TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 370, Number 6, June 2018, Pages 4323–4349
http://dx.doi.org/10.1090/tran/7155
Article electronically published on February 21, 2018

EXTRAPOLATION AND INTERPOLATION IN GENERALIZED ORLICZ SPACES

DAVID CRUZ-URIBE AND PETER HÄSTÖ

Abstract. We prove versions of the Rubio de Francia extrapolation theorem in generalized Orlicz spaces. As a consequence, we obtain boundedness results for several classical operators as well as a Sobolev inequality in this setting. We also study complex interpolation in the same setting and use it to derive a compact embedding theorem. Our results include as special cases classical Lebesgue and Sobolev space estimates and their variable exponent and double phase growth analogs.

1. Introduction

Generalized Orlicz spaces, also known as Musielak–Orlicz spaces and Nakano spaces, are a class of Banach spaces that include a number of spaces of interest in harmonic analysis and PDEs as special cases. They were introduced by Nakano [45, 46] and others, following Orlicz [48]. We refer to the monograph of Musielak [44] for a comprehensive synthesis of this earlier work. Intuitively, given a function $\varphi(\cdot) : \Omega \times [0,\infty) \to [0,\infty]$, the generalized Orlicz space $L^{\varphi(\cdot)}$ consists of all measurable functions f such that

$$\int_{\mathbb{R}^n} \varphi(x, |f(x)|)\, dx < \infty.$$

If $\varphi(x,t) = t^p$, $1 \leqslant p < \infty$, then we get the classical Lebesgue spaces L^p. If $\varphi(x,t) = t^p w(x)$, we get the weighted Lebesgue spaces $L^p(w)$. If $\varphi(x,t) = \Phi(t)$, where Φ is a Young function, then we get the classical Orlicz spaces [38, 50]. When $\varphi(x,t) = t^{p(x)}$ we get the variable Lebesgue spaces which have been extensively studied in the past 25 years [18, 24, 33], and when $\varphi(x,t) = t^p + a(x)t^q$, $p < q$, we get the double phase functional recently studied in [5, 6, 13].

The generalized Orlicz spaces are of interest not only as the natural generalization of these important examples but also in their own right. They have appeared in many problems in PDEs and the calculus of variations [4–6, 13, 27] and have applications to image processing [10, 32, 37] and fluid dynamics [53].

The operators of classical harmonic analysis (e.g., convolution operators, maximal operators, fractional and singular integrals) and generalized Sobolev spaces in the variable exponent setting have been studied for many years. Subsequently,

Received by the editors July 25, 2016, and, in revised form, December 5, 2016.
2010 *Mathematics Subject Classification.* Primary 46E35; Secondary 46E30, 42B20, 42B25.
Key words and phrases. Non-standard growth, variable exponent, Musielak–Orlicz spaces, extrapolation, weighted norm inequalities.
The first author was supported by NSF grant DMS-1362425 and research funds from the Dean of the College of Arts & Sciences, the University of Alabama.
The authors would also like to thank the anonymous referee for several useful comments and additional references.

results were proved in specific generalized Orlicz spaces: for instance, $\varphi(x,t) = t^{p(x)}\log(e+t)^{q(x)}$ [17,36,40]. Recently, however, attention has shifted to studying these problems in more generality; see [41,42,47]. The second author and his collaborators [30,31,34,35] have systematically studied these questions and established a very broad theory that unites and extends previous work.

The goal of this paper is to develop harmonic analysis on generalized Orlicz spaces by extending the theory of Rubio de Francia extrapolation to this setting. Extrapolation is a powerful tool in the study of weighted norm inequalities: roughly, it shows that if an operator T is bounded on the weighted spaces $L^2(w)$, where the weights belong to the Muckenhoupt class A_2, then for all p, $1 < p < \infty$, the operator T is bounded on $L^p(w)$ when $w \in A_p$. Extrapolation was originally discovered by Rubio de Francia [51]; for more information on the history of extrapolation and an extensive bibliography, see [20].

The power of extrapolation is that it can be used to prove norm inequalities on Banach spaces, provided that the maximal operator is bounded on the dual space (or, more precisely, the associate space). This approach was first explored in [19], where extrapolation was used to prove norm inequalities on variable Lebesgue spaces. (See also [18,20,22].)

The main results of this paper are the following three extrapolation theorems, which show that weighted norm inequalities for operators imply norm inequalities on generalized Orlicz spaces. For the definition of the notation used, please see Sections 2 and 4 below. Our first result is the natural generalization of Rubio de Francia extrapolation to generalized Orlicz spaces. In the variable Lesbesgue spaces, i.e., when $\varphi(x,t) = t^{p(x)}$, Theorem 1.1 was proved in [19].

Theorem 1.1 (Extrapolation). *Given an operator T, suppose that for some p, $1 \leqslant p < \infty$, and all $w \in A_1$,*

$$\|Tf\|_{L^p(w)} \leqslant C(T,n,p,[w]_{A_1})\|f\|_{L^p(w)}.$$

Then, given any weak Φ-function φ such that simple functions belong to $L^{\varphi_p^*(\cdot)}$ and the Hardy–Littlewood maximal operator is bounded on $L^{\varphi_p^*(\cdot)}$, where $\varphi_p(x,t) = \varphi(x,t^{1/p})$, we have that

$$\|Tf\|_{L^{\varphi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

Remark 1.2. In Theorem 1.1 we implicitly assume that T is defined on $L^{\varphi(\cdot)}$ and that Tf is a measurable function. However, we do not assume that T is linear or sublinear. We make the analogous assumptions in the next two results.

Our second result generalizes the off-diagonal extrapolation theorem of Harboure, Macias, and Segovia [28]. In the variable Lebesgue spaces, this result was also proved in [19].

Theorem 1.3 (Off-diagonal extrapolation). *Given an operator T, suppose that for some p, q, $1 \leqslant p \leqslant q < \infty$, and all $w \in A_1$,*

$$\|Tf\|_{L^q(w)} \leqslant C(T,n,p,q,[w]_{A_1})\|f\|_{L^p(w^{p/q})}.$$

Let $\beta := \frac{1}{p} - \frac{1}{q}$. Then, given weak Φ-functions φ and ψ with $\psi^{-1}(x,t) = t^{-\beta}\varphi^{-1}(x,t)$ such that simple functions belong to $L^{\psi_q^*(\cdot)}$ and the Hardy–Littlewood maximal operator is bounded on $L^{\psi_q^*(\cdot)}$, where $\psi_q(x,t) := \psi(x,t^{1/q})$, we have that

$$\|Tf\|_{L^{\psi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

Our final result generalizes the limited range extrapolation theorem of Auscher and Martell [3]. In the particular case of the variable Lebesgue spaces, this was proved very recently in [22].

Theorem 1.4 (Limited range extrapolation). *Given an operator T, suppose that for some p, $1 < q_- < p < q_+ < \infty$, and all $w \in A_{p/q_-} \cap RH_{(q_+/p)'}$,*

$$\|Tf\|_{L^p(w)} \leqslant C(T, n, p, [w]_{A_{p/q_-}}, [w]_{RH_{(q_+/p)'}})\|f\|_{L^p(w)}.$$

Let $\alpha = (q_+/p)'$. Then, given any weak Φ-function φ such that simple functions belong to $L^{\varphi_p^*(\cdot)}$ and such that the Hardy–Littlewood maximal operator is bounded on $L^{\psi(\cdot)}$, where $\psi(x,t) = \varphi_p^*(x, t^{1/\alpha})$, we have that

$$\|Tf\|_{L^{\varphi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

One drawback of these results is that they are stated abstractly for generalized Orlicz spaces that contain the simple functions and where the Hardy–Littlewood maximal operator is bounded; moreover, these assumptions pertain to generalized Orlicz spaces that are related to, but not the same as, the spaces on which we want to prove our norm inequalities. By using the recent work in [31, 34] we can give sufficient conditions on the Φ-functions φ for the hypotheses of each of these theorems to hold. These conditions are somewhat technical, so we defer their precise statement: see Corollaries 4.8 and 4.21 below. Here, however, we want to emphasize that they are easy to check and sufficiently general as to encompass almost all of the examples of generalized Orlicz spaces discussed above. The one exception is that our corollaries do not recapture results for weighted Lebesgue spaces: see the discussion after Theorem 3.3.

As immediate consequences of our extrapolation results, we derive norm inequalities for a number of operators on generalized Orlicz spaces: in particular, for

- the maximal operator,
- Calderón–Zygmund singular integrals and commutators,
- the Riesz potential and fractional maximal operators,
- the spherical maximal operator.

These results are not intended to be exhaustive but to show the versatility of extrapolation in proving inequalities with very little additional work. While extrapolation allows us to obtain easily a vast array of tools, it is worth noting that there is a small price: since the technique uses the maximal operator in the dual space, it means that we must assume the Φ-function is doubling. Thus results such as regularity under exponential growth (e.g., [39]) are out of our reach. Consequently, it is of interest also to find direct proofs for the above-mentioned operators.

We also use extrapolation to prove a Sobolev embedding theorem. We are not aware of any method to obtain compact embeddings directly by extrapolation. Therefore, for this purpose we prove a complex interpolation theorem (Theorem 5.1), which combined with the Sobolev embedding theorem allows us to extend the Rellich–Kondratchov Theorem to generalized Orlicz spaces.

Remark 1.5. As we were completing this paper we learned that an extrapolation theorem in the scale of generalized Orlicz spaces had been proved independently by Maeda et al. [43]. Their result is analogous to Corollary 4.10 but with more complicated hypotheses.

The remainder of this paper is organized as follows. In Section 2 we give the necessary definitions and preliminary results about generalized Orlicz spaces. In Section 3 we discuss sufficient conditions for the Hardy–Littlewood maximal operator to be bounded on generalized Orlicz spaces. We will show that these conditions hold for a wide variety of examples. In Section 4 we first give some necessary definitions about weights, and we introduce the abstract formalism of families of extrapolation pairs that we use to state and prove our extrapolation results. We then prove Theorems 1.1, 1.3, and 1.4. In fact, we will actually prove generalizations of these results and then deduce a number of immediate corollaries that follow from the theory of extrapolation and the sufficient conditions in Section 3. In Section 5 we prove our complex interpolation theorem. Finally, in Section 6 we give our applications of extrapolation.

Throughout this paper, C, c, etc., will denote constants whose value may change at each appearance. If we write $C(X, Y, \ldots)$, we mean that the constant depends on the parameters X, Y, etc. The notation $f \lesssim g$ means that there exists a constant $C > 0$ such that $f \leqslant Cg$. The notation $f \approx g$ means that $f \lesssim g \lesssim f$. By $L^0 = L^0(\mathbb{R}^n)$ we denote the set of (Lebesgue) measurable functions on $\mathbb{R}^n$.

2. Φ-functions and generalized Orlicz spaces

We recall some definitions pertaining to generalized Orlicz spaces. For proofs and further properties, see [24, Chapter 2] and [44]. Our approach follows the development in [31].

Hereafter, we say that a function f is almost increasing if there exists $L \geqslant 1$ such that for all $s \leqslant t$, $f(s) \leqslant Lf(t)$. Almost decreasing is defined analogously. If we can take $L = 1$, we say that f is increasing/decreasing.

Definition 2.1. Let $\varphi\colon [0,\infty) \to [0,\infty]$ be an increasing function such that $\varphi(0) = \lim_{t\to 0^+} \varphi(t) = 0$ and $\lim_{t\to\infty} \varphi(t) = \infty$. Such a function φ is called a *Φ-prefunction.* Furthermore, we say that φ is:

(1) a *weak Φ-function*, denoted $\varphi \in \Phi_w$, if, additionally, $t \mapsto \frac{\varphi(t)}{t}$ is almost increasing on $(0,\infty)$;
(2) a *Φ-function*, denoted $\varphi \in \Phi$, if, additionally, it is left-continuous and convex;
(3) a *strong Φ-function*, denoted $\varphi \in \Phi_s$, if, additionally, it is continuous in $\overline{\mathbb{R}}$ and convex.

Two Φ-(pre)functions φ and ψ are *equivalent*, $\varphi \simeq \psi$, if there exists $L \geqslant 1$ such that $\psi(\frac{t}{L}) \leqslant \varphi(t) \leqslant \psi(Lt)$ for all t. Equivalent Φ-functions give rise to the same space with comparable norms. The converse, however, is false: there exist Φ-functions that induce comparable norms but are not equivalent; cf. [24, Theorem 2.8.1]. We say that φ is doubling if $\varphi(2t) \leqslant A\varphi(t)$ for every $t > 0$. For doubling Φ-functions, $\simeq$ and $\approx$ are equivalent.

While it is common in the literature to work with Φ-functions, it is also convenient to work at times with either weak or strong Φ-functions. We can do so because every weak Φ-function is equivalent to a strong one: the following result was proved in [31, Proposition 2.3].

Lemma 2.2. *Every weak Φ-function is equivalent to a strong Φ-function.*

Given $\varphi \in \Phi_w$, we let φ^{-1} denote the left-inverse of φ:

$$\varphi^{-1}(\tau) := \inf\{t \geqslant 0 \,:\, \varphi(t) \geqslant \tau\}.$$

Then $\varphi(\varphi^{-1}(\tau)) \leqslant \tau$ and $\varphi^{-1}(\varphi(t)) \leqslant t$. Equality holds in the former when φ is continuous, in the latter when φ is strictly increasing. Note that $\varphi \simeq \psi$ if and only if $\varphi^{-1} \approx \psi^{-1}$.

The conjugate Φ-function φ^* is defined by the formula

$$\varphi^*(t) := \sup_{s \geqslant 0} st - \varphi(s).$$

In [24, (2.6.12)], it was shown that $\varphi^{-1}(t)(\varphi^*)^{-1}(t) \approx t$ when φ satisfies additional growth conditions (more precisely, when it is an N-function; see [24, Definition 2.4.4]). Here we show that this holds without these additional assumptions.

Lemma 2.3. *If $\varphi \in \Phi_w$, then $\varphi^{-1}(t)(\varphi^*)^{-1}(t) \approx t$.*

Proof. The claim is invariant under equivalence of Φ-functions, so by Lemma 2.2 we may assume that $\varphi \in \Phi_s$.

Since φ is convex, $t \mapsto \frac{\varphi(t)}{t}$ is increasing. If we combine this with the definition of the conjugate function, we get, for $s > 0$, that

$$\varphi^*\Big(\frac{\varphi(s)}{s}\Big) = \sup_{t \geqslant 0} \Big(\frac{\varphi(s)}{s} - \frac{\varphi(t)}{t}\Big)t = \sup_{t \in [0,s]} \Big(\frac{\varphi(s)}{s} - \frac{\varphi(t)}{t}\Big)t \leqslant \sup_{t \in [0,s]} \frac{\varphi(s)}{s}t = \varphi(s).$$

On the other hand, choosing $t = s$ in the supremum, we find that

$$\varphi^*\Big(2\frac{\varphi(s)}{s}\Big) = \sup_{t \geqslant 0} 2\frac{\varphi(s)}{s}t - \varphi(t) \geqslant 2\varphi(s) - \varphi(s) = \varphi(s).$$

Thus we have shown that

$$\varphi^*\Big(\frac{\varphi(s)}{s}\Big) \leqslant \varphi(s) \leqslant \varphi^*\Big(2\frac{\varphi(s)}{s}\Big). \tag{2.4}$$

The claim of the lemma is immediate for $t = 0$, so we may assume that $0 < t < \infty$. In this case we can find $s > 0$ such that $t = \varphi(s)$ and $\varphi^{-1}(t) = s$ since $\varphi \in \Phi_s$; then (2.4) gives

$$\varphi^*\left(\frac{t}{\varphi^{-1}(t)}\right) \leqslant t \leqslant \varphi^*\left(\frac{2t}{\varphi^{-1}(t)}\right).$$

Since $(\varphi^*)^{-1}$ is increasing and $(\varphi^*)^{-1}(\varphi^*(t)) \leqslant t$, we obtain from this that

$$(\varphi^*)^{-1}\left(\varphi^*\left(\frac{t}{\varphi^{-1}(t)}\right)\right) \leqslant (\varphi^*)^{-1}(t) \leqslant \frac{2t}{\varphi^{-1}(t)}.$$

For the left-hand side we distinguish two cases. If $\varphi^*(s) > 0$, then $(\varphi^*)^{-1}(\varphi^*(s)) = s$. If $\varphi^*(s) = 0$, then $s \leqslant (\varphi^*)^{-1}(t)$ for any $t > 0$ (cf. [30, (2.2)]). In either case, we obtain that

$$\frac{t}{\varphi^{-1}(t)} \leqslant (\varphi^*)^{-1}(t) \leqslant \frac{2t}{\varphi^{-1}(t)},$$

from which the claim follows when we multiply by $\varphi^{-1}(t) > 0$. $\square$

Generalized Orlicz spaces. To define generalized Orlicz spaces, we extend the definition of Φ-functions to depend on the location in space.

Definition 2.5. The set $\Phi(\Omega)$ of *generalized Φ-functions* consists of those functions $\varphi\colon \Omega \times [0,\infty) \to [0,\infty]$ such that

(1) $\varphi(y,\cdot) \in \Phi$ for every $y \in \Omega$;
(2) $\varphi(\cdot,t) \in L^0(\Omega)$ for every $t \geqslant 0$.

The families $\Phi_s(\Omega)$ and $\Phi_w(\Omega)$ are defined analogously.

For simplicity, weak, strong, and generalized Φ-functions will all be called Φ-functions. In sub- and superscripts the dependence on x will be emphasized by $\varphi(\cdot)$: L^φ (Orlicz) versus $L^{\varphi(\cdot)}$ (generalized Orlicz). The properties and definitions of Φ-functions carry over to generalized Φ-functions pointwise.

We can now define generalized Orlicz spaces. By Lemma 2.2 we can take weak Φ-functions in our definitions, though in the references above these definitions are made for Φ-functions.

Definition 2.6. Let $\varphi \in \Phi_w(\Omega)$ and define the semimodular $\varrho_{\varphi(\cdot)}$ for $f \in L^0(\Omega)$ by

$$\varrho_{\varphi(\cdot)}(f) := \int_\Omega \varphi(x,|f(x)|)\,dx.$$

The *generalized Orlicz space*, also called a Musielak–Orlicz space, is defined as the set

$$L^{\varphi(\cdot)}(\Omega) = \{f \in L^0(\Omega)\colon\ \lim_{\lambda\to 0} \varrho_{\varphi(\cdot)}(\lambda f) = 0\}$$

equipped with the (Luxemburg) norm

$$\|f\|_{L^{\varphi(\cdot)}(\Omega)} := \inf\Big\{\lambda > 0\colon \varrho_{\varphi(\cdot)}\Big(\frac{f}{\lambda}\Big) \leqslant 1\Big\}.$$

If the set is clear from the context we abbreviate $\|f\|_{L^{\varphi(\cdot)}(\Omega)}$ by $\|f\|_{L^{\varphi(\cdot)}}$ or $\|f\|_{\varphi(\cdot)}$.

Hölder's inequality holds in generalized Orlicz spaces, with constant 2, without restrictions on the Φ-function [24, Lemma 2.6.3]:

$$\int_\Omega |f|\,|g|\,dx \leqslant 2\|f\|_{\varphi(\cdot)}\|g\|_{\varphi^*(\cdot)}.$$

Moreover, the following general norm conjugate formula, in a sense the opposite of Hölder's inequality, is also true. It was proved in [24, Corollary 2.7.5] for $\varphi \in \Phi(\mathbb{R}^n)$ with $c(\varphi) = 1$. However, by Lemma 2.2 it extends to weak Φ-functions.

Lemma 2.7 (Norm conjugate formula)**.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$, let $f, g \in L^0(\Omega)$, and suppose that simple functions belong to $L^{\varphi^*(\cdot)}$. Then*

$$c(\varphi)\|f\|_{\varphi(\cdot)} \leqslant \sup_{\|g\|_{\varphi^*(\cdot)}\leqslant 1} \int_\Omega |f|\,|g|\,dx \leqslant 2\|f\|_{\varphi(\cdot)}.$$

Remark 2.8. The norm conjugate formula is not directly related to the dual space $(L^{\varphi(\cdot)})^*$, only to the conjugate modular φ^*. Consequently, the formula is also useful in situations where the space is not reflexive.

Remark 2.9. After completing this paper, we found in [29] that it is possible to prove the norm conjugate formula without the assumption that simple functions belong to $L^{\varphi^*(\cdot)}$. This would allow us to remove some of the auxiliary assumptions in our results. Details are left to the reader.

Rescaling. Suppose that $\varphi \in \Phi_w(\mathbb{R}^n)$ and $\varphi_p(x,t) := \varphi(x,t^{1/p})$ is also a weak Φ-function. Then it follows directly from the definition of the Luxemburg norm that

$$\|v^p\|_{\varphi_p(\cdot)} = \|v\|^p_{\varphi(\cdot)}.$$

This identity will be referred to as *rescaling*. The next lemma describes how rescaling behaves under conjugation.

Lemma 2.10. *Let $1 \leqslant p \leqslant q < \infty$, $\beta := \frac{1}{p} - \frac{1}{q}$, and $\varphi, \psi \in \Phi_w(\mathbb{R}^n)$ be such that $\varphi^{-1}(x,t) = t^\beta \psi^{-1}(x,t)$. Define $\varphi_p(x,t) := \varphi(x,t^{1/p})$ and $\psi_q(x,t) := \psi(x,t^{1/q})$, and assume they are in $\Phi_w(\mathbb{R}^n)$. Then*

$$\|v^{\frac{p}{q}}\|_{\varphi_p^*(\cdot)} \approx \|v\|^{\frac{p}{q}}_{\psi_q^*(\cdot)}.$$

Proof. Our proof is based on a pointwise estimate; for simplicity of notation, we thus drop the "x" for the rest of the proof. By Lemma 2.3, $\varphi_p^{-1}(t)(\varphi_p^*)^{-1}(t) \approx t$. By definition, $\varphi_p^{-1}(t) = \varphi^{-1}(t)^p$; hence, we have that $(\varphi_p^*)^{-1}(t) \approx t\big(\varphi^{-1}(t)\big)^{-p}$. Analogously, $\psi^{-1}(t) \approx (t/(\psi_q^*)^{-1}(t))^{1/q}$.

By assumption, $\varphi^{-1}(t) = t^\beta \psi^{-1}(t)$. Therefore,

$$(\varphi_p^*)^{-1}(t) \approx t\big(\varphi^{-1}(t)\big)^{-p} = t^{1-p\beta}\psi^{-1}(t)^{-p} = t^{1-p\beta-\frac{p}{q}}(\psi_q^*)^{-1}(t)^{\frac{p}{q}}.$$

Since $1-p\beta-\frac{p}{q} = 0$, we have shown that $\varphi_p^*(t) \approx \psi_q^*(t^{q/p})$, and so the claim follows by rescaling. □

3. The maximal operator in generalized Orlicz spaces

We begin by recalling the definition of the maximal operator. The *Hardy–Littlewood maximal operator* is defined for $f \in L^0(\mathbb{R}^n)$ by

$$Mf(x) := \sup_{r>0} \fint_{B(x,r)} |f(y)|\,dy,$$

where $B(x,r)$ is the ball with center x and radius r, and $\fint$ denotes the average integral. Equivalently, the averages can be taken over all balls (or all cubes) that contain x. For the general theory of the maximal operator, see [26].

We now give a family of hypotheses that are closely related to the boundedness of the maximal operator on generalized Orlicz spaces.

Definition 3.1. Given $\varphi \in \Phi_w(\mathbb{R}^n)$ and $0 < p < \infty$, we define the following conditions:

- (A0) $\varphi^{-1}(x,1) \approx 1$ uniformly in $x \in \Omega$.
- (A1) There exists $\beta \in (0,1)$ such that $\beta\varphi^{-1}(x,t) \leqslant \varphi^{-1}(y,t)$ for every $t \in \big[1, \frac{1}{|x-y|^n}\big]$ and every $x,y \in \Omega$ with $|x-y| \leqslant 1$.
- (A2) $L^{\varphi(\cdot)}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) = L^{\varphi_\infty}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, with $\varphi_\infty(t) := \limsup_{|x|\to\infty} \varphi(x,t)$ and $\varphi_\infty \in \Phi_w$.
- $(\mathrm{Inc})_p$ $s \mapsto s^{-p}\varphi(x,s)$ is increasing for all $x \in \Omega$.
- $(\mathrm{Dec})_p$ $s \mapsto s^{-p}\varphi(x,s)$ is decreasing for all $x \in \Omega$.
- $(\mathrm{aInc})_p$ $s \mapsto s^{-p}\varphi(x,s)$ is almost increasing uniformly in $x \in \Omega$.
- $(\mathrm{aDec})_p$ $s \mapsto s^{-p}\varphi(x,s)$ is almost decreasing uniformly in $x \in \Omega$.

We say that φ satisfies (aInc) if it satisfies $(\mathrm{aInc})_p$ for some $p > 1$ and (aDec) if it satisfies $(\mathrm{aDec})_p$ for some $p < \infty$.

There are three facts we want to observe about these definitions. First, (aDec) is equivalent to doubling; see [31]. Second, by a change of variables, $\varphi \in \Phi_w$ satisfies $(\mathrm{aInc})_p$ for $p > 1$ if and only if $\varphi_p(x,t) = \varphi(x, t^{1/p})$ satisfies $(\mathrm{aInc})_1$, i.e., if and only if $\varphi_p \in \Phi_w$. The third is that (A0) yields the inclusion of simple functions in a generalized Orlicz space. We state this as a lemma; we will use it below to apply Lemma 2.7.

Lemma 3.2. *Let $\varphi \in \Phi_w(\Omega)$ satisfy* (A0). *Then simple functions belong to $L^{\varphi(\cdot)}(\Omega)$.*

Proof. Since $L^{\varphi(\cdot)}(\Omega)$ is a vector space, it suffices to show that $\chi_A \in L^{\varphi(\cdot)}(\Omega)$ for every $A \subset \Omega$ of finite measure. By (A0) and the properties of φ^{-1}, there exists $\beta > 0$ such that $\varphi(x,\beta) \leqslant 1$ for every $x \in \Omega$. Then

$$\varrho_{\varphi(\cdot)}(\lambda\beta\chi_A) \lesssim \lambda \int_A \varphi(x,\beta)\,dx \leqslant \lambda\,|A| \to 0$$

as $\lambda \to 0$, so $\chi_A \in L^{\varphi(\cdot)}(\Omega)$ by the definition of the space. □

The importance of these conditions is that we can use them to give sufficient conditions for the maximal operator to be bounded on generalized Orlicz spaces. The following result was proved in [34, Theorem 4.7].

Theorem 3.3. *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy conditions* (A0)–(A2) *and* (aInc). *Then*

$$M : L^{\varphi(\cdot)}(\mathbb{R}^n) \to L^{\varphi(\cdot)}(\mathbb{R}^n)$$

is bounded.

Examples. To better understand these conditions, we consider them in several special cases. In the classical Lebesgue spaces, $\varphi(x,t) = t^p$, (A0)–(A2) hold trivially, and (aInc) is equivalent to $p > 1$. This corresponds to the well-known fact that the maximal operator is not bounded on L^1. The fact that we do not need to assume (aDec) corresponds to the fact that the maximal operator is bounded on L^∞. Similarly, if $\varphi(x,t) = \Phi(t)$ for some Young function Φ, then (aInc) is equivalent to the lower Boyd index of Φ being greater than 1, which is again necessary for the maximal operator to be bounded on L^Φ. (See [7].)

In the variable Lebesgue spaces, the (aInc) condition is equivalent to $p^- = \operatorname{ess\,inf} p(x) > 1$, which is necessary for the maximal operator to be bounded [24, Theorem 4.7.1]. The (A1) condition is the local log-Hölder continuity for $\frac{1}{p}$, whereas (A2) is the Nekvinda decay condition N_∞. (For definitions, see [17] or [24].) These conditions are close to optimal, as it is known that one cannot replace log-Hölder continuity by any weaker modulus of continuity and still guarantee that the maximal operator is bounded. On the other hand, there exist examples of exponents that do not satisfy these conditions—indeed, which are not even continuous—but for which the maximal operator is still bounded on $L^{p(\cdot)}$ [24, Example 5.1.8].

In the double phase case, the critical issue is the behavior of a around the zero set $\{x : a(x) = 0\}$. Colombo and Mingione [13] found that the critical Hölder exponent with which a must approach zero is $\frac{n}{p}(q-p)$, which also gives a sufficient condition for (A1) to hold [34]. A similar observation holds in the limiting double phase case [5, 6]. Double phase regularity has been studied only in bounded domains; however, one can use Lemma 3.4 to show that (A2) imposes no additional constraint compared to (A0).

$\varphi(x,t)$	(A0)	(A1)	(A2)	(aInc)	(aDec)
$t^{p(x)}a(x)$	$a \approx 1$	$\frac{1}{p} \in C^{\log}_{\rm loc}$	$p \in N_\infty$	$p^- > 1$	$p^+ < \infty$
$t^{p(x)}\log(e+t)$	–	$\frac{1}{p} \in C^{\log}_{\rm loc}$	$p \in N_\infty$	$p^- > 1$	$p^+ < \infty$
$t^p + a(x)t^q$	$a \in L^\infty$	$a \in C^{\frac{n}{p}(q-p)}$	$a \in L^\infty$	$p > 1$	$q < \infty$
$t^p + a(x)t^p\log(e+t)$	$a \in L^\infty$	$a \in C^{\log}_{\rm loc}$	$a \in L^\infty$	$p > 1$	$p < \infty$

Finally, we note that the (A0) condition precludes weighted norm inequalities, both in the classical case, $\varphi(x,t) = t^p a(x)$, and in the variable exponent case, $\varphi(x,t) = t^{p(x)}a(x)$. In either case the (A0) condition requires the weight to be essentially constant. Therefore, Theorem 3.3 does not capture these very important cases. A "weighted" theory in the setting of generalized Orlicz spaces remains an open problem.

The maximal operator and conjugate Φ-functions. In our extrapolation theorems, the hypotheses are not given on the space $L^{\varphi(\cdot)}$ for the original Φ-function φ, but rather on $L^{\varphi^*(\cdot)}$ or on some rescaling of this space. In the scale of the variable Lebesgue spaces these conditions are essentially equivalent; see Diening [23] or [24, Theorem 5.7.2] for a precise statement. It is not known if this is true for generalized Orlicz spaces. However, using our assumptions we can give sufficient conditions on φ for the maximal operator to be bounded on $L^{\varphi^*(\cdot)}$ or on a rescaling of this space. To do so, we need to consider the effect of conjugation and rescaling on the above conditions. This is done in Propositions 3.5 and 3.6 below. To prove the first, we give some additional characterizations of the (A2) condition.

Lemma 3.4. *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ and let $\varphi_\infty(t) := \limsup_{|x|\to\infty} \varphi(x,t)$, $\varphi_\infty \in \Phi_w$. Then the following are equivalent:*

(1) *φ satisfies* (A2).

(2) *There exist $h \in L^1$ and $\beta > 0$ such that for every $t \in [0,1]$,*

$$\varphi(x,\beta t) \leqslant \varphi_\infty(t) + h(x) \quad \textit{and} \quad \varphi_\infty(\beta t) \leqslant \varphi(x,t) + h(x).$$

(3) *For any $s > 0$ there exist $h \in L^1$ and $\beta > 0$ such that for every $t \in [0,s]$,*

$$\varphi(x,\beta t) \leqslant \varphi_\infty(t) + h(x) \quad \textit{and} \quad \varphi_\infty(\beta t) \leqslant \varphi(x,t) + h(x).$$

Proof. It is clear that (3.4) implies (3.4). For the converse, we assume that (3.4) holds and $s > 1$. Let $\beta' := \frac{\beta}{s}$ and $t \in [0,s]$. Since $\frac{t}{s} \in [0,1]$, by (3.4) we see that

$$\varphi(x,\beta' t) = \varphi(x,\beta \tfrac{t}{s}) \leqslant \varphi_\infty(\tfrac{t}{s}) + h(x) \leqslant \varphi_\infty(t) + h(x).$$

The same argument works for the other inequality as well; hence, (3.4) holds. Therefore (3.4) and (3.4) are equivalent.

We now consider (3.4) and (3.4). Let

$$\xi(x,t) := \max\{\varphi(x,t), \infty\chi_{(1,\infty)}(t)\} \quad \text{and} \quad \psi(t) := \max\{\varphi_\infty(t), \infty\chi_{(1,\infty)}(t)\}.$$

Then $L^{\varphi(\cdot)} \cap L^\infty = L^{\xi(\cdot)}$ and $L^{\varphi_\infty} \cap L^\infty = L^\psi$ so that (3.4) becomes $L^{\xi(\cdot)} = L^\psi$. By [24, Theorem 2.8.1], $L^{\xi(\cdot)} = L^\psi$ if and only if there exist $\beta > 0$ and $h \in L^1$ such that

$$\xi(x,\beta t) \leqslant \psi(t) + h(x) \quad \text{and} \quad \psi(\beta t) \leqslant \xi(x,t) + h(x).$$

We may assume without loss of generality that $\beta \leqslant 1$. When $t \in [0,1]$, $\xi(x,t) = \varphi(x,t)$, $\xi(x,\beta t) = \varphi(x,\beta t)$, $\psi(t) = \varphi_\infty(t)$, and $\psi(\beta t) = \varphi_\infty(\beta t)$, whereas for $t > 1$ the inequalities are trivial since the right-hand side is infinite. Thus, (3.4) and (3.4) are equivalent. $\square$

To describe the impact of rescaling and conjugation we define two operators on Φ-functions: given $\varphi \in \Phi_w$ and $\alpha > 0$, let

$$T_*(\varphi)(x,t) := \varphi^*(x,t) \quad \text{and} \quad T_\alpha(\varphi)(x,t) := \varphi\big(x, t^{\frac{1}{\alpha}}\big).$$

We study the behavior of our conditions under these transformations. We first consider (A0)–(A2).

Proposition 3.5. *Conditions* (A0), (A1), *and* (A2) *are invariant under* T_* *and* T_α.

Proof. By Lemma 2.3, $\varphi^{-1}(x,t)(\varphi^*)^{-1}(x,t) \approx t$. Furthermore, $(T_\alpha\varphi)^{-1}(x,t) = \varphi^{-1}(x,t)^\alpha$. When $t = 1$, we see from these that (A0) is invariant. Likewise, we see that (A1) is invariant.

For (A2), we use Lemma 3.4. We find that

$$T_\alpha\varphi(x, \beta^\alpha t) = \varphi\big(x, \beta t^{\frac{1}{\alpha}}\big) \leqslant \varphi_\infty\big(t^{\frac{1}{\alpha}}\big) + h(x) = T_\alpha\varphi_\infty(t) + h(x)$$

and similarly for the other inequality in Lemma 3.4(3.4). Hence (A2) is invariant under T_α. By [24, Lemma 2.6.4], the inequality

$$\varphi(x, \beta t) \leqslant \varphi_\infty(t) + h(x)$$

is equivalent to

$$T_*\varphi(x, \tfrac{t}{\beta}) \geqslant T_*\varphi_\infty(t) - h(x),$$

and similarly for the other inequality. Hence we see by Lemma 3.4(3.4) that (A2) is invariant under T_* as well. □

We next consider (aDec) and (aInc). The following result is a generalization of [31, Lemma 2.4].

Proposition 3.6. *Let* $\varphi \in \Phi_w$. *Then* φ *satisfies* $(\mathrm{aInc})_p$ *if and only if* $T_*(\varphi) = \varphi^*$ *satisfies* $(\mathrm{aDec})_{p'}$, *where* $\frac{1}{p} + \frac{1}{p'} = 1$. *Further,* T_α *maps* $(\mathrm{aInc})_p$ *to* $(\mathrm{aInc})_{p/\alpha}$ *and* $(\mathrm{aDec})_p$ *to* $(\mathrm{aDec})_{p/\alpha}$.

Proof. We first consider the special case of (Inc) and (Dec). We have that φ satisfies $(\mathrm{Inc})_p$ if and only if $\frac{\varphi(t^{1/p})}{t}$ is increasing, similarly for φ^* and $(\mathrm{Dec})_{p'}$. From the definition of the conjugate function,

$$\frac{\varphi^*(s^{\frac{1}{p'}})}{s} = \frac{1}{s}\sup_{t\geqslant 0} ts^{\frac{1}{p'}} - \varphi(t) = \sup_{v\geqslant 0}\Big(v^{-\frac{1}{p'}} - \frac{\varphi\big((sv)^{\frac{1}{p}}\big)}{sv}\Big)v,$$

where we used the change of variables $t := (sv)^{\frac{1}{p}}$. From this expression, we see that φ^* satisfies $(\mathrm{Dec})_{p'}$ and $(\mathrm{Inc})_{p'}$ if φ satisfies $(\mathrm{Inc})_p$ and $(\mathrm{Dec})_p$, respectively. Since $(\varphi^*)^* = \varphi$ [24, Corollary 2.6.3], we conclude that φ satisfies $(\mathrm{Inc})_p$ if and only if φ^* satisfies $(\mathrm{Dec})_{p'}$.

Suppose now that φ satisfies $(\mathrm{aInc})_p$. Then $\psi(s) := s^p \inf_{t\geqslant s} t^{-p}\varphi(t)$ satisfies $(\mathrm{Inc})_p$ and $\varphi \approx \psi$. By the above argument, ψ^* satisfies $(\mathrm{Dec})_{p'}$ and, by [24, Lemma 2.6.4], $\varphi^* \simeq \psi^*$; hence, φ^* satisfies $(\mathrm{aDec})_{p'}$. Conversely, suppose φ^* satisfies $(\mathrm{aDec})_{p'}$. Then we can argue in the same way with the auxiliary function $\psi(s) := s^{p'} \sup_{t\geqslant s} t^{-p'}\varphi^*(t)$ (since $\varphi = (\varphi^*)^*$).

It remains to consider T_p. As before, φ satisfies $(\mathrm{aInc})_p$ if and only if $\frac{\varphi(t^{1/p})}{t}$ is almost increasing. Since $T_\alpha\varphi(x, t^{\frac{\alpha}{p}}) = \varphi(x, t^{\frac{1}{p}})$, $T_\alpha\varphi$ then satisfies $(\mathrm{aInc})_{p/\alpha}$. The case of (aDec) is proved analogously. □

4. Extrapolation in generalized Orlicz spaces

Weights and classical extrapolation. We give some preliminary definitions and results about weights and the classical theory of Rubio de Francia extrapolation, as well as more recent generalizations. For more information and proofs, we refer the reader to [20, 26] and the references they contain.

By a weight we mean a non-negative, locally integrable function w such that $0 < w(x) < \infty$ almost everywhere. For $1 \leqslant p < \infty$, the weighted Lebesgue space $L^p(w)$ consists of all $f \in L^0$ such that

$$\|f\|_{L^p(w)} = \left(\int_{\mathbb{R}^n} |f|^p w\,dx\right)^{1/p} < \infty.$$

For $1 < p < \infty$, a weight w is in the Muckenhoupt class A_p, denoted $w \in A_p$, if

$$[w]_{A_p} = \sup_Q \left(\fint_Q w\,dx\right)\left(\fint_Q w^{1-p'}\,dx\right)^{p-1} < \infty,$$

where the supremum is taken over all cubes with sides parallel to the coordinate axes. (Equivalently, we can replace cubes by balls.) When $p = 1$, we say $w \in A_1$ if

$$[w]_{A_1} = \sup_Q \left(\fint_Q w\,dx\right) \operatorname*{ess\,sup}_{x\in Q} \frac{1}{w(x)} < \infty.$$

Given $1 < p < q < \infty$, $A_1 \subsetneq A_p \subsetneq A_q$. A simple example of a Muckenhoupt A_p weight is $w(x) = |x|^a$, $-n < a < n(p-1)$.

A weight w satisfies the reverse Hölder condition with exponent $s > 1$, denoted $w \in RH_s$, if

$$[w]_{RH_s} = \sup_Q \left(\fint_Q w^s\,dx\right)^{1/s}\left(\fint_Q w\,dx\right)^{-1} < \infty.$$

By Hölder's inequality, if $s > t$, then $RH_s \subsetneq RH_t$. There is a close connection between the A_p and RH_s classes:

$$\bigcup_{p\geqslant 1} A_p = \bigcup_{s>1} RH_s =: A_\infty.$$

However, it is important to note that while $w \in A_p$ for some p if and only if $w \in RH_s$ for some s, there is no connection between the size of s and p. This can be seen by considering the reverse Hölder exponents of the power weights $|x|^a$.

While the theory of Rubio de Francia extrapolation is usually applied to prove norm inequalities for operators, the properties of the operator itself play no direct role in the statement or proof of extrapolation. Moreover, we will want to use extrapolation to prove more general inequalities, such as the Coifman–Fefferman type inequalities relating pairs of operators, e.g., inequalities of the form

$$\|Tf\|_{\varphi(\cdot)} \lesssim \|Mf\|_{\varphi(\cdot)},$$

where T is a singular integral operator and M the Hardy–Littlewood maximal operator. We also want to be able to use extrapolation to prove weak-type and vector-valued inequalities. For more on this approach in the classical setting, see [20].

Therefore, rather than consider inequalities relating the norms of Tf and f, we will consider families of pairs of non-negative measurable functions,

$$\mathcal{F} = \{(f,g)\},$$

with (implicitly) additional restrictions on f and g. The pairs (f, g) are called extrapolation pairs. In our extrapolation theorems we will assume that we have weighted norm inequalities

$$\|f\|_{L^p(w)} \leqslant C(n, p, [w]_{A_q})\|g\|_{L^p(w)} \tag{4.1}$$

and use them to deduce generalized Orlicz space inequalities

$$\|f\|_{\varphi(\cdot)} \leqslant C(n, \varphi)\|g\|_{\varphi(\cdot)}. \tag{4.2}$$

More precisely, in order to prove (4.2) for a fixed pair (f, g), we need to have that (4.1) holds for a weight $w \in A_q$ that we construct in the course of the proof. The problem is that in this abstract setting do not know that this will be the case: e.g., we cannot rule out *a priori* that $\|f\|_{L^p(w)} = \infty$ but $\|g\|_{L^p(w)}$ is finite.

To avoid this problem we adopt the following convention. Given a family $\mathcal{F}$ of extrapolation pairs, if we write

$$\|f\|_{L^p(w)} \leqslant C(n, p, [w]_{A_q})\|g\|_{L^p(w)}, \qquad (f, g) \in \mathcal{F},$$

then we mean this inequality holds for a given weight $w \in A_q$ for all pairs (f, g) such that the left-hand side is finite. If we write

$$\|f\|_{\varphi(\cdot)} \leqslant C(n, \varphi)\|g\|_{\varphi(\cdot)}, \qquad (f, g) \in \mathcal{F},$$

we mean the same thing: this inequality holds for all pairs such that $f \in L^{\varphi(\cdot)}$. In the proof we will use this latter assumption to prove that $\|f\|_{L^p(w)}$ is finite for a specific weight w. Note that we do not assume that $\|g\|_{L^p(w)}$ is finite, though if $g \in L^{\varphi(\cdot)}$, then in the course of the proof we will show that it is. (If $\|g\|_{L^p(w)} = \infty$, then there is nothing to prove for this particular pair.)

To apply extrapolation to prove norm inequalities for an operator T (as in our main results in the Introduction), we would consider a family of extrapolation pairs of the form $\mathcal{F} = \{(|Tf|, |f|)\}$. If T is defined on $L^{\varphi(\cdot)}$, then we can take $f \in L^{\varphi(\cdot)}$, but then we need to check that the above conventions hold for all such pairs. This is the approach we take to prove Theorems 1.1–1.4; see below. Alternatively, if T is defined on a dense subset, then we can use approximation arguments to prove norm inequalities. We will discuss this approach in further detail in Section 6 below.

To put our extrapolation results in context and because we will need them below in our proofs, we state two versions of extrapolation. The first is the classical result of Rubio de Francia. For a proof, see [20, Theorem 3.9, Corollary 3.12].

Theorem 4.3. *Given a family of extrapolation pairs $\mathcal{F}$, suppose that for some $p_0 \in [1, \infty)$ and every $w_0 \in A_{p_0}$,*

$$\|f\|_{L^{p_0}(w_0)} \leqslant C(n, p_0, [w_0]_{A_{p_0}})\|g\|_{L^{p_0}(w_0)}, \qquad (f, g) \in \mathcal{F}.$$

Then for every $p \in (1, \infty)$ and every $w \in A_p$,

$$\|f\|_{L^p(w)} \leqslant C(n, p, [w]_{A_p})\|g\|_{L^p(w)}, \qquad (f, g) \in \mathcal{F}.$$

Moreover, for every $p, q \in (1, \infty)$ and $w \in A_p$,

$$\Big\|\Big(\sum_k f_k^q\Big)^{1/q}\Big\|_{L^p(w)} \leqslant C(n, p, q, [w]_{A_p})\Big\|\Big(\sum_k g_k^q\Big)^{1/q}\Big\|_{L^p(w)}, \qquad \{(f_k, g_k)\}_k \subset \mathcal{F}.$$

The second result we need is the limited range extrapolation theorem of Auscher and Martell [3]. (See also [20, Theorem 3.23].)

Theorem 4.4. *Given a family of extrapolation pairs $\mathcal{F}$ and $1 < q_- < q_+ < \infty$, suppose that for some $p_0 \in (q_-, q_+)$ and every $w_0 \in A_{p_0/q_-} \cap RH_{(q_+/p_0)'}$,*

$$\|f\|_{L^{p_0}(w_0)} \leqslant C(n, p_0, [w_0]_{A_{p_0/q_-}}, [w_0]_{RH_{(q_+/p_0)'}})\|g\|_{L^{p_0}(w_0)}, \qquad (f,g) \in \mathcal{F}.$$

Then for every $p \in (q_-, q_+)$ and every $w \in A_{p/q_-} \cap RH_{(q_+/p)'}$,

$$\|f\|_{L^p(w)} \leqslant C(n, p, [w]_{A_{p/q_-}}, [w]_{RH_{(q_+/p)'}})\|g\|_{L^p(w)}, \qquad (f,g) \in \mathcal{F}.$$

Diagonal and off-diagonal extrapolation. We now prove Theorems 1.1 and 1.3. First note that Theorem 1.1 is actually a special case of Theorem 1.3 when $p = q$, so it will suffice to prove the latter. In turn, Theorem 1.3 is a consequence of the following more general result expressed in terms of extrapolation pairs.

Theorem 4.5. *Given a family of extrapolation pairs $\mathcal{F}$, suppose that for some p, q, $1 \leqslant p \leqslant q < \infty$, and all $w \in A_1$,*

$$\|f\|_{L^q(w)} \leqslant C(n, p, q, [w]_{A_1})\|g\|_{L^p(w^{p/q})}, \qquad (f,g) \in \mathcal{F}. \tag{4.6}$$

Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy $(\mathrm{aInc})_p$ (i.e., $\varphi_p \in \Phi_w$). Define $\beta := \frac{1}{p} - \frac{1}{q}$, $\psi^{-1}(x,t) := t^{-\beta}\varphi^{-1}(x,t)$, and $\psi_q(x,t) := \psi(x, t^{1/q})$. If simple functions belong to $L^{\psi_q^*(\cdot)}$ and the Hardy–Littlewood maximal operator is bounded on $L^{\psi_q^*(\cdot)}$, then

$$\|f\|_{L^{\psi(\cdot)}} \leqslant C\|g\|_{L^{\varphi(\cdot)}}, \qquad (f,g) \in \mathcal{F}. \tag{4.7}$$

Before proving Theorem 4.5 we first state and prove three corollaries. In the first two, we use the results of Section 3 to immediately get sufficient conditions on φ, ψ for the conclusions of Theorem 4.5 to hold.

Corollary 4.8. *Let p, q, β, and $\mathcal{F}$ be as in Theorem 4.5 and suppose (4.6) holds. Let $\psi \in \Phi_w(\mathbb{R}^n)$ and define $\varphi^{-1}(x,t) := t^{\beta}\psi^{-1}(x,t)$. If ψ satisfies assumptions (A0)–(A2) and there exist $q_+ > q$ such that ψ satisfies $(\mathrm{aInc})_q$ and $(\mathrm{aDec})_{q_+}$, then (4.7) holds.*

Proof. Since ψ satisfies $(\mathrm{aInc})_q$, $\frac{\varphi_q(x,t)}{t}$ is almost increasing and so $\psi_q \in \Phi_w$. Since ψ satisfies (A0)–(A2) and $(\mathrm{aDec})_{q_+}$, and since $\psi_q^* = T_* T_q(\psi)$, ψ_q^* satisfies (A0)–(A2) and $(\mathrm{aInc})_{(q_+/q)'}$ by Propositions 3.6 and 3.5. Hence, by Theorem 3.3, the maximal operator is bounded on $L^{\psi_q^*(\cdot)}$. Finally, since ψ_q^* satisfies (A0), by Lemma 3.2, simple functions are contained in $L^{\psi_q^*(\cdot)}$. Therefore, we can apply Theorem 4.5 to get the desired result. □

Alternatively, we can state the assumptions in terms of φ.

Corollary 4.9. *Let p, q, β, and $\mathcal{F}$ be as in Theorem 4.5 and suppose (4.6) holds. Let $\varphi \in \Phi_w(\mathbb{R}^n)$ and define $\psi^{-1}(x,t) := t^{-\beta}\varphi^{-1}(x,t)$. If φ satisfies (A0)–(A2), $(\mathrm{aInc})_p$, and $(\mathrm{aDec})_{p_+}$ for some $p_+ > p$, then (4.7) holds.*

Proof. It follows from the definition of ψ that it also satisfies (A0)–(A2). Furthermore, φ satisfies $(\mathrm{aInc})_p$ and $(\mathrm{aDec})_{p_+}$ if and only if φ^{-1} satisfies $(\mathrm{aDec})_{1/p}$ and $(\mathrm{aInc})_{1/p_+}$, respectively. Thus, ψ satisfies $(\mathrm{aInc})_q$ and $(\mathrm{aDec})_{q_+}$. Hence, the result follows from Corollary 4.9. □

By using the full strength of Rubio de Francia extrapolation, we can also prove the following result, which holds for a large class of Φ-functions.

Corollary 4.10. *Given a family of extrapolation pairs $\mathcal{F}$, suppose that for some $p \in [1,\infty)$ and all $w \in A_p$,*

$$\|f\|_{L^p(w)} \leqslant C(n,p,[w]_{A_p})\|g\|_{L^p(w)}, \qquad (f,g) \in \mathcal{F}. \tag{4.11}$$

Suppose φ is a weak Φ-function that satisfies assumptions (A0)–(A2) *and* (aDec). *If $p>1$, then we also assume* (aInc). *Then*

$$\|f\|_{L^{\varphi(\cdot)}} \leqslant C\|g\|_{L^{\varphi(\cdot)}}, \qquad (f,g) \in \mathcal{F}. \tag{4.12}$$

Moreover, we have that for any q, $1<q<\infty$,

$$\Big\|\Big(\sum_k f_k^q\Big)^{1/q}\Big\|_{L^{\varphi(\cdot)}} \leqslant C(n,q,\varphi)\Big\|\Big(\sum_k g_k^q\Big)^{1/q}\Big\|_{L^{\varphi(\cdot)}}, \qquad \{(f_k,g_k)\}_k \subset \mathcal{F}. \tag{4.13}$$

Proof. If $p>1$, then φ satisfies $(\mathrm{aInc})_{p_-}$ for some $p_->1$ by assumption. By Theorem 4.3, we have that (4.11) holds also for some $p \leqslant p_-$. (If $p=1$, then this is automatically true without using Theorem 4.3.) Since $A_1 \subset A_p$ we satisfy the hypotheses of Corollary 4.8, and so we get (4.12). To prove (4.13) we repeat this argument, starting from the weighted vector-valued inequality in Theorem 4.3. □

Remark 4.14. An off-diagonal version of Corollary 4.10 holds, using the off-diagonal extrapolation theorem [20, Theorem 3.23]. Details are left to the interested reader.

Proof of Theorem 4.5. We begin the proof by using the Rubio de Francia iteration algorithm. Let $m := \|M\|_{L^{\psi_q^*(\cdot)} \to L^{\psi_q^*(\cdot)}}$ and define $\mathcal{R} : L^0(\mathbb{R}^n) \to [0,\infty]$ by

$$\mathcal{R}h(x) := \sum_{k=0}^{\infty} \frac{M^k h(x)}{2^k m^k},$$

where for $k \geqslant 1$, M^k denotes k iterations of the maximal operator, and $M^0 h = |h|$. Then the following properties hold:

(A) $|h| \leqslant \mathcal{R}h$,
(B) $\|\mathcal{R}h\|_{L^{\psi_q^*(\cdot)}} \leqslant 2\|h\|_{L^{\psi_q^*(\cdot)}}$,
(C) $\mathcal{R}h \in A_1$ and $[\mathcal{R}h]_{A_1} \leqslant 2m$.

Property (A) holds since $\mathcal{R}h \geqslant M^0 h = |h|$, (B) holds since

$$\Big\|\frac{M^k h}{2^k m^k}\Big\|_{L^{\psi_q^*(\cdot)}} = \frac{\|M^k h\|_{L^{\psi_q^*(\cdot)}}}{2^k m^k} \leqslant \frac{m\|M^{k-1}h\|_{L^{\psi_q^*(\cdot)}}}{2^k m^k} \leqslant \cdots \leqslant \frac{\|h\|_{L^{\psi_q^*(\cdot)}}}{2^k},$$

and (C) holds since $M(\mathcal{R}h) \leqslant 2m\mathcal{R}h$ by the sublinearity of the maximal operator.

Fix $(f,g) \in \mathcal{F}$ and define $\mathcal{H} := \{h : \|h\|_{\varphi_q^*(\cdot)} \leqslant 1\}$. By rescaling, Lemma 2.7, and (A),

$$\|f\|_{\psi(\cdot)}^q = \big\|f^q\big\|_{\psi_q(\cdot)} \lesssim \sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} f^q h\,dx \leqslant \sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} f^q\,\mathcal{R}h\,dx.$$

To apply our hypothesis, by our convention on families of extrapolation pairs we need to show that the right-hand term is finite. But this follows at once by Hölder's inequality and (B): for all $h \in \mathcal{H}$,

$$\int_{\mathbb{R}^n} f^q\,\mathcal{R}h\,dx \leqslant 2\|f^q\|_{\psi_q(\cdot)}\|\mathcal{R}h\|_{\psi_q^*(\cdot)} \leqslant 4\|f\|_{\psi(\cdot)}^q\|h\|_{\psi_q^*(\cdot)} < \infty; \tag{4.15}$$

the last inequality holds since, again by our convention, $f \in L^{\psi(\cdot)}$. Given this and (C) we can apply our hypothesis (4.6) to get that

$$\|f\|_{\psi(\cdot)} \leqslant \sup_{h\in\mathcal{H}} \left(\int_{\mathbb{R}^n} f^q \mathcal{R}h\,dx\right)^{\frac{1}{q}} \leqslant C \sup_{h\in\mathcal{H}} \left(\int_{\mathbb{R}^n} g^p (\mathcal{R}h)^{\frac{p}{q}}\,dx\right)^{\frac{1}{p}}.$$

Let $\varphi_p(x,t) := \varphi(x,t^{1/p})$; then for any $h \in \mathcal{H}$, by Hölder's inequality, rescaling, Lemma 2.10, and property (B),

$$\int g^p(\mathcal{R}h)^{\frac{p}{q}}, dx \leqslant 2\big\| g^p\big\|_{\varphi_p(\cdot)}\big\|(\mathcal{R}h)^{\frac{p}{q}}\big\|_{\varphi_p^*(\cdot)} \lesssim \|g\|^p_{\varphi(\cdot)}\|Rh\|^{\frac{p}{q}}_{\psi_q^*(\cdot)} \lesssim \|g\|^p_{\varphi(\cdot)}. \tag{4.16}$$

If we combine the last two inequalities we get that $\|f\|_{\psi(\cdot)} \lesssim \|g\|_{\varphi(\cdot)}$, as desired. ☐

Proof of Theorem 1.3. We will derive this result as a consequence of the proof of Theorem 4.5. Define the family of extrapolation pairs

$$\mathcal{F} = \{(|Tg|, |g|) : g \in L^{\varphi(\cdot)}\}.$$

(Recall that we assume that T is defined on $L^{\varphi(\cdot)}$ and Tg is measurable.) By inequality (4.16) we have that $g \in L^p((\mathcal{R}h)^{\frac{p}{q}})$ for every $h \in \mathcal{H}$; therefore, by assumption $Tg \in L^q(\mathcal{R}h)$. This gives us inequality (4.15) (with $f = |Tg|$) without the *a priori* assumption that $Tg \in L^{\psi(\cdot)}$. Therefore, the proof goes through and we get the desired conclusion. ☐

Remark 4.17. It is also possible to derive Theorem 4.5 from a similar result for Banach function spaces that was proved in [20, Theorem 4.6]. To do so, we must first prove that our hypotheses imply that $L^{\varphi(\cdot)}$ is a Banach function space. We sketch the proof of this fact. By our assumption that the maximal operator is bounded on $L^{\varphi_p^*(\cdot)}$, we have that $L^{\varphi_p^*(\cdot)} \hookrightarrow L^1_{\mathrm{loc}}$. This inclusion is equivalent to χ_E being contained in the associate space $(L^{\varphi_p^*(\cdot)})'$ for every measurable set E, $|E| < \infty$ [24, Remark 2.7.10]. By our assumption that simple functions are contained in $L^{\varphi_p^*(\cdot)}$, $(L^{\varphi_p^*(\cdot)})' = L^{\varphi_p(\cdot)}$ ([24, Remark 2.7.4], since $(\varphi^*)^* = \varphi$). Thus, we have that $\chi_E \in L^{\varphi_p(\cdot)}$, and so by rescaling, $\chi_E \in L^{\varphi(\cdot)}$. On the other hand, by Lemma 2.10 (replacing p, q with 1, p), the fact that $\chi_E \in L^{\varphi_p^*(\cdot)}$ implies that $\chi_E \in L^{\varphi^*(\cdot)}$. Thus simple functions are contained in $L^{\varphi(\cdot)} \cap L^{\varphi^*(\cdot)}$. Therefore, by [24, Corollary 2.7.9], $L^{\varphi(\cdot)}$ is a Banach function space.

The remainder of the hypotheses of [20, Theorem 4.6] can be checked similarly. However, if we take this approach, the proof of Theorem 1.3 is more complicated. Therefore, it seemed more straightforward to give a direct proof of Theorem 4.5.

Limited range extrapolation. We now turn to Theorem 1.4. As before, this theorem will be a consequence of the following result stated in terms of extrapolation pairs. The details of the proof of Theorem 1.4 are essentially the same as the proof of Theorem 1.3 above and so are omitted.

Theorem 4.18. *Given a family of extrapolation pairs $\mathcal{F}$ and $1 < q_- < q_+ < \infty$, suppose that for some $p \in (q_-, q_+)$ and every $w \in A_{p/q_-} \cap RH_{(q_+/p)'}$,*

$$\|f\|_{L^p(w)} \leqslant C(n,p,[w]_{A_{p/q_-}},[w]_{RH_{(q_+/p)'}})\|g\|_{L^p(w)}, \qquad (f,g) \in \mathcal{F}. \tag{4.19}$$

Let $\varphi \in \Phi_w(\mathbb{R}^n)$ *satisfy* $(\mathrm{aInc})_p$ *and* $(\mathrm{aDec})_{q_+}$. *If simple functions belong to* $L^{\varphi_p^*(\cdot)}$ *and the Hardy–Littlewood maximal operator is bounded on* $L^{\psi(\cdot)}$, *where* $\psi(x,t) := \varphi_p^*(x, t^{1/\alpha})$, $\alpha := (q_+/p)'$, *and* $\varphi_p(x,t) = \varphi(x, t^{1/p})$, *then*

$$\|f\|_{L^{\varphi(\cdot)}} \leqslant C\|g\|_{L^{\varphi(\cdot)}}, \qquad (f,g) \in \mathcal{F}. \tag{4.20}$$

Again by using the results in Section 3 we can give sufficient conditions on the Φ-function φ for Theorem 4.18 to hold.

Corollary 4.21. *Let* p, q_-, q_+, *and* $\mathcal{F}$ *be as in Theorem* 4.18 *and suppose* (4.19) *holds. Let* $\varphi \in \Phi_w(\mathbb{R}^n)$ *satisfy assumptions* (A0)–(A2), $(\mathrm{aInc})_{p_-}$, *and* $(\mathrm{aDec})_{p_+}$ *for some* $q_- < p_- \leqslant p_+ < q_+$. *Then* (4.20) *holds.*

Remark 4.22. To compare the hypotheses of Corollary 4.21 to those of Corollary 4.10, note that the latter can be restated as φ satisfies $(\mathrm{aInc})_{p_-}$ and $(\mathrm{aDec})_{p_+}$ for some $1 \leqslant p_- \leqslant p_+ < \infty$.

Proof. By Theorem 4.4 we may assume that (4.19) holds for some $p \in (q_-, p_-)$. Since $p < p_-$, it follows that φ satisfies $(\mathrm{aInc})_p$. Since φ satisfies (A0)–(A2) and $(\mathrm{aDec})_{p_+}$, by Propositions 3.5 and 3.6, $\psi = T_\alpha T_* T_p(\varphi)$ satisfies (A0)–(A2) and $(\mathrm{aInc})_{(p_+/p)'/\alpha}$. Note that this makes sense because

$$\frac{(p_+/p)'}{\alpha} = \frac{(p_+/p)'}{(q_+/p)'} > 1.$$

Hence, by Theorem 3.3, the maximal operator is bounded on $L^{\psi(\cdot)}$. Finally, $\varphi_p^* = T_* T_p(\varphi)$ satisfies (A0), and so by Lemma 3.2, simple functions are contained in $L^{\varphi_p^*(\cdot)}$. Therefore, we can use Theorem 4.18 to get the desired conclusion. ☐

Proof of Theorem 4.18. First note that $\varphi_p \in \Phi_w(\mathbb{R}^n)$ since φ satisfies $(\mathrm{aInc})_p$. Moreover, the calculations in the proof of Corollary 4.21 show that $\psi \in \Phi_w(\mathbb{R}^n)$ since φ satisfies $(\mathrm{aDec})_{q_+}$.

As in the proof of Theorem 4.5, we now define an iteration algorithm. Let $m := \|M\|_{L^{\psi(\cdot)} \to L^{\psi(\cdot)}}$ and with the same notation as before let

$$\mathcal{R}h(x) := \sum_{k=0}^{\infty} \frac{M^k h(x)}{2^k m^k}.$$

Assume that $h \geqslant 0$ and define $H := \mathcal{R}(h^\alpha)^{1/\alpha}$. Then we have that

(A) $h \leqslant H$,
(B) $\|H\|_{\varphi_p^*(\cdot)} \leqslant 2\|h\|_{\varphi_p^*(\cdot)}$,
(C) $H \in A_1 \cap RH_{(q_+/p)'} \subset A_{p/q_-} \cap RH_{(q_+/p)'}$.

These properties are proved in much the same way as the analogous properties for $\mathcal{R}h$ in the proof of Theorem 4.5. As before, we have $h^\alpha \leqslant \mathcal{R}(h^\alpha)$, so Property (A) holds. Similarly, we have $\|\mathcal{R}h\|_{\psi(\cdot)} \leqslant 2\|h\|_{\psi(\cdot)}$, and so by rescaling

$$\|H\|_{\varphi_p^*(\cdot)} = \|\mathcal{R}(h^\alpha)^{1/\alpha}\|_{\varphi_p^*(\cdot)} = \|\mathcal{R}(h^\alpha)\|_{\psi(\cdot)}^{1/\alpha} \leqslant 2^{1/\alpha}\|h^\alpha\|_{\psi(\cdot)}^{1/\alpha} \leqslant 2\|h\|_{\varphi_p^*(\cdot)}.$$

Finally, we have $\mathcal{R}(h^\alpha) \in A_1$ and $[\mathcal{R}(h^\alpha)]_{A_1} \leqslant 2m$, and so $H^\alpha \in A_1$. Thus, by the rescaling properties of A_1 weights (see [21, Theorems 2.2, 2.3]) we have that $H \in A_1 \cap RH_{(q_+/p)'} \subset A_{p/q_-} \cap RH_{(q_+/p)'}$.

We can now estimate as follows. Let $\mathcal{H} := \{h : \|h\|_{\varphi_p^*(\cdot)} \leqslant 1\}$. Then by rescaling, Lemma 2.7, and (A),

$$\|f\|_{\varphi(\cdot)}^p = \|f^p\|_{\varphi_p(\cdot)} \lesssim \sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} f^p h\, dx \leqslant \sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} f^p H\, dx.$$

As before, to apply our hypothesis, we check to see if our convention about extrapolation pairs holds. By Hölder's inequality and (B) we have that for any $h \in \mathcal{H}$,

$$\int_{\mathbb{R}^n} f^p H\, dx \leqslant 2\|f^p\|_{\varphi_p(\cdot)}\|H\|_{\varphi_p^*(\cdot)} \leqslant 4\|f\|_{\varphi(\cdot)}^p\|h\|_{\varphi_p^*(\cdot)} < \infty;$$

the last inequality holds since we assume that $\|f\|_{\varphi(\cdot)} < \infty$. Therefore, by (C) we can apply our hypothesis (4.19) to get

$$\sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} f^p H\, dx \lesssim \sup_{h\in\mathcal{H}} \int_{\mathbb{R}^n} g^p H\, dx.$$

By Hölder's inequality, rescaling, and (B) we have that for all $h \in \mathcal{H}$,

$$\int_{\mathbb{R}^n} g^p H\, dx \leqslant 2\|g^p\|_{\varphi_p(\cdot)}\|H\|_{\varphi_p^*(\cdot)} \leqslant 4\|g\|_{\varphi(\cdot)}\|h\|_{\varphi_p^*(\cdot)} \leqslant 4\|g\|_{\varphi(\cdot)}.$$

If we combine these estimates we get $\|f\|_{\varphi(\cdot)} \lesssim \|g\|_{\varphi}$. □

5. Complex interpolation

In this section we prove a complex interpolation theorem in the scale of generalized Orlicz spaces. Note that real interpolation has, for the most part, not been especially useful even in the variable exponent setting, since the primary and secondary parameter (i.e., p and θ in $(A, B)_{p,\theta}$) do not co-vary (but see [1] for an exception). Therefore, it is natural to first consider complex interpolation in the more general setting of generalized Orlicz spaces.

Previously, Musielak [44, Theorem 14.16] proved complex interpolation results, but his proofs were longer and more complicated; a simpler proof was given in [25]. However, in both cases the results apply only to N-functions which are proper. Here we eliminate the first restriction.[1]

We recall the definition of the norm in the interpolation space $[L^{\varphi_0(\cdot)}, L^{\varphi_1(\cdot)}]_{[\theta]}$. Let $S := \{z \in \mathbb{C} :\ 0 < \operatorname{Re} z < 1\}$, so that $\overline{S} = \{z \in \mathbb{C}\ :\ 0 \leqslant \operatorname{Re} z \leqslant 1\}$, where $\operatorname{Re} z$ is the real part of z. Let $\mathcal{G}$ be the space of functions on $\overline{S}$ with values in $L^{\varphi_0(\cdot)} + L^{\varphi_1(\cdot)}$ which are analytic on S and bounded and continuous on $\overline{S}$ such that $F(it)$ and $F(1 + it)$ tend to zero as $|t| \to \infty$. (Recall that i denotes the imaginary unit. Also, F is analytic with values in a Banach space means that $\frac{d}{d\bar{z}}F = 0$ in the Banach space.) For $F \in \mathcal{G}$ we set

$$\|F\|_{\mathcal{G}} := \sup_{t\in\mathbb{R}} \max\Big\{\big\|F(it)\big\|_{\varphi_0(\cdot)}, \big\|F(1+it)\big\|_{\varphi_1(\cdot)}\Big\}.$$

Then we define the norm of $[L^{\varphi_0(\cdot)}, L^{\varphi_1(\cdot)}]_{[\theta]}$ by

$$\|f\|_{[\theta]} := \inf\big\{\|F\|_{\mathcal{G}}\ :\ F \in \mathcal{G} \text{ and } f = F(\theta)\big\}.$$

For $\varphi_0, \varphi_1 \in \Phi_w(\mathbb{R}^n)$ and $\theta \in (0, 1)$ we define the θ-intermediate function φ_θ by

$$\varphi_\theta^{-1}(x,\cdot) = \big(\varphi_0^{-1}(x,\cdot)\big)^{1-\theta}\big(\varphi_1^{-1}(x,\cdot)\big)^{\theta}.$$

[1] Using the results of [29], it is possible to eliminate also the latter restriction.

Then φ_θ is also a weak Φ-function.

Theorem 5.1 (Complex interpolation)**.** *Let* $\varphi_0, \varphi_1 \in \Phi_w(\Omega)$ *satisfy* (A0)*. Then*

$$\big[L^{\varphi_0(\cdot)}(\Omega), L^{\varphi_1(\cdot)}(\Omega)\big]_{[\theta]} = L^{\varphi_\theta(\cdot)}(\Omega)$$

for all $0 < \theta < 1$*.*

Proof. We proceed along the lines of [8]. By Lemma 2.2 we may assume without loss of generality that $\varphi_0, \varphi_1 \in \Phi_s(\Omega)$. We extend $\varphi_j : \Omega \times [0,\infty) \to [0,\infty]$, $j = 1$, 2, to $\varphi_j : \Omega \times \mathbb{C} \to [0,\infty]$ via $\varphi_j(x,t) = \varphi_j(x,|t|)$. For $z \in \mathbb{C}$ with $0 \leqslant \operatorname{Re} z \leqslant 1$ define φ_z by

$$\varphi_z^{-1}(x,t) = \big(\varphi_0^{-1}(x,t)\big)^{1-z}\big(\varphi_1^{-1}(x,t)\big)^{z}.$$

Then φ_z^{-1} is holomorphic in z on S and continuous on $\overline{S}$.

For $g \in L^{\varphi_\theta(\cdot)}$ with $\|g\|_{\varphi_\theta} \leqslant 1$ define

$$f_\varepsilon(z;x) := \exp(-\varepsilon + \varepsilon z^2 - \varepsilon\theta^2)\, \varphi_z^{-1}\big(x, \varphi_\theta(x, g(x))\big) \operatorname{sgn} g(x).$$

Then $f(\theta) = \exp(-\varepsilon) g$ when $\varphi_\theta(x, g(x)) \in (0,\infty)$ and

$$|f_\varepsilon(1+it, x)| = \exp(-\varepsilon t^2 - \varepsilon\theta^2)\, \big|\varphi_{1+it}^{-1}\big(x, \varphi_\theta(x,g)\big)\big| \leqslant \varphi_1^{-1}\big(x, \varphi_\theta(x,g)\big),$$

$$|f_\varepsilon(it, x)| = \exp(-\varepsilon - \varepsilon t^2 - \varepsilon\theta^2)\, \big|\varphi_{it}^{-1}\big(x, \varphi_\theta(x,g)\big)\big| \leqslant \varphi_0^{-1}\big(x, \varphi_\theta(x,g)\big).$$

Since $\varphi_1(\varphi_1^{-1}(t)) \leqslant t$ and $\int \varphi_\theta(x, g(x))\, dx \leqslant 1$ we conclude that $\varrho_{\varphi_1(\cdot)}(f_\varepsilon(1+it,\cdot)) \leqslant 1$, similarly for φ_0. Thus $\|f_\varepsilon\|_{\mathcal{G}} = \sup_{t\in\mathbb{R}} \max\big\{\|f_\varepsilon(it,\cdot)\|_{\varphi_0(\cdot)}, \|f_\varepsilon(1+it,\cdot)\|_{\varphi_1(\cdot)}\big\} \leqslant 1$. This and $f(\theta) = \exp(-\varepsilon) g$ imply that $\|\exp(-\varepsilon) g\|_{[\theta]} \leqslant 1$. Letting $\varepsilon \to 0$, we find by a scaling argument that $\|g\|_{[\theta]} \leqslant \|g\|_{\varphi_\theta(\cdot)}$.

We now prove the opposite inequality. Since φ_0 and φ_1 satisfy (A0), so does φ_z. By Lemma 2.3, $\varphi_z^* := (\varphi_z)^*$ also satisfies (A0). Be Lemma 2.2, we may assume that $\varphi_z^* \in \Phi_s(\mathbb{R}^n)$. By Lemma 2.7,

$$\|g\|_{\varphi_\theta(\cdot)} \lesssim \sup_{\|b\|_{\varphi_\theta^*(\cdot)} \leqslant 1} \int_{\mathbb{R}^n} |g|\,|b|\, dx. \tag{5.2}$$

For $\|g\|_{[\theta]} \leqslant 1$ and b as above put

$$h_\varepsilon(z;x) := \exp(-\varepsilon + \varepsilon z^2 - \varepsilon\theta^2)\, \psi_z\big(x, \varphi_\theta^*\big(x, b(x)\big)\big) \operatorname{sgn} g(x),$$

where ψ_z is the *right*-inverse of φ_z^*. Since $\varphi_z^* \in \Phi_s(\mathbb{R}^n)$, the right-inverse agrees with the left-inverse, except possibly at the origin:

$$\psi_z(x,t) = \begin{cases} (\varphi_z^*)^{-1}(x,t) & \text{if } t > 0, \\ t_z(x) & \text{if } t = 0. \end{cases}$$

Here $t_z(x) := \sup\{t \geqslant 0 : \varphi_z^*(x,t) = 0\}$. Since $\varphi_z^* \in \Phi_s(\Omega)$, it follows that $\varphi_z^*(x, \psi_z(t)) = t$ when $t > 0$. But since φ_z^* is left-continuous, also $\varphi_z^*(x, \psi_z(0)) = \varphi_z^*(x, t_z(x)) = 0$, so that $\varphi_z^*(x, \psi_z(t)) = t$ for all t.

Writing $F_\varepsilon(z) := \int_{\mathbb{R}^n} |f_\varepsilon|\,|h_\varepsilon|\,dx$ we find by Young's inequality that

$$\begin{aligned}F_\varepsilon(it) &\leqslant \int_\Omega \psi_{it}\big(x, \varphi_\theta^*(x, b(x))\big)\varphi_{it}^{-1}\big(x, \varphi_\theta\big(x, g(x)\big)\big)\,dx \\ &\leqslant \int_\Omega \varphi_{it}^*\big[x, \psi_{it}\big(x, \varphi_\theta^*(x, b(x))\big)\big] + \varphi_{it}\big[\varphi_{it}^{-1}\big(x, \varphi_\theta\big(x, g(x)\big)\big)\big]\,dx \\ &= \int_\Omega \varphi_\theta^*(x, b(x)) + \varphi_\theta(x, g(x))\,dx \\ &\leqslant 2.\end{aligned}$$

Analogously, $F_\varepsilon(1+it) \leqslant 2$, so the three-line theorem implies that $F_\varepsilon(z) \leqslant 2$ for all $z \in S$.

When $z = \theta$, ψ_θ is the right inverse of φ_θ^*. Then by the definition, $\psi_\theta(x, \varphi_\theta^*(x,t)) \geqslant t$. Thus, we obtain that

$$F_\varepsilon(\theta) = \exp(-2\varepsilon)\int_\Omega |g(x)|\,\psi_z\big(x, \varphi_\theta^*\big(x, b(x)\big)\big)\,dx \geqslant \exp(-2\varepsilon)\int_\Omega |g(x)|\,b(x)\,dx.$$

Taking the supremum over b and letting $\varepsilon \to 0$, we get from this and (5.2) that $\|g\|_{\varphi_\theta(\cdot)} \leqslant c$; hence, $\|g\|_{\varphi_\theta(\cdot)} \lesssim \|g\|_{[\theta]}$. □

Remark 5.3. Section 7.1 of [24] contains a proof of the complex interpolation theorem without the N-function assumption (for variable exponent Lebesgue spaces). However, that proof contains an error since it is based on the inequality $\varphi^{-1}(\varphi(t)) \geqslant t$, which is in general false. This problem is overcome above by the use of the right-inverse.

The following result is proved using Theorem 5.1 by means of the Riesz–Thorin Interpolation Theorem and the Hahn–Banach Theorem; cf. [24, Corollary 7.1.4] and [25, Corollary A.5].

Corollary 5.4. *Let $\varphi_0, \varphi_1 \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0) *and let T be a sublinear operator that is bounded from $L^{\varphi_j(\cdot)}(\Omega)$ to $L^{\varphi_j(\cdot)}(\Omega)$ for $j = 0, 1$. Then for $0 < \theta < 1$, T is also bounded from $L^{\varphi_\theta(\cdot)}(\Omega)$ to $L^{\varphi_\theta(\cdot)}(\Omega)$.*

The next result is proved using Calderón's interpolation theorem; cf. [24, Corollary 7.1.6].

Corollary 5.5. *Let $\varphi_0, \varphi_1 \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0)*, let X be a Banach space, and let T be a linear operator that is bounded from X to $L^{\varphi_0(\cdot)}(\Omega)$ and compact from X to $L^{\varphi_1(\cdot)}(\Omega)$. Then for $0 < \theta < 1$, T is also compact from X to $L^{\varphi_\theta(\cdot)}(\Omega)$.*

6. Applications of extrapolation and interpolation

In this section we give some representative applications of extrapolation to prove norm inequalities in generalized Orlicz spaces. The key to such inequalities is the existence of weighted norm inequalities, and there is a vast literature on this subject. For additional examples in the context of variable Lebesgue spaces that can be easily extended to generalized Orlicz spaces, see [18–20, 22].

As we noted in Section 4, to apply extrapolation to prove norm inequalities we either need that the operator is *a priori* defined on $L^{\varphi(\cdot)}$ or we need to use density and approximation arguments. In this case our conditions from Section 3 are very useful. For instance, if $\varphi \in \Phi_w(\Omega)$ satisfies (A0) and (aDec), then $L^\infty_c(\Omega)$ and $C^\infty_c(\Omega)$ are both dense in $L^{\varphi(\cdot)}(\Omega)$. (See [31, Theorems 4.3, 4.5].)

To apply extrapolation via density and approximation, we consider a common special case. Suppose T is a linear operator that is defined on a dense subset $X \subset L^{\varphi(\cdot)}$, and suppose further that $X \subset L^p(w)$ for all $p \geqslant 1$ and $w \in A_1$. (This is the case if $X = L^\infty_c$, C^∞_c.) If T satisfies weighted norm inequalities on $L^p(w)$, then we can take as our extrapolation pairs the family

$$\mathcal{F} = \{(\min(|Tf|, k)\chi_{B(0,k)}, |f|) : f \in X\}.$$

Arguing as in Remark 4.17, we have that simple functions, and so L^∞_c, are contained in $L^{\varphi(\cdot)}$; hence, $\min(|Tf|, k)\chi_{B(0,k)} \in L^{\varphi(\cdot)}$, and so $\mathcal{F}$ satisfies the convention for families of extrapolation pairs. Thus, we can apply Theorem 4.5 (when $p = q$) to prove that for all $f \in X$, $\|Tf\|_{\varphi(\cdot)} \leqslant C\|f\|_{\varphi(\cdot)}$. Since T is linear, given an arbitrary $f \in L^{\varphi(\cdot)}$, if we take any sequence $\{f_j\} \subset X$ converging to f, $\{Tf_j\}$ is Cauchy and we can define Tf as the limit. This extends the norm inequality to all of $L^{\varphi(\cdot)}$.

If T is not linear, then this argument does not work. However, suppose T has the property that $|Tf(x)| \leqslant T(|f|)(x)$ and if f is non-negative and $\{f_j\}$ is any non-negative sequence that increases pointwise to f, then

$$Tf(x) \leqslant \liminf_{k\to\infty} Tf_j(x).$$

(This is the case, for instance, if T is the maximal operator.) Given this, the above argument can essentially be repeated, since given non-negative $f \in L^\varphi$, $f_j = \min(f, k)\chi_{B(0,k)} \in L^{\varphi(\cdot)} \cap L^\infty_c$.

In the following examples we will state our hypotheses in terms of the assumptions used in Corollaries 4.10 and 4.21. The necessary families of extrapolation pairs can be constructed using the above arguments; we leave the details to the interested reader. (In the case of variable exponent spaces, several examples are worked out in detail in [18, Chapter 5].) Obviously, weaker assumptions can be used; again, we leave the precise statements to the interested reader.

The maximal operator. Though we assume the boundedness of the maximal operator in order to apply extrapolation, one important consequence is that we get vector-valued inequalities for the maximal operator. For $1 < p, q < \infty$, $w \in A_p$, and sequence $\{f_k\}_k \subset L^0$,

$$\Big\| \Big(\sum_k (Mf_k)^q \Big)^{1/q} \Big\|_{L^p(w)} \leqslant C \Big\| \Big(\sum_k f_k^q \Big)^{1/q} \Big\|_{L^p(w)}.$$

See [2]. Therefore, we have the following result.

Corollary 6.1. *Suppose that $\varphi \in \Phi_w$ satisfies* (A0)–(A2), (aInc), *and* (aDec). *Then for $1 < q < \infty$,*

$$\Big\| \Big(\sum_k (Mf_k)^q \Big)^{1/q} \Big\|_{L^{\varphi(\cdot)}} \leqslant C \Big\| \Big(\sum_k |f_k|^q \Big)^{1/q} \Big\|_{L^{\varphi(\cdot)}}.$$

Calderón–Zygmund singular integrals. Let Δ be the diagonal in $\mathbb{R}^n \times \mathbb{R}^n$, that is, $\Delta := \{(x, x) : x \in \mathbb{R}^n\}$. A bounded linear operator $T : L^2 \to L^2$ is a Calderón–Zygmund singular integral operator if there exists a kernel $K : \mathbb{R}^n \times \mathbb{R}^n \setminus \Delta \to \mathbb{R}$ such that for all $f \in C^\infty_c$ and $x \notin \operatorname{supp}(f)$,

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y)\, dy,$$

and, moreover, for some $\varepsilon > 0$ the kernel satisfies

$$|K(x,y)| \leqslant \frac{C}{|x-y|^n},$$

$$|K(x,y) - K(x,y+h)| + |K(x,y) - K(x+h,y)| \leqslant C\frac{|h|^\varepsilon}{|x-y|^{n+\varepsilon}}, \quad 2|h| \leqslant |x-y|.$$

Calderón–Zygmund operators satisfy weighted norm inequalities: for $1 < p < \infty$ and $w \in A_p$,

$$\|Tf\|_{L^p(w)} \leqslant C\|f\|_{L^p(w)}.$$

See [26]. Therefore, we get the following result.

Corollary 6.2. *Suppose that* $\varphi \in \Phi_w$ *satisfies* (A0)–(A2), (aInc)*, and* (aDec)*. Then*

$$\|Tf\|_{L^{\varphi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

Moreover, for $1 < q < \infty$*,*

$$\bigg\|\bigg(\sum_k |Tf_k|^q\bigg)^{1/q}\bigg\|_{L^{\varphi(\cdot)}} \leqslant C\bigg\|\bigg(\sum_k |f_k|^q\bigg)^{1/q}\bigg\|_{L^{\varphi(\cdot)}}.$$

We can also extend the Coifman–Fefferman inequality [11] relating singular integrals and the Hardy–Littlewood maximal function. Given $w \in A_\infty$ and $0 < p < \infty$,

$$\|Tf\|_{L^p(w)} \leqslant C\|Mf\|_{L^p(w)}.$$

By extrapolation we can extend this to generalized Orlicz spaces.

Corollary 6.3. *Suppose that* $\varphi \in \Phi_w$ *satisfies* (A0)–(A2) *and* (aDec)*. Then*

$$\|Tf\|_{L^{\varphi(\cdot)}} \leqslant C\|Mf\|_{L^{\varphi(\cdot)}}.$$

Remark 6.4. One feature of the Coifman–Fefferman inequality is that it holds for $0 < p < 1$. The analogous condition in the generalized Orlicz spaces would be for it to hold for a "quasi Φ-function", that is, a function φ such that for some $r > 1$, $\varphi_r(x,t) = \varphi(x,t^r)$ is a Φ-function. Our proof of extrapolation can be generalized to this context; details are left to the interested reader. For such a result in the context of variable Lebesgue spaces, see [19].

Commutators of singular integrals. Given a Calderón–Zygmund singular integral T and a function $b \in BMO$, the space of functions of bounded mean oscillation, we define the commutator $[b,T]$ to be the operator

$$[b,T]f(x) = b(x)Tf(x) - T(bf)(x) = \int_{\mathbb{R}^n} \big(b(x) - b(y)\big)K(x,y)f(y)\,dy.$$

Commutators were introduced by Coifman, Rochberg, and Weiss [12], who proved that they are bounded on L^p, $1 < p < \infty$, if and only if $b \in BMO$. Weighted norm inequalities were proved by Pérez [49]: if $b \in BMO$ and $w \in A_p$, then

$$\|[b,T]f\|_{L^p(w)} \leqslant C\|f\|_{L^p(w)}.$$

Therefore, we get the following result.

Corollary 6.5. *Suppose that* $\varphi \in \Phi_w$ *satisfies* (A0)–(A2), (aInc)*, and* (aDec)*. Then*

$$\|[b,T]f\|_{L^{\varphi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

Remark 6.6. It was recently shown in [15] that in the variable exponent case, given an exponent $p(\cdot)$ such that the maximal operator is bounded on $L^{p(\cdot)}$ and a singular integral T with sufficiently smooth kernel K, if the commutator $[b,T]$ is bounded on $L^{p(\cdot)}$, then $b \in BMO$. The same argument extends to the setting of generalized Orlicz spaces.

The Riesz potential and fractional maximal operators. Given $0 < \alpha < n$, we define the Riesz potential (also referred to as the fractional integral operator) to be the positive integral operator

$$I_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy.$$

The associated fractional maximal operator is defined by

$$M_\alpha f(x) = \sup_{r>0} |B(x,r)|^{\frac{\alpha}{n}} \fint_{B(x,r)} |f(y)|\, dy.$$

These operators satisfy the following weighted norm inequalities: for $w \in A_1$ and p, q such that $1 < p < n/\alpha$ and $\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}$,

$$\|I_\alpha\|_{L^q(w)} \leqslant C\|f\|_{L^p(w^{p/q})}$$

and

$$\|M_\alpha\|_{L^q(w)} \leqslant C\|f\|_{L^p(w^{p/q})}.$$

(These inequalities are usually stated in terms of the A_{pq} condition of Muckenhoupt and Wheeden, but this special case is sufficient for our purposes. See [18] for further details.)

Moreover, for $w \in A_\infty$ and $0 < p < \infty$ we have the Coifman–Fefferman type inequality

$$\|I_\alpha\|_{L^p(w)} \leqslant C\|M_\alpha f\|_{L^p(w)}.$$

Therefore, by extrapolation we get the following results.

Corollary 6.7. *Given $0 < \alpha < n$, suppose φ, $\psi \in \Phi_w$ are such that $\varphi^{-1}(x,t) = t^{\frac{\alpha}{n}}\psi^{-1}(x,t)$, ψ satisfies* (A0)–(A2), *and there exist $\frac{n}{n-\alpha} < p_- < p_+ < \infty$ such that ψ satisfies* $(\mathrm{aInc})_{p_-}$ *and* $(\mathrm{aDec})_{p_+}$. *Then*

$$\|I_\alpha f\|_{\psi(\cdot)} \leqslant C\|f\|_{\varphi(\cdot)},$$

$$\|M_\alpha f\|_{\psi(\cdot)} \leqslant C\|f\|_{\varphi(\cdot)}.$$

Corollary 6.8. *Suppose that $\varphi \in \Phi_w$ satisfies* (A0)–(A2), (aInc), *and* (aDec) *and* $0 < \alpha < n$. *Then*

$$\|I_\alpha f\|_{\varphi(\cdot)} \leqslant C\|M_\alpha f\|_{\varphi(\cdot)}.$$

Remark 6.9. Given $b \in BMO$, it is also possible to define commutators $[b, I_\alpha]$. These operators were introduced by Chanillo [9], and weighted inequalities analogous to those for I_α were proved in [16]. We can therefore prove estimates on generalized Orlicz spaces for $[b, I_\alpha]$. Details are left to the interested reader.

The spherical maximal operator. The spherical maximal operator is defined for $f \in L^0$ by

$$\mathcal{M}f(x) = \sup_{t>0} \left| \int_{S^{n-1}} f(x-ty)\, d\sigma(y) \right|,$$

where S^{n-1} is the unit sphere in $\mathbb{R}^n$ and σ is the surface measure. Stein [52] proved that for $n \geqslant 3$, the spherical maximal operator is bounded on L^p for $p > \frac{n}{n-1}$. Weighted norm inequalities hold for p in the same range, but with strong conditions on the weight. It follows from a result of Cowling et al. [14] that

$$\|\mathcal{M}f\|_{L^p(w)} \leqslant C\|f\|_{L^p(w)},$$

provided $w \in A_{p/q_-} \cap RH_{(q_+/p)'}$, where

$$q_- = \frac{p}{(n-1)(1-\delta)}, \quad q_+ = (n-1)q_-, \quad \max(0, 1-p/n) \leqslant \delta \leqslant \frac{n-2}{n-1}.$$

See [22, Corollary 3.12] for details. By taking δ close to $\frac{n-2}{n-1}$ and arguing as in the previous reference, we get the following result.

Corollary 6.10. *Given $n \geqslant 3$, suppose $\varphi \in \Phi_w$ satisfies* (A0)–(A2) *and for $\frac{n}{n-1} < p_- < p_+ = (n-1)p_-$, φ satisfies* $(\mathrm{aInc})_{p_-}$ *and* $(\mathrm{aDec})_{p_+}$. *Then*

$$\|\mathcal{M}f\|_{L^{\varphi(\cdot)}} \leqslant C\|f\|_{L^{\varphi(\cdot)}}.$$

Remark 6.11. Even though weighted norm inequalities hold for all $p > \frac{n}{n-1}$, the restriction that $p_+ = (n-1)p_-$ is close to optimal. See [22] and the references it contains.

The Sobolev embedding theorem. Given $\varphi \in \Phi_w(\mathbb{R}^n)$ and an open set $\Omega \in \mathbb{R}^n$, we define the generalized Orlicz–Sobolev space $W^{1,\varphi(\cdot)}(\Omega)$ to be the set of all $f \in W^{1,1}_{loc}(\Omega)$ such that $f, |\nabla f| \in L^{\varphi(\cdot)}(\Omega)$. This is a Banach space with norm $\|f\|_{W^{1,\varphi(\cdot)}(\Omega)} = \|f\|_\varphi + \|\nabla f\|_\varphi$. We define $W^{1,\varphi(\cdot)}_0(\Omega)$ to be the closure of $C^\infty_c(\Omega)$ in $W^{1,\varphi(\cdot)}(\Omega)$. For more information on these spaces, see [31].

We can use extrapolation to prove the Sobolev embedding theorem and then combine this with interpolation to prove a version of the Rellich–Kondratchov Theorem. We begin with the following weighted norm inequality: for all $f \in C^\infty_c(\Omega)$, $w \in A_1$, and $p \in [1, n)$,

$$\|f\|_{L^{p^*}(w)} \leqslant C\|\nabla f\|_{L^p(w^{p/p^*})},$$

where $p^* = \frac{np}{n-p}$ is the Sobolev exponent of p. See [20, Lemma 4.31] or [18, Lemma 6.32].

We use extrapolation (Corollary 4.9) and the above inequality to prove the Sobolev embedding theorem. This improves [30, Corollary 6.9] by removing the extraneous assumptions that φ is an N-function and satisfies (aInc).

Corollary 6.12 (Sobolev embedding)**.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0)–(A2) *and* $(\mathrm{aDec})_{p_+}$ *for some $p_+ < n$. Define $\psi^{-1}(x,t) := t^{-\frac{1}{n}}\varphi^{-1}(x,t)$. Then*

$$W^{1,\varphi(\cdot)}_0(\mathbb{R}^n) \hookrightarrow L^{\psi(\cdot)}(\mathbb{R}^n).$$

To prove our main compact embedding theorem, we first give a preliminary result.

Theorem 6.13. *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0)–(A2) *and* (aDec)*. Then*

$$W_0^{1,\varphi(\cdot)}(\Omega) \hookrightarrow\hookrightarrow L^{\varphi(\cdot)}(\Omega).$$

Theorem 6.13 is proved in essentially the same way as in the case of variable Lebesgue spaces; see [24, Theorem 8.4.2] for details. The proof requires the following lemma, which can be proved in two ways, both analogous to the proofs in variable Lebesgue spaces. First, it can be proved directly, arguing as in [24, Lemma 8.4.1] (based on Young's convolution inequality [24, Lemma 4.6.3]). Alternatively, it can be proved via extrapolation and a density argument as in [18, Theorem 5.11].

Lemma 6.14. *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0)–(A2) *and let Ψ be a standard mollifier. Then*

$$\|\Psi_\varepsilon * u - u\|_{\varphi(\cdot)} \lesssim \varepsilon \|u\|_{\varphi(\cdot)}$$

for every $u \in W^{1,\varphi(\cdot)}(\mathbb{R}^n)$. Here $\Psi_\varepsilon(t) := \varepsilon^{-n}\Psi(\frac{t}{\varepsilon})$.

If we now use Corollary 5.5 to interpolate between the inequalities in Corollary 6.12 and Theorem 6.13, we get the Rellich–Kondrachov Theorem for generalized Orlicz–Sobolev spaces. The analogous result for variable Sobolev spaces was proved in [24, Corollary 8.4.4].

Theorem 6.15 (Compact Sobolev embedding)**.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfy* (A0)–(A2) *and* $(\mathrm{aDec})_{p_+}$ *for $p_+ < n$. Define $\psi^{-1}(x,t) := t^{-\alpha}\varphi^{-1}(x,t)$, with $\alpha \in [0, \frac{1}{n})$. Then*

$$W_0^{1,\varphi(\cdot)}(\Omega) \hookrightarrow\hookrightarrow L^{\psi(\cdot)}(\Omega).$$

Remark 6.16. In Theorems 6.13 and 6.15 we have assumed that $\varphi \in \Phi_w(\mathbb{R}^n)$ even though the results hold in domain Ω. Alternatively, we could assume $\varphi \in \Phi_w(\Omega)$ if Ω is bounded and quasiconvex or if we replace assumption (A1) by $(\mathrm{A1})_\Omega$. See [31] for more details, including the definition of the last condition.

References

[1] A. Almeida, P. Harjulehto, P. Hästö, and T. Lukkari, *Riesz and Wolff potentials and elliptic equations in variable exponent weak Lebesgue spaces*, Ann. Mat. Pura Appl. (4) **194** (2015), no. 2, 405–424, DOI 10.1007/s10231-013-0382-2. MR3322428

[2] Kenneth F. Andersen and Russel T. John, *Weighted inequalities for vector-valued maximal functions and singular integrals*, Studia Math. **69** (1980/81), no. 1, 19–31. MR604351

[3] Pascal Auscher and José María Martell, *Weighted norm inequalities, off-diagonal estimates and elliptic operators. I. General operator theory and weights*, Adv. Math. **212** (2007), no. 1, 225–276, DOI 10.1016/j.aim.2006.10.002. MR2319768

[4] M. Avci and A. Pankov, *Multivalued elliptic operators with nonstandard growth*, Advances in Nonlinear Analysis, March 2016, 1–14.

[5] Paolo Baroni, Maria Colombo, and Giuseppe Mingione, *Harnack inequalities for double phase functionals*, Nonlinear Anal. **121** (2015), 206–222, DOI 10.1016/j.na.2014.11.001. MR3348922

[6] P. Baroni, M. Colombo, and G. Mingione, *Nonautonomous functionals, borderline cases and related function classes*, Algebra i Analiz **27** (2015), no. 3, 6–50; English transl., St. Petersburg Math. J. **27** (2016), no. 3, 347–379. MR3570955

[7] Colin Bennett and Robert Sharpley, *Interpolation of operators*, Pure and Applied Mathematics, vol. 129, Academic Press, Inc., Boston, MA, 1988. MR928802

[8] Jöran Bergh and Jörgen Löfström, *Interpolation spaces. An introduction*, Grundlehren der Mathematischen Wissenschaften, No. 223, Springer-Verlag, Berlin-New York, 1976. MR0482275

[9] S. Chanillo, *A note on commutators*, Indiana Univ. Math. J. **31** (1982), no. 1, 7–16, DOI 10.1512/iumj.1982.31.31002. MR642611

[10] Yunmei Chen, Stacey Levine, and Murali Rao, *Variable exponent, linear growth functionals in image restoration*, SIAM J. Appl. Math. **66** (2006), no. 4, 1383–1406, DOI 10.1137/050624522. MR2246061

[11] R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. **51** (1974), 241–250. MR0358205

[12] R. R. Coifman, R. Rochberg, and Guido Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635, DOI 10.2307/1970954. MR0412721

[13] Maria Colombo and Giuseppe Mingione, *Bounded minimisers of double phase variational integrals*, Arch. Ration. Mech. Anal. **218** (2015), no. 1, 219–273, DOI 10.1007/s00205-015-0859-9. MR3360738

[14] Michael Cowling, José García-Cuerva, and Hendra Gunawan, *Weighted estimates for fractional maximal functions related to spherical means*, Bull. Austral. Math. Soc. **66** (2002), no. 1, 75–90, DOI 10.1017/S0004972700020694. MR1922609

[15] Lucas Chaffee and David Cruz-Uribe, *Necessary conditions for the boundedness of linear and bilinear commutators on Banach function spaces*, Math. Inequal. Appl. **21** (2018), no. 1, 1–16. MR3716205

[16] D. Cruz-Uribe and A. Fiorenza, *Endpoint estimates and weighted norm inequalities for commutators of fractional integrals*, Publ. Mat. **47** (2003), no. 1, 103–131, DOI 10.5565/PUBLMAT_47103_05. MR1970896

[17] D. Cruz-Uribe and A. Fiorenza, $L\log L$ *results for the maximal operator in variable* L^p *spaces*, Trans. Amer. Math. Soc. **361** (2009), no. 5, 2631–2647, DOI 10.1090/S0002-9947-08-04608-4. MR2471932

[18] David V. Cruz-Uribe and Alberto Fiorenza, *Variable Lebesgue spaces*, Foundations and harmonic analysis, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, Heidelberg, 2013. MR3026953

[19] D. Cruz-Uribe, A. Fiorenza, J. M. Martell, and C. Pérez, *The boundedness of classical operators on variable* L^p *spaces*, Ann. Acad. Sci. Fenn. Math. **31** (2006), no. 1, 239–264. MR2210118

[20] David V. Cruz-Uribe, José Maria Martell, and Carlos Pérez, *Weights, extrapolation and the theory of Rubio de Francia*, Operator Theory: Advances and Applications, vol. 215, Birkhäuser/Springer Basel AG, Basel, 2011. MR2797562

[21] David Cruz-Uribe and C. J. Neugebauer, *The structure of the reverse Hölder classes*, Trans. Amer. Math. Soc. **347** (1995), no. 8, 2941–2960, DOI 10.2307/2154763. MR1308005

[22] David Cruz-Uribe and Li-An Daniel Wang, *Extrapolation and weighted norm inequalities in the variable Lebesgue spaces*, Trans. Amer. Math. Soc. **369** (2017), no. 2, 1205–1235, DOI 10.1090/tran/6730. MR3572271

[23] Lars Diening, *Maximal function on Musielak-Orlicz spaces and generalized Lebesgue spaces* (English, with English and French summaries), Bull. Sci. Math. **129** (2005), no. 8, 657–700, DOI 10.1016/j.bulsci.2003.10.003. MR2166733

[24] Lars Diening, Petteri Harjulehto, Peter Hästö, and Michael Růžička, *Lebesgue and Sobolev spaces with variable exponents*, Lecture Notes in Mathematics, vol. 2017, Springer, Heidelberg, 2011. MR2790542

[25] L. Diening, P. Hästö, and A. Nekvinda, Open problems in variable exponent Lebesgue and Sobolev spaces. In *FSDONA04 Proceedings*, Drabek and Rakosnik (eds.), Milovy, Czech Republic, Academy of Sciences of the Czech Republic, Prague, 2005, 38–58.

[26] Javier Duoandikoetxea, *Fourier analysis,* translated and revised from the 1995 Spanish original by David Cruz-Uribe, Graduate Studies in Mathematics, vol. 29, American Mathematical Society, Providence, RI, 2001. MR1800316

[27] Flavia Giannetti and Antonia Passarelli di Napoli, *Regularity results for a new class of functionals with non-standard growth conditions*, J. Differential Equations **254** (2013), no. 3, 1280–1305, DOI 10.1016/j.jde.2012.10.011. MR2997371

[28] Eleonor Harboure, Roberto A. Macías, and Carlos Segovia, *Extrapolation results for classes of weights*, Amer. J. Math. **110** (1988), no. 3, 383–397, DOI 10.2307/2374616. MR944321

[29] P. Harjulehto and P. Hästö, *Reflexivity and uniform convexity of generalized Orlicz spaces*, preprint.

[30] Petteri Harjulehto and Peter Hästö, *The Riesz potential in generalized Orlicz spaces*, Forum Math. **29** (2017), no. 1, 229–244, DOI 10.1515/forum-2015-0239. MR3592600

[31] Petteri Harjulehto, Peter Hästö, and Riku Klén, *Generalized Orlicz spaces and related PDE*, Nonlinear Anal. **143** (2016), 155–173, DOI 10.1016/j.na.2016.05.002. MR3516828

[32] P. Harjulehto, P. Hästö, V. Latvala, and O. Toivanen, *Critical variable exponent functionals in image restoration*, Appl. Math. Lett. **26** (2013), no. 1, 56–60, DOI 10.1016/j.aml.2012.03.032. MR2971400

[33] Petteri Harjulehto, Peter Hästö, Út V. Lê, and Matti Nuortio, *Overview of differential equations with non-standard growth*, Nonlinear Anal. **72** (2010), no. 12, 4551–4574, DOI 10.1016/j.na.2010.02.033. MR2639204

[34] Peter A. Hästö, *The maximal operator on generalized Orlicz spaces*, J. Funct. Anal. **269** (2015), no. 12, 4038–4048, DOI 10.1016/j.jfa.2015.10.002. MR3418078

[35] Peter A. Hästö, *Corrigendum to "The maximal operator on generalized Orlicz spaces" [J. Funct. Anal. 269 (2015) 4038–4048] [MR3418078]*, J. Funct. Anal. **271** (2016), no. 1, 240–243, DOI 10.1016/j.jfa.2016.04.005. MR3494250

[36] Peter Hästö, Yoshihiro Mizuta, Takao Ohno, and Tetsu Shimomura, *Sobolev inequalities for Orlicz spaces of two variable exponents*, Glasg. Math. J. **52** (2010), no. 2, 227–240, DOI 10.1017/S0017089509990292. MR2610969

[37] Mohammed Kbiri Alaoui, Tamer Nabil, and Mohamed Altanji, *On some new non-linear diffusion models for the image filtering*, Appl. Anal. **93** (2014), no. 2, 269–280, DOI 10.1080/00036811.2013.769132. MR3169204

[38] M. A. Krasnosel′skiĭ and Ja. B. Rutickiĭ, *Convex functions and Orlicz spaces*, translated from the first Russian edition by Leo F. Boron, P. Noordhoff Ltd., Groningen, 1961. MR0126722

[39] Gary M. Lieberman, *On the regularity of the minimizer of a functional with exponential growth*, Comment. Math. Univ. Carolin. **33** (1992), no. 1, 45–49. MR1173745

[40] Fumi-Yuki Maeda, Yoshihiro Mizuta, and Takao Ohno, *Approximate identities and Young type inequalities in variable Lebesgue-Orlicz spaces $L^{p(\cdot)}(\log L)^{q(\cdot)}$*, Ann. Acad. Sci. Fenn. Math. **35** (2010), no. 2, 405–420, DOI 10.5186/aasfm.2010.3526. MR2731699

[41] Fumi-Yuki Maeda, Yoshihiro Mizuta, Takao Ohno, and Tetsu Shimomura, *Approximate identities and Young type inequalities in Musielak-Orlicz spaces*, Czechoslovak Math. J. **63(138)** (2013), no. 4, 933–948, DOI 10.1007/s10587-013-0063-8. MR3165506

[42] Fumi-Yuki Maeda, Yoshihiro Mizuta, Takao Ohno, and Tetsu Shimomura, *Boundedness of maximal operators and Sobolev's inequality on Musielak-Orlicz-Morrey spaces*, Bull. Sci. Math. **137** (2013), no. 1, 76–96, DOI 10.1016/j.bulsci.2012.03.008. MR3007101

[43] Fumi-Yuki Maeda, Yoshihiro Sawano, and Tetsu Shimomura, *Some norm inequalities in Musielak-Orlicz spaces*, Ann. Acad. Sci. Fenn. Math. **41** (2016), no. 2, 721–744, DOI 10.5186/aasfm.2016.4148. MR3525396

[44] Julian Musielak, *Orlicz spaces and modular spaces*, Lecture Notes in Mathematics, vol. 1034, Springer-Verlag, Berlin, 1983. MR724434

[45] Hidegorô Nakano, *Modulared semi-ordered linear spaces*, Maruzen Co., Ltd., Tokyo, 1950. MR0038565

[46] Hidegorô Nakano, *Topology of linear topological spaces*, Maruzen Co., Ltd., Tokyo, 1951. MR0046560

[47] Takao Ohno and Tetsu Shimomura, *Trudinger's inequality for Riesz potentials of functions in Musielak-Orlicz spaces*, Bull. Sci. Math. **138** (2014), no. 2, 225–235, DOI 10.1016/j.bulsci.2013.05.007. MR3175020

[48] W. Orlicz, *Über konjugierte Exponentenfolgen*, Stud. Math. **3** (1931), 200–211.

[49] Carlos Pérez, *Sharp estimates for commutators of singular integrals via iterations of the Hardy-Littlewood maximal function*, J. Fourier Anal. Appl. **3** (1997), no. 6, 743–756, DOI 10.1007/BF02648265. MR1481632

[50] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 146, Marcel Dekker, Inc., New York, 1991. MR1113700

[51] José Luis Rubio de Francia, *Factorization and extrapolation of weights*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), no. 2, 393–395, DOI 10.1090/S0273-0979-1982-15047-9. MR663793

[52] Elias M. Stein, *Maximal functions. I. Spherical means*, Proc. Nat. Acad. Sci. U.S.A. **73** (1976), no. 7, 2174–2175. MR0420116

[53] Agnieszka Świerczewska-Gwiazda, *Nonlinear parabolic problems in Musielak-Orlicz spaces*, Nonlinear Anal. **98** (2014), 48–65, DOI 10.1016/j.na.2013.11.026. MR3158445

Department of Mathematics, University of Alabama, Tuscaloosa, Alabama 35487-0350

Email address: dcruzuribe@ua.edu

Department of Mathematical Sciences, P.O. Box 3000, FI-90014 University of Oulu, Finland – and – Department of Mathematics and Statistics, FI-20014 University of Turku, Finland

Email address: peter.hasto@oulu.fi

URL: http://cc.oulu.fi/∼phasto/