F}$-valued symbols associated with S and T, and we have utilized the commutativity of $\mathbb{F}$.

Chapter 5

The calculus of variations

5.1 Introduction

The field calculus of variations has a long history in mathematics and shares a close connection with classical mechanics. Its importance and development throughout the centuries are intimately related to many famous problems, such as the brachistochrone curve problem, isoperimetric problems, the Fermat principle in geometrical optics, minimal surface problems and the Dirichlet principle to name a few. In the 18th century, Euler and Lagrange introduced a systematic procedure for tackling problems in variational calculus through the so-called Euler-Lagrange equation. From the work of Weierstrass in the 19th century these methods then emerged into a more rigorous theory, and important contributors in the 20th century include Hilbert, Noether, Tonelli, Lebesgue and Hadamard. Today variational calculus finds widespread applications in the study of differential equations and constitutes for example a key component in the success of finite element methods in numerical mathematics. The calculus of variations shares intimate links to optimization theory and functional analysis, and it is also vital to mention its modern and extending companion known as optimal control theory.

In essence the calculus of variations deals with optimization of functionals and the challenge of finding the associated extrema (see item i) below). This extends the well-known techniques in ordinary calculus and has nowadays simple, abstract formulations. To this end, let $\mathscr{E}\colon C \to \mathbb{R}$ be a functional defined on a subset C of a normed space. We then study the following abstract optimization problem.

i) Prove the existence of a *local extreme point (minimizer/maximizer) $\overline{x}$ for $\mathscr{E}$ on C*, that is, an $\overline{x} \in C$ satisfying

$$\mathscr{E}(\overline{x}) = \inf_{x \in C \cap U} \mathscr{E}(x) \qquad \text{or} \qquad \mathscr{E}(\overline{x}) = \sup_{x \in C \cap U} \mathscr{E}(x),$$

where U is a neighborhood of $\overline{x}$. When there is just one extreme point in U, it is called *strict* or *unique*. If $U = C$, then $\overline{x}$ is a *global extreme point (minimizer/maximizer) for $\mathscr{E}$ on C*; and

ii) characterize and possibly find explicit expressions for such point(s) $\overline{x}$.

Convention. In this chapter X, Y and Z denote generic normed spaces, while U, V and W symbolize open sets. If $Y = \mathbb{R}$ or $Z = \mathbb{R}$, relevant in functional optimization, then both X, Y and Z are assumed to be real normed spaces. □

All maximization problems can equivalently be formulated as minimization problems by relabeling $\mathscr{E}$ into $-\mathscr{E}$, and we consider primarily the latter cases.

From classical calculus in $\mathbb{R}^n$ we recall that there are several optimality conditions for local extrema based on information about the gradient and the Hessian. These ideas extend to the abstract setting, and in Section 5.2 we shall for example see that interior local extreme points $\overline{x}$ for $\mathscr{E}$ are *critical points*, that is, $\mathscr{E}'(\overline{x}) = 0$, as expected. Suggestively $\mathscr{E}'$ denotes the derivative of $\mathscr{E}$, to be interpreted in the sense of Gâteaux (Definition 5.1) and Fréchet (Definition 5.2) for general normed spaces.

Many applications seek to minimize integral functionals $\mathscr{E}\colon C \to \mathbb{R}$ of the form

$$\mathscr{E}(u) = \int_\Omega f\big(x, u(x), u'(x)\big)\,\mathrm{d}x, \tag{5.1}$$

where

$$f : \Omega \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (x, y, z) \mapsto f(x, y, z)$$

is a given function, $\Omega \subset \mathbb{R}$ is bounded and open and C constitutes a subset of a Sobolev space of functions $u\colon \Omega \to \mathbb{R}$. Often $\mathscr{E}$ defines some kind of energy in physics and is commonly known as an "energy functional." Based on du Bois Reymond's lemma (Theorem 2.20) and Section 5.2 it can be shown that critical points for $\mathscr{E}$ must satisfy

$$\frac{\partial f}{\partial y}\big(x, u(x), u'(x)\big) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial z}\big(x, u(x), u'(x)\big)\right) = 0 \quad \text{for} \quad x \in \Omega,$$

along with boundary conditions inherited from C. This second-order, elliptic ODE is known as the *Euler-Lagrange equation* associated with the variational problem, and in higher dimensions and/or several variables it becomes a system of ODEs and/or PDEs. The minimization problem is thus transferred to solving a differential equation plus an analysis of properties of f which guarantee that critical points are local minimizers.

Solving the Euler-Lagrange equation, however, can be a difficult or even impossible task. But by flipping the viewpoint we can use the above idea as follows (Yan [69, Section 4.1]). Let $\mathscr{A}\colon X \to X'$ denote an abstract differential operator and consider the differential problem

$$\mathscr{A}(u) = f, \tag{5.2}$$

where $f \in X'$ is some given datum. If we can find a differentiable functional $\mathscr{E}\colon X \to \mathbb{R}$ with $\mathscr{E}' = \mathscr{A} - f$, then interior local extreme points of $\mathscr{E}$ satisfy the Euler-Lagrange equation (5.2). Differential problems admitting such a *variational formulation* are called *variational* or *potential problems*.

A standard approach for proving the existence of extrema is the so-called *direct method in the calculus of variations*. This is an extension of Weierstrass' extreme value theorem and will be a key focus in Section 5.3. Other useful techniques, not discussed here, include the energy method of Lax-Milgram and the mountain pass method for saddle-point critical points; see Yan [69, Introduction] for examples.

In many problems it is natural to optimize functionals under various constraints. Just as in standard calculus, these situations are connected to the *method of Lagrange multipliers*, which will be our final concern in Section 5.4.

The calculus of variations is a broad subject and we only scratch the surface. Introductory aspects of variational methods can be found in the monographs by Dacorogna [14] and Troutman [64], both of which contain numerous examples along with historical comments. Other mathematical resources include the texts by Evans [1, Chapter 8], Clarke [12], Ekeland and Témam [25], Luenberger [46, Chapters 7-9] and Yan [69], but a more physics-oriented book is given by Gelfand and Fomin [31]. For advanced topics the reader is referred to Dacorogna [13] and Struwe [61].

5.2 DIFFERENTIABILITY IN NORMED SPACES AND OPTIMALITY CONDITIONS FOR LOCAL EXTREMA

In this section we look at the notion of differentiability of operators $X \to Y$ and first-order optimality characterizations of local extrema. We start with the *Gâteaux derivative*, which generalizes the directional derivative in multivariable calculus, and then focus on the more restrictive *Fréchet derivative*.

Definition 5.1 (Gâteaux derivative). A function $f : U \subseteq X \to V \subseteq Y$ is said to be *Gâteaux differentiable at $x \in U$* if there exists $g_x : X \to Y$ such that

$$g_x(h) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon h) - f(x)}{\epsilon} = \frac{\mathrm{d}}{\mathrm{d}\epsilon} f(x+\epsilon h)\Big|_{\epsilon=0} \tag{5.3}$$

for all $h \in X$, where $\epsilon \in \mathbb{F} \setminus \{0\}$ and $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is the scalar field of X and Y. The operator g_x, written $f'(x)$, is called the *Gâteaux derivative of f at x*. Moreover, f is referred to as *Gâteaux differentiable on U* if it is differentiable at every $x \in U$. □

The conditions in Definition 5.1 can be relaxed considerably; it suffices merely that X and Y are vector spaces. Our choice, however, provides enough flexibility for this text. The limit in (5.3) is well-defined since $x + \epsilon h \in U$ for all sufficiently small $|\epsilon|$. Observe also that the Gâteaux derivative may be a nonlinear map.

Definition 5.2 (Fréchet derivative and C^k-function). A map $f : U \subseteq X \to V \subseteq Y$ is said to be *Fréchet differentiable at* $x \in U$ if there exists $T_x \in \mathscr{B}(X,Y)$, called the *Fréchet derivative of f at x* and written as $f'(x)$, such that

$$\frac{\|f(x+h) - f(x) - f'(x)h\|_Y}{\|h\|_X} \to 0 \qquad \text{as} \qquad \|h\|_X \to 0 \tag{5.4}$$

for all $h \in X \setminus \{0\}$ at which $x + h \in U$. Equivalently,

$$f(x+h) = f(x) + f'(x)h + o(\|h\|_X). \tag{5.5}$$

If the map

$$f' : U \to \mathscr{B}(X,Y), \qquad x \mapsto f'(x),$$

named the *Fréchet derivative of f*, is continuous in the operator norm, then f is said to be a *C^1-function* and we write $f \in \mathrm{C}^1(U,V)$. Now, if possible, iterate the process to obtain higher-order derivatives

$$\begin{aligned} f'' &: U \to \mathscr{B}(X, \mathscr{B}(X,Y)) \\ &\vdots \\ f^{(k)} &: U \to \mathscr{B}(X, \mathscr{B}(X, \cdots, \mathscr{B}(X,Y)) \cdots), \end{aligned}$$

so that f is a *C^k-function*, symbolized as $f \in \mathrm{C}^k(U,V)$. □

Both types of derivatives are unique and satisfy the usual linearity of differentiation. The common notation $f'(x)$ is unproblematic as the type will be clear from the context, and we also have the following result directly from (5.5).

Proposition 5.3. *If $f : U \subseteq X \to V \subseteq Y$ is Fréchet differentiable at $x \in U$, then it is also Gâteaux differentiable there and the derivatives agree.* □

The converse is in general false. Nevertheless, (5.3) is usually the starting point in the process of calculating either derivative.

Example 5.4. i) If $f \in \mathscr{B}(X,Y)$, then $f'(x) = f$ and evidently $f \in \mathrm{C}^k(X,Y)$ for all $k \in \mathbb{Z}_+$. For constant operators the Fréchet derivative equals the zero map.

ii) In finite dimensional spaces the Fréchet derivative coincides with the classical derivative. If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is Fréchet differentiable at $x \in U$, then $f'(x)$ equals (in a one-to-one correspondence) the $m \times n$ Jacobian matrix of f at x.

iii) If H is a Hilbert space, then $f = \|\cdot\|^2$ is a bounded C^1-functional with $f'(x) = \langle 2x, \cdot\rangle$. This is clear from (5.5) and the calculation

$$f(x+h) - f(x) - f'(x)h = \langle x+h, x+h\rangle - \langle x, x\rangle - \langle 2x, h\rangle = \|h\|^2,$$

combined with continuity of inner products and the Cauchy-Schwarz inequality.

iv) We now take a step up from case iii). Assume that H and K are real Hilbert spaces with $K \hookrightarrow H$ and let $L \in \mathscr{B}(H)$ be symmetric. Then $\mathscr{L}: K \to \mathbb{R}$ defined by $\mathscr{L}(u) = \frac{1}{2}\langle Lu, u\rangle_H$ is a bounded C^1-functional with Fréchet derivative

$$\mathscr{L}'(u) = \langle Lu, \cdot\rangle_H \quad \text{for each} \quad u \in K.$$

Indeed, the boundedness of $\mathscr{L}$ comes straight from the Cauchy-Schwarz inequality and we next calculate

$$\mathscr{L}'(u)v = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathscr{L}(u+\epsilon v)\Big|_{\epsilon=0} = \frac{1}{2}\Big(\langle Lu, v\rangle_H + \langle Lv, u\rangle_H\Big) = \langle Lu, v\rangle_H$$

by linearity and symmetry of L. Therefore

$$|\mathscr{L}'(u)v| \leq \|Lu\|_H\|v\|_H \lesssim \|u\|_H\|v\|_K \lesssim \|u\|_K\|v\|_K,$$

and so $\|\mathscr{L}'(u)\| \lesssim \|u\|_K$. Thus $\mathscr{L}'$ is bounded, and since L is linear, the same holds true for $\mathscr{L}'$. In particular, $\mathscr{L}'$ is continuous. □

Fréchet differentiation extends several well-known results from calculus. For example, (5.4) shows that if a function is differentiable at point, it is also Lipschitz continuous there. Moreover, the classical chain rule holds.

Theorem 5.5 (Chain rule). *Suppose that $f : U \subseteq X \to V \subseteq Y$ is Fréchet differentiable at $x \in U$ and $g : V \to W \subseteq Z$ is Fréchet differentiable at $f(x)$. Then $g \circ f$ is Fréchet differentiable at x and*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x). \tag{5.6}$$

Additionally, $g \circ f \in \mathrm{C}^1(U, W)$ whenever $f \in \mathrm{C}^1(U, V)$ and $g \in \mathrm{C}^1(V, W)$. □

See Liu [45, Theorem 2.4] for a detailed proof of the chain rule and other basic facts such as the product rule, where we in the last case require a suitable algebra or some sort of product structure. In passing we mention that Liu [45, Chapter 2] deals with many aspects of generalized differentiation in Banach spaces, including Taylor's theorem and the fundamental theorem of calculus, all in the abstract setting.

Switching focus to functional optimization with $Y = \mathbb{R}$, we next deduce that local extrema in the interior of a set are critical points.

Theorem 5.6 (First-order necessary condition for local extrema). *Assume that $\overline{x}$ is a local extreme point for $f : C \subseteq X \to \mathbb{R}$. If $\overline{x}$ lies in the interior of C and f is Gâteaux differentiable at $\overline{x}$, then $f'(\overline{x}) = 0$.*

Proof. Assume without loss of generality that $\overline{x}$ is a local minimizer. Note that differentiability at $\overline{x}$ is meaningful because the interior of C is open. Moreover, $\overline{x} + \epsilon h$ is in the interior for all $h \in X$ and sufficiently small $|\epsilon| > 0$. In particular, these points satisfy

$$f(\overline{x} + \epsilon h) - f(\overline{x}) \geq 0$$

by minimality of $\overline{x}$, provided $|\epsilon| > 0$ is sufficiently small. Dividing by ϵ and letting $\epsilon \to 0$, we conclude from (5.3) that $f'(\overline{x}) = 0$. □

If extreme points lie on the boundary of C, however, we only know that the Gâteaux derivative is negative (minimizers)/positive (maximizers) in the possible directions, expressed as a *variational inequality*.

Proposition 5.7 (Variational inequality). *Let $C \subseteq X$ be convex. If $f : C \to \mathbb{R}$ is Gâteaux differentiable at a global minimizer $\overline{x}$ for f on C, then the* variational inequality

$$f'(\overline{x})(x - \overline{x}) \geq 0 \quad \textit{for all} \quad x \in C$$

holds. Conversely, if f is convex and $\overline{x} \in C$ satisfies the variational inequality, then $\overline{x}$ minimizes f on C.

Remark. The result similarly holds for global maximizers with a reversed inequality and concave f. □

Proof (of Proposition 5.7). Let $x \in C$. Since C is convex, $\overline{x} + \epsilon(x - \overline{x}) \in C$ for all $\epsilon \in (0, 1)$. By minimality of $\overline{x}$ we then get $f(\overline{x} + \epsilon(x - \overline{x})) \geq f(\overline{x})$, which yields that

$$\frac{f(\overline{x} + \epsilon(x - \overline{x})) - f(\overline{x})}{\epsilon} \geq 0 \quad \text{for each} \quad \epsilon \in (0, 1).$$

Now let $\epsilon \searrow 0$ to obtain $f'(\overline{x})(x - \overline{x}) \geq 0$.

Conversely, from the convexity of f we calculate

$$\begin{aligned} f'(\overline{x})(x - \overline{x}) &= \lim_{\epsilon \searrow 0} \frac{f(\overline{x} + \epsilon(x - \overline{x})) - f(\overline{x})}{\epsilon} \\ &\leq \lim_{\epsilon \searrow 0} \frac{(1 - \epsilon) f(\overline{x}) + \epsilon f(x) - f(\overline{x})}{\epsilon} = f(x) - f(\overline{x}) \end{aligned}$$

for each $x \in C$. Thus $f(\overline{x}) \leq f(x)$ for all $x \in C$ because $f'(\overline{x})(x - \overline{x}) \geq 0$. □

Of course, if $\overline{x}$ is an interior point, then the possible directions equal all directions and the variational inequality reduces to $f'(\overline{x}) = 0$.

A natural next step for the characterization of local extrema is to examine the second-order derivative. It is well-known from ordinary calculus that symmetric positive/negative semi-definiteness of the Hessian in a neighborhood of a local minimizer/maximizer is a necessary second-order optimality condition. Moreover, if the Hessian is continuous and symmetric positive/negative definite at a critical point, then this point must be a strict local minimizer/maximizer, giving rise to a second-order sufficient condition. Equivalently, the conditions can be formulated in terms of local (strict) convexity/concavity. Generalizations exist for Banach spaces; we omit the details, but refer to Liu [45, Section 2.5].

5.3 Existence of extrema via the direct method

While the second-order sufficient conditions guarantee existence of extrema, they require a certain smoothness of the functional. From Weierstrass' extreme value theorem, however, we can detect extrema using much less regularity of the functional.

Theorem 5.8 (Extreme value theorem). *Any continuous functional $f : C \to \mathbb{R}$ defined on a non-empty, compact subset $C \subseteq X$ admits a global minimizer in C.* □

Recall (see Evgrafov [27]) that one way of proving the extreme value theorem is to start with a *minimizing sequence* $\{x_n\} \subseteq C$ satisfying $f(x_n) \to \inf\{f(x) : x \in C\}$. Compactness and continuity then yield a convergent subsequence $\{x_{n_k}\}$ with limit $\overline{x} \in C$ and $f(\overline{x}) = \inf\{f(x) : x \in C\}$. This procedure is an example of the direct method in the calculus of variations. Note that the extreme value theorem in fact holds for any topological space X.

The continuity of f can easily be weakened to *lower semi-continuity (l.s.c.)*, in the sense that

$$\liminf_{n\to\infty} f(x_n) \geq f(x) \qquad \text{whenever} \qquad x_n \to x \text{ in } C.$$

Moreover, remember that when $X = \mathbb{R}^n$, the *Heine-Borel property* holds, which states that bounded, closed subsets are compact; the converse is always true in normed spaces. This property allows us to trade the boundedness of C with growth of f at infinity, called (weak) coercivity.

Definition 5.9 (Weak coercivity). A functional $f : C \subseteq X \to \mathbb{R}$ is *(weakly) coercive on C* if

$$f(x) \to \infty \quad \text{as} \quad \|x\| \to \infty \text{ while } x \in C. \qquad \square$$

In general infinite-dimensional spaces, however, the Heine-Borel property is false. For example, the closed unit ball of ℓ^∞ is not compact. Indeed, the sequence $\{x_n\} \subset B_1(0)$ defined by $x_n^{(m)} = \delta_{mn}$, where δ_{mn} is the Kronecker-delta, has no convergent subsequences.

This is a serious drawback for the extreme value theorem, but by strengthening the compactness assumption, Weierstrass' result can be generalized to apply in so-called reflexive Banach spaces (see Definition 5.20). Underhood we coarsen the topology on X' so that convergence in X is easier, while sequential continuity and compactness are harder. This leads to a class of "weak" notions, by which we mean that a weakly compact subset of X is compact, but not necessarily conversely. The generalized Weierstrass theorem (Theorem 5.23) will essentially appear with the word "weak" in front of the above properties on f and C. First out on this path is the new concept of convergence.

Definition 5.10 (Weak convergence). A sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is said to be *weakly convergent* if there exists an $x \in X$, called the *weak limit*, such that

$$f(x_n) \to f(x) \quad \text{for all} \quad f \in X'.$$

We symbolize this as $x_n \rightharpoonup x$. □

When X is a Hilbert space, Riesz' representation theorem (Theorem 2.5) also gives an alternate characterization of weak convergence:

$$x_n \rightharpoonup x \qquad \text{if and only if} \qquad \langle x_n, y\rangle \to \langle x, y\rangle \quad \text{for all} \quad y \in X.$$

Proposition 5.11 (Basic properties of weak convergence). *Assume that $\{x_n\}$ and $\{y_n\}$ are two sequences in X and let c be a scalar. Then the following are true.*

i) *Uniqueness of limits: If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$, then $x = y$;*

ii) *(strong) convergence implies weak convergence: If $x_n \to x$, then $x_n \rightharpoonup x$;*

iii) *linearity: If $x_n \rightharpoonup x$ and $y_n \rightharpoonup y$, then $cx_n + y_n \rightharpoonup cx + y$;*

iv) *subsequences converge to the same limit: If $x_n \rightharpoonup x$, then $x_{n_k} \rightharpoonup x$ also; and*

v) *if $x_n \rightharpoonup x$, then $\{x_n\}$ is bounded and*

$$\liminf_{n\to\infty} \|x_n\| \geq \|x\|. \tag{5.7}$$

Proof. Item i) is a consequence of the Hahn-Banach theorem (Theorem 2.4) and we refer to Yan [69, Theorem 1.19] for details on both case i) and v). Properties ii), iii) and iv) are immediate by linearity and continuity of functionals in X'. □

Definition 5.12 (Weak continuity). Let $f : C \subseteq X \to Y$ and an $x \in C$ be given. Moreover, let $\{x_n\}_{n\in\mathbb{N}}$ be any sequence in C converging weakly to x. We then distinguish between three types of sequential continuity:

i) f is *weak-weak continuous at x* if $f(x_n) \rightharpoonup f(x)$;

ii) f is *weak-strong continuous at x* if $f(x_n) \to f(x)$; and

iii) when $Y = \mathbb{R}$ and X is a real normed space, f is *weakly lower semi-continuous (w.l.s.c.) at x* if

$$\liminf_{n\to\infty} f(x_n) \geq f(x).$$

If these properties hold for all $x \in C$, we call f *weak-weak continuous*, *weak-strong continuous* or *w.l.s.c. on C*, respectively. □

Standard (norm) continuity is naturally characterized as strong-strong continuity according to Definition 5.12. It is clear that weak-strong implies weak-weak and strong-strong; weak-weak implies w.l.s.c. when relevant; and w.l.s.c. yields l.s.c. Additionally, linear combinations of weak-weak/strong maps remain weak-weak/strong, but we must require positive linear combinations (nonnegative scalars) in the case of w.l.s.c. functions.

Proposition 5.13. *Every $T \in \mathscr{B}(X,Y)$ is weak-weak continuous, and if T additionally is compact, it is weak-strong continuous.*

Proof. Let $x_n \rightharpoonup x$. Then $Tx_n \rightharpoonup Tx$ in Y. Indeed, if $f \in Y'$, we have $f \circ T \in X'$ because $T \in \mathscr{B}(X,Y)$. Thus, by assumption, $f(Tx_n) \to f(Tx)$.

Switching to the second case, note first that $\{x_n\}$ is bounded from Proposition 5.11 v). Hence, by compactness of T there exists a subsequence $\{x_{n_k}\}_k$ such that $Tx_{n_k} \to y$ for some $y \in Y$. In particular, $Tx_{n_k} \rightharpoonup y$ and from the first case we get $y = Tx$ by uniqueness of weak limits.

We now prove that $Tx_{n_k} \to Tx$ for *all* subsequences of $\{x_n\}$. Assume to the contrary that there is some subsequence $\{x_{n_k}\}_k$ with $Tx_{n_k} \not\to Tx$. Without loss of generality we may assume that $\|Tx_{n_k} - Tx\|_Y \geq \delta > 0$ for all k. Since $x_{n_k} \rightharpoonup x$, the previous paragraph shows that $Tx_{n_{k_\ell}} \to Tx$ for some subsequence of $\{x_{n_k}\}_k$. This is a contradiction and so $Tx_n \to Tx$. □

Example 5.14. i) All $f \in X'$ are weak-strong continuous on any subset of X.

ii) Assume that there is a compact embedding $X \hookrightarrow Y$ and let f be a continuous (not necessarily linear) functional on Y. Then $f\restriction_X$ is weak-strong continuous, because if $x_n \rightharpoonup x$ in X, we get $x_n \to x$ in Y by compactness and Proposition 5.13, and subsequently $f(x_n) \to f(x)$.

iii) From Theorem 3.35 it is known that $\mathrm{H}_p^t \hookrightarrow \mathrm{H}_p^s$ is compact for $t > s \geq 0$. Thus $u_n \rightharpoonup u$ in H_p^t implies that $u_n \to u$ in H_p^s. Similarly, $u_n \rightharpoonup u$ in H_p^s gives $u_n \to u$ in C_p^k whenever $s > k + \frac{1}{2}$. □

Proposition 5.15. *If $g : C \subseteq X \to \mathbb{R}$ is w.l.s.c. and $f : g(C) \to \mathbb{R}$ is increasing and continuous, then the composition $f \circ g$ is w.l.s.c. on C.*

Proof. Let $x_n \rightharpoonup x$ in C and notice that

$$\liminf_{n\to\infty} f(g(x_n)) = \lim_{n\to\infty} \inf_{k\geq n} f(g(x_k)) \geq \lim_{n\to\infty} f\left(\inf_{k\geq n} g(x_k)\right) = f\left(\liminf_{n\to\infty} g(x_n)\right),$$

because the assumptions for f. We now appeal to the w.l.s.c. of g and the fact that f is increasing:

$$\liminf_{n\to\infty} g(x_n) \geq g(x) \qquad \text{and so} \qquad f\left(\liminf_{n\to\infty} g(x_n)\right) \geq f(x).$$

In conclusion $f \circ g$ is w.l.s.c. on C. □

Example 5.16. From Proposition 5.11 v) we know that norms are w.l.s.c. on any subset $C \subseteq X$. Since $t \mapsto t^p$ is continuous and increasing on $[0, \infty)$ for $p > 0$, Proposition 5.15 yields w.l.s.c. of the composition $\|\cdot\|^p : C \to [0, \infty)$. □

We next apply the new notion of convergence to sets.

Definition 5.17 (Weakly closed and compact sets). A subset $C \subseteq X$ is called

i) *weakly (sequentially) closed* if it contains all its weak limits, that is, if $\{x_n\}_n \subseteq C$ and $x_n \rightharpoonup x$ imply $x \in C$; and

ii) *weakly (sequentially) compact* if every sequence in C has a weakly convergent subsequence with limit in C. □

Weakly closed/compact sets are (strongly) closed/compact. Moreover, weak compactness implies weak closedness by Proposition 5.11 iv), and a quick argument shows that the intersection of weakly closed sets remains weakly closed. From a version of the Hahn-Banach theorem (Theorem 2.4) we also get the following result; see Ekeland and Témam [25, Corllary 1.4 and Mazur's Lemma] for a proof.

Proposition 5.18. *A closed and convex subset of a normed space is weakly closed.* □

Example 5.19. Let $t > s \geq 0$ and $\mu \geq 0$. Then the subset $\left\{u \in \mathrm{H}_p^t : \|u\|_{\mathrm{H}_p^s} = \mu\right\}$ is weakly closed in H_p^t. This is a direct consequence of Example 5.14 iii). Indeed, if $u_n \rightharpoonup u$ in H_p^t with $\|u_n\|_{\mathrm{H}_p^s} = \mu$, then $u_n \to u$ in H_p^s and so $\|u\|_{\mathrm{H}_p^s} = \mu$ by continuity norms. □

Recall that X' is a Banach space when endowed with the operator norm. Analogously to the construction of X', we may introduce the dual space $(X')'$ of X'. This space, denoted for simplicity by X'', is called the *bidual space of X* and constitutes again a Banach space. There exists also an injective canonical embedding $\pi\colon X \to X''$ defined by

$$\pi(x) = F_x, \qquad \text{where} \qquad F_x(f) = f(x) \quad \text{for all} \quad f \in X'. \tag{5.8}$$

Definition 5.20 (Reflexivity). A Banach space X is said to be *reflexive* whenever the canonical map (5.8) is surjective and thus provides an isomorphism $X \cong X''$. □

Note that all Hilbert spaces are reflexive according to Riesz' representation theorem (Theorem 2.5). One reason for studying reflexive Banach spaces is that they satisfy the Heine-Borel property in the weak sense. That is, bounded, weakly closed subsets are weakly compact. This is a consequence of the following characterizations.

Theorem 5.21 (Eberlein–Šmulian theorem). *A Banach space is reflexive if and only if every bounded sequence has a weakly convergent subsequence.* □

Theorem 5.22 (Kakutani's theorem). *A Banach space is reflexive if and only if the closed unit ball (equivalently, every closed ball) is weakly compact.* □

Proofs of these results can be found in Albiac and Kalton [2, Theorem 1.6.3 and Corollary 1.6.4] and Fabian *et al.* [28, Theorem 3.31], respectively.

Theorem 5.23 (Generalized Weierstrass theorem). *Let $f : C \to \mathbb{R}$ be w.l.s.c. on a nonempty, weakly closed subset C of a reflexive Banach space X. If either C is bounded or f is weakly coercive on C, then f is bounded from below and has a global minimizer in C.*

Proof. The argument follows Yan [69, Theorem 1.38] and Struwe [61, 1.2 Theorem].

Suppose that first C is bounded and let $\{x_n\} \subseteq C$ be a minimizing sequence. Then C is weakly compact by the Eberlein–Šmulian theorem (Theorem 5.21), so that there is a subsequence $\{x_{n_k}\}$ converging weakly to some $\overline{x} \in C$. Since f is w.l.s.c. on C, we get

$$f(\overline{x}) \le \liminf_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = \inf_{x\in C} f(x).$$

Hence, f is bounded from below and admits a global minimizer $\overline{x} \in C$.

Consider now f to be weakly coercive on C. By assumption there is an $x \in C$, so coercivity implies the existence of a closed ball $B = \overline{B}_R(0) \subseteq X$ with sufficiently large radius, satisfying

$$x \in B \cap C \qquad \text{and} \qquad f(y) \ge f(x) \quad \text{for all} \quad y \in C \setminus (B \cap C).$$

Note that B is weakly closed according to Proposition 5.18. Since the intersection of weakly closed sets is weakly closed, $B \cap C$ fits the condition of the first part of the theorem. Thus there is some $\overline{x} \in B \cap C$ satisfying $f(z) \geq f(\overline{x})$ for all $z \in B \cap C$. In particular, $f(x) \geq f(\overline{x})$, and so by construction of x it follows that $\overline{x}$ minimizes f in C. □

5.4 CONSTRAINED PROBLEMS AND THE METHOD OF LAGRANGE MULTIPLIERS

When optimizing $f: C \to \mathbb{R}$, we have up till now encapsulated constraints in the unspecified subset $C \subseteq X$, but if possible, some of them may be written as functional constraints $g: C \to \mathbb{R}$. Up to translation and reflection, the constraints can always be posed as either an equality constraint (level set) $g(x) = 0$ or an inequality constraint (sublevel set) $g(x) \leq 0$. The benefit of introducing g is for example that, "as was known to Lagrange and Euler, it may suffice to [optimize] an augmented functional *without* constraints" (Troutman [64, §2.3]).

Proposition 5.24. *Let $f, g: C \subseteq X \to \mathbb{R}$ and suppose that there exists a constant $\lambda \in \mathbb{R}$ such that $\overline{x}$ is a [strict] global minimizer for $f + \lambda g$ on C. Then $\overline{x}$ is also a [strict] global minimizer for f on the sublevel set $\{x \in C : \lambda g(x) \leq \lambda g(\overline{x})\}$.*

Proof. As in Troutman [64, (2.4) Corollary], note by assumption that

$$f(\overline{x}) + \lambda g(\overline{x}) \leq f(x) + \lambda g(x)$$

for all $x \in C$, so that $f(\overline{x}) \leq f(x)$ on $\{x \in C : \lambda g(x) \leq \lambda g(\overline{x})\}$ [with $<$ for $x \in C \setminus \{\overline{x}\}$ if $\overline{x}$ is a strict minimizer]. □

The parameter λ in Proposition 5.24 is known as a *Lagrange multiplier*. We shall also prove a partial converse of this result, roughly saying that if $\overline{x}$ is a local extreme point for f restricted to the constraint $g(x) = g(\overline{x})$, then $\overline{x}$ is a either critical point for g or the augmented functional $f + \lambda g$ for some $\lambda \in \mathbb{R}$. The hope is that $g'(\overline{x}) \neq 0$.

In applications the *method of Lagrange multipliers* may thus be viewed from two angles. Either optimize with respect to a constraint and get a multiplier, or fix a multiplier in an augmented functional and obtain a constraint.

Theorem 5.25 (Inverse function theorem). *Assume that $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a C^1-function. If $x \in U$ and $f'(x)$ is invertible, then there exists open sets $V \subseteq U$ and $W \subseteq f(U)$ such that $x \in V$ and the restriction $f: V \to W$ has a continuously differentiable inverse $f^{-1}: W \to V$. Furthermore,*

$$(f^{-1})'(f(x)) = f'(x)^{-1}.$$ □

See Liu [45, Theorem 2.45] for a proof of the inverse function theorem.

Lemma 5.26. *Let $f,g \in C^1(U,\mathbb{R})$ be given, where $U \subseteq X$, and suppose $\overline{x} \in U$. If there exist $y,z \in X$ such that*

$$\begin{vmatrix} f'(\overline{x})\,y & f'(\overline{x})z \\ g'(\overline{x})\,y & g'(\overline{x})z \end{vmatrix} \neq 0, \tag{5.9}$$

then $\overline{x}$ cannot be a local extremal point for f on the constraint set $\{x \in U : g(x) = g(\overline{x})\}$.

Proof. The argument is mainly based on Troutman [64, §5.7 and (5.14) Proposition]. Yan [69, Lemma 1.42] uses a similar technique via the implicit function theorem.

Note first that $\overline{x}$ cannot be such a local extremal point for f if there exist $x,x' \in U$ arbitrarily close to $\overline{x}$ with

$$f(x) < f(\overline{x}) < f(x') \qquad \text{and} \qquad g(x) = g(\overline{x}) = g(x'). \tag{5.10}$$

We establish a neighborhood of $\big(f(\overline{x}), g(\overline{x})\big) \in \mathbb{R}^2$ containing such points.

Let $y,z \in X$ satisfy (5.9) and introduce the auxiliary function φ defined by

$$\varphi(s,t) = \begin{bmatrix} \varphi_1(s,t) \\ \varphi_2(s,t) \end{bmatrix} = \begin{bmatrix} f(\overline{x}+sy+tz) \\ g(\overline{x}+sy+tz) \end{bmatrix}.$$

Observe that φ is well-defined in a neighborhood $N \subseteq \mathbb{R}^2$ of $(0,0)$, with $\varphi(0,0) = \big(f(\overline{x}), g(\overline{x})\big)$. Moreover, the Jacobian equals

$$\varphi'(0,0) = \begin{bmatrix} f'(\overline{x})\,y & f'(\overline{x})z \\ g'(\overline{x})\,y & g'(\overline{x})z \end{bmatrix},$$

because, for example,

$$\frac{\partial \varphi_1}{\partial s}(s,t) = \frac{\partial}{\partial s} f(\overline{x}+sy+tz) = f'(\overline{x}+sy+tz)\,y.$$

Thus $\varphi'(0,0)$ is invertible by (5.9). If $\varphi \in C^1(N,\mathbb{R}^2)$, then Theorem 5.25 implies the existence of open sets $V \subseteq N$ and $W \subseteq \varphi(N)$, such that $(0,0) \in V$ and the restriction $\varphi : V \to W$ has a continuous inverse $\varphi^{-1} : W \to V$. Since W is a neighborhood of $\varphi(0,0) = \big(f(\overline{x}), g(\overline{x})\big)$, there are $x = \overline{x}+sy+tz$ and $x' = \overline{x}+s'y+t'z$ satisfying (5.10) for some $(s,t),(s',t') \in V$. By the continuity of φ^{-1} we may choose $(s,t),(s',t')$ arbitrarily close to $(0,0)$, which subsequently yields $x,x' \in U$ arbitrarily close to $\overline{x}$.

It remains to establish that $\varphi \in C^1(N,\mathbb{R}^2)$, which will follow provided φ_1 and φ_2 have continuous partial derivatives in N. To this end, let $(s_n,t_n) \to (s,t)$ in N. Then $(\overline{x}+s_ny+t_nz) \to \overline{x}+sy+tz$ in U. The estimate

$$\begin{aligned} \left| \frac{\partial \varphi_1}{\partial s}(s,t) - \frac{\partial \varphi_1}{\partial s}(s_n,t_n) \right| &= \left| f'(\overline{x}+sy+tz)\,y - f'(\overline{x}+s_ny+t_nz)\,y \right| \\ &\leq \left\| f'(\overline{x}+sy+tz) - f'(\overline{x}+s_ny+t_nz) \right\| \end{aligned}$$

then gives continuity of $\partial\varphi_1/\partial s$ at (s,t), because $f' \in \mathrm{C}(U,\mathbb{R})$ by assumption. The other cases are similar. □

Theorem 5.27 (Lagrange multiplier theorem). *Let $f,g \in \mathrm{C}^1(U,\mathbb{R})$, where $U \subseteq X$, and suppose that $\overline{x}$ is a local extreme point for f on the level set $\{x \in U : g(x) = g(\overline{x})\}$. Then either*

i) $g'(\overline{x}) = 0$; or

ii) there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that $f'(\overline{x}) = \lambda g'(\overline{x})$.

Proof. We follow Yan [69, Theorem 1.43], but refer to Liu [45, Lemma 2.58 and Theorem 2.59] for a different proof strategy.

Suppose that i) does not hold. Then there is a $y \in X$ with $g'(\overline{x})y \neq 0$, which by Lemma 5.26 yields

$$\begin{vmatrix} f'(\overline{x})y & f'(\overline{x})z \\ g'(\overline{x})y & g'(\overline{x})z \end{vmatrix} = 0 \quad \text{for all} \quad z \in X.$$

Now put $\lambda = f'(\overline{x})y/g'(\overline{x})y$ to obtain $f'(\overline{x})z = \lambda g'(\overline{x})z$ for all $z \in X$. □

Both Proposition 5.24 and Theorem 5.27 can be extended to any finite number of equality constraints; see Troutman [64, (2.4) Proposition and (5.16) Theorem] for details. More generally the Lagrange multiplier may also constitute a function. We note that for problems involving both equality and inequality constraints, the famous Karush-Kuhn-Tucker optimality conditions for Euclidean spaces extend to real Banach spaces. Tröltzsch [63, Chapter 6] gives an overview of this theory and associated constraint qualifications.

Chapter 6

Existence of periodic traveling waves to the generalized Whitham equation

In this final chapter we study the Whitham equation (1.3) in a more general setting and show that it admits periodic traveling waves using the tools of Fourier methods, Sobolev spaces and the calculus of variations.

The upcoming sections are structured as follows. In Section 6.1 we set up the generalized Whitham equation, while Section 6.2 discusses its version for periodic traveling waves. Section 6.3 next focuses on a variational formulation of the "traveling Whitham equation" and outlines our proof technique. Section 6.4 proceeds by establishing the existence of solutions to a penalized problem, and in Section 6.5 we demonstrate via *a priori* estimates that those solutions solve the original problem, that is, are periodic traveling waves to the generalized Whitham equation. Section 6.6 lastly provides suggestions for further work.

As said in Section 1.3, the existence argument builds on the first part of a work by Ehrnström *et al.* [21]. Especially, we consider the same constrained, penalized variational formulation and argue with similar *a priori* estimates, but remove the comparisons with the solitary-wave problem. Additionally, we lower the regularity of the nonlinearity and extend the results from H^1_P to the fractional-order setting with help of interpolation and a natural estimate of the nonlinearity. References are stated throughout for results that originates in some way from [21], and if proofs are not cited, they have been established independently by the author.

6.1 The generalized Whitham equation

Before we generalize Whitham's model equation (1.3) it is helpful to rewrite it into a simpler form and collect some properties of its kernel and linear phase velocity. Via the scalings

$$\eta \mapsto \frac{4h_0}{3}\,\eta, \quad x \mapsto h_0\,x \quad \text{and} \quad t \mapsto \frac{h_0}{c_0}\,t,$$

the nondimensionalized version

$$\eta_t + (\eta^2)_x + K * \eta_x = 0.$$

appears. The function $K = \frac{1}{\sqrt{2\pi}}\mathscr{F}^{-1}(m)$, with the Fourier transform applied distributionally, is called the *Whitham kernel* and m constitutes the nondimensionalized variant of c_{Euler} (1.2) given as

$$m(\xi) = \begin{cases} \sqrt{\tanh(\xi)/\xi} & \text{for } \xi \neq 0; \\ 1 & \text{for } \xi = 0, \end{cases} \tag{6.1}$$

with the dimensionless variable $\xi = kh$. Based on [21, Assumptions (A1)] we list some attributes of m.

Proposition 6.1. *The linear phase speed m (6.1) is an even, real analytic function with Taylor expansion*

$$m(\xi) = 1 - \tfrac{1}{6}\xi^2 + \mathcal{O}(\xi^4) \tag{6.2}$$

and has a strict global maximum at the origin. Additionally, m satisfies

$$m(\xi) \lesssim \big(1 + |\xi|\big)^{-1/2}. \tag{6.3}$$

Proof. Since tanh is odd, m is even. It can be seen that m is real analytic and strictly decreasing on $(0, \infty)$; see Ehrnström and Kalisch [22, Section 2]. In particular, m has a strict global maximum at $\xi = 0$. We note that l'Hôpital's rule is relevant for calculating the Taylor coefficients, but omit the details. As regards (6.3), observe that

$$\big(1 + |\xi|\big)\frac{\tanh \xi}{\xi} < 1 + \tanh \xi < 2. \qquad \square$$

Moreover, the following result by Ehrnström and Kalisch [22, Sections 2 and 4] and Ehrnström and Wahlén [19, 24] states fundamental properties of the kernel. Note that the integrability of K is in symmetry with that of m, as $m \in \mathrm{L}^p(\mathbb{R})$ only for $p > 2$ by (6.3).

Proposition 6.2. *The Whitham kernel K is real, even and lies in $\mathrm{C}^\infty(\mathbb{R} \setminus \{0\}) \cap \mathrm{L}^p(\mathbb{R})$ for $1 \leq p < 2$. It features a singularity of order $\mathcal{O}\big(|\xi|^{-1/2}\big)$ at the origin and has exponential decay at infinity.* $\square$

We next generalize the nonlinear term $(\eta^2)_x$ and consider power nonlinearities $(n(\eta))_x$, where $n\colon \mathbb{R} \to \mathbb{R}$ is defined either as

$$n(x) = |x|^q \qquad \text{or} \qquad n(x) = x|x|^{q-1}. \tag{6.4}$$

Here $n(\eta)$ is understood as the composition

$$n(\eta)(x) = n\big(\eta(x)\big) \tag{6.5}$$

induced by n and $q \in (1, \infty)$ is some fixed exponent. In fact, it is natural to consider all $q \in (0, \infty)$, but our current approach needs $q > 1$.

Since convolution commutes with differentiation, the *(generalized) Whitham equation* then reads

$$\eta_t + \big(L\eta + n(\eta)\big)_x = 0, \tag{6.6}$$

where we have introduced the operator L defined by $L\eta = K * \eta$. By the convolution theorem (Theorem 4.7), L is a Fourier multiplier operator with symbol m, that is,

$$\widehat{L\eta} = \sqrt{2\pi}\,\widehat{K} \cdot \widehat{\eta} = m \cdot \widehat{\eta}.$$

6.2 The traveling wave equation for the periodic case

Traveling wave solutions to (6.6) take the form $\eta(x,t) = u(x - ct)$ for some function $u\colon \mathbb{R} \to \mathbb{R}$, where c is the constant wave speed; see Appendix A for details. By introducing the "steady" variable $x - ct$ and immediately denoting this again by x, the Whitham equation reduces to

$$-cu_x + \big(Lu + n(u)\big)_x = 0. \tag{6.7}$$

Integration gives

$$-cu + Lu + n(u) = C \tag{6.8}$$

for some constant $C \in \mathbb{R}$ which is set equal to 0. The choice of C is practical and does no harm when it comes to establishing the existence of traveling waves. Indeed, just differentiate (6.8) to get (6.7) regardless of the value of C. Ehrnström and Kalisch [23] remark also that both the original Whitham equation ($q = 2$) and the KdV equation (1.1) are invariant under the Galilean transformation

$$u \mapsto u + D, \quad c \mapsto c + 2D \quad \text{and} \quad C \mapsto C + D(1 - c - D)$$

for any $D \in \mathbb{R}$. In total the "*traveling Whitham equation*" equals

$$Lu + n(u) - cu = 0. \tag{6.9}$$

Our interest is weak P-periodic solutions to (6.9) residing in a real Sobolev space H^s_P, and we now investigate how L and n affect the order s.

Remark. Note that since the focus is on real waves $u\colon \mathbb{R} \to \mathbb{R}$, this chapter only treats the real variants of the relevant function spaces ($\mathbb{F} = \mathbb{R}$). Fourier coefficients indexed by $\mathbb{Z}$ may nonetheless be complex and thus the sequence spaces are still complex. □

How does L act on periodic functions? To understand this we follow Ehrnström and Kalisch [22, Section 3] and suppose that $u \in \mathrm{L}^\infty_P$. Since $K \in \mathrm{L}^1(\mathbb{R})$ via Proposition 6.2, the dominated convergence theorem (Theorem 2.14) gives

$$\begin{aligned} L(u) &= \int_{\mathbb{R}} K(x-y)u(y)\,\mathrm{d}y \\ &= \sum_{j\in\mathbb{Z}} \int_0^P K\big(x-(y-jP)\big)u(y-jP)\,\mathrm{d}y \\ &= \int_0^P \sum_{j\in\mathbb{Z}} K\big(x-(y-jP)\big)u(y)\,\mathrm{d}y \\ &= K_P *_P u(x), \end{aligned}$$

where $K_P = \sum_{j\in\mathbb{Z}} K(\cdot - jP)$ is the periodized kernel. The triangle inequality and Proposition 6.2 imply that $K_P \in \mathrm{L}^p_P$ for $1 \le p < 2$; see [22, Section 3]. Moreover,

$$\begin{aligned} (\widehat{K_P})_k &= \frac{1}{\sqrt{P}} \int_0^P \sum_{j\in\mathbb{Z}} K(x-jP)\,\mathrm{e}^{-2\pi \mathrm{i} kx/P}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{P}} \sum_{j\in\mathbb{Z}} \int_0^P K(x-jP)\,\mathrm{e}^{-2\pi \mathrm{i} k(x-jP)/P}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{P}} \int_{\mathbb{R}} K(x)\,\mathrm{e}^{-2\pi \mathrm{i} kx/P}\,\mathrm{d}x \\ &= \sqrt{\tfrac{2\pi}{P}}\,\widehat{K}\big(\tfrac{2\pi k}{P}\big), \end{aligned}$$

and so

$$(\widehat{Lu})_k = \sqrt{P}\,(\widehat{K_P})_k \cdot \widehat{u}_k = m\Big(\tfrac{2\pi k}{P}\Big)\,\widehat{u}_k \tag{6.10}$$

by the convolution theorem (Theorem 3.24). Hence, at least for $u \in \mathrm{L}^\infty_P$, we conclude that L is a multiplier operator also in the periodic setting, with the same symbol as the solitary-wave problem.

In particular, L has the same smoothing effect [21, Proposition 2.1 (i)] in the periodic case.

Proposition 6.3. *For all $s > \frac{1}{2}$ the map L defines a continuous linear operator $\mathrm{H}^s_P \to \mathrm{H}^{s+\frac{1}{2}}_P$. Additionally, if $r \in [0,s]$, then L satisfies the symmetry*

$$\langle Lu, v\rangle_r = \langle u, Lv\rangle_r \quad \textit{for} \quad u, v \in \mathrm{H}^s_P.$$

Convention. To simplify notation we write $\|\cdot\|_s$ for $\|\cdot\|_{\mathrm{H}^s_P}$ and $\langle\cdot,\cdot\rangle_s$ for $\langle\cdot,\cdot\rangle_{\mathrm{H}^s_P}$ throughout the chapter ($\|\cdot\|_\infty$ is still the supremum norm). □

Proof (of Proposition 6.3). The linearity of L is evident. Note next from (6.3) that

$$m\left(\tfrac{2\pi k}{P}\right) \lesssim \left(1 + \left|\tfrac{2\pi k}{P}\right|^2\right)^{-1/4} = \langle k\rangle_P^{-1/2} \tag{6.11}$$

because $1 + k^2 \le (1+|k|)^2$. Since $\mathrm{H}^s_P \hookrightarrow \mathrm{C}_P \subset \mathrm{L}^\infty_P$ by the Sobolev embedding theorem (Theorem 3.35 ii)), we use (6.10) and calculate

$$\begin{aligned}\|Lu\|^2_{s+\frac{1}{2}} &= \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s+1} \left|(\widehat{Lu})_k\right|^2 = \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s+1} \left|m\left(\tfrac{2\pi k}{P}\right)\right|^2 |\widehat{u}_k|^2 \\ &\lesssim \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s} |\widehat{u}_k|^2 = \|u\|_s^2.\end{aligned}$$

Therefore L is a bounded linear operator $\mathrm{H}^s_P \to \mathrm{H}^{s+\frac{1}{2}}_P$. Lastly, $\mathrm{H}^s_P \hookrightarrow \mathrm{H}^r_P$, and the symmetry statement on H^r_P is a consequence of m being real:

$$(\widehat{Lu})_k\, \overline{\widehat{v}_k} = m\left(\tfrac{2\pi k}{P}\right) \widehat{u}_k\, \overline{\widehat{v}_k} = \widehat{u}_k\, \overline{m\left(\tfrac{2\pi k}{P}\right) \widehat{v}_k} = \widehat{u}_k\, \overline{(\widehat{Lv})_k}$$

so that

$$\langle Lu, v\rangle_r = \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2r} (\widehat{Lu})_k\, \overline{\widehat{v}_k} = \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2r}\, \widehat{u}_k\, \overline{(\widehat{Lv})_k} = \langle u, Lv\rangle_r. \qquad \square$$

Theorem 6.4. *For all $q > 1$ the nonlinearity (6.4) induces a bounded composition operator*

$$n\colon \mathrm{H}^s_P \to \mathrm{H}^s_P, \qquad \textit{where} \qquad s \in \begin{cases} \left(\frac{1}{2}, \infty\right) & \textit{if } n(x) = x^q \textit{ with } 2 \le q \in \mathbb{Z}_+; \quad (6.12a) \\ \left(\frac{1}{2}, q\right) & \textit{for general } n \text{ (6.4)}, \quad (6.12b) \end{cases}$$

defined as in (6.5) and still denoted by n.

Remark. We refer to Runst [53, Theorem 2], Sickel [59, Theorem 2] and Runst and Sickel [54, Chapter 5 (Section 5.4)] for a proof of case (6.12b) in more general spaces of Triebel-Lizorkin and Besov type, including the real version of $\mathrm{H}^s(\mathbb{R})$. As noted in [53, Remark 3], the result also holds in bounded domains with smooth boundary and, in particular, for H^s_P since $n(u)$ is P-periodic whenever u is. Their methods are based upon the Taylor expansion of n, estimates of certain integral means and maximal-function techniques (Fefferman-Stein-Petre and Hardy-Littlewood maximal inequalities). The upper bound $s < q$ originates from the finite Hölder regularity of n.

Moreover, the estimate

$$\|n(u)\|_s \lesssim_{q,s,P} \|u\|_s^q, \tag{6.13}$$

where the estimation constant decreases with increasing P, likely holds for all s in (6.12), and gives boundedness (and well-definedness) of the operator. Our arguments for the traveling Whitham equation (6.9) will use (6.13) to guarantee that the waves are of small amplitude, measured by the H^s_P norm; see Section 6.5.

In case (6.12a) the estimate is immediate from Theorem 3.36. Runst and Sickel have established (6.13) for the finite case (6.12b), but it is not clear how the estimation constant depends on P, and it is beyond the scope of this text to study the P-dependence in their approach. Calculating directly and using simple estimates, however, we give a proof of (6.13) for the subcases

$$s \in \left(\tfrac{1}{2}, 1\right] \cup \cdots \cup \left(\lfloor q \rfloor - \tfrac{1}{2}, \lfloor q \rfloor\right], \tag{6.14}$$

excluding $s = \lfloor q \rfloor$ if $2 \leq q \in \mathbb{Z}_+$, of (6.12b). □

Lemma 6.5. *Let $n \in \mathrm{C}^{m+1}(\mathbb{R})$ for some $m \in \mathbb{Z}_+$. Then n induces a composition operator $n \colon \mathrm{H}^m_P \to \mathrm{H}^m_P$ defined as in (6.5) and still denoted by n. Moreover, the higher-order chain rule*

$$\big(n(u)\big)^{(\ell)} = \sum c_{k_1,\ldots,k_\ell}\, n^{(k_1+\cdots+k_\ell)}(u) \cdot \prod_{j=1}^{\ell} \big(u^{(j)}\big)^{k_j}, \tag{6.15}$$

known as Faà di Bruno's formula, holds a.e. for each $\ell = 1, \ldots, m$, where the sum runs through all $(k_1, \ldots, k_\ell) \in \mathbb{N}^\ell$ satisfying $\sum_{j=1}^{\ell} j\, k_j = \ell$ and $c_{k_1,\ldots,k_\ell} \in \mathbb{Z}_+$ are some coefficients.

Proof. See Yan [69, Example 1.32] and Johnson [42]. □

Proof (of Theorem 6.4 for s in (6.14)). We show (6.13) and first note that $n \in \mathrm{C}^{\lfloor q \rfloor}(\mathbb{R})$, with

$$\left|n^{(\ell)}(u)\right| \lesssim_q |u|^{q-\ell} \quad \text{for} \quad \ell = 0, \ldots, \lfloor q \rfloor, \tag{6.16}$$

which follows from the differentiation formula

$$n'(u) = \frac{q\, n(u)}{u} = \begin{cases} q\, u|u|^{q-2} & \text{if } n(x) = |x|^q; \\ q\, |u|^{q-1} & \text{if } n(x) = x|x|^{q-1} \end{cases}$$

when $q > 1$. The trick is now to repeatedly apply the inequalities

$$|ab - cd| \leq |a||b - d| + |d||a - c| \qquad \text{and} \qquad \left|\sum_{i=1}^{r} a_i\right|^2 \lesssim_r \sum_{i=1}^{r} |a_i|^2, \tag{6.17}$$

the Sobolev embedding theorem (Theorem 3.35) and Faà di Bruno's formula (6.15). Additionally, we use any of the equivalent norms

$$\|u\|_m^2 \asymp_m \sum_{j=1}^m \left\|u^{(j)}\right\|_0^2 \asymp_m \|u\|_0^2 + \left\|u^{(m)}\right\|_0^2$$

from (3.11) when $s = m \in \mathbb{Z}_+$ or the Sobolev–Slobodeckij norm (3.16) of H_P^s if $s \notin \mathbb{Z}_+$.

The case $s = m \in \mathbb{Z}_+$. We have

$$\|n(u)\|_0^2 \le \|u\|_\infty^{2(q-1)} \|u\|_0^2 \lesssim_{m,P} \|u\|_m^{2q}$$

from Theorem 3.35 ii). Next, (6.15), (6.16) and (6.17) yield

$$\begin{aligned}\left\|(n(u))^{(m)}\right\|_0^2 &\lesssim_m \sum \left\|n^{(k_1+\cdots+k_m)}(u)\right\|_\infty^2 \left\|\prod_{j=1}^m \left(u^{(j)}\right)^{k_j}\right\|_0^2 \\ &\lesssim_q \sum \|u\|_\infty^{2\left(q-(k_1+\cdots+k_m)\right)} \left(\prod_{j=1}^{\widetilde{m}-1} \left\|u^{(j)}\right\|_\infty^{2k_j}\right) \left\|u^{(\widetilde{m})}\right\|_\infty^{2(k_{\widetilde{m}}-1)} \left\|u^{(\widetilde{m})}\right\|_0^2, \end{aligned} \tag{6.18}$$

where the summation is as in Lemma 6.5, $\widetilde{m} = \max\left\{j : k_j > 0\right\}$ and the parenthesized product equals 1 whenever $\widetilde{m} = 1$. Notice how the exponents always add up to

$$2\left(q - (k_1 + \cdots + k_m)\right) + \sum_{j=1}^{\widetilde{m}-1} 2k_j + 2(k_{\widetilde{m}} - 1) + 2 = 2q.$$

Moreover, the constraint $\sum_{j=1}^m j\,k_j = m$ from Lemma 6.5 implies that $k_m \le 1$. In particular, when $\widetilde{m} = m$, there is no troublesome $\left\|u^{(m)}\right\|_\infty$ term. Since $\left\|u^{(j)}\right\|_\infty \lesssim_{m,P} \|u\|_m$ for each $j = 0, \ldots, m-1$ by Theorem 3.35 ii) and $\left\|u^{(\widetilde{m})}\right\|_0 \le \|u\|_m$ for all $\widetilde{m}$, the estimate (6.18) then becomes

$$\left\|(n(u))^{(m)}\right\|_0^2 \lesssim_{q,m,P} \sum \|u\|_m^{2q} \lesssim_m \|u\|_m^{2q}.$$

Hence,

$$\|n(u)\|_m^2 \asymp_m \|n(u)\|_0^2 + \left\|(n(u))^{(m)}\right\|_0^2 \lesssim_{q,m,P} \|u\|_m^{2q}. \tag{6.19}$$

The case $s \notin \mathbb{Z}_+$. Let $m = \lfloor s \rfloor$ and $\sigma = s - m$. Since

$$\|n(u)\|_m \lesssim_{q,m,P} \|u\|_m^q \le \|u\|_s^q$$

from (6.19), it remains to estimate the semi-norm $\left|(n(u))^{(m)}\right|_{\mathrm{H}_P^\sigma}$.

When $s = \sigma \in \left(\frac{1}{2}, 1\right)$, observe that

$$\left|n\left(u(x)\right) - n\left(u(y)\right)\right| \le \sup_{|z| \le \|u\|_\infty} |n'(z)| \left|u(x) - u(y)\right| \lesssim_q \|u\|_\infty^{q-1} \left|u(x) - u(y)\right|,$$

because $n \in \mathrm{C}^1(\mathbb{R})$ as $q > 1$. This gives

$$\left|n(u)\right|^2_{\mathrm{H}^\sigma_P} = \int_0^P \int_0^P \frac{\left|n\big(u(x)\big) - n\big(u(y)\big)\right|^2}{|x-y|^{1+2\sigma}} \,\mathrm{d}x\,\mathrm{d}y \lesssim_q \|u\|_\infty^{2(q-1)} |u|^2_{\mathrm{H}^\sigma_P} \lesssim_{s,P} \|u\|_s^{2q},$$

and we are done.

When $s > 1$, we split (6.15) into

$$\big(n(u)\big)^{(m)} = \left[\sum c_{k_1,\dots,k_{m-1}}\, n^{(\kappa)}(u) \cdot \prod_{j=1}^{m-1} \big(u^{(j)}\big)^{k_j}\right] + n'(u)\, u^{(m)},$$

where the sum runs through all $(k_1,\dots,k_{m-1}) \in \mathbb{N}^{m-1}$ satisfying $\sum_{j=1}^{m-1} j\,k_j = m$ and $\kappa = k_1 + \cdots + k_{m-1}$. This is valid when $m \geq 2$, and if $m = 1$, only the last term is present. Now estimate

$$\left|\big(n(u)\big)^{(m)}(x) - \big(n(u)\big)^{(m)}(y)\right|^2 \lesssim_m A + B$$

using (6.17), where

$$A = \sum \left| n^{(\kappa)}\big(u(x)\big) \cdot \prod_{j=1}^{m-1} \big(u^{(j)}(x)\big)^{k_j} \;-\; n^{(\kappa)}\big(u(y)\big) \cdot \prod_{j=1}^{m-1} \big(u^{(j)}(y)\big)^{k_j} \right|^2$$

is present only if $m \geq 2$, and

$$B = \left| n'\big(u(x)\big)\, u^{(m)}(x) \;-\; n'\big(u(y)\big)\, u^{(m)}(y) \right|^2.$$

From (6.17) again, we get

$$\begin{aligned} A \lesssim \sum \Bigg\{ &\left| n^{(\kappa)}\big(u(x)\big) - n^{(\kappa)}\big(u(y)\big) \right|^2 \cdot \prod_{j=1}^{m-1} \left|u^{(j)}(y)\right|^{2k_j} \\ &+ \left| n^{(\kappa)}\big(u(x)\big) \right|^2 \cdot \left| \prod_{j=1}^{m-1} \big(u^{(j)}(x)\big)^{k_j} - \prod_{j=1}^{m-1} \big(u^{(j)}(y)\big)^{k_j} \right|^2 \Bigg\}. \end{aligned} \tag{6.20}$$

Next,

$$\left| n^{(\kappa)}\big(u(x)\big) - n^{(\kappa)}\big(u(y)\big) \right|^2 \lesssim_q \|u\|_\infty^{2(q-\kappa-1)} \left|u(x) - u(y)\right|^2 \tag{6.21}$$

by (6.16) because $n \in \mathrm{C}^{m+1}(\mathbb{R})$ as $\lfloor q \rfloor \geq m+1$. Thus the first part in (6.20) is bounded by

$$\begin{aligned} &\sum \|u\|_\infty^{2(q-\kappa-1)} \left|u(x) - u(y)\right|^2 \cdot \prod_{j=1}^{m-1} \left\|u^{(j)}\right\|_\infty^{2k_j} \\ \lesssim_{s,P} &\sum \|u\|_s^{2\left(q-\kappa-1+(k_1+\cdots+k_{m-1})\right)} \left|u(x) - u(y)\right|^2 \\ \lesssim &\; \|u\|_s^{2(q-1)} \left|u(x) - u(y)\right|^2 \end{aligned} \tag{6.22}$$

with help of Theorem 3.35 ii). For the second part in (6.20) we similarly estimate

$$\left|n^{(\kappa)}\big(u(x)\big)\right|^2 \overset{(6.16)}{\lesssim_q} \|u\|_\infty^{2(q-\kappa)} \lesssim_{s,P} \|u\|_s^{2(q-\kappa)} \tag{6.23}$$

and

$$\left|\prod_{j=1}^{m-1}\big(u^{(j)}(x)\big)^{k_j} - \prod_{j=1}^{m-1}\big(u^{(j)}(y)\big)^{k_j}\right|^2 \overset{(6.17)}{\lesssim} C_1 + D_1, \tag{6.24}$$

where, if $m \geq 3$,

$$C_i = \left|\big(u^{(i)}(x)\big)^{k_i} - \big(u^{(i)}(y)\big)^{k_i}\right|^2 \cdot \prod_{j=i+1}^{m-1} \left|u^{(j)}(y)\right|^{2k_j}$$

and

$$D_i = \left|u^{(i)}(x)\right|^{2k_i} \cdot \left|\prod_{j=i+1}^{m-1}\big(u^{(j)}(x)\big)^{k_j} - \prod_{j=i+1}^{m-1}\big(u^{(j)}(y)\big)^{k_j}\right|^2$$

for $i = 1, \ldots, m-2$, and

$$C_{m-1} = \left|\big(u^{(m-1)}(x)\big)^{k_{m-1}} - \big(u^{(m-1)}(y)\big)^{k_{m-1}}\right|^2 \qquad \text{and} \qquad D_{m-1} = 0,$$

which is also valid when $m = 2$. The pattern of D_i is recursively given by

$$D_i \overset{(6.17)}{\lesssim} \left|u^{(i)}(x)\right|^{2k_i}\big(C_{i+1} + D_{i+1}\big) \lesssim_{s,P} \|u\|_s^{2k_i}\big(C_{i+1} + D_{i+1}\big),$$

so that

$$D_1 \lesssim_{s,P} \sum_{i=2}^{m-1} \|u\|_s^{2\sum_{h=1}^{i-1} k_h} C_i. \tag{6.25}$$

As regards C_i, observe that

$$\begin{aligned}\left|\big(u^{(i)}(x)\big)^{k_i} - \big(u^{(i)}(y)\big)^{k_i}\right|^2 &\overset{(6.17)}{\lesssim} \left|u^{(i)}(x) - u^{(i)}(y)\right|^2 \left|u^{(i)}(x)\right|^{2(k_i-1)} \\ &\quad + \left|u^{(i)}(y)\right|^2 \left|\big(u^{(i)}(x)\big)^{k_i-1} - \big(u^{(i)}(y)\big)^{k_i-1}\right|^2\end{aligned}$$

if $k_i \geq 2$. Again we have a recursive pattern, and hence, by induction,

$$\begin{aligned}\left|\big(u^{(i)}(x)\big)^{k_i} - \big(u^{(i)}(y)\big)^{k_i}\right|^2 &\overset{(6.17)}{\lesssim} \sum_{\ell=1}^{k_i-1} \|u^{(i)}\|_\infty^{2(k_i-1)} \left|u^{(i)}(x) - u^{(i)}(y)\right|^2 \\ &\lesssim_{s,P} \|u\|_s^{2(k_i-1)} \left|u^{(i)}(x) - u^{(i)}(y)\right|^2,\end{aligned}$$

which is also valid for $k_i = 1$. Therefore

$$C_i \lesssim_{s,P} \|u\|_s^{2(k_i-1)} \big|u^{(i)}(x)-u^{(i)}(y)\big|^2 \cdot \prod_{j=i+1}^{m-1} \big\|u^{(j)}\big\|_\infty^{2k_j}$$

$$\lesssim_{s,P} \|u\|_s^{2\left(-1+\sum_{h=i}^{m-1} k_h\right)} \big|u^{(i)}(x)-u^{(i)}(y)\big|^2$$

whenever $k_i \geq 1$, and if $k_i = 0$, then clearly $C_i = 0$.

Insertion into (6.25) then gives

$$C_1 + D_1 \lesssim_{s,P} \|u\|_s^{2(\kappa-1)} \sum_{i=1}^{m-1} \big|u^{(i)}(x)-u^{(i)}(y)\big|^2,$$

where we remember that $\kappa = k_1 + \cdots + k_{m-1}$. Combining (6.22), (6.23) and (6.24) this yields

$$A \lesssim_{s,P} \|u\|_s^{2(q-1)} \big|u(x)-u(y)\big|^2 + \|u\|_s^{2(q-\kappa)} \big(C_1 + D_1\big)$$

$$\lesssim_{s,P} \|u\|_s^{2(q-1)} \sum_{i=0}^{m-1} \big|u^{(i)}(x)-u^{(i)}(y)\big|^2.$$

For part B we use that $\big\|u^{(m)}\big\|_\infty \lesssim_{s,P} \|u\|_s$ by Theorem 3.35 ii) since $s \in \big(\lfloor q \rfloor - \frac{1}{2}, \lfloor q \rfloor\big)$ from (6.14). Thus (6.17), (6.21) and (6.23) imply that

$$B \lesssim \big|n'\big(u(x)\big)\big|^2 \big|u^{(m)}(x)-u^{(m)}(y)\big|^2 + \big|u^{(m)}(y)\big|^2 \big|n'\big(u(x)\big)-n'\big(u(y)\big)\big|^2$$

$$\lesssim_{q,s,P} \|u\|_s^{2(q-1)} \big|u^{(m)}(x)-u^{(m)}(y)\big|^2 + \big\|u^{(m)}\big\|_\infty^2 \|u\|_s^{2(q-2)} \big|u(x)-u(y)\big|^2$$

$$\lesssim_{s,P} \|u\|_s^{2(q-1)} \Big[\big|u^{(m)}(x)-u^{(m)}(y)\big|^2 + \big|u(x)-u(y)\big|^2\Big].$$

In total we have

$$\Big|\big(n(u)\big)^{(m)}(x) - \big(n(u)\big)^{(m)}(y)\Big|^2 \lesssim_s A + B \lesssim_{q,s,P} \|u\|_s^{2(q-1)} \sum_{i=0}^{m} \big|u^{(i)}(x)-u^{(i)}(y)\big|^2,$$

and so

$$\big|(n(u))^{(m)}\big|_{\mathrm{H}_P^\sigma}^2 \lesssim_{q,s,P} \|u\|_s^{2(q-1)} \sum_{i=0}^{m} \big|u^{(i)}\big|_{\mathrm{H}_P^\sigma}^2 \lesssim_s \|u\|_s^{2q}.$$

Since all P-dependent estimates have utilized Theorem 3.35 ii), the estimation constant has an upper bound which decreases with increasing P. □

Remark. The composition operator is likely continuous in all cases (6.12), which at least is easy to see for (6.12a), because

$$n(u) - n(v) = u^q - v^q = (u-v) \sum_{j=0}^{q-1} u^{q-j-1} v^j. \qquad \square$$

Based on Proposition 6.3 and Theorem 6.4 we study (6.9) for n and s given in (6.12).

6.3 Constrained minimization via a variational formulation

Suppose for the moment that some $u \in \mathrm{H}^s_P$ satisfies the traveling Whitham equation (6.9). In particular,

$$\langle Lu + n(u) - cu, v\rangle_s = 0 \quad \text{for all} \quad v \in \mathrm{H}^s_P. \tag{6.26}$$

By uniqueness of Fourier coefficients (Theorem 3.7 iii)) the relation (6.26) is actually equivalent to (6.9). The inner product may be replaced with that of H^r_P for all $r \in [0,s]$ because $\mathrm{H}^s_P \hookrightarrow \mathrm{H}^r_P$, and we pick henceforth simply the L^2_P inner product. In accordance with Section 5.1, observe from (6.26) that u can be viewed as a critical point for a functional $\tilde{\mathscr{E}} \colon \mathrm{H}^s_P \to \mathbb{R}$ with Fréchet derivative $\tilde{\mathscr{E}}'(u) = \langle Lu + n(u) - cu, \cdot\rangle_0$.

There will however be one optimization problem for $\tilde{\mathscr{E}}$ for each wave speed c. Due to the structure of the traveling Whitham equation (6.9) we may instead think of c as a Lagrange multiplier in the following constrained optimization problem. Construct functionals $\mathscr{E}, \mathscr{U} : \mathrm{H}^s_P \to \mathbb{R}$ with Fréchet derivatives

$$\mathscr{E}'(u) = \langle Lu + n(u), \cdot\rangle_0 \qquad \text{and} \qquad \mathscr{U}'(u) = \langle u, \cdot\rangle_0, \tag{6.27}$$

and establish the existence of a local extremizer for $\mathscr{E}$ under the constraint $\mathscr{U}(u) = \mu$, where $\mu > 0$ is a fixed parameter. The Lagrange multiplier theorem (Theorem 5.27 ii)) then gives the Euler-Lagrange equation $\mathscr{E}'(u) = c\,\mathscr{U}'(u)$, which is (6.26).

We pursue this last approach and shall prove the existence of a local maximizer for $\mathscr{E}$ on the constraint set $\{u \in \mathrm{H}^s_P : \mathscr{U}(u) = \mu\}$ with help of the generalized Weierstrass theorem (Theorem 5.23). Equivalently we consider minimization of $-\mathscr{E}$, from now on relabeled to $\mathscr{E}$, and the functionals $\mathscr{E}$ and $\mathscr{U}$ take the form

$$\mathscr{E}(u) = -\frac{1}{2}\langle Lu, u\rangle_0 - \frac{1}{q+1}\langle n(u), u\rangle_0 \qquad \text{and} \qquad \mathscr{U}(u) = \frac{1}{2}\|u\|_0^2. \tag{6.28}$$

For clarity we write $\mathscr{E} = \mathscr{L} + \mathscr{N}$, with the functionals $\mathscr{L}$ and $\mathscr{N}$ defined as

$$\mathscr{L}(u) = -\frac{1}{2}\langle Lu, u\rangle_0 \qquad \text{and} \qquad \mathscr{N}(u) = -\frac{1}{q+1}\langle n(u), u\rangle_0 = -\int_0^P N(u)\,\mathrm{d}x, \tag{6.29}$$

where $N : \mathbb{R} \to \mathbb{R}$ is the primitive of n (6.4) vanishing at the origin:

$$N(x) = \frac{x\,n(x)}{q+1} = \frac{1}{q+1}\begin{cases} x|x|^q & \text{if } n(x) = |x|^q; \\ |x|^{q+1} & \text{if } n(x) = x|x|^{q-1}. \end{cases} \tag{6.30}$$

The choice of $\mathscr{E}$ and $\mathscr{U}$ agrees with (6.27), as is verified by the following extension of [21, Proposition 2.2].

Proposition 6.6. *Let $s > \frac{1}{2}$. The maps $\mathscr{L}$ and $\mathscr{N}$ from (6.29) and $\mathscr{U}$ from (6.28) constitute bounded* C^1*-functionals*

$$\mathscr{L}, \mathscr{U} : \mathrm{H}^s_P \to \mathbb{R} \qquad \textit{and} \qquad \mathscr{N} : \mathrm{H}^s_P \to \mathbb{R} \quad \textit{for } s \textit{ as in (6.12)},$$

with Fréchet derivatives

$$\mathscr{L}'(u) = -\langle Lu, \cdot \rangle_0, \qquad \mathscr{N}'(u) = -\langle n(u), \cdot \rangle_0 \qquad \textit{and} \qquad \mathscr{U}'(u) = \langle u, \cdot \rangle_0. \tag{6.31}$$

In particular, $\mathscr{E}$ is a bounded C^1*-functional $\mathrm{H}^s_P \to \mathbb{R}$ for s as in (6.12).*

Proof. The arguments for $\mathscr{L}$ and $\mathscr{U}$ are direct consequences of Example 5.4 iii) and iv) combined with Proposition 6.3. As regards $\mathscr{N}$, we get boundedness by Theorem 3.35 ii):

$$|\mathscr{N}(u)| < \int_0^P |u|^{q+1} \,\mathrm{d}x \leq \|u\|_\infty^{q-1} \|u\|_0^2 \lesssim_{s,P} \|u\|_s^{q+1}. \tag{6.32}$$

Moreover, n is continuous $\mathrm{H}^s_P \to \mathrm{L}^2_P$, because, similar to the proof of Theorem 6.4,

$$\begin{aligned}\|n(u) - n(v)\|_0 &\leq \sup_{|z| \leq \max\left\{\|u\|_\infty, \|v\|_\infty\right\}} |n'(z)| \, \|u - v\|_0 \\ &\lesssim_{q,s,P} \max\left\{\|u\|_s^{q-1}, \|v\|_s^{q-1}\right\} \|u - v\|_s.\end{aligned}$$

We next calculate

$$\begin{aligned}\mathscr{N}'(u)\, v &= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \mathscr{N}(u + \epsilon v)\Big|_{\epsilon=0} \\ &= \int_0^P \frac{\partial}{\partial \epsilon} N(u + \epsilon v) \,\mathrm{d}x\Big|_{\epsilon=0} \\ &= \int_0^P n(u + \epsilon v)\, v \,\mathrm{d}x\Big|_{\epsilon=0} \\ &= \langle n(u), v \rangle_0\end{aligned}$$

with help of Leibniz' integral rule (Theorem 2.15) and the continuity of integrals (Remark 2). Linearity and continuity of $\mathscr{N}'(u)$ is inherited from the inner product. This implies continuity of the nonlinear map $\mathscr{N}'$, because n is continuous $\mathrm{H}^s_P \to \mathrm{L}^2_P$. □

6.3.1 A proof technique through penalization

Although minimization of $\mathscr{E}$ constrained to $\mathscr{U}(u) = \mu$ can be studied on all of H^s_P, the following result excludes global minimizers.

Proposition 6.7. *For all $q \in (1,\infty)$ the functional $\mathscr{E}$ is unbounded from below on the constraint set $\{u \in \mathrm{H}^s_P : \mathscr{U}(u) = \mu\}$.*

Proof. Note first by (6.10) that

$$\mathscr{L}(u) = -\frac{1}{2}\langle Lu, u\rangle_0 = -\frac{1}{2}\sum_{k\in\mathbb{Z}} m\big(\tfrac{2\pi k}{P}\big)\,|\widehat{u}_k|^2. \tag{6.33}$$

Since m is bounded by 1 via Proposition 6.1, we get

$$|\mathscr{L}(u)| \le \frac{1}{2}\|u\|_0^2 = \mathscr{U}(u) = \mu. \tag{6.34}$$

In order to see that $\mathscr{N}$ is unbounded, take a nonnegative $u \in \mathrm{C}^\infty(\mathbb{R})$ with support in $(0,P)$ and satisfying $\frac{1}{2}\|u\|^2_{\mathrm{L}^2(\mathbb{R})} = \mu$. Let $u_P \in \mathrm{C}^\infty_P \subset \mathrm{H}^s_P$ be its P-periodic extension with $\mathscr{U}(u_P) = \mu$ and introduce the scaling function w_λ through

$$w_\lambda(x) = \lambda^\alpha u_P(\lambda^\beta x),$$

where $\lambda, \alpha, \beta > 0$ are parameters. We require $\lambda \ge 1$ so that $\operatorname{supp} w_\lambda \subseteq \operatorname{supp} u_P$. Moreover, $\mathscr{U}(w_\lambda) = \lambda^{2\alpha-\beta}\mathscr{U}(u_P)$, and thus $\mathscr{U}(w_\lambda) = \mu$ if $2\alpha - \beta = 0$. With these choices we get

$$\mathscr{N}(w_\lambda) = \lambda^{(q+1)\alpha-\beta}\mathscr{N}(u_P) = \lambda^{(q-1)\alpha}\mathscr{N}(u_P).$$

Observe that $\mathscr{N}(u_P) < 0$ by the nonnegativity of u_P. Hence, because $q > 1$, we obtain

$$\mathscr{E}(w_\lambda) = \mathscr{L}(w_\lambda) + \mathscr{N}(w_\lambda) \le -\mu + \mathscr{N}(w_\lambda) \to -\infty \quad \text{as} \quad \lambda \to \infty. \qquad \square$$

According to Proposition 6.7 we restrict the search for (local) minimizers for $\mathscr{E}$ to some bounded subset of H^s_P. Two conflicting demands now arise: The generalized Weierstrass theorem requires a weakly closed subset, but we need an open subset for the Lagrange multiplier theorem. As was done in [21], based on techniques by Buffoni [9], our solution to this problem is as follows.

i) Desire the existence of a local minimizer for $\mathscr{E}$ in an open ball $B_R(0) \subset \mathrm{H}^s_P$ where we can use the Lagrange multiplier theorem.

ii) Look for a local minimizer in an open ball containing $B_R(0)$, say $B_{2R}(0)$, but penalize $\mathscr{E}$ as $R \le \|u\|_s \nearrow 2R$, leading to $\mathscr{E}(u) \to \infty$. This forces any such minimizer for $\mathscr{E}$ to lie in a weakly closed subset W, with $B_R(0) \subset W \subset B_{2R}(0)$, where we apply the generalized Weierstrass theorem. The Lagrange multiplier theorem can now be applied on $B_{2R}(0)$, but the Euler-Lagrange equation may include a penalty term.

iii) Show via *a priori* estimates that the minimizer actually lies in the ball $B_R(0)$ unaffected by the penalization. The minimizer then satisfies the original Euler-Lagrange equation.

We employ the above strategy, introducing a penalty function ϱ and considering constrained minimization of the augmented functional $\mathscr{E}_\varrho : V \to \mathbb{R}$, where

$$\mathscr{E}_\varrho(u) = \mathscr{E}(u) + \varrho\left(\|u\|_s^2\right) \tag{6.35}$$

and

$$V = \left\{u \in \mathrm{H}^s_P : \|u\|_s < 2R \text{ and } \mathscr{U}(u) = \mu\right\}. \tag{6.36}$$

Finally we establish via *a priori* estimates that minimizers for $\mathscr{E}_\varrho$ lie in the set

$$U = \left\{u \in \mathrm{H}^s_P : \|u\|_s < R \text{ and } \mathscr{U}(u) = \mu\right\}. \tag{6.37}$$

The estimation relies on the (fixed) parameters R and μ being sufficiently small and P being sufficiently large. In Assumption 6.8 below we collect the required specifications of ϱ; see [21, Section 3]. Property iii) comes first into use in Section 6.5. Note that the squared norm in (6.35) is chosen to guarantee C^1-regularity of $\mathscr{E}_\varrho$.

Assumption 6.8 (Properties of ϱ). *The penalty function $\varrho : \left[0, (2R)^2\right) \to [0, \infty)$ is chosen as a continuously differentiable map satisfying*

i) no penalty in the original region U: $\varrho(t) = 0$ when $t \in [0, R^2)$;

ii) weak coercivity: $\varrho(t) \nearrow \infty$ as $t \nearrow (2R)^2$, so that ϱ is strictly increasing on $\left[R^2, (2R)^2\right)$; and

iii) a bound on the derivative: For every constant $a \in (0,1)$ there exist constants $A, B > 0$ and $b > 1$ such that

$$\varrho'(t) \le A\varrho(t)^a + B\varrho(t)^b$$

for all $t \in \left[R^2, (2R)^2\right)$. □

A scaled variant of

$$t \mapsto \begin{cases} \dfrac{\mathrm{e}^{-1/(t-R^2)}}{(2R)^2 - t} & \text{if } t \in \left(R^2, (2R)^2\right); \\ 0 & \text{if } t \in \left[0, R^2\right] \end{cases}$$

serves as an example of such a penalty function; see [21, Section 3].

6.3.2 Exclusion of trivial solutions

It is clear that any constant function u_{const} satisfies the traveling Whitham equation (6.9) for some wave speed c, and from the constraint $\mathscr{U}(u_{\mathrm{const}}) = \mu$ they must take the form $u_{\mathrm{const}} = \pm\sqrt{2\mu/P}$. We now want to exclude these trivial waves as possible minimizers for $\mathscr{E}_\varrho$ on V and will show that

$$\inf_{u \in V} \mathscr{E}_\varrho(u) < \mathscr{E}_\varrho(u_{\mathrm{const}}).$$

It suffices to only consider $u_{\mathrm{const}} = +\sqrt{2\mu/P}$, as $\mathscr{E}_\varrho(\sqrt{2\mu/P}) \le \mathscr{E}_\varrho(-\sqrt{2\mu/P})$.

The upper bound on $\inf_{u\in V}\mathscr{E}_\varrho(u)$ also plays a crucial role in the rest of the existence argument. In particular, it gives an *a priori* lower bound on the wave speed c; see Lemma 6.15. Ehrnström *et al.* [21, Corollary 3.4] established a similar bound by comparing the periodic case with the solitary-wave problem. Their arguments, however, are more involved and may not be applicable for fractional orders s. We suggest here a different, simpler approach by constructing a concrete candidate for $\inf_{u\in V}\mathscr{E}_\varrho(u)$.

Lemma 6.9. *The inequality*

$$I_q = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sqrt{\tfrac{2}{3}}(1+\sin x)\right)^{q+1}\mathrm{d}x > 1$$

holds for all $q>1$.

Proof. Let

$$f(x)=\left(\sqrt{\tfrac{2}{3}}(1+\sin x)\right)^{2} \qquad \text{and} \qquad \varphi(x)=x^{(q+1)/2}$$

and observe from a version of Jensen's integral inequality [16, page 121] that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi\big(2\pi f(x)\big)\,\mathrm{d}x > \varphi\left(\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x\right),$$

because φ is strictly convex for $q>1$ and f is nonconstant. Written out each side becomes

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi\big(2\pi f(x)\big)\,\mathrm{d}x = (2\pi)^{\frac{q+1}{2}} I_q \qquad \text{and} \qquad \varphi\left(\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x\right) = (2\pi)^{\frac{q+1}{2}}. \qquad \square$$

Proposition 6.10 (Upper bound on the infimum of $\mathscr{E}_\varrho$). *Let $q\in(1,5)$ and a radius $R>0$ be given. Then there exist constants $\mu_{s,R}>0$ and $P_{q,\mu}>0$ such that*

$$\inf_{u\in V}\mathscr{E}_\varrho(u) < -\mu\left[1+\frac{2}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}\right] \tag{6.38}$$

for all $\mu\in(0,\mu_{s,R}]$ and $P\ge P_{q,\mu}$. In particular, trivial solutions do not minimize $\mathscr{E}_\varrho$ on V.

Proof. We show that, depending on μ and P, the smooth, P-periodic sinusoidal function

$$u(x)=\sqrt{\frac{4\mu}{3P}}\left(1+\sin\left(\frac{2\pi x}{P}\right)\right)$$

satisfies $\mathscr{E}_\varrho(u)<\mathscr{E}_\varrho(u_{\text{const}})$, where $u_{\text{const}}=\sqrt{2\mu/P}$ and the factor $\sqrt{4\mu/3P}$ is determined by the constraint $\mathscr{U}(u)=\mu$. The parameters μ and P are chosen such that $u,u_{\text{const}}\in U$, implying that $\mathscr{E}_\varrho$ coincides with $\mathscr{E}$ for these functions.

Evidently $\|u_{\text{const}}\|_s=(\widehat{u_{\text{const}}})_0=\sqrt{2\mu}$ and therefore $u_{\text{const}}\in U$ if $\mu<\frac{1}{2}R^2$. Two routine calculations, the first one using (6.33) and $m(0)=1$ from (6.1), give

$$\mathscr{L}(u_{\mathrm{const}}) = -\mu \qquad \text{and} \qquad \mathscr{N}(u_{\mathrm{const}}) = -\frac{2\mu}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}. \tag{6.39}$$

The nonzero Fourier coefficients of u are

$$\widehat{u}_0 = 2\sqrt{\frac{\mu}{3}} \qquad \text{and} \qquad \widehat{u}_{\pm 1} = \mp \mathrm{i}\sqrt{\frac{\mu}{3}},$$

and so

$$\|u\|_s^2 = \langle -1 \rangle_P^{2s} |\widehat{u}_{-1}|^2 + |\widehat{u}_0|^2 + \langle 1 \rangle_P^{2s} |\widehat{u}_1|^2 = \frac{2\mu}{3}\left(2 + \langle 1 \rangle_P^{2s}\right),$$

which implies that $u \in U$ provided $\mu \le \widetilde{\mu}_{s,R}$ and $P \ge \widetilde{P}_\mu$ for some sufficiently small $\widetilde{\mu}_{s,R} > 0$ and large $\widetilde{P}_\mu > 0$. From (6.33) and the evenness of m (Proposition 6.1) we next obtain

$$\mathscr{L}(u) = -\frac{1}{2}\left[m\left(-\tfrac{2\pi}{P}\right)|\widehat{u}_{-1}|^2 + m(0)\,|\widehat{u}_0|^2 + m\left(\tfrac{2\pi}{P}\right)|\widehat{u}_1|^2\right] = -\frac{\mu}{3}\left(2 + m\left(\tfrac{2\pi}{P}\right)\right),$$

and

$$\mathscr{N}(u) = -\frac{2\mu}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}} I_q$$

by a change of variables, where I_q is the integral in Lemma 6.9. Since $q > 1$, we have $I_q > 1$. A combination of these expressions with (6.39) and the Taylor expansion of m (6.2) then result in

$$\begin{aligned}
\mathscr{E}(u) - \mathscr{E}(u_{\mathrm{const}}) &= \frac{\mu}{3}\left(1 - m\left(\frac{2\pi}{P}\right)\right) - \frac{2\mu}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}(I_q - 1) \\
&= \mu\left[\frac{1}{18}\left(\frac{2\pi}{P}\right)^2 + \mathscr{O}\left(\left(\frac{2\pi}{P}\right)^4\right) - \frac{2}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}(I_q - 1)\right] \\
&= \tfrac{2}{9}\pi^2\mu\left[P^{-2} + \mathscr{O}\left(P^{-4}\right) - C_{q,\mu}\,P^{-(q-1)/2}\right],
\end{aligned}$$

where $C_{q,\mu} > 0$ is a constant. If $(q-1)/2 < 2$, that is, if $q < 5$, we infer that

$$\mathscr{E}(u) - \mathscr{E}(u_{\mathrm{const}}) < 0$$

for all sufficiently large periods $P \ge \widetilde{P}_{q,\mu} > 0$. In conclusion,

$$\inf_{v \in V} \mathscr{E}_\varrho(v) \le \mathscr{E}(u) < \mathscr{E}(u_{\mathrm{const}})$$

for $\mu \le \mu_{s,R} = \min\left\{\frac{1}{2}R^2, \widetilde{\mu}_{s,R}\right\}$ and $P \ge P_{q,\mu} = \max\left\{\widetilde{P}_\mu, \widetilde{P}_{q,\mu}\right\}$, which is (6.38). □

In accordance with Proposition 6.10 we focus on exponents $q \in (1,5)$ from now on. We remark, however, that Ehrnström and Kalisch [22, Remark 3.3] established periodic traveling waves with $n(x) = x^q$ for every $2 \le q \in \mathbb{Z}_+$ via bifurcation techniques. It is not known by the author if $q<5$ is necessary in this constrained variational setting, but the bound in [21, Corollary 3.4], which excludes trivial solutions similarly as Proposition 6.10, required $q<5$ in its proof for the periodic case. This limit agrees for example also with results by Iorio and Iorio [40, Theorem 5.32], who considered a generalized nonlinear Schrödinger equation with $n(x)=x|x|^{q-1}$ for $q\in\mathbb{Z}_+$.

6.4 Existence of a solution to the penalized problem

By construction $\mathscr{E}_\varrho$ is weakly coercive on V and the possible minimizers for $\mathscr{E}_\varrho$ will lie in a weakly closed subset $W\subset V$. Combined with weak continuity of the considered functionals, we can then apply the generalized Weierstrass theorem (Theorem 5.23) and the Lagrange multiplier theorem (Theorem 5.27). The result is based on a short argument in [21, Lemma 3.1] for H^1_P, but we provide all details in the fractional setting.

Lemma 6.11 (Weak continuity). *The functionals $\mathscr{L}$, $\mathscr{U}$, $\mathscr{N}$ and $\mathscr{E}$ are weak-strong continuous on H^s_P. In particular, $\mathscr{E}_\varrho$ is w.l.s.c. on $B_R(0)$.*

Proof. From Theorem 3.35 there is a compact embedding $\mathrm{H}^s_P\hookrightarrow\mathrm{H}^r_P$ whenever $r\in\left(\frac{1}{2},s\right)$. Moreover, Proposition 6.6 gives continuity of both $\mathscr{L}$ and $\mathscr{U}$ on H^r_P and thus they are weak-strong continuous on H^s_P by Example 5.14 ii).

Switching focus to $\mathscr{N}$, suppose that $u_n\rightharpoonup u$ in H^s_P. As $s>\frac{1}{2}$, Example 5.14 iii) shows that $u_n\to u$ (uniformly) in C_P. Then also $N(u_n)\to N(u)$ in C_P, and the dominated convergence theorem (Theorem 2.14) or even simpler, the uniform limit theorem for Riemann integrable functions, yields

$$\int_0^P N(u_n)\,\mathrm{d}x\to\int_0^P N(u)\,\mathrm{d}x\quad\text{as}\quad n\to\infty.$$

Therefore $\mathscr{N}(u_n)\to\mathscr{N}(u)$ and so $\mathscr{N}$ is weak-strong continuous on H^s_P. The same conclusion applies to $\mathscr{E}=\mathscr{L}+\mathscr{N}$.

Since ϱ is continuous and increasing and $\|\cdot\|_s^2$ is w.l.s.c. by Example 5.16, it follows from Proposition 5.15 that the composition $\varrho\big(\|\cdot\|_s^2\big)$ is w.l.s.c. on $B_R(0)$. Thus $\mathscr{E}_\varrho$ also has this property, being the sum of two w.l.s.c. functionals. □

Theorem 6.12. *The functional $\mathscr{E}_\varrho$ in (6.35) has a minimizer u on V which satisfies the Euler-Lagrange equation*

$$\mathscr{E}_\varrho'(u)+c\,\mathscr{U}'(u)=0 \tag{6.40}$$

for some Lagrange multiplier $c\in\mathbb{R}$.

Proof. Recall that H^s_P is a reflexive Banach space. The composition $\varrho\big(\|\cdot\|_s^2\big)$ is weakly coercive on V, because ϱ is weakly coercive on $\big[0,(2R)^2\big)$ by Assumption 6.8 ii). Hence, the boundedness of $\mathscr{E}$ on V (Proposition 6.6) yields weak coercivity of $\mathscr{E}_\varrho$ on V. The search for minimizers may therefore be limited to the subset

$$\begin{aligned} W &= \{u \in V : \|u\|_s \leq S\} \\ &= \overline{B}_S(0) \cap \big\{u \in \mathrm{H}^s_P : \mathscr{U}(u) = \mu\big\}, \end{aligned}$$

where $S \geq R$ is sufficiently close to but strictly less than $2R$.

It is clear that W is bounded. Moreover, the closed and convex ball $\overline{B}_S(0)$ is weakly closed (Proposition 5.18), and a trivial modification of Example 5.19 implies weak closedness of $\{u \in \mathrm{H}^s_P : \mathscr{U}(u) = \mu\}$. Thus W is weakly closed, being the intersection of two weakly closed sets. Since $\mathscr{E}_\varrho$ is w.l.s.c. on $B_{2R}(0)$ by Lemma 6.11, the same property holds true for $\mathscr{E}_\varrho\restriction_W$ and we deduce the existence of a minimizer $u \in W \subset V$ for $\mathscr{E}_\varrho$ on V via Theorem 5.23. (To ease the notation we drop the bar above u.)

The Euler-Lagrange equation (6.40) is now a consequence of Theorem 5.27 applied at the extremal point u. Indeed, note from Proposition 6.6 that $\mathscr{E}, \mathscr{U} \in \mathrm{C}^1(B_{2R}(0), \mathbb{R})$. The composition $\varrho\big(\|\cdot\|_s^2\big)$ is likewise continuously differentiable by Assumption 6.8 and the chain rule for Fréchet derivatives (Theorem 5.5). Thus $\mathscr{E}_\varrho \in \mathrm{C}^1(B_{2R}(0), \mathbb{R})$ as well. Lastly, the case $\mathscr{U}'(u) = 0$ is excluded by

$$\mathscr{U}'(u)u = \langle u, u\rangle_0 = 2\mathscr{U}(u) = 2\mu > 0,$$

and so Theorem 5.27 ii) implies the existence of $c \in \mathbb{R}$ such that (6.40) holds. □

Based on (6.31) the Euler-Lagrange equation (6.40) takes the form

$$\langle Lu + n(u) - cu, v\rangle_0 - 2\varrho'\big(\|u\|_s^2\big)\langle u, v\rangle_s = 0 \quad \text{for all} \quad v \in \mathrm{H}^s_P, \tag{6.41}$$

or, equivalently,

$$\sum_{k\in\mathbb{Z}} \Big[(\widehat{Lu})_k + \widehat{n(u)}_k - c\,\widehat{u}_k - 2\varrho'\big(\|u\|_s^2\big)\langle k\rangle_P^{2s}\,\widehat{u}_k\Big]\overline{\widehat{v}_k} = 0 \quad \text{for all} \quad v \in \mathrm{H}^s_P.$$

This relation holds especially for all v in the Fourier basis of complex exponentials $\{\varphi_k\}_{k\in\mathbb{Z}}$ for H^s_P, and so orthogonality yields

$$(\widehat{Lu})_k + \widehat{n(u)}_k - c\,\widehat{u}_k - 2\varrho'\big(\|u\|_s^2\big)\langle k\rangle_P^{2s}\,\widehat{u}_k = 0 \quad \text{for all} \quad k \in \mathbb{Z}. \tag{6.42}$$

If $\varrho'\big(\|u\|_s^2\big) = 0$, then

$$(\widehat{Lu})_k + \widehat{n(u)}_k - c\,\widehat{u}_k = 0 \quad \text{for all} \quad k \in \mathbb{Z}, \tag{6.43}$$

which, by completeness of the Fourier basis (Theorem 3.7 iii)), means that u solves the traveling Whitham equation (6.9) in H^s_P.

The remaining possibility is that $\varrho'\big(\|u\|_s^2\big) > 0$. Via *a priori* estimates we shall, similarly as [21], exclude this case by establishing the small-amplitude bound

$$\|u\|_s^2 \lesssim_s \mu \tag{6.44}$$

for all sufficiently small R and μ and large periods P, provided $\varrho'\big(\|u\|_s^2\big) > 0$. Eventually $\|u\|_s < R$, so that $u \in U$ and ϱ vanishes. Since ϱ is strictly increasing on $\big[R^2, (2R)^2\big)$, it follows that $\varrho'(t) = 0$ if and only if $\varrho(t) = 0$. Thus ϱ' also vanishes and this case never occurs.

Note that for the case $\varrho'\big(\|u\|_s^2\big) = 0$ we as of yet only know that $u \in V$ with $\|u\|_s \leq R$. In Section 6.5, however, we get (6.44) via (6.13) at least for s as in (6.12a) and (6.14).

6.5 A priori estimates bounding the H^s_P norm

As a motivation for the estimation we first deduce an inequality from the Euler-Lagrange equation (6.40), inspired by [21, (20)–(21) and first part of Lemma 3.6].

Lemma 6.13. *Each minimizer u of $\mathcal{E}_\varrho$ on V with $\varrho'\big(\|u\|_s^2\big) > 0$ is in C^∞_P.*

Proof. By construction $u \in V \subset \mathrm{H}^s_P$ and in particular $Lu, n(u) \in \mathrm{H}^s_P$. Consequently the sequence $\widehat{Lu} + \widehat{n(u)} - c\,\widehat{u}$ is in the Sobolev sequence space h^s_P. From (6.42) we therefore also get that $\langle\cdot\rangle_P^{2s}\,\widehat{u} \in h^s_P$, which by Definition 3.33 implies $\widehat{u} \in h^{3s}_P$, or, equivalently, $u \in \mathrm{H}^{3s}_P$. Iterating the argument then gives $u \in \mathrm{C}^\infty_P$. □

Corollary 6.14. *Each minimizer u of $\mathcal{E}_\varrho$ on V satisfies*

$$c\|u\|_s^2 \lesssim \|u\|^2_{s-\frac{1}{4}} + \langle n(u), u\rangle_s. \tag{6.45}$$

Proof. If $\varrho'\big(\|u\|_s^2\big) > 0$, the function $v = \mathcal{F}^{-1}\left(\langle\cdot\rangle_P^{2s}\,\widehat{u}\right)$ is in H^s_P by Lemma 6.13, because

$$\|v\|_s = \|\widehat{v}\|_{h^s_P} = \left\|\langle\cdot\rangle_P^{2s}\,\widehat{u}\right\|_{h^s_P} = \|u\|_{3s} < \infty.$$

Inserting this v into the Euler-Lagrange equation (6.41) yields

$$c\|u\|_s^2 = \langle Lu + n(u), u\rangle_s - 2\varrho'\big(\|u\|_s^2\big)\|u\|_{2s}^2 \leq \langle Lu + n(u), u\rangle_s, \tag{6.46}$$

where it is used that $\langle\cdot, v\rangle_0 = \langle\cdot, u\rangle_s$ and $\varrho' > 0$. If, however, $\varrho'\big(\|u\|_s^2\big) = 0$, we obtain (6.46) with equality by multiplying (6.43) by $\langle k\rangle_P^{2s}\,\overline{\widehat{u}_k}$ and summing over $k \in \mathbb{Z}$.

The estimate (6.11) of m next gives

$$\langle Lu,u\rangle_s = \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2s}\, m\left(\tfrac{2\pi k}{P}\right) |\widehat{u}_k|^2 \lesssim \sum_{k\in\mathbb{Z}} \langle k\rangle_P^{2\left(s-\frac{1}{4}\right)} |\widehat{u}_k|^2 = \|u\|_{s-\frac{1}{4}}^2,$$

and thus (6.46) becomes (6.45). □

Similar to [21, Lemma 3.6], suppose now that the Lagrange multiplier c is bounded strictly away from 0, that is,

$$c \geq M > 0. \tag{6.47}$$

If we have

$$\langle n(u),u\rangle_s \leq M_{q,s,R,P,\mu}\, \|u\|_s^2, \tag{6.48}$$

where $M_{q,s,R,P,\mu}$ is a constant satisfying $M_{q,s,R,P,\mu} < M$ for all sufficiently small radii R, small μ and/or large periods P, depending on q and s, then (6.45) implies that

$$\|u\|_s^2 \lesssim \frac{1}{c - M_{q,s,R,P,\mu}} \|u\|_{s-\frac{1}{4}}^2 \lesssim \|u\|_{s-\frac{1}{4}}^2 \tag{6.49}$$

for such R, μ and P. The Sobolev interpolation inequality (Theorem 3.37) with exponent $\theta = 1/4s$ results in

$$\|u\|_{s-\frac{1}{4}} \leq \|u\|_0^{\frac{1}{4s}}\, \|u\|_s^{1-\frac{1}{4s}},$$

which inserted in (6.49) with rearrangement gives

$$\|u\|_s^2 \lesssim_s \|u\|_0^2 = 2\mu.$$

Our search for (6.44) thus hinges on the validity of (6.47) and (6.48).

The following result confirms (6.47) and is an extension of [21, Lemma 3.5] using interpolation. Additionally, the first part of the bound is explicit in terms of μ and P.

Lemma 6.15 (Lower bound on the wave speed). *Let the constants $\mu_{s,R} > 0$ and $P_{q,\mu} > 0$ be as in Proposition 6.10. Then for all $\mu \in (0,\mu_{s,R}]$ and $P \geq P_{q,\mu}$ the Lagrange multiplier satisfies*

$$c > 1 + \left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}} - \Lambda_\varrho\, \mu^\delta\, R^{2+\epsilon}, \tag{6.50}$$

where $\delta = \delta_{q,s} > 0$ and $\epsilon = \epsilon_{q,s} > 0$, and $\Lambda_\varrho = \Lambda_{\varrho,q,s} \geq 0$ vanishes when $\varrho\left(\|u\|_s^2\right) = 0$.

Proof. The Euler-Lagrange equation (6.46) with $v = u$ yields

$$2c\mu = c\langle u,u\rangle_0 = \langle Lu + n(u),u\rangle_0 - 2\varrho'\left(\|u\|_s^2\right)\|u\|_s^2. \tag{6.51}$$

Using estimate (6.38), valid for all $\mu \in (0, \mu_{s,R}]$ and $P \geq P_{q,\mu}$, the following trick gives

$$\begin{aligned}\langle Lu + n(u), u\rangle_0 &= (q+1)\left\langle \frac{1}{2}Lu + \frac{1}{q+1}n(u), u\right\rangle_0 - \frac{q-1}{2}\langle Lu, u\rangle_0 \\ &= (q+1)\Big[-\mathscr{E}_\varrho(u) + \varrho\big(\|u\|_s^2\big)\Big] + (q-1)\mathscr{L}(u) \\ &> (q+1)\mu\left[1 + \frac{2}{q+1}\left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}\right] - (q-1)\mu \\ &= 2\mu\left[1 + \left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}\right],\end{aligned}$$

where we have used that $\varrho \geq 0$ and $-\mathscr{L}(u) \leq \mu$ via (6.33). Thus (6.51) becomes

$$c > 1 + \left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}} - \mu^{-1}\varrho'\big(\|u\|_s^2\big) \cdot 4R^2,$$

with the last part nonpresent when $\varrho\big(\|u\|_s^2\big) = 0$. If this is not the case, we want

$$\varrho'\big(\|u\|_s^2\big) \lesssim \mu^{1+\delta} \tag{6.52}$$

for some $\delta \geq 0$ in order to control this last part.

To this end, notice that (6.38) in particular gives

$$\mathscr{L}(u) + \mathscr{N}(u) + \varrho\big(\|u\|_s^2\big) < -\mu,$$

and so (6.32) and (6.34) imply

$$\varrho\big(\|u\|_s^2\big) < -\mathscr{N}(u) < \|u\|_\infty^{q-1}\|u\|_0^2.$$

Based on an idea from [39], we next bound $\|u\|_\infty$ by $\|u\|_0^\theta$ for some $\theta \in (0,1)$ via interpolation. Since $s > \frac{1}{2}$, define $\theta = \theta_s$ such that $r_s = (1-\theta)s \in \big(\frac{1}{2}, s\big)$. Then Theorem 3.35 and Theorem 3.37 result in

$$\|u\|_\infty \lesssim_{r_s} \|u\|_{r_s} \leq \|u\|_0^\theta \|u\|_s^{1-\theta} \lesssim_s \|u\|_0^\theta R^{1-\theta},$$

and the first estimation constant is independent of $P \geq P_{q,\mu}$. Thus

$$\begin{aligned}\varrho\big(\|u\|_s^2\big) &\lesssim_s \|u\|_0^{2+(q-1)\theta} R^{(q-1)(1-\theta)} \\ &\lesssim_{q,s} \mu^{1+\tilde{\delta}} R^{(q-1)(1-\theta)},\end{aligned} \tag{6.53}$$

where $\tilde{\delta} = (q-1)\theta/2 > 0$ since $q > 1$. The validity of (6.52), considering only the μ-term in (6.53), now follows from Assumption 6.8 iii). Indeed, if $\varrho \geq 1$, we have $\varrho' \lesssim \varrho^b$ for some $b > 1$, which gives (6.52) with $\delta = b(1+\widetilde{\delta}) - 1 > 0$. If, however, $\varrho < 1$, then $\varrho' \lesssim \varrho^a$ with $a \in (0,1)$. Choosing a sufficiently close to 1 such that $a(1+\widetilde{\delta}) > 1$, establishes (6.52) with $\delta = a(1+\widetilde{\delta}) - 1 > 0$.

Estimate (6.50) now appears with

$$\delta_{q,s} = \delta = x(1+\widetilde{\delta}) - 1 = \frac{x}{2}\big(2 + (q-1)\theta_s\big) - 1 > 0$$

and

$$\epsilon_{q,s} = 2 + x(q-1)(1-\theta_s) > 0,$$

where $x = a$ or b. □

Lemma 6.16. *There exists a constant $\gamma = \gamma_q > 0$ such that*

$$\langle n(u), u\rangle_s \lesssim_{q,s,P} R^\gamma \|u\|_s^2$$

uniformly over all minimizers u of $\mathcal{E}_\varrho$ on V with $\varrho'\big(\|u\|_s^2\big) > 0$, or regardless of the value of $\varrho'\big(\|u\|_s^2\big)$ for s in (6.12a) and (6.14). The estimation constant decreases with increasing P.

Proof. If $\varrho'\big(\|u\|_s^2\big) > 0$, we know that $u \in \mathrm{H}_P^{2s}$ from Lemma 6.13, and thus

$$\langle n(u), u\rangle_s = \sum_{k\in\mathbb{Z}} \widehat{n(u)}_k \, \overline{\langle k\rangle_P^{2s}\, \widehat{u}_k} \leq \|n(u)\|_0 \|u\|_{2s}$$

by the Cauchy-Schwarz inequality. Theorem 3.35 ii) next yields

$$\|n(u)\|_0 \leq \|u\|_\infty^{q-1} \|u\|_0 \lesssim_{s,P} \|u\|_s^q,$$

and with Theorem 3.37 we interpolate

$$\|u\|_{2s} \leq \|u\|_s^{\theta_q} \|u\|_{t_{q,s}}^{1-\theta_q} \lesssim_{q,s} \|u\|_s^{\theta_q},$$

where $t_{q,s} = (2-\theta_q)s/(1-\theta_q) > 2s$ and $\|u\|_{t_{q,s}} < \infty$ by Lemma 6.13. Since $q > 1$, we can choose $\theta_q \in (0,1)$ such that $q + \theta_q > 2$. Hence, as $u \in V$, put $\gamma = q + \theta_q - 2 > 0$ to arrive at

$$\langle n(u), u\rangle_s \lesssim_{s,P} \|u\|_s^q \|u\|_{2s} \lesssim_{q,s} \|u\|_s^{q+\theta_q} \lesssim_q R^\gamma \|u\|_s^2.$$

For s in (6.12a) and (6.14), we have

$$\langle n(u), u\rangle_s \leq \|n(u)\|_s \|u\|_s \lesssim_{q,s,P} R^\gamma \|u\|_s^2$$

from (6.13) with $\gamma = q - 1 > 0$. □

We can now deduce the existence result, which is based on [21, Theorem 3.7], but we do a detailed proof with respect to the parameters q, s, P, R and μ.

Theorem 6.17 (Existence of periodic traveling waves). *Let $q \in (1,5)$ and s be as in (6.12). Then for each sufficiently small $\mu = \mu_{q,s} > 0$ there exists a period $P_{q,\mu} > 0$, such that for all $P \geq P_{q,\mu}$ the traveling Whitham equation (6.9) admits a nontrivial solution $u \in H_P^s$, with $\|u\|_0^2 = 2\mu$. The wave speed satisfies*

$$c > 1 + \left(\frac{2\mu}{P}\right)^{\frac{q-1}{2}}, \tag{6.54}$$

and at least for s given by (6.12a) and (6.14) the estimates

$$\|u\|_s^2 \lesssim_s \mu \qquad \text{and} \qquad c < M_{q,s}$$

hold uniformly over all $P \geq P_{q,\mu}$, where $M_{q,s}$ is a constant.

Remark. Without improved regularity, the periodic traveling wave in Theorem 6.17 is a classical solution to the generalized Whitham equation (6.6) when $s > \frac{3}{2}$, which follows by the embedding $H_P^s \hookrightarrow C_P^1$ from Theorem 3.35 ii). It is, however, possible to gain regularity depending on the smoothness of the nonlinearity; see [21, Lemma 2.3] in the solitary-wave setting and Hur and Johnson [37, Lemma 2]. In particular, traveling waves for the original Whitham equation are smooth. □

Proof (of Theorem 6.17). We choose $R = R_{q,s} > 0$ sufficiently small such that (6.50) in Lemma 6.15 implies (6.47) for some $M > 0$, valid for all $\mu \in (0, \mu_{s,R}]$ and $P \geq P_{q,\mu}$. Then, if necessary, we let R be even smaller and apply Lemma 6.16 to get (6.48) with $M_{q,s,R,P,\mu} < M$ for all $\mu \in (0, \mu_{s,R}]$ and $P \geq P_{q,\mu}$. This is true if $\varrho'\big(\|u\|_s^2\big) > 0$ or regardless of the value of $\varrho'\big(\|u\|_s^2\big)$ for s in (6.12a) and (6.14). In fact, $M_{q,s,R,P,\mu}$ can be chosen to depend only on $R_{q,s}$.

Thus $\|u\|_s^2 \lesssim_s \mu$ and we now choose $\mu = \mu_s \in (0, \mu_{s,R}]$ so that $\|u\|_s < R$. Therefore both $\varrho\big(\|u\|_s^2\big)$ and $\varrho'\big(\|u\|_s^2\big)$ vanish and u is a nontrivial (by Proposition 6.10) solution to (6.9). In total, μ depends on s and R, that is, on q and s.

Since $\varrho\big(\|u\|_s^2\big) = 0$, Lemma 6.15 next gives (6.54). Moreover, with (6.13), (6.31) and (6.33) we estimate

$$-\mathcal{E}'(u)\,u = \langle Lu, u\rangle_0 + \langle n(u), u\rangle_0 \lesssim_{q,s} 2\,\|u\|_0^2 + R^{q-1}\|u\|_s^2 \lesssim_{q,s} \|u\|_s^2 \lesssim_s \mu,$$

where the first estimation constant is independent of $P \geq P_{q,\mu}$. Hence,

$$c = -\frac{\mathcal{E}'(u)\,u}{\mathcal{U}'(u)\,u} \lesssim_{q,s} \frac{\mu}{2\mu} \lesssim 1$$

by the Euler-Lagrange equation (6.40). □

6.6 SUGGESTED FURTHER WORK

We lastly list some possible improvements of Theorem 6.17 and directions for further study.

- ⋆ Obtain the estimate (6.13) for all s in (6.12b);
- ⋆ expand our results to a more general class of nonlocal evolution equations, including those considered in [21, Assumptions]. In particular, consider a class of Fourier multiplier operators of Whitham-type and nonhomogeneous nonlinearities;
- ⋆ investigate the necessity of $q > 1$ and/or $q < 5$ for the nonlinearity exponent in this constrained variational setting;
- ⋆ study the traveling Whitham equation (6.9) for $s \leq \frac{1}{2}$, possibly with nonlinearity exponents $q \leq 1$;
- ⋆ establish regularity of the solution(s) and its dependence on the regularity of the nonlinearity;
- ⋆ prove uniqueness of the solution; and
- ⋆ get estimates for the sizes of $\mu_{q,s}$, the minimal period $P_{q,\mu}$ and $M_{q,s}$.

Appendix A

FUNDAMENTAL CHARACTERISTICS OF WAVES

The rudimentary features of waves seem best described by the prototype of a wave function: A time-dependent, onedimensional sine wave η of the form

$$\eta(x,t) = h + A\sin\left[2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right) + \phi\right] = h + A\sin(kx - \omega t + \phi) \tag{A.1}$$

at position x on the real line at time t modeling simple harmonic motion. Often it is also beneficial to consider η as a complex exponential

$$\eta(x,t) = h + \tilde{A}\mathrm{e}^{\mathrm{i}\left(kx - \omega t + \tilde{\phi}\right)}.$$

The factor A denotes the *amplitude* of the wave and determines in particular the height of crests and depth of troughs relative to the normal *height h*. Moreover, by inspection we see that

$$\eta(x + \lambda, t) = \eta(x,t)$$

for all x and t, and λ is therefore naturally called the *wavelength*, most easily measured as the length between any two consecutive crests/troughs. The associated parameter $k = 2\pi/\lambda$ is known as the *wavenumber*, which more generally (in higher dimensions) defines a vector pointing in the direction of propagation. Similarly the *period T* satisfies

$$\eta(x, t + T) = \eta(x,t)$$

and quantifies the time it takes for the wave to repeat itself. Next comes the *(angular) frequency $\omega = 2\pi/T$*. The argument $kx - \omega t + \phi$ in (A.1) is called the *phase* of η, with ϕ being the *phase angle* corresponding to a horizontal shift of the wave. We mention also that more complicated waves can for example look like

$$\eta(x,t) = A(x,t)\,\mathrm{e}^{\mathrm{i}\theta(x,t)},$$

with position and time-dependence on the amplitude and involved expressions for the phase θ, while even more general waves $\eta(x,t)$ arising as solutions to differential equations may not possess a closed-form expression.

The sinusoidal wave (A.1) is the archetype of a *periodic traveling wave*. We say that a time-dependent wave $\eta(x,t)$ is *traveling* or *steady* if it can be written on the form $\eta(x,t) = u(x-ct)$ for some function u and real parameter c named the *phase velocity*. Observe that the phase velocity of the sine wave (A.1) equals

$$c = \frac{\lambda}{T} = \frac{\omega}{k}. \tag{A.2}$$

If $c > 0$, the wave is right-going, while $c < 0$ determines left-going motion. Steady waves therefore propagate with constant shape in time with speed $|c|$. Furthermore, when the function u in addition is *periodic with period $P > 0$*, or *P-periodic*, that is, if

$$u(x+P) = u(x) \quad \text{for all} \quad x \in \mathbb{R},$$

then we call η a (P-)*periodic traveling wave*. Note that P should not be confused with T; the sine wave (A.1) has for example P equal to (an arbitrary multiple of) $\lambda = cT$.

The other important kind of traveling waves is the class of *solitary waves* satisfying the decay

$$|\eta(x,t)| = |u(x-ct)| \to 0 \quad \text{as} \quad x-ct \to \pm\infty.$$

Physically these waves are localized in nature and eventually die out. Famous for his reported observation in 1844 of a solitary water wave on a canal, the naval engineer Russell called it the *wave of translation* [55].

Notice from (A.2) that the phase velocity of a sinusoidal wave (A.1) with a given frequency ω increases with decreasing wavenumbers k. We call this fundamental effect *dispersion* and infer that waves with large wavelengths λ travel faster than small ones. The *dispersion relation* of a wave expresses the relationship between ω and k, and for the sine wave (A.1) we have $\omega(k) = ck$. For more general waves the dispersion relation can be more complex, and the frequency ω may in fact also depend on other quantities such as the amplitude. In these cases we define the phase velocity through $c = \omega(k,A)/k$.

The last basic wave characteristic we consider is that of *group velocity*. In the sine wave (A.1) all points of the wave travel with equal speed. However, if we for example have a group of sinusoidal waves with different wavenumbers, then the phase velocity may vary locally within the group. It is nevertheless possible to give a notion of the average velocity of the group. As an illustration we consider a pack of two sine waves

$$\eta(x,t) = \sin(k_1 x - \omega_1 t) + \sin(k_2 x - \omega_2 t)$$

with $k_1 \approx k_2$ and $\omega_1 \approx \omega_2$, but neither identical. By a trigonometric identity η is equivalent to

$$\eta(x,t) = 2\cos\left(\frac{k_1-k_2}{2}x - \frac{\omega_1-\omega_2}{2}t\right)\sin\left(\frac{k_1+k_2}{2}x - \frac{\omega_1+\omega_2}{2}t\right).$$

The last term is a standard traveling sine wave with wavenumber $k \approx k_1 \approx k_2$ and frequency $\omega \approx \omega_1 \approx \omega_2$, whereas the first factor plays the role of a slowly changing amplitude. The phase speed of this leftmost term is

$$\frac{\omega_1 - \omega_2}{k_1 - k_2} \approx \frac{\mathrm{d}\omega}{\mathrm{d}k}$$

and expresses the speed of the whole wave's envelope. We call therefore

$$c_\mathrm{g} = \frac{\mathrm{d}\omega}{\mathrm{d}k}$$

the *group velocity* associated with η and note that for simple sine waves (A.1) this concept coincides with the usual phase velocity, as expected.

More details of all the above wave features can be found in standard books on wave physics or fluid mechanics with emphasis on surface waves. We followed, however, partly the mathematically-oriented lecture notes by Ehrnström [17].

References

[1] L. C. Adams, *Partial Differential Equaions*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, 2 ed., March 2010.

[2] F. Albiac and N. J. Kalton, *Topics in Banach Space Theory*, vol. 233 of Graduate Texts in Mathematics, Springer New York, 2006.

[3] T. B. Benjamin, J. L. Bona, and J. J. Mahony, *Model Equations for Long Waves in Nonlinear Dispersive Systems*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 272 (1972), pp. 47–78.

[4] T. B. Benjamin and K. Hasselmann, *Instability of Periodic Wavetrains in Nonlinear Dispersive Systems [and Discussion]*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 299 (1967), pp. 59–76.

[5] H. Borluk, H. Kalisch, and D. P. Nicholls, *A Numerical Study of the Whitham Equation as a Model for Steady Surface Water Waves*, to appear, Journal of Computational and Applied Mathematics, (2013).

[6] N. Bottman and B. Deconinck, *KdV cnoidal waves are spectrally stable*, Discrete and Continuous Dynamical Systems, 25 (2009), pp. 1163–1180.

[7] J. C. Bronski and M. A. Johnson, *The Modulational Instability for a Generalized Korteweg-de Vries Equation*, Archive for Rational Mechanics and Analysis, 197 (2010), pp. 357–400.

[8] G. Brüll, M. Ehrnström, and L. Pei, *Symmetry and decay properties of travelling wave solutions to the Whitham equation*. Url: `https://wiki.math.ntnu.no/_media/dispersive15/bruell.pdf` [accessed Nov 21, 2015], 2015. Work in preparation.

[9] B. Buffoni, *Existence and Conditional Energetic Stability of Capillary-Gravity Solitary Water Waves by Minimisation*, Archive for Rational Mechanics and Analysis, 173 (2004), pp. 25–68.

[10] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Mathematica, 116 (1966), pp. 135–157.

[11] P. R. CHERNOFF, *Pointwise Convergence of Fourier Series*, The American Mathematical Monthly, 87 (1980), pp. 399–400.

[12] F. CLARKE, *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264 of Graduate Texts in Mathematics, Springer, London, 2013.

[13] B. DACOROGNA, *Direct Methods in the Calculus of Variations*, vol. 78 of Applied Mathematical Sciences, Springer-Verlag, 2 ed., 2008.

[14] ———, *Introduction to the Calculus of Variations*, Imperial College Press, 2 ed., 2009.

[15] F. DEMENGEL AND G. DEMENGEL, *Functional Spaces for the Theory of Elliptic Partial Differential Equations*, Universitext, Springer-Verlag London, 2012. translated by R. Erné.

[16] P. DUREN, *Invitation to Classical Analysis*, vol. 17 of Pure and Applied Undergraduate Texts, American Mathematical Society, University of Michigan, Ann Arbor, MI, 2012.

[17] M. EHRNSTRÖM, *Waves and solitons*. Lecture notes, Leibniz Universität Hannover, February 2010.

[18] ———, *Lecture Notes in TMA4145 - Linear Methods*. Url: `http://www.math.ntnu.no/emner/TMA4145/2013h/lecturenotes27aug.pdf` [accessed Feb 18, 2015], 2013.

[19] ———, *On Whitham's conjecture of a highest cusped wave for a nonlocal shallow water wave equation*, Report 19/2015, Mathematisches Forschungsinstitut Oberwolfach, 2015.

[20] M. EHRNSTRÖM, J. ESCHER, AND L. PEI, *A Note on the Local Well-Posedness for the Whitham Equation*, in Elliptic and Parabolic Equations, J. Escher, E. Schrohe, J. Seiler, and C. Walker, eds., vol. 119 of Springer Proceedings in Mathematics & Statistics, Springer International Publishing, 2015, pp. 63–75.

[21] M. EHRNSTRÖM, M. D. GROVES, AND E. WAHLÉN, *On the existence and stability of solitary-wave solutions to a class of evolution equations of Whitham type*, Nonlinearity, 25 (2012), pp. 2903–2936.

[22] M. EHRNSTRÖM AND H. KALISCH, *Traveling waves for the Whitham equation*, Differential and Integral Equations, 22 (2009), pp. 1193–1210.

[23] M. EHRNSTRÖM AND H. KALISCH, *Global Bifurcation for the Whitham Equation*, Mathematical Modelling of Natural Phenomena, 8 (2013), pp. 13–30.

[24] M. EHRNSTRÖM AND E. WAHLÉN, *On the highest wave for the Whitham equation*. Url: https://wiki.math.ntnu.no/_media/dispersive15/wahlen.pdf [accessed Nov 21, 2015], 2015. Work in preparation.

[25] I. EKELAND AND R. TÉMAM, *Convex Analysis and Variational Problems*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1999.

[26] M. EMERTON, *What is the relationship between Fourier transformation and Fourier series?* Url: http://math.stackexchange.com/q/536519 [version 1, accessed Mar 19, 2015], October 2013.

[27] A. EVGRAFOV, *Lower semi-continuity, compactness and existence of solutions*. Url: http://www.math.ntnu.no/emner/TMA4180/2014v/Notes/note01.pdf [accessed Oct 12, 2015], 2014.

[28] M. FABIAN, P. HABALA, P. HÁJEK, V. MONTESINOS SANTALUCÍA, J. PELANT, AND V. ZIZLER, *Functional Analysis and Infinite-Dimensional Geometry*, CMS Books in Mathematics, Springer-Verlag, New York, 2001.

[29] S. A. GABOV, *On Whitham's equation*, Sov. Math. Dokl., 19 (1978), pp. 1225–1229.

[30] C. GASQUET AND P. WITOMSKI, *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets*, Texts in Applied Mathematics, Springer, November 1998.

[31] I. GELFAND AND S. FOMIN, *Calculus of Variations*, Dover Books on Mathematics, Dover Publications, 2000.

[32] A. GRIGIS AND J. SJÖSTRAND, *Microlocal Analysis for Differential Operators: An Introduction*, vol. 196 of London Mathematical Society Lecture Note Series, Cambridge University Press, 1994.

[33] R. A. HUNT, *On the convergence of Fourier series*, in Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255.

[34] J. K. HUNTER, *Notes on Partial Differential Equations*. Url: https://www.math.ucdavis.edu/~hunter/pdes/pde_notes.pdf [accessed Nov 9, 2015], June 2014.

[35] J. K. HUNTER AND B. NACHTERGAELE, *Applied Analysis*, World Scientific Publishing Company, 2 ed., 2005.

[36] V. M. HUR, *Breaking in the Whitham equation for shallow water waves*, ArXiv e-prints, (2015).

[37] V. M. HUR AND M. A. JOHNSON, *Modulational Instability in the Whitham Equation for Water Waves*, Studies in Applied Mathematics, 134 (2015), pp. 120–143.

[38] V. M. HUR AND L. TAO, *Wave breaking for the Whitham equation with fractional dispersion*, Nonlinearity, 27 (2014), pp. 2937–2949.

[39] J. ILMAVIRTA, *A bound on the* L^p *norm in terms of the* L^2 *norm in periodic Sobolev spaces*. Url: `http://mathoverflow.net/q/211217` [version 1, accessed: Jul 13, 2015], July 2015.

[40] R. J. IORIO AND V. DE MAGALHÃES IORIO, *Fourier Analysis and Partial Differential Equations*, vol. 70 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.

[41] M. A. JOHNSON, *On the stability of periodic solutions of the generalized Benjamin-Bona-Mahony equation*, Physica D Nonlinear Phenomena, 239 (2010), pp. 1892–1908.

[42] W. P. JOHNSON, *The Curious History of Faà di Bruno's Formula*, The American Mathematical Monthly, 109 (2002), pp. 217–234.

[43] E. KREYSZIG, *Introductory Functional Analysis with Applications*, Wiley, New York, 1989.

[44] D. LANNES, *The Water Waves Problem: Mathematical Analysis and Asymptotics*, vol. 188 of Mathematical Surveys and Monographs, American Mathematical Society, Ecole Normale Supérieure et CNRS, Paris, France, 2013.

[45] W. J. LIU, *General Analysis*. Url: `http://wj32.org/wp/wp-content/uploads/2015/03/General-Analysis.pdf` [accessed Apr 14, 2015], March 2015.

[46] D. G. LUENBERGER, *Optimization by Vector Space Methods*, Wiley Professional Paperback Series, Wiley-Interscience, 1969.

[47] J. N. MCDONALD AND N. A. WEISS, *A Course in Real Analysis*, Academic Press, 2 ed., 2012.

[48] D. MOLDABAYEV, H. KALISCH, AND D. DUTYKH, *The Whitham Equation as a Model for Surface Water Waves*, ArXiv e-prints, (2014).

[49] P. I. NAUMKIN AND I. A. SHISHMAREV, *Nonlinear Nonlocal Equations in the Theory of Waves*, vol. 133 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, February 1994.

[50] NONLOCAL EQUATIONS WIKI, *Introduction to nonlocal equations*. Url: `http://www.ma.utexas.edu/mediawiki/index.php?title=Introduction_to_nonlocal_equations&oldid=1175` [accessed Apr 27, 2015], May 2014.

[51] M. ROSENZWEIG, *Double integral definition of periodic Sobolev spaces*. Url: `http://math.stackexchange.com/q/1414035` [version 3, accessed: Oct 21, 2015], September 2015.

[52] W. RUDIN, *Fourier Analysis on Groups*, vol. 12 of Interscience Tracts in Pure & Applied Mathematics, Wiley, New York, 1962.

[53] T. RUNST, *Mapping properties of non-linear operators in spaces of Triebel-Lizorkin and Besov type*, Analysis Mathematica, 12 (1986), pp. 313–346.

[54] T. RUNST AND W. SICKEL, *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*, vol. 3 of De Gruyter Series in Nonlinear Analysis and Applications, Walter de Gruyter, September 1996.

[55] J. S. RUSSELL, *Report on Waves*, Report of the fourteenth meeting of the British Association for the Advancement of Science, (1844), pp. 311–390, plates XLVII–LVII.

[56] N. SANFORD, K. KODAMA, J. D. CARTER, AND H. KALISCH, *Stability of traveling wave solutions to the Whitham equation*, Physics Letters A, 378 (2014), pp. 2100–2107.

[57] R. SERVADEI AND G. BISCI, *Lower semicontinuity of functionals of fractional type and applications to nonlocal equations with critical Sobolev exponent*. Preprint. Url: `http://www.ma.utexas.edu/mp_arc/c/13/13-24.pdf` [accessed Feb 18, 2015], March 2013.

[58] S. SHKOLLER, *Lecture Notes on L^p and Sobolev Spaces*, University of Oxford, January 2014.

[59] W. SICKEL, *Superposition of functions in Sobolev spaces of fractional order. A survey*, Banach Center Publications, 27 (1992), pp. 481–497.

[60] K. SOCHA, *Circles in Circles: Creating a Mathematical Model of Surface Water Waves*, The American Mathematical Monthly, 114 (2007), pp. 202–216.

[61] M. STRUWE, *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, A Series of Modern Surveys in Mathematics, Springer, 4 ed., June 2008.

[62] M. E. TAYLOR, *Partial Differential Equations I*, vol. 115 of Applied Mathematical Sciences, Springer-Verlag New York, 2 ed., 2011. Basic Theory.

[63] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010. Theory, methods and applications.

[64] J. L. TROUTMAN, *Variational Calculus and Optimal Control: Optimization with Elementary Convexity*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 2 ed., 1996.

[65] G. B. WHITHAM, *Variational Methods and Applications to Water Waves*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 299 (1967), pp. 6–25.

[66] ———, *Linear and Nonlinear Waves*, Pure and Applied Mathematics, Wiley-Interscience, August 1974.

[67] WIKIPEDIA, THE FREE ENCYCLOPEDIA, *Fourier series*. Url: `https://en.wikipedia.org/w/index.php?title=Fourier_series&oldid=690956051` [accessed Nov 27, 2015], November 2015.

[68] ———, *Multiplier (fourier analysis)*. Url: `https://en.wikipedia.org/w/index.php?title=Multiplier_%28Fourier_analysis%29&oldid=684813794` [accessed Nov 27, 2015], October 2015.

[69] B. YAN, *Introduction to Variational Methods in Partial Differential Equations and Applications*. Url: `http://www.math.msu.edu/~yan/yan-lecture.pdf` [accessed Feb 18, 2015], 2013.

[70] A. ZAITSEV, *Stationary Whitham waves and their dispersion relation*, Dokl. Akad. Nauk SSSR, 286 (1986), pp. 1364–1369.

[71] V. E. ZAKHAROV AND L. A. OSTROVSKY, *Modulation instability: The beginning*, Physica D: Nonlinear Phenomena, 238 (2009), pp. 540–548.