Nonlinear Approximation and Function Spaces of Dominating Mixed Smoothness

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der Fakultät für Mathematik und Informatik
der Friedrich-Schiller-Universität Jena

von Dipl.-Math. Markus Hansen
geboren am 19. November 1982 in Eisenach

Gutachter

1. Prof. Dr. Winfried Sickel - *Summa cum laude*
2. Prof. Dr. Stephan Dahlke - *Summa cum laude*
3. Prof. Dr. Erich Novak - *Summa cum laude*

Tag der öffentlichen Verteidigung: 16.08.2010

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Dr. Winfried Sickel. During the times of my Diploma thesis and my Ph.D. studies he was always supporting, encouraging, attentive and willing to discuss every minor detail or question. Moreover, I would like to thank the members of the research groups "Function spaces" and "Theoretical numerics" in Jena, in particular Prof. Dr. Erich Novak, PD Aicke Hinrichs, Prof. Dr. Hans-Jürgen Schmeißer and Prof. Dr. Hans Triebel for valuable discussions and numerous hints and comments during the preparation of this work.

Zusammenfassung

Motiviert durch zahlreiche Anwendungen beispielsweise aus der Finanzmathematik oder der Quantenmechanik, rückte die Approximation hochdimensionaler Probleme in den letzten Jahrzehnten immer mehr in den Blickpunkt analytischer und numerischer Untersuchungen. Hochdimensional bedeutet in diesem Zusammenhang, dass die betrachteten Funktionen von sehr vielen Variablen abhängen, typischerweise mehrere 100 oder 1000. Ein immer wieder kehrendes Problem stellt dabei die Beobachtung dar, dass die Komplexität und der zeitliche Aufwand der Verfahren sehr schnell mit der Anzahl der Variablen ansteigt, in vielen Fällen exponentiell. Dieses wird zusammengefasst bezeichnet als *Fluch der Dimension*. Trotz der in den letzten Jahren enorm angewachsenen Rechenleistung von Computern scheinen Probleme mit 15 Variablen oder mehr nach heutigen Stand außer Reichweite.
In der theoretischen und numerischen Analysis verfolgt man daher zwei Wege, um die Auswirkungen dieses Fluches entweder zu umgehen, oder sie zumindest zu reduzieren. Einerseits schränkt man die Klasse der betrachteten Funktionen (in geeigneter Weise) ein, und zum anderen sucht man nach „besseren“ (angepassteren) Algorithmen.
Diese Dissertation betrachtet dazu jeweils einen Ansatz. Der erste Teil (die Kapitel 1–4) beschäftigt sich dazu mit Tensorprodukten von Funktionenräumen und mit Besov- und Triebel-Lizorkin-Funktionenräumen mit gemischt dominanter Glattheit. Der zweite Teil der Arbeit (Kapitel 5 und 6) schließlich behandelt eine spezielle Variante nichtlinearer Approximation, die sogenannte *beste* m*-Term Approximation*.

Bei der Behandlung hochdimensionaler Probleme erscheint ein Typ von Funktionen als besonders einfach und erlaubt in Rechnungen oftmals drastische Vereinfachungen der Probleme. Dabei handelt es sich um Tensorprodukt-Funktionen, also jene Funktionen, welche eine Darstellung als Produkt niederdimensionaler Funktionen gestatten, also etwa $f(x_1, \ldots, x_n) = g_1(x_1) \cdots g_n(x_n)$. Einige Vorteile solcher Funktionen beispielsweise für das Abspeichern von Funktionswerten sind evident. Will man etwa die Funktionswerte einer Funktion f, definiert auf $[0,1]^2$, in den Gitterpunkten $(\frac{i}{n}, \frac{j}{n})$, $0 \leq i, j \leq n$, abspeichern, so benötigt dies für allgemeine Funktionen $(n+1)^2$ Funktionsauswertungen, im Gegensatz zu $2(n+1)$ für Tensorprodukte. In höheren Dimensionen wird dieser Unterschied noch deutlicher ($(n+1)^d$ gegenüber $d(n+1)$).
Davon ausgehend erscheint die Betrachtungen von Tensorprodukträumen ein natürlicher Ansatz. Tensorprodukte von Vektorräumen und anderen Strukturen sind in der Algebra wohlbekannt, ebenso im Zusammenhang mit der topologischen Struktur von Banachräumen. Für letzteres gehen einige der grundlegenden Begriffe auf Schatten [67] zurück, und entscheidend vorangetrieben wurde die Entwicklung nach Grothendieck [31]. Grob gesprochen enthalten die Tensorprodukträume alle Tensorproduktfunktionen und deren Linearkombinationen. Allerdings ist nur in wenigen Situationen eine explizitere Beschreibung der topologischen Struktur dieser Räume bekannt, d.h. ausgedrückt in Form von Integrabilität oder Glattheit. Außer an den Tensorprodukten selbst ist man daher auch an solchen Räume interessiert, die in gewisser (nicht näher spezifizierter) Hinsicht „dicht“ an solchen Tensorprodukträumen dran sind. Die Hoffnung ist dann die, dass solche Räume ähnliche Eigenschaften beispielsweise für approximationstheoretische Betrachtungen zeigen, etwa für Fehlerabschätzungen bis auf zusätzliche logarithmische Fakoren.
Wie schon angedeutet sind diese Fragen eng verknüpft mit der Suche nach besseren Modellen für die zu approximierenden Objekte. Ein einfaches Beispiel: Wir betrachten für zwei k-

fach stetig differenzierbare Funktionen deren Tensorprodukt h, wobei $h(x,y) = f(x)g(y)$. Nach der Produktregel wird dieses wiederum k-mal differenzierbar sein, aber es gilt noch mehr. Da die Ableitugen $f^{(i)}$ und $g^{(j)}$, $0 \leq i,j \leq k$, stetig sein sollen, sind auch die partiellen Ableitungen $\frac{\partial^{i+j}h}{\partial x^i \partial y^j}(x,y) = f^{(i)}(x)g^{(j)}(y)$ der Ordnung $i+j$ stetig, also auch für Ordnungen höher als k bis zu $2k$. Andererseits ist sicher nicht jede partielle Ableitung der Ordnung $2k$ stetig, z.B. $\frac{\partial^i h}{\partial x^i}(x,y) = f^{(i)}(x)g(y)$ oder $\frac{\partial^j h}{\partial y^j}(x,y) = f(x)g^{(j)}(y)$, $k+1 \leq i,j \leq 2k$, müssen nicht existieren. Man verliert also Informationen, wenn man h lediglich als k-mal differenzierbare Funktion behandelt, aber man kann sie auch nicht ohne weiteres als $2k$-mal differenzierbar betrachten.
Dieses Beispiel verdeutlicht, dass in Zusammenhang mit Tensorprodukten eine *isotrope* Theorie ungeeignet ist, bei der alle Variablen in der gleichen Weise behandelt werden. Vielmehr muss man ein gerichtetes, eigenständigeres Verhalten der Variablen gestatten. Ein mögliches Modell in dieser Richtung ist gegeben durch Funktionenräume gemischt dominanter Glattheit.

Funktionenräume sind ein wichtiges Hilfsmittel in vielen Bereichen der Analysis, insbesondere bei der Behandlung von partiellen Differentialgleichungen. Eines der bekanntesten Beispiele solcher Funktionenräume sind Sobolev-Räume $W_p^m(\mathbb{R}^d)$, eingeführt in den 1930er Jahren von S.L. Sobolev. Diese sind charakterisiert durch ihre Norm,

$$\| f \,|W_p^m(\mathbb{R}^d)\| = \sum_{|\alpha|\leq m} \| D^\alpha f \,|L_p(\mathbb{R}^d)\| \,, \qquad 1 < p < \infty \,, \quad m \in \mathbb{N}_0 \,,$$

d.h. man fordert, dass alle verallgemeinerten Ableitungen der Funktion $f \in L_p(\mathbb{R}^d)$ bis zur Ordnung m ebenfalls wieder zu $L_p(\mathbb{R}^d)$ gehört. Die Skala der Triebel-Lizorkin-Räume $F^s_{p,q}(\mathbb{R}^d)$ kann als Verallgemeinerung der Sobolev-Skala verstanden werden.
Die Besov-Räume $B^s_{p,q}(\mathbb{R}^d)$ sind mit den Triebel-Lizorkin-Räumen eng verwandt. Ursprünglich eingeführt durch S.M. Nikol'skij (1951) und O.V. Besov (1959/60) stellte sich schnell heraus, dass diese Räume eng verknüpft sind mit einigen zentralen Problemen der Approximationstheorie wie Approximation periodischer Funktionen durch Partialsummen der zugeordneten Fourierreihe.
Beide Skalen von Funktionenräumen erlauben eine Untersuchung mit Hilfe fourieranalytischer Techniken. Als eine wesentliche Eigenschaft erweist sich dabei ihre Charakterisierung mit Hilfe von Wavelet-Systemen. Eine Funktion oder Distribution gehört demzufolge zu $B^s_{p,q}(\mathbb{R}^d)$ bzw. $F^s_{p,q}(\mathbb{R}^d)$ genau dann, wenn die zugehörige Folge von Wavelet-Koeffizienten in einem zugeordneten Folgenraum liegt.
Funktionenräume gemischt-dominanter Glattheit wurden erstmals in den 1960er Jahren von Nikol'skij definiert. Auch in diesem Fall begann das Studium mit Räumen vom Sobolev-Typ. Man betrachtet dabei die Norm

$$S_p^{(k_1,k_2)}W(\mathbb{R}^2) = \Bigg\{ f \in L_p(\mathbb{R}^2) \,:\, \| f \,|S_p^{(k_1,k_2)}W(\mathbb{R}^2)\| = \| f \,|L_p(\mathbb{R}^2)\| + \left\| \frac{\partial^{k_1} f}{\partial x_1^{k_1}} \middle| L_p(\mathbb{R}^2) \right\| + \left\| \frac{\partial^{k_2} f}{\partial x_2^{k_2}} \middle| L_p(\mathbb{R}^2) \right\| + \left\| \frac{\partial^{k_1+k_2} f}{\partial x_1^{k_1} \partial x_2^{k_2}} \middle| L_p(\mathbb{R}^2) \right\| < \infty \Bigg\} ,$$

wobei $1 < p < \infty$ und $k_i = 0,1,2,...$ $(i = 1,2)$. In der Folgezeit wurden ebenfalls Räume vom Besov- ($S^{(r_1,r_2)}_{p,q}B(\mathbb{R}^2)$) und Triebel-Lizorkin-Typ ($S^{(r_1,r_2)}_{p,q}F(\mathbb{R}^2)$) betrachtet. In den letzten Jahren wurde auch für diese Räume eine entsprechende Wavelet-Charakterisierung bewiesen. Darüber hinaus konnten die Sobolev-Räume und die Besov-Räume $S^{(r_1,r_2)}_{p,p}B(\mathbb{R}^2)$ als Tensorprodukte identifiziert werden,

$$S^{(r_1,r_2)}_{p,p}B(\mathbb{R}^2) = B^{r_1}_{p,p}(\mathbb{R}) \otimes_p B^{r_2}_{p,p}(\mathbb{R}) \,, \qquad S^{(k_1,k_2)}_{p}W(\mathbb{R}^2) = W^{k_1}_{p}(\mathbb{R}) \otimes_p W^{k_2}_{p}(\mathbb{R}) \,.$$

Neben solchen Tensorprodukten von Funktionen in einer Variablen treten in Anwendungen auch „mehrdimensionale" Varianten auf. Beispielsweise kann man bei der elektronischen Schrödingergleichung zeigen, dass ihre Eigenfunktionen zu Räumen der Form $\bigcap_{i=1}^{N} W^1_2(\mathbb{R}^3) \otimes_2 \cdots W^2_2(\mathbb{R}^3) \otimes_2 \cdots \otimes W^1_2(\mathbb{R}^3)$ gehören, wobei $W^2_2(\mathbb{R}^3)$ der $i$te Faktor des Tensorproduktes ist, siehe [98]. Ein Ziel dieser Arbeit war ein Gegenstück für die angegebenen Tensorprodukt-Identitäten mit Hilfe einer Wavelet-Charakterisierung für eine entsprechende Modifikation der Räume gemischt dominanter Glattheit.
Nachdem also Tensorprodukträume und Räume gemischt dominanter Glattheit geeignete Modelle für hochdimensionale Probleme liefern, bleibt die Frage nach einem passenden Approximationsverfahren als dem zweiten Teil der beschriebenen Ansätze zur Reduktion des Fluchs der Dimension.

Der zweite Teil der Dissertation beschäftigt daher mit der (nichtlinearen) besten m-Term Approximation und auch kurz mit der (linearen) Approximation vom hyperbolischen Kreuz. Während früher lineare Approximationsverfahren, beschrieben durch lineare Operatoren, favorisiert wurden, rückten in den letzten Jahren und Jahrzehnten nichtlineare Verfahren zunehmend in den Fokus. Die Idee dahinter ist ziemlich einfach: Man erhofft sich ein besseres Fehlerverhalten dadurch, dass man nicht mehr eine ganze Klasse durch einen festen Algorithmus approximiert, sondern den Algorithmus an die zu approximierende Funktion anpasst.
In dieser Arbeit habe ich mich auf ein spezielles solches Verfahren konzentriert, die sogenannte m-Term Approximation. Während viele lineare Verfahren dadurch beschieben werden können, dass die vorgegebene Funktion durch Elemente eines (fixierten) Unterraumes angenähert wird, passt man bei der m-Term Approximation den Unterraum an die Funktion an. Dazu gibt man sich eine Menge von Funktionen in dem betrachteten Raum vor, genannt *dictionary*, und betrachtet alle Unterräume, die von höchstens m Elementen des dictionary aufgespannt werden. Anschließend optimiert man über alle diese Unterräume, wodurch der optimale Unterraum von der gegebenen Funktion abhängt.
Gerade aufgrund dieses Optimierungsprozesses ist m-Term Approximation vor allem ein theoretisches Verfahren, da typischerweise zur Berechnung der optimalen Approximation vollständige Kenntnis der Funktion benötigt wird.
Unmittelbar an der Definition der m-Term Approximation wird deutlich, dass diese entscheidend auch vom verwendeten dictionary abhängt. Die natürliche Wahl wäre eine Basis im betrachteten Funktionenraum. Dabei haben seit den 80er Jahren vor allem die bereits erwähnten Wavelet-Basen große Aufmerksamkeit erregt. Da diese auch hervorragend geeignet sind im Zusammenhang mit den beschriebenen Funktionenräumen gemischt dominanter Glattheit, motivierte dies das Studium der besten m-Term Approximation bezüglich Wavelet-Basen in solchen Funktionenräumen.

Die beschriebene Zweiteilung der Aufgabenstellung spiegelt sich auch im Aufbau der Dissertation wieder. Der erste Teil (die Kapitel 1–4) beschäftigt sich mit Tensorprodukten und mit den Funktionenräumen gemischt dominanter Glattheit. Im ersten Abschnitt wiederholen wir zunächst die Definitionen und Wavelet-Charakterisierungen der isotropen Funktionenräume und der Räume gemischt-dominanter Glattheit. Darüber hinaus wer-

den die grundlegenden Begriffe und Aussagen für Tensorprodukte von Banach- und auch Quasi-Banachräumen besprochen. Insbesondere werden Sobolev-Räume und auch gewisse Besov-Räume als Tensorprodukträume identifiziert.
Danach folgt die fourieranalytische Behandlung der Sobolev-, Besov- und Triebel-Lizorkin-Räume gemischt dominanter Glattheit $S_p^{\overline{r}}W(\mathbb{R}^{\overline{d}})$, $S_{p,q}^{\overline{r}}B(\mathbb{R}^{\overline{d}})$ und $S_{p,q}^{\overline{r}}F(\mathbb{R}^{\overline{d}})$ bezüglich allgemeiner Variablenunterteilungen gemäß $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}$. Nach den grundlegenden Aussagen, die weitestgehend parallel zu denen für die isotropen Räume sind, behandeln die Abschnitte 3 und 4 weitere wesentliche Hilfsmittel für diese Räume, insbesondere Charakterisierungen durch Lokale Mittel und Atome. Hauptresultat dieses Teils der Dissertation ist die Charakterisierung durch Tensorprodukt-Wavelets, welche grob aussagt, dass eine Funktion (bzw. eine temperierte Distribution) zu einem Funktionenraum $S_{p,q}^{\overline{r}}B(\mathbb{R}^{\overline{d}})$ oder $S_{p,q}^{\overline{r}}F(\mathbb{R}^{\overline{d}})$ gehört genau dann, wenn die zugehörige Folge der Wavelet-Koeffizienten in einem zugeordneten Folgenraum liegt. Darüber hinaus liefert diese Zuordnung einen Isomorphismus vom Funktionen- auf den Folgenraum. Als eine erste Folgerung aus dieser Charakterisierung erhalten wir die angestrebte Darstellung von Sobolev- und Besov-Räumen als Tensorprodukte ihrer isotropen Gegenstücke,

$$S_{p,p}^{(r_1,r_2)}B(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) = B_{p,p}^{r_1}(\mathbb{R}^{d_1}) \otimes_p B_{p,p}^{r_2}(\mathbb{R}^{d_2}) \,,$$
$$S_p^{(k_1,k_2)}W(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) = W_p^{k_1}(\mathbb{R}^{d_1}) \otimes_p W_p^{k_2}(\mathbb{R}^{d_2}) \,.$$

Der genannte Isomorphismus ermöglicht es in den Kapiteln 5 und 6 das Studium von Einbettungen und der m-Term Approximation von den Funktionenräume auf entsprechende Probleme für Folgenräume zu übertragen. Kapitel 5 behandelt zunächst notwendige und hinreichende Bedingungen für stetige und kompakte Einbettungen, da diese wiederum eine notwendige Bedingung für die Untersuchungen der m-Term Approximation darstellen.
Der sechste Abschnitt ist der zweite zentrale Bestandteil dieser Arbeit. Nach der Bereitstellung weiterer Hilfsmittel, insbesondere von Approximationsräumen, und einiger a priori Vereinfachungen, die die Betrachtung der Folgenräume ermöglicht, folgt zunächst die explizite Konstruktion nahezu optimaler Approximationen unter gewissen Zusatzvoraussetzungen an die beteiligten Parameter. Dabei ist „explizite“ Konstruktion dahingehend zu verstehen, dass man dafür die vollständige Kenntnis der zu approximierenden Folge voraussetzt. Der zweite Schritt ist dann die Charakterisierung des asymptotischen Fehlerverhaltens der m-Term Approximation. Ausgehend von den Resultaten, die die expliziten Konstruktionen liefern, können wir diese auf fast alle anderen möglichen Parameterkonstellationen mit Hilfe von reeller Interpolation und Reiterationsaussagen ausdehnen.
Der Schlussabschnitt 7 fasst die auf diese Weise erhaltenen Ergebnisse nochmals zusammen. Außerdem werden dabei die Resultate von den Folgenräume rückübertragen auf die zugeordneten Funktionenräume. Neben Räumen auf dem ganzen $\mathbb{R}^d$ können dabei auch Funktionenräume auf Gebieten Ω wie dem Einheitswürfel $[0,1]^d$ behandelt werden. Das Hauptresultat über beste m-Term Approximation in L_p-Räumen lässt sich dann wie folgt beschreiben: Seien $1 < p_0, p_1 < \infty$ und $t \in \mathbb{R}$ derart, dass $t > \max(0\,, \frac{1}{p_0} - \frac{1}{p_1})$. Dann gilt

$$\sup_{\|f|F\| \le 1} \inf_{\#\Lambda \le m} \inf_{c_j \in \mathbb{C}} \Big\| f - \sum_{j \in \Lambda} c_j \Psi_j \Big| L_{p_1}(\mathbb{R}^d) \Big\| \overset{\log}{\asymp} m^{-t} \,, \qquad m \ge 2 \,,$$

wobei die exakten logarithmischen Ordnungen für fast alle möglichen Parameter bekannt sind. Dabei ist F einer der Räume $S_{p_0}^{\overline{t}}H(\mathbb{R}^{\overline{d}})$ oder $S_{p_0,p_0}^{\overline{t}}B(\mathbb{R}^{\overline{d}})$, $\overline{t} = t\overline{d}$.
Den Abschluss der Arbeit bildet ein Vergleich der erzielten Resultate mit solchen, die in den letzten Jahren von Temlyakov und Dinh Dung publiziert worden sind.

Preface

Over the course of the last three decades highdimensional approximation, i.e. approximation of functions of many variables, became an important topic in several fields of mathematics and its applications. To name only some prominent examples, various problems in financial mathematics deal with PDEs or integration problems in a large number of variables (typically 360 or multiples thereof). One of the basic objects in Quantum mechanics and related fields of theoretical physics and chemistry is the electronic Schrödinger equation, where the number of variables is a multiple of the number of particles involved which generally increases the higher the demands on the materials which are to be designed. As a last example may serve the large field of stochastic processes and stochastic differential equations whose deterministic numerical treatment often produces highdimensional problems.
We want to deal with these types of problems from a more theoretical point of view. In the past many types of approximation problems were investigated and solved, determining the (asymptotic behaviour of the) error of different methods and describing optimal solutions (approximants). However, most of these results have one thing in common which is nowadays summarized by the term *"Curse of Dimension"*. This describes the observation that the asymptotic convergence rate for many methods is of the form $n^{-k/d}$, where n stands for the degrees of freedom for the approximant, d is the number of variables, and k is some characteristic parameter of the function which shall be approximated, typically some smoothness parameter. Though theoretically sufficient, numerically this turns out to be a major drawback. Such results imply that in order to approximate the function to within some prescribed error ε the computational cost increases exponentially in the number of variables, which even with modern equipment cannot be handled, thus problems with $d > 15$ or 20 seem out of range.
Hence one mainly has two possibilities in trying to circumvent this obstacle: Either one finds "better" approximation methods, i.e. with better error behaviour, or one shrinks the class of objects to be approximated. A lot of effort has been put into both strategies, and this thesis deals with two particular approaches.

When talking about highdimensional functions one type of functions appears particularly simple and leads in many cases to a drastical reduction of the complexity. These are tensor product functions, where a function f can be written as the product of lower dimensional ones, e.g. $f(x_1, \ldots, x_n) = g_1(x_1) \cdots g_n(x_n)$. Some advantages of such functions are obvious. If the function f is defined on $[0,1]^2$ and one wants to store the function values at the lattice points $(\frac{i}{n}, \frac{j}{n})$, $0 \leq i, j \leq n$, for general functions this would mean $(n+1)^2$ values, opposed to $2(n+1)$ values for tensor product functions. In higher dimension this generalizes to n^d or dn, respectively. Similar comparisons can be made for the numerical solution of differential equations etc. A simple example where tensor product functions lead to an enormous simplification of the problem is given by the separation ansatz for analytical solutions of linear partial differential equations like the heat or the wave equation.
Unfortunately not every highdimensional function can be identified as a tensor product, hence we are back with another approximation problem when asking: Which (classes of) functions can be approximated well by tensor products? The most immediate answer is both simple and theoretically demanding: tensor product spaces.

Tensor product constructions for vector spaces and other types of algebraic structures are known to algebraists for a long time. In connection with Banach spaces most of the basic concepts were introduced by Schatten in 1943 [67], and their importance for Banach space theory became gradually clear after Grothendieck's groundbreaking paper [31], see also the recent monograph by Pietsch [62] for an historic overview.

Roughly spoken, these spaces collect linear combinations of tensor products and their limits. Hence it would be interesting to know which types of spaces can be identified which such tensor product spaces, and moreover, one should try to find spaces which are "close" to tensor product spaces. The latter stems from the fact that in many situations the tensor products itself are difficult to handle, whereas one might expect that, as far as most properties like performance for corresponding approximation methods, sufficiently "close" spaces (whatever the precise meaning of this phrase might be) should behave similarly, i.e. for error estimates possibly up to additional logarithmic terms.

This question is closely related to the above mentioned problem of finding a better model for the objects which shall be approximated, i.e. smaller classes of functions. A simple example: Suppose two functions f and g are both k-times continuously differentiable. Their tensor product h, where $h(x,y) = f(x)g(y)$, will of course again be k-times continuously differentiable, but some more is true. Since $f^{(i)}$ and $g^{(j)}$, $0 \leq i,j \leq k$, are assumed to be continuous, h will have continuous mixed partial derivatives $\frac{\partial^{i+j}h}{\partial x^i \partial y^j}(x,y) = f^{(i)}(x)g^{(j)}(y)$ of order $i+j$, which is possibly higher than k up to $2k$. But clearly not every partial derivative of h of order $2k$ will be continuous, e.g. $\frac{\partial^i h}{\partial x^i}(x,y) = f^{(i)}(x)g(y)$ and $\frac{\partial^j h}{\partial y^j}(x,y) = f(x)g^{(j)}(y)$, $k+1 \leq i,j \leq 2k$, need not even exist. In other words on the one hand we lose information about h when treating it as a k-times differentiable function, on the other hand we can not treat it as $2k$-times differentiable.

This simple example shows that tensor product functions do not fit into a classical "isotropic" theory, where all variables are treated alike. In some sense we have to allow for the variables to have some "separate behaviour", though clearly they are not totally independent of each other.

Hence, while such tensor product spaces might be a good approach towards the treatment of high-dimensional problems, it needs some more work beforehand. We have to determine suitable spaces from which tensor products are constructed. And moreover, we still have to find a proper approximation method.

While in former times linear methods were preferred over the course of the last three decades nonlinear methods came more into focus. The idea behind this is quite simple: One no longer wants to approximate a whole class of functions by the same method, and by this unified treatment potentially limiting the performance, but one adjusts the construction of the approximant (out of a prescribed class of methods) according to the given function. This concept has proved quite successful in numerous numerical applications such as adaptive finite element schemes.

In the sequel we concentrate on one particular nonlinear method, the so-called m-term approximation. The basic idea behind this concept seems quite natural. Most linear methods could be understood as: We fix the system of possible approximants as a finite-dimensional subspace of the given space of functions, and consider linear mappings into that subspace. When dealing with m-term approximation we instead fix a certain system of functions, called dictionary, and afterwards we consider finite linear combinations of elements of that dictionary. Though of course every finite selection of elements of the

dictionary spans a finite-dimensional subspace, this subspace may be a different one for every function we wish to approximate. In other words the approximating subspace depends on the given function.
However, best m-term approximation first of all is a theoretical concept. Though in many situations explicit constructions and even continuous mappings for near best approximations, i.e. optimal up to constant factors, are known there are currently no implementations realizing such constructions. The reason for this lies in the fact that those mappings require complete knowledge of the given function, e.g. when expanding it with respect to the dictionary all coefficients are needed for the construction of the approximant. Nevertheless this concept has proven to be useful, e.g. as a benchmark for implemented algorithms.

Having a closer look on the concept of m-term approximation it becomes quite clear that it depends heavily firstly on the spaces involved, i.e. which functions shall be approximated and the norm in which the error is measured, and secondly on the chosen dictionary. A natural choice would be to select a basis in the given space. But then immediately one might ask: What are "good" bases?
Since the 1980s one particular type of bases has attracted a lot of attention, namely wavelet-type bases. This term refers to systems which are constructed out of a single function (the *wavelet*) by dilations and translations. In recent years this type of bases has proved to be quite useful in numerous applications, perhaps the most prominent one being image compression. In particular one of the basic ideas behind the JPEG 2000-algorithm, representing the image by certain wavelet expansions and afterwards taking only the terms with the largest coefficients, can be interpreted as an m-term approximation of the given image.
Another important advantage of wavelet bases is given by the fact that many classical function spaces allow a characterization by such wavelet systems, usually in terms of decay conditions for the sequence of coefficients appearing in the corresponding wavelet expansions. This makes it possible to consider m-term approximation with respect to wavelet-type systems for a wide range of function classes.
Nowadays the approximative powers of m-term approximation for many combinations of function spaces and different dictionaries are well-understood. One result in this direction which is of particular interest for our further considerations is due to DeVore, Jawerth and Popov [15]. Without going into details at this point, some approximation classes of best m-term approximation with respect to wavelet bases in $L_p(\mathbb{R}^d)$, i.e. classes of functions with a common asymptotic behaviour of its error, are identified as Besov spaces $B^s_{q,q}(\mathbb{R}^d)$. Keeping in mind our interest in tensor product spaces, when dealing with m-term approximation tensor products of such Besov spaces are a reasonable starting point.
Another type of spaces is of particular interest in lots of applications, namely Sobolev spaces. For instance, spaces of this type are the natural framework for many boundary value problems for PDEs like the Laplace-Poisson equation. Moreover, the spaces $H^s_p(\mathbb{R}^d)$ are another scale which allows a characterization by wavelets. Additionally, they have been studied in connection with m-term approximation. The case $p = 2$ is of particular interest, since the spaces H^s_2 are Hilbert spaces. In this situation also corresponding tensor products were studied before, and it is a well-known result that these tensor product spaces can be identified with Sobolev spaces of dominating mixed smoothness.

Historically Sobolev spaces of such type on $\mathbb{R}^2$ were first introduced in the early 1960s by

S. M. Nikol'skij. More precisely, he proposed to consider spaces

$$S_p^{(r_1,r_2)}W(\mathbb{R}^2) = \Bigg\{ f \in L_p(\mathbb{R}^2) \,:\, \| f \,|S_p^{(r_1,r_2)}W(\mathbb{R}^2)\| = \| f \,|L_p(\mathbb{R}^2)\| + \left\| \frac{\partial^{r_1} f}{\partial x_1^{r_1}} \,\Big| L_p(\mathbb{R}^2)\right\| + \left\| \frac{\partial^{r_2} f}{\partial x_2^{r_2}} \,\Big| L_p(\mathbb{R}^2)\right\| + \left\| \frac{\partial^{r_1+r_2} f}{\partial x_1^{r_1}\partial x_2^{r_2}} \,\Big| L_p(\mathbb{R}^2)\right\| < \infty \Bigg\},$$

where $1 < p < \infty$ and $r_i = 0, 1, 2, ...$ $(i = 1, 2)$. The mixed derivative $\frac{\partial^{r_1+r_2} f}{\partial x_1^{r_1}\partial x_2^{r_2}}$ turned out to play a dominant role, which gave these scales of function spaces their name. Though not intended at that time a close connection to tensor products of functions is indicated by the observation that this norm applied to some tensor product factorizes, we remind on the example above.
Later on the definition was extended to include non-integral parameters r_1 and r_2 by Fourier analytic methods to obtain the spaces $S_p^{(r_1,r_2)}H(\mathbb{R}^2)$ and their respective multivariate analoga. In recent time W. Sickel and T. Ullrich [74] were able to give precise results concerning the mentioned connection of these Sobolev-type spaces with tensor products of the usual isotropic ones. In particular, every such space $S_p^{(r_1,r_2)}H(\mathbb{R}^2)$ can be identified with the tensor product space $H_p^{r_1}(\mathbb{R}) \otimes_p H_p^{r_2}(\mathbb{R})$.
In the same article they furthermore were able to deal with tensor products of Besov spaces. Starting with Nikol'skij's definition above a theory of spaces with dominating mixed smoothness was developed by many authors, primarily in the former Soviet Union. Important contributions, including generalizations to Besov-type spaces, were made by Amanov, Besov, Lizorkin, Nikol'skij and Potapov, to name only some of them. A first systematical treatment can be found in the monograph [1].
The mentioned result by Sickel and Ullrich now states that also the tensor product space $B_{p,p}^{r_1}(\mathbb{R}) \otimes_p B_{p,p}^{r_2}(\mathbb{R})$ can be identified with a function space of dominating mixed smoothness, namely the Besov space $S_{p,p}^{(r_1,r_2)}B(\mathbb{R}^2)$. This motivates having a closer look on these function spaces of dominating mixed smoothness first, prior to the treatment of corresponding m-term approximation problems.

Apart from the classical one via derivatives and differences, preferred by most authors from the former Soviet Union, there is another main approach towards function spaces with the help of Fourier analytic methods. Using this approach related scales of Besov- and Triebel-Lizorkin-type spaces of dominating mixed smoothness were introduced. For a detailed treatment of the spaces on $\mathbb{R}^2$ we refer to [71]. The Fourier analytical approach is based on a representation of functions and distributions by entire analytic functions,

$$f = \sum_{\bar{k}\in\mathbb{N}_0^N} \mathcal{F}^{-1}\big[\varphi_{k_1}^1 \otimes \cdots \otimes \varphi_{k_N}^N \mathcal{F} f\big]\,, \qquad \text{convergence in } \mathcal{S}'(\mathbb{R}^d)\,,$$

where $(\varphi_j^i)_{j\in\mathbb{N}_0}$, $i = 1, \ldots, N$, are decompositions of unity, known from the study of isotropic function spaces, and $\varphi_{k_1}^1 \otimes \cdots \otimes \varphi_{k_N}^N$ denotes their tensor product. This construction is a first example of the importance of tensor product constructions in the treatment of these function spaces, and it once more indicates a close connection to tensor product spaces.
The main advantage of this strategy lies in the possibly unified treatment of the spaces $S_{p,p}^{\bar{r}}B(\mathbb{R}^d)$ and $S_p^{\bar{r}}H(\mathbb{R}^d) = S_{p,2}^{\bar{r}}F(\mathbb{R}^d)$, and moreover, in many situations we obtain additional information via the introduction of a further fine index.

So far when talking about tensor products we dealt with the product of functions in one variable, accordingly for the function spaces. The generalization of this situation is obvious, and it is motivated by a result of H. Yserentant [98]. He considered the electronic Schrödinger equation and proved that the eigenfunctions of the corresponding Hamilton operator belong to spaces $\bigcap_{i=1}^{N} H_2^1(\mathbb{R}^3) \otimes \cdots \otimes H_2^2(\mathbb{R}^3) \otimes \cdots \otimes H_2^1(\mathbb{R}^3)$, where $H_2^2(\mathbb{R}^3)$ is the ith factor of the tensor product. Hence the idea is to split the set of d variables into N groups, where each group may behave differently, but within each group all variables are treated alike. This approach leads to a slight modification of the function spaces of dominating mixed smoothness whose treatment shall be one of the main objectives for our considerations.

According to the above considerations this thesis consists of two main parts. The first one is devoted to the study of function spaces. To begin with we recall in Chapter 1 the definitions and wavelet characterizations of first the isotropic Sobolev, Besov and Triebel-Lizorkin spaces, and later on this is done for the spaces of dominating mixed smoothness. Moreover, we present the basic concepts and notions for tensor products of Banach spaces, and have a closer look at their extension to quasi-Banach spaces. At the end of that chapter we state the precise formulation of the previously mentioned result establishing a connection between tensor product spaces and function spaces of dominating mixed smoothness.
Chapter 2 then presents the definition and basic properties of our main objects of study, the function spaces of dominating mixed smoothness with respect to general variable splittings $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}$. The treatment is restricted to those facts needed in the later considerations, though a greater number of results could easily be obtained with the help of the methods presented, following either the approaches for the isotropic spaces or those ones for the usual spaces of dominating mixed smoothness.
In Chapter 3 we derive a characterization of our function spaces in terms of the Peetre maximal operator and local means. This characterization and its corollaries will be the main tool to establish theorems for atomic and wavelet decompositions in Chapter 4. That characterization by tensor products of Daubechies-type wavelets in Theorem 4.3.1 is the main result of the first part of this thesis. As a corollary we can prove the identities

$$S^{\overline{r}}_{p,p}B(\mathbb{R}^{\overline{d}}) = B^{r_1}_{p,p}(\mathbb{R}^{d_1}) \otimes_p \cdots \otimes_p B^{r_N}_{p,p}(\mathbb{R}^{d_N}) \,, \qquad 0 < p < \infty \,,$$

and

$$S^{\overline{r}}_{p}H(\mathbb{R}^{\overline{d}}) = H^{r_1}_{p}(\mathbb{R}^{d_1}) \otimes_p \cdots \otimes_p H^{r_N}_{p}(\mathbb{R}^{d_N}) \,, \qquad 1 < p < \infty \,,$$

and hence verify the aspired relation between tensor products of isotropic Sobolev and Besov spaces in arbitrary dimensions on the one hand, and function spaces of dominating mixed smoothness for general variable splittings on the other hand.
The second part of the thesis, Chapters 5 and 6, then deals with the problem of best m-term approximation. More precisely, we study this problem for sequence spaces which are related to the Besov and Triebel-Lizorkin spaces. This reduction can be done, because the mentioned wavelet characterization establishes an isomorphism from the function spaces onto those sequence spaces. As a first step, in Chapter 5 results for continuous and compact embeddings are presented since the boundedness of the embedding is a necessary condition for the m-term width to be finite.

After introducing further notions and tools at the beginning of Chapter 6, the calculation of the asymptotic behaviour of the m-term widths is split into two steps. The first one consists in explicit constructions for near best m-term approximation, establishing said asymptotics for a restricted range of parameters. As mentioned before, though those constructions are explicit they can't be reformulated as algorithms since they require the complete knowledge of the given sequence. Afterwards, in the second step we extend the results obtained from those explicit constructions. The main tool in this step are approximation spaces related to m-term approximation. With the help of embedding, interpolation and reiteration results for such approximation spaces we were able to characterize the asymptotic behaviour of the error of the best m-term approximation for almost all possible constellations of parameters.
Finally, the last chapter of this thesis presents our main results on m-term approximation. After introducing function spaces on domains and deriving a description in terms of wavelets, we can apply said description to transfer the previously obtained results for the asymptotics from the sequence spaces to related function spaces. The main result on best m-term approximation in L_p-spaces then reads as follows: Let $1 < p_0, p_1 < \infty$ and $t \in \mathbb{R}$, such that $t > \max(0\,, \frac{1}{p_0} - \frac{1}{p_1})$. Then it holds

$$\sup_{\|f|F\|\leq 1} \inf_{\#\Lambda \leq m} \inf_{c_j \in \mathbb{C}} \left\| f - \sum_{j \in \Lambda} c_j \Psi_j \middle| L_{p_1}(\mathbb{R}^d) \right\| \overset{\log}{\sim} m^{-t}\,, \qquad m \geq 2\,,$$

where the exact orders of the logarithm are known for almost all possible parameters. Here F is one of the spaces $S^{\bar{t}}_{p_0}H(\mathbb{R}^{\bar{d}})$ or $S^{\bar{t}}_{p_0,p_0}B(\mathbb{R}^{\bar{d}})$, $\bar{t} = t\bar{d}$.
At the end of Chapter 7 we furthermore compare our results to related work by Temlyakov [79] and Dinh Dung [21, 22]. Concerning our starting point, high-dimensional approximation, our results are both positive and negative. On the one hand, the obtained asymptotic rates m^{-t} show considerable progress compared to the rate $m^{-t/d}$ for isotropic spaces. However, the curse of dimension cannot be overcome in this way, since the behaviour of the constants involved remains an open problem. While for most situations on sequence space level explicit constants could be derived (on some occasions we did just that) this information gets lost upon applying the wavelet isomorphism. Apart from the constants also the occurring logarithmic terms depend on the dimension (their exponent being proportional to $N - 1$), which has a major influence in numerical applications.
Another open problem remains the question, whether the tensor product structure of the function spaces is particularly helpful when considering m-term approximation. On the one hand we used a dictionary with tensor product structure (the employed wavelet system consisted of tensor product functions), on the other hand the applied techniques made very limited use of this additional structure, even more so after the transfer to sequence spaces. Exploiting this property might possibly lead to (considerable) simplifications of the proofs.

Contents

1 Preliminaries/Introduction

In this chapter, we review the definitions and wavelet-characterizations of the isotropic Sobolev, Besov and Triebel-Lizorkin spaces and their counterparts for dominating mixed smoothness. Moreover, we study tensor product spaces, at first in their abstract formulation, and afterwards we apply that abstract theory to tensor products of Besov and Sobolev spaces.

1.1 Notation

As usual, we denote by $\mathbb{R}^d$ the d-dimensional real Euclidean space, $\mathbb{Z}$ is the set of all integers, $\mathbb{N}$ are the natural numbers, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ are all non-negative integers. Moreover, $\mathbb{C}$ denotes the complex numbers, and $\mathbb{R}_+$ stands for the collection of all non-negative real numbers. Finally, $\mathbb{R}^n_+ = (\mathbb{R}_+)^n$ denotes the set of all vectors with solely non-negative components.
Points of the underlying Euclidean space are denoted by $x, y, z, \ldots$, and their components are numbered from 1 to d, i.e. $x = (x_1, \ldots, x_d)$. Moreover, in later subsections we will use a splitting $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N} =: \mathbb{R}^{\overline{d}}$, where $N, d_1, \ldots, d_N \in \mathbb{N}$, $d = d_1 + \cdots + d_N$, $\overline{d} \in \mathbb{N}^N$ and $\overline{d} = (d_1, \ldots, d_N)$. Accordingly, we will split the components of $x \in \mathbb{R}^d$ as per $x = (x^1, \ldots, x^N)$, $x^i \in \mathbb{R}^{d_i}$, where $x^i = (x^i_1, \ldots, x^i_{d_i}) = (x_{d_1+\cdots+d_{i-1}+1}, \ldots, x_{d_1+\cdots+d_{i-1}+d_i})$. Any other d-tuple will be dealt with analogously, in particular lattice points from $\mathbb{Z}^d$.
Besides d-tuples, we will need N-dimensional vectors. For distinction from d-dimensional ones these will be denoted with a bar, i.e. we will write $\overline{r} \in \mathbb{R}^N$, $\overline{\nu} \in \mathbb{N}_0^N$ etc.
The notation $a > b$ for n-tuples a and b (where $n = d_i$, $n = N$ or $n = d$, according to the concrete case) will be used, if $a_i > b_i$ holds for every $i = 1, \ldots, n$. The relations $a \geq b$, $a \leq b$ and $a < b$ are understood similarly. Moreover, the expression $a \not< b$ means the negation of $a < b$, i.e. there is at least one index $i \in \{1, \ldots, n\}$ with $a_i \geq b_i$. Similar notations are used for the other relations. Finally, despite a slight abuse of notation, by $a > \lambda$ for an n-tuple a and a real number $\lambda \in \mathbb{R}$ we mean $a_i > \lambda$ for every $i = 1, \ldots, n$.
For n-tuples we will use three different norms. If not indicated otherwise, with $|a|$ the usual Euclidean norm is meant, i.e. $|a| \equiv |a|_2 = \left(\sum_{i=1}^n |a_i|^2\right)^{1/2}$. In particular for integer parameters, we will use $|a|_1 = \sum_{i=1}^n |a_i|$. The last one is the usual maximum-norm, i.e. $|a|_\infty = \max_{i=1,\ldots,n} |a_i|$.
For every multiindex $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, its length is given by $|\alpha| \equiv |\alpha|_1 = \alpha_1 + \cdots + \alpha_d$. The derivatives $D^\alpha = \partial^{|\alpha|}/\big(\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}\big) = D^{\alpha^1} \cdots D^{\alpha^N}$ are understood in the distributional (weak) sense. Moreover, we put $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d} = \big(x^1\big)^{\alpha^1} \cdots \big(x^N\big)^{\alpha^N}$. Finally, we define for $\alpha \in \mathbb{N}_0^d$ a vector $\overline{\alpha} \in \mathbb{N}_0^N$ by $\overline{\alpha} = (|\alpha^1|, \ldots, |\alpha^N|)$, according to the chosen splitting of variables.

Let $\mathcal{S}(\mathbb{R}^d)$ be the Schwartz space of all complex-valued, rapidly decaying, infinitely differentiable functions on $\mathbb{R}^d$. By $\mathcal{F}\varphi$, $\mathcal{F}(\varphi)$ or $\widehat{\varphi}$ we denote the d-dimensional Fourier transform of $\varphi \in \mathcal{S}(\mathbb{R}^d)$, i.e.

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

where $x \cdot \xi = x_1\xi_1 + \cdots x_d\xi_d$ is the standard scalar product on $\mathbb{R}^d$. Accordingly, we will

use $a \cdot b = a_1 b_1 + \cdots a_n b_n$ for arbitrary n-tuples. The inverse Fourier transform will be denoted by $\mathcal{F}^{-1}\varphi$, $\mathcal{F}^{-1}(\varphi)$ or $\varphi^\vee$. Both $\mathcal{F}$ and $\mathcal{F}^{-1}$ are extended in the usual way to the dual space of $\mathcal{S}(\mathbb{R}^d)$, the space of tempered distributions $\mathcal{S}'(\mathbb{R}^d)$.
Occasionally, we have to distinguish between the d-dimensional, the d_i-dimensional and the one-dimensional Fourier transform. In that case, we will write $\mathcal{F}_d$, $\mathcal{F}_{d_i}$ and $\mathcal{F}_1$, respectively, and their inverses will be denoted by $\mathcal{F}_d^{-1}$, $\mathcal{F}_{d_i}^{-1}$ and $\mathcal{F}_1^{-1}$. We would like to point out, that for tensor product functions $\varphi(x) = \varphi_1(x^1)\cdots\varphi_N(x^N) = \big(\varphi_1 \otimes \cdots \otimes \varphi_N\big)(x)$, $\varphi_i \in \mathcal{S}(\mathbb{R}^{d_i})$, these transformations are connected:

$$(\mathcal{F}_d\varphi)(\xi) = (\mathcal{F}_{d_1}\varphi_1)(\xi^1)\cdots(\mathcal{F}_{d_N}\varphi_N)(\xi^N) = \big((\mathcal{F}_{d_1}\varphi_1)\otimes\cdots\otimes(\mathcal{F}_{d_N}\varphi_N)\big)(\xi)\,, \tag{1.1.1}$$

valid for all $\xi = (\xi^1, \ldots, \xi^N) \in \mathbb{R}^d$. In particular, for functions $\varphi(x) = \varphi_1(x_1)\cdots\varphi_n(x_n) = \big(\varphi_1 \otimes \cdots \otimes \varphi_n\big)(x)$, $\varphi_i \in \mathcal{S}(\mathbb{R})$, we find

$$(\mathcal{F}_n\varphi)(\xi) = (\mathcal{F}_1\varphi_1)(\xi_1)\cdots(\mathcal{F}_1\varphi_n)(\xi_n) = \big((\mathcal{F}_1\varphi_1)\otimes\cdots\otimes(\mathcal{F}_1\varphi_n)\big)(\xi)\,. \tag{1.1.2}$$

Now let $0 < p, q \le \infty$. As usual, the space $L_p(\mathbb{R}^d)$ consists of all (equivalence classes of) Lebesgue-measurable functions, such that

$$\big\| f \,\big|L_p(\mathbb{R}^d)\big\| = \left(\int_{\mathbb{R}^d} |f(x)|^p dx\right)^{1/p}$$

is finite. If $p \ge 1$, then these spaces can be interpreted as subsets of $\mathcal{S}'(\mathbb{R}^d)$. Any statement that a distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ belongs to $L_p(\mathbb{R}^d)$ hence includes that f is regular, where such a distribution and its generator are identified.
For a sequence of (complex-valued) measurable functions $\big(f_k\big)_{k\in A}$ on $\mathbb{R}^d$, where A is an arbitrary countable index set, we put

$$\big\|f_k\big|\ell_q(L_p)\big\| = \left(\sum_{k\in A}\big\|f_k\big|L_p(\mathbb{R}^d)\big\|^q\right)^{1/q} = \left(\sum_{k\in A}\left(\int_{R^d}|f_k(x)|^p dx\right)^{q/p}\right)^{1/q} \tag{1.1.3}$$

as well as

$$\big\|f_k\big|L_p(\ell_q)\big\| = \left\|\left(\sum_{k\in A}|f_k(\cdot)|^q\right)^{1/q}\middle|L_p(\mathbb{R}^d)\right\| = \left(\int_{R^d}\left(\sum_{k\in A}|f_k(x)|^q\right)^{p/q}dx\right)^{1/p}, \tag{1.1.4}$$

with the usual modification in case p and/or $q = \infty$. If there is no danger of confusion, we won't explicitly mention the index set for ℓ_q-norms.
Now, let X and Y be quasi-Banach spaces. Then we denote by $\mathcal{L}(X,Y)$ the class of all linear bounded operators $T : X \longrightarrow Y$, equipped with the usual operator (quasi-)norm

$$\big\| T \,\big|\mathcal{L}(X,Y)\big\| = \sup_{\|f|X\|\le 1}\big\| Tf \,\big|\, Y \,\big\|, \qquad T \in \mathcal{L}(X,Y)\,,$$

which turns $\mathcal{L}(X,Y)$ again into a quasi-Banach space.

We will write $a_+ = \max(a, 0)$ for an arbitrary real number $a \in \mathbb{R}$. Furthermore, let $\overline{\sigma}_p$ and $\overline{\sigma}_{p,q}$ be defined by

$$\sigma_p^i = d_i\left(\frac{1}{p} - 1\right)_+ \qquad \text{and} \qquad \sigma_{p,q}^i = d_i\left(\frac{1}{\min(p,q)} - 1\right)_+, \qquad i = 1, \ldots, N\,,$$

or in a shorter way $\overline{\sigma}_p = \overline{d}\big(\frac{1}{p} - 1\big)_+$, $\overline{\sigma}_{p,q} = \overline{d}\big(\frac{1}{\min(p,q)} - 1\big)_+$, for all $0 < p,q \le \infty$. Furthermore, let the symbol denote $[x]$ the integer part of the real number $x \in \mathbb{R}$, that is, the uniquely determined integer $m \in \mathbb{Z}$, such that $m \le x < m+1$.

All unimportant constants will be denoted by c, c', C etc. The concrete value of these constants may vary from one formula to the next, but remains the same within one chain of (in)equalities. We will write $A \lesssim B$, if there is a constant $c > 0$, independent of the relevant parameters, such that $A \le cB$. Finally, the notation $A \sim B$ is an abbreviation of $A \lesssim B \lesssim A$.

1.2 Isotropic spaces

We begin with a short repetition of the definitions and wavelet characterizations for the isotropic function spaces.

1.2.1 Definitions of the isotropic function spaces

Definition 1.2.1. Let $1 < p < \infty$.

(i) For any $m \in \mathbb{N}_0$ we put

$$W_p^m(\mathbb{R}^n) := \Big\{ f \in L_p(\mathbb{R}^n) : \big\| f \,\big| W_p^m(\mathbb{R}^n) \big\| := \sum_{|\alpha| \le m} \big\| D^\alpha f \big| L_p(\mathbb{R}^n) \big\| < \infty \Big\},$$

where the derivatives have to be understood in the distributional sense.

(ii) For an arbitrary $s \in \mathbb{R}$, we define

$$H_p^s(\mathbb{R}^n) := \Big\{ f \in \mathcal{S}'(\mathbb{R}^n) : \big\| f \,\big| H_p^s(\mathbb{R}^n) \big\| < \infty \Big\},$$

where

$$\big\| f \,\big| H_p^s(\mathbb{R}^n) \big\| := \big\| \mathcal{F}^{-1} (1 + |\xi|^2)^{s/2} \mathcal{F} f \big| L_p(R^n) \big\|.$$

The spaces $H_p^s(\mathbb{R}^n)$ are the Sobolev spaces of fractional order (Bessel potential spaces).

Remark 1.2.1. Obviously, we have $W_p^0(\mathbb{R}^n) = H_p^0(\mathbb{R}^n) = L_p(\mathbb{R}^n)$. Furthermore, it is a well known fact that for any $m \in \mathbb{N}_0$ it holds $W_p^m(\mathbb{R}^n) = H_p^m(\mathbb{R}^n)$ in the sense of equivalent norms.

Definition 1.2.2. We define the set $\Phi(\mathbb{R}^n)$ to be the collection of all systems $\big(\varphi_j\big)_{j=0}^\infty \subset \mathcal{S}(\mathbb{R}^n)$, such that

$$\begin{cases} \operatorname{supp} \varphi_0 \subset \big\{ t \in \mathbb{R}^n : |t| \le 2 \big\}, \\ \operatorname{supp} \varphi_j \subset \big\{ t \in \mathbb{R}^n : 2^{j-1} \le |t| \le 2^{j+1} \big\}, \qquad \text{if } j = 1, 2, \dots, \end{cases} \tag{1.2.1}$$

for every $\alpha \in \mathbb{N}_0^n$ exist positive constants c_α with

$$2^{j|\alpha|} |D^\alpha \varphi_j(t)| \le c_\alpha \qquad \text{for all } j \in \mathbb{N}_0 \text{ and all } t \in \mathbb{R}^n \tag{1.2.2}$$

and

$$\sum_{j=0}^{\infty} \varphi_j(t) = 1 \qquad \text{for every } t \in \mathbb{R}^n. \tag{1.2.3}$$

Any such system $\left(\varphi_j\right)_{j=0}^{\infty}$ is called a smooth dyadic decomposition of unity.

Remark 1.2.2. We shall give an example of such a decomposition, whose special structure will be helpful in proofs. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be a function with the following properties:

(i) $\varphi(x) = 1 \qquad$ for all $|x| \leq 1$,
(ii) $\varphi(x) = 0 \qquad$ for all $|x| \geq 2$.

Now we put $\varphi_0 = \varphi$, $\varphi_1 = \varphi(\cdot/2) - \varphi$ and

$$\varphi_j(x) = \varphi_1(2^{-j+1}x), \qquad x \in \mathbb{R}^n, j \in \mathbb{N}.$$

Then it can be easily verified that this system $\left(\varphi_j\right)_{j=0}^{\infty}$ satisfies the conditions (1.2.1)–(1.2.3), that is, for every such system we have $\left(\varphi_j\right)_{j\in\mathbb{N}_0} \in \Phi(\mathbb{R}^n)$.

Now we can proceed to the definition of the isotropic Besov and Triebel-Lizorkin spaces.

Definition 1.2.3. Let $s \in \mathbb{R}$, $0 < q \leq \infty$, and let $\varphi = \left(\varphi_j\right)_{j\in\mathbb{N}_0} \in \Phi(\mathbb{R}^n)$.

(i) Let $0 < p \leq \infty$. Then $B^s_{p,q}(\mathbb{R}^n)$ is defined as the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^n)$, such that

$$\big\| f \,|B^s_{p,q}(\mathbb{R}^n)\big\| := \left(\sum_{j=0}^{\infty} 2^{jsq}\big\|\big(\varphi_j\widehat{f}\big)^{\vee}\big|L_p(\mathbb{R}^n)\big\|^q\right)^{1/q} = \big\|2^{js}\big(\varphi_j\widehat{f}\big)^{\vee}\big|\ell_q(L_p)\big\|$$

is finite.

(ii) Let $0 < p < \infty$. Then $F^s_{p,q}(\mathbb{R}^n)$ is the collection of all $f \in \mathcal{S}'(\mathbb{R}^n)$, such that

$$\big\| f \,|F^s_{p,q}(\mathbb{R}^n)\big\| := \left\|\left(\sum_{j=0}^{\infty} 2^{jsq}\big|\big(\varphi_j\widehat{f}\big)^{\vee}(\cdot)\big|^q\right)^{1/q}\middle|L_p(\mathbb{R}^n)\right\| = \big\|2^{js}\big(\varphi_j\widehat{f}\big)^{\vee}\big|L_p(\ell_q)\big\|$$

is finite.

Remark 1.2.3. It is a well-known fact, that the spaces $B^s_{p,q}(\mathbb{R}^n)$ and $F^s_{p,q}(\mathbb{R}^n)$ are independent of the system φ, in the sense that different decompositions of unity generate equivalent (quasi-)norms, see e.g. [83, Proposition 2.3.2/1]

Remark 1.2.4. The above function spaces are closely connected, though apart from the obvious identity $B^s_{p,p}(\mathbb{R}^n) = F^s_{p,p}(\mathbb{R}^n)$, it holds $B^s_{p,q}(\mathbb{R}^n) \neq F^s_{p,q}(\mathbb{R}^n)$ whenever $p \neq q$ (see [83, Section 2.3.9]). As usual, we will use the notation $A^s_{p,q}(\mathbb{R}^n)$ to refer to both Besov and Triebel-Lizorkin spaces.

Both scales of function spaces have been studied extensively in the last fifty years, since the original definition of Besov in 1959/60. They cover many classical scales of function spaces like Sobolev spaces, (real) Hardy spaces or Hölder-Zygmund spaces. In particular, there is a well-known corollary from Littlewood-Paley theory, stating that for $1 < p < \infty$ we have $H^s_p(\mathbb{R}^n) = F^s_{p,2}(\mathbb{R}^n)$ in the sense of equivalent norms. For the basic (Fourier analytical) investigation of these spaces we refer mainly to the work of Triebel, in particular in [84, 86] many historical remarks are to be found, and Peetre, particularly the monograph [56].

Here we shall only be concerned with basic results on wavelet characterizations for these spaces which will be described in the next subsection.

1.2.2 Wavelet characterizations

We start with the following key assertion of the wavelet theory which is due to Daubechies. It can be found in [13] or [95].

Theorem 1.2.1. For every $s \in \mathbb{N}$, there are real-valued functions

$$\psi_0, \psi_1 \in C^s(\mathbb{R}) \tag{1.2.4}$$

with compact support and

$$\int_{\mathbb{R}} t^\alpha \psi_1(t)dt = 0, \qquad \alpha = 0, 1 \ldots, s, \tag{1.2.5}$$

such that

$$\big\{\psi_0(\cdot - m) : m \in \mathbb{Z}\big\} \cup \big\{\psi_{j,m} : j \in \mathbb{N}_0, m \in \mathbb{Z}\big\} \tag{1.2.6}$$

with

$$\psi_{j,m}(t) = 2^{j/2}\psi_1(2^j t - m)\,, \qquad j \in \mathbb{N}_0, m \in \mathbb{Z}\,,$$

is an orthonormal basis of $L_2(\mathbb{R})$.

The function ψ_0 is called (orthogonal) scaling function, and ψ_1 is the associated wavelet. There are now two standard constructions to obtain bases for $L_2(\mathbb{R}^n)$ based on the system (1.2.6). The first one will be used here to obtain bases for isotropic spaces, the other one will be applied in Section 1.4.4 to construct bases for function spaces related to tensor product spaces.
Let ψ_0 and ψ_1 be functions as in Theorem 1.2.1. We define index sets Γ_n and γ_n by

$$\Gamma_n := \{0,1\}^n \setminus \{(0,\ldots,0)\} = \big\{G^1, \ldots, G^{2^n-1}\big\}\,, \qquad \gamma_n := \big\{1, \ldots, 2^n - 1\big\}\,.$$

Then we put

$$\psi^0_{0,m}(x) := \psi^0(x - m) := \psi_0(x_1 - m_1) \cdots \psi_0(x_n - m_n) \tag{1.2.7}$$

and

$$\psi^i_{j,m}(x) := 2^{jn/2}\psi^i(2^j x - m) := 2^{jn/2}\psi_{G^i_1}(2^j x_1 - m_1) \cdots \psi_{G^i_n}(2^j x_n - m_n)\,, \tag{1.2.8}$$

where

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad m = (m_1, \dots, m_n) \in \mathbb{Z}^n\,, \quad j \in \mathbb{N}_0, \quad i \in \gamma_n\,.$$

Finally, we put $I_j = \gamma_n$, $j \in \mathbb{N}$, and $I_0 = \gamma_n \cup \{0\}$. The n-dimensional counterpart of Theorem 1.2.1 then reads as follows.

Proposition 1.2.1. For every $s \in \mathbb{N}$, there are real-valued functions $\psi_0, \psi_1 \in C^s(\mathbb{R})$ with compact support and property (1.2.5), such that the system

$$\Psi = \Big\{ \psi^i_{j,m} : j \in \mathbb{N}_0, m \in \mathbb{Z}^n, i \in I_j \Big\}, \tag{1.2.9}$$

where the functions $\psi^i_{j,m}$ are defined as in (1.2.7) and (1.2.8), respectively, forms an orthonormal basis of $L_2(\mathbb{R}^n)$.

Thus, in addition to the scaling function ψ^0 we now use $2^n - 1$ associated wavelets ψ^i. The aspired characterization of the isotropic Besov and Triebel-Lizorkin spaces is formulated in terms of the wavelet coefficients and certain sequence spaces.

Definition 1.2.4. Let $s \in \mathbb{R}$ and $0 < p, q \leq \infty$.

(i) The space $b^s_{p,q}$ is defined as the collection of all sequences

$$\lambda = \left\{ \lambda^i_{j,m} \in \mathbb{C} : j \in \mathbb{N}_0, m \in \mathbb{Z}^n, i \in I_j \right\}, \tag{1.2.10}$$

such that

$$\| \lambda \,| b^s_{p,q} \| := \left(\sum_{j=0}^{\infty} 2^{j(s+\frac{n}{2}-\frac{n}{p})q} \sum_{i \in I_j} \left(\sum_{m \in \mathbb{Z}^n} |\lambda^i_{j,m}|^p \right)^{p/q} \right)^{1/q} \tag{1.2.11}$$

is finite.

(ii) Let $0 < p < \infty$. Moreover, we denote by $\mathcal{X}_{j,m}$ the characteristic function of the cube $Q_{j,m} = 2^{-j}([0,1]^n + m)$. Then the space $f^s_{p,q}$ is defined as the collection of all sequences as in (1.2.10), such that

$$\| \lambda \,| f^s_{p,q} \| := \left\| \left(\sum_{j=0}^{\infty} \sum_{i \in I_j} \sum_{m \in \mathbb{Z}^n} 2^{j(s+\frac{n}{2})q} |\lambda^i_{j,m}|^q \mathcal{X}_{j,m}(\cdot) \right)^{1/q} \middle| L_p(\mathbb{R}^n) \right\| \tag{1.2.12}$$

is finite.

If p and/or $q = \infty$ the (quasi-)norms have to be modified in the usual way.

Now we can present the theorem on the wavelet characterization.

Theorem 1.2.2. Let $s \in \mathbb{R}$ and $0 < p, q \leq \infty$. Moreover, let $\Psi \subset C^u(\mathbb{R}^n)$ be a wavelet system according to Proposition 1.2.1. If $u \in \mathbb{N}$ is chosen sufficiently large then the following statements are true:

(i) The space $B^s_{p,q}(\mathbb{R}^n)$ is the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^n)$, such that

$$f = \sum_{j=0}^{\infty} \sum_{i \in I_j} \sum_{m \in \mathbb{Z}^n} \lambda^i_{j,m} \psi^i_{j,m} , \qquad \text{convergence in } \mathcal{S}'(\mathbb{R}^n) , \tag{1.2.13}$$

where $\lambda = \big(\lambda^i_{j,m}\big)_{j \in \mathbb{N}_0, i \in I_j, m \in \mathbb{Z}^n} \in b^s_{p,q}$.

(ii) Let $0 < p < \infty$. Then the space $F^s_{p,q}(\mathbb{R}^n)$ is the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^n)$ that can be represented as in (1.2.13), where $\lambda = \big(\lambda^i_{j,m}\big)_{j \in \mathbb{N}_0, i \in I_j, m \in \mathbb{Z}^n} \in f^s_{p,q}$.

(iii) The coefficients in (1.2.13) are uniquely determined. It holds

$$\lambda^i_{j,m} = \langle f, \psi^i_{j,m} \rangle , \qquad j \in \mathbb{N}_0, i \in I_j, m \in \mathbb{Z}^n ,$$

where $\langle \cdot \,, \cdot \rangle$ denotes a dual pairing. Moreover, the mapping J, defined by

$$f \longmapsto \big(\langle f, \psi^i_{j,m} \rangle\big)_{j \in \mathbb{N}_0, i \in I_j, m \in \mathbb{Z}^n} ,$$

is an isomorphism from $B^s_{p,q}(\mathbb{R}^n)$ onto $b^s_{p,q}$ and from $F^s_{p,q}(\mathbb{R}^n)$ onto $f^s_{p,q}$.

(iv) If $\max(p, q) < \infty$, then the system Ψ in (1.2.9) is a basis in $A^s_{p,q}(\mathbb{R}^n)$.

Remark 1.2.5. The dual pairing in part (iii) needs some further explanation. Since we are working with compactly supported Daubechies wavelets instead of Meyer wavelets, the functions $\psi^i_{j,m}$ do not belong to $\mathcal{S}(\mathbb{R}^n)$, hence we cannot use the dual pairing in $\mathcal{S}'(\mathbb{R}^n)$. However, with suitable assumptions on u and with the help of embedding assertions and the characterizations of the dual spaces of Besov spaces we can always either interpret $\psi^i_{j,m}$ as a linear functional on some Besov space containing $f \in A^s_{p,q}(\mathbb{R}^n)$, or vice versa we can interpret f as a linear functional on a Besov space containing $\psi^i_{j,m}$. For details we refer to [86] or [89].
In both cases the expressions $\langle f, \psi^i_{j,m} \rangle$ then have to be understood in the sense of that respective dual pairing. Finally, for $f \in A^s_{p,q}(\mathbb{R}^n)$ it is possible to justify such a pairing directly in the sense of local means, see [86, Section 5.1.7].

For more details concerning these wavelet decompositions as well as proofs, we refer to the literature, e.g. Meyer [46], Kahane and Lemarié-Rieusset [42], or Triebel [87].

1.3 Tensor product spaces

In this section, we introduce some general notions on tensor products of (quasi-)Banach spaces, and afterwards we consider tensor products of weighted sequence spaces of ℓ_p-type and Besov spaces $B^s_{p,p}(\mathbb{R})$. For a brief overview concerning the basic notions and constructions we refer to [48]. A more detailed treatment, including the deep interrelation with the theory of operator ideals, can be found in [14]. Finally, for historic remarks see [62, Section 5.7.2].

1.3.1 Abstract background

In algebra tensor product constructions are known for several different structures. The starting point for the explicit construction for vector spaces X and Y (with respect to the same field; here we concentrate on real or complex vector spaces) is the free vector space $F(X,Y)$ on $X \times Y$, i.e. the set

$$\begin{aligned} F(X,Y) &:= \operatorname{span}\Big\{x \otimes y : x \in X, y \in Y\Big\} \\ &= \left\{\sum_{j=1}^{n} \lambda_j x_j \otimes y_j : x_j \in X, y_j \in Y, \lambda_j \in \mathbb{C}, j = 1, \dots, n, n \in \mathbb{N}\right\}. \end{aligned}$$

Afterwards the algebraic tensor product $X \otimes Y$ is defined as the quotient space of $F(X,Y)$ with respect to the subspace

$$\begin{aligned} U := \operatorname{span}\Big(&\big\{(x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y : x_1, x_2 \in X, y \in Y\big\} \\ &\cup \big\{x \otimes (y_1 + y_2) - x \otimes y_1 - x \otimes y_2 : x_1, x_2 \in X, y \in Y\big\} \\ &\cup \big\{\lambda(x \otimes y) - (\lambda x) \otimes y, \lambda(x \otimes y) - x \otimes (\lambda y) : x \in X, y \in Y, \lambda \in \mathbb{C}\big\}\Big). \end{aligned}$$

In this way some intuitive calculational rules are ensured.
The usual functional analytic approach for normed spaces X and Y is slightly different. Once more one starts with $F(X,Y)$, but this times this space is equipped with the following equivalence relation. We say $f = \sum_{j=1}^{n} \lambda_j x_j \otimes y_j \in F(X,Y)$ generates an operator $A_f : X' \longrightarrow Y$ by the determination

$$A_f \psi := \sum_{j=1}^{n} \lambda_j \psi(x_j) y_j \,, \qquad \psi \in X' \,.$$

Then we define for $f, g \in F(X,Y)$, $f = \sum_{j=1}^{n} \lambda_j^1 x_j^1 \otimes y_j^1$, $g = \sum_{j=1}^{m} \lambda_j^2 x_j^2 \otimes y_j^2$

$$f \simeq g \iff A_f(\psi) = A_g(\psi) \qquad \text{for all } \psi \in X' \,,$$

i.e. f and g generate the same operator from the dual space X' of X to Y. Of interest now is the quotient space $T(X,Y) = F(X,Y)/\simeq$, which is found to coincide as a vector space with $X \otimes Y$.
This approach applies to quasi-normed spaces as well, but since the dual space is possibly trivial, this equivalence relation as well as the respective quotient space might become trivial. To avoid this, i.e. to ensure the equivalence of both approaches, we have to impose certain restrictions on the quasi-normed spaces. This situation is clarified by the following lemma.

Lemma 1.3.1. Let X and Y be two quasi-normed spaces. Then it holds $T(X,Y) = X \otimes Y$ if, and only if, X' separates the points in X, i.e. for every $x \in X \setminus \{0\}$ there exists a functional $\varphi_x \in X'$, such that $\varphi_x(x) \neq 0$.

Proof. In order to show the coincidence of both spaces we have to show that $U = V := \{f \in F(X,Y) : A_f = 0\}$ holds. The inclusion $U \subset V$ is obvious. For the reverse inclusion

we remark that the condition on X' is equivalent to $A_{x\otimes y} \neq 0$ for all $x \neq 0$ and $y \neq 0$. To show now $V \subset U$, we show instead, that from $f \notin U$ follows $f \notin V$.
We shall use the fact, that for every $f \notin U$ there exists an (algebraically) equivalent representation $f = \sum_{i=1}^{n} x_i \otimes y_i$, where $\{x_1, \dots, x_n\} \subset X$ and $\{y_1, \dots, y_n\} \subset Y$ are linearly independent (this can be seen analogously to [48, Lemma 1.1]). The linearity of $f \longmapsto A_f$, the linear independency of $\{y_1, \dots, y_n\}$ and the assumption for X' (applied to the vectors $x_i \neq 0$) now yield $A_f \neq 0$. □

In case of Banach spaces, this condition is always fulfilled. On the other hand in those cases when X' and hence also $T(X,Y)$ is trivial the algebraic tensor product is of little use since many functional analytic methods used for tensor products fail. Hence there is no hope of a general abstract theory for quasi-Banach spaces.
However, for those quasi-Banach spaces of sequences and distributions we are interested in the dual spaces have the property described in the above lemma and are thus sufficiently rich to provide meaningful results, see Lemma 1.3.3. We shall add a few more remarks about quasi-Banach spaces in Section 1.3.4. For the rest of this subsection, we shall be concerned with Banach spaces only.
To derive Banach spaces from $X \otimes Y$, we equip it additionally with various norms, adapted to tensor products. Here, we shall use the following ones.

Definition 1.3.1. Let X and Y be Banach spaces.

(i) Let $[f] \in X \otimes Y$ with a representative $f \in F(X,Y)$,

$$f = \sum_{j=1}^{n} x_j \otimes y_j, \qquad x_j \in X, y_j \in Y, n \in \mathbb{N}. \tag{1.3.1}$$

Then the *injective tensor norm* $\lambda(\,\cdot\,, X, Y)$ is defined by

$$\begin{aligned}\lambda\big([f], X, Y\big) &:= \big\|A_f \big| \mathcal{L}(X', Y)\big\| \\ &= \sup\left\{ \left\| \sum_{j=1}^{n} \psi(x_j) y_j \right| Y \right\| \ : \ \psi \in X', \ \big\|\psi \big| X'\big\| \leq 1 \right\}.\end{aligned}$$

(ii) Let $1 \leq p \leq \infty$ and $1 = \frac{1}{p} + \frac{1}{p'}$. Then the *p-nuclear tensor norm* $\alpha_p(\,\cdot\,, X, Y)$ is defined by

$$\alpha_p\big([f], X, Y\big) := \inf\left\{ \left(\sum_{j=1}^{n} \big\|x_j \big| X\big\|^p \right)^{1/p} \cdot \sup_{\|\psi|Y'\| \leq 1} \left(\sum_{j=1}^{n} |\psi(y_j)|^{p'} \right)^{1/p'} \right\},$$

where the infimum is taken over all representatives $f \in [f]$ as in (1.3.1).

(iii) The *tensor norm* $\gamma(\,\cdot\,, X, Y)$ is defined as

$$\gamma\big([f], X, Y\big) := \inf\left\{ \sum_{j=1}^{n} \big\|x_j \big| X\big\| \cdot \big\|y_j \big| Y\big\| \ : f \in [f] \text{ as in (1.3.1)} \right\}.$$

Remark 1.3.1. Let α be one of the functionals λ, α_p ($1 \le p \le \infty$) or γ. If X and Y are Banach spaces, then α is indeed a norm on $X \otimes Y$. In particular, by the definition of the equivalence relation $\simeq$, $\lambda([f], X, Y)$ is independent of the representative $f \in [f]$. We denote by $X \otimes_\alpha Y$ the completion of $X \otimes Y$ with respect to this norm. Hence $X \otimes_\alpha Y$ is again a Banach space.
As usual, in the sequel we will always identify equivalence classes $[f]$ and their representatives f.
A vector $f \in F(X, Y)$, that can be represented as $f = x \otimes y$, is called *simple tensor* or *pure tensor* or *dyad*.
These norms α have some additional properties. First of all, they all are *crossnorms*, i.e. for any dyad $f = x \otimes y$ it holds

$$\alpha(x \otimes y, X, Y) = \big\|x\big|X\big\| \cdot \big\|y\big|Y\big\| , \qquad x \in X, y \in Y . \tag{1.3.2}$$

Moreover, they are so-called *uniform tensor norms*. Let $T_i : X_i \longrightarrow Y_i$, $i = 1, 2$, be bounded linear operators mapping Banach spaces X_i into Banach spaces Y_i. We define a linear mapping $T_1 \otimes T_2$ on $F(X_1, X_2)$ by the property

$$(T_1 \otimes T_2)(x_1 \otimes x_2) = (T_1 x_1) \otimes (T_2 x_2) , \qquad x_1 \in X_1, x_2 \in X_2 , \tag{1.3.3}$$

and linear extension. Then a crossnorm α is called a uniform tensor norm, if

$$\alpha\big((T_1 \otimes T_2)h, Y_1, Y_2\big) \le \big\|T_1\big|\mathcal{L}(X_1, Y_1)\big\| \cdot \big\|T_2\big|\mathcal{L}(X_2, Y_2)\big\|\alpha(h, X_1, X_2) \tag{1.3.4}$$

holds for all $h \in X_1 \otimes X_2$. Then there is a unique continuous extension T of $T_1 \otimes T_2$ to $X_1 \otimes_\alpha X_2$, such that

$$T : X_1 \otimes_\alpha X_2 \longrightarrow Y_1 \otimes_\alpha Y_2 \qquad \text{and} \qquad T \in \mathcal{L}(X_1 \otimes_\alpha X_2, Y_1 \otimes_\alpha Y_2) .$$

We will denote the extension T by $T_1 \otimes_\alpha T_2$. Finally, the norms λ, α_p and γ are *reasonable crossnorms*. For functionals $\varphi \in X'$ and $\psi \in Y'$, we can define a functional $\varphi \otimes \psi$ on $F(X, Y)$ via

$$(\varphi \otimes \psi)(x \otimes y) = \varphi(x) \cdot \psi(y) , \qquad x \in X, y \in Y ,$$

and linear extension. A crossnorm α is called reasonable, if $\varphi \otimes \psi$ is bounded on $X \otimes Y$ with respect to α, and its continuous extension $\varphi \otimes_\alpha \psi$ to $X \otimes_\alpha Y$ satisfies

$$\big\|\varphi \otimes_\alpha \psi\big|\big(X \otimes_\alpha Y\big)'\big\| = \big\|\varphi\big|X'\big\| \cdot \big\|\psi\big|Y'\big\| .$$

Though the next lemma is quite simple, it is important for our considerations. It can be found, e.g., in [74].

Lemma 1.3.2. Let X_1, X_2, Y_1, Y_2 be Banach spaces, and let α be a uniform tensor norm. Suppose that $T_1 \in \mathcal{L}(X_1, Y_1)$ and $T_2 \in \mathcal{L}(X_2, Y_2)$ are isomorphisms. Then also $T_1 \otimes_\alpha T_2$ is an isomorphism from $X_1 \otimes_\alpha X_2$ onto $Y_1 \otimes_\alpha Y_2$.

Remark 1.3.2. We add a remark on tensor products for more than two spaces. In this work this will be understood as iterated tensor products, i.e. by $X \otimes Y \otimes Z \equiv X \otimes (Y \otimes Z)$ we mean the tensor product of the space X with the space $Y \otimes Z$, accordingly for more factors and for completions with respect to tensor norms. Hence for a precise statement we have to add brackets to clarify the order of the iteration. Only if the resulting tensor product spaces coincide for every order of iteration we will instead just write $X \otimes Y \otimes Z$ etc.

1.3.2 Tensor products of Hilbert spaces

Before we turn our attention to more concrete examples, we shall deal with one more aspect of the abstract theory. The case of Hilbert spaces is well-investigated in the literature for many years. In fact these tensor products were the starting point of that theory (first considered by Murray and von Neumann in the late 1930s). In this case there exists a canonical construction of the tensor product space. Moreover, it has been shown by Grothendieck [32] that for the class of nuclear spaces (which Hilbert spaces are a special case of) all reasonable crossnorms are equivalent, hence there is de facto only one tensor product.

Hence, let G and H be Hilbert spaces with scalar products $\langle \cdot , \cdot \rangle_G$ and $\langle \cdot , \cdot \rangle_H$, respectively. Then we define a functional $\langle \cdot , \cdot \rangle$ on $F(G, H)$ be putting

$$\langle g_1 \otimes h_1 , g_2 \otimes h_2 \rangle := \langle g_1 , g_2 \rangle_G \cdot \langle h_1 , h_2 \rangle_H$$

and subsequent bilinear (or sesquilinear, respectively) extension. Clearly, this defines a scalar product on the algebraic tensor product $G \otimes H$. Moreover, this scalar product induces a norm on $G \otimes H$ in the usual way, and the completion $\mathcal{H}$ of $G \otimes H$ then becomes a Hilbert space as well. With a slight abuse of notation (as explained above), we will denote this completion again by $G \otimes H$. We refer to [48] for further details.

1.3.3 Tensor products of sequences, functions and distributions

For sequences, functions or distributions the (so far formal) expressions $f = \sum_{j=1}^{n} x_j \otimes y_j$ have a more immediate meaning. At first, for some sequences $a = (a_i)_{i \in I}$ and $b = (b_j)_{j \in J}$ of complex numbers, indexed by some arbitrary index sets I and J, we define

$$a \otimes^s b = \big((a \otimes^s b)_{i,j} \big)_{i \in I, j \in J} , \qquad (a \otimes^s b)_{i,j} = a_i \cdot b_j .$$

For functions $x : \Omega_1 \longrightarrow \mathbb{C}$ and $y : \Omega_2 \longrightarrow \mathbb{C}$, defined on some arbitrary sets Ω_1 and Ω_2, we have

$$x \otimes^f y : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{C}, \qquad (x \otimes^f y)(s,t) = x(s) \cdot y(t) , \quad s \in \Omega_1, t \in \Omega_2 .$$

Of course, this definition coincides with the one for sequences, if we interpret the sequence $a = (a_i)_{i \in I}$ as a mapping from I into $\mathbb{C}$. Finally, in the theory of distributions there exists a calculus for tensor products as well. While originally developed for distributions from $\mathcal{D}'(\mathbb{R}^n)$, the topological dual of $\mathcal{D}(\mathbb{R}^n) = C_0^\infty(\mathbb{R}^n)$ (equipped with the standard topology), the main assertions still hold true for tempered distributions. The calculus is based on the following proposition, which can be found in [74, Appendix B], see also [72] and [38].

Proposition 1.3.1. Let $S \in \mathcal{S}'(\mathbb{R}^{d_1})$ and $T \in \mathcal{S}'(\mathbb{R}^{d_2})$. Then there exists a uniquely determined distribution $U \in \mathcal{S}'(\mathbb{R}^{d_1+d_2})$, such that for all functions $\varphi \in \mathcal{S}(\mathbb{R}^{d_1})$ and $\psi \in \mathcal{S}(\mathbb{R}^{d_2})$

$$U(\varphi \otimes^f \psi) = S(\varphi) \cdot T(\psi) \tag{1.3.5}$$

holds. Furthermore, U is given explicitly by the formula

$$U\big(\rho(x,y)\big) = T_y\big(S_x(\rho(x,y))\big) = S_x\big(T_y(\rho(x,y))\big), \qquad \rho \in \mathcal{S}(\mathbb{R}^{d_1+d_2}) .$$

The distribution U is called the tensor product of S and T and is denoted by $S \otimes^D T$.

This proposition can immediately extended to (finite) linear combinations.

Proposition 1.3.2. Let $S_i \in \mathcal{S}'(\mathbb{R}^{d_1})$ and $T_i \in \mathcal{S}'(\mathbb{R}^{d_2})$, $i = 1, \ldots, n$. Then there exists a uniquely determined distribution $U \in \mathcal{S}'(\mathbb{R}^{d_1+d_2})$, such that for all functions $\varphi \in \mathcal{S}(\mathbb{R}^{d_1})$ and $\psi \in \mathcal{S}(\mathbb{R}^{d_2})$

$$U(\varphi \otimes^f \psi) = \sum_{i=1}^{n} S_i(\varphi) \cdot T_i(\psi)$$

holds. Furthermore, U is given explicitly by the formula

$$U = \sum_{i=1}^{n} S_i \otimes^D T_i \,.$$

This last proposition ensures, that $\sum_{i=1}^n S_i \otimes^D T_i$ is again a well-defined tempered distribution. Moreover, if S and T are regular distributions, generated by functions $f : \mathbb{R}^{d_1} \longrightarrow \mathbb{C}$ and $g : \mathbb{R}^{d_2} \longrightarrow \mathbb{C}$, then it can be easily seen, that also $S \otimes^D T$ is a regular distribution, generated by $f \otimes^f g : \mathbb{R}^{d_1+d_2} \longrightarrow \mathbb{C}$.

At the end we intend to apply the theory to tensor products of Sobolev and Besov spaces. This motivates a closer look on spaces of tempered distribution.
Hence, let X and Y be quasi-Banach spaces of tempered distributions. The first question to be addressed is whether their dual spaces are rich enough to provide meaningful results for tensor products.

Lemma 1.3.3. Let X be a topological vector space, such that $X \hookrightarrow \mathcal{S}'(\mathbb{R}^n)$. Then X' separates the points in X, and hence $T(X, Y) = X \otimes Y$ for every vector space Y.

Proof. We consider the natural injection $\mathcal{J} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}''(\mathbb{R}^n)$, which is defined by $(\mathcal{J}\varphi)(f) = f(\varphi)$, $\varphi \in \mathcal{S}(\mathbb{R}^n)$, $f \in \mathcal{S}'(\mathbb{R}^n)$. Due to the assumed topological embedding $X \hookrightarrow \mathcal{S}'(\mathbb{R}^n)$ we immediately find $\mathcal{J}\varphi \in X'$ for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$.
Now let $f \in X$, $f \neq 0$. Then we also have $f \neq 0$ in the sense of $\mathcal{S}'(\mathbb{R}^{d_1})$. This means there is some function $\varphi \in \mathcal{S}(\mathbb{R}^{d_1})$ such that $f(\varphi) \neq 0$. This immediately implies $(\mathcal{J}\varphi)(f) \neq 0$, which yields the desired functional from X'. □

Thus when dealing with tensor products of spaces of tempered distributions the minimal assumption is a continuous topological embedding into the space $\mathcal{S}'(\mathbb{R}^n)$. In particular, all types of Fourier analytical Besov and Triebel-Lizorkin spaces satisfy this condition.
Another interesting aspect arises from Proposition 1.3.2. Due to the uniqueness assertion the set

$$X \otimes^D Y = \left\{ \sum_{i=1}^{n} f_i \otimes^D g_i : \ f_i \in X,\ g_i \in Y,\ i = 1, \ldots, n,\ n \in \mathbb{N} \right\}$$

is a well-defined subspace of $\mathcal{S}'(\mathbb{R}^{d_1+d_2})$. Moreover, Proposition 1.3.2 motivates the following definition for all $h = \sum_{i=1}^n \lambda_i f_i \otimes g_i$ and $w = \sum_{j=1}^m \mu_j u_j \otimes v_j$ from $F(X, Y)$:

$$\sum_{i=1}^{n} \lambda_i f_i \otimes g_i \cong \sum_{j=1}^{m} \mu_j u_j \otimes v_j$$

$$\iff \sum_{i=1}^{n} \lambda_i f_i(\varphi) \cdot g_i(\psi) = \sum_{j=1}^{m} \mu_j u_j(\varphi) \cdot v_j(\psi) \quad \text{for all} \quad \varphi \in \mathcal{S}(\mathbb{R}^{d_1}), \psi \in \mathcal{S}(\mathbb{R}^{d_2}).$$

The relation $\cong$ then turns out to be an equivalence relation on $F(X,Y)$, and we have $X \otimes^D Y = F(X,Y)/\cong$. This yields another approach towards tensor product spaces which is applicable also for quasi-Banach spaces. The immediate question regarding a comparison with the algebraic tensor product $X \otimes Y$ is dealt with in the next lemma.

Lemma 1.3.4. Let X, Y be topological vector spaces of tempered distributions, such that $X \hookrightarrow \mathcal{S}'(\mathbb{R}^{d_1})$. Then $f \simeq g$ implies $f \cong g$ for all $f, g \in F(X,Y)$, i.e. we find $X \otimes^D Y \subset X \otimes Y$.
Now suppose the continuous embeddings $\mathcal{S}(\mathbb{R}^{d_1}) \hookrightarrow X \hookrightarrow \mathcal{S}'(\mathbb{R}^{d_1})$ and $\mathcal{S}(\mathbb{R}^{d_2}) \hookrightarrow Y \hookrightarrow \mathcal{S}'(\mathbb{R}^{d_2})$ to hold. Moreover, we assume $\mathcal{J}\big(\mathcal{S}(\mathbb{R}^{d_1})\big)$ to be dense in X', where $\mathcal{J}$ is the natural injection from $\mathcal{S}(\mathbb{R}^{d_1})$ into $\mathcal{S}''(\mathbb{R}^{d_1})$. Then the vector spaces $X \otimes Y$ and $X \otimes^D Y$ are isomorphic.

Proof. A proof of this lemma, though in a slightly different formulation, can be found in [74, Appendix B]. □

Remark 1.3.3. It remains an open problem to what extent the assumed density of $\mathcal{J}\big(\mathcal{S}(\mathbb{R}^{d_1})\big)$ in X' is necessary. A more natural condition seems to assume that $\mathcal{S}(\mathbb{R}^{d_1})$ is dense in X.
Moreover, for isotropic Besov spaces this assumption is not always satisfied, since $\mathcal{S}(\mathbb{R}^n)$ is dense in $\big(B^s_{p,p}(\mathbb{R}^n)\big)'$ if, and only if, $1 < p < \infty$, i.e. for $p \le 1$ we could end up with considerably smaller tensor product spaces when working with $B^{s_1}_{p,p}(\mathbb{R}^{d_1}) \otimes^D B^{s_2}_{p,p}(\mathbb{R}^{d_2})$.
The main advantage of this approach towards tensor products of distributions is the more immediate meaning of the expressions $\sum_{i=1}^n S_i \otimes^D T_i$ in comparison with the purely algebraic definition. However, since most of the abstract results for tensor products are based on the algebraic tensor product we will not further pursue this angle at this point.
We shall finally add that for the concrete case of Besov spaces the method used in Sections 1.4.5 and 4.4 provides the respective identifications automatically.

Remark 1.3.4. We add another remark on dense subsets. Assume F and G to be dense subsets of given quasi-Banach spaces X and Y, respectively. We equip $X \otimes Y$ with some crossnorm α. Then it is immediately clear, that $(F \otimes Y) + (X \otimes G)$ is a dense subset of $X \otimes Y$. On the other hand one finds at once that $F \otimes G$ is a dense subset of $(F \otimes Y) + (X \otimes G)$ with respect to the same crossnorm α. Altogether, it follows that $F \otimes G$ is dense in $X \otimes_\alpha Y$.
In particular, this argument applies to dense embeddings $\mathcal{S}(\mathbb{R}^{d_1}) \hookrightarrow X$ and $\mathcal{S}(\mathbb{R}^{d_2}) \hookrightarrow Y$ with quasi-Banach spaces of functions or distributions X and Y. Conversely, this observation will be helpful in identifying tensor product spaces $X \otimes_\alpha Y$ with known spaces Z. If the dense embedding $\mathcal{S}(\mathbb{R}^{d_1+d_2}) \hookrightarrow Z$ is known, it suffices to show that $X \otimes_\alpha Y \subset Z$ is a closed subspace with equivalent quasi-norms. The desired identity $X \otimes_\alpha Y = Z$ then follows from the density of $\mathcal{S}(\mathbb{R}^{d_1}) \otimes \mathcal{S}(\mathbb{R}^{d_2})$ in $\mathcal{S}(\mathbb{R}^{d_1+d_2})$, which in turn is a consequence of the dense topological embedding $\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathcal{S}(\mathbb{R}^n)$ and the density of $\mathcal{D}(\mathbb{R}^{d_1}) \otimes \mathcal{D}(\mathbb{R}^{d_2})$ in $\mathcal{D}(\mathbb{R}^{d_1+d_2})$, see e.g. [81] for details.

1.3.4 Tensor products of quasi-Banach spaces

Before we return to tensor product spaces, we shall recall a well-known notion for quasi-Banach spaces.

Definition 1.3.2. Let $0 < p \le 1$, and let X be a quasi-Banach space. Then X is called a *p-Banach space* and its quasi-norm *p-norm*, respectively, if

$$\| f + g \,|\, X \|^p \le \| f \,|\, X \|^p + \| g \,|\, X \|^p \qquad \text{for all} \quad f, g \in X \,.$$

It is clear, that every Banach space is a 1-Banach space and every norm is a 1-norm. Furthermore, it can be shown, that for every quasi-Banach space $(X, \|\cdot\|)$ there exists a $p \in (0,1]$ and a p-norm $\|\cdot\|^*$ on X, which is equivalent to $\|\cdot\|$, i.e. $(X, \|\cdot\|^*)$ is a p-Banach space. We refer to [25] and [62] for details and further references.

Definition 1.3.3. Let X and Y be quasi-Banach spaces, and let $0 < p \le 1$. Then we define the *p-nuclear tensor norm* $\gamma_p(\,\cdot\,, X, Y)$ as

$$\gamma_p\big([f], X, Y\big) := \inf\left\{ \left(\sum_{j=1}^{n} \|x_j|X\|^p \cdot \|y_j|Y\|^p \right)^{1/p} : f \in [f] \text{ as in (1.3.1)} \right\}.$$

These quasi-norms have been originally introduced by Grothendieck in [31]. It can be shown that the norm $\gamma_1 = \gamma$ is always equivalent to α_1 (which justifies the above notion in case $p = 1$). Moreover, it can be checked easily that the p-nuclear norms γ_p, $p \le 1$, are always p-norms.
In Section 1.3.1, the concepts of reasonable crossnorms and uniform tensor norms were introduced for Banach spaces only. Both notions can directly be extended to quasi-Banach spaces and to quasi-norms. The following lemma deals with these properties in connection with γ_p.

Lemma 1.3.5. Let X and Y be quasi-Banach spaces, such that X' separates the points in X. Then $\gamma_p(\cdot, X, Y)$, $0 < p \le 1$, is a reasonable quasi-norm on $X \otimes Y$. Furthermore, let (X_1, X_2) and (Y_1, Y_2) be two pairs of quasi-Banach spaces such that X_i' separates the points in X_i, $i = 1, 2$. Then it holds (1.3.4) for every $T_1 \in \mathcal{L}(X_1, Y_1)$, $T_2 \in \mathcal{L}(X_2, Y_2)$ and $h \in X_1 \otimes X_2$, i.e. γ_p is uniform.

Proof. We shall follow closely the corresponding proofs for $\gamma_1 \equiv \gamma$ in [48].
Due to the assumptions Lemma 1.3.1 is applicable, hence the functional analytic tensor product coincides with the algebraic one. Now let $\phi \in X'$, $\psi \in Y'$ and $z = \sum_{i=1}^n x_i \otimes y_i \in X \otimes Y$. Then we find

$$\begin{aligned} |(\phi \otimes \psi)(z)| &\le \sum_{i=1}^{n} |(\phi \otimes \psi)(x_i \otimes y_i)| = \sum_{i=1}^{n} |\phi(x_i)| \cdot |\psi(y_i)| \\ &\le \| \phi \,|\, X' \| \cdot \| \psi \,|\, Y' \| \;\sum_{i=1}^{n} \| x_i \,|\, X \| \cdot \| y_i \,|\, Y \| \end{aligned}$$

Taking the infimum over all representation of z we obtain

$$|(\phi \otimes \psi)(z)| \le \| \phi \,|\, X' \| \cdot \| \psi \,|\, Y' \| \gamma(z, X, Y) \le \| \phi \,|\, X' \| \cdot \| \psi \,|\, Y' \| \gamma_p(z, X, Y)\,,$$

where we additionally used the monotonicity of the ℓ_p-quasi-norms. If we denote by γ_p^* the induced operator quasi-norm on $(X \otimes Y, \gamma_p)'$ we have shown

$$\gamma_p^*(\phi \otimes \psi) \leq \| \phi \,|\, X' \| \cdot \| \psi \,|\, Y' \| \tag{1.3.6}$$

In particular, this implies for all $x \in X$ and $y \in Y$

$$|\phi(x)\psi(y)| = |(\phi \otimes \psi)(x \otimes y)| \leq \gamma_p^*(\phi \otimes \psi, X, Y)\gamma_p(x \otimes y, X, Y)$$

Using the obvious estimate

$$\gamma_p(x \otimes y, X, Y) \leq \|x|X\| \cdot \|y|Y\| \qquad \text{for all} \quad x \in X\,, y \in Y\,, \tag{1.3.7}$$

and taking a supremum over the unit balls of X and Y we find

$$\| \phi \,|\, X' \| \cdot \| \psi \,|\, Y' \| \leq \gamma_p^*(\phi \otimes \psi) \tag{1.3.8}$$

Combining (1.3.6) and (1.3.8) proves that γ_p^* is a crossnorm on $X' \otimes Y'$. On the other hand, this is obviously equivalent to γ_p being reasonable.
Finally, let $T_1 \in \mathcal{L}(X_1, Y_1)$, $T_2 \in \mathcal{L}(X_2, Y_2)$ and $z = \sum_{i=1}^n u_i \otimes v_i \in X_1 \otimes X_2$. Then it holds

$$\begin{aligned}\gamma_p\Bigg(\sum_{i=1}^n T_1 u_i \otimes T_2 v_i\Bigg) &\leq \Bigg(\sum_{i=1}^n \| T_1 u_i \,|\, Y_1 \|^p \cdot \| T_2 v_i \,|\, Y_2 \|^p\Bigg)^{1/p} \\ &\leq \big\|T_1\big|\mathcal{L}(X_1, Y_1)\big\| \cdot \big\|T_2\big|\mathcal{L}(X_2, Y_2)\big\| \Bigg(\sum_{i=1}^n \big\|u_i\big|X_1\big\|^p \big\|v_i\big|X_2\big\|^p\Bigg)^{1/p}\end{aligned}$$

Taking the infimum over all representations of z yields (1.3.4). □

Remark 1.3.5. We want to point out that so far we have not shown whether γ_p is a crossnorm. As mentioned in the proof, we have the one-sided estimate (1.3.7). Comparing with the proofs in [48] the reverse inequality can be shown using norming functionals, i.e. given $x \in X$ these are functionals $\varphi_x \in X'$ such that $\|\varphi_x|X'\| = 1$ and $\varphi_x(x) = \|x|X\|$. However, this well-known consequence of the Hahn-Banach theorem fails to be true for general quasi-Banach spaces.
On the other hand, explicitly assuming the property

$$\sup_{\|\phi|X'\| \leq 1} |\phi(x)| = \|x|X\| \quad \text{for all} \quad x \in X$$

turns out to be a severe restriction, because even quite simple and prominent examples of quasi-Banach spaces such as the spaces $\ell_p(I)$, $0 < p \leq 1$, and their weighted versions (for a definition see the next subsection) do not satisfy this condition. This follows from the well-known characterization $(\ell_p(I))' = \ell_\infty(I)$, which is provided by the usual isometric isomorphism. As a consequence of this characterization every functional $\varphi \in (\ell_p)'$ can be extended to a functional in $(\ell_1)'$ with equal norm, and vice versa every functional on ℓ_1 can be restricted to ℓ_p without changing the norm. In particular, this implies

$$\sup_{\|\phi|(\ell_p)'\| \leq 1} |\phi(x)| = \sup_{\|\phi|(\ell_1)'\| \leq 1} |\phi(x)| = \|x|\ell_1\| \quad \text{for all} \quad x \in \ell_p\,.$$

Since for every $x \neq \lambda e^i$, $i \in I$, $\lambda \in \mathbb{C}$, we have $\|x|\ell_p\| > \|x|\ell_1\|$ this shows the non-existence of norming functionals for such $x \in \ell_p$.
However, the mentioned one-sided estimate (1.3.7) for γ_p is sufficient for most purposes, in particular all assertions needed in our investigations remain valid under this weakened assumption.

The next lemma is the counterpart of Lemma 1.3.2 for quasi-Banach spaces, see [74, Section 5.1].

Lemma 1.3.6. Let X_1, X_2, Y_1, Y_2 be quasi-Banach spaces, which fulfil the assumptions of Lemma 1.3.5, and let $0 < p < 1$. Suppose that $T_1 \in \mathcal{L}(X_1, Y_1)$ and $T_2 \in \mathcal{L}(X_2, Y_2)$ are isomorphisms. Then also $T_1 \otimes_{\gamma_p} T_2$ is an isomorphism from $X_1 \otimes_{\gamma_p} X_2$ onto $Y_1 \otimes_{\gamma_p} Y_2$.

Remark 1.3.6. Several further attempts to deal with tensor products of quasi-Banach spaces can be found in the literature, we refer to Turpin [90] and Nitsche [55].
Instead of working with the topological dual, Nitsche stated restrictions for the algebraical dual to circumvent the failure of the Hahn-Banach theorem, which lead to the notion of *placid q-Banach spaces*. Furthermore, he introduced a version of the p-nuclear tensor norm for values $p < 1$. His main result reads as $\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) = \ell_q(\mathbb{N}^{d+1})$ with equal quasi-norms (with the interpretation of tensor products of sequences as above).
Moreover, his approach can be applied to Besov spaces $B^{1/q-1/2}_{q,q}(\mathbb{R})$ as well. The results obtained in this way correspond very well to special cases of those ones which will be presented in the sequel (see Sections 1.3.5, 1.4.5 and 4.4).

1.3.5 Tensor products of weighted sequence spaces

In this section, we shall cite some results on tensor products of weighted sequence spaces, which were proven in [74]. For this, let I be a countable index set, a sequence $w = (w_i)_{i\in I}$ of positive real numbers, and let $0 < p < \infty$. Then $\ell_p(w, I)$ consists of all sequences $a = (a_i)_{i\in I}$ of complex numbers, such that the quasi-norm

$$\big\| \, a \, \big| \ell_p(w, I) \big\| := \big\| \, a \cdot w \, \big| \ell_p(I) \big\| = \left(\sum_{i\in I} |a_i w_i|^p \right)^{1/p}$$

is finite. Moreover, according to our definitions in Section 1.3.3, the space $\ell_p(w_1 \otimes w_2, I \times J)$ is the collection of all sequences $a = (a_{i,j})_{i\in I, j\in J}$, such that

$$\big\| \, a \, \big| \ell_p(w_1 \otimes w_2, I \times J) \big\| := \left(\sum_{i\in I} \sum_{j\in J} |a_{i,j} w_{1,i} w_{2,j}|^p \right)^{1/p} < \infty \, .$$

Here J is another arbitrary countable index set, and w_1 and w_2 are two weight-sequences. Finally, we denote by $c_0(w, I)$ the closure of the set of finite sequences with respect to the norm

$$\big\| \, a \, \big| c_0(w, I) \big\| := \big\| \, a \, \big| \ell_\infty(w, I) \big\| := \sup_{i\in I} |a_i w_i| \, ,$$

i.e. $c_0(w, I)$ consists of all sequences $a = (a_i)_{i\in I}$, such that $a \cdot w = (a_i w_i)_{i\in I}$ is a null sequence.

Proposition 1.3.3. Let w_1 and w_2 be two arbitrary weight sequences.

(i) Let $1 < p < \infty$. Then it holds

$$\ell_p(w_1, \mathbb{N}) \otimes_{\alpha_p} \ell_p(w_2, \mathbb{N}) = \ell_p(w_1 \otimes w_2, \mathbb{N}^2) \, .$$

(ii) If $0 < p \leq 1$, it follows

$$\ell_p(w_1, \mathbb{N}) \otimes_{\gamma_p} \ell_p(w_2, \mathbb{N}) = \ell_p(w_1 \otimes w_2, \mathbb{N}^2) .$$

(iii) It holds

$$c_0(w_1, \mathbb{N}) \otimes_\lambda c_0(w_2, \mathbb{N}) = c_0(w_1 \otimes w_2, \mathbb{N}^2) .$$

In (i)–(iii), the identities hold with equality of (quasi-)norms.

Remark 1.3.7. As an immediate consequence of this proposition we find that γ_p is a crossnorm for these particular class of quasi-Banach spaces. This follows from the explicit meaning of the tensor product of sequences and the above mentioned equality of the quasi-norms.

We now want to apply these results to the sequence spaces $b^s_{p,p}$ from Section 1.2.2. To this purpose, we define another scale of sequence spaces first.

Definition 1.3.4. Let $r_1, \ldots, r_d \in \mathbb{R}$ and $0 < p < \infty$. Then $\ell_p^{r_1,\ldots,r_d}$ is defined as the collection of all sequences $a = (a_{\nu,k})_{\nu \in \mathbb{N}_0^d, k \in \mathbb{Z}^d}$, such that

$$\|a|\ell_p^{r_1,\ldots,r_d}\| := \left(\sum_{\nu \in \mathbb{N}_0^d} \sum_{k \in \mathbb{Z}^d} 2^{(k_1 r_1 + \cdots k_d r_d)p} 2^{|k|_1(\frac{1}{2}-\frac{1}{p})p} |a_{\nu,k}|^p \right)^{1/p}$$

is finite.

The next corollary follows from Proposition 1.3.3 for appropriate weight sequences simply by renumbering the index sets. The sequence spaces $b^{r_i}_{p,p}$ have to be understood for $d = 1$. Moreover, we shall use the notation

$$\delta_p = \begin{cases} \alpha_p, & 1 < p < \infty , \\ \gamma_p, & 0 < p \leq 1 , \end{cases}$$

to combine these two cases.

Corollary 1.3.1. Let $r_1, \ldots, r_d, r_{d+1} \in \mathbb{R}$ and let $0 < p < \infty$. Then it holds

$$b^{r_1}_{p,p} \otimes_{\delta_p} \ell_p^{r_2,\ldots,r_{d+1}} = \ell_p^{r_1,\ldots,r_d} \otimes_{\delta_p} b^{r_{d+1}}_{p,p} = \ell_p^{r_1,\ldots,r_{d+1}}$$

with coinciding (quasi-)norms.

We define $J^d := J_1 \otimes \cdots \otimes J_d$, where $J_i : B^{r_i}_{p,p} \longrightarrow b^{r_i}_{p,p}$, $i = 1, \ldots, d$ are the isomorphisms according to Theorem 1.2.2. Then from Lemmas 1.3.2 and 1.3.6 we obtain the following corollary. We shall use the temporary notation

$$T_p^{r_1,\ldots,r_d}(\mathbb{R}^d) = B^{r_1}_{p,p}(\mathbb{R}) \otimes_{\delta_p} T_p^{r_2,\ldots,r_d}(\mathbb{R}^{d-1}) = B^{r_1}_{p,p}(\mathbb{R}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B^{r_d}_{p,p}(\mathbb{R}) .$$

Corollary 1.3.2. Let $d \geq 2$, $r_1, \dots, r_d \in \mathbb{R}$, and let $0 < p < \infty$. Then J^d is an isomorphism from $T_p^{r_1,\dots,r_d}(\mathbb{R}^d)$ onto $\ell_p^{r_1,\dots,r_d}$. Moreover, the iterated tensor product $T_p^{r_1,\dots,r_d}(\mathbb{R}^d) = B_{p,p}^{r_1}(\mathbb{R}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B_{p,p}^{r_d}(\mathbb{R})$ is independent of the order of iteration.

In Section 1.4.5, we will identify the tensor product spaces $T_p^{r_1,\dots,r_d}(\mathbb{R}^d)$ as Besov spaces of dominating mixed smoothness.

1.4 Spaces of dominating mixed smoothness

In this section, we will recall the definitions and wavelet characterizations of Sobolev, Besov and Triebel-Lizorkin spaces of dominating mixed smoothness. In special cases, we can identify these spaces as tensor products of spaces of functions/distributions defined on $\mathbb{R}$.

1.4.1 Sobolev spaces of dominating mixed smoothness

As before, we begin with the definition of Sobolev spaces.

Definition 1.4.1. Let $1 < p < \infty$.

(i) For any $l \in \mathbb{N}_0^d$ we define the Sobolev spaces $S_p^l W(\mathbb{R}^d)$ by

$$S_p^l W(\mathbb{R}^d) := \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \| f \,|S_p^l W(\mathbb{R}^d)\| := \sum_{\alpha \leq l} \| D^\alpha f | L_p(\mathbb{R}^d) \| < \infty \right\},$$

where $D^\alpha f$ denotes the weak derivative of f of order $\alpha = (\alpha_1, \dots, \alpha_d)$.

(ii) For $r \in \mathbb{R}^d$ we put

$$S_p^r H(\mathbb{R}^d) := \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \| f \,|S_p^r H(\mathbb{R}^d)\| < \infty \right\},$$

where

$$\| f \,|S_p^r H(\mathbb{R}^d)\| := \| \mathcal{F}^{-1}(1+|\xi_1|^2)^{r_1/2} \cdots (1+|\xi_d|^2)^{r_d/2} \mathcal{F} f | L_p(\mathbb{R}^d) \| .$$

These spaces are Sobolev spaces of dominating mixed smoothness of fractional order, or Bessel potential spaces of dominating mixed smoothness, respectively.

Remark 1.4.1. Similar to the case of isotropic function spaces, we have by definition $S_p^0 W(\mathbb{R}^d) = S_p^0 H(\mathbb{R}^d) = L_p(\mathbb{R}^d)$. Furthermore, also for these spaces it can be shown, that for $m \in \mathbb{N}_0^d$ it holds $S_p^m W(\mathbb{R}^d) = S_p^m H(\mathbb{R}^d)$ in the sense of equivalent norms.

1.4.2 Tensor products of Sobolev spaces

In this section we concentrate on Sobolev spaces with fractional smoothness. However, the results can be transferred immediately to the spaces $S_p^m W(\mathbb{R}^d)$.
A first indication of the connection of the spaces $S_p^m H(\mathbb{R}^d)$ with tensor products of Sobolev spaces on $\mathbb{R}$ is given by the fact, that the norm in $S_p^m H(\mathbb{R}^d)$ is a crossnorm. More precisely,

let $f = f_1 \otimes^D \cdots \otimes^D f_d \in \mathcal{S}'(\mathbb{R}^d)$ with $f_i \in H_p^{r_i}(\mathbb{R})$, $i = 1, \ldots, d$, then it holds with $r = (r_1, \ldots, r_d)$

$$\left\| f_1 \otimes^D \cdots \otimes^D f_d \middle| S_p^r H(\mathbb{R}^d) \right\| = \left\| f_1 \middle| H_p^{r_1}(\mathbb{R}) \right\| \cdots \left\| f_d \middle| H_p^{r_d}(\mathbb{R}) \right\| .$$

This follows directly from the definition of the tensor product of distributions and the tensor product property of $\mathcal{F}_d$ on $\mathcal{S}(\mathbb{R}^d)$, see (1.1.2). For two distributions $S \in \mathcal{S}'(\mathbb{R}^{d_1})$ and $T \in \mathcal{S}'(\mathbb{R}^{d_2})$ and their tensor product $U = S \otimes^D T \in \mathcal{S}'(\mathbb{R}^{d_1+d_2})$ the counterpart of (1.3.5) for the respective Fourier transformed distributions is immediate:

$$\begin{aligned}\big(\mathcal{F}_{d_1+d_2} U\big)\big(\varphi \otimes^f \psi\big) &= U\big(\mathcal{F}_{d_1+d_2}(\varphi \otimes^f \psi)\big) = U\big((\mathcal{F}_{d_1}\varphi) \otimes^f (\mathcal{F}_{d_2}\psi)\big) \\ &= S(\mathcal{F}_{d_1}\varphi) \cdot T(\mathcal{F}_{d_2}\psi) = (\mathcal{F}_{d_1} S)(\varphi) \cdot (\mathcal{F}_{d_2} T)(\psi)\end{aligned}$$

for arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^{d_1})$ and $\psi \in \mathcal{S}(\mathbb{R}^{d_2})$. Hence, by Proposition 1.3.1 we have found

$$\mathcal{F}_{d_1+d_2}\big(S \otimes^D T\big) = \big(\mathcal{F}_{d_1} S\big) \otimes^D \big(\mathcal{F}_{d_2} T\big) . \tag{1.4.1}$$

That the spaces $S_p^m H(\mathbb{R}^d)$ are even tensor product spaces is the content of the next proposition.

Proposition 1.4.1. Let $d \geq 2$, $1 < p < \infty$, and let $r = (r_1, \ldots, r_d) \in \mathbb{R}^d$. Then it holds

$$\begin{aligned} S_p^r H(\mathbb{R}^d) &= H_p^{r_1}(\mathbb{R}) \otimes_{\alpha_p} S_p^{(r_2,\ldots,r_d)} H(\mathbb{R}^{d-1}) \\ &= S_p^{(r_1,\ldots,r_{d-1})} H(\mathbb{R}^{d-1}) \otimes_{\alpha_p} H_p^{r_d}(\mathbb{R}) = H_p^{r_1}(\mathbb{R}) \otimes_{\alpha_p} \cdots \otimes_{\alpha_p} H_p^{r_d}(\mathbb{R}) \end{aligned}$$

with coinciding norms.

The proof can be found in [74, Section 3.1]. In case of Hilbert spaces, the corresponding assertion, usually written as $H^s(\mathbb{R}) \otimes H^s(\mathbb{R}) = H^s_{\text{mix}}(\mathbb{R}^2)$, has been known for a long time.

1.4.3 Besov and Triebel-Lizorkin spaces of dominating mixed smoothness

We choose arbitrary systems $\varphi^i = \big(\varphi^i_j\big)_{j=0}^{\infty} \in \Phi(\mathbb{R})$, $i = 1, \ldots, d$, and put

$$\varphi_k(x) := \varphi^1_{k_1}(x_1) \cdots \varphi^d_{k_d}(x_d). \tag{1.4.2}$$

where $k \in \mathbb{N}_0^d$ and $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. The system $\varphi = (\varphi_k)_{k \in \mathbb{N}_0^d}$ can be viewed as the tensor product of the systems φ^i, $i = 1, \ldots, d$. Using (1.2.3), we find at once

$$\sum_{k \in \mathbb{N}_0^d} \varphi_k(x) = \left(\sum_{k_1=0}^{\infty} \varphi^1_{k_1}(x_1) \right) \cdots \left(\sum_{k_d=0}^{\infty} \varphi^d_{k_d}(x_d) \right) = 1 \tag{1.4.3}$$

for all $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. In this sense the system φ again is a decomposition of unity on $\mathbb{R}^d$.

Equipped with this decomposition of unity, we proceed to the definition of Besov and Triebel-Lizorkin spaces of dominating mixed smoothness.

Definition 1.4.2. Let $r \in \mathbb{R}^d$, $0 < q \le \infty$ and $\varphi = \big(\varphi_k\big)_{k\in\mathbb{N}_0^d}$ as above.

(i) Let $0 < p \le \infty$. Then $S^r_{p,q}B(\mathbb{R}^d)$ is the collection of all $f \in \mathcal{S}'(\mathbb{R}^d)$, such that

$$\big\| f \,|S^r_{p,q}B(\mathbb{R}^d)\big\|_\varphi := \Bigg(\sum_{k\in\mathbb{N}_0^d} 2^{k\cdot rq} \big\| \big(\varphi_k \widehat{f}\big)^\vee \big| L_p(\mathbb{R}^d)\big\|^q\Bigg)^{1/q}$$

is finite.

(ii) Let $0 < p < \infty$. Then we define $S^r_{p,q}F(\mathbb{R}^d)$ as the set of all $f \in \mathcal{S}'(\mathbb{R}^d)$, such that

$$\big\| f \,|S^r_{p,q}F(\mathbb{R}^d)\big\|_\varphi := \Bigg\| \Bigg(\sum_{k\in\mathbb{N}_0^d} 2^{k\cdot rq} \big| \big(\varphi_k \widehat{f}\big)^\vee(\cdot)\big|^q\Bigg)^{1/q} \Bigg| L_p(\mathbb{R}^d)\Bigg\|$$

is finite.

These quasi-norms have to be modified in the usual way, if p and/or $q = \infty$.

Remark 1.4.2. Spaces of this type together with several modifications and generalizations have been studied by many authors, in particular Amanov, Besov, Nikol'skij, Schmeißer and Triebel. For a detailed systematical treatment of these spaces and further references we refer to [1], [71] and the recent survey [69]. As for the isotropic spaces, we mention here only some basic facts.
Both scales of spaces are independent of the chosen decomposition of unity φ up to equivalence of quasi-norms, see e.g. [71, Theorem 2.2.4]. Hence we omit this index for the quasi-norms.
It obviously holds $S^r_{p,p}B(\mathbb{R}^d) = S^r_{p,p}F(\mathbb{R}^d)$. Moreover, the notation $S^r_{p,q}A(\mathbb{R}^d)$ is used in the usual way to refer to both scales of spaces at once.
These scales contain many classical spaces as special cases. The most important result in this direction is a variant of the Littlewood-Paley-Theorem for tensor product decompositions as above, stating that for $1 < p < \infty$ we have $S^r_{p,2}F(\mathbb{R}^d) = S^r_pH(\mathbb{R}^d)$ in the sense of equivalent quasi-norms.

The characterization by wavelets will be presented in the next section.

1.4.4 Characterization by wavelets

For the aspired characterization, we will use another construction of wavelet bases for $L_2(\mathbb{R}^d)$, which is different from the one used in Section 1.2.2. To this purpose, consider again the wavelet systems according to Theorem 1.2.1. We renumber the system (1.2.6) by

$$\psi_{j,m}(x) = \begin{cases} \psi_0(x-m), & j = 0, m \in \mathbb{Z}\,, \\ 2^{\frac{j-1}{2}}\psi_1(2^{j-1}t - m), & j \in \mathbb{N}, m \in \mathbb{Z}\,. \end{cases}$$

Now, let Ψ be the d-fold tensor product of the system (1.2.6), i.e.

$$\Psi := \Big\{ \psi_{\nu,k} = \psi_{\nu_1,k_1} \otimes^f \cdots \otimes^f \psi_{\nu_d,k_d} : \nu \in \mathbb{N}_0^d, k \in \mathbb{Z}^d \Big\}\,. \tag{1.4.4}$$

Then the next proposition is a consequence of well-known results in classical harmonical analysis. It provides a counterpart to Theorem 1.2.1 and Proposition 1.2.1.

Proposition 1.4.2. For every $s \in \mathbb{N}$, there are real-valued functions $\psi_0, \psi_1 \in C^s(\mathbb{R})$ with compact support and property (1.2.5), such that the system Ψ from (1.4.4) forms an orthonormal basis of $L_2(\mathbb{R}^d)$.

For the wavelet characterization of the distribution spaces, we need again certain associated sequence spaces.

Definition 1.4.3. Let $r \in \mathbb{R}^d$, $0 < p, q \leq \infty$, and let $e = (1, \ldots, 1) \in \mathbb{N}^d$.

(i) The space $s^r_{p,q}b$ is defined as the collection of all sequences

$$\lambda = \left\{\lambda_{\nu,k} \in \mathbb{C} : \nu \in \mathbb{N}^d_0, k \in \mathbb{Z}^d\right\}, \tag{1.4.5}$$

such that

$$\| \lambda \,|s^r_{p,q}b\| := \left(\sum_{\nu\in\mathbb{N}^d_0} 2^{\nu\cdot(r+\frac{1}{2}e-\frac{1}{p}e)q}\left(\sum_{k\in\mathbb{Z}^d}|\lambda_{\nu,k}|^p\right)^{p/q}\right)^{1/q} \tag{1.4.6}$$

is finite.

(ii) Let $0 < p < \infty$. Moreover, we denote by $\mathcal{X}_{\nu,k}$ the characteristic function of the rectangle

$$Q_{\nu,k} = \left[2^{-\nu_1}m_1, 2^{-\nu_1}(m_1+1)\right) \times \cdots \times \left[2^{-\nu_d}m_d, 2^{-\nu_d}(m_d+1)\right). \tag{1.4.7}$$

Then the space $s^r_{p,q}f$ is defined as the collection of all sequences as in (1.4.5), such that

$$\| \lambda \,|s^r_{p,q}f\| := \left\| \left(\sum_{\nu\in\mathbb{N}^d_0}\sum_{k\in\mathbb{Z}^d} 2^{\nu\cdot(r+\frac{1}{2}e)q}|\lambda_{\nu,k}|^q\mathcal{X}_{\nu,k}(\cdot)\right)^{1/q}\middle| L_p(\mathbb{R}^d)\right\| \tag{1.4.8}$$

is finite.

If p and/or $q = \infty$ the (quasi-)norms have to be modified in the usual way.

Now the desired wavelet characterization for the spaces $S^r_{p,q}A(\mathbb{R}^d)$ reads as follows.

Theorem 1.4.1. Let $r \in \mathbb{R}^d$ and $0 < q \leq \infty$. Moreover, we assume $\Psi \subset C^u(\mathbb{R}^d)$. Then there exists some (sufficiently large) $u \in \mathbb{N}$ such that the following statements hold:

(i) Let $0 < p \leq \infty$. Then the space $S^r_{p,q}B(\mathbb{R}^d)$ is the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^d)$, such that

$$f = \sum_{\nu\in\mathbb{N}^d_0}\sum_{k\in\mathbb{Z}^d}\lambda_{\nu,k}\psi_{\nu,k}, \qquad \text{convergence in } \mathcal{S}'(\mathbb{R}^d), \tag{1.4.9}$$

and $\lambda = \left(\lambda_{\nu,k}\right)_{\nu\in\mathbb{N}^d_0, k\in\mathbb{Z}^d} \in s^r_{p,q}b$.

(ii) Let $0 < p < \infty$. Then a distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ belongs to the space $S^r_{p,q}F(\mathbb{R}^d)$ if, and only if, it can be represented as in (1.4.9) with $\lambda = \big(\lambda_{\nu,k}\big)_{\nu\in\mathbb{N}_0^d, k\in\mathbb{Z}^d} \in s^r_{p,q}f$.

(iii) The coefficients in (1.4.9) are uniquely determined. It holds

$$\lambda_{\nu,k} = \langle f, \psi_{\nu,k}\rangle\,, \qquad \nu \in \mathbb{N}_0^d, k \in \mathbb{Z}^d\,,$$

where $\langle \cdot\,,\cdot\rangle$ denotes a dual pairing. Moreover, the mapping J, defined by

$$f \longmapsto \big(\langle f, \psi_{\nu,k}\rangle\big)_{k\in\mathbb{N}_0^d, k\in\mathbb{Z}^d}\,,$$

is an isomorphism from $S^r_{p,q}B(\mathbb{R}^d)$ onto $s^r_{p,q}b$ and from $S^r_{p,q}F(\mathbb{R}^d)$ onto $s^r_{p,q}f$.

(iv) If $\max(p, q) < \infty$, then the system Ψ in (1.4.4) is a basis in $S^r_{p,q}A(\mathbb{R}^d)$.

For a proof of this theorem as well as more details on these wavelet decompositions we refer to [94, Section 2.4]. Similar results can be found in [5].

Remark 1.4.3. The dual pairing in part (iii) has to be justified, since the wavelet system Ψ from (1.4.4) is no subset of $\mathcal{S}(\mathbb{R}^d)$. This can be done in complete analogy to the isotropic case (see Remark 1.2.5), i.e. $\langle f, \psi_{\nu,k}\rangle$ is interpreted as a dual pairing in some Besov space $S^s_{\widetilde{p},\widetilde{p}}B(\mathbb{R}^d)$ with suitably chosen parameters s and $\widetilde{p}$. For further details we refer to [94, Section 2.4.1].

1.4.5 Tensor products of Besov spaces

In this section we have a closer look at the tensor product structure of the spaces $S^r_{p,p}B(\mathbb{R}^d)$. First of all, we note that the quasi-norm in $S^r_{p,q}B(\mathbb{R}^d)$ is a crossnorm. Let $f = f_1 \otimes^D \cdots \otimes^D f_d \in \mathcal{S}'(\mathbb{R}^d)$ with $f_i \in B^{r_i}_{p,q}(\mathbb{R})$, $i = 1, \ldots, d$, then it holds with $r = (r_1, \ldots, r_d)$

$$\big\|f_1 \otimes^D \cdots \otimes^D f_d\big|S^r_{p,q}B(\mathbb{R}^d)\big\| = \big\|\, f_1\,\big|B^{r_1}_{p,q}(\mathbb{R})\big\| \cdots \big\|\, f_d\,\big|B^{r_d}_{p,q}(\mathbb{R})\big\|\,. \tag{1.4.10}$$

Moreover, we now are able to identify the spaces $S^r_{p,p}B(\mathbb{R}^d)$ as tensor product spaces.

Theorem 1.4.2. Let $d \geq 2$, $r = (r_1, \ldots, r_d) \in \mathbb{R}^d$, and let $0 < p < \infty$. Then it holds

$$S^r_{p,p}B(\mathbb{R}^d) = B^{r_1}_{p,p}(\mathbb{R}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B^{r_d}_{p,p}(\mathbb{R})$$

in the sense of equivalent quasi-norms.

This theorem follows immediately from Theorem 1.4.1, Corollary 1.3.2, the identification $s^r_{p,p}b = \ell_p^{r_1,\ldots,r_d}$ and the observation, that the isomorphism J from Theorem 1.4.1 coincides with the isomorphism J^d from Corollary 1.3.2.

2 Function spaces of dominating mixed smoothness with respect to general splittings $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}$

One of the main aims of the first part of this thesis is to generalize Theorem 1.4.2 to tensor products of spaces $B^{r_i}_{p,p}(\mathbb{R}^{d_i})$. To this purpose we introduce a generalization of the spaces $S^r_{p,q}A(\mathbb{R}^d)$ to spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$. This spaces will again be referred to as spaces of dominating mixed smoothness. We also define related scales of Sobolev spaces. Afterwards we will study the basic properties of these spaces, where we will prove modified versions of results in the monograph [71]. Subsequent chapters will be devoted to characterizations of these spaces by local means, atoms, and eventually wavelets.

2.1 Sobolev-type spaces

We begin again with the definition of Sobolev spaces of dominating mixed smoothness and their counterparts with fractional smoothness.

Definition 2.1.1. Let $1 < p < \infty$.

(i) For every $\overline{l} \in \mathbb{N}_0^N$ we define the Sobolev space $S^{\overline{l}}_pW(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = S^{\overline{l}}_pW(\mathbb{R}^{\overline{d}})$ by

$$S^{\overline{l}}_pW(\mathbb{R}^{\overline{d}}) := \Big\{ f \in L_p(\mathbb{R}^d) : \| f \,|S^{\overline{l}}_pW(\mathbb{R}^{\overline{d}})\| := \sum_{\overline{\alpha} \leq \overline{l}} \| D^\alpha f | L_p(\mathbb{R}^d)\| < \infty \Big\},$$

where as before $D^\alpha f$ denotes the weak (distributional) derivative of f.

(ii) For $\overline{r} \in \mathbb{R}^N$ we define the Sobolev spaces of fractional order (Bessel-potential spaces) $S^{\overline{r}}_pH(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})$ by

$$S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}}) := \Big\{ f \in \mathcal{S}'(\mathbb{R}^d) : \| f \,|S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})\| < \infty \Big\},$$

where

$$\| f \,|S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})\| := \| \mathcal{F}^{-1}(1+|\xi^1|^2)^{r_1/2} \cdots (1+|\xi^N|^2)^{r_N/2} \mathcal{F}f | L_p(\mathbb{R}^d)\|.$$

Remark 2.1.1. From Definition 2.1.1 it follows at once $S^{\overline{0}}_pW(\mathbb{R}^{\overline{d}}) = S^{\overline{0}}_pH(\mathbb{R}^{\overline{d}}) = L_p(\mathbb{R}^d)$. Moreover, for $N = 1$ one re-obtains the isotropic spaces $W^l_p(\mathbb{R}^d)$ and $H^r_p(\mathbb{R}^d)$ from Definition 1.2.1, while for $\overline{d} = (1, \ldots, 1)$, i.e. $d = N$, we get Sobolev spaces of dominating mixed smoothness $S^l_pW(\mathbb{R}^d)$ and $S^r_pH(\mathbb{R}^d)$ as in Definition 1.4.1.

Remark 2.1.2. Both norms, $\| \cdot |S^l_pW(\mathbb{R}^{\overline{d}})\|$ and $\| \cdot |S^r_pH(\mathbb{R}^{\overline{d}})\|$, are crossnorms. For functions $\varphi = \varphi_1 \otimes \cdots \otimes \varphi_N$ with $\varphi_i \in W^{l_i}_p(\mathbb{R}^{d_i})$, $i = 1, \ldots, N$, it holds

$$\| \varphi \,|S^{\overline{l}}_pW(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\| = \prod_{i=1}^N \| \varphi_i \,|W^{l_i}_p(\mathbb{R}^{d_i})\|$$

where $\overline{l} = (l_1, \ldots, l_N) \in \mathbb{N}_0^N$. Analogously we find

$$\| \varphi \otimes \psi \,|S^{(\overline{l},\overline{m})}_pW(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{N_1}} \times \mathbb{R}^{d_{N_1+1}} \times \cdots \times \mathbb{R}^{d_{N_1+N_2}})\|$$

$$= \big\| \varphi \,\big| S_p^{\bar{l}} W(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\big\| \cdot \big\| \psi \,\big| S_p^{\overline{m}} W(\mathbb{R}^{d_{N_1+1}} \times \cdots \times \mathbb{R}^{d_{N_1+N_2}})\big\|$$

for $\bar{l} \in \mathbb{N}_0^{N_1}$ and $\overline{m} \in \mathbb{N}_0^{N_2}$. Similar equations are true for $S_p^{\bar{r}} H(\mathbb{R}^{\bar{d}})$.

Remark 2.1.3. In Section 4.4 we will prove that indeed these Sobolev spaces are tensor product spaces. More precisely, we will prove

$$S_p^{\bar{r}} H(\mathbb{R}^{\bar{d}}) = H_p^{r_1}(\mathbb{R}^{d_1}) \otimes_{\alpha_p} \cdots \otimes_{\alpha_p} H_p^{r_N}(\mathbb{R}^{d_N}) \,, \qquad \bar{r} \in \mathbb{R}^N \,, 1 < p < \infty \,,$$

with equality of norms. In view of Theorem 2.1.1 this transfers to spaces $S_p^{\overline{m}} W(\mathbb{R}^{\bar{d}})$ and $W_p^{m_i}(\mathbb{R}^{d_i})$, $i = 1, \ldots, N$, where $\overline{m} \in \mathbb{N}_0^N$ and $1 < p < \infty$.

The connection between the function spaces $S_p^{\bar{l}} W(\mathbb{R}^{\bar{d}})$ and $S_p^{\bar{r}} H(\mathbb{R}^{\bar{d}})$ is given by the following theorem.

Theorem 2.1.1. Let $1 < p < \infty$ and $\overline{m} \in \mathbb{N}_0^N$. Then it holds

$$S_p^{\overline{m}} W(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = S_p^{\overline{m}} H(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$$

in the sense of equivalent norms.

This result is the generalization of the relations mentioned in Remark 1.2.1 and Remark 1.4.1, respectively. For its proof we need a Fourier multiplier theorem of Lizorkin [50].

Theorem 2.1.2. Let the function Φ have continuous strictly mixed derivatives of order $j \le d$, i.e. $D^\alpha \Phi$ is continuous for all $\alpha \in \{0,1\}^d$. Furthermore, let

$$\Big| \xi_{k_1} \cdots \xi_{k_j} \frac{\partial^j \Phi}{\partial \xi_{k_1} \cdots \partial \xi_{k_j}} \Big| \le A$$

for all $1 \le k_1 < \cdots < k_j \le d$ and all $0 \le j \le d$. Then Φ is a Fourier multiplier in $L_p(\mathbb{R}^d)$ for $1 < p < \infty$.

Proof of Theorem 2.1.1.

Step 1: We show $\|D^\alpha f | L_p(\mathbb{R}^d)\| \le c_\alpha \| f \,| S_p^{\overline{m}} H(\mathbb{R}^{\bar{d}})\|$.

We put $h(\xi) := \prod_{i=1}^N (1 + |\xi^i|^2)^{m_i/2}$. As $\mathcal{S}(\mathbb{R}^d)$ is a dense subset of $L_p(\mathbb{R}^d)$ for $1 < p < \infty$, it is sufficient to consider $f \in \mathcal{S}(\mathbb{R}^d)$. Let $\alpha \in \mathbb{N}_0^d$ with $\overline{\alpha} \le \overline{m}$. Then we find

$$\begin{aligned}
\big\| D^\alpha f \big| L_p(\mathbb{R}^d) \big\| &= \big\| \mathcal{F}^{-1} \mathcal{F} D^\alpha f \big| L_p(\mathbb{R}^d) \big\| = \big\| \mathcal{F}^{-1} \xi^\alpha \mathcal{F} f \big| L_p(\mathbb{R}^d) \big\| \\
&= \Big\| \mathcal{F}^{-1} \frac{\xi^\alpha}{h(\xi)} \mathcal{F} \Big(\mathcal{F}^{-1} h(\xi) \mathcal{F} f \Big) \Big| L_p(\mathbb{R}^d) \Big\| \\
&\le c \, \big\| \mathcal{F}^{-1} h(\xi) \mathcal{F} f \big| L_p(\mathbb{R}^d) \big\| = c \, \big\| f \,\big| S_p^{\overline{m}} H(\mathbb{R}^{\bar{d}}) \big\| \,.
\end{aligned}$$

Here we used Theorem 2.1.2 in order to show, that $M(\xi) := \frac{\xi^\alpha}{h(\xi)}$ is a Fourier multiplier in $L_p(\mathbb{R}^d)$. Obviously, $M \in C^\infty(\mathbb{R}^d)$, in particular, all partial derivatives are continuous. It remains to show, that

$$|D^\beta M(\xi)| \le c_\beta |\xi^\beta|^{-1} \iff |\xi^\beta D^\beta M(\xi)| \le c_\beta$$

for all $\beta \in \{0,1\}^d$. First of all, $|M(\xi)| \le 1$ due to $\overline{\alpha} \le \overline{m}$, thus M is bounded. Moreover, w.l.o.g. consider the case $j \le d_1$, all the other cases can be treated completely analogous. Then we obtain

$$\begin{aligned}
\partial_j M(\xi) &= \frac{\alpha_j\, \xi^{\alpha'} h(\xi) - \xi^\alpha \partial_j h(\xi)}{h(\xi)^2} \qquad (\alpha' = \alpha - e_j, \text{if } \alpha_j \neq 0) \\
&= \alpha_j \frac{\xi^{\alpha'}}{h(\xi)} - \frac{\xi^\alpha}{h(\xi)} \cdot \frac{\partial_j (1+|\xi^1|^2)^{m_1/2}}{(1+|\xi^1|^2)^{m_1/2}} = \alpha_j \frac{\xi^{\alpha'}}{h(\xi)} - m_1 \frac{\xi^\alpha}{h(\xi)} \cdot \frac{\xi_j}{1+|\xi^1|^2} \\
&= \frac{\xi^\alpha}{h(\xi)} \left(\alpha_j \frac{1}{\xi_j} - m_1 \frac{\xi_j}{1+|\xi^1|^2} \right) = M(\xi) \left(\alpha_j \frac{1}{\xi_j} - m_1 \frac{\xi_j}{1+|\xi^1|^2} \right),
\end{aligned}$$

and hence, because of $|M(\xi)| \le 1$,

$$|\partial_j M(\xi)| \le |M(\xi)| \left(\alpha_j \frac{1}{|\xi_j|} + m_1 \frac{|\xi_j|}{1+|\xi^1|^2} \right) \le \frac{\alpha_j}{|\xi_j|} + \frac{m_1 |\xi_j|}{1+|\xi_j|^2} \le \frac{\alpha_j + m_1}{|\xi_j|}\,.$$

In particular, we can read off

$$|\partial_j M(\xi)| \le c\, |\xi_j|^{-1} \iff |\xi_j \partial_j M(\xi)| \le c\,.$$

The desired estimate for arbitrary β now follows inductively. Assume that for every $\beta \in \mathbb{N}_0^d$, $|\beta| \le k$, the partial derivative $D^\beta M$ can be written in the following form

$$D^\beta M(\xi) = \frac{\xi^\alpha}{h(\xi)} r_\beta(\xi)\,,$$

where $r_\beta(\xi)$ is a rational function such that $\xi^\beta r_\beta(\xi)$ can be written as a linear combination of products of (up to $|\beta|$) factors $\frac{\xi_j^2}{1+|\xi^{i(j)}|^2}$, where $i(j) \in \{1,\dots,N\}$ is the uniquely determined index, such that $d_1 + \dots + d_{i-1} + 1 \le j \le d_1 + \dots + d_i$. In particular, this at once yields $|\xi^\beta r_\beta(\xi)| \le c_\beta$. Above, this has been shown for $|\beta| = 1$. Then we obtain (let again w.l.o.g. $j \le d_1$)

$$\partial_j D^\beta M = \partial_j \left(\frac{\xi^\alpha}{h(\xi)} r_\beta(\xi) \right) = \frac{\xi^\alpha}{h(\xi)} \left(\alpha_j \frac{1}{\xi_j} - m_1 \frac{\xi_j}{1+|\xi^1|^2} \right) r_\beta(\xi) + \frac{\xi^\alpha}{h(\xi)} \partial_j r_\beta(\xi).$$

The term in brackets can be estimated by $c\,|\xi_j|^{-1}$ as before. Moreover, one can verify easily, that $\partial_j r_\beta$ is again a rational function and that $\xi_j \xi^\beta \partial_j r_\beta$ is of the described form as well (one only has to consider derivatives of the "elementary factors"). Hence we conclude $|\xi_j \xi^\beta \partial_j r(\xi)| \le c'$. Altogether, we get

$$\partial_j D^\beta M = \frac{\xi^\alpha}{h(\xi)} \widetilde{r}_\beta(\xi)\,, \quad \text{where} \quad \widetilde{r}_\beta(\xi) = \alpha_j \frac{1}{\xi_j} - m_1 \frac{\xi_j}{1+|\xi^1|^2} r_\beta(\xi) + \partial_j r_\beta(\xi)\,,$$

as well as the estimate $|\xi_j \xi^\beta \widetilde{r}_\beta(\xi)| \le c'$. This shows the induction step and thus the desired estimate $|\xi^\beta D^\beta M(\xi)| \le C_\beta$. Because only finitely many derivatives are relevant for Theorem 2.1.2, the functions $\xi^\beta D^\beta M$ are uniformly bounded for all $\beta \in \{0,1\}^d$, and the assumptions of that theorem are fulfilled.

This completes the proof of the estimate announced at the beginning. Together with the definition of the norm on $S_p^{\overline{m}} W(\mathbb{R}^{\overline{d}})$ we now conclude $\| f \,|S_p^{\overline{m}} W(\mathbb{R}^{\overline{d}})\| \le c \,\| f \,|S_p^{\overline{m}} H(\mathbb{R}^{\overline{d}})\|$.

Step 2: We show the reverse estimate $\| f \,|S_p^{\overline{m}}H(\mathbb{R}^{\overline{d}})\| \le c \,\| f \,|S_p^{\overline{m}}W(\mathbb{R}^{\overline{d}})\|$.

Let $\rho \in C^\infty(\mathbb{R})$ be a monotone function with $\rho(0) = 0$ and $\rho(t) = \mathrm{sgn}(t)$ for $|t| \ge 1$. Then the function $\xi^\alpha \rho(\xi^\alpha) \in C^\infty(\mathbb{R}^d)$ is non-negative on the whole of $\mathbb{R}^d$. The derivatives of $\rho(\xi^\alpha)$ can be written in the form $\rho_\beta(\xi^\alpha)\xi^{-\beta}$, where the functions ρ_β are linear combinations of terms $\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^k$, $k = 0, \dots, |\beta|$. This can be seen inductively with the help of

$$\partial_j \big[\rho(\xi^\alpha)\big] = \alpha_j \rho'(\xi^\alpha)\xi^{\alpha - e_j}\,, \qquad \partial_j \xi^{-\beta} = -\beta_j \xi^{-(\beta + e_j)}\,, \qquad \alpha_j, \beta_j \ge 0\,,$$

and

$$\begin{aligned}\partial_j \big[\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^k\big] &= \rho^{(k+1)}(\xi^\alpha)(\xi^\alpha)^k \alpha_j \xi^{\alpha - e_j} + \rho^{(k)}(\xi^\alpha) k (\xi^\alpha)^{k-1} \alpha_j \xi^{\alpha - e_j} \\ &= \alpha_j \rho^{(k+1)}(\xi^\alpha)(\xi^\alpha)^{k+1}\xi^{-e_j} + k\alpha_j \rho^{(k)}(\xi^\alpha)(\xi^\alpha)^k \xi^{-e_j}\,.\end{aligned}$$

Moreover, $\rho^{(k)}(\xi^\alpha)$ vanishes for $|\xi^\alpha| > 1$ for all α and all $k > 0$ because of $\rho^{(k)}(t) = 0$ for all $|t| > 1$. For all $|\xi^\alpha| \le 1$ all (finitely many) terms $\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^k$, $k = 0, \dots, |\beta|$, are bounded due to continuity. Hence, by Lizorkin's multiplier theorem it follows, that $\rho(\xi^\alpha)$ is a Fourier multiplier for $L_p(\mathbb{R}^d)$ for every $\alpha \in \mathbb{N}_0^d$.
Now, define the function $M \in C^\infty(\mathbb{R}^d)$ by

$$M(\xi) := \frac{h(\xi)}{g(\xi)} = \frac{h(\xi)}{1 + \sum_{\overline{\alpha} \le \overline{m}} \rho(\xi^\alpha)\xi^\alpha}\,. \tag{2.1.1}$$

From $h^2(\xi) \sim 1 + \sum_{\overline{\alpha} \le \overline{m}} |\xi^\alpha|^2 \sim \big(1 + \sum_{\overline{\alpha} \le \overline{m}} |\xi^\alpha|\big)^2$ we conclude the boundedness of M, see also Remark 2.1.4. Moreover, M is non-negative. For any multiindex $\beta \in \{0,1\}^d$ it holds $D^\beta M(\xi) = \sum_{\gamma \le \beta} c_{\beta,\gamma} D^\gamma h(\xi) D^{\beta - \gamma} \frac{1}{g(\xi)}$. The derivatives $D^\gamma h$ can be calculated as before, and these can be written as a product of $h(\xi)$ with a linear combination of products of $|\gamma|$ terms $\frac{\xi_j}{1 + |\xi^{i(j)}|^2}$ with $\gamma_j = 1$. Thereby index $i(j) \in \{1, \dots, N\}$ is defined as in Step 1. Hence, we find $|\xi^\gamma D^\gamma h(\xi)| \le c_\gamma\, h(\xi)$.
The derivatives $D^{\beta-\gamma}\frac{1}{g(\xi)}$ can likewise be written as products of $\xi^{-(\beta-\gamma)} g(\xi)^{-1}$ with linear combinations of products of terms $\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^{k+1} g(\xi)^{-1}$, $k = 0, \dots, |\beta - \gamma|$. This follows inductively from

$$\partial_j \frac{1}{g(\xi)} = -\frac{\sum_{\overline{\alpha} \le \overline{m}} \big(\rho'(\xi^\alpha)\xi^\alpha - \rho(\xi^\alpha)\big)\xi^{\alpha - e_j}\alpha_j}{g^2(\xi)}\,,$$

and

$$\begin{aligned}\partial_j \frac{\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^{k+1}}{g(\xi)} &= \frac{\rho^{(k+1)}(\xi^\alpha)(\xi^\alpha)^{k+1} + (k+1)\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^k}{g(\xi)} \alpha_j \xi^{\alpha - e_j} \\ &\quad - \frac{\rho^{(k)}(\xi^\alpha)(\xi^\alpha)^{k+1}}{g(\xi)} \cdot \frac{\sum_{\overline{\delta} \le \overline{m}} \big(\rho'(\xi^\delta)\xi^\delta - \rho(\xi^\delta)\big)\delta_j \xi^{\delta - e_j}}{g(\xi)}\,.\end{aligned}$$

Altogether, this yields $|\xi^{\beta-\gamma} D^{\beta-\gamma}\frac{1}{g(\xi)}| \le c\,\frac{1}{g(\xi)}$, since $\rho^{(k)}$ has compact support for all $k > 0$ and $\rho(\xi^\alpha)\xi^\alpha \le g(\xi)$.
Eventually, we obtain for $1 \le j_1 < \dots < j_n \le d$

$$\big|\xi_{j_i} \cdots \xi_{j_n} \partial_{j_1} \cdots \partial_{j_n} M(\xi)\big| = \big|\xi^\beta D^\beta M(\xi)\big| \le c\, M(\xi) \le c'\,.$$

From this, it follows by Theorem 2.1.2, that M is a Fourier multiplier as well. Altogether, we find

$$\begin{aligned}
\big\| f \,|S^{\overline{m}}_p H(\mathbb{R}^{\overline{d}})\big\| &= \big\|\mathcal{F}^{-1}h(\xi)\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\| \\
&= \Big\|\mathcal{F}^{-1}\frac{h(\xi)}{1+\sum_{\overline{\alpha}\le\overline{m}}\rho(\xi^\alpha)\xi^\alpha}\mathcal{F}\mathcal{F}^{-1}\Big(1+\sum_{\overline{\alpha}\le\overline{m}}\rho(\xi^\alpha)\xi^\alpha\Big)\mathcal{F}f\Big|L_p(\mathbb{R}^d)\Big\| \\
&\le c\,\Big\|\mathcal{F}^{-1}\Big(1+\sum_{\overline{\alpha}\le\overline{m}}\rho(\xi^\alpha)\xi^\alpha\Big)\mathcal{F}f\Big|L_p(\mathbb{R}^d)\Big\| \\
&\le c\,\Big(\big\| f \,|L_p(\mathbb{R}^d)\big\| + \sum_{\overline{\alpha}\le\overline{m}}\big\|\mathcal{F}^{-1}\rho(\xi^\alpha)\mathcal{F}\big(\mathcal{F}^{-1}\xi^\alpha\mathcal{F}f\big)\,|L_p(\mathbb{R}^d)\big\|\Big) \\
&\le c'\,\Big(\big\| f \,|L_p(\mathbb{R}^d)\big\| + \sum_{\overline{\alpha}\le\overline{m}}\big\|\mathcal{F}^{-1}\xi^\alpha\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|\Big) \\
&= c'\,\Big(\big\| f \,|L_p(\mathbb{R}^d)\big\| + \sum_{\overline{\alpha}\le\overline{m}}\big\|D^\alpha f\big|L_p(\mathbb{R}^d)\big\|\Big) \le 2c'\,\big\|f\big|S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})\big\|.
\end{aligned}$$

The proof is complete. □

Remark 2.1.4. From the second step of the proof it can be seen, that with the same arguments we can show the equivalence of norms of the form

$$\big\| f \,|S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})\big\|_A := \big\| f \,|L_p(\mathbb{R}^d)\big\| + \sum_{\alpha\in A}\big\|D^\alpha f\big|L_p(\mathbb{R}^d)\big\|\,,$$

where $A \subset A_{\overline{m}} := \big\{\alpha\in\mathbb{N}_0^d : \overline{\alpha}\le\overline{m}\big\}$, to the norm $\big\|f\big|S^{\overline{m}}_p H(\mathbb{R}^{\overline{d}})\big\|$. The only fact needed is the boundedness of the function M_A, which is a modification of the Fourier multiplier from (2.1.1),

$$M_A(\xi) := \frac{h(\xi)}{g_A(\xi)} = \frac{h(\xi)}{1+\sum_{\alpha\in A}\rho(\xi^\alpha)\xi^\alpha}\,.$$

In the next corollary we apply this consideration to derive further equivalent norms on $S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})$.

Corollary 2.1.1. Let $1<p<\infty$ and $\overline{m}\in\mathbb{N}_0^N$. Then the following functionals define equivalent norms on $S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})$:

$$\begin{aligned}
\big\| f \,|S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})\big\|^{\clubsuit} &:= \sum_{I\subset\{1,\dots,N\}}\;\sum_{\substack{\alpha\in\mathbb{N}_0^d:\\ |\alpha^i|=m_i,\, i\in I}}\big\|D^\alpha f\big|L_p(\mathbb{R}^d)\big\|\,,\\
\big\| f \,|S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})\big\|^{\spadesuit} &:= \sum_{\substack{I\subset\{1,\dots,N\},\\ I=\{i_1,\dots,i_n\}}}\;\sum_{\substack{j_1,\dots,j_n:\\ i_k=i(j_k),\,k=1,\dots,n}}\big\|D^{\alpha(j_1,\dots,j_n)} f\big|L_p(\mathbb{R}^d)\big\|\,,
\end{aligned}$$

where $i(j)$ is defined as in the second step of the proof of Theorem 2.1.1, and moreover we have $\alpha(j_1,\dots,j_n) = m_{i_1}e_{j_1}+\dots+m_{i_n}e_{j_n}$.

Proof. Let $B := \{\xi \in \mathbb{R}^d : |\xi^i| \le \sqrt{d_i}, i = 1, \dots, N\}$. As B is compact, the continuous function $M_\spadesuit = M_{A(\spadesuit)}$ is bounded on B, where $M_{A(\spadesuit)}$ is defined as in Remark 2.1.4, and the set $A(\spadesuit)$ corresponds to the derivatives used in the definition of $\|\cdot|S_p^{\overline{m}}W(\mathbb{R}^{\overline{d}})\|^\spadesuit$. For every index $i \in \{1, \dots, N\}$ such that $|\xi^i| > \sqrt{d_i}$ we define the index $k(i)$ by

$$|\xi_{k(i)}| = |\xi^i|_\infty = \max_{d_1+\cdots+d_{i-1}<k\le d_1+\cdots+d_i} |\xi_k| > 1\,.$$

For every $\xi \notin B$ such an index i exists. Now define the multiindex $\beta = \beta(\xi)$ by $\beta = \sum_{i:|\xi^i|>\sqrt{d_i}} m_i e_{k(i)}$. Then it holds $|\xi^\beta| > 1$ by choice of $k(i)$. Furthermore, as $|\xi^i|_\infty \sim |\xi^i| \equiv |\xi^i|_2$ and $|\beta^i| = m_i$, we find $1 + |\xi^{\beta(\xi)}| \sim h(\xi)$, where the constants can be chosen independent of $\xi \notin B$. Since $\beta(\xi) \in A(\spadesuit)$ this implies the boundedness of $M_\spadesuit$ on $\mathbb{R}^d \setminus B$. To complete the proof, we add the obvious estimate $M \le M_\clubsuit \le M_\spadesuit$. $\square$

Remark 2.1.5. Both norms in the above corollary are crossnorms. More precisely, these are the tensorized versions of the following well-known norms on the isotropic spaces $W_p^m(\mathbb{R}^n)$:

$$\|\,g\,|W_p^m(\mathbb{R}^n)\|^\clubsuit := \|\,g\,|L_p(\mathbb{R}^n)\| + \sum_{\beta\in\mathbb{N}_0^n:|\beta|=m} \|D^\beta g|L_p(\mathbb{R}^n)\|\,,$$

$$\|\,g\,|W_p^m(\mathbb{R}^n)\|^\spadesuit := \|\,g\,|L_p(\mathbb{R}^n)\| + \sum_{j=1}^n \|\partial_j^m g|L_p(\mathbb{R}^d)\|\,.$$

A connection of these spaces to the spaces of dominating mixed smoothness of Besov and Triebel-Lizorkin type will be discussed in Section 2.3.7.

2.2 Besov- and Triebel-Lizorkin-type spaces

We now proceed to the definition of the main objects of our investigations, compare with the corresponding definitions in the Sections 1.2 and 1.4.3.
For $i = 1, \dots, N$ we choose systems $\varphi^i = \big(\varphi^i_j\big)_{j=0}^\infty \in \Phi(\mathbb{R}^{d_i})$, and put for $\overline{k} \in \mathbb{N}_0^N$ and $x = (x^1, \dots, x^N) \in \mathbb{R}^d$

$$\varphi_{\overline{k}}(x) := \varphi^1_{k_1}(x^1)\cdots\varphi^N_{k_N}(x^N). \tag{2.2.1}$$

As in Section 1.4.3 the system $\varphi = (\varphi_{\overline{k}})_{\overline{k}\in\mathbb{N}_0^N}$ can be interpreted as the tensor product of the decompositions φ^i, $i = 1, \dots, N$. Similar to equation (1.4.3), we have

$$\sum_{k\in\mathbb{N}_0^N} \varphi_{\overline{k}}(x) = \left(\sum_{k_1=0}^\infty \varphi^1_{k_1}(x^1)\right)\cdots\left(\sum_{k_N=0}^\infty \varphi^N_{k_N}(x^N)\right) = 1 \tag{2.2.2}$$

for all $x = (x^1, \dots, x^N) = (x_1, \dots, x_d) \in \mathbb{R}^d$. In this sense the system φ again is a decomposition of unity on $\mathbb{R}^d$ with respect to the splitting $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}$.

Equipped with this notation, we can introduce the function spaces of dominating mixed smoothness $S^{\overline{r}}_{p,q}B(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$ and $S^{\overline{r}}_{p,q}F(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$.

Definition 2.2.1. Let $\overline{r} \in \mathbb{R}^N$, $0 < q \leq \infty$, and let $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be defined as above.

(i) Let $0 < p \leq \infty$. Then we define $S^{\overline{r}}_{p,q}B(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$ to be the collection of all $f \in \mathcal{S}'(\mathbb{R}^d)$, such that

$$\big\| f \,\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\big\|_{\varphi} := \left(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \big\|\big(\varphi_{\overline{k}}\widehat{f}\big)^{\vee}\big|L_p(\mathbb{R}^d)\big\|^q\right)^{1/q}$$

is finite.

(ii) Let $0 < p < \infty$. Then $S^{\overline{r}}_{p,q}F(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$ is the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^d)$, such that

$$\big\| f \,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\big\|_{\varphi} := \left\|\left(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \big|\big(\varphi_{\overline{k}}\widehat{f}\big)^{\vee}(\cdot)\big|^q\right)^{1/q}\middle| L_p(\mathbb{R}^d)\right\|$$

is finite.

Remark 2.2.1. With the notations introduced in (1.1.3) and (1.1.4), the above quasi-norms can be reformulated as

$$\big\| f \,\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\big\|_{\varphi} = \big\|2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\big(\varphi_{\overline{k}}\mathcal{F}f\big)\big|\ell_q(L_p)\big\|$$

and

$$\big\| f \,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})\big\|_{\varphi} = \big\|2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\big(\varphi_{\overline{k}}\mathcal{F}f\big)\big|L_p(\ell_q)\big\|$$

Similar formulations can be obtained for the spaces $A^s_{p,q}(\mathbb{R}^n)$ and $S^r_{p,q}A(\mathbb{R}^d)$.

Remark 2.2.2. As before, we will use the notations $S^{\overline{r}}_{p,q}A(\mathbb{R}^{d_1}\times\cdots\times\mathbb{R}^{d_N})$ and $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ to refer to both scales of spaces.
As mentioned before, the introduced scales of function spaces generalize the classical case of isotropic Besov and Triebel-Lizorkin spaces $A^s_{p,q}(\mathbb{R}^d)$, which corresponds to the case $N = 1$, as well as the Besov and Triebel-Lizorkin spaces of dominating mixed smoothness $S^{\overline{r}}_{p,q}A(\mathbb{R}^d)$, which can be obtained for $d_1 = \cdots = d_N = 1$, or $d = N$ respectively. Hence, most results of the next sections are natural generalizations of results for these well-known spaces.

Remark 2.2.3. Spaces of this type had been introduced before, and were discussed primarily by several authors from the former Soviet Union including Amanov, Nikol'skij, Lizorkin and Besov. We refer to [51] and [8] and the references given there for an overview of their work. Moreover, we shall add that Bazarkhanov dealt with characterizations by differences and related topics for periodic and non-periodic spaces, see [3, 4].
In the Fourier analytic framework spaces of this type have been investigated before by Schmeißer in the 1980s, most often in connection with integral operators, whose kernels were assumed to be elements of vector-valued Besov or Triebel-Lizorkin spaces, taking values in another Besov or Triebel-Lizorkin space. In that framework some further generalizations for mixed (quasi-)norms were discusssed. We refer to [68, 69] and the references given there.

Remark 2.2.4. As for the Sobolev-type spaces in Remark 2.1.2 the quasi-norms of the Besov and Triebel-Lizorkin spaces are crossnorms. Due to the tensor product properties of the Fourier transform (equations (1.1.1) and (1.4.1)) and the tensor product structure of the decompositions φ we find immediately

$$\begin{aligned}&\big\| f \otimes g \,\big|S^{(\overline{r},\overline{s})}_{p,q}A(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{N_1}} \times \mathbb{R}^{d_{N_1+1}} \times \cdots \times \mathbb{R}^{d_{N_1+N_2}})\big\| \\ &\qquad = \big\| f \,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{N_1}})\big\| \cdot \big\| g \,\big|S^{\overline{s}}_{p,q}A(\mathbb{R}^{d_{N_1+1}} \times \cdots \times \mathbb{R}^{d_{N_1+N_2}})\big\|\end{aligned}$$

for $\overline{r} \in \mathbb{R}^{N_1}$, $\overline{s} \in \mathbb{R}^{N_2}$. This is the direct generalization of equation (1.4.10).

Remark 2.2.5. We will not give an exhaustive treatment of the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. Instead we firstly aim at a wavelet characterization of these spaces, and secondly we will identify certain Besov spaces as tensor products. More precisely, we will show in Section 4.4

$$S^{\overline{r}}_{p,p}B(\mathbb{R}^{\overline{d}}) = B^{r_1}_{p,p}(\mathbb{R}^{d_1}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B^{r_N}_{p,p}(\mathbb{R}^{d_N}) , \qquad \overline{r} \in \mathbb{R}^N , 0 < p < \infty ,$$

in the sense of equivalent quasi-norms. This will be the counterpart of Theorem 1.4.2.

2.3 Basic facts and inequalities

The results of this section and their proofs are based on the approach in [71]. Our theorems and proofs are direct generalizations of those ones presented there for the case $N = 2$, $d_1 = d_2 = 1$.

A basic question in the Fourier analytical approach to Besov- and Triebel-Lizorkin-type spaces is the independency of the definition of the decomposition of unity, i.e. in our case, whether Definition 2.2.1 depends on the system $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$. The answer is given by the following theorem, its counterparts may be found in [71, Proposition 2.2.3/1] and [83, Theorem 2.3.3].

Theorem 2.3.1. Let $\big(\varphi^i_j\big)_{j=0}^{\infty}, \big(\psi^i_j\big)_{j=0}^{\infty} \in \Phi(\mathbb{R}^{d_i})$, $i = 1, \ldots, N$, and define systems $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ and $\psi = \big(\psi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ according to equation (2.2.1). Furthermore, let $\overline{r} \in \mathbb{R}^N$ and $0 < p, q \leq \infty$ ($p < \infty$ in the F-case).
Then $\big\| \cdot \,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|_{\varphi}$ and $\big\| \cdot \,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|_{\psi}$ are equivalent quasi-norms.

Remark 2.3.1. As a consequence of this theorem, we may omit the index φ or ψ in $\big\| \cdot \,\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|$ or $\big\| \cdot \,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|$, referring to one of these equivalent quasi-norms.

Before we proceed to the proof of Theorem 2.3.1 we need some auxiliary results. In the next sections, we will deal with some maximal inequalities, and then prove a multiplier-theorem for L_p-spaces of analytic functions as well as its vector-valued analogon.

2.3.1 Maximal operators

Maximal operators (and their boundedness on appropriate function spaces) play an important role in harmonic analysis and the theory of function spaces. Our constructions in later sections will make use of the Hardy-Littlewood maximal operator, and are based essentially on the maximal operator of Peetre. The definition and some boundedness results for the former one are the subject of this section, for the latter one we refer to Section 3.2.

For any measurable and locally integrable function $f \in L_1^{\mathrm{loc}}(\mathbb{R}^d)$ the classical Hardy-Littlewood maximal operator is defined by

$$\big(Mf\big)(x) = \sup_Q \frac{1}{|Q|} \int_Q |f(y)| dy, \qquad x \in \mathbb{R}^d,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^d$, which are centred in x and whose sides are parallel to the coordinate axes. The symbol $|Q|$ denotes the Lebesgue measure of the cube Q. The famous Hardy-Littlewood maximal inequality now states, that for every p with $1 < p \leq \infty$ exists a positive constant c_p, such that

$$\big\| Mf \big| L_p(\mathbb{R}^d) \big\| \leq c \big\| f \big| L_p(\mathbb{R}^d) \big\|, \qquad f \in L_p(\mathbb{R}^d) . \tag{2.3.1}$$

A vector-valued generalization of this assertion goes back to C. Fefferman and E. M. Stein [27].
In our considerations tensor product constructions (will) play an important role, as it could be seen for the decompositions of unity $\varphi = \big(\varphi_{\overline{k}}\big)$ in Definition 2.2.1. In order to take this structure into account, we regard the following "directional" maximal operators. We define for $i = 1, \ldots, N$

$$\begin{aligned} \big(M_i f\big)(x) &= \Big(M\big[f(x^1, \ldots, x^{i-1}, \cdot, x^{i+1}, \ldots, x^N)\big]\Big)(x^i) \\ &= \sup_Q \frac{1}{|Q|} \int_Q \big| f(x^1, \ldots, x^{i-1}, y, x^{i+1}, \ldots, x^N) \big| \, dy, \end{aligned} \tag{2.3.2}$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^{d_i}$, centred in x^i and with sides parallel to the coordinate axes. The composition of all these operators will be denoted by $\overline{M} = M_N \circ \cdots \circ M_1$. The following maximal inequality goes back to a result of R. J. Bagby [2] (indeed, this result is only a special case of much more general result presented in this reference).

Proposition 2.3.1. Let $1 < p < \infty$ and $1 < q \leq \infty$. Then there exists a positive constant c, depending only on p, q, d_i, and N, such that

$$\big\| \big(M_i f_{\overline{k}}\big)_{\overline{k} \in \mathbb{N}_0^N} \big| L_p(\ell_q) \big\| \leq c \big\| \big(f_{\overline{k}}\big)_{\overline{k} \in \mathbb{N}_0^N} \big| L_p(\ell_q) \big\|, \qquad i = 1, \ldots, N, \tag{2.3.3}$$

for all sequences $\big(f_{\overline{k}}\big)_{\overline{k} \in \mathbb{N}_0^N} \in L_p(\ell_q)$ of measurable functions on $\mathbb{R}^d$.

Iterated applications of this proposition show, that an analogon of (2.3.3) holds true for the operator $\overline{M}$ as well. In addition to this proposition we shall need another corollary of Bagby's result. That one will be concerned with double-indexed sequences of functions. We introduce the notation

$$\Big\| \big(\lambda_{k,m}\big)_{k \in A, m \in B} \Big| \ell_q(\ell_v) \Big\| := \Bigg(\sum_{k \in A} \bigg(\sum_{m \in B} |\lambda_{k,m}|^v \bigg)^{q/v} \Bigg)^{1/q}$$

for sequences $\lambda = \big(\lambda_{k,m}\big)_{k\in A, m\in B}$ of complex numbers and arbitrary countable index sets A and B. This is a special case of so-called mixed norms, which are treated, e.g., in [71]. The notation $L_p\big(\ell_q(\ell_v)\big)$ is to be understood accordingly as in (1.1.4).

Proposition 2.3.2. Let $1 < p < \infty$, $1 < q < \infty$, and $1 < v \le \infty$. Then there exists a positive constant c, independent of f, such that

$$\big\|\overline{M} f_{\overline{k},m} \big| L_p\big(\ell_q(\ell_v)\big)\big\| \le c\,\big\| f_{\overline{k},m} \big| L_p\big(\ell_q(\ell_v)\big)\big\| \tag{2.3.4}$$

for all sequences $\big(f_{\overline{k},m}\big)_{\overline{k}\in\mathbb{N}_0^N, m\in\mathbb{Z}^d} \subset L_p(\mathbb{R}^d)$ of measurable functions on $\mathbb{R}^d$.

Remark 2.3.2. Both, Propositions 2.3.1 and 2.3.2, were originally formulated in [27] and [2], respectively, for finite parameters only. However, the assertion for $L_p(\ell_\infty)$ follows immediately from the scalar case using the monotonicity of $\overline{M}$ (i.e. for functions $f \le g$ it holds $\overline{M} f \le \overline{M} g$). Similarly, the result for $L_p(\ell_q(\ell_\infty))$ follows from the usual vector-valued one.

Finally, we shall cite a lemma connecting the Hardy-Littlewood maximal operator to convolution of functions.

Lemma 2.3.1. Let f be a non-negative measurable function. Moreover, let $\varphi \in L_1(\mathbb{R}^n)$ be a function of the form $\varphi(t) = \psi(|t|)$ for some non-negative, non-increasing function ψ on $[0,\infty)$. Then it holds

$$\big(f * \varphi\big)(x) \le c\,\big(Mf\big)(x) \big\|\,\varphi\,\big|L_1(\mathbb{R}^n)\big\|\,, \quad x \in \mathbb{R}^n\,,$$

where the constant $c > 0$ depends only on n.

This result can be found, e.g., in [75], Chapter 2 (3.9). In case the function φ is given as the tensor product $\varphi = \varphi^1 \otimes \cdots \otimes \varphi^N$ of radially symmetric functions $\varphi^i \in L_1(\mathbb{R}^{d_i})$, $i = 1,\ldots,N$, as in Lemma 2.3.1, then an analogous statement holds true for $\overline{M}$.

2.3.2 Fourier multipliers

In this section we will simply write $\mathcal{F}^{-1} m \mathcal{F} f$ instead of $\mathcal{F}^{-1}\big[m(\mathcal{F}f)\big]$, if there is no danger of confusion.

Lemma 2.3.2. Let $\overline{\rho} \in \mathbb{R}^N_+$, and let $0 < p \le 2$. Moreover, let $\overline{r} \in \mathbb{R}^N$ with

$$r_i > \kappa_i := \rho_i + d_i\left(\frac{1}{p} - \frac{1}{2}\right), \quad i = 1,\ldots,N\,. \tag{2.3.5}$$

Then there exists a positive constant c, such that

$$\big\|(1+|x^1|^2)^{\rho_1/2}\cdots(1+|x^N|^2)^{\rho_N/2}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\| \le c\,\big\|f\big|S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})\big\|$$

holds for all $f \in S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})$.

Proof. Define for $i = 1, \ldots, N$

$$Q_0^i := \left\{ t \in \mathbb{R}^{d_i} : |t| \le 1 \right\},$$
$$Q_{k_i}^i := \left\{ t \in \mathbb{R}^{d_i} : 2^{k_i - 1} \le |t| \le 2^{k_i} \right\}, \quad k_i \in \mathbb{N},$$

and let $\mathcal{X}_{k_i}^i$ be the characteristic functions of $Q_{k_i}^i$. Then we have $1 + |x^i|^2 \sim 2^{2k_i}$ for all $x^i \in Q_{k_i}^i$. Hence, we obtain

$$\begin{aligned}
&\left\| (1+|x^1|^2)^{\rho_1/2} \cdots (1+|x^N|^2)^{\rho_N/2} \mathcal{F} f \middle| L_p(\mathbb{R}^d) \right\| \\
&= \left(\sum_{k_N=0}^{\infty} \int_{Q_{k_N}^N} \cdots \sum_{k_1=0}^{\infty} \int_{Q_{k_1}^1} (1+|x^1|^2)^{\rho_1 p/2} \cdots (1+|x^N|^2)^{\rho_N p/2} |\mathcal{F} f(x)|^p dx^1 \cdots dx^N \right)^{1/p} \\
&\le c \left(\sum_{k_N=0}^{\infty} 2^{k_N \rho_N p} \int_{Q_{k_N}^N} \cdots \sum_{k_1=0}^{\infty} 2^{k_1 \rho_1 p} \int_{Q_{k_1}^1} |\mathcal{F} f(x)|^p dx^1 \cdots dx^N \right)^{1/p}.
\end{aligned}$$

An application of Hölder's inequality for integrals with respect to $\frac{p}{2} + \frac{2-p}{2} = 1$ yields

$$\begin{aligned}
2^{k_1 \rho_1 p} \int_{Q_{k_1}^1} |\mathcal{F} f(x)|^p dx^1 &\le 2^{k_1 \rho_1 p} |Q_{k_1}^1|^{(2-p)/2} \left(\int_{Q_{k_1}^1} |\mathcal{F} f(x)|^{p \cdot \frac{2}{p}} dx^1 \right)^{p/2} \\
&= c\, 2^{k_1 \rho_1 p} \left(2^{k_1 d_1} - 2^{(k_1 - 1) d_1} \right)^{1-p/2} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^p \\
&\le c'\, 2^{k_1 \rho_1 p} 2^{k_1 d_1 (1-p/2)} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^p \\
&= c'\, 2^{k_1 (\rho_1 + d_1 (1/p - 1/2)) p} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^p \\
&= c'\, 2^{k_1 \kappa_1 p} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^p .
\end{aligned}$$

Correspondingly one estimates the integrations over $x^2, \ldots, x^N$. The notation $L_{2|x^i}$ indicates, that the L_2-norm is taken with respect to the variables x^i. Altogether, we find

$$\begin{aligned}
&\left(\sum_{k_N=0}^{\infty} 2^{k_N \rho_N p} \int_{Q_{k_N}} \cdots \sum_{k_1=0}^{\infty} 2^{k_1 \rho_1 p} \int_{Q_{k_1}} |\mathcal{F} f(x)|^p dx^1 \cdots dx^N \right)^{\frac{1}{p}} \\
&\quad \le c \left(\sum_{k_N=0}^{\infty} 2^{k_N \kappa_N p} \left\| \mathcal{X}_{k_N}^N (x^N) \cdots \left(\sum_{k_2=0}^{\infty} 2^{k_2 \kappa_2 p} \right. \right. \right. \\
&\qquad \times \left. \left. \left. \left\| \mathcal{X}_{k_2}^2 (x^2) \left(\sum_{k_1=0}^{\infty} 2^{k_1 \kappa_1 p} \left\| \mathcal{X}_{k_1}^1 (x^1) (\mathcal{F} f)(x) \middle| L_{2|x^1} \right\|^p \right)^{\frac{1}{p}} \middle| L_{2|x^2} \right\|^p \right)^{\frac{1}{p}} \cdots \middle| L_{2|x^N} \right\|^p \right)^{\frac{1}{p}}.
\end{aligned}$$

Now we apply Hölder's inequality once more, this time for series, and use $\kappa_i - r_i < 0$. In this way, a convergent geometric series arises. Again, we demonstrate it for k_1 only, the other sums can be treated similarly (iteratively for $k_2, \ldots, k_N$):

$$\begin{aligned}
&\left(\sum_{k_1=0}^{\infty} 2^{k_1 \kappa_1 p} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^p \right)^{1/p} \\
&\qquad \le \left(\sum_{k_1=0}^{\infty} 2^{k_1 (\kappa_1 - r_1) p \cdot \frac{2}{2-p}} \right)^{\frac{2-p}{2} \cdot \frac{1}{p}} \left(\sum_{k_1=0}^{\infty} 2^{k_1 r_1 p \cdot \frac{2}{p}} \left\| \mathcal{X}_{k_1}^1 \mathcal{F} f \middle| L_{2|x^1} \right\|^{p \cdot \frac{2}{p}} \right)^{\frac{1}{p} \cdot \frac{p}{2}}
\end{aligned}$$

$$\leq c \left(\sum_{k_1=0}^{\infty} 2^{2k_1 r_1} \left\| \mathcal{X}^1_{k_1} \mathcal{F} f \middle| L_{2|x^1} \right\|^2 \right)^{1/2} .$$

This has to be multiplied by $\mathcal{X}^2_{k_2}$, and then the $L_{2|x^2}$-norm is formed etc. The arising iterative integrals (after interchanging summation and integration) can be comprised to an $L_2(\mathbb{R}^d)$-norm. Arguing as at the beginning of the proof, we finally obtain

$$\begin{aligned}
&\left\|(1+|x^1|^2)^{\rho_1/2} \cdots (1+|x^N|^2)^{\rho_N/2} \mathcal{F} f | L_p(\mathbb{R}^d)\right\| \\
&\leq c \left(\sum_{k_N=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} 2^{2k_1 r_1 + \cdots + 2k_N r_N} \left\| (\mathcal{X}^1_{k_1} \otimes \cdots \otimes \mathcal{X}^N_{k_N}) \mathcal{F} f \middle| L_2(\mathbb{R}^d) \right\|^2 \right)^{1/2} \\
&\leq c' \left\|(1+|x^1|^2)^{r_1/2} \cdots (1+|x^N|^2)^{r_N/2} \mathcal{F} f \middle| L_2(\mathbb{R}^d)\right\| = c' \left\| f \middle| S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})\right\| ,
\end{aligned}$$

what finally proves the assertion. □

Remark 2.3.3. The lemma and its proof are based on [71, Proposition 1.8.3]. One can even prove a slightly stronger version of this result for $f \in S^{\overline{\kappa}}_{2,p} B(\mathbb{R}^{\overline{d}})$, see [71], Section 1.7.5, as well as J. Peetre [57], pages 9–11, and the references given there. Nevertheless, the above lemma suffices for our purposes.

Definition 2.3.1. Let $0 < p \leq \infty$, and let $\Omega \subset \mathbb{R}^n$ be compact. Then we put

$$L^{\Omega}_p(\mathbb{R}^n) := \left\{ f \in \mathcal{S}'(\mathbb{R}^n) : \operatorname{supp} \mathcal{F} f \subset \Omega, \left\| f \middle| L_p(\mathbb{R}^n) \right\| < \infty \right\}.$$

By the Paley-Wiener-Schwartz theorem the spaces $L^{\Omega}_p(\mathbb{R}^n)$ consist of analytic functions.

Lemma 2.3.3. Let Ω and Γ be compact subsets of $\mathbb{R}^n$. Let $0 < p \leq \infty$ and $u = \min(1, p)$. Then there exists a positive constant c, such that

$$\left\|\mathcal{F}^{-1} M \mathcal{F} f \middle| L_p(\mathbb{R}^n)\right\| \leq c \left\|\mathcal{F}^{-1} M | L_u(\mathbb{R}^n)\right\| \cdot \left\| f \middle| L_p(\mathbb{R}^n)\right\|$$

holds for all $f \in L^{\Omega}_p$ and all $\mathcal{F}^{-1} M \in L^{\Gamma}_u$.

A proof for this lemma can be found in [83], Proposition 1.5.1. We shall add a corollary dealing with the behaviour of the appearing constant $c = c(\Gamma, \Omega)$ for some special cases of sets Γ and Ω.

Corollary 2.3.1. We define sets

$$\begin{aligned}
&\Gamma_j := \{ x \in \mathbb{R}^d \,:\, |x| \leq 2^{j+1} \} , \quad j \in \mathbb{N}_0 , \\
&\Omega_{\overline{k}} := \{ x \in \mathbb{R}^d \,:\, |x^i| \leq 2^{k_i+1},\ i = 1, \ldots, N \} , \quad \overline{k} \in \mathbb{N}^N_0 .
\end{aligned}$$

Then we obtain

$$\left\|\mathcal{F}^{-1} M \mathcal{F} f \middle| L_p(\mathbb{R}^d)\right\| \leq c_1 \, 2^{\overline{k} \cdot \overline{d}(1/u-1)} \left\|\mathcal{F}^{-1} M \middle| L_u(\mathbb{R}^d)\right\| \cdot \left\| f \middle| L_p(\mathbb{R}^d)\right\|$$

for every $f \in L^{\Omega_{\overline{k}}}_p$ and $\operatorname{supp} M \subset \Omega_{\overline{k}}$ and

$$\left\|\mathcal{F}^{-1} M \mathcal{F} f \middle| L_p(\mathbb{R}^d)\right\| \leq c_2 \, 2^{jd(1/u-1)} \left\|\mathcal{F}^{-1} M \middle| L_u(\mathbb{R}^d)\right\| \cdot \left\| f \middle| L_p(\mathbb{R}^d)\right\|$$

for every $f \in L^{\Gamma_j}_p$ and $\operatorname{supp} M \subset \Gamma_j$ with constants c_1, c_2 independent of j and $\overline{k}$.

Proof. Let at first be $f \in L_p^{\Omega_{\overline{k}}}$ and $\operatorname{supp} M \subset \Omega_{\overline{k}}$. Then clearly $f(2^{-\overline{k}} \cdot) \in L_p^{\Omega_{\overline{0}}}$ and $\operatorname{supp} M(2^{\overline{k}} \cdot) \subset \Omega_{\overline{0}}$ for every $\overline{k} \in \mathbb{N}_0^N$. Then we find by standard calculational arguments

$$\begin{aligned}\left\|\mathcal{F}^{-1}M\mathcal{F}f\middle|L_p(\mathbb{R}^d)\right\| &= 2^{-\overline{k}\cdot\overline{d}/p}\left\|\left(\mathcal{F}^{-1}M\mathcal{F}f\right)(2^{-\overline{k}}\cdot)\middle|L_p(\mathbb{R}^d)\right\| \\ &= 2^{-\overline{k}\cdot\overline{d}/p}2^{\overline{k}\cdot\overline{d}}\left\|\mathcal{F}^{-1}\left[(M\mathcal{F}f)(2^{\overline{k}}\cdot)\right]\middle|L_p(\mathbb{R}^d)\right\| \\ &\leq c\,2^{-\overline{k}\cdot\overline{d}/p}2^{\overline{k}\cdot\overline{d}}2^{-\overline{k}\cdot\overline{d}}\left\|\mathcal{F}^{-1}\left[M(2^{\overline{k}}\cdot)\right]\middle|L_u(\mathbb{R}^d)\right\|\cdot\left\|f(2^{-\overline{k}}\cdot)\middle|L_p(\mathbb{R}^d)\right\| \\ &\leq c\,2^{-\overline{k}\cdot\overline{d}/p}2^{-\overline{k}\cdot\overline{d}}\left\|(\mathcal{F}^{-1}M)(2^{-\overline{k}}\cdot)\middle|L_u(\mathbb{R}^d)\right\|\cdot 2^{\overline{k}\cdot\overline{d}/p}\left\|f\middle|L_p(\mathbb{R}^d)\right\| \\ &\leq c\,2^{-\overline{k}\cdot\overline{d}}2^{\overline{k}\cdot\overline{d}/u}\left\|\mathcal{F}^{-1}M\middle|L_u(\mathbb{R}^d)\right\|\cdot\left\|f\middle|L_p(\mathbb{R}^d)\right\|.\end{aligned}$$

The proof of the second assertion follows by similar arguments. □

Theorem 2.3.2 (Minkowski's inequality). Let $(\Omega_1, \mathfrak{M}_1, \mu_1)$ and $(\Omega_2, \mathfrak{M}_2, \mu_2)$ be σ-finite measure spaces, and let f be a $\mu_1 \otimes \mu_2$-measurable function on $\Omega_1 \times \Omega_2$. Then it holds for $1 \leq p \leq \infty$

$$\left\|\int_{\Omega_2} f(\cdot, y)d\mu_2(y)\middle|L_p(\Omega_1)\right\| \leq \int_{\Omega_2}\left\|f(\cdot, y)\middle|L_p(\Omega_1)\right\|d\mu_2(y).$$

Remark 2.3.4. For more details and a proof for this version of Minkowski's inequality see either the literature on (classical) measure theory and integration, e.g. [47] or [65], or on vector measures and Bochner integration, e.g. [20] or [24]. In the framework of Bochner-integration the theorem is a consequence of the elementary inequality $\left\|\int_\Omega f d\mu\right\|_X \leq \int_\Omega \|f(\cdot)\|_X d\mu$, applied to $X = L_p$, where $1 \leq p < \infty$.

Remark 2.3.5. We will apply this inequality in two special cases. In each case, one of the measure spaces will be the standard one for the Lebesgue measure, $(\mathbb{R}^d, \mathfrak{B}_d, \lambda^d)$. The other one is either $(\mathbb{N}, \mathfrak{P}\mathbb{N}, \mu)$, where μ is the counting measure on the power set $\mathfrak{P}\mathbb{N}$ of $\mathbb{N}$, analogous for any other countable set Ω, or $(\mathbb{R}^d, \mathfrak{B}_d, w\lambda^d)$, where $w : \mathbb{R}^d \longrightarrow [0, \infty)$ is some density function (weight function). In the first situation the case $p = \infty$ is an almost trivial consequence of the monotonicity properties of the Lebesgue integral.

Lemma 2.3.4. Let $\overline{r} \in \mathbb{R}^N$ with $\overline{r} > \overline{d}/2$. Moreover, let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be compactly supported, and let $M \in S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})$. Then it holds

$$\left\|\psi M\middle|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\right\| \leq c\left\|M\middle|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\right\|$$

with a constant $c > 0$ independent of M.

Proof. The assumption $\overline{r} > \overline{d}/2$ makes it possible to apply Lemma 2.3.2 with $\overline{\rho} = \overline{0}$ and $p = 1$. This yields $\mathcal{F}M \in L_1(\mathbb{R}^d)$, and it holds

$$\mathcal{F}\big(\psi M\big)(\xi) = (2\pi)^{d/2}\big(\mathcal{F}\psi * \mathcal{F}M\big)(\xi)\,.$$

Hence we obtain

$$\left\|\psi M\middle|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\right\|^2 = \int_{\mathbb{R}^d}\left(\prod_{i=1}^N(1+|\xi^i|^2)^{r_i}\right)(2\pi)^d\left|\big(\mathcal{F}\psi * \mathcal{F}M\big)(\xi)\right|^2 d\xi$$

$$= (2\pi)^d \int_{\mathbb{R}^d} \left(\prod_{i=1}^N (1+|\xi^i|^2)^{r_i} \right) \left| \int_{\mathbb{R}^d} \mathcal{F}\psi(\eta)\mathcal{F}M(\xi-\eta)d\eta \right|^2 d\xi. \qquad (2.3.6)$$

Now we put $w(\xi) = \prod_{i=1}^N (1+|\xi^i|^2)^{r_i}$, and consider

$$\big\|f\big|L_{2,w}(\mathbb{R}^d)\big\| := \left(\int_{\mathbb{R}^d} |f(\xi)|^2 w(\xi) d\xi \right)^{1/2},$$

i.e. $\| \cdot |L_{2,w}(\mathbb{R}^d)\|$ is the L_2-norm with respect to the measure $w\lambda^d$. First of all, we find

$$\begin{aligned} w(\xi) &\le \prod_{i=1}^N \Big(1 + \big(|\xi^i - \eta^i| + |\eta^i| \big)^2 \Big)^{r_i} \\ &\le \prod_{i=1}^N \Big(2\big(1+|\xi^i-\eta^i|^2\big)\big(1+|\eta^i|^2\big) \Big)^{r_i} = 2^{|\overline{r}|_1} w(\xi-\eta)w(\eta) \end{aligned}$$

for all $\xi, \eta \in \mathbb{R}^d$. Applying Minkowski's inequality, we get

$$\begin{aligned} &\left\| \int_{\mathbb{R}^d} \mathcal{F}M(\xi-\eta)\mathcal{F}\psi(\eta)d\eta \middle| L_{2,w}(\mathbb{R}^d) \right\| \\ &= \left\| \int_{\mathbb{R}^d} \mathcal{F}M(\xi-\eta)\left(\frac{w(\xi-\eta)}{w(\xi)}\right)^{1/2} \mathcal{F}\psi(\eta)\left(\frac{w(\xi)}{w(\xi-\eta)}\right)^{1/2} d\eta \middle| L_{2,w}(\mathbb{R}^d) \right\| \\ &\le \int_{\mathbb{R}^d} \left\| \mathcal{F}M(\xi-\eta)\left(\frac{w(\xi-\eta)}{w(\xi)}\right)^{1/2} \mathcal{F}\psi(\eta)\left(\frac{w(\xi)}{w(\xi-\eta)}\right)^{1/2} \middle| L_{2,w}(\mathbb{R}^d) \right\| d\eta\,. \qquad (2.3.7) \end{aligned}$$

Apparently, it holds

$$\begin{aligned} &\int_{\mathbb{R}^d} \big|\mathcal{F}M(\xi-\eta)\big|^2 \frac{w(\xi-\eta)}{w(\xi)} \big|\mathcal{F}\psi(\eta)\big|^2 \frac{w(\xi)}{w(\xi-\eta)} w(\xi)d\xi \\ &\le \sup_{\tau\in\mathbb{R}^d} \left| \mathcal{F}\psi(\eta)\left(\frac{w(\tau)}{w(\tau-\eta)}\right)^{1/2} \right|^2 \int_{\mathbb{R}^d} \big|\mathcal{F}M(\xi-\eta)\big|^2 \frac{w(\xi-\eta)}{w(\xi)} w(\xi)d\xi \\ &\le \sup_{\tau\in\mathbb{R}^d} \left| \mathcal{F}\psi(\eta)\left(\frac{2^{|\overline{r}|_1}w(\tau-\eta)w(\eta)}{w(\tau-\eta)}\right)^{1/2} \right|^2 \big\|\mathcal{F}M\big|L_{2,w}(\mathbb{R}^d)\big\|^2 \\ &= \left| \mathcal{F}\psi(\eta)\Big(2^{|\overline{r}|_1}w(\eta)\Big)^{1/2} \right|^2 \big\|M\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|^2. \qquad (2.3.8) \end{aligned}$$

Inserting (2.3.8) in (2.3.6) and (2.3.7), respectively, finally yields

$$\begin{aligned} \big\|\psi M\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\| &\le (2\pi)^{d/2} \int_{\mathbb{R}^d} \big|\mathcal{F}\psi(\eta)\big| \Big(2^{|\overline{r}|_1}w(\eta)\Big)^{1/2} \big\|M\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\| d\eta \\ &= (2\pi)^{d/2} 2^{|\overline{r}|_1/2} \big\|M\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\| \int_{\mathbb{R}^d} \big|\mathcal{F}\psi(\eta)\big| \prod_{i=1}^N \big(1+|\eta^i|^2\big)^{r_i/2} d\eta \\ &= c\,\big\|M\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|\,. \end{aligned}$$

The last integral is finite as $\psi \in \mathcal{S}(\mathbb{R}^d)$ if, and only if, its Fourier transform $\mathcal{F}\psi$ belongs to $\mathcal{S}(\mathbb{R}^d)$, and hence $\sup_{\eta\in\mathbb{R}^d} |\mathcal{F}\psi(\eta)|(1+|\eta|^2)^s < \infty$ for every $s \in \mathbb{R}$. □

Proposition 2.3.3. Let $0 < p \le \infty$, and let Ω be a compact subset of $\mathbb{R}^d$ with

$$\Omega \subset Q_{\overline{b}} = \{x \in \mathbb{R}^d : |x^i| \le b_i, i = 1, \dots, N\} \quad \text{for some} \quad \overline{b} > 0\,.$$

Moreover, let $\overline{r} \in \mathbb{R}^N$ with

$$r_i > \sigma_i = d_i\left(\frac{1}{\min(1,p)} - \frac{1}{2}\right), \quad i = 1\dots, N\,. \tag{2.3.9}$$

Then, every function $M \in S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})$ is a Fourier multiplier for L_p^Ω. Furthermore, there exists a positive constant c, such that

$$\|\mathcal{F}^{-1}M\mathcal{F}f|L_p(\mathbb{R}^d)\| \le c\,\|M(b_1\cdot,\dots,b_N\cdot)|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\| \cdot \|\,f\,|L_p(\mathbb{R}^d)\|$$

holds for all $M \in S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})$, all $\overline{b} \in \mathbb{R}_+^N$, all $\Omega \subset Q_{\overline{b}}$ and all $f \in L_p^\Omega$.

Proof. As $\Omega_1 \subset \Omega_2$ always implies $L_p^{\Omega_1} \subset L_p^{\Omega_2}$, it suffices to prove the assertion for $\Omega = Q_{\overline{b}}$. To this purpose, we shall begin with the case $b_1 = \dots = b_N = 1$.
Therefore, let $\Omega = Q_{\overline{1}}$, and let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be a compactly supported function, $\operatorname{supp}\psi =: \Gamma$, satisfying

$$\psi(x) = 1 \quad \text{for all } x \in \Omega. \tag{2.3.10}$$

Applying Lemma 2.3.2 with $\overline{\rho} = 0$ and $\widetilde{p} := \min(1,p)$ (the assumptions of that lemma are assured by (2.3.9)) as well as Lemma 2.3.4 (by (2.3.9) it holds $\overline{r} > \overline{\sigma} \ge \overline{d}/2$) yields

$$\begin{aligned}\|\mathcal{F}^{-1}(\psi M)|L_{\widetilde{p}}(\mathbb{R}^d)\| &= \|\mathcal{F}(\psi M)|L_{\widetilde{p}}(\mathbb{R}^d)\| \\ &\le c\,\|\psi M|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\| \le c'\,\|M|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\|\,.\end{aligned} \tag{2.3.11}$$

Since $\operatorname{supp}\psi M \subset \operatorname{supp}\psi$, it follows $\mathcal{F}^{-1}(\psi M) \in L_{\widetilde{p}}^\Gamma$. Hence, we can apply Lemma 2.3.3 to ψM instead of M in order to obtain

$$\|\mathcal{F}^{-1}\psi M\mathcal{F}f\|L_p(\mathbb{R}^d)\| \le c\,\|\mathcal{F}^{-1}(\psi M)\|L_{\widetilde{p}}(\mathbb{R}^d)\| \cdot \|f|L_p(\mathbb{R}^d)\|\,.$$

From this, together with (2.3.10) and (2.3.11), it finally follows

$$\|\mathcal{F}^{-1}M\mathcal{F}f|L_p(\mathbb{R}^d)\| = \|\mathcal{F}^{-1}\psi M\mathcal{F}f|L_p(\mathbb{R}^d)\| \le c\,\|M|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\| \cdot \|f|L_p(\mathbb{R}^d)\|$$

for all $f \in L_p^\Omega$.

The general case now follows by a simple scaling argument: Since $\operatorname{supp}\mathcal{F}f \subset Q_{\overline{b}} \iff \operatorname{supp}(\mathcal{F}f)(b_1\cdot,\dots,b_N\cdot) \subset Q_{\overline{1}}$, it holds $f \in L_p^{Q_{\overline{b}}} \iff f\big(\frac{\cdot}{b_1},\dots,\frac{\cdot}{b_N}\big) \in L_p^{Q_{\overline{1}}}$. But then we find for $\widetilde{f} := b_1^{-d_1}\cdots b_N^{-d_N} f\big(\frac{\cdot}{b_1},\dots,\frac{\cdot}{b_N}\big)$ and $\widetilde{M} := M(b_1\cdot,\dots,b_N\cdot)$

$$\begin{aligned}&\|\mathcal{F}^{-1}M\mathcal{F}f|L_p(\mathbb{R}^d)\| \\ &\quad = \Big\|\big(b_1^{-d_1}\cdots b_N^{-d_N}\big)^{1/p}\big(\mathcal{F}^{-1}M\mathcal{F}f\big)\big(\tfrac{\cdot}{b_1},\dots,\tfrac{\cdot}{b_N}\big)\Big|L_p(\mathbb{R}^d)\Big\| \\ &\quad = \Big\|\mathcal{F}^{-1}\Big[\big(M\mathcal{F}f\big)\big(b_1\cdot,\dots,b_N\cdot\big)\Big]\Big|L_p(\mathbb{R}^d)\Big\|\big(b_1^{-d_1}\cdots b_N^{-d_N}\big)^{1/p-1}\end{aligned}$$

$$
\begin{aligned}
&= \left\| \mathcal{F}^{-1}\left[\widetilde{M}\mathcal{F}\widetilde{f}\right]\Big|L_p(\mathbb{R}^d)\right\| \left(b_1^{-d_1}\cdots b_N^{-d_N}\right)^{1/p-1} \\
&\le c\,\big\|\widetilde{M}\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|\big\|\widetilde{f}\big|L_p(\mathbb{R}^d)\big\|\left(b_1^{-d_1}\cdots b_N^{-d_N}\right)^{1/p-1} \\
&= c\,\big\|M\big(b_1\cdot,\ldots,b_N\cdot\big)\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|\cdot\big\|f\big(\tfrac{\cdot}{b_1},\ldots,\tfrac{\cdot}{b_N}\big)\big|L_p(\mathbb{R}^d)\big\|\left(b_1^{-d_1}\cdots b_N^{-d_N}\right)^{1/p} \\
&= c\,\big\|M\big(b_1\cdot,\ldots,b_N\cdot\big)\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|\cdot\big\|f\big|L_p(\mathbb{R}^d)\big\|\,,
\end{aligned}
$$

where the constant c is independent of $M \in S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})$, $f \in L_p^{Q_{\overline{b}}}$, and $\overline{b} \in \mathbb{R}_+^N$. This finally proves the assertion. □

Remark 2.3.6. For spaces of dominating mixed smoothness $S_{p,q}^r A(\mathbb{R}^d)$ a corresponding result can be found in [71, Theorem 1.8.3], and for its isotropic counterpart we refer to [83, Theorem 1.5.2].

We shall need a vector-valued version of the above proposition as well. But at first, we will prove some auxiliary means. The next lemma is a maximal inequality for a variant of the Peetre maximal operator. To this end, let $\Omega = \big(\Omega_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be a sequence of compact subsets of $\mathbb{R}^d$, defined by

$$\Omega_{\overline{k}} = \left\{x \in \mathbb{R}^d : |x^1| \le a_{1,k_1},\ldots,|x^N| \le a_{N,k_N}\right\}$$

for some $a_{1,k_1},\ldots,a_{N,k_N} > 0$. The result is based on [71, Theorem 1.6.4].

Lemma 2.3.5. Let $\Gamma \subset \mathbb{R}^d$ be compact, and let $0 < p \le \infty$.

(i) Let $\overline{r} \in \mathbb{R}^N$ with $\overline{r} > 0$. Then there is a positive constant c, such that

$$\sup_{z\in\mathbb{R}^d} \frac{|f(x-z)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})} \le c\left(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\right)^{1/r_N}(x) \qquad (2.3.12)$$

holds for all $f \in L_p^\Gamma$ and all $x \in \mathbb{R}^d$.

(ii) Let $0 < r_i < p < \infty$, $i = 1,\ldots,N$. Then there exists a positive constant c, such that

$$\left\|\sup_{z\in\mathbb{R}^d} \frac{|f(\cdot-z)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})}\right|L_p(\mathbb{R}^d)\Bigg\| \le c\,\big\|f\big|L_p(\mathbb{R}^d)\big\| \qquad (2.3.13)$$

holds for all $f \in L_p^\Gamma$.

Proof. **Step 1:**
Let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be a function satisfying $(\mathcal{F}\psi)(x) = 1$ for all $x \in \Gamma$. In case $\operatorname{supp}\mathcal{F}f \subset \Gamma$, it follows $\mathcal{F}f = \mathcal{F}f\cdot\mathcal{F}\psi$ and $f = \mathcal{F}^{-1}(\mathcal{F}f\cdot\mathcal{F}\psi)$. Then known properties of the Fourier transform yield

$$f(x) = (2\pi)^{d/2}\int_{\mathbb{R}^d} f(y)\psi(x-y)dy, \quad x \in \mathbb{R}^d\,.$$

Now we replace in this equation x by $x-z$ and apply ∇_x to obtain

$$\big|(\nabla_x f)(x-z)\big| \le (2\pi)^{d/2} d^{1/2} \int_{\mathbb{R}^d} |f(y)|\big|(\nabla_x \psi)(x-z-y)\big| dy\,.$$

We divide both sides by $\prod_{i=1}^N (1+|z^i|^{d_i/r_i})$ and use the inequality

$$\frac{1+|x^i-y^i|^{d_i/r_i}}{1+|z^i|^{d_i/r_i}} \le c\,(1+|x^i-y^i-z^i|^{d_i/r_i}), \qquad x^i,y^i,z^i \in \mathbb{R}^{d_i}, i=1,\dots,N.$$

This follows from the triangle inequality in $\mathbb{R}^{d_i}$. Thus we get

$$\begin{aligned}
&\sup_{z\in\mathbb{R}^d} \frac{|\nabla_x f(x-z)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})} \\
&\quad \le c \sup_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|f(y)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})} \big|(\nabla_x\psi)(x-z-y)\big| dy \\
&\quad \le c' \sup_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} |f(y)|\big|(\nabla_x\psi)(x-z-y)\big| \prod_{i=1}^N \frac{1+|x^i-y^i-z^i|^{d_i/r_i}}{1+|x^i-y^i|^{d_i/r_i}} dy \\
&\quad \le c' \sup_{w\in\mathbb{R}^d} \frac{|f(w)|}{\prod_{i=1}^N (1+|x^i-w^i|^{d_i/r_i})} \\
&\qquad\qquad \times \sup_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} \big|(\nabla_x\psi)(x-y-z)\big| \prod_{i=1}^N (1+|x^i-y^i-z^i|^{d_i/r_i}) dy \\
&\quad = c' \sup_{w\in\mathbb{R}^d} \frac{|f(w)|}{\prod_{i=1}^N (1+|x^i-w^i|^{d_i/r_i})} \int_{\mathbb{R}^d} \big|(\nabla_v\psi)(v)\big| \prod_{i=1}^N (1+|v^i|^{d_i/r_i}) dv\,.
\end{aligned}$$

The integral is finite due to general properties of functions from $\mathcal{S}(\mathbb{R}^d)$, in particular $\sup_{v\in\mathbb{R}^d}(1+|v|^2)^k|(D^\alpha\psi)(v)| < \infty$ for all $\alpha \in \mathbb{N}_0^d$ and all $k \in \mathbb{N}_0$. Altogether we have found

$$\sup_{z\in\mathbb{R}^d} \frac{|\nabla_x f(x-z)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})} \le c_0 \sup_{z\in\mathbb{R}^d} \frac{|f(x-z)|}{\prod_{i=1}^N (1+|z^i|^{d_i/r_i})}\,, \tag{2.3.14}$$

where the constant c_0 is independent of $f \in L_p^\Gamma$ and $x \in \mathbb{R}^d$.

Step 2:
We need another auxiliary result. Let g be a complex-valued continuously differentiable function on the set $B_\delta = \{y \in \mathbb{R}^d : |y^i| \le \delta, i=1,\dots,N\}$ for some arbitrary fixed $\delta > 0$. Then the mean value theorem yields

$$|g(z)| \le \min_{w\in B_\delta} |g(w)| + \sqrt{2N\delta} \sup_{w\in B_\delta} |\nabla g(w)|\,, \qquad z \in B_\delta\,.$$

The minimum can be further estimated by

$$\prod_{i=1}^N |B_\delta^i|^{-\frac{1}{r_i}} \left(\int_{B_\delta^N} \cdots \left(\int_{B_\delta^2} \left(\int_{B_\delta^1} \Big(\min_{w\in B_\delta} |g(w)| \Big)^{r_1} du^1 \right)^{r_2/r_1} du^2 \right)^{r_3/r_2} \cdots du^N \right)^{1/r_N}$$

$$\leq \prod_{i=1}^{N}\Big(|B_1^i|^{-\frac{1}{r_i}}\delta^{-\frac{d_i}{r_i}}\Big)\left(\int_{B_\delta^N}\cdots\left(\int_{B_\delta^2}\left(\int_{B_\delta^1}|g(u)|^{r_1}du^1\right)^{r_2/r_1}du^2\right)^{r_3/r_2}\cdots du^N\right)^{1/r_N},$$

where we used $B_\delta = B_\delta^1\times\cdots\times B_\delta^N$, $B_\delta^i=\{t\in\mathbb{R}^{d_i}:|t|\leq\delta\}$. Hence, we obtain for $z\in B_\delta$

$$|g(z)|\leq c\prod_{i=1}^{N}\delta^{-d_i/r_i}\left(\int_{B_\delta^N}\cdots\left(\int_{B_\delta^2}\left(\int_{B_\delta^1}|g(u)|^{r_1}du^1\right)^{\frac{r_2}{r_1}}du^2\right)^{\frac{r_3}{r_2}}\cdots du^N\right)^{\frac{1}{r_N}} + \sqrt{2N\delta}\sup_{w\in B_\delta}|\nabla g(w)| \tag{2.3.15}$$

with some positive constant c independent of δ.

Step 3:
Now we apply the inequality (2.3.15) to the function $f(x-y-\cdot)\in L_p^{-\Gamma}$ with respect to the point $0\in B_\delta$ in order to obtain

$$|f(x-y)|\leq c\,\delta^{-\overline{d}\cdot\frac{1}{\overline{r}}}\left(\int_{B_\delta^N}\cdots\left(\int_{B_\delta^2}\left(\int_{B_\delta^1}|f(x-y-u)|^{r_1}du^1\right)^{\frac{r_2}{r_1}}du^2\right)^{\frac{r_3}{r_2}}\cdots du^N\right)^{\frac{1}{r_N}} + \sqrt{2N\delta}\sup_{w\in B_\delta}\big|\nabla f(x-y-w)\big|. \tag{2.3.16}$$

The integral in (2.3.16) can be estimated from above for $0<\delta<1$ as follows:

$$\begin{aligned}
&\left(\int_{B_\delta^N}\cdots\left(\int_{B_\delta^2}\left(\int_{B_\delta^1}|f(x-y-u)|^{r_1}du^1\right)^{r_2/r_1}du^2\right)^{r_3/r_2}\cdots du^N\right)^{1/r_N}\\
&\leq\left(\int_{B^N_{\delta+|y^N|}}\cdots\left(\int_{B^2_{\delta+|y^2|}}\left(\int_{B^1_{\delta+|y^1|}}|f(x-u)|^{r_1}du^1\right)^{r_2/r_1}du^2\right)^{r_3/r_2}\cdots du^N\right)^{1/r_N}\\
&\leq c\prod_{i=1}^{N}(\delta+|y^i|)^{d_i/r_i}\big(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big)^{1/r_N}(x)\\
&\leq c'\prod_{i=1}^{N}\big(1+|y^i|^{d_i/r_i}\big)\big(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big)^{1/r_N}(x).
\end{aligned}$$

Now we insert the last inequality in (2.3.16). Afterwards we divide by $\prod_{i=1}^{N}\big(1+|y^i|^{d_i/r_i}\big)$ and take a supremum over $y\in\mathbb{R}^d$. Altogether we find

$$\sup_{y\in\mathbb{R}^d}\frac{|f(x-y)|}{\prod_{i=1}^{N}(1+|y^i|^{d_i/r_i})}\leq c\,\delta^{-\overline{d}\cdot\frac{1}{\overline{r}}}\big(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big)^{1/r_N}(x) + \sup_{y\in\mathbb{R}^d}\sqrt{2N\delta}\,\frac{\sup_{w\in B_\delta}|\nabla f(x-y-w)|}{\prod_{i=1}^{N}(1+|y^i|^{d_i/r_i})}. \tag{2.3.17}$$

Moreover, we shall use the estimate

$$\frac{\prod_{i=1}^{N}(1+|z^i|^{d_i/r_i})}{\prod_{i=1}^{N}(1+|z^i-w^i|^{d_i/r_i})}\leq c\,\frac{\prod_{i=1}^{N}(1+|z^i-w^i|^{d_i/r_i}+|w^i|^{d_i/r_i})}{\prod_{i=1}^{N}(1+|z^i-w^i|^{d_i/r_i})}$$

$$\leq c\,\frac{\Big(\prod_{i=1}^N(1+|z^i-w^i|^{d_i/r_i})\Big)\Big(\prod_{i=1}^N(1+|w^i|^{d_i/r_i})\Big)}{\prod_{i=1}^N(1+|z^i-w^i|^{d_i/r_i})}$$
$$= c\prod_{i=1}^N(1+|w^i|^{d_i/r_i}) \leq c\prod_{i=1}^N(1+\delta^{d_i/r_i}) \leq c\,2^N,$$

for all $0<\delta\leq 1$ and $w\in B_\delta$. By putting $z=y+w$ we further obtain

$$\frac{\sup_{w\in B_\delta}|\nabla f(x-y-w)|}{\prod_{i=1}^N(1+|y^i|^{d_i/r_i})} = \sup_{w\in B_\delta}\frac{|\nabla f(x-z)|}{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})}\frac{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})}{\prod_{i=1}^N(1+|z^i-w^i|^{d_i/r_i})}$$
$$\leq c\sup_{z-y\in B_\delta}\frac{|\nabla f(x-z)|}{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})} \leq c\sup_{z\in\mathbb{R}^d}\frac{|\nabla f(x-z)|}{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})}\,.$$

If we insert this into (2.3.17), we end up with

$$\sup_{y\in\mathbb{R}^d}\frac{|f(x-y)|}{\prod_{i=1}^N(1+|y^i|^{d_i/r_i})} \leq c\,\delta^{-\overline{d}\cdot\frac{1}{\overline{r}}}\Big(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\Big)^{1/r_N}(x)$$
$$+\,c_1\sqrt{2N\delta}\sup_{y\in\mathbb{R}^d}\frac{|\nabla f(x-y)|}{\prod_{i=1}^N(1+|y^i|^{d_i/r_i})}\,.$$

The assertion (2.3.12) now follows from this last estimate together with (2.3.14), if we choose δ sufficiently small. Of course this requires of the supremum on right hand side to be finite, but this is a consequence of $f\in L_p^\Gamma$ and Nikol'skij's inequality, which is the subject of the next subsection (Proposition 2.3.6). This inequality yields the boundedness of all partial derivatives of the analytic function f, and hence also of the norm of the gradient.

Step 4:
We finally prove (2.3.13). From (2.3.12) and the maximal inequality (2.3.3) from Proposition 2.3.1 (by assumption, we have $p/r_i>1$ for all $i=1,\ldots,N$) we conclude

$$\left\|\sup_{z\in\mathbb{R}^d}\frac{|f(\cdot-z)|}{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})}\,\middle|\,L_p(\mathbb{R}^d)\right\|$$
$$\leq c\left\|\Big(M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\Big)^{1/r_N}\,\middle|\,L_p(\mathbb{R}^d)\right\|$$
$$= c\left\|M_N\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\,\middle|\,L_{p/r_N}(\mathbb{R}^d)\right\|^{1/r_N}$$
$$\leq c'\left\|\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\,\middle|\,L_{p/r_N}(\mathbb{R}^d)\right\|^{1/r_N}$$
$$= c'\left\|M_{N-1}\big(\cdots M_2\big(M_1|f|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_{N-1}/r_{N-2}}\,\middle|\,L_{p/r_{N-1}}(\mathbb{R}^d)\right\|^{1/r_{N-1}}.$$

Iterating this argument eventually yields

$$\left\|\sup_{z\in\mathbb{R}^d}\frac{|f(\cdot-z)|}{\prod_{i=1}^N(1+|z^i|^{d_i/r_i})}\,\middle|\,L_p(\mathbb{R}^d)\right\|$$
$$\leq c\left\|M_1|f|^{r_1}\,\middle|\,L_{p/r_1}(\mathbb{R}^d)\right\|^{1/r_1} \leq c'\left\||f|^{r_1}\,\middle|\,L_{p/r_1}(\mathbb{R}^d)\right\|^{1/r_1} = c'\,\|\,f\,|L_p(\mathbb{R}^d)\|\,.$$

Now the proof is complete. □

The next proposition is a vector-valued counterpart of Lemma 2.3.5(ii). It is based on [71, Theorem 1.10.2(ii)].

Proposition 2.3.4. Let $0 < p < \infty$, $0 < q \leq \infty$, and let $\Omega = \big(\Omega_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ as before. Furthermore, let

$$0 < r_i < \min(p,q), \quad i = 1,\ldots,N\,.$$

Then there exists a positive constant c, such that

$$\left\| \sup_{z\in\mathbb{R}^d} \frac{|f_{\overline{k}}(\cdot - z)|}{\prod_{i=1}^N (1+|a_{i,k_i} z^i|^{d_i/r_i})} \middle| L_p(\ell_q) \right\| \leq c \, \big\| f_{\overline{k}} \big| L_p(\ell_q) \big\|$$

holds for all sequences of functions $\big(f_{\overline{k}}\big)_{k\in\mathbb{N}_0^N} \subset L_p(\mathbb{R}^d)$ with $\operatorname{supp}\mathcal{F} f_{\overline{k}} \subset \Omega_{\overline{k}}$.

Proof. We apply Lemma 2.3.5(i) to $f_{\overline{k}}\big(\frac{\cdot}{a_{1,k_1}},\ldots,\frac{\cdot}{a_{N,k_N}}\big)$ (i.e. in this case, Γ is the unit cube). At first, one makes sure of

$$\begin{aligned}\Big(M_N\Big(\cdots M_2\Big(M_1\Big|f_{\overline{k}}\Big(\frac{\cdot}{a_{1,k_1}},\ldots,\frac{\cdot}{a_{N,k_N}}\Big)\Big|^{r_1}\Big)^{r_2/r_1}\cdots\Big)^{r_N/r_{N-1}}\Big)^{1/r_N}(x) \\ = \big(M_N\big(\cdots M_2\big(M_1|f_{\overline{k}}|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big)^{1/r_N}\Big(\frac{x^1}{a_{1,k_1}},\ldots,\frac{x^N}{a_{N,k_N}}\Big)\,.\end{aligned}$$

That way, we obtain

$$\sup_{z\in\mathbb{R}^d} \frac{|f_{\overline{k}}(x - z)|}{\prod_{i=1}^N (1+|a_{i,k_i} z^i|^{d_i/r_i})} \leq c\,\big(M_N\big(\cdots M_2\big(M_1|f_{\overline{k}}|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big)^{1/r_N}(x)\,,$$

where the constant c is independent of $x\in\mathbb{R}^d$, $\overline{k}$ and $f_{\overline{k}}$. Moreover, similarly to the proof of Lemma 2.3.5 we find

$$\begin{aligned}&\left\| \sup_{z\in\mathbb{R}^d} \frac{|f_{\overline{k}}(\cdot - z)|}{\prod_{i=1}^N (1+|a_{i,k_i} z^i|^{d_i/r_i})} \middle| L_p(\ell_q) \right\| \\ &\quad \leq c\,\big\|M_N\big(M_{N-1}\cdots M_2\big(M_1|f_{\overline{k}}|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big|L_{p/r_N}(\ell_{q/r_N})\big\|^{1/r_N}\,.\end{aligned}$$

The right hand side of this last inequality can be further estimated with the help of Proposition 2.3.1 (by assumption it holds $p/r_i > 1$ and $q/r_i > 1$) by

$$\begin{aligned}&c'\,\big\|\big(M_{N-1}\cdots M_2\big(M_1|f_{\overline{k}}|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_N/r_{N-1}}\big|L_{p/r_N}(\ell_{q/r_N})\big\|^{1/r_N} \\ &\quad = c'\,\big\|M_{N-1}\big(\cdots M_2\big(M_1|f_{\overline{k}}|^{r_1}\big)^{r_2/r_1}\cdots\big)^{r_{N-1}/r_{N-2}}\big|L_{p/r_{N-1}}(\ell_{q/r_{N-1}})\big\|^{1/r_{N-1}}.\end{aligned}$$

Iterated use of this argument eventually yields

$$\left\| \sup_{z\in\mathbb{R}^d} \frac{|f_{\overline{k}}(\cdot - z)|}{\prod_{i=1}^N (1+|a_{i,k_i} z^i|^{d_i/r_i})} \middle| L_p(\ell_q) \right\| \leq C\,\big\||f_{\overline{k}}|^{r_1}\big|L_{p/r_1}(\ell_{q/r_1})\big\|^{1/r_1} = C\,\big\|\,f_{\overline{k}}\,\big|L_p(\ell_q)\big\|.$$

This proves the assertion. □

Proposition 2.3.5. Let $\Omega = \big(\Omega_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ and $a_{1,k_1},\dots,a_{N,k_N} > 0$ be as before. Furthermore, let $0 < p < \infty$, $0 < q \le \infty$, and

$$\overline{r} = (r_1,\dots,r_N) > \overline{d}\bigg(\frac{1}{\min(p,q)} + \frac{1}{2}\bigg). \tag{2.3.18}$$

Then there exists a positive constant c, such that

$$\Big\|\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\Big|L_p(\ell_q)\Big\| \le c\,\Big(\sup_{\overline{k}\in\mathbb{N}_0^N}\big\|\rho_{\overline{k}}(a_{1,k_1}\cdot,\dots,a_{N,k_N}\cdot)\big|S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})\big\|\Big)\big\|f_{\overline{k}}\big|L_p(\ell_q)\big\|$$

holds for all sequences $\big(f_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \in L_p(\ell_q)$ with $\operatorname{supp}\mathcal{F}f_{\overline{k}} \subset \Omega_{\overline{k}}$, $\overline{k}\in\mathbb{N}_0^N$, and all sequences $\big(\rho_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \subset S_2^{\overline{r}}H(\mathbb{R}^{\overline{d}})$.

This proposition is the aspired vector-valued counterpart of Proposition 2.3.3. Its forerunner for multipliers from $S_2^rH(\mathbb{R}^d)$ can be found in [71, Theorem 1.10.3].

Proof. By Lemma 2.3.2 we have $\mathcal{F}^{-1}\rho_{\overline{k}} \in L_1(\mathbb{R}^d)$, hence the expressions $\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}$ make sense pointwise, and it holds

$$\big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)(x) = (2\pi)^{d/2}\int_{\mathbb{R}^d}\big(\mathcal{F}^{-1}\rho_{\overline{k}}\big)(y)f_{\overline{k}}(x-y)dy, \quad x\in\mathbb{R}^d.$$

Abbreviatory we put

$$f_{\overline{k}}^*(x) := \sup_{z\in\mathbb{R}^d}\frac{|f_{\overline{k}}(x-z)|}{\prod_{i=1}^N(1+|a_{i,k_i}z^i|^{d_i/s_i})} = \sup_{y\in\mathbb{R}^d}\frac{|f_{\overline{k}}(y)|}{\prod_{i=1}^N(1+|a_{i,k_i}(x^i-y^i)|^{d_i/s_i})}, \tag{2.3.19}$$

where the s_i will be chosen later on. Then it holds

$$\begin{aligned}
&\big|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)(x-z)\big| \\
&\quad\le (2\pi)^{d/2}\int_{\mathbb{R}^d}\big|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\big)(x-z-y)\big||f_{\overline{k}}(y)|dy \\
&\quad\le f_{\overline{k}}^*(x)(2\pi)^{d/2}\int_{\mathbb{R}^d}\big|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\big)(x-z-y)\big|\prod_{i=1}^N\big(1+|a_{i,k_i}(x^i-y^i)|^{d_i/s_i}\big)dy \\
&\quad\le c\,f_{\overline{k}}^*(x)\prod_{i=1}^N\big(1+|a_{i,k_i}z^i|^{d_i/s_i}\big)\frac{1}{\prod_{i=1}^N a_{i,k_i}^{d_i}} \\
&\qquad\quad\times\int_{\mathbb{R}^d}\Big|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\big)\Big(\frac{y^1}{a_{1,k_1}},\dots,\frac{y^N}{a_{N,k_N}}\Big)\Big|\prod_{i=1}^N\big(1+|y^i|^{d_i/s_i}\big)dy.
\end{aligned} \tag{2.3.20}$$

The last step follows from the triangle inequality in $\mathbb{R}^{d_i}$ and a substitution $x-z-y \longmapsto \big(\frac{y^1}{a_{1,k_1}},\dots,\frac{y^N}{a_{N,k_N}}\big)$. Furthermore, we find

$$\frac{1}{\prod_{i=1}^N a_{i,k_i}^{d_i}}\big(\mathcal{F}^{-1}\rho_{\overline{k}}\big)\Big(\frac{y^1}{a_{1,k_1}},\dots,\frac{y^N}{a_{N,k_N}}\Big) = \big(\mathcal{F}^{-1}\rho_{\overline{k}}(a_{1,k_1}\cdot,\dots,a_{N,k_N}\cdot)\big)(y).$$

Moreover, for $\xi \in \mathbb{R}^n$ and arbitrary $\kappa \in \mathbb{R}$ it always holds $1 + |\xi|^\kappa \sim (1+|\xi|^2)^{\kappa/2}$. Hence, we can apply Lemma 2.3.2 (with $p = 1$ and $\rho_i = \frac{d_i}{s_i}$) to the integral in (2.3.20). Together with (2.3.19) (for $f_{\overline{k}}$ replaced by $\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}$), we obtain

$$\big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)^*(x) \le c\, f^*_{\overline{k}}(x) \big\|\rho_{\overline{k}}(a_{1,k_1}\cdot, \ldots, a_{N,k_N}\cdot)\big|S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})\big\| \tag{2.3.21}$$

as long as

$$r_i > \frac{d_i}{s_i} + d_i\Big(1 - \frac{1}{2}\Big) = d_i\Big(\frac{1}{s_i} + \frac{1}{2}\Big)\,, \qquad i = 1, \ldots, N\,. \tag{2.3.22}$$

We additionally mention, that

$$\big|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)(x)\big| \le \big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)^*(x) \tag{2.3.23}$$

holds for all $x \in \mathbb{R}^d$. Under the assumption (2.3.18) we can find real numbers s_i with $0 < s_i < \min(p, q)$ and (2.3.22). From Proposition 2.3.4 (with s_i instead of r_i), together with (2.3.21) and (2.3.23), it finally follows

$$\begin{aligned}\big\|\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big|L_p(\ell_q)\big\| &\le \big\|\big(\mathcal{F}^{-1}\rho_{\overline{k}}\mathcal{F}f_{\overline{k}}\big)^*\big|L_p(\ell_q)\big\| \\ &\le c\,\big\|\, f^*_{\overline{k}}\,\big|L_p(\ell_q)\big\| \sup_{\overline{k}\in\mathbb{N}_0^N} \big\|\rho_{\overline{k}}(a_{1,k_1}\cdot, \ldots, a_{N,k_N}\cdot)\big|S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})\big\| \\ &\le C \sup_{\overline{k}\in\mathbb{N}_0^N} \big\|\rho_{\overline{k}}(a_{1,k_1}\cdot, \ldots, a_{N,k_N}\cdot)\big|S^{\overline{r}}_2 H(\mathbb{R}^{\overline{d}})\big\| \cdot \big\|\, f_{\overline{k}}\,\big|L_p(\ell_q)\big\|\,.\end{aligned}$$

This completes the proof. □

Proof of Theorem 2.3.1. We begin with the case of F-spaces. Hence, let $0 < p < \infty$. At first we remark, that

$$\varphi^i_j(t) = \big(\psi^i_{j-1}(t) + \psi^i_j(t) + \psi^i_{j+1}(t)\big)\varphi^i_j(t), \quad t \in \mathbb{R}^{d_i}, j = 0, 1, 2, \ldots,$$

holds with $\psi^i_{-1}(t) \equiv 0$. Therefore, it immediately follows

$$\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f = \sum_{\overline{l}\in\{-1,0,1\}^N} \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}\Big[\mathcal{F}^{-1}\psi_{\overline{k}+\overline{l}}\mathcal{F}f\Big]\,. \tag{2.3.24}$$

Therefore we put abbreviatory

$$f_{\overline{k}+\overline{l}} := \mathcal{F}^{-1}\psi_{\overline{k}+\overline{l}}\mathcal{F}f, \qquad \overline{k} \in \mathbb{N}_0^N, \overline{l} \in \{-1, 0, 1\}^N\,.$$

Then it follows from (2.3.24)

$$\big\|2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big|L_p(\ell_q)\big\| \le c \sum_{\overline{l}\in\{-1,0,1\}^N} \big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}\big[2^{\overline{k}\cdot\overline{r}}f_{\overline{k}+\overline{l}}\big]\big|L_p(\ell_q)\big\|\,. \tag{2.3.25}$$

We want to apply Proposition 2.3.5 to the right hand side of the last estimate (with $a_{i,k_i} = 2^{k_i+2}$). The properties of $\varphi^i \in \Phi(\mathbb{R}^{d_i})$, in particular (1.2.1) and (1.2.2), imply

$$\big\|\varphi^i_{k_i}(2^{k_i+2}\cdot)\big|W^m_2(\mathbb{R}^{d_i})\big\| \le c_i\,, \qquad i = 1, \ldots, N\,, \quad k_i \in \mathbb{N}_0\,,$$

where the constants c_i are independent of k_i. Thus we find

$$\sup_{\bar{k}\in\mathbb{N}_0^N} \big\| \varphi_{\bar{k}}(2^{k_1+2}\cdot, \ldots, 2^{k_N+2}\cdot) \big| S_2^{\overline{m}} W(\mathbb{R}^{\overline{d}}) \big\|$$

$$= \sup_{\bar{k}\in\mathbb{N}_0^N} \prod_{i=1}^{N} \big\| \varphi^i_{k_i}(2^{k_i+2}\cdot) \big| W_2^m(\mathbb{R}^{d_i}) \big\| \le \prod_{i=1}^{N} c_i < \infty \,. \tag{2.3.26}$$

Here we have used Remark 2.1.2. One needs to bear in mind that due to (1.2.1) we have $\operatorname{supp} \varphi^i_{k_i}(2^{k_i+2}\cdot) \subset \{t \in \mathbb{R}^{d_i} : |t| \le 1/2\}$, hence its Lebesgue measure can be estimated independent of k_i. Thus Proposition 2.3.5 is applicable because of Theorem 2.1.1. From (2.3.25) and (2.3.26) we obtain

$$\big\| 2^{\bar{k}\cdot\bar{r}} \mathcal{F}^{-1} \varphi_{\bar{k}} \mathcal{F} f \big| L_p(\ell_q) \big\| \le c \sum_{\bar{l}\in\{-1,0,1\}^N} \big\| 2^{\bar{k}\cdot\bar{r}} f_{\bar{k}+\bar{l}} \big| L_p(\ell_q) \big\| \,.$$

By Definition 2.2.1 this results in

$$\big\| f \,\big| S^{\bar{r}}_{p,q} F(\mathbb{R}^{\overline{d}}) \big\|_{\varphi} \le c \, \big\| f \,\big| S^{\bar{r}}_{p,q} F(\mathbb{R}^{\overline{d}}) \big\|_{\psi} \,.$$

Interchanging of φ and ψ finally yields the equivalence of the quasi-norms. The case of B-spaces can be treated analogously with Proposition 2.3.3 instead of Proposition 2.3.5. □

Remark 2.3.7. We have not defined the spaces $S^{\bar{r}}_{p,q} F(\mathbb{R}^{\overline{d}})$ for $p = \infty$. The reason is the same one as for the isotropic Triebel-Lizorkin spaces. Of course it would be possible to extend Definition 2.2.1 to $p = \infty$, but there would be no counterpart of Theorem 2.3.1. On the contrary, this spaces indeed do depend on the choice of the systems $\varphi^j \in \Phi(\mathbb{R}^{d_j})$. This can be seen by tensorizing the corresponding examples for the spaces $\widetilde{F}^{r_i}_{\infty,q}(\mathbb{R}^{d_i})$ and using the crossnorm property for $S^{\bar{r}}_{\infty,q} \widetilde{F}(\mathbb{R}^{\overline{d}})$ (the tilde refers to the usage of the "wrong" definition, i.e. the "wrong" quasi-norm).
M. Frazier and B. Jawerth defined in [29] the spaces $F^s_{\infty,q}$ for $s \in \mathbb{R}$ and all $0 < q \le \infty$ with the help of atomic representations. Though this could be carried over to our situation, this shall not be executed here. See also [84], Section 1.5.2, and the references given there.

2.3.3 A Nikol'skij-type inequality

For the further treatment of the function spaces of dominating mixed smoothness, in particular for the Besov-type spaces, we shall need an adapted version of the Nikol'skij inequality. Generally, it says

$$\big\| D^\alpha f \big| L_u(\mathbb{R}^d) \big\| \le c \, \big\| f \big| L_p(\mathbb{R}^d) \big\| \tag{2.3.27}$$

for all $f \in L_p^\Omega$, $\Omega \subset \mathbb{R}^d$, a multiindex $\alpha \in \mathbb{N}_0^d$ and $0 < p \le u \le \infty$. In this form, the inequality goes back to B. Stöckert [76] and A. P. Uninskij [93].
For special domains Ω more precise statements about the constant are possible. One obtains

$$\big\| D^\alpha f \big| L_u(\mathbb{R}^d) \big\| \le c \, b^{|\alpha|+d\left(\frac{1}{p}-\frac{1}{u}\right)} \big\| f \big| L_p(\mathbb{R}^d) \big\| \tag{2.3.28}$$

for $f \in L_p^{B_b}$, where $B_b = \big\{ x \in \mathbb{R}^d : |x| \le b \big\}$. For further details and references compare with [83], Sections 1.3.2 and 1.4.1. We shall seek an analogon of the inequality (2.3.28).

Proposition 2.3.6. Let $0 < p \le u \le \infty$, and let $\alpha \in \mathbb{N}_0^d$ be a multiindex. Moreover, let $\Omega = Q_{\overline{b}}$ be defined by

$$Q_{\overline{b}} = \left\{ x \in \mathbb{R}^d : |x^i| \le b_i, i = 1, \ldots, N \right\}. \tag{2.3.29}$$

Then there exists a constant $c > 0$, independent of $\overline{b}$, such that

$$\left\| D^\alpha f \middle| L_u(\mathbb{R}^d) \right\| \le c\, b_1^{|\alpha^1| + d_1 \left(\frac{1}{p} - \frac{1}{u}\right)} \cdots b_N^{|\alpha^N| + d_N \left(\frac{1}{p} - \frac{1}{u}\right)} \left\| f \middle| L_p(\mathbb{R}^d) \right\| \tag{2.3.30}$$

holds for all $f \in L_p^\Omega(\mathbb{R}^d)$.

Proof. The inequality (2.3.27) holds true for arbitrary compact sets $\Omega \subset \mathbb{R}^d$ and $f \in L_p^\Omega$. In particular, it covers the assertion (2.3.30) for the case $b_1 = \cdots = b_N = 1$ with some constant c_1. The general assertion then follows by a simple scaling argument
With the notation $\overline{1} = (1, \ldots, 1)$ it holds $f \in L_p^{Q_{\overline{b}}} \iff f(b_1^{-1}\cdot, \ldots, b_N^{-1}\cdot) \in L_p^{Q_{\overline{1}}}$. Hence, we obtain

$$\begin{aligned}
\left\| D^\alpha f \middle| L_u(\mathbb{R}^d) \right\| &= b_1^{-d_1/u} \cdots b_N^{-d_N/u} \left\| \left[D^\alpha f \right] (b_1^{-1}\cdot, \ldots, b_N^{-1}\cdot) \middle| L_u(\mathbb{R}^d) \right\| \\
&= b_1^{-d_1/u} \cdots b_N^{-d_N/u} \left\| b_1^{|\alpha^1|} \cdots b_N^{|\alpha^N|} D^\alpha \left[f(b_1^{-1}\cdot, \ldots, b_N^{-1}\cdot) \right] \middle| L_u(\mathbb{R}^d) \right\| \\
&= b_1^{|\alpha^1| - d_1/u} \cdots b_N^{|\alpha^N| - d_N/u} \left\| D^\alpha \left[f(b_1^{-1}\cdot, \ldots, b_N^{-1}\cdot) \right] \middle| L_u(\mathbb{R}^d) \right\| \\
&\le c_1 b_1^{|\alpha^1| - d_1/u} \cdots b_N^{|\alpha^N| - d_N/u} \left\| f(b_1^{-1}\cdot, \ldots, b_N^{-1}\cdot) \middle| L_p(\mathbb{R}^d) \right\| \\
&= c_1 b_1^{|\alpha^1| - d_1/u} \cdots b_N^{|\alpha^N| - d_N/u} b_1^{d_1/p} \cdots b_N^{d_N/p} \left\| f \middle| L_p(\mathbb{R}^d) \right\| \\
&= c_1 b_1^{|\alpha^1| + d_1(\frac{1}{p} - \frac{1}{u})} \cdots b_N^{|\alpha^N| + d_N(\frac{1}{p} - \frac{1}{u})} \left\| f \middle| L_p(\mathbb{R}^d) \right\|.
\end{aligned}$$

This proves the assertion. □

2.3.4 Elementary embeddings

We now return to the investigation of the basic properties of the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. After having shown the independency of the decomposition of unity, we now consider some elementary embeddings of these function spaces. Furthermore, we show their completeness.

Again we put abbreviatory $f_{\overline{k}} = \mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} f$ for $f \in \mathcal{S}'(\mathbb{R}^d)$, where $\left(\varphi_{\overline{k}} \right)_{\overline{k} \in \mathbb{N}_0^N}$ is some fixed decomposition of unity according to Definition 1.2.2 and equation (2.2.1).

Proposition 2.3.7. Let $0 < p, q_0, q_1, q_2 \le \infty$ ($p < \infty$ for F-spaces), and let $\overline{r}, \overline{t} \in \mathbb{R}^N$. Then it holds for $q_0 \le q_1$ and $\overline{t} > 0$

$$S^{\overline{r}}_{p,q_0}A(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}}_{p,q_1}A(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}-\overline{t}}_{p,q_2}A(\mathbb{R}^{\overline{d}}). \tag{2.3.31}$$

Moreover, for $0 < p < \infty$ and every $0 < q \le \infty$ we find

$$S^{\overline{r}}_{p,\min(p,q)}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}}_{p,\max(p,q)}B(\mathbb{R}^{\overline{d}}). \tag{2.3.32}$$

Proof. **Step 1:** The left hand embedding immediately follows from the monotonicity of the ℓ_q-quasi-norms, because for $q_0 \le q_1$ it holds $\left\| (c_j)_{j \in \mathbb{N}_0} \middle| \ell_{q_1} \right\| \le \left\| (c_j)_{j \in \mathbb{N}_0} \middle| \ell_{q_0} \right\|$ for an arbitrary sequence $(c_j)_{j \in \mathbb{N}_0} \in \ell_{q_0}$.

By the same argument for the right hand embedding it suffices to prove

$$S^{\overline{r}}_{p,\infty}A(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}-\overline{t}}_{p,q_2}A(\mathbb{R}^{\overline{d}}) ,$$

where $q_2 < \infty$. We start with the B-case. For $f \in S^{\overline{r}}_{p,\infty}B(\mathbb{R}^{\overline{d}})$ we get

$$\begin{aligned}\big\|f\big|S^{\overline{r}-\overline{t}}_{p,q_2}B(\mathbb{R}^{\overline{d}})\big\| &= \Bigg(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{t})q_2}\big\|f_{\overline{\nu}}\big|L_p(\mathbb{R}^d)\big\|^{q_2}\Bigg)^{1/q_2}\\ &\le \Bigg(\sup_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q_2}\big\|f_{\overline{k}}\big|L_p(\mathbb{R}^d)\big\|^{q_2}\Bigg(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot\overline{t}q_2}\Bigg)\Bigg)^{1/q_2}\\ &= c\sup_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}}\big\| f_{\overline{k}}\,\big|L_p(\mathbb{R}^d)\big\| = c\,\big\|\, f\,\big|S^{\overline{r}}_{p,\infty}B(\mathbb{R}^{\overline{d}})\big\| .\end{aligned}$$

The F-spaces can be treated analogously. We obtain for $f \in S^{\overline{r}}_{p,\infty}F(\mathbb{R}^{\overline{d}})$

$$\begin{aligned}\big\|f\big|S^{\overline{r}-\overline{t}}_{p,q_2}F(\mathbb{R}^{\overline{d}})\big\| &= \Bigg\|\Bigg(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{t})q_2}\big|f_{\overline{\nu}}(\cdot)\big|^{q_2}\Bigg)^{1/q_2}\Bigg|L_p(\mathbb{R}^d)\Bigg\|\\ &\le \Bigg\|\Bigg(\sup_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q_2}\big|f_{\overline{k}}(\cdot)\big|^{q_2}\Bigg(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot\overline{t}q_2}\Bigg)\Bigg)^{1/q_2}\Bigg|L_p(\mathbb{R}^d)\Bigg\|\\ &= c\,\Bigg\|\sup_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}}\big|f_{\overline{k}}(\cdot)\big|\Bigg|L_p(\mathbb{R}^d)\Bigg\| = c\,\big\|\, f\,\big|S^{\overline{r}}_{p,\infty}F(\mathbb{R}^{\overline{d}})\big\| .\end{aligned}$$

Both times we used, that the series $\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot\overline{t}q_2}$ is a convergent geometric series due to $\overline{t} > 0$, whose value depends on $\overline{t}$ and q_2 only.

Step 2: To prove (2.3.32), it is sufficient to consider the situation for $L_p(\ell_q)$ and $\ell_q(L_p)$, see Remark 2.2.1. The monotonicity of the ℓ_q-quasi-norms immediately yields the left hand embedding in case $p \le q$ and the right hand embedding in case $q \le p$. The other two embeddings follow from Minkowski's inequality (Theorem 2.3.2). Exemplary we show the left hand embedding for the case $q \le p$. Let $f = (f_j)_{j\in A} \in \ell_q(L_p)$. Then we obtain due to $p/q \ge 1$

$$\begin{aligned}\big\|\, f\,\big|L_p(\ell_q)\big\| &= \Bigg\|\sum_{j\in A} |f_j|^q\Bigg|L_{p/q}(\mathbb{R}^d)\Bigg\|^{1/q} \le \Bigg(\sum_{j\in A}\big\|\,|f_j|^q\,\big|L_{p/q}(\mathbb{R}^d)\big\|\Bigg)^{1/q}\\ &= \Bigg(\sum_{j\in A}\big\|\, f_j\,\big|L_p(\mathbb{R}^d)\big\|^q\Bigg)^{1/q} = \big\|\, f\,\big|\ell_q(L_p)\big\| .\end{aligned}$$

□

The forerunners of the next theorem can be found in [83, Theorem 2.3.3] and [71, Theorem 2.2.4(ii)]. It describes some basic topological properties of the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. In the sequel it will be frequently used, mostly without explicitly mentioning it.

Theorem 2.3.3. Let $\overline{r} \in \mathbb{R}^N$ and $0 < p, q \le \infty$ ($p < \infty$ for F-spaces). Then it holds

$$\mathcal{S}(\mathbb{R}^d) \hookrightarrow S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \hookrightarrow \mathcal{S}'(\mathbb{R}^d) \tag{2.3.33}$$

in the sense of continuous topological embeddings. If we additionally have $\max(p,q) < \infty$, then $\mathcal{S}(\mathbb{R}^d)$ is dense in $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$.

While these embeddings could be proved in complete analogy to the above mentioned references, we rely on the following proposition, which establishes a connection of the spaces of dominating mixed smoothness to their isotropic counterparts.

Proposition 2.3.8.

(i) Let $1 < p < \infty$ and $\overline{m} \in \mathbb{N}_0^N$. Then it holds

$$W^K_p(\mathbb{R}^d) \hookrightarrow S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}}) \hookrightarrow W^M_p(\mathbb{R}^d)\,,$$

where $K, M \in \mathbb{N}_0$ satisfy $M \leq \min_{i=1,\dots,N} m_i$ and $K \geq \sum_{i=1}^N m_i$.

(ii) Let $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces) and $\overline{r} \in \mathbb{R}^N$. Then we have

$$A^s_{p,q}(\mathbb{R}^d) \hookrightarrow S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \hookrightarrow A^t_{p,q}(\mathbb{R}^d)\,,$$

where $s, t \in \mathbb{R}$ satisfy

$$t < \max\big(0, \min(r_1, \dots, r_N)\big) + \sum_{i=1}^N \min\big(0, r_i\big)$$

and

$$s > \min\big(0, \max(r_1, \dots, r_N)\big) + \sum_{i=1}^N \max\big(0, r_i\big)\,.$$

Remark 2.3.8. The above conditions for s and t are almost optimal as can be seen when comparing the case $\overline{r} \geq 0$ with the conditions on K and M in part (i) (we remind on $F^m_{p,2}(\mathbb{R}^d) = W^m_p(\mathbb{R}^d)$ and $S^{\overline{m}}_{p,2}F(\mathbb{R}^{\overline{d}}) = S^{\overline{m}}_p W(\mathbb{R}^{\overline{d}})$ for $1 < p < \infty$ and $m \in \mathbb{N}_0$, $\overline{m} \in \mathbb{N}_0^N$).

Proof. The embeddings in (i) are immediate consequences of the definitions of the corresponding norms, i.e. of the partial derivatives involved. Hence we may concentrate on (ii). We fix some decompositions of unity $\psi \in \Phi(\mathbb{R}^d)$ and $\varphi^i \in \Phi(\mathbb{R}^{d_i})$, $i = 1, \dots, N$, constructed as in Remark 1.2.2, and define $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k} \in \mathbb{N}_0^N}$ as in (2.2.1).

Step 1: At first, we shall be concerned with the support of products $\psi_j \varphi_{\overline{k}}$. To this purpose, define for every $j \in \mathbb{N}_0$ sets

$$\Lambda_j = \big\{\overline{k} \in \mathbb{N}_0^N \,:\, \operatorname{supp} \varphi_{\overline{k}} \cap \operatorname{supp} \psi_j \neq \emptyset\big\}\,,$$

and conversely, for every $\overline{k} \in \mathbb{N}_0^N$ we define

$$\Lambda_{\overline{k}} = \big\{j \in \mathbb{N}_0 \,:\, \operatorname{supp} \psi_j \cap \operatorname{supp} \varphi_{\overline{k}} \neq \emptyset\big\}\,.$$

Immediate consequences of these definitions and the properties (1.2.3) and (2.2.1), respectively, are the pointwise identities

$$\psi_j(x) = \sum_{\overline{k}\in\Lambda_j} \varphi_{\overline{k}}(x)\psi_j(x) \quad \text{and} \quad \varphi_{\overline{k}}(x) = \sum_{j\in\Lambda_{\overline{k}}} \psi_j(x)\varphi_{\overline{k}}(x)\,, \tag{2.3.34}$$

$j \in \mathbb{N}_0$, $\overline{k} \in \mathbb{N}_0^N$, $x \in \mathbb{R}^d$. Next, we want to estimate the count of the elements of the above index sets. To begin with let $x \in \operatorname{supp} \varphi_{\overline{k}}$. Then we have $|x^i| \le 2^{k_i+1}$, and hence

$$|x|^2 = \sum_{i=1}^N |x^i|^2 \le \sum_{i=1}^N 2^{2(k_i+1)} \le N 2^{2(1+\max k_i)} \le 2^{2(1+\eta+\max k_i)}\,,$$

where $\eta \in \mathbb{N}$ is the smallest number, such that $\eta \ge \log_2 \sqrt{N}$. Altogether, since $x \in \operatorname{supp} \psi_j$ implies $|x| \ge 2^{j-1}$, we find $j \le \max k_i + \eta + 2$ for every $j \in \Lambda_{\overline{k}}$. On the other hand, for every $i \in \{1, \ldots, N\}$ we obtain from $|x^i| \ge 2^{k_i-1}$ the condition $j \ge k_i - 2$, since otherwise we have the contradiction

$$x \in \operatorname{supp} \psi_j \Longrightarrow |x| \le 2^{j+1} < 2^{k_i-1} \le |x^i| \le |x|\,.$$

This implies $j \ge \max k_i - 2$.
Now let $x \in \operatorname{supp} \psi_j$. Then the property $|x| \ge 2^{j-1}$ implies that for some $i \in \{1, \ldots, N\}$ we have $|x^i| \ge \frac{1}{\sqrt{N}} 2^{j-1} \ge 2^{j-\eta-1}$. Thus we find the condition $j - \eta - 2 < k_i$ for that i, which in particular is fulfilled, if $j - \eta - 2 \le \max k_i$. Finally, $|x| \le 2^{j+1}$ yields $|x^i| \le 2^{j+1}$, thus the condition $j \ge \max k_i - 2$ is necessary for $x \in \operatorname{supp} \varphi_{\overline{k}}$.
Altogether we have found, that for $\operatorname{supp} \psi_j \cap \operatorname{supp} \varphi_{\overline{k}} \ne \emptyset$ the condition

$$\max k_i - 2 \le j \le \max k_i + \eta + 2 \tag{2.3.35}$$

is necessary. This implies

$$\#\Lambda_{\overline{k}} \le \eta + 5 \sim 1 \quad \text{and} \quad \#\Lambda_j \le (j+2)^N - (j-\eta-2)^N \sim j^{N-1}\,. \tag{2.3.36}$$

Step 2:
We now proceed similarly to the proof of Theorem 2.3.1. This time we obtain from the identities (2.3.34)

$$\mathcal{F}^{-1}\psi_j\mathcal{F}f = \sum_{\overline{k}\in\Lambda_j} \mathcal{F}^{-1}\varphi_{\overline{k}}\psi_j\mathcal{F}f \qquad \text{and} \qquad \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f = \sum_{j\in\Lambda_{\overline{k}}} \mathcal{F}^{-1}\psi_j\varphi_{\overline{k}}\mathcal{F}f\,,$$

and hence (appropriate modifications in case $q = \infty$)

$$\begin{aligned}\big\| f \,\big|B^t_{p,q}(\mathbb{R}^d)\big\|^q &= \sum_{j=0}^\infty 2^{jtq} \Big\| \sum_{\overline{k}\in\Lambda_j} \mathcal{F}^{-1}\varphi_{\overline{k}}\psi_j\mathcal{F}f \Big| L_p(\mathbb{R}^d)\Big\|^q \\ &\le \sum_{j=0}^\infty 2^{jtq} \Big(\sum_{\overline{k}\in\Lambda_j} \big\|\mathcal{F}^{-1}\psi_j\mathcal{F}\big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)\big|L_p(\mathbb{R}^d)\big\|^u\Big)^{q/u}\,,\end{aligned}$$

where $u = \min(1,p)$. Similarly, it follows

$$\| f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\|^q \le \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\Big(\sum_{j\in\Lambda_{\overline{k}}} \|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}\big(\mathcal{F}^{-1}\psi_j\mathcal{F}f\big)\,|L_p(\mathbb{R}^d)\|^u\Big)^{q/u}.$$

Hence we have to consider the Fourier multiplier properties of the functions ψ_j and $\varphi_{\overline{k}}$ next.

Step 3: We shall use Corollary 2.3.1 due to the support properties of ψ_j and $\varphi_{\overline{k}}$. To simplify notation we will restrict ourselves to the case $j \ge 1$ and $\overline{k} \ge \overline{1}$, the other cases can be treated analogously.
Keeping in mind the condition (2.3.35), we find $\operatorname{supp}\psi_j(2^{\overline{k}-\overline{1}}\,\cdot) \subset \Gamma^0$ for some compact set Γ^0, which is independent of j and $\overline{k}$. By the same arguments as in the proof of Corollary 2.3.1 we find

$$\big\|\mathcal{F}^{-1}\psi_j\mathcal{F}\big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)\,|L_p(\mathbb{R}^d)\big\| \le c\,2^{(\overline{k}-\overline{1})\cdot\overline{d}(\frac{1}{u}-1)}\big\|\mathcal{F}^{-1}\psi_j\,|L_u(\mathbb{R}^d)\big\|\cdot\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\,|L_p(\mathbb{R}^d)\big\|.$$

Similarly, we obtain

$$\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}\big(\mathcal{F}^{-1}\psi_j\mathcal{F}f\big)\,|L_p(\mathbb{R}^d)\big\| \le c\,2^{(j-1)d(\frac{1}{u}-1)}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\,|L_u(\mathbb{R}^d)\big\|\cdot\big\|\mathcal{F}^{-1}\psi_j\mathcal{F}f\,|L_p(\mathbb{R}^d)\big\|,$$

since $\operatorname{supp}\varphi_{\overline{k}}(2^{j-1}\,\cdot) \subset \Omega^0$ for some compact set Ω^0.

Step 4: We now calculate the L_u-quasi-norms of the functions $\mathcal{F}^{-1}\psi_j$ and $\mathcal{F}^{-1}\varphi_{\overline{k}}$. Here we restrict ourselves to the cases $j \ge 1$ and $\overline{k} \ge \overline{1}$, all the other cases yield similar estimates. By straightforward calculations, we obtain

$$\begin{aligned}\big\|\mathcal{F}^{-1}\psi_j\,|L_u(\mathbb{R}^d)\big\| &= \big\|\mathcal{F}\big(\psi_1(2^{-j+1}\,\cdot)\big)\,|L_u(\mathbb{R}^d)\big\| = 2^{(j-1)d}\big\|(\mathcal{F}\psi_1)(2^{j-1}\,\cdot)\,|L_u(\mathbb{R}^d)\big\|\\ &= 2^{(j-1)d}2^{(-j+1)d/u}\big\|\,\mathcal{F}\psi_1\,|L_u(\mathbb{R}^d)\big\| = c_\psi 2^{-(j-1)d(1/u-1)}.\end{aligned}$$

The quasi-norm is finite, since $\mathcal{F}\psi_1 \in \mathcal{S}(\mathbb{R}^d)$. Using the tensor product properties of the Fourier transform and the crossnorm-property of the L_u-quasi-norm, we further find

$$\begin{aligned}\big\|\mathcal{F}_d^{-1}\varphi_{\overline{k}}\,|L_u(\mathbb{R}^d)\big\| &= \prod_{i=1}^N \big\|\mathcal{F}_{d_i}\varphi^i_{k_i}\,|L_u(\mathbb{R}^{d_i})\big\| = \prod_{i=1}^N 2^{(k_i-1)d_i}2^{(-k_i+1)d_i/u}\big\|\,\mathcal{F}_{d_i}\varphi^i_1\,|L_u(\mathbb{R}^{d_i})\big\|\\ &= 2^{-(\overline{k}-\overline{1})\cdot\overline{d}(1/u-1)}\big\|\,\mathcal{F}_d\varphi_{\overline{1}}\,|L_u(\mathbb{R}^d)\big\| = c_\varphi 2^{-(\overline{k}-\overline{1})\cdot\overline{d}(1/u-1)}.\end{aligned}$$

Recalling the results of Step 3, applying Lemma 2.3.3 now results in

$$\begin{aligned}\| f\,|B^t_{p,q}(\mathbb{R}^d)\|^q &\le c\sum_{j=0}^\infty 2^{jtq}\Big(\sum_{\overline{k}\in\Lambda_j}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\,|L_p(\mathbb{R}^d)\big\|^u\Big)^{q/u}\\ &\le c'\sum_{\overline{k}\in\mathbb{N}_0^N}\sum_{j\in\Lambda_{\overline{k}}} 2^{jtq}j^{(N-1)v}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\,|L_p(\mathbb{R}^d)\big\|^q\\ &\le c''\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\,|L_p(\mathbb{R}^d)\big\|^q,\end{aligned}$$

where $v = (1-q/u)_+$ (use Hölder's inequality or the monotonicity of ℓ_p-quasi-norms, respectively). Furthermore, we used (2.3.35), hence $2^j \sim 2^{\max k_i}$ for all pairs $(j,\overline{k})$ under

consideration, as well as (2.3.36). Eventually, the assumption on t implies $(t+\varepsilon)\max k_i \le \overline{k}\cdot\overline{r}$ for some sufficiently small $\varepsilon > 0$. This proves the right-hand-embedding in (ii).
For the left-hand-embedding we obtain analogously

$$\begin{aligned}\big\| f \,\big| S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|^q &\le c \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\bigg(\sum_{j\in\Lambda_{\overline{k}}} \big\|\mathcal{F}^{-1}\psi_j\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^u\bigg)^{q/u} \\ &\le c' \sum_{j=0}^{\infty}\sum_{\overline{k}\in\Lambda_j} 2^{\overline{k}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\psi_j\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q \\ &\le c'' \sum_{j=0}^{\infty} 2^{jsq}\big\|\mathcal{F}^{-1}\psi_j\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q\,,\end{aligned}$$

where we used $\#\Lambda_{\overline{k}} \sim 1$ and $\#\Lambda_j \sim j^{N-1}$ as well as the estimate $\overline{k}\cdot\overline{r} \le (s-\varepsilon)\max k_i$ for some $\varepsilon > 0$, which is implied by the condition on s.

Step 5: The corresponding embeddings for the F-scale follow immediately from the embeddings in Proposition 2.3.7. □

Proof of Theorem 2.3.3. The embeddings (2.3.33) are immediate consequences of Proposition 2.3.8 and the corresponding embeddings for the isotropic spaces in [83, Theorem 2.3.3]. We shall only mention, that the right hand embedding in (2.3.33) means an inequality of the form

$$\big|f(\psi)\big| \le c\,\big\| f \,\big| S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|_{\varphi}\big\| \psi \,\big\|_{n_1,n_2}\,, \qquad \psi\in\mathcal{S}(\mathbb{R}^d)\,, \tag{2.3.37}$$

for suitable $n_1, n_2 \in \mathbb{N}_0$. Hence, it only remains to prove the density assertion.

Step 1: At first, let $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$, $\max(p,q) < \infty$. We put for $n \in \mathbb{N}_0$

$$f_n := \sum_{\overline{k}\le n} \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\,.$$

Let $\big(\psi_{\overline{\nu}}\big)_{\overline{\nu}\in\mathbb{N}_0^N}$ be an arbitrary further decomposition of unity. Then by Lemma 2.3.3, applied with $\widetilde{p} = \min(1,p)$, it follows

$$\big\|\mathcal{F}^{-1}\psi_{\overline{\nu}}\mathcal{F}\big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)\big|L_p(\mathbb{R}^d)\big\| \le c\,\big\|\mathcal{F}^{-1}\psi_{\overline{\nu}}\big|L_{\widetilde{p}}(\mathbb{R}^d)\big\|\cdot\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|\,,$$

and hence $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$ as well. One has to take into account, that because of property (1.2.1) only the (finitely many) terms with $\overline{\nu} = \overline{k} + \overline{l}$ for $\overline{l} \in \{-1,0,1\}^N$ are of relevance. Again due to (1.2.1) and (1.2.3) it holds $f_n \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \cap L_p^{\Omega}$, where $\Omega = \big\{x\in\mathbb{R}^d : |x^i| \le 2^{n+1}, i = 1,\ldots,N\big\}$. We now show

$$\big\|f - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \longrightarrow 0 \qquad \text{for} \quad n\longrightarrow\infty\,. \tag{2.3.38}$$

In order to see this, we observe that due to (2.2.2) we have $f = \sum_{\overline{k}\in\mathbb{N}_0^N}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ with convergence in $\mathcal{S}'(\mathbb{R}^d)$. Hence we obtain

$$f - f_n = \sum_{\overline{k}\in\mathbb{N}_0^N}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f - \sum_{\overline{k}\le n}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f = \sum_{\overline{k}\not\le n}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$$

in the sense of $\mathcal{S}'(\mathbb{R}^d)$. With $\varphi_{\overline{\nu}}\varphi_{\overline{k}} \not\equiv 0 \iff \overline{\nu} - \overline{k} \in \{-1,0,1\}^N$ in mind, we find

$$
\begin{aligned}
\big\|f - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| &= \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot\overline{r}q}\left\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\sum_{\overline{k}\nleq n}\varphi_{\overline{k}}\mathcal{F}f\right|L_p(\mathbb{R}^d)\right\|^q\Bigg)^{1/q} \\
&\leq c\left(\sum_{\overline{\nu}\nleq n-1}\ \sum_{\overline{l}\in\{-1,0,1\}^N} 2^{\overline{\nu}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\varphi_{\overline{\nu}+\overline{l}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q\right)^{1/q} \\
&\leq c'\left(\sum_{\overline{\nu}\nleq n-1} 2^{\overline{\nu}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q\right)^{1/q},
\end{aligned}
$$

where the last estimate follows as in the proof of Theorem 2.3.1 from Proposition 2.3.3, Theorem 2.1.1 and (1.2.2). Altogether we have shown

$$
\begin{aligned}
&\big\|f - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \\
&\quad\leq c\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q - \sum_{\overline{\nu}\leq n-1} 2^{\overline{\nu}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|^q\right)^{1/q}.
\end{aligned}
$$

From $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$ we obtain, that the right hand side of the last inequality converges to 0 as $n \longrightarrow \infty$. Thus it finally follows (2.3.38). For F-spaces one uses an analogous argument with Proposition 2.3.5 instead of Proposition 2.3.3, and eventually Lebesgue's theorem on dominated convergence.

Step 2: We construct an approximation $g \in \mathcal{S}(\mathbb{R}^d)$ of f_n in the space L_p^Γ, where $\Gamma = \big\{x \in \mathbb{R}^d : |x^i| \leq 2^{n+2},\, i = 1,\dots,N\big\}$. To this purpose, let $\psi \in \mathcal{S}(\mathbb{R}^d)$ with $\psi(0) = 1$ and $\operatorname{supp}\mathcal{F}\psi \subset \{x \in \mathbb{R}^d : |x^i| \leq 1,\, i = 1,\dots,N\}$. Then it holds $\psi(\varepsilon\cdot)f_n \in L_p^\Gamma \cap \mathcal{S}(\mathbb{R}^d)$ for an arbitrary $0 < \varepsilon < 1$, because f_n is an infinitely often differentiable and bounded function, whose every partial derivative is bounded as well, and for the support of the Fourier transform of $\psi(\varepsilon\cdot)f_n$ we find

$$
\begin{aligned}
\operatorname{supp}\mathcal{F}\big[\psi(\varepsilon\cdot)f_n\big] &\subset \operatorname{supp}\mathcal{F}\big[\psi(\varepsilon\cdot)\big] + \operatorname{supp}\mathcal{F}f_n = \operatorname{supp}(\mathcal{F}\psi)(\tfrac{\cdot}{\varepsilon}) + \operatorname{supp}\mathcal{F}f_n \\
&\subset \{x \in \mathbb{R}^d : |x^i| \leq \varepsilon + 2^{n+1},\, i = 1,\dots,N\} \subset \Gamma\,.
\end{aligned}
$$

The boundedness of f_n is a consequence of the boundedness of $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ for all $\overline{k} \in \mathbb{N}_0^N$, which in turn follows from Nikol'skij's inequality. More precisely, an application of Proposition 2.3.6 with $u = \infty$, $\alpha = 0$, and $b_i = 2^{k_i+1}$ (we remind of property (1.2.1)) yields

$$
\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big|L_\infty(\mathbb{R}^d)\big\| \leq c\,2^{\overline{k}\cdot\overline{d}/p}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big|L_p(\mathbb{R}^d)\big\|\,,
$$

which is finite, since $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \in L_p(\mathbb{R}^d)$ due to $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. This argument immediately extends to the partial derivatives of f_n. Moreover, for every $x \in \mathbb{R}^d$ we get

$$
\psi(\varepsilon x)f_n(x) \longrightarrow f_n(x), \qquad \varepsilon \longrightarrow 0\,,
$$

and $|\psi(\varepsilon x)f_n(x)| \leq |\psi(\varepsilon x)|\big\|f_n\big|L_\infty(\mathbb{R}^d)\big\|$. Since $\psi(\varepsilon\cdot)$ is integrable due to $\psi \in \mathcal{S}(\mathbb{R}^d)$, we conclude from Lebesgue's theorem on dominated convergence

$$
\big\|f_n - \psi(\varepsilon\,\cdot)f_n\big|L_p(\mathbb{R}^d)\big\| \longrightarrow 0, \qquad \varepsilon \longrightarrow 0\,.
$$

Hence, for all $\delta > 0$ there exists a function $g \in \mathcal{S}(\mathbb{R}^d)$, such that $\big\|g - f_n\big|L_p(\mathbb{R}^d)\big\| \le \delta$.

Step 3: Now, let an arbitrary $\varepsilon > 0$ be given. By Step 1 there exists an $n_0 \in \mathbb{N}$, such that $\big\|f - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \le \frac{\varepsilon}{2c_1}$ holds true for all $n \ge n_0$. Here c_1 is the constant from the quasi-triangle inequality in $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. We fix such an n, say n_1. Moreover, let $g \in \mathcal{S}(\mathbb{R}^d) \cap L_p^\Gamma$ be as in Step 2, such that $\big\|g - f_n\big|L_p(\mathbb{R}^d)\big\| \le \delta$. Then it follows from Lemma 2.3.3 with $\widetilde{p} = \min(1, p)$

$$\begin{aligned}
&\big\|g - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \\
&\quad = \Bigg(\sum_{\overline{\nu} \le n+2} 2^{\overline{\nu}\cdot\overline{r}q}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\mathcal{F}(g - f_n)\big|L_p(\mathbb{R}^d)\big\|^q\Bigg)^{1/q} \\
&\quad \le \Bigg(\sum_{\overline{\nu} \le n+2} 2^{\overline{\nu}\cdot\overline{r}q}c_{\overline{\nu}}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\big|L_{\widetilde{p}}(\mathbb{R}^d)\big\|^q\big\|g - f_n\big|L_p(\mathbb{R}^d)\big\|^q\Bigg)^{1/q} \\
&\quad \le \delta\Bigg(\sum_{\overline{\nu} \le n+2} 2^{\overline{\nu}\cdot\overline{r}q}c_{\overline{\nu}}\big\|\mathcal{F}^{-1}\varphi_{\overline{\nu}}\big|L_{\widetilde{p}}(\mathbb{R}^d)\big\|^q\Bigg)^{1/q} =: c_2\delta
\end{aligned}$$

for some constant $c_2 = c_2(n_1)$ independent of δ (because of $\varphi_{\overline{\nu}} \in \mathcal{S}(\mathbb{R}^d)$ we have $\mathcal{F}^{-1}\varphi_{\overline{\nu}} \in \mathcal{S}(\mathbb{R}^d) \subset L_{\widetilde{p}}(\mathbb{R}^d)$). With $\delta \le \frac{\varepsilon}{2c_1c_2}$ we finally obtain

$$\begin{aligned}
\big\|g - f\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| &\le c_1\Big(\big\|g - f_n\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| + \big\|f_n - f\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|\Big) \\
&\le c_1\Big(\delta c_2 + \frac{\varepsilon}{2c_1}\Big) \le \frac{\varepsilon}{2c_1c_2}\,c_1c_2 + \frac{\varepsilon}{2} = \varepsilon\,.
\end{aligned}$$

Thus, we have proved the density of $\mathcal{S}(\mathbb{R}^d)$ in $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. In case of F-spaces one proceeds analogously. □

Remark 2.3.9. Indeed there exist functions $\varphi \in \mathcal{S}(\mathbb{R}^d)$ with the properties required in Step 2, $\varphi(0) = 1$ and $\operatorname{supp}\mathcal{F}\varphi \subset \big\{y \in \mathbb{R}^d : |y| \le 1\big\}$. To this purpose let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be some real-valued function with $\psi(x) > 0$ for $|x| < 1$ and $\psi(x) = 0$ for $|x| \ge 1$. Then it holds for $\varphi := \mathcal{F}^{-1}\psi$

$$\varphi(0) = \big(\mathcal{F}^{-1}\psi\big)(0) = (2\pi)^{-d/2}\int_{\mathbb{R}^d}\psi(x)dx = (2\pi)^{-d/2}\int_{|x|<1}\psi(x)dx > 0\,.$$

A renormalization of φ yields a function with the desired properties.

We finish these considerations of basic properties with the next proposition. It is based on [83, Theorem 2.3.3].

Proposition 2.3.9. Let $0 < p, q \le \infty$ ($p < \infty$ for F-spaces) and $\overline{r} \in \mathbb{R}^N$. Then the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ are quasi-Banach spaces, and they are Banach spaces, if $\min(p, q) \ge 1$.

Proof. In the previous considerations we already used that the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ are quasi-normed. This can widely be derived from Remark 2.2.1 and the quasi-norm properties of $\|\cdot|L_p(\ell_q)\|$ and $\|\cdot|\ell_q(L_p)\|$. Here we will only show the property $\big\|\,f\,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| = 0 \Longrightarrow f = 0$.

To this purpose let $\left(\varphi_{\overline{k}}\right)_{\overline{k}\in\mathbb{N}_0^N}$ be an arbitrary decomposition of unity according to Definition 1.2.2 and (2.2.1). Furthermore, we suppose $\| f \,|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\| = 0$. By Definition 2.2.1 this implies at first

$$\sum_{\overline{k}\in\mathbb{N}_0^N} \left|2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x)\right|^q = 0 \quad \text{for almost every } x\in\mathbb{R}^d\,.$$

But this clearly yields, that for all $\overline{k}\in\mathbb{N}_0^N$ it holds $2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x) = 0$ almost everywhere. Hence we have found, that the (regular) distribution $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ is nothing else than the Null-distribution in $\mathcal{S}'(\mathbb{R}^d)$ for all $\overline{k}\in\mathbb{N}_0^N$. Then we obtain with the help of (2.2.2)

$$\mathcal{F}^{-1}\mathcal{F}f = \mathcal{F}^{-1}\sum_{\overline{k}\in\mathbb{N}_0^N}\varphi_{\overline{k}}\mathcal{F}f = \sum_{\overline{k}\in\mathbb{N}_0^N}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f = 0 \quad \text{ in } \mathcal{S}'(\mathbb{R}^d)\,.$$

So we have shown $f=0$ in the sense of $\mathcal{S}'(\mathbb{R}^d)$, and thus it holds $f=0$ in $S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})$.
Analogously from the assumption $\| f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\| = 0$ we obtain $\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\middle|L_p(\mathbb{R}^d)\right\| = 0$ for all $\overline{k}\in\mathbb{N}_0^N$. But this clearly is equivalent to $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x) = 0$ almost everywhere for all $\overline{k}\in\mathbb{N}_0^N$. Now we can conclude as before $f=0$ in the sense of $\mathcal{S}'(\mathbb{R}^d)$ and hence $f=0$ in $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$.

It remains the proof of the completeness. At first we remark, that $\mathcal{S}'(\mathbb{R}^d)$ is complete, as it is the dual space of the Fréchet space $\mathcal{S}(\mathbb{R}^d)$, and being equipped with the strong topology. Let $\left(f_l\right)_{l=1}^\infty$ be a Cauchy sequence in $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. By (2.3.33) and (2.3.37), respectively, this sequence is a Cauchy sequence in $\mathcal{S}'(\mathbb{R}^d)$ as well. Due to the completeness of $\mathcal{S}'(\mathbb{R}^d)$ we find a limit element $f\in\mathcal{S}'(\mathbb{R}^d)$. Hence, also $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f_l$ converges to $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ in $\mathcal{S}'(\mathbb{R}^d)$ for $l\longrightarrow\infty$ for every $\overline{k}\in\mathbb{N}_0^N$, as $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous transformations on $\mathcal{S}'(\mathbb{R}^d)$, as well as multiplication with $\varphi_{\overline{k}}$. On the other hand, because of

$$\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\middle|L_p(\mathbb{R}^d)\right\| \le 2^{-\overline{k}\cdot\overline{r}}\| g\,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\| \qquad \text{for all } g\in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\,,$$

the sequence $\left\{\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f_l\right\}_{l=1}^\infty$ is a Cauchy sequence in $L_p(\mathbb{R}^d)$, and by Proposition 2.3.6 in $L_\infty(\mathbb{R}^d)$ as well. Likewise by Proposition 2.3.6 follows, that if $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f_l\longrightarrow g_{\overline{k}}$ for $l\longrightarrow\infty$ in $L_p(\mathbb{R}^d)$, this convergence holds in $L_\infty(\mathbb{R}^d)$ as well. As for arbitrary regular distribution $h_1,h_2\in\mathcal{S}'(\mathbb{R}^d)$ and a test function $\psi\in\mathcal{S}(\mathbb{R}^d)$ always holds

$$|h_1(\psi)-h_2(\psi)| \le \int_{\mathbb{R}^d}|h_1(x)-h_2(x)||\psi(x)|dx \le \left\|h_1-h_2\middle|L_\infty(\mathbb{R}^d)\right\|\cdot\left\|\psi\middle|L_1(\mathbb{R}^d)\right\|, \quad (2.3.39)$$

we additionally obtain $g_{\overline{k}} = \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ in $\mathcal{S}'(\mathbb{R}^d)$. This follows from the fact that $\mathcal{S}'(\mathbb{R}^d)$ is a Hausdorff space, hence limiting elements are unique. In particular, we find $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\in L_p(\mathbb{R}^d)$ for all $\overline{k}\in\mathbb{N}_0^N$.
Now, let $M\in\mathbb{N}$. Then for every $\varepsilon>0$ there exists some $l\in\mathbb{N}$, such that

$$\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f_l - \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\middle|L_p(\mathbb{R}^d)\right\| < \varepsilon 2^{-\overline{k}\cdot(\overline{r}+\overline{d})} \quad \text{for all } |\overline{k}|\le M\,. \qquad (2.3.40)$$

Hence, we find

$$\sum_{|\overline{k}|\le M} 2^{\overline{k}\cdot\overline{r}q}\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\middle|L_p(\mathbb{R}^d)\right\|^q$$

$$
\begin{aligned}
&\le \sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} c_1 \Big(\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F} f_l \big| L_p(\mathbb{R}^d)\big\|^q + \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F} f_l - \mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F} f \big| L_p(\mathbb{R}^d)\big\|^q \Big) \\
&\le c_1 \sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F} f_l \big| L_p(\mathbb{R}^d)\big\|^q + c_1 \sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} \varepsilon^q 2^{-\bar{k}\cdot(\bar{r}+\bar{d})q} \\
&\le c_1 \sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}q} \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F} f_l \big| L_p(\mathbb{R}^d)\big\|^q + c_1 \sum_{\bar{k}\in\mathbb{N}_0^N} \varepsilon^q 2^{-\bar{k}\cdot\bar{d}q} \\
&= c_1 \big\| f_l \big| S^{\bar{r}}_{p,q}B(\mathbb{R}^{\bar{d}})\big\|^q + c_1 c_2 \varepsilon^q \le C ,
\end{aligned}
$$

because Cauchy sequences are always bounded. Since the constant C does not depend on M, we can conclude $f \in S^{\bar{r}}_{p,q}B(\mathbb{R}^{\bar{d}})$.
Let again $M \in \mathbb{N}$. Then by the definition of a Cauchy sequence for every $\varepsilon > 0$ there is some $l_0(\varepsilon) \in \mathbb{N}$, such that

$$
\sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}\big(f_l - f_m\big) \big| L_p(\mathbb{R}^d)\big\|^q < \varepsilon^q, \quad \text{if} \quad l, m \ge l_0(\varepsilon) . \tag{2.3.41}
$$

As the sum is a finite one, it follows from (2.3.41)

$$
\sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}\big(f_l - f\big) \big| L_p(\mathbb{R}^d)\big\|^q < \varepsilon^q, \quad \text{if} \quad l \ge l_0(\varepsilon) . \tag{2.3.42}
$$

Letting $M \longrightarrow \infty$, we finally obtain from (2.3.42)

$$
\big\| f_l - f \big| S^{\bar{r}}_{p,q}B(\mathbb{R}^{\bar{d}})\big\| = \Bigg(\sup_{M\in\mathbb{N}} \sum_{|\bar{k}|\le M} 2^{\bar{k}\cdot\bar{r}q} \big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}\big(f_l - f\big) \big| L_p(\mathbb{R}^d)\big\|^q \Bigg)^{1/q} \le \varepsilon
$$

for every $l \ge l_0(\varepsilon)$. Thus, we have proved, that f_l converges to f in the quasi-norm of $S^{\bar{r}}_{p,q}B(\mathbb{R}^{\bar{d}})$.
For F-spaces one argues similarly with the help of Lebesgue's theorem on dominated convergence. □

Remark 2.3.10. Most often, the notion of Cauchy sequences is used in connection with metric spaces. Since $\mathcal{S}'(\mathbb{R}^d)$ is no metric space, we want to recall for sake of completeness what is meant by the notion of a Cauchy sequence in this particular case. In the weak* topology (which coincides in this case with the strong topology) the open neighbourhoods of 0 are of the form

$$
U_{A,\varepsilon} = \big\{ f \in \mathcal{S}'(\mathbb{R}^d) : |f(\varphi_i)| < \varepsilon, i = 1, \ldots, n \big\} ,
$$

where $A = \{\varphi_1, \ldots, \varphi_n\} \subset \mathcal{S}(\mathbb{R}^d)$ and $\varepsilon > 0$. We put

$$
\begin{aligned}
&\mathfrak{U}_0 = \big\{ U \supset U_{A,\varepsilon} : A \subset \mathcal{S}(\mathbb{R}^d), \#A < \infty, \varepsilon > 0 \big\} , \\
&\mathfrak{B} = \big\{ U_{A,\varepsilon} : A = \{\varphi\}, \varphi \in \mathcal{S}(\mathbb{R}^d), \varepsilon > 0 \big\} .
\end{aligned}
$$

Then $\mathfrak{B}$ is a local basis of the neighbourhood system $\mathfrak{U}_0$ (thereby a system $\mathfrak{C} \subset \mathfrak{U}_0$ is called local basis, if it holds $\forall\, U \in \mathfrak{U}_0 \, \exists\, C \in \mathfrak{C} \,:\, C \subset U$). Now, a sequence $(f_k)_{k\in\mathbb{N}}$ is called a Cauchy sequence, if

$$
\forall B \in \mathfrak{B} \;\; \exists M \in \mathbb{N} \;\; \forall n, m > M : \; f_n - f_m \in B .
$$

In our concrete case, this means for distributions: A sequence $(f_k)_{k=1}^{\infty} \subset \mathcal{S}'(\mathbb{R}^d)$ is a Cauchy sequence if, and only if,

$$\forall \varepsilon > 0 \;\; \forall \varphi \in \mathcal{S}(\mathbb{R}^d) \;\; \exists M \in \mathbb{N} \;\; \forall n, m > M \, : \, |f_n(\varphi) - f_m(\varphi)| < \varepsilon \, .$$

In the above proof, this is ensured by an estimate as in (2.3.37), compare also to (2.3.39).

2.3.5 Sobolev embeddings for Besov spaces

Among the most important properties of function and distribution spaces like $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ are embeddings connected to these spaces, both into each other and into other scales of function spaces, in particular L_p-spaces. At this point, we shall restrict ourselves to two particular results. However, we will return to the question of embeddings later, after we proved the wavelet characterization.
In this section, we shall at first clarify when the Besov spaces consist of continuous functions. Thereafter, we shall present a result on so-called Sobolev embeddings, which will be of importance later on.

Lemma 2.3.6. With the notation $\overline{0} = (0, \ldots, 0) \in \mathbb{R}^N$ it holds

$$S^{\overline{0}}_{\infty,1}B(\mathbb{R}^{\overline{d}}) \hookrightarrow C(\mathbb{R}^d) \, .$$

Here, the space $C(\mathbb{R}^d)$ is the Banach space of all complex-valued, uniformly continuous and bounded functions on $\mathbb{R}^d$, equipped with the usual norm

$$\big\| f \,\big|\, C(\mathbb{R}^d)\big\| := \sup_{x\in\mathbb{R}^d} |f(x)| = \big\| f \,\big|L_\infty(\mathbb{R}^d)\big\| \, .$$

Proof. Let $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be a dyadic decomposition of unity according to Definition 1.2.2 and (2.2.1). Then for every $f \in S^{\overline{0}}_{\infty,1}B(\mathbb{R}^{\overline{d}})$, the function $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$ is uniformly continuous. This follows from the Nikol'skij inequality (Proposition 2.3.6), because the boundedness of $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \in L_\infty(\mathbb{R}^d)$ implies the boundedness of all of its partial derivatives. Then we get from the definition of the norm in $S^{\overline{0}}_{\infty,1}B(\mathbb{R}^{\overline{d}})$, that the series

$$\sum_{\overline{k}\in\mathbb{N}_0^N} \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \tag{2.3.43}$$

converges at first absolutely in $C(\mathbb{R}^d)$. Hence, the series converges also uniformly to some uniformly continuous function. On the other hand, by (2.2.2) the series (2.3.43) converges in $\mathcal{S}'(\mathbb{R}^d)$ to the limit element f.
Hence f is equivalent to the function defined by (2.3.43), in other words the equivalence class of f contains a continuous representative. In this sense, the statement $f \in C(\mathbb{R}^d)$ is to be understood, and the corresponding estimate for the norms follows from the triangle inequality in $L_\infty(\mathbb{R}^d)$ and once again from the definition of the norm in $S^{\overline{0}}_{\infty,1}B(\mathbb{R}^{\overline{d}})$. □

Remark 2.3.11. For L_p-spaces a similar assertion holds true, compare to [71, Proposition 2.2.3/4], where the corresponding result for the spaces $S^r_{p,q}A(\mathbb{R}^d)$ can be found.

Proposition 2.3.10. Let $0 < p_0 \le p_1 \le \infty$, $0 < q_0, q_1 \le \infty$ and $\bar{r}^{\,0}, \bar{r}^{\,1} \in \mathbb{R}^N$. Assume either

$$r_i^0 - r_i^1 > d_i\left(\frac{1}{p_0} - \frac{1}{p_1}\right), \quad i = 1, \ldots, N\,, \tag{2.3.44}$$

or

$$r_i^0 - r_i^1 \ge d_i\left(\frac{1}{p_0} - \frac{1}{p_1}\right), \quad i = 1, \ldots, N\,, \quad \text{and} \quad q_0 \le q_1\,. \tag{2.3.45}$$

Then it holds

$$S^{\bar{r}^{\,0}}_{p_0,q_0}B(\mathbb{R}^{\bar{d}}) \hookrightarrow S^{\bar{r}^{\,1}}_{p_1,q_1}B(\mathbb{R}^{\bar{d}})\,.$$

Proof. Let $f \in S^{\bar{r}^{\,0}}_{p_0,q_0}B(\mathbb{R}^{\bar{d}})$, and let $\big(\varphi_{\bar{k}}\big)_{\bar{k}\in\mathbb{N}_0^N}$ be a dyadic decomposition of unity according to Definition 1.2.2 and (2.2.1). We apply the Nikol'skij inequality (Proposition 2.3.6) to $\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f$ with $\alpha = 0$ and $b_i = 2^{k_i+1}$, $i = 1, \ldots, N$. Then we obtain

$$\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_1}(\mathbb{R}^d)\big\| \le c\, 2^{(\bar{k}+\bar{1})\cdot\bar{d}(1/p_0-1/p_1)}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|\,. \tag{2.3.46}$$

From this we can conclude

$$\begin{aligned}\big\|\, f\,\big|S^{\bar{r}^{\,1}}_{p_1,q_1}B(\mathbb{R}^{\bar{d}})\big\| &= \left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}^{\,1}q_1}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_1}(\mathbb{R}^d)\big\|^{q_1}\right)^{1/q_1}\\ &\le c\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}^{\,1}q_1}2^{\bar{k}\cdot\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})q_1}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|^{q_1}\right)^{1/q_1}.\end{aligned} \tag{2.3.47}$$

In case $q_0 \le q_1$ we use the monotonicity of the ℓ_p-norms and obtain together with (2.3.45)

$$\begin{aligned}\big\|\, f\,\big|S^{\bar{r}^{\,1}}_{p_1,q_1}B(\mathbb{R}^{\bar{d}})\big\| &\le c\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\big(\bar{r}^{\,1}+\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})\big)q_0}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|^{q_0}\right)^{1/q_0}\\ &\le c\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}^{\,0}q_0}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|^{q_0}\right)^{1/q_0} = c\,\big\|\, f\,\big|S^{\bar{r}^{\,0}}_{p_0,q_0}B(\mathbb{R}^{\bar{d}})\big\|\,.\end{aligned}$$

On the other hand, if $q_0 > q_1$ we apply Hölder's inequality to (2.3.47) with respect to $\frac{q_1}{q_0} + \frac{q_0-q_1}{q_0} = 1$. Then we find

$$\begin{aligned}&\big\|\, f\,\big|S^{\bar{r}^{\,1}}_{p_1,q_1}B(\mathbb{R}^{\bar{d}})\big\|\\ &\quad\le c\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}^{\,0}q_0}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|^{q_0}\right)^{\frac{1}{q_0}}\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\big(-\bar{r}^{\,0}+\bar{r}^{\,1}+\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})\big)\frac{q_1q_0}{q_0-q_1}}\right)^{\frac{q_0-q_1}{q_0q_1}}\\ &\quad\le c'\left(\sum_{\bar{k}\in\mathbb{N}_0^N} 2^{\bar{k}\cdot\bar{r}^{\,0}q_0}\big\|\mathcal{F}^{-1}\varphi_{\bar{k}}\mathcal{F}f\big|L_{p_0}(\mathbb{R}^d)\big\|^{q_0}\right)^{1/q_0} = c'\,\big\|\, f\,\big|S^{\bar{r}^{\,0}}_{p_0,q_0}B(\mathbb{R}^{\bar{d}})\big\|\,.\end{aligned}$$

The arising geometric series is convergent due to the assumption (2.3.44). □

2.3.6 Lifting property

Similarly to the case of classical Besov and Triebel-Lizorkin spaces, we now define a lifting operator. It will be an important tool in later considerations.

Definition 2.3.2. Let $\overline{\rho} \in \mathbb{R}^N$. Then we define the lifting operator $I_{\overline{\rho}}$ by

$$I_{\overline{\rho}} f = \mathcal{F}^{-1}(1+|\xi^1|^2)^{\rho_1/2} \cdots (1+|\xi^N|^2)^{\rho_N/2} \mathcal{F} f, \qquad f \in \mathcal{S}'(\mathbb{R}^d)\,. \tag{2.3.48}$$

It is an immediate consequence of the respective definitions, that the lifting operator $I_{\overline{\rho}}$ is an isomorphism from $\mathcal{S}(\mathbb{R}^d)$ onto itself and from $\mathcal{S}'(\mathbb{R}^d)$ onto itself. However, at this point we are much more interested in its mapping properties for spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. The answer is given by the following proposition.

Proposition 2.3.11. Let $0 < p, q \le \infty$ ($p < \infty$ for F-spaces) and $\overline{\rho}, \overline{r} \in \mathbb{R}^N$. Then the lifting operator $I_{\overline{\rho}}$ maps the space $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ isomorphically onto the space $S^{\overline{r}-\overline{\rho}}_{p,q}A(\mathbb{R}^{\overline{d}})$. Furthermore, $\big\| I_{\overline{\rho}}(\,\cdot\,) \big| S^{\overline{r}-\overline{\rho}}_{p,q}A(\mathbb{R}^{\overline{d}}) \big\|$ defines an equivalent quasi-norm on $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$.

Proof. At first, we consider the case of F-spaces. Let $\big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be an arbitrary decomposition of unity according to Definition 1.2.2 and (2.2.1). Furthermore, we put $\widetilde{\varphi}^i_{k_i} = \varphi^i_{k_i-1} + \varphi^i_{k_i} + \varphi^i_{k_i+1}$, $i = 1, \dots, N$, where $\varphi^i_{-1} \equiv 0$. Thus we obtain

$$\big\| I_{\overline{\rho}} f \big| S^{\overline{r}-\overline{\rho}}_{p,q}F(\mathbb{R}^{\overline{d}}) \big\| = \big\| \mathcal{F}^{-1} m_{\overline{k}} \mathcal{F}\big(2^{\overline{k}\cdot\overline{r}} f_{\overline{k}}\big) \big| L_p(\ell_q) \big\|\,,$$

where

$$m_{\overline{k}}(x) = \prod_{i=1}^N m^i_{k_i}(x^i) = \prod_{i=1}^N 2^{-k_i\rho_i}(1+|x^i|^2)^{\rho_i/2} \widetilde{\varphi}^i_{k_i}(x^i)$$

and $f_{\overline{k}} = \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f$. Now consider first the functions $m^i_{k_i}(2^{k_i+1}\,\cdot\,)$. By property (1.2.1) it follows at once, that the support of these functions is contained in $\{t \in \mathbb{R}^{d_i} : |t| \le 2\}$ for all $k_i \in \mathbb{N}_0$, $i = 1, \dots, N$. From property (1.2.2) we get

$$D^\alpha\big[\widetilde{\varphi}^i_{k_i}(2^{k_i+1}\,\cdot\,)\big](t) = 2^{(k_i+1)|\alpha|}(D^\alpha \widetilde{\varphi}^i_{k_i})(2^{k_i+1}t) \le c_\alpha 2^{(k_i+1)|\alpha|} 2^{-k_i|\alpha|} \le c'_\alpha\,. \tag{2.3.49}$$

Moreover, derivatives $D^\beta(1+|x^i|^2)^{\rho_i/2}$ can obviously be written as linear combinations of terms $(1+|x^i|^2)^{\rho_i/2-l}\pi_{l,n}(x^i)$, $l = 1, \dots, |\beta|$, $n = 1, \dots, n_l$, where $\pi_{l,n}$ are monomials of total degree at most $(2l-|\beta|)_+$, their count n_l being bounded depending on β only. This implies

$$\big| D^\beta\big[(1+|2^{k_i+1}\cdot|^2)^{\rho_i/2}\big](t) \big| \le \max_{l\colon 2l \ge |\beta|} 2^{|\beta|(k_i+1)} 2^{k_i(\rho_i-2l)} 2^{k_i(2l-|\beta|)} C_\beta = C'_\beta 2^{k_i\rho_i}\,, \tag{2.3.50}$$

where only those $t \in \operatorname{supp} \widetilde{\varphi}^i_{k_i}(2^{k_i+1}\,\cdot\,)$ are of relevance. Thus it follows, that the functions $D^\alpha[m^i_{k_i}(2^{k_i+1}\,\cdot\,)]$ are uniformly bounded independent of k_i and α. From this, we eventually obtain for $\overline{\sigma} \in \mathbb{N}_0^N$

$$\big\| m_{\overline{k}}(2^{k_1+1}\,\cdot, \dots, 2^{k_N+1}\,\cdot\,) \big| S^{\overline{\sigma}}_2 H(\mathbb{R}^{\overline{d}}) \big\|$$

$$\sim \big\| m_{\overline{k}}(2^{k_1+1}\cdot,\dots,2^{k_N+1}\cdot)\big| S_2^{\overline{\sigma}}W(\mathbb{R}^{\overline{d}})\big\| = \prod_{i=1}^{N} \big\| m^i_{k_i}(2^{k_i+1}\cdot)\big| W_2^{\sigma_i}(\mathbb{R}^{d_i})\big\| \le \prod_{i=1}^{N} C_i\,,$$

where we used Theorem 2.1.1 and Remark 2.1.2. Moreover, though the constants C_i may depend on ρ_i, σ_i, and φ^i, they are independent of k_i.
If we choose $\overline{\sigma}$ sufficiently large, such that the condition (2.3.18) is satisfied, we can finally apply Proposition 2.3.5 to obtain

$$\begin{aligned}
&\big\|\mathcal{F}^{-1}m_{\overline{k}}\mathcal{F}\big(2^{\overline{k}\cdot\overline{r}}f_{\overline{k}}\big)\big|L_p(\ell_q)\big\| \\
&\quad\le c \sup_{\overline{k}\in\mathbb{N}_0^N} \big\| m_{\overline{k}}(2^{k_1+1}\cdot,\dots,2^{k_N+1}\cdot)\big| S_2^{\overline{\sigma}}H(\mathbb{R}^{\overline{d}})\big\| \cdot \big\|2^{\overline{k}\cdot\overline{r}}f_{\overline{k}}\big|L_p(\ell_q)\big\| \\
&\quad\le C\,\big\|2^{\overline{k}\cdot\overline{r}}f_{\overline{k}}\big|L_p(\ell_q)\big\| = C\,\big\|f\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,.
\end{aligned}$$

Hence we have shown

$$\big\|\,I_{\overline{\rho}}f\,\big|S^{\overline{r}-\overline{\rho}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \le c\,\big\|\,f\,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,. \tag{2.3.51}$$

The inverse estimate now follows from (2.3.51) and the observation, that on $\mathcal{S}'(\mathbb{R}^d)$ the inverse of the lifting operator $I_{\overline{\rho}}$ is given by $I_{-\overline{\rho}}$. Consequently we find

$$\big\|\,f\,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| = \big\|I_{-\overline{\rho}}I_{\overline{\rho}}f\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \le c\,\big\|\,I_{\overline{\rho}}f\,\big|S^{\overline{r}-\overline{\rho}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,.$$

This also means, that the restriction of $I_{-\overline{\rho}}$ to the space $S^{\overline{r}-\overline{\rho}}_{p,q}F(\mathbb{R}^{\overline{d}})$ is the inverse of the corresponding lifting operator $I_{\overline{\rho}}$ on $S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})$, and that this restriction is continuous. Altogether we have proved, that the mapping $I_{\overline{\rho}}: S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}}) \longrightarrow S^{\overline{r}-\overline{\rho}}_{p,q}F(\mathbb{R}^{\overline{d}})$ is an isomorphism.
For B-spaces one proceeds analogously, and applies Proposition 2.3.3 instead of Proposition 2.3.5. □

Proposition 2.3.12. Let $0 < p, q \le \infty$ ($p < \infty$ for F-spaces), $\overline{r} \in \mathbb{R}^N$, and $\overline{l} \in \mathbb{N}_0^N$. Then it holds

$$\sum_{\alpha\in\mathbb{N}_0^d:\,\overline{\alpha}\le\overline{l}} \big\|D^\alpha f\big|S^{\overline{r}-\overline{l}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \le c\,\big\|\,f\,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \tag{2.3.52}$$

for all $f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$.

Proof. We begin with the case of B-spaces. Let $\big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be a dyadic decomposition of unity according to Definition 1.2.2 and (2.2.1), and again we define $\widetilde{\varphi}^i_{k_i} = \varphi^i_{k_i-1} + \varphi^i_{k_i} + \varphi^i_{k_i+1}$, where $\varphi^i_{-1} \equiv 0$. Let $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. We put for $\alpha \in \mathbb{N}_0^d$ with $\overline{\alpha} \le \overline{l}$

$$m_\alpha(x) = \prod_{i=1}^N m^i_{\alpha^i}(x^i) = \prod_{i=1}^N x^{\alpha^i}(1+|x^i|^2)^{-l_i/2}, \quad x \in \mathbb{R}^d$$

and $m_{\alpha,\overline{k}}(x) = m_\alpha(x)\widetilde{\varphi}^1_{k_1}(x^1)\cdots\widetilde{\varphi}^N_{k_N}(x^N) = m^1_{\alpha^1,k_1}(x^1)\cdots m^N_{\alpha^N,k_N}(x^N)$ for $\overline{k} \in \mathbb{N}_0^N$. Then we obtain firstly for $\beta \in \mathbb{N}_0^{d_i}$

$$\Big|D^\beta\big[m^i_{\alpha^i,k_i}(2^{k_i+1}\cdot)\big](x^i)\Big| \le c\sum_{\gamma\le\beta}\Big|D^\gamma\big[m^i_{\alpha^i}(2^{k_i+1}\cdot)\big](x^i)\Big|\cdot\Big|D^{\beta-\gamma}\big[\widetilde{\varphi}^i_{k_i}(2^{k_i+1}\cdot)\big](x^i)\Big|\,.$$

The second factor can be estimated as before in (2.3.49), and also for the first factor we can argue similarly as in the proof of Proposition 2.3.11 or Theorem 2.1.1. More precisely, the derivatives $D^\beta m^i_{\alpha^i}(x^i)$ can be written as linear combinations of terms $(1+|x^i|^2)^{-l_i/2-j}\pi_{j,n}$, $j=1,\ldots,|\beta|$, $n=1,\ldots,n_j$, where $\pi_{j,n}$ are monomials of degree at most $(|\alpha^i|+2j-|\beta|)_+$, their count n_j being bounded independently of j. Altogether we obtain for every $i\in\{1,\ldots,N\}$

$$\begin{aligned}
&\Big|D^\beta\big[m^i_{\alpha^i,k_i}(2^{k_i+1}\cdot)\big](x^i)\Big| \\
&\quad\le \sum_{\gamma\le\beta} c_{\beta,\gamma}\Big|D^\gamma\big[m^i_{\alpha^i}(2^{k_i+1}\cdot)\big](x^i)\Big| \le c\sum_{\gamma\le\beta}2^{|\gamma|(k_i+1)}\Big|\big[D^\gamma m^i_{\alpha^i}\big](2^{k_i+1}x^i)\Big| \\
&\quad\le c'\sum_{\gamma\le\beta}2^{|\gamma|(k_i+1)}\max_{j:\,2j+|\alpha^i|\ge|\beta|}2^{k_i(|\alpha^i|+2j-|\beta|)}2^{k_i(-l_i-2j)} \\
&\quad\le c''\,2^{|\beta|(k_i+1)}2^{k_i(|\alpha^i|-l_i-|\beta|)} = C_\beta\,2^{k_i(|\alpha^i|-l_i)}\le C_\beta\,.
\end{aligned}$$

One has to keep in mind, that only the $x^i\in\operatorname{supp}\widetilde{\varphi}^i_{k_i}$ are of relevance, i.e. $2^{k_i+1}|x^i|\sim 2^{k_i+1}$. As in the proof of Proposition 2.3.11 we can conclude from this

$$\big\|m_{\alpha,\overline{k}}\big|S^{\overline{\sigma}}_2H(\mathbb{R}^{\overline{d}})\big\|\le C(\alpha,\overline{\sigma})\quad\text{for all }\overline{k}\in\mathbb{N}_0^N\,,\overline{\sigma}\in\mathbb{N}_0^N\,,$$

where we used Theorem 2.1.1 and the crossnorm-property of the Sobolev spaces (see Remark 2.1.2). Eventually, by the observation $\varphi_{\overline{k}}\widetilde{\varphi}_{\overline{k}}=\varphi_{\overline{k}}$, Proposition 2.3.3 (applicable, if $\overline{\sigma}\in\mathbb{N}_0^N$ is chosen sufficiently large) and the lifting property for Besov spaces (Proposition 2.3.11) we obtain

$$\begin{aligned}
\big\|D^\alpha f\big|S^{\overline{r}-\overline{l}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| &= \big\|\mathcal{F}^{-1}x^\alpha\mathcal{F}f\big|S^{\overline{r}-\overline{l}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| = \big\|\mathcal{F}^{-1}m_{\alpha,\overline{k}}\mathcal{F}\big(I_{\overline{l}}f\big)\big|S^{\overline{r}-\overline{l}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \\
&\le c\,\big\|\,I_{\overline{l}}f\,\big|S^{\overline{r}-\overline{l}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|\le c'\,\big\|\,f\,\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|
\end{aligned}$$

for every multiindex α with $\overline{\alpha}\le\overline{l}$. This finally implies the assertion.
The proof for the F-case uses an identical argumentation, only at the end one has to replace the usage of Proposition 2.3.3 by Proposition 2.3.5. □

Remark 2.3.12. In analogy to [71, Theorem 2.2.6/2] one can show, that the left hand side of (2.3.52) even defines an equivalent quasi-norm on $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$. However, for our purposes the above proposition is sufficient.

Corollary 2.3.2. For every $\overline{K}\in\mathbb{N}_0^N$ it holds

$$S^{\overline{K}}_{\infty,1}B(\mathbb{R}^{\overline{d}})\hookrightarrow S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\,.$$

Here, the space $S^{\overline{K}}C(\mathbb{R}^{\overline{d}})$ is the Banach space of all complex-valued differentiable functions on $\mathbb{R}^d$ with classical partial derivatives $D^\alpha f\in C(\mathbb{R}^d)$ for every multiindex $\alpha\in\mathbb{N}_0^d$ with $\overline{\alpha}\le\overline{K}$. This space is equipped with the norm

$$\big\|f\big|S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\big\| := \sum_{\alpha\in\mathbb{N}_0^d:\,\overline{\alpha}\le\overline{K}}\big\|D^\alpha f\big|C(\mathbb{R}^d)\big\|\,. \tag{2.3.53}$$

We will call such functions $\overline{K}$-times continuously differentiable.

Proof. We apply Proposition 2.3.12 with $p = \infty$, $q = 1$, and $\overline{r} = \overline{l} = \overline{K}$, as well as Lemma 2.3.6 to obtain from (2.3.52)

$$\sum_{\alpha \in \mathbb{N}_0^d:\, \overline{\alpha} \le \overline{l}} \big\| D^\alpha f \big| C(\mathbb{R}^d) \big\| \le \sum_{\alpha \in \mathbb{N}_0^d:\, \overline{\alpha} \le \overline{l}} \big\| D^\alpha f \big| S^{\overline{0}}_{\infty,1} B(\mathbb{R}^{\overline{d}}) \big\| \le c \,\big\| f \,\big| S^{\overline{K}}_{\infty,1} B(\mathbb{R}^{\overline{d}}) \big\| .$$

Bearing in mind the norm introduced in (2.3.53) this proves the assertion. □

2.3.7 Littlewood-Paley theory

We want to establish a connection between the scale of Triebel-Lizorkin spaces $S^{\overline{r}}_{p,q} F(\mathbb{R}^{\overline{d}})$ and the Sobolev spaces of fractional order of smoothness $S^{\overline{r}}_p H(\mathbb{R}^{\overline{d}})$ discussed in Section 2.1. The result is the following theorem.

Theorem 2.3.4. Let $1 < p < \infty$ and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$S^{\overline{r}}_{p,2} F(\mathbb{R}^{\overline{d}}) = S^{\overline{r}}_p H(\mathbb{R}^{\overline{d}})$$

in the sense of equivalent norms.

Hence Sobolev spaces turn out to be contained in the scale of Triebel-Lizorkin spaces. The essential step is the case $\overline{r} = 0$, which is an extension of the classical Littlewood-Paley theorem.

Proposition 2.3.13. Let $1 < p < \infty$. Moreover, let $\big(\varphi_{\overline{k}}\big)_{\overline{k} \in \mathbb{N}_0^N}$ be a decomposition of unity according to Definition 1.2.2 and equation (2.2.1). Then there exist constants $B_p > A_p > 0$, such that

$$A_p \big\| f \,\big| L_p(\mathbb{R}^d) \big\| \le \big\| \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} f\big)_{\overline{k} \in \mathbb{N}_0^N} \big| L_p(\ell_2) \big\| \le B_p \big\| f \,\big| L_p(\mathbb{R}^d) \big\| \qquad (2.3.54)$$

holds for every $f \in L_p(\mathbb{R}^d)$.

Remark 2.3.13. This result is (in a more general form for parabolic metrics) due to Yamazaki [97]. In the literature, there exists various modifications and generalizations, see e.g. Hytönen and Portal [37] for a vector-valued version or Nagel and Stein [54] for functions defined on products of manifolds, as well as more details and further references. However, all these assertions are based on results on singular integrals on product domains, which are essentially due to Fefferman and Stein [28].

Remark 2.3.14. The above Littlewood-Paley decomposition seems to be the first occurrence, where we have an essentially distinct behaviour, on the one hand for the case $\overline{d} = \overline{1}$ (corresponding results can already be found in articles of Lizorkin [50] or Triebel [88]) and on the other hand the case of a general splitting $\mathbb{R}^{\overline{d}}$. Proposition 2.3.13 is not a straightforward generalization, neither of the isotropic case, nor of the case $\overline{d} = \overline{1}$, though it is traced back by a clever induction argument to the classical case.

Remark 2.3.15. Yamazaki's result is formulated using a very specific decomposition of unity, based on decompositions on $\mathbb{R}^{d_i}$ as in Remark 1.2.2, where the generating function is additionally assumed to be a radial function, i.e. $\varphi^i(x^i) = \widetilde{\varphi}(|x^i|)$ for some function $\varphi \in \mathcal{S}(\mathbb{R})$ with analogous properties. Nevertheless, in view of Theorem 2.3.1 the above version follows at once.

Proof of Theorem 2.3.4. We obtain directly from the lifting property (Proposition 2.3.11), the Littlewood-Paley theorem (Proposition 2.3.13) and the definition of the spaces $S_p^{\overline{r}}H(\mathbb{R}^{\overline{d}})$ in Definition 2.1.1(ii)

$$\| f \,|S^{\overline{r}}_{p,2}F(\mathbb{R}^{\overline{d}})\| \sim \|I_{\overline{r}}f\,|S^{\overline{0}}_{p,2}F(\mathbb{R}^{\overline{d}})\| \sim \|I_{\overline{r}}f\,|L_p(\mathbb{R}^d)\| = \| f \,|S^{\overline{r}}_{p}H(\mathbb{R}^{\overline{d}})\|$$

for every $f \in \mathcal{S}'(\mathbb{R}^d)$. This proves the assertion. □

2.3.8 Dual spaces of Besov spaces

By Theorem 2.3.3 for $\max(p,q) < \infty$ the space $\mathcal{S}(\mathbb{R}^d)$ is a dense subset of $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. Hence every functional on $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ can be interpreted as an element of the dual space of $\mathcal{S}(\mathbb{R}^d)$, i.e. as an element of $\mathcal{S}'(\mathbb{R}^d)$. More precisely, this means that $g \in \mathcal{S}'(\mathbb{R}^d)$ belongs to the dual space $\big[S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big]'$ of the space $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, where $0 < p,q < \infty$ and $\overline{r} \in \mathbb{R}^N$, if, and only if, there is a constant $c > 0$, such that

$$|g(\varphi)| \le c \,\| \varphi \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\| \qquad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^d). \tag{2.3.55}$$

All the subsequent statements have to be understood in that way.
At first, we quote a result on the dual spaces of $L_p(\ell_q)$ and $\ell_q(L_p)$. For $1 < q < \infty$ the conjugate exponent q' is determined in the usual way by $\frac{1}{q} + \frac{1}{q'} = 1$. Moreover, for $0 < q \le 1$ we put $q' = \infty$. Correspondingly, p' has to be understood. A proof of the following proposition can be found in [83], Proposition 2.11.1, and in [26], Theorem 8.20.5.

Proposition 2.3.14. Let $1 \le p < \infty$ and $0 < q < \infty$.

(i) We have $g \in \big(\ell_q(L_p)\big)'$ if, and only if, it can be represented uniquely as

$$g(f) = \sum_{\overline{k}\in\mathbb{N}_0^N} \int_{\mathbb{R}^d} g_{\overline{k}}(x) f_{\overline{k}}(x)dx\,, \qquad f = \big(f_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \in \ell_q(L_p)\,,$$

where $\big(g_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \in \ell_{q'}(L_{p'})$. Furthermore, it holds

$$\| g \| = \| g_{\overline{k}} \,|\ell_{q'}(L_{p'})\|$$

for the usual operator norm of $\big(\ell_q(L_p)\big)'$.

(ii) It holds $g \in \big(L_p(\ell_q)\big)'$ if, and only if, it can be represented uniquely as

$$g(f) = \sum_{\overline{k}\in\mathbb{N}_0^N} \int_{\mathbb{R}^d} g_{\overline{k}}(x) f_{\overline{k}}(x)dx\,, \qquad f = \big(f_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \in L_p(\ell_q)\,,$$

where $\big(g_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \in L_{p'}(\ell_{q'})$. Moreover, it holds

$$\| \, g \, \| = \| \, g_{\overline{k}} \, | L_{p'}(\ell_{q'}) \|$$

for the usual operator norm on $\big(L_p(\ell_q)\big)'$.

We restrict our investigations on the dual spaces to the B-scale. Since for our later considerations on decomposition theorems we only need results for Besov spaces, this will be sufficient. The F-spaces could be treated similarly, but this will be postponed to Section 5.5, where we will use an alternative approach.
We will begin with the case $1 \le p < \infty$ and remind on the determination of q' by $\frac{1}{q}+\frac{1}{q'} = 1$ for $1 < q < \infty$ and $q' = \infty$ for $0 < q \le 1$.

Proposition 2.3.15. Let $\overline{r} \in \mathbb{R}^N$, $0 < q < \infty$, and $1 \le p < \infty$. Then it holds

$$\big[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big]' = S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}})$$

in the sense of the interpretation (2.3.55).

Proof. **Step 1: We show $S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}}) \hookrightarrow \big[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big]'$.**
Let $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $\big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ be a decomposition of unity according to Definition 1.2.2 and (2.2.1). Once again, we put $\widetilde{\varphi}^i_{k_i} = \varphi^i_{k_i-1} + \varphi^i_{k_i} + \varphi^i_{k_i+1}$ and $\widetilde{\varphi}_{\overline{k}} = \widetilde{\varphi}_{k_i} \otimes \cdots \otimes \widetilde{\varphi}_{k_N}$, where $\varphi^i_{-1} \equiv 0$. Then we obtain for every $f \in S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}})$

$$\begin{aligned}
|f(\psi)| &= \Big| \sum_{\overline{k}\in\mathbb{N}_0^N} \big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)(\psi) \Big| = \Big| \sum_{\overline{k}\in\mathbb{N}_0^N} \big(\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{k}}\mathcal{F}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)(\psi) \Big| \\
&= \Big| \sum_{\overline{k}\in\mathbb{N}_0^N} \big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)\big(\mathcal{F}\widetilde{\varphi}_{\overline{k}}\mathcal{F}^{-1}\psi\big) \Big| \\
&\le \big\| 2^{-\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \big| \ell_{q'}(L_{p'}) \big\| \cdot \big\| 2^{\overline{k}\cdot\overline{r}}\mathcal{F}\widetilde{\varphi}_{\overline{k}}\mathcal{F}^{-1}\psi \big| \ell_q(L_p) \big\| \,.
\end{aligned}$$

At the end we used Proposition 2.3.14 and interpreted the sequence $\big(2^{-\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)_{\overline{k}\in\mathbb{N}_0^N}$ as an element of $\big(\ell_q(L_p)\big)'$. Since for every decomposition of unity $\big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ the system $\big(\varphi_{\overline{k}}(-\cdot)\big)_{\overline{k}\in\mathbb{N}_0^N}$ is an admissible decomposition as well, we further find by Theorem 2.3.1

$$\begin{aligned}
\big\| 2^{\overline{k}\cdot\overline{r}}\mathcal{F}\widetilde{\varphi}_{\overline{k}}\mathcal{F}^{-1}\psi \big| \ell_q(L_p) \big\| &\le c \sum_{\overline{l}\in\{-1,0,1\}^N} \big\| \big(2^{\overline{k}\cdot\overline{r}}\mathcal{F}\varphi_{\overline{k}+\overline{l}}\mathcal{F}^{-1}\psi\big)_{\overline{k}\in\mathbb{N}_0^N} \big| \ell_q(L_p) \big\| \\
&= c \sum_{\overline{l}\in\{-1,0,1\}^N} \big\| \big(2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}+\overline{l}}(-\cdot)\mathcal{F}\psi\big)_{\overline{k}\in\mathbb{N}_0^N} \big| \ell_q(L_p) \big\| \\
&\le c' \sum_{\overline{l}\in\{-1,0,1\}^N} \big\| \big\{2^{\overline{k}\cdot\overline{r}}\mathcal{F}^{-1}\varphi_{\overline{k}+\overline{l}}\mathcal{F}\psi\big\}_{\overline{k}\in\mathbb{N}_0^N} \big| \ell_q(L_p) \big\| \\
&\le c''\, 3^N \big\| \, \psi \, \big| S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \big\| \,.
\end{aligned}$$

If we combine both estimates, we have shown

$$\big|f(\psi)\big| \le c \, \big\| \, f \, \big| S^{-\overline{r}}_{p',q'} B(\mathbb{R}^{\overline{d}}) \big\| \cdot \big\| \, \psi \, \big| S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}}) \big\| \, ,$$

which means nothing else than $f \in \big[S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}})\big]'$, and for the operator norm of the functional f we find $\| \, f \, \| \le c \, \big\| \, f \, \big| S^{-\overline{r}}_{p',q'} B(\mathbb{R}^{\overline{d}}) \big\|$.

Step 2: We prove the reverse embedding in the case $1 \le q < \infty$.
Since $f \longmapsto J(f) = \big(2^{\overline{k}\cdot\overline{r}} \mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} f\big)_{\overline{k}\in\mathbb{N}_0^N}$ is an isometric bijective mapping from $S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}})$ onto a closed subspace of $\ell_q(L_p)$, every functional $g \in \big[S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}})\big]'$ can be interpreted as a functional on this subspace. By the Hahn-Banach theorem there exists a norm-preserving extension of $g \circ J^{-1}$ to a continuous linear functional $\widetilde{g}$ on the Banach space $\ell_q(L_p)$. Now for every $\psi \in \mathcal{S}(\mathbb{R}^d) \subset S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}})$ follows from Proposition 2.3.14

$$g(\psi) = \widetilde{g}\big(J(\psi)\big) = \sum_{\overline{k}\in\mathbb{N}_0^N} \int_{\mathbb{R}^d} 2^{-\overline{k}\cdot\overline{r}} g_{\overline{k}}(x) 2^{\overline{k}\cdot\overline{r}} \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} \psi\big)(x) dx, \tag{2.3.56}$$

and it holds

$$\big\| \, g \, \big| \big[S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}})\big]' \big\| = \big\| \, \widetilde{g} \, \big| \big[\ell_q(L_p)\big]' \big\| = \big\| 2^{-\overline{k}\cdot\overline{r}} g_{\overline{k}} \, \big| \ell_{q'}(L_{p'}) \big\| \, . \tag{2.3.57}$$

Equation (2.3.56) can be rewritten as

$$g(\psi) = \sum_{\overline{k}\in\mathbb{N}_0^N} g_{\overline{k}}\big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} \psi\big) = \sum_{\overline{k}\in\mathbb{N}_0^N} \big(\mathcal{F} \varphi_{\overline{k}} \mathcal{F}^{-1} g_{\overline{k}}\big)(\psi) \, ,$$

where $g_{\overline{k}}$ is identified with the generated regular distribution in $\mathcal{S}'(\mathbb{R}^d)$. With the help of the support properties of the functions $\varphi_{\overline{k}}$, see (1.2.1), it follows

$$\varphi_{\overline{k}} \mathcal{F} g = \sum_{\overline{l}\in\{-1,0,1\}:\, \overline{k}+\overline{l}\ge 0} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \mathcal{F} g_{\overline{k}+\overline{l}} \tag{2.3.58}$$

as an equation in $\mathcal{S}'(\mathbb{R}^d)$. Hence in the case $p = 1$, i.e. $p' = \infty$, we find

$$\big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} g \big| L_\infty(\mathbb{R}^d) \big\| \le \sum_{\overline{l}\in\{-1,0,1\}:\, \overline{k}+\overline{l}\ge 0} \big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \mathcal{F} g_{\overline{k}+\overline{l}} \big| L_\infty(\mathbb{R}^d) \big\| \, .$$

For a single summand on the right hand side of the last estimate we obtain

$$\begin{aligned}
&\big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \mathcal{F} g_{\overline{k}+\overline{l}} \big| L_\infty(\mathbb{R}^d) \big\| \\
&\qquad = \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \Big| g_{\overline{k}+\overline{l}} \Big(\big(\mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot)\big)(x - \cdot) \Big) \Big| \\
&\qquad \le \sup_{x\in\mathbb{R}^d} \big\| g_{\overline{k}+\overline{l}} \big| L_\infty(\mathbb{R}) \big\| \cdot \big\| \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot)\big)(x - \cdot) \big| L_1(\mathbb{R}^d) \big\| \\
&\qquad = \big\| g_{\overline{k}+\overline{l}} \big| L_\infty(\mathbb{R}) \big\| \cdot \big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \big| L_1(\mathbb{R}^d) \big\| \, .
\end{aligned}$$

We now additionally assume, that the decompositions $\varphi^i = (\varphi^i_{k_i})_{k_i\in\mathbb{N}_0}$, $i = 1, \ldots, N$, are of the form described in Remark 1.2.2, i.e. we have $\varphi^i_{k_i}(x) = \varphi^i_1(2^{-k_i+1}x)$ for $k_i \ge 1$. Then we can calculate further

$$\big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \big| L_1(\mathbb{R}^d) \big\|$$

$$= \int_{\mathbb{R}^d} \big|\big(\mathcal{F}^{-1}\varphi_{\overline{k}}\varphi_{\overline{k}+\overline{l}}(-\cdot)\big)(x)\big|dx = \prod_{i=1}^N \int_{\mathbb{R}^{d_i}} \big|\big(\mathcal{F}_{d_i}^{-1}\varphi^i_{k_i}\varphi^i_{k_i+l_i}(-\cdot)\big)(x^i)\big|dx^i$$
$$= \prod_{i:\,k_i\geq 1} \int_{\mathbb{R}^{d_i}} \big|2^{(k_i-1)d_i}\big(\mathcal{F}_{d_i}^{-1}\varphi^i_1\varphi^i_{1+l_i}(-\cdot)\big)(2^{k_i-1}x^i)\big|dx^i \prod_{i:\,k_i=0} \int_{\mathbb{R}^{d_i}} \big|\big(\mathcal{F}_{d_i}^{-1}\varphi^i_0\varphi^i_{l_i}(-\cdot)\big)(x^i)\big|dx^i$$
$$= \prod_{i:\,k_i\geq 1} \int_{\mathbb{R}^{d_i}} \big|\big(\mathcal{F}_{d_i}^{-1}\varphi^i_1\varphi^i_{1+l_i}(-\cdot)\big)(y^i)\big|dy^i \prod_{i:\,k_i=0} \int_{\mathbb{R}^{d_i}} \big|\big(\mathcal{F}_{d_i}^{-1}\varphi^i_0\varphi^i_{l_i}(-\cdot)\big)(x^i)\big|dx^i = c_{\overline{l}}\,.$$

In particular, the result $c_{\overline{l}}$ is independent of $\overline{k}$. Altogether, the last considerations yield

$$\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_\infty(\mathbb{R}^d)\big\| \leq c \sum_{\overline{l}\in\{-1,0,1\}:\,\overline{k}+\overline{l}\geq 0} \big\|g_{\overline{k}+\overline{l}}\big|L_\infty(\mathbb{R}^d)\big\|\,. \tag{2.3.59}$$

In the case $1 < p < \infty$ (and thus $1 < p' < \infty$) one obtains an analogous result by a considerably simpler calculation. Again we start with (2.3.58) and apply $\mathcal{F}^{-1}$ and the $L_{p'}(\mathbb{R}^d)$-norm to obtain

$$\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_{p'}(\mathbb{R}^d)\big\| \leq \sum_{\overline{l}\in\{-1,0,1\}:\,\overline{k}+\overline{l}\geq 0} \big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\varphi_{\overline{k}+\overline{l}}(-\cdot)\mathcal{F}g_{\overline{k}+\overline{l}}\big|L_{p'}(\mathbb{R}^d)\big\|\,.$$

With the help of Proposition 2.3.3 we now find

$$\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_{p'}(\mathbb{R}^d)\big\| \leq c \sum_{\overline{l}\in\{-1,0,1\}:\,\overline{k}+\overline{l}\geq 0} \big\|\,g_{\overline{k}+\overline{l}}\,\big|L_{p'}(\mathbb{R}^d)\big\|\,, \tag{2.3.60}$$

where the constant c is independent of $\overline{k}$ (compare to the argumentation in the proof of Theorem 2.3.1). From (2.3.59) and (2.3.60) we can conclude by multiplying with $2^{-\overline{k}\cdot\overline{r}}$ and applying the $\ell_{q'}$-norm

$$\begin{aligned}
\big\|\,g\,\big|S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}})\big\| &\leq c\,\Bigg\|\sum_{\overline{l}\in\{-1,0,1\}:\,\overline{k}+\overline{l}\geq 0} 2^{-\overline{k}\cdot\overline{r}}\big\|g_{\overline{k}+\overline{l}}\big|L_{p'}(\mathbb{R}^d)\big\|\Bigg|\,\ell_{q'}\,\Bigg\| \\
&\leq c' \sum_{\overline{l}\in\{-1,0,1\}:\,\overline{k}+\overline{l}\geq 0} \Big\|2^{-(\overline{k}+\overline{l})\cdot\overline{r}}\big\|g_{\overline{k}+\overline{l}}\big|L_{p'}(\mathbb{R}^d)\big\|\,\Big|\,\ell_{q'}\,\Big\| \\
&\leq c'\,3^N\,\big\|2^{-\overline{k}\cdot\overline{r}}g_{\overline{k}}\,\big|\ell_{q'}(L_{p'})\big\|\,.
\end{aligned}$$

Together with (2.3.57) this eventually proves

$$\big\|\,g\,\big|S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}})\big\| \leq C\,\big\|\,g\,\big|\big[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big]'\big\|\,,$$

i.e. $g \in S^{-\overline{r}}_{p',q'}B(\mathbb{R}^{\overline{d}})$, as well as the asserted estimate for the norms.

Step 3: We treat the case $0 < q < 1$.
In this case the inclusion from Step 1 can also be obtained by the following argument: By Proposition 2.3.7 we have $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}}_{p,1}B(\mathbb{R}^{\overline{d}})$, and hence we find

$$S^{-\overline{r}}_{p',\infty}B(\mathbb{R}^{\overline{d}}) = \big[S^{\overline{r}}_{p,1}B(\mathbb{R}^{\overline{d}})\big]' \hookrightarrow \big[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big]'\,.$$

For the reverse inclusion assume again $g \in \left[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right]'$. Then it holds

$$\begin{aligned}
\left|\left(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\right)(\psi)\right| &= \left|g\left(\mathcal{F}\varphi_{\overline{k}}\mathcal{F}^{-1}\psi\right)\right| \le \|\, g\,\|\left\|\mathcal{F}\varphi_{\overline{k}}\mathcal{F}^{-1}\psi\middle|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right\| \\
&\le c\,\|\, g\,\|\left(\sum_{\overline{l}\in\{-1,0,1\}^N} 2^{\overline{k}\cdot\overline{r}q}\left\|\mathcal{F}^{-1}\varphi_{\overline{k}+\overline{l}}\varphi_{\overline{k}}(-\cdot)\mathcal{F}\psi\middle|L_p(\mathbb{R}^d)\right\|^q\right)^{1/q} \\
&\le c'\,\|\, g\,\|\left(\sum_{\overline{l}\in\{-1,0,1\}^N} 2^{\overline{k}\cdot\overline{r}q}\left\|\,\psi\,\middle|L_p(\mathbb{R}^d)\right\|^q\right)^{1/q} \\
&= 3^N c'\,\|\, g\,\|2^{\overline{k}\cdot\overline{r}}\left\|\,\psi\,\middle|L_p(\mathbb{R}^d)\right\|.
\end{aligned}$$

Here we used (1.2.1) and Proposition 2.3.3 as in the second step (compare to (2.3.60)). It is a well-known fact that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_p(\mathbb{R}^d)$ for $1 \le p < \infty$, and hence $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g$ can be extended to a linear functional on $L_p(\mathbb{R}^d)$. Thus it follows from the usual (isometric) identification $\left(L_p(\mathbb{R}^d)\right)' = L_{p'}(\mathbb{R}^d)$ that the (regular) distribution $\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g$ belongs to $L_{p'}(\mathbb{R}^d)$ and

$$\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\middle|L_{p'}(\mathbb{R}^d)\right\| \le c\,2^{\overline{k}\cdot\overline{r}}\|\, g\,\|,$$

from which we finally conclude

$$\left\|\, g\,\middle|S^{-\overline{r}}_{p',\infty}B(\mathbb{R}^{\overline{d}})\right\| = \sup_{\overline{k}\in\mathbb{N}_0^N} 2^{-\overline{k}\cdot\overline{r}}\left\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\,\middle|L_{p'}(\mathbb{R}^d)\right\| \le c\left\|\, g\,\middle|\left[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right]'\right\|.$$

This proves the assertion in the case $0 < q < 1$. □

Eventually, we shall treat the case $0 < p < 1$ as well.

Proposition 2.3.16. Let $0 < p < 1$, $0 < q < \infty$, and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$\left[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right]' = S^{-\overline{r}+\overline{d}(\frac{1}{p}-1)}_{\infty,q'}B(\mathbb{R}^{\overline{d}})$$

in the sense of the interpretation (2.3.55).

Proof. **Step 1:**
By Proposition 2.3.10 we find

$$S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}-\overline{d}(\frac{1}{p}-1)}_{1,q}B(\mathbb{R}^{\overline{d}})$$

for $0 < p < 1$ and $0 < q < \infty$. Thus we conclude from Proposition 2.3.15

$$S^{-\overline{r}+\overline{d}(\frac{1}{p}-1)}_{\infty,q'}B(\mathbb{R}^{\overline{d}}) \hookrightarrow \left[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right]'. \tag{2.3.61}$$

Step 2:
Let $\left(\varphi_{\overline{k}}\right)_{\overline{k}\in\mathbb{N}_0^N}$ be the same decomposition of unity as in the proof of Proposition 2.3.15. Furthermore, let $g \in \left[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right]'$. Then it holds

$$\begin{aligned}
\left|\left(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\right)(x)\right| &= \left|g\left((\mathcal{F}^{-1}\varphi_{\overline{k}})(x-\cdot)\right)\right| \\
&\le \|g\|\left\|(\mathcal{F}^{-1}\varphi_{\overline{k}})(x-\cdot)\middle|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\right\|.
\end{aligned} \tag{2.3.62}$$

Due to $\varphi^j_{k_j}(x) = \varphi^j_1(2^{-k_j+1}x)$ for $k_j \geq 1$ and the support properties (1.2.1) we find for $j = 1, \dots, N$ and $k_j \geq 2$

$$\begin{aligned}
&\big\|(\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j})(x^j - \cdot)\big|B^{r_j}_{p,q}(\mathbb{R}^{d_j})\big\| \\
&\qquad = \left(\sum_{l_j=-1}^{1} 2^{r_j k_j q}\big\|\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j+l_j}\mathcal{F}_{d_j}\big[(\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j})(x^j - \cdot)\big]\big|L_p(\mathbb{R}^{d_j})\big\|^q\right)^{1/q} \\
&\qquad \leq c \sum_{l_j=-1}^{1} 2^{r_j k_j}\big\|\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j+l_j}\mathcal{F}_{d_j}\big[(\mathcal{F}^{-1}_{d_j}[\varphi^j_{k_j}(-\cdot)])(\cdot - x^j)\big]\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}\big\|\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j+l_j}\varphi^j_{k_j}(-\cdot)e^{ix^j\cdot}\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}\big\|\big(\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j+l_j}\varphi^j_{k_j}(-\cdot)\big)(x^j + \cdot)\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}\big\|\mathcal{F}^{-1}_{d_j}\varphi^j_{k_j+l_j}\varphi^j_{k_j}(-\cdot)\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}\big\|2^{-(-k_j+1)d_j}\big[\mathcal{F}^{-1}_{d_j}\varphi^j_1(2^{-l_j}\cdot)\varphi^j_1(-\cdot)\big](2^{-(-k_j+1)}\cdot)\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}2^{(k_j-1)d_j}\left(\int_{\mathbb{R}^{d_j}} 2^{-(k_j-1)d_j}\big|\big[\mathcal{F}^{-1}_{d_j}\varphi^j_1(2^{-l_j}\cdot)\varphi^j_1(-\cdot)\big](y)\big|^p dy\right)^{1/p} \\
&\qquad = c \sum_{l_j=-1}^{1} 2^{r_j k_j}2^{(k_j-1)d_j(1-\frac{1}{p})}\big\|\mathcal{F}^{-1}_{d_j}\varphi^j_1(2^{-l_j}\cdot)\varphi^j_1(-\cdot)\big|L_p(\mathbb{R}^{d_j})\big\| \\
&\qquad = c' \sum_{l_j=-1}^{1} 2^{r_j k_j}2^{k_j d_j(1-\frac{1}{p})}c_{j,l_j} = c''\, 2^{r_j k_j}2^{k_j d_j(1-\frac{1}{p})}\,.
\end{aligned}$$

For the cases $k_j = 0$ and $k_j = 1$ analogous estimates hold true. With the help of the crossnorm-property (see equation (1.4.10) and Remark 2.2.4) it follows

$$\big\|(\mathcal{F}^{-1}\varphi_{\overline{k}})(x - \cdot)\big|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \leq c\, 2^{\overline{k}\cdot(\overline{r}+\overline{d}(1-\frac{1}{p}))}\,. \tag{2.3.63}$$

In particular, the constant c is independent of $\overline{k}$ and x. Inserting (2.3.63) into (2.3.62) eventually yields

$$2^{-\overline{k}\cdot(\overline{r}+\overline{d}(1-\frac{1}{p}))}\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_\infty(\mathbb{R}^d)\big\| \leq c\,\|g\|\,. \tag{2.3.64}$$

If we now take the supremum with respect to $\overline{k} \in \mathbb{N}^N_0$, then this together with (2.3.61) proves the assertion in the case $0 < q \leq 1$.

Step 3: Now let $1 < q < \infty$.
Let $g \in \big[S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big]'$. We choose points $x^{(\overline{k})} \in \mathbb{R}^d$, such that

$$\tfrac{1}{2}\,\big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_\infty(\mathbb{R}^d)\big\| \leq \big|\big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big)(x^{(\overline{k})})\big| \leq \big\|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}g\big|L_\infty(\mathbb{R}^d)\big\|\,.$$

Thereby the estimate (2.3.64) shows, that the essential supremum is finite. Now let $a_{\overline{k}}$ for $|\overline{k}| \le n$, $n \in \mathbb{N}_0$, be arbitrary complex numbers. Then we put

$$\psi(x) = \sum_{|\overline{k}| \le n} a_{\overline{k}} \big(\mathcal{F}^{-1} \varphi_{\overline{k}}\big)(x^{(\overline{k})} - x) 2^{\overline{k}\cdot(-\overline{r}+\overline{d}(\frac{1}{p}-1))} .$$

Obviously, we have $\psi \in \mathcal{S}(\mathbb{R}^d)$. Using the inequality (2.3.63) and similar arguments as in the second step we further obtain for this function

$$\begin{aligned}
&\big\| \psi \,\big| S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}}) \big\|^q \\
&= \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \Bigg\| \sum_{\substack{\overline{l} \in \{-1,0,1\}^N, \\ |\overline{k}+\overline{l}| \le n}} 2^{(\overline{k}+\overline{l})\cdot(-\overline{r}+\overline{\sigma}_p)} a_{\overline{k}+\overline{l}} \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) e^{ix^{(\overline{k}+\overline{l})}\cdot} \Bigg| L_p(\mathbb{R}^d) \Bigg\|^q \\
&\le c \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \sum_{\substack{\overline{l} \in \{-1,0,1\}^N, \\ |\overline{k}+\overline{l}| \le n}} \Bigg\| 2^{(\overline{k}+\overline{l})\cdot(-\overline{r}+\overline{\sigma}_p)} a_{\overline{k}+\overline{l}} \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) e^{ix^{(\overline{k}+\overline{l})}\cdot} \Bigg| L_p(\mathbb{R}^d) \Bigg\|^q \\
&\le c \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \Bigg(\sum_{\substack{\overline{l} \in \{-1,0,1\}^N, \\ |\overline{k}+\overline{l}| \le n}} \big| a_{\overline{k}+\overline{l}} \big|^q \Bigg) \sup_{\overline{l} \in \{-1,0,1\}^N} 2^{(\overline{k}+\overline{l})\cdot(-\overline{r}+\overline{\sigma}_p)q} \Big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \varphi_{\overline{k}+\overline{l}}(-\cdot) \Big| L_p(\mathbb{R}^d) \Big\|^q \\
&= c \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \Bigg(\sum_{\substack{\overline{l} \in \{-1,0,1\}^N, \\ |\overline{k}+\overline{l}| \le n}} \big| a_{\overline{k}+\overline{l}} \big|^q \Bigg) \big\| \mathcal{F}^{-1} \varphi_{\overline{k}} \big| S^{-\overline{r}+\overline{\sigma}_p}_{p,\infty} B(\mathbb{R}^{\overline{d}}) \big\|^q_{\varphi(-\cdot)} \\
&\le c' \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \Bigg(\sum_{\substack{\overline{l} \in \{-1,0,1\}^N, \\ |\overline{k}+\overline{l}| \le n}} \big| a_{\overline{k}+\overline{l}} \big|^q \Bigg) 2^{-\overline{k}\cdot\overline{r}q} \le c'' \sum_{|\overline{k}| \le n} \big| a_{\overline{k}} \big|^q ,
\end{aligned}$$

we remind on $\varphi^i_{-1} \equiv 0$, $i = 1, \ldots, N$. This last estimate then yields

$$\begin{aligned}
\Bigg| \sum_{|\overline{k}| \le n} a_{\overline{k}} 2^{\overline{k}\cdot(-\overline{r}+\overline{d}(\frac{1}{p}-1))} \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} g\big)(x^{(\overline{k})}) \Bigg| &= \Bigg| \sum_{|\overline{k}| \le n} a_{\overline{k}} 2^{\overline{k}\cdot(-\overline{r}+\overline{d}(\frac{1}{p}-1))} \big(g * \mathcal{F}^{-1} \varphi_{\overline{k}}\big)(x^{(\overline{k})}) \Bigg| \\
&= \big| g(\psi) \big| \le \| g \| \cdot \big\| \psi \,\big| S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}}) \big\| \le c \, \| g \| \Bigg(\sum_{|\overline{k}| \le n} |a_{\overline{k}}|^q \Bigg)^{1/q} .
\end{aligned}$$

Here the constant c is independent of g, n, and the numbers $a_{\overline{k}}$. Letting $n \longrightarrow \infty$ and $(a_{\overline{k}})_{\overline{k} \in \mathbb{N}_0^N} \in \ell_q$ this inequality can be re-interpreted in such a way, that the sequence $\Big(2^{\overline{k}\cdot(-\overline{r}+\overline{d}(\frac{1}{p}-1))} \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} g\big)(x^{(\overline{k})})\Big)_{\overline{k} \in \mathbb{N}_0^N}$ generates a continuous linear functional on ℓ_q, whose norm can be estimated by $c\|g\|$ (we remind on the fact that finite sequences are dense in ℓ_q, $q < \infty$). By the usual (isometric) identification $(\ell_q)' = \ell_{q'}$ we then obtain

$$\big\| g \,\big| S^{-\overline{r}+\overline{\sigma}_p}_{\infty,q'} B(\mathbb{R}^{\overline{d}}) \big\| \le 2 \bigg\| \Big(2^{\overline{k}\cdot(-\overline{r}+\overline{d}(\frac{1}{p}-1))} \big(\mathcal{F}^{-1} \varphi_{\overline{k}} \mathcal{F} g\big)(x^{(\overline{k})})\Big)_{\overline{k} \in \mathbb{N}_0^N} \bigg| \ell_{q'} \bigg\| \le 2c \, \| g \| .$$

Together with (2.3.61) this proves the case $1 < q < \infty$. □

3 The Peetre maximal operator and local means

This chapter is devoted to the investigation of an essential tool in the treatment of function spaces of Besov and Triebel-Lizorkin type, that is a characterization of these spaces with the help of local means. The basis of this discussion is an article of V. S. Rychkov [63] about a theorem of Bui, Paluszynski and Taibleson, dealing with a characterization of the isotropic spaces in terms of the Peetre maximal operator. His work in turn was based on techniques presented in the book of Strömberg and Torchinsky [77]. The case of function spaces with dominating mixed smoothness as in Definition 1.4.2 was treated by J. Vybral in his dissertation [94]. Moreover, he extended Rychkov's method to non-smooth kernels. Similar results can be found in the work of D. B. Bazarkhanov [3] and H. Triebel [84].

3.1 Preliminaries

Before we introduce the Peetre maximal operator, we present two technical lemmata. These correspond to Lemma 1 and 2 in [63] and Lemma 1.17 and 1.18 in [94], respectively.

Lemma 3.1.1. Let $K \in \mathbb{N}_0$, and let $g, h \in L_1(\mathbb{R}^n)$ with $\mathcal{F}g, \mathcal{F}h \in C^{K+1}(\mathbb{R}^n)$. Moreover, let $-1 \le M \le K$ be a fixed integer, such that

$$(D^\alpha \mathcal{F}g)(0) = 0 \quad \text{for all multiindices } \alpha \in \mathbb{N}_0^n \text{ with } |\alpha| \le M\,. \tag{3.1.1}$$

Then it holds for every $N \in \mathbb{N}_0$ with $0 \le N \le K$

$$\sup_{z\in\mathbb{R}^n} |(g_b * h)(z)|(1+|z|^N) \le C_N b^{M+1}\,, \qquad 0<b<1\,, \tag{3.1.2}$$

where $g_b(t) = b^{-n} g(t/b)$ and

$$C_N = c_N \big\| \,\mathcal{F}g \,\big|\, C^{M+1}(\mathbb{R}^n)\big\| \sum_{|\beta|\le N+1} \sum_{|\gamma|=(M+1-|\beta|)_+} \big\| \,\xi^\gamma D^\beta \mathcal{F}h \,\big|\, L_1(\mathbb{R}^n)\big\|\,. \tag{3.1.3}$$

Proof. For every function $f \in L_1(\mathbb{R}^n)$ we have

$$\big\|\mathcal{F}^{-1} f \big| L_\infty(\mathbb{R}^n)\big\| \le (2\pi)^{-n/2} \big\| \,f \,\big| L_1(\mathbb{R}^n)\big\|\,.$$

From this and further elementary properties of the Fourier transform we obtain

$$\begin{aligned} \sup_{z\in\mathbb{R}^n} |(g_b * h)(z)|(1+|z|^N) &\le \sup_{z\in\mathbb{R}^n} |(g_b * h)(z)|(1+|z|^{2[\frac{N+1}{2}]}) \\ &\le (2\pi)^{-n/2} \big\|\mathcal{F}\big[(g_b * h)(z)(1+|z|^{2[\frac{N+1}{2}]})\big]\big| L_1(\mathbb{R}^n)\big\| \\ &\le c \max_{0\le|\alpha|\le N+1} \|D^\alpha[(g_b*h)^\wedge]|L_1(\mathbb{R}^n)\|\,, \end{aligned} \tag{3.1.4}$$

and it holds $\mathcal{F}g_b = (\mathcal{F}g)(b\,\cdot)$. Furthermore, the Leibniz-formula gives

$$|D^\alpha[\widehat{g}(b\,\cdot)\widehat{h}(\cdot)](\xi)| \le c \sum_{0\le\beta\le\alpha} b^{|\beta|} |(D^\beta \widehat{g})(b\xi)(D^{\alpha-\beta}\widehat{h})(\xi)|, \quad \xi\in\mathbb{R}^n\,, \tag{3.1.5}$$

where $|\alpha| \le N+1 \le K+1$. Since by assumption $\widehat{g} \in C^{K+1}(\mathbb{R}^n) \subset C^{M+1}(\mathbb{R}^n)$ we can apply Taylor's theorem and obtain

$$\begin{aligned}(D^\beta \widehat{g})(b\xi) &= \sum_{|\gamma| \le M-|\beta|} \frac{(D^{\beta+\gamma}\widehat{g})(0)}{\gamma!}(b\xi)^\gamma + \sum_{|\gamma|=M+1-|\beta|} \frac{(D^{\beta+\gamma}\widehat{g})(\theta b\xi)}{\gamma!}(b\xi)^\gamma \\ &= \sum_{|\gamma|=M+1-|\beta|} \frac{(D^{\beta+\gamma}\widehat{g})(\theta b\xi)}{\gamma!}(b\xi)^\gamma\end{aligned}$$

for some $\theta \in (0,1)$, taking into account (3.1.1). We further conclude

$$\big|(D^\beta \widehat{g})(b\xi)\big| \le \max_{|\lambda|=M+1} \big\| D^\lambda \widehat{g} \,\big|L_\infty(\mathbb{R}^n)\big\| \sum_{|\gamma|=M+1-|\beta|} \frac{|(b\xi)^\gamma|}{\gamma!}, \quad 0 \le |\beta| \le M\,,$$

for all $b \in \mathbb{R}_+$ and $\xi \in \mathbb{R}^n$. Hence we have for $0 \le |\beta| \le M$ and $\xi \in \mathbb{R}^n$

$$b^{|\beta|}|(D^\beta \widehat{g})(b\xi)(D^{\alpha-\beta}\widehat{h})(\xi)| \le b^{M+1}\big\| \widehat{g} \,\big|C^{M+1}(\mathbb{R}^n)\big\| \big|(D^{\alpha-\beta}\widehat{h})(\xi)\big| \sum_{|\gamma|=M+1-|\beta|} \frac{|\xi^\gamma|}{\gamma!}\,.$$

In case $M < |\beta| \le K+1$ and $0 < b < 1$ it holds $b^{|\beta|} \le b^{M+1}$, thus together with $D^\beta \widehat{g} \in C(\mathbb{R}^n)$ we find for every $\xi \in \mathbb{R}^n$ and all $\beta \le \alpha$

$$\begin{aligned}&b^{|\beta|}|(D^\beta \widehat{g})(b\xi)(D^{\alpha-\beta}\widehat{h})(\xi)| \\ &\qquad \le b^{M+1}\big\| \widehat{g} \,\big|C^{M+1}(\mathbb{R}^n)\big\| |(D^{\alpha-\beta}\widehat{h})(\xi)| \sum_{|\gamma|=(M+1-|\beta|)_+} \frac{|\xi^\gamma|}{\gamma!}\,, \qquad (3.1.6)\end{aligned}$$

where we put $\xi_i^{\gamma_i} = 1$ if $\gamma_i = 0$ or $\xi_i = 0$. Inserting (3.1.6) into (3.1.5) results in

$$\begin{aligned}|D^\alpha[\widehat{g}(b\,\cdot)\widehat{h}(\cdot)](\xi)| &\le c\, b^{M+1}\big\| \widehat{g} \,\big|C^{M+1}(\mathbb{R}^n)\big\| \sum_{\substack{0 \le \beta \le \alpha \\ |\gamma|=(M+1-|\beta|)_+}} \frac{|(D^{\alpha-\beta}\widehat{h})(\xi)\xi^\gamma|}{\gamma!} \\ &\le c\, b^{M+1}\big\| \widehat{g} \,\big|C^{M+1}(\mathbb{R}^n)\big\| \sum_{\substack{0 \le |\beta| \le N+1 \\ |\gamma|=(M+1-|\beta|)_+}} |(D^\beta \widehat{h})(\xi)\xi^\gamma|\end{aligned}$$

for all $\xi \in \mathbb{R}^n$. Integrating this estimate over $\mathbb{R}^n$, we obtain together with (3.1.4) the assertion (3.1.2). □

Lemma 3.1.2. Let $0 < p, q \le \infty$ and $\delta > 0$. Moreover, let $(g_{\overline{k}})_{\overline{k} \in \mathbb{N}_0^N}$ be a sequence of non-negative measurable functions on $\mathbb{R}^d$, and let

$$G_{\overline{\nu}}(x) = \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{-|\overline{\nu}-\overline{k}|\delta} g_{\overline{k}}(x)\,, \quad x \in \mathbb{R}^d, \quad \overline{\nu} \in \mathbb{N}_0^N\,.$$

Then there is a constant $C = C(p,q,\delta)$, such that

$$\big\| G_{\overline{k}} \,\big|\ell_q(L_p)\big\| \le C\, \big\| g_{\overline{k}} \,\big|\ell_q(L_p)\big\|\,, \quad \text{and} \qquad (3.1.7)$$

$$\big\| G_{\overline{k}} \,\big|L_p(\ell_q)\big\| \le C\, \big\| g_{\overline{k}} \,\big|L_p(\ell_q)\big\|\,. \qquad (3.1.8)$$

Remark 3.1.1. A proof of this lemma can be found in Rychkov's article [63], pages 6–7. We merely note, that the "dimension" of the summation domain (in the above version N) is not connected with the dimension d of the domain of the functions $g_{\overline{k}}$.

3.2 The Peetre maximal operator

In this section we prove an analog of a theorem of Bui, Paluszynski and Taibleson in a formulation given by Rychkov (Theorem BPT in [63]), an equivalent characterization of the isotropic Besov and Triebel-Lizorkin spaces with the help of a maximal operator originally defined by Peetre in [58]. Our result generalizes Theorem 1.23 in [94] as well.

3.2.1 Definition of the maximal operator for non-smooth kernels

By Theorem 2.3.1 the function spaces introduced in Definition 2.2.1 are independent of the used decomposition of unity. Hence we may fix a special decomposition for the subsequent considerations. We choose a system as described in Remark 1.2.2, i.e. for $i = 1, \dots, N$ we choose (real-valued) non-negative functions $\varphi^i \in \mathcal{S}(\mathbb{R}^{d_i})$, where $\varphi^i(x^i) = 1$ for $|x^i| \le \frac{4}{3}$ and $\operatorname{supp} \varphi^i \subset \left\{ t \in \mathbb{R}^{d_i} : |t| \le \frac{7}{5} \right\}$. Then we put $\varphi_0^i = \varphi^i$, $\varphi_1^i(x^i) = \varphi^i(x^i/2) - \varphi^i(x^i)$ and

$$
\begin{aligned}
&\varphi^i_{k_i}(x^i) = \varphi^i_1(2^{-k_i+1}x^i)\,, \quad k_i \in \mathbb{N}\,, \quad x^i \in \mathbb{R}^{d_i}\,, \quad i = 1, \dots, N\,;\\
&\varphi_{\overline{k}}(x) = \varphi^1_{k_1}(x^1) \cdots \varphi^N_{k_N}(x^N)\,, \quad \overline{k} \in \mathbb{N}_0^N\,, \quad x \in \mathbb{R}^d\,.
\end{aligned}
$$

Next we want to transfer the definition of the maximal operator in [58] from the isotropic spaces to our spaces of dominating mixed smoothness. To this purpose we assign to every system $\left(\psi_{\overline{k}}\right)_{\overline{k} \in \mathbb{N}_0^N} \subset \mathcal{S}(\mathbb{R}^d)$, every distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ and every vector $\overline{a} \in \mathbb{R}_+^N$ the following quantities:

$$
\sup_{y \in \mathbb{R}^d} \frac{|(\psi_{\overline{k}} \widehat{f})^\vee(y)|}{\prod_{i=1}^N (1 + |2^{k_i}(y^i - x^i)|^{a_i})}\,, \quad x \in \mathbb{R}^d, \quad \overline{k} \in \mathbb{N}_0^N\,.
$$

Since $\psi_{\overline{k}} \in \mathcal{S}(\mathbb{R}^d)$ for every $\overline{k} \in \mathbb{N}_0^N$ the product $\psi_{\overline{k}} \widehat{f}$ is well-defined for every $f \in \mathcal{S}'(\mathbb{R}^d)$, and by the Paley-Wiener-Schwartz theorem (see e.g. [83], Section 1.2.1, and the references given there for further details) we conclude that $(\psi_{\overline{k}} \widehat{f})^\vee$ is an entire analytic function. In particular, $(\psi_{\overline{k}} \widehat{f})^\vee(y)$ makes sense pointwise.
On the other hand eventually we aspire a characterization of the function spaces with the help of Daubechies wavelets (see Section 4.3). Since these wavelets do not belong to $\mathcal{S}(\mathbb{R}^d)$, hence we will consider non-smooth kernels as well. To that purpose we weaken the definition of the Schwartz space in a rather natural way to obtain the scale of function spaces $X^{\overline{S}}(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = X^{\overline{S}}(\mathbb{R}^{\overline{d}})$. These spaces are defined for every $\overline{S} \in \mathbb{N}_0^N$ by

$$
\begin{aligned}
&X^{\overline{S}}(\mathbb{R}^{\overline{d}}) := \left\{ \psi \in S_2^{\overline{S}} W(\mathbb{R}^{\overline{d}}) \,:\, \left\| \psi \,\middle|\, X^{\overline{S}}(\mathbb{R}^{\overline{d}}) \right\| < \infty \right\},\\
&\left\| \psi \,\middle|\, X^{\overline{S}}(\mathbb{R}^{\overline{d}}) \right\| := \left(\sum_{0 \le \overline{\alpha}, \overline{\beta} \le \overline{S}} \left\| x^\beta D^\alpha \psi(x) \,\middle|\, L_2(\mathbb{R}^d) \right\|^2 \right)^{1/2}.
\end{aligned}
$$

That this definition is indeed well-chosen and adapted to our Fourier analytical methods can be seen from the next lemma.

Lemma 3.2.1. For every $\overline{S} \in \mathbb{N}_0^N$ the Fourier transform is an isomorphism from $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ onto itself, where $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ is interpreted as a subspace of $\mathcal{S}'(\mathbb{R}^d)$.

Proof. The result follows from well-known facts for the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ and $L_2(\mathbb{R}^d)$,

$$\begin{aligned} &D^\alpha\big(\mathcal{F}f\big) = \mathcal{F}\big[(-i\xi)^\alpha f\big]\,,\\ &\mathcal{F}\big(D^\alpha f\big) = (ix)^\alpha \mathcal{F}f\,, \qquad \alpha\in\mathbb{N}_0^d\,,\quad f\in\mathcal{S}'(\mathbb{R}^d)\,. \end{aligned} \tag{3.2.1}$$

By definition we have $\psi\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ if, and only if, $x^\beta D^\alpha \mathcal{F}\big(\mathcal{F}\psi\big)$ belongs to $L_2(\mathbb{R}^d)$ for all $\overline{\alpha},\overline{\beta}\le\overline{S}$. It obviously follows $x^{\beta-\gamma}D^{\alpha-\gamma}\mathcal{F}\big(\mathcal{F}\psi\big)\in L_2(\mathbb{R}^d)$ for all multiindices γ with $\gamma\le\alpha$ and $\gamma\le\beta$.
Now we conclude by the Leibniz-rule $D^\alpha\big(x^\beta\mathcal{F}\big[\mathcal{F}\psi\big]\big)\in L_2(\mathbb{R}^d)$, and from (3.2.1) we obtain $D^\alpha\big(\mathcal{F}\big[D^\beta\mathcal{F}\psi\big]\big)\in L_2(\mathbb{R}^d)$ and $\mathcal{F}\big[x^\alpha D^\beta\mathcal{F}\psi\big]\in L_2(\mathbb{R}^d)$. Finally, Plancherel's equation yields $x^\alpha D^\beta\mathcal{F}\psi\in L_2(\mathbb{R}^d)$ for all $\overline{\alpha},\overline{\beta}\le\overline{S}$, which in turn means just $\mathcal{F}\psi\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$. The boundedness of the Fourier transform follows along the same lines. Hence we have shown $\psi\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})\Longrightarrow\mathcal{F}\psi\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$, and since $\mathcal{F}^2 f = f(-\cdot)$ and $f\in X^{\overline{S}}(\mathbb{R}^{\overline{d}}) \iff f(-\cdot)\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$, this proves the asserted isomorphy. □

Other facts which will be of importance later on are embeddings for these spaces. From Theorem 2.1.1, Theorem 2.3.4 and Proposition 2.3.10 we conclude

$$X^{\overline{S}}(\mathbb{R}^{\overline{d}})\hookrightarrow S_2^{\overline{S}}W(\mathbb{R}^{\overline{d}}) = S_2^{\overline{S}}H(\mathbb{R}^{\overline{d}}) = S_{2,2}^{\overline{S}}F(\mathbb{R}^{\overline{d}}) = S_{2,2}^{\overline{S}}B(\mathbb{R}^{\overline{d}})\hookrightarrow S_{\infty,\infty}^{\overline{S}-\overline{d}/2}B(\mathbb{R}^{\overline{d}})\,.$$

(one has to keep in mind $\overline{S}\in\mathbb{N}_0^N$). Together with Proposition 2.3.7 and Corollary 2.3.2 we obtain for $\overline{S}>\overline{K}+\frac{\overline{d}}{2}$

$$X^{\overline{S}}(\mathbb{R}^{\overline{d}})\hookrightarrow S_{\infty,1}^{\overline{K}}B(\mathbb{R}^{\overline{d}})\hookrightarrow S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\,, \tag{3.2.2}$$

Furthermore, on the one hand we obviously have $\mathcal{S}(\mathbb{R}^d)\subset X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ for all $\overline{S}\in\mathbb{N}_0^N$. On the other hand, since by Theorem 2.3.3 $\mathcal{S}(\mathbb{R}^d)$ is dense in $S_{2,2}^{\overline{S}}B(\mathbb{R}^{\overline{d}})$, the embeddings yield that this remains true for the spaces $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$.
We now put $\omega(x)=\prod_{i=1}^N(1+|x^i|^2)^{S_i/2}$. Due to $\omega^2(x)=\sum_{\overline{\beta}\le\overline{S}}c_\beta x^{2\beta}$ it follows that $\psi\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ if, and only if, $\omega D^\alpha\psi\in L_2(\mathbb{R}^d)$ for all $0\le\overline{\alpha}\le\overline{S}$. A simple calculation, similar to those in the proof of Theorem 2.1.1, shows $|D^\beta\omega(x)|\le c'_\beta\omega(x)$, and one obtains from the Leibniz-rule $D^\alpha(\omega\psi)\in L_2(\mathbb{R}^d)$ for all α with $\overline{\alpha}\le\overline{S}$, if only $\omega D^\beta\psi\in L_2(\mathbb{R}^d)$ for all β with $\overline{\beta}\le\overline{S}$.
Now let $D^\alpha(\omega\psi)\in L_2(\mathbb{R}^d)$ for all α with $\overline{\alpha}\le\overline{S}$. We show inductively, that under this assumption $\omega D^\beta\psi$ belongs to $L_2(\mathbb{R}^d)$ for all multiindices β with $\overline{\beta}\le\overline{S}$. At first, we have by assumption $\omega\psi\in L_2(\mathbb{R}^d)$, which serves as the induction basis. Furthermore, let β be a multiindex with $\overline{\beta}\le\overline{S}$. Then the induction assumption can be formulated as $\omega D^\gamma\psi\in L_2(\mathbb{R}^d)$ for all γ, such that $\gamma\lneq\beta$ (i.e. $\gamma\le\beta$, and for at least one component we have $\gamma_i<\beta_i$). Since $|D^\alpha\omega(x)|\le c'_\alpha\omega(x)$ it immediately follows that $D^\alpha\omega D^\gamma\psi$ belongs to $L_2(\mathbb{R}^d)$ for every α with $\overline{\alpha}\le\overline{S}$ and all admissible γ. Moreover, the Leibniz-rule implies

$$\omega D^\beta\psi = D^\beta(\omega\psi) - \sum_{\gamma\lneq\beta} c_{\gamma,\beta}D^{\beta-\gamma}\omega\cdot D^\gamma\psi\,,$$

and hence we obtain $\omega D^\beta\psi\in L_2(\mathbb{R}^d)$.

Eventually, the statement $D^\alpha(\omega\psi) \in L_2(\mathbb{R}^d)$ for all $0 \leq \overline{\alpha} \leq \overline{S}$ can be rewritten as $\omega\psi$ belonging to $S_2^{\overline{S}}W(\mathbb{R}^{\overline{d}})$. Altogether we have shown

$$\psi \in X^{\overline{S}}(\mathbb{R}^{\overline{d}}) \iff \omega\psi \in S_2^{\overline{S}}W(\mathbb{R}^{\overline{d}})\,. \tag{3.2.3}$$

Thus the spaces $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ turn out to be Banach spaces (this will be clear by the considerations in the next section as well), and (3.2.3) enables us to characterize their dual spaces. We obtain together with Proposition 2.3.15

$$f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)' \iff \tfrac{1}{\omega}f \in \big(S_2^{\overline{S}}W(\mathbb{R}^{\overline{d}})\big)' = \big(S_{2,2}^{\overline{S}}B(\mathbb{R}^{\overline{d}})\big)' = S_{2,2}^{-\overline{S}}B(\mathbb{R}^{\overline{d}})\,. \tag{3.2.4}$$

In the style of the situation for $\mathcal{S}'(\mathbb{R}^d)$ we will not use the norm topology for those dual spaces, but the strong topology. Hence by defining $(\mathcal{F}f)(\psi) = f(\mathcal{F}\psi)$ for $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ and $\psi \in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ the Fourier transform becomes an isomorphism from $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ onto itself. This definition is consistent with the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$. As mentioned before, it holds $\mathcal{S}(\mathbb{R}^d) \subset X^{\overline{S}}(\mathbb{R}^{\overline{d}})$, and hence for $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ we have $f\big|_{\mathcal{S}(\mathbb{R}^d)} \in \mathcal{S}'(\mathbb{R}^d)$. Moreover, its Fourier transform $\mathcal{F}\big(f\big|_{\mathcal{S}(\mathbb{R}^d)}\big)$ coincides with $(\mathcal{F}f)\big|_{\mathcal{S}(\mathbb{R}^d)}$.
Finally, we define the convolution of a function $\psi \in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ and a distribution $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ by

$$\big(f * \psi\big)(y) = \int_{\mathbb{R}^d} f(x)\psi(y-x)dx = f\big(\psi(y-\cdot)\big), \qquad y \in \mathbb{R}^d\,,$$

again similar to the situation for $\mathcal{S}(\mathbb{R}^d)$. To include the case of smooth kernels and the space $\mathcal{S}(\mathbb{R}^d)$ in the subsequent considerations we put $X^{\overline{S}}(\mathbb{R}^{\overline{d}}) = \mathcal{S}(\mathbb{R}^d)$ for $\overline{S} = \infty$.

Remark 3.2.1. In case $S \geq K+1+n$ the assumptions in Lemma 3.1.1 on the functions g and h can be replaced by $g, h \in X^S(\mathbb{R}^n)$. Under this condition g and h as well as $\mathcal{F}g$ and $\mathcal{F}h$ are $(K+1)$-times continuously differentiable functions, and the constant C_N in (3.1.3) can be estimated by $c'_N\big\|g\big|X^S(\mathbb{R}^n)\big\| \cdot \big\|h\big|X^S(\mathbb{R}^n)\big\|$. In particular, the choice of S and the embedding (3.2.2) imply $\big\|\, g\,\big|C^{K+1}(\mathbb{R}^n)\big\| \leq c\,\big\|\, g\,\big|X^S(\mathbb{R}^n)\big\|$. Moreover, for $v(x) := \prod_{i=1}^n(1+|x_i|)$ we conclude from the Cauchy-Schwarz inequality

$$\begin{aligned}
&\sum_{0\leq|\beta|\leq N+1}\ \sum_{|\gamma|=(M+1-|\beta|)_+} \big\|\xi^\gamma D^\beta\mathcal{F}h\big|L_1(\mathbb{R}^n)\big\| \\
&\qquad\leq \sum_{0\leq|\beta|\leq N+1}\ \sum_{|\gamma|=(M+1-|\beta|)_+} \big\|v\xi^\gamma D^\beta\mathcal{F}h\big|L_2(\mathbb{R}^n)\big\| \cdot \big\|v^{-1}\big|L_2(\mathbb{R}^n)\big\| \\
&\qquad\leq c \sum_{0\leq|\beta|\leq N+1}\ \sum_{|\gamma|\leq M+1} \big\|v\xi^\gamma D^\beta\mathcal{F}h\big|L_2(\mathbb{R}^n)\big\| \\
&\qquad\leq c' \sum_{0\leq|\beta|\leq N+1}\ \sum_{|\gamma|\leq M+1+n} \big\|\xi^\gamma D^\beta\mathcal{F}h\big|L_2(\mathbb{R}^n)\big\| \\
&\qquad\leq c' \sum_{0\leq|\beta|\leq K+1+n}\ \sum_{|\gamma|\leq K+1+n} \big\|\xi^\gamma D^\beta\mathcal{F}h\big|L_2(\mathbb{R}^n)\big\| \leq c''\,\big\|\, h\,\big|X^S(\mathbb{R}^n)\big\|\,.
\end{aligned}$$

The above considerations finally justify the following definition.

Definition 3.2.1. For every system of functions $\big(\Psi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N} \subset X^{\overline{S}}(\mathbb{R}^{\overline{d}})$, every distribution $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ and every vector $\overline{a} \in \mathbb{R}_+^N$ we define the Peetre maximal function by

$$(\Psi_{\overline{k}}^* f)_{\overline{a}}(x) = \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{k}} * f)(y)|}{\prod_{i=1}^N (1+|2^{k_i}(y^i - x^i)|^{a_i})} = \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{k}} * f)(y-x)|}{\prod_{i=1}^N (1+|2^{k_i} y^i|^{a_i})}$$

for every $x \in \mathbb{R}^d$ and $\overline{k} \in \mathbb{N}_0^N$.

3.2.2 Weighted Besov- and Triebel-Lizorkin spaces

Before we turn our attention to properties of the maximal function defined above, we need a relation between the spaces $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ and the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ for arbitrary parameters. This will be done in the context of weighted Besov and Triebel-Lizorkin spaces.

Definition 3.2.2. Let $w \in C^\infty(\mathbb{R}^d)$ be a non-negative function with the properties

$$\big|D^\gamma w(x)\big| \le c_\gamma w(x) \quad \text{for all } \gamma \in \mathbb{N}_0^d \text{ and } x \in \mathbb{R}^d\,,$$
$$w(x) \le c\, w(y)\big(1+|x-y|^2\big)^{\alpha/2}$$

for some $\alpha \ge 0$ and constants $c, c_\gamma > 0$. In this case we write $w \in W^d$. For such weight functions we define weighted function spaces by

$$L_p(\mathbb{R}^d, w) := \Big\{ f : \mathbb{R}^d \longrightarrow \mathbb{C} : f \text{ is measurable}\,,\ \big\| f \,\big| L_p(\mathbb{R}^d, w)\big\| < \infty \Big\}$$

as well as

$$S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}, w) := \Big\{ f \in \mathcal{S}'(\mathbb{R}^d) \ :\ \big\| f \,\big| S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}, w)\big\|_\varphi < \infty \Big\}\,.$$

The respective quasi-norms are given by

$$\big\| f \,\big| L_p(\mathbb{R}^d, w)\big\| := \big\| wf \big| L_p(\mathbb{R}^d)\big\| = \left(\int_{\mathbb{R}^d} \big|f(x)\big|^p w^p(x)dx \right)^{1/p},$$

$$\big\| f \,\big| S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}, w)\big\|_\varphi := \left(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \big\| \mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f \big| L_p(\mathbb{R}^d, w)\big\|^q \right)^{1/q},$$

$$\big\| f \,\big| S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}}, w)\big\|_\varphi := \left\| \left(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q} \big| \big(\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f\big)(\cdot)\big|^q \right)^{1/q} \middle| L_p(\mathbb{R}^d, w) \right\|,$$

where $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ is a smooth decomposition of unity as in Definition 2.2.1.

Remark 3.2.2. In the isotropic case, there exists a well-developed theory and extensive literature for weighted function spaces, for smooth weights as above, as well as for non-smooth weights, e.g. Muckenhoupt weights. We refer to [71], where the case of smooth weights is treated extensively in the framework of ultradistributions.
Though most of the treatment in the previous sections for the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ could be transfered to weighted spaces with nearly no additional effort this will not be done here. However, it is clear, that the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}, w)$ are quasi-Banach spaces independent of the decomposition of unity.

For our further proceedings we need the following lemma. Its counterpart for isotropic spaces (which holds in far more generality than needed here) is one of the basic assertions for weighted spaces. We refer to [25, Theorem 4.2.2(ii)] and the references given there.

Lemma 3.2.2. Let $w_i \in W^{d_i}$, $i = 1, \dots, N$, and $\overline{r} \in \mathbb{R}^N$. We define another weight $w \in W^d$ by $w = w_1 \otimes \cdots \otimes w_N$. Then the operator $J : f \longmapsto wf$ is an isomorphism from $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$ onto $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}})$. In particular, we have

$$\| \, wf \, | S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}) \| \sim \| \, f \, | S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w) \| \, .$$

Proof. **Step 1:** The proof relies on tensor product arguments for Hilbert spaces (we remind on Section 1.3.2). At first we observe that the space $B^{r_i}_{2,2}(\mathbb{R}^{d_i}, w_i)$ is a Hilbert space with respect to the scalar product

$$\langle \, f \, , g \, \rangle_i := \sum_{j=0}^{\infty} 2^{2jr_i} \langle \, w_i \mathcal{F}^{-1}_{d_i} \varphi^i_j \mathcal{F}_{d_i} f \, , w_i \mathcal{F}^{-1}_{d_i} \varphi^i_j \mathcal{F}_{d_i} g \, \rangle_{L_2(\mathbb{R}^{d_i})} \, ,$$

where $\varphi^i = (\varphi^i_j)_{j=0}^{\infty} \in \Phi(\mathbb{R}^{d_i})$, $i = 1, \dots, N$. Similarly $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$ turns out to be a Hilbert space with respect to the scalar product

$$\langle \, f \, , g \, \rangle := \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{2\overline{k} \cdot \overline{r}} \langle \, w \mathcal{F}^{-1}_{d} \varphi_{\overline{k}} \mathcal{F}_{d} f \, , w \mathcal{F}^{-1}_{d} \varphi_{\overline{k}} \mathcal{F}_{d} g \, \rangle_{L_2(\mathbb{R}^{d})} \, ,$$

where $(\varphi_{\overline{k}})_{\overline{k} \in \mathbb{N}_0^N}$ is defined as in (2.2.1). We immediately find for arbitrary $f_i, g_i \in B^{r_i}_{2,2}(\mathbb{R}^{d_i}, w_i)$, $i = 1, \dots, N$,

$$\langle \, f_1 \otimes \cdots \otimes f_N \, , \, g_1 \otimes \cdots \otimes g_N \, \rangle = \prod_{i=1}^{N} \langle \, f_i \, , \, g_i \, \rangle_i \, .$$

Hence we obtain that the tensor product $\mathcal{T} = B^{r_1}_{2,2}(\mathbb{R}^{d_1}, w_1) \otimes \cdots \otimes B^{r_N}_{2,2}(\mathbb{R}^{d_N}, w_N)$ is a closed subspace of $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$.
We further note that $\mathcal{S}(\mathbb{R}^d)$ is dense in $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$. The proof of this assertion is completely analogous to the one of Theorem 2.3.3. One has to use a weighted counterpart of Lemma 2.3.3, which can be found in [71, Theorem 1.7.2].
Now we conclude from Remark 1.3.4 and the density of $\mathcal{S}(\mathbb{R}^d)$ in $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$, that the mentioned closed subspace is the space $S^{\overline{r}}_{2,2}B(\mathbb{R}^{\overline{d}}, w)$ itself. Finally, since constant functions belong to W^d this argumentation applies to unweighted spaces as well.

Step 2: We now define linear operators $J_i : f \longmapsto w_i f$. The aforementioned theorem in [25] now states that $J_i : B^{r_i}_{2,2}(\mathbb{R}^{d_i}, w_i) \longrightarrow B^{r_i}_{2,2}(\mathbb{R}^{d_i})$ is an isomorphism (the respective inverse is given by the mapping $f \longmapsto w_i^{-1} f$, since the assumption $w_i \in W^{d_i}$ implies that weights $w_i \not\equiv 0$ are non-vanishing on $\mathbb{R}^{d_i}$). The assertion now follows from Lemma 1.3.2 and the observation $J = J_1 \otimes \cdots \otimes J_N$. This identity in turn follows from the uniformness of the Hilbert space tensor norm, i.e. the uniqueness of the extension from dyads to the whole tensor product. □

We now consider the function $\omega = \prod_{i=1}^{N}(1+|x^i|^2)^{S_i/2}$ defined in the last section. Obviously, we have $\omega \in W^d$ for all $\overline{S} \in \mathbb{N}_0^N$, and it can be written in the product form required for Lemma 3.2.2. Hence due to that lemma we can reformulate (3.2.3) as $X^{\overline{S}}(\mathbb{R}^{\overline{d}}) = S^{\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}}, \omega)$ (this is to be understood in the sense of equivalent norms). Now we can state the aspired relation between the considered scales of function spaces.

Proposition 3.2.1. Let $\overline{r} \in \mathbb{R}^N$ and $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces). Then it exists a vector $\overline{S} \in \mathbb{N}_0^N$, such that it holds

$$S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \hookrightarrow \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'. \tag{3.2.5}$$

Proof. We start with the case $p \leq 2$. At first, because of $X^{\overline{S}}(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}})$ it follows $S^{-\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}}) = \big(S^{\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}})\big)' \hookrightarrow \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$. Now the asserted embedding follows directly from the Propositions 2.3.7 and 2.3.10.
In case $p > 2$ and $q > 2$ we apply Hölder's inequality twice with s and t defined by $\frac{2}{q} + \frac{1}{s} = 1$ and $\frac{2}{p} + \frac{1}{t} = 1$, respectively. In this way, we obtain for every $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$

$$\begin{aligned}
&\big\| f \,|S^{-\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}}, \omega^{-1})\big\| \\
&= \Bigg(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{-2\overline{k}\cdot\overline{S}} \int_{\mathbb{R}^d} \big|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x)\big|^2 \omega^{-2}(x)dx\Bigg)^{1/2} \\
&\leq \Bigg(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\bigg(\int_{\mathbb{R}^d} \big|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x)\big|^2 \omega^{-2}(x)dx\bigg)^{q/2}\Bigg)^{1/q} \Bigg(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{2s\overline{k}\cdot(-\overline{r}-\overline{S})}\Bigg)^{1/2s} \\
&\leq c\Bigg(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\bigg(\int_{\mathbb{R}^d} \big|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x)\big|^p dx\bigg)^{q/p} \bigg(\int_{\mathbb{R}^d} \omega^{-2t}(x)dx\bigg)^{q/2t}\Bigg)^{1/q} \\
&\leq c'\Bigg(\sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot\overline{r}q}\bigg(\int_{\mathbb{R}^d} \big|\mathcal{F}^{-1}\varphi_{\overline{k}}\mathcal{F}f(x)\big|^p dx\bigg)^{q/p}\Bigg)^{1/q} = c' \,\big\| f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\|\,.
\end{aligned}$$

The occurring geometric series is convergent, if $\overline{S} > -\overline{r}$. Moreover, we have to fulfil $2t\overline{S} > \overline{d}$ in order to assure the integrability of ω^{-2t}. Both conditions can be satisfied by choosing $\overline{S}$ sufficiently large.
If $q < 2$, we can replace the first application of Hölder's inequality by the usage the monotonicity of the ℓ_q-quasi-norms. Moreover, we once again choose $\overline{S} > -\overline{r}$. From this it follows $2^{-\overline{k}\cdot\overline{S}} \leq 2^{\overline{k}\cdot\overline{r}}$, and the rest of the calculation remains unchanged.
In all these case we have shown $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{-\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}}, \omega^{-1})$, and by (3.2.4) and Lemma 3.2.2 the spaces $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ and $S^{-\overline{S}}_{2,2}B(\mathbb{R}^{\overline{d}}, \omega^{-1})$ coincide in the sense of equivalent norms. Altogether this proves (3.2.5).
The case of F-spaces can be traced back to the B-case with the help of Proposition 2.3.7. □

3.2.3 Essential properties of the maximal operator

Let functions $\psi_0^i, \psi_1^i : \mathbb{R}^{d_i} \longrightarrow \mathbb{C}$, $i = 1, \ldots, N$, be given. We further put

$$\psi_j^i(t) = \psi_1^i(2^{-j+1}t)\,, \quad t \in \mathbb{R}^{d_i}\,, j = 2, 3, \ldots\,, \tag{3.2.6}$$

$$\psi_{\overline{k}}(x) = \prod_{i=1}^{N} \psi_{k_i}^i(x^i), \quad x \in \mathbb{R}^d, x^i \in \mathbb{R}^{d_i}\,, x = (x^1, \ldots, x^N)\,, \overline{k} \in \mathbb{N}_0^N\,, \tag{3.2.7}$$

$$\Psi_{\overline{k}} = \widehat{\psi}_{\overline{k}}\,, \quad \overline{k} \in \mathbb{N}_0^N\,. \tag{3.2.8}$$

To likewise given functions ϕ_0^i, ϕ_1^i, $i = 1, \ldots, N$, we associate functions $\phi_{\overline{k}}, \Phi_{\overline{k}}$ in an analogous way. Moreover, we assume that the functions $\psi_{\overline{k}}$ and $\phi_{\overline{k}}$ obtained in this way belong to $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ for some $\overline{S} \in \mathbb{N}_0^N$.

Proposition 3.2.2. Let $\overline{a}, \overline{r} \in \mathbb{R}^N$, $\overline{R} \in \mathbb{N}_0^N$, and $0 < p, q \leq \infty$, where $\overline{a} > 0$ and $\overline{r} < \overline{R} + \overline{1}$. Furthermore, we suppose

$$D^\alpha \psi_1^i(0) = 0\,, \quad \alpha \in \mathbb{N}_0^{d_i}\,, \quad |\alpha| \leq R_i\,, \quad i = 1, \ldots, N\,, \tag{3.2.9}$$

as well as for every $i = 1, \ldots, N$ and some $\varepsilon > 0$

$$|\phi_0^i(t)| > 0 \quad \text{on} \quad \{t \in \mathbb{R}^{d_i} \,:\, |t| < \varepsilon\}\,, \tag{3.2.10}$$

$$|\phi_1^i(t)| > 0 \quad \text{on} \quad \{t \in \mathbb{R}^{d_i} \,:\, \varepsilon/2 < |t| < 2\varepsilon\}\,. \tag{3.2.11}$$

If $\overline{S} > \overline{R}$ is large enough, then it holds

$$\big\| \, 2^{\overline{k}\cdot\overline{r}}(\Psi_{\overline{k}}^* f)_{\overline{a}} \, \big| \ell_q(L_p) \big\| \leq c \, \big\| \, 2^{\overline{k}\cdot\overline{r}}(\Phi_{\overline{k}}^* f)_{\overline{a}} \, \big| \ell_q(L_p) \big\|\,, \tag{3.2.12}$$

$$\big\| \, 2^{\overline{k}\cdot\overline{r}}(\Psi_{\overline{k}}^* f)_{\overline{a}} \, \big| L_p(\ell_q) \big\| \leq c \, \big\| \, 2^{\overline{k}\cdot\overline{r}}(\Phi_{\overline{k}}^* f)_{\overline{a}} \, \big| L_p(\ell_q) \big\| \tag{3.2.13}$$

for all $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$.

Proof. **Step 1:** Formal calculations.

We define functions $\big(\lambda_j^i\big)_{j=0}^\infty$, $i = 1, \ldots, N$, by

$$\lambda_j^i(t) = \begin{cases} \varphi_j^i(3t/2\varepsilon)/\phi_j^i(t)\,, & t \in \operatorname{supp} \varphi_j^i(\frac{3\cdot}{2\varepsilon})\,, \\ 0, & \text{else}\,. \end{cases} \tag{3.2.14}$$

Here the functions φ_j^i are the smooth dyadic decompositions of unity on $\mathbb{R}^{d_i}$ fixed at the beginning of the section. The assumptions (3.2.10) and (3.2.11) ensure, that ϕ_j^i is non-vanishing on $\operatorname{supp} \varphi_j^i(\frac{3\cdot}{2\varepsilon})$, hence λ_j^i is well-defined. Moreover, the functions λ_j^i have the following properties:

$$\sum_{j=0}^{\infty} \lambda_j^i(t)\phi_j^i(t) = 1\,, \quad t \in \mathbb{R}^{d_i}\,, \tag{3.2.15}$$

$$\lambda_j^i(t) = \lambda_1^i(2^{-j+1}t)\,, \quad t \in \mathbb{R}^{d_i}\,, j \in \mathbb{N}\,, \tag{3.2.16}$$

$$\operatorname{supp} \lambda_0^i \subset \big\{t \in \mathbb{R}^{d_i} \,:\, |t| \leq \varepsilon\big\}\,, \tag{3.2.17}$$

$$\operatorname{supp}\lambda_j^i \subset \left\{t \in \mathbb{R}^{d_i} \,:\, 2^{j-2}\varepsilon \le |t| \le 2^j\varepsilon\right\}, \quad j \in \mathbb{N}\,. \tag{3.2.18}$$

Eventually we define as usual $\lambda_{\overline{k}}(x) = \lambda^1_{k_1}(x^1)\cdots\lambda^N_{k_N}(x^N)$ for every $\overline{k} \in \mathbb{N}_0^N$, $x = (x^1,\dots,x^N) \in \mathbb{R}^d$. Then we obtain from (3.2.15)

$$\sum_{\overline{k}\in\mathbb{N}_0^N} \lambda_{\overline{k}}(x)\phi_{\overline{k}}(x) = 1\,, \quad x \in \mathbb{R}^d\,.$$

Besides, we put $\Lambda_{\overline{k}} = \widehat{\lambda}_{\overline{k}}, \overline{k} \in \mathbb{N}_0^N$. This leads us to the following identities:

$$f = (2\pi)^d \sum_{\overline{k}\in\mathbb{N}_0^N} \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f\,, \tag{3.2.19}$$

$$\Psi_{\overline{\nu}} * f = (2\pi)^d \sum_{\overline{k}\in\mathbb{N}_0^N} \Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f\,, \quad \overline{\nu} \in \mathbb{N}_0^N\,. \tag{3.2.20}$$

The problems concerning convergence connected with these equations will be discussed in the second step. For the moment, we further obtain

$$\begin{aligned} 2^{\overline{\nu}\cdot\overline{r}}(\Psi^*_{\overline{\nu}}f)_{\overline{a}}(x) &\le \sup_{y\in\mathbb{R}^d} \frac{\sum_{\overline{k}\in\mathbb{N}_0^N} |(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f)(y)|}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{a_i})} \\ &\le \sum_{\overline{k}\in\mathbb{N}_0^N} \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f)(y)|}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{a_i})}\,. \end{aligned} \tag{3.2.21}$$

Moreover, it holds

$$\begin{aligned} |(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f)(y)| &\le \int_{\mathbb{R}^d} |(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}})(z)| \cdot |(\Phi_{\overline{k}} * f)(y-z)| dz \\ &\le (\Phi^*_{\overline{k}}f)_{\overline{a}}(y) \int_{\mathbb{R}^d} |(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}})(z)| \prod_{i=1}^N (1 + |2^{k_i}z^i|^{a_i}) dz \\ &\equiv (\Phi^*_{\overline{k}}f)_{\overline{a}}(y) I_{\overline{\nu}\overline{k}} = (\Phi^*_{\overline{k}}f)_{\overline{a}}(y) \prod_{i=1}^N I^i_{\nu_i k_i}\,, \end{aligned} \tag{3.2.22}$$

where

$$I^i_{\nu_i k_i} = \int_{\mathbb{R}^{d_i}} |(\Psi^i_{\nu_i} * \Lambda^i_{k_i})(z^i)| (1 + |2^{k_i}z^i|^{a_i}) dz^i\,.$$

From Lemma 3.1.1 (taking into account Remark 3.2.1) it follows

$$I^i_{\nu_i k_i} \le C \begin{cases} 2^{(k_i-\nu_i)(R_i+1)}\,, & \text{if } k_i \le \nu_i\,, \\ 2^{(\nu_i-k_i)(a_i+|r_i|+1)}\,, & \text{if } k_i \ge \nu_i\,. \end{cases} \tag{3.2.23}$$

This can be shown in complete analogy to [63], page 5, where S_i is to be chosen large enough, such that Lemma 3.1.1 is applicable with $N \ge a_i + d_i + 1$ and $M > 2a_i + |r_i|$ or $M = R_i$, respectively. Furthermore, it holds

$$(\Phi^*_{\overline{k}}f)_{\overline{a}}(y) \le c\,(\Phi^*_{\overline{k}}f)_{\overline{a}}(x) \prod_{i=1}^N (1 + |2^{k_i}(x^i - y^i)|^{a_i})$$

$$\leq c\,(\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i=1}^{N}(1+|2^{\nu_i}(x^i-y^i)|^{a_i})\max(1,2^{(k_i-\nu_i)a_i})\,.$$

Inserting this into (3.2.22) and using (3.2.23) we obtain

$$\begin{aligned}\sup_{y\in\mathbb{R}^d}\frac{|(\Psi_{\overline{\nu}}*\Lambda_{\overline{k}}*\Phi_{\overline{k}}*f)(y)|}{\prod_{i=1}^N(1+|2^{\nu_i}(x^i-y^i)|^{a_i})} &\leq \sup_{y\in\mathbb{R}^d}\frac{(\Phi^*_{\overline{k}}f)_{\overline{a}}(y)\prod_{i=1}^N I^i_{\nu_i k_i}}{\prod_{i=1}^N(1+|2^{\nu_i}(x^i-y^i)|^{a_i})}\\
&\leq \sup_{y\in\mathbb{R}^d}(\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i=1}^N I^i_{\nu_i k_i}\max(1,2^{(k_i-\nu_i)a_i})\\
&= (\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i=1}^N I^i_{\nu_i k_i}\max(1,2^{(k_i-\nu_i)a_i})\\
&\leq (\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i=1}^N\begin{cases}2^{(k_i-\nu_i)(R_i+1)}, & \text{if } k_i\leq\nu_i\,,\\ 2^{(\nu_i-k_i)(|r_i|+1)}, & \text{if } k_i\geq\nu_i\,.\end{cases}\end{aligned}$$

Together with (3.2.21) and

$$\delta=\min\{1,R_i+1-r_i;\ i=1,\ldots,N\}>0$$

this last estimate results in

$$\begin{aligned}2^{\overline{\nu}\cdot\overline{r}}(\Psi^*_{\overline{\nu}}f)_{\overline{a}}(x) &\leq c\sum_{\overline{k}\in\mathbb{N}_0^N}\prod_{i:k_i\leq\nu_i}2^{(k_i-\nu_i)(R_i+1-r_i)}2^{(k_i-\nu_i)r_i}2^{\nu_i r_i}\\
&\qquad\times\prod_{i:k_i>\nu_i}2^{(\nu_i-k_i)(|r_i|+1+r_i)}2^{-(\nu_i-k_i)r_i}2^{\nu_i r_i}(\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\\
&\leq c\sum_{\overline{k}\in\mathbb{N}_0^N}(\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i:k_i\leq\nu_i}2^{(k_i-\nu_i)(R_i+1-r_i)}2^{k_i r_i}\prod_{i:k_i>\nu_i}2^{(\nu_i-k_i)1}2^{k_i r_i}\\
&\leq c\sum_{\overline{k}\in\mathbb{N}_0^N}(\Phi^*_{\overline{k}}f)_{\overline{a}}(x)\prod_{i:k_i\leq\nu_i}2^{-|k_i-\nu_i|\delta}2^{k_i r_i}\prod_{i:k_i>\nu_i}2^{-|k_i-\nu_i|\delta}2^{k_i r_i}\\
&= c\sum_{\overline{k}\in\mathbb{N}_0^N}2^{-|\overline{k}-\overline{\nu}|\delta}2^{\overline{k}\cdot\overline{r}}(\Phi^*_{\overline{k}}f)_{\overline{a}}(x),\quad \overline{\nu}\in\mathbb{N}_0^N,\quad x\in\mathbb{R}^d\,.\end{aligned}$$

The asserted inequalities now follow directly with the help of Lemma 3.1.2.

Step 2: Concerning the identities (3.2.19) and (3.2.20).

At first we deal with the fact, that the expression $\Lambda_{\overline{k}}*\Phi_{\overline{k}}*f$ is a well-defined function for every $\overline{k}\in\mathbb{N}_0^N$. We already have seen, that for the application of Lemma 3.1.1 we have to ensure $\overline{S}\geq\overline{R}+\overline{1}+\overline{d}$. Hence it follows from (3.2.2) that $\phi_{\overline{k}}\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ is a continuous function. This yields that the functions $\lambda_{\overline{k}}$ are measurable, bounded and compactly supported. From this we conclude $\lambda_{\overline{k}}\in L_1(\mathbb{R}^d)\cap L_2(\mathbb{R}^d)$. Thus $\Lambda_{\overline{k}}=\widehat{\lambda_{\overline{k}}}$ is well-defined and continuous (it is even analytic). Moreover, it follows $\Lambda_{\overline{k}}\in L_2(\mathbb{R}^d)$, and since $\Phi_{\overline{k}}\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})\hookrightarrow L_2(\mathbb{R}^d)$ we obtain from the Cauchy-Schwarz inequality that $\Lambda_{\overline{k}}*\Phi_{\overline{k}}$ is a well-defined integrable function. Due to the choice of λ^i_j it holds $\lambda_{\overline{k}}\phi_{\overline{k}}=\varphi_{\overline{k}}(\frac{3\cdot}{2\varepsilon})$, and consequently $\lambda_{\overline{k}}\phi_{\overline{k}}\in\mathcal{S}(\mathbb{R}^d)\subset X^{\overline{S}}(\mathbb{R}^{\overline{d}})$. This finally implies

$$\Lambda_{\overline{k}}*\Phi_{\overline{k}}=(2\pi)^{-d/2}(\lambda_{\overline{k}}\phi_{\overline{k}})^\wedge\in X^{\overline{S}}(\mathbb{R}^{\overline{d}})\,.$$

Similar arguments apply to $\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}}$.
Next we consider the convergence of the series in (3.2.20) in $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ for every $f \in \big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ and every $\overline{\nu} \in \mathbb{N}_0^N$. Due to the choice of the strong topology and since the Fourier transform is an isomorphism from $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ onto itself it is sufficient to show, that

$$\sum_{|\overline{k}|_1 \le M} \psi_{\overline{\nu}} \lambda_{\overline{k}} \phi_{\overline{k}} \mu \longrightarrow \psi_{\overline{\nu}} \mu\,, \quad (M \longrightarrow \infty)\,, \quad \overline{\nu} \in \mathbb{N}_0^N\,,$$

in $X^{\overline{S}}(\mathbb{R}^{\overline{d}})$ for every $\mu \in X^{\overline{S}}(\mathbb{R}^{\overline{d}})$, compare with the corresponding arguments for $\mathcal{S}'(\mathbb{R}^d)$. Hence we have to prove

$$\left\| x^\alpha D^\beta \left(\sum_{|\overline{k}|_1 \le M} \psi_{\overline{\nu}} \lambda_{\overline{k}} \phi_{\overline{k}} \mu - \psi_{\overline{\nu}} \mu \right) \middle| L_2(\mathbb{R}^d) \right\| \longrightarrow 0$$

for all multiindices α and β, where $0 \le \overline{\alpha}, \overline{\beta} \le \overline{S}$. For this we recall the fact, that for every $g \in L_2(\mathbb{R}^d)$ we have $g\mathcal{X}_{Q_{\overline{k}}} \longrightarrow g$ and $g\mathcal{X}_{Q_{\overline{k}} \setminus Q_{\overline{k}-\overline{1}}} \longrightarrow 0$ in $L_2(\mathbb{R}^d)$ for $\min_{i=1,\dots,N} k_i \longrightarrow \infty$, $\overline{k} \in \mathbb{N}_0^N$, and $Q_{\overline{k}} = \big\{x \in \mathbb{R}^d : |x^i| \le 2^{k_i},\ i = 1, \dots, N\big\}$. In both cases this follows immediately from Lebesgue's dominated convergence theorem. Since $\psi_{\overline{\nu}}$ and $\psi_{\overline{\nu}} \lambda_{\overline{k}} \phi_{\overline{k}}$ are continuous and bounded, see (3.2.2), we find that $\psi_{\overline{\nu}} \lambda_{\overline{k}} \phi_{\overline{k}} \mu$ and $\psi_{\overline{\nu}} \mu$ belong to $L_2(\mathbb{R}^d)$ for all $\overline{\nu}, \overline{k} \in \mathbb{N}_0^N$. But now the asserted convergence follows immediately from (3.2.15), (3.2.17), (3.2.18), and the above mentioned fact. The convergence of the series in (3.2.19) follows similarly.
Finally, for the step from (3.2.20) to (3.2.21) we need that

$$|(\Psi_{\overline{\nu}} * f)(y)| \le \sum_{\overline{k} \in \mathbb{N}_0^N} |(\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f)(y)| < \infty \tag{3.2.24}$$

holds for all $\overline{\nu} \in \mathbb{N}_0^N$ and almost every $y \in \mathbb{R}^d$, i.e. we need the pointwise convergence of the series. For this fix $\overline{\nu} \in \mathbb{N}_0^N$ and put $f_{\overline{k}}(y) := (\Psi_{\overline{\nu}} * \Lambda_{\overline{k}} * \Phi_{\overline{k}} * f)(y)$. Then it follows from (3.2.22)

$$|f_{\overline{k}}(y)| \le \big(\Phi_{\overline{k}}^* f\big)_{\overline{a}}(y) I_{\overline{\nu},\overline{k}}\,, \quad y \in \mathbb{R}^d\,.$$

From (3.2.23), the monotonicity of the ℓ_q-quasi-norms for $q \le 1$ or Hölder's inequality for $q > 1$ we obtain

$$\sum_{\overline{k} \in \mathbb{N}_0^N} \big\| f_{\overline{k}} \big| L_p(\mathbb{R}^d) \big\| \le c \,\big\| 2^{\overline{k} \cdot \overline{r}} (\Phi_{\overline{k}}^* f)_{\overline{a}} \big| \ell_q(L_p) \big\|\,,$$

where c may depend on $\overline{\nu}$. Hence, if the right hand side of (3.2.12) is finite, the series $\sum_{\overline{k} \in \mathbb{N}_0^N} |f_{\overline{k}}|$ converges in $L_p(\mathbb{R}^d)$.
As a series of non-negative functions it converges also pointwise almost everywhere. In those points, the series can be interpreted as an absolute convergent series of complex numbers, which implies that $\sum_{\overline{k} \in \mathbb{N}_0^N} f_{\overline{k}}$ converges pointwise almost everywhere. Together with the convergence of this series in $\big(X^{\overline{S}}(\mathbb{R}^{\overline{d}})\big)'$ it follows (3.2.24). Whereas if the right hand side of (3.2.13) is finite then we use

$$\big\| 2^{\overline{k} \cdot \overline{r}} (\Phi_{\overline{k}}^* f)_{\overline{a}} \big| \ell_{\max(p,q)}(L_p) \big\| \le \big\| 2^{\overline{k} \cdot \overline{r}} (\Phi_{\overline{k}}^* f)_{\overline{a}} \big| L_p(\ell_q) \big\|\,.$$

This follows for $p \geq q$ from the monotonicity of the ℓ_q-quasi-norm, and for $p \leq q$ it is a consequence of Minkowski's inequality (Theorem 2.3.2). With the help of the same arguments of before we obtain again (3.2.24). □

Remark 3.2.3. The conditions (3.2.9) are usually called *moment conditions*, while (3.2.10) and (3.2.11) are referred to as *Tauberian conditions*.

Proposition 3.2.3. Let $\overline{a}, \overline{r} \in \mathbb{R}^N$ and $0 < p, q \leq \infty$. Further, let

$$|\psi_0^i(t)| > 0 \quad \text{on} \quad \left\{t \in \mathbb{R}^{d_i} \,:\, |t| < \varepsilon\right\}, \tag{3.2.25}$$

$$|\psi_1^i(t)| > 0 \quad \text{on} \quad \left\{t \in \mathbb{R}^{d_i} \,:\, \varepsilon/2 < |t| < 2\varepsilon\right\} \tag{3.2.26}$$

for every $i = 1, \dots, N$ for some $\varepsilon > 0$.

(i) If $\overline{a} > \overline{d}/p$, and $\overline{S}$ is large enough, then it holds

$$\left\| \, 2^{\overline{k}\cdot\overline{r}} (\Psi_{\overline{k}}^* f)_{\overline{a}} \,\middle|\, \ell_q(L_p) \right\| \leq c \left\| \, 2^{\overline{k}\cdot\overline{r}} \Psi_{\overline{k}} * f \,\middle|\, \ell_q(L_p) \right\| \tag{3.2.27}$$

for all $f \in \left(X^{\overline{S}-\overline{d}-[(\overline{d}+\overline{1})/2]-\overline{1}}(\mathbb{R}^{\overline{d}})\right)'$.

(ii) If $\overline{a} > \overline{d}/\min(p,q)$, $p < \infty$, and $\overline{S}$ is large enough, then it holds

$$\left\| \, 2^{\overline{k}\cdot\overline{r}} (\Psi_{\overline{k}}^* f)_{\overline{a}} \,\middle|\, L_p(\ell_q) \right\| \leq c \left\| \, 2^{\overline{k}\cdot\overline{r}} \Psi_{\overline{k}} * f \,\middle|\, L_p(\ell_q) \right\| \tag{3.2.28}$$

for all $f \in \left(X^{\overline{S}-\overline{d}-[(\overline{d}+\overline{1})/2]-\overline{1}}(\mathbb{R}^{\overline{d}})\right)'$.

Proof. **Step 1:** In analogy to (3.2.14) we once again define functions $\left\{\lambda_j^i\right\}_{j=0}^\infty$, $i = 1, \dots, N$, $j \in \mathbb{N}_0$, by

$$\lambda_j^i(t) = \varphi_j^i(3t/2\varepsilon)/\psi_j^i(t)\,.$$

These functions possess the properties (3.2.16)–(3.2.18) as well, and it holds

$$\sum_{j=0}^\infty \lambda_j^i(t)\psi_j^i(t) = 1\,, \quad t \in \mathbb{R}^{d_i}\,. \tag{3.2.29}$$

Instead of (3.2.19) we now obtain the identity

$$f = (2\pi)^d \sum_{\overline{k}\in\mathbb{N}_0^N} \Lambda_{\overline{k}} * \Psi_{\overline{k}} * f\,. \tag{3.2.30}$$

We now put

$$\Lambda_{\overline{k},\overline{\nu}}(\xi) = \mathcal{F}[\lambda_{\overline{k}}(2^{-\overline{\nu}}\cdot)] \wedge (\xi) = 2^{\overline{\nu}\cdot\overline{d}} \Lambda_{\overline{k}}(2^{\overline{\nu}}\xi)\,, \quad \overline{k}, \overline{\nu} \in \mathbb{N}_0^N\,.$$

In the same way $\Psi_{\overline{k},\overline{\nu}}$ is to be understood. Here and in subsequent considerations we will use the abbreviated notation $2^{\overline{\nu}}\xi = (2^{\nu_1}\xi^1, \dots, 2^{\nu_N}\xi^N)$. Obviously we have for $\overline{k} \geq 1$ and arbitrary $\overline{\nu} \in \mathbb{N}_0^N$ always $\Psi_{\overline{k},\overline{\nu}} = \Psi_{\overline{k}+\overline{\nu}}$. To simplify notation we put

$$\psi_{\overline{k}}(2^{-\overline{\nu}}x)\psi_{\overline{\nu}} = \sigma_{\overline{k},\overline{\nu}}(x)\psi_{\overline{k}+\overline{\nu}}(x)\,, \quad \overline{k}, \overline{\nu} \in \mathbb{N}_0^N\,.$$

We find at once by the tensor product structure of the functions $\psi_{\overline{k}}$

$$\sigma_{\overline{k},\overline{\nu}}(x) = \prod_{i=1}^{N} \sigma^i_{k_i,\nu_i}(x^i)\,, \qquad \sigma^i_{k_i,\nu_i}(x^i) = \begin{cases} \psi^i_{\nu_i}(x^i)\,, & \text{if } k_i > 0\,, \\ \psi^i_0(2^{-\nu_i}x^i)\,, & \text{if } k_i = 0\,. \end{cases}$$

With the help of a dilatation $t \longmapsto 2^{-\nu_i}t$ in (3.2.29) we can rewrite (3.2.30) as follows:

$$\begin{aligned} \Psi_{\overline{\nu}} * f &= (2\pi)^{-d/2}\mathcal{F}\big[\psi_{\overline{\nu}}\mathcal{F}^{-1}f\big] = (2\pi)^{-d/2}\mathcal{F}\Bigg[\sum_{\overline{k}\in\mathbb{N}_0^N} \lambda_{\overline{k}}(2^{-\overline{\nu}}\cdot)\psi_{\overline{k}}(2^{-\overline{\nu}}\cdot)\psi_{\overline{\nu}}\mathcal{F}^{-1}f\Bigg] \\ &= (2\pi)^{-d/2}\sum_{\overline{k}\in\mathbb{N}_0^N} \mathcal{F}\big[\lambda_{\overline{k}}(2^{-\overline{\nu}}\cdot)\psi_{\overline{k}}(2^{-\overline{\nu}}\cdot)\psi_{\overline{\nu}}\mathcal{F}^{-1}f\big] \\ &= (2\pi)^{-d/2}(2\pi)^{d}\sum_{\overline{k}\in\mathbb{N}_0^N} \mathcal{F}\left[\lambda_{\overline{k}}(2^{-\overline{\nu}}\cdot)\right] * \mathcal{F}\left[\psi_{\overline{k}}(2^{-\overline{\nu}}\cdot)\psi_{\overline{\nu}}\right] * f \\ &= (2\pi)^{d/2}\sum_{\overline{k}\in\mathbb{N}_0^N} \Lambda_{\overline{k},\overline{\nu}} * \mathcal{F}\left[\sigma_{\overline{k},\overline{\nu}}\psi_{\overline{k}+\overline{\nu}}\right] * f \\ &= (2\pi)^{d/2}(2\pi)^{d/2}\sum_{\overline{k}\in\mathbb{N}_0^N} \Lambda_{\overline{k},\overline{\nu}} * \widehat{\sigma}_{\overline{k},\overline{\nu}} * \Psi_{\overline{k}+\overline{\nu}} * f\,, \quad \overline{\nu}\in\mathbb{N}_0^N\,. \end{aligned} \tag{3.2.31}$$

The issues concerning convergence, which occur in the above equations, can be treated as in the second step of the proof of Proposition 3.2.2. We now discuss the case $k_i > 0$ and $\nu_i > 0$ first. It holds

$$\begin{aligned} &\big|\big(\big[\lambda^i_{k_i}(2^{-\nu_i}\cdot)\big]^\wedge * \widehat{\sigma}^i_{k_i,\nu_i}\big)(z^i)\big| \\ &\qquad = |(\Lambda^i_{k_i+\nu_i} * \Psi^i_{\nu_i})(z^i)| = 2^{-\nu_i d_i}|(\Lambda^i_{k_i+\nu_i}(2^{-\nu_i}\cdot) * \Psi^i_{\nu_i}(2^{-\nu_i}\cdot))(2^{\nu_i}z^i)| \\ &\qquad = 2^{-\nu_i d_i}|(2^{\nu_i d_i}[\lambda^i_{k_i+\nu_i}(2^{\nu_i}\cdot)]^\wedge * 2^{\nu_i d_i}[\psi^i_{\nu_i}(2^{\nu_i}\cdot)]^\wedge)(2^{\nu_i}z^i)| \\ &\qquad = 2^{\nu_i d_i}|([\lambda^i_{k_i}]^\wedge * [\psi^i_1(2\,\cdot)]^\wedge)(2^{\nu_i}z^i)| = 2^{\nu_i d_i}2^{-d_i}|(\Lambda^i_{k_i} * \Psi^i_1(\tfrac{\cdot}{2}))(2^{\nu_i}z^i)| \\ &\qquad \le C_{N_i}2^{-d_i}2^{\nu_i d_i}\frac{2^{-k_iN_i}}{1+|2^{\nu_i}z^i|^{N_i}}\,. \end{aligned}$$

In the last line we used Lemma 3.1.1 with $K \mathrel{\widehat{=}} S_i - d_i - 1$, $M+1 = N \mathrel{\widehat{=}} N_i$, $b \mathrel{\widehat{=}} 2^{-k_i}$, $g \mathrel{\widehat{=}} \Lambda^i_1$ and $h \mathrel{\widehat{=}} \Psi^i_1(\frac{\cdot}{2})$. Due to its construction the function $\mathcal{F}g = \lambda^i_1(-\cdot)$ satisfies arbitrarily many moment conditions (see (3.2.18)). Because of the tensor product structure we further obtain

$$|(\Lambda_{\overline{k},\overline{\nu}} * \widehat{\sigma}_{\overline{k},\overline{\nu}})(z)| \le C_{\overline{N}}2^{\overline{\nu}\cdot\overline{d}}\frac{2^{-\overline{k}\cdot\overline{N}}}{\prod_{i=1}^N(1+|2^{\nu_i}z^i|^{N_i})} \tag{3.2.32}$$

for every $\overline{N} \le \overline{S} - \overline{d} - \overline{1}$, at first for $\overline{k},\overline{\nu}\in\mathbb{N}^N$. In all the other cases one gets an estimate similar to (3.2.32) with the help of analog arguments. Together with (3.2.31) we further obtain

$$\begin{aligned} |(\Psi_{\overline{\nu}} * f)(y)| &\le \sum_{\overline{k}\in\mathbb{N}_0^N}\int_{\mathbb{R}^d} |(\Lambda_{\overline{k},\overline{\nu}} * \widehat{\sigma}_{\overline{k},\overline{\nu}})(y-z)|\cdot|(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|\,dz \\ &\le C_{\overline{N}}2^{\overline{\nu}\cdot\overline{d}}\sum_{\overline{k}\in\mathbb{N}_0^N}\int_{\mathbb{R}^d}\frac{2^{-\overline{k}\cdot\overline{N}}|(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|}{\prod_{i=1}^N(1+|2^{\nu_i}(y^i-z^i)|^{N_i})}\,dz\,. \end{aligned} \tag{3.2.33}$$

Step 2: We define another maximal function by putting

$$\mathfrak{M}_{\overline{\nu},\overline{N}}(x) := \sup_{\overline{k}\in\mathbb{N}_0^N} \sup_{y\in\mathbb{R}^d} 2^{-\overline{k}\cdot\overline{N}} \frac{|(\Psi_{\overline{k+\nu}} * f)(y)|}{\prod_{i=1}^N (1+|2^{\nu_i}(y^i-x^i)|^{N_i})}$$

Now we fix an arbitrary $s \in (0,1]$. We replace $\overline{\nu}$ by $\overline{\nu}+\bar{l}$ in (3.2.33) and apply the following inequalities:

$$(1+|2^{\nu_i}(y^i-z^i)|^{N_i})(1+|2^{\nu_i}(x^i-y^i)|^{N_i}) \geq c_{N_i}^{-1}\,(1+|2^{\nu_i}(x^i-z^i)|^{N_i})\,, \tag{3.2.34}$$

$$\begin{aligned}\big|(\Psi_{\overline{k+\nu}} * f)(z)\big| &\leq \big|(\Psi_{\overline{k+\nu}} * f)(z)\big|^s \prod_{i=1}^N (1+|2^{\nu_i}(x^i-z^i)|^{N_i})^{1-s} \\ &\quad\times \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{k+\nu}} * f)(y)|^{1-s}}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})^{1-s}}\,.\end{aligned}$$

In this way we obtain

$$\begin{aligned}
&2^{-\bar{l}\cdot\overline{N}}|(\Psi_{\overline{\nu}+\bar{l}} * f)(y)| \\
&\quad\leq C_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{-\bar{l}\cdot\overline{N}} 2^{(\overline{\nu}+\bar{l})\cdot\bar{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}}|(\Psi_{\overline{k+\nu+l}} * f)(z)|}{\prod_{i=1}^N (1+|2^{\nu_i+l_i}(y^i-z^i)|^{N_i})}\,dz \\
&\quad= C_{\overline{N}} \sum_{\overline{m}\geq\bar{l}} 2^{(\overline{\nu}+\bar{l})\cdot\bar{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}}|(\Psi_{\overline{m+\nu}} * f)(z)|}{\prod_{i=1}^N (1+|2^{\nu_i+l_i}(y^i-z^i)|^{N_i})}\,dz \\
&\quad\leq C_{\overline{N}} \sum_{\overline{m}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{m})\cdot\bar{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}}|(\Psi_{\overline{m+\nu}} * f)(z)|}{\prod_{i=1}^N (1+|2^{\nu_i}(y^i-z^i)|^{N_i})}\,dz \\
&\quad\leq C_{\overline{N}} \sum_{\overline{m}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{m})\cdot\bar{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}}|(\Psi_{\overline{m+\nu}} * f)(z)|^s \prod_{i=1}^N (1+|2^{\nu_i}(x^i-z^i)|^{N_i})^{1-s}}{\prod_{i=1}^N (1+|2^{\nu_i}(y^i-z^i)|^{N_i})} \\
&\qquad\qquad\times \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{m+\nu}} * f)(y)|^{1-s}}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})^{1-s}}\,dz \\
&\quad\leq C'_{\overline{N}} \sum_{\overline{m}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{m})\cdot\bar{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}s}|(\Psi_{\overline{m+\nu}} * f)(z)|^s \prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-z^i)|^{N_i})^s} \\
&\qquad\qquad\times 2^{-\overline{m}\cdot\overline{N}(1-s)} \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{m+\nu}} * f)(y)|^{1-s}}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})^{1-s}}\,dz
\end{aligned}$$

Dividing by $\prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})$ and taking the supremum over $y\in\mathbb{R}^d$ and $\bar{l}\in\mathbb{N}_0^N$, we can further estimate

$$\begin{aligned}\mathfrak{M}_{\overline{\nu},\overline{N}}(x) &\leq C'_{\overline{N}} \sum_{\overline{m}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{m})\cdot\bar{d}} \left(\sup_{y\in\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}}|(\Psi_{\overline{m+\nu}} * f)(y)|}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-y^i)|^{N_i})}\right)^{1-s} \\ &\qquad\times \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}s}|(\Psi_{\overline{m+\nu}} * f)(z)|^s}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-z^i)|^{N_i})^s}\,dz\end{aligned}$$

$$\leq C'_{\overline{N}} \sum_{\overline{m}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{m})\cdot\overline{d}} \mathfrak{M}_{\overline{\nu},\overline{N}}(x)^{1-s} \int_{\mathbb{R}^d} \frac{2^{-\overline{m}\cdot\overline{N}s}|(\Psi_{\overline{m}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1+|2^{\nu_i}(x^i-z^i)|^{N_i})^s}\, dz\,. \qquad (3.2.35)$$

In order to be allowed to divide by $\mathfrak{M}_{\overline{\nu},\overline{N}}(x)^{1-s}$ we need to show, that this maximal function is finite.

Step 3: We now assume $f \in \big(X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big)'$, where $\overline{T} \leq \overline{S}$, and at first we have a closer look on the convolution. We find for $\overline{\nu} \geq 1$

$$\begin{aligned}\big|(\Psi_{\overline{\nu}} * f)(x)\big| &= \big|f\big((\Psi_{\overline{\nu}}(y-\cdot)\big)\big| \leq \big\| f \,\big|\big(X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big)'\big\| \cdot \big\|\mathcal{F}\psi_{\overline{\nu}}(-\cdot+y)\big|X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big\| \\ &= \big\| f \,\big|\big(X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big)'\big\| 2^{(\overline{\nu}-\overline{1})\cdot\overline{d}} \big\|\mathcal{F}\psi_{\overline{1}}\big(2^{\overline{\nu}-\overline{1}}(\cdot-y)\big)\big|X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big\|\,.\end{aligned}$$

For the last factor we further obtain

$$\begin{aligned}&\big\|x^\alpha D^\beta\big[\mathcal{F}\psi_{\overline{1}}(2^{\overline{\nu}-\overline{1}}(x-y))\big]\big|L_2(\mathbb{R}^d)\big\| \\ &\quad= 2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}\big\|x^\alpha (D^\beta\mathcal{F}\psi_{\overline{1}})\big(2^{\overline{\nu}-\overline{1}}(x-y)\big)\big|L_2(\mathbb{R}^d)\big\| \\ &\quad= 2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}2^{(-\overline{\nu}+\overline{1})\cdot\overline{d}/2}\big\|\big(2^{-\overline{\nu}+\overline{1}}z+y\big)^\alpha D^\beta\mathcal{F}\psi_{\overline{1}}\big|L_2(\mathbb{R}^d)\big\| \\ &\quad\leq c\,2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}2^{(-\overline{\nu}+\overline{1})\cdot\overline{d}/2}\sum_{\gamma\leq\alpha}\big\|2^{(-\overline{\nu}+\overline{1})\cdot\overline{\gamma}}z^\gamma y^{\alpha-\gamma}D^\beta\mathcal{F}\psi_{\overline{1}}\big|L_2(\mathbb{R}^d)\big\| \\ &\quad= c\,2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}2^{(-\overline{\nu}+\overline{1})\cdot\overline{d}/2}\sum_{\gamma\leq\alpha}\big|y^{\alpha-\gamma}\big|\,2^{(-\overline{\nu}+\overline{1})\cdot\overline{\gamma}}\,\big\|z^\gamma D^\beta\mathcal{F}\psi_{\overline{1}}\big|L_2(\mathbb{R}^d)\big\| \\ &\quad\leq c\,2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}2^{(-\overline{\nu}+\overline{1})\cdot\overline{d}/2}\sum_{\gamma\leq\alpha}\prod_{i=1}^N|y^i|^{|\alpha^i-\gamma^i|}\cdot\big\|\,\mathcal{F}\psi_{\overline{1}}\,\big|X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big\| \\ &\quad\leq c'\,2^{(\overline{\nu}-\overline{1})\cdot\overline{\beta}}2^{(-\overline{\nu}+\overline{1})\cdot\overline{d}/2}\prod_{i=1}^N|y^i|^{T_i}\cdot\big\|\,\mathcal{F}\psi_{\overline{1}}\,\big|X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big\|\,,\end{aligned}$$

where α and β are multiindices with $\overline{\alpha},\overline{\beta} \leq \overline{T}$. Combining these two estimates we end up with

$$\big|(\Psi_{\overline{\nu}} * f)(x)\big| \leq c\,2^{\overline{\nu}\cdot(\overline{T}+\overline{d}/2)}\big\| f \,\big|\big(X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big)'\big\| \cdot \big\|\,\mathcal{F}\psi_{\overline{1}}\,\big|X^{\overline{T}}(\mathbb{R}^{\overline{d}})\big\| \prod_{i=1}^N |y^i|^{T_i}\,.$$

We further find

$$\sup_{y\in\mathbb{R}^d}\frac{|(\Psi_{\overline{k}+\overline{\nu}} * f)(y)|}{\prod_{i=1}^N(1+|2^{\nu_i}(y^i-x^i)|^{N_i})} \leq c_f \sup_{y\in\mathbb{R}^d}\prod_{i=1}^N\frac{2^{(\overline{k}+\overline{\nu})\cdot(\overline{T}+\overline{d}/2)}|y^i|^{T_i}}{(1+|2^{\nu_i}(y^i-x^i)|^{N_i})}\,,$$

which is finite if, and only if, $\overline{N} \geq \overline{T}$. Moreover, the maximal function $\mathfrak{M}_{\overline{\nu},\overline{N}}(x)$ turns out to be finite if, and only if, $\overline{N} \geq \overline{T} + \overline{d}/2$. Though the constant does depend on f, the condition on $\overline{N}$ does not, and this is sufficient at this point. Altogether we find

$$\mathfrak{M}_{\overline{\nu},\overline{N}}(x)^s \leq C_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{k})\cdot\overline{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s}|(\Psi_{\overline{k}+\overline{\nu}} * f)(y)|^s}{\prod_{i=1}^N(1+|2^{\nu_i}(x^i-y^i)|^{N_i})^s}\, dy \qquad (3.2.36)$$

for all $\overline{N} \geq \overline{T} + \overline{d}/2$. Keeping in mind the condition $\overline{N} \leq \overline{S} - \overline{d} - \overline{1}$ from the beginning, this leads to the restriction $\overline{T} + \overline{d}/2 \leq \overline{S} - \overline{d} - \overline{1}$ and hence $f \in \left(X^{\overline{S}-\overline{d}-[(\overline{d}+\overline{1})/2]-\overline{1}}(\mathbb{R}^{\overline{d}})\right)'$. If now we apply (3.2.36) for $\overline{N}^* = \max(\overline{N}, \overline{T} + \overline{d}/2)$, then we further get

$$\left|(\Psi_{\overline{\nu}} * f)(x)\right|^s \leq C_{\overline{N}^*} \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{k})\cdot\overline{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s} |(\Psi_{\overline{k}+\overline{\nu}} * f)(y)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{N_i})^s} \, dy \,, \tag{3.2.37}$$

and it follows that we can drop the lower condition on $\overline{N}$. This follows from the observation, that the right hand side of (3.2.36) increases if the components of $\overline{N}$ decrease. Moreover, we shall mention that (3.2.36) has a counterpart also in case $s > 1$, and indeed with a much more direct proof. To see this we consider (3.2.33) with $\overline{N} + \overline{d}$ instead of $\overline{N}$, divide by $\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{N_i})$, and apply Hölder's inequality, first for series and afterwards for integrals, both times with respect to $\frac{1}{s} + \frac{1}{s'} = 1$. In this way we obtain for all $\overline{N} \in \mathbb{N}_0^N$

$$\begin{aligned}
\left|(\Psi_{\overline{\nu}} * f)(x)\right| &\leq C_{\overline{N}} \sum_{\overline{k} \in \mathbb{N}_0^N} \int_{\mathbb{R}^d} \frac{2^{\overline{\nu}\cdot\overline{d}} 2^{-\overline{k}\cdot\overline{N}} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i + d_i})} \, dz \\
&\leq C_{\overline{N}} \sum_{\overline{k} \in \mathbb{N}_0^N} \left(\int_{\mathbb{R}^d} \frac{2^{\overline{\nu}\cdot\overline{d}s} 2^{-\overline{k}\cdot\overline{N}s} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i})^s} \, dz \right)^{1/s} \\
&\qquad\qquad \times \left(\int_{\mathbb{R}^d} \frac{dz}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{d_i})^{s'}} \right)^{1/s'} \\
&= C'_{\overline{N}} \sum_{\overline{k} \in \mathbb{N}_0^N} \left(\int_{\mathbb{R}^d} \frac{2^{\overline{\nu}\cdot\overline{d}s} 2^{-\overline{k}\cdot\overline{N}s} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i})^s} \, dz \right)^{1/s} 2^{-\overline{\nu}\cdot\overline{d}/s'} \\
&\leq C'_{\overline{N}} \sum_{\overline{k} \in \mathbb{N}_0^N} \left(\int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s} 2^{-\overline{k}\cdot\overline{d}} 2^{\overline{k}\cdot\overline{d}} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i})^s} \, dz \right)^{1/s} 2^{\overline{\nu}\cdot\overline{d}/s} \\
&\leq C''_{\overline{N}} \left(\sum_{\overline{k} \in \mathbb{N}_0^N} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s} 2^{(\overline{k}+\overline{\nu})\cdot\overline{d}} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i})^s} \, dz \right)^{1/s} \left(\sum_{\overline{k} \in \mathbb{N}_0^N} 2^{-\overline{k}\cdot\overline{d}s'/s} \right)^{1/s'} \\
&\leq \widetilde{C}_{\overline{N}} \left(\sum_{\overline{k} \in \mathbb{N}_0^N} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s} 2^{(\overline{k}+\overline{\nu})\cdot\overline{d}} |(\Psi_{\overline{k}+\overline{\nu}} * f)(z)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{N_i})^s} \, dz \right)^{1/s} .
\end{aligned}$$

Step 4: First we remark, that (3.2.37) can immediately be strengthened by dividing by $\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - z^i)|^{a_i})$, taking the supremum over $x \in \mathbb{R}^d$ and using the inequality (3.2.34). We obtain for all $\overline{a} \leq \overline{N}$

$$\left|(\Psi^*_{\overline{\nu}} f)_{\overline{a}}(x)\right|^s \leq C_{\overline{N}^*} \sum_{\overline{k} \in \mathbb{N}_0^N} 2^{(\overline{\nu}+\overline{k})\cdot\overline{d}} \int_{\mathbb{R}^d} \frac{2^{-\overline{k}\cdot\overline{N}s} |(\Psi_{\overline{k}+\overline{\nu}} * f)(y)|^s}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{a_i})^s} \, dy \,, \tag{3.2.38}$$

after renaming $z \mapsto x$. Now we choose some $s > 0$, such that $d_i/a_i < s < p$ (or

$d_i/a_i < s < \min(p,q)$ for F-spaces, respectively) for every $i = 1, \ldots, N$. Then it holds

$$\frac{1}{\prod_{i=1}^N (1+|z^i|^{a_i})^s} \in L_1(\mathbb{R}^d)\,,$$

and from Lemma 2.3.1 for the Hardy-Littlewood maximal operator $\overline{M}$ we obtain

$$(\Psi^*_{\overline{\nu}} f)_{\overline{a}}(x)^s \leq c_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{-\overline{k}\cdot\overline{N}s}\, \overline{M}(|\Psi_{\overline{k}+\overline{\nu}} * f|^s)(x)\,. \tag{3.2.39}$$

We now choose $\overline{N} > 0$, such that $\overline{N} > -\overline{r}$. This is possible, if we ensure $\overline{S}$ to be large enough, in particular $\overline{S} - \overline{d} - \overline{1} + \overline{r} > 0$. Then we put

$$g_{\overline{k}}(x) = 2^{\overline{k}\cdot\overline{r}s}\, \overline{M}(|\Psi_{\overline{k}} * f|^s)(x)\,.$$

Now it follows from (3.2.39)

$$\begin{aligned}
G_{\overline{\nu}}(x) &= 2^{\overline{\nu}\cdot\overline{r}s} (\Psi^*_{\overline{\nu}} f)_{\overline{a}}(x)^s \\
&\leq c_{\overline{N}} 2^{\overline{\nu}\cdot\overline{r}s} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{-\overline{k}\cdot\overline{L}s}\, \overline{M}(|\Psi_{\overline{k}+\overline{\nu}} * f|^s)(x) = c_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot\overline{r}s} 2^{-\overline{k}\cdot\overline{L}s} 2^{-(\overline{k}+\overline{\nu})\cdot\overline{r}s} g_{\overline{k}+\overline{\nu}}(x) \\
&= c_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{\overline{k}\cdot(-\overline{r}-\overline{L})s} g_{\overline{k}+\overline{\nu}}(x) = c_{\overline{N}} \sum_{\overline{k}\geq\overline{\nu}} 2^{s(\overline{k}-\overline{\nu})\cdot(-\overline{L}-\overline{r})} g_{\overline{k}}(x)\,.
\end{aligned}$$

We choose some δ, such that $0 < \delta < \min\{N_i + r_i : i = 1, \ldots, N\}$. Then it follows $(\overline{k}-\overline{\nu})(-\overline{N}-\overline{r}) \leq -|\overline{k}-\overline{\nu}|\delta$ for $\overline{k} \geq \overline{\nu}$, and hence we obtain

$$G_{\overline{\nu}}(x) \leq c_{\overline{N}} \sum_{\overline{k}\geq\overline{\nu}} 2^{-s|\overline{k}-\overline{\nu}|\delta} g_{\overline{k}}(x) \leq c_{\overline{N}} \sum_{\overline{k}\in\mathbb{N}_0^N} 2^{-s|\overline{k}-\overline{\nu}|\delta} g_{\overline{k}}(x)\,.$$

An application of Lemma 3.1.2 for the $\ell_{q/s}(L_{p/s})$-norms now results in

$$\left\| 2^{\overline{k}\cdot\overline{r}s} \left[(\Psi^*_{\overline{k}} f)_{\overline{a}}(x)\right]^s \middle| \ell_{q/s}(L_{p/s}) \right\| \leq c \left\| 2^{\overline{k}\cdot\overline{r}s}\, \overline{M}(|\Psi_{\overline{k}} * f|^s)(x) \middle| \ell_{q/s}(L_{p/s}) \right\|, \tag{3.2.40}$$

and for the $L_{p/s}(\ell_{q/s})$-norms we find correspondingly

$$\left\| 2^{\overline{k}\cdot\overline{r}s} \left[(\Psi^*_{\overline{k}} f)_{\overline{a}}(x)\right]^s \middle| L_{p/s}(\ell_{q/s}) \right\| \leq c \left\| 2^{\overline{k}\cdot\overline{r}s}\, \overline{M}(|\Psi_{\overline{k}} * f|^s)(x) \middle| L_{p/s}(\ell_{q/s}) \right\|. \tag{3.2.41}$$

From these two estimates we obtain the assertion by standard arguments. In particular, in the first case we rewrite the left hand side of (3.2.40), and use the classical Hardy-Littlewood maximal inequality (see (2.3.1)) we remind on $s < p$). In this way we find

$$\left\| 2^{\overline{k}\cdot\overline{r}} (\Psi^*_{\overline{k}} f)_{\overline{a}}(x) \middle| \ell_q(L_p) \right\| \leq c \left\| 2^{\overline{k}\cdot\overline{r}} (\Psi_{\overline{k}} * f)(x) \middle| \ell_q(L_p) \right\|.$$

In the second case we rewrite the left hand side of (3.2.41) accordingly, and apply Proposition 2.3.1 (here we remind on $s < \min(p,q)$) to finally obtain

$$\left\| 2^{\overline{k}\cdot\overline{r}} (\Psi^*_{\overline{k}} f)_{\overline{a}}(x) \middle| L_p(\ell_q) \right\| \leq c \left\| 2^{\overline{k}\cdot\overline{r}} (\Psi_{\overline{k}} * f)(x) \middle| L_p(\ell_q) \right\|,$$

which completes the proof. □

Remark 3.2.4. Steps 2 and 3 differ essentially from the proofs in [63] and [94], since both had a gap in their argumentation. Concerning the finiteness of the maximal function they only referred to the order of the given distribution f. Though this referral is not altogether wrong, the argument is far from being complete. This can be easily seen using (in $d = 1$) $\Psi_0 = \Psi_1 = e^{-t^2} \in \mathcal{S}(\mathbb{R})$ and the tempered distributions $|t|^n \in \mathcal{S}'(\mathbb{R})$, where $n > a$. Then $(\Psi_0^* |t|^n)_a(x) = \infty$ for all $x \in \mathbb{R}$.
However, this oversight can be corrected using arguments as above. Though the case $\overline{S} = \infty$ is different from the case $\overline{S} < \infty$, we won't repeat the arguments. The crucial estimates were presented in the proof, and for the arguments leading to the finiteness of the maximal function we refer to [64] and [92]. While their setting is slightly different, the arguments remain identical.

3.3 Local means

The following theorem is a direct consequence of the last two propositions. The characterization of the Besov- and Triebel-Lizorkin spaces (in terms of the maximal operator of Peetre) contained therein is the main result of this section. Thereafter we will discuss reformulations of this theorem and state some corollaries.

Theorem 3.3.1. Let $0 < p, q \leq \infty$, $\overline{r}, \overline{a} \in \mathbb{R}^N$, and $\overline{R}, \overline{S} \in \mathbb{N}_0^N$, where $\overline{r} \leq \overline{R} + \overline{1}$.

(i) Moreover, we assume $\overline{a} > \overline{d}/p$. If $\overline{S} > \overline{R}$ is large enough and

$$D^\alpha \psi_1^i(0) = 0\,, \quad \alpha \in \mathbb{N}_0^{d_i}\,, |\alpha| \leq R_i\,, \quad i = 1, \dots, N\,, \tag{3.3.1}$$

and if additionally

$$|\psi_0^i(t)| > 0 \quad \text{on} \quad \big\{t \in \mathbb{R}^{d_i} \,:\, |t| < \varepsilon\big\}\,, \tag{3.3.2}$$
$$|\psi_1^i(t)| > 0 \quad \text{on} \quad \big\{t \in \mathbb{R}^{d_i} \,:\, \varepsilon/2 < |t| < 2\varepsilon\big\} \tag{3.3.3}$$

is satisfied for some $\varepsilon > 0$, then it holds

$$\big\|\, f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\big\| \sim \big\|\, 2^{\overline{k}\cdot\overline{r}}(\Psi^*_{\overline{k}}f)_{\overline{a}}\,|\ell_q(L_p)\big\| \sim \big\|\, 2^{\overline{k}\cdot\overline{r}}(\Psi_{\overline{k}} * f)\,|\ell_q(L_p)\big\| \tag{3.3.4}$$

for all $f \in \big(X^{\overline{S}-\overline{d}-[(\overline{d}+\overline{1})/2]-\overline{1}}(\mathbb{R}^{\overline{d}})\big)'$.

(ii) Now we suppose $p < \infty$ and $\overline{a} > \overline{d}/\min(p,q)$. If $\overline{S} > \overline{R}$ is large enough, and if the conditions (3.3.1)–(3.3.3) are satisfied, then it holds

$$\big\|\, f \,|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \sim \big\|\, 2^{\overline{k}\cdot\overline{r}}(\Psi^*_{\overline{k}}f)_{\overline{a}}\,|L_p(\ell_q)\big\| \sim \big\|\, 2^{\overline{k}\cdot\overline{r}}(\Psi_{\overline{k}} * f)\,|L_p(\ell_q)\big\| \tag{3.3.5}$$

for all $f \in \big(X^{\overline{S}-\overline{d}-[(\overline{d}+\overline{1})/2]-\overline{1}}(\mathbb{R}^{\overline{d}})\big)'$.

Proof. We restrict ourselves to the B-case, the argumentation for F-spaces is identical. The decomposition of unity $\varphi = \big(\varphi_{\overline{k}}\big)_{\overline{k}\in\mathbb{N}_0^N}$ fixed at the beginning of the section fulfills arbitrarily many moment conditions due to its construction. On the other hand, we cannot satisfy the Tauberian conditions with this system. Hence we consider another decomposition of the same type, where we require from the basic function $\widetilde{\varphi}_0^i \in \mathcal{S}(\mathbb{R}^{d_i})$, that

$$\begin{aligned}
&\widetilde{\varphi}_0^i(t) = 1 \quad \text{on } \left\{t \in \mathbb{R}^{d_i} \,:\, |t| \le \tfrac{3}{4}\right\},\\
&0 < \widetilde{\varphi}_0^i(t) < 1 \quad \text{on } \left\{t \in \mathbb{R}^{d_i} \,:\, \tfrac{3}{4} < |t| < \tfrac{3}{2}\right\},\\
&\operatorname{supp} \widetilde{\varphi}_0^i \subset \left\{t \in \mathbb{R}^{d_i} \,:\, |t| \le \tfrac{3}{2}\right\}
\end{aligned}$$

holds for $i = 1, \ldots, N$. Then we find that $\widetilde{\varphi}_0^i(-\,\cdot)$ and $\widetilde{\varphi}_1^i(-\,\cdot)$, $i = 1, \ldots, N$, satisfy the Tauberian conditions for $\varepsilon = \frac{3}{2}$.
The right hand side equivalence in (3.3.4) follows immediately from Proposition 3.2.3 and the definition of the maximal function. For the left hand part we first apply Proposition 3.2.2 for $\phi_j^i \mathrel{\widehat{=}} \widetilde{\varphi}_j^i(-\,\cdot)$, i.e. $\Phi_j^i \mathrel{\widehat{=}} \mathcal{F}_{d_i}^{-1}\widetilde{\varphi}_j^i$, and obtain

$$\left\|2^{\overline{\nu}\cdot\overline{r}}\left(\Psi_{\overline{\nu}}^* f\right)_{\overline{a}}\right\|\ell_q(L_p)\right\| \le c\left\|2^{\overline{\nu}\cdot\overline{r}}\left(\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{\nu}}^* f\right)_{\overline{a}}\right\|\ell_q(L_p)\right\|.$$

Now Proposition 3.2.3, used with $\psi_j^i \mathrel{\widehat{=}} \widetilde{\varphi}_j^i(-\cdot)$, yields

$$\begin{aligned}
\left\|2^{\overline{\nu}\cdot\overline{r}}\left(\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{\nu}}^* f\right)_{\overline{a}}\right|\ell_q(L_p)\right\| &\le c\left\|2^{\overline{\nu}\cdot\overline{r}}\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{\nu}} * f\right|\ell_q(L_p)\right\|\\
&= c\,(2\pi)^{-d/2}\left\|2^{\overline{\nu}\cdot\overline{r}}\mathcal{F}^{-1}\left[\widetilde{\varphi}_{\overline{\nu}}\mathcal{F}f\right]\right|\ell_q(L_p)\right\| = c'\left\|\,f\,\right|S_{p,q}^{\overline{r}}B(\mathbb{R}^{\overline{d}})\right\|.
\end{aligned}$$

On the other hand, by the definition of the maximal function and once more Proposition 3.2.2, this time applied with $\psi_j^i \mathrel{\widehat{=}} \widetilde{\varphi}_j^i(-\cdot)$ and $\phi_j^i \mathrel{\widehat{=}} \psi_j^i$, we have

$$\left\|2^{\overline{\nu}\cdot\overline{r}}\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{\nu}} * f\right|\ell_q(L_p)\right\| \le \left\|2^{\overline{\nu}\cdot\overline{r}}\left(\mathcal{F}^{-1}\widetilde{\varphi}_{\overline{\nu}}^* f\right)_{\overline{a}}\right|\ell_q(L_p)\right\| \le c\left\|2^{\overline{\nu}\cdot\overline{r}}\left(\Psi_{\overline{\nu}}^* f\right)_{\overline{a}}\right|\ell_q(L_p)\right\|.$$

This proves the assertion. □

Remark 3.3.1. If we interpret the convolution appropriately,

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

then the last theorem can be seen as a characterization of Besov and Triebel-Lizorkin spaces with the help of local means, the above assumptions correspond to conditions of the Fourier transform of the kernels. On this formulation of Theorem 3.3.1 the subsequent decomposition theorems in Chapter 4 are based.

The rest of this section deals with the construction of suitable kernels of local means. This will be done in two different ways. At first, we derive from Theorem 3.3.1 a statement on a type of local means, which were considered before, e.g. in [84]. Later on, we follow up with a second construction based on [85], Section 12.8.

Theorem 3.3.2. Let $0 < p, q \le \infty$ ($p < \infty$ for F-spaces) and $\overline{r} \in \mathbb{R}^N$. Furthermore, let $\overline{S}^1, \overline{S}^2 \in \mathbb{N}_0^N$, where $\overline{S}^1 - \overline{S}^2 > \frac{3}{2}\overline{d} + \overline{1}$. Moreover, let $\overline{M}, \overline{R} \in \mathbb{N}_0^N$ vectors of non-negative integers, where $2\overline{M} > \overline{R} > \overline{r}$. Let $k_0^1, \ldots, k_0^N, k^1, \ldots, k^N$ complex-valued functions, where $k_0^i, k^i \in X^{S_i^1}(\mathbb{R}^{d_i})$ and $\operatorname{supp} k_0^i, \operatorname{supp} k^i \subset \{t \in \mathbb{R}^{d_i} : |t| < 1\}$, which additionally satisfy

$$(\mathcal{F}_{d_i}k_0^i)(0) \ne 0, \qquad (\mathcal{F}_{d_i}k^i)(0) \ne 0, \qquad i = 1, \ldots, N. \tag{3.3.6}$$

We define for $x^i = (x^i_1, \ldots, x^i_{d_i})$ and $\Delta_i = \sum_{j=1}^{d_i} \frac{\partial^2}{\partial {x^i_j}^2}$

$$k^i_m(t) = 2^{md_i}(\Delta_i^{M_i} k^i)(2^m t), \quad i = 1, \ldots, N, \quad m \in \mathbb{N}, \quad t \in \mathbb{R}^{d_i}.$$

As usual we denote by $k_{\overline{\nu}}(x) = k^1_{\nu_1}(x^1) \cdots k^N_{\nu_N}(x^N)$, $\overline{\nu} = (\nu_1, \ldots, \nu_N) \in \mathbb{N}_0^N$, the tensor product of these functions. The corresponding local means are defined by

$$k_{\overline{\nu}}(f)(x) = \int_{\mathbb{R}^d} k_{\overline{\nu}}(y) f(x+y) dy\,, \quad \overline{\nu} \in \mathbb{N}_0^N\,, x \in \mathbb{R}^d\,, \tag{3.3.7}$$

appropriately interpreted for arbitrary $f \in (X^{\overline{S}^1}(\mathbb{R}^{\overline{d}}))'$. If $\overline{S}^2$ is large enough, then it holds

$$\left\| \, 2^{\overline{\nu}\cdot\overline{r}} k_{\overline{\nu}}(f) \, \middle| L_p(\ell_q) \right\| \sim \left\| \, f \, \middle| S^{\overline{r}}_{p,q} F(\mathbb{R}^{\overline{d}}) \right\|, \quad f \in \left(X^{\overline{S}^2}(\mathbb{R}^{\overline{d}})\right)', \tag{3.3.8}$$

and

$$\left\| \, 2^{\overline{\nu}\cdot\overline{r}} k_{\overline{\nu}}(f) \, \middle| \ell_q(L_p) \right\| \sim \left\| \, f \, \middle| S^{\overline{r}}_{p,q} B(\mathbb{R}^{\overline{d}}) \right\|, \quad f \in \left(X^{\overline{S}^2}(\mathbb{R}^{\overline{d}})\right)'. \tag{3.3.9}$$

Proof. We put for $i = 1, \ldots, N$

$$\psi^i_0 = \mathcal{F}^{-1}_{d_i} k^i_0 \qquad \text{and} \qquad \psi^i_1 = \left(\mathcal{F}^{-1}_{d_i}(\Delta_i^{M_i} k^i)\right)(\tfrac{\cdot}{2}) = 2^{d_i} \mathcal{F}^{-1}_{d_i}\left[(\Delta_i^{M_i} k^i)(2\,\cdot)\right].$$

If $S^1_i > d_i/2$ then $\mathcal{F}_{d_i} k^i_0$ is a continuous function due to (3.2.2), and the condition (3.3.2) follows from (3.3.6). The condition (3.3.3) can be obtained from

$$|\psi^i_1(t)| = \left|\mathcal{F}^{-1}_{d_i}\left[\Delta_i^{M_i} k^i\right](t/2)\right| = \left(|t/2|^2\right)^{M_i} \left|\left(\mathcal{F}^{-1}_{d_i} k^i\right)(t/2)\right|.$$

The first factor is strictly positive outside of the origin, the second one is non-vanishing in a (small) neighbourhood of $t = 0$ due to continuity. Hence the Tauberian conditions are fulfilled for sufficiently small $\varepsilon > 0$. Finally, the moment condition (3.3.1) for $\alpha \in \mathbb{N}_0^{d_i}$, $|\alpha| \le R_i$, follows from

$$\begin{aligned} D^\alpha \psi^i_1(0) &= D^\alpha \left[\mathcal{F}^{-1}_{d_i}\left(\Delta_i^{M_i} k^i\right)\right](0) = D^\alpha \left[\left(-|t|^2\right)^{M_i} \mathcal{F}^{-1}_{d_i} k^i\right](0) \\ &= (-1)^{M_i} \sum_{\beta \le \alpha} c_{\alpha,\beta} D^\beta \left[|t|^{2M_i}\right] D^{\alpha-\beta}\left[\mathcal{F}^{-1}_{d_i} k^i\right](0)\,. \end{aligned}$$

On the one hand we have $|\beta| \le |\alpha| \le R_i$, on the other hand $|t|^{2M_i}$ is a linear combination of monomials of total degree $2M_i > R_i$. Their derivatives are monomials of total degree $2M_i - |\beta| > 0$, hence all derivatives $D^\beta\left[|t|^{2M_i}\right]$ vanish for $t = 0$.
If we now define $\psi_{\overline{\nu}}$, $\overline{\nu} \in \mathbb{N}_0^N$, as in (3.2.6)–(3.2.8), we obtain

$$\begin{aligned} (2\pi)^{-d/2}(\psi_{\overline{\nu}} \widehat{f})^\vee(x) &= \int_{\mathbb{R}^d} (\mathcal{F}^{-1}\psi_{\overline{\nu}})(y) f(x-y) dy = \int_{\mathbb{R}^d} (\mathcal{F}\psi_{\overline{\nu}})(y) f(x+y) dy \\ &= \int_{\mathbb{R}^d} \left(\prod_{i=1}^N (\mathcal{F}_{d_i} \psi^i_{\nu_i})(y^i) \right) f(x+y) dy\,. \end{aligned} \tag{3.3.10}$$

Moreover, for $\nu_i = 0$ we re-obtain $(\mathcal{F}_{d_i}\psi_0^i)(y^i) = k_0^i(y^i)$, and for $\nu_i \geq 1$ we get

$$
\begin{aligned}
(\mathcal{F}_{d_i}\psi_{\nu_i}^i)(y^i) &= (\mathcal{F}_{d_i}[\psi_1^i(2^{-\nu_i+1}\cdot)])(y^i) = 2^{(\nu_i-1)d_i}(\mathcal{F}_{d_i}\psi_1^i)(2^{\nu_i-1}y^i) \\
&= 2^{\nu_i d_i}(\Delta_i^{M_i}k^i)(2^{\nu_i}y^i) = k_{\nu_i}^i(y^i)\,.
\end{aligned}
$$

Inserting this calculation into (3.3.10) results in

$$
(\psi_{\overline{\nu}}\widehat{f})^\vee(x) = (2\pi)^{d/2}\int_{\mathbb{R}^d} k_{\overline{\nu}}(y)f(x+y)dy\,, \qquad \overline{\nu} \in \mathbb{N}_0^N\,, x \in \mathbb{R}^d\,.
$$

The assertion of the theorem now follows immediately from Theorem 3.3.1. □

Before we proceed to other version of Theorem 3.3.1 we introduce some notation first which will be of importance in the sequel. For $\overline{\nu} \in \mathbb{N}_0^N$ and $m \in \mathbb{Z}^d$, $m = (m^1, \ldots, m^N)$, $m_i \in \mathbb{Z}^{d_i}$, we denote by $Q_{\overline{\nu},m}$ the rectangle with centre in $2^{-\overline{\nu}}m = (2^{-\nu_1}m^1, \ldots, 2^{-\nu_N}m^N)$, sides parallel to the coordinate axes and with side lengths $2^{-\nu_1}, \ldots, 2^{-\nu_N}$. Explicitly, this means

$$
Q_{\overline{\nu},m} = \left\{x \in \mathbb{R}^d \,:\, |x^i - 2^{-\nu_i}m^i|_\infty \leq 2^{-\nu_i-1}\,,\, i = 1, \ldots, N\right\}, \qquad \overline{\nu} \in \mathbb{N}_0^N\,, m \in \mathbb{Z}^d\,.
$$

An important observation about these rectangles is the fact that they can be written as products of lower-dimensional cubes,

$$
Q_{\overline{\nu},m} = Q^1_{\nu_1,m^1} \times \cdots \times Q^N_{\nu_N,m^N}\,, \quad Q^i_{\nu_i,m^i} = \left\{t \in \mathbb{R}^{d_i} \,:\, |t - 2^{-\nu_i}m^i|_\infty \leq 2^{-\nu_i-1}\right\}. \quad (3.3.11)
$$

Moreover, for $\overline{\gamma} > 0$ we denote by $\overline{\gamma}Q_{\overline{\nu},m}$ the rectangle concentric with $Q_{\overline{\nu},m}$ with side lengths $\gamma_1 2^{-\nu_1}, \ldots, \gamma_N 2^{-\nu_N}$ and sides parallel to the coordinate axes. Finally, the notion $\gamma Q_{\overline{\nu},m}$ refers to the case $\overline{\gamma} = \gamma\overline{1}$, $\gamma > 0$.
From the chosen definition of the Peetre maximal function we now conclude

$$
\begin{aligned}
(\Psi_{\overline{\nu}}^* f)_{\overline{a}}(x) &= \sup_{y\in\mathbb{R}^d} \frac{|(\Psi_{\overline{\nu}} * f)(y)|}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{a_i})} \\
&\geq \sup_{x-y\in\overline{\gamma}Q_{\overline{\nu},0}} \frac{|(\Psi_{\overline{\nu}} * f)(y)|}{\prod_{i=1}^N (1 + |2^{\nu_i}(x^i - y^i)|^{a_i})} \geq \sup_{x-y\in\overline{\gamma}Q_{\overline{\nu},0}} \frac{|(\Psi_{\overline{\nu}} * f)(y)|}{\prod_{i=1}^N (1 + (\frac{1}{2}\gamma_i\sqrt{d_i})^{a_i})} \\
&= c(\overline{a}, \overline{\gamma}, \overline{d}) \sup_{x-y\in\overline{\gamma}Q_{\overline{\nu},0}} |(\Psi_{\overline{\nu}} * f)(y)| \geq c\,|(\Psi_{\overline{\nu}} * f)(x)|
\end{aligned}
$$

with some constant c independent of x and $\overline{\nu}$. Thereby we used the observation $x - y \in \overline{\gamma}Q_{\overline{\nu},0} \iff |x^i - y^i|_\infty \leq \gamma_i 2^{-\nu_i-1}$, and hence $|2^{\nu_i}(x^i - y^i)| \leq \frac{1}{2}\gamma_i\sqrt{d_i}$. This simple observation together with the Theorems 3.3.1 and 3.3.2 now yields the following proposition.

Proposition 3.3.1. Let $\overline{r} \in \mathbb{R}^N$ and $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces). Furthermore, let $\overline{M}, \overline{R} \in \mathbb{N}_0^N$, where $2\overline{M} > \overline{R} > \overline{r}$,, and let $\overline{S}^1, \overline{S}^2 \in \mathbb{N}_0^N$ and $k_{\overline{\nu}}$ be as in Theorem 3.3.2. Then it holds for every $\gamma > 0$

$$
\left\|\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{q\overline{\nu}\cdot\overline{r}} \sup_{x-y\in\gamma Q_{\overline{\nu},0}} |k_{\overline{\nu}}(f)(y)|^q\right)^{1/q}\middle| L_p(\mathbb{R}^d)\right\| \sim \|\, f \,|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\|
$$

and

$$
\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{q\overline{\nu}\cdot\overline{r}} \left\|\sup_{x-y\in\gamma Q_{\overline{\nu},0}} \left|k_{\overline{\nu}}(f)(y)\right| \middle| L_p(\mathbb{R}^d)\right\|^q\right)^{1/q} \sim \|\, f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\|\,,
$$

respectively, for every $f \in \left(X^{\overline{S}^2}(\mathbb{R}^{\overline{d}})\right)'$.

Remark 3.3.2. If we keep in mind Remark 3.3.1, then this proposition holds true accordingly also for kernels which are not generated as in Theorem 3.3.2, but still satisfy all the assumptions as in Theorem 3.3.1. In the sequel, we will not distinguish between those two variants.

Theorem 3.3.3. Let $\overline{r} \in \mathbb{R}^N$ and $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces). Moreover, let $\overline{M}, \overline{R} \in \mathbb{N}_0^N$ be as in Theorem 3.3.2. Then it holds:

(i) There exist functions $k_0^1, \ldots, k_0^N, k^1, \ldots, k^N$, such that $k_0^i, k^i \in \mathcal{S}(\mathbb{R}^{d_i})$ and

$$\operatorname{supp} k^i, \operatorname{supp} k_0^i \subset \left\{t \in \mathbb{R}^{d_i} \,:\, |t| < 1\right\}; \tag{3.3.12}$$

$$(\mathcal{F}_{d_i} k_0^i)(0) = c_i \neq 0\,; \tag{3.3.13}$$

$$c_i = (\mathcal{F}_{d_i} k_0^i)(\xi) + \sum_{\nu_i=1}^{\infty} (\mathcal{F}_{d_i} k^i)(2^{-\nu_i}\xi)\,; \tag{3.3.14}$$

$$(\mathcal{F}_{d_i} k^i)(\xi) = (\mathcal{F}_{d_i} k_0^i)(\xi) - (\mathcal{F}_{d_i} k_0^i)(2\xi)\,; \tag{3.3.15}$$

$$D^{\alpha}(\mathcal{F}_{d_i} k^i)(0) = 0\,, \quad 0 \leq \overline{\alpha} \leq \overline{R}\,, \tag{3.3.16}$$

each time for all $\xi \in \mathbb{R}^{d_i}$ and $i = 1, \ldots, N$, where $c_i \neq 0$ may be given arbitrarily.

(ii) With the help of the functions $k_0^1, \ldots, k_0^N$, $k^1, \ldots, k^N$ we define functions $k_{\overline{\nu}}$ for $\overline{\nu} \in \mathbb{N}_0^N$ as in Theorem 3.3.2. Then it holds for every $\gamma > 0$

$$\left\| \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{q\overline{\nu}\cdot\overline{r}} \sup_{x-y \in \gamma Q_{\overline{\nu},0}} \left| k_{\overline{\nu}}(f)(y) \right|^q \right)^{1/q} \middle| L_p(\mathbb{R}^d) \right\| \sim \| f \,|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\|$$

and, respectively,

$$\left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{q\overline{\nu}\cdot\overline{r}} \left\| \sup_{x-y \in \gamma Q_{\overline{\nu},0}} \left| k_{\overline{\nu}}(f)(y) \right| \middle| L_p(\mathbb{R}^d) \right\|^q \right)^{1/q} \sim \| f \,|S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})\|\,.$$

Proof. **Step 1: Construction of the kernels.**
Obviously it suffices to present the construction for one pair $k_0^i, k^i \in \mathcal{S}(\mathbb{R}^{d_i})$, hence we will drop the index i and consider the situation for $\mathbb{R}^n$.
We start with a function $\kappa_0 \in \mathcal{S}(\mathbb{R}^n)$ with $\operatorname{supp} \kappa_0 \subset \{t \in \mathbb{R}^n \,:\, |t| \leq 1/2\}$ and $\mathcal{F}_n\kappa_0(0) = c \neq 0$. In particular, it follows that $\mathcal{F}_n\kappa_0$ is an entire analytic function. Let

$$(\mathcal{F}_n\kappa_0)(\xi) = c + \sum_{|\alpha| \geq 1} c_\alpha \xi^\alpha\,, \quad \xi \in \mathbb{R}^n\,,$$

be its Taylor expansion. The construction will now be given iteratively. In each step we will obtain functions, which are supported in $\{t \in \mathbb{R}^n \,:\, |t| \leq 1/2\}$. Hence, the sequence of the Fourier transformed functions will be a sequence of entire analytic functions. Now let κ_m be the function we obtained after m steps, and let

$$(\mathcal{F}_n\kappa_m)(\xi) = c + \sum_{|\alpha| \geq m+1} c_\alpha^m \xi^\alpha\,, \quad \xi \in \mathbb{R}^n\,,$$

be the Taylor expansion of the corresponding Fourier transformed function. Now there exist real numbers λ_1 and λ_2, such that

$$\lambda_1 + \lambda_2 = 1\,, \qquad \lambda_1 + 2^{-m-1}\lambda_2 = 0\,.$$

One easily verifies, that the coefficient matrix is invertible for every $m \in \mathbb{N}_0$. Then we put

$$\kappa_{m+1}(x) = \lambda_1 \kappa_m(x) + \lambda_2 2^n \kappa_m(2x)\,,$$

thus we have $\operatorname{supp}\kappa_{m+1} \subset \operatorname{supp}\kappa_m \subset \operatorname{supp}\kappa_0 \subset \{t \in \mathbb{R}^n : |t| \leq 1/2\}$. Consequently, $\mathcal{F}_n\kappa_{m+1}$ is an entire analytic function as well. Indeed in its Taylor expansion all terms ξ^α with $1 \leq |\alpha| \leq m+1$ are vanishing. This follows directly from the choice of λ_1 and λ_2. We find

$$\begin{aligned}
(\mathcal{F}_n\kappa_{m+1})(\xi) &= \lambda_1(\mathcal{F}_n\kappa_m)(\xi) + \lambda_2\big(2^n\mathcal{F}_n\kappa_m(2\,\cdot)\big)(\xi) = \lambda_1(\mathcal{F}_n\kappa_m)(\xi) + \lambda_2\big(\mathcal{F}_n\kappa_m\big)(\xi/2)\\
&= \lambda_1\left(c + \sum_{|\alpha|\geq m+1} c_\alpha^m \xi^\alpha\right) + \lambda_2\left(c + \sum_{|\alpha|\geq m+1} c_\alpha^m (\xi/2)^\alpha\right)\\
&= \underbrace{(\lambda_1+\lambda_2)}_{=1}c + \sum_{|\alpha|=m+1} c_\alpha^m \underbrace{(\lambda_1 + 2^{-|\alpha|}\lambda_2)}_{=\lambda_1+2^{-m-1}\lambda_2=0}\xi^\alpha + \sum_{|\alpha|\geq m+2} \underbrace{c_\alpha^m(\lambda_1+2^{|\alpha|}\lambda_2)}_{=:c_\alpha^{m+1}}\xi^\alpha\\
&= c + \sum_{|\alpha|\geq m+2} c_\alpha^{m+1}\xi^\alpha\,.
\end{aligned}$$

By this procedure we obtain a sequence of functions $\big(\kappa_m\big)_{m=0}^\infty$ in such a way that $c_\alpha^m = 0$ holds for all $1 \leq |\alpha| \leq m$, and hence

$$\kappa^m(x) = \kappa_m(x) - 2^{-n}\kappa_m(x/2)\,, \quad x \in \mathbb{R}^n\,,$$

has the property

$$(\mathcal{F}_n\kappa^m)(\xi) = (\mathcal{F}_n\kappa_m)(\xi) - (\mathcal{F}_n\kappa_m)(2\xi) = o(|\xi|^m)\,, \quad \xi \in \mathbb{R}^n\,. \tag{3.3.17}$$

Hence $k_0 = \kappa_{2M}$ and $k = \kappa^{2M}$ satisfy the desired properties. The equation (3.3.17) corresponds to (3.3.15), and the vanishing of the derivatives follows from this definition as well, more precisely from the absence of according terms in the Taylor expansion.
Property (3.3.15) yields, that the series in (3.3.14) is a telescoping sum. Since from (3.3.15) or (3.3.17), respectively, we conclude $\mathcal{F}_nk(0) = 0$, the pointwise convergence of the series follows from the continuity of $\mathcal{F}_nk$, and hence we obtain (3.3.14).

Step 2: Regarding the norm equivalences.
Similar to the proof of Theorem 3.3.2 these equivalences will be traced back to Theorem 3.3.1 by putting $\psi_0^i = \mathcal{F}_{d_i}^{-1}k_0^i$ and $\psi_1^i = \mathcal{F}_{d_i}^{-1}k^i$. Due to their construction in Step 1 these functions satisfy the conditions (3.3.1) and (3.3.2). Though by (3.3.15) we have $(\mathcal{F}_{d_i}k^i)(0) = 0$, since this function is an analytic one different from the nullfunction there is a neighbourhood of the origin containing no further zero. Hence we can always fulfil condition (3.3.3) as well. □

One last modification of Theorem 3.3.2 is rather technical. It refers to "directional" local means. By this we mean means of the form ($N = 2, d_1 = d_2 = 1$)

$$\int_{\mathbb{R}} k^1_{\nu_1}(y_1) f(x_1 + y_1, x_2) dy_1 \,.$$

To introduce the local means in general dimensions, we define for every index set $I = \{i_1, \ldots, i_L\} \subset \{1, \ldots, N\}$, $1 \le i_1 < \cdots < i_L \le N$, $L = |I|$, mappings

$$\sigma_I : \mathbb{R}^{d_{i_1}} \times \cdots \times \mathbb{R}^{d_{i_L}} \longrightarrow \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N} \,, \qquad \sigma_I(x^{i_1}, \ldots, x^{i_L}) = (y^1, \ldots, y^N) \,.$$

Thereby we put $y^j = 0$, if $j \notin I$, and $y^j = x^{i_l}$, if $j = i_l \in I$. Then we define the directional local means by

$$k_{\overline{\nu},I}(f)(x) = \int_{\mathbb{R}^{d_{i_1}+\cdots+d_{i_L}}} \left(\prod_{j\in I} k^{i_j}_{\nu_{i_j}}(y^{i_j})\right) f\big(x + \sigma_I(y^{i_1}, \ldots, y^{i_L})\big) \left(\prod_{j\in I} dy^{i_j}\right) . \qquad (3.3.18)$$

This means, we restrict the integration in (3.3.7) to the variables y^i, for which we have $i \in I$, in all the other directions the function remains unchanged.

With the help of this notation we can present the announced modification of Theorem 3.3.2.

Lemma 3.3.1. Let $0 < p, q \le \infty$ ($p < \infty$ for F-spaces), $I \subset \{i, \ldots, N\}$, and $\gamma > 0$. Furthermore, let $\overline{r} \in \mathbb{R}^d$, such that for $i \notin I$ we have $r_i > d_i/p$ in the B-case and $r_i > d_i/\min(p, q)$ in the F-case. Moreover, let $M_i, R_i \in \mathbb{N}_0$ and $k^i_{\nu_i} \in \mathcal{S}(\mathbb{R}^{d_i})$ be as in Theorem 3.3.2 or Theorem 3.3.3 for every $i \in I$. Finally, let $k_{\overline{\nu},I}(f)$ be defined as in (3.3.18). Then it holds

$$\left\| \left(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i=0, i\notin I}} 2^{q\overline{\nu}\cdot\overline{r}} \sup_{x-y\in\gamma Q_{\overline{\nu},0}} \left|k_{\overline{\nu},I}(f)(y)\right|^q \right)^{1/q} \middle| L_p(\mathbb{R}^d) \right\| \le c \, \| f \,| S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}}) \|$$

for every $f \in S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})$, and

$$\left(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i=0, i\notin I}} \left\| 2^{q\overline{\nu}\cdot\overline{r}} \sup_{x-y\in\gamma Q_{\overline{\nu},0}} \left|k_{\overline{\nu},I}(f)(y)\right| \middle| L_p(\mathbb{R}^d) \right\|^q \right)^{1/q} \le c \, \| f \,| S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \|$$

for every $f \in S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$.

Remark 3.3.3. The proof follows along the lines of the one of Theorem 3.3.1 (or Proposition 3.3.1, respectively). Moreover, one uses

$$f = \sum_{\substack{\overline{k}\in\mathbb{N}_0^N, \\ k_i=0, i\notin I}} \mathcal{F}^{-1}\big(\varphi_{\overline{k}} \mathcal{F} f\big) = (2\pi)^{d/2} \sum_{\substack{\overline{k}\in\mathbb{N}_0^N, \\ k_i=0, i\notin I}} \mathcal{F}^{-1}\varphi_{\overline{k}} * f$$

as an equation in $\mathcal{S}'(\mathbb{R}^d)$, and for the corresponding maximal functions $\big(\mathcal{F}^{-1}\varphi^*_{\overline{k}} f\big)_{\overline{a}}$ additionally $r_i > a_i > d_i/p$ for $i \notin I$ is required. We omit details.

4 Decomposition theorems

This chapter is devoted to characterizations of the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ via certain types of decompositions. After discussing the necessary adaptions of the sequence spaces $a^s_{p,q}$ and $s^r_{p,q}a$, we present a result for atomic decompositions. The main result of this chapter can be found in Section 4.3, where we prove a characterization in terms of wavelets, which will be the generalization of the corresponding results for the isotropic spaces $A^s_{p,q}(\mathbb{R}^n)$ (Theorem 1.2.2) and the spaces of dominating mixed smoothness $S^r_{p,q}A(\mathbb{R}^d)$ (Theorem 1.4.1).

4.1 Sequence spaces

Definition 4.1.1. Let $0 < p, q \leq \infty$ and $\overline{r} \in \mathbb{R}^N$. For sequences

$$\lambda = \left\{\lambda_{\overline{\nu},m} \in \mathbb{C} \,:\, \overline{\nu} \in \mathbb{N}_0^N, m \in \mathbb{Z}^d\right\} \tag{4.1.1}$$

we define

$$s^{\overline{r}}_{p,q}b^* := \left\{\lambda : \|\,\lambda\,|s^{\overline{r}}_{p,q}b^*\| < \infty\right\},$$

$$\|\,\lambda\,|s^{\overline{r}}_{p,q}b^*\| := \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)q}\left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}|^p\right)^{q/p}\right)^{1/q},$$

and for $p < \infty$ we put

$$s^{\overline{r}}_{p,q}f^* := \left\{\lambda : \|\,\lambda\,|s^{\overline{r}}_{p,q}f^*\| < \infty\right\},$$

$$\|\,\lambda\,|s^{\overline{r}}_{p,q}f^*\| := \left\|\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d} |2^{\overline{\nu}\cdot\overline{r}}\lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}(\cdot)|^q\right)^{1/q}\Big|L_p(\mathbb{R}^d)\right\|,$$

where $\mathcal{X}_{\overline{\nu},m}$ denotes the characteristic function of the rectangle (generalized cube) $Q_{\overline{\nu},m}$. Moreover, in case in case p and/or q are infinite one has to use the usual modifications.

Remark 4.1.1. As before we shall use the notation $s^{\overline{r}}_{p,q}a^*$ for $a \in \{b, f\}$ to refer to both scales of sequence spaces
The notation $s^{\overline{r}}_{p,q}a$ is reserved for a slight modification of the spaces, which will be needed in connection with wavelet decompositions, see Definition 4.3.1.

Remark 4.1.2. For a given sequence λ as in (4.1.1), we put $g_{\overline{\nu},m} = \lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}(x)$ and $g_{\overline{\nu}} = \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}(x)$. Then the above quasi-norms can be rewritten as follows:

$$\|\,\lambda\,|s^{\overline{r}}_{p,q}b^*\| = \|\,2^{\overline{\nu}\cdot\overline{r}}g_{\overline{\nu}}\,|\ell_q(L_p)\|,$$

$$\|\,\lambda\,|s^{\overline{r}}_{p,q}f^*\| = \|\,2^{\overline{\nu}\cdot\overline{r}}g_{\overline{\nu},m}\,|L_p\big(\ell_q(\mathbb{N}_0^N \times \mathbb{Z}^d)\big)\| = \|\,2^{\overline{\nu}\cdot\overline{r}}g_{\overline{\nu}}\,|L_p\big(\ell_q(\mathbb{Z}^d)\big)\|.$$

In particular, we have

$$\|\,g_{\overline{\nu}}\,|L_p(\mathbb{R}^d)\| = \|(\lambda_{\overline{\nu},m})_{m\in\mathbb{Z}^d}|\ell_p(\mathbb{Z}^d)\| \quad \text{for all} \quad \overline{\nu} \in \mathbb{N}_0^N. \tag{4.1.2}$$

Hence in view of Remark 2.2.1 these sequence spaces can be regarded as discrete counterparts of the spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N})$. This impression is amplified by the decomposition theorems in the next sections.

4.2 Atomic decomposition

We remind on the notation $\overline{\alpha} := (|\alpha^1|, \ldots, |\alpha^N|) \in \mathbb{N}_0^N$ for some multiindex $\alpha \in \mathbb{N}_0^d$, $\alpha = (\alpha^1, \ldots, \alpha^N)$. With it we can specify the definition of the considered atoms.

Definition 4.2.1. Let $\overline{K} \in \mathbb{N}_0^N$, $\overline{L} + \overline{1} \in \mathbb{N}_0^N$, and let $\gamma > 1$. A $\overline{K}$-times continuously differentiable complex-valued function $a \in S^{\overline{K}}C(\mathbb{R}^{\overline{d}})$ is called $[\overline{K}, \overline{L}]$-atom centred in $Q_{\overline{\nu},m}$, if

$$\operatorname{supp} a \subset \gamma Q_{\overline{\nu},m}\,, \tag{4.2.1}$$

$$|D^\alpha a(x)| \le 2^{\overline{\alpha}\cdot\overline{\nu}} \text{ for all } \alpha \in \mathbb{N}_0^d \text{ with } 0 \le \overline{\alpha} \le \overline{K}, x \in \mathbb{R}^d\,, \tag{4.2.2}$$

and for $i = 1, \ldots, N,\ \beta \in \mathbb{N}_0^{d_i}$ with $0 \le |\beta| \le L_i$ and $\nu_i \ge 1$

$$\int_{\mathbb{R}^{d_i}} (x^i)^\beta a(x^1, \ldots, x^i, \ldots, x^N)\, dx^i = 0 \qquad \text{for all } (x^1, \ldots, x^N) \in \mathbb{R}^d\,. \tag{4.2.3}$$

Here $L_i = -1$ means that no moment conditions are necessary.

Before we deal with atomic decompositions, we need an auxiliary result on the maximal operator $\overline{M}$.

Lemma 4.2.1. Let $\overline{\nu} \in \mathbb{N}_0^N$ and $m \in \mathbb{Z}^d$.

(i) It holds for all $x \in \mathbb{R}^d$

$$\big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)(x) \sim \frac{2^{-\overline{\nu}\cdot\overline{d}}}{\prod_{i=1}^N (|x^i - 2^{-\nu_i} m^i|_\infty + 2^{-\nu_i})^{d_i}}\,,$$

where the equivalence constants depend on $d_1, \ldots, d_N$ only.

(ii) Consider $\overline{\gamma}\, Q_{\overline{\nu},m}$ for real numbers $\gamma_i > 1$, $i = 1, \ldots, N$. Then it holds

$$\big(\overline{M}\mathcal{X}_{\overline{\gamma}\, Q_{\overline{\nu},m}}\big)(x) \le c\, \gamma_1^{d_1} \cdots \gamma_N^{d_N} \big(\overline{M}\mathcal{X}_{Q_{\overline{\nu},m}}\big)(x)$$

with some constant $c > 0$, depending only on $d_1, \ldots, d_N$.

Proof. Part (i) immediately follows from the known result for the characteristic function of the interval $[-1, 1]$,

$$\big(M\mathcal{X}_{[-1,1]}\big)(x) \sim \frac{1}{1 + |x|}\,,$$

from which we can derive for some cube $Q_R \subset \mathbb{R}^n$ with sidelength $2R$ and centred at the origin

$$\big(M\mathcal{X}_{Q_R}\big)(x) \sim \frac{R^n}{(|x|_\infty + R)^n}\,. \tag{4.2.4}$$

The assertion now follows at once using the product structure of the generalized cube $Q_{\overline{\nu},m}$, see (3.3.11).

For part (ii) it is sufficient to consider the isotropic case, i.e. $N = 1$ and M instead of $\overline{M}$, again due to the product structure of the generalized cube $Q_{\overline{\nu},m}$. We obtain from (4.2.4) for $\gamma > 1$

$$\begin{aligned}\big(M\mathcal{X}_{\gamma Q_{\nu,m}}\big)(x) &\leq c\,\frac{(\gamma 2^{-\nu-1})^n}{(|x-2^{-\nu}m|_\infty + \gamma 2^{-\nu-1})^n} \\ &\leq c\,\gamma^n \frac{2^{-(\nu+1)n}}{(|x-2^{-\nu}m|_\infty + 2^{-\nu-1})^n} \leq c'\gamma^n \big(M\mathcal{X}_{Q_{\nu,m}}\big)(x)\,.\end{aligned}$$

Since the constants in (4.2.4) only depend on n, this holds for c' as well. □

Lemma 4.2.2. Let $0 < p, q \leq \infty$ and $\overline{r} \in \mathbb{R}^N$. Moreover, let $\overline{K} \in \mathbb{N}_0^N$ and $\overline{L} \in \mathbb{Z}^N$ be fixed, where

$$\overline{L} + \overline{1} \in \mathbb{N}_0^N\,, \qquad L_i \geq \max(-1, [\sigma_p^i - r_i]), \qquad i = 1, \ldots, N\,. \tag{4.2.5}$$

(i) It holds

$$s_{p,q}^{\overline{r}}a^* \hookrightarrow s_{p,\infty}^{\overline{r}}b^*\,, \qquad \big\|\,\lambda\,\big|s_{p,\infty}^{\overline{r}}b^*\big\| \leq \big\|\,\lambda\,\big|s_{p,q}^{\overline{r}}a^*\big\|\,, \tag{4.2.6}$$

where $p < \infty$ for $a = f$.

(ii) For every sequence $\lambda \in s_{p,\infty}^{\overline{r}}b^*$ and every family $(a_{\overline{\nu},m})_{\overline{\nu}\in\mathbb{N}_0^N, m\in\mathbb{Z}^d}$ of $[\,\overline{K},\overline{L}\,]$-atoms centred in $Q_{\overline{\nu},m}$ the series

$$\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\lambda_{\overline{\nu},m}a_{\overline{\nu},m} \tag{4.2.7}$$

converges unconditionally in $\mathcal{S}'(\mathbb{R}^d)$.

Proof. **Step 1: We prove** (4.2.6)**.**
For b-spaces the assertion follows immediately from the monotonicity of the ℓ_q-spaces and the corresponding quasi-norm estimates. In the f-case we use the functions $g_{\overline{\nu},m}$ from Remark 4.1.2 and equation (4.1.2) and obtain

$$\big\|\,\lambda\,\big|s_{p,q}^{\overline{r}}f\,\big\| \geq \big\|\,2^{\overline{\nu}\cdot\overline{r}}g_{\overline{\nu}}\,\big|L_p(\mathbb{R}^d)\big\| = 2^{\overline{\nu}\cdot\overline{r}}\big\|(\lambda_{\overline{\nu},m})_{m\in\mathbb{Z}^d}\big|\,\ell_p\,\big\|$$

for all $\overline{\nu} \in \mathbb{N}_0^N$. Taking the supremum over $\overline{\nu}$ gives (4.2.6).

Step 2: We prove the convergence of (4.2.7) **in** $\mathcal{S}'(\mathbb{R}^d)$**.**
To this purpose, let $\varphi \in \mathcal{S}(\mathbb{R}^d)$. We use the Taylor expansion of φ with respect to the first variables, i.e. with respect to x^1:

$$\begin{aligned}\varphi(y) = &\sum_{\alpha^1\in\mathbb{N}_0^{d_1}:\,|\alpha^1|\leq L_1}\frac{D^{(\alpha^1,0,\ldots,0)}\varphi(2^{-\nu_1}m^1,y^2,\ldots,y^N)}{\alpha^1!}\big(y^1-2^{-\nu_1}m^1\big)^{\alpha^1} \\ &+\sum_{|\alpha^1|=L_1+1}\frac{L_1+1}{\alpha^1!}\big(y^1-2^{-\nu_1}m^1\big)^{\alpha^1}\end{aligned} \tag{4.2.8}$$

$$\times \int_0^1 (D^{\alpha^1}\varphi)((1-t_1)2^{-\nu_1}m^1 + t_1 y^1, y^2, \dots, y^N)(1-t_1)^{L_1} dt_1 \,.$$

Moreover, we employ (4.2.3) to get

$$\begin{aligned} &\int_{\mathbb{R}^d} a_{\overline{\nu},m}(y)\varphi(y)dy \qquad (4.2.9)\\ &\quad = \int_{\mathbb{R}^d} a_{\overline{\nu},m}(y) \sum_{\substack{\alpha=(\alpha^1,0,\dots,0)\in\mathbb{N}_0^d,\\ |\alpha^1|=L_1+1}} \frac{L_1+1}{\alpha^1!}\big(y^1-2^{-\nu_1}m^1\big)^{\alpha^1} \\ &\qquad\qquad \times \int_0^1 (D^{\alpha}\varphi)((1-t_1)2^{-\nu_1}m^1 + t_1 y^1, y^2, \dots, y^N)(1-t_1)^{L_1} dt_1\, dy \,. \end{aligned}$$

By iterated applications of Taylor expansions as in (4.2.8) and the moment condition (4.2.3) to the other variables we obtain from the right hand side of (4.2.9)

$$\begin{aligned} &\int_{\mathbb{R}^d} a_{\overline{\nu},m}(y) \sum_{\substack{\alpha=(\alpha^1,\dots,\alpha^N)\in\mathbb{N}_0^d,\\ |\alpha^i|=L_i+1}} \frac{1}{\alpha!}\prod_{i=1}^N (L_i+1)(y^i-2^{-\nu_i}m^i)^{\alpha^i} \\ &\qquad\qquad \times \int_{[0,1]^N} \big(D^\alpha\varphi\big)\big((\overline{1}-\overline{t})2^{-\overline{\nu}}m + \overline{t}y\big)\, d\overline{t}\ dy \,. \qquad (4.2.10) \end{aligned}$$

Here we used the abbreviations $2^{-\overline{\nu}}m = (2^{-\nu_1}m^1, \dots, 2^{-\nu_N}m^N)$ and $\overline{t}y = (t_1y^1, \dots, t_Ny^N)$. Using the support property (4.2.1) of the atoms $a_{\overline{\nu},m}$, we can further estimate the absolute value of (4.2.10). We obtain from $y \in \gamma Q_{\overline{\nu},m}$

$$\big|y^i - 2^{-\nu_i}m^i\big| \le \gamma 2^{-\nu_i} \qquad \text{and} \qquad \big|(y^i-2^{-\nu_i}m^i)^{\alpha^i}\big| \le (\gamma 2^{-\nu_i})^{L_i+1} \,,$$

hence it follows for the integrand in (4.2.10)

$$\begin{aligned} &\left|\prod_{i=1}^N \big(y^i - 2^{-\nu_i}m^i\big)^{\alpha^i} \int_{[0,1]^N} \big(D^\alpha\varphi\big)\big((\overline{1}-\overline{t})2^{-\overline{\nu}}m + \overline{t}y\big)\, d\overline{t}\right| \\ &\qquad \le \gamma^{|\overline{L}|+N} 2^{-\overline{\nu}\cdot(\overline{L}+\overline{1})} \sup_{x\in\gamma Q_{\overline{\nu},m}} \langle x\rangle^M \big|(D^\alpha\varphi)(x)\big| \int_{[0,1]^N} \big\langle(\overline{1}-\overline{t})2^{-\overline{\nu}}m + \overline{t}y\big\rangle^{-M} d\overline{t} \\ &\qquad \le c_M 2^{-\overline{\nu}\cdot(\overline{L}+\overline{1})} \langle y\rangle^{-M} \sup_{x\in\gamma Q_{\overline{\nu},m}} \langle x\rangle^M \big|(D^\alpha\varphi)(x)\big| \,. \end{aligned}$$

Here we used the shortened notation $\langle x\rangle = (1+|x|^2)^{1/2}$ for $x \in \mathbb{R}^d$, the parameter M is still at our disposition and will be chosen later. The last estimate is a consequence of the observation $\langle 2^{-\overline{\nu}}m\rangle \sim \langle\xi\rangle$ for all $\xi \in \gamma Q_{\overline{\nu},m}$, where the constants are independent of $\overline{\nu}$ and m. This in turn follows from

$$\begin{aligned} \langle\xi\rangle^2 &\lesssim 1 + |2^{-\overline{\nu}}m|_2^2 + |\xi - 2^{-\overline{\nu}}m|_2^2 \ \le\ 1 + |2^{-\overline{\nu}}m|_2^2 + \sum_{i=1}^N d_i |2^{-\nu_i}m^i - \xi^i|_\infty^2 \\ &\le 1 + |2^{-\overline{\nu}}m|_2^2 + \sum_{i=1}^N d_i \gamma^2 2^{-2\nu_i-2} \ \lesssim\ 1 + |2^{-\overline{\nu}}m|_2^2 \ =\ \langle 2^{-\overline{\nu}}m\rangle^2 \end{aligned}$$

$$\lesssim 1 + |\xi|_2^2 + |2^{-\overline{\nu}} m - \xi|_2^2 \;\lesssim\; 1 + |\xi|_2^2 \;=\; \langle \xi \rangle^2 .$$

Now we suppose at first $p > 1$, and use $|a_{\overline{\nu},m}(y)| \le \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(y)$, which is a consequence of (4.2.1) and (4.2.2) for $\alpha = 0$. We find

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}(y) \varphi(y) dy \right| \\
&\quad \le c\, 2^{-\overline{\nu}\cdot(\overline{L}+\overline{1})} \int_{\mathbb{R}^d} \sum_{m\in\mathbb{Z}^d} \sum_{\alpha:\,\overline{\alpha}=\overline{L}+\overline{1}} \left(\sup_{x\in\gamma Q_{\overline{\nu},m}} \langle x\rangle^M |(D^\alpha \varphi)(x)| \right) |\lambda_{\overline{\nu},m}| \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(y) \langle y\rangle^{-M} dy \\
&\quad \le C\, 2^{-\overline{\nu}\cdot(\overline{r}+\overline{L}+\overline{1})} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}|^p \right)^{1/p} \left(\sup_{x\in\mathbb{R}^d} \langle x\rangle^M \sum_{\alpha:\,\overline{\alpha}=\overline{L}+\overline{1}} \big|(D^\alpha\varphi)(x)\big| \right) \\
&\quad \le C\, 2^{-\overline{\nu}\cdot(\overline{r}+\overline{L}+\overline{1})} \big\| \lambda \,|s^{\overline{r}}_{p,\infty} b^* \big\| \cdot \big\| \varphi \big\|_{M,|\overline{L}|+N} .
\end{aligned} \tag{4.2.11}$$

The intermittently appearing integral can be estimated by using Hölder's inequality for integrals with respect to $1 = \frac{1}{p} + \frac{1}{p'}$ and choosing $Mp' > d$,

$$\int_{\mathbb{R}^d} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}| \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(y) \right) \langle y\rangle^{-M} dy \le c_{\gamma,M}\, 2^{-\overline{\nu}\cdot\overline{d}/p} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}|^p \right)^{1/p} ,$$

where additionally the fact is of importance, that every $x \in \mathbb{R}^d$ is contained in only finitely many of the sets $\gamma Q_{\overline{\nu},m}$, the count being bounded depending only on γ. Concerning the full series (4.2.7) we finally arrive at

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}(y) \right) \varphi(y) dy \right| \\
&\qquad \le c \sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot(\overline{r}+\overline{L}+\overline{1})} \big\| \lambda \,|s^{\overline{r}}_{p,\infty} b^* \big\| \cdot \big\| \varphi \big\|_{M,|\overline{L}|+N} \\
&\qquad \le C \, \big\| \lambda \,|s^{\overline{r}}_{p,\infty} b^* \big\| \cdot \big\| \varphi \big\|_{M,|\overline{L}|+N} .
\end{aligned} \tag{4.2.12}$$

Hence the series (4.2.7) converges in $\mathcal{S}'(\mathbb{R}^d)$, if only $\overline{r} + \overline{L} + \overline{1} > 0$, since then the $\overline{\nu}$-summation results in a convergent geometric series. But this restriction is assured by the assumption (4.2.5), as $L_i \ge \max(-1, [\sigma_p^i - r_i])$ implies $L_i + 1 > \sigma_p^i - r_i$, and $\sigma_p^i = 0$ due to $p > 1$.

In case $p \le 1$ we obtain by similar arguments (this corresponds to the choice $M = 0$)

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}(y) \varphi(y) dy \right|^p \\
&\le c\, 2^{-\overline{\nu}\cdot(\overline{L}+\overline{1})p} \sum_{m\in\mathbb{Z}^d} \left(\sum_{\alpha\,:\,\overline{\alpha}=\overline{L}+\overline{1}} \sup_{x\in\gamma Q_{\overline{\nu},m}} |(D^\alpha\varphi)(x)| \right)^p \left(\int_{\mathbb{R}^d} |\lambda_{\overline{\nu},m}| \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(y) dy \right)^p \\
&\le c'\, 2^{-\overline{\nu}\cdot(\overline{r}+\overline{L}+\overline{1}-\overline{d}(1/p-1))p} \big\| \varphi \big\|^p_{0,|\overline{L}|+N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)p} \sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}|^p
\end{aligned}$$

$$\leq C\, 2^{-\overline{\nu}\cdot(\overline{r}+\overline{L}+1-\overline{d}(1/p-1))p} \big\| \lambda \,|s^{\overline{r}}_{p,\infty}b^*\big\|^p \cdot \big\| \varphi \big\|^p_{0,|\overline{L}|+N} \,.$$

Adding finally the $\overline{\nu}$-summation, this results once more in a converging geometric series due to the assumption on L_i.

Step 3: Unconditional convergence.
Unconditional convergence of (4.2.7) in $\mathcal{S}'(\mathbb{R}^d)$ is equivalent to the unconditional convergence of every series

$$\sum_{\overline{\nu}\in\mathbb{N}_0^N} \sum_{m\in\mathbb{Z}^d} \big(\lambda_{\overline{\nu},m} a_{\overline{\nu},m}\big)(\varphi)\,, \qquad \varphi \in \mathcal{S}(\mathbb{R}^d)\,. \tag{4.2.13}$$

This follows immediately from the choice of the strong topology on $\mathcal{S}'(\mathbb{R}^d)$. Hence we only have to consider unconditional convergence of series of complex numbers.
For series of complex numbers it is well-known that absolute convergence implies unconditional convergence. In our case, a closer inspection of Step 2 shows, that we have in fact proven the existence of

$$\sum_{\overline{\nu}\in\mathbb{N}_0^N} \sum_{m\in\mathbb{Z}^d} \Big|\big(\lambda_{\overline{\nu},m} a_{\overline{\nu},m}\big)(\varphi)\Big| = \lim_{n\to\infty} \sum_{|\overline{\nu}|\leq n} \sum_{m\in\mathbb{Z}^d} \Big|\big(\lambda_{\overline{\nu},m} a_{\overline{\nu},m}\big)(\varphi)\Big| \tag{4.2.14}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, i.e. convergence as an iterated series, see (4.2.12). Thereby, the convergence of the inner series is obvious from (4.2.11), at least for $p < \infty$. For $p = \infty$, this convergence follows from

$$\int_{\mathbb{R}^d} \sum_{|m|_\infty \geq n} \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(y) \langle y\rangle^{-M} dy \leq c_\gamma \int_{y\,:\,|y^i|_\infty \geq 2^{-\nu_i}(n-\frac{\gamma}{2})} \langle y\rangle^{-M} dy \overset{n\to\infty}{\longrightarrow} 0\,.$$

From (4.2.14) we obtain by Fubini's theorem the absolute convergence of the series (4.2.13), which implies its unconditional convergence. □

Theorem 4.2.1. Let $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces) and $\overline{r} \in \mathbb{R}^N$. Moreover, let $\overline{K} \in \mathbb{N}_0^N$ and $\overline{L} + \overline{1} \in \mathbb{N}_0^N$ be fixed, where for $i = 1, \ldots, N$ we assume

$$\begin{aligned} K_i &\geq (1 + [r_i])_+\,, \\ L_i &\geq \begin{cases} \max(-1, [\sigma^i_p - r_i])\,, & \text{for } B\text{-spaces}, \\ \max(-1, [\sigma^i_{pq} - r_i])\,, & \text{for } F\text{-spaces, } 0 < q < \infty\,, \\ \max(-1, d_i + [\sigma^i_p - r_i])\,, & \text{for } F\text{-spaces, } q = \infty\,. \end{cases} \end{aligned} \tag{4.2.15}$$

(i) For every sequence $\lambda \in s^{\overline{r}}_{p,q}a^*$ and every family $(a_{\overline{\nu},m})_{\overline{\nu}\in\mathbb{N}_0^N, m\in\mathbb{Z}^d}$ of $[\,\overline{K}, \overline{L}\,]$-atoms centred in $Q_{\overline{\nu},m}$ the limit f of the series (4.2.7) belongs to the space $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, and it holds

$$\big\| f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \leq c \,\big\| \lambda \,|s^{\overline{r}}_{p,q}a^*\big\|\,, \tag{4.2.16}$$

where $c > 0$ is some universal constant, independent of all admissible λ and $a_{\overline{\nu},m}$.

(ii) For every tempered distribution $f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ exists a sequence $\lambda \in s^{\overline{r}}_{p,q}a^*$ and a family of $[\,\overline{K}, \overline{L}\,]$-atoms centred in $Q_{\overline{\nu},m}$ (with some sufficiently large $\gamma > 1$), denoted again by $(a_{\overline{\nu},m})_{\overline{\nu}\in\mathbb{N}_0^N, m\in\mathbb{Z}^d}$, such that the series (4.2.7) converges in $\mathcal{S}'(\mathbb{R}^d)$ to f. Moreover, the sequence λ satisfies

$$\big\|\, \lambda \,|s^{\overline{r}}_{p,q}a^*\big\| \le c \,\big\|\, f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|\,, \tag{4.2.17}$$

where the constant $c > 0$ is independent of $f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$.

Proof. For most parts of this proof we shall only be concerned with the F-case, the one for B-spaces is similar. The convergence of the series (4.2.7) for $\lambda \in s^{\overline{r}}_{p,q}f^*$ follows from Lemma 4.2.2(ii) combined with the embedding (4.2.6) in Lemma 4.2.2(i).

Step 1: A pointwise estimate for local means.
For the proof of the estimate (4.2.16) we intent to use the equivalent quasi-norms on $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, defined by (3.3.8) and (3.3.9). We choose $\overline{R}, \overline{S} \in \mathbb{N}_0^N$ with $2\overline{S} > \overline{R} > \overline{K}$, and define functions $k_{\overline{l}} \in \mathcal{S}(\mathbb{R}^d)$ for $\overline{l} \in \mathbb{N}_0^N$ as in Theorem 3.3.2. Then we obtain for all $\overline{l}, \overline{\nu} \in \mathbb{N}_0^N$ and all $m \in \mathbb{Z}^d$

$$2^{\overline{l}\cdot\overline{r}} k_{\overline{l}}(a_{\overline{\nu},m})(x) = 2^{\overline{l}\cdot\overline{r}} \int_{\mathbb{R}^d} k^1_{l_1}(y^1)\cdots k^N_{l_N}(y^N) a_{\overline{\nu},m}(x+y)dy\,. \tag{4.2.18}$$

The further calculations depend on the size of the supports of $k_{\overline{l}}$ and $a_{\overline{\nu},m}$. Thus we have to distinguish between $l_i \ge \nu_i$ and $l_i < \nu_i$, hence a total of 2^N different cases. In the sequel we will discuss first the case $\overline{l} \ge \overline{\nu}$ and afterwards the case $\overline{l} < \overline{\nu}$ in detail, and finally sketch the "mixed" cases. For every $\overline{l} \in \mathbb{N}_0^N$ we put abbreviative $A_{\overline{l}} := A^1_{l_1} \times A^2_{l_2} \times \cdots \times A^N_{l_N}$ and $A^i_j := \{y \in \mathbb{R}^{d_i} : |y| \le 2^{-j}\}$.

Substep 1.1: $\overline{l} \ge \overline{\nu}$ **.**
We assume $\overline{l} > 0$ to simplify the notation, the cases $l_i = \nu_i = 0$ can be treated similarly and will be re-included afterwards. We apply the definition of $k^i_{l_i}$ and integrate partially K_i-times with respect to the variables y^i to obtain

$$\begin{aligned}
2^{\overline{l}\cdot\overline{r}} k_{\overline{l}}(a_{\overline{\nu},m})(x) &= 2^{\overline{l}\cdot(\overline{r}+\overline{d})} \int_{\mathbb{R}^d} \prod_{i=1}^N \Big(\Delta^{S_i} k^i\Big)(2^{l_i}y^i) a_{\overline{\nu},m}(x+y)dy \\
&= 2^{\overline{l}\cdot\overline{r}} \int_{\mathbb{R}^d} \prod_{i=1}^N \Big(\Delta^{S_i} k^i\Big)(y^i) a_{\overline{\nu},m}(x^1 + 2^{-l_1}y^1, \ldots, x^N + 2^{-l_N}y^N)dy \\
&= 2^{\overline{l}\cdot(\overline{r}-\overline{K})} \int_{\mathbb{R}^d} \sum_{\alpha:\,\overline{\alpha}=\overline{S}} \frac{c_\alpha}{c_{\alpha,\overline{K}}} \sum_{\substack{\beta:\,\overline{\beta}=\overline{K},\\ \beta\le 2\alpha}} \Big(D^{2\alpha-\beta}_y \prod_{i=1}^N k^i\Big)(y) \Big(D^\beta_y a_{\overline{\nu},m}\Big)(x + 2^{-\overline{l}}y)dy\,.
\end{aligned}$$

Next, we use that $k^i \in \mathcal{S}(\mathbb{R}^{d_i})$; in particular, we have $\operatorname{supp} D^{2\alpha^i-\beta^i} k^i \subset \operatorname{supp} k^i$ and all partial derivatives are bounded. Moreover, we shall factor in the corresponding properties of the atoms $a_{\overline{\nu},m}$, i.e. properties (4.2.1) and (4.2.2). Then we can further estimate

$$2^{\overline{l}\cdot\overline{r}} \big|k_{\overline{l}}(a_{\overline{\nu},m})(x)\big| \le c_{k^i,\overline{S},\overline{K}}\, 2^{\overline{l}\cdot(\overline{r}-\overline{K})} 2^{\overline{K}\cdot\overline{\nu}} \int_{\mathbb{R}^d} \bigg(\prod_{i=1}^N \mathcal{X}_{\operatorname{supp} k^i}(y^i)\bigg) \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x+2^{-\overline{l}}y)\, dy\,. \tag{4.2.19}$$

By the definition of the rectangle $\gamma Q_{\overline{\nu},m}$ it holds

$$\mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x^1+2^{-l_1}y^1,\dots,x^N+2^{-l_N}y^N)\neq 0$$
$$\iff |x^i-2^{-\nu_i}(m^i-2^{\nu_i-l_i}y^i)|_\infty\le\gamma 2^{-\nu_i-1}\,,\quad i=1,\dots,N\,.$$

Hence x has to belong to a moved rectangle. Since $\operatorname{supp}k^i\subset\{t\in\mathbb{R}^{d_i}:|t|\le 1\}$, $i=1,\dots,N$, we have $|2^{\nu_i-l_i}y^i|_\infty\le 1$. This means the additional shift can be compensated by enlarging the set $\gamma Q_{\overline{\nu},m}$ by a factor 3, thus including all neighbouring rectangles. In this way we find

$$\mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x^1+2^{-l_1}y^1,\dots,x^N+2^{-l_N}y^N)\le\mathcal{X}_{3\gamma Q_{\overline{\nu},m}}(x^1,\dots,x^N) \tag{4.2.20}$$

for all $y^i\in\operatorname{supp}k^i$. Inserting this into (4.2.19) we obtain

$$\begin{aligned}2^{\overline{l}\cdot\overline{r}}\big|k_{\overline{l}}(a_{\overline{\nu},m})(x)\big|&\le 2^{\overline{l}\cdot(\overline{r}-\overline{K})}2^{\overline{K}\cdot\overline{\nu}}\int_{\mathbb{R}^d}\left(\prod_{i=1}^N\mathcal{X}_{\operatorname{supp}k^i}(y^i)\right)\mathcal{X}_{3\gamma Q_{\overline{\nu},m}}(x^1,\dots,x^N)dy\\&\le c\,2^{-(\overline{K}-\overline{r})\cdot(\overline{l}-\overline{\nu})}2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}2^{\overline{\nu}\cdot\overline{d}/p}\mathcal{X}_{3\gamma Q_{\overline{\nu},m}}(x)\,.\end{aligned}\tag{4.2.21}$$

From Lemma 4.2.1(ii) we conclude

$$\mathcal{X}_{3\gamma Q_{\overline{\nu},m}}(x)\le\overline{M}\mathcal{X}_{3\gamma Q_{\overline{\nu},m}}(x)\le c\,\overline{M}\mathcal{X}_{\overline{\nu},m}(x)\,,\qquad x\in\mathbb{R}^d\,. \tag{4.2.22}$$

Passing to L_p-normalized characteristic functions this finally results in

$$2^{\overline{l}\cdot\overline{r}}\big|k_{\overline{l}}(a_{\overline{\nu},m})(x)\big|\le C\,2^{-(\overline{K}-\overline{r})\cdot(\overline{l}-\overline{\nu})}2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}\big(\overline{M}\mathcal{X}^{(p)}_{Q_{\overline{\nu},m}}\big)(x)\,. \tag{4.2.23}$$

Substep 1.2: $\overline{l}<\overline{\nu}$ **.**
The integration in (4.2.18) can be restricted to $A^i_{l_i}$, since the smoothness and support properties of k^i imply $\operatorname{supp}k^i_{l_i}\subset A^i_{l_i}$. We apply the Taylor expansion to $k^i_{l_i}(y^i)$ with respect to the points $2^{-\nu_i}m^i-x^i$ up to the order L_i,

$$\begin{aligned}2^{-l_id_i}k^i_{l_i}(y^i)=&\sum_{0\le|\beta^i|\le L_i}c^i_{\beta^i}(x^i)(y^i-2^{-\nu_i}m^i+x^i)^{\beta^i}\\&+2^{l_i(L_i+1)}O\big(|x^i+y^i-2^{-\nu_i}m^i|^{L_i+1}\big)\,,\end{aligned}\tag{4.2.24}$$

where $c^i_{\beta^i}$ are some coefficient functions, not depending on y^i. Hence property (4.2.3) yields

$$2^{\overline{l}\cdot\overline{r}}k_{\overline{l}}(a_{\overline{\nu},m})(x)=2^{\overline{l}\cdot(\overline{r}+\overline{d})}\int_{A_{\overline{l}}}a_{\overline{\nu},m}(x+y)\prod_{i=1}^N 2^{l_i(L_i+1)}O\Big(\big|x^i+y^i-2^{-\nu_i}m^i\big|^{L_i+1}\Big)dy\,.$$

Moreover, it holds $|a_{\overline{\nu},m}(x+y)|\le\mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x+y)$, due to (4.2.1) and (4.2.2). In particular, we have $|x^i+y^i-2^{-\nu_i}m^i|^{L_i+1}\le(\gamma 2^{-\nu_i-1})^{L_i+1}$. Hence we find

$$2^{\overline{l}\cdot\overline{r}}\big|k_{\overline{l}}(a_{\overline{\nu},m})(x)\big|\le c\,2^{\overline{l}\cdot(\overline{r}+\overline{d})}2^{\overline{l}\cdot(\overline{L}+\overline{1})}2^{-\overline{\nu}\cdot(\overline{L}+\overline{1})}\int_{A_{\overline{l}}}\mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x+y)\,dy\,. \tag{4.2.25}$$

The last integral is always at most $\left|\gamma Q_{\overline{\nu},m}\right| = \gamma^d 2^{-\overline{\nu}\cdot\overline{d}}$, and it even vanishes, if $\{y : x+y \in \gamma Q_{\overline{\nu},m}\} \cap \{y : |y^i| \le 2^{-l_i}\} = \emptyset$. Hence it follows

$$\int_{A_{\overline{l}}} \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x+y)dy \le c\, 2^{-\overline{\nu}\cdot\overline{d}} \mathcal{X}_{\gamma' 2^{\overline{\nu}-\overline{l}} Q_{\overline{\nu},m}}(x)\,. \tag{4.2.26}$$

This is a consequence of

$$\begin{aligned} |x^i - 2^{-\nu_i} m^i|_\infty &\le |x^i + y^i - 2^{-\nu_i} m^i|_\infty + |y^i|_\infty \le \gamma 2^{-\nu_i - 1} + 2^{-l_i} \\ &\le (\gamma+2) 2^{-l_i - 1} = \gamma' 2^{\nu_i - l_i} 2^{-\nu_i - 1}\,, \end{aligned}$$

i.e. $\{z :\ x+z \in \gamma Q_{\overline{\nu},m}\} \cap \{z :\ |z^i| \le 2^{-l_i}\} \ne \emptyset$ implies $x \in \gamma'\, 2^{\overline{\nu}-\overline{l}} Q_{\overline{\nu},m}$. The characteristic function on the right hand side of (4.2.26) can be estimated further by using the maximal operator $\overline{M}$ and Lemma 4.2.1(ii). We obtain

$$2^{-(\overline{\nu}-\overline{l})\cdot\overline{d}} \mathcal{X}_{\gamma'\, 2^{\overline{\nu}-\overline{l}} Q_{\overline{\nu},m}}(x) \le c\, \gamma'^d \left(\overline{M} \mathcal{X}_{\overline{\nu},m}\right)(x)\,. \tag{4.2.27}$$

Now let $0 < \omega < \min(1,p,q)$. By inserting the $(1/\omega)$th power of (4.2.27) into (4.2.26), we find

$$\int_{A_{\overline{l}}} \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x+y)dy \le c\, 2^{-\overline{\nu}\cdot\overline{d}} 2^{(\overline{\nu}-\overline{l})\cdot\overline{d}/\omega} \left(\overline{M} \mathcal{X}_{\overline{\nu},m}\right)^{1/\omega}(x)\,. \tag{4.2.28}$$

Next, we replace $\mathcal{X}_{\overline{\nu},m}$ in equation (4.2.28) by $\mathcal{X}^{(p)}_{\overline{\nu},m}$, and insert this into (4.2.25):

$$\begin{aligned} 2^{\overline{l}\cdot\overline{r}} \left|k_{\overline{l}}(a_{\overline{\nu},m})(x)\right| &\le c\, 2^{\overline{l}\cdot(\overline{r}+\overline{d})} 2^{(\overline{l}-\overline{\nu})\cdot(\overline{L}+\overline{1})} 2^{-\overline{\nu}\cdot\overline{d}} 2^{(\overline{\nu}-\overline{l})\cdot\overline{d}/\omega} 2^{-\overline{\nu}\cdot\overline{d}/p} \left(\overline{M} \mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}(x) \\ &= c\, 2^{(\overline{l}-\overline{\nu})\cdot(\overline{r}+\overline{L}+\overline{1}+\overline{d}-\overline{d}/\omega)} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)} \left(\overline{M} \mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}(x)\,. \end{aligned} \tag{4.2.29}$$

From the restriction $\omega < \min(1,p,q)$ and the definition of $\overline{\sigma}_{p,q}$ it follows at once, that $\overline{d}(1/\omega - 1) > \overline{\sigma}_{p,q}$. Hence the assumption (4.2.15) implies that we can choose ω, such that $\overline{\kappa} = \overline{r} + \overline{L} + \overline{1} - \overline{d}(1/\omega - 1) > 0$, or $\overline{r} + \overline{L} + \overline{1} > \overline{d}(1/\omega - 1) > \overline{\sigma}_{p,q}$, respectively.

Substep 1.3: Mixed cases.
Exemplary we treat the terms with $l_1 \ge \nu_1$ and $l_i < \nu_i$, $i = 2, \ldots, N$, for all other cases the calculation is very similar and can be transferred correspondingly.
At first, we apply the expansion (4.2.24) for $i = 2, \ldots, N$ and use property (4.2.3) to get rid of the terms with $\overline{\beta} \le \overline{L}$. Afterwards we integrate K_1-times partially with respect to the variables y^1, similar to Step 1.1. In the resulting expression we use once more the support properties of the occurring functions together with (4.2.2). We finally obtain

$$\begin{aligned} 2^{\overline{l}\cdot\overline{r}} \left|k_{\overline{l}}(a_{\overline{\nu},m})(x)\right| \le 2^{\overline{\nu}\cdot\overline{r}} 2^{(l_1-\nu_1)(r_1-K_1)} \prod_{i=2}^{N} 2^{l_i(r_i+d_i)+(l_i-\nu_i)(L_i+1)-\nu_i r_i} \\ \times \int_{\widetilde{A}_{\overline{l}}} \mathcal{X}_{\gamma Q_{\overline{\nu},m}}(x^1 + 2^{-l_1}y^1, x^2+y^2, \ldots, x^N + y^N)dy\,, \end{aligned} \tag{4.2.30}$$

where $\widetilde{A}_{\overline{l}} := A^1_0 \times A^2_{l_2} \times \cdots \times A^N_{l_N}$.
Due to the tensor product structure of both the integrand (the set $\gamma Q_{\overline{\nu},m}$ is a product, see (3.3.11)) and the integration domain we can use Fubini's theorem to split the integral in

a d_1-dimensional one with respect to y^1 and a $(d-d_1)$-dimensional integral with respect to $y^2,\dots,y^N$.
The first one can be estimated by $c\,\mathcal{X}_{\gamma' Q^1_{\nu_1,m^1}} = c\,\mathcal{X}_{\{t:|t-2^{-\nu_1}m^1|_\infty \le \gamma' 2^{-\nu_1-1}\}}(x^1)$, where $\gamma' = \gamma+2$. This can be seen as follows: On the one hand, due to the integration domain we have that the integral is at most $|A^1_0| = c_{d_1}$. Moreover, we obtain by a consideration similar to the one leading to (4.2.26) for every $|y^1|\le 1$ and $x^1 \notin \gamma'\, Q^1_{\nu_1,m^1}$

$$
\begin{aligned}
\big|x^1 + 2^{-l_1}y^1 - 2^{-\nu_1}m^1\big|_\infty &\ge \Big|\big|x^1 - 2^{-\nu_1}m^1\big|_\infty - 2^{-l_1}\big|y^1\big|_\infty\Big| \\
&\ge \big|x^1 - 2^{-\nu_1}m^1\big|_\infty - 2^{-l_1}\big|y^1\big|_\infty \\
&> \gamma' 2^{-\nu_1-1} - 2^{-l_1} \ge (\tfrac{1}{2}\gamma+1)2^{-\nu_1} - 2^{-\nu_1} = \gamma 2^{-\nu_1-1}\,.
\end{aligned}
$$

In other words we have $\{y^1\in\mathbb{R}^{d_1} : x^1 + 2^{-l_1}y^1 \in \gamma\, Q^1_{\nu_1,m^1}\} \cap \{y^1 : |y^1|\le 1\} = \emptyset$ for every $x^1 \notin \gamma'\, Q^1_{\nu_1,m^1}$. From this we deduce the announced estimate.
For the $(d-d_1)$-dimensional integral, we use once more the estimates (4.2.26)–(4.2.28) and Lemma 4.2.1(ii) (or more precisely their respective isotropic counterparts). In this way we find for every $i=2,\dots,N$

$$
\begin{aligned}
\int_{A^i_{l_i}} \mathcal{X}_{\gamma Q^i_{\nu_i,m^i}}(x^i+y^i)dy^i &\le c\,2^{-\nu_i d_i}\mathcal{X}_{\gamma' 2^{\nu_i-l_i} Q^i_{\nu_i,m^i}}(x^i) \\
&\le C\,2^{-\nu_i d_i} 2^{(\nu_i-l_i)d_i/\omega}\big(M_i\mathcal{X}_{\nu_i,m^i}\big)^{1/\omega}(x^i)\,. \qquad (4.2.31)
\end{aligned}
$$

Inserting both parts into the above estimate and using Lemma 4.2.1(ii) now yields

$$
\begin{aligned}
&2^{\bar{l}\cdot\bar{r}}\big|k_{\bar{l}}(a_{\overline{\nu},m})(x)\big| \\
&\quad \le c\,2^{\overline{\nu}\cdot\overline{r}}2^{(l_1-\nu_1)(r_1-K_1)}\mathcal{X}_{\gamma' Q^1_{\nu_1,m^1}}(x^1) \\
&\qquad\quad \times \prod_{i=2}^N 2^{l_i(r_i+d_i)+(l_i-\nu_i)(L_i+1)-\nu_i r_i} 2^{-\nu_i d_i} 2^{(\nu_i-l_i)d_i/\omega}\Big(M_i\mathcal{X}_{\nu_i,m^i}\Big)^{1/\omega}(x^i) \\
&\quad \le C\,2^{\overline{\nu}\cdot\overline{r}}2^{-|l_1-\nu_1|(K_1-r_1)}\Big(\overline{M}\mathcal{X}_{\overline{\nu},m}\Big)^{1/\omega}(x)\prod_{i=2}^N 2^{(l_i-\nu_i)(r_i+d_i+L_i+1-d_i/\omega)} \\
&\quad = C\,2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}2^{-|l_1-\nu_1|(K_1-r_1)}\Big(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\Big)^{1/\omega}(x)\prod_{i=2}^N 2^{-|l_i-\nu_i|(r_i+d_i+L_i+1-d_i/\omega)}\,.
\end{aligned}
$$

As in Substep 1.2, the choice of ω and assumption (4.2.15) imply $r_i+d_i+L_i+1-d_i/\omega>0$ for all $i=2,\dots,N$. Furthermore, the same assumption ensures $K_1-r_1>0$, since $K_i \ge (1+[r_i])_+$ yields $K_i \ge 1+[r_i]$ and thus $K_i > r_i$ for all $i=1,\dots,N$. Altogether it follows, that there exists a vector $\overline{\rho}>0$, such that for all $x\in\mathbb{R}^d$ holds

$$
2^{\bar{l}\cdot\bar{r}}\big|k_{\bar{l}}(a_{\overline{\nu},m})(x)\big| \le c\,2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}\Big(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\Big)^{1/\omega}(x)\prod_{i=1}^N 2^{-|l_i-\nu_i|\rho_i}\,. \qquad (4.2.32)
$$

Moreover, we find that the results of the Substeps 1.1 and 1.2, estimates (4.2.23) and (4.2.29), can be rewritten in the same way. This can be achieved by taking the $1/\omega$th power of (4.2.22) and inserting this in (4.2.23). Additionally, one has to keep in mind

$K_i > r_i$ for all $i = 1, \dots, N$ by assumption (4.2.15). Hence we conclude, that the estimate (4.2.32) is valid for all $\bar{l}, \bar{\nu} \in \mathbb{N}_0^N$.

Step 2: We prove the estimate (4.2.16)**.**

Subtep 2.1: The case $q \le 1$.
An application of (4.2.32) and the monotonicity of ℓ_q-quasi-norms yield

$$\left|2^{\bar{l}\cdot\bar{r}} k_{\bar{l}}\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\lambda_{\bar{\nu},m}a_{\bar{\nu},m}\right)(x)\right|^q$$
$$\le c \sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d} 2^{\bar{\nu}\cdot(\bar{r}-\bar{d}/p)q}|\lambda_{\bar{\nu},m}|^q\left(\overline{M}\mathcal{X}_{\bar{\nu},m}^{(p)\omega}\right)^{q/\omega}(x)\prod_{i=1}^N 2^{-q|l_i-\nu_i|\rho_i}.$$

If we put $g_{\bar{\nu},m}(x) = 2^{\bar{\nu}\cdot(\bar{r}-\bar{d}/p)}\lambda_{\bar{\nu},m}\mathcal{X}_{\bar{\nu},m}^{(p)}$, compare to Remark 4.1.2, we obtain

$$\left\|\left(\sum_{\bar{l}\in\mathbb{N}_0^N}\left|2^{\bar{l}\cdot\bar{r}}k_{\bar{l}}\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\lambda_{\bar{\nu},m}a_{\bar{\nu},m}\right)\right|^q\right)^{1/q}\middle|L_p(\mathbb{R}^d)\right\|$$
$$\le c\left\|\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}2^{\bar{\nu}\cdot(\bar{r}-\bar{d}/p)q}|\lambda_{\bar{\nu},m}|^q\left(\overline{M}\mathcal{X}_{\bar{\nu},m}^{(p)\omega}\right)^{q/\omega}\sum_{\bar{l}\in\mathbb{Z}^N}\prod_{i=1}^N 2^{-q|l_i-\nu_i|\rho_i}\right)^{1/q}\middle|L_p(\mathbb{R}^d)\right\|$$
$$= c'\left\|\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}2^{\bar{\nu}\cdot(\bar{r}-\bar{d}/p)q}|\lambda_{\bar{\nu},m}|^q\left(\overline{M}\mathcal{X}_{\bar{\nu},m}^{(p)\omega}\right)^{q/\omega}\right)^{1/q}\middle|L_p(\mathbb{R}^d)\right\|$$
$$= c'\left\|\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\left(\overline{M}g_{\bar{\nu},m}^{\omega}\right)^{q/\omega}\right)^{1/q}\middle|L_p(\mathbb{R}^d)\right\| = c'\left\|\overline{M}g_{\bar{\nu},m}^{\omega}\middle|L_{p/\omega}(\ell_{q/\omega})\right\|^{1/\omega}.$$

Due to the choice of ω we have $p/\omega > 1$ and $q/\omega > 1$, hence Proposition 2.3.1 is applicable. This finally yields

$$\left\|\overline{M}g_{\bar{\nu},m}^{\omega}\left|L_{p/\omega}(\ell_{q/\omega})\right.\right\|^{1/\omega} \le c\left\|g_{\bar{\nu},m}^{\omega}\left|L_{p/\omega}(\ell_{q/\omega})\right.\right\|^{1/\omega} = c\left\|g_{\bar{\nu},m}\left|L_p(\ell_q)\right.\right\| = c\left\|\lambda\left|s_{p,q}^{\bar{r}}f^*\right.\right\|.$$

Together with (3.3.8) this proves the estimate (4.2.16).

Substep 2.2: The case $1 < q < \infty$.
Using (4.2.32) and applying Hölder's inequality yields for some arbitrary real number ε with $0 < \varepsilon < \rho_i$, $i = 1, \dots, N$,

$$\left|2^{\bar{l}\cdot\bar{r}}k_{\bar{l}}\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\lambda_{\bar{\nu},m}a_{\bar{\nu},m}\right)(x)\right|^q$$
$$\le c\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}2^{\bar{\nu}\cdot(\bar{r}-\bar{d}/p)}|\lambda_{\bar{\nu},m}|\left(\overline{M}\mathcal{X}_{\bar{\nu},m}^{(p)\omega}\right)^{1/\omega}(x)\prod_{i=1}^N 2^{-|l_i-\nu_i|\rho_i}\right)^q$$
$$\le c\left(\sum_{\bar{\nu}\in\mathbb{N}_0^N}\prod_{i=1}^N 2^{-|l_i-\nu_i|\varepsilon q'}\right)^{q/q'}$$

$$\times \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)q} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}| \left(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}(x)\right)^q \prod_{i=1}^N 2^{-|l_i-\nu_i|(\rho_i-\varepsilon)q}\right)^{q/q}$$

$$= c' \sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)q} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}| \left(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}(x)\right)^q \prod_{i=1}^N 2^{-|l_i-\nu_i|(\rho_i-\varepsilon)q} .$$

By choice of ω we have $1 < p/\omega < \infty$, $1 < q/\omega < \infty$ and $1 < 1/\omega < \infty$, hence we can apply the maximal inequality for $\overline{M}$ in the version for mixed sequence space norms (Proposition 2.3.2). Together with the abbreviation $g_{\overline{\nu},m}(x) = 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}\lambda_{\overline{\nu},m}\mathcal{X}^{(p)}_{\overline{\nu},m}$ we finally obtain

$$\begin{aligned}
&\left\| 2^{\overline{l}\cdot\overline{r}} k_{\overline{l}}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}\right) \middle| L_p(\ell_q)\right\| \\
&\le c \left\| \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)q} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}| \left(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}\right)^q \sum_{\overline{l}\in\mathbb{N}_0^N}\prod_{i=1}^N 2^{-|l_i-\nu_i|(\rho_i-\varepsilon)q}\right)^{1/q} \middle| L_p(\mathbb{R}^d)\right\| \\
&\le c' \left\| \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)q} \left(\sum_{m\in\mathbb{Z}^d} |\lambda_{\overline{\nu},m}| \left(\overline{M}\mathcal{X}^{(p)\omega}_{\overline{\nu},m}\right)^{1/\omega}\right)^q \right)^{1/q} \middle| L_p(\mathbb{R}^d)\right\| \\
&= c' \left\| \sum_{m\in\mathbb{Z}^d} \left(\overline{M} g^{\omega}_{\overline{\nu},m}\right)^{1/\omega} \middle| L_p(\ell_q)\right\| = c' \left\| \left(\sum_{m\in\mathbb{Z}^d} \left(\overline{M} g^{\omega}_{\overline{\nu},m}\right)^{1/\omega}\right)^{\omega} \middle| L_{p/\omega}(\ell_{q/\omega})\right\|^{1/\omega} \\
&= c' \left\| \overline{M} g^{\omega}_{\overline{\nu},m} \middle| L_{p/\omega}\big(\ell_{q/\omega}(\ell_{1/\omega})\big)\right\|^{1/\omega} \le c'' \left\| g^{\omega}_{\overline{\nu},m} \middle| L_{p/\omega}\big(\ell_{q/\omega}(\ell_{1/\omega})\big)\right\|^{1/\omega} \\
&= c'' \left\| g_{\overline{\nu},m} \middle| L_p(\ell_q)\right\| = c'' \left\| \lambda \middle| s^{\overline{r}}_{p,q} f^*\right\| ,
\end{aligned}$$

compare to Remark 4.1.2. This proves (4.2.16) for $1 < q < \infty$.

Substep 2.3: The case $q = \infty$.

Unfortunately, the method used above does not cover the case $q = \infty$, as no counterpart for the maximal inequality is known. We return to the estimates (4.2.21), (4.2.25)–(4.2.26) and (4.2.30)–(4.2.31). These can be summarized by

$$\begin{aligned}
2^{\overline{l}\cdot\overline{r}} \left| k_{\overline{l}}(a_{\overline{\nu},m})(x)\right| \le c\, 2^{\overline{\nu}\cdot\overline{r}} &\prod_{i:\, l_i\ge\nu_i} 2^{(l_i-\nu_i)(r_i-K_i)} \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}}(x^i) \\
&\times \prod_{i:\, l_i<\nu_i} 2^{(l_i-\nu_i)(r_i+d_i+L_i+1)} \mathcal{X}_{(\gamma+2)2^{\nu_i-l_i}Q^i_{\nu_i,m^i}}(x^i) .
\end{aligned}$$

We are going to use this product structure, in order to trace the desired estimate back to the isotropic case. To that end, we obtain for every fixed $x \in \mathbb{R}^d$

$$\begin{aligned}
&\sup_{\overline{l}\in\mathbb{N}_0^N} 2^{\overline{l}\cdot\overline{r}} \left| k_{\overline{l}}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}\right)(x)\right| \\
&\qquad \le \sum_{A\subset\{1,\dots,N\}} \sup_{\overline{l}\in\mathbb{N}_0^N} \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N:\\ \nu_i\le l_i,\, i\in A,\\ \nu_i>l_i,\, i\notin A}} \sum_{m\in I_{\overline{\nu},\overline{l}}(x)} 2^{\overline{l}\cdot\overline{r}} |\lambda_{\overline{\nu},m}| \cdot \left| k_{\overline{l}}(a_{\overline{\nu},m})(x)\right| , \qquad (4.2.33)
\end{aligned}$$

where the sets $I_{\overline{\nu},\overline{l}}(x)$ are defined by

$$I_{\overline{\nu},\overline{l}}(x) := \left\{ m \in \mathbb{Z}^d : A_{\overline{l}} \cap \left(x + \gamma Q_{\overline{\nu},m} \right) \neq \emptyset \right\} .$$

Since the kernel of the local mean $k_{\overline{l}}$ has compact support, we know that for every $x \in \mathbb{R}^d$ and $\overline{\nu}, \overline{l} \in \mathbb{N}_0^N$ the terms $k_{\overline{l}}(a_{\overline{\nu},m})(x)$ are non-vanishing for finitely many $m \in \mathbb{Z}^d$ only, i.e. these index sets $I_{\overline{\nu},\overline{l}}(x)$ are always finite. Clearly, these sets are again cross products,

$$I_{\overline{\nu},\overline{l}}(x) = I^1_{\nu_1,l_1}(x^1) \times \cdots \times I^N_{\nu_N,l_N}(x^N) .$$

With this knowledge we now can estimate (4.2.33) iteratively, where the results of one iteration serve as the coefficients in the next one, see below. Hence we treat the isotropic situation (i.e. $N = 1$) first. In those considerations we will drop the indices i.

Substep 2.3.1: We begin with the case $\nu \leq l$. Here we have to consider

$$\begin{aligned}
&\sup_{l \in \mathbb{N}_0} \sum_{\nu=0}^{l} \sum_{m \in I_{\nu,l}(x)} 2^{\nu r} 2^{(l-\nu)(r-K)} |\lambda_{\nu,m}| \mathcal{X}_{(\gamma+2)Q_{\nu,m}}(x) \\
&\quad \leq \sup_{l \in \mathbb{N}_0} \left(\sup_{\mu \leq l} \sup_{k \in I_{\mu,l}(x)} 2^{\mu r} |\lambda_{\mu,k}| \mathcal{X}_{(\gamma+2)Q_{\mu,k}}(x) \right) \sum_{\nu=0}^{l} \sum_{m \in I_{\nu,l}(x)} 2^{(l-\nu)(r-K)} \\
&\quad \leq c_0 \left(\sup_{\mu \in \mathbb{N}_0} \sup_{k \in \mathbb{Z}^d} 2^{\mu r} |\lambda_{\mu,k}| \mathcal{X}_{(\gamma+2)Q_{\mu,k}}(x) \right) \sup_{l \in \mathbb{N}_0} \sum_{\nu=0}^{l} 2^{\nu(r-K)} \\
&\quad = c_1 \sup_{\nu \in \mathbb{N}_0} \sup_{m \in \mathbb{Z}^d} 2^{\nu r} |\lambda_{\nu,m}| \mathcal{X}_{(\gamma+2)Q_{\nu,m}}(x) .
\end{aligned}$$

When estimating the count of the elements of $\# I_{\nu,l}(x)$, we immediately find $\# I_{\nu,l}(x) \leq c_0$, where c_0 depends on γ only. Moreover, we used the assumption $(1 + [r])_+ \leq K$, which yields $r < K$.

Substep 2.3.2: Now consider $\nu > l$. Then we find $\# I_{\nu,l}(x) \sim 2^{(\nu-l)d}$, and we further estimate with the help of $\mathcal{X}_{(\gamma+2)2^{\nu-l}Q_{\nu,m}}(x) \leq \left(M \mathcal{X}_{(\gamma+2)2^{\nu-l}Q_{\nu,m}} \right)^{1/\omega}(x)$ for every $x \in \mathbb{R}^d$, where $0 < \omega < \min(1,p)$, and Lemma 4.2.1

$$\begin{aligned}
&\sup_{l \in \mathbb{N}_0} \sum_{\nu=l+1}^{\infty} \sum_{m \in I_{\nu,l}(x)} 2^{(l-\nu)(r+d+L+1)} 2^{\nu r} |\lambda_{\nu,m}| \mathcal{X}_{(\gamma+2)2^{\nu-l}Q_{\nu,m}}(x) \\
&\quad \leq c \sup_{l \in \mathbb{N}_0} \sum_{\nu=l+1}^{\infty} \sum_{m \in I_{\nu,l}(x)} 2^{(l-\nu)(r+d+L+1)} 2^{\nu r} |\lambda_{\nu,m}| 2^{(\nu-l)d/\omega} \left(M \mathcal{X}_{(\gamma+2)Q_{\nu,m}} \right)^{1/\omega}(x) \\
&\quad \leq c \sup_{l \in \mathbb{N}_0} \left(\sup_{\mu > l} \sup_{k \in I_{\mu,l}(x)} 2^{\mu r} |\lambda_{\mu,k}| \left(M \mathcal{X}_{(\gamma+2)Q_{\mu,k}} \right)^{1/\omega}(x) \right) \sum_{\nu=l+1}^{\infty} \sum_{m \in I_{\nu,l}(x)} 2^{(l-\nu)(r+L+1+d-d/\omega)} \\
&\quad \leq c' \left(\sup_{\mu \in \mathbb{N}_0} \sup_{k \in \mathbb{Z}^d} 2^{\mu r} |\lambda_{\mu,k}| \left(M \mathcal{X}_{(\gamma+2)Q_{\mu,k}} \right)^{1/\omega}(x) \right) \sup_{l \in \mathbb{N}_0} \sum_{\nu=l+1}^{\infty} 2^{(l-\nu)(r+L+1-d/\omega)} \\
&\quad = c' \left(\sup_{\mu \in \mathbb{N}_0} \sup_{k \in \mathbb{Z}^d} 2^{\mu r} |\lambda_{\mu,k}| \left(M \mathcal{X}_{(\gamma+2)Q_{\mu,k}} \right)^{1/\omega}(x) \right) \sum_{\nu=1}^{\infty} 2^{-\nu(r+L+1-d/\omega)}
\end{aligned}$$

$$= c_2 \sup_{\nu\in\mathbb{N}_0} \sup_{m\in\mathbb{Z}^d} 2^{\nu r}|\lambda_{\nu,m}|\big(M\mathcal{X}_{(\gamma+2)Q_{\nu,m}}\big)^{1/\omega}(x)\,.$$

At the end we used the assumption $L > d + \sigma_p - r$, which yields that we can choose ω such that $L+1+r > d/\omega > d/\min(1,p)$. This implies $r + L + 1 - d/\omega > 0$, hence the geometrical series converges.

Substep 2.3.3: We finally prove the estimate (4.2.16)**.** As announced above this will be done iteratively. Moreover, we fix some set $A \subset \{1,\ldots,N\}$ and write it as $A = \{i_1,\ldots,i_n\}$. We will explicitly demonstrate the first iterative. With $j = i_1$ we find

$$\begin{aligned}
&\sup_{\bar{l}\in\mathbb{N}_0^N} \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N:\\ \nu_i\le l_i\,,i\in A\,,\\ \nu_i>l_i\,,i\notin A}} \sum_{m\in I_{\overline{\nu},\bar{l}}(x)} |\lambda_{\overline{\nu},m}| 2^{\overline{\nu}\cdot\overline{r}} \prod_{i\in A} 2^{(l_i-\nu_i)(r_i-K_i)} \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}}(x^i)\\
&\qquad\times \prod_{i\notin A} 2^{(l_i-\nu_i)(r_i+d_i+L_i+1)} \mathcal{X}_{(\gamma+2)2^{\nu_i-l_i}Q^i_{\nu_i,m^i}}(x^i)\\
&= \sup_{l_i\in\mathbb{N}_0,\,i\neq j} \sum_{\substack{\nu_i\le l_i,\,i=i_2,\ldots,i_n,\\ \nu_i>l_i,\,i\notin A}} \sum_{m^i\in I^i_{\nu_i,l_i}(x^i),\,i\neq j} \prod_{i\in A,\,i\neq j} 2^{\nu_i r_i} 2^{(l_i-\nu_i)(r_i-K_i)} \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}}(x^i)\\
&\qquad\times \prod_{i\notin A} 2^{\nu_i r_i} 2^{(l_i-\nu_i)(r_i+d_i+L_i+1)} \mathcal{X}_{(\gamma+2)2^{\nu_i-l_i}Q^i_{\nu_i,m^i}}(x^i)\\
&\qquad\times \left(\sup_{l_j\in\mathbb{N}_0} \sum_{\nu_j=0}^{l_j} \sum_{m^j\in I^j_{\nu_j,l_j}(x^j)} 2^{\nu_j r_j} |\lambda_{\overline{\nu},m}| 2^{(l_j-\nu_j)(r_j-K_j)} \mathcal{X}_{(\gamma+2)Q^j_{\nu_j,m^j}}(x^j) \right)\\
&\le c \sup_{l_i\in\mathbb{N}_0,\,i\neq j} \sum_{\substack{\nu_i\le l_i,\,i=i_2,\ldots,i_n,\\ \nu_i>l_i,\,i\notin A}} \sum_{m^i\in I^i_{\nu_i,l_i}(x^i),\,i\neq j} \prod_{i\in A,\,i\neq j} 2^{\nu_i r_i} 2^{(l_i-\nu_i)(r_i-K_i)} \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}}(x^i)\\
&\qquad\times \prod_{i\notin A} 2^{\nu_i r_i} 2^{(l_i-\nu_i)(r_i+d_i+L_i+1)} \mathcal{X}_{(\gamma+2)2^{\nu_i-l_i}Q^i_{\nu_i,m^i}}(x^i)\\
&\qquad\times \left(\sup_{\nu_j\in\mathbb{N}_0} \sup_{m^j\in\mathbb{Z}^{d_j}} 2^{\nu_j r_j} |\lambda_{\overline{\nu},m}| \mathcal{X}_{(\gamma+2)Q^j_{\nu_j,m^j}}(x^j) \right).
\end{aligned}$$

At the end we used the isotropic estimate together with the assumption $K_j > r_j$. For the next step, the terms in brackets serve as coefficients, they replace $\lambda_{\nu,m}$ in Substeps 2.3.1 or 2.3.2, respectively. The final result is given by

$$\begin{aligned}
&\sup_{\bar{l}\in\mathbb{N}_0^N} 2^{\bar{l}\cdot\bar{r}} \left| k_{\bar{l}}\Big(\sum_{\overline{\nu}\in\mathbb{N}_0^N} \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m} \Big)(x) \right|\\
&\qquad\le c \sup_{\overline{\nu}\in\mathbb{N}_0^N} \sup_{m\in\mathbb{Z}^d} 2^{\overline{\nu}\cdot\overline{r}} |\lambda_{\overline{\nu},m}| \prod_{i\in A} \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}}(x^i) \prod_{i\notin A} \big(M_i \mathcal{X}_{(\gamma+2)Q^i_{\nu_i,m^i}} \big)(x^i)\\
&\qquad\le c' \sup_{\overline{\nu}\in\mathbb{N}_0^N} \sup_{m\in\mathbb{Z}^d} 2^{\overline{\nu}\cdot\overline{r}} |\lambda_{\overline{\nu},m}| \big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)(x)\,.
\end{aligned}$$

The last estimate is a consequence of Lemma 4.2.1. Due to $\big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)(x) \le 1$ for all $x \in \mathbb{R}^d$ we can replace the maximal function by $\big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)^{1/\omega}$, $0 < \omega < \min(1,p,q)$. Now we apply the $L_p(\mathbb{R}^d)$-quasi-norm and use the maximal inequality for $\overline{M}$ in $L_{p/\omega}(\ell_\infty)$ (Proposition 2.3.1). Together with the abbreviation $g_{\overline{\nu},m} = 2^{\overline{\nu}\cdot\overline{r}}\lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}$ this results in

$$\left\|2^{\overline{l}\cdot\overline{r}}k_{\overline{l}}\Big(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}\lambda_{\overline{\nu},m}a_{\overline{\nu},m}\Big)\Big|L_p(\ell_\infty)\right\| \le c\left\|2^{\overline{\nu}\cdot\overline{r}}\lambda_{\overline{\nu},m}\big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)^{1/\omega}\Big|L_p(\ell_\infty)\right\|$$
$$= c\left\|\big(\overline{M}g^\omega_{\overline{\nu},m}\big)^{1/\omega}\Big|L_p(\ell_\infty)\right\| = c\left\|\overline{M}g^\omega_{\overline{\nu},m}\Big|L_{p/\omega}(\ell_\infty)\right\|^{1/\omega}$$
$$\le c'\left\|g^\omega_{\overline{\nu},m}\big|L_{p/\omega}(\ell_\infty)\right\|^{1/\omega} = c'\left\|2^{\overline{\nu}\cdot\overline{r}}\lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}\big|L_p(\ell_\infty)\right\| = c'\left\|\lambda\,\big|s^{\overline{r}}_{p,\infty}f^*\right\|.$$

This completes the proof of (4.2.16) for F-spaces.

Substep 2.4: The case $p=\infty$ for B-spaces.
The proof for the B-scale is based on the estimate (4.2.32) as well. In case $0<p<\infty$, one uses afterwards the triangle inequality to get the $L_{p/\omega}$-norm inside the $\overline{\nu}$-summation (observe $p/\omega > 1$). Finally, the maximal inequality for $L_{p/\omega}(\ell_{1/\omega})$ (Proposition 2.3.1) is applied to derive the desired estimate.
The case $p=\infty$ has to be treated separately. The essential step here is the derivation of a replacement for the (non-existing) maximal inequality in $L_\infty(\ell_{1/\omega})$. We will show the estimate

$$\left\|\sum_{m\in\mathbb{Z}^d}\Big(\overline{M}\mathcal{X}_{\overline{\nu},m}\Big)^{1/\omega}\Big|L_\infty(\mathbb{R}^d)\right\| \le C_0\,, \tag{4.2.34}$$

where C_0 is some positive constant independent of $\overline{\nu}$. At first from the tensor product structure of the characteristic functions it follows

$$\big(\overline{M}\mathcal{X}_{\overline{\nu},m}\big)(x) = (M\mathcal{X}_{Q^1_{\nu_1,m^1}})(x^1)\cdots(M\mathcal{X}_{Q^N_{\nu_N,m^N}})(x^N)\,,$$

where $Q^i_{\nu_i,m^i} = \big\{y\in\mathbb{R}^{d_i} : |y-2^{-\nu_i}m^i|_\infty \le 2^{-\nu_i-1}\big\}$, $i=1,\ldots,N$.
Hence it is sufficient to prove (4.2.34) in the isotropic setting, i.e. $N=1$. With the help of Lemma 4.2.1(i) we obtain

$$\sum_{m\in\mathbb{Z}^n}\Big(M\mathcal{X}_{\nu,m}\Big)^{1/\omega}(x) \le c\sum_{m\in\mathbb{Z}^n}\left(\frac{2^{-\nu n}}{(|x-2^{-\nu}m|_\infty+2^{-\nu})^n}\right)^{1/\omega} = c\sum_{m\in\mathbb{Z}^n}\frac{1}{(|y-m|_\infty+1)^{n/\omega}}\,.$$

Here we substituted $y=2^\nu x$. Since the last series defines a function which is 1-periodic in every direction, we may assume $|y|_\infty\le\frac{1}{2}$. Then we find for every $m\neq 0$ by means of the triangle inequality $|y-m|_\infty \ge |m|_\infty-\frac{1}{2}$. Hence we can further estimate

$$\sum_{m\in\mathbb{Z}^n}\frac{1}{(|y-m|_\infty+1)^{n/\omega}} \le 1+\sum_{m\neq 0}\frac{1}{(|m|_\infty+\frac{1}{2})^{n/\omega}} = 1+\sum_{k=1}^\infty\frac{(2k+1)^n-(2k-1)^n}{(k+\frac{1}{2})^{n/\omega}}$$
$$= \sum_{k=1}^\infty\frac{2n(2k-1+2\theta_k)^{n-1}}{2^{-n/\omega}(2k+1)^{n/\omega}} \le c\sum_{k=1}^\infty\frac{1}{(2k+1)^{n/\omega-n+1}} = C_0\,.$$

In the first line we used, that the number of n-tuples from $\mathbb{Z}^n$ with $|m|_\infty=k$ for $k\ge 1$ is given by $(2k+1)^n-(2k-1)^n$. The second line follows from the mean value theorem

for certain $\theta_k \in (0,1)$, the last series being convergent due to $\omega < 1$. Furthermore, we easily see that the constant C_0 does depend on n and ω only. Thus the estimate (4.2.34) is verified.
With the help of similar arguments than before (at first using triangle inequality, and afterwards either Hölder's inequality in case $q > 1$ or the monotonicity of ℓ_q-quasi-norms in case $q \le 1$ with respect to the $\overline{\nu}$-summation) we obtain from (4.2.32) and (4.2.34)

$$\left(\sum_{\overline{l}\in\mathbb{N}_0^N}\left\|2^{\overline{l}\cdot\overline{r}}k_{\overline{l}}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\mathbb{Z}^d}a_{\overline{\nu},m}\lambda_{\overline{\nu},m}\right)\middle|L_\infty(\mathbb{R}^d)\right\|^q\right)^{1/q}$$
$$\le c\left(\sum_{\overline{l}\in\mathbb{N}_0^N}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\left\|\sum_{m\in\mathbb{Z}^d}2^{\overline{\nu}\cdot\overline{r}}\left(\overline{M}\mathcal{X}_{\overline{\nu},m}\right)^{1/\omega}|\lambda_{\overline{\nu},m}|\prod_{i=1}^N 2^{-|l_i-\nu_i|\rho_i}\middle|L_\infty(\mathbb{R}^d)\right\|\right)^q\right)^{1/q}$$
$$\le c\left(\sum_{\overline{l}\in\mathbb{N}_0^N}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}2^{\overline{\nu}\cdot\overline{r}}\left\|\sum_{m\in\mathbb{Z}^d}\left(\overline{M}\mathcal{X}_{\overline{\nu},m}\right)^{1/\omega}\middle|L_\infty(\mathbb{R}^d)\right\|\sup_{m\in\mathbb{Z}^d}|\lambda_{\overline{\nu},m}|\prod_{i=1}^N 2^{-|l_i-\nu_i|\rho_i}\right)^q\right)^{1/q}$$
$$\le c'\left(\sum_{\overline{l}\in\mathbb{N}_0^N}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}2^{\overline{\nu}\cdot\overline{r}}\sup_{m\in\mathbb{Z}^d}|\lambda_{\overline{\nu},m}|\prod_{i=1}^N 2^{-|l_i-\nu_i|\rho_i}\right)^q\right)^{1/q}$$
$$\le c''\left(\sum_{\overline{l}\in\mathbb{N}_0^N}\sum_{\overline{\nu}\in\mathbb{N}_0^N}2^{\overline{\nu}\cdot\overline{r}q}\sup_{m\in\mathbb{Z}^d}|\lambda_{\overline{\nu},m}|^q\prod_{i=1}^N 2^{-|l_i-\nu_i|(\rho_i-\varepsilon)q}\right)^{1/q}$$
$$\le C\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}2^{\overline{\nu}\cdot\overline{r}q}\sup_{m\in\mathbb{Z}^d}|\lambda_{\overline{\nu},m}|^q\right)^{1/q} = C\left\|\lambda\left|s^{\overline{r}}_{\infty,q}b^*\right\|\right. .$$

This finally proves the estimate (4.2.16).

Step 3: It remains the proof of (ii)**.**
We prove an important special case first. To this purpose let $A \subset \{1,\dots,N\}$. We assume that additionally to (4.2.15) we have

$$L_i = -1 \quad \text{for all} \quad i\in A, \quad \text{i.e.} \quad r_i > \sigma^i_{pq}\,, \quad \text{and} \quad f\in S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\,. \tag{4.2.35}$$

In case $q=\infty$ this is complemented by $r_i > d_i$, $i=1,\dots,N$. Moreover, let $\overline{n}\in\mathbb{N}_0^N$ be a vector with $2\overline{n} > \overline{r}$. Finally, we choose functions $k_0^1,\dots,k_0^N$, $k^1,\dots,k^N$ with properties as in Theorem 3.3.3, where $c_i \mathrel{\widehat{=}} (2\pi)^{-d_i/2}$ and $\overline{M}\mathrel{\widehat{=}}\overline{n}$. Let $k_{\overline{l}}(x)$ and $k_{\overline{l}}(f)(x)$ be defined as in Theorem 3.3.2 as well as $k_{\overline{l},A}(f)(x)$ as in Lemma 3.3.1. Here we put $k_{\overline{l},A}(f)(x) = f(x)$ if $A=\emptyset$, and the occurring sums and products have to be treated accordingly. We claim, that under these assumptions it holds

$$f = \sum_{\substack{\overline{l}\in\mathbb{N}_0^N\\ l_i=0,i\notin A}} k_{\overline{l},A}(f)(x) = \lim_{P\longrightarrow\infty}\sum_{\substack{\overline{l}\le P\\ l_i=0,i\notin A}} k_{\overline{l},A}(f)(x) \tag{4.2.36}$$

with convergence in $\mathcal{S}'(\mathbb{R}^d)$. To prove this, we fix $\varphi\in\mathcal{S}(\mathbb{R}^d)$. From the definition of the local means, we have

$$(k_{\overline{l},A}(f))^\wedge(\xi) = \left(\prod_{i\in A}(2\pi)^{d_i/2}\mathcal{F}_{d_i}(k^i_{l_i})(-\xi^i)\right)\widehat{f}(\xi)\,.$$

As the Fourier transform is an isomorphic mapping from $\mathcal{S}'(\mathbb{R}^d)$ onto itself, it suffices to show

$$\varphi(\xi) \sum_{\substack{\bar{l}\leq P \\ l_i=0, i\notin A}} \Big(\prod_{i\in A} (2\pi)^{d_i/2} \mathcal{F}_{d_i}(k^i_{l_i})(-\xi^i)\Big) \longrightarrow \varphi(\xi) \qquad \text{in } \mathcal{S}(\mathbb{R}^d)\,. \tag{4.2.37}$$

The last sum can be rewritten using (3.3.15):

$$\begin{aligned}
&\sum_{\substack{\bar{l}\leq P \\ l_i=0, i\notin A}} \Big(\prod_{i\in A} (2\pi)^{d_i/2} \mathcal{F}_{d_i}(k^i_{l_i})(-\xi^i)\Big) \\
&\qquad = \prod_{i\in A} (2\pi)^{d_i/2} \sum_{l_i=0}^{P} \mathcal{F}_{d_i}(k^i_{l_i})(-\xi^i) \\
&\qquad = \prod_{i\in A} (2\pi)^{d_i/2} \Big(\mathcal{F}_{d_i}(k^i_0)(-\xi^i) + \sum_{l_i=1}^{P} \mathcal{F}_{d_i}\Big(2^{l_i d_i} k^i(2^{l_i}\cdot)\Big)(-\xi^i)\Big) \\
&\qquad = \prod_{i\in A} (2\pi)^{d_i/2} \Big(\mathcal{F}_{d_i}(k^i_0)(-\xi^i) + \sum_{l_i=1}^{P} (\mathcal{F}_{d_i} k^i)(-2^{-l_i}\xi^i)\Big) \\
&\qquad = \prod_{i\in A} (2\pi)^{d_i/2} \Big(\mathcal{F}_{d_i}(k^i_0)(-\xi^i) + \sum_{l_i=1}^{P} \Big(\mathcal{F}_{d_i} k^i_0(-2^{-l_i}\xi^i) - \mathcal{F}_{d_i} k^i_0(-2\cdot 2^{-l_i}\xi^i)\Big)\Big) \\
&\qquad = \prod_{i\in A} (2\pi)^{d_i/2} \Big(\mathcal{F}_{d_i}(k^i_0)(-2^{-P}\xi^i)\Big)\,.
\end{aligned}$$

We denote the last expression by $1-\Phi(2^{-P}\xi)$, and fix $M\in\mathbb{N}$. Bearing in mind $\varphi\in\mathcal{S}(\mathbb{R}^d)$ we find

$$\begin{aligned}
&p_M\Big(\varphi(\xi)\Phi(2^{-P}\xi)\Big) \\
&\quad = \sup_{\xi\in\mathbb{R}^d} \sup_{0\leq|\alpha|\leq M} \Big| D^\alpha\Big(\varphi\,\Phi(2^{-P}\cdot)\Big)(\xi)\Big| \langle\xi\rangle^M \\
&\quad \leq c \sup_{\xi\in\mathbb{R}^d} \sup_{0\leq|\alpha|,|\beta|\leq M} 2^{-P|\beta|} \Big|\Big(D^\alpha\varphi\Big)(\xi)\Big|\, \Big|\Big(D^\beta\Phi\Big)(2^{-P}\xi)\Big| \langle\xi\rangle^M \\
&\quad \leq c \Big(\sup_{\xi\in\mathbb{R}^d} \sup_{0\leq|\beta|\leq M} 2^{-P|\beta|} \Big|\Big(D^\beta\Phi\Big)(2^{-P}\xi)\Big| \langle\xi\rangle^{-1}\Big) \Big(\sup_{\xi\in\mathbb{R}^d} \sup_{0\leq|\alpha|\leq M} \Big|\Big(D^\alpha\varphi\Big)(\xi)\Big|\, \Big|\langle\xi\rangle^{M+1}\Big) \\
&\quad \leq c' \sup_{0\leq|\beta|\leq M} \Big(2^{-P|\beta|} \sup_{\xi\in\mathbb{R}^d} \Big|\Big(D^\beta\Phi\Big)(2^{-P}\xi)\Big| \langle\xi\rangle^{-1}\Big)\,,
\end{aligned}$$

where the constant c does not depend on P (but on M). The functionals p_M are defined by $p_M(\varphi) = \sup_{0\leq|\alpha|\leq M}\sup_{x\in\mathbb{R}^d}|D^\alpha\varphi(x)|\langle x\rangle^M$. These are a another family of norms generating the topology of $\mathcal{S}(\mathbb{R}^d)$.
In case $\beta_i>0$ for at least one β_i, i.e. $|\beta|>0$, the term in brackets in the last expression tends to 0 for $P\to\infty$ (observe that because of $k^i_0\in\mathcal{S}(\mathbb{R}^{d_i})$ the function Φ and all its derivatives are bounded, hence the supremum with respect to ξ is finite and bounded independent of P).

For $\beta = 0$ we split the supremum in $\sup_{|\xi|\geq 2^P}$ and $\sup_{|\xi|<2^P}$. The first one can be estimated by $c\,2^{-P}$ due to $\langle\xi\rangle^{-1} \leq 2^{-P}$. To estimate the other one we remark, that from (3.3.13) follows $\Phi(0) = 0$, and hence the mean value theorem yields $|\Phi(\xi)| \leq c\,|\xi|$ in $\{\xi : |\xi| \leq 1\}$. Thus it holds

$$\sup_{|\xi|\leq 2^P} |\Phi(2^{-P}\xi)|\langle\xi\rangle^{-1} \leq c \sup_{|\xi|\leq 2^P} |2^{-P}\xi|\langle\xi\rangle^{-1} \leq c\,2^{-P} \sup_{\xi\in\mathbb{R}^d} |\xi|\langle\xi\rangle^{-1} \leq c\,2^{-P} .$$

Altogether, this shows $p_M\Big(\varphi(\xi)\Phi(2^{-P}\xi)\Big) \leq C\,2^{-P} \longrightarrow 0$ for $P \longrightarrow \infty$. Consequently (4.2.37) is proved as well as (4.2.36).
Next we choose a compactly supported non-negative function $\psi \in \mathcal{S}(\mathbb{R}^d)$ with the property

$$\sum_{m\in\mathbb{Z}^d} \psi(x-m) = 1 \quad \text{for all} \quad x \in \mathbb{R}^d . \tag{4.2.38}$$

Furthermore, for every $\overline{\nu} \in \mathbb{N}_0^N$ and $m \in \mathbb{Z}^d$ we define $\psi_{\overline{\nu},m}(x) := \psi(2^{\overline{\nu}}x - m)$. Then there exists some $\gamma > 1$, such that

$$\operatorname{supp}\psi_{\overline{\nu},m} \subset \gamma Q_{\overline{\nu},m} \quad \text{for all} \quad \overline{\nu} \in \mathbb{N}_0^N\,, m \in \mathbb{Z}^d . \tag{4.2.39}$$

We multiply equation (4.2.36) with these decompositions of unity to obtain

$$\begin{aligned} f &= \sum_{\substack{\overline{l}\in\mathbb{N}_0^N \\ l_i=0, i\notin A}} k_{\overline{l},A}(f)(x) = \sum_{\substack{\overline{l}\in\mathbb{N}_0^N \\ l_i=0, i\notin A}} k_{\overline{l},A}(f)(x)\Bigg(\sum_{m\in\mathbb{Z}^d} \psi_{\overline{l},m}(x)\Bigg) \\ &= \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i=0, i\notin A}} \sum_{m\in\mathbb{Z}^d} k_{\overline{\nu},A}(f)(x)\psi_{\overline{\nu},m}(x) =: \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i=0, i\notin A}} \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{\nu},m} a_{\overline{\nu},m}(x)\,, \end{aligned} \tag{4.2.40}$$

where

$$\begin{aligned} \lambda_{\overline{\nu},m} &= \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} \sup_{y\in\gamma Q_{\overline{\nu},m}} \big|D^\alpha[k_{\overline{\nu},A}(f)](y)\big|\,, \qquad \big(\overline{K}_A\big)_i = \mathcal{X}_A(i)K_i\,, \\ a_{\overline{\nu},m}(x) &= \lambda_{\overline{\nu},m}^{-1}\psi_{\overline{\nu},m}(x)k_{\overline{\nu},A}(f)(x)\,. \end{aligned}$$

In case $\lambda_{\overline{\nu},m} = 0$ we define $a_{\overline{\nu},m} = 0$. On the other hand, these coefficients are always finite since $\big|D^\alpha[k_{\overline{\nu},A}(f)](y)\big| \leq c\,\|k_{\overline{\nu},A}\|_{d+1,|\alpha|}\|f|S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\|$. Moreover, we put for every $\overline{\nu} \in \mathbb{N}_0^N$ with $\nu_i \neq 0$ for some $i \notin A$ just $\lambda_{\overline{\nu},m} = 0$ and $a_{\overline{\nu},m} = 0$. It follows, that the functions $a_{\overline{\nu},m}$ are $[\,\overline{K}, \overline{L}\,]$-atoms centred at $Q_{\overline{\nu},m}$.
The required differentiability follows from smoothness properties of the convolution and from the assumption $f \in S^{\overline{K}}C(\mathbb{R}^d)$. The support property (4.2.1) is an immediate consequence of the choice of $\psi_{\overline{\nu},m}$. Furthermore, by assumption (4.2.35) we always have either $L_i = -1$, $i \in A$, and $\nu_i = 0$, $i \notin A$, or $a_{\overline{\nu},m} = 0$. Hence, no moment conditions need to be checked. Eventually, property (4.2.2) follows from the definition of the coefficients $\lambda_{\overline{\nu},m}$ up to some constant C independent of $\overline{\nu}$, m and x:

$$\begin{aligned} \big|D^\alpha a_{\overline{\nu},m}(x)\big| &= \lambda_{\overline{\nu},m}^{-1}\big|D^\alpha\big(k_{\overline{\nu},A}(f)\psi_{\overline{\nu},m}\big)(x)\big| \\ &\leq c\,\lambda_{\overline{\nu},m}^{-1} \sum_{\beta+\gamma=\alpha} \big|D^\beta\big(k_{\overline{\nu},A}(f)\big)(x)\big|\big|\big(D^\gamma\psi_{\overline{\nu},m}\big)(x)\big| \end{aligned}$$

$$\leq c' \sum_{\gamma\leq\alpha} \big|\big(D^\gamma \psi_{\overline{\nu},m}\big)(x)\big| = c' \sum_{\gamma\leq\alpha} \big|D^\gamma\big(\psi(2^{\overline{\nu}}\cdot -m)\big)(x)\big|$$
$$= c' \sum_{\gamma\leq\alpha} 2^{\overline{\gamma}\cdot\overline{\nu}} \big|\big(D^\gamma\psi\big)(2^{\overline{\nu}}x-m)\big| \leq c'\, 2^{\overline{\alpha}\cdot\overline{\nu}} \sum_{\gamma\leq\alpha} \big|\big(D^\gamma\psi\big)(2^{\overline{\nu}}x-m)\big|$$
$$\leq c'\, 2^{\overline{\alpha}\cdot\overline{\nu}} \sum_{|\gamma|\leq|\overline{K}|} \sup_{x\in\mathbb{R}^d} \big|\big(D^\gamma\psi\big)(x)\big| = c'\, 2^{\overline{\alpha}\cdot\overline{\nu}} \|\psi\|_{0,|\overline{K}|}\,.$$

In order to prove that this decomposition indeed fulfills (4.2.17), we estimate

$$\big\|\, \lambda \,\big|s^{\overline{r}}_{p,q}f^*\big\| = \Bigg\| \Bigg(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N\\ \nu_i=0, i\notin A}} \sum_{m\in\mathbb{Z}^d} 2^{\overline{\nu}\cdot\overline{r}q} |\mathcal{X}_{\overline{\nu},m}\lambda_{\overline{\nu},m}|^q \Bigg)^{\frac{1}{q}} \Bigg| L_p(\mathbb{R}^d) \Bigg\|$$
$$\leq c \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} \Bigg\| \Bigg(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N\\ \nu_i=0, i\notin A}} \sum_{m\in\mathbb{Z}^d} 2^{\overline{\nu}\cdot\overline{r}q} \mathcal{X}_{\overline{\nu},m}(x) \sup_{y\in\gamma Q_{\overline{\nu},m}} \big|D^\alpha[k_{\overline{\nu},A}(f)](y)\big|^q \Bigg)^{\frac{1}{q}} \Bigg| L_p(\mathbb{R}^d) \Bigg\|$$
$$\leq c \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} \Bigg\| \Bigg(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N\\ \nu_i=0, i\notin A}} \sum_{m\in\mathbb{Z}^d} 2^{\overline{\nu}\cdot\overline{r}q} \mathcal{X}_{\overline{\nu},m}(x) \sup_{x-y\in\gamma' Q_{\overline{\nu},0}} \big|D^\alpha[k_{\overline{\nu},A}(f)](y)\big|^q \Bigg)^{\frac{1}{q}} \Bigg| L_p(\mathbb{R}^d) \Bigg\|$$
$$= c \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} \Bigg\| \Bigg(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N\\ \nu_i=0, i\notin A}} 2^{\overline{\nu}\cdot\overline{r}q} \sup_{x-y\in\gamma' Q_{\overline{\nu},0}} \big|D^\alpha[k_{\overline{\nu},A}(f)](y)\big|^q \Bigg)^{\frac{1}{q}} \Bigg| L_p(\mathbb{R}^d) \Bigg\|$$

and apply Lemma 3.3.1 with $D^{\alpha^i}k_0^i$ and $D^{\alpha^i}k^i$ instead of k_0^i and k^i.
On the one hand these new kernels do no longer satisfy the Tauberian conditions (3.3.6), but by Proposition 3.2.2 these are not necessary for the proof of (4.2.17), compare also with the proof of Theorem 3.2.2 (observe that at this point we only need a one-sided estimate). We obtain

$$\big\|\, \lambda \,\big|s^{\overline{r}}_{p,q}f^*\big\| \leq c \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} \Bigg\| \Bigg(\sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N\\ \nu_i=0, i\notin A}} 2^{\overline{\nu}\cdot\overline{r}q} \sup_{x-y\in\gamma' Q_{\overline{\nu},0}} |[D^\alpha k_{\overline{\nu},A}](f)(y)|^q \Bigg)^{1/q} \Bigg| L_p(\mathbb{R}^d) \Bigg\|$$
$$\leq c \sum_{0\leq\overline{\alpha}\leq\overline{K}_A} c_\alpha \big\|\, f \,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| = C \,\big\|\, f \,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,.$$

Step 4: The general case.
We now prove the existence of an optimal decomposition for all $\overline{r}\in\mathbb{R}^N$ and all $\overline{L}$ satisfying (4.2.15). At first we remark, that instead of the lifting operator from Definition 2.3.2 we can define another operator upon replacing the Fourier multiplier $\prod_{i=1}^N (1+|x^i|^2)^{\rho_i/2}$ by $\prod_{i=1}^N (1+|x^i|^{\rho_i})$ for some $\rho\in\mathbb{R}^N$. Using Proposition 2.3.5 twice we obtain a counterpart of Proposition 2.3.11. We shall use the abbreviations

$$\Delta_A := \prod_{i\in A}\Delta_i\,, \quad \Delta_A^{\overline{M}_A} = \prod_{i\in A}\Delta_i^{M_i}\,, \quad \overline{k}\in\mathbb{N}_0^N \Longrightarrow \big(\overline{k}_A\big)_i := \begin{cases} k_i, & i\in A\\ 0, & i\notin A\end{cases}$$

for some set $A \subset \{1,\dots,N\}$. Then we can decompose f as

$$f = g + \sum_{\emptyset \subsetneq A \subset \{1,\dots,N\}} \Delta_A^{\overline{M}_A} g = \sum_{A \subset \{1,\dots,N\}} \Delta_A^{\overline{M}_A} g\,, \tag{4.2.41}$$

where $\overline{M} \in 2\mathbb{N}_0^N$ is still at our disposal and can be chosen arbitrarily large. Moreover, we have $g = I_{-2\overline{M}} f \in S^{\overline{r}+2\overline{M}}_{p,q}F(\mathbb{R}^{\overline{d}})$, and by Proposition 2.3.11 it holds $\big\| g \,\big| S^{\overline{r}+2\overline{M}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \sim \big\| f \,\big| S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|$. We now decompose every summand in (4.2.41). To begin with, we choose $\overline{M}$, such that it holds for some $\varepsilon > 0$

$$\begin{aligned}\big\| g \,\big| S^{\overline{K}}C(\mathbb{R}^{\overline{d}})\big\| &\le \big\| g \,\big| S^{\overline{K}}_{\infty,1}B(\mathbb{R}^{\overline{d}})\big\| \le c \,\big\| g \,\big| S^{\overline{K}+\overline{d}/p+\varepsilon}_{p,p}B(\mathbb{R}^{\overline{d}})\big\| \\ &\le c' \,\big\| g \,\big| S^{\overline{K}+\overline{d}/p+2\varepsilon}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \le c' \,\big\| g \,\big| S^{\overline{r}+2\overline{M}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,,\end{aligned}$$

compare with Corollary 2.3.2, Proposition 2.3.10 and Proposition 2.3.7. For an arbitrary set A with $\emptyset \subset A \subset \{1,\dots,N\}$ we use the decomposition

$$g(x) = \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i = 0, i\notin A}} \sum_{m\in\mathbb{Z}^d} \psi_{\overline{\nu},m}(x) k_{\overline{\nu},A}(g)(x) = \sum_{\substack{\overline{\nu}\in\mathbb{N}_0^N \\ \nu_i = 0, i\notin A}} \sum_{m\in\mathbb{Z}^d} \lambda^A_{\overline{\nu},m} a^A_{\overline{\nu},m}(x)\,,$$

as in (4.2.36) and (4.2.40), respectively. Here, it is

$$\begin{aligned}\lambda^A_{\overline{\nu},m} &= c_1 \sum_{\overline{\beta} \le \overline{K}_A + 2\overline{M}_A} \sup_{y\in c_2 Q_{\overline{\nu},m}} \Big| D^\beta \Big(k_{\overline{\nu},A}(g)\Big)(y)\Big|\,, \\ a^A_{\overline{\nu},m}(x) &= \big(\lambda^A_{\overline{\nu},m}\big)^{-1} \psi_{\overline{\nu},m}(x) k_{\overline{\nu},A}(g)(x)\,.\end{aligned}$$

For the remaining $\overline{\nu}$, i.e. those with $\nu_i \neq 0$ for some $i \notin A$, we put $\lambda^A_{\overline{\nu},m} = 0$ and $a^A_{\overline{\nu},m} = 0$. We now further assume $\overline{M}$ to be large enough, such that $r_i + 2M_i > \sigma^i_{p,q}$ for all $i = 1,\dots,N$, and for $q = \infty$ additionally $r_i + 2M_i > d_i$. Choosing also c_1, c_2 large enough, then the functions $a^A_{\overline{\nu},m}$ are $[\,\overline{K} + 2\overline{M}, -\overline{1}\,]$-atoms, and we obtain by Step 3 the estimate

$$\big\|\lambda^A \big| s^{\overline{r}+2\overline{M}}_{p,q} f^*\big\| \le c \,\big\| g \,\big| S^{\overline{r}+2\overline{M}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| \le c' \,\big\| f \,\big| S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,.$$

Moreover, it follows that the functions $2^{-2\overline{\nu}\cdot\overline{M}_A} \Delta_A^{\overline{M}_A} a^A_{\overline{\nu},m}$ are $[\,\overline{K}, 2\overline{M} - \overline{1}\,]$-atoms, where $\overline{L} = 2\overline{M} - \overline{1}$ satisfies (4.2.15).
The support and differentiability properties follow as in Step 3. The moment conditions for $\Delta_A^{\overline{M}_A} a^A_{\overline{\nu},m}$ can be obtained from

$$\int_{\mathbb{R}^{d_i}} (x^i)^\alpha D^{2\beta} a^A_{\overline{\nu},m}(x)\, dx^i = 0\,, \qquad \beta \in \mathbb{N}_0^d, \overline{\beta} = \overline{M}_A\,, \quad \alpha \in \mathbb{N}_0^{d_i}\,, |\alpha^i| \le 2M_i - 1\,, \quad i \in A\,.$$

This in turn follows by partial integration together with the compact support of the functions $a^A_{\overline{\nu},m}$. Since for $i \notin A$ we either have $\nu_i = 0$, or $\nu_i \neq 0$ and hence $a^A_{\overline{\nu},m} \equiv 0$, no further moment conditions need to be checked. Moreover, property (4.2.2) for $a^A_{\overline{\nu},m}$ and $2^{-2\overline{\nu}\cdot\overline{M}_A} \Delta_A^{\overline{M}_A} a^A_{\overline{\nu},m}$, respectively, follows by a calculation similar to the one in Step 3. Due to the continuity of the differential operators on $\mathcal{S}'(\mathbb{R}^d)$ together with Lemma 4.2.2 this yields an atomic decomposition for $\Delta_A^{\overline{M}_A} g$.

Finally, because of (4.2.41) the sum of all these (finitely many) decompositions gives a decomposition of f. In order to see that this one has the desired properties we put

$$\widetilde{\lambda}^A = 2^{2\overline{\nu}\cdot\overline{M}_A}\lambda^A \qquad \text{and} \qquad \lambda = \sum_{\emptyset\subset A\subset\{1,\dots,N\}} \widetilde{\lambda}^A\,.$$

Then we immediately find $\big\|\,\widetilde{\lambda}^A\,\big|s^{\overline{r}}_{p,q}f^*\big\| = \big\|\,\lambda^A\,\big|s^{\overline{r}+2\overline{M}}_{p,q}f^*\big\|$ and hence

$$\big\|\,\lambda\,\big|s^{\overline{r}}_{p,q}f^*\big\| \le c \sum_{\emptyset\subset A\subset\{1,\dots,N\}} \big\|\,\lambda^A\,\big|s^{\overline{r}+2\overline{M}}_{p,q}f^*\big\| \le c'\,\big\|\,f\,\big|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\|\,.$$

This completes the proof. □

Remark 4.2.1. The proof uses essentially the same methods as the according results for atomic decompositions of spaces $A^s_{p,q}(\mathbb{R}^n)$ in [85] and for spaces $S^r_{p,q}A(\mathbb{R}^d)$ in [94]. However, the estimates in Substeps 2.2–2.4 differ slightly from those proofs, because both contained a minor gap in their argumentation which could be closed here. For finite parameters this was done with very little additional effort, but for F-spaces and $q=\infty$ this happened at the cost of additional moment conditions.
Another way to circumvent such difficulties is the investigation of *molecules*. In contrast to atoms these possess only polynomial decay (instead of compact support), but they allow similar decomposition theorems, see e.g. [30] (the φ-transform yields a particular molecular decomposition), [9] or [66]. In that framework one would prove part (i) of Theorem 4.2.1 for molecules, while the proof of part (ii) remains the same. This strategy then would yield both, an atomic and a molecular characterization, simultaneously, see also [43] for a realization of that approach.

Corollary 4.2.1. Under the assumptions of Theorem 4.2.1

$$\big\|\,f\,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|^* := \inf\big\|\,\lambda\,\big|s^{\overline{r}}_{p,q}a^*\big\|$$

defines an equivalent quasi-norm is $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. Here the infimum is taken over all sequences $\lambda \in s^{\overline{r}}_{p,q}a^*$ and all families $\big(a_{\overline{\nu},m}\big)_{\overline{\nu}\in\mathbb{N}_0^N,m\in\mathbb{Z}^d}$ of admissible atoms, such that the series (4.2.16) converges in $\mathcal{S}'(\mathbb{R}^d)$ to f.

4.3 Wavelets

We remind of the constructions of wavelet bases in Sections 1.2.2 and 1.4.4. In this section we will combine these constructions to obtain further bases adapted to the splitting $\mathbb{R}^d = \mathbb{R}^{d_1}\times\cdots\times\mathbb{R}^{d_N}$. Let

$$\Psi^j = \Big\{\psi^{i,j}_{k,m} : k\in\mathbb{N}_0, m\in\mathbb{Z}^n, i\in I_k\Big\} \subset C^{s_j}(\mathbb{R}^{d_j}), \quad j=1,\dots,N,$$

be wavelet bases according to Proposition 1.2.1. We then define functions

$$\Psi_{\overline{i},\overline{k},m}(x) := 2^{\overline{k}\cdot\overline{d}/2}\Psi_{\overline{i}}(2^{\overline{k}}x-m) := 2^{\overline{k}\cdot\overline{d}/2}\psi^{i_1,1}_{k_1,m^1}(x^1)\cdots\psi^{i_N,N}_{k_N,m^N}(x^N)\,, \tag{4.3.1}$$

where

$$x = (x^1,\dots,x^N)\in\mathbb{R}^d\,, \qquad x^j\in\mathbb{R}^{d_j}\,,$$

$$m = (m^1, \dots, m^N) \in \mathbb{Z}^d\,, \qquad m^j \in \mathbb{Z}^{d_j}\,,$$
$$\bar{i} = (i_1, \dots, i_N) \in \{0, \dots, 2^{d_1} - 1\} \times \dots \times \{0, \dots, 2^{d_N} - 1\} =: I_{\overline{d}}\,,$$
$$\overline{k} = (k_1, \dots, k_n) \in \mathbb{N}_0^N \quad \text{with} \quad k_j = 0 \quad \text{if} \quad i_j = 0\,.$$

This results in the following proposition.

Proposition 4.3.1. For arbitrary $s_1, \dots, s_N \in \mathbb{N}$ there exist real-valued and compactly supported functions $\psi^{0,j}, \psi^{1,j} \in C^{s_j}(\mathbb{R})$, $j = 1, \dots, N$, with property (1.2.5), such that

$$\Psi = \left\{ \Psi_{\bar{i},\overline{k},m} : \bar{i} \in I_{\overline{d}}, \overline{k} \in \mathbb{N}_0^N \text{ with } k_j > 0 \text{ if } i_j > 0, m \in \mathbb{Z}^d \right\}, \tag{4.3.2}$$

is an orthonormal basis in $L_2(\mathbb{R}^d)$, where $\Psi_{\bar{i},\overline{k},m}$ is defined as in (4.3.1).

For the characterizations of the Besov and Triebel-Lizorkin spaces the indices $\bar{i}$ and m play only a minor role. Hence we will combine them to one index by putting

$$\Gamma_0^j := \left\{0, \dots, 2^{d_j} - 1\right\} \times \mathbb{Z}^{d_j}\,, \qquad \Gamma_k^j := \left\{1, \dots, 2^{d_j} - 1\right\} \times \mathbb{Z}^{d_j}, \quad j \in \mathbb{N}\,,$$
$$\Gamma_{\overline{k}} := \Gamma_{k_1}^1 \times \dots \times \Gamma_{k_N}^N\,, \quad \overline{k} \in \mathbb{N}_0^N\,.$$

The elements $\gamma \in \Gamma_{\overline{k}}$ will simply be written as $\gamma = \big(\bar{i}, m\big)$, where $(i_j, m^j) \in \Gamma_{k_j}^j$, $j = 1, \dots, N$. Then the wavelet system (4.3.2) can be rewritten as

$$\Psi = \left\{ \Psi_{\overline{k},\gamma} \;:\; \gamma \in \Gamma_{\overline{k}}\,, \overline{k} \in \mathbb{N}_0^N \right\}, \tag{4.3.3}$$

where we used the abbreviation $\Psi_{\overline{k},\gamma} := \Psi_{\bar{i},\overline{k},m}$ for every $\gamma = \big(\bar{i}, m\big) \in \Gamma_{\overline{k}}$.
To formulate the theorem on the wavelet decomposition, we need to modify the sequence spaces introduced in Definition 4.1.1.

Definition 4.3.1. Let $0 < p, q \le \infty$ and $\overline{r} \in \mathbb{R}^N$. For sequences

$$\lambda = \left\{ \lambda_{\overline{\nu},\gamma} \in \mathbb{C} \;:\; \overline{\nu} \in \mathbb{N}_0^N, \gamma \in \Gamma_{\overline{\nu}} \right\}$$

we define

$$s_{p,q}^{\overline{r}} b := \left\{ \lambda \;:\; \| \lambda \,|\, s_{p,q}^{\overline{r}} b \| < \infty \right\},$$
$$\| \lambda \,|\, s_{p,q}^{\overline{r}} b \| := \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu} \cdot (\overline{r} + \overline{d}/2 - \overline{d}/p) q} \left(\sum_{\gamma \in \Gamma_{\overline{\nu}}} |\lambda_{\overline{\nu},\gamma}|^p \right)^{q/p} \right)^{1/q}$$

as well as for $p < \infty$

$$s_{p,q}^{\overline{r}} f := \left\{ \lambda \;:\; \| \lambda \,|\, s_{p,q}^{\overline{r}} f \| < \infty \right\},$$
$$\| \lambda \,|\, s_{p,q}^{\overline{r}} f \| = \left\| \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} \sum_{\gamma \in \Gamma_{\overline{\nu}}} \left| 2^{\overline{\nu} \cdot (\overline{r} + \overline{d}/2)} \lambda_{\overline{\nu},\gamma} \mathcal{X}_{\overline{\nu},\gamma}(\cdot) \right|^q \right)^{1/q} \middle| L_p(\mathbb{R}^d) \right\|$$

with the usual modifications in case p and/or q are infinite. Here it is $\mathcal{X}_{\overline{\nu},\gamma} = \mathcal{X}_{\overline{\nu},m}$ for $\gamma = \big(\bar{i}, m\big)$. We use the notation $s_{p,q}^{\overline{r}} a$ as before.

Remark 4.3.1. For atomic decompositions of the form $\sum_{\overline{k}\in\mathbb{N}_0^N}\sum_{\gamma\in\Gamma_{\overline{k}}}\lambda_{\overline{k},\gamma}a_{\overline{k},\gamma}$, together with sequences as in Definition 4.3.1 and according sequence spaces, an analogue of Theorem 4.2.1 holds true. Replacing the m-summation by the γ-summation has no consequence for the used arguments, only the constants involved are altered. Hence, if we refer to Theorem 4.2.1 in the sequel this modification is meant.

Remark 4.3.2. Besides the modified summation domain for the inner sum the sequence spaces $s^{\overline{r}}_{p,q}a$ from Definition 4.3.1 and the spaces $s^{\overline{r}}_{p,q}a^*$ differ also in their normalization, i.e. in the exponent of the weight. This is caused by the different normalizations of atoms and wavelets. On the one hand the atoms are L_∞-normalized bounded functions due to property (4.2.2), on the other hand the wavelets are L_2-normalized since they are assumed to form an orthonormal basis.
Moreover, for these sequence spaces the fact $S^{\overline{0}}_{2,2}F(\mathbb{R}^{\overline{d}}) = L_2(\mathbb{R}^d)$ (compare to Proposition 2.3.13) has its counterpart in the observation $s^{\overline{0}}_{2,2}f = \ell_2$.

Now we have the necessary definitions to state the result on the wavelet-decomposition. The decomposition described in the theorem and the succeeding corollary on the discretization of the function spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ are the main results of the first part of this thesis.

Theorem 4.3.1. Let $\overline{r}\in\mathbb{R}^N$, $0<p,q\le\infty$ ($p<\infty$ for F-spaces), and $\overline{s}\in\mathbb{N}_0$. Then the following assertions hold true:

(i) Let $\lambda\in s^{\overline{r}}_{p,q}a$, and let

$$\begin{aligned} s_j &> \max\big(r_j, \sigma_p^j - r_j\big) && \text{for } B\text{-spaces}, \\ s_j &> \max\big(r_j, \sigma_{p,q}^j - r_j\big) && \text{for } F\text{-spaces}, 0<q<\infty, \\ s_j &> \max\big(r_j, d_j + \sigma_p^j - r_j\big) && \text{for } F\text{-spaces}, q=\infty, \end{aligned} \tag{4.3.4}$$

respectively, for all $j=1,\dots,N$. Then it holds:

(a) The series

$$\sum_{\overline{k}\in\mathbb{N}_0^N}\sum_{\gamma\in\Gamma_{\overline{k}}}\lambda_{\overline{k},\gamma}\Psi_{\overline{k},\gamma} \tag{4.3.5}$$

converges in $\mathcal{S}'(\mathbb{R}^d)$ to some distribution f.

(b) The distribution f belongs to $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, and we have

$$\big\| f\,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \le c\,\big\|\,\lambda\,\big|\,s^{\overline{r}}_{p,q}a\,\big\|, \tag{4.3.6}$$

where the constant c does not depend on λ.

(c) The series (4.3.5) converges unconditionally in $S^{\overline{r}-\varepsilon}_{p,q}A(\mathbb{R}^{\overline{d}})$ for every $\varepsilon>0$.

(d) If additionally we have $\max(p,q)<\infty$, then (4.3.5) converges unconditionally in $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ as well.

(ii) Let $f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$. We define the sequence λ by

$$\lambda_{\overline{k},\gamma} := \big(f, \Psi_{\overline{k},\gamma}\big)\,, \qquad \overline{k} \in \mathbb{N}_0^N\,, \gamma \in \Gamma_{\overline{k}}\,. \tag{4.3.7}$$

If $\overline{s} \in \mathbb{N}_0^N$ is large enough, then it follows:

(a) The sequence λ belongs to $s^{\overline{r}}_{p,q}a$, and it holds

$$\big\|\,\lambda\,\big|\,s^{\overline{r}}_{p,q}a\,\big\| \le c\,\big\|\,f\,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\|\,,$$

where the constant c does not depend on f.

(b) The series (4.3.5) converges to f in $\mathcal{S}'(\mathbb{R}^d)$.

(c) If the series $\sum_{\overline{k}\in\mathbb{N}_0^N}\sum_{\gamma\in\Gamma_{\overline{k}}}\eta_{\overline{k},\gamma}\Psi_{\overline{k},\gamma}$ converges in $\mathcal{S}'(\mathbb{R}^d)$ to f for some sequence $\eta \in s^{\overline{r}}_{p,q}a$, then we have $\eta = \lambda$.

Remark 4.3.3. Before we turn to the proof of this theorem we have to remark on the problems caused by the limited smoothness of the functions $\Psi_{\overline{k},\gamma}$.
With limited smoothness we refer to the fact, that all functions $\Psi_{\overline{k},\gamma}$ do *not* belong to $\mathcal{S}(\mathbb{R}^d)$. According to Theorem 1.2.1 and the constructions in (1.2.8) and (4.3.1) we only have $\Psi_{\overline{k},\gamma} \in S^{\overline{s}}C(\mathbb{R}^{\overline{d}})$, $\overline{s} = (s_1, \ldots, s_N) \in \mathbb{N}_0^N$. Hence, the expressions $\big(f, \Psi_{\overline{k},\gamma}\big)$ in (ii) cannot be interpreted in the distributional sense as a dual pairing $\big(\mathcal{S}(\mathbb{R}^d), \mathcal{S}'(\mathbb{R}^d)\big)$ right from the beginning, at least not for arbitrary $f \in \mathcal{S}'(\mathbb{R}^d)$. To give the symbol $\big(f, \Psi_{\overline{k},\gamma}\big)$ a meaning anyhow, we use the assertions on the dual spaces of $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$ in Propositions 2.3.15 and 2.3.16 in Section 2.3.8.
The functions $D^\alpha\Psi_{\overline{k},\gamma}$, $0 \le \overline{\alpha} \le \overline{s}$, are bounded, continuous, and compactly supported. Hence it follows at once

$$\big\|D^\alpha\Psi_{\overline{k},\gamma}\big|L_{\widetilde{p}}(\mathbb{R}^d)\big\| < \infty\,, \qquad 0 \le \overline{\alpha} \le \overline{s}, \quad 0 < \widetilde{p} \le \infty\,.$$

In particular, this yields (see Theorems 2.1.1 and 2.3.4)

$$\Psi_{\overline{k},\gamma} \in S^{\overline{s}}_{\widetilde{p}}W(\mathbb{R}^{\overline{d}}) = S^{\overline{s}}_{\widetilde{p},2}F(\mathbb{R}^{\overline{d}}) \quad \text{for all} \quad 1 < \widetilde{p} < \infty\,.$$

Together with Proposition 2.3.7, the Sobolev embedding (Proposition 2.3.10), and the assumption (4.3.4) we obtain for a suitably chosen $\widetilde{p}$

$$S^{\overline{s}}_{\widetilde{p},2}F(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{s}-\varepsilon\overline{1}}_{\widetilde{p},\widetilde{p}}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{-\overline{r}+\varepsilon\overline{1}+\overline{\sigma}_p}_{p',p'}B(\mathbb{R}^{\overline{d}}) = \big[S^{\overline{r}-\varepsilon\overline{1}}_{p,p}B(\mathbb{R}^{\overline{d}})\big]'\,,$$

for some $\varepsilon > 0$. Thus, for every $f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}-\varepsilon\overline{1}}_{p,p}B(\mathbb{R}^{\overline{d}})$ (compare to Proposition 2.3.7) we can interpret $\Psi_{\overline{k},\gamma}$ as a bounded linear functional defined on a space containing f. The expression $\big(f, \Psi_{\overline{k},\gamma}\big)$ then means the value of this functional upon inserting f.
We can reverse these arguments. As before, the functions $\Psi_{\overline{k},\gamma}$ belong to $S^{\overline{s}}_{\widetilde{p},2}F(\mathbb{R}^{\overline{d}})$, $1 < \widetilde{p} < \infty$, and $S^{\overline{s}}_{\widetilde{p},2}F(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{s}-\varepsilon\overline{1}}_{\widetilde{p},\widetilde{p}}B(\mathbb{R}^{\overline{d}})$. On the other hand, we have $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}}) \hookrightarrow \big[S^{\overline{s}-\varepsilon\overline{1}}_{\widetilde{p},\widetilde{p}}B(\mathbb{R}^{\overline{d}})\big]'$ for some suitably chosen $\widetilde{p}$. Hence we can interpret f as a bounded linear functional defined on a space containing $\Psi_{\overline{k},\gamma}$. Then $\big(f, \Psi_{\overline{k},\gamma}\big)$ means the value of this functional for $\Psi_{\overline{k},\gamma}$.

Proof of Theorem 4.3.1(i). Let $\lambda \in s^{\overline{r}}_{p,q}f$. The functions $2^{-\overline{k}\cdot\overline{d}/2}\Psi_{\overline{k},\gamma}$ are $[\overline{s},\overline{s}]$-atoms centred at $Q_{\overline{k},m}$ (possibly up to some normalizing constant), where $\gamma = \big(\overline{i}, m\big)$ and $\overline{s} = (s_1,\ldots,s_N) \in \mathbb{N}_0^N$. Moreover, the condition (4.3.4) implies

$$s_j \geq \max\Big\{\big(1+[r_j]\big)_+ , \big[\sigma^j_{p,q} - r_j\big]\Big\}, \quad j = 1,\ldots,N,$$

thus all assumptions of Lemma 4.2.2 and Theorem 4.2.1 are fulfilled. From Lemma 4.2.2 we conclude the convergence of the series (4.3.5) in $\mathcal{S}'(\mathbb{R}^d)$. We denote its limit by f. Observe, that $\lambda \in s^{\overline{r}}_{p,q}f$ implies $\big(2^{\overline{k}\cdot\overline{d}/2}\lambda_{k,\gamma}\big)_{k,\gamma} \in s^{\overline{r}}_{p,q}f^*$ with equal norm. Moreover, Theorem 4.2.1 yields $f \in S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})$ and the estimate (4.3.6). Thus the assertions (a) and (b) are already proven. Analog arguments apply to the B-case.
Now let at first $p < \infty$. For every sequence $\lambda \in s^{\overline{r}}_{p,q}a$ and a natural number μ we define sequences λ^μ by

$$\lambda^\mu_{\overline{k},\gamma} = \begin{cases} \lambda_{\overline{k},\gamma} & \text{if} \quad |\overline{k}| > \mu\,, \\ 0 & \text{else}\,. \end{cases}$$

In case $q < \infty$, we find for these sequences

$$\lim_{\mu\to\infty} \big\| \lambda^\mu \,|s^{\overline{r}}_{p,q}a\big\| = 0\,.$$

This is immediately clear for b-spaces, the f-case follows from Lebesgue's theorem on dominated convergence. Moreover, the sequences $\lambda^\mu_{\overline{k}}$, defined by

$$\lambda^\mu_{\overline{k},(\overline{i},m)} = \begin{cases} \lambda_{\overline{k},(\overline{i},m)} & \text{if} \quad |m| > \mu\,, \\ 0 & \text{else}\,, \end{cases}$$

converge to 0 in $s^{\overline{r}}_{p,q}a$ (or in ℓ_p, respectively) for $\mu \longrightarrow \infty$ for every $\overline{k}$. Together with the previously proven estimate (4.3.6) this yields the convergence of (4.3.4) in $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$.
With the same argument as before together with an additional application of Hölder's inequality we obtain (also in the case $q = \infty$)

$$\lim_{\mu\to\infty} \big\| \lambda^\mu \,|s^{\overline{r}-\varepsilon\overline{1}}_{p,q}a\big\| = 0\,.$$

As before, this implies the convergence of (4.3.4) in $S^{\overline{r}-\varepsilon\overline{1}}_{p,q}A(\mathbb{R}^{\overline{d}})$. The statements concerning the unconditional convergence of the series (4.3.4) follows now directly from the unconditional convergence of the series

$$\lambda = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\Gamma_{\overline{k}}} \lambda_{\overline{k},\gamma} e^{\overline{k},\gamma}\,,$$

where the sequences $e^{\overline{k},\gamma}$ are the canonical orthonormal basis vectors of $s^{\overline{0}}_{2,2}b$. This in turn follows from the convergence of $(\lambda^\mu)_{\mu\in\mathbb{N}}$ and $(\lambda^\mu_{\overline{k}})_{\mu\in\mathbb{N}}$ as stated above. We remind of the fact, that ℓ_u-quasi-norms, $0 < u < \infty$, always form unconditional convergent series. □

Unfortunately the above argumentation fails in case $p = \infty$. In this case there is only local convergence, i.e. on given balls or bounded domains, which corresponds to a restriction of the γ-summation (compare with Section 7.1.1).

Proof of Theorem 4.3.1(ii). The meaning of the expressions $\big(f, \Psi_{\overline{\nu},\gamma}\big)$ was discussed before in Remark 4.3.3. In this proof, we treat only the F-case, the proof for B-spaces is once more very similar.
As before, we can rewrite the quasi-norm in $s^{\overline{r}}_{p,q}f$ as

$$\big\| \lambda \,|s^{\overline{r}}_{p,q}f\big\| = \big\|2^{\overline{k}\cdot(\overline{r}+\overline{d}/2)} g_{\overline{k},\overline{i}}\big|L_p(\ell_q)\big\| ,$$

where

$$g_{\overline{k},\overline{i}}(x) = \sum_{m\in\mathbb{Z}^d} \lambda_{\overline{k},(\overline{i},m)} \mathcal{X}_{\overline{k},m}(x) . \tag{4.3.8}$$

Thereby, the ℓ_q-quasi-norm is formed with respect to $\overline{k}$ and $\overline{i}$. If $x \in Q_{\overline{k},m}$, and λ is defined by (4.3.7), we use (4.3.8) to obtain

$$g_{\overline{k},\overline{i}}(x) = \lambda_{\overline{k},(\overline{i},m)} = \int_{\mathbb{R}^d} \Psi_{\overline{k},(\overline{i},m)}(y) f(y) dy = \int_{\mathbb{R}^d} \psi^{i_1,1}_{k_1,m^1}(y^1)\cdots\psi^{i_N,N}_{k_N,m^N}(y^N) f(y) dy .$$

We now insert the definitions (4.3.1) and (1.2.8) and substitute $z^i = y^i - 2^{-k_i}m^i$. This results in

$$\begin{aligned} g_{\overline{k},\overline{i}}(x) &= \int_{\mathbb{R}^d} \psi^{i_1,1}(2^{k_1}z^1)\cdots\psi^{i_N,N}(2^{k_N}z^N) f(2^{-k_1}m^1 + z^1, \ldots, 2^{-k_n}m^N + z^N) dz \\ &= \mathcal{K}_{\overline{k},\overline{i}}(f)(2^{-\overline{k}}m) . \end{aligned}$$

Here $\mathcal{K}_{\overline{k},\overline{i}}(f)(2^{-\overline{k}}m)$ denote the local means

$$\mathcal{K}_{\overline{k},\overline{i}}(f)(y) = \int_{\mathbb{R}^d} \mathcal{K}_{\overline{k},\overline{i}}(z) f(y+z) dz , \qquad y \in \mathbb{R}^d ,$$

with respect to the kernels

$$\mathcal{K}_{\overline{k},\overline{i}}(z) = \psi^{i_1,1}(2^{k_1}z^1)\cdots\psi^{i_N,N}(2^{k_N}z^N) .$$

We remark that all integrals have to be understood in the distributional sense. Thus it follows for $x \in Q_{\overline{k},m}$ (observe $x \in Q_{\overline{k},m} \iff x - 2^{-\overline{k}}m \in Q_{\overline{k},0}$)

$$|g_{\overline{k},\overline{i}}(x)| \le \sup_{x-y\in Q_{\overline{k},0}} |\mathcal{K}_{\overline{k},\overline{i}}(f)(y)| .$$

If $\overline{s}$ is large enough, then the kernels satisfy the assumptions of Proposition 3.3.1. The moment conditions follow from property (1.2.5) in Theorem 1.2.1, and the Tauberian conditions are due to general results about scaling functions (in particular $\mathcal{F}\psi_0 = \frac{1}{2\pi} \neq 0$) and the compact support of ψ_1. Though $\mathcal{F}\psi_1(0) = 0$, the Tauberian condition can be fulfilled since this function is analytic by the Paley-Wiener-Schwartz theorem. The compactness of the supports further implies, that the kernels belong to $X^{\overline{s}}(\mathbb{R}^{\overline{d}})$. Now an application of Proposition 3.3.1 or Lemma 3.3.1, respectively, results in

$$\big\| \lambda \,|s^{\overline{r}}_{p,q}f\big\| \le c \sum_{\overline{i}\in I_{\overline{d}}} \big\|2^{\overline{k}\cdot\overline{r}} g_{\overline{k},\overline{i}}\big|L_p(\ell_q)\big\| \le C \,\big\| f \,|S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})\big\| ,$$

where only those $\overline{k} \in \mathbb{N}_0^N$ with $k_j = 0$ if $i_j = 0$, $j = 1, \ldots, N$, are considered. In particular, for $\overline{i} = \overline{0}$ there is in fact no ℓ_q-summation, and the estimate for $g_{\overline{0},\overline{0}}$ follows already from the interpretation of $\Psi_{\overline{0},\overline{0},m}$ as linear functionals in Remark 4.3.3. This completes the proof of (a).
For the proof of the second statement we define a new distribution g by

$$g = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\Gamma_{\overline{k}}} \lambda_{\overline{k},\gamma} \Psi_{\overline{k},\gamma}\,, \tag{4.3.9}$$

where the coefficients $\lambda_{\overline{k},\gamma}$ are given by (4.3.7). The convergence of this series is assured by $\lambda \in s^{\overline{r}}_{p,q}f$ (which we just proved) and part (i) of the theorem. This even shows $g \in S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}})$. Hence we have to show $g = f$, or equivalently

$$(g, \varphi) = (f, \varphi) \quad \text{for all} \quad \varphi \in \mathcal{S}(\mathbb{R}^d)\,.$$

To this purpose we consider the expressions $\big(g, \Psi_{\overline{k'},\gamma'}\big)$. Since $\lambda \in s^{\overline{r}}_{p,q}f$ the series (4.3.9) converges in $S^{\overline{r}-\varepsilon\overline{1}}_{p,q}F(\mathbb{R}^{\overline{d}})$, where $\varepsilon > 0$ is arbitrary. By Remark 4.3.3 we have $\Psi_{\overline{k'},\gamma'} \in \big[S^{\overline{r}-\varepsilon\overline{1}}_{p,q}F(\mathbb{R}^{\overline{d}})\big]'$. Hence it follows

$$\begin{aligned}\big(g, \Psi_{\overline{k'},\gamma'}\big) &= \lim_{\mu\to\infty} \Bigg(\sum_{|\overline{k}|\le\mu}\ \sum_{\gamma=(\overline{i},m)\in\Gamma_{\overline{k}}:\,|m|\le\mu} \lambda_{\overline{k},\gamma}\Psi_{\overline{k},\gamma}\ ,\ \Psi_{\overline{k'},\gamma'} \Bigg) \\ &= \lim_{\mu\to\infty} \sum_{|\overline{k}|\le\mu}\ \sum_{\gamma=(\overline{i},m)\in\Gamma_{\overline{k}}:\,|m|\le\mu} \big(f, \Psi_{\overline{k},\gamma}\big)\big(\Psi_{\overline{k},\gamma}, \Psi_{\overline{k'},\gamma'}\big)\,.\end{aligned}$$

Using the orthonormality of the system (4.3.3) we obtain from this

$$\big(g, \Psi_{\overline{k'},\gamma'}\big) = \big(f, \Psi_{\overline{k'},\gamma'}\big) \quad \text{for all} \quad \overline{k} \in \mathbb{N}_0^N, \gamma \in \Gamma_{\overline{k}}\,.$$

This argument can be extended to arbitrary linear combinations of functions $\Psi_{\overline{k},\gamma}$. For a general function $\varphi \in \mathcal{S}(\mathbb{R}^d) \hookrightarrow L_2(\mathbb{R}^d)$ we consider its Fourier series expansion with respect to the system (4.3.3):

$$\varphi = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\Gamma_{\overline{k}}} \big(\varphi, \Psi_{\overline{k},\gamma}\big)\Psi_{\overline{k},\gamma}\,. \tag{4.3.10}$$

Since $\mathcal{S}(\mathbb{R}^d)$ is a subset of all (Fourier-analytic) Besov and Triebel-Lizorkin spaces, it follows from the previously proven assertions (ii.a) and (i.c), that (4.3.10) converges in $\big[S^{\overline{r}-\varepsilon\overline{1}}_{p,q}F(\mathbb{R}^{\overline{d}})\big]'$ as well. Thus we find

$$\begin{aligned}\big(g, \varphi\big) &= \lim_{\mu\to\infty} \sum_{|\overline{k}|\le\mu}\ \sum_{\gamma=(\overline{i},m)\in\Gamma_{\overline{k}}:\,|m|\le\mu} \big(\varphi, \Psi_{\overline{k},\gamma}\big)\big(g, \Psi_{\overline{k},\gamma}\big) \\ &= \lim_{\mu\to\infty} \sum_{|\overline{k}|\le\mu}\ \sum_{\gamma=(\overline{i},m)\in\Gamma_{\overline{k}}:\,|m|\le\mu} \big(\varphi, \Psi_{\overline{k},\gamma}\big)\big(f, \Psi_{\overline{k},\gamma}\big) = \big(f, \varphi\big)\,.\end{aligned}$$

This shows that the series (4.3.5) converges to f.

The last step, that is the proof of the third statement, follows now quite easily. We assume that η satisfies the named assumptions. Furthermore, we define the coefficients $\lambda_{\overline{k},\gamma}$ once more by (4.3.7). Then the same duality arguments as before and (i.c) yield

$$\lambda_{\overline{k},\gamma} = \big(f, \Psi_{\overline{k},\gamma}\big) = \Bigg(\sum_{\overline{k'}\in\mathbb{N}_0^N} \sum_{\gamma'\in\Gamma_{\overline{k'}}} \eta_{\overline{k'},\gamma'} \Psi_{\overline{k'},\gamma'} \ , \ \Psi_{\overline{k},\gamma}\Bigg) = \eta_{\overline{k},\gamma}$$

for all $\overline{k} \in \mathbb{N}_0^N$ and $\gamma \in \Gamma_{\overline{k}}$. The proof of Theorem 4.3.1 is now complete. □

Corollary 4.3.1. Let $\overline{r}, p, q, \overline{s}$ be as in Theorem 4.3.1. Then the mapping

$$J : S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \longrightarrow s^{\overline{r}}_{p,q}a \,, \qquad f \longmapsto \Big(\big(f, \Psi_{\overline{k},\gamma}\big)\Big)_{\overline{k}\in\mathbb{N}_0^N,\gamma\in\Gamma_{\overline{k}}} \,,$$

is an isomorphism from the function space $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ onto the sequence space $s^{\overline{r}}_{p,q}a$. In particular, it holds

$$\big\| \, f \,\big|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \sim \big\|\big((f, \Psi_{\overline{k},\gamma})\big)_{\overline{k},\gamma}\big|s^{\overline{r}}_{p,q}a\big\| \,.$$

Proof. That the mapping J is a bounded operator follows from Theorem 4.3.1(ii.a). Likewise the operator J^{-1}, given by

$$J^{-1} \ : \ s^{\overline{r}}_{p,q}a \longrightarrow S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \,, \qquad \lambda \longmapsto \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\Gamma_{\overline{k}}} \lambda_{\overline{k},\gamma} \Psi_{\overline{k},\gamma} \,,$$

is bounded due to (i.b). Finally, the fact that J^{-1} is indeed the inverse operator of J follows from (ii.b) and (ii.c). □

4.4 Tensor products of Sobolev and Besov spaces

In the previous sections tensor product constructions were of exceptional importance, in particular the constructions of local means and wavelets. We also remind on Proposition 1.4.1 and Theorem 1.4.2, where the spaces $S^r_pH(\mathbb{R}^d)$ and $S^r_{p,p}B(\mathbb{R}^d)$ were identified as tensor product spaces. This section now is devoted to the study of the respective counterparts for the spaces $S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})$ and $S^{\overline{r}}_{p,p}B(\mathbb{R}^{\overline{d}})$. This complements the results from the proof of Lemma 3.2.2.
In Remark 2.1.2 we already mentioned, that the norms in $S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})$ and $S^{\overline{r}}_pH(\mathbb{R}^{\overline{d}})$ are crossnorms. The next proposition is the counterpart of Proposition 1.4.1. Its proof is an immediate corollary of [74, Proposition 3.1] as well.

Proposition 4.4.1. Let $N \geq 2$, $1 < p < \infty$, and let $\overline{r} = (r_1, \dots, r_N) \in \mathbb{R}^N$. Then it holds

$$\begin{aligned} S^{\overline{r}}_pH(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) &= H^{r_1}_p(\mathbb{R}^{d_1}) \otimes_{\alpha_p} S^{(r_2,\dots,r_N)}_p H(\mathbb{R}^{d_2} \times \cdots \times \mathbb{R}^{d_N}) \\ &= S^{(r_1,\dots,r_{N-1})}_p H(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{N-1}}) \otimes_{\alpha_p} H^{r_N}_p(\mathbb{R}^{d_N}) \\ &= H^{r_1}_p(\mathbb{R}^{d_1}) \otimes_{\alpha_p} \cdots \otimes_{\alpha_p} H^{r_N}_p(\mathbb{R}^{d_N}) \end{aligned}$$

with coinciding norms. Similarly it holds for $\overline{m} \in \mathbb{N}_0^N$

$$S^{\overline{m}}_pW(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = W^{m_1}_p(\mathbb{R}^{d_1}) \otimes_{\alpha_p} \cdots \otimes_{\alpha_p} W^{m_N}_p(\mathbb{R}^{d_N})$$

in the sense of equivalent norms.

Next we turn to tensor products of Besov spaces. To this purpose we consider again tensor products of the corresponding sequence spaces first.

Lemma 4.4.1. Let $N \geq 2$, $r_1, \ldots, r_N \in \mathbb{R}$ and let $0 < p < \infty$. Then it holds

$$s^{r_1,\ldots,r_N}_{p,p}b = b^{r_1}_{p,p} \otimes_{\delta_p} s^{r_2,\ldots,r_N}_{p,p}b = s^{r_1,\ldots,r_{N-1}}_{p,p}b \otimes_{\delta_p} b^{r_N}_{p,p} = b^{r_1}_{p,p} \otimes_{\delta_p} \cdots \otimes_{\delta_p} b^{r_N}_{p,p}$$

with coinciding (quasi-)norms. The sequence spaces $b^{r_i}_{p,p}$ are those from Definition 1.2.4, where $n \mathrel{\widehat{=}} d_i$, $i = 1, \ldots, N$.

This lemma is an immediate corollary of Proposition 1.3.3 for suitably chosen weight sequences. Now let $J_i : B^{r_i}_{p,p}(\mathbb{R}^{d_i}) \longrightarrow b^{r_i}_{p,p}$, $i = 1, \ldots, N$, be wavelet isomorphisms as in Theorem 1.2.2(iii). Then Lemmas 1.3.2 and 1.3.6 yield that $J^N = J_1 \otimes \cdots \otimes J_N$ is an isomorphism from $b^{r_1}_{p,p} \otimes_{\delta_p} \cdots \otimes_{\delta_p} b^{r_N}_{p,p}$ onto $B^{r_1}_{p,p}(\mathbb{R}^{d_1}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B^{r_N}_{p,p}(\mathbb{R}^{d_N})$. Moreover, for dyads this isomorphism obviously coincides with the isomorphism J from Corollary 4.3.1. Hence after linear and (uniquely determined) continuous extension we find $J^N = J$. Now we conclude from Lemma 4.4.1 the following theorem.

Theorem 4.4.1. Let $N \geq 2$, $\overline{r} = (r_1, \ldots, r_N) \in \mathbb{R}^N$, and let $0 < p < \infty$. Then it holds

$$S^{\overline{r}}_{p,p}B(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_N}) = B^{r_1}_{p,p}(\mathbb{R}^{d_1}) \otimes_{\delta_p} \cdots \otimes_{\delta_p} B^{r_N}_{p,p}(\mathbb{R}^{d_N})$$

in the sense of equivalent quasi-norms.

5 Embeddings of sequence spaces

In this chapter we shall investigate many of the basic properties of the sequence spaces $s^{\overline{r}}_{p,q}a$. A particular focus lies on embeddings. Besides directly calculating these embeddings, it is possible to transfer embedding results for the corresponding function spaces to the sequence spaces using the wavelet isomorphism just proven. While for a certain number of results this could be done (since we had to prove some of the embeddings to obtain the wavelet characterization) we intend to use the wavelet isomorphism the other way round: We prove embeddings (and properties thereof) on sequence space level, and translate them afterwards into statements for function spaces.

5.1 Preliminary remarks

For our subsequent considerations we need another slight modification of the sequence spaces introduced in Definition 4.1.1. Moreover, we shall discuss several conventions which can be transformed easily into each other, and eventually fix our choice.
As mentioned in Section 4.3, the exact form of the inner summation is necessary only when considering both function and sequence spaces in connection with the associated isomorphism. Otherwise, one can always reduce the sequence of wavelet coefficients by defining $\lambda_{\overline{\nu},m} := \sum_{\overline{i}\,:\,(\overline{i},m)\in\nabla_{\overline{\nu}}} |\lambda_{\overline{\nu},(\overline{i},m)}|$. This reduction leads to equivalent (quasi-)norms. Moreover, the approximation results from Section 6 can be transferred immediately to the full sequence, resulting only in additional constants.
We explained in Remark 4.3.2 that the exact form of the weight factors corresponds to the normalization of the functions used in the decomposition. Besides L_∞- and L_2-normalized basis functions some applications work with L_p-normalization. In our calculations the version as for atoms will be most convenient. However, all these normalizations can be transformed into each other by simple lifting arguments (see Section 5.2, Proposition 5.2.1).

One last modification will be helpful. When discussing function spaces, and in particular (local) means, the abbreviation $Q_{\overline{\nu},m} = 2^{-\overline{\nu}-\overline{1}}[-1,1]^d + 2^{-\overline{\nu}}m$ is well adapted. On the other hand, when concentrating on sequence spaces (as we will do in the sequel) $\widetilde{Q}_{\overline{\nu},m} = 2^{-\overline{\nu}}([0,1]^d + m)$ is better suited. The meaning of $\mathcal{X}_{\overline{\nu},m}$ changes accordingly. However, both variants lead to equivalent quasi-norms.
To see this, we observe

$$[-1,1]^d = \bigcup_{e\in\{0,1\}^d} \big([0,1]^d - e\big) \quad \text{and} \quad [0,1]^d = \bigcup_{e\in\{\frac{1}{4},\frac{3}{4}\}^d} \big([-\tfrac{1}{4},\tfrac{1}{4}]^d + e\big)\,,$$

and consequently for all $\overline{\nu}\in\mathbb{N}_0^N$ and $m\in\mathbb{Z}^d$

$$Q_{\overline{\nu},m} = \bigcup_{e\in\{0,1\}^d} \widetilde{Q}_{\overline{\nu}+\overline{1}\,,\,m-e} \quad \text{and} \quad \widetilde{Q}_{\overline{\nu},m} = \bigcup_{e\in\{1,3\}^d} 2\,Q_{\overline{\nu}+\overline{2}\,,\,m+e}\,.$$

The equivalence of the quasi-norms now follows directly from the triangle inequality, the observations $\mathcal{X}_{\overline{\nu},m}(2\,\cdot) = \mathcal{X}_{\overline{\nu}+\overline{1}\,,\,m}$ and $\widetilde{\mathcal{X}}_{\overline{\nu},m}(2\,\cdot) = \widetilde{\mathcal{X}}_{\overline{\nu}+\overline{1}\,,\,m}$, the estimates $\mathcal{X}_{2Q_{\overline{\nu},m+e}} \le c_e\big(\overline{M}\mathcal{X}^\omega_{\overline{\nu},m}\big)^{1/\omega}$ and $\widetilde{\mathcal{X}}_{\overline{\nu},m-e} \le c'_e\big(\overline{M}\widetilde{\mathcal{X}}^\omega_{\overline{\nu},m}\big)^{1/\omega}$ for some fixed $0<\omega<\min(1,p,q)$, and eventually the maximal inequality in $L_{p/\omega}(\ell_{q/\omega})$ (Proposition 2.3.1; we remind on $p<\infty$ for f-spaces).

Definition 5.1.1. Let $0 < p, q \leq \infty$ and $\overline{r} \in \mathbb{R}^N$. Moreover, let $\nabla = (\nabla_{\overline{k}})_{\overline{k} \in \mathbb{N}_0^N}$ be a sequence of subspaces of $\mathbb{Z}^d$. For sequences

$$\lambda = \left\{ \lambda_{\overline{\nu},m} \in \mathbb{C} \, : \, \overline{\nu} \in \mathbb{N}_0^N \, , m \in \nabla_{\overline{\nu}} \right\} \tag{5.1.1}$$

we define

$$s^{\overline{r}}_{p,q}b(\nabla) := \left\{ \lambda \, : \, \left\| \lambda \left| s^{\overline{r}}_{p,q}b(\nabla) \right\| < \infty \right\},$$

$$\left\| \lambda \left| s^{\overline{r}}_{p,q}b(\nabla) \right\| := \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} 2^{\overline{\nu} \cdot (\overline{r} - \overline{d}/p) q} \left(\sum_{m \in \nabla_{\overline{\nu}}} |\lambda_{\overline{\nu},m}|^p \right)^{q/p} \right)^{1/q} .$$

Furthermore, in case $p < \infty$ we put

$$s^{\overline{r}}_{p,q}f(\nabla) := \left\{ \lambda \, : \, \left\| \lambda \left| s^{\overline{r}}_{p,q}f(\nabla) \right\| < \infty \right\},$$

$$\left\| \lambda \left| s^{\overline{r}}_{p,q}f(\nabla) \right\| := \left\| \left(\sum_{\overline{\nu} \in \mathbb{N}_0^N} \sum_{m \in \nabla_{\overline{\nu}}} |2^{\overline{\nu} \cdot \overline{r}} \lambda_{\overline{\nu},m} \widetilde{\mathcal{X}}_{\overline{\nu},m}(\,\cdot\,)|^q \right)^{1/q} \right| L_p(\mathbb{R}^d) \right\| .$$

In case p and/or q are infinite, one uses the usual modifications.

In the sequel we will drop the tilde and write again $\mathcal{X}_{\overline{\nu},m}$. As usual, we will also use the notation $s^{\overline{r}}_{p,q}a(\nabla)$. In case $\nabla_{\overline{\nu}} = \mathbb{Z}^d$ for all $\overline{\nu} \in \mathbb{N}_0^N$ we will simply write $s^{\overline{r}}_{p,q}a$.

In addition to it, another special case of sequences ∇ is of great importance. For that case, we will use the notation $s^{\overline{r}}_{p,q}a(\Omega)$, where the sequence $\nabla = \nabla(\Omega) = (\nabla_{\overline{\nu}})_{\overline{\nu} \in \mathbb{N}_0^N}$ is defined by

$$\nabla_{\overline{\nu}} = \left\{ m \in \mathbb{Z}^d \, : \, Q_{\overline{\nu},m} \cap \Omega \neq \emptyset \right\} . \tag{5.1.2}$$

Here Ω is an arbitrary open subset of $\mathbb{R}^d$. Later on, we will concentrate on bounded domains. In case Ω is the Cartesian product of subsets of $\mathbb{R}^{d_i}$, then also ∇ possesses an according product structure. One particular example for such domains will be $\Omega = [0,1]^d$. At first, the case $\Omega = \mathbb{R}^d$ is allowed. Clearly, we have $s^{\overline{r}}_{p,q}a(\mathbb{R}^d) = s^{\overline{r}}_{p,q}a$.
The following lemma presents a significant property of these sequences $\nabla = \nabla(\Omega)$ for bounded domains Ω.

Lemma 5.1.1. Let Ω be a bounded open subset of $\mathbb{R}^d$, and define $\nabla = \left(\nabla_{\overline{\nu}}\right)_{\overline{\nu} \in \mathbb{N}_0^N}$ as in (5.1.2). Then there exist constants $C_1, C_2 > 0$, such that it holds

$$C_1 \, 2^{\overline{\nu} \cdot \overline{d}} \leq \# \nabla_{\overline{\nu}} \leq C_2 \, 2^{\overline{\nu} \cdot \overline{d}} \quad \text{for all} \quad \overline{\nu} \in \mathbb{N}_0^N . \tag{5.1.3}$$

Remark 5.1.1. The proof is obvious, we only mention that due to the boundedness there exists a cube $\Gamma_2 \supset \Omega$, and the openness implies the existence of another cube $\Gamma_1 \subset \Omega$. More precisely, these cubes can be chosen such that

$$\Gamma_1 \subset \bigcup_{m \in \nabla_{\overline{\nu}}} Q_{\overline{\nu},m} \subset \Gamma_2 \qquad \text{for all} \quad \overline{\nu} \in \mathbb{N}_0^N . \tag{5.1.4}$$

From this, the estimates of the count of the index set $\nabla_{\overline{\nu}}$ follows easily.

Additionally, with the help of appropriate shifting and rescaling we can always achieve $[0,1]^d \subset \Omega$. This will be a convenient assumption in several proofs.

In the upcoming calculations it will often be necessary to interchange summations and integrations. This can always be justified by the following theorem on monotone convergence, since all occurring functions are non-negative. In most instances we will apply the theorem without further mentioning it.

Theorem 5.1.1. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions, $f_n : \Omega \longrightarrow [0,\infty]$, defined on an arbitrary measure space $(\Omega, \mathcal{A}, \mu)$. Then it holds

$$\sum_{n=1}^{\infty} \int_\Omega f_n d\mu = \lim_{N\to\infty} \sum_{n=1}^{N} \int_\Omega f_n d\mu = \int_\Omega \lim_{N\longrightarrow\infty} \sum_{n=1}^{N} f_n d\mu = \int_\Omega \sum_{n=1}^{\infty} f_n d\mu \,.$$

In particular, limits $+\infty$ are admissible.

Remark 5.1.2. Further properties of the sequence spaces $s^{\overline{r}}_{p,q}a(\nabla)$ will be proven over the course of the next sections. At this point we shall only mention, that clearly these spaces are quasi-Banach spaces, which can be shown as for weighted vector-valued ℓ_p- and L_p-spaces. Moreover, using Minkowski's inequality (Theorem 2.3.2) we find that they are u-Banach spaces (see Definition 1.3.2), where $u = \min(1, p, q)$.

Remark 5.1.3. As for the function spaces we did not define sequence spaces $s^{\overline{r}}_{\infty,q}f(\nabla)$. Of course, the definition can be extended to that effect, on the other hand we no longer had the wavelet isomorphism for these spaces. As mentioned, our aim is to study embeddings and approximative properties for sequence spaces, as the arguments and calculations are considerably simpler. Afterwards these results shall be transferred to the function spaces. Hence, while many proofs could be carried over to that case we will not amplify this fact. Moreover, the isotropic case suggests that the extended definition would not yield the correct spaces $s^{\overline{r}}_{\infty,q}f$ for $q < \infty$, see the famous article of Frazier and Jawerth [30].

5.2 Lifting operator for sequence spaces

In Section 2.3.6 we defined an operator $I_{\overline{\rho}}$, which turned out to be an isomorphism from $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ onto $S^{\overline{r}-\overline{\rho}}_{p,q}A(\mathbb{R}^{\overline{d}})$, see Proposition 2.3.11. Of even greater importance for our purposes will be the analogue for the sequence spaces $s^{\overline{r}}_{p,q}a(\nabla)$, which shall be considered next.

Definition 5.2.1. Let $\overline{s} \in \mathbb{R}^N$. Then we define the lifting operator $L_{\overline{s}}$ for sequences λ of the form (5.1.1) by

$$\big(L_{\overline{s}}\lambda\big)_{\overline{\nu},m} := 2^{\overline{\nu}\cdot\overline{s}}\lambda_{\overline{\nu},m}\,, \qquad \overline{\nu} \in \mathbb{N}_0^N\,, m \in \nabla_{\overline{\nu}}\,. \tag{5.2.1}$$

The mapping properties of this operator are the content of the next proposition. Its proof is obvious from the definition of the components of the lifted sequence in (5.2.1).

Proposition 5.2.1. Let $0 < p, q \le \infty$ ($p < \infty$ for f-spaces), $\overline{r}, \overline{s} \in \mathbb{R}^N$, and let ∇ be a sequence of subsets of $\mathbb{Z}^d$. Then the operator $L_{\overline{s}}$ is an isometric isomorphism from $s^{\overline{r}}_{p,q}a(\nabla)$ onto $s^{\overline{r}-\overline{s}}_{p,q}a(\nabla)$.

We only add that clearly we have $L_{\overline{r}}^{-1} = L_{-\overline{r}}$. The following corollary yields considerable simplifications for investigations of embeddings.

Corollary 5.2.1. Let $0 < p_0, p_1, q_0, q_1 \leq \infty$ ($p_0, p_1 < \infty$ for f-spaces) and $\overline{r}, \overline{s}, \overline{t} \in \mathbb{R}^N$. Then the continuous embedding $s_{p,q}^{\overline{r}+\overline{t}} a(\nabla) \hookrightarrow s_{p,q}^{\overline{r}} a^\dagger(\nabla)$ holds if, and only if, it holds $s_{p,q}^{\overline{s}+\overline{t}} a(\nabla) \hookrightarrow s_{p,q}^{\overline{s}} a^\dagger(\nabla)$, where $a, a^\dagger \in \{b, f\}$.

In other words, whenever one investigates the continuity or compactness of embeddings and several related properties (like approximation by linear methods), rather than the smoothness vectors themselves only their difference is a relevant parameter. This allows us to choose $\overline{r} = \overline{0}$ without loss of generality, i.e. we may restrict our investigations to embeddings $s_{p,q}^{\overline{t}} a(\nabla) \hookrightarrow s_{p,q}^{\overline{0}} a^\dagger(\nabla)$.

5.3 Continuous embeddings

In connection with sequence space defined over some index set I, the notation e^i, $i \in I$, is reserved for the corresponding canonical unit sequences, i.e. for the sequences whose components are given by Kronecker symbols, $(e^i)_j = \delta_{i,j}$, $i, j \in I$. Accordingly, the notations $\overline{e}^{\,i}$ and e^j are used for the canonical unit vectors in $\mathbb{R}^N$ and $\mathbb{R}^d$, respectively.

5.3.1 Necessary conditions

Due to the next lemma we may restrict our considerations to sufficient conditions for embeddings. The sequence spaces $a_{p,q}^s(\nabla)$ occurring in the formulation of the lemma are variants of the spaces $a_{p,q}^s$ from Definition 1.2.4, modified according to the discussion in the previous section.

Lemma 5.3.1. Let $0 < p, p_0, p_1, q, q_0, q_1 \leq \infty$ ($p, p_0, p_1 < \infty$ for f-spaces) and $\overline{r}, \overline{s} \in \mathbb{R}^N$.

(i) Assume the sequence ∇ to have product structure, i.e. let sequences $\nabla^i = (\nabla_j^i)_{j=0}^\infty$, $i = 1, \ldots, N$, of subsets of $\mathbb{Z}^{d_i}$ be given, and define for $\overline{\nu} \in \mathbb{N}_0^N$ sets $\nabla_{\overline{\nu}} = \nabla_{\nu_1}^1 \times \nabla_{\nu_N}^N \subset \mathbb{Z}^d$. Then the quasi-norms of the spaces $s_{p,q}^{\overline{r}} a(\nabla)$ are crossnorms. More precisely, if $\lambda^i \in a_{p,q}^{r_i}(\nabla^i)$, then $\lambda = \lambda^1 \otimes \cdots \otimes \lambda^N \in s_{p,q}^{\overline{r}} a(\nabla)$, and it holds

$$\left\| \lambda \,|\, s_{p,q}^{\overline{r}} a(\nabla) \right\| = \left\| \lambda^1 \,|\, a_{p,q}^{r_1}(\nabla^1) \right\| \cdots \left\| \lambda^N \,|\, a_{p,q}^{r_N}(\nabla^N) \right\| .$$

(ii) If the embedding $s_{p_0,q_0}^{\overline{r}} x(\nabla) \hookrightarrow s_{p_1,q_1}^{\overline{s}} y(\nabla)$ is continuous, then for appropriate sequences $\widetilde{\nabla}^i$ the embeddings $x_{p_0,q_0}^{r_i}(\widetilde{\nabla}^i) \hookrightarrow y_{p_1,q_1}^{s_i}(\widetilde{\nabla}^i)$, $i = 1, \ldots, N$, are continuous as well.

Proof. Part (i) is an immediate consequence of Fubini's theorem and the factorization $2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)} = 2^{\nu_1(r_1-d_1/p)} \cdots 2^{\nu_N(r_N-d_N/p)}$ for the weight factor.
For sequences ∇ with product structure part (ii) follows at once from the crossnorm-property (tensorize sequences $\eta \in a_{p,q}^{r_i}(\nabla^i)$ with fixed sequences $\lambda^j \in a_{p,q}^{r_j}(\nabla^j)$, $j \neq i$). In case of arbitrary sequences ∇ we identify the spaces $a_{p,q}^{r_i}(\widetilde{\nabla}^i)$ with subspaces of $s_{p,q}^{\overline{r}} a(\nabla)$ with coinciding quasi-norms.

We fix some $i \in \{1,\dots,N\}$ and put $\widetilde{\nabla}_j^i := \{k \in \mathbb{Z}^{d_i} : k = m^i \text{ for some } m \in \nabla_{j\bar{e}_i}\}$. In case $\nabla = \nabla(\Omega)$, these sets fulfil $C_1 2^{jd_i} \le \#\widetilde{\nabla}_j^i \le C_2 2^{jd_i}$, $j \in \mathbb{N}_0$, similar to (5.1.3). This follows as in Lemma 5.1.1 from (5.1.4). Now given a sequence $\lambda \in a^{r_i}_{p,q}(\widetilde{\nabla}^i)$ we can associate a sequence $\sigma_i\lambda \in s^{\overline{r}}_{p,q}a(\nabla)$ by defining

$$\big(\sigma_i\lambda\big)_{\overline{\nu},m} = \begin{cases} \lambda_{j,k}\,, & \overline{\nu} = j\bar{e}_i\,,\ m = m(k)\,,\ j \in \mathbb{N}_0\,,\ k \in \widetilde{\nabla}_j^i\,, \\ 0, & \text{else}\,, \end{cases}$$

where $m(k) \in \nabla_{j\bar{e}_i}$ is chosen such that $\big(m(k)\big)^i = k$. If follows at once $\|\,\sigma_i\lambda\,|s^{\overline{r}}_{p,q}a(\nabla)\| = \|\,\lambda\,|a^{r_i}_{p,q}(\widetilde{\nabla}^i)\|$. Vice versa, to every given sequence $\eta \in s^{\overline{r}}_{p,q}a(\nabla)$ we can associate a sequence $\pi_i\eta \in a^{r_i}_{p,q}(\widetilde{\nabla}^i)$ via restriction. The mapping $\sigma_i \circ \pi_i$ is a projection which yields the mentioned subspace. □

Part (ii) of the above lemma particularly states, that the necessary conditions for embeddings in the isotropic case are necessary (componentwise) for the dominating mixed spaces as well. Since those conditions do not depend on the exact form of $\widetilde{\nabla}^i$ (if at all an estimate on $\#\widetilde{\nabla}_j^i$ is needed), the chosen formulation suffices. For an overview about embeddings in the isotropic situation we refer to [73].
In the next section we will show, that the necessary conditions obtained in this way are also sufficient. This additionally implies, that as in the isotropic case we have continuous embeddings if, and only if, it holds set theoretic inclusion.

5.3.2 Elementary embeddings

In Section 2.3.1, we introduced iterated sequence spaces $\ell_q(\ell_p)$. Using this notation we find $s^{\overline{d}/p}_{p,q}b = \ell_q(\ell_p)$. When dealing with spaces $s^{\overline{r}}_{p,q}b(\nabla)$ we shall need the following generalization.

Definition 5.3.1. Let $0 < p, q \le \infty$. Let I be an arbitrary countable index set, and let $J = (J_i)_{i\in I}$ be a family of countable index sets. Then the space $\ell_q(I, \ell_p(J))$ is defined as the collection of all sequences $\lambda = (\lambda_{i,j})_{i\in I, j\in J_i}$, such that

$$\|\,\lambda\,|\ell_q\big(I, \ell_p(J)\big)\| := \left(\sum_{i\in I}\left(\sum_{j\in J_i}|\lambda_{i,j}|^p\right)^{q/p}\right)^{1/p}$$

is finite (modification if p and/or q is infinite).

As a consequence, we find the relation $\ell_q\big(\mathbb{N}_0^N, \ell_p(\nabla)\big) = L_{\overline{r}-\overline{d}/p}\big(s^{\overline{r}}_{p,q}b(\nabla)\big)$, which will be helpful numerous times. The abbreviations $I \times J$ and $\mathbb{N}_0^N \times \nabla$ have to be understood according to $\ell_p(I, \ell_p(J)) = \ell_p(I \times J)$, i.e. $I \times J = \{(i,j) : i \in I, j \in J : i\}$. We now begin with some elementary embeddings.

Lemma 5.3.2. Let $0 < p, q \le \infty$ ($p < \infty$ for f-spaces) and $\overline{r} \in \mathbb{R}^N$. Then it holds

(i) $s^{\overline{r}}_{p,q}a(\nabla) \hookrightarrow s^{\overline{r}}_{p,q'}a(\nabla)$ for all $q \le q' \le \infty$,

(ii) $s^{\overline{\varrho}}_{p,q}a(\nabla) \hookrightarrow s^{\overline{r}}_{p,q}a(\nabla)$ for all $\overline{r} \le \overline{\varrho} \in \mathbb{R}^N$.

(iii) Let $\lambda \in s^{\overline{r}}_{p,q}b(\nabla)$, and consider $\widetilde{\lambda} = L_{\overline{r}-\overline{d}/p}\lambda$. Then it holds

$$\| \widetilde{\lambda} \,|\ell_{\max(p,q)}(\mathbb{N}_0^N \times \nabla)\| \le \| \lambda \,| s^{\overline{r}}_{p,q}b(\nabla)\| \le \| \widetilde{\lambda} \,|\ell_{\min(p,q)}(\mathbb{N}_0^N \times \nabla)\| \,.$$

Proof. The assertion (i) follows at once from the monotonicity of the ℓ_q-quasi-norms. (ii) is a trivial consequence of $2^{\overline{\nu}\cdot\overline{r}} \le 2^{\overline{\nu}\cdot\overline{\varrho}}$ for all $\overline{\nu} \in \mathbb{N}_0^N$. Finally, (iii) follows from the mentioned relation $\ell_q\big(\mathbb{N}_0^N, \ell_p(\nabla)\big) = L_{\overline{r}-\overline{d}/p}\big(s^{\overline{r}}_{p,q}b(\nabla)\big)$ and again the monotonicity of the ℓ_q-quasi-norms. □

Lemma 5.3.3. Let $0 < p < \infty$, $0 < q \le \infty$ and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$\| \lambda \,| s^{\overline{r}}_{p,\max(p,q)}b(\nabla)\| \le \| \lambda \,| s^{\overline{r}}_{p,q}f(\nabla)\| \le \| \lambda \,| s^{\overline{r}}_{p,\min(p,q)}b(\nabla)\|$$

for all sequences as in (5.1.1). Accordingly, we have the embedding

$$s^{\overline{r}}_{p,\min(p,q)}b(\nabla) \hookrightarrow s^{\overline{r}}_{p,q}f(\nabla) \hookrightarrow s^{\overline{r}}_{p,\max(p,q)}b(\nabla)\,.$$

Proof. As in Remark 4.1.2, the quasi-norms in $s^{\overline{r}}_{p,q}a(\nabla)$ can be reformulated using the functions $g_{\overline{\nu}} = 2^{\overline{\nu}\cdot\overline{r}} \sum_{m\in\nabla_{\overline{\nu}}} \lambda_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}$, $\overline{\nu} \in \mathbb{N}_0^N$. In this way we obtain

$$\| \lambda \,| s^{\overline{r}}_{p,q}b(\nabla)\| = \| \big(g_{\overline{\nu}}\big)_{\overline{\nu}} | \ell_q(L_p)\| \quad \text{and} \quad \| \lambda \,| s^{\overline{r}}_{p,q}f(\nabla)\| = \| \big(g_{\overline{\nu}}\big)_{\overline{\nu}} | L_p(\ell_q)\| \,.$$

The embeddings now follow by exactly the same arguments as in the proof of Proposition 2.3.7 (Step 2). □

Remark 5.3.1. In case $p = q$ we re-obtain the obvious identity $s^{\overline{r}}_{p,p}f(\nabla) = s^{\overline{r}}_{p,p}b(\nabla)$ which holds with equality of quasi-norms. Moreover, together with Lemma 5.3.2(i) we find $s^{\overline{r}}_{p,q}f(\nabla) \hookrightarrow s^{\overline{r}}_{p,\infty}b(\nabla)$, the counterpart of Lemma 4.2.2(i).

The embeddings in Lemma 5.3.3 are sharp in the following sense.

Lemma 5.3.4. Let $0 < p < \infty$, $0 < q_0, q, q_1 \le \infty$, and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$s^{\overline{r}}_{p,q_0}b(\Omega) \hookrightarrow s^{\overline{r}}_{p,q}f(\Omega) \hookrightarrow s^{\overline{r}}_{p,q_1}b(\Omega)$$

if, and only if, $q_0 \le \min(p,q)$ and $\max(p,q) \le q_1$.

5.3.3 Embeddings of Sobolev-type

Proposition 5.3.1. Let $0 < p_0 \le p_1 \le \infty$, $0 < q_0, q_1 \le \infty$ and $\overline{r}, \overline{t} \in \mathbb{R}^N$. Then the following assertions are equivalent:

(i) $s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\nabla) \subset s^{\overline{r}}_{p_1,q_1}b(\nabla)$

(ii) $s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q_1}b(\nabla)$

(iii) We have either

$$t_i > d_i\left(\frac{1}{p_0} - \frac{1}{p_1}\right), \quad i = 1, \dots, N\,,$$

or

$$t_i \ge d_i\left(\frac{1}{p_0} - \frac{1}{p_1}\right), \quad i = 1, \dots, N\,, \quad \text{and} \quad q_0 \le q_1\,.$$

Proof. At first, let $q_0 \le q_1$. Then due to Lemma 5.3.2(ii) it suffices to treat the case $\overline{t} = \overline{d}(\frac{1}{p_0} - \frac{1}{p_1})$. Then we obtain directly from the monotonicity of the ℓ_p-quasi-norms and the assumptions $p_0 \le p_1$ and $q_0 \le q_1$

$$\begin{aligned}\left\| a \,|s^{\overline{r}}_{p_1,q_1} b(\nabla)\right\| &= \left\| L_{\overline{r}-\overline{d}/p_1} a \,|\ell_{q_1}\left(\mathbb{N}_0^N, \ell_{p_1}(\nabla)\right)\right\| = \left\| L_{\overline{r}+\overline{t}-\overline{d}/p_0} a \,|\ell_{q_1}\left(\mathbb{N}_0^N, \ell_{p_1}(\nabla)\right)\right\| \\ &\le \left\| L_{\overline{r}+\overline{t}-\overline{d}/p_0} a \,|\ell_{q_0}\left(\mathbb{N}_0^N, \ell_{p_0}(\nabla)\right)\right\| = \left\| a \,|s^{\overline{r}+\overline{t}}_{p_0,q_0} b(\nabla)\right\| .\end{aligned}$$

This proves the embedding $s^{\overline{r}+\overline{t}}_{p_0,q_0} b(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q_1} b(\nabla)$ in this case. In case $\overline{t} - \overline{d}(\frac{1}{p_0} - \frac{1}{p_1}) > 0$ we proceed similarly. For simplicity we assume $0 < p_0 \le p_1 < \infty$, otherwise one has to use the usual modifications. To begin with we further assume $0 < q < q_1 \le \infty = q_0$. Then we obtain once more from the monotonicity of the ℓ_p-quasi-norms

$$\begin{aligned}\left\| a \,|s^{\overline{r}}_{p_1,q} b(\nabla)\right\| &\le \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}(1/p_0-1/p_1))q} 2^{\overline{\nu}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)q} \left(\sum_{m\in\nabla_{\overline{\nu}}} |a_{\overline{\nu},m}|^{p_0} \right)^{q/p_0} \right)^{1/q} \\ &\le \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}(1/p_0-1/p_1))q} \right)^{1/q} \sup_{\overline{\lambda}\in\mathbb{N}_0^N} 2^{\overline{\lambda}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} \left(\sum_{m\in\nabla_{\overline{\lambda}}} |a_{\overline{\lambda},m}|^{p_0} \right)^{1/p_0} \\ &= c \sup_{\overline{\lambda}\in\mathbb{N}_0^N} 2^{\overline{\lambda}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} \left(\sum_{m\in\nabla_{\overline{\lambda}}} |a_{\overline{\lambda},m}|^{p_0} \right)^{1/p_0} = c \left\| a \,|s^{\overline{r}+\overline{t}}_{p_0,\infty} b(\nabla)\right\| .\end{aligned}$$

The general situation follows from the monotonicity of the ℓ_q-quasi-norms,

$$s^{\overline{r}+\overline{t}}_{p_0,q_0} b(\nabla) \hookrightarrow s^{\overline{r}+\overline{t}}_{p_0,\infty} b(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q} b(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q_1} b(\nabla)$$

for arbitrary $0 < q_0, q_1 \le \infty$, $0 < q < q_1$. □

For sequences ∇ with further properties, the condition $p_0 \le p_1$ might not be necessary. In particular, the sets $\nabla_{\overline{\nu}}$ have to be finite for every $\overline{\nu} \in \mathbb{N}_0^N$, compare to the situation for $\ell_p(I)$. In case of sequence spaces $s^{\overline{r}}_{p,q} b(\Omega)$, we obtain the following result.

Proposition 5.3.2. Let Ω be a bounded open subset of $\mathbb{R}^d$. Furthermore, let $\overline{r}, \overline{t} \in \mathbb{R}^N$, $0 < p_1 < p_0 \le \infty$ and $0 < q_0, q_1 \le \infty$. Then the following assertions are equivalent:

(i) $s^{\overline{r}+\overline{t}}_{p_0,q_0} b(\Omega) \subset s^{\overline{r}}_{p_1,q_1} b(\Omega)$

(ii) $s^{\overline{r}+\overline{t}}_{p_0,q_0} b(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,q_1} b(\Omega)$

(iii) It either holds $\overline{t} > 0$, or $\overline{t} \ge 0$ and $q_0 \le q_1$.

Proof. Let at first $a \in s^{\overline{r}+\overline{t}}_{p_0,\infty} b(\Omega)$, $0 < p_1 < p_0 < \infty$ and $0 < q_1 < \infty$. Then we obtain from Hölder's inequality for $\frac{p_1}{p_0} + \frac{p_0-p_1}{p_0} = 1$ and (5.1.3)

$$\begin{aligned}\left\| a \,|s^{\overline{r}}_{p_1,q_1} b(\Omega)\right\| &= \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)q_1} \left(\sum_{m\in\nabla_{\overline{\nu}}} |a_{\overline{\nu},m}|^{p_1} \right)^{q_1/p_1} \right)^{1/q_1} \\ &\le \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)q_1} \left(\sum_{m\in\nabla_{\overline{\nu}}} |a_{\overline{\nu},m}|^{p_0} \right)^{\frac{p_1}{p_0}\cdot\frac{q_1}{p_1}} \left(\sum_{m\in\nabla_{\overline{\nu}}} 1 \right)^{\frac{p_0-p_1}{p_0}\cdot\frac{q_1}{p_1}} \right)^{1/q_1}\end{aligned}$$

$$\leq c \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)q_1} \left(\sum_{m\in\nabla_{\overline{\nu}}} |a_{\overline{\nu},m}|^{p_0} \right)^{\frac{q_1}{p_0}} 2^{\overline{\nu}\cdot\overline{d}(1/p_1-1/p_0)q_1} \right)^{1/q_1}$$

$$\leq c \left(\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_0)q_1} \left(\sum_{m\in\nabla_{\overline{\nu}}} |a_{\overline{\nu},m}|^{p_0} \right)^{\frac{q_1}{p_0}} \right)^{1/q_1} .$$

Now we obtain the embedding by exactly the same arguments as in the proof of Proposition 5.3.1, i.e. either one uses the assumption $q_0 \leq q_1$ and the monotonicity of the ℓ_p-quasi-norms, or one uses the assumption $\overline{t} > 0$ and the convergence of $\sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot\overline{t}q_1}$. In case $p_0 = \infty$ or $q_1 = \infty$ one has to use the usual modifications. □

The treatment of embeddings between f-spaces is a little more involved. We start with one more elementary result, which is the counterpart of (2.3.31) (Proposition 2.3.7) for sequence spaces.

Lemma 5.3.5. Let $0 < p < \infty$, $0 < q_0, q_1 \leq \infty$ and $\overline{r}, \overline{t} \in \mathbb{R}^N$. If $\overline{t} > 0$, then it holds

$$s^{\overline{r}+\overline{t}}_{p,q_0} f(\nabla) \hookrightarrow s^{\overline{r}}_{p,q_1} f(\nabla) .$$

Proof. Let at first $a \in s^{\overline{r}+\overline{t}}_{p,\infty} f(\nabla)$ and $q_1 < \infty$. As in Lemma 5.3.3 we use the functions $g_{\overline{\nu}} = \sum_{m\in\nabla_{\overline{\nu}}} a_{\overline{\nu},m}\mathcal{X}_{\overline{\nu},m}$, $\overline{\nu} \in \mathbb{N}_0^N$, to reformulate the $s^{\overline{r}}_{p,q} f(\nabla)$-quasi-norm. The embedding $s^{\overline{r}+\overline{t}}_{p,\infty} f(\nabla) \hookrightarrow s^{\overline{r}}_{p,q_1} f(\nabla)$ now follows by the same arguments as in the proof of the mentioned Proposition 2.3.7.
In the general case, Lemma 5.3.2(i) yields for every $0 < q < q_1 \leq \infty$

$$s^{\overline{r}+\overline{t}}_{p,q_0} f(\nabla) \hookrightarrow s^{\overline{r}+\overline{t}}_{p,\infty} f(\nabla) \hookrightarrow s^{\overline{r}}_{p,q} f(\nabla) \hookrightarrow s^{\overline{r}}_{p,q_1} f(\nabla) ,$$

which completes the proof. □

Proposition 5.3.3. Let $0 < p_0 < p_1 < \infty$, $0 < q_0, q_1 \leq \infty$, and $\overline{r}, \overline{t} \in \mathbb{R}^N$. Furthermore, let $\nabla = (\nabla_{\overline{\nu}})_{\overline{\nu}\in\mathbb{N}_0^N}$ be a sequence of subsets of $\mathbb{Z}^d$. Then the following assertions are equivalent:

(i) $s^{\overline{r}+\overline{t}}_{p_0,q_0} f(\nabla) \subset s^{\overline{r}}_{p_1,q_1} f(\nabla)$

(ii) $s^{\overline{r}+\overline{t}}_{p_0,q_0} f(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q_1} f(\nabla)$

(iii) $t_i \geq d_i\left(\frac{1}{p_0} - \frac{1}{p_1}\right) > 0, \quad i = 1, \ldots, N.$

Remark 5.3.2. This result is the counterpart of the classical Sobolev-type embedding $F^{d(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}(\mathbb{R}^d) \hookrightarrow F^0_{p_1,q_1}(\mathbb{R}^d)$, see e.g. [39] or [83, Theorem 2.7.1(ii)]. The according embedding for sequence spaces can either be proved using the wavelet isomorphisms from Section 1.2.2, or it can be proved directly following the lines of the proof of the function space result in [83] (essentially one has to replace $f_j = \mathcal{F}^{-1}\varphi_j\mathcal{F}f$ by $g_j = \sum_{m\in\nabla_j} a_{j,m}\mathcal{X}_{j,m}$). The corresponding result for function spaces in the case $N = 2$, $d_1 = d_2 = 1$, can already be found in [71], Proposition 2.4.1 and Theorem 2.4.1. From that proof we took the idea for the proof below.

Proof. In case $\overline{t} > \overline{d}(\frac{1}{p_0} - \frac{1}{p_1})$ the embedding (ii) already follows from Proposition 5.3.1 and Lemma 5.3.3. On the other hand, the condition $\overline{t} \geq \overline{d}(\frac{1}{p_0} - \frac{1}{p_1})$ can always be reduced to $\overline{t} = \overline{d}(1/p_0 - 1/p_1)$ because of Lemma 5.3.2(ii). Finally, it is sufficient to consider the case $\nabla_{\overline{\nu}} = \mathbb{Z}^d$, since the spaces $s^{\overline{r}}_{p,q}f(\nabla)$ can be identified with a closed subspaces of $s^{\overline{r}}_{p,q}f$ with coinciding quasi-norms, and the restriction of the embedding operator is a mapping between the respective subspaces.
We prove the result by induction over N. The induction basis $N = 1$ corresponds to spaces $f^s_{p,q}$. In the induction step we write every N-tuple as in $\overline{\nu} = (\nu', \nu_N)$, similar for d-tuples $m = (m', m^N)$. Using the embedding $\ell_{\min(p,q)}(L_p) \hookrightarrow L_p(\ell_q) \hookrightarrow \ell_{\max(p,q)}(L_p)$, see Proposition 2.3.7, we find for $a \in s^{\overline{t}}_{p_0,q_0}f$

$$\left\| g_{\overline{\nu},m} \left| L_{p_1}(\ell_{q_1}) \right\| \leq c \left\| g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_1|x^N} \Big(\ell_{\min(p_1,q_1)|\nu_N,m^N} \big(L_{p_1|x'}(\ell_{q_1|\nu',m'}) \big) \Big) \right\|$$

and

$$\left\| 2^{\overline{\nu}\cdot\overline{t}} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_0|x^N} \Big(\ell_{\max(p_0,q_0)|\nu_N,m^N} \big(L_{p_0|x'}(\ell_{q_0|\nu',m'}) \big) \Big) \Big\| \leq c \left\| 2^{\overline{\nu}\cdot\overline{t}} g_{\overline{\nu},m} \left| L_{p_0}(\ell_{q_0}) \right\| , \right.$$

where $g_{\overline{\nu},m}(x) = 2^{\overline{\nu}\cdot\overline{r}} |a_{\overline{\nu},m}| \mathcal{X}_{\overline{\nu},m}(x)$, $\overline{\nu} \in \mathbb{N}_0^N$, $m \in \mathbb{Z}^d$. The notation $\ell_{q_1|\nu',m'}$ indicates, that the ℓ_{q_1}-norm is applied to the indices ν' and m', accordingly for the other norms. By the ℓ_q-monotonicity it hence suffices to show

$$\begin{aligned} &\left\| g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_1|x^N} \Big(\ell_{q_1|\nu_N,m^N} \big(L_{p_1|x'}(\ell_{q_1|\nu',m'}) \big) \Big) \Big\| \\ &\qquad \leq c \left\| 2^{\overline{\nu}\cdot\overline{t}} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_0|x^N} \Big(\ell_{\infty|\nu_N,m^N} \big(L_{p_0|x'}(\ell_{\infty|\nu',m'}) \big) \Big) \Big\| , \end{aligned} \tag{5.3.1}$$

i.e. we consider the case $q_1 < p_1 < \infty$ and $q_0 = \infty$. Furthermore, we put

$$h_{\nu_N,m^N}(x^N) = \left\| 2^{\nu'\cdot t'} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_0|x'}(\ell_{\infty|\nu',m'}) \Big\| , \qquad \overline{\nu} \in \mathbb{N}_0^N , \quad m \in \mathbb{Z}^d .$$

These functions are constant on the interior of every cube $Q_{\nu_N,m^N} \subset \mathbb{R}^{d_N}$, since it can be seen easily that for $x^N \in \operatorname{int} Q_{\nu_N,m^N}$ we have $h_{\nu_N,m^N}(x^N)\mathcal{X}_{\nu_N,m^N}(x^N) = h_{\nu_N,m^N}(x^N)$. This means we can identify these functions with the sequence η, given by

$$\eta_{j,k} = h_{j,k}(t) , \qquad j \in \mathbb{N}_0 , \quad k \in \mathbb{Z}^{d_N} , \quad t \in 2^{-j}\big((0,1)^{d_N} + k\big) .$$

The finiteness of the right hand side of (5.3.1) then implies $\eta \in f^{d_N(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,\infty}$. From the embedding $f^{d_N(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,\infty} \hookrightarrow f^0_{p_1,q_1}$ we now conclude

$$\begin{aligned} &\left\| 2^{\nu'\cdot t'} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_1|x^N} \Big(\ell_{q_1|\nu_N,m^N} \big(L_{p_0|x'}(\ell_{\infty|\nu',m'}) \big) \Big) \Big\| \\ &\qquad \leq c \left\| 2^{\overline{\nu}\cdot\overline{t}} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_0|x^N} \Big(\ell_{\infty|\nu_N,m^N} \big(L_{p_0|x'}(\ell_{\infty|\nu',m'}) \big) \Big) \Big\| . \end{aligned} \tag{5.3.2}$$

Moreover, for fixed $x^N \in \mathbb{R}^{d_N}$, $m^N \in \mathbb{Z}^{d_N}$ and $\nu_N \in \mathbb{N}_0$ the induction hypothesis implies

$$\left\| g_{(\nu',\nu_N),(m',m^N)} \left| L_{p_1|x'}(\ell_{q_1|\nu',m'}) \right\| \leq c \left\| 2^{\nu'\cdot t'} g_{(\nu',\nu_N),(m',m^N)} \right| L_{p_0|x'}(\ell_{\infty|\nu',m'}) \right\| ,$$

since the families of functions $\big(g_{(\nu',\nu_N),(m',m^N)}(\,\cdot\,,x^N)\big)_{\nu'\in\mathbb{N}_0^{N-1},\,m'\in\mathbb{Z}^{d-d_1}}$ (defined on $\mathbb{R}^{d-d_N}$) correspond to sequences in $s^{t'}_{p_0,\infty}f$, where $\mathcal{X}_{\nu_N,m^N}(x^N)$ appears as a constant factor. Applying the $L_{p_1|x^N}\big(\ell_{q_1|\nu_N,m^N}\big)$-quasi-norm further yields

$$\begin{aligned}\Big\|\, g_{(\nu',\nu_N),(m',m^N)}\,\Big|L_{p_1|x^N}\Big(\ell_{q_1|\nu_N,m^N}\big(L_{p_1|x'}(\ell_{q_1|\nu',m'})\big)\Big)\Big\|&\\ \le c\,\Big\|\, 2^{\nu'\cdot t'} g_{(\nu',\nu_N),(m',m^N)}\,\Big|L_{p_1|x^N}\Big(\ell_{q_1|\nu_N,m^N}\big(L_{p_0|x'}(\ell_{\infty|\nu',m'})\big)\Big)\Big\|\,.&\qquad(5.3.3)\end{aligned}$$

Combining (5.3.2) and (5.3.3) proves (5.3.1).
Finally, together with Lemma 5.3.2 we find for arbitrary $\overline{t}\ge\overline{d}(1/p_0-1/p_1)>0$

$$s^{\overline{r}+\overline{t}}_{p_0,q_0}f(\nabla)\hookrightarrow s^{\overline{r}+\overline{d}(1/p_0-1/p_1)}_{p_0,\infty}f(\nabla)\hookrightarrow s^{\overline{r}}_{p_1,q}f(\nabla)\hookrightarrow s^{\overline{r}}_{p_1,q_1}f(\nabla)\,,$$

for arbitrary $0<q_0,q_1\le\infty$ and $0<q<\min(p_1,q_1)<\infty$, which finally proves (ii). □

As before, we can overcome the condition $p_0<p_1$ by considering sequences $\nabla=\nabla(\Omega)$ for bounded domains.

Proposition 5.3.4. Let Ω be a bounded open subset of $\mathbb{R}^d$. Furthermore, let $\overline{r},\overline{t}\in\mathbb{R}^N$, $0<p_1\le p_0<\infty$, and $0<q_0,q_1\le\infty$. Then the following assertions are equivalent:

(i) $s^{\overline{r}+\overline{t}}_{p_0,q_0}f(\Omega)\subset s^{\overline{r}}_{p_1,q_1}f(\Omega)$

(ii) $s^{\overline{r}+\overline{t}}_{p_0,q_0}f(\Omega)\hookrightarrow s^{\overline{r}}_{p_1,q_1}f(\Omega)$

(iii) It either holds $\overline{t}>0\,,$ or $\overline{t}\ge 0$ and $q_0\le q_1\,.$

Proof. As discussed in Remark 5.1.1 for every sequence $\nabla(\Omega)$ generated by a bounded set Ω there exists a bounded set Γ, such that

$$\operatorname{supp}\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\nabla_{\overline{\nu}}}|a_{\overline{\nu},m}|^q\mathcal{X}_{\overline{\nu},m}(\,\cdot\,)\right)\subset\Gamma$$

for all sequences $a\in s^{\overline{r}}_{p,q}f(\Omega)$ and all $0<q\le\infty$. It is a well-known consequence of Hölder's inequality that for measurable sets with finite Lebesgue measure $\Gamma\subset\mathbb{R}^d$ it holds $\|\,f\,|L_p(\Gamma)\|\le|\Gamma|^{1/p-1/q}\|\,f\,|L_q(\Gamma)\|$ for all $0<p\le q\le\infty$.
The assumptions in (iii) ensure the embedding $s^{\overline{r}+\overline{t}}_{p_1,q_0}f(\Omega)\hookrightarrow s^{\overline{r}}_{p_1,q_1}f(\Omega)$. This follows either from Lemma 5.3.5 (in case $\overline{t}>0$) or from the monotonicity of the ℓ_q-quasi-norms (in case $q_0\le q_1$; see Lemma 5.3.2). Both times we can further conclude

$$\begin{aligned}\big\|\,a\,\big|s^{\overline{r}}_{p_1,q_1}f(\Omega)\big\|&\le c\,\big\|\,a\,\big|s^{\overline{r}+\overline{t}}_{p_1,q_0}f(\Omega)\big\|\\&=c\,\left\|\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\nabla_{\overline{\nu}}}2^{\overline{\nu}\cdot(\overline{r}+\overline{t})q_0}|a_{\overline{\nu},m}|^{q_0}\mathcal{X}_{\overline{\nu},m}(x)\right)^{1/q_0}\right|L_{p_1}(\Gamma)\Bigg\|\\&\le c'\,\left\|\left(\sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{m\in\nabla_{\overline{\nu}}}2^{\overline{\nu}\cdot(\overline{r}+\overline{t})q_0}|a_{\overline{\nu},m}|^{q_0}\mathcal{X}_{\overline{\nu},m}(x)\right)^{1/q_0}\right|L_{p_0}(\Gamma)\Bigg\|\\&=c'\,\big\|\,a\,\big|s^{\overline{r}+\overline{t}}_{p_0,q_0}f(\Omega)\big\|\,.\end{aligned}$$

This proves the assertion (modification for $q_0=\infty$). □

To complete our overview on embeddings between the considered scales of Besov and Triebel-Lizorkin spaces we cite the following result from [33], which deals with the mixed cases. In that article one can find some further results and references.

Proposition 5.3.5. Let $0 < p_0 < p < p_1 \leq \infty$, $0 < q_0, q, q_1 \leq \infty$ and $\overline{r}^{\,0}, \overline{r}, \overline{r}^{\,1} \in \mathbb{R}^N$, such that

$$\overline{r}^{\,0} - \overline{d}/p_0 = \overline{r} - \overline{d}/p = \overline{r}^{\,1} - \overline{d}/p_1 \,.$$

Then it holds

$$S^{\overline{r}^{\,0}}_{p_0,q_0}B(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}}_{p,q}F(\mathbb{R}^{\overline{d}}) \hookrightarrow S^{\overline{r}^{\,1}}_{p_1,q_1}B(\mathbb{R}^{\overline{d}})$$

if, and only if, $q_0 \leq p \leq q_1$.

Remark 5.3.3. We would like to point out that also in this situation we have a continuous embedding if, and only if, it holds set theoretic inclusion. This follows immediately from the counterexamples used in [33] to proof the "only if"-part.
The proof in [33] itself is done for sequence spaces, and a closer examination yields, that the sufficiency of the above conditions remains valid even for arbitrary sequences ∇. This can be seen either by direct arguments, or via interpreting the spaces $s^{\overline{r}}_{p,q}a(\nabla)$ as subspaces of $s^{\overline{r}}_{p,q}a$.

5.4 Compactness of embeddings

Our next aim is to investigate which of the embeddings of the last section are not only continuous but compact. To this purpose we want to approximate the embedding operator by a sequence of finite rank operators with respect to the operator norm in $\mathcal{L}\big(s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\Omega), s^{\overline{r}}_{p_1,q_1}a^{\dagger}(\Omega)\big)$.
We define operators $\mathrm{re}_{\overline{\nu}}$ by putting

$$\big(\mathrm{re}_{\overline{\nu}}\,\eta\big)_{\overline{\lambda},m} = \begin{cases} \eta_{\overline{\nu},m}\,, & \overline{\lambda} = \overline{\nu}\,, m \in \nabla_{\overline{\nu}}\,, \\ 0, & \text{else}\,, \end{cases} \tag{5.4.1}$$

for every sequence $\eta \in s^{\overline{r}}_{p,q}a(\nabla)$. We find at once

$$2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)}\big\|\mathrm{re}_{\overline{\nu}}\,\eta\big|\ell_p(\mathbb{N}_0^N \times \nabla)\big\| = \big\|\mathrm{re}_{\overline{\nu}}\,\eta\big|s^{\overline{r}}_{p,q}a(\nabla)\big\| \leq \big\|\,\eta\,\big|s^{\overline{r}}_{p,q}a(\nabla)\big\|. \tag{5.4.2}$$

As a first step, we now collect some further properties of the operators $\mathrm{re}_{\overline{\nu}}$.

Lemma 5.4.1. Consider sequences η as in (5.1.1), and define spaces

$$\mathcal{R}^{\overline{r},\overline{\nu}}_{p,q}(\nabla) = \big\{\eta \in s^{\overline{r}}_{p,q}b(\nabla) \,:\, \eta_{\overline{\lambda},m} = 0 \quad \text{for all} \quad \overline{\lambda} \neq \overline{\nu}\big\}\,,$$

where $0 < p, q \leq \infty$, $\overline{r} \in \mathbb{R}^N$, and $\overline{\nu} \in \mathbb{N}_0^N$.

(i) The operator $\mathrm{re}_{\overline{\nu}}$ is the identical mapping on $\mathcal{R}^{\overline{r},\overline{\nu}}_{p,q}(\nabla)$ for all admissible parameters.

(ii) Let $0 < p_0 \leq p_1 \leq \infty$, $0 < q_0, q_1 \leq \infty$, and $\overline{r}, \overline{t} \in \mathbb{R}^N$. Furthermore, let $\overline{\nu} \in \mathbb{N}_0^N$. Then it holds

$$\left\| \mathrm{re}_{\overline{\nu}} : \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\nabla) \longrightarrow \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\nabla) \right\| = 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}(1/p_0-1/p_1))} . \tag{5.4.3}$$

(iii) Let Ω be a bounded open subset of $\mathbb{R}^d$. Moreover, let $0 < p_1 \leq p_0 \leq \infty$ and $\overline{r}, \overline{t} \in \mathbb{R}^N$. Then it holds for all $\overline{\nu} \in \mathbb{N}_0^N$

$$c_1^{1/p_1-1/p_0} 2^{-\overline{\nu}\cdot\overline{t}} \leq \left\| \mathrm{re}_{\overline{\nu}} : \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \longrightarrow \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega) \right\| \leq c_2^{1/p_1-1/p_0} 2^{-\overline{\nu}\cdot\overline{t}} , \tag{5.4.4}$$

where the constants c_1, c_2 have the same meaning as in (5.1.3).

Proof. **Step 1:** Let $p_0 \leq p_1$, and let $\eta \in \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\nabla)$. Then it follows from (5.4.2) and the monotonicity of the ℓ_p-quasi-norms

$$\begin{aligned}
\| \eta \,|\mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\nabla) \| &= \| \eta \,| s^{\overline{r}}_{p_1,q_1} b(\nabla) \| = 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)} \| \eta \,| \ell_{p_1}(\mathbb{N}_0^N \times \nabla) \| \\
&= 2^{\overline{\nu}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}/p_0+\overline{d}/p_1)} \| \eta \,| \ell_{p_1}(\mathbb{N}_0^N \times \nabla) \| \\
&\leq 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}/p_0+\overline{d}/p_1)} 2^{\overline{\nu}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} \| \eta \,| \ell_{p_0}(\mathbb{N}_0^N \times \nabla) \| \\
&= 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}(1/p_0-1/p_1))} \| \eta \,| \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\nabla) \| .
\end{aligned}$$

The estimate from below can be obtained from the sequences $\eta = e^{\overline{\nu},m}$ for arbitrary $m \in \nabla_{\overline{\nu}}$. It follows

$$\left\| e^{\overline{\nu},m} \middle| \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\nabla) \right\| = 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)} \quad \text{and} \quad \left\| e^{\overline{\nu},m} \middle| \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\nabla) \right\| = 2^{\overline{\nu}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} ,$$

hence their quotient yields the desired estimate. This proves (i).

Step 2: Let $p_1 < p_0$ and $\eta \in \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega)$. Then it follows from Hölder's inequality with respect to $1 = \frac{p_1}{p_0} + \frac{p_0-p_1}{p_0}$ and (5.1.3)

$$\begin{aligned}
\| \eta \,|\mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega) \| &= 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)} \Big(\sum_{m\in\nabla_{\overline{\nu}}} |\eta_{\overline{\nu},m}|^{p_1} \Big)^{1/p_1} \\
&\leq 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)} \Big(\sum_{m\in\nabla_{\overline{\nu}}} |\eta_{\overline{\nu},m}|^{p_0} \Big)^{1/p_0} \Big(\#\nabla_{\overline{\nu}} \Big)^{1/p_1-1/p_0} \\
&= \Big(\#\nabla_{\overline{\nu}} \Big)^{1/p_1-1/p_0} 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}/p_0+\overline{d}/p_1)} \| \eta \,| \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \| \\
&\leq 2^{-\overline{\nu}\cdot\overline{t}} c_2^{1/p_1-1/p_0} \| \eta \,| \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \| .
\end{aligned}$$

This yields the estimate from above in (5.4.4). The estimate from below now follows from the constant sequence, i.e. the sequence $\eta^c \in \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega)$ with $\eta^c_{\overline{\nu},m} = 1$ for all $m \in \nabla_{\overline{\nu}}$. For that sequence we obtain

$$\| \eta^c \,| \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \| = 2^{\overline{\nu}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)} \big(\#\nabla_{\overline{\nu}} \big)^{1/p_0}$$

and on the other hand

$$\| \eta^c \,| \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega) \| = 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p_1)} \big(\#\nabla_{\overline{\nu}} \big)^{1/p_1} .$$

Together, their quotient yields

$$\left\|\mathrm{re}_{\overline{\nu}} : \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \longrightarrow \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega)\right\| \geq 2^{-\overline{\nu}\cdot\overline{t}} 2^{-\overline{\nu}\cdot\overline{d}(1/p_1-1/p_0)} \big(\#\nabla_{\overline{\nu}}\big)^{1/p_1-1/p_0} \geq c_1^{1/p_1-1/p_0} 2^{-\overline{\nu}\cdot\overline{t}}.$$

This completes the proof of (ii). □

With the help of the operators $\mathrm{re}_{\overline{\nu}}$ we can decompose every sequence $\eta \in s^{\overline{r}}_{p,q}a(\nabla)$ into pieces belonging to $\mathcal{R}^{\overline{r},\overline{\nu}}_{p,q}(\nabla)$. Likewise this can be seen as a decomposition of the identical map (i.e. of embedding operators). This decomposition is the crucial step for proving compactness of embedding operators.

Theorem 5.4.1. Let $0 < p_0, p_1, q_0, q_1 \leq \infty$ and $\overline{r}, \overline{t} \in \mathbb{R}^N$.

(i) Let $p_0 \leq p_1$, and let the condition (iii) in Proposition 5.3.1 be satisfied. Moreover, we assume that there exists some $\overline{\lambda} \in \mathbb{N}_0^N$, such that $\#\nabla_{\overline{\lambda}} = \infty$. Then the embedding $s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\nabla) \hookrightarrow s^{\overline{r}}_{p_1,q_1}b(\nabla)$ is not compact.

(ii) Let Ω be a bounded open subset of $\mathbb{R}^d$. Then the continuous embedding $s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,q_1}a^\dagger(\Omega)$ is compact for every combination of $a, a^\dagger \in \{b, f\}$ if, and only if, it holds

$$\overline{\alpha} := \overline{t} - \overline{d}\left(\frac{1}{p_0} - \frac{1}{p_1}\right)_+ > 0\,, \tag{5.4.5}$$

where $p_0 < \infty$ if $a = f$, and $p_1 < \infty$ if $a^\dagger = f$.

Proof. **Step 1:** Let $\overline{\lambda} \in \mathbb{N}_0^N$, such that $\#\nabla_{\overline{\lambda}} = \infty$. Then we obtain for the sequences $e^{\overline{\lambda},m}$ for every $m \in \nabla_{\overline{\lambda}}$

$$\left\|e^{\overline{\lambda},m}\middle|s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\nabla)\right\| = 2^{\overline{\lambda}\cdot(\overline{r}+\overline{t}-\overline{d}/p_0)}\,,$$

independently of $m \in \nabla_{\overline{\lambda}}$, and on the other hand we find

$$\left\|e^{\overline{\lambda},m} - e^{\overline{\lambda},m'}\middle|s^{\overline{r}}_{p_1,q_1}b(\nabla)\right\| = 2^{1/p_1} 2^{\overline{\lambda}\cdot(\overline{r}-\overline{d}/p_1)}\,, \quad m \neq m'\,.$$

Hence the sequence $(e^{\overline{\lambda},m})_{m\in\nabla_{\overline{\lambda}}} \subset s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\nabla)$ is a bounded one, but one cannot pick a subsequence convergent in $s^{\overline{r}}_{p_1,q_1}b(\nabla)$. This proves (i).

Step 2: We start with the case $a = a^\dagger = b$. Let $\overline{\alpha} > 0$. To prove the compactness we decompose the identity $\mathrm{id} : s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q_1}b(\Omega)$ in

$$\mathrm{id} = \sum_{\overline{\nu}\in\mathbb{N}_0^N} \mathrm{re}_{\overline{\nu}}\,. \tag{5.4.6}$$

We now want to use the properties of the operators $\mathrm{re}_{\overline{\nu}}$ proved in Lemma 5.4.1 to prove the absolute convergence of this series in $\mathcal{L}\big(s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega), s^{\overline{r}}_{p_1,q_1}b(\Omega)\big)$. It follows at once, keeping in mind the relation $\mathrm{re}_{\overline{\nu}}\big(s^{\overline{s}}_{p,q}b(\Omega)\big) = \mathcal{R}^{\overline{s},\overline{\nu}}_{p,q}(\Omega)$,

$$\left\|\mathrm{re}_{\overline{\nu}} : s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q_1}b(\Omega)\right\| = \left\|\mathrm{re}_{\overline{\nu}} : \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \longrightarrow \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega)\right\|.$$

Now Lemma 5.4.1, equations (5.4.3) and (5.4.4), yield

$$\sum_{\overline{\nu}\in\mathbb{N}_0^N} \left\|\mathrm{re}_{\overline{\nu}} : s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q_1}b(\Omega)\right\| \leq \sum_{\overline{\nu}\in\mathbb{N}_0^N} \left\|\mathrm{re}_{\overline{\nu}} : \mathcal{R}^{\overline{r}+\overline{t},\overline{\nu}}_{p_0,q_0}(\Omega) \longrightarrow \mathcal{R}^{\overline{r},\overline{\nu}}_{p_1,q_1}(\Omega)\right\| \leq c \sum_{\overline{\nu}\in\mathbb{N}_0^N} 2^{-\overline{\nu}\cdot\overline{\alpha}} < \infty\,,$$

where in the last estimate we used the assumption (5.4.5). Thus the series in (5.4.6) converges in the operator norm, since absolute convergence implies norm convergence. On the other hand, for every sequence $\eta \in s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega)$ the only possible limit for $\sum_{\overline{\nu}\in\mathbb{N}_0^N} \mathrm{re}_{\overline{\nu}}\,\eta$ is η itself, since for every component at most one summand is non-vanishing. This implies, that the series (5.4.6) indeed converges to the identical mapping.
Finally, since we have $\dim \mathcal{R}^{\overline{r},\overline{\nu}}_{p,q}(\Omega) = \#\nabla_{\overline{\nu}} \sim 2^{\overline{\nu}\cdot\overline{d}} < \infty$ for all parameters $\overline{r}$, p and q (see (5.1.3)), the mappings $\mathrm{re}_{\overline{\nu}}$ have finite rank for all $\overline{\nu} \in \mathbb{N}_0^N$. Altogether we have shown, that $\mathrm{id} : s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q_1}b(\Omega)$ can be approximated by finite rank operators in $\mathcal{L}\big(s^{\overline{r}+\overline{t}}_{p_0,q_0}b(\Omega), s^{\overline{r}}_{p_1,q_1}b(\Omega)\big)$, which yields the compactness of the embedding.

Step 3: The compactness of the embedding $s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,q_1}a^{\dagger}(\Omega)$ now will be traced back to step 2 with the help of Lemma 5.3.2(i) and Lemma 5.3.3. It holds

$$s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\Omega) \hookrightarrow s^{\overline{r}+\overline{t}}_{p_0,\max(p_0,q_0)}b(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,\min(p_1,q_1)}b(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,q_1}a^{\dagger}(\Omega)\,,$$

where all embeddings are continuous, and the middle one is even compact. Hence their composition is compact as well, which proves the assertion. □

Remark 5.4.1. Similar non-compactness assertions as in (i) hold for the other continuous embeddings in Lemma 5.3.2, Lemma 5.3.3, Lemma 5.3.5, Proposition 5.3.3 and Proposition 5.3.5, whenever at least one of the sets $\nabla_{\overline{\nu}}$ is infinite.

5.5 Dual spaces of sequence spaces

In Section 2.3.8 we proved duality assertions for the Besov spaces of dominating mixed smoothness. Using the wavelet isomorphism from Section 4.3, these results could be transferred to the sequence spaces $s^{\overline{r}}_{p,q}b$. Since we will need results for spaces $s^{\overline{r}}_{p,q}b(\nabla)$ for more general ∇, we will give an alternative argument using the dual spaces of $\ell_q(I, \ell_p(J))$. Moreover, in this section we will characterize the dual spaces for $s^{\overline{r}}_{p,q}f(\nabla)$. However, throughout this section we will stick to the case of Banach spaces, i.e. we will consider the case $1 \leq p, q < \infty$ only. The restriction $\max(p,q) < \infty$ has the same background as in Section 2.3.8, because the counterpart of the density of $\mathcal{S}(\mathbb{R}^d)$ in $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ consists in the density of the finite sequences.

Lemma 5.5.1. Let $1 \leq p, q < \infty$ and I, J as before. Then

$$\big(\ell_q(I, \ell_p(J))\big)' = \ell_{q'}(I, \ell_{p'}(J))\,, \tag{5.5.1}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} + \frac{1}{q'} = 1$. More precisely, there is a canonical isometric isomorphism $T : \ell_{q'}(I, \ell_{p'}(J)) \longrightarrow \big(\ell_q(I, \ell_p(J))\big)'$, which admits the representation

$$\big(Ta\big)(b) = \sum_{i\in I}\sum_{j\in J_i} a_{i,j}b_{i,j} \tag{5.5.2}$$

for $a \in \ell_{q'}(I, \ell_{p'}(J))$ and $b \in \ell_q(I, \ell_p(J))$. If the sets J_i are finite for all $i \in I$, then this result extends to spaces $\ell_q(I, \ell_\infty(J))$.

Proof. The convergence of the series in (5.5.2) follows by using Hölder's inequality twice. More precisely, it follows

$$|(Ta)(b)| \le \big\| \, a \,\big|\ell_{q'}(I, \ell_{p'}(J))\big\| \cdot \big\| \, b \,\big|\ell_q(I, \ell_p(J))\big\| \,,$$

hence T is well-defined and $\big\| \, Ta \,\big|\big(\ell_q(I, \ell_p(J))\big)'\big\| \le \big\| \, a \,\big|\ell_{q'}(I, \ell_{p'}(J))\big\|$. That T is injective follows from $(Ta)(e^{i,j}) = a_{i,j}$.
It remains to show, that T is surjective and isometric. Let $y \in \big(\ell_q(I, \ell_p(J))\big)'$ and define $x_{i,j} = y(e^{i,j})$. Thus we shall show $x \in \ell_{q'}(I, \ell_{p'}(J))$, where $x = (x_{i,j})_{i,j}$, and $y = Tx$ with $\big\| \, x \,\big|\ell_{q'}(I, \ell_{p'}(J))\big\| \le \big\| \, y \,\big|\big(\ell_q(I, \ell_p(J))\big)'\big\|$.
For this purpose, let $1 < p, q < \infty$ first. Then, for some fixed $i \in I$ consider the restriction of y to sequences of the form $a = \sum_{j \in J_i} a_{i,j} e^{i,j}$. We denote this restriction by y_i. We immediately find $y_i \in \big(\ell_p(J_i)\big)'$, if we interpret $\ell_p(J_i)$ as a subspace of $\ell_q(I, \ell_p(J))$. By the well-known characterization of the spaces $\big(\ell_p(J_i)\big)'$ we then know, that y_i is generated by some sequence $x_i \in \ell_{p'}(J_i)$ in the usual way. But as $(x_i)_j = y_i(e^j) = y(e^{i,j}) = x_{i,j}$, this can be rewritten as $(x_{i,j})_{j \in J_i} \in \ell_{p'}(J_i)$ and

$$\alpha_i := \sum_{j \in J_i} |x_{i,j}|^{p'} < \infty \,.$$

Next, put

$$\beta_{i,j} := \begin{cases} \frac{|x_{i,j}|^{p'}}{x_{i,j}} \,, & x_{i,j} \neq 0 \,, \\ 0 \,, & \text{else} \,, \end{cases} \qquad \text{and} \qquad \gamma_i := \begin{cases} \frac{|\alpha_i|^{q'/p'}}{\alpha_i} \,, & \alpha_i \neq 0 \,, \\ 0 \,, & \text{else} \,. \end{cases}$$

We now write the countable index set I as $I = \{i_1, i_2, \ldots\}$, and define further index sets $I_N := \{i_1, \ldots, i_N\}$ for $N \in \mathbb{N}$. This gives

$$\begin{aligned}
\sum_{i \in I_N} \Big(\sum_{j \in J_i} |x_{i,j}|^{p'} \Big)^{\frac{q'}{p'}} &= \sum_{i \in I_N} \gamma_i \sum_{j \in J_i} |x_{i,j}|^{p'} = \sum_{i \in I_N} \gamma_i \sum_{j \in J_i} \beta_{i,j} x_{i,j} \\
&= \sum_{i \in I_N} \gamma_i \sum_{j \in J_i} \beta_{i,j} y(e^{i,j}) = y \Big(\sum_{i \in I_N} \gamma_i \sum_{j \in J_i} \beta_{i,j} e^{i,j} \Big) \\
&\le \big\| \, y \,\big|\big(\ell_q(I, \ell_p(J))\big)'\big\| \Big(\sum_{i \in I_N} \gamma_i^q \Big(\sum_{j \in J_i} |\beta_{i,j}|^p \Big)^{\frac{q}{p}} \Big)^{\frac{1}{q}} \\
&= \big\| \, y \,\big|\big(\ell_q(I, \ell_p(J))\big)'\big\| \Big(\sum_{i \in I_N} \gamma_i^q \Big(\sum_{j \in J_i} |x_{i,j}|^{p'} \Big)^{\frac{q}{p}} \Big)^{\frac{1}{q}} \\
&= \big\| \, y \,\big|\big(\ell_q(I, \ell_p(J))\big)'\big\| \Big(\sum_{i \in I_N} \Big(\sum_{j \in J_i} |x_{i,j}|^{p'} \Big)^{\frac{q'}{p'}} \Big)^{\frac{1}{q}} ,
\end{aligned}$$

where we used $p(p'-1)=p'$ and $\frac{q'}{p'}-\frac{q}{p}=q\big(\frac{q'}{p'}-1\big)$. Hence

$$\left(\sum_{i\in I_N}\left(\sum_{j\in J_i}|x_{i,j}|^{p'}\right)^{\frac{q'}{p'}}\right)^{\frac{1}{q'}} \le \big\| \, y\,\big|\big(\ell_q(I,\ell_p(J))\big)'\big\| \qquad \text{for all} \quad N\in\mathbb{N}\,.$$

Letting $N\longrightarrow\infty$ (i.e. taking the supremum over $N\in\mathbb{N}$) yields $x\in\ell_{q'}(I,\ell_{p'}(J))$ and $\big\|\,x\,\big|\ell_{q'}(I,\ell_{p'}(J))\big\| \le \big\|\,y\,\big|\big(\ell_q(I,\ell_p(J))\big)'\big\|$. Finally, $Tx=y$ follows from the respective definitions and the density of the finite sequences.
The case $q=1$ follows from $\big\|\,y_i\,\big|\big(\ell_p(J_i)\big)'\big\| \le \big\|\,y\,\big|\big(\ell_q(I,\ell_p(J))\big)'\big\|$ for all $i\in I$. For the case $p=1$ and $q>1$, we define for every $i\in I$ sets $J_i^N:=\{j_i^1,\ldots,j_i^N\}$, where $J_i=\{j_i^1,j_i^2,\ldots,\}$. Then we always find an index j_i^l, such that $|x_{i,j_i^l}|=\sup_{j\in J_i^N}|x_{i,j}|$. If we put

$$\beta_{i,j}=\begin{cases}\frac{|x_{i,j}|}{x_{i,j}}\,, & j=j_i^l\,,\; x_{i,j_i^l}\neq 0\,,\\ 0\,, & \text{else}\,,\end{cases} \qquad \text{and} \qquad \gamma_i=\sup_{j\in J_i^N}|x_{i,j}|\,,$$

then we have $\sum_{j\in J_i^N}\beta_{i,j}x_{i,j}=\sup_{j\in J_i^N}|x_{i,j}|$, and the rest of the argument remains the same.
If the sets J_i are finite for every $i\in I$, then the finite sequences are dense in $\ell_q(I,\ell_\infty(J))$, and the characterization of the dual space follows by the same arguments as in the case $1<p<\infty$. We only note the necessary modifications in the proof of the surjectivity of T. We put $\beta_{i,j}=\operatorname{sgn}x_{i,j}$ and $\gamma_i=\alpha_i^{q'-1}$. Then we find

$$\begin{aligned}
\sum_{i\in I_N}\left(\sum_{j\in J_i}|x_{i,j}|\right)^{q'} &= \sum_{i\in I_N}\gamma_i\sum_{j\in J_i}|x_{i,j}| = \sum_{i\in I_N}\gamma_i\sum_{j\in J_i}\beta_{i,j}x_{i,j}\\
&= \sum_{i\in I_N}\gamma_i\sum_{j\in J_i}\beta_{i,j}y(e^{i,j}) = y\left(\sum_{i\in I_N}\gamma_i\sum_{j\in J_i}\beta_{i,j}e^{i,j}\right)\\
&\le \big\|\,y\,\big|\big(\ell_q(I,\ell_\infty(J))\big)'\big\|\left(\sum_{i\in I_N}\gamma_i^q\left(\sup_{j\in J_i}|\beta_{i,j}|\right)^q\right)^{\frac{1}{q}}\\
&\le \big\|\,y\,\big|\big(\ell_q(I,\ell_\infty(J))\big)'\big\|\left(\sum_{i\in I_N}\alpha_i^{q'}\right)^{\frac{1}{q}}\\
&= \big\|\,y\,\big|\big(\ell_q(I,\ell_\infty(J))\big)'\big\|\left(\sum_{i\in I_N}\left(\sum_{j\in J_i}|x_{i,j}|\right)^{q'}\right)^{\frac{1}{q}}\,,
\end{aligned}$$

and the rest remains the same. □

The following lemma is an immediate corollary of Lemma 5.5.1.

Lemma 5.5.2. Let $1\le p,q<\infty$ and $\overline{r}\in\mathbb{R}^N$. Then it holds

$$\big(s^{\overline{r}}_{p,q}b(\nabla)\big)'=s^{-\overline{r}}_{p',q'}b(\nabla)$$

in the sense that there is a canonical isometric isomorphism $\widetilde{T} : s^{-\overline{r}}_{p',q'}b(\nabla) \longrightarrow \big(s^{\overline{r}}_{p,q}b(\nabla)\big)'$, which admits the representation

$$\big(\widetilde{T}x\big)(y) = \sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{k\in\nabla_{\overline{\nu}}} 2^{-\overline{\nu}\cdot\overline{d}} x_{\overline{\nu},k} y_{\overline{\nu},k} \tag{5.5.3}$$

for $x \in s^{-\overline{r}}_{p',q'}b(\nabla)$ and $y \in s^{\overline{r}}_{p,q}b(\nabla)$, where $\frac{1}{p}+\frac{1}{p'}=1$ and $\frac{1}{q}+\frac{1}{q'}=1$. This characterization remains true for spaces $s^{\overline{r}}_{\infty,q}b(\Omega)$, where Ω is a bounded open subset of $\mathbb{R}^d$.

Proof. We remind on the identification $\ell_q\big(\mathbb{N}_0^N, \ell_p(\nabla)\big) = L_{\overline{r}-\overline{d}/p}\big(s^{\overline{r}}_{p,q}b(\nabla)\big)$. Together with the operator $T : \ell_{q'}(\mathbb{N}_0^N, \ell_{p'}(\nabla)) \longrightarrow \big(\ell_q(\mathbb{N}_0^N, \ell_p(\nabla))\big)'$ from Lemma 5.5.1 we obtain

$$\widetilde{T} := \big(L_{\overline{r}-\overline{d}/p}\big)' T L_{-\overline{r}-\overline{d}/p'} \;:\; s^{-\overline{r}}_{p',q'}b(\nabla) \longrightarrow \big(s^{\overline{r}}_{p,q}b(\nabla)\big)'.$$

Hence, it remains to determine the dual operator $\big(L_{\overline{r}-\overline{d}/p}\big)' = L'_{\overline{r}-\overline{d}/p}$ of the lift-operator $L_{\overline{r}-\overline{d}/p}$. Inserting the respective definitions yields the following: if $y' \in \big(\ell_q(\mathbb{N}_0^N, \ell_p(\nabla))\big)'$ is generated by $y = (y_{\overline{\nu},k})_{\overline{\nu},k} \in \ell_{q'}(\mathbb{N}_0^N, \ell_{p'}(\nabla))$, then $L'_{\overline{r}-\overline{d}/p}y'$ can be represented as

$$\big(L'_{\overline{r}-\overline{d}/p}y'\big)(z) = y'\big(L_{\overline{r}-\overline{d}/p}z\big) = \sum_{\overline{\nu},k\in\mathbb{Z}} y_{\overline{\nu},k} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)} z_{\overline{\nu},k}\,,$$

where $z = (z_{\overline{\nu},k})_{\overline{\nu},k} \in s^{\overline{r}}_{p,q}b(\nabla)$.
As all operators involved are isometric isomorphisms, this also holds for $\widetilde{T}$, and we have the representation

$$\big(\widetilde{T}x\big)(y) = \sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{k\in\nabla_{\overline{\nu}}} 2^{\overline{\nu}\cdot(\overline{r}-\overline{d}/p)} 2^{\overline{\nu}\cdot(-\overline{r}-\overline{d}/p')} x_{\overline{\nu},k} y_{\overline{\nu},k} = \sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{k\in\nabla_{\overline{\nu}}} 2^{-\overline{\nu}\cdot\overline{d}} x_{\overline{\nu},k} y_{\overline{\nu},k}\,,$$

where $x \in s^{-\overline{r}}_{p',q'}b(\nabla)$ and $y \in s^{\overline{r}}_{p,q}b(\nabla)$. □

Lemma 5.5.3. Let $1 < p < \infty$, $1 \le q < \infty$ and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$\big(s^{\overline{r}}_{p,q}f(\nabla)\big)' = s^{-\overline{r}}_{p',q'}f(\nabla) \tag{5.5.4}$$

in the sense that there is a canonical isomorphism $T : s^{-\overline{r}}_{p',q'}f(\nabla) \longrightarrow \big(s^{\overline{r}}_{p,q}f(\nabla)\big)'$ with operator norm at most 1, which admits the representation

$$\big(Ta\big)(b) = \int_{\mathbb{R}^d} \sum_{\overline{\nu}\in\mathbb{N}_0^N}\sum_{k\in\nabla_{\overline{\nu}}} a_{\overline{\nu},k} b_{\overline{\nu},k} \mathcal{X}_{\overline{\nu},k}(x)\,dx\,, \tag{5.5.5}$$

for $a \in s^{-\overline{r}}_{p',q'}f(\nabla)$ and $b \in s^{\overline{r}}_{p,q}f(\nabla)$, where $\frac{1}{p}+\frac{1}{p'}=1$ and $\frac{1}{q}+\frac{1}{q'}=1$.

Proof. As $\mathcal{X}_{\overline{\nu},k}(x) = \mathcal{X}_{\overline{\nu},k}(x)^2$ for all $x \in \mathbb{R}^d$, the convergence of the right hand side of (5.5.5) follows by using Hölder's inequality twice. More precisely, it follows

$$\big|(Ta)(b)\big| \le \big\|\, a\,\big|s^{-\overline{r}}_{p',q'}f(\nabla)\big\| \cdot \big\|\, b\,\big|s^{\overline{r}}_{p,q}f(\nabla)\big\|\,,$$

hence T is well-defined and $\| Ta \,|\, (s^{\overline{r}}_{p,q}f(\nabla))' \| \leq \| a \,|\, s^{-\overline{r}}_{p',q'}f(\nabla)\|$. That T is injective follows from $(Ta)(e^{\overline{\nu},k}) = |Q_{\overline{\nu},k}| a_{\overline{\nu},k} = 2^{-\overline{\nu}\cdot\overline{d}} a_{\overline{\nu},k}$.
For the proof of the surjectivity of T we begin with the case $\overline{r} = \overline{0}$. Then it is clear, that we can interpret $s^{\overline{0}}_{p,q}f(\nabla)$ as a closed subspace of $L_p(\ell_q)$ (the corresponding index set being $I = \mathbb{N}_0^N$), see Remark 4.1.2. Now let $y \in \big(s^{\overline{0}}_{p,q}f(\nabla)\big)'$. Then by the Hahn-Banach theorem there is an extension of y to a functional $\widetilde{y}$ on $L_p(\ell_q)$ with equal norm. By the known characterization of the dual of $L_p(\ell_q)$, see Proposition 2.3.14, there is an element $g = (g_{\overline{\nu}})_{\overline{\nu}\in I} \in L_{p'}(\ell_{q'})$, such that for all $f = (f_{\overline{\nu}})_{\overline{\nu}\in I} \in L_p(\ell_q)$

$$\widetilde{y}(f) = \int_{\mathbb{R}} \sum_{\overline{\nu}\in I} f_{\overline{\nu}}(x) g_{\overline{\nu}}(x) dx\,.$$

Furthermore, it holds

$$\| \widetilde{y} \,|\, \big(L_p(\ell_q)\big)' \| = \| g \,|\, L_{p'}(\ell_{q'}) \|\,.$$

Now, put $x_{\overline{\nu},k} = |Q_{\overline{\nu},k}|^{-1} y(e^{\overline{\nu},k}) = |Q_{\overline{\nu},k}|^{-1} \widetilde{y}\big((\delta_{\overline{\lambda},\overline{\nu}} \mathcal{X}_{\overline{\nu},k})_{\overline{\lambda}\in I}\big)$. Then we find

$$|x_{\overline{\nu},k}| = |Q_{\overline{\nu},k}|^{-1} \left| \int_{Q_{\overline{\nu},k}} g_{\overline{\nu}}(s) ds \right| \leq \frac{1}{|Q_{\overline{\nu},k}|} \int_{Q_{\overline{\nu},k}} |g_{\overline{\nu}}(s)| ds \leq \big(\overline{M} g_{\overline{\nu}}\big)(t) \tag{5.5.6}$$

for every $t \in Q_{\overline{\nu},k}$. We remind on the directional maximal operators from Section 2.3.1. From (5.5.6) we conclude $\sum_{k\in\nabla_{\overline{\nu}}} |x_{\overline{\nu},k}| \mathcal{X}_{\overline{\nu},k} \leq \overline{M} g_{\overline{\nu}}$ for all $\overline{\nu}$, and by the vector-valued maximal inequality for $\overline{M}$, Proposition 2.3.1, we obtain

$$\begin{aligned} \| x \,|\, s^{\overline{0}}_{p',q'}f(\nabla) \| &= \Big\| \Big(\sum_{k\in\nabla_{\overline{\nu}}} x_{\overline{\nu},k} \mathcal{X}_{\overline{\nu},k} \Big)_{\overline{\nu}\in I} \Big| L_{p'}(\ell_{q'}) \Big\| \leq \Big\| \big(\overline{M} g_{\overline{\nu}}\big)_{\overline{\nu}\in I} \Big| L_{p'}(\ell_{q'}) \Big\| \\ &\leq c \, \big\| (g_{\overline{\nu}})_{\overline{\nu}\in I} \big| L_{p'}(\ell_{q'}) \big\| = c \, \| \widetilde{y} \,|\, \big(L_p(\ell_q)\big)' \| = c \, \| y \,|\, \big(s^{\overline{0}}_{p,q}f(\nabla)\big)' \|\,. \end{aligned}$$

We only note $1 < q' \leq \infty$. This proves $x \in s^{\overline{0}}_{p',q'}f(\nabla)$, and as p and q are finite by the density of the finite sequences in $s^{\overline{0}}_{p,q}f(\nabla)$ we easily find $Tx = y$.
Finally, the case of general $\overline{r}$ is traced back to the case $\overline{r} = \overline{0}$ by lifting arguments analog to the proof of Lemma 5.5.2. □

Remark 5.5.1. Since we proved the (absolute) convergence of the representations (5.5.3) and (5.5.5) we can show that these two operators are identical using the density of finite sequences and Lebesgue's theorem on dominated convergence.

6 Nonlinear approximation

The approximation of "complicated" functions by "easier" ones is a classical topic in pure and applied analysis. Roughly this can be divided in linear and nonlinear methods. Linear methods can be described as linear operators, and the approximative powers are most often measured in certain operator norms.

This chapter is devoted to the study of one particular type of nonlinear approximation, namely so-called m-term approximation. Nonlinearity means the approximant may depend in a nonlinear or even discontinuous way on the given function (or sequence). The precise definitions of this method and several related notions are given in Section 6.1. As in the last chapter our considerations are done for sequence spaces, since these provide some a priori simplifications of the problem.

When dealing with m-term approximation there are two quite different objectives. On the one hand one is interested in the (asymptotic) behaviour of the error, preferably in terms of parameters which characterize the (classes of) functions or sequences which are to be approximated as well as the quasi-norm in which the error is measured. On the other hand one looks for explicit constructions of near best approximants, i.e. approximants which realize the best possible error up to some constant factor. Typically not for all possible parameters both aims can be accomplished. Therefore our strategy consists in providing explicit construction for a certain range of parameters (see Section 6.7), and afterwards the obtained results for asymptotics are extended. The main tools in this context are approximation spaces (Section 6.3) and real interpolation (Section 6.9). In this way the asymptotic error behaviour for almost all possible parameter constellations are established, see Section 6.10.

6.1 General definitions

Definition 6.1.1. A dictionary $\mathcal{D}$ is a countable subset of a complex-valued quasi-normed space X, whose linear span is dense in X. For such sets $\mathcal{D} = \{h_1, h_2, \ldots\}$ we define

$$\Sigma_m = \Sigma_m(\mathcal{D}) = \left\{ \sum_{j\in\Lambda} c_j h_j \, : \, \Lambda \subset \mathbb{N}\, , \#\Lambda \le m\, , \; c_j \in \mathbb{C}\, , j \in \Lambda \right\}.$$

Remark 6.1.1. In general, the sets Σ_m are no linear sets. More precisely, it holds $\Sigma_m + \Sigma_m = \Sigma_{2m}$, if the dictionary consists of at least $2m$ linearly independent vectors. Furthermore, we obtain the representation

$$\Sigma_m = \bigcup_{i_1,\ldots,i_m} \operatorname{span}\{h_{i_1}, \ldots, h_{i_m}\}\, .$$

Hence Σ_m is the union over all subspaces spanned by at most m vectors from the dictionary Φ.

Since we are mainly interested in the asymptotical error of approximation methods, we henceforth assume that the considered quasi-Banach spaces are infinite-dimensional, and Φ consists of infinitely many linearly independent vectors.

Our main interest lies in approximation from such sets Σ_m. This is called m-term approximation. The quantities defined next measure the error of this approximation procedure.

Definition 6.1.2. Let X and Y be two quasi-normed spaces, and let $\mathcal{D} \subset X$ be a dictionary. Then the quantity

$$\sigma_m(a,Y) \equiv \sigma_m(a)_Y \equiv \sigma_m(a,\mathcal{D})_Y := \inf\Big\{ \| a - u \,|Y\| \;:\; u \in \Sigma_m \Big\}$$

is called the error of the best m-term approximation of $a \in X$ with respect to the quasi-norm of the space Y. Moreover, we define the m-term width of X and Y with respect to the dictionary $\mathcal{D}$ by

$$\sigma_m(X,Y) \equiv \sigma_m(X,Y;\mathcal{D}) := \sup\Big\{ \sigma_m(a)_Y : \| a \,|X\| \le 1 \Big\}.$$

Though the assumed density of $\operatorname{span}\mathcal{D}$ is not necessary in the above definition it ensures that for every $a \in X$ we have $\sigma_m(a)_X \longrightarrow 0$ for $m \longrightarrow \infty$. Our aim is to estimate the rate of convergence of $\sigma_m(a)_Y$ in terms of properties of the spaces X and Y.

Remark 6.1.2. Since the sets Σ_m are nonlinear, this approximation method is a special case of nonlinear approximation. However, we will compare our results to linear approximation widths, in particular to linear widths as in the next definition.

Definition 6.1.3. Let X and Y be two quasi-normed spaces, and let $T \in \mathcal{L}(X,Y)$ be a bounded linear operator from X to Y. Then we put

$$a_n(T) := \inf\Big\{ \| T - A \,|\mathcal{L}(X,Y)\| \;:\; \operatorname{rank}(A) < n \Big\}.$$

$a_n(T)$ is called n-th approximation number of T.

Of particular interest are embedding operators. In that case the approximation numbers are often referred to as linear widths. Together with similar approximation quantities they describe important properties of the embeddings and the spaces itself.

Remark 6.1.3. For a first comparison of the defined linear and nonlinear approximation methods may serve the following reformulation of the m-term approximation:

$$\sigma_m(a)_Y = \inf\Big\{ \| a - A(a) \,|Y\| \;\Big|\; A : X \longrightarrow \Sigma_m \Big\}.$$

In contrast to the situation for approximation numbers here arbitrary (i.e. nonlinear, and even non-continuous) mappings A are allowed. Another difference between a_n and σ_m is given by the order of infimum and supremum: For a_n the supremum over the unit ball of X is taken first, and only then the infimum over all operators, conversely for σ_m. Hence the optimal approximation algorithm for σ_m may be a different one for every $a \in X$.

The following lemma will be quite useful in the sequel. We will use it mostly without explicitly mentioning it.

Lemma 6.1.1. Let X, Y be quasi-normed spaces.

(i) The quantity $\sigma_m(a)_Y$ is homogeneous, i.e. it holds

$$\sigma_m(\lambda a)_Y = |\lambda| \sigma_m(a)_Y \,, \quad a \in X \,, \quad \lambda \in \mathbb{C} \,.$$

(ii) Let X_0 be a further quasi-normed space such that $X \hookrightarrow X_0$. Then it holds

$$\sigma_m(X, Y) \leq \big\| \mathrm{id} : X \longrightarrow X_0 \,\big|\mathcal{L}(X, X_0)\big\| \; \sigma_m(X_0, Y) \,.$$

(iii) Let Y_0 be a quasi-normed space such that $Y_0 \hookrightarrow Y$. Moreover, let $\Phi \subset Y_0$ be a dictionary for both Y_0 and Y. Then it holds

$$\sigma_m(X, Y) \leq \big\| \mathrm{id} : Y_0 \longrightarrow Y \,\big|\mathcal{L}(Y_0, Y)\big\| \; \sigma_m(X, Y_0) \,.$$

Proof. All assertions are immediate corollaries of the respective definitions. In particular, we have $\lambda \Sigma_m = \Sigma_m$ for all such sets Σ_m and all $\lambda \in \mathbb{C}$. □

6.2 m-term approximation in sequence spaces: Preliminary remarks

In the sequel the spaces X and Y will be quasi-Banach sequence spaces of either (vector-valued) ℓ_p-type or of $s^{\bar{r}}_{p,q}a$-type. The dictionary $\mathcal{D}$ will always be some set $\mathcal{B} \equiv \mathcal{B}_I$,

$$\mathcal{B}_I = \Big\{ e^i \,:\, i \in I \Big\} \,, \qquad (e^i)_j = \delta_{i,j} \,, \quad i, j \in I \,,$$

i.e. $\mathcal{B}_I$ is the set of canonical sequences with respect to some suitable countable index set I. For example, in case of sequence spaces $s^{\bar{r}}_{p,q}a$ we have $I = \mathbb{N}_0^N \times \mathbb{Z}^d$.
In principle, the calculation of an optimal m-term approximation consists in two tasks: On the one hand, one has to determine the right elements of the dictionary, from which the approximant is formed, and on the other hand one has to determine the corresponding coefficients. However, due to special properties of the quasi-norms in the sequence spaces under consideration, one part of the problem is quite easy to solve.

Lemma 6.2.1. Let X be a sequence space either of ℓ_p-, Besov- or Triebel-Lizorkin-type.

(i) The space X is a Banach lattice, i.e. for any two sequences a and b with $b \in X$ and $|a| \leq |b|$ (componentwise) it follows $a \in X$ and $\|a|X\| \leq \|b|X\|$.

(ii) For every sequence $a = (a_j)_{j \in \mathcal{J}} \in X$ it holds

$$\sigma_m(a, X, \mathcal{B}_{\mathcal{J}}) = \inf \Big\{ \Big\| a - \sum_{j \in \Lambda} a_j e^j \,\Big|\, X \Big\| \,:\, \Lambda \subset \mathcal{J} \,,\; \#\Lambda \leq m \Big\} .$$

So far when considering $\sigma_m(X, Y)$, the spaces X and Y were allowed to be arbitrary, as long as they possessed a common dictionary $\mathcal{D}$. However, since we are interested in its decay rate we get a priori restrictions on the parameters for the sequence spaces $s^{\bar{r}}_{p,q}a$ just by excluding the cases, where the m-term widths are infinite. To this end the following lemma is helpful.

Lemma 6.2.2. Let $0 < p_0, p_1, q_0, q_1 \le \infty$ ($p_0, p_1 < \infty$ for f-spaces), and let $\overline{r}, \overline{s} \in \mathbb{R}^N$. Then the quantity $\sigma_m\big(s^{\overline{r}}_{p_0,q_0}a(\nabla)\,,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big)$ is finite if, and only if, we have a continuous embedding $s^{\overline{r}}_{p_0,q_0}a(\nabla) \hookrightarrow s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)$, $a, a^\dagger \in \{b, f\}$.

Proof. Since it trivially holds

$$\sigma_m\big(\lambda\,,s^{\overline{s}}_{p_1,q_1}a(\nabla)\,,\mathcal{B}\big) \le \big\|\,\lambda\,\big|s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\big\|\,, \qquad \lambda \in s^{\overline{r}}_{p_0,q_0}a(\nabla)\,,$$

the sufficiency of a continuous embedding is immediately clear. The necessity follows from the observation, that for these sequence spaces a continuous embedding holds if, and only if, it holds set theoretic inclusion. Hence, if we assume $s^{\overline{r}}_{p_0,q_0}a(\nabla) \not\hookrightarrow s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)$ there exists a sequence $\lambda \in s^{\overline{r}}_{p_0,q_0}a(\nabla)$, such that $\big\|\,\lambda\,\big|s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\big\| = \infty$. But for this sequence it immediately follows from the triangle inequality that $\sigma_m(\lambda\,,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}) = \infty$ and thus also $\sigma_m\big(s^{\overline{r}}_{p_0,q_0}a(\nabla)\,,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big) = \infty$. □

As a consequence of this lemma, we will always assume restrictions on the parameters which ensure a continuous embedding, according to the results of the last sections.

The following lemma will considerably simplify the following calculations, since it allows a reduction of parameters. The property described therein, the behaviour of the m-term width in connection with the lifting operator, is the counterpart of well-known assertions in the theory of s-numbers, see e.g. the monograph of Pietsch [61].

Lemma 6.2.3. Let $0 < p_0, p_1, q_0, q_1 \le \infty$ ($p_0, p_1 < \infty$ for f-spaces), and let $\overline{r}, \overline{s}, \overline{t} \in \mathbb{R}^N$. Then it holds for every combination of $a, a^\dagger \in \{b, f\}$

$$\sigma_m\big(s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)\,,s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big) = \sigma_m\big(s^{\overline{s}+\overline{t}}_{p_0,q_0}a(\nabla)\,,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big)$$

for all $m \in \mathbb{N}_0$.

In other words, the behaviour of the (nonlinear) m-term approximation depends only on the difference of the smoothness vectors. Of course, for these quantities to be finite the condition $\overline{t} \ge \overline{d}\big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+$ is necessary.

Proof. According to Proposition 5.2.1 the lifting operator $L_{\overline{s}}$ is an isometry from $s^{\overline{r}}_{p,q}a(\nabla)$ onto $s^{\overline{r}-\overline{s}}_{p,q}a(\nabla)$ for all parameters p and q. In particular, it follows

$$\big\|\,\lambda\,\big|s^{\overline{s}+\overline{t}}_{p_0,q_0}a(\nabla)\big\| = \big\|L_{\overline{s}-\overline{r}}\lambda\,\big|s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)\big\|\,,$$

hence $L_{\overline{s}-\overline{r}}$ maps the unit ball of $s^{\overline{s}+\overline{t}}_{p_0,q_0}a(\nabla)$ onto the unit ball of $s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)$. Similarly we have

$$\big\|\,\lambda - S_m\lambda\,\big|s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\big\| = \big\|L_{\overline{s}-\overline{r}}\big(\lambda - S_m\lambda\big)\big|s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\big\| = \big\|L_{\overline{s}-\overline{r}}\lambda - \widetilde{S_m\lambda}\,\big|s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\big\|$$

for every approximation $S_m\lambda \in \Sigma_m$. Moreover, we obviously have $L_{\overline{s}-\overline{r}}\Sigma_m = \Sigma_m$, i.e. $\widetilde{S_m\lambda} = L_{\overline{s}-\overline{r}}(S_m\lambda) \in \Sigma_m$. This yields

$$\sigma_m\big(\lambda\,,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big) = \sigma_m\big(L_{\overline{s}-\overline{r}}\lambda\,,s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big)\,. \tag{6.2.1}$$

Altogether we conclude

$$
\begin{aligned}
&\sup\Big\{\sigma_m\big(\lambda\,,\,s^{\overline{s}}_{p_1,q_1}a^\dagger(\nabla)\big)\;:\;\big\|\,\lambda\,\big|s^{\overline{s}+\overline{t}}_{p_0,q_0}a(\nabla)\big\|\le 1\Big\}\\
&\qquad=\sup\Big\{\sigma_m\big(L_{\overline{s}-\overline{r}}\lambda\,,\,s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big)\;:\;\big\|\,L_{\overline{s}-\overline{r}}\lambda\,\big|s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)\big\|\le 1\Big\}\\
&\qquad=\sup\Big\{\sigma_m\big(\lambda\,,\,s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)\,,\mathcal{B}\big)\;:\;\big\|\,\lambda\,\big|s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)\big\|\le 1\Big\}\,,
\end{aligned}
$$

which completes the proof. □

According to this lemma, when considering the embedding $s^{\overline{r}+\overline{t}}_{p_0,q_0}a(\nabla)\hookrightarrow s^{\overline{r}}_{p_1,q_1}a^\dagger(\nabla)$ it is sufficient to consider the case $\overline{r}=\overline{0}$. This will be done in the subsequent sections.
Finally, since geometric series play an important role in the upcoming calculations, we shall state the following lemma on estimates for polynomially perturbed geometric series. It is proved by standard straightforward calculations, hence we omit the details.

Lemma 6.2.4. Let $\alpha,\beta\in\mathbb{R}$ and $L\in\mathbb{N}$, where $\alpha>0$. Then it holds

$$
\sum_{j=L}^{\infty} j^\beta 2^{-j\alpha}\sim L^\beta 2^{-L\alpha}\,,
$$

with equivalence constants independent of L.

6.3 Approximation spaces

Approximation spaces are a well-known tool in approximation theory. They proved particularly useful in connection with several types of approximation methods including m-term approximation. Though they can be defined in the far more general context of approximation schemes, we are only interested in the approximation spaces relative to σ_m.
Originally approximation spaces were introduced in the framework of interpolation theory, see [11, 59]. The formulation which we will use in the sequel is due to Pietsch [60]. However, when dealing with spaces of dominating mixed smoothness, the spaces $\mathcal{A}^s_q(X,\mathcal{D})$ defined there are not general enough, as they don't reflect the behaviour of logarithmic terms. Recently, Luther and Almira worked on generalizations of these approximation spaces, which led to the notion of *generalized approximation spaces*, cf. [52, 53]. In our present situation, we only need one particular case of these generalizations.

Definition 6.3.1. Let $s,t\in\mathbb{R}$ and $0<u\le\infty$, where $s>0$. We define spaces $\mathcal{A}^{s,t}_u(X,\mathcal{D})$ to be the collection of all elements $a\in X$, such that

$$
\big\|\,a\,\big|\mathcal{A}^{s,t}_u(X,\mathcal{D})\big\|:=\begin{cases}\Big(\sum\limits_{m=1}^{\infty}\Big[m^s\big(1+\log m\big)^t\,\sigma_{m-1}\big(a,X,\mathcal{D}\big)\Big]^u\dfrac{1}{m}\Big)^{1/u} & \text{if}\quad 0<u<\infty\,,\\ \sup\limits_{m=1,2,\dots}\; m^s\big(1+\log m\big)^t\,\sigma_{m-1}\big(a,X,\mathcal{D}\big) & \text{if}\quad u=\infty\,,\end{cases}
$$

is finite. If we put $t=0$ in this definition, we re-obtain the classical approximation spaces, which will be denoted by $\mathcal{A}^s_u(X,\mathcal{D})$.

In [52, 53] most properties of these approximation spaces can be found. At this point we shall collect some of the basic properties which will be of use in the sequel. For proofs, we refer to the mentioned literature.
We begin with some almost trivial facts. Since $\sigma_0(a, X, \mathcal{D}) = \| a \,|X\|$, the functionals $\| \cdot |\mathcal{A}^{s,t}_q(X, \mathcal{D})\|$ are quasi-norms, and the spaces $\mathcal{A}^{s,t}_q(X, \mathcal{D})$ become quasi-Banach spaces embedded in X. Moreover, for $0 < u_0 \le u_1 \le \infty$ we have the embedding

$$\mathcal{A}^{s,t}_{u_0}(X, \mathcal{D}) \hookrightarrow \mathcal{A}^{s,t}_{u_1}(X, \mathcal{D}) \,.$$

Furthermore, if X and Y are two quasi-Banach spaces with $\mathcal{D} \subset X$ and $X \hookrightarrow Y$, then it holds

$$\mathcal{A}^{s,t}_{u}(X, \mathcal{D}) \hookrightarrow \mathcal{A}^{s,t}_{u}(Y, \mathcal{D}) \tag{6.3.1}$$

for all admissible parameters. The following lemma is the main reason for our interest in approximation spaces.

Lemma 6.3.1. Let $Y \hookrightarrow X$ be quasi-Banach spaces, and let $\mathcal{D} \subset X$. Moreover, let $s > 0$ and $t \in \mathbb{R}$. Then the following assertions are equivalent:

(i) The Jackson-type inequality

$$\sigma_m(f, X, \mathcal{D}) \le c\, m^{-s} \big(\log m\big)^{-t} \| \, f \, | \, Y \, \|$$

holds with some constant $c > 0$ independent of $f \in Y$ and $m \ge 2$.

(ii) The m-term width satisfies

$$\sigma_m\big(Y, X, \mathcal{D}\big) \le c\, m^{-s} \big(\log m\big)^{-t}$$

with some constant $c > 0$ independent of $m \ge 2$.

(iii) We have the continuous embedding

$$Y \hookrightarrow \mathcal{A}^{s,t}_\infty\big(X, \mathcal{D}\big) \,.$$

In other words, this lemma allows us to reformulate estimates from above for the error of the best m-term approximation as embedding problems for quasi-Banach spaces. This reformulation is particularly useful, because the following proposition shows that the above approximation spaces are compatible with real interpolation. For the basics in real interpolation we refer to [7, 82].

Proposition 6.3.1.

(i) Let $0 < u, q \le \infty$, $s > 0$, $t \in \mathbb{R}$ and $0 < \Theta < 1$. Let X be a quasi-Banach space and $\mathcal{D}$ a subset of X. Then it holds

$$\big(X, \mathcal{A}^{s,t}_u(X, \mathcal{D})\big)_{\Theta,q} = \mathcal{A}^{\Theta s, \Theta t}_q(X, \mathcal{D}) \,. \tag{6.3.2}$$

(ii) Let $0 < u, u_0, u_1 \le \infty$ and $0 < \Theta < 1$. Further we assume $s_0, s_1 > 0$, $s_0 \ne s_1$, and $t_0, t_1 \in \mathbb{R}$. Let X be a quasi-Banach space and $\mathcal{D}$ a subset of X. Then, with $s := (1 - \Theta)\, s_0 + \Theta\, s_1$ and $t := (1 - \Theta)\, t_0 + \Theta\, t_1$, it holds

$$\big(\mathcal{A}^{s_0,t_0}_{u_0}(X, \mathcal{D}), \mathcal{A}^{s_1,t_1}_{u_1}(X, \mathcal{D})\big)_{\Theta,u} = \mathcal{A}^{s,t}_u(X, \mathcal{D}) \,. \tag{6.3.3}$$

The counterparts of the above identities for classical approximation spaces are well-known. Proves for those counterparts can be found in [18, 17] and [59, 11], respectively. A proof of (a more general version of) the above proposition can be found in Luther [52].
We shall cite one last result. As one can easily check, we have $\mathcal{D} \subset \mathcal{A}_u^s(X, \mathcal{D})$ for all parameters s and u. Hence, it makes sense to consider approximation spaces $\mathcal{A}_v^t\big(\mathcal{A}_u^s(X, \mathcal{D}), \mathcal{D}\big)$, and the respective analogue for generalized approximation spaces.

Proposition 6.3.2. Let $0 < u, v \leq \infty$, $s_1, s_2 > 0$, and $t_1, t_2 \in \mathbb{R}$. Let X be a quasi-Banach space and $\mathcal{D}$ a fixed subset of X. Then we have

$$\mathcal{A}_u^{s_1,t_1}\big(\mathcal{A}_v^{s_2,t_2}(X, \mathcal{D}), \mathcal{D}\big) = \mathcal{A}_u^{s_1+s_2,t_1+t_2}(X, \mathcal{D}) \tag{6.3.4}$$

in the sense of equivalent quasi-norms.

This proposition, known as the *Reiteration theorem*, can be found in its classical version in [60, Section 3.2]. The general assertion (6.3.4) is due to Luther [52].

Remark 6.3.1. The idea of the Theorem is quite simple: If one can't calculate the approximation error of $f \in Y$ in the norm of X directly, then find some space Z between Y and X, i.e. $Y \hookrightarrow Z \hookrightarrow X$. Afterwards, the approximation consists of two steps: First approximate f by $S_m f$ in the norm of Z, and then approximate $f - S_m f \in Z$ in the norm of the target space X. Of course, this procedure can be iterated involving several intermediate spaces.

6.4 Building blocks

In Section 5.4 we decomposed sequences λ as in (5.1.1) according to

$$\lambda = \sum_{\overline{\nu} \in \mathbb{N}_0^N} \mathrm{re}_{\overline{\nu}}\, \lambda \qquad \text{(convergence at least componentwise)}\,.$$

This was used to prove compactness of embeddings. Now we shall consider another decomposition,

$$\lambda = \sum_{\mu=0}^{\infty} \mathrm{re}_\mu\, \lambda = \sum_{\mu=0}^{\infty} \sum_{\overline{\nu}\,:\,\overline{\nu}\cdot\overline{d}=\mu} \mathrm{re}_{\overline{\nu}}\, \lambda\,. \tag{6.4.1}$$

This section will be devoted to the study of properties of the modified restriction operators re_μ, $\mu \in \mathbb{N}_0$, and their images, the *building blocks* $\mathrm{re}_\mu\, \lambda$. We begin with a number theoretic result.

Lemma 6.4.1. For every number $\mu \in \mathbb{N}_0$ we put

$$M\big(\mu, \overline{d}\big) := \Big\{\overline{\nu} \in \mathbb{N}_0^N \ :\ \mu = \overline{\nu}\cdot\overline{d}\Big\} \qquad \text{and} \qquad S\big(\mu, \overline{d}\big) := \# M\big(\mu, \overline{d}\big)\,.$$

Then it holds

$$S\big(\mu, \overline{d}\big) \leq c_1\, \mu^{N-1}\,.$$

Moreover, for every $\mu \in d_1\mathbb{N}_0$ we find

$$S\big(\mu, \overline{d}\big) \geq c_2\, \mu^{N-1}\,.$$

Proof. The estimate from above follows from the trivial observation that for every solution $\overline{\nu}$ of the equation $\mu = \overline{\nu} \cdot \overline{d}$ we have $0 \leq \nu_i \leq \mu$ for $i = 1, \ldots, N-1$, and for fixed $\nu_1, \ldots, \nu_{N-1}$ there is only at most one choice possible for ν_N. Hence $S\big(\mu, \overline{d}\big) \leq (\mu+1)^{N-1}$.
The estimate from below follows by an induction argument with respect to N. For the induction basis $N = 1$ we have to consider the equation $\mu = \nu_1 d_1$. Clearly, this one is solvable if, and only if, μ is divisible by d_1, which is ensured by the assumption $\mu \in d_1\mathbb{N}$. It follows $S(\mu, d_1) = 1 = \mu^0$, as we have claimed.
The induction step follows from the observation

$$\mu - \nu_{N+1}d_{N+1} = \nu_1 d_1 + \cdots + \nu_N d_N \,,$$

which yields the formula

$$S\big(\mu, (\overline{d}, d_{N+1})\big) = \sum_{j=0}^{[\mu/d_{N+1}]} S\big(\mu - j d_{N+1}, \overline{d}\big) \,.$$

In other words, we sum over the counts of the solutions of corresponding N-dimensional equations, and the index runs over all possible values for ν_{N+1}, since we have to keep in mind $\nu_1 d_1 + \cdots + \nu_N d_N \geq 0$. From this, more precise estimates from above could be derived, but we will concentrate on estimates from below. Hence we suppose $S\big(\mu, \overline{d}\big) \gtrsim \mu^{N-1}$ for all $\mu \in d_1\mathbb{N}_0$ to further obtain

$$\begin{aligned} S\big(\mu, (\overline{d}, d_{N+1})\big) &= \sum_{j=0}^{[\mu/d_{N+1}]} S\big(\mu - j d_{N+1}, \overline{d}\big) \geq \sum_{\substack{0 \leq j \leq [\mu/d_{N+1}],\\ d_1 | j}} S\big(\mu - j d_{N+1}, \overline{d}\big) \\ &= \sum_{j=0}^{[\mu/d_1 d_{N+1}]} S\big(\mu - j d_1 d_{N+1}, \overline{d}\big) \gtrsim \sum_{j=0}^{[\mu/d_1 d_{N+1}]} (\mu - j d_1 d_{N+1})^{N-1} \\ &\sim \sum_{j=0}^{[\mu/d_1 d_{N+1}]} (j d_1 d_{N+1})^{N-1} \sim \int_0^\mu x^{N-1} dx \sim \mu^N. \end{aligned}$$

We used that due to $d_1 | \mu$ we have also $d_1 | (\mu - j d_1 d_{N+1})$, thus it follows $\mu - j d_1 d_{N+1} \in d_1\mathbb{N}_0$ for all $0 \leq j \leq [\mu/d_1 d_{N+1}]$. □

This lemma corresponds to the well-known result in the case $\overline{d} = \overline{1}$, i.e. $N = d$. Then it holds

$$S\big(\mu, \overline{d}\big) \equiv S(\mu, d) = \binom{\mu + d - 1}{\mu} \sim \mu^{d-1} \quad \text{for all} \quad \mu \in \mathbb{N} \,,$$

which can be proved using the same induction argument as above.

Remark 6.4.1. Since the solvability and the number of solutions of the equation $\mu = \overline{\nu} \cdot \overline{d}$ obviously are independent from the numeration of the d_i, the assumption $d_1 | \mu$ for the estimate from below can be weakened to $d_i | \mu$ for some $i \in \{1, \ldots, N\}$, which can be reformulated as $\mu \in \bigcup_{i=1}^N d_i\mathbb{N}_0 =: \mathcal{N}$. In particular, if $d_i = 1$ for some index i, then we already obtain $S\big(\mu, \overline{d}\big) \sim \mu^{N-1}$ for all $\mu \in \mathbb{N}$.
Of further interest will also be the special case $\mu \in d_0\mathbb{N}_0$, where $d_0 = \operatorname{lcm}(d_1, \ldots, d_N)$. Then it holds $\frac{\mu}{d_i}\overline{e}_i \in M\big(\mu, \overline{d}\big)$ for all $i = 1, \ldots, N$.
However, as one of the main corollaries of all these contemplations we conclude that the set $M\big(\mu, \overline{d}\big)$ is non-empty for sufficiently many μ.

Remark 6.4.2. Since only solutions $\overline{\nu} \in \mathbb{N}_0^N$ of the equation $\mu = \overline{\nu} \cdot \overline{d}$ are of interest to us, the well-known condition $\gcd(d_1, \dots, d_N)|\mu$ for the solvability of diophantic equations cannot be applied. In particular, even in the case $\gcd(d_1, \dots, d_N) = 1$ it does not automatically follow $M(\mu, \overline{d}) \neq \emptyset$ for all $\mu \in \mathbb{N}$ (of course, in this situation solutions in $\mathbb{Z}^N$ do always exist).

As a next step we collect some properties of the building blocks. The following lemma can be seen as a counterpart of Lemma 5.4.1.

Lemma 6.4.2. Consider sequences η as in (5.1.1), and let $\nabla = \nabla(\Omega)$ for some bounded domain $\Omega \subset \mathbb{R}^d$. We define spaces

$$s^{\overline{r}}_{p,q} a_\mu(\Omega) = \Big\{ \eta \in s^{\overline{r}}_{p,q} a(\Omega) \,:\, \eta_{\overline{\nu},m} = 0 \quad \text{for all} \quad \overline{\nu} \notin M(\mu, \overline{d}) \Big\},$$

where $0 < p, q \leq \infty$ ($p < \infty$ for f-spaces), $\overline{r} \in \mathbb{R}^N$, and $\overline{\nu} \in \mathbb{N}_0^N$.

(i) The operator re_μ is the identical mapping on $s^{\overline{r}}_{p,q} a_\mu(\Omega)$ for all admissible parameters.

(ii) For $0 < p, q \leq \infty$ the spaces $s^{\overline{r}}_{p,q} b_\mu(\Omega)$ can be represented as

$$s^{\overline{r}}_{p,q} b_\mu(\Omega) = L^{-1}_{\overline{r} - \overline{d}/p} \Big(\ell_q\big(M(\mu, \overline{d}), \ell_p(\nabla)\big) \Big).$$

Moreover, we have for $0 < p < \infty$

$$s^{\overline{r}}_{p,p} b_\mu(\Omega) = s^{\overline{r}}_{p,p} f_\mu(\Omega) = L^{-1}_{\overline{r} - \overline{d}/p} \Big(\ell_p\big(M(\mu, \overline{d}) \times \nabla\big) \Big).$$

(iii) For $0 < p_0, p_1 \leq \infty$ and $0 < q \leq \infty$ it holds

$$\Big\| \mathrm{re} \,:\, s^{\overline{r}}_{p_0,q} a_\mu(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q} a_\mu(\Omega) \Big\| \sim 2^{\mu(1/p_0 - 1/p_1)_+}.$$

Here the equivalence constants are independent of $\mu \in \mathbb{N}_0$.

(iv) Let $0 < p \leq \infty$ and $0 < q_0, q_1 \leq \infty$. Then we find for every $\mu \in \mathbb{N}$

$$\Big\| \mathrm{re} \,:\, s^{\overline{r}}_{p,q_0} a_\mu(\Omega) \longrightarrow s^{\overline{r}}_{p,q_1} a_\mu(\Omega) \Big\| = S\big(\mu, \overline{d}\big)^{(1/q_1 - 1/q_0)_+}.$$

We shall also use the notations

$$\nabla_\mu = M(\mu, \overline{d}) \times \nabla = \Big\{ (\overline{\nu}, k) \,:\, \overline{\nu} \in M(\mu, \overline{d}), k \in \nabla_{\overline{\nu}} \Big\},$$

as well as $D_{\overline{\nu}} = \#\nabla_{\overline{\nu}}$ and $D_\mu = \#\nabla_\mu = \dim s^{\overline{r}}_{p,q} a_\mu(\Omega)$.

Proof. For b-spaces the assertions (iii) and (iv) follow immediately from the representation in (ii), the known fact

$$\big\| \mathrm{id} \,:\, \ell^M_{p_0} \longrightarrow \ell^M_{p_1} \big\| = \begin{cases} 1, & 0 < p_0 \leq p_1 \leq \infty\,, \\ M^{1/p_1 - 1/p_0}, & 0 < p_1 < p_0 \leq \infty\,, \end{cases}$$

as well as $D_{\overline{\nu}} \sim 2^{\overline{\nu}\cdot\overline{d}} = 2^{\mu}$ for every $\overline{\nu} \in M\big(\mu, \overline{d}\big)$.
In the f-case we obtain (iv) in the same way, keeping in mind

$$\Bigg(\sum_{\overline{\nu}\in M(\mu,\overline{d})}\sum_{m\in\nabla_{\overline{\nu}}} 2^{\overline{\nu}\cdot\overline{r}q}|\eta_{\overline{\nu},m}|^q\mathcal{X}_{\overline{\nu},m}(\,\cdot\,)\Bigg)^{1/q} = \Bigg(\sum_{\overline{\nu}\in M(\mu,\overline{d})} 2^{\overline{\nu}\cdot\overline{r}q}\Big(\sum_{m\in\nabla_{\overline{\nu}}}|\eta_{\overline{\nu},m}|\mathcal{X}_{\overline{\nu},m}(\,\cdot\,)\Big)^q\Bigg)^{1/q},$$

since the supports of the functions $\mathcal{X}_{\overline{\nu},m}$ are pairwise disjoint for fixed $\overline{\nu} \in \mathbb{N}_0^N$.
For the proof of (iii) in the case $p_1 \le p_0$ we use the boundedness of Ω and Hölder's inequality with respect to $\frac{p_1}{p_0} + \frac{1}{s} = 1$. We find

$$\begin{aligned}
\big\|\,\eta\,\big|s^{\overline{r}}_{p_1,q}f_\mu(\Omega)\big\| &= \Bigg\|\Bigg(\sum_{\overline{\nu}\in M(\mu,\overline{d})}\sum_{m\in\nabla_{\overline{\nu}}} 2^{\overline{\nu}\cdot\overline{r}q}|\lambda_{\overline{\nu},m}|^q\mathcal{X}_{\overline{\nu},m}(\,\cdot\,)\Bigg)^{1/q}\Bigg|L_{p_1}(\mathbb{R}^d)\Bigg\| \\
&\le |\Gamma_2|^{\frac{1}{sp_1}}\Bigg\|\Bigg(\sum_{\overline{\nu}\in M(\mu,\overline{d})}\sum_{m\in\nabla_{\overline{\nu}}} 2^{\overline{\nu}\cdot\overline{r}q}|\lambda_{\overline{\nu},m}|^q\mathcal{X}_{\overline{\nu},m}(\,\cdot\,)\Bigg)^{1/q}\Bigg|L_{p_0}(\mathbb{R}^d)\Bigg\| \\
&= c\,\big\|\,\eta\,\big|s^{\overline{r}}_{p_0,q}f_\mu(\Omega)\big\|\,,
\end{aligned}$$

where Γ_2 has the same meaning as in (5.1.4). The estimates from below follow for some fixed $\overline{\lambda} \in M\big(\mu, \overline{d}\big)$ with the help of sequences $\eta^{\overline{\lambda}}$, defined by

$$\big(\eta^{\overline{\lambda}}\big)_{\overline{\nu},m} = \begin{cases} 2^{-\overline{\nu}\cdot\overline{r}}\,, & \overline{\nu} = \overline{\lambda}\,, m \in \nabla_{\overline{\nu}}\,, \\ 0\,, & \text{else}\,. \end{cases}$$

We immediately find

$$\big\|\,\eta^{\overline{\lambda}}\,\big|s^{\overline{r}}_{p_0,q}f_\mu(\Omega)\big\| = \big\|\,\eta^{\overline{\lambda}}\,\big|s^{\overline{r}}_{p_1,q}f_\mu(\Omega)\big\| \sim |\Omega| \sim 1\,.$$

In case $p_0 < p_1$ we use the embedding from Proposition 5.3.3. The embedding operator $\mathrm{id} : s^{\overline{r}+\overline{d}/p_0-\overline{d}/p_1}_{p_0,q}f(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q}f(\Omega)$ also maps $s^{\overline{r}+\overline{d}/p_0-\overline{d}/p_1}_{p_0,q}f_\mu(\Omega)$ to $s^{\overline{r}}_{p_1,q}f_\mu(\Omega)$, and it holds

$$\Big\|\mathrm{id} : s^{\overline{r}+\overline{d}/p_0-\overline{d}/p_1}_{p_0,q}f_\mu(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q}f_\mu(\Omega)\Big\| \le \Big\|\mathrm{id} : s^{\overline{r}+\overline{d}/p_0-\overline{d}/p_1}_{p_0,q}f(\Omega) \longrightarrow s^{\overline{r}}_{p_1,q}f(\Omega)\Big\|$$

for all $\mu \in \mathbb{N}_0$. From this we obtain

$$\big\|\,\eta\,\big|s^{\overline{r}}_{p_1,q}f_\mu(\Omega)\big\| \le c\,\big\|\,\eta\,\big|s^{\overline{r}+\overline{d}/p_0-\overline{d}/p_1}_{p_0,q}f_\mu(\Omega)\big\| = c\,2^{\mu(1/p_0-1/p_1)}\big\|\,\eta^{\overline{\lambda}}\,\big|s^{\overline{r}}_{p_0,q}f_\mu(\Omega)\big\|$$

for all $\eta \in s^{\overline{r}}_{p_0,q}f_\mu(\Omega)$, which proves the estimate from above.
The corresponding estimates from below follow from $\eta^{\overline{\nu},m} = 2^{-\overline{\nu}\cdot\overline{r}}e^{\overline{\nu},m}$ for arbitrary $\overline{\nu} \in M\big(\mu, \overline{d}\big)$ and $m \in \nabla_{\overline{\nu}}$. For those sequences we obtain

$$\big\|\,\eta^{\overline{\nu},m}\,\big|s^{\overline{r}}_{p_0,q}f_\mu(\Omega)\big\| = 2^{-\overline{\nu}\cdot\overline{d}/p_0} = 2^{-\mu/p_0}\,, \qquad \big\|\,\eta^{\overline{\nu},m}\,\big|s^{\overline{r}}_{p_1,q}f_\mu(\Omega)\big\| = 2^{-\mu/p_1}\,.$$

This proves the desired assertion. □

Remark 6.4.3. The results of Lemma 6.4.2(iv) remain valid also for general sequences ∇. The same holds true for (iii) in case $p_0 \le p_1$.
We previously explained that we will not consider sequence spaces $s^{\overline{r}}_{\infty,q}f(\nabla)$. However, for later use it will be important that (iii) and (iv) remain valid (without changes in the proof) for parameters $0 < p_1 < \infty < p_0 = \infty$ and $p = \infty$, i.e. for accordingly extended definitions of the respective norms. We further mention at this point the identification $s^{\overline{r}}_{\infty,\infty}f(\nabla) = s^{\overline{r}}_{\infty,\infty}b(\nabla)$.

We are not interested in any questions concerning the convergence of the decomposition (6.4.1), apart from the obvious componentwise convergence. However, we need some further estimates connecting the quasi-norm of the building blocks with the one of the full sequence. We immediately find for $0 < q < \infty$

$$\big\| \, a \, \big| s^{\overline{r}}_{p,q} b(\nabla) \big\|^q = \sum_{\mu=0}^{\infty} \big\| \mathrm{re}_\mu a \big| s^{\overline{r}}_{p,q} b(\nabla) \big\|^q , \tag{6.4.2}$$

and similarly for $q = \infty$

$$\big\| \, a \, \big| s^{\overline{r}}_{p,\infty} b(\nabla) \big\| = \sup_{\mu \in \mathbb{N}_0} \big\| \mathrm{re}_\mu a \big| s^{\overline{r}}_{p,q} b(\nabla) \big\| . \tag{6.4.3}$$

These equations have a counterpart for f-spaces as well. It holds

$$\big\| \, a \, \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^u \leq \sum_{\mu=0}^{\infty} \big\| \mathrm{re}_\mu a \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^u , \tag{6.4.4}$$

where $u = \min(p, q)$. In case $p/q \geq 1$ this follows from the triangle inequality in $L_{p/q}(\mathbb{R}^d)$. With the functions $g_\mu = \sum_{(\overline{\nu},k) \in \nabla_\mu} |a_{\overline{\nu},k}|^q \mathcal{X}_{\overline{\nu},k}$ we find

$$\big\| \, a \, \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^q = \bigg\| \sum_{\mu=0}^{\infty} g_\mu \bigg| L_{p/q}(\mathbb{R}^d) \bigg\| \leq \sum_{\mu=0}^{\infty} \big\| \, g_\mu \, \big| L_{p/q}(\mathbb{R}^d) \big\| = \sum_{\mu=0}^{\infty} \big\| \mathrm{re}_\mu a \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^q$$

Similarly, from $p/q \leq 1$ we obtain by the monotonicity of the ℓ_q-quasi-norms

$$\big\| \, a \, \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^p = \int_{\mathbb{R}^d} \bigg(\sum_{\mu=0}^{\infty} g_\mu(x) \bigg)^{p/q} dx \leq \int_{\mathbb{R}^d} \sum_{\mu=0}^{\infty} \big(g_\mu(x) \big)^{p/q} dx = \sum_{\mu=0}^{\infty} \big\| \mathrm{re}_\mu a \big| s^{\overline{r}}_{p,q} f(\nabla) \big\|^p .$$

In particular, we have found (see also Lemma 6.2.1)

$$\big\| \, \mathrm{re}_\mu a \, \big| s^{\overline{r}}_{p,q} y(\nabla) \big\| \leq \big\| \, a \, \big| s^{\overline{r}}_{p,q} y(\nabla) \big\| , \qquad y \in \{b, f\} , \tag{6.4.5}$$

for all $\mu \in \mathbb{N}_0$.

6.5 Estimates from below

We will begin our investigation of the behaviour of m-term approximation by discussing several estimates from below, which can be applied to various combinations of spaces $s^{\overline{t}}_{p_0,q_0} a(\Omega)$ and $s^{\overline{0}}_{p_1,q_1} a^\dagger(\Omega)$.

Example 1: Define sequences $a^m := (a^m_{\overline{\nu},\lambda})_{\overline{\nu},\lambda}$, $m \in \mathbb{N}$, by

$$a^m_{\overline{\nu},\lambda} := \begin{cases} 2^{\overline{\nu} \cdot \overline{d}/p_1} , & \overline{\nu} = \overline{\nu}_m , \quad \lambda \in \Lambda_m , \\ 0 & \text{otherwise} . \end{cases} \tag{6.5.1}$$

where $\overline{\nu}_m$ is chosen such that $\# \nabla_{\overline{\nu}_m} \geq 2m$ and Λ_m is a subset of $\nabla_{\overline{\nu}_m}$ satisfying $\# \Lambda_m = 2m$. An easy calculation shows

$$\big\| \, a^m \big| s^{\overline{d}(\frac{1}{p_0} - \frac{1}{p_1})}_{p_0,q_0} b(\Omega) \big\| = \big\| \, a^m \, \big| s^{\overline{d}(\frac{1}{p_0} - \frac{1}{p_1})}_{p_0,q_0} f(\Omega) \big\| = (2m)^{1/p_0} .$$

Due to the special structure of the sequences, the best m-term approximation is easy to determine. We obtain

$$\sigma_m\big(a^m\,,s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{B}\big) = \sigma_m\big(a^m\,,s^{\overline{0}}_{p_1,q_1}f(\Omega)\,,\mathcal{B}\big) = m^{1/p_1}\,.$$

Hence we find

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}x(\Omega), s^{\overline{0}}_{p_1,q_1}y(\Omega), \mathcal{B}\Big) \geq c\, m^{-1/p_0+1/p_1}\,, \tag{6.5.2}$$

where $x, y \in \{b, f\}$ and c is independent of m.

Example 2: This time the construction is a little bit more sophisticated. Let $m \in \mathbb{N}$ be fixed. We choose a sequence of pairwise disjoint cubes $Q_{\overline{\nu}_j,k^j}$, $j = 1,\ldots,2m$, where the vectors $\overline{\nu}_j$ are pairwise distinct. Now define sequences $b^m = (b^m_{\overline{\nu},\lambda})_{\overline{\nu},\lambda}$ by

$$b^m_{\overline{\nu},\lambda} = \begin{cases} 2^{\overline{\nu}_j\cdot\overline{d}/p_1}\,, & 1 \leq j \leq m\,, \quad \lambda = k^j\,,\\ 0\,, & \text{otherwise}\,. \end{cases} \tag{6.5.3}$$

Similarly b^{2m} is defined (taking the same sequence of cubes). As a consequence of this construction we get

$$\big\|\, b^m\,\big|s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\Omega)\big\| = m^{1/q_0} \qquad \text{and} \qquad \big\|\, b^m\,\big|s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}f(\Omega)\big\| = m^{1/p_0}$$

as well as

$$\big\|\, b^{2m} - b^m\,\big|s^{\overline{0}}_{p_1,q_1}f(\Omega)\big\| = m^{1/p_1} \qquad \text{and} \qquad \big\|b^{2m} - b^m\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\| = m^{1/q_1}\,.$$

Furthermore, b^m is a best m-term approximation for b^{2m}. This implies the estimates

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\Omega)\,,\ s^{\overline{0}}_{p_1,q_1}f(\Omega)\,,\ \mathcal{B}\Big) \geq\ m^{-1/q_0+1/p_1}\,, \tag{6.5.4}$$

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}f(\Omega)\,,\ s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\ \mathcal{B}\Big) \geq\ m^{-1/p_0+1/q_1}\,, \tag{6.5.5}$$

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\Omega)\,,\ s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\ \mathcal{B}\Big) \geq\ m^{-1/q_0+1/q_1}\,. \tag{6.5.6}$$

This proves the needed estimates from below for the widths σ_m related to most pairs of spaces $\big(s^{\overline{r}}_{p_0,q_0}x(\Omega), s^{\overline{s}}_{p_1,q_1}y(\Omega)\big)$ with $x, y \in \{b, f\}$ and $\overline{r} - \overline{s} = \overline{d}(\frac{1}{p_0} - \frac{1}{p_1})$. The required sequence of disjoint cubes can always be found for appropriate $\overline{\nu}_j$, which is a consequence of the standard assumption that Ω is an open set. Of course, the same constructions work for both bounded and unbounded domains, particularly for $\Omega = \mathbb{R}^d$.

When dealing with estimates for spaces associated to bounded domains we need further constructions with slightly more complicated calculations. We remind on the notations $M\big(\mu, \overline{d}\big)$, $S\big(\mu, \overline{d}\big)$, ∇_μ, D_μ, $D_{\overline{\nu}}$ and $\mathcal{N}$ introduced in Section 6.4.

Proposition 6.5.1. Let $0 < p_0, p_1, q_0, q_1 \leq \infty$ and $t \in \mathbb{R}^N$. We put $\overline{t} = t\overline{d}$. Moreover, let Ω be bounded. Then we have

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}x(\Omega)\,, s^{\overline{0}}_{p_1,q_1}y(\Omega)\,, \mathcal{B}\Big) \gtrsim m^{-t}\big(\log m\big)^{(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)_+} \tag{6.5.7}$$

for all $x, y \in \{b, f\}$, where $p_0 < \infty$ if $x = f$ and $p_1 < \infty$ if $y = f$.

Proof. **Step 1:** Due to the obvious monotonicity properties of σ_m it is sufficient to consider $m = 2^M$ for some $M \in \mathbb{N}$. Let $\overline{\nu} \in \mathbb{N}_0^N$ be some arbitrary vector, such that $\overline{\nu}\cdot\overline{d} = M$. Additionally let K be the smallest natural number such that $C_1 2^{Kd_1} \geq 2$. Then we have $\#\nabla_{\overline{\nu}+K\overline{e}^1} \geq 2m$ (here C_1 and C_2 are the constants in (5.1.3), and $\overline{e}^1 = (1,0,\ldots,0)$). Now consider the sequence

$$\alpha_M = \sum_{k\in\nabla_{\overline{\nu}+K\overline{e}^1}} \#(\nabla_{\overline{\nu}+K\overline{e}^1})^{-1/p_0} 2^{-(\overline{\nu}+K\overline{e}^1)\cdot\overline{d}(t-1/p_0)} e^{\overline{\nu}+K\overline{e}^1,k} .$$

Consequently for any $0 < q_0 \leq \infty$ we find $\big\|\alpha_M\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\| = 1$. On the other hand for some arbitrary $\Gamma \subset \nabla_{\overline{\nu}+K\overline{e}^1}$ with $\#\Gamma = m$ and for all q_1 we get

$$\begin{aligned}
&\sigma_m\big(\alpha_M\,, s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{B}\big)\\
&\quad= \Big\|\sum_{k\in\nabla_{\overline{\nu}+K\overline{e}^1}\setminus\Gamma} \#(\nabla_{\overline{\nu}+K\overline{e}^1})^{-1/p_0} 2^{-(\overline{\nu}+K\overline{e}^1)\cdot\overline{d}(t-1/p_0)} e^{\overline{\nu}+K\overline{e}^1,k}\Big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\Big\|\\
&\quad= \#(\nabla_{\overline{\nu}+K\overline{e}^1})^{-1/p_0} 2^{-(\overline{\nu}+K\overline{e}^1)\cdot\overline{d}(t-1/p_0+1/p_1)} \#(\nabla_{\overline{\nu}+K\overline{e}^1}\setminus\Gamma)^{1/p_1}\\
&\quad\geq \big(C_2 2^{M+Kd_1}\big)^{-1/p_0}\big(2^M(C_1 2^{Kd_1}-1)\big)^{1/p_1} 2^{-(\overline{\nu}+K\overline{e}^1)\cdot\overline{d}(t-1/p_0+1/p_1)}\\
&\quad\geq C_2^{-1/p_0} 2^{-Kd_1(t+1/p_1)} 2^{-Mt} = c(C_1,C_2) m^{-t} .
\end{aligned}$$

Moreover, since α_M has non-vanishing entries for only one level $\overline{\nu}$ we immediately see $\sigma_m\big(\alpha_M\,, s^{\overline{0}}_{p_1,q_1}f(\Omega)\,,\mathcal{B}\big) = \sigma_m\big(\alpha_M\,, s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{B}\big)$.

Step 2: Now let $m = \big[\frac{D_\mu}{2}\big]$ for some $\mu \in \mathcal{N}$. Then we put

$$\beta^\mu := S\big(\mu,\overline{d}\big)^{-1/q_0} \sum_{(\overline{\nu},k)\in\nabla_\mu} 2^{-\overline{\nu}\cdot\overline{d}(t-1/p_0)} D_{\overline{\nu}}^{-1/p_0} e^{\overline{\nu},k} .$$

We immediately find $\big\|\,\beta^\mu\,\big|s^{\overline{t}}_{p_0,q_0}b(\Omega)\big\| = 1$, and with the help of (5.1.4) we further conclude $\big\|\,\beta^\mu\,\big|s^{\overline{t}}_{p_0,q_0}f(\Omega)\big\| \sim 1$. On the other hand for any set $\Gamma \subset \nabla_\mu$ with $\#\Gamma = m$ we obtain with the abbreviation $\gamma_1 = \min(p_1,q_1)$

$$\begin{aligned}
&\big\|\beta^\mu - S_\Gamma\beta^\mu\big|s^{\overline{0}}_{p_1,q_1}y(\Omega)\big\|\\
&\quad= \Big\|\sum_{(\overline{\nu},k)\in\nabla_\mu\setminus\Gamma} 2^{-\overline{\nu}\cdot\overline{d}(t-1/p_0)} S\big(\mu,\overline{d}\big)^{-1/q_0} D_{\overline{\nu}}^{-1/p_0} e^{\overline{\nu},k}\Big|s^{\overline{0}}_{p_1,q_1}y(\Omega)\Big\|\\
&\quad\sim 2^{-\mu(t-1/p_0)} S\big(\mu,\overline{d}\big)^{-1/q_0} 2^{-\mu/p_0}\Big\|\sum_{(\overline{\nu},k)\in\nabla_\mu\setminus\Gamma} e^{\overline{\nu},k}\Big|s^{\overline{0}}_{p_1,q_1}y_\mu(\Omega)\Big\|\\
&\quad\gtrsim 2^{-\mu t} S\big(\mu,\overline{d}\big)^{-1/q_0} S\big(\mu,\overline{d}\big)^{-(1/\gamma_1-1/q_1)}\Big\|\sum_{(\overline{\nu},k)\in\nabla_\mu\setminus\Gamma} e^{\overline{\nu},k}\Big|s^{\overline{0}}_{\gamma_1,\gamma_1}y_\mu(\Omega)\Big\|\\
&\quad= 2^{-\mu t} S\big(\mu,\overline{d}\big)^{-1/q_0-1/\gamma_1+1/q_1} 2^{-\mu/\gamma_1}\big(\#(\nabla_\mu\setminus\Gamma)\big)^{1/\gamma_1}\\
&\quad\geq 2^{-\mu(t+1/\gamma_1)} S\big(\mu,\overline{d}\big)^{-1/q_0-1/\gamma_1+1/q_1}\big(\tfrac{1}{2}D_\mu\big)^{1/\gamma_1}\\
&\quad\gtrsim 2^{-\mu(t+1/\gamma_1)} S\big(\mu,\overline{d}\big)^{-1/q_0-1/\gamma_1+1/q_1} S\big(\mu,\overline{d}\big)^{1/\gamma_1} 2^{\mu/\gamma_1} = 2^{-\mu t} S\big(\mu,\overline{d}\big)^{-1/q_0+1/q_1} .
\end{aligned}$$

We applied the estimates from Lemma 6.4.2, parts (iii) and (iv), and the observation $s^{\overline{0}}_{\gamma_1,\gamma_1}y_\mu(\Omega) = 2^{-\mu/\gamma_1}\ell_{\gamma_1}(\nabla_\mu)$. If we further use $S\big(\mu,\overline{d}\big) \asymp \mu^{N-1}$ (Lemma 6.4.1) and $m \sim$

$S\big(\mu,\overline{d}\big)2^\mu$, i.e insert $\mu \sim \log m$ and $2^\mu \sim m(\log m)^{-(N-1)}$ into the last estimate, we obtain

$$\sigma_m\big(\beta^\mu, s^{\overline{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big) = \inf_{\Gamma\subset\nabla_\mu:\ \#\Gamma=m}\big\|\beta^m - S_\Gamma\beta^m\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\| \\ \gtrsim m^{-t}\big(\log m\big)^{(N-1)(t-1/q_0+1/q_1)}.$$

Eventually, for general m, the result again follows by monotonicity. □

Remark 6.5.1. These estimates are indeed valid for all parameters, but of course they are meaningful only when $s^{\overline{t}}_{p_0,q_0}x(\Omega) \hookrightarrow s^{\overline{0}}_{p_1,q_1}y(\Omega)$. Moreover, the same sequences may be considered for arbitrary $\overline{t} \in \mathbb{R}^N$. The outcome would be $m^{-\varrho}$ in Step 1 and $m^{-\tau}\big(\log m\big)^{(N-1)(\tau-\frac{1}{q_0}+\frac{1}{q_1})}$ in Step 2, where $\varrho = \min\big\{\frac{t_i}{d_i} : i=1,\ldots,N\big\}$ and $\tau = \max\big\{\frac{t_i}{d_i} : i=1,\ldots,N\big\}$. One only has to replace $\overline{e}^{\,1}$ in α_M by $\overline{e}^{\,i_0}$, where i_0 is some index with $\frac{t_{i_0}}{d_{i_0}} = \varrho$. However, since in that case our methods for estimates from above do not yield matching results, those estimates from below are of minor interest and thus omitted here. Finally, we add that the estimate from Step 2 remains valid for spaces $s^{\overline{r}}_{p,q}a(\widetilde{\nabla})$ for every sequence $\widetilde{\nabla}$, for which there is a sequence $\nabla = \nabla(\Omega)$ with $\Omega \subset \mathbb{R}^d$ a bounded domain, such that $\nabla_{\overline{\nu}} \subset \widetilde{\nabla}_{\overline{\nu}}$, $\overline{\nu} \in \mathbb{N}_0^N$. Particularly this applies to $\widetilde{\nabla} = \widetilde{\nabla}(\Gamma)$, where Γ is an unbounded domain, and to $\nabla_{\overline{\nu}} = \mathbb{Z}^d$.

6.6 m-term approximation for unbounded domains

In this section we will demonstrate the importance of approximation spaces to obtain results on the behaviour of $\sigma_m\big(s^{\overline{r}}_{p_0,q_0}a\,, s^{\overline{s}}_{p_1,q_1}a^\dagger\,,\mathcal{B}\big)$, where $0 < p_0 \le p_1 \le \infty$ and $\overline{r}-\overline{s} > \overline{d}\big(\frac{1}{p_0}-\frac{1}{p_1}\big)$.

To begin with, we study the problem of m-term approximation in spaces $\ell_p(I)$ first, where I is some fixed arbitrary countable index set. By $\ell_{p,u}(I)$ we denote the Lorentz sequence spaces. These are the collection of all sequences $a = (a_j)_{j\in I}$, such that

$$\big\|\,a\,\big|\ell_{p,u}(I)\big\| := \Big\|\big(n^{\frac{1}{p}-\frac{1}{u}}a_n^*\big)_{n\in\mathbb{N}}\Big|\ell_u(\mathbb{N})\Big\| < \infty, \qquad 0<p,u\le\infty\,,$$

where $a^* = (a_n^*)_n$ denotes the non-increasing rearrangement of a. Then we have the following result, which gives a complete characterization for all approximation spaces $\mathcal{A}^s_u(\ell_{p_1}(I),\mathcal{B}_I)$. The proposition goes back to Pietsch [60, Ex. 1].

Proposition 6.6.1. Let $0 < p_1, u \le \infty$. Let I be a fixed index set. Then $a \in \ell_{p_1}(I)$ belongs to the approximation space $\mathcal{A}^s_u(\ell_{p_1}(I),\mathcal{B}_I)$, if and only if $a \in \ell_{p_0,u}(I)$, where $1/p_0 := s + 1/p_1$. Furthermore,

$$\big\|\,a\,\big|\mathcal{A}^s_u\big(\ell_{p_1}(I),\mathcal{B}_I\big)\big\| \asymp \big\|\,a\,\big|\ell_{p_0,u}(I)\big\|\,, \tag{6.6.1}$$

where the constants of equivalence do not depend on I.

This proposition will be of use at various places when trying to characterize approximation spaces. We note some immediate consequences of Pietsch's result.

Corollary 6.6.1. Let $0 < p_0 \le p_1 \le \infty$ and $\overline{r} \in \mathbb{R}^N$. Then it holds

$$\mathcal{A}_{p_0}^{\frac{1}{p_0}-\frac{1}{p_1}}\Big(s_{p_1,p_1}^{\overline{r}}b(\nabla)\,,\mathcal{B}\Big) = s_{p_0,p_0}^{\overline{r}+\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}b(\nabla)$$

in the sense of equivalent quasi-norms.

Proof. The obvious identification

$$\ell_p(\mathbb{N}_0^N \times \nabla) = s_{p,p}^{\overline{d}/p}b(\nabla) = s_{p,p}^{\overline{d}/p}f(\nabla) \tag{6.6.2}$$

and Proposition 6.6.1 yield

$$\mathcal{A}_{p_0}^{\frac{1}{p_0}-\frac{1}{p_1}}\Big(s_{p_1,p_1}^{\overline{d}/p_1}b(\nabla)\,,\mathcal{B}\Big) = s_{p_0,p_0}^{\overline{d}/p_0}b(\nabla)\,.$$

The result now follows from a lifting argument, see Lemma 6.2.3 (equation (6.2.1)). □

A first application will be given in the following theorem.

Theorem 6.6.1. Let $0 < p_0 \le p_1 \le \infty$ and $0 < q_0, q_1 \le \infty$, and let $\overline{t} \in \mathbb{R}^N$, where $\overline{t} > \overline{d}\big(\frac{1}{p_0} - \frac{1}{p_1}\big)$. Then it holds

$$\sigma_m\Big(s_{p_0,q_0}^{\overline{t}}a\,,\; s_{p_1,q_1}^{\overline{0}}a^\dagger\,,\mathcal{B}\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \qquad m \ge 1\,,$$

where $a, a^\dagger \in \{b, f\}$, $p_0 < \infty$ if $a = f$ and $p_1 < \infty$ if $a^\dagger = f$.

Remark 6.6.1. We wish to draw attention to several remarkable aspects of this result. First of all, the approximation rate does not depend on $\overline{t}$, hence increasing the smoothness parameter does not improve the approximation properties. Secondly, the estimate is independent of the microscopic parameters q_0 and q_1. Lastly, we wish to emphasize that the approximation error does tend to zero, though the embedding is non-compact. Hence in this case nonlinear approximation is always superior to linear approximation. All these aspects are in sharp contrast to the compact case, i.e. approximation in spaces associated to functions on bounded domains.

Proof. We begin with the estimate from above, and define

$$\overline{\varepsilon} := \frac{1}{2}\Big(\overline{t} - \overline{d}\Big(\frac{1}{p_0} - \frac{1}{p_1}\Big)\Big) > \overline{0}\,.$$

Now we apply Corollary 6.6.1, the embedding property (6.3.1), and the embeddings in Proposition 5.3.1 and Lemma 5.3.5 to obtain

$$\begin{aligned} s_{p_0,q_0}^{\overline{t}}a \hookrightarrow s_{p_0,p_0}^{\overline{t}-\overline{\varepsilon}}a = \mathcal{A}_{p_0}^{\frac{1}{p_0}-\frac{1}{p_1}}\Big(s_{p_1,p_1}^{\overline{t}-\overline{\varepsilon}-\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}b\,,\mathcal{B}\Big) \\ = \mathcal{A}_{p_0}^{\frac{1}{p_0}-\frac{1}{p_1}}\Big(s_{p_1,p_1}^{\overline{\varepsilon}}b\,,\mathcal{B}\Big) \hookrightarrow \mathcal{A}_{p_0}^{\frac{1}{p_0}-\frac{1}{p_1}}\Big(s_{p_1,q_1}^{\overline{0}}a^\dagger\,,\mathcal{B}\Big) \end{aligned}$$

Eventually, Lemma 6.3.1 is used to transform this embedding into the desired Jackson-type inequality.

For the estimate from below, we use sequences a^m as in (6.5.1), where $\overline{\nu}_m = \overline{0}$ for all $m \in \mathbb{N}$. One easily obtains

$$\| a^m \,|s^{\overline{t}}_{p_0,q_0}a\| = (2m)^{1/p_0} \qquad \text{and} \qquad \sigma_m\big(a^m, s^{\overline{0}}_{p_1,q_1}a^\dagger, \mathcal{B}\big) = m^{1/p_1}$$

with the same arguments as given there. □

Remark 6.6.2. The used arguments clearly remain valid for spaces $s^{\overline{r}}_{p,q}a(\Omega)$ for arbitrary unbounded domains Ω. This follows immediately, since the estimate from above is valid even for arbitrary sequences ∇, and the estimate from below only uses $\#\nabla_{\overline{0}} = \infty$.

Remark 6.6.3. Instead of the above rather indirect proof of the Jackson-type inequality, one can also give an explicit construction of an approximant which yields the optimal approximation order.
Since the result is independent of q_0 and q_1, in view of Lemma 5.3.3 it is sufficient to consider $a = a^\dagger = b$. We choose some $\overline{\alpha} > 0$, such that $(1 - p_0/p_1)\overline{\alpha} < \overline{t} - \overline{d}(1/p_0 - 1/p_1)$, and define

$$\varepsilon_{\overline{\nu}} = n^{-1/p_0} 2^{-\overline{\nu}\cdot(\overline{t}-\overline{d}/p_0-\overline{\alpha})} \qquad \text{and} \qquad \Lambda_{\overline{\nu}} = \left\{k \in \nabla_{\overline{\nu}} \;:\; |\lambda_{\overline{\nu},k}| \geq \varepsilon_{\overline{\nu}}\right\},$$

where $\lambda \in s^{\overline{t}}_{p_0,q_0}b(\Omega)$, $n \in \mathbb{N}$, $\overline{\nu} \in \mathbb{N}_0^N$. Then the approximant is given by

$$S_n\lambda := \sum_{\overline{\nu}\in\mathbb{N}_0^N} \sum_{k\in\Lambda_{\overline{\nu}}} \lambda_{\overline{\nu},k} e^{\overline{\nu},k}.$$

The estimate $\sum_{\overline{\nu}\in\mathbb{N}_0^N} \#\Lambda_{\overline{\nu}} \leq c\,n$ and the error estimate now use basically the same arguments as the proof of Theorem 6.7.4 with slightly simpler calculations.

6.7 Estimates from above: Explicit constructions

In this section we consider explicit constructions for (order-optimal) m-term approximants. This will be done in two quite different situations. We begin with the limiting case for Besov-type sequences spaces and afterwards for Triebel-Lizorkin-type sequence spaces, i.e. we treat the case $\overline{t} = \overline{d}\big(\frac{1}{p_0} - \frac{1}{p_1}\big)$, $p_0 \leq p_1$. The final subsection is devoted to the case of high smoothness, i.e. we consider a compact embedding with additional assumptions on the smoothness vector $\overline{t}$.

6.7.1 The limiting case for b-spaces

As discussed in Section 5.3.2, the sequence spaces $s^{\overline{r}}_{p,q}b(\nabla)$ can be interpreted as lifted versions of the iterated sequence spaces $\ell_q(\mathbb{N}_0^N, \ell_p(\nabla))$. This formulation turns out to be useful also in connection with certain situations for m-term approximation. Hence as a first step, we investigate approximation in general spaces $\ell_q(I, \ell_p(J))$ as introduced in Definition 5.3.1. We will prove a result which is even slightly stronger than needed for our purposes.

Proposition 6.7.1. Let I be a countable index set, and let $J = (J_i)_{i\in I}$ be a family of countable index sets. Moreover, let $0 < p, q \leq \infty$ and $r > 0$. We define parameters p_r and q_r by

$$\frac{1}{p_r} = r + \frac{1}{p} \qquad \text{and} \qquad \frac{1}{q_r} = r + \frac{1}{q}. \tag{6.7.1}$$

Then we can identify the approximation spaces $\mathcal{A}^r_{q_r}\big(\ell_q(I,\ell_p(J))\,,\mathcal{B}_{I\times J}\big)$ as interpolation spaces. For every $0<\Theta<1$ and $0<u\le\infty$ it holds

$$\mathcal{A}^{\Theta r}_u\big(\ell_q(I,\ell_p(J))\,,\mathcal{B}_{I\times J}\big)=\Big(\ell_q\big(I,\ell_p(J)\big)\,,\ell_{q_r}\big(I,\ell_{p_r}(J)\big)\Big)_{\Theta,u}\,, \tag{6.7.2}$$

and for $0<q_0<q_r<q_1<q$, $0<p_0<p_r<p_1<p$ with $r_0=\frac{1}{p_0}-\frac{1}{p}=\frac{1}{q_0}-\frac{1}{q}$, $r_1=\frac{1}{p_1}-\frac{1}{p}=\frac{1}{q_1}-\frac{1}{q}$, $\frac{1}{q_r}=\frac{1-\Theta}{q_0}+\frac{\Theta}{q_1}$ and $\frac{1}{p_r}=\frac{1-\Theta}{p_0}+\frac{\Theta}{p_1}$, we find

$$\mathcal{A}^r_u\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big)=\Big(\ell_{q_0}(I,\ell_{p_0}(J)),\ell_{q_1}(I,\ell_{p_1}(J))\Big)_{\Theta,u}\,. \tag{6.7.3}$$

In particular, we obtain

$$\mathcal{A}^r_{q_r}\big(\ell_q(I,\ell_p(J))\,,\mathcal{B}_{I\times J}\big)=\ell_{q_r}\big(I,\ell_{p_r,q_r}(J)\big)\,, \tag{6.7.4}$$

where the space $\ell_{q_r}\big(I,\ell_{p_r,q_r}(J)\big)$ is defined as in Definition 5.3.1, replacing the ℓ_{p_r}-quasi-norm by a Lorentz sequence space quasi-norm ℓ_{p_r,q_r}. The identity (6.7.4) has to be understood in the sense of equivalent quasi-norms, where the equivalence constants do not depend on I or J.

Remark 6.7.1. This proposition, particularly the identity (6.7.4), is a counterpart of a result by Jawerth and Milman in [40] for sequence spaces instead of Besov spaces.

Proof. **Step 1: Jackson-type inequality.**
Under the above assumptions it holds

$$\sigma_m\Big(\ell_{q_r}\big(I,\ell_{p_r,\infty}(J)\big)\,,\ell_q\big(I,\ell_p(J)\big)\,,\mathcal{B}_{I\times J}\Big)\sim m^{-r}\,. \tag{6.7.5}$$

This result can be found in [35].

Step 2: Bernstein-type inequality.
In this step we will prove a further inequality, which is in some sense a reverse version of the above Jackson-type inequality. Let $a\in\Sigma_m(\mathcal{B}_{I\times J})$, $a=\sum_{i\in I}\sum_{j\in\Lambda_i}a_{i,j}e^{i,j}$, $\sum_{i\in I}\#\Lambda_i\le m$. Then we obtain from Hölder's inequality with respect to $1=\frac{p_r}{p}+\frac{p-p_r}{p}$ and $1=\frac{q_r}{q}+\frac{q-q_r}{q}$

$$\begin{aligned}\big\|\,a\,|\ell_{q_r}\big(I,\ell_{p_r}(J)\big)\big\| &= \Bigg(\sum_{i\in I}\Big(\sum_{j\in\Lambda_i}|a_{i,j}|^{p_r}\Big)^{q_r/p_r}\Bigg)^{1/q_r}\\ &\le \Bigg(\sum_{i\in I}\Big(\sum_{j\in\Lambda_i}|a_{i,j}|^{p}\Big)^{q_r/p}\Big(\#\Lambda_i\Big)^{q_r r}\Bigg)^{1/q_r}\\ &\le \Bigg(\sum_{i\in I}\Big(\sum_{j\in\Lambda_i}|a_{i,j}|^{p}\Big)^{q/p}\Bigg)^{1/q}\Bigg(\sum_{i\in I}\Big(\#\Lambda_i\Big)\Bigg)^{r}.\end{aligned}$$

We used the definition of p_r and q_r, in particular $\frac{p-p_r}{p}=rp_r$ and $\frac{q-q_r}{q}=rq_r$. Consequently, we have shown

$$\big\|\,a\,|\ell_{q_r}\big(I,\ell_{p_r}(J)\big)\big\|\le m^r\big\|\,a\,|\ell_q\big(I,\ell_p(J)\big)\big\|\,.$$

It is well-known, that this kind of inequality implies embeddings for approximation spaces. Either using the embedding theorem in [60, Section 3.4], or using a direct calculation similar to [45, Section 6] yields

$$\mathcal{A}^r_{\min(1,p,q)}\big(\ell_q(I,\ell_p(J)),\mathcal{B}_{I\times J}\big) \hookrightarrow \ell_{q_r}(I,\ell_{p_r}(J))\,, \tag{6.7.6}$$

since $\ell_q(I,\ell_p(J))$ is a $\min(1,p,q)$-Banach space (compare with Remark 5.1.2).

Step 3: Real interpolation.
The characterization (6.7.2) of the spaces $\mathcal{A}^r_u\big(\ell_q(I,\ell_p(J))\,,\mathcal{B}_{I\times J}\big)$ now is a consequence of Theorem 6.7.1 below.
For the proof of (6.7.3) we note that due to the assumptions on the parameters we find $r = (1-\Theta)r_0 + \Theta r_1$. Then we get from (6.7.5), Lemma 6.3.1 and (6.7.6) with $v = \min(1,p,q)$

$$\begin{aligned}
&\Big(\mathcal{A}^{r_0}_v\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big),\mathcal{A}^{r_1}_v\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big)\Big)_{\Theta,u}\\
&\quad\hookrightarrow \Big(\ell_{q_0}(I,\ell_{p_0}(J)),\ell_{q_1}(I,\ell_{p_1}(J))\Big)_{\Theta,u} \hookrightarrow \Big(\ell_{q_0}(I,\ell_{p_0,\infty}(J)),\ell_{q_1}(I,\ell_{p_1,\infty}(J))\Big)_{\Theta,u}\\
&\quad\hookrightarrow \Big(\mathcal{A}^{r_0}_\infty\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big),\mathcal{A}^{r_1}_\infty\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big)\Big)_{\Theta,u}\,.
\end{aligned}$$

Proposition 6.3.1(ii) shows, that the first as well as the last interpolation space coincides with $\mathcal{A}^r_u\big(\ell_q(I,\ell_p(J)),\mathcal{B}\big)$. The identity (6.7.4) now follows by applying the interpolation theorem of Peetre and Lions (see Section 6.9, Theorem 6.9.2) to either (6.7.2) (replace r by r/Θ) or (6.7.3).
Finally, the statement concerning the equivalence constants follows from the fact, that the constant in the Jackson-type inequality is independent of I and J, and for the Bernstein-type inequality we do not have any constant. □

Remark 6.7.2. Under some light additional restrictions on the index set, both the Jackson- and the Bernstein-type inequality can be seen to be optimal. We either have to assume that either I or one of the sets J_i is infinite. In both cases it follows

$$\sup\Big\{\big\|\, a\,|\ell_{q_r}\big(I,\ell_{p_r}(J)\big)\big\| \ :\ a\in\Sigma_m(\mathcal{B}_{I\times J})\,,\big\|\,a\,|\ell_q\big(I,\ell_p(J)\big)\big\| = 1\Big\} = m^r\,.$$

Theorem 6.7.1. Let X and Y be two quasi-Banach spaces. Furthermore, we assume the Jackson-type inequality

$$\sigma_m(f,X,\mathcal{D}) \le c\,m^{-r}\big\|\,f\,|Y\big\|\,,\qquad f\in Y\,,$$

and the Bernstein-type inequality

$$\big\|\,f\,|Y\big\| \le c\,m^r\big\|\,f\,|X\big\|\,,\qquad f\in\Sigma_m(\mathcal{D})\,,$$

to be fulfilled for some $r>0$. Then it holds

$$\mathcal{A}^{\Theta r}_u(X,\mathcal{D}) = \big(X,Y\big)_{\Theta,u}$$

for every $0<\Theta<1$ and $0<u\le\infty$.

This well-known theorem is due to DeVore and Popov [18], see also [17]. It remains valid in the more general framework of *approximation schemes*, and a version for generalized approximation spaces exists as well (see [52]).

Corollary 6.7.1. Let $0 < p, q \le \infty$, $r > 0$, and define p_r and q_r as before. Then it holds

$$\ell_{q_r}\big(I, \ell_{p_r}(J)\big) \hookrightarrow \mathcal{A}^r_{\max(p_r,q_r)}\big(\ell_q(I, \ell_p(J))\,, \mathcal{B}_{I\times J}\big)\,.$$

The proof already uses interpolation arguments whose formulation will be given in Section 6.9.

Proof. We start with the case $p_r \le q_r$, i.e. $p \le q$. Then we find by Theorem 6.9.1(i) and the well-known monotonicity properties of Lorentz sequence spaces

$$\begin{aligned}\ell_{q_r}(I, \ell_{p_r}(J)) &\hookrightarrow \ell_{q_r}(I, \ell_{p_r,q_r}(J)) = \Big(\ell_{q_0}(I, \ell_{p_0}(J)), \ell_{q_1}(I, \ell_{p_1}(J))\Big)_{\Theta,q_r} \\ &\hookrightarrow \Big(\mathcal{A}^{r_0}_\infty\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big), \mathcal{A}^{r_1}_\infty\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big)\Big)_{\Theta,q_r} = \mathcal{A}^r_{q_r}\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big)\,,\end{aligned}$$

where $0 < r_0 < r < r_1 < \infty$, $0 < \Theta < 1$, $r = (1-\Theta)r_0 + \Theta r_1$, $\frac{1}{q_0} = r_0 + \frac{1}{q}$, $\frac{1}{q_1} = r_1 + \frac{1}{q}$, $\frac{1}{p_0} = r_0 + \frac{1}{p}$, and $\frac{1}{p_1} = r_1 + \frac{1}{p}$, and hence also $\frac{1}{q_r} = \frac{1-\Theta}{q_0} + \frac{\Theta}{q_1}$ and $\frac{1}{p_r} = \frac{1-\Theta}{p_0} + \frac{\Theta}{p_1}$.
In case $q \le p$, i.e. $q_r \le p_r$, we proceed similarly, now using Theorem 6.9.2(iii). We get

$$\begin{aligned}\ell_{q_r}(I, \ell_{p_r}(J)) &\hookrightarrow \Big(\ell_{q_0}(I, \ell_{p_0}(J)), \ell_{q_1}(I, \ell_{p_1}(J))\Big)_{\Theta,p_r} \\ &\hookrightarrow \Big(\mathcal{A}^{r_0}_\infty\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big), \mathcal{A}^{r_1}_\infty\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big)\Big)_{\Theta,p_r} = \mathcal{A}^r_{p_r}\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big)\,,\end{aligned}$$

where $q_0, q_1, p_0, p_1, r_0, r_1$ have the same meaning as before. □

Proposition 6.7.1 and Corollary 6.7.1 now are the keys to the calculation of m-term widths for Besov spaces in the limiting case.

Theorem 6.7.2. Let $0 < p_0 \le p_1 \le \infty$ and $0 < q_0 \le q_1 \le \infty$, and define

$$r = \min\left(\frac{1}{p_0} - \frac{1}{p_1}\,,\, \frac{1}{q_0} - \frac{1}{q_1}\right).$$

(i) Let I be a countable index set, and let $J = (J_i)_{i\in I}$ be a family of countable index sets satisfying $\sup_{i\in I} \#J_i = \infty$. Then it holds

$$\sigma_m\Big(\ell_{p_0}\big(I, \ell_{q_0}(J)\big)\,, \ell_{p_1}\big(I, \ell_{q_1}(J)\big)\,, \mathcal{B}_{I\times J}\Big) \sim m^{-r}\,,$$

where the equivalence constants do not depend on I or J.

(ii) Let $\overline{t} = \overline{d}\big(\frac{1}{p_0} - \frac{1}{p_1}\big)$, and assume $\sup_{\overline{\nu}\in\mathbb{N}_0^N} \#\nabla_{\overline{\nu}} = \infty$. Then it holds

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}b(\nabla)\,, s^{\overline{0}}_{p_1,q_1}b(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \sim m^{-r}\,,$$

where the equivalence constants do not depend on $d_1, \ldots, d_N$ or N.

The prove coincides essentially with the proof of Theorem 4 in [34].

Proof. The identification $\ell_q\big(\mathbb{N}_0^N, \ell_p(\nabla)\big) = L_{\overline{r}-\overline{d}/p}\big(s^{\overline{r}}_{p,q}b(\nabla)\big)$ mentioned at the start of the section (see Section 5.3.2), the assumption $\overline{t} - \overline{d}/p_0 = -\overline{d}/p_1$ and Lemma 6.2.3 imply

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}b(\nabla)\,, s^{\overline{0}}_{p_1,q_1}b(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) = \sigma_m\Big(\ell_{p_0}\big(\mathbb{N}_0^N, \ell_{q_0}(\nabla)\big)\,, \ell_{p_1}\big(\mathbb{N}_0^N, \ell_{q_1}(\nabla)\big)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big)\,.$$

Hence it suffices to prove (i). Moreover, the assumptions on the parameters ensure $\ell_{p_0}\big(\mathbb{N}_0^N, \ell_{q_0}(\nabla)\big) \hookrightarrow \ell_{p_1}\big(\mathbb{N}_0^N, \ell_{q_1}(\nabla)\big)$ (which follows from the ℓ_p-monotonicity), where the norm of the embedding operator equals one. Hence we can estimate the m-term width by 1 as well. It remains to treat the case $r > 0$.
The inequality (6.7.5) proves the case $\frac{1}{p_0} - \frac{1}{p_1} = \frac{1}{q_0} - \frac{1}{q_1}$. Now let at first be $r = \frac{1}{p_0} - \frac{1}{p_1} < \frac{1}{q_0} - \frac{1}{q_1}$. We then define q_* by $\frac{1}{q_*} := \frac{1}{p_0} - \frac{1}{p_1} + \frac{1}{q_1} < \frac{1}{q_0}$. Hence $q_0 < q_*$, and the ℓ_q-monotonicity implies

$$\ell_{q_0}\big(I, \ell_{p_0}(J)\big) \hookrightarrow \ell_{q_*}\big(I, \ell_{p_0}(J)\big) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{p_0}-\frac{1}{p_1}}\big(\ell_{q_1}\big(I, \ell_{p_1}(J)\big), \mathcal{B}_{I\times J}\big)\,,$$

where the second embedding follows from Corollary 6.7.1.
In case $r = \frac{1}{q_0} - \frac{1}{q_1} < \frac{1}{p_0} - \frac{1}{p_1}$, we define p_* by $\frac{1}{p_*} := \frac{1}{q_0} - \frac{1}{q_1} + \frac{1}{p_1} < \frac{1}{p_0}$. Hence $p_0 < p_*$, and from the ℓ_q-monotonicity and Corollary 6.7.1 we conclude this time

$$\ell_{q_0}\big(I, \ell_{p_0}(J)\big) \hookrightarrow \ell_{q_0}\big(I, \ell_{p_*}(J)\big) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{q_0}-\frac{1}{q_1}}\big(\ell_{q_1}\big(I, \ell_{p_1}(J)\big), \mathcal{B}_{I\times J}\big)\,.$$

Thus the estimates from above are proved in view of Lemma 6.3.1. The estimates from below follow from obvious modifications of the examples in (6.5.1) and (6.5.3). □

6.7.2 Some Bernstein-type inequalities

In this section we want to generalize an inequality for L_p-normalized characteristic functions which is due to Temlyakov [78] in the isotropic case and Wojtaszczyk [96] in the tensor product case, originally proved for the Haar system.
We need some further notations. The set of all dyadic cubes will be denoted by $\mathcal{D}(n)$, i.e.

$$\mathcal{D}(n) = \Big\{C = 2^{-j}\big([0,1]^n + k\big) \;:\; j \in \mathbb{N}_0\,, k \in \mathbb{Z}^n\Big\}\,.$$

The following result is well-known, see the Lemmas 2.1 and 2.2 in [78], Lemma 1 in [19] or Theorem 11.2 in [36].

Lemma 6.7.1. Let $0 < p < \infty$, and let $\mathcal{I} \subset \mathcal{D}(n)$ be a set with $|\mathcal{I}| = m$. Then it holds

$$\Bigg\|\sum_{C\in\mathcal{I}} \mathcal{X}_C^{(p)}\Bigg|L_p(\mathbb{R}^d)\Bigg\| = \left(\int_{\mathbb{R}^d}\Big(\sum_{C\in\mathcal{I}} \mathcal{X}_C^{(p)}(s)\Big)^p ds\right)^{1/p} \sim m^{1/p}\,.$$

In case $1 < p < \infty$ the functions $\mathcal{X}_C^{(p)}$ on the left hand side of this lemma can be replaced by functions from a general L_p-normalized wavelet system as in Proposition 1.2.1. Moreover, the estimate remains true upon replacing the ℓ_1-summation in the middle term by an ℓ_q-quasi-norm, $0 < q \leq \infty$.

For our considerations we will need a tensorized version of the above lemma. We define the set of dyadic rectangles $\mathcal{D}$ by

$$\mathcal{D} = \mathcal{D}(d_1,\ldots,d_N) = \Big\{ Q = Q_1 \times \cdots \times Q_N \,:\, Q_l \in \mathcal{D}(d_l)\,, l = 1,\ldots,N \Big\}\,.$$

In other words, the set $\mathcal{D}$ contains just the rectangles $Q_{\overline{\nu},k}$, $\overline{\nu} \in \mathbb{N}_0^N$, $k \in \mathbb{Z}^d$.
We then put for sequences $a = (a_Q)_{Q\in\mathcal{D}}$ of complex numbers

$$\|\, a \,\|_{\overline{d}} := \left(\int_{\mathbb{R}^d} \left(\sum_{Q\in\mathcal{D}} \big| a_Q \mathcal{X}_Q^{(p)}(s) \big|^q \right)^{p/q} ds \right)^{1/p} = \big\|\, a \,|\, s_{p,q}^{\overline{d}/p} f \big\|\,.$$

The first step toward the desired generalization of Lemma 6.7.1 is the following estimate for finite sequences, i.e. sequences with only finitely many non-vanishing components a_Q.

Lemma 6.7.2. Let a be a finite sequence, $a = \sum_{Q\in\mathcal{I}} a_Q e^Q$, $\mathcal{I} \subset \mathcal{D}$ with $\#\mathcal{I} = m \ge 2$. Then it holds for $0 < p \le q < \infty$

$$\big(\log m\big)^{N(1/q-1/p)} \left(\sum_{Q\in\mathcal{I}} |a_Q|^p \right)^{1/p} \lesssim \|\, a \,\|_{\overline{d}} \le \left(\sum_{Q\in\mathcal{I}} |a_Q|^p \right)^{1/p}, \qquad (6.7.7)$$

and for $0 < q \le p < \infty$ we obtain

$$\left(\sum_{Q\in\mathcal{I}} |a_Q|^p \right)^{1/p} \le \|\, a \,\|_{\overline{d}} \lesssim \big(\log m\big)^{N(1/q-1/p)} \left(\sum_{Q\in\mathcal{I}} |a_Q|^p \right)^{1/p}. \qquad (6.7.8)$$

All occurring constants depend on p, q and $\overline{d}$ only.

Remark 6.7.3. The proofs of this lemma and the successive proposition follow closely the arguments given in [96], Section 4.
Since the assumption on a can be reformulated as $a \in \Sigma_m(\mathcal{B}_{\mathcal{D}})$, these results can be interpreted as Bernstein-type inequalities for the spaces $s_{p,q}^{\overline{d}/p} f$ and $\ell_p(\mathcal{D})$, compare with Theorem 6.7.1.

Proof. **Step 1: We start with the case** $q = 1$**.**
The prove of the right hand side estimate in (6.7.7) for $0 < p \le 1$ follows immediately from the monotonicity of ℓ_p-quasi-norms. We obtain

$$\begin{aligned} \|\, a \,\|_{\overline{d}} &= \left(\int_{\mathbb{R}^d} \left(\sum_{Q\in\mathcal{I}} |a_Q \mathcal{X}_Q^{(p)}(s)| \right)^p ds \right)^{1/p} \\ &\le \left(\int_{\mathbb{R}^d} \sum_{Q\in\mathcal{I}} \Big(|a_Q \mathcal{X}_Q^{(p)}(s)| \Big)^p ds \right)^{1/p} = \left(\sum_{Q\in\mathcal{I}} |a_Q|^p \right)^{1/p}. \end{aligned}$$

For the prove of the right hand side inequality in (6.7.8) for $1 \le p < \infty$ we consider the case $N = 1$ first. Let $\pi : \{1,\ldots,m\} \longrightarrow \mathcal{I}$ be a bijection, such that $|a_{\pi(j)}|$ is a non-increasing sequence. Furthermore, let M be the uniquely determined integer, such that

$2^{M-1} \le m < 2^M$, and define $g_k = \sum_{j=2^{k-1}}^{2^k-1} |a_{\pi(j)}\mathcal{X}^{(p)}_{\pi(j)}|$, $k = 1, \ldots, M$. Then the triangle inequality yields

$$\| a \|_d = \left\| \sum_{k=1}^{M} g_k \Big| L_p(\mathbb{R}^d) \right\| \le \sum_{k=1}^{M} \| g_k \,| L_p(\mathbb{R}^d) \| = \sum_{k=1}^{M} \left\| \sum_{j=2^{k-1}}^{2^k-1} a_{\pi(j)} e^{\pi(j)} \right\|_d$$
$$\le \sum_{k=1}^{M} \left\| \sum_{j=2^{k-1}}^{2^k-1} a_{\pi(2^{k-1})} e^{\pi(j)} \right\|_d \lesssim \sum_{k=1}^{M} 2^{(k-1)/p} \big| a_{\pi(2^{k-1})} \big| \,.$$

The last two estimates follow from the lattice structure of $\| \cdot \|_d$ and from Lemma 6.7.1. On the other hand, we obtain from Hölder's inequality with respect to $1 = \frac{1}{p} + (1 - \frac{1}{p})$

$$\sum_{Q \in \mathcal{I}} |a_Q|^p = \sum_{j=1}^{m} |a_{\pi(j)}|^p \ge \sum_{k=1}^{M} 2^{k-1} |a_{\pi(2^k-1)}|^p \ge M^{1-p} \left(\sum_{k=1}^{M} 2^{(k-1)/p} |a_{\pi(2^k-1)}| \right)^p$$
$$\ge 2^{-p} M^{1-p} \left(\sum_{k=1}^{M-1} 2^{(k-1)/p} \big| a_{\pi(2^k)} \big| + \big| a_{\pi(1)} \big| + \sum_{k=2}^{M} 2^{(k-1)/p} \big| a_{\pi(2^k-1)} \big| \right)^p$$
$$\ge 2^{1-p} M^{1-p} \left(\sum_{k=1}^{M} 2^{(k-1)/p} \big| a_{\pi(2^{k-1})} \big| \right)^p .$$

Combining both estimates eventually yields

$$\| a \|_d \lesssim M^{1-1/p} \left(\sum_{Q \in \mathcal{I}} |a_Q|^p \right)^{1/p} \lesssim \big(\log m \big)^{1-1/p} \left(\sum_{Q \in \mathcal{I}} |a_Q|^p \right)^{1/p} .$$

Step 2: Now the case $N \ge 2$ will be proven by induction over N. Given a finite set of rectangles $\mathcal{I} \subset \mathcal{D}(d_1, \ldots, d_N)$, we can rewrite every $Q \in \mathcal{I}$ as in $Q = Q' \times Q''$ with $Q' \in \mathcal{D}(d_1)$ and $Q'' \in \mathcal{D}(d_2, \ldots, d_N)$, and accordingly $\mathcal{X}^{(p)}_Q = \mathcal{X}^{(p)}_{Q'} \otimes \mathcal{X}^{(p)}_{Q''}$. Note that for $\#\mathcal{I} = m \ge 2$ there are at most m different cubes Q' and at most m rectangles Q'' occurring in this way. Then we find

$$\| a \|_d^p = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d-d_1}} \left(\sum_{Q = Q' \times Q'' \in \mathcal{I}} |a_Q \mathcal{X}^{(p)}_{Q'}(t)| \cdot |\mathcal{X}^{(p)}_{Q''}(s)| \right)^p ds\, dt$$
$$= \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d-d_1}} \left(\sum_{Q'' \in \mathcal{D}(d_2, \ldots, d_N)} \left(\sum_{Q' \,:\, Q' \times Q'' \in \mathcal{I}} |a_Q \mathcal{X}^{(p)}_{Q'}(t)| \right) |\mathcal{X}^{(p)}_{Q''}(s)| \right)^p ds\, dt \,.$$

Now we can apply the induction hypothesis to the $(d - d_1)$-dimensional integral (the inner sums are treated as coefficients for fixed $t \in \mathbb{R}^{d_1}$). In this way we obtain (observe $(N-1)(p-1) \ge 0$)

$$\| a \|_d^p \lesssim \big(\log m \big)^{(N-1)(p-1)} \int_{\mathbb{R}^{d_1}} \sum_{Q'' \in \mathcal{D}(d_2, \ldots, d_N)} \left(\sum_{Q' \,:\, Q' \times Q'' \in \mathcal{I}} |a_Q \mathcal{X}^{(p)}_{Q'}(t)| \right)^p dt \,.$$

At this point we further apply the result for the case $N = 1$. We end up with

$$\| a \|_d^p \lesssim \big(\log m \big)^{(N-1)(p-1)} \big(\log m \big)^{(p-1)} \sum_{Q'' \in \mathcal{D}(d_2, \ldots, d_N)} \sum_{Q' \,:\, Q' \times Q'' \in \mathcal{I}} \big| a_Q \big|^p$$

$$= \big(\log m\big)^{N(p-1)} \sum_{Q\in\mathcal{I}} \big|a_Q\big|^p .$$

This proves the right hand side of (6.7.8).

Step 3: Now consider general q.
We obtain for $0 < p \le q$ from Step 1, applied to $0 < p/q \le 1$,

$$\left(\int_{\mathbb{R}^d}\left(\sum_{Q\in\mathcal{I}} \big|a_Q \mathcal{X}_Q^{(p)}(s)\big|^q\right)^{\frac{p}{q}} ds\right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}^d}\left(\sum_{Q\in\mathcal{I}} |a_Q|^q \mathcal{X}_Q^{(p/q)}(s)\right)^{\frac{p}{q}} ds\right)^{\frac{q}{p}\cdot\frac{1}{q}}$$
$$\le \left(\sum_{Q\in\mathcal{I}} \big(|a_Q|^q\big)^{p/q}\right)^{\frac{q}{p}\cdot\frac{1}{q}} = \left(\sum_{Q\in\mathcal{I}} |a_Q|^p\right)^{1/p} .$$

Similarly we find for $0 < q \le p$ from Step 2, applied to $1 \le p/q < \infty$,

$$\left(\int_{\mathbb{R}^d}\left(\sum_{Q\in\mathcal{I}} |a_Q \mathcal{X}_Q^{(p)}(s)|^q\right)^{\frac{p}{q}} ds\right)^{\frac{1}{p}} \lesssim \left(\big(\log m\big)^{N(1-q/p)}\left(\sum_{Q\in\mathcal{I}} \big(|a_Q|^q\big)^{p/q}\right)^{\frac{q}{p}}\right)^{\frac{1}{q}}$$
$$= \big(\log m\big)^{N(1/q-1/p)}\left(\sum_{Q\in\mathcal{I}} |a_Q|^p\right)^{1/p} .$$

Step 4: We prove the estimates from below in the case $1 < p, q < \infty$.
For every finite sequence $a = \sum_{Q\in\mathcal{I}} a_Q e^Q$, $\#\mathcal{I} = m$, with $a_Q \ne 0$, $Q \in \mathcal{I}$, define another finite sequence b by $b_Q = \frac{|a_Q|^p}{a_Q}$ for $Q \in \mathcal{I}$ and zero otherwise. Then we have $f_a = \big(a_Q \mathcal{X}_Q^{(p)}\big)_{Q\in\mathcal{D}} \in L_p(\ell_q)$ and $f_b = \big(b_Q \mathcal{X}_Q^{(p')}\big)_{Q\in\mathcal{D}} \in L_{p'}(\ell_{q'})$, where p', q' are the usual conjugated indices.
We begin with the case $1 < q \le p < \infty$ and hence $1 < p' \le q' < \infty$. In view of Proposition 2.3.14 we find that f_b generates a functional on $L_p(\ell_q)$. Applying the characterization of these functionals we obtain

$$\sum_{Q\in\mathcal{I}} |a_Q|^p = \left|\sum_{Q\in\mathcal{I}} a_Q b_Q\right| = \left|\int_{\mathbb{R}^d} \sum_{Q\in\mathcal{I}} |Q|^{-1} a_Q b_Q \mathcal{X}_Q(s) ds\right|$$
$$= \left|\int_{\mathbb{R}^d} \sum_{Q\in\mathcal{D}} \Big(a_Q \mathcal{X}_Q^{(p)}(s)\Big)\Big(b_Q \mathcal{X}_Q^{(p')}(s)\Big) ds\right| \le \big\| f_a \,\big|L_p(\ell_q)\big\| \cdot \big\| f_b \,\big|L_{p'}(\ell_{q'})\big\| .$$

Moreover, Step 3 yields

$$\big\| f_b \,\big|L_{p'}(\ell_{q'})\big\|^{p'} = \int_{\mathbb{R}^d}\left(\sum_{Q\in\mathcal{I}} \Big| b_Q \mathcal{X}_Q^{(p')}(s) \Big|^{q'}\right)^{p'/q'} ds \le \sum_{Q\in\mathcal{I}} |b_Q|^{p'} = \sum_{Q\in\mathcal{I}} |a_Q|^p ,$$

where we used $(p-1)p' = p$. Combining both estimates we now obtain

$$\big\| f_a \,\big|L_p(\ell_q)\big\| \equiv \left(\int_{\mathbb{R}^d}\left(\sum_{Q\in\mathcal{I}} \Big| a_Q \mathcal{X}_Q^{(p)}(s) \Big|^q\right)^{p/q} ds\right)^{1/p}$$

$$\geq \big\| \, f_b \,\big|L_{p'}(\ell_{q'})\big\|^{-1} \sum_{Q\in\mathcal{I}} |a_Q|^p \geq \bigg(\sum_{Q\in\mathcal{I}} |a_Q|^p\bigg)^{1/p} .$$

For the estimate from above in (6.7.8), consider the same sequence b. Now the condition $1 < p \leq q < \infty$ implies $1 < q' \leq p' < \infty$, and due to Step 3 we find this time

$$\big\| \, f_b \,\big|L_{p'}(\ell_{q'})\big\|^{p'} \lesssim \big(\log m\big)^{N(1/q'-1/p')p'} \sum_{Q\in\mathcal{I}} |b_Q|^{p'} = \big(\log m\big)^{N(1/p-1/q)p'} \sum_{Q\in\mathcal{I}} |a_Q|^p ,$$

and consequently

$$\bigg(\int_{\mathbb{R}^d}\bigg(\sum_{Q\in\mathcal{I}}\Big|\, a_Q\mathcal{X}_Q^{(p)}(s)\,\Big|^q\bigg)^{p/q} ds\bigg)^{1/p} \geq \big(\log m\big)^{N(1/q-1/p)}\bigg(\sum_{Q\in\mathcal{I}} |a_Q|^p\bigg)^{1/p} .$$

Step 5: The results for the case $q = 1$ and $0 < p < \infty$ now follow from Step 4 by choosing parameters $1 < \widetilde{p}, \widetilde{q} < \infty$, such that $p = \widetilde{p}/\widetilde{q}$. Afterwards we use the same arguments as in Step 3, but in a reversed way. Exemplary, in case $p < 1$ we find

$$\begin{aligned}
\big(\log m\big)^{N(1-1/p)}\bigg(\sum_{Q\in\mathcal{I}} |a_Q|^p\bigg)^{1/p} &= \Bigg(\big(\log m\big)^{N(1/\widetilde{q}-1/\widetilde{p})}\bigg(\sum_{Q\in\mathcal{I}}\big(|a_Q|^{1/\widetilde{q}}\big)^{\widetilde{p}}\bigg)^{1/\widetilde{p}}\Bigg)^{\widetilde{q}} \\
&\lesssim \bigg(\int_{\mathbb{R}^d}\bigg(\sum_{Q\in\mathcal{I}}\big(|a_Q|^{1/\widetilde{q}}\mathcal{X}_Q^{(\widetilde{p})}(s)\big)^{\widetilde{q}}\bigg)^{\widetilde{p}/\widetilde{q}} ds\bigg)^{\widetilde{q}/\widetilde{p}} \\
&= \bigg(\int_{\mathbb{R}^d}\bigg(\sum_{Q\in\mathcal{I}} |a_Q|\mathcal{X}_Q^{(p)}(s)\bigg)^{p} ds\bigg)^{1/p} .
\end{aligned}$$

Finally, the case of arbitrary parameters $0 < q < \infty$ follows once more from the arguments in Step 3 with the help of the case $q = 1$. This yields the lower estimates in (6.7.7) and (6.7.8) for all parameters $0 < p \leq q < \infty$ and $0 < q \leq p < \infty$, respectively. □

Proposition 6.7.2. Let $\mathcal{I}$ be a finite non-empty subset of $\mathbb{N}_0^N \times \mathbb{Z}^d$. Then it holds for $0 < p \leq q < \infty$

$$\big(1+\log m\big)^{(N-1)(1/q-1/p)} m^{1/p} \lesssim \bigg(\int_{\mathbb{R}^d}\bigg(\sum_{(\overline{\nu},k)\in\mathcal{I}} 2^{\overline{\nu}\cdot\overline{d}q/p}\mathcal{X}_{\overline{\nu},k}(s)\bigg)^{p/q} ds\bigg)^{1/p} \lesssim m^{1/p} ,$$

and for $0 < q \leq p < \infty$ we find

$$m^{1/p} \lesssim \bigg(\int_{\mathbb{R}^d}\bigg(\sum_{(\overline{\nu},k)\in\mathcal{I}} 2^{\overline{\nu}\cdot\overline{d}q/p}\mathcal{X}_{\overline{\nu},k}(s)\bigg)^{p/q} ds\bigg)^{1/p} \lesssim \big(1+\log m\big)^{(N-1)(1/q-1/p)} m^{1/p} .$$

Moreover, all estimates are order-optimal.

Proof. We only treat the case $q = 1$, the general case can be obtained as in Step 3 of the proof of Lemma 6.7.2.

Step 1: The proof of the estimates from above follows by the same arguments (ℓ_p-monotonicity or induction, respectively) as in Steps 1–3 of the previous proof. The induction basis now is given by Lemma 6.7.1, and at the end of the induction step when applying the result for $N = 1$ one uses once more Lemma 6.7.1 instead of Step 1 of that proof. Moreover, the proof of the estimates from below follows by the same duality arguments as in Steps 4 and 5 of the proof of Lemma 6.7.2.

Step 2: Optimality. If the rectangles $Q_{\overline{\nu},k}$, $(\overline{\nu}, k) \in \mathcal{I}$, are pairwise disjoint, then clearly we have $\big\|\sum_{(\overline{\nu},k)\in\mathcal{I}} \mathcal{X}^{(p)}_{\overline{\nu},k} \big| L_p(\mathbb{R}^d)\big\| = m^{1/p}$. For the other estimates, we define index sets $\mathcal{I}_\mu$ by

$$\mathcal{I}_\mu := \Big\{(\overline{\nu}, k) \,:\, Q_{\overline{\nu},k} \subset [0,1]^d \,, \overline{\nu} \in M\big(\mu, \overline{d}\big)\Big\} \,, \qquad \mu \in \mathbb{N}_0 \,.$$

Then we find $\#\mathcal{I}_\mu = S\big(\mu, \overline{d}\big)2^\mu$ and $\sum_{k\,:\,(\overline{\nu},k)\in\mathcal{I}_\mu} \mathcal{X}_{\overline{\nu},k} = \mathcal{X}_{[0,1]^d}$ for every fixed $\overline{\nu} \in M\big(\mu, \overline{d}\big)$. Moreover, we now obtain

$$\int_{\mathbb{R}^d}\bigg(\sum_{(\overline{\nu},k)\in\mathcal{I}_\mu} 2^{\overline{\nu}\cdot\overline{d}/p}\mathcal{X}_{\overline{\nu},k}(s)\bigg)^p ds = \int_{\mathbb{R}^d}\bigg(\sum_{\overline{\nu}\in M(\mu,\overline{d})} 2^{\mu/p}\mathcal{X}_{[0,1]^d}(s)\bigg)^p ds = S\big(\mu, \overline{d}\big)^p 2^\mu \,.$$

If we additionally assume $\mu \in \mathcal{N}$, then it holds $m = \#\mathcal{I}_\mu \sim 2^\mu \mu^{N-1}$ and hence

$$\Bigg(\int_{\mathbb{R}^d}\bigg(\sum_{(\overline{\nu},k)\in\mathcal{I}_\mu} 2^{\overline{\nu}\cdot\overline{d}/p}\mathcal{X}_{\overline{\nu},k}(s)\bigg)^p ds\Bigg)^{1/p} \sim \big(\log m\big)^{(N-1)(1-1/p)} m^{1/p} \,.$$

Altogether this means that neither of the assertions can be improved. □

Remark 6.7.4. The argument used in the induction step in the proof additionally shows, that also the estimates in Lemma 6.7.2 are order-optimal, since improved estimates there would imply better results in Proposition 6.7.2, in contradiction to the proven optimality.

Remark 6.7.5. The proof furthermore shows, that the result remains valid, if we replace the characteristic functions by Haar functions (constructed in the same way as the wavelets in Proposition 4.3.1), since the estimate only depends on the absolute value of these functions. Though there is possibly more than one Haar function with the same support Q involved, their count is bounded independent of Q, hence the asymptotic behaviour remains unchanged.

Remark 6.7.6. If we slightly change the interpretation in Step 4 of the proof of Lemma 6.7.2, i.e. if we consider $f_a \in L_p(\ell_q)$ as the generator of a functional on $L_{p'}(\ell_{q'}) \ni f_b$, then we see that the estimates from below in both, Lemma 6.7.2 and Proposition 6.7.2, remain valid also for $q = \infty$ (where, as usual, the summation has to be replaced by a supremum).

Remark 6.7.7. Proposition 6.7.2 has some interesting consequences, apart from its importance for our further considerations. If $p \neq q$ then the respective estimates are no longer purely determined by $\#\mathcal{I}$, as it was in the isotropic case (Lemma 6.7.1), i.e. there are two sequences of index sets $(\mathcal{I}_m^1)_{m\geq 2}$ and $(\mathcal{I}_m^2)_{m\geq 2}$ such that the associated sequences

$$\left(\left(\int_{\mathbb{R}^d}\left(\sum_{(\overline{\nu},k)\in\mathcal{I}_m^i} 2^{\overline{\nu}\cdot\overline{d}q/p}\mathcal{X}_{\overline{\nu},k}(s)\right)^{p/q} ds\right)^{1/p}\right)_{m\geq 2}, \qquad i=1,2,$$

are non-equivalent. The consequence of this observation is the fact that the best m-term approximation in spaces $s^{\overline{0}}_{p,q}f(\nabla)$ generally is not given by simply selecting the m largest coefficients (in a normalized sense). We refer to [41] and the survey [80] for more information in that direction. In the terms of those references, the space $s^{\overline{0}}_{p,q}f(\nabla)$ is not weakly rearrangement invariant, and the system $\mathcal{B}$ is not democratic. This behaviour of tensor product systems had been investigated before in [44].

6.7.3 The limiting case for f-spaces

In this section we will consider estimates for the asymptotic behaviour of the m-term width $\sigma_m\big(s^{\overline{t}}_{p_0,q_0}f(\nabla)\,,s^{\overline{0}}_{p_1,q_1}f(\nabla)\,,\mathcal{B}_{\mathbb{N}_0^N\times\nabla}\big)$. Unfortunately, our method works only in some special cases. Further ones will be treated with the help of real interpolation in Section 6.10. However, at the end of this section we will formulate a conjecture for the general result.

Theorem 6.7.3. Let $0<p_0<p_1<\infty$, $0<q_0,q_1\leq\infty$, and put $\overline{t}=\overline{d}\big(\frac{1}{p_0}-\frac{1}{p_1}\big)$. Furthermore, assume $\sup_{\overline{\nu}\in\mathbb{N}_0^N}\#\nabla_{\overline{\nu}}=\infty$.

(i) Let $q_1\leq p_1$ and $p_0\leq q_0$. Then it holds for $m\geq 2$

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}f(\nabla)\,,s^{\overline{0}}_{p_1,q_1}f(\nabla)\,,\mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big)\sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})}.$$

(ii) If $q_0\leq p_0<p_1\leq q_1$ then we find

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}f(\nabla)\,,s^{\overline{0}}_{p_1,q_1}f(\nabla)\,,\mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big)\sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}, \qquad m\in\mathbb{N}.$$

Proof. **Step 1:** We begin with the case $q_1<p_1$ and $p_0\leq q_0$. Let $a\in s^{\overline{t}}_{p_0,q_0}f(\nabla)$ with $\big\|\,a\,\big|s^{\overline{t}}_{p_0,q_0}f(\nabla)\big\|=1$. For $j\in\mathbb{Z}$ we define

$$\Lambda_{\mu,j}:=\big\{(\overline{\nu},k)\in\nabla_\mu\,:\,2^{-j}<2^{-\overline{\nu}\cdot\overline{d}/p_1}|a_{\overline{\nu},k}|\leq 2^{-j+1}\big\}, \qquad \mu\in\mathbb{N}_0.$$

In case $j\leq 0$ these sets are empty. This follows from Lemma 6.2.1 and

$$2^{-\overline{\nu}\cdot\overline{d}p_0/p_1}|a_{\overline{\nu},k}|^{p_0}=2^{\overline{\nu}\cdot(\overline{t}-\overline{d}/p_0)p_0}|a_{\overline{\nu},k}|^{p_0}=\int_{\mathbb{R}^d}2^{\overline{\nu}\cdot\overline{t}p_0}|a_{\overline{\nu},k}|^{p_0}\mathcal{X}_{\overline{\nu},k}(x)dx\leq\big\|\,a\,\big|s^{\overline{t}}_{p_0,q_0}f(\nabla)\big\|^{p_0}$$

Furthermore, we put for $M\in\mathbb{N}$

$$\Lambda^M=\bigcup_{j=1}^M\Lambda_j\,, \qquad \Lambda_j=\bigcup_{\mu=0}^M\Lambda_{\mu,j} \qquad \text{as well as} \qquad T_Ma:=\sum_{(\overline{\nu},k)\in\Lambda^M}a_{\overline{\nu},k}e^{\overline{\nu},k}.$$

We begin with estimates for the counts of the sets Λ_j and Λ^M. Using eventually Lemma 6.4.2 we find

$$\begin{aligned}
\#\Lambda_{\mu,j} &= \sum_{(\overline{\nu},k)\in\Lambda_{\mu,j}} 1 \le \sum_{(\overline{\nu},k)\in\Lambda_{\mu,j}} \left(\frac{|a_{\overline{\nu},k}|2^{-\overline{\nu}\cdot\overline{d}/p_1}}{2^{-j}}\right)^{p_0} \\
&\le 2^{jp_0}\int_{\mathbb{R}^d}\sum_{(\overline{\nu},k)\in\nabla_\mu} 2^{\overline{\nu}\cdot\overline{d}(1/p_0-1/p_1)p_0}|a_{\overline{\nu},k}|^{p_0}\mathcal{X}_{\overline{\nu},k}(x)dx \\
&\le 2^{jp_0}S\big(\mu,\overline{d}\big)^{(1/p_0-1/q_0)p_0}\big\|\mathrm{re}_\mu a\big|s_{p_0,q_0}^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}f(\nabla)\big\|^{p_0} < \infty\,.
\end{aligned}$$

Thus also the sets Λ_j are finite ones. Because of $p_0 \le q_0 \le \infty$ it now follows from Proposition 6.7.2

$$\begin{aligned}
\#\Lambda_j &= \#\Lambda_j\big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(1/q_0-1/p_0)p_0}\big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(1-p_0/q_0)} \\
&\lesssim \big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(1-p_0/q_0)}\int_{\mathbb{R}^d}\Bigg(\sum_{(\overline{\nu},k)\in\Lambda_j}\big(2^{\overline{\nu}\cdot\overline{d}/p_0}\mathcal{X}_{\overline{\nu},k}(x)\big)^{q_0}\Bigg)^{p_0/q_0}dx \\
&\le \big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(1-\frac{p_0}{q_0})}2^{jp_0}\int_{\mathbb{R}^d}\Bigg(\sum_{(\overline{\nu},k)\in\Lambda_j}\Big(2^{\overline{\nu}\cdot\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}|a_{\overline{\nu},k}|\mathcal{X}_{\overline{\nu},k}(x)\Big)^{q_0}\Bigg)^{\frac{p_0}{q_0}}dx \\
&\le \big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(1-\frac{p_0}{q_0})}2^{jp_0}\big\|\,a\,\big|s_{p_0,q_0}^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}f(\nabla)\big\|\,.
\end{aligned}$$

For $q_0 = \infty$ one uses the usual modification, keeping in mind Remark 6.7.6. This means for non-empty sets Λ_j we have found

$$2^{jp_0} \gtrsim \#\Lambda_j\big(1+\log(1+\#\Lambda_j)\big)^{(N-1)(p_0/q_0-1)}\,,$$

which trivially remains true for $\Lambda_j = \emptyset$. Because

$$\Lambda^M = \big\{(\overline{\nu},k)\ :\ 2^{-\overline{\nu}\cdot\overline{d}/p_1}|a_{\overline{\nu},k}| > 2^{-M}\,,\ \overline{\nu}\cdot\overline{d} \le M\big\}\,,$$

this estimate also applies to Λ^M. Moreover, the estimate for Λ_j can be reformulated for $j \ge 2$ as

$$\log(1+\#\Lambda_j) \lesssim j \qquad \text{and} \qquad \#\Lambda_j \lesssim 2^{jp_0}j^{(N-1)(1-p_0/q_0)}\,. \tag{6.7.9}$$

Hence we conclude that $T_M a$ is an m-term approximation of a, where m is given by $m = [c_0 2^{Mp_0}M^{(N-1)(1-p_0/q_0)}]$. Thus it is sufficient to show

$$\big\|\,a - T_M a\,\big|s_{p_1,q_1}^{\overline{0}}f(\nabla)\big\| \le c\,\big(2^{Mp_0}\big)^{-\frac{1}{p_0}+\frac{1}{p_1}}M^{(N-1)(\frac{1}{q_1}-\frac{p_0}{q_0 p_1})}\,, \qquad M \ge 1\,.$$

As before, for the other $m \in \mathbb{N}$ the result follows by monotonicity arguments.

Step 2: Initially, we get

$$\big\|\,a - T_M a\,\big|s_{p_1,q_1}^{\overline{0}}f(\nabla)\big\|^{p_1} = \int\Bigg(\sum_{j\ge M+1}\sum_{(\overline{\nu},k)\in\Lambda_j}\big(|a_{\overline{\nu},k}|\mathcal{X}_{\overline{\nu},k}(x)\big)^{q_1}\Bigg)^{p_1/q_1}dx$$

$$\leq c \int \Big(\sum_{j \geq M+1} \sum_{(\overline{\nu},k)\in\Lambda_j} \big(2^{\overline{\nu}\cdot\overline{d}/p_1} 2^{-j} \mathcal{X}_{\overline{\nu},k}(x)\big)^{q_1} \Big)^{p_1/q_1} dx\,.$$

Because of $p_1 > q_1$ there exists a $\delta > 0$, such that $p_1(q_1-\delta)/q_1 > p_0$. Applying Hölder's inequality with respect to $1 = \frac{p_1-q_1}{p_1} + \frac{q_1}{p_1}$ to the integrand yields

$$\begin{aligned}
\sum_{j \geq M+1} \sum_{(\overline{\nu},k)\in\Lambda_j} \big(2^{\overline{\nu}\cdot\overline{d}/p_1} 2^{-j} \mathcal{X}_{\overline{\nu},k}\big)^{q_1} &= \sum_{j \geq M+1} 2^{-j\delta} \sum_{(\overline{\nu},k)\in\Lambda_j} 2^{\overline{\nu}\cdot\overline{d}q_1/p_1} 2^{-j(q_1-\delta)} \mathcal{X}_{\overline{\nu},k} \\
&\leq \Big(\sum_{j \geq M+1} 2^{-j\delta \frac{p_1}{p_1-q_1}} \Big)^{\frac{p_1-q_1}{p_1}} \Big(\sum_{j \geq M+1} \Big(\sum_{(\overline{\nu},k)\in\Lambda_j} 2^{\overline{\nu}\cdot\overline{d}\,\frac{q_1}{p_1}} 2^{-j(q_1-\delta)} \mathcal{X}_{\overline{\nu},k} \Big)^{\frac{p_1}{q_1}} \Big)^{\frac{q_1}{p_1}} \\
&\leq c\, 2^{-M\delta} \Big(\sum_{j \geq M+1} \Big(\sum_{(\overline{\nu},k)\in\Lambda_j} 2^{\overline{\nu}\cdot\overline{d}\,\frac{q_1}{p_1}} 2^{-j(q_1-\delta)} \mathcal{X}_{\overline{\nu},k} \Big)^{\frac{p_1}{q_1}} \Big)^{\frac{q_1}{p_1}}.
\end{aligned}$$

Hence, with Proposition 6.7.2, (6.7.9) and by the choice of δ we finally obtain

$$\begin{aligned}
&\big\| a - T_M a \,\big|s^{\overline{0}}_{p_1,q_1} f(\nabla)\big\|^{p_1} \\
&\quad\lesssim 2^{-M\delta p_1/q_1} \int \sum_{j \geq M+1} \Big(\sum_{(\overline{\nu},k)\in\Lambda_j} 2^{\overline{\nu}\cdot\overline{d}q_1/p_1} 2^{-j(q_1-\delta)} \mathcal{X}_{\overline{\nu},k}(x) \Big)^{p_1/q_1} dx \\
&\quad= 2^{-M\delta p_1/q_1} \sum_{j \geq M+1} 2^{-jp_1(q_1-\delta)/q_1} \int \Big(\sum_{(\overline{\nu},k)\in\Lambda_j} 2^{\overline{\nu}\cdot\overline{d}q_1/p_1} \mathcal{X}_{\overline{\nu},k}(x) \Big)^{p_1/q_1} dx \\
&\quad\lesssim 2^{-M\delta p_1/q_1} \sum_{j \geq M+1} 2^{-jp_1(q_1-\delta)/q_1} \Big((\#\Lambda_j)^{1/p_1} \big(1 + \log \#\Lambda_j\big)^{(N-1)(1/q_1-1/p_1)} \Big)^{p_1} \\
&\quad\lesssim 2^{-M\delta p_1/q_1} \sum_{j \geq M+1} 2^{-jp_1(q_1-\delta)/q_1} 2^{jp_0} j^{(N-1)(1-p_0/q_0)} j^{(N-1)(p_1/q_1-1)} \\
&\quad\lesssim 2^{-M\delta\frac{p_1}{q_1}} 2^{-M(p_1\frac{q_1-\delta}{q_1}-p_0)} M^{(N-1)(\frac{p_1}{q_1}-\frac{p_0}{q_0})} = 2^{-M(p_1-p_0)} M^{(N-1)(\frac{p_1}{q_1}-\frac{p_0}{q_0})}\,.
\end{aligned}$$

Step 3: The result extends to the case $p_1 = q_1$ with mostly the same arguments. For the same approximant $T_M a$ we now find from (6.7.9)

$$\begin{aligned}
\big\| a - T_M a \,\big|s^{\overline{0}}_{p_1,p_1} f(\nabla)\big\|^{p_1} &= \big\| a - T_M a \,\big|s^{\overline{0}}_{p_1,p_1} b(\nabla)\big\|^{p_1} \\
&= \sum_{j \geq M+1} \sum_{(\overline{\nu},k)\in\Lambda_j} 2^{-\overline{\nu}\cdot\overline{d}} |a_{\overline{\nu},k}|^{p_1} \lesssim \sum_{j \geq M+1} \sum_{(\overline{\nu},k)\in\Lambda_j} 2^{-jp_1} \\
&\lesssim \sum_{j \geq M+1} 2^{-j(p_1-p_0)} j^{(N-1)(1-p_0/q_0)}\,.
\end{aligned}$$

Inserting the choice of m, we end up with

$$\big\| a - T_M a \,\big|s^{\overline{0}}_{p_1,p_1} f(\nabla)\big\|^{p_1} \lesssim m^{-1/p_0+1/p_1} \big(\log m\big)^{(N-1)(1/p_0-1/q_0)}\,.$$

Step 4: The case $q_0 \leq p_0 < p_1 \leq q_1$ follows simply by monotonicity arguments from Corollary 6.6.1:

$$s^{\overline{t}}_{p_0,q_0} f(\nabla) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0} \big(s^{\overline{0}}_{p_1,p_1} f(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\big) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{\infty} \big(s^{\overline{0}}_{p_1,q_1} f(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\big)\,.$$

Remember $s^{\overline{r}}_{p,p}b(\nabla) = s^{\overline{r}}_{p,p}f(\nabla)$ for all admissible parameters. In both cases, the estimates from below follow from Proposition 6.5.1. $\square$

Conjecture 6.7.1. Let $0 < p_0 < p_1 < \infty$ and $0 < q_0, q_1 \le \infty$. Then it holds

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}f(\Omega), s^{\overline{0}}_{p_1,q_1}f(\Omega), \mathcal{B}\Big) \lesssim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)\left(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1}\right)_+}$$

for all $m \ge 2$.

6.7.4 Further results and mixed embeddings

In this section we shall deal with mixed-type embeddings (see Proposition 5.3.5). For these, we will estimate the m-term width for several cases of the parameters. Moreover, we will formulate conjectures for the remaining ones, which are immediate corollaries of Conjecture 6.7.1 (upon its validity) and Lemma 5.3.3.

Proposition 6.7.3. Let $0 < p_0 < p_1 < \infty$ and $0 < q_0, q_1 \le \infty$. Furthermore, let Ω be an open subset of $\mathbb{R}^d$.

(i) Let $q_0 < p_0$, and let $\frac{1}{p_0} - \frac{1}{p_1} \le \frac{1}{q_0} - \frac{1}{q_1}$. Then it holds

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\Omega)\,, s^{\overline{0}}_{p_1,q_1}f(\Omega)\,, \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \qquad m \ge 2\,.$$

(ii) Now let $p_0 \le q_0$. Then it holds

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\Omega)\,, s^{\overline{0}}_{p_1,q_1}f(\Omega)\,, \mathcal{B}\Big) \asymp m^{-\frac{1}{q_0}+\frac{1}{p_1}}$$

for every natural number $m \in \mathbb{N}$.

Proof. **Step 1:** We shall begin with the case $q_0 < p_0$ and $\frac{1}{p_0} - \frac{1}{p_1} = \frac{1}{q_0} - \frac{1}{q_1}$. These assumptions immediately imply $q_1 < p_1$, and hence by Theorem 6.7.2(ii), Lemma 6.3.1 and Lemma 5.3.3 we have the embeddings

$$s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\nabla) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_\infty\big(s^{\overline{0}}_{p_1,q_1}b(\nabla)\,, \mathcal{B}\big) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_\infty\big(s^{\overline{0}}_{p_1,q_1}f(\nabla)\,, \mathcal{B}\big)\,.$$

Step 2: If $\frac{1}{p_0} - \frac{1}{p_1} < \frac{1}{q_0} - \frac{1}{q_1}$ we define q_* by $\frac{1}{q_*} = \frac{1}{p_0} - \frac{1}{p_1} + \frac{1}{q_1}$. Then we have $q_0 < q_*$, and it holds $q_* < p_0$ if, and only if, it holds $q_1 < p_1$. In this case, we can conclude from Step 1

$$s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0}b(\nabla) \hookrightarrow s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_*}b(\nabla) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_\infty\big(s^{\overline{0}}_{p_1,q_1}f(\nabla)\,, \mathcal{B}\big)\,.$$

On the other hand, in case $q_0 < p_0$ and $p_1 \le q_1$, the result follows from Corollary 6.6.1 as in the fourth step of the proof of Theorem 6.7.3.

Step 3: We now assume $p_0 \le q_0 \le p_1$, and define

$$\frac{1}{p_*} = \frac{1}{p_0} - \frac{1}{q_0} + \frac{1}{p_1} > 0\,.$$

Then we obtain from Theorem 6.7.2 together with Lemma 6.2.3 and eventually the Jawerth-Franke embedding (Proposition 5.3.5)

$$s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} b(\nabla) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{q_0}-\frac{1}{p_1}}\big(s^{\bar{d}(\frac{1}{p_0}-\frac{1}{q_0})}_{p_*,p_1} b(\nabla), \mathcal{B}\big) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{q_0}-\frac{1}{p_1}}\big(s^{\bar{0}}_{p_1,q_1} f(\nabla), \mathcal{B}\big) .$$

In case $\nabla = \nabla(\Omega)$, all the estimates from below are given by (6.5.2) and (6.5.4), respectively. □

Conjecture 6.7.2. Let $0 < p_0 < p_1 < \infty$ and $0 < q_0, q_1 \leq \infty$, where $q_0 \leq p_0$. Furthermore, let Ω be an open subset of $\mathbb{R}^d$. Then it holds

$$\sigma_m\Big(s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} b(\Omega)\,, s^{\bar{0}}_{p_1,q_1} f(\Omega)\,, \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)\big(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1}\big)_+}$$

for every natural number $m \geq 2$.

Similarly, we can treat the limiting situation for the other mixed embedding.

Proposition 6.7.4. Let $0 < p_0 < p_1 \leq \infty$ and $0 < q_0, q_1 \leq \infty$. Furthermore, let Ω be an open subset of $\mathbb{R}^d$.

(i) Let $p_0 \leq q_1 < p_1$. Then it holds

$$\sigma_m\Big(s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\Omega)\,, s^{\bar{0}}_{p_1,q_1} b(\Omega)\,, \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{q_1}}\,, \qquad m \in \mathbb{N}\,.$$

(ii) Now let $p_1 \leq q_1$, and let $\frac{1}{p_0}-\frac{1}{p_1} \leq \frac{1}{q_0}-\frac{1}{q_1}$. Then it holds

$$\sigma_m\Big(s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\Omega)\,, s^{\bar{0}}_{p_1,q_1} b(\Omega)\,, \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}$$

for every integer $m \in \mathbb{N}$.

Proof. **Step 1:** We start with the case $p_1 \leq q_1$ and $\frac{1}{p_0}-\frac{1}{p_1} = \frac{1}{q_0}-\frac{1}{q_1}$. Then we find $p_0 \leq q_0$, and hence by Theorem 6.7.2(ii), Lemma 6.3.1 and Lemma 5.3.3 we have the embeddings

$$s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\nabla) \hookrightarrow s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} b(\nabla) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{p_0}-\frac{1}{p_1}}\big(s^{\bar{0}}_{p_1,q_1} b(\nabla)\,, \mathcal{B}\big) .$$

Step 2: If $\frac{1}{p_0}-\frac{1}{p_1} < \frac{1}{q_0}-\frac{1}{q_1}$ we define q_* this time by $\frac{1}{q_*} = \frac{1}{q_0}-\frac{1}{p_0}+\frac{1}{p_1} > 0$. Then it holds $p_1 \leq q_*$ if, and only if, it holds $p_0 \leq q_0$. In this case, we can conclude from Step 1

$$s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\nabla) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{p_0}-\frac{1}{p_1}}\big(s^{\bar{0}}_{p_1,q_*} b(\nabla)\,, \mathcal{B}\big) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{p_0}-\frac{1}{p_1}}\big(s^{\bar{0}}_{p_1,q_1} b(\nabla)\,, \mathcal{B}\big) .$$

On the other hand, in case $q_0 < p_0$ and $p_1 \leq q_1$, the result once more follows from Corollary 6.6.1 as in the fourth step of the proof of Theorem 6.7.3.

Step 3: Now let $p_0 \leq q_1 < p_1$, and define

$$\frac{1}{p_*} = \frac{1}{p_0} - \frac{1}{q_1} + \frac{1}{p_1} \geq 0\ .$$

Then we obtain from the Franke-Jawerth embedding and Theorem 6.7.2

$$s^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\nabla) \hookrightarrow s^{\bar{d}(\frac{1}{p_0}-\frac{1}{q_1})}_{p_*,p_0} b(\nabla) \hookrightarrow \mathcal{A}_\infty^{\frac{1}{p_0}-\frac{1}{q_1}}\big(s^{\bar{0}}_{p_1,q_1} b(\nabla), \mathcal{B}\big) .$$

The estimates from below follow from (6.5.2) and (6.5.5), respectively. □

Conjecture 6.7.3. Let $0 < p_0 < p_1 \le \infty$ and $0 < q_0, q_1 \le \infty$, where $p_1 \le q_1$. Furthermore, let Ω be an open subset of $\mathbb{R}^d$. Then it holds

$$\sigma_m\Big(s^{\overline{d}(\frac{1}{p_0}-\frac{1}{p_1})}_{p_0,q_0} f(\Omega)\,,\, s^{\overline{0}}_{p_1,q_1} b(\Omega)\,,\mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}} \big(\log m\big)^{(N-1)\big(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1}\big)_+}$$

for every natural number $m \ge 2$.

For completeness we shall add some last limiting cases.

Proposition 6.7.5.

(i) Let $\overline{r} \in \mathbb{R}^N$, $0 < p \le \infty$ ($p < \infty$ for f-spaces) and $0 < q_0, q_1 \le \infty$. For every pair of spaces $s^{\overline{r}}_{p,q_0} x(\nabla)$ and $s^{\overline{r}}_{p,q_1} y(\nabla)$, $x, y \in \{b, f\}$, such that $s^{\overline{r}}_{p,q_0} x(\nabla) \hookrightarrow s^{\overline{r}}_{p,q_1} y(\nabla)$ it holds

$$\sigma_m\Big(s^{\overline{r}}_{p,q_0} x(\nabla)\,,\, s^{\overline{r}}_{p,q_1} y(\nabla)\,,\mathcal{B}\Big) \asymp 1\,, \qquad m \in \mathbb{N}_0\,.$$

(ii) Let $\overline{r} \in \mathbb{R}^N$, $0 < p_1 \le p_0 \le \infty$ ($p_0, p_1 < \infty$ for f-spaces) and $0 < q_0 \le q_1 \le \infty$. Furthermore, let Ω be a bounded open subset of $\mathbb{R}^d$. Then it holds

$$\sigma_m\Big(s^{\overline{r}}_{p_0,q_0} x(\Omega)\,,\, s^{\overline{r}}_{p_1,q_1} x(\Omega)\,,\mathcal{B}\Big) \asymp 1\,, \qquad m \in \mathbb{N}_0\,, \quad x \in \{b, f\}\,.$$

Proof. The conditions in (ii) ensure the embedding $s^{\overline{r}}_{p_0,q_0} x(\Omega) \hookrightarrow s^{\overline{r}}_{p_1,q_1} x(\Omega)$, see Propositions 5.3.2 and 5.3.4. Since $\sigma_m(a, Y, \mathcal{D}) \le \|a|Y\|$ for every $m \in \mathbb{N}_0$, every $a \in X \hookrightarrow Y$, every two quasi-Banach spaces X and Y and every dictionary $\mathcal{D} \subset X$, the estimate from above follows directly from the boundedness of the embedding.
On the other hand, the estimates from below follow from (6.5.2) and Proposition 6.5.1, respectively. □

6.7.5 The case of high smoothness

In this section, we concentrate on the case $\overline{t} = t\overline{d}$. The case of general $\overline{t}$ can be traced back to this one using the elementary embedding from Lemma 5.3.2(ii). The decisive parameter is $\varrho = \min\big\{\frac{t_i}{d_i} : i = 1, \dots, N\big\}$, i.e. our estimates do not change as long as ϱ remains the same. In other words, additional smoothness in only some directions does not improve the approximation quality. The notion "high smoothness" corresponds to the fact that we have to impose certain additional restrictions on t, apart from $t > \big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+$, which is necessary for a compact embedding.

Theorem 6.7.4. Let $t \in \mathbb{R}$ and $0 < p_0, p_1, q_0, q_1 \le \infty$. Furthermore, let

$$\min(p_0, q_0) < \max(p_1, q_1) \qquad \text{and} \qquad t > \frac{1}{\min(p_0, q_0)} - \frac{1}{\max(p_1, q_1)}\,. \tag{6.7.10}$$

Then for all combinations $x, y \in \{b, f\}$, where $p_0 < \infty$ if $x = f$ and $p_1 < \infty$ if $y = f$, every $a \in s^{\overline{t}}_{p_0,q_0} x(\Omega)$ and every natural number $m \ge 2$ there exists an approximation $S_m a \in \Sigma_{c_0 m}(\mathcal{B})$ of a, such that

$$\big\|\, a - S_m a \,|s^{\overline{0}}_{p_1,q_1} y(\Omega)\big\| \le c_1\, m^{-t} \big(\log m\big)^{(N-1)\big(t-\frac{1}{q_0}+\frac{1}{q_1}\big)} \big\|\, a \,|s^{\overline{t}}_{p_0,q_0} x(\Omega)\big\|\,.$$

The constants c_0 and c_1 do not depend on a or m.

Proof. **Step 1.** Let $a \in s^{\bar{t}}_{p_0,q_0}x(\Omega)$ with $\| a \,|s^{\bar{t}}_{p_0,q_0}x(\Omega)\| \le 1$ and $m = \lambda^{N-1}2^\lambda$ for some $\lambda \in \mathcal{N}$. Furthermore we put

$$\begin{aligned} \Lambda_\mu &:= \Big\{(\overline{\nu},k) \in \nabla_\mu :\ |a_{\overline{\nu},k}| \ge \varepsilon_\mu\Big\}\,, \\ \varepsilon_\mu &:= \begin{cases} 0\,, & \mu \le \lambda\,, \\ 2^{\mu\alpha}2^{\lambda\beta}S\big(\lambda,\overline{d}\big)^\eta\,, & \mu > \lambda\,, \end{cases} \qquad \mu \in \mathbb{N}_0\,. \end{aligned} \tag{6.7.11}$$

The parameters α, β and η will be chosen later on. Moreover, we use the abbreviations $\gamma_0 := \min(p_0,q_0)$ and $\delta_1 := \max(p_1,q_1)$. Then we find for $\mu > \lambda$

$$\begin{aligned} \#\Lambda_\mu &= \sum_{(\overline{\nu},k)\in\Lambda_\mu} 1 \le \sum_{(\overline{\nu},k)\in\Lambda_\mu} \frac{|a_{\overline{\nu},k}|^{\gamma_0}}{\varepsilon_\mu^{\gamma_0}} \\ &\le \varepsilon_\mu^{-\gamma_0} \sum_{\overline{\nu}\in M(\mu,\overline{d})} 2^{-\mu(t-1/\gamma_0)\gamma_0}2^{\overline{\nu}\cdot\overline{d}(t-1/\gamma_0)\gamma_0} \sum_{k\in\nabla_{\overline{\nu}}} \big|a_{\overline{\nu},k}\big|^{\gamma_0} \\ &= \varepsilon_\mu^{-\gamma_0}2^{-\mu(t-1/\gamma_0)\gamma_0}\big\|\mathrm{re}_\mu a\big|s^{\bar{t}}_{\gamma_0,\gamma_0}b(\Omega)\big\|^{\gamma_0} \\ &\lesssim \varepsilon_\mu^{-\gamma_0}2^{-\mu(t-1/\gamma_0)\gamma_0}S\big(\mu,\overline{d}\big)^{\gamma_0(1/\gamma_0-1/q_0)}\big\|\mathrm{re}_\mu a\big|s^{\bar{t}}_{p_0,q_0}x(\Omega)\big\|^{\gamma_0} \\ &\le \varepsilon_\mu^{-\gamma_0}2^{-\mu(t-1/\gamma_0)\gamma_0}S\big(\mu,\overline{d}\big)^{\gamma_0(1/\gamma_0-1/q_0)}\big\|\, a\,\big|s^{\bar{t}}_{p_0,q_0}x(\Omega)\big\|^{\gamma_0}\,. \end{aligned}$$

Here we used Lemma 6.4.2, (iii) and (iv), and the estimate (6.4.5). Now we put

$$S_m a := \sum_{\mu=0}^{\infty} \sum_{(\overline{\nu},k)\in\Lambda_\mu} a_{\overline{\nu},k}e^{\overline{\nu},k}\,, \tag{6.7.12}$$

and show, that with suitably chosen parameters α, β, and η the sequence $S_m a$ is a near-best m-term approximation of a.

Step 2. Summing up the result from Step 1 and inserting (6.7.11) we obtain

$$\begin{aligned} \sum_{\mu=0}^{\infty}\#\Lambda_\mu &= \sum_{\mu=0}^{\lambda}\#\nabla_\mu + \sum_{\mu=\lambda+1}^{\infty}\#\Lambda_\mu \lesssim \sum_{\mu=0}^{\lambda} D_\mu + \sum_{\mu=\lambda+1}^{\infty} \varepsilon_\mu^{-\gamma_0}2^{-\mu(t-\frac{1}{\gamma_0})\gamma_0}S\big(\mu,\overline{d}\big)^{\gamma_0(\frac{1}{\gamma_0}-\frac{1}{q_0})} \\ &\lesssim \sum_{\mu=0}^{\lambda} S(\mu,\overline{d})2^\mu + \sum_{\mu=\lambda+1}^{\infty} 2^{-\mu(\alpha+t-\frac{1}{\gamma_0})\gamma_0}2^{-\lambda\beta\gamma_0}S\big(\lambda,\overline{d}\big)^{-\eta\gamma_0}S\big(\mu,\overline{d}\big)^{\gamma_0(\frac{1}{\gamma_0}-\frac{1}{q_0})}. \end{aligned}$$

Now we choose

$$\begin{aligned} &\alpha = -t + \tfrac{1}{\gamma_0} + \vartheta\,, \qquad \vartheta(1-\tfrac{\gamma_0}{\delta_1}) = \tfrac{1}{2}\big(t - \tfrac{1}{\gamma_0} + \tfrac{1}{\delta_1}\big)\,, \\ &\beta = -\tfrac{1}{\gamma_0} - \vartheta \iff -\beta\gamma_0 = 1 + \vartheta\gamma_0\,, \\ &\eta = -\tfrac{1}{q_0} \iff -\eta\gamma_0 = 1 - \gamma_0\big(\tfrac{1}{\gamma_0} - \tfrac{1}{q_0}\big)\,. \end{aligned}$$

In particular, we have $\vartheta > 0$. Then we get from Lemmas 6.4.1 and 6.2.4

$$\sum_{\mu=0}^{\infty}\#\Lambda_\mu \lesssim S\big(\lambda,\overline{d}\big)2^\lambda + 2^{\lambda(1+\vartheta\gamma_0)}S\big(\lambda,\overline{d}\big)^{1-\gamma_0(\frac{1}{\gamma_0}-\frac{1}{q_0})} \sum_{\mu=\lambda+1}^{\infty} 2^{-\mu\vartheta\gamma_0}S\big(\mu,\overline{d}\big)^{\gamma_0(\frac{1}{\gamma_0}-\frac{1}{q_0})}$$

$$\lesssim S\big(\lambda,\overline{d}\big)2^{\lambda} + 2^{\lambda}S\big(\lambda,\overline{d}\big) \le c_0\lambda^{N-1}2^{\lambda}.$$

Hence, S_m is a c_0m-term approximation of a. But for the investigation of the asymptotics of $\sigma_m(a)$ this is sufficient.

Step 3. Now consider first $T_\mu := \sum_{(\overline{\nu},k)\in\nabla_\mu\setminus\Lambda_\mu} a_{\overline{\nu},k}e^{\overline{\nu},k} = \mathrm{re}_\mu(a - S_m a)$. Due to the assumption $\gamma_0 < \delta_1$, we find for $\delta_1 < \infty$

$$|a_{\overline{\nu},k}|^{\delta_1} = |a_{\overline{\nu},k}|^{\delta_1-\gamma_0}|a_{\overline{\nu},k}|^{\gamma_0} \le \varepsilon_\mu^{\delta_1-\gamma_0}|a_{\overline{\nu},k}|^{\gamma_0}\,, \qquad (\overline{\nu},k)\in\nabla_\mu\setminus\Lambda_\mu\,. \tag{6.7.13}$$

Using this estimate, we now obtain with the help of Lemma 6.4.2, Remark 6.4.3 and (6.4.5)

$$\begin{aligned}
\big\|\, T_\mu\,\big|s^{\overline{0}}_{p_1,q_1}y(\Omega)\big\| &\lesssim S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}\big\|\, T_\mu\,\big|s^{\overline{0}}_{\delta_1,\delta_1}b(\Omega)\big\| \\
&= S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}\Bigg(\sum_{(j,k)\in\nabla_\mu\setminus\Lambda_\mu}|a_{j,k}|^{\delta_1}\Bigg)^{1/\delta_1} \\
&\le S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}\Bigg(\sum_{(\overline{\nu},k)\in\nabla_\mu\setminus\Lambda_\mu}2^{-\mu(t-\frac{1}{\gamma_0})\gamma_0}2^{\overline{\nu}\cdot\overline{d}(t-\frac{1}{\gamma_0})\gamma_0}|a_{\overline{\nu},k}|^{\gamma_0}\Bigg)^{\frac{1}{\delta_1}} \\
&\le S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}2^{-\mu(t-\frac{1}{\gamma_0})\frac{\gamma_0}{\delta_1}}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}\Bigg(\sum_{(\overline{\nu},k)\in\nabla_\mu}2^{\overline{\nu}\cdot\overline{d}(t-\frac{1}{\gamma_0})\gamma_0}|a_{\overline{\nu},k}|^{\gamma_0}\Bigg)^{\frac{1}{\gamma_0}\cdot\frac{\gamma_0}{\delta_1}} \\
&= S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}\big\|\mathrm{re}_\mu a\big|s^{\overline{t}}_{\gamma_0,\gamma_0}b(\Omega)\big\|^{\frac{\gamma_0}{\delta_1}} \\
&\lesssim S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}S\big(\mu,\overline{d}\big)^{(\frac{1}{\gamma_0}-\frac{1}{q_0})\frac{\gamma_0}{\delta_1}}\big\|\mathrm{re}_\mu a\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\|^{\frac{\gamma_0}{\delta_1}} \\
&\le S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}\big\|\, a\,\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\|^{\frac{\gamma_0}{\delta_1}}\,.
\end{aligned}$$

In case $\delta_1 = \infty$ we arrive at a similar result upon using the usual modifications and replacing the estimate (6.7.13) by $|a_{\overline{\nu},k}| \le \varepsilon_\mu$.
Now assume $q_1 < \infty$ and $\big\|\, a\,\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\| = 1$. With $T_\mu = 0$ for $\mu \le \lambda$ we find

$$\begin{aligned}
\big\|a - S_m a\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\|^{q_1} &= \Bigg\|\sum_{\mu=0}^{\infty}T_\mu\Bigg|s^{\overline{0}}_{p_1,q_1}b(\Omega)\Bigg\|^{q_1} = \sum_{\mu=\lambda+1}^{\infty}\big\|\, T_\mu\,\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\|^{q_1} \\
&\lesssim \sum_{\mu=\lambda+1}^{\infty}\Big(S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}\varepsilon_\mu^{1-\frac{\gamma_0}{\delta_1}}\Big)^{q_1} \\
&= 2^{\lambda\beta(1-\frac{\gamma_0}{\delta_1})q_1}S\big(\lambda,\overline{d}\big)^{\eta(1-\frac{\gamma_0}{\delta_1})q_1}\sum_{\mu=\lambda+1}^{\infty}\Big(S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}2^{\mu\alpha(1-\frac{\gamma_0}{\delta_1})}\Big)^{q_1} \\
&= 2^{-\lambda(\frac{1}{\gamma_0}+\vartheta)(1-\frac{\gamma_0}{\delta_1})q_1}S\big(\lambda,\overline{d}\big)^{-(1-\frac{\gamma_0}{\delta_1})\frac{q_1}{q_0}} \\
&\qquad\times\sum_{\mu=\lambda+1}^{\infty}\Big(S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}}2^{-\mu t\frac{\gamma_0}{\delta_1}}2^{\mu(-t+\frac{1}{\gamma_0}+\vartheta)(1-\frac{\gamma_0}{\delta_1})}\Big)^{q_1} \\
&= 2^{-\lambda(\frac{1}{\gamma_0}+\vartheta)(1-\frac{\gamma_0}{\delta_1})q_1}S\big(\lambda,\overline{d}\big)^{-(1-\frac{\gamma_0}{\delta_1})\frac{q_1}{q_0}} \\
&\qquad\times\sum_{\mu=\lambda+1}^{\infty}\Big(S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}}2^{-\mu\big(t-\frac{1}{\gamma_0}+\frac{1}{\delta_1}-\vartheta(1-\frac{\gamma_0}{\delta_1})\big)}\Big)^{q_1}
\end{aligned}$$

$$
\begin{aligned}
&\lesssim 2^{-\lambda(\frac{1}{\gamma_0}+\vartheta)(1-\frac{\gamma_0}{\delta_1})q_1} S\big(\lambda,d\big)^{-(1-\frac{\gamma_0}{\delta_1})\frac{q_1}{q_0}} \Big(S\big(\lambda,d\big)^{\frac{1}{q_1}-\frac{\gamma_0}{q_0\delta_1}} 2^{-\frac{1}{2}\lambda(t-\frac{1}{\gamma_0}+\frac{1}{\delta_1})}\Big)^{q_1} \\
&= 2^{-\lambda(\frac{1}{\gamma_0}+\vartheta)(1-\frac{\gamma_0}{\delta_1})q_1} S\big(\lambda,d\big)^{1-\frac{q_1}{q_0}} 2^{-\lambda\vartheta(1-\frac{\gamma_0}{\delta_1})q_1} \\
&= 2^{-\lambda(\frac{1}{\gamma_0}-\frac{1}{\delta_1})q_1} S\big(\lambda,d\big)^{1-\frac{q_1}{q_0}} 2^{-\lambda(t-\frac{1}{\gamma_0}+\frac{1}{\delta_1})q_1} = 2^{-\lambda t q_1} S\big(\lambda,d\big)^{1-\frac{q_1}{q_0}} .
\end{aligned}
$$

In case $y=f$, we replace the application of (6.4.2) at the beginning by (6.4.4). Moreover, the modification in case $q_1=\infty$ is obvious. Finally, inserting $\lambda \sim \log m$, $2^\lambda \sim m(\log m)^{-(N-1)}$ and $S(\lambda,\overline{d}) \sim \lambda^{N-1}$, we obtain

$$
\big\| a - S_m a \big| s^{\overline{0}}_{p_1,q_1} b(\Omega)\big\| \le c\, m^{-t}\big(\log m\big)^{(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})} .
$$

For all other m, the result now follows by monotonicity arguments. □

Remark 6.7.8. The assumption (6.7.10) implies in particular $t > 1/p_0 - 1/p_1$, and from $\min(p_0,q_0) < \max(p_1,q_1)$ follows $t>0$. Hence we recover the condition $t > \big(\frac{1}{p_0}-\frac{1}{p_1}\big)_+$ for the compact embedding. Moreover, we obtain $t - 1/q_0 + 1/q_1 > 0$, so there is no contradiction to the estimates from below in Proposition 6.5.1.

Remark 6.7.9. With a standard soft thresholding argument we can always find a modification $\widetilde{S_m a}$ of $S_m a$, such that the components depend continuously on a. More precisely, one has to replace the numbers $a_{\overline{\nu},k}$ in (6.7.12) by $\widetilde{a}_{\overline{\nu},k}$, where

$$
\widetilde{a}_{\overline{\nu},k} = g\big(a_{\overline{\nu},k}\big) , \qquad g(\xi) = \begin{cases} 0\,, & \xi \le 0\,, \\ 2(\xi-1)\,, & 1<\xi<2\,, \\ \xi\,, & \xi \ge 2\,. \end{cases}
$$

For this approximant one can prove the same error estimate even with the same constant c_1 by a slight modification of the estimate of $\big\|\, T_\mu \,\big| s^{\overline{0}}_{p_1,q_1} y(\Omega)\big\|$ in the above proof.

Remark 6.7.10. The construction and the corresponding estimates are based on results for the isotropic case in [12]. Moreover, the idea of using building blocks goes back to [94].

Theorem 6.7.5. Let $t \in \mathbb{R}$ and $0 < p_0,p_1,q_0,q_1 \le \infty$. Furthermore, let

$$
\min(p_0,q_0) \ge \max(p_1,q_1) \qquad \text{and} \qquad t>0\,. \tag{6.7.14}
$$

Consider the operators P_λ, $\lambda \in \mathcal{N}$, defined by

$$
P_\lambda a = \sum_{\mu=0}^{\lambda} \mathrm{re}_\mu a = \sum_{\mu=0}^{\lambda} \sum_{\overline{\nu}\in M(\mu,\overline{d})} \sum_{k\in\nabla_{\overline{\nu}}} a_{\overline{\nu},k} e^{\overline{\nu},k} , \qquad a = \big(a_{\overline{\nu},k}\big)_{\overline{\nu}\in\mathbb{N}_0^N, k\in\nabla_{\overline{\nu}}} .
$$

Then for all combinations $x,y\in\{b,f\}$, where $p_0<\infty$ if $x=f$ and $p_1<\infty$ if $y=f$, it holds

$$
\Big\| \mathrm{id} - P_\lambda \,:\, s^{\overline{t}}_{p_0,q_0} x(\Omega) \longrightarrow s^{\overline{0}}_{p_1,q_1} y(\Omega)\Big\| \lesssim 2^{-\lambda t} S\big(\lambda,\overline{d}\big)^{-\frac{1}{q_0}+\frac{1}{q_1}} . \tag{6.7.15}
$$

Proof. From Lemma 6.4.2, (iii) and (iv), Hölder's inequality and the assumption $\gamma_0 := \min(p_0, q_0) \geq \max(p_1, q_1) =: \delta_1$ we obtain at first

$$\begin{aligned}\big\|\mathrm{re}_\mu a\big|s^{\overline{0}}_{p_1,q_1}y(\Omega)\big\| &\lesssim S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}\big\|\mathrm{re}_\mu a\big|s^{\overline{0}}_{\delta_1,\delta_1}b(\Omega)\big\| = S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}\big\|\mathrm{re}_\mu a\big|\ell^{D_\mu}_{\delta_1}\big\| \\ &\leq S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}D_\mu^{\frac{1}{\delta_1}-\frac{1}{\gamma_0}}\big\|\mathrm{re}_\mu a\big|\ell^{D_\mu}_{\gamma_0}\big\| \\ &= S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}2^{-\mu/\delta_1}D_\mu^{\frac{1}{\delta_1}-\frac{1}{\gamma_0}}2^{-\mu(t-\frac{1}{\gamma_0})}\big\|\mathrm{re}_\mu a\big|s^{\overline{t}}_{\gamma_0,\gamma_0}b(\Omega)\big\| \\ &\lesssim S\big(\mu,d\big)^{\frac{1}{q_1}-\frac{1}{\delta_1}}D_\mu^{\frac{1}{\delta_1}-\frac{1}{\gamma_0}}2^{-\mu(t-\frac{1}{\gamma_0}+\frac{1}{\delta_1})}S\big(\mu,d\big)^{\frac{1}{\gamma_0}-\frac{1}{q_0}}\big\|\mathrm{re}_\mu a\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\|\,.\end{aligned}$$

As $D_\mu \asymp S(\mu,d)2^\mu$, this simplifies to

$$\big\|\mathrm{re}_\mu a\big|s^{\overline{0}}_{p_1,q_1}y(\Omega)\big\| \lesssim S\big(\mu,d\big)^{\frac{1}{q_1}-\frac{1}{q_0}}2^{-\mu t}\big\|\,a\,\big|s^{\overline{t}}_{p_0,q_0}x(\Omega)\big\|\,.$$

Summing up this estimate we find using (6.4.2) and Lemma 6.4.1

$$\begin{aligned}\big\|a - P_\lambda a\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\|^{q_1} &= \Big\|\sum_{\mu=\lambda+1}^{\infty}\mathrm{re}_\mu a\Big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\Big\|^{q_1} = \sum_{\mu=\lambda+1}^{\infty}\big\|\mathrm{re}_\mu a\big|s^{\overline{0}}_{p_1,q_1}b(\Omega)\big\|^{q_1} \\ &\lesssim \sum_{\mu=\lambda+1}^{\infty}\Big(S\big(\mu,\overline{d}\big)^{\frac{1}{q_1}-\frac{1}{q_0}}2^{-\mu t}\Big)^{q_1}\big\|\,a\,\big|s^{\overline{t}}_{p_0,q_0}b(\Omega)\big\|^{q_1} \\ &\lesssim 2^{-\lambda t q_1}S\big(\lambda,\overline{d}\big)^{(\frac{1}{q_1}-\frac{1}{q_0})q_1}\big\|\,a\,\big|s^{\overline{t}}_{p_0,q_0}b(\Omega)\big\|^{q_1}\,.\end{aligned}$$

In case $y = f$, we argue similarly using (6.4.4). □

Remark 6.7.11. The assumption (6.7.14) implies in particular $q_0 \geq q_1$. Hence, the exponent of the logarithmic term is again positive, so there is no contradiction to Proposition 6.5.1.

Remark 6.7.12. Partial sum operators for wavelets and other basis systems are well-studied objects in approximation theory. The operators defined in Theorem 6.7.5 correspond to so-called hyperbolic cross approximation. This notion had been introduced in connection with approximation by multivariate trigonometric polynomials, but it has a non-periodic counterpart for tensor product systems. Previous results for tensor product wavelet systems in $L_p(\mathbb{R}^d)$ can be found in [16].

We add a reformulation of the above theorems, which combines both results.

Corollary 6.7.2. Let $t \in \mathbb{R}$ and $0 < p_0, p_1, q_0, q_1 \leq \infty$, where

$$t > \left(\frac{1}{\min(p_0,q_0)} - \frac{1}{\max(p_1,q_1)}\right)_+ . \tag{6.7.16}$$

Then for all combinations $x, y \in \{b, f\}$, where $p_0 < \infty$ if $x = f$ and $p_1 < \infty$ if $y = f$, it holds

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}x(\Omega)\,,s^{\overline{0}}_{p_1,q_1}y(\Omega)\,,\mathcal{B}\Big) \lesssim m^{-t}\big(\log m\big)^{(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}\,, \qquad m \geq 2\,.$$

Notice, that we do not need any assumption on the relation of $\min(p_0,q_0)$ and $\max(p_1,q_1)$.

Proof. From $\#\nabla_\mu \le c\,\mu^{N-1}2^\mu$ we obtain at once, that $P_\lambda a$ is a linear combination of at most $c_0'\lambda^{N-1}2^\lambda$ elements of $\mathcal{B}$, and hence it is also suited for the estimate of the asymptotic behaviour of $\sigma_m\big(a, s^{\overline{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)$. The estimate for the m-term width now follows immediately from the results in Theorems 6.7.4 and 6.7.5, where (6.7.15) is extended to arbitrary $m \ge 2$ using monotonicity arguments. $\square$

The following corollary re-interprets the result of Theorem 6.7.5 as an estimate for the approximation numbers. We remind on their definition in Section 6.1.

Corollary 6.7.3. Let $t \in \mathbb{R}$ and $0 < p_0, p_1, q_0, q_1 \le \infty$ satisfying (6.7.14). Then it holds for $m \ge 2$

$$a_m\big(\,\mathrm{id}\,:\, s^{\overline{t}}_{p_0,q_0}x(\Omega) \longrightarrow s^{\overline{0}}_{p_1,q_1}y(\Omega)\,\big) \lesssim m^{-t}\big(\log m\big)^{(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}\,.$$

Proof. As in the proof of the last corollary, we obtain from $\#\nabla_\mu \le c\,\mu^{N-1}2^\mu$ that the rank of P_λ is at most $c_0'\lambda^{N-1}2^\lambda$. The estimate for general m then follows once more by monotonicity arguments. $\square$

Remark 6.7.13. Similar results as in the Theorems 6.7.4 and 6.7.5 can be obtained without the additional restriction on t, i.e. for every $t > \big(\frac{1}{p_0}-\frac{1}{p_1}\big)_+$. On the other hand this modification yields worse exponents for the logarithmic factor. The same approximants $S_m a$ and $P_\lambda a$ show the error estimate

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}x(\Omega)\,,\, s^{\overline{0}}_{p_1,q_1}y(\Omega)\,,\mathcal{B}\Big) \le c\,m^{-t}\big(\log m\big)^{(N-1)\left(t-\frac{1}{\max(p_0,q_0)}+\frac{1}{\min(p_1,q_1)}\right)}\,,$$

for $m \ge 2$ and all $0 < p_0, p_1, q_0, q_1 \le \infty$. An according result holds for the approximation numbers in case $p_1 \le p_0$.

6.8 Gagliardo-Nirenberg-type inequalities

With the notion "Gagliardo-Nirenberg-type inequality" we refer to an estimate of the form

$$\big\|\, f\,|X\big\| \le c\,\big\|\, f\,|X_0\big\|^{1-\Theta}\big\|\, f\,|X_1\big\|^{\Theta}\,, \qquad f \in X_0 \cap X_1\,,$$

for some $0 < \Theta < 1$. Here $\{X_0, X_1\}$ is an interpolation couple of quasi-Banach spaces and X is an intermediate space, i.e. $X_0 \cap X_1 \hookrightarrow X \hookrightarrow X_0 + X_1$.

Proposition 6.8.1. Let $0 < p_0, p_1 < \infty$, $0 < q_0, q_1 \le \infty$, $r_0, r_1 \in \mathbb{R}$ and $0 < \Theta < 1$. Furthermore, let Ω be an open subset of $\mathbb{R}^d$. We put

$$\frac{1}{p} = \frac{1-\Theta}{p_0} + \frac{\Theta}{p_1} \qquad \text{and} \qquad \overline{r} = (1-\Theta)\,\overline{r}^{\,0} + \Theta\,\overline{r}^{\,1}\,,$$

where $\overline{r}^{\,0} = r_0\overline{d}$, $\overline{r}^{\,1} = r_1\overline{d}$ and $\overline{r} = r\overline{d}$. Then the following assertions are equivalent for $N \ge 2$:

(i) It holds

$$\frac{1}{q} \le \frac{1-\Theta}{q_0} + \frac{\Theta}{q_1}\,, \qquad 0 < q \le \infty\,. \tag{6.8.1}$$

(ii) There is some positive constant c_1, such that

$$\big\| \,\lambda\, \big|\ell_q^{\overline{r}}(\mathbb{N}_0^N)\big\| \le c_1 \,\big\| \,\lambda\, \big|\ell_{q_0}^{\overline{r}^0}(\mathbb{N}_0^N)\big\|^{1-\Theta} \big\| \,\lambda\, \big|\ell_{q_1}^{\overline{r}^1}(\mathbb{N}_0^N)\big\|^{\Theta} \tag{6.8.2}$$

holds for all $\lambda \in \ell_{q_0}^{\overline{r}^0}(\mathbb{N}_0^N) \cap \ell_{q_1}^{\overline{r}^1}(\mathbb{N}_0^N)$.

(iii) There is some positive constant c_2, such that

$$\big\| \,\lambda\, \big|\ell_q\big(\mathbb{N}_0^N, \ell_p(\nabla)\big)\big\| \le c_2 \,\big\| \,\lambda\, \big|\ell_{q_0}\big(\mathbb{N}_0^N, \ell_{p_0}(\nabla)\big)\big\|^{1-\Theta} \big\| \,\lambda\, \big|\ell_{q_1}\big(\mathbb{N}_0^N, \ell_{p_1}(\nabla)\big)\big\|^{\Theta} \tag{6.8.3}$$

holds for all $\lambda \in \ell_{q_0}(\mathbb{N}_0^N, \ell_{p_0}(\nabla)) \cap \ell_{q_1}(\mathbb{N}_0^N, \ell_{p_1}(\nabla))$.

(iv) There is some positive constant c_3, such that

$$\big\| \,\lambda\, \big|s_{p,q}^{\overline{r}} f(\Omega)\big\| \le c_3 \,\big\| \,\lambda\, \big|s_{p_0,q_0}^{\overline{r}^0} f(\Omega)\big\|^{1-\Theta} \big\| \,\lambda\, \big|s_{p_1,q_1}^{\overline{r}^1} f(\Omega)\big\|^{\Theta} \tag{6.8.4}$$

holds for all $\lambda \in s_{p_0,q_0}^{\overline{r}^0} f(\Omega) \cap s_{p_1,q_1}^{\overline{r}^1} f(\Omega)$.

(v) There is some positive constant c_4, such that

$$\big\| \,\lambda\, \big|s_{p,q}^{\overline{r}} b(\Omega)\big\| \le c_4 \,\big\| \,\lambda\, \big|s_{p_0,q_0}^{\overline{r}^0} b(\Omega)\big\|^{1-\Theta} \big\| \,\lambda\, \big|s_{p_1,q_1}^{\overline{r}^1} b(\Omega)\big\|^{\Theta} \tag{6.8.5}$$

holds for all $\lambda \in s_{p_0,q_0}^{\overline{r}^0} b(\Omega) \cap s_{p_1,q_1}^{\overline{r}^1} b(\Omega)$.

Proof. We show two chains of implications, at first (i)$\Longrightarrow$(iv)$\Longrightarrow$(ii)$\Longrightarrow$(i) and then (i)$\Longrightarrow$(v)$\Longrightarrow$(ii). Finally, we show (i) $\Longleftrightarrow$ (iii).

Step 1: (i)$\Longrightarrow$(ii) follows by using the monotonicity of the ℓ_q-spaces and applying Hölder's inequality twice. Now let $b = (b_{\overline{\nu}})_{\overline{\nu}\in\mathbb{N}_0^N}$ be an arbitrary sequence of complex numbers. Then define a by

$$a_{\overline{\nu},k} = \begin{cases} b_{\overline{\nu}}\,, & \overline{\nu} \in \mathbb{N}_0^N\,,\ Q_{\overline{\nu},k} \subset [0,1]^d\,, \\ 0\,, & \text{else}\,. \end{cases}$$

By suitable dilation and translation we can always achieve $[0,1]^d \subset \Omega$, because as an open set Ω contains a dyadic rectangle. A simple calculation shows $\big\| \,a\, \big|s_{p,q}^{\overline{r}} f(\Omega)\big\| = \big\| \,b\, \big|\ell_q^{\overline{r}}(\mathbb{N}_0^N)\big\|$, similarly for the other spaces under consideration, and hence (ii)$\Longrightarrow$(iv). Finally, consider sequences a^n, defined by

$$(a^n)_{\overline{\nu},k} = \begin{cases} 1\,, & \overline{\nu} \in \mathbb{N}_0^N\,, \overline{\nu}\cdot\overline{d} = n\,, \\ 0\,, & \text{else}\,. \end{cases}$$

An easy calculation shows $\big\|a^n\big|\ell_q^{\overline{r}}(\mathbb{N}_0^N)\big\| = 2^{rn} S(n,\overline{d})^{1/q}$. Hence (6.8.2) implies $S(n,\overline{d})^{1/q} \le c_1 S(n,\overline{d})^{\frac{1-\Theta}{q_0}} S(n,\overline{d})^{\frac{\Theta}{q_1}}$. Considering only $n \in \mathcal{N}$, we know $S(n,\overline{d}) \sim n^{N-1}$, and thus for this inequality to hold the condition (6.8.1) is required.

Step 2: The conclusion (i)$\Longrightarrow$(v) is again a matter of Hölder's inequality. For the same sequence a as in Step 1 we also find $\| a \,|s^{\overline{r}}_{p,q}b(\Omega)\| = \| b \,|\ell^{\overline{r}}_q(\mathbb{N}^N_0)\|$ for all $\overline{r}, p, q$, and hence (v)$\Longrightarrow$(ii).

Step 3: Once more we immediately obtain (i)$\Longrightarrow$(iii) by applying Hölder's inequality twice. The reverse implication follows by similar arguments as at the end of Step 1 for sequences $\alpha^n = \sum_{\overline{\nu}\in M(n,\overline{d})} e^{\overline{\nu},k_{\overline{\nu}}}$, $n \in \mathcal{N}$, for arbitrary $k_{\overline{\nu}} \in \nabla_{\overline{\nu}}$. $\square$

Remark 6.8.1. If we assume (6.8.1), then the inequalities (6.8.2)–(6.8.5) hold with $c_1 = c_2 = c_3 = c_4 = 1$, and this remains valid for spaces $\ell_q(\mathbb{N}^N_0, \ell_p(\nabla))$, $s^{\overline{r}}_{p,q}b(\nabla)$ and $s^{\overline{r}}_{p,q}f(\nabla)$ for arbitrary ∇ and even arbitrary $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$.

Remark 6.8.2. The statement (i) $\Longleftrightarrow$ (iv) is in sharp contrast to the isotropic case (i.e. $N = 1$), where the counterpart of (6.8.4) is valid for all parameters $0 < q, q_0, q_1 \leq \infty$. We refer e.g. to Brezis and Mironescu [10].

The consequences of these Gagliardo-Nirenberg-type inequalities for real interpolation of the sequence spaces under consideration as well as the problem of m-term approximation will be discussed in the next sections.

6.9 Real interpolation of sequence spaces

Our aim for the remainder of this chapter is to weaken the restriction (6.7.16) in Corollary 6.7.2 for the estimate of the m-term approximation. One of the tools in this attempt will be real interpolation. To this purpose we have to deal with interpolation assertions for the various types of sequence spaces under consideration.
We first collect some known results on interpolation of weighted and vector-valued ℓ_p-spaces. We refer to Triebel [82, Section 1.18] and Bergh/Löfström [7], for proofs and further references.

Definition 6.9.1.

(i) Let I be a countable index set, and let A_i, $i \in I$, be Banach spaces. Then we define $\ell_p(A_j)$, $1 \leq p \leq \infty$, to be the collection of all sequences $a = (a_i)_{i\in I}$, $a_i \in A_i$, such that
$$\| a \,|\ell_p(A_i)\| := \Big\| \big(\| a_i \,|A_i\|\big)_{i\in I} \Big| \ell_p \Big\| < \infty \,.$$

(ii) Let $1 \leq p \leq \infty$ and $\overline{s} \in \mathbb{R}^N$. Furthermore, let $A_{\overline{\nu}}$, $\overline{\nu} \in \mathbb{N}^N_0$, be Banach spaces. Then $\ell^{\overline{s}}_p(A_{\overline{\nu}})$ is the collection of all sequences $a = (a_{\overline{\nu}})_{\overline{\nu}\in\mathbb{N}^N_0}$, $a_{\overline{\nu}} \in A_{\overline{\nu}}$, such that
$$\| a \,|\ell^{\overline{s}}_p(A_{\overline{\nu}})\| := \Big\| \big(2^{\overline{\nu}\cdot\overline{s}}\| a_{\overline{\nu}} \,|A_{\overline{\nu}}\|\big)_{\overline{\nu}\in\mathbb{N}^N_0} \Big| \ell_p \Big\| < \infty \,.$$

(iii) If $A_{\overline{\nu}} = A$ for all $\overline{\nu} \in \mathbb{N}^N_0$, we shall write $\ell^{\overline{s}}_p(A)$. In particular, if $A = \mathbb{C}$, then $\ell^{\overline{s}}_p(\mathbb{C}) \equiv \ell^{\overline{s}}_p$.

(iv) Let A be a Banach space and $1 \le p \le \infty$. Let $(\Omega, \mathfrak{B}, \mu)$ be a σ-finite complete measure space. Then $L_p(A) = L_p(A, \Omega, \mathfrak{B}, \mu)$ is the space of all A-valued measurable functions f on Ω, such that

$$\| f \,|L_p(A)\| := \Big\| \|f(\cdot)|A\| \Big| L_p(\Omega) \Big\| = \left(\int_\Omega \|f(x)|A\|^p d\mu(x) \right)^{\frac{1}{p}} < \infty\,.$$

Theorem 6.9.1.

(i) Let $\{A_i, B_i\}$ be interpolation couples for all $i \in I$, where I is a countable index set. Moreover, let $1 \le p_0, p_1 < \infty$, $0 < \Theta < 1$, and put

$$\frac{1}{p} = \frac{1-\Theta}{p_0} + \frac{\Theta}{p_1}\,.$$

Then it holds

$$\Big(\ell_{p_0}(A_i)\,, \ell_{p_1}(B_i)\Big)_{\Theta,p} = \ell_p\Big(\big(A_i, B_i\big)_{\Theta,p}\Big)\,. \tag{6.9.1}$$

(ii) Let A_j, $j \in \mathbb{N}_0$, be Banach spaces, $0 < q_0, q_1, q \le \infty$, $s_0, s_1 \in \mathbb{R}$, $s_0 \neq s_1$, and $0 < \Theta < 1$. Then we find

$$\Big(\ell^{s_0}_{q_0}(A_j)\,, \ell^{s_1}_{q_1}(A_j)\Big)_{\Theta,q} = \ell^s_q(A_j)\,, \tag{6.9.2}$$

where $s := (1-\Theta)s_0 + \Theta s_1$.

Remark 6.9.1. In [82], part (i) is formulated only for $I = \mathbb{N}$, whereas part (ii) can be found in [83], where it is stated only in the case $A_j = A$ for all j. But both proofs can be extended to the above generalizations.

Theorem 6.9.2.

(i) Let $\{A, B\}$ be an interpolation couple of Banach spaces. Let $1 \le p_0, p_1 < \infty$, $0 < \Theta < 1$ and

$$\frac{1}{p} = \frac{1-\Theta}{p_0} + \frac{\Theta}{p_1}\,. \tag{6.9.3}$$

Then it holds

$$\Big(L_{p_0}(A)\,, L_{p_1}(B)\Big)_{\Theta,p} = L_p\big((A,B)_{\Theta,p}\big)\,. \tag{6.9.4}$$

(ii) Let additionally $1 \le p \le q < \infty$. Then we have the embedding

$$L_p\big((A,B)_{\Theta,q}\big) \hookrightarrow \Big(L_{p_0}(A)\,, L_{p_1}(B)\Big)_{\Theta,q}\,. \tag{6.9.5}$$

(iii) Let $\{A_i, B_i\}$ be interpolation couples for all $i \in I$, where I is a countable index set. Moreover, let $1 \leq p_0, p_1 < \infty$, $0 < \Theta < 1$, and define p be as before. Then it holds for every $p \leq q < \infty$

$$\ell_p\big((A_i, B_i)_{\Theta,q}\big) \hookrightarrow \Big(\ell_{p_0}(A_i)\,,\ell_{p_1}(B_i)\Big)_{\Theta,q}\,. \tag{6.9.6}$$

Proof. Part (i) is the famous interpolation theorem of Peetre and Lions [49]. A proof for (ii) can be found in [34]. (iii) is some discrete version of (ii) and follows analogously (where now the proof of Theorem 1.18.1 in [82] has to be modified in a similar way). □

After these preparations, we are able to formulate our results on interpolation of Besov- and Triebel-Lizorkin-type sequence spaces.

Theorem 6.9.3.

(i) Let $A_{\overline{\nu}}$, $\overline{\nu} \in \mathbb{N}_0^N$, be Banach spaces, $0 < q_0, q_1 < \infty$, $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$, and $0 < \Theta < 1$. Furthermore, let

$$\frac{1}{q} := \frac{1-\Theta}{q_0} + \frac{\Theta}{q_1} \qquad \text{and} \qquad \overline{r} := (1-\Theta)\overline{r}^{\,0} + \Theta\overline{r}^{\,1}\,. \tag{6.9.7}$$

Then it holds

$$\Big(\ell_{q_0}^{\overline{r}^{\,0}}(A_{\overline{\nu}})\,,\ell_{q_1}^{\overline{r}^{\,1}}(A_{\overline{\nu}})\Big)_{\Theta,q} = \ell_q^{\overline{r}}(A_{\overline{\nu}})\,. \tag{6.9.8}$$

(ii) Let $0 < q_0, q_1 < \infty$, $0 < p \leq \infty$, $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$, and $0 < \Theta < 1$. Furthermore, let q and $\overline{r}$ as in (6.9.7). Then

$$\Big(s_{p,q_0}^{\overline{r}^{\,0}}b(\nabla)\,, s_{p,q_1}^{\overline{r}^{\,1}}b(\nabla)\Big)_{\Theta,q} = s_{p,q}^{\overline{r}}b(\nabla)\,. \tag{6.9.9}$$

(iii) Let $0 < p_0, p_1 \leq \infty$, $p_0 \neq p_1$, $0 < q_0, q_1 < \infty$, and $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$. Furthermore, let $0 < \Theta < 1$ and define p as in (6.9.3), and q and $\overline{r}$ as in (6.9.7). Then it holds

$$s_{p,q}^{\overline{r}}b(\nabla) \hookrightarrow \Big(s_{p_0,q_0}^{\overline{r}^{\,0}}b(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}b(\nabla)\Big)_{\Theta,q} \iff p \leq q\,, \tag{6.9.10}$$

$$\Big(s_{p_0,q_0}^{\overline{r}^{\,0}}b(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}b(\nabla)\Big)_{\Theta,q} \hookrightarrow s_{p,q}^{\overline{r}}b(\nabla) \iff q \leq p\,. \tag{6.9.11}$$

Additionally, for $q \leq p$ we obtain the embedding

$$s_{p,q}^{\overline{r}}b(\nabla) \hookrightarrow \Big(s_{p_0,q_0}^{\overline{r}^{\,0}}b(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}b(\nabla)\Big)_{\Theta,p}\,. \tag{6.9.12}$$

(iv) Let $0 < p_0, p_1, q_0, q_1 < \infty$, $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$, and $0 < \Theta < 1$. Furthermore, define p as in (6.9.3), and q and $\overline{r}$ as in (6.9.7). Then it holds

$$s_{p,q}^{\overline{r}}f(\nabla) \hookrightarrow \Big(s_{p_0,q_0}^{\overline{r}^{\,0}}f(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}f(\nabla)\Big)_{\Theta,p} \iff q \leq p\,, \tag{6.9.13}$$

$$\Big(s_{p_0,q_0}^{\overline{r}^{\,0}}f(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}f(\nabla)\Big)_{\Theta,p} \hookrightarrow s_{p,q}^{\overline{r}}f(\nabla) \iff p \leq q\,, \tag{6.9.14}$$

and for $p \leq q$ we obtain the embedding

$$s_{p,q}^{\overline{r}}f(\nabla) \hookrightarrow \Big(s_{p_0,q_0}^{\overline{r}^{\,0}}f(\nabla)\,, s_{p_1,q_1}^{\overline{r}^{\,1}}f(\nabla)\Big)_{\Theta,q}\,. \tag{6.9.15}$$

Slightly weaker versions of (i) and (ii) (for function spaces) in case $N = 2$ can be found in [70].

Proof. **Step 1:** At first, we consider Banach spaces only, i.e. we additionally assume $1 \le p_0, p_1, p, q_0, q_1, q \le \infty$. Then (6.9.8) follows from $\ell_q^{r_1,\ldots,r_N}(A_{\overline{\nu}}) = \ell_q^{r_N}\big(\ell_q^{r_1,\ldots,r_{N-1}}(A_{\overline{\nu}})\big)$ and (6.9.1), we remind on the identities

$$K(t, a; 2^{js_0}X, 2^{js_1}Y) = 2^{js_0}K(2^{j(s_1-s_0)}t, a; X, Y)\,, \qquad t > 0\,, s_0, s_1 \in \mathbb{R}\,,$$

for the K-functional of Peetre, and hence

$$\big(2^{js_0}X, 2^{js_1}Y\big)_{\Theta,q} = 2^{js}\big(X, Y\big)_{\Theta,q}\,, \qquad s = (1-\Theta)s_0 + \Theta s_1\,,$$

for an arbitrary interpolation couple $\{X, Y\}$ of Banach spaces. The identity (6.9.9) is a special case of (i) for $A_{\overline{\nu}} = \ell_p(\nabla_{\overline{\nu}})$.
The identities (6.9.10) and (6.9.11) follow from (6.9.1) (with $q \mathrel{\widehat{=}} p$, $q_0 \mathrel{\widehat{=}} p_0$, $q_1 \mathrel{\widehat{=}} p_1$) by an argument similar to the one for (i), together with the well-known facts $(\ell_{p_0}(I), \ell_{p_1}(I))_{\Theta,q} = \ell_{p,q}(I)$ and $\ell_p(I) \hookrightarrow \ell_{p,q}(I) \iff p \le q$ as well as $\ell_{p,q}(I) \hookrightarrow \ell_p(I) \iff q \le p$. The embedding (6.9.12) follows as in the proof of (i) by iterated usage of (6.9.6).
The proof of (iv) follows by similar arguments. For $q \le p$ we use the elementary embedding $\big(\ell_{q_0}^{\overline{r}^0}, \ell_{q_1}^{\overline{r}^1}\big)_{\Theta,q} \hookrightarrow \big(\ell_{q_0}^{\overline{r}^0}, \ell_{q_1}^{\overline{r}^1}\big)_{\Theta,p}$ for interpolation spaces. Now the Peetre-Lions-formula (6.9.4) together with (6.9.8) imply (6.9.13), using the usual retraction-coretraction arguments (for the general results, see e.g. [7, Theorem 6.4.2] or [82, Theorem 1.2.4]), since directly from the definition we conclude that $s^{\overline{r}}_{p,q}f(\nabla)$ is isometrically isomorphic to a closed subspace of $L_p(\ell_q^{\overline{r}})$. Likewise (6.9.14) can be obtained from $\big(\ell_{q_0}^{\overline{r}^0}, \ell_{q_1}^{\overline{r}^1}\big)_{\Theta,p} \hookrightarrow \big(\ell_{q_0}^{\overline{r}^0}, \ell_{q_1}^{\overline{r}^1}\big)_{\Theta,q}$ for $p \le q$. Finally, (6.9.15) follows from (6.9.5), using once more (6.9.8) with $A_{\overline{\nu}} = \mathbb{C}$.

Step 2: We remove the restrictions on the parameters.
In this step the lattice property (see Lemma 6.2.1) of the sequence spaces will be crucial. Given any sequence $a = (a_{\overline{\nu},k})_{\overline{\nu}\in\mathbb{N}_0^N, k\in\nabla_{\overline{\nu}}}$ we define

$$|a| = \big(|a_{\overline{\nu},k}|\big)_{\overline{\nu}\in\mathbb{N}_0^N, k\in\nabla_{\overline{\nu}}} \qquad \text{and} \qquad |a|^\varepsilon = \big(|a_{\overline{\nu},k}|^\varepsilon\big)_{\overline{\nu}\in\mathbb{N}_0^N, k\in\nabla_{\overline{\nu}}}\,.$$

These definitions together with the definitions of the respective quasi-norms immediately yield

$$\big\|\, |a|\, \big| s^r_{p,q}x(\nabla)\big\| = \big\|\, a\, \big| s^r_{p,q}x(\nabla)\big\|$$

as well as for every $\varepsilon > 0$

$$\big\|\, |a|^\varepsilon\, \big| s^{\overline{r}\varepsilon}_{p/\varepsilon,q/\varepsilon}x(\nabla)\big\| = \big\|\, |a|\, \big| s^{\overline{r}}_{p,q}x(\nabla)\big\|^\varepsilon = \big\|\, a\, \big| s^{\overline{r}}_{p,q}x(\nabla)\big\|^\varepsilon\,.$$

In [34], we derived identities for the K-functional with respect to two of the isotropic sequence spaces on the basis of the lattice property. Since no other property of these sequence spaces was used in that proof, the results can immediately be extended to the spaces $s^{\overline{r}}_{p,q}x(\nabla)$, $x \in \{b, f\}$. In this way we obtain

$$K\big(t, a\,;\, s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big) = K\big(t, |a|\,;\, s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big)$$

for every $t > 0$, and for every $0 < \varepsilon \le 1$ we find

$$K\big(|a|^\varepsilon, t\,;\, s^{\bar{r}^0\varepsilon}_{p_0/\varepsilon,q_0/\varepsilon}x(\nabla)\,,\, s^{\bar{r}^1\varepsilon}_{p_1/\varepsilon,q_1/\varepsilon}x(\nabla)\big) \asymp K\big(t^{1/\varepsilon}, |a|\,;\, s^{\bar{r}^0}_{p_0,q_0}x(\nabla)\,,\, s^{\bar{r}^1}_{p_1,q_1}x(\nabla)\big)^\varepsilon\,, \qquad t > 0\,.$$

Choosing $0 < \varepsilon < \min(1, p, p_0, p_1, q_0, q_1, q)$ and inserting the identities for the K-functional into the definition of the quasi-norm of the respective interpolation spaces, the embedding results in (ii)–(iv) for parameters $0 < p, p_0, p_1, q_0, q_1, q \le \infty$ now follow from the results of the first step for $1 < p/\varepsilon, p_0/\varepsilon, p_1/\varepsilon, q_0/\varepsilon, q_1/\varepsilon, q/\varepsilon \le \infty$. □

Although Theorem 6.9.3 suffices for most purposes, its restriction to finite parameters is often inconvenient. Hence we look for an embedding result, which will be only slightly weaker than the previous theorem for finite parameters, but which will also admit infinite parameters. At this point the Gagliardo-Nirenberg-type inequalities (Proposition 6.8.1) come into play. We need two preparatory results first. The first one deals with interrelations between interpolation theory and duality, see [82, Theorem 1.11.2].

Lemma 6.9.1. Let $\{A_0, A_1\}$ be an interpolation couple of Banach spaces, such that $A_0 \cap A_1$ is dense in both A_0 and A_1. Moreover, let $0 < \Theta < 1$ and $1 \le q < \infty$. Then it holds

$$\Big(\big(A_0\,, A_1\big)_{\Theta,q}\Big)' = \big(A_1'\,, A_0'\big)_{1-\Theta,q'} = \big(A_0'\,, A_1'\big)_{\Theta,q'}\,.$$

The next lemma is another well-known assertion in interpolation theory and can be found, e.g., in [6, Proposition 5.2.10].

Lemma 6.9.2. Let $\{X_0, X_1\}$ be an interpolation couple of Banach spaces, and let X be an intermediate space. Furthermore, let $0 < \Theta < 1$. Then the embedding

$$\big(X_0, X_1\big)_{\Theta,1} \hookrightarrow X \hookrightarrow X_0 + X_1$$

holds if, and only if, for some constant $c > 0$ the inequality

$$\|\, f\,|X\| \le c\,\|\, f\,|X_0\|^{1-\Theta}\|\, f\,|X_1\|^{\Theta} \qquad (6.9.16)$$

is fulfilled for all $f \in X_0 \cap X_1$.

Remark 6.9.2. Having a closer look at the proof of this lemma we also find, that the estimate

$$\|\, f\,|\big(X_0, X_1\big)_{\Theta,1}\| \le c\,\|\, f\,|X\|, \qquad f \in X,$$

holds for exactly the same constant c as in the Gagliardo-Nirenberg-type inequality (6.9.16).

Lemma 6.9.2 is the key to further interpolation results.

Theorem 6.9.4. Let $0 < p_0, p_1, q, q_0, q_1 \le \infty$ and $\overline{r}^0, \overline{r}^1 \in \mathbb{R}$. Moreover, let $0 < \Theta < 1$ and define

$$\frac{1}{p} = \frac{1-\Theta}{p_0} + \frac{\Theta}{p_1}\,, \qquad \overline{r} = (1-\Theta)\overline{r}^0 + \Theta\overline{r}^1\,, \qquad \frac{1}{q_\Theta} = \frac{1-\Theta}{q_0} + \frac{\Theta}{q_1}\,. \tag{6.9.17}$$

Then we have for $q \le q_\Theta$ the embedding

$$s^{\overline{r}}_{p,q}x(\nabla) \hookrightarrow \Big(s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\Big)_{\Theta,\infty}\,, \qquad x \in \{b, f\}\,. \tag{6.9.18}$$

Thereby we assume $0 < p_0, p_1 < \infty$ if $x = f$. Moreover, the norm of the embedding operator for b-spaces does not depend on $\overline{d}$.

Proof. By the monotonicity of the ℓ_q-spaces the case $q = q_\Theta$ is sufficient.

Step 1: First, we prove that $s^{\overline{r}}_{p,q}x(\nabla)$ is an intermediate space with respect to the interpolation couple $\big\{s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big\}$. In view of the intended application in Step 2, we will do so only for parameters $1 \le p_0, p_1, q_0, q_1 < \infty$. We recall the definition of the norm in $s^{\overline{r}^0}_{p_0,q_0}x(\nabla) \cap s^{\overline{r}^1}_{p_1,q_1}x(\nabla)$,

$$\big\| a \,|s^{\overline{r}^0}_{p_0,q_0}x(\nabla) \cap s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big\| := \max\Big(\big\| a \,|s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\big\|\,,\, \big\| a \,|s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big\|\Big)\,.$$

Due to (6.9.17) we find by Proposition 6.8.1

$$\begin{aligned}\big\| a \,|s^{\overline{r}}_{p,q}x(\nabla)\big\| &\le \big\| a \,|s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\big\|^{1-\Theta}\big\| a \,|s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big\|^{\Theta} \\ &\le \max\Big(\big\| a \,|s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\big\|\,,\, \big\| a \,|s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\big\|\Big)\,,\end{aligned}$$

and hence $s^{\overline{r}^0}_{p_0,q_0}x(\nabla) \cap s^{\overline{r}^1}_{p_1,q_1}x(\nabla) \hookrightarrow s^{\overline{r}}_{p,q}x(\nabla)$. Moreover, we note that from (6.9.9), (6.9.10) and (6.9.12) (b-spaces), or (6.9.13) and (6.9.15) (f-spaces), respectively, follows

$$s^{\overline{r}}_{p,q}x(\nabla) \hookrightarrow \Big(s^{\overline{r}^0}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^1}_{p_1,q_1}x(\nabla)\Big)_{\Theta,\max(p,q)}\,.$$

We remind on the assumed finiteness of all parameters. Since all interpolation spaces particularly are intermediate spaces by definition, this proves the embedding $s^{\overline{r}}_{p,q}x(\nabla) \hookrightarrow s^{\overline{r}^0}_{p_0,q_0}x(\nabla) + s^{\overline{r}^1}_{p_1,q_1}x(\nabla)$.

Step 2: Suppose now $1 < p_0, p_1, q_0, q_1 \le \infty$. Then we find for the respective conjugated indices $1 \le p', p'_0, p'_1, q', q'_0, q'_1 < \infty$. By (6.9.17) and Proposition 6.8.1 we then obtain

$$\big\| a \,|s^{-\overline{r}}_{p',q'}x(\nabla)\big\| \le \big\| a \,|s^{-\overline{r}^0}_{p'_0,q'_0}x(\nabla)\big\|^{1-\Theta}\big\| a \,|s^{-\overline{r}^1}_{p'_1,q'_1}x(\nabla)\big\|^{\Theta}\,.$$

Due to Lemma 6.9.2 this inequality is equivalent to the embedding

$$\Big(s^{-\overline{r}^0}_{p'_0,q'_0}x(\nabla)\,, s^{-\overline{r}^1}_{p'_1,q'_1}x(\nabla)\Big)_{\Theta,1} \hookrightarrow s^{-\overline{r}}_{p',q'}x(\nabla)\,.$$

The lemma is applicable due to Step 1. From this embedding we obtain from Lemma 6.9.1 with the help of Lemmas 5.5.2 and 5.5.3 (the required density properties are valid due to the density of the finite sequences) the assertion (6.9.18).

Step 3: The additional restriction $1 < p_0, p_1, q_0, q_1 \le \infty$ can be removed by the same arguments as in the second step of the proof of Theorem 6.9.3.

Step 4: Finally, the statement concerning the $\overline{d}$-dependence of the constants follows from Remark 6.9.2, Lemma 5.5.2, $\|\mathrm{id} : Y' \longrightarrow X'\| \le \|\mathrm{id} : X \longrightarrow Y\|$, and the observation, that the equivalence constants in the duality assertion Lemma 6.9.1 do not depend on the interpolation couple. □

Remark 6.9.3. The complex interpolation method represents an alternative approach which yields (almost) the same result. It is well-known, see [82, Theorem 1.10.3], that the embeddings

$$\big(X\,,Y\big)_{\Theta,1} \hookrightarrow \big[X\,,Y\big]_{\Theta} \hookrightarrow \big(X\,,Y\big)_{\Theta,\infty}\,, \qquad 0 < \Theta < 1\,, \tag{6.9.19}$$

hold for every interpolation couple $\{X, Y\}$ of Banach spaces. For Banach lattices of functions (and under some additional restrictions also for quasi-Banach spaces) these complex interpolation spaces can be calculated as Calderón products $X^{1-\Theta}Y^{\Theta}$. This has been done in [94] for sequence spaces $s^{\overline{r}}_{p,q}x(\Omega)$ in the case $d_1 = \cdots = d_N$ with the outcome

$$\big[s^{\overline{r}^0}_{p_0,q_0}x(\Omega)\,, s^{\overline{r}^1}_{p_1,q_1}x(\Omega)\big]_{\Theta} = s^{\overline{r}^0}_{p_0,q_0}x(\Omega)^{1-\Theta}s^{\overline{r}^1}_{p_1,q_1}x(\Omega)^{\Theta} = s^{\overline{r}}_{p,q}x(\Omega)$$

for the same set of parameters as in Theorem 6.9.4, where $q = q_\Theta$, $\Omega \subset \mathbb{R}^d$ a bounded domain, and additionally $\min(q_0, q_1) < \infty$. This means if we restrict ourselves first to Banach spaces, we obtain the embedding (6.9.19), in particular we re-obtain (6.9.18). Afterwards, the assertion is extended to quasi-Banach spaces in the same way as before.

6.10 Estimates from above: Closing some gaps

In this section we will apply the interpolation formulas and the Reiteration theorem to obtain further estimates from above for the m-term approximation. Concerning compact embeddings we want to get rid of the assumption (6.7.16). At least in the case of two Besov-type spaces we achieve full generality, matching the estimate from below in Proposition 6.5.1. Moreover, we shall deal with the case of two Triebel-Lizorkin-type spaces. But to begin with we want to weaken the assumptions in the limiting case.

Theorem 6.10.1. Let $0 < p_0 < p_1 < \infty$ and $0 < q_0, q_1 \le \infty$, and put $\bar{t} = \overline{d}\big(\frac{1}{p_0} - \frac{1}{p_1}\big)$.

(i) Let $\frac{1}{p_0} - \frac{1}{p_1} < \frac{1}{q_0} - \frac{1}{q_1}$ and $p_1 \le q_1$. Then it holds

$$\sigma_m\Big(s^{\bar{t}}_{p_0,q_0}f(\nabla)\,, s^{\bar{0}}_{p_1,q_1}f(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \quad m \in \mathbb{N}\,.$$

(ii) Let $\frac{1}{p_0} - \frac{1}{p_1} > \frac{1}{q_0} - \frac{1}{q_1}$ and $q_1 \le p_1$. Then it holds

$$\sigma_m\Big(s^{\bar{t}}_{p_0,q_0}f(\nabla)\,, s^{\bar{0}}_{p_1,q_1}f(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)\left(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1}\right)}$$

for every natural number $m \ge 2$.

Proof. **Step 1: We prove (i).** Due to the assumptions we have $p_0 < p_1$ and $q_0 < q_1$, hence for every $0 < \Theta < 1$ we find parameters $0 < u < \infty$ and $0 < v \le \infty$, such that

$$\frac{1}{p_0} = \frac{1-\Theta}{p_1} + \frac{\Theta}{u} \qquad \text{and} \qquad \frac{1}{q_0} = \frac{1-\Theta}{q_1} + \frac{\Theta}{v} .$$

In order to apply Theorem 6.7.3(ii) to the space $s_{u,v}^{\bar{d}(\frac{1}{u}-\frac{1}{p_1})} f(\nabla)$ we have to ensure the condition $v \le u < p_1$. Firstly, $u < p_1$ follows immediately from the definition of u and $p_0 < p_1$. Secondly, $v \le u$ is equivalent to

$$\frac{\Theta}{u} \le \frac{\Theta}{v} \iff \frac{1}{p_0} - \frac{1-\Theta}{p_1} \le \frac{1}{q_0} - \frac{1-\Theta}{q_1} \iff \Theta\Big(\frac{1}{p_1} - \frac{1}{q_1}\Big) \le \frac{1}{q_0} - \frac{1}{q_1} - \frac{1}{p_0} + \frac{1}{p_1} ,$$

thus the assumptions of Theorem 6.7.3(ii) can be fulfilled by choosing Θ sufficiently small. We obtain in this way

$$s_{u,v}^{\bar{d}(\frac{1}{u}-\frac{1}{p_1})} f(\nabla) \hookrightarrow \mathcal{A}_{\infty}^{\frac{1}{u}-\frac{1}{p_1}} \big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , \mathcal{B}\big) .$$

Now we conclude from Theorem 6.9.4 and Proposition 6.3.1

$$\begin{aligned} s_{p_0,q_0}^{\bar{d}(\frac{1}{p_0}-\frac{1}{p_1})} f(\nabla) &\hookrightarrow \Big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , s_{u,v}^{\bar{d}(\frac{1}{u}-\frac{1}{p_1})} f(\nabla)\Big)_{\Theta,\infty} \\ &\hookrightarrow \Big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , \mathcal{A}_{\infty}^{\frac{1}{u}-\frac{1}{p_1}} \big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , \mathcal{B}\big)\Big)_{\Theta,\infty} = \mathcal{A}_{\infty}^{\frac{1}{p_0}-\frac{1}{p_1}} \big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , \mathcal{B}\big) . \end{aligned}$$

Step 2: We prove (ii). Suppose at first $q_0 \le q_1$. Then we can choose the parameters u and v as above. This time we want to apply Theorem 6.7.3(i), hence we have to check $u \le v$. Similar than before we find

$$u \le v \iff \frac{\Theta}{v} \le \frac{\Theta}{u} \iff \Theta\Big(\frac{1}{q_1} - \frac{1}{p_1}\Big) \le \frac{1}{p_0} - \frac{1}{p_1} - \frac{1}{q_0} + \frac{1}{q_1} ,$$

which is satisfied for Θ sufficiently small. Now we conclude

$$s_{u,v}^{\bar{d}(\frac{1}{u}-\frac{1}{p_1})} f(\nabla) \hookrightarrow \mathcal{A}_{\infty}^{\frac{1}{u}-\frac{1}{p_1}, -(N-1)(\frac{1}{u}-\frac{1}{p_1}-\frac{1}{v}+\frac{1}{q_1})} \big(s_{p_1,q_1}^{\bar{0}} f(\nabla) , \mathcal{B}\big) ,$$

and the result follows by the same interpolation argument as in Step 1.
Now assume $q_1 < q_0$ and $q_1 < p_1$. We once more want to apply Theorem 6.7.3(i) to the pair $\big(s_{u,v}^{\bar{d}(\frac{1}{u}-\frac{1}{p_1})} f(\nabla) , s_{p_1,q_1}^{\bar{0}} f(\nabla)\big)$. Since $\frac{1}{v} \ge 0$, the definition of v implies $\Theta \ge 1 - \frac{q_1}{q_0} > 0$. On the other hand, the condition $u \le v$ can be reformulated as before, which yields an upper bound for Θ. However, the assumptions $p_0 < p_1$ and $q_1 < q_0$ ensure the existence of parameters $\Theta \in (0,1)$ which satisfy both inequalities simultaneously. This follows from

$$1 - \frac{q_1}{q_0} < \frac{\frac{1}{p_0} - \frac{1}{p_1} - \frac{1}{q_0} + \frac{1}{q_1}}{\frac{1}{q_1} - \frac{1}{p_1}} \iff \frac{q_1}{q_0 p_1} < \frac{1}{p_0} .$$

Now we obtain (ii) by interpolation as before. □

The following theorem deals with the case of two Besov-type sequence spaces.

Theorem 6.10.2. Let $0 < p_0, p_1, q_0, q_1 \le \infty$ and $t \in \mathbb{R}$, where $t > \big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+$. We put $\overline{t} = t\overline{d}$. Furthermore, let Ω be a bounded open subset of $\mathbb{R}^d$. Then it holds

$$\sigma_m\Big(s^{\overline{t}}_{p_0,q_0}b(\Omega)\,,\, s^{\overline{0}}_{p_1,q_1}b(\Omega)\,,\, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-t}\big(\log m\big)^{(N-1)\big(t-\frac{1}{q_0}+\frac{1}{q_1}\big)_+} \tag{6.10.1}$$

for all natural numbers $m \ge 2$.

Proof. We shall distinguish several cases, according to the relations of t, $\frac{1}{p_0} - \frac{1}{p_1}$ and $\frac{1}{q_0} - \frac{1}{q_1}$.

Step 1: We start with the case $\frac{1}{q_0} - \frac{1}{q_1} \ge t > \frac{1}{p_0} - \frac{1}{p_1}$. Let p_* be defined by

$$\frac{1}{p_*} = t + \frac{1}{p_1}\,, \qquad \text{i.e.} \quad \overline{t} = \overline{d}\bigg(\frac{1}{p_*} - \frac{1}{p_1}\bigg)\,.$$

Then by the assumption on t we find $p_* > p_0$, and hence by Proposition 5.3.2 it holds

$$s^{\overline{t}}_{p_0,q_0}b(\Omega) \hookrightarrow s^{\overline{t}}_{p_*,q_0}b(\Omega)\,.$$

As $t = \frac{1}{p_*} - \frac{1}{p_1} \le \frac{1}{q_0} - \frac{1}{q_1}$, from Theorem 6.7.2(ii) and Lemma 6.3.1 we conclude

$$s^{\overline{t}}_{p_*,q_0}b(\Omega) \hookrightarrow \mathcal{A}^t_\infty\big(s^{\overline{0}}_{p_1,q_1}b(\Omega), \mathcal{B}\big)\,.$$

Combining both embeddings yields the desired estimate in view of Lemma 6.3.1.

Step 2: Now let $t > \frac{1}{q_0} - \frac{1}{q_1} > \frac{1}{p_0} - \frac{1}{p_1} > 0$. Put $r = \frac{1}{p_0} - \frac{1}{p_1}$, and define u_0 by $\frac{1}{u_0} = r + \frac{1}{q_1} < \frac{1}{q_0}$. Then choose $0 < \Theta < 1$ and $0 < u_1 < q_0$, such that $\frac{1}{q_0} = \frac{1-\Theta}{u_0} + \frac{\Theta}{u_1}$. Finally, define $r' \in \mathbb{R}$ by $t = (1-\Theta)r + \Theta r'$. By choosing u_1 (and hence also Θ) sufficiently small, in particular $u_1 \le \min(p_0, p_1, q_1)$, also the condition

$$r' > \frac{1}{\min(p_0,u_1)} - \frac{1}{\max(p_1,q_1)} = \frac{1}{u_1} - \frac{1}{\max(p_1,q_1)} \ge 0 \tag{6.10.2}$$

can be fulfilled. To see this, we have to check $r' > \frac{1}{u_1} - \frac{1}{p_1}$, if $p_1 \ge q_1$. But this follows from

$$\begin{aligned} r' > \frac{1}{u_1} - \frac{1}{p_1} &\iff t - (1-\Theta)r > \frac{1}{q_0} - \frac{1-\Theta}{u_0} - \frac{\Theta}{p_1} \\ &\iff t - \frac{1-\Theta}{p_0} + \frac{1-\Theta}{p_1} > \frac{1}{q_0} - \frac{1-\Theta}{p_0} + \frac{1-\Theta}{p_1} - \frac{1-\Theta}{q_1} - \frac{\Theta}{p_1} \\ &\iff t > \frac{1}{q_0} - \frac{1-\Theta}{q_1} - \frac{\Theta}{p_1} = \frac{1}{q_0} - \frac{1}{q_1} + \Theta\bigg(\frac{1}{q_1} - \frac{1}{p_1}\bigg)\,, \end{aligned}$$

what is valid for sufficiently small Θ. The case $p_1 \le q_1$ can be obtained by a similar calculation. More precisely, we find

$$r' > \frac{1}{u_1} - \frac{1}{q_1} \iff t > \frac{1}{q_0} - \frac{1}{q_1}\,.$$

Now we know by Theorem 6.7.2(ii)

$$s^{r\overline{d}}_{p_0,u_0}b(\Omega) \hookrightarrow \mathcal{A}^{r,0}_\infty\big(s^{\overline{0}}_{p_1,q_1}b(\Omega), \mathcal{B}\big)\,,$$

and Corollary 6.7.2 yields

$$s^{r'\bar{d}}_{p_0,u_1}b(\Omega) \hookrightarrow \mathcal{A}^{r',-(N-1)(r'-\frac{1}{u_1}+\frac{1}{q_1})}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,,$$

Then we obtain by interpolation (Theorem 6.9.3(ii) is applicable, since $u_0<\infty$ due to $r>0$)

$$\begin{aligned} s^{t\bar{d}}_{p_0,q_0}b(\Omega) &= \Big(s^{r\bar{d}}_{p_0,u_0}b(\Omega), s^{r'\bar{d}}_{p_0,u_1}b(\Omega)\Big)_{\Theta,q_0} \\ &\hookrightarrow \Big(\mathcal{A}^{r,0}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,, \mathcal{A}^{r',-(N-1)(r'-\frac{1}{u_1}+\frac{1}{q_1})}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\Big)_{\Theta,q_0} \\ &= \mathcal{A}^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}_{q_0}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{B}\big)\,. \end{aligned}$$

The last line follows from Proposition 6.3.1(ii).

Step 3: Let $t>\frac{1}{p_0}-\frac{1}{p_1}>\frac{1}{q_0}-\frac{1}{q_1}>0$. Put $r=\frac{1}{q_0}-\frac{1}{q_1}$, and define u_0 by $\frac{1}{u_0}=r+\frac{1}{p_1}<\frac{1}{p_0}$. Then choose $0<\Theta<1$ and $0<u_1<p_0$, such that $\frac{1}{p_0}=\frac{1-\Theta}{u_0}+\frac{\Theta}{u_1}$. Finally, define $r'\in\mathbb{R}$ by $t=(1-\Theta)r+\Theta r'$. By choosing u_1 (and hence also Θ) sufficiently small, in particular $u_1\le\min(q_0,p_1,q_1)$, also the condition

$$r' > \frac{1}{\min(u_1,q_0)}-\frac{1}{\max(p_1,q_1)}=\frac{1}{u_1}-\frac{1}{\max(p_1,q_1)} \tag{6.10.3}$$

can be satisfied. This can be seen as in Step 2.
Now we know by Theorem 6.7.2

$$s^{r\bar{d}}_{u_0,q_0}b(\Omega) \hookrightarrow \mathcal{A}^{r,0}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,.$$

Moreover, Corollary 6.7.2 (or Theorem 6.7.4, respectively) yields

$$s^{r'\bar{d}}_{u_1,q_0}b(\Omega) \hookrightarrow \mathcal{A}^{r',-(N-1)(r'-\frac{1}{q_0}+\frac{1}{q_1})}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,.$$

From this we conclude in case $p_0\le q_0$ from Theorem 6.9.3(iii) and Proposition 6.3.1(ii)

$$\begin{aligned} s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \Big(s^{r\bar{d}}_{u_0,q_0}b(\Omega)\,, s^{r'\bar{d}}_{u_1,q_0}b(\Omega)\Big)_{\Theta,q_0} \\ &\hookrightarrow \Big(\mathcal{A}^{r,0}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,, \mathcal{A}^{r',-(N-1)(r'-\frac{1}{q_0}+\frac{1}{q_1})}_{\infty}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\Big)_{\Theta,q_0} \\ &= \mathcal{A}^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}_{q_0}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,. \end{aligned}$$

On the other hand, if $q_0\le p_0$, we get similarly

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow \mathcal{A}^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}_{p_0}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)$$

by interpolation with respect to the $(.\,,.)_{\Theta,p_0}$-functor. In both cases Theorem 6.9.3(iii) is applicable, since the assumptions ensure $r=\frac{1}{q_0}-\frac{1}{q_1}>0$ and $0<u_0,u_1,q_0,p_0<\infty$.

Step 4: Now consider the case $t>\frac{1}{q_0}-\frac{1}{q_1}>0>\frac{1}{p_0}-\frac{1}{p_1}$ and $q_1<\infty$. Then we choose $u_1>q_0$, $r'>t$, and $0<\Theta<1$, such that

$$t=\Theta r' \qquad \text{and} \qquad \frac{1}{q_0}=\frac{1-\Theta}{q_1}+\frac{\Theta}{u_1}\,.$$

By taking u_1 and hence also Θ sufficiently small, particularly $u_1 \leq p_0$, we can ensure

$$r' > \frac{1}{\min(p_0, u_1)} - \frac{1}{\max(p_1, q_1)} = \frac{1}{u_1} - \frac{1}{\max(p_1, q_1)}$$

as in Step 2, and $r' > 0$ follows from $t > 0$. Thus we find by Corollary 6.7.2

$$s^{r'\overline{d}}_{p_0,u_1} b(\Omega) \hookrightarrow \mathcal{A}_\infty^{r',-(N-1)(r'-\frac{1}{u_1}+\frac{1}{q_1})}\big(s^{\overline{0}}_{p_1,q_1} b(\Omega), \mathcal{B}\big)\,.$$

Then we obtain from Theorem 6.9.3(ii) and Proposition 6.3.1(i)

$$\begin{aligned} s^{t\overline{d}}_{p_0,q_0} b(\Omega) &= \Big(s^{\overline{0}}_{p_0,q_1} b(\Omega)\,, s^{r'\overline{d}}_{p_0,u_1} b(\Omega)\Big)_{\Theta,q_0} \\ &\hookrightarrow \Big(s^{\overline{0}}_{p_1,q_1} b(\Omega)\,, \mathcal{A}_\infty^{r',-(N-1)(r'-\frac{1}{u_1}+\frac{1}{q_1})}\big(s^{\overline{0}}_{p_1,q_1} b(\Omega), \mathcal{B}\big)\Big)_{\Theta,q_0} \\ &= \mathcal{A}_{q_0}^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}\big(s^{\overline{0}}_{p_1,q_1} b(\Omega), \mathcal{B}\big)\,. \end{aligned}$$

Step 5: Finally, all cases left open so far (where we still assume $q_0 \leq q_1$) could be treated similarly by carefully choosing parameters and applying real interpolation. However, since most cases would have to be treated separately, we choose a slightly different method (which could have been applied also instead of Steps 2–4).
Let $t > \frac{1}{q_0} - \frac{1}{q_1} \geq 0$. Consider first the case $p_0 \leq p_1$. Then we can choose $0 < u, v \leq \infty$, $r > 0$, and $0 < \Theta < 1$, such that

$$t = \Theta r\,, \qquad \frac{1}{p_0} = \frac{1-\Theta}{p_1} + \frac{\Theta}{u}\,, \qquad \frac{1}{q_0} = \frac{1-\Theta}{q_1} + \frac{\Theta}{v}\,.$$

This is possible due to $q_0 \leq q_1$ and $p_0 \leq p_1$, i.e. these assumptions ensure $\frac{1}{u} \geq 0$ and $\frac{1}{v} \geq 0$. Moreover, we still have Θ completely at our disposal. By choosing Θ sufficiently small, also the condition

$$r > \frac{1}{\min(u, v)} - \frac{1}{\max(p_1, q_1)}$$

can be fulfilled. This follows from

$$\begin{aligned} r > \frac{1}{u} - \frac{1}{p_1} &\iff t > \frac{\Theta}{u} - \frac{\Theta}{p_1} = \frac{1}{p_0} - \frac{1-\Theta}{p_1} - \frac{\Theta}{p_1} = \frac{1}{p_0} - \frac{1}{p_1} \\ r > \frac{1}{u} - \frac{1}{q_1} &\iff t > \frac{\Theta}{u} - \frac{\Theta}{q_1} = \frac{1}{p_0} - \frac{1-\Theta}{p_1} - \frac{\Theta}{q_1} = \frac{1}{p_0} - \frac{1}{p_1} + \Theta\left(\frac{1}{p_1} - \frac{1}{q_1}\right) \\ r > \frac{1}{v} - \frac{1}{p_1} &\iff t > \frac{\Theta}{v} - \frac{\Theta}{p_1} = \frac{1}{q_0} - \frac{1-\Theta}{q_1} - \frac{\Theta}{p_1} = \frac{1}{q_0} - \frac{1}{q_1} + \Theta\left(\frac{1}{q_1} - \frac{1}{p_1}\right) \\ r > \frac{1}{v} - \frac{1}{q_1} &\iff t > \frac{\Theta}{v} - \frac{\Theta}{q_1} = \frac{1}{q_0} - \frac{1-\Theta}{q_1} - \frac{\Theta}{q_1} = \frac{1}{q_0} - \frac{1}{q_1} \end{aligned}$$

and the assumptions on t. Now we know by Corollary 6.7.2

$$s^{r\overline{d}}_{u,v} b(\Omega) \hookrightarrow \mathcal{A}_\infty^{r,-(N-1)(r-\frac{1}{v}+\frac{1}{q_1})}\big(s^{\overline{0}}_{p_1,q_1} b(\Omega), \mathcal{B}\big)\,,$$

and we find by interpolation (Theorem 6.9.4)

$$\begin{aligned}
s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \Big(s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,s^{r\bar{d}}_{u,v}b(\Omega)\Big)_{\Theta,\infty}\\
&\hookrightarrow \Big(s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{A}_\infty^{r,-(N-1)(r-\frac{1}{v}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\Big)_{\Theta,\infty}\\
&= \mathcal{A}_\infty^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,.
\end{aligned}$$

The last line follows from Proposition 6.3.1(ii) and the choice of the parameters.

Step 6: In case $p_0 > p_1$ and $q_0 \le q_1$, we can argue similarly. By choosing v and r as before, but now complemented by $u = p_0$, we end up with a similar system of conditions for r. The condition for $r > \frac{1}{v} - \frac{1}{\max(p_1,q_1)}$ remains the same, and $r > \frac{1}{p_0} - \frac{1}{\max(p_1,q_1)}$ follows this time from $t > 0$.
Hence, for Θ sufficiently small we obtain

$$\begin{aligned}
s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \Big(s^{\bar{0}}_{p_0,q_1}b(\Omega)\,,s^{r\bar{d}}_{p_0,v}b(\Omega)\Big)_{\Theta,\infty}\\
&\hookrightarrow \Big(s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,\mathcal{A}_\infty^{r,-(N-1)(r-\frac{1}{v}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\Big)_{\Theta,\infty}\\
&= \mathcal{A}_\infty^{t,-(N-1)\left(t-\frac{1}{q_0}+\frac{1}{q_1}\right)}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,,
\end{aligned}$$

where we additionally used $s^{\bar{0}}_{p_0,q_1}b(\Omega) \hookrightarrow s^{\bar{0}}_{p_1,q_1}b(\Omega)$ from Proposition 5.3.2.

Step 7: The Steps 1–6 can be summarized by

$$s^{r_0\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow \mathcal{A}_\infty^{r_0-r_1,-(N-1)\left(r_0-r_1-\frac{1}{q_0}+\frac{1}{q_1}\right)_+}\big(s^{r_1\bar{d}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,, \tag{6.10.4}$$

where $r_0, r_1 \in \mathbb{R}$, $r_0 - r_1 > \big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+$ and $0 < q_0 \le q_1 \le \infty$. We additionally applied a lifting argument (see Lemma 6.2.3).

Step 8: Now consider the case $0 < q_1 \le p_0 \le q_0 \le \infty$. We choose $\alpha > 0$, such that $t - \big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+ > \alpha$. Then we have the embeddings

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow s^{(t-\alpha)\bar{d}}_{p_0,q_1}b(\Omega) \hookrightarrow s^{\bar{0}}_{p_1,q_1}b(\Omega)\,.$$

By choice, we have $\alpha - \frac{1}{q_0} + \frac{1}{q_1} > 0$ and

$$\alpha > 0 = \frac{1}{p_0} - \frac{1}{p_0} = \frac{1}{\min(p_0,q_0)} - \frac{1}{\max(p_0,q_1)}\,,$$

hence the assumptions of Corollary 6.7.2 are satisfied for the left hand embedding. Thus we find by an additional lifting argument, see Lemma 6.2.3,

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow \mathcal{A}_\infty^{\alpha,-(N-1)(\alpha-\frac{1}{q_0}+\frac{1}{q_1})}\big(s^{(t-\alpha)\bar{d}}_{p_0,q_1}b(\Omega),\mathcal{B}\big)\,,$$

and on the other hand we obtain from (6.10.4)

$$s^{(t-\alpha)\bar{d}}_{p_0,q_1}b(\Omega) \hookrightarrow \mathcal{A}_\infty^{t-\alpha,-(N-1)(t-\alpha)}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,.$$

Now the Reiteration theorem (Proposition 6.3.2) yields

$$\begin{aligned} s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \mathcal{A}_\infty^{\alpha,-(N-1)(\alpha-\frac{1}{q_0}+\frac{1}{q_1})}\big(s^{(t-\alpha)\bar{d}}_{p_0,q_1}b(\Omega),\mathcal{B}\big) \\ &\hookrightarrow \mathcal{A}_\infty^{\alpha,-(N-1)(\alpha-\frac{1}{q_0}+\frac{1}{q_1})}\Big(\mathcal{A}_\infty^{t-\alpha,-(N-1)(t-\alpha)}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big),\mathcal{B}\Big) \\ &= \mathcal{A}_\infty^{t,-(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,. \end{aligned}$$

The same type of argument can be applied in the case $0 < q_1 \le p_1 \le q_0 \le \infty$ to the embedding

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow s^{\alpha\bar{d}}_{p_1,q_0}b(\Omega) \hookrightarrow s^{\bar{0}}_{p_1,q_1}b(\Omega)\,.$$

Moreover, the cases $p_0 \le q_1 \le q_0 \le p_1$ and $p_1 \le q_1 \le q_0 \le p_0$ are covered by Corollary 6.7.2.

Step 9: It remains to treat the cases $\max(p_0,p_1) \le q_1 \le q_0$ and $q_1 \le q_0 \le \min(p_0,p_1)$. For the first one, we consider the embeddings

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow s^{\alpha\bar{d}}_{p_1,p_1}b(\Omega) \hookrightarrow s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,$$

where once more $t-\big(\frac{1}{p_0}-\frac{1}{p_1}\big)_+ > \alpha$. Then the left hand embedding is covered by Corollary 6.7.2 due to $p_0 \le q_0$ and the choice of α, and the right hand embedding corresponds to (6.10.4). More precisely, we find

$$\begin{aligned} s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \mathcal{A}_\infty^{t-\alpha,-(N-1)(t-\alpha-\frac{1}{q_0}+\frac{1}{p_1})}\big(s^{\alpha\bar{d}}_{p_1,p_1}b(\Omega),\mathcal{B}\big)\,, \\ s^{\alpha\bar{d}}_{p_1,p_1}b(\Omega) &\hookrightarrow \mathcal{A}_\infty^{\alpha,-(N-1)(\alpha-\frac{1}{p_1}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,. \end{aligned}$$

Applying the Reiteration theorem we end up with

$$\begin{aligned} s^{t\bar{d}}_{p_0,q_0}b(\Omega) &\hookrightarrow \mathcal{A}_\infty^{t-\alpha,-(N-1)(t-\alpha-\frac{1}{q_0}+\frac{1}{p_1})}\big(s^{\alpha\bar{d}}_{p_1,p_1}b(\Omega),\mathcal{B}\big) \\ &\hookrightarrow \mathcal{A}_\infty^{t,-(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}b(\Omega),\mathcal{B}\big)\,. \end{aligned}$$

Finally, the case $q_1 \le q_0 \le \min(p_0,p_1)$ follows by the same arguments, applied to the embeddings

$$s^{t\bar{d}}_{p_0,q_0}b(\Omega) \hookrightarrow s^{(t-\alpha)\bar{d}}_{p_0,p_0}b(\Omega) \hookrightarrow s^{\bar{0}}_{p_1,q_1}b(\Omega)\,,$$

whereas the left hand embedding now is covered by (6.10.4), and the result for right hand embedding follows from Corollary 6.7.2. □

The result for f-spaces is similar, but not quite as satisfactory, since not for all parameters the upper bounds for the m-term width match the lower ones in Proposition 6.5.1.

Theorem 6.10.3. Let $0 < p_0, p_1 < \infty$ and $0 < q_0, q_1 \le \infty$. Moreover, let $t \in \mathbb{R}$ and put $\bar{t} = t\bar{d}$. Furthermore, let Ω be a bounded open subset of $\mathbb{R}^d$.

(i) We assume

$$t > \max\left(\frac{1}{p_0} - \frac{1}{p_1}\,,\, \frac{1}{q_0} - \frac{1}{q_1}\,,\, 0\right).$$

Then it holds

$$\sigma_m\Big(s^{\bar{t}}_{p_0,q_0}f(\Omega)\,,\, s^{\bar{0}}_{p_1,q_1}f(\Omega)\,,\, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-t}\big(\log m\big)^{(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})}$$

for all natural numbers $m \geq 2$.

(ii) Let now

$$\max\left(\frac{1}{p_0} - \frac{1}{p_1}\,,\, 0\right) < t \leq \frac{1}{q_0} - \frac{1}{q_1}\,.$$

Then it holds for all $\varepsilon > 0$

$$\sigma_m\Big(s^{\bar{t}}_{p_0,q_0}f(\Omega)\,,\, s^{\bar{0}}_{p_1,q_1}f(\Omega)\,,\, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-t}\big(\log m\big)^{\varepsilon}$$

for all natural numbers $m \geq 2$.

(iii) Now suppose either

$$\max\left(\frac{1}{p_0} - \frac{1}{p_1}\,,\, 0\right) < t < \frac{1}{q_0} - \frac{1}{q_1} \qquad \text{and} \qquad p_1 < q_1\,, \tag{6.10.5}$$

or

$$\max\left(\frac{1}{p_0} - \frac{1}{p_1}\,,\, 0\right) < t \leq \frac{1}{q_0} - \frac{1}{q_1} \qquad \text{and} \qquad p_1 = q_1\,. \tag{6.10.6}$$

Then it holds

$$\sigma_m\Big(s^{\bar{t}}_{p_0,q_0}f(\Omega)\,,\, s^{\bar{0}}_{p_1,q_1}f(\Omega)\,,\, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big) \lesssim m^{-t}\,, \qquad m \in \mathbb{N}. \tag{6.10.7}$$

Proof. The proof of part (i) uses exactly the same methods as in the Steps 5–6 and Steps 8–9 of the proof of Theorem 6.10.2, where one has to replace applications of embedding and interpolation results for b-spaces by their corresponding f-counterparts.

For the proof of part (ii) choose $q_0 \leq q_* \leq \infty$ in such a way, that

$$t - \frac{1}{q_0} + \frac{1}{q_1} \leq 0 < t - \frac{1}{q_*} + \frac{1}{q_1} < \varepsilon\,.$$

We then obtain from Lemma 5.3.2(i) and part (i)

$$\begin{aligned} s^{\bar{t}}_{p_0,q_0}f(\Omega) \hookrightarrow s^{\bar{t}}_{p_0,q_*}f(\Omega) &\hookrightarrow \mathcal{A}_\infty^{t\,,\,-(N-1)(t-\frac{1}{q_*}+\frac{1}{q_1})}\big(s^{\bar{0}}_{p_1,q_1}f(\Omega),\mathcal{B}\big) \\ &\hookrightarrow \mathcal{A}_\infty^{t\,,\,-(N-1)\varepsilon}\big(s^{\bar{0}}_{p_1,q_1}f(\Omega),\mathcal{B}\big)\,. \end{aligned}$$

The last embedding follows from the trivial fact $\mathcal{A}_u^{\varrho,\tau}(X,\mathcal{D}) \hookrightarrow \mathcal{A}_u^{\varrho,\tau'}(X,\mathcal{D})$ for every $\tau \geq \tau'$ and every scale of generalized approximation spaces. From these embeddings and Lemma 6.3.1 we obtain the desired estimate, because $\varepsilon > 0$ was arbitrary.

Finally, for the proof of (iii) we use an embedding argument as in Step 1 of the proof of Theorem 6.10.2, i.e. we use the embedding $s^{\bar{t}}_{p_0,q_0}f(\Omega) \hookrightarrow s^{\bar{t}}_{p_*,q_0}f(\Omega)$, where $\frac{1}{p_*} = t + \frac{1}{p_1}$. Its continuity is due to Proposition 5.3.4. Then we find $\bar{t} = \bar{d}\left(\frac{1}{p_*} - \frac{1}{p_1}\right)$, hence we can apply the results for the limiting case. Now (6.10.7) follows either from Theorem 6.7.3 in case of (6.10.6), or from Theorem 6.10.1 for parameters as in (6.10.5). □

Remark 6.10.1. If Conjecture 6.7.1 holds true, the $(\log m)^{\varepsilon}$-gap can be closed. In that case, one uses the same embedding argument as in the above proof to obtain the upper bound m^{-t} for all parameters p_1, q_1 and $t \leq \frac{1}{q_0} - \frac{1}{q_1}$.
Even more is true: If the conjectured result is valid then the limiting case completely determines the result for the compact case, once more by exactly the same embedding argument. This behaviour could already be observed for the isotropic spaces (compare with the results in [34]), but it is in sharp contrast to other combinations for b and f-spaces (we refer e.g. to the results obtained in case of two Besov-type sequence spaces).

7 Conclusion

In the first four chapters of this thesis we dealt with the function spaces $S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, which were defined as subsets of $\mathcal{S}'(\mathbb{R}^d)$. This investigation culminated in the characterization in terms of wavelet decompositions in Section 4.3, establishing a connection between these function spaces and certain sequence spaces $s^{\overline{r}}_{p,q}a$. Sections five and six were devoted to the study of these sequence spaces, in particular continuous and compact embeddings as well as best m-term approximation with respect to the canonical basic system. In this chapter finally we will once more use the wavelet isomorphisms to reformulate the results obtained for sequence spaces into results for the associated function spaces.
In many applications one is also interested in functions and distributions which are not defined on the whole $\mathbb{R}^d$, but on certain bounded or unbounded domains Ω. Since we previously dealt with the slightly more general sequence spaces $s^{\overline{r}}_{p,q}a(\nabla)$ and $s^{\overline{r}}_{p,q}a(\Omega)$, we can formulate results for function spaces $S^{\overline{r}}_{p,q}A(\Omega)$ as well.
Finally, we will compare our results with those to be found in the literature. Though in most cases the setting is slightly different, particularly in the Russian school function spaces on the torus are more popular, nevertheless they usually behave very similar as far as approximative properties are concerned. Hence they are often called on for a comparison of results.

7.1 Transfer to function spaces on domains

In this section we will define function and distribution spaces on domains Ω, and establish a connection to the previously discussed sequence spaces $s^{\overline{r}}_{p,q}a(\Omega)$. Moreover, we will formulate the main results of this thesis.

7.1.1 Function spaces on domains

First, we define function spaces on domains by restriction of function spaces on $\mathbb{R}^d$. Let $\mathcal{D}(\Omega)$ denote the locally convex vector space of all infinitely differentiable functions with compact support in Ω, where Ω is an arbitrary non-empty open subset (domain) of $\mathbb{R}^d$. Moreover, we denote by $\mathcal{D}'(\Omega)$ its topological dual.

Definition 7.1.1. Let Ω be an arbitrary domain in $\mathbb{R}^d$, and let $0 < p, q \leq \infty$ ($p < \infty$ for F-spaces) and $\overline{r} \in \mathbb{R}^N$. Then the spaces $S^{\overline{r}}_{p,q}A(\Omega)$ are defined as

$$S^{\overline{r}}_{p,q}A(\Omega) := \Big\{ f \in \mathcal{D}'(\Omega) \,:\, f = g_{|\Omega} \quad \text{for some} \quad g \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}}) \Big\} ,$$
$$\big\| f \,|S^{\overline{r}}_{p,q}A(\Omega)\big\| := \inf \big\| g \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| ,$$

where the infimum is taken over all $g \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$, such that $f = g_{|\Omega}$.

As mentioned before we are interested in a connection between these spaces on domains and sequence spaces with the help of wavelets. To this purpose, denote by $\mathcal{E}f \in S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})$ an extension of $f \in S^{\overline{r}}_{p,q}A(\Omega)$, such that

$$\big\| f \,|S^{\overline{r}}_{p,q}A(\Omega)\big\| \leq \big\| \mathcal{E}f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| \leq 2 \,\big\| f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\big\| .$$

By Theorem 4.3.1 we know that $\mathcal{E}f$ can be represented as

$$\mathcal{E}f = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\Gamma_{\overline{k}}} \langle \mathcal{E}f, \Psi_{\overline{k},\gamma}\rangle \Psi_{\overline{k},\gamma}\,, \qquad \big(\langle \mathcal{E}f, \Psi_{\overline{k},\gamma}\rangle\big)_{\overline{k},\gamma} \in s^{\overline{r}}_{p,q}a\,.$$

Since the construction of the basis functions $\Psi_{\overline{k},\gamma}$ is based on the scaling function ψ_0 and the associated wavelet ψ_1 and both were assumed to be compactly supported, say

$$\operatorname{supp}\psi_0 \cup \operatorname{supp}\psi_1 \subset [-M, M] \qquad \text{for some} \quad M > 0\,,$$

also the functions $\Psi_{\overline{k},\gamma}$ are compactly supported, and the distribution

$$\mathcal{E}^*f = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\substack{\gamma\in\Gamma_{\overline{k}}:\\ \operatorname{supp}\Psi_{\overline{k},\gamma}\cap\Omega\neq\emptyset}} \langle f, \Psi_{\overline{k},\gamma}\rangle \Psi_{\overline{k},\gamma}$$

is an extension of f as well. The uniqueness of the wavelet-coefficients and the lattice property of the sequence spaces $s^{\overline{r}}_{p,q}a$ further imply

$$\big\|\big(\langle \mathcal{E}^*f, \Psi_{\overline{k},\gamma}\rangle\big)_{\overline{k},\gamma}\,\big|\, s^{\overline{r}}_{p,q}a\,\big\| \le \big\|\big(\langle \mathcal{E}f, \Psi_{\overline{k},\gamma}\rangle\big)_{\overline{k},\gamma}\,\big|\, s^{\overline{r}}_{p,q}a\,\big\|\,,$$

which in view of Theorem 4.3.1 yields

$$\|\, f \,|S^{\overline{r}}_{p,q}A(\Omega)\| \sim \|\, \mathcal{E}^*f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\|\,.$$

Immediately from the definition of $\mathcal{E}^*f$ if follows that in this case the sequence of wavelet-coefficients $\big(\langle \mathcal{E}^*f, \Psi_{\overline{k},\gamma}\rangle\big)_{\overline{k},\gamma}$ may be interpreted as an element of $s^{\overline{r}}_{p,q}a(\nabla)$, where

$$\nabla_{\overline{k}} = \big\{\gamma \in \Gamma_{\overline{k}} \,:\, \operatorname{supp}\Psi_{\overline{k},\gamma} \cap \Omega \neq \emptyset\big\}\,. \tag{7.1.1}$$

Moreover, from the support property of ψ_0 and ψ_1 we conclude

$$\operatorname{supp}\mathcal{E}^*f \subset \Gamma = \Big\{x \in \mathbb{R}^d \,:\, \operatorname{dist}(x, \Omega) < 2M\Big\}\,.$$

It follows that the mentioned space $s^{\overline{r}}_{p,q}a(\nabla)$ can be identified as a subspace of $s^{\overline{r}}_{p,q}a(\Gamma)$. Both interpretations motivate in hindsight the definition of $\nabla(\Omega)$ in (5.1.2).
We want to emphasize the point that, in contrast to the characterization in Theorem 4.3.1, we no longer have an isomorphism mapping the function spaces onto sequence spaces. This stems from the fact that in general the mappings $f \longmapsto \mathcal{E}f$ and $f \longmapsto \mathcal{E}^*f$ are nonlinear. However, since these mappings are bounded we can derive sufficient conditions for embeddings and estimates from above for the error of the best m-term approximation in this way directly from the results for sequence spaces.
As pointed out before, the necessary conditions for embeddings can directly be obtained from the ones in the isotropic setting by tensor product arguments, compare with Section 5.3.1. For estimates from below for m-term approximation we have to argue slightly differently. We now define the sequence $\widetilde{\nabla}$ by $\widetilde{\nabla}_{\overline{k}} = \big\{\gamma \in \Gamma_{\overline{k}} \,:\, \operatorname{supp}\Psi_{\overline{k},\gamma} \subset \Omega\big\}$. Then we find for every $f = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\widetilde{\nabla}_{\overline{k}}} \lambda_{\overline{k},\gamma}\Psi_{\overline{k},\gamma}$ with $\lambda \in s^{\overline{r}}_{p,q}a(\widetilde{\nabla})$

$$\|\, f \,|S^{\overline{r}}_{p,q}A(\Omega)\| \sim \|\, f \,|S^{\overline{r}}_{p,q}A(\mathbb{R}^{\overline{d}})\| \sim \|\, \lambda \,|s^{\overline{r}}_{p,q}a(\widetilde{\nabla})\|\,.$$

Though the sets $\widetilde{\nabla}_{\overline{k}}$ are possibly empty, it can easily be seen that for some suitably chosen set $\widetilde{\Gamma}$ we have $\nabla_{\overline{k}}(\widetilde{\Gamma}) \subset \widetilde{\nabla}_{\overline{k}}$ for every $\overline{k} \in \mathbb{N}_0^N$ with $|\overline{k}| \ge K$. This follows from the assumed openness of Ω.

The following lemma summarizes these constructions.

Lemma 7.1.1. For every domain Ω in $\mathbb{R}^d$ there are domains Γ and $\widetilde{\Gamma}$, such that $\widetilde{\Gamma} \subset \Omega \subset \Gamma$. Moreover, there is a (nonlinear) bounded mapping $\mathcal{I} : S^{\overline{r}}_{p,q}A(\Omega) \longrightarrow s^{\overline{r}}_{p,q}a(\Gamma)$ and a linear injective mapping $\mathcal{J} : s^{\overline{r}}_{p,q}a(\widetilde{\Gamma}) \longrightarrow S^{\overline{r}}_{p,q}A(\Omega)$, such that

$$\| f \,|S^{\overline{r}}_{p,q}A(\Omega)\| \sim \| \mathcal{I}f \,|\, s^{\overline{r}}_{p,q}a(\Gamma) \| , \qquad \mathcal{I}f = \big(\langle \mathcal{E}^* f, \Psi_{\overline{k},\gamma}\rangle\big)_{\overline{k},\gamma} ,$$

where $\mathcal{E}^*$ is defined as above, and

$$\| \mathcal{J}\lambda \,|S^{\overline{r}}_{p,q}A(\Omega)\| \sim \| \lambda \,|\, s^{\overline{r}}_{p,q}a(\widetilde{\Gamma}) \| , \qquad \mathcal{J}\lambda = \sum_{\overline{k}\in\mathbb{N}_0^N} \sum_{\gamma\in\widetilde{\nabla}_{\overline{k}}} \lambda_{\overline{k},\gamma}\Psi_{\overline{k},\gamma} .$$

Note that this method does not yield intrinsic characterizations of the spaces $S^{\overline{r}}_{p,q}A(\Omega)$. Concerning such intrinsic wavelet characterizations, at least for isotropic spaces $A^s_{p,q}(\Omega)$, we refer to the monograph [87]. However, this lemma is sufficient for our purposes, since our results for the sequence spaces $s^{\overline{r}}_{p,q}a(\Omega)$ do not depend directly on Ω or $\nabla = \nabla(\Omega)$, as long as we have the properties (5.1.3) and (5.1.4).

7.1.2 m-term approximation in function spaces

After the preparations in the previous section we are now in the position to transfer our results for the best m-term approximation from the sequence spaces $s^{\overline{r}}_{p,q}a(\Omega)$ to function spaces. This is based on the following lemma, which itself is an immediate consequence of Lemma 7.1.1.

Lemma 7.1.2. Let $\Omega \subset \mathbb{R}^d$ be an arbitrary domain. Moreover, let $0 < p_0, p_1, q_0, q_1 \le \infty$ ($p_0, p_1 < \infty$ for F-spaces) and $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$ be parameters, such that we have a continuous embedding $S^{\overline{r}^{\,0}}_{p_0,q_0}X(\Omega) \hookrightarrow S^{\overline{r}^{\,1}}_{p_1,q_1}Y(\Omega)$, $X, Y \in \{B, F\}$. Then it holds

$$\begin{aligned}\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}X(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}Y(\Omega)\,, \Psi_\Omega\Big) &\lesssim \Big(s^{\overline{r}^{\,0}}_{p_0,q_0}x(\nabla)\,, s^{\overline{r}^{\,1}}_{p_1,q_1}y(\nabla)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla}\Big)\\ &\le \Big(s^{\overline{r}^{\,0}}_{p_0,q_0}x(\Gamma)\,, s^{\overline{r}^{\,1}}_{p_1,q_1}y(\Gamma)\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla(\Gamma)}\Big) ,\end{aligned}$$

where ∇ is defined as in (7.1.1), and the system Ψ_Ω is given accordingly by

$$\Psi_\Omega = \Big\{ \Psi_{\overline{k},\gamma} \in \Psi \,:\, \operatorname{supp} \Psi_{\overline{k},\gamma} \cap \Omega \ne \emptyset \Big\} .$$

On the other hand, it holds

$$\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}X(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}Y(\Omega)\,, \Psi_\Omega\Big) \gtrsim \Big(s^{\overline{r}^{\,0}}_{p_0,q_0}x(\widetilde{\Gamma})\,, s^{\overline{r}^{\,1}}_{p_1,q_1}y(\widetilde{\Gamma})\,, \mathcal{B}_{\mathbb{N}_0^N\times\nabla(\widetilde{\Gamma})}\Big) .$$

The sets Γ and $\widetilde{\Gamma}$ have the same meaning as in Lemma 7.1.1.

We now begin with the limiting case.

Theorem 7.1.1. Let $0 < p_0 \le p_1 \le \infty$ ($p_0, p_1 < \infty$ for F-spaces), $0 < q_0, q_1 \le \infty$ and $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$, such that

$$\overline{r}^{\,0} - \overline{d}/p_0 = \overline{r}^{\,1} - \overline{d}/p_1 .$$

Moreover, let Ω be an arbitrary domain in $\mathbb{R}^d$.

(i) Let additionally $q_0 \le q_1$, and put

$$\alpha = \min\Big(\frac{1}{p_0} - \frac{1}{p_1}\,,\, \frac{1}{q_0} - \frac{1}{q_1}\Big)$$

Then it holds

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}B(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}B(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\alpha}\,, \qquad m \ge 1\,,$$

where Ψ_Ω is the wavelet system defined in Lemma 7.1.2.

(ii) Let $p_0 < p_1$ and either

(a) $\frac{1}{p_0} - \frac{1}{p_1} > \frac{1}{q_0} - \frac{1}{q_1}$ and $q_1 \le p_1$, or

(b) $\frac{1}{p_0} - \frac{1}{p_1} < \frac{1}{q_0} - \frac{1}{q_1}$ and $q_1 \ge p_1$, or

(c) $q_1 = p_1$.

Then it holds for $m \ge 2$

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}F(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}F(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})_+}\,.$$

(iii) Let $p_0 \le q_0 \le p_1$. Then it holds for $m \ge 1$

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}B(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}F(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\frac{1}{q_0}+\frac{1}{p_1}}\,.$$

Moreover, in case $q_0 \le p_0$ let either $\frac{1}{p_0} - \frac{1}{p_1} \le \frac{1}{q_0} - \frac{1}{q_1}$, or $\frac{1}{p_0} - \frac{1}{p_1} > \frac{1}{q_0} - \frac{1}{q_1}$ and $q_1 \le p_1$. Then it holds for $m \ge 2$

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}B(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}F(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\big(\log m\big)^{(N-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})_+}\,.$$

(iv) Let $p_0 \le q_1 \le p_1$. Then it holds for $m \ge 1$

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}F(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}B(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{q_1}}\,.$$

Furthermore, in case $p_1 \le q_1$ and $\frac{1}{p_0} - \frac{1}{p_1} \le \frac{1}{q_0} - \frac{1}{q_1}$ it holds

$$\sigma_m\Big(S^{\overline{r}^0}_{p_0,q_0}F(\Omega)\,,S^{\overline{r}^1}_{p_1,q_1}B(\Omega)\,,\Psi_\Omega\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \qquad m \ge 1\,.$$

Proof. Part (i) is the counterpart of Theorem 6.7.2. Moreover, part (ii) follows from Theorems 6.7.3 and 6.10.1, and part (iii) is either a consequence of Proposition 6.7.3, or of (ii) and the elementary embedding $S^{\overline{r}^0}_{p_0,q_0}B(\Omega) \hookrightarrow S^{\overline{r}^0}_{p_0,q_0}F(\Omega)$. Finally, (iv) corresponds to Proposition 6.7.4. Those results are complemented by the Conjectures 6.7.1–6.7.3. We shall add that in the above theorem the case $\Omega = \mathbb{R}^d$ is admitted. □

Now we turn to the non-limiting case. We begin with the result for spaces on the whole of $\mathbb{R}^d$. The next theorem is the immediate counterpart of Theorem 6.6.1 using the wavelet isomorphism from Corollary 4.3.1.

Theorem 7.1.2. Let $0 < p_0 \le p_1 \le \infty$ and $0 < q_0, q_1 \le \infty$, and let $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$ such that

$$\overline{r}^{\,0} - \overline{r}^{\,1} > \overline{d}\Big(\frac{1}{p_0} - \frac{1}{p_1}\Big) .$$

Then it holds for all $X, Y \in \{B, F\}$

$$\sigma_m\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}X(\mathbb{R}^{\overline{d}})\,, S^{\overline{r}^{\,1}}_{p_1,q_1}Y(\mathbb{R}^{\overline{d}})\,, \Psi\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \qquad m \ge 1\,,$$

where $p_0 < \infty$ if $X = F$ and $p_1 < \infty$ if $Y = F$.

The situation on bounded domains, i.e. for a compact embedding, is slightly more complicated.

Theorem 7.1.3. Let $0 < p_0, p_1, q_0, q_1 \le \infty$, and let $\overline{r}^{\,0}, \overline{r}^{\,1} \in \mathbb{R}^N$ such that

$$\overline{r}^{\,0} - \overline{r}^{\,1} = t\overline{d} \qquad \text{for some } t \in \mathbb{R} \text{ with } \quad t > \Big(\frac{1}{p_0} - \frac{1}{p_1}\Big)_+ .$$

Finally, let Ω be a bounded domain in $\mathbb{R}^d$.

(i) Then it holds for $m \ge 2$

$$\sigma_m\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}B(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}B(\Omega)\,, \Psi_\Omega\Big) \sim m^{-t}\big(\log m\big)^{(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})_+} .$$

(ii) Let $p_0, p_1 < \infty$, and let either

(a) $t > \frac{1}{q_0} - \frac{1}{q_1}$, or

(b) $\big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+ < t < \frac{1}{q_0} - \frac{1}{q_1}$ and $p_1 < q_1$, or

(c) $\big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+ < t \le \frac{1}{q_0} - \frac{1}{q_1}$ and $p_1 = q_1$.

Then it holds for $m \ge 2$

$$\sigma_m\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}F(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}F(\Omega)\,, \Psi_\Omega\Big) \sim m^{-t}\big(\log m\big)^{(N-1)\big(t-\frac{1}{q_0}+\frac{1}{q_1}\big)_+} .$$

Moreover, in the remaining cases for $t \le \frac{1}{q_0} - \frac{1}{q_1}$ it holds for every $\varepsilon > 0$

$$m^{-t} \lesssim \sigma_m\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}F(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}F(\Omega)\,, \Psi_\Omega\Big) \lesssim m^{-t}\big(\log m\big)^{\varepsilon} .$$

(iii) Now assume

$$t > \Big(\frac{1}{\min(p_0, q_0)} - \frac{1}{\max(p_1, q_1)}\Big)_+ .$$

Then it holds for all $X, Y \in \{B, F\}$ and $m \ge 2$

$$\sigma_m\Big(S^{\overline{r}^{\,0}}_{p_0,q_0}X(\Omega)\,, S^{\overline{r}^{\,1}}_{p_1,q_1}Y(\Omega)\,, \Psi_\Omega\Big) \sim m^{-t}\big(\log m\big)^{(N-1)(t-\frac{1}{q_0}+\frac{1}{q_1})} ,$$

where $p_0 < \infty$ if $X = F$ and $p_1 < \infty$ if $Y = F$. Moreover, for every $f \in S^{\overline{r}^{\,0}}_{p_0,q_0}X(\Omega)$ a near best approximation can be constructed explicitly.

To aid our comparisons in the next section, we shall specialize the above results to the most interesting case in applications, approximation in L_p. We remind on the Littlewood-Paley-type assertion $S^{\overline{r}}_{p,2}F(\mathbb{R}^{\overline{d}}) = S^{\overline{r}}_{p}H(\mathbb{R}^{\overline{d}})$. If we define Sobolev spaces on domains, $S^{\overline{r}}_{p}H(\Omega)$, likewise via restriction, this identity immediately carries over.

Theorem 7.1.4. Let $1 < p_0, p_1 < \infty$ and $t \in \mathbb{R}$, where $t > \big(\frac{1}{p_0} - \frac{1}{p_1}\big)_+$. Moreover, let Ω be a bounded domain in $\mathbb{R}^d$. Then it holds for $m \geq 2$

$$\sigma_m\Big(S^{t\overline{d}}_{p_0}H(\Omega)\,, L_{p_1}(\Omega)\,, \Psi_\Omega\Big) \sim m^{-t}\big(\log m\big)^{(N-1)t}\,.$$

Furthermore, if additionally either $t > \big(\frac{1}{\min(p_0,q_0)} - \frac{1}{\max(p_1,2)}\big)_+$ or $p_1 \geq 2$, then it holds

$$\sigma_m\Big(S^{t\overline{d}}_{p_0,q_0}B(\Omega)\,, L_{p_1}(\Omega)\,, \Psi_\Omega\Big) \sim m^{-t}\big(\log m\big)^{(N-1)(t-\frac{1}{q_0}+\frac{1}{2})_+}\,.$$

7.2 Comparison to the literature

In this section we want to compare our results on m-term approximation with those ones obtained by Temlyakov [79] and Dinh Dung [21, 22]. Unfortunately the classes of functions studied by these authors differ slightly from our scales $S^{\overline{r}}_{p,q}A(\Omega)$. Though our approach via decomposition techniques and henceforth discretizing the function spaces hides this effect, the results on m-term approximation clearly depend heavily also on the dictionary used. This is made obvious when comparing the results of m-term approximation for different classes of periodic functions with so-called m-term trigonometric approximation, i.e. m-term approximation with respect to the multivariate trigonometric system $(e^{ik\cdot x})_{k\in\mathbb{Z}^d}$.
To begin with we describe the setting used by Temlyakov in [79]. In this and related articles he mainly considers two scales of spaces of periodic functions defined on the d-dimensional torus $\mathbb{T}^d$, which are denoted by MH^r_q and MW^r_q, $r > 0$, $1 < q < \infty$, with the error of approximation being measured in the $L_p(\mathbb{T}^d)$-norm, $1 < p < \infty$. The Besov-Nikol'skij-type spaces MH^r_q are introduced via iterated differences, and are most closely connected to (the unit ball of) $S^{\overline{r}}_{q,\infty}B([0,1]^d)$, $\overline{r} = (r,\ldots,r) \in \mathbb{R}^d$, $\overline{d} = (1,\ldots,1)$, in our notation. On the other hand, the Sobolev-type spaces MW^r_q are defined via convolutions with multivariate Bernoulli kernels. These spaces turn out to be closely related to (the unit ball of) the Sobolev spaces $S^{\overline{r}}_{q}H([0,1]^d) = S^{\overline{r}}_{q,2}F([0,1]^d)$, where $\overline{r}$ and $\overline{d}$ are as before. For a more detailed comparison of the spaces MH^r_q and MW^r_q to the scale of periodic Besov and Triebel-Lizorkin spaces with dominating mixed smoothness we refer to [91, Section 2.7]. Finally, the basis U^d which is considered consists of translates of tensor products of Dirichlet kernels.
We now present the main results of [79]. With subsequent comparisons in mind we shall use the notations $\widetilde{S}^r_{q,\infty}B(\mathbb{T}^d)$ and $\widetilde{S}^r_{q,2}F(\mathbb{T}^d)$ instead of MH^r_q and MW^r_q, respectively.

Theorem 7.2.1 (Temlyakov [79]). Let $1 < q, p < \infty$.

(i) We put

$$r_1(q,p) = \begin{cases} (1/q - 1/p)_+\,, & p \geq 2\,, \\ \big(\max(2/p, 2/q) - 1\big)/p\,, & p < 2\,. \end{cases}$$

Then it holds for all $r > r_1(q,p)$

$$\sigma_m\Big(\widetilde{S}^r_{q,\infty}B(\mathbb{T}^d)\,, L_p(\mathbb{T}^d)\,, U^d\Big) \sim m^{-r}\big(\log m\big)^{(d-1)(r+1/2)}\,, \qquad m \geq 2\,.$$

(ii) Now we put

$$r_2(q,p) = \begin{cases} \max(1/q, 1/2) - 1/p\,, & p \geq 2\,, \\ \big(\max(2/p, 2/q) - 1\big)/p\,, & p < 2\,. \end{cases}$$

Then it holds for all $r > r_2(q,p)$

$$\sigma_m\Big(\widetilde{S}^r_{q,2}F(\mathbb{T}^d)\,, L_p(\mathbb{T}^d)\,, U^d\Big) \sim m^{-r}\big(\log m\big)^{(d-1)r}\,, \qquad m \geq 2\,.$$

Comparing Temlyakov's conditions $r > r_1(q,p)$ and $r > r_2(q,p)$ with our condition $t > \big(\frac{1}{\min(p_0,q_0)} - \frac{1}{\max(p_1,q_1)}\big)_+$ for according parameters we find that they coincide for $p \geq 2$, but for $p < 2$ the latter condition is the weaker one. In other words, Temlyakov's results correspond to the constructions in Section 6.7.5, but under more restrictive assumptions. However, examining Temlyakov's proofs it becomes clear that the imposed restrictions are rather artificial and caused by the techniques applied. In [79] and other related articles he concentrated on a particular class of explicit constructions of approximants, so-called greedy-type approximations, and studied their efficiency in comparison to the best m-term approximation, see also the recent survey [80].

Dinh Dung's setting is similar, but there are also several important differences. He also works with scales of Besov-type spaces $B^r_{p,\theta}$ and Sobolev-type spaces W^r_p, $0 < p, \theta \leq \infty$, $r \in \mathbb{R}$, which correspond to spaces $S^{\overline{r}}_{p,\theta}B([0,1]^d)$ and $S^{\overline{r}}_p H([0,1]^d)$ in our scale, where as before $\overline{r} = (r, \ldots, r) \in \mathbb{R}^d$ and $\overline{d} = (1, \ldots, 1)$. Using Weyl-derivatives both scales of spaces are defined as subsets of L^0_p, which consist of those periodic functions from L_p, whose integrals with respect to every variable x_i, $1 \leq i \leq d$, vanish, i.e. $\int_{\mathbb{T}} f(x)dx_i = 0$. Hence these spaces differ from the ones used by Temlyakov as well. The dictionary $\mathbb{V}$ used by Dinh Dung consists of translates of tensor products of de la Vallée Poussin kernels. In contrast to all previously occurring dictionaries this one is linearly dependent. However, it admits discretization techniques similar to wavelet-type bases.
We now state Dinh Dung's results, first in the case of two Besov-type spaces, and afterwards for approximation in L_q. Similar than before, we denote the occurring spaces by $\widehat{S}^r_{p,\theta}B$ and $\widehat{S}^r_p W$ instead of $B^r_{p,\theta}$ and W^r_p, respectively.

Theorem 7.2.2 (Dinh Dung [21, 22]). Let $0 < p, q, \theta, \tau \leq \infty$ and $r \in \mathbb{R}$.

(i) Then it holds

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, \widehat{S}^r_{q,\tau}B(\mathbb{T}^d)\,, \mathbb{V}\Big) \gtrsim m^{-r}\big(\log m\big)^{(d-1)(r+1/\tau-1/\theta)}\,, \quad m \geq 2\,.$$

(ii) Let $1 \leq \tau \leq \theta \leq \infty$ and $r > (1/p - 1/q)_+$. Then it holds

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, \widehat{S}^r_{q,\tau}B(\mathbb{T}^d)\,, \mathbb{V}\Big) \sim m^{-r}\big(\log m\big)^{(d-1)(r+1/\tau-1/\theta)}\,, \quad m \geq 2\,.$$

(iii) If $1 \leq \tau \leq \infty$ and $r > 1/p$, then it holds for $m \geq 2$

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, \widehat{S}^r_{q,\tau}B(\mathbb{T}^d)\,, \mathbb{V}\Big) \lesssim m^{-r}\big(\log m\big)^{(d-1)(r+1/\tau-1/\max(p,\theta))}\,.$$

The estimate in (ii) coincides with the one in Theorem 6.10.2, the restriction $\tau \leq \theta$ corresponds to the case $q_1 \leq q_0$. Since Dinh Dung constructs near best approximations explicitly this complements our constructions from Section 6.7.5.
In case $\theta \geq p$ the restriction $r > 1/p$ in (iii) is stronger than (6.7.10), i.e. this result corresponds to the case of high smoothness. On the other hand, in case $\theta < p$ the result is no longer sharp, and it corresponds once more to the case of high smoothness, combined with the elementary embedding $\widehat{S}^r_{p,\theta}B(\mathbb{T}^d) \hookrightarrow \widehat{S}^r_{p,p}B(\mathbb{T}^d)$.

Theorem 7.2.3 (Dinh Dung [21, 22]). Let $1 < p, q < \infty$, $0 < \theta \leq \infty$ and $r \in \mathbb{R}$.

(i) Then it holds

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, L_q(\mathbb{T}^d)\,, \mathbb{V}\Big) \gtrsim m^{-r}\big(\log m\big)^{(d-1)(r+1/2-1/\theta)}\,, \quad m \geq 2\,.$$

(ii) Let $r > (1/p - 1/q)_+$ and $\theta \geq \min(q,2)$. Then it holds for $m \geq 2$

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, L_q(\mathbb{T}^d)\,, \mathbb{V}\Big) \lesssim m^{-r}\big(\log m\big)^{(d-1)(r+1/\min(q,2)-1/\theta)}\,.$$

(iii) Let $r > \max\big(0, 1/p - 1/q, 1/p - 1/2,\big)$ and $2 \leq \theta \leq \infty$. Then it holds

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, L_q(\mathbb{T}^d)\,, \mathbb{V}\Big) \sim m^{-r}\big(\log m\big)^{(d-1)(r+1/2-1/\theta)}\,, \quad m \geq 2\,.$$

(iv) Let $r > 1/p$. Then it holds for $m \geq 2$

$$\sigma_m\Big(\widehat{S}^r_{p,\theta}B(\mathbb{T}^d)\,, L_q(\mathbb{T}^d)\,, \mathbb{V}\Big) \lesssim m^{-r}\big(\log m\big)^{(d-1)(r+1/2-1/\max(p,\theta))}\,.$$

(v) Now let $r > \max\big(0, 1/p - 1/q, 1/p - 1/2, 1/2 - 1/q\big)$. Then we have

$$\sigma_m\Big(\widehat{S}^r_{p,2}F(\mathbb{T}^d)\,, L_q(\mathbb{T}^d)\,, \mathbb{V}\Big) \sim m^{-r}\big(\log m\big)^{(d-1)r}\,, \quad m \geq 2\,.$$

In both theorems, the estimates from below coincide with the ones in Proposition 6.5.1 (at least the one in Step 2 of its proof).
The result in (ii) corresponds to Theorem 6.10.2, combined with the elementary embedding $L_q(\mathbb{T}^d) \hookrightarrow \widehat{S}^r_{p,\min(q,2)}B(\mathbb{T}^d)$. However, the resulting estimate is sharp only for $q \geq 2$.
The condition on r in (iii) coincides with (6.7.10) in case $p \leq \theta$. Only in case $2 \leq \theta \leq p$, the conditions in (iii) are weaker than those for Theorem 6.7.4. However, in that case one can use the elementary embedding $\widehat{S}^r_{p,\theta}B(\mathbb{T}^d) \hookrightarrow \widehat{S}^r_{p,\theta}F(\mathbb{T}^d)$, and the result is covered by Theorem 6.10.3 (Theorem 7.1.3(ii)).
Part (iv) can be discussed as part (iii) of the preceding theorem. Finally, observe that the condition in (v) can be reformulated as $r > \frac{1}{\min(p,2)} - \frac{1}{\max(q,2)}$, which coincides with the one for the explicit construction in Theorem 6.7.4.

Remark 7.2.1. We have to mention that the parts (ii) and (v) in Theorem 7.2.3 differ slightly from the results formulated by Dinh Dung. The results in Theorem 7.2.3 are consequences of Theorem 7.2.2 using elementary embeddings. However, an application of these embeddings either leads to additional restrictions on r as in (v), or increases the exponent of the logarithm as in (ii). In both cases this was not completely taken into account.

Apart from the situation in the above theorems Dinh Ding further dealt with several generalizations of the spaces $\widehat{S}^{r}_{p,q}B(\mathbb{T}^d)$, commonly denoted by $B^{A}_{p,q}(\mathbb{T}^d)$ and $B^{\Omega}_{p,q}(\mathbb{T}^d)$ (we will not give definitions here). Concerning those results we only want to mention that these spaces cover the case $S^{\overline{r}}_{p,q}B(\mathbb{T}^{\overline{d}})$, which would be the immediate periodic counterpart of $S^{\overline{r}}_{p,q}B(\mathbb{R}^{\overline{d}})$, even for general vectors $\overline{r} > 0$. In the asymptotics r would have to be replaced by $\rho = \min\big\{r_i/d_i \,:\, i = 1, \ldots, N\big\}$ (with some lower bound for ρ), and $N-1$ is replaced by $\nu = \#\big\{1 \leq i \leq N \,: r_i/d_i = \rho\big\} - 1$. We refer to [21, 23].
In our situation the treatment of general smoothness vectors $\overline{t} \neq t\overline{d}$ remains an interesting open problem.

References

[1] T. I. Amanov, *Spaces of differentiable functions with dominating mixed derivatives.* Nauka Kaz. SSR, Alma-Ata, 1976.

[2] R. J. Bagby, An extended inequality for the maximal function. *Proc. Amer. Math. Soc.* **48** (1975), 419–422.

[3] D. B. Bazarkhanov, Characterizations of the Nikol'skii-Besov and Lizorkin-Triebel function spaces of mixed smoothness. *Proc. Stekolv Inst. Math.* **243** (2003), 53–65.

[4] D. B. Bazarkhanov, Equivalent (quasi)normings of some function spaces of generalized mixed smoothness. *Proc. Steklov Inst. Math.* **248** (2005), 21–34. Translation from *Trudy Mat. Inst. Steklova* **248** (2005), 26–39.

[5] D. B. Bazarkhanov, Wavelet representations and equivalent normings of some function spaces of generalized mixed smoothness, *Math. Zh.*, **5** (2005), 12–16.

[6] C. Bennett, R. Sharpley, *Interpolation of operators.* Academic Press, Boston, 1988.

[7] J. Bergh, J. Löfström, *Interpolation Spaces.* Springer, Berlin, 1976.

[8] O. V. Besov, L. D. Kudryavtsev, P. I. Lizorkin, S. M. Nikol'skij, Investigations in the theory of spaces of differentiable functions of several variables. *Trudy Mat. Inst. Steklova* **182** (1988), 68–127 (Russian). Translated in *Proc. Steklov Inst. Math.* **182** (1990), 73–129.

[9] M. Bownik, Atomic and molecular decompositions of anisotropic Besov spaces. *Math. Zeitschrift* **250** (2005), 539–571.

[10] H. Brezis, P. Mironescu, Gagliardo-Nirenberg, composition and products in fractional Sobolev spaces. *J. Evol. Eq.* **1** (2001), 387–404.

[11] P. L. Butzer, K. Scherer, *Approximationsprozesse und Interpolationsmethoden.* Mannheim, Zürich, 1968.

[12] S. Dahlke, E. Novak, W. Sickel, Optimal approximation of elliptic problems by linear and nonlinear mappings II. *J. Complexity* **22** (2006), 549–603.

[13] I. Daubechies, *Ten lectures on wavelets.* Society for Industrial and Applied Math., Philadelphia, 1992.

[14] A. Defant, K. Floret, *Tensor Norms and Operator Ideals.* North Holland, Amsterdam, 1993.

[15] R. A. DeVore, B. Jawerth, V. Popov, Compression of wavelet decompositions. *Amer. J. Math.* **114** (1992), 737–785.

[16] R.A. DeVore, S.V. Konyagin, V.N. Temlyakov, Hyperbolic wavelet approximation. *Constr. Approx.* **14** (1998), 1–26.

[17] R. A. DeVore, G. G. Lorentz, *Constructive Approximation.* Grundlehren vol. 303, Springer-Verlag, Heidelberg, 1993.

[18] R. A. DeVore, V. Popov, Interpolation spaces and nonlinear approximation. In: *Function Spaces and Applications.* M. Cwikel et al. (eds.), Lecture Notes in Mathematics, vol. 1302. Springer, Berlin, 1988, 191–205.

[19] R. A. DeVore, Nonlinear Approximation. *Acta numerica* **7** (1998), 51–150.

[20] J. Diestel, J. J. Uhl, *Vector measures.* Mathematical Surveys and Monographs 15, AMS, Providence, Rhode Island, 1977.

[21] D. Dung, Continuous Algorithms in n-Term Approximation and Non-Linear Widths. *J. Approx. Theory* **102** (2000), 217–242.

[22] D. Dung, Asymptotic orders of optimal non-linear approximation. *East J. Approx.* **7** (2001), 55–76.

[23] D. Dung, Non-linear n-term approximations of smooth functions using wavelet decompositions. Proceedings of the third Asian Mathematical Conference 2000 (Eds. T. Sunada, P. W. Sy, Y. Lo). World Scientific, New Jersey, London, Singapore, Hongkong, 2002, 153–174.

[24] N. Dunford, J. Schwartz, *Linear operators, Vol. 1: General theory.* Wiley, New York, 1988.

[25] D. Edmunds, H. Triebel, *Function spaces, Entropy numbers, Differential operators.* Cambridge University Press, 1996.

[26] R. E. Edwards, *Functional Analysis. Theory and Applications.* New York, Chicago, San Francisco, Toronto, London: Holt, Rinehardt and Winston, 1965.

[27] C. Fefferman, E. M. Stein, Some maximal inequalities. *Amer. J. Math.* **93** (1971), 107–115.

[28] R. Fefferman, E. M. Stein, Singular integrals on product spaces. *Adv. Math.* **45** (1982), 117–143.

[29] M. Frazier, B. Jawerth, The φ-transform and applications to distribution spaces. In: *Function spaces and applications*, Proc. Conf. Lund 1986, Lect. Notes Math. 1302. Berlin, Springer, 1988, 223–246.

[30] M. Frazier, B. Jawerth, A discrete transform and decomposition of distribution spaces. *J. Functional Anal.* **93** (1990), 34–170.

[31] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo* **8** (1956), 1–79.

[32] A. Grothendieck, Produits tensoriels topologiques et espaces nucleaires. *Memoirs Amer. Math. Soc.* **16** (1955).

[33] M. Hansen, J. Vybíral, The Jawerth-Franke embedding of spaces of dominating mixed smoothness. *Georg. J. Math.* **16** (4) (2009), 667–682.

[34] M. Hansen, W. Sickel, Best m-term approximation and Lizorkin-Triebel spaces. Preprint **22**, DFG-SPP 1324, Marburg, 2009.

[35] M. Hansen, W. Sickel, Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint **39**, DFG-SPP 1324, Marburg, 2010.

[36] C.C. Hsiao, B. Jawerth, B.J. Lucier, X.M. Yu, Near optimal compression of orthonormal wavelet expansions. In: *Wavelets: mathematics and applications*, pp. 425–446, CRC, Boca Raton, 1994.

[37] T. Hytönen, P. Portal, Vector-valued multiparameter singular integrals and pseudodifferential operators. *Adv. Math.* **217** (2008), 519–536.

[38] L. Jantscher, *Distributionen.* Walter de Gruyter, Berlin, New York, 1971.

[39] B. Jawerth, Some observations on Besov and Triebel-Lizorkin spaces. *Math. Scand.* **40** (1977), 94–104.

[40] B. Jawerth, M. Milman, Wavelets and best approximation in Besov spaces. *Isr. Math. Conf. Proc.* **5** (1992), 107–112.

[41] B. Jawerth, M. Milman, Weakly rearrangement invariant spaces and approximation by largest elements. *Contemporary Math.* **445** (2007), 103–110.

[42] J. P. Kahane, P. G. Lemarié-Rieusset, *Fourier series and wavelets.* Gordon and Breach, London, 1995.

[43] H. Kempka, Atomic, molecular and wavelet decomposition for generalized 2-microlocal Besov spaces. *J. Function Spaces Appl.*, to appear.

[44] G. Kerkyacharian, D. Picard, V. N. Temlyakov, Some inequalities for the tensor product of greedy bases and weight-greedy bases. *East J. Approx.* **12** (2006), 103–118.

[45] G. Kyriazis, Non-linear approximation and interpolation spaces. *J. Approx. Theory* **113** (2001), 110–126.

[46] Y. Meyer, *Ondelettes et opérateurs I: Ondelettes.* Hermann, Paris, 1990; Engl. Transl. *Wavelets and operators.* Cambridge Univ. Press, Cambridge, 1992.

[47] E. H. Lieb, M. Loss, *Analysis, Second Edition.* Graduate Studies in Mathematics 14, 2001.

[48] W. A. Light, E. W. Cheney, *Approximation theory in tensor product spaces.* Lecture Notes in Math. 1169, Springer, Berlin, 1985.

[49] J. L. Lions, J. Peetre, Sur une classe d'espaces d'interpolation. *Inst. Hautes Études Sci. Publ. Math.* **19** (1964), 5–68.

[50] P. I. Lizorkin, Multipliers of Fourier integrals and bounds of convolution in spaces with mixed norms. Applications. *Izv. Akad. Nauk SSSR Ser. Mat. Tom* **34** (1970), No. 1, 225–256.

[51] P. I. Lizorkin, S. M. Nikol'skij, Function spaces of mixed smoothness from the decomposition point of view. *Trudy Mat. Inst. Steklova* **187** (1989), 143–161 (Russian). Translation: Studies in the theory of differentiable functions of several variables and its applications. *Proc. Steklov Inst. Math.* **187** (1990), 163–184.

[52] U. Luther, Representation, interpolation, and reiteration theorems for generalized approximation spaces. *Ann. Math. Pura Appl.* **182** (2003), 161–200.

[53] U. Luther, J. M. Almira, Generalized approximation spaces and applications. *Math. Nachr.* **263-264** (2004), 3–35.

[54] A. Nagel, E. M. Stein, On the product theory of singular integrals. *Rev. Mat. Iberoamericana* **20** (2004), 531–561.

[55] P.-A. Nitsche, Best N-term approximation spaces for tensor product wavelet bases. *Constr. Approx.* **24** (2006), 49–70.

[56] J. Peetre, *New thoughts on Besov spaces.* Duke Univ. Math. Series I, Durham, N.C., 1976.

[57] J. Peetre, Applications de la théorie des espaces d'interpolation dans l'analyse harmonique. *Ricerche Mat.* **15** (1966), 1–34.

[58] J. Peetre, On spaces of Triebel-Lizorkin type. *Ark. Mat.* **13** (1975), 123–130.

[59] J. Peetre, G. Sparr, Interpolation of normed abelian groups. *Ann. Mat. Pura Appl.* **92** (1972), 217–262.

[60] A. Pietsch, Approximation spaces. *J. Approx. Theory* **32** (1980), 115–134.

[61] A. Pietsch, *Operator ideals.* Deutsch. Verlag Wiss., Berlin, 1978; North-Holland, Amsterdam-London-New York-Tokyo, 1980.

[62] A. Pietsch, *History of Banach spaces and linear operators.* Birkhäuser, Boston, 2007.

[63] V. S. Rychkov, On a theorem of Bui, Paluszynski and Taibleson. *Proc. Steklov Inst. Math.* **227** (1999), 280–292.

[64] V. S. Rychkov, Littlewood-Paley theory and Function spaces with A_p^{loc} weights. *Math. Nachr.* **224** (2001), 145–180.

[65] S. G. Samko, A. A. Kilbas, O. I. Marichev, *Fractional integrals and derivatives: theory and applications.* Gordon and Breach Science Publishers, Yverdon, 1993.

[66] Y. Sawano, H. Tanaka, Decompositions of BesovMorrey spaces and TriebelLizorkinMorrey spaces. *Math. Zeitschrift* **257** (2007), 871–905.

[67] R. Schatten, On the direct product of Banach spaces. *Trans. Amer. Math. Soc.* **53** (1943), 195–217.

[68] H. J. Schmeisser, *Über Räume von Funktionen und Distributionen mit dominierenden gemischten Glattheitseigenschaften vom Besov-Triebel-Lizorkin Typ.* Habilitation Thesis, Jena, 1982.

[69] H.-J. Schmeisser, Recent developments in the theory of function spaces with dominating mixed smoothness. Proc. Conf. NAFSA-8, Prague 2006, (ed. J. Rakosnik), Inst. of Math. Acad. Sci., Czech Republic, pp. 145–204, Prague, 2007.

[70] H.-J. Schmeisser, W. Sickel, Spaces of functions of mixed smoothness and approximation from hyperbolic crosses. *J. Approx. Theory* **128** (2004), 115–150.

[71] H.-J. Schmeisser, H. Triebel, *Topics in Fourier Analysis and Function Spaces.* Wiley, Chichester, 1987.

[72] L. Schwartz, *Théorie des Distributions.* Hermann, Paris, 1957, 1959.

[73] W. Sickel, H. Triebel, Hölder inequalities and sharp embeddings in function spaces of $B^s_{p,q}$ and $F^s_{p,q}$ type. *Z. Anal. Anw.* **14** (1995), 105–140.

[74] W. Sickel, T. Ullrich, Tensor products of Sobolev-Besov spaces and applications to approximation from the hyperbolic cross. *J. Approx. Theory* **161** (2009), 748–786.

[75] E. M. Stein, G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton, 1971.

[76] B. Stöckert, Ungleichungen vom Plancherel-Polya-Nikol'skij-Typ in gewichteten L^Ω_p-Räumen mit gemischten Normen. *Math. Nachr.* **86** (1978), 19–32.

[77] J.-O. Strömberg, A. Torchinsky, *Weighted Hardy spaces.* Lecture Notes in Math. 1381, Springer, New York, Berlin, Heidelberg, 1989.

[78] V. N. Temlyakov, The best m-term approximations and greedy algorithms. *Adv. Comp. Math.* **8** (1998), 249–265.

[79] V. N. Temlyakov, Greedy Algorithms with regard to Multivariate Systems with Special Structure. *Constr. Approx.* **16** (2000), 399–425.

[80] V. N. Temlyakov, Greedy approximation. *Acta Numerica* **17** (2008), 235–409.

[81] H. Triebel, *Höhere Analysis.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1972.

[82] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators.* Amsterdam, New York, Oxford: North-Holland Publ. Company, 1978, and: VEB Deutscher Verlag der Wissenschaften, Berlin, 1978.

[83] H. Triebel, *Theory of Function Spaces.* Birkhäuser, Basel, 1983.

[84] H. Triebel, *Theory of Function Spaces II.* Birkhäuser, Basel, 1992.

[85] H. Triebel, *Fractals and Spectra.* Birkhäuser, Basel, 1997.

[86] H. Triebel, *Theory of Function Spaces III.* Birkhäuser, Basel, 2006.

[87] H. Triebel, *Function spaces and wavelets on domains.* EMS Publishing House, Zürich, 2008.

[88] H. Triebel, General function spaces III. *Anal. Math.* **3** (1977), 221–249.

[89] H. Triebel, Local Means and Wavelets in Function spaces. Proc. Conf. Function Spaces 8, Bêdlewo, 2006. *Banach Center Publications*, Volume **79** (2007), 215–234.

[90] P. Turpin, Représentation fonctionelle des espaces vectorielles topologiques. *Studia Math.* **73** (1982), 1–10.

[91] T. Ullrich, Smolyak's Algorithm, Sparse Grid Approximation and Periodic Function Spaces with Dominating Mixed Smoothness, PhD-Thesis, FSU Jena, 2007.

[92] T. Ullrich, Local mean characterization of Besov-Triebel-Lizorkin type spaces with dominating mixed smoothness on rectangular domains. Preprint, Bonn, 2008.

[93] A. P. Uninskij, Embedding theorems for a class of functions with mixed norm. *Dokl. Akad. Nauk SSSR* **166** (1966), 806–808 (Russian).

[94] J. Vybral, Function spaces with dominating mixed smoothness. *Dissertationes Mathematicae* **436** (2006), 1–73.

[95] P. Wojtaszczyk, *A mathematical introduction to wavelets.* Cambridge Univ. Press, Cambridge, 1997.

[96] P. Wojtaszczyk, Greedy algorith for general biorthogonal systems. *J. Approx. Theory* **107** (2000), 293–314.

[97] M. Yamazaki, The L^p-boundedness of pseudodifferential operators with estimates of parabolic type and product type. *J. Math. Soc. Japan* **38** (2) (1986), 199–225.

[98] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.* **98** (2004), 731–759.

(iii) M. Hansen and W. Sickel, *Best m-term approximation and Lizorkin-Triebel spaces*, DFG-SPP 1324 Preprint 22, submitted to JAT.

(iv) M. Hansen and W. Sickel, *Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings*, DFG-SPP 1324 Preprint 39, submitted to Constr. Approx.

(v) M. Hansen and W. Sickel, *Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness – the Case of Compact Embeddings*, DFG-SPP 1324 Preprint 44, submitted to Constr. Approx.

Participation at conferences and invitations

2007	Rhein-Ruhr-Workshop, Darmstadt (D)
	Summer school on Signal analysis, Jacobs University Bremen (D)
	Invitation of Prof. T. Strobl to Erwin-Schrödinger-Institut, Wien (A)
	Summer school on Time-frequency analysis, Inzell (D)
2008	FOCM conference, Hongkong
	Workshop "Function spaces and applications", Freyburg (D)
	Summer school on applied analysis, TU Chemnitz (D)
	Workshop "Spaces between us", Trest (CZ)
2009	Workshop "Approximation theory and signal analysis", Lindau (D)
	Spring School "Function spaces, inequalities and interpolation", Paseky (CZ)
	Conference on Time-Frequency, Strobl (A)
	Workshop "Algorithms and complexity for continuous problems", Dagstuhl (D)
	Workshop "Nonlinear and adaptive approximation", Günzburg (D)
	Annual meeting of the DFG-SPP 1324, Berlin (D)
2010	13th International conference on Approximation theory, San Antonio (USA)
	Workshop on Nonlinear Approximation, Marburg (D)
	Conference "Constructive theory of functions", Sozopol (Bulgaria)
	Summer school "Mini-courses in Mathematical Analysis", Padova (Italy)

Teaching

Since October 2005 I taught several tutorials to lectures on Calculus I–III to students of Mathematics, Physics and Geological and material sciences. During that time I also represented single lectures on various occasions and different topics.

Ehrenwörtliche Erklärung

Hiermit erkläre ich,

- dass mir die Promotionsordnung der Fakultät bekannt ist,
- dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte oder Ergebnisse eines Dritten oder eigenen Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe,
- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Disseration stehen,
- dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe.

Es waren keine weiteren Personen an der Auswahl und Auswertung von Materialien und der Herstellung des Manuskripts beteiligt. Ich habe weder eine gleiche, in wesentlichen Teilen ähnliche noch eine andere Abhandlung bereits bei einer anderen Hochschule als Dissertation eingereicht.

Jena, den 27.08.2010 Markus Hansen