

# A HANDBOOK OF HARMONIC ANALYSIS

YOSHIHIRO SAWANO

## CONTENTS

Preface	10
Acknowledgement	10
Orientation of this book	10
Notations in this book	13
<b>Part 1. A bird's-eye-view of this book</b>	<b>16</b>
1. Introduction	16
1.1. Maximal operator on $\partial\mathbb{D}$	16
1.2. Conjugate functions on $\partial\mathbb{D}$	22
1.3. Alternate version of $L^1(\partial\mathbb{D})$ -boundedness and Calderón-Zygmund operators	23
1.4. Concluding remarks	28
<b>Part 2. Fundamental facts of Fourier analysis</b>	<b>30</b>
<b>Part 3. Measure theory</b>	<b>30</b>
2. A quick review of general topology	30
3. Integration theory	32
3.1. Measures and outer measures	32
3.2. Construction of measures starting from a content	36
3.3. Measurable functions	40
3.4. Definition of the integral	46
3.5. Convergence theorems	54
3.6. Product measures and Fubini's theorem	58
4. Lebesgue spaces	62
4.1. $L^p(\mu)$ -spaces	62

4.2.	Integration formulae	66
4.3.	Convergence in measure	68
4.4.	Radon Nikodym derivative	70
4.5.	Convolution and mollifier	75
4.6.	Density	78
4.7.	Dual spaces in connection with measure theory	79
4.8.	Some calculus formulas	81
<b>Part 4. The Fourier transform</b>		<b>88</b>
5.	$\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$	88
5.1.	Schwartz Space $\mathcal{S}(\mathbb{R}^d)$	88
5.2.	Schwartz distribution $\mathcal{S}'$	90
5.3.	Convolution of $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$	96
6.	Schwartz distribution on the torus $\mathbb{T}^d$	99
6.1.	$2\pi$ -periodic functions	99
6.2.	Fourier series	100
7.	Fourier transform	107
7.1.	Definition and elementary properties	107
7.2.	Fourier transform for Schwartz distributions	110
7.3.	Examples	114
7.4.	Fourier transform of measures	119
8.	Sobolev spaces	128
8.1.	Weak derivative	128
8.2.	Sobolev space $W^{m,p}(\mathbb{R}^d)$	129
<b>Part 5. Elementary facts of functional analysis</b>		<b>138</b>
9.	Normed spaces	138
9.1.	Linear operators on normed spaces	138
9.2.	Resolvent	143
9.3.	Quotient topology and quotient vector spaces	144
9.4.	Banach Alaoglu theorem	146

9.5.	Hahn-Banach theorem	146
10.	Banach spaces and quasi-Banach spaces	150
10.1.	Elementary properties	150
10.2.	Baire category theorem and its applications	156
11.	Hilbert spaces	158
11.1.	Definitions and elementary properties	158
11.2.	Complete orthonormal system	162
11.3.	Bounded linear operators defined on Hilbert spaces	166
11.4.	Reproducing kernel Hilbert spaces	173
<b>Part 6.</b>	<b>Maximal operators and singular integral operators</b>	<b>177</b>
<b>Part 7.</b>	<b>Maximal operators</b>	<b>177</b>
12.	Maximal operators	178
12.1.	Definition of Hardy-Littlewood maximal operators and elementary properties	178
12.2.	$5r$ -covering lemma	180
12.3.	Weak-(1,1) boundedness of the Hardy-Littlewood maximal operators	182
13.	Applications and related topics	184
13.1.	Density argument	184
13.2.	Application to the Lebesgue differentiation theorem	185
13.3.	Application to the approximation to the unit	186
13.4.	Dyadic maximal operator	189
13.5.	Other covering lemmas and some related exercises	192
13.6.	Non-tangential maximal operator	198
<b>Part 8.</b>	<b>Singular integral operators</b>	<b>205</b>
14.	Hilbert transform	205
14.1.	Strong-(2,2) estimate	207
14.2.	Weak-(1,1) estimate	208
14.3.	Truncated Hilbert transform	210
14.4.	Cotlar's inequality and almost everywhere convergence	211
14.5.	Cotlar's lemma and its application	213

15.	Rotation method and the Riesz transform	215
15.1.	Rotation method	215
15.2.	Riesz transform	219
16.	Generalized singular integral operators	221
16.1.	Weak-(1,1) boundedness	224
16.2.	$L^p(\mathbb{R}^d)$ -boundedness	226
16.3.	Truncation and pointwise convergence	226
16.4.	Fourier multipliers	230
17.	Fractional integral operators	230
<b>Part 9. More about maximal and CZ-operators</b>		236
18.	The Hardy space $H^1(\mathbb{R}^d)$	236
19.	The space BMO	238
19.1.	Definition	238
19.2.	The John-Nirenberg inequality	242
19.3.	Duality $H^1$ -BMO( $\mathbb{R}^d$ )	245
20.	Sharp-maximal operators	247
20.1.	Definition	247
20.2.	Good $\lambda$ -inequality	248
20.3.	Sharp-maximal inequality	249
21.	Weighted norm estimates	251
21.1.	$A_1$ -weights	252
21.2.	$A_p$ -weights	257
21.3.	$A_\infty$ -weights	259
21.4.	$A_p$ -weights and singular integral operators	264
<b>Part 10. Probability theory, martingale and ergodicity</b>		271
<b>Part 11. Probability theory</b>		271
22.	Some elementary notions	271
22.1.	Probability spaces	271
22.2.	The Characteristic functions	274

22.3.	Conditional expectation	278
23.	Martingales with discrete time	283
23.1.	Martingales	283
23.2.	Decomposition of martingales	285
23.3.	Stopping time	287
23.4.	Elementary properties	287
24.	Properties of martingales	289
24.1.	The optional sampling theorem	289
24.2.	Doob's maximal inequality	290
24.3.	Convergence theorems of martingales	291
24.4.	Applications of convergence theorems	292
24.5.	The strong law of large numbers	293
24.6.	Uniform integrability	295
24.7.	Upcrossing time and almost sure convergence	297
<b>Part 12.</b>	<b>Ergodic theory</b>	<b>300</b>
25.	Ergodicity	300
26.	Ergodic maximal function	300
<b>Part 13.</b>	<b>Functional analysis and harmonic analysis</b>	<b>304</b>
<b>Part 14.</b>	<b>More about functional analysis</b>	<b>304</b>
27.	Bochner integral	304
27.1.	Measurable functions	304
27.2.	Definition of the Bochner integral	305
27.3.	Convergence theorems	310
27.4.	Fubini's theorem for Bochner integral	312
27.5.	$L^p(X; B)$ -spaces	312
28.	Semigroups	314
28.1.	Continuous semigroups	314
28.2.	Sectorial operators and semigroups	315
29.	Banach and $C^*$ -algebra	320

29.1.	Banach algebras	320
29.2.	$C^*$ -algebras	329
30.	Spectral decomposition of bounded self-adjoint operators	335
30.1.	Spectral decomposition	335
30.2.	Compact self-adjoint operators	339
<b>Part 15. Topological vector spaces</b>		<b>342</b>
31.	Nets and topology	342
32.	Topological vector spaces	344
32.1.	Definition	344
32.2.	Locally convex spaces	352
33.	Examples of locally convex spaces : $D(\Omega)$ and $E(\Omega)$	359
33.1.	$D(\Omega)$ and $D'(\Omega)$	359
33.2.	$E(\Omega)$ and $E'(\Omega)$	364
<b>Part 16. Interpolation</b>		<b>367</b>
34.	Interpolation	367
34.1.	Compatible couple	367
34.2.	Weak-type function spaces	369
34.3.	Interpolation techniques	370
35.	Interpolation functors	373
35.1.	Real interpolation functors	373
35.2.	Complex interpolation functors	382
36.	Interpolation of $L^p(\mu)$ -spaces	393
36.1.	Real interpolation of $L^p(\mu)$ -spaces	393
36.2.	Complex interpolation of $L^p(\mu)$ -spaces	398
<b>Part 17. Wavelets</b>		<b>403</b>
37.	Wavelets and scaling functions	403
37.1.	Definition	403
37.2.	Construction of wavelets for $d = 1$	405
37.3.	Examples when $d = 1$	408

38. Unconditional basis	412
38.1. Unconditional convergence	412
38.2. Bases for Banach spaces	416
38.3. Unconditional bases and applications to wavelets	417
39. Existence of unconditional basis	419
<b>Part 18. Vector-valued norm inequalities</b>	424
<b>Part 19. Vector-valued norm inequalities on <math>\mathbb{R}^d</math></b>	424
40. Vector-valued inequalities	424
41. Vector-valued inequalities for maximal operators	426
42. Vector-valued inequalities for CZ-operators	428
43. Kintchine's inequality and its applications	430
<b>Part 20. Littlewood-Paley theory</b>	434
44. Littlewood-Paley theory	434
44.1. $G$ -functionals	434
44.2. Discrete Littlewood-Paley theory	436
45. Burkholder-Gundy-Davis inequality	437
<b>Part 21. Function spaces appearing in harmonic analysis</b>	443
<b>Part 22. Functions on <math>\mathbb{R}</math></b>	444
Overview	444
46. 1-variable functions and their differentiability a.e.	444
46.1. Monotone functions	444
46.2. Functions of bounded variations	446
46.3. Absolutely continuous functions	447
46.4. Convex functions	450
47. $N$ -functions	451
47.1. Right-inverse	451
47.2. Definition and fundamental properties of $N$ -functions	452
48. Orlicz spaces	458

48.1.	Definition	458
48.2.	Convergence theorems	460
48.3.	Hölder inequality for Orlicz spaces	461
48.4.	Maximal operators and singular integral operators	463
49.	Reference for Chapter 22	465
<b>Part 23. Function spaces with one or two parameters</b>		468
50.	Hardy spaces $H^p(\mathbb{R}^d)$ with $0 < p \leq 1$	468
50.1.	Definition by means of atomic decomposition	469
50.2.	Maximal characterization of the Hardy space $H^p(\mathbb{R}^d)$	470
50.3.	Atomic decomposition vs. the grand maximal operator - I	477
50.4.	Atomic decomposition vs. the grand maximal operator - II	479
50.5.	Summary of Sections 50.3 and 50.4	490
51.	Lipschitz space	491
52.	Hölder-Zygmund spaces	493
52.1.	Hölder-Zygmund space $\mathcal{C}^\theta(\mathbb{R}^d)$	493
52.2.	Hölder-Zygmund spaces with higher regularity	495
52.3.	Interpolation of Hölder spaces	498
53.	Morrey spaces	499
53.1.	Boundedness of maximal operators	500
53.2.	Morrey's lemma	501
53.3.	Fractional integral operators	501
53.4.	Singular integral operators	503
54.	Commutators	505
54.1.	Commutators generated by BMO and singular integral operators	506
54.2.	Compactness	508
54.3.	Another type of commutators	510
<b>Part 24. Besov and Triebel-Lizorkin spaces</b>		513
55.	Band-limited distributions	513
55.1.	Maximal operator control	513



55.2. Multiplier theorems	517
55.3. Application to singular integral operators	519
56. Besov spaces and Triebel-Lizorkin spaces	520
56.1. Definition	520
56.2. Elementary properties	522
56.3. Elementary inclusions	526
56.4. Lift operators for nonhomogeneous spaces	528
57. The space $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$	530
57.1. Definition	530
57.2. $\mathcal{S}(\mathbb{R}^d)_0$ and $\mathcal{S}'(\mathbb{R}^d)'_0$	531
58. Spaces of homogeneous type	533
59. Concrete spaces	536
59.1. Potential spaces and Sobolev spaces	536
59.2. Lipschitz spaces	538
60. Other related function spaces	539
60.1. Modulation spaces	539
60.2. Herz spaces	540
60.3. Amalgam spaces	541
<b>Part 25. Applications to partial differential equations</b>	<b>543</b>
61. The heat semigroup	543
62. Pseudo-differential operators	545
62.1. Some heuristics	545
62.2. Pseudo-differential operators on $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$	545
63. $L^p(\mathbb{R}^d)$ -boundedness of pseudo-differential operators	547
63.1. $L^2(\mathbb{R}^d)$ -boundedness	547
63.2. $L^p(\mathbb{R}^d)$ -boundedness	550
<b>Part 26. Supplemental facts on measure theory</b>	<b>554</b>
<b>Part 27. Supplemental facts on measure theory</b>	<b>554</b>

Overview	554
64. Supplemental of the construction of measures	554
64.1. Topological groups and Haar measures	554
64.2. Hausdorff measures	558
64.3. Hausdorff metric and fractals	560
65. Countable product of probability spaces	566
Textbooks	570
References	570
Research papers	571
Personal communications	584

## PREFACE

This book is intended to serve as a comprehensive textbook of harmonic analysis with two goals; the first is to present typical arguments for readers to feel the flavor of the real-variable method. The other is to introduce various function spaces. Anyone that has even just scratched the surface of the theory of integration, general topology and functional analysis can start this book without much difficulty.

I wrote this book based on my teaching experience as well. The experience of my teaching assistant at the University of Tokyo helped a lot.

Throughout this book, we usually place ourselves in  $\mathbb{R}^d$  equipped with the Lebesgue measure. However, after Chapter 11, we sometimes place ourselves in the setting of the probability space  $(\Omega, \mathcal{F}, P)$ . My decision to deal with probability theory is to show that harmonic analysis has a lot to do with probability theory.

I hope that this book will be of service to the students wishing to specialize in harmonic analysis or who wish to scratch the surface of harmonic analysis. I intended to publish a book that contains topics when he struggled to study in 2002 and 2003.

Yoshihiro Sawano, Sagamihara.

## ACKNOWLEDGEMENT

This book is originally based on a seminar given at the University of Tokyo, Graduate School of Mathematical Sciences. However, after deciding to publish it, much content was added and the new material is based mostly on other lectures given there.

## Orientation of this book.

In Chapter 1 we have selected some typical problems and shown the readers the flavor of harmonic analysis.

After this bird's-eye-view of this book, we shall briefly review Lebesgue spaces in Chapter 3. In Chapter, 4 we take up Schwartz distributions as a review of elementary Fourier analysis, and collect some elementary topics of functional analysis in Chapter 5.

Chapter 7 is devoted to presenting elementary facts of maximal operator theories that are frequently applied to analysis, in general. In Chapter 8, we investigate the boundedness properties of singular integral operators from many points of view. In Chapter 9, we integrate Chapters 7 and 8 and investigate the Hardy-Littlewood maximal operator and singular integral operators in more depth.

Chapter 11 deals with the theory of martingales. It is the author's hope that the reader will come to see that there is a close connection between harmonic analysis and probability theory. Moreover, for the reader who is interested in only harmonic analysis, the discussion on probability theory can be skipped. In Chapter 12 we consider ergodic theory as an application of probability theory.

Chapter 14 deals with more specialized topics related to harmonic analysis, as well as we review elementary notions such as resolvent sets and compact operators. This chapter deals with the Bochner integral, semigroups and Banach algebra. Chapter 15 deals with more complicated structures: the space of tempered distributions  $\mathcal{S}'(\mathbb{R}^d)$  is a typical example of topological vector spaces. As a concrete example, we present a theory of the distribution space  $D'(\Omega)$  for open sets  $\Omega \subset \mathbb{R}^d$  and take up topological linear spaces which cannot be endowed with any (quasi-)norm. In Chapter 16 we reconsider interpolation, which somehow appeared in places like Chapters 7 and 8, and aim to develop a systematic theory of interpolation. Chapter 17 will serve as an introduction to wavelet theory, as well as an example of the usage of the Calderón-Zygmund theory. One of the purposes of this chapter is to consider the notion of basis.

Part 18 can be read independently of Part 13. In Chapter 19, we develop the  $\ell^q$ -valued extension. By Theorem 1.7 in Section 1, we know that

$$(0.1) \quad \|Mf\|_{L^p(\mathbb{T})} \leq c_p \|f\|_{L^p(\mathbb{T})}.$$

By "a vector-valued extension" we mean

$$(0.2) \quad \left\| \left( \sum_{j=1}^{\infty} Mf_j^q \right)^{\frac{1}{q}} \right\|_{L^p(\mathbb{T})} \leq c_{p,q} \left\| \left( \sum_{j=1}^{\infty} |f_j|^q \right)^{\frac{1}{q}} \right\|_{L^p(\mathbb{T})},$$

where  $0 < q < \infty$ . Here we modify (0.2) to define

$$\left\| \sup_{j \in \mathbb{N}} Mf_j \right\|_{L^p(\mathbb{T})} \leq c_p \left\| \sup_{j \in \mathbb{N}} |f_j| \right\|_{L^p(\mathbb{T})}.$$

The aim of this chapter is to consider such extensions. In Chapter 20, we present a powerful tool called the Littlewood-Paley theory. As well as the theory on  $\mathbb{R}^d$ , we consider its counterpart to probability theory, called the Burkholder-Gundy-Davis inequality.

In Part 21, we consider function spaces appearing in Fourier analysis and partial differential equations. In actual analysis, such as investigation of a particular differential equation, it is not enough to use the Sobolev space  $W^{m,p}(\mathbb{R}^d)$  with  $m \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$  and  $1 \leq p \leq \infty$ . We will have to measure fractional order differentiability. In Chapter 23, we consider function spaces carrying 1 or 2 parameters, while in Chapter 24, we exhibit some examples of applications of our results from functional analysis, function spaces, Littlewood-Paley theory and theory of singular integral operators. As an application of the maximal theory we take up the functions on

$\mathbb{R}$ . There are many properties describing the functions on  $\mathbb{R}$ : monotonicity, convexity, Lipschitz continuity, etc. Here covering lemmas that appeared in Chapter 7 play a crucial role. After setting down the key properties of such functions, in particular convex functions, we take up Orlicz spaces. We can say that the space  $L^p$  are made of the function  $\varphi(t) = t^p$ ,  $t > 0$ . From this viewpoint Orlicz spaces reflect behavior near  $t = 0$  and  $t = \infty$ . For example the Orlicz spaces contain a class of the function space corresponding to  $\min(t, t^2)$ .

Finally, in Part 26, we collect some other materials and supplement the topics in this book. Chapter 27 contains some further facts on measure theory.

Exercise problems are scattered throughout this book but, most of them are presented with some clues. Some problems are solved easily by re-examining or mimicking the proofs of other theorems. However, the author expects the reader to solve them so that they can be familiar with the arguments of harmonic analysis, more than just the statement of the propositions.

**Notations in this book.**

*Notation* (Sets and set functions).

- (1) We usually use the  $\ell^2$ -norm or the  $\ell^\infty$ -norm to define a ball in  $\mathbb{R}^d$ .
- (2) The metric ball defined by  $\ell^2$  is usually called a *ball*. We denote by  $B(x, r)$  the ball centered at  $x$  of radius  $r$ . Given a ball  $B$ , we denote by  $c(B)$  its center and by  $r(B)$  its radius. We write  $B(r)$  instead of  $B(o, r)$ , where  $o := (0, 0, \dots, 0)$ .
- (3) By “cube” we mean a compact cube whose edges are parallel to the coordinate axes. The metric ball defined by  $\ell^\infty$  is called a *cube*. If a cube has center  $x$  and radius  $r$ , we denote it by  $Q(x, r)$ . From the definition of  $Q(x, r)$ , its volume is  $(2r)^d$ . We write  $Q(r)$  instead of  $Q(o, r)$ . Given a cube  $Q$ , we denote by  $c(Q)$  the center of  $Q$  and by  $\ell(Q)$  the sidelength of  $Q$ :  $\ell(Q) = |Q|^{1/d}$ , where  $|Q|$  denotes the volume of the cube  $Q$ .
- (4) Given a cube  $Q$  and  $k > 0$ ,  $kQ$  means the cube concentric to  $Q$  with sidelength  $k\ell(Q)$ . Given a ball  $B$  and  $k > 0$ , we denote by  $kB$  the ball concentric to  $B$  with radius  $kr(B)$ .
- (5) Let  $E$  be a measurable set. Then we denote its indicator function by  $\chi_E$ . If  $E$  has positive measure and  $E$  is integrable over  $f$ , then denote by  $m_E(f)$  the average of  $f$  over  $E$ .  $|E|$  denotes the volume of  $E$ .
- (6) If we are working on  $\mathbb{R}^d$ , then  $\mathcal{B}$  denotes the set of all balls in  $\mathbb{R}^d$ , while  $\mathcal{Q}$  denotes the set of all cubes in  $\mathbb{R}^d$ . Be careful because  $\mathcal{B}$  can be used for a different purpose: When we are working on a measure space  $(X, \mathcal{B}, \mu)$ , then  $\mathcal{B}$  stands for the set of all Borel sets.
- (7) The symbol  $\sharp A$  means the cardinality of the set  $A$ .
- (8) A family of sets  $\{X_\lambda\}_{\lambda \in \Lambda}$  is said to be *almost disjoint*, if there exists a constant  $c > 0$  depending only on the underlying space  $X$  so that

$$(0.3) \quad \sum_{\lambda \in \Lambda} X_\lambda \leq c.$$

- (9) The symbol  $2^X$  denotes the set of all subsets in  $X$ .
- (10) Let  $X$  be a topological space. Then  $\mathcal{K}_X$  is the set of all compact subsets of  $X$ , and  $\mathcal{O}_X$  is the set of all open subsets of  $X$ .
- (11) The set  $\mathcal{I}(\mathbb{R})$  denotes the set of all closed intervals in  $\mathbb{R}$ .

*Notation* (Numbers).

- (1) Let  $a \in \mathbb{R}$ . Then write  $a_+ := \max(a, 0)$  and  $a_- := \min(a, 0)$ . Correspondingly, given an  $\mathbb{R}$ -valued function  $f$ ,  $f_+$  and  $f_-$  are function given by  $f_+(x) := \max(f(x), 0)$  and  $f_-(x) := \min(f(x), 0)$ , respectively.
- (2) Let  $a, b \in \mathbb{R}$ . Then denote  $a \vee b = \max(a, b)$  and  $a \wedge b = \min(a, b)$ . Correspondingly, given  $\mathbb{R}$ -valued functions  $f, g$ ,  $f \vee g$  and  $f \wedge g$  are functions given by  $f \vee g(x) = \max(f(x), g(x))$  and  $f \wedge g(x) = \min(f(x), g(x))$ , respectively.
- (3) Let  $A, B \geq 0$ . Then  $A \lesssim B$  means that there exists a constant  $C > 0$  such that  $A < CB$ , where  $C$  depends only on the parameters of importance.
- (4) We define

$$(0.4) \quad \mathbb{N} := \{1, 2, \dots\}, \mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}, \mathbb{N}_0 := \{0, 1, \dots\}.$$

- (5) We denote by  $\mathbb{K}$  either  $\mathbb{R}$  or  $\mathbb{C}$ , the coefficient field under consideration.

*Notation* (Function spaces).

- (1) Let  $X$  be a Banach space. We denote its norm by  $\|\cdot\|_X$ . However, we sometimes denote the  $L^p(\mu)$ -norm of functions by  $\|\cdot\|_p$ .
- (2) Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . Then  $C_c^\infty(\Omega)$  denotes the set of smooth function with compact support in  $\Omega$ .
- (3) Let  $1 \leq j \leq d$ . The symbol  $x_j$  denotes not only the  $j$ -th coordinate but also the function  $x = (x_1, \dots, x_d) \mapsto x_j$ .

(4) Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of measurable functions. Then we write

$$(0.5) \quad \|f_j\|_{L^p(\ell^q)} := \left( \int_{\mathbb{R}^d} \left( \sum_{j=1}^{\infty} |f_j(x)|^q \right)^{\frac{p}{q}} dx \right)^{\frac{1}{p}}, \quad 0 < p, q \leq \infty.$$

(5) We denote uninteresting positive constants by  $c$ . Even if they are different we denote them by the same letter. Therefore the inequality  $2c \leq c$  makes sense.

(6) The space  $L^2(\mathbb{R}^d)$  is the Hilbert space of square integrable functions on  $\mathbb{R}^d$  whose inner product is given by

$$(0.6) \quad \langle f, g \rangle = \int_{\mathbb{R}^d} f(x) \cdot \bar{g}(x) dx.$$

(7) In view of (0.6), the inner product of  $L^2(\mathbb{R}^d)$ , it seems appropriate that we define the embedding  $L^1_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$  by:

$$(0.7) \quad f \in L^1_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d) \mapsto F_f := \left[ g \in \mathcal{S}(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} g(x) \overline{f(x)} dx \right].$$

However, in order that  $f \mapsto F_f$  be linear, we shall define it later by:

$$(0.8) \quad f \in L^1_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d) \mapsto F_f := \left[ g \in \mathcal{S}(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} g(x) f(x) dx \right].$$

(8) Let  $E$  be a measurable set and  $f$  be a measurable function with respect to the Lebesgue measure. Then denote  $m_E(f) := \frac{1}{|E|} \int_E f$ .

(9) Let  $0 < \eta < \infty$ ,  $E$  be a measurable set, and  $f$  be a positive measurable function with respect to the Lebesgue measure. Then denote  $m_E^{(\eta)}(f) := m_E(f^\eta)^{\frac{1}{\eta}}$ .

(10) Let  $0 < \eta < \infty$ . We define the powered Hardy-Littlewood maximal operator  $M^{(\eta)}$  by

$$(0.9) \quad M^{(\eta)} f(x) := \sup_{R>0} \left( \frac{1}{|B(x, R)|} \int_{B(x, R)} |f(y)|^\eta dy \right)^{\frac{1}{\eta}}.$$

(11) For  $x \in \mathbb{R}^d$ , we define  $\langle x \rangle := \sqrt{1 + |x|^2}$ .

(12) The space  $C$  denotes the set of all continuous functions on  $\mathbb{R}^d$ .

(13) The space  $\text{BC}(\mathbb{R}^d)$  denotes the set of all bounded continuous functions on  $\mathbb{R}^d$ .

(14) The space  $\text{BUC}(\mathbb{R}^d)$  denotes the set of all bounded uniformly continuous functions on  $\mathbb{R}^d$ .

(15) Occasionally we identify the value of functions with functions. For example  $\sin x$  denotes the function on  $\mathbb{R}$  defined by  $x \mapsto \sin x$ .

(16) Given a Banach space  $X$ , we denote by  $X^*$  its dual space. The set  $X_1$  is the closed unit ball in  $X$ .

(17) Let  $\mu$  be a measure on a measure space  $(X, \mathcal{B}, \mu)$ . Given a  $\mu$ -measurable set  $A$  with positive  $\mu$ -measure and a function  $f$ , we denote  $m_Q(f) = \frac{1}{\mu(A)} \int_A f(x) d\mu(x)$ . Let  $0 < \eta < \infty$ . Then define  $m_Q^{(\eta)}(f) = m_Q(f^\eta)^{\frac{1}{\eta}}$  whenever  $f$  is positive.

(18) For  $x \in \mathbb{R}^d$ , we define  $\mathcal{Q}_x$  to be the set of all cubes containing  $x$ . Given a measurable function,  $Mf$  denotes the uncentered Hardy-Littlewood maximal operator and  $M'f$  denotes the centered Hardy-Littlewood maximal operator.

$$Mf(x) := \sup_{Q \in \mathcal{Q}_x} m_Q(|f|),$$

$$M'f(x) := \sup_{r>0} m_{Q(x, r)}(|f|).$$

(19) If notational confusion seems likely, then we use  $[ \ ]$  to write  $Mf(x) = M[f](x)$ ,  $\mathcal{F}\varphi(\xi) = \mathcal{F}[\varphi](\xi)$ , etc.

Finally, we admit the choice of axiom. We remark that it is equivalent to admitting that  $\prod_{\lambda \in \Lambda} K_\lambda$  is compact whenever we are given a collection of compact sets  $\{K_\lambda\}_{\lambda \in \Lambda}$ .

## Part 1. A bird's-eye-view of this book

What is harmonic analysis ? Roughly speaking:

- (1) To study the property of the Fourier series and the Fourier transform.
- (2) To study the property of integral transforms in general.
- (3) To study the behaviour of functions.

In this chapter, we shall make a quick overview of the main ideas used throughout this book. As a starting material, we take up the Dirichlet problem on the unit disk  $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$  in the complex plane  $\mathbb{C}$  and shall start by introducing about what a harmonic analytic argument looks like.

### 1. INTRODUCTION

Although this section is meant for those who have learnt about complex analysis, the Lebesgue integral and functional analysis, it should not be too difficult for those we are acquainted with these fields. The results in this section will not be used later: we wish only to describe the key ideas of this book. From Chapter 3 we get into the main theory of this field, but Chapter 1 will help us to get into the remaining parts.

#### 1.1. Maximal operator on $\partial\mathbb{D}$ .

First let us recall the Dirichlet problem on the complex plane  $\mathbb{C}$ :

$$(1.1) \quad \begin{cases} -\Delta u = 0 & \text{on } \mathbb{D}, \\ u|_{\partial\mathbb{D}} = f & \text{on } \partial\mathbb{D}, \end{cases}$$

where  $f$  is a given function and  $\partial\mathbb{D}$  is the boundary of the unit disk on the complex plane  $\mathbb{C}$ , that is,  $\partial\mathbb{D} := \{z \in \mathbb{C} : |z| = 1\}$ . The unknown is the function  $u : \mathbb{D} \rightarrow \mathbb{C}$ . Of course, it is ideal that  $-\Delta u$  exists in the sense of the usual partial derivative. For example, in engineering, it is not preferable that we consider the weak solution. This is why we prefer to postulate that  $u \in C^2(\mathbb{D})$ . From this point of view, it is natural that  $f \in C(\partial\mathbb{D})$ . Thus, a natural conclusion for (1.1) is the following:

**Theorem 1.1.** *Given  $f \in C(\partial\mathbb{D})$ , we can find a unique solution  $u \in C^2(\mathbb{D}) \cap C(\overline{\mathbb{D}})$  of (1.1).*

*Outline of the proof.* To prove the uniqueness, we use, for example, the maximum principle. For details, we refer to [2, 50]; we shall not go into the details because it is not of much importance in this chapter.

To prove the existence, it turns out that we can write the solution out in full:

$$(1.2) \quad u(z) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta, \quad z \in \mathbb{D}.$$

Observe that the function  $u$  inherits harmonicity from the kernel  $z \in \mathbb{C} \mapsto \frac{1 - |z|^2}{|z - e^{i\theta}|^2} \in \mathbb{R}$ .

Up to this point, we have made a very quick review of the proof of uniqueness and harmonicity of the constructed solution (1.2). The interested reader may find it worthwhile to fill in the details by themselves.



However, they are not what we want to shed light on. Indeed, we are left with the task of showing the boundary condition, for which give a detailed proof. It suffices to estimate  $|u(z) - f(e^{i\theta_0})|$  for fixed  $\theta$ :

$$\begin{aligned} & \text{Let } 0 < \eta < \frac{\pi}{6} \text{ be fixed. Since } f(e^{i\theta_0}) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta, \text{ it follows that:} \\ |u(z) - f(e^{i\theta_0})| & \leq \frac{1}{2\pi} \int_0^{2\pi} |f(e^{i\theta}) - f(e^{i\theta_0})| \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta \\ & = \frac{1}{2\pi} \int_{\theta_0 - \eta}^{\theta_0 + \eta} |f(e^{i\theta}) - f(e^{i\theta_0})| \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta + \frac{1}{2\pi} \int_{\theta_0 - \pi}^{\theta_0 - \eta} + \int_{\theta_0 + \eta}^{\theta_0 + \pi} |f(e^{i\theta}) - f(e^{i\theta_0})| \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta \\ & \leq \frac{1}{2\pi} \left( \sup_{\theta \in [\theta_0 - \eta, \theta_0 + \eta]} |f(e^{i\theta}) - f(e^{i\theta_0})| \right) \int_{\theta_0 - \eta}^{\theta_0 + \eta} \frac{1 - |z|^2}{|z - e^{i\theta}|^2} d\theta + \frac{4(1 - |z|)}{|1 - e^{i\eta}|^2} \sup_{w \in \partial\mathbb{D}} |f(w)| \\ & \leq \sup_{\theta \in [\theta_0 - \eta, \theta_0 + \eta]} |f(e^{i\theta}) - f(e^{i\theta_0})| + \frac{4(1 - |z|)}{|1 - e^{i\eta}|^2} \sup_{w \in \partial\mathbb{D}} |f(w)|. \end{aligned}$$

Letting  $z \rightarrow e^{i\theta_0}$ , we obtain

$$(1.3) \quad \limsup_{z \in \mathbb{D} \rightarrow e^{i\theta_0}} |u(z) - f(e^{i\theta_0})| \leq \sup_{\theta \in [\theta_0 - \eta, \theta_0 + \eta]} |f(e^{i\theta}) - f(e^{i\theta_0})|.$$

Since  $\partial\mathbb{D}$  is compact,  $f$  is uniformly continuous. Therefore, letting  $\eta \downarrow 0$ , we obtain

$$(1.4) \quad \lim_{z \in \mathbb{D} \rightarrow e^{i\theta_0}} |u(z) - f(e^{i\theta_0})| = 0.$$

□

*Exercise 1.* Show that the function  $u$  given by (1.2) is harmonic.

We are now happy because (1.1) was solved completely. However, after learning partial differential equations, we come to feel that perhaps the function space  $C(\partial\mathbb{D})$  is not so good. Instead,  $L^2(\partial\mathbb{D})$  seems to be a nice candidate for the space of initial data. More advanced learners may feel it is still good to replace  $C(\partial\mathbb{D})$  with the Hölder class; the Hölder class is not taken up in this part, but it will appear later in this book. Furthermore in view of (1.2) we are tempted to make an excursion to other function spaces, say  $L^1(\partial\mathbb{D})$ ,  $L^2(\partial\mathbb{D})$  and  $L^\infty(\partial\mathbb{D})$ . Let  $1 \leq p < \infty$ . We say that  $f \in L^p(\partial\mathbb{D})$  if

$$\|f\|_{L^p(\partial\mathbb{D})} = \left( \int_0^{2\pi} |f(e^{i\theta})|^p d\theta \right)^{1/p} < \infty.$$

When  $p = \infty$ , just let us say that  $L^p(\partial\mathbb{D})$  is the set of all essentially bounded functions. The problem is that in these function spaces, we are considering some equivalence relation, that is, we disregard the difference for the set of Lebesgue measure 0, so the boundary problem (1.1) does not make sense. However, it is still possible for (1.1) to make sense by stopping identifying the functions as above.

A skillful use of maximal operators, defined just below (see (1.7)), will lead us to the following theorem:

**Theorem 1.2.** *Suppose that  $f \in L^1(\partial\mathbb{D})$  and that  $u : \mathbb{D} \rightarrow \mathbb{C}$  is defined for  $t \in \partial\mathbb{D}$  by (1.1). Then we have*

$$(1.5) \quad \lim_{r \uparrow 1} u(rt) = f(t) \text{ for a.e. } t \in \partial\mathbb{D}.$$

Before we come to the proof of this theorem, let us make some preparatory observations. First, letting  $t = e^{i\theta}$ , we write  $u(re^{i\theta_0})$  out in full:

$$(1.6) \quad u(re^{i\theta_0}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{1-r^2}{|1-re^{i(\theta-\theta_0)}|^2} f(e^{i\theta}) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1-r^2}{|1-re^{i\theta}|^2} f(e^{i(\theta+\theta_0)}) d\theta.$$

If we just look at it, there is no way to proceed further. However, the Hardy-Littlewood maximal operator paves a way: For a measurable function  $f$ , let us define the Hardy-Littlewood maximal function  $Mf$ , which is a function of  $t \in \partial\mathbb{D}$ , by

$$(1.7) \quad Mf(t) := \sup \left\{ \frac{1}{|I|} \int_I |f(y)| dy : t \in I \subset \partial\mathbb{D} \right\}$$

where  $I$  runs over all closed intervals in  $\partial\mathbb{D}$  containing  $t$ . The mappint  $f \mapsto Mf$  is referred to as the Hardy-Littlewood maximal operator. Then, from (1.6), we have

$$(1.8) \quad |u(re^{i\theta_0})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i(\theta+\theta_0)})| \frac{1-r^2}{|1-re^{i\theta}|^2} d\theta.$$

An elementary calculation shows  $|1-rt|^2 = 1+r^2-2\cos\theta$ , so if we let  $P_r(\theta) := \frac{1-r^2}{2\pi|1-re^{i\theta}|^2}$ , then an approximation procedure gives us a sequence of functions  $\{P_{r,N}\}_{N \in \mathbb{N}}$  of the form

$$(1.9) \quad P_{r,N} = \sum_{j=1}^{2^N} a_{j,N} \chi_{[-j2^{-N}, j2^{-N}]}$$

with  $a_{j,N} \geq 0$  and  $P_{r,N}(\theta) \uparrow P_r(\theta)$ . Inserting (1.9) to (1.8), we obtain

$$(1.10) \quad |u(re^{i\theta_0})| \leq \lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} |f(e^{i(\theta+\theta_0)})| P_{r,N}(\theta) d\theta \leq Mf(t).$$

Let us summarize (1.10) as a lemma. Here and below given a function  $F$  and a measurable set  $E$ , let us write  $\{t \in \partial\mathbb{D} : F > \lambda\} = \{F > \lambda\}$ ,  $\int_E F(t) dt = \int_E F$  for simplicity. From (1.10) we obtain the following pointwise estimate.

**Lemma 1.3.** *Let  $f \in L^1_{\text{loc}}(\partial\mathbb{D})$  and define:*

$$(1.11) \quad Mf(t) := \sup \left\{ \frac{1}{|I|} \int_I |f(s)| ds : t \in I \subset \partial\mathbb{D} \right\} \quad t \in \partial\mathbb{D},$$

where  $I$  runs over all closed intervals in  $\partial\mathbb{D}$  containing  $t$ . Then;

$$(1.12) \quad |u(rt)| \leq Mf(t) \quad (r \in (0,1), t \in \partial\mathbb{D}).$$

Now we need information of the operator  $M$ . To see this, since  $M$  is defined by using intervals on  $\partial\mathbb{D}$ , it will be helpful to summarize a key property of intervals on  $\partial\mathbb{D}$ .

**Lemma 1.4** (A special geometric structure of intervals in  $\partial\mathbb{D}$ ). *Suppose that  $\{I_\lambda\}_{\lambda \in \Lambda}$  is a finite family of intervals in  $\partial\mathbb{D}$ . Then we can select  $\Lambda_0 \subset \Lambda$  so that:*

$$(1.13) \quad \chi_{\bigcup_{\lambda \in \Lambda} I_\lambda} \leq \sum_{\lambda \in \Lambda_0} \chi_{I_\lambda} \leq 2.$$

Let us clarify what the conclusion (1.13) says. First, the left inequality says  $\{I_\lambda\}_{\lambda \in \Lambda_0}$  and  $\{I_\lambda\}_{\lambda \in \Lambda}$  covers exactly the same set. Next, the right inequality says the overlapping at each point never exceeds 2. With this observation in mind, let us prove Lemma 1.4.

*Proof.* Suppose that we are given three intervals  $I_1, I_2, I_3$  meeting in a point. Then one of them is not necessary because the remaining two intervals cover it. The proof can be obtained and is easy to understand, once we notice this geometric observation. If there is a superfluous interval, we have only to throw it away. By induction on the number of intervals we can prove the proposition.  $\square$

Now we prove the following boundedness property of  $M$ , which is called the weak-(1,1) boundedness of  $M$ .

**Theorem 1.5.** *For all  $f \in L^1(\partial\mathbb{D})$  we have*

$$(1.14) \quad |\{Mf > \lambda\}| \leq \frac{2}{\lambda} \int_{\partial\mathbb{D}} |f(t)| dt.$$

*Proof.* By the inner regularity of the Lebesgue measure, we have only to prove

$$(1.15) \quad |K| \leq \frac{2}{\lambda} \int_{\partial\mathbb{D}} |f(t)| dt$$

for any compact set  $K$  contained in  $\{Mf > \lambda\}$ .

From the definition of  $M$ , if  $t \in \partial\mathbb{D}$  satisfies  $Mf(t) > \lambda$ , then there exists an interval  $I$  containing  $t$  such that  $\frac{1}{|I|} \int_I |f(s)| ds > \lambda$ . By compactness of  $K$ , together with Lemma 1.4, we can find  $I_1, I_2, \dots, I_k$  so that

$$\chi_K \leq \sum_{j=1}^k \chi_{I_j} \leq 2, \quad \frac{1}{|I_j|} \int_{I_j} |f(s)| ds > \lambda, \quad j = 1, 2, \dots, k.$$

Thus, we have

$$(1.16) \quad |K| \leq \sum_{j=1}^k |I_j| \leq \frac{1}{\lambda} \int_{I_j} |f(s)| ds \leq \frac{2}{\lambda} \int_{\partial\mathbb{D}} |f(s)| ds.$$

This is the inequality we wish to prove.  $\square$

With (1.14) in mind, let us prove the following theorem concerning with the boundary value of (1.1).

**Theorem 1.6.** *Let  $f \in L^1(\partial\mathbb{D})$ . Then we have:*

$$(1.17) \quad \lim_{r \uparrow 1} u(r t) = f(t) \quad \text{a.e. } t \in \partial\mathbb{D}.$$

*Proof.* , Observe that what we have to prove can be rephrased as follows: The set

$$(1.18) \quad E_\varepsilon := \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(r t) - f(t)| > \varepsilon \right\}$$

has measure 0 in terms of  $\limsup$ . However, since  $(0, \infty) = \bigcup_{j=1}^{\infty} (j^{-1}, \infty)$ , it suffices to establish that:

$$\left| \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(r t) - f(t)| > \varepsilon \right\} \right| = 0$$

for  $\varepsilon > 0$  by virtue of the monotonicity of the measures. However, as is easily verified, (1.12) is true if  $f$  is continuous. As a consequence:

$$\left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(rt) - f(t)| > \varepsilon \right\} = \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |(u-v)(rt) - (f-g)(t)| > \varepsilon \right\},$$

where  $v$  is a solution corresponding to the initial value of  $g \in C(\partial\mathbb{D})$ . If we invoke the Hardy-Littlewood maximal operator control for  $f - g$  instead of  $f$ , we obtain

$$(1.19) \quad |u(rt) - v(rt)| \leq M[f - g](t).$$

Therefore, it follows that

$$(1.20) \quad \left| \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(rt) - f(t)| > \varepsilon \right\} \right| \leq |\{ M[f - g] + |f - g| > \varepsilon \}|.$$

The set in the right-hand side is contained in the union of  $\{ M[f - g] > \frac{\varepsilon}{2} \}$  and  $\{ |f - g| > \frac{\varepsilon}{2} \}$ . One convenient way to check this is to observe that

$$(1.21) \quad M[f - g](t) + |f(t) - g(t)| \leq \varepsilon,$$

whenever  $M[f - g](t) \leq \frac{\varepsilon}{2}$  and  $|f(t) - g(t)| \leq \frac{\varepsilon}{2}$ . (Note that (1.21) follows immediately from the definition of sup and the triangle inequality.)

Now we note that

$$(1.22) \quad \left| \left\{ |f - g| > \frac{\varepsilon}{2} \right\} \right| = \int_{\{2|f-g|>\varepsilon\}} 1 \, ds \leq \frac{2}{\varepsilon} \|f - g\|_1,$$

because  $1 \leq \frac{2}{\varepsilon} |f - g|$  on the set  $\{2|f - g| > \varepsilon\}$ . By virtue of the Hardy-Littlewood maximal inequality (1.14) established above, we obtain

$$(1.23) \quad \left| \left\{ M[f - g] > \frac{\varepsilon}{2} \right\} \right| \leq \frac{4}{\varepsilon} \|f - g\|_1.$$

Putting (1.20)–(1.23) together, we conclude

$$\left| \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(rt) - f(t)| > \varepsilon \right\} \right| \leq \frac{6}{\varepsilon} \|f - g\|_1$$

for all continuous functions  $g$ . Since  $g \in C(\partial\mathbb{D})$  is arbitrary and the set of all continuous functions forms a dense subset in  $L^1(\partial\mathbb{D})$ , it follows that:

$$(1.24) \quad \left| \left\{ t \in \partial\mathbb{D} : \limsup_{r \uparrow 1} |u(rt) - f(t)| > \varepsilon \right\} \right| = 0.$$

This is the result we wish to prove. □

Before we conclude this section let us provide some supplemental information on the Hardy-Littlewood maximal operator  $M$ .

**Theorem 1.7.** *Let  $1 < p \leq \infty$ . Then we have*

$$(1.25) \quad \|Mf\|_p \leq \left( \frac{4p}{p-1} \right)^{\frac{1}{p}} \|f\|_p$$

for all  $f \in L^p(\partial\mathbb{D})$ .

*Proof.* First, let us fix  $t \in \partial\mathbb{D}$ . Then we have

$$(1.26) \quad Mf(t)^p = \int_0^{Mf(t)} p \lambda^{p-1} d\lambda$$

thanks to the fundamental theorem of calculus. Seemingly, (1.26) yields nothing, but with the help of Fubini's theorem, we are led to an expression where the level set  $\{Mf > \lambda\}$  appears:

$$(1.27) \quad \|Mf\|_p^p = \int_{\partial\mathbb{D}} \left( \int_0^{Mf(t)} p \lambda^{p-1} d\lambda \right) dt = \int_{\partial\mathbb{D}} \left( \int_0^\infty p \lambda^{p-1} \chi_{\{Mf > \lambda\}}(t, \lambda) d\lambda \right) dt.$$

If we change the order of integration by using the Fubini theorem and we factor out the function independent of  $t$ , then the distribution function, which we have considered in Theorem 1.5, appears.

$$(1.28) \quad \int_{\partial\mathbb{D}} \left( \int_0^\infty p \lambda^{p-1} \chi_{\{Mf > \lambda\}}(t, \lambda) d\lambda \right) dt = \int_0^\infty p \lambda^{p-1} |\{Mf > \lambda\}| d\lambda.$$

Recall that  $|\{Mf > \lambda\}|$  is exactly the right-hand side of Theorem 1.5, which is bounded by  $\frac{2}{\lambda} \|f\|_1$ . However, it turns out that direct usage of this inequality directly will not work. The crux of the proof is to combine Theorem 1.5 and the  $L^\infty(\partial\mathbb{D})$ -boundedness of  $M$ , which we have not been alluding to, for each  $\lambda$ . To be precise, observe that

$$\{Mf > \lambda\} = \{M[\chi_{\{|f| > \lambda/2\}}f + \chi_{\{|f| \leq \lambda/2\}}f] > \lambda\} \subset \{M[\chi_{\{|f| > \lambda/2\}}f] + M[\chi_{\{|f| \leq \lambda/2\}}f] > \lambda\},$$

by virtue of the subadditive inequality

$$(1.29) \quad M[F + G] \leq MF + MG.$$

Since the maximal operator  $M$  deals with the average of the functions,

$$M[\chi_{\{|f| \leq \lambda/2\}}f] \leq \frac{\lambda}{2}$$

holds. The estimate is called the  $L^\infty(\partial\mathbb{D})$ -boundedness. Thus, it follows that

$$(1.30) \quad \{Mf > \lambda\} \subset \{M[\chi_{\{|f| > \lambda/2\}}f] > \lambda/2\}.$$

Now we invoke the Hardy-Littlewood maximal inequality (1.14) obtained above. Then we obtain the key estimate which is valid for our present situation.

$$(1.31) \quad |\{Mf > \lambda\}| \leq \frac{4}{\lambda} \|\chi_{\{|f| > \lambda/2\}}f\|_1.$$

If we insert this key inequality, then it follows that

$$(1.32) \quad \|Mf\|_p^p \leq 4p \int_0^\infty \lambda^{p-2} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| > \lambda/2\}}(\lambda, t) |f(t)| dt \right) d\lambda.$$

Changing the order of integrations once more, we are led to

$$(1.33) \quad 4p \int_0^\infty \lambda^{p-2} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| > \lambda/2\}}(\lambda, t) |f(t)| dt \right) d\lambda = 4p \int_{\partial\mathbb{D}} \left( |f(t)| \int_0^{|f(t)|} \lambda^{p-2} d\lambda \right) dt.$$

Since we are assuming  $p > 1$ , the integral

$$\int_0^{|f(t)|} \lambda^{p-2} d\lambda = \frac{1}{p-1} |f(t)|^{p-1}.$$

is finite for a.e.  $t \in \partial\mathbb{D}$ .

Thus, inserting (1.33) to (1.32), we finally see

$$(1.34) \quad \|Mf\|_p^p \leq \frac{4p}{p-1} \|f\|_p^p.$$

□

*Exercise 2.* Prove (1.29).

## 1.2. Conjugate functions on $\partial\mathbb{D}$ .

In this section, we take up the Hilbert transform on  $\partial\mathbb{D}$ , which is concerned indirectly with the Fourier series. In the previous subsection, we dealt with harmonic functions with boundary value given beforehand: now, we shall consider their harmonic conjugate.

**Definition 1.8** (Harmonic conjugate). Let  $\Omega$  be a domain in  $\mathbb{C}$ . Then the harmonic conjugate of the harmonic function  $u : \Omega \rightarrow \mathbb{C}$  is a harmonic function  $v : \Omega \rightarrow \mathbb{C}$  such that  $u + iv$  is holomorphic.

Since  $\mathbb{D}$  is simply connected, we see that, for any  $u$  which is harmonic on  $\mathbb{D}$ , its harmonic conjugate always exists. For details we refer to [2, 50]. It is not so hard to see that two different harmonic conjugates differ by an additive constant, if we use the maximal principle and so on. In what follows, we say that the conjugate function  $v$  is *canonical* if  $v(0) = 0$ .

In this section, starting from a boundary function  $f : \partial\mathbb{D} \rightarrow \mathbb{C}$ , we consider its harmonic extension  $u$  given by (1.1) and then the harmonic conjugate  $v$  of  $u$ . We are now going to show

**Theorem 1.9.** *The boundary value of  $v$  exists almost everywhere and belongs to  $L^p(\partial\mathbb{D})$ , provided  $f \in L^p(\partial\mathbb{D})$ .*

$L^2(\partial\mathbb{D})$ -boundedness by way of Fourier series. The case when  $p = 2$  is considerably easy and transparent. We begin with starting a canonical value.

**Proposition 1.10.** *The mapping  $f \mapsto \tilde{f}$ , taking  $f$  to the boundary value of the canonical conjugate function, is  $L^2(\partial\mathbb{D})$ -bounded.*

*Proof.* We shall calculate  $\tilde{E}_k$ , where  $E_k(e^{i\theta}) = e^{ik\theta}$  by expanding the kernel. The result is:

$$\frac{2 \sin \theta}{1 - 2r \cos \theta + r^2} = \frac{z - \bar{z}}{i(1 - z)(1 - \bar{z})} = \frac{z - \bar{z}}{i} \sum_{j,k=0}^{\infty} z^j \bar{z}^k = \frac{1}{i} \left( \sum_{j,k=0}^{\infty} z^{j+1} \bar{z}^k - \sum_{j,k=0}^{\infty} z^j \bar{z}^{k+1} \right).$$

Here, we used  $|z| < 1$  to obtain the absolute convergence of the series. We write  $z = r e^{i\theta}$ . Then:

$$(1.35) \quad \frac{2 \sin \theta}{1 - 2r \cos \theta + r^2} = i \left( \sum_{j=1}^{\infty} z^j - \sum_{j=1}^{\infty} \bar{z}^j \right).$$

Since  $\int_0^{2\pi} e^{ij\rho} \cdot e^{ik(\theta-\rho)} d\rho = 2\pi \delta_{jk} \cdot e^{ik\theta}$ , it follows that

$$(1.36) \quad \tilde{E}_k = \begin{cases} E_k, & k \geq 1, \\ -E_k, & k \leq -1. \end{cases}$$

Let us denote

$$\text{sign}(k) := \begin{cases} 1, & k \geq 1, \\ -1, & k \leq -1, \\ 0 & k = 0. \end{cases}$$

Since any  $f \in L^2(\partial\mathbb{D})$  can be expressed as  $f = \sum_{k=-\infty}^{\infty} a_k E_k$  with  $\|\{a_k\}_{k=-\infty}^{\infty}\|_2 = \frac{1}{2\pi}\|f\|_2$ , we conclude that:

$$(1.37) \quad f = \sum_{k=-\infty}^{\infty} a_k E_k \in L^2(\partial\mathbb{D}) \mapsto \tilde{f} = \sum_{k=-\infty}^{\infty} \text{sign}(k) a_k E_k \in L^2(\partial\mathbb{D})$$

is bounded. □

**1.3. Alternate version of  $L^1(\partial\mathbb{D})$ -boundedness and Calderón-Zygmund operators.** In this section, we shall prove a kind of  $L^1(\partial\mathbb{D})$ -boundedness. In order to explain what we mean, consider the following:

**Definition 1.11** (Dyadic interval on  $\partial\mathbb{D}$ ).

- (1) Let  $l \in \mathbb{N}_0$ . Let the integer  $k$  satisfy  $0 \leq k < 2^l$ . Define  $I_{k,l}$  as the arc connecting  $p_{k,l}$  and  $p_{k,l+1}$ , where  $p_{k,l} = \exp(2\pi i k 2^{-l}) \in \partial\mathbb{D}$ . Define  $\mathcal{D}_l := \{I_{k,l} : 0 \leq k < 2^l\}$ .
- (2) Let  $f \in L^1(\partial\mathbb{D})$  and  $l \in \mathbb{N}_0$ . Then define

$$(1.38) \quad E_l f := E_l[f] := \sum_{k=1}^{2^l} m_{I_{k,l}}(f) \cdot \chi_{I_{k,l}}, \quad l \in \mathbb{Z}^d, \quad M_{\text{dyadic}} f(t) := \sup_{l \in \mathbb{N}} E_l[|f|](t).$$

**Lemma 1.12.** *Suppose that  $\lambda > 0$  satisfies*

$$(1.39) \quad \lambda > \frac{1}{|\partial\mathbb{D}|} \int_{\partial\mathbb{D}} |f(t)| dt.$$

*Then there exist disjoint dyadic intervals  $I_1, I_2, \dots, I_k, \dots$  such that the following are satisfied:*

$$(1.40) \quad \{M_{\text{dyadic}} f > \lambda\} = \bigcup_{j=1}^{\infty} I_j,$$

$$(1.41) \quad \lambda \leq \frac{1}{|I_k|} \int_{|I_k|} M_{\text{dyadic}} f(t) ds \leq 2\lambda.$$

*Proof.* By definition of the dyadic maximal operator, if  $M_{\text{dyadic}} f(t) > \lambda$ , we can choose a dyadic interval  $I_t$  such that

$$\frac{1}{|I_t|} \int_{I_t} |f(s)| ds > \lambda.$$

If  $I_t$  is not maximal, that is, there exists another dyadic interval  $J$  strictly larger than  $I_t$  such that  $\frac{1}{|J|} \int_J |f(s)| ds > \lambda$ , then replace  $I_t$  with  $J$ . Thus, if we replace  $I_t$  with a larger one, then we may assume

$$\frac{1}{|I_t|} \int_{I_t} |f(s)| ds \leq 2\lambda;$$

as a result, (1.41) is satisfied. A geometric observation shows (1.40). □

We decompose the functions into the good part and the bad part.

**Definition 1.13** (Good part and bad part of functions). Keeping the same notation as in Lemma 1.12, define:

$$\begin{aligned} b_k(t) &:= \left( f(t) - \frac{1}{|I_k|} \int_{|I_k|} f(s) ds \right) \chi_{I_k}(t) \quad (t \in \partial\mathbb{D}), \\ b(t) &:= \sum_k b_k(t) \quad (t \in \partial\mathbb{D}), \\ g(t) &:= f(t) - b(t) \quad (t \in \partial\mathbb{D}) \end{aligned}$$

for  $k \in \mathbb{N}$ .

**Lemma 1.14.** *We have  $|f(t)| \leq M_{\text{dyadic}} f(t)$  for almost every  $t \in \partial\mathbb{D}$ .*

*Proof.* Assume that  $f$  is continuous. It is elementary to prove that:

$$(1.42) \quad f(t) = \lim_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \frac{1}{|I|} \int_I |f(s)| ds \quad \text{for all } t \in \partial\mathbb{D},$$

where  $\lim_{\substack{I \in \mathcal{D} \\ I \downarrow t}}$  is a symbolical notation meaning that, for every  $\varepsilon > 0$ , there exists a dyadic interval

$I_0$  containing  $t$  such that  $\left| \frac{1}{|I|} \int_I f(s) ds - f(t) \right| < \varepsilon$ , whenever  $I \subset I_0$  is a dyadic interval containing  $t$ .

Let us pass to the general case. Fix  $f \in L^1(\partial\mathbb{D})$ . Then, our strategy is to prove that:

$$(1.43) \quad E_{\varepsilon, f} := \left\{ t \in \partial\mathbb{D} : \limsup_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \left| \frac{1}{|I|} \int_I f(s) ds - f(t) \right| > \varepsilon \right\}$$

has zero Lebesgue measure for all  $\varepsilon > 0$ . Choose  $g \in C(\partial\mathbb{D})$  arbitrarily. As we have seen above,  $g$  satisfies

$$(1.44) \quad g(t) = \lim_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \frac{1}{|I|} \int_I g(s) ds \quad \text{for all } t \in \partial\mathbb{D}.$$

Therefore, it follows that

$$(1.45) \quad E_{\varepsilon, f} = E_{\varepsilon, f-g}.$$

Next, we claim that

$$(1.46) \quad E_{\varepsilon, f-g} \subset \left\{ t \in \partial\mathbb{D} : \limsup_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \left| \frac{1}{|I|} \int_I (f-g)(s) ds \right| > \frac{\varepsilon}{2} \right\} \cup \left\{ |f-g| > \frac{\varepsilon}{2} \right\}$$

To see this, suppose that  $t \in \partial\mathbb{D}$  does not belong to the right-hand side; i.e., assume that  $t$  satisfies

$$\limsup_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \left| \frac{1}{|I|} \int_I (f-g)(s) ds \right| \leq \frac{\varepsilon}{2}$$

and that

$$|f(t) - g(t)| \leq \frac{\varepsilon}{2}.$$

Then we have

$$\limsup_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \left| \frac{1}{|I|} \int_I (f(u) - g(u)) du - (f(t) - g(t)) \right| \leq \varepsilon,$$

establishing (1.46).



Putting together (1.45) and (1.46), we obtain

$$|E_{\varepsilon, f}| \leq \left| \left\{ t \in \partial\mathbb{D} : \limsup_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \left| \frac{1}{|I|} \int_I (f - g)(s) ds \right| > \frac{\varepsilon}{2} \right\} \right| + \left| \left\{ |f - g| > \frac{\varepsilon}{2} \right\} \right|.$$

We estimate the first term on the right-hand side by the Hardy-Littlewood maximal operator  $M$  and use the weak-(1, 1) boundedness:

$$|E_{\varepsilon, f}| \leq \left| \left\{ M[f - g] > \frac{\varepsilon}{2} \right\} \right| + \left| \left\{ |f - g| > \frac{\varepsilon}{2} \right\} \right| \leq \frac{6}{\varepsilon} \|f - g\|_1.$$

Since  $g \in C(\partial\mathbb{D})$  is arbitrary, we see that  $|E_{\varepsilon, f}| = 0$ . From this, we deduce that:

$$(1.47) \quad f(t) = \lim_{\substack{I \in \mathcal{D} \\ I \downarrow t}} \frac{1}{|I|} \int_I f(s) ds \text{ for almost all } t \in \partial\mathbb{D}.$$

Proceeding in the same way, we obtain that  $|f(t)| \leq M_{\text{dyadic}} f(t)$  if  $t \in \partial\mathbb{D}$  satisfies (1.47).  $\square$

Now that Lemma 1.14 has been proved, the following is immediate:

**Corollary 1.15.** *Using the notation for Lemma 1.14, one has  $|g(t)| \leq 2\lambda$  for almost every  $t \in \partial\mathbb{D}$ .*

*Proof.* If  $t \in Q_j$  for some  $j$ , then

$$(1.48) \quad g(t) = \left| \frac{1}{|I_j|} \int_{I_j} f(s) ds \right| \leq \frac{1}{|I_j|} \int_{I_j} |f(s)| ds \leq 2\lambda.$$

If  $t$  belongs to none of the  $I_j$ 's, then this means

$$(1.49) \quad M_{\text{dyadic}} f(t) \leq \lambda.$$

Consequently, applying Lemma 1.14, we obtain  $|g(t)| \leq M_{\text{dyadic}} f(t) \leq 2\lambda$ .  $\square$

*Exercise 3.* Prove (1.49).

**Theorem 1.16** (Riesz, 1927). *There exists  $c > 0$  such that:*

$$(1.50) \quad \left| \left\{ \tilde{f} > \lambda \right\} \right| \leq \frac{c}{\lambda} \|f\|_{L^1(\mathbb{D})}$$

for every  $\lambda > 0$  and  $f \in L^1(\partial\mathbb{D}) \cap L^2(\partial\mathbb{D})$ , and:

$$(1.51) \quad \|\tilde{f}\|_{L^p(\partial\mathbb{D})} \leq c \|f\|_{L^p(\partial\mathbb{D})}$$

for every  $\lambda > 0$  and  $f \in L^p(\partial\mathbb{D}) \cap L^2(\partial\mathbb{D})$ .

*Proof.* Observe that we can assume  $\frac{1}{|\partial\mathbb{D}|} \int_{\partial\mathbb{D}} |f(s)| ds > \lambda$ . For, if not, then we have:

$$(1.52) \quad \left| \left\{ \tilde{f} > \lambda \right\} \right| \leq |\partial\mathbb{D}| \leq \frac{1}{\lambda} \|f\|_1.$$

Thus, let us assume  $\frac{1}{|\partial\mathbb{D}|} \int_{\partial\mathbb{D}} |f(s)| ds > \lambda$  and let  $f = g + b$  as above. Observe that  $\frac{4}{\lambda^2} \leq |\tilde{g}(t)|^2$  if  $t \in \partial\mathbb{D}$  satisfies  $|\tilde{g}(t)| > \frac{\lambda}{2}$ . Thus,

$$(1.53) \quad \left| \left\{ |\tilde{f}| > \lambda \right\} \right| \leq \left| \left\{ |\tilde{g}| > \frac{\lambda}{2} \right\} \right| + \left| \left\{ |\tilde{b}| > \frac{\lambda}{2} \right\} \right|,$$

and:

$$(1.54) \quad \left| \left\{ |\tilde{g}| > \frac{\lambda}{2} \right\} \right| \leq \frac{4}{\lambda^2} \int_{\partial\mathbb{D}} |\tilde{g}(t)|^2 dt.$$

Now that  $\tilde{g}$  is  $L^2(\partial\mathbb{D})$ -bounded, we have

$$(1.55) \quad \frac{4}{\lambda^2} \int_{\partial\mathbb{D}} |\tilde{g}(s)|^2 ds \leq \frac{c}{\lambda^2} \int_{\partial\mathbb{D}} |g(s)|^2 ds.$$

Since  $|g(t)| \leq 2\lambda$  for a.e.  $t \in \partial\mathbb{D}$ , it follows from (1.55) that

$$(1.56) \quad \frac{1}{\lambda^2} \int_{\partial\mathbb{D}} |g(t)|^2 dt \leq \frac{2}{\lambda} \int_{\partial\mathbb{D}} |g(t)| dt.$$

Finally let us recall how  $g$  was constructed:  $g$  was obtained by taking the average of  $f$  on each interval  $I_k$ , and  $g$  remains intact outside such intervals. Therefore,

$$(1.57) \quad \int_{\partial\mathbb{D}} |g(t)| dt \leq \int_{\partial\mathbb{D}} |f(t)| dt.$$

Putting together (1.54), (1.56) and (1.57), we finally have:

$$(1.58) \quad \left| \left\{ |\tilde{g}| > \frac{\lambda}{2} \right\} \right| \leq \frac{c}{\lambda} \int_{\partial\mathbb{D}} |g(t)| dt.$$

Turning to the estimate for  $b$ , we first separate the influence of the intervals  $\{I_k\}_k$ .

$$(1.59) \quad \left| \left\{ |\tilde{b}| > \frac{\lambda}{2} \right\} \right| \leq \sum_k |2I_k| + \left| \left( \partial\mathbb{D} \setminus \bigcup_k 2I_k \right) \cap \left\{ |\tilde{b}| > \frac{\lambda}{2} \right\} \right|.$$

Here,  $2I_k$  is the double of  $I_k$  which is concentric to  $I_k$ , that is, if

$$I_k = \{e^{i\theta} : \theta \in (a_k - h_k, a_k + h_k)\},$$

then  $2I_k := \{e^{i\theta} : \theta \in (a_k - 2h_k, a_k + 2h_k)\}$  and we have

$$(1.60) \quad \sum_k |2I_k| = 2 \sum_k |I_k| = 2 \left| \bigcup_k I_k \right| = 2 |\{M_{\text{dyadic}} f > \lambda\}| \leq 2 |\{Mf > \lambda\}|.$$

If we use the boundedness of  $M$ , then we obtain

$$(1.61) \quad \sum_k |2I_k| \leq 2 |\{Mf > \lambda\}| \leq \frac{c}{\lambda} \|f\|_1.$$

Therefore, the estimate of  $\sum_k |2I_k|$  is now valid.

For the estimate of the remaining term, as usual, we have:

$$\begin{aligned} \left| \left( \partial\mathbb{D} \setminus \bigcup_k 2I_k \right) \cap \left\{ |\tilde{b}| > \frac{\lambda}{2} \right\} \right| &\leq \frac{2}{\lambda} \int_{\partial\mathbb{D} \setminus \bigcup_k 2I_k} |\tilde{b}(s)| ds \\ &\leq \sum_j \frac{2}{\lambda} \int_{\partial\mathbb{D} \setminus \bigcup_k 2I_k} |\tilde{b}_j(s)| ds \\ &\leq \sum_j \frac{2}{\lambda} \int_{\partial\mathbb{D} \setminus 2I_j} |\tilde{b}_j(s)| ds. \end{aligned}$$

Let  $j$  be fixed and  $e^{i\theta}$  be a point outside  $2I_j$ . Then:

$$\begin{aligned} \tilde{b}_j(e^{i\theta}) &= \int_0^{2\pi} \frac{\sin \rho}{1 - \cos \rho} b_j(e^{i(\theta-\rho)}) d\rho \\ &= \int_0^{2\pi} \left( \frac{\sin \rho}{1 - \cos \rho} - \frac{\sin(\theta - c_I)}{1 - \cos(\theta - c_I)} \right) b_j(e^{i(\theta-\rho)}) d\rho \\ (1.62) \quad &= \int_0^{2\pi} \left( \cot\left(\frac{\rho}{2}\right) - \cot\left(\frac{\theta - c_I}{2}\right) \right) b_j(e^{i(\theta-\rho)}) d\rho. \end{aligned}$$

By the mean value theorem:

$$(1.63) \quad \left| \cot\left(\frac{\rho}{2}\right) - \cot\left(\frac{\theta - c_I}{2}\right) \right| \lesssim \frac{|I|}{|\theta - c_I|^2}.$$

so by inserting this estimate to (1.62), we find that:

$$(1.64) \quad \int_{\partial\mathbb{D} \setminus 2I} |\tilde{b}_j(e^{i\theta})| d\theta \lesssim |I| \int_{\partial\mathbb{D} \setminus 2I} \left( \int_{\partial\mathbb{D}} \frac{|b(e^{i(\theta-\rho)})|}{|\theta - c_I|^2} d\rho \right) d\theta \simeq \int_{I_j} |b(t)| dt.$$

If we put together (1.57), (1.59), (1.61) and (1.64) as well as the fact that  $f = g + b$ , then we obtain the desired estimate (1.50).  $\square$

*Exercise 4.* Prove (1.63).

$L^p(\partial\mathbb{D})$ -boundedness with  $1 < p < 2$  by means of interpolation. Here, we shall not make use of the intrinsic expression of the conjugate operation, but instead, what we will need is the so called weak  $L^1(\partial\mathbb{D})$ -boundedness and  $L^2(\partial\mathbb{D})$ -boundedness. Using the fundamental theorem of calculus, we obtain:

$$(1.65) \quad |\tilde{f}(t)|^p = p 2^p \int_0^{\frac{|\tilde{f}(t)|}{2}} \lambda^{p-1} d\lambda = p 2^p \int_0^\infty \chi_{\{(t,\lambda) \in \partial\mathbb{D} \times [0,\infty) : |\tilde{f}(t)| > 2\lambda\}}(t, \lambda) \lambda^{p-1} d\lambda,$$

so

$$\begin{aligned} \int_{\partial\mathbb{D}} |\tilde{f}(t)|^p dt &= p 2^p \int \left( \int_0^\infty \chi_{\{(t,\lambda) \in \partial\mathbb{D} \times [0,\infty) : |\tilde{f}(t)| > 2\lambda\}}(t, \lambda) \lambda^{p-1} d\lambda \right) dt \\ &= p 2^p \int_0^\infty \lambda^{p-1} \left( \int \chi_{\{(t,\lambda) \in \partial\mathbb{D} \times [0,\infty) : |\tilde{f}(t)| > 2\lambda\}}(t, \lambda) dt \right) d\lambda \\ &= p 2^p \int_0^\infty \lambda^{p-1} \left| \left\{ |\tilde{f}| > 2\lambda \right\} \right| d\lambda. \end{aligned}$$

This leads to estimate the measure of the level set  $\left\{ |\tilde{f}| > 2\lambda \right\}$ , so we decompose  $f \in L^p(\partial\mathbb{D})$  at height  $\lambda$ ; that is, we split  $f$  by  $f = f_1 + f_2$  with  $f_1 := \chi_{\{|f| \leq \lambda\}} \cdot f \in L^2(\partial\mathbb{D})$  and  $f_2 := \chi_{\{|f| > \lambda\}} \cdot f \in L^1(\partial\mathbb{D})$ . Then a similar argument to (1.46) gives us:

$$\left| \left\{ |\tilde{f}| > 2\lambda \right\} \right| \leq \left| \left\{ |\tilde{f}_1| > \lambda \right\} \right| + \left| \left\{ |\tilde{f}_2| > \lambda \right\} \right|.$$

To deal with  $\tilde{f}_2$ , observe that  $1 \leq \frac{|\tilde{f}_2|^2}{\lambda^2}$  on  $\left\{ |\tilde{f}_2| > \lambda \right\}$ , which yields:

$$(1.66) \quad \left| \left\{ |\tilde{f}_2| > \lambda \right\} \right| \leq \frac{1}{\lambda^2} \int_{\partial\mathbb{D}} |\tilde{f}_2(t)|^2 dt.$$

As was established before, the operation  $g \mapsto \tilde{g}$  is  $L^2(\partial\mathbb{D})$ -bounded, so as a consequence, we obtain

$$(1.67) \quad \left| \left\{ |\tilde{f}_2| > \lambda \right\} \right| \leq \frac{c}{\lambda^2} \int_{\partial\mathbb{D}} |f_2(t)|^2 dt = \frac{c}{\lambda^2} \int_{\partial\mathbb{D}} \chi_{\{|f| > \lambda\}}(t) |f(t)|^2 dt.$$

Now that we established an alternative version of the  $L^1(\partial\mathbb{D})$ -boundedness as well, we have

$$(1.68) \quad \left| \left\{ |\tilde{f}_1| > \lambda \right\} \right| \leq \frac{c}{\lambda} \int_{\partial\mathbb{D}} |f_1(t)| dt = \frac{c}{\lambda} \int_{\partial\mathbb{D}} \chi_{\{|f| \leq \lambda\}}(t) |f(t)| dt.$$

(1.67) and (1.68) are valid. Here we explain how we use (1.67) and (1.68). It is convenient to calculate some integrals beforehand. Since  $1 < p < 2$ , we have

$$(1.69) \quad \int_0^\infty \lambda^{p-2} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| \leq \lambda\}}(t) \cdot |f(t)| dt \right) d\lambda = \frac{1}{p-1} \int_{\partial\mathbb{D}} |f(t)|^p dt$$

$$(1.70) \quad \int_0^\infty \lambda^{p-3} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| > \lambda\}}(t) |f(t)|^2 dt \right) d\lambda = \frac{1}{2-p} \int_{\partial\mathbb{D}} |f(t)|^p dt.$$

Inserting (1.69) and (1.70) to (1.67) and (1.68), we finally get:

$$\begin{aligned} & \int_{\partial\mathbb{D}} |\tilde{f}(t)|^p dt \\ & \leq C \left( \int_0^\infty \lambda^{p-1} \left| \left\{ |\tilde{f}_1| > \lambda \right\} \right| d\lambda + \int_0^\infty \lambda^{p-1} \left| \left\{ |\tilde{f}_2| > \lambda \right\} \right| d\lambda \right) \\ & \leq C \left\{ \int_0^\infty \lambda^{p-2} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| \leq \lambda\}}(t) |f(t)| dt \right) d\lambda + \int_0^\infty \lambda^{p-3} \left( \int_{\partial\mathbb{D}} \chi_{\{|f| > \lambda\}}(t) |f(t)|^2 dt \right) d\lambda \right\} \\ & = C \int_{\partial\mathbb{D}} |f(t)|^p dt. \end{aligned}$$

The  $L^p(\partial\mathbb{D})$ -boundedness for  $1 < p < 2$  is therefore established.

*Exercise 5.* Prove (1.69) and (1.70).

$L^p(\partial\mathbb{D})$ -boundedness with  $2 < p < \infty$  by way of duality. Let us conclude the proof of the theorem for the case when  $2 < p < \infty$ . We note that:

$$(1.71) \quad \int_{\partial\mathbb{D}} f(t) \tilde{g}(t) dt = \int_{\partial\mathbb{D}} \tilde{f}(t) g(t) dt$$

for all  $f, g \in C(\partial\mathbb{D})$ . By the duality  $L^p(\partial\mathbb{D})$ - $L^{p'}(\partial\mathbb{D})$  we also have

$$(1.72) \quad \|\tilde{f}\|_p = \sup_{g \in C(\partial\mathbb{D}) \setminus \{0\}} \frac{1}{\|g\|_{p'}} \left| \int_{\mathbb{D}} \tilde{f}(t) g(t) dt \right|.$$

If we put (1.71) and (1.72) together and use the Hölder inequality, we obtain

$$(1.73) \quad \|\tilde{f}\|_p = \sup_{g \in C(\partial\mathbb{D}) \setminus \{0\}} \frac{1}{\|g\|_{p'}} \left| \int_{\partial\mathbb{D}} f(t) \tilde{g}(t) dt \right| \leq \sup_{g \in C(\partial\mathbb{D}) \setminus \{0\}} \frac{\|f\|_p \cdot \|\tilde{g}\|_{p'}}{\|g\|_{p'}}$$

Since  $1 < p' < 2$ , we have  $\|\tilde{g}\|_{p'} \lesssim \|g\|_p$  for all  $g \in C(\partial\mathbb{D})$ . Inserting (1.73), we see

$$\|\tilde{f}\|_p \leq C \sup_{g \in C(\partial\mathbb{D}) \setminus \{0\}} \frac{\|f\|_p \cdot \|g\|_{p'}}{\|g\|_{p'}} = C \|f\|_p.$$

Therefore, the proof is now complete.

#### 1.4. Concluding remarks.

Summarize our observations so far, we see that even from this chapter, we have learnt several things.

- (1) The Hardy-Littlewood maximal operator plays a key role when we want to deduce results of a.e. convergence.
- (2) A wonderful geometric observation paved the way to boundedness of maximal operators.
- (3) The  $L^1(\partial\mathbb{D})$ -boundedness was really difficult to obtain in comparison with the  $L^p(\partial\mathbb{D})$ -boundedness with  $1 < p < \infty$ .
- (4) A “substitute” of  $L^1(\partial\mathbb{D})$ -boundedness exists and was proved by a feat, called the CZ-decomposition.

- (5) The space  $L^2(\partial\mathbb{D})$ , along with the Hilbert space structure, seems nicest among other  $L^p(\partial\mathbb{D})$ -spaces.
- (6) Once we are given two different  $L^p(\partial\mathbb{D})$ -estimates, we can interpolate them to yield boundedness for other parameters.
- (7) To obtain deep results, it is not sufficient to rely only on the triangle inequality: We have to take into account the distribution functions.
- (8) Duality  $L^p(\partial\mathbb{D})$ - $L^{p'}(\partial\mathbb{D})$  with  $1 \leq p < \infty$  also helped us obtain the boundedness of operators.
- (9) A weaker variant of the differentiation theorem still holds for  $L^1(\partial\mathbb{D})$ .

Notes and references for Chapter 1.

Section 1. Theorem 1.2 Theorem 1.5 Theorem 1.6 Theorem 1.9

M. Riesz established Theorem 1.7, which asserts that the conjugation operator  $f \mapsto \tilde{f}$  is  $L^p(\mathbb{T})$ -bounded in 1927 [409].

Theorem 1.16

## Part 2. Fundamental facts of Fourier analysis

## Part 3. Measure theory

The aim of this chapter is to deal with the fundamental material on harmonic analysis. That is, we are going to build up a theory on measure theory and Fourier analysis. In Section 3 we deal with measure theory on general measure spaces based on Section 2. In Section 4 the function space  $L^p(\mathbb{R}^d)$  is taken up, where we will give some fundamental inequalities used throughout this book. Section 5 is devoted to dealing with key material, the properties of Schwartz distributions. For example, we discuss the integrability, the differential features of function spaces and so on. Section 8 is the first example of discussing smoothness. Sobolev spaces are widely used in partial differential equations. Since this chapter is intended for those who are not familiar with these topics, the readers who are accustomed to these fields may skip this part.

### 2. A QUICK REVIEW OF GENERAL TOPOLOGY

First let us review fundamental facts on topological spaces. The proofs are omitted.

Metric space.

As important examples of topological spaces, we have metric spaces.

**Definition 2.1** (Metric spaces). Let  $X$  be a set and  $d : X \times X \rightarrow [0, \infty)$  be a mapping. The function  $d$  is said to be a distance function, if for all  $x, y, z \in X$

- (1)  $d(x, y) \geq 0$  and  $d(x, y) = 0$  implies  $x = y$ ,
- (2)  $d(x, y) = d(y, x)$ ,
- (3)  $d(x, z) \leq d(x, y) + d(y, z)$ .

Let  $x \in X$  and  $r > 0$ . Then define  $B(x, r) := \{y \in X : d(x, y) < r\}$ .

**Definition 2.2** (Open sets in metric spaces). Let  $(X, d)$  be a metric space. A subset  $O$  of  $X$  is said to be an open set, if for all  $x \in O$  we can find a real number  $r_x > 0$  such that  $B(x, r_x) \subset O$ .

Topological space. Now we make a quick review topological spaces.

**Definition 2.3** (Topological space). A collection of subsets  $\mathcal{O}_X$  of a set  $X$  is said to be a topology on  $X$ , if it satisfies three axioms listed below.

- (1)  $\emptyset, X \in \mathcal{O}_X$ .
- (2)  $\mathcal{O}_X$  is closed under finite intersection, that is,  $U, V \in \mathcal{O}_X$  implies  $U \cap V \in \mathcal{O}_X$ .
- (3)  $\mathcal{O}_X$  is closed under union, that is,  $\{U_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{O}_X$  implies  $\bigcup_{\lambda \in \Lambda} U_\lambda \in \mathcal{O}_X$ .

Here and below in this section we assume that  $X$  is a set and that  $\mathcal{O}_X$  is a system of open sets in  $X$ .

**Definition 2.4** (Closed set). A subset  $A$  of  $X$  is closed, if  $A^c \in \mathcal{O}_X$ .

*Exercise 6.* Denote by  $\mathcal{F}_X$  the set of all closed sets. Prove the following.

- (1)  $\emptyset, X \in \mathcal{F}_X$ .
- (2)  $\mathcal{F}_X$  is closed under finite union, namely,  $U, V \in \mathcal{F}_X$  implies  $U \cup V \in \mathcal{F}_X$ .

(3)  $\mathcal{F}_X$  is closed under intersection, namely,  $\{U_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{F}_X$  implies  $\bigcap_{\lambda \in \Lambda} U_\lambda \in \mathcal{F}_X$ .

Here we shall collect some elementary facts used in this book concerning general topology.

*Exercise 7.* Let  $X$  be a set and  $\mathcal{X} \subset 2^X$ . Then there exists the weakest topology that contains  $\mathcal{X}$ . Prove that the open sets  $U$  in this topology are characterized by the following condition: For all  $x \in U$  there exist  $N_x \in \mathbb{N}$  and  $U_1, U_2, \dots, U_{N_x} \in \mathcal{X}$  such that  $x \in U_1 \cap U_2 \cap \dots \cap U_{N_x} \subset U$ .

*Exercise 8.* Let  $X$  be a set and  $Y$  be a topological space. Suppose that we are given a family of mappings  $\{f_\lambda\}_{\lambda \in \Lambda}$  from  $X$  to  $Y$ . Then prove that the weakest topology under which each  $f_\lambda$  is continuous is generated by the following family of subsets.

$$(2.1) \quad \{f_\lambda^{-1}(U) : \lambda \in \Lambda, U \text{ is an open set in } Y\}.$$

Let us recall the definition of compact sets.

**Definition 2.5** (Compact set). A subset  $A$  of a topological space  $X$  is said to be compact, if every open covering of  $A$  has a finite subcovering.

A subset in  $\mathbb{R}^d$  is compact precisely when it is bounded and closed, by the Heine-Borel theorem.

**Definition 2.6** (Hausdorff space). A topological space  $X$  is said to be Hausdorff, if for every pair of distinct points  $x, y \in X$  there exist disjoint open sets  $U, V \in \mathcal{O}_X$  with  $x \in U$  and  $y \in V$ .

In this book, we use the following definition and notation about the closure and the interior of sets.

**Definition 2.7** (Closure and interior). Let  $A$  be a subset of a topological space  $X$ .

- (1) The closure of  $A$  is the smallest closed set containing  $A$ , that is, the intersection of all closed sets containing  $A$ . In this book one writes  $\bar{A}$  for the closure of  $A$ .
- (2) The interior of  $A$  is the largest open set contained in  $A$ , that is, the union of all open sets contained in  $A$ . In this book we write  $\text{Int}(A)$  for the interior of  $A$ .

We fix the terminology “neighborhood”.

**Definition 2.8** (Neighborhood). A neighborhood of a set  $A$  is an open set containing  $A$ .

Some authors use the word neighborhood to mean a subset whose interior contains  $A$ . However, for the sake of consistency, we shall use it only for open sets.

Next, we recall dense subsets.

**Definition 2.9** (Dense subset). A subset  $A$  of a topological space  $X$  is said to be dense in  $X$ , if  $\bar{A} = X$ .

Sometimes, we need the notion of separability.

**Definition 2.10** (Separable topological space). A topological space  $X$  is said to be separable, if there exists a countable subset that is dense in  $X$ .

The following is a good criterion for non-separability.

**Proposition 2.11.** *Let  $A$  be a subset of a metric space  $(X, d)$ . If the distance between any two distinct points  $A$  is more than 1, then any dense set of  $X$  has cardinality greater than  $A$ .*

*Proof.* Let  $Y$  be a dense subset of  $X$ . Then, there exists a mapping  $x : A \rightarrow Y$  such that  $d(a, x(a)) < \frac{1}{3}$  for each  $a \in A$ . If  $a$  and  $b$  are distinct, then so are  $x(a)$  and  $x(b)$ . Therefore  $a \in A \mapsto x(a) \in Y$  is an injection and as a result  $Y$  is greater than  $A$  in cardinality.  $\square$

*Exercise 9.* Let  $\ell^\infty(\mathbb{N})$  be the set of all bounded sequence indexed by  $\mathbb{N}$ . Define

$$(2.2) \quad d(a, b) := \sup_{j \in \mathbb{N}} |a_j - b_j|, \quad a = \{a_j\}_{j \in \mathbb{N}}, \quad b = \{b_j\}_{j \in \mathbb{N}} \in \ell^\infty(\mathbb{N}).$$

Then show that the metric space  $(\ell^\infty(\mathbb{N}), d)$  is not separable.

### 3. INTEGRATION THEORY

In this section we give some elementary facts on integration, along with some warm-up exercises. The readers who are familiar with the elementary integration theory can skip this section.

#### 3.1. Measures and outer measures.

$\sigma$ -algebra and measure. We just define the integral for functions. The starting point is to define the integral for  $\chi_A$ , the function that equals 1 on  $A$  and vanishes outside of  $A$ . Recall that the volume is obtained by “height”  $\times$  “area”. in  $\mathbb{R}^3$ . Therefore, if the height equals 1, then the volume and the area agree up to scale. Thus, to determine the value of the integral of  $\chi_A$  is the same as to determine the volume of  $A$ . First we begin with determining what sets are measurable. The  $\sigma$ -algebra, whose definition will be presented below, is our answer. That is, we assign to every element in a  $\sigma$ -algebra its volume. A measure will be a tool to do this.

**Definition 3.1** ( $\sigma$ -algebra). Let  $X$  be a set.  $\mathcal{B} \subset 2^X$  is said to be a  $\sigma$ -algebra if the following three properties hold.

- (1)  $\emptyset, X \in \mathcal{B}$ .
- (2)  $\mathcal{B}$  is closed under complement, that is,  $A \in \mathcal{B} \implies A^c \in \mathcal{B}$ .
- (3)  $\mathcal{B}$  is closed under countable union, that is,  $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$ , whenever  $A_1, A_2, \dots, A_j, \dots \in \mathcal{B}$ .

The measurable sets are sets which can consider their volumes. We do not consider volumes for other sets. Now we consider the problem of what the “volume” should satisfy.

**Definition 3.2** (measure). Let  $\mathcal{B}$  be a  $\sigma$ -algebra over  $X$ .  $\mu : \mathcal{B} \rightarrow [0, \infty]$  is a measure if the following two properties hold.

- (1)  $\mu(\emptyset) = 0$ .
- (2) If  $A_1, A_2, \dots, A_j, \dots \in \mathcal{B}$  are disjoint, that is,  $A_i \cap A_j = \emptyset$  for all  $i \neq j$ , then

$$(3.1) \quad \sum_{j=1}^{\infty} \mu(A_j) = \mu \left( \bigcup_{j=1}^{\infty} A_j \right).$$

Let us give some examples of additive measures.

**Example 3.3** (Counting measure). Let  $X$  be a set and  $\mathcal{B} = 2^X$ . Then if we define  $\mu(A) = \#A$ , then  $(X, \mathcal{B}, \mu)$  is a measure space. We call  $\mu$  the counting measure (of  $X$ ).



We would like to construct a measure  $\mu$  on  $\mathbb{R}^d$  such that  $\mu(I) = |I|$ . However, the existence of such a measure is not trivial at all. Outer measures can be used to construct measures.

**Definition 3.4** (Outer measure).  $\Gamma : 2^X \rightarrow [0, \infty]$  is an outer measure, if

- (1)  $\Gamma(\emptyset) = 0$ .
- (2)  $A \subset B$  implies  $\mu(A) \leq \mu(B)$ .
- (3) For any countable family  $A_1, A_2, \dots, A_j, \dots \in 2^X$  we have  $\sum_{j=1}^{\infty} \Gamma(A_j) \geq \Gamma\left(\bigcup_{j=1}^{\infty} A_j\right)$ .

Note that the outer measure can be defined on  $2^X$ . After fixing the outer measure  $\Gamma$ , we can define the measurability and  $\mathcal{B}$ .

**Definition 3.5** (Measurable set for an outer measure  $\Gamma$ ). Suppose that  $\Gamma$  is an outer measure on  $X$ . A subset  $A$  of  $X$  is said to be  $\Gamma$ -measurable, if  $\Gamma(E) = \Gamma(A \cap E) + \Gamma(A^c \cap E)$  for all  $E \in 2^X$ .

We remark that  $A \subset X$  is  $\Gamma$ -measurable precisely when  $\Gamma(E) \geq \Gamma(A \cap E) + \Gamma(A^c \cap E)$  for all  $E \in 2^X$ . Because the reverse inequality always holds from the definition of outer measures.

**Theorem 3.6.** *Suppose that  $\Gamma$  is an outer measure. Then the set of all  $\Gamma$ -measurable sets forms a  $\sigma$ -algebra.*

*Proof.* Denote by  $\mathcal{B}$  the set of all  $\Gamma$ -measurable sets. It is easy to show  $\emptyset \in \mathcal{B}$ ,  $A \in \mathcal{B} \implies A^c \in \mathcal{B}$ .

$\mathcal{B}$  is closed under finite intersection. Let us prove  $A_1 \cap A_2 \in \mathcal{B}$ , if  $A_1, A_2 \in \mathcal{B}$ . As we have remarked, our present task is to show that  $\Gamma(E) \geq \Gamma(A_1 \cap A_2 \cap E) + \Gamma((A_1 \cap A_2)^c \cap E)$  for all  $E \subset X$ . By virtue of the measurability of  $A_1$  and  $A_2$  we have

$$\begin{aligned} \Gamma(E) &= \Gamma(A_1 \cap E) + \Gamma(A_1^c \cap E) \\ &= \Gamma(A_1 \cap A_2 \cap E) + \Gamma(A_1 \cap A_2^c \cap E) + \Gamma(A_1^c \cap A_2 \cap E) + \Gamma(A_1^c \cap A_2^c \cap E) \\ &= \Gamma(A_1 \cap A_2 \cap E) + \Gamma(A_1^c \cap A_2 \cap E) + \Gamma(A_1^c \cap A_2^c \cap E). \end{aligned}$$

The subadditivity of  $\Gamma$  yields

$$\begin{aligned} \Gamma(E) &\geq \Gamma(A_1 \cap A_2 \cap E) + \Gamma(E \cap ((A_1 \cap A_2^c) \cup (A_1^c \cap A_2) \cup (A_1^c \cap A_2^c))) \\ &= \Gamma(A_1 \cap A_2 \cap E) + \Gamma((A_1 \cap A_2)^c \cap E). \end{aligned}$$

Therefore, we obtain  $\Gamma(E) \geq \Gamma(A_1 \cap A_2 \cap E) + \Gamma((A_1 \cap A_2)^c \cap E)$ , which shows that  $A_1 \cap A_2 \in \mathcal{B}$ .

$\mathcal{B}$  is closed under countably many intersections. Now that  $\mathcal{B}$  is shown to be closed under finite intersection, the matter is reduced to establishing that  $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$ , as long as  $A_1, A_2, \dots, A_j, \dots \in \mathcal{B}$  are mutually disjoint. Our target is to show  $\Gamma(E) \geq \Gamma\left(\bigcup_{j=1}^{\infty} A_j \cap E\right) + \Gamma\left(\bigcap_{j=1}^{\infty} A_j^c \cap E\right)$  for all  $E \subset X$ .

Keeping the disjointness of  $\{A_j\}_{j \in \mathbb{N}}$  in mind, we proceed as follows.

$$\begin{aligned} \Gamma(E) &= \Gamma(A_1 \cap E) + \Gamma(A_1^c \cap E) \\ &= \Gamma(A_1 \cap E) + \Gamma(A_2 \cap E) + \Gamma(A_1^c \cap A_2^c \cap E) \\ &\quad \dots \\ &= \sum_{j=1}^J \Gamma(A_j \cap E) + \Gamma\left(\bigcap_{j=1}^J A_j^c \cap E\right) \\ &\geq \sum_{j=1}^J \Gamma(A_j \cap E) + \Gamma\left(\bigcap_{j=1}^{\infty} A_j^c \cap E\right). \end{aligned}$$

Note that the extreme right-hand side of the above formula is increasing with respect to  $J \in \mathbb{N}$ . Thus, letting  $J \rightarrow \infty$ , we obtain

$$(3.2) \quad \Gamma(E) \geq \sum_{j=1}^{\infty} \Gamma(A_j \cap E) + \Gamma\left(\bigcap_{j=1}^{\infty} A_j^c \cap E\right).$$

Countable subadditivity of  $\Gamma$  leads us to

$$(3.3) \quad \Gamma(E) = \sum_{j=1}^{\infty} \Gamma(A_j \cap E) + \Gamma\left(\bigcap_{j=1}^{\infty} A_j^c \cap E\right) \geq \Gamma\left(\bigcup_{j=1}^{\infty} A_j \cap E\right) + \Gamma\left(\bigcap_{j=1}^{\infty} A_j^c \cap E\right).$$

As we have remarked above, (3.3) is sufficient to show measurability of  $E$ . Thus, the proof is finished.  $\square$

**Theorem 3.7.** *Under the same setting as before, the mapping  $\Gamma|_{\mathcal{B}} : \mathcal{B} \rightarrow [0, \infty]$  is a measure on  $X$ .*

*Proof.* Indeed, we have shown (3.2) for any disjoint family  $A_1, A_2, \dots$  and for all  $E \in 2^X$ . If we put  $E = \bigcup_{j=1}^{\infty} A_j$ , we obtain the countable additivity.  $\square$

Examples. Before we finish this section, we give examples of outer measures.

**Example 3.8.** Let  $\Gamma : 2^{\mathbb{R}^d} \rightarrow [0, \infty]$  be given by

$$(3.4) \quad \Gamma(E) := \sup \left\{ \sum_{j=1}^{\infty} |R_j| : R_j \in \mathcal{R}, E \subset \bigcup_{j=1}^{\infty} R_j \right\},$$

where  $\mathcal{R}$  is the set of open rectangles in  $\mathbb{R}^d$ . Then  $\Gamma$  is an outer measure.  $\Gamma$  is said to be the Lebesgue outer measure, which is a prototype of measures.

*Proof.* It is easy to see that  $\Gamma$  is monotone and  $\Gamma(\emptyset) = 0$ . Thus, it remains to show that  $\Gamma$  enjoys the subadditivity.

$$(3.5) \quad \Gamma\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \Gamma(E_j)$$

for all  $E_1, E_2, \dots \in 2^{\mathbb{R}^d}$ . Suppose that the right-hand side of (3.5) is infinite. In this case the inequality is obvious, no matter what the right-hand side of (3.5) is. Assume that the right-hand side of (3.5) is finite. Let  $\varepsilon > 0$ . Then we can choose a sequence of open rectangles

$\{R_j^k\}_{j \in \mathbb{N}}$  so that

$$(3.6) \quad E_k \subset \bigcup_{j=1}^{\infty} R_j^k, \quad \sum_{j=1}^{\infty} |R_j^k| \leq \Gamma(E_k) + \frac{\varepsilon}{2^k}.$$

Adding the above relation over  $k \in \mathbb{N}$ , we have

$$(3.7) \quad \bigcup_{k=1}^{\infty} E_k \subset \bigcup_{j,k=1}^{\infty} R_j^k, \quad \sum_{k=1}^{\infty} \Gamma(E_k) \leq \left( \sum_{j,k=1}^{\infty} |R_j^k| \right) + \varepsilon,$$

which shows

$$(3.8) \quad \Gamma \left( \bigcup_{k=1}^{\infty} E_k \right) \leq \left( \sum_{j,k=1}^{\infty} |R_j^k| \right) + \varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, we see that  $\Gamma$  is subadditive. □

**Example 3.9.** This example generalizes Example 3.8 with  $d = 1$ . Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a right continuous function. Let  $\Gamma : \mathbb{R}^d \rightarrow [0, \infty]$  be given by

$$(3.9) \quad \Gamma(E) := \inf \left\{ \sum_{j=1}^{\infty} (f(b_j) - f(a_j)) : E \subset \bigcup_{j=1}^{\infty} (a_j, b_j] \right\}.$$

Then  $\Gamma$  is an outer measure. This example is used to define the Stielsties integral.

*Exercise 10.* Reexamine the proof of Example 3.8 to prove that  $\Gamma$ , given by (3.9), is an outer measure.

We are working on something connected with countable sets. So we are naturally led to the following notion:

**Definition 3.10** ( $\sigma$ -finiteness). Let  $(X, \mathcal{B}, \mu)$  be a measure space. A measure  $\mu$  is said to be  $\sigma$ -finite, if  $X$  can be partitioned into a countable collection of disjoint subsets with finite  $\mu$ -measure.

*Exercise 11.* Let  $(X, \mathcal{B}, \mu)$  be a measure space. Then  $\mu$  is  $\sigma$ -finite, if and only if there exists an exhausting sequence  $\{X_j\}_{j \in \mathbb{N}}$  with finite  $\mu$ -measure of  $X$ , that is,  $X$  can be expressed as follows:

$$(3.10) \quad X = \bigcup_{j \in \mathbb{N}} X_j, \quad X_j \subset X_{j+1} \text{ for all } j \in \mathbb{N}, \quad \mu(X_j) < \infty.$$

$\pi$ - $\lambda$  system. To familiarize ourselves with  $\sigma$ -algebras, we deal with  $\pi$ -systems and  $\lambda$ -systems. They are of much importance in their own right.

**Definition 3.11** ( $\pi$ -system,  $\lambda$ -system). Let  $\mathcal{N}$  be a subset of  $2^X$ , where  $X$  is a set.

- (1)  $\mathcal{N}$  is said to be  $\pi$ -system if it is closed under finite intersection.
- (2)  $\mathcal{N}$  is said to be  $\lambda$ -system if it satisfies the following conditions.
  - (a)  $\bigcup_{j=1}^{\infty} N_j \in \mathcal{N}$  whenever  $N_1, N_2, \dots$  is an increasing sequence in  $\mathcal{N}$
  - (b)  $N, M \in \mathcal{N}$  and  $N \subset M$  imply  $M \setminus N \in \mathcal{N}$ .

**Theorem 3.12** ( $\pi$ - $\lambda$  system). Let  $\mathcal{C}$  be a  $\pi$ -system and  $\mathcal{A}$  be a  $\lambda$ -system. Assume that  $\mathcal{C} \subset \mathcal{A}$ . Then we have

$$(3.11) \quad \sigma(\mathcal{C}) \subset \mathcal{A}.$$

*Proof.* Let  $\mathcal{D}$  be the smallest  $\lambda$ -system that contains  $\mathcal{C}$ . Then it suffices to show  $\mathcal{D}$  is a  $\sigma$ -algebra for the purpose of (3.11). To verify that  $\mathcal{D}$  is a  $\sigma$ -algebra, it suffices to show  $\mathcal{D}$  is closed under finite intersection.

Let  $B \in \mathcal{C}$ . Then define

$$(3.12) \quad \mathcal{D}_B := \{C \in \mathcal{D} : B \cap C \in \mathcal{D}\}.$$

Then  $\mathcal{D}_B$  is a  $\lambda$ -system that contains  $\mathcal{C}$ . Therefore, taking into account the minimality, we obtain  $\mathcal{D}_B \supset \mathcal{D}$ . Or equivalently, for all  $B \in \mathcal{C}$  and  $C \in \mathcal{D}$ , we obtain  $B \cap C \in \mathcal{D}$ .

Let  $C \in \mathcal{D}$ . Now consider

$$(3.13) \quad \mathcal{D}_C := \{B \in \mathcal{D} : B \cap C \in \mathcal{D}\}.$$

Then by the above paragraph, we see that  $\mathcal{D}_C$  contains  $\mathcal{C}$ . Furthermore, it is the same as before that  $\mathcal{D}_C$  is a  $\lambda$ -system. Hence we obtain  $\mathcal{D}_C \supset \mathcal{D}$ . Hence for all  $B, C \in \mathcal{D}$ , we have  $B \cap C \in \mathcal{D}$ .  $\square$

### 3.2. Construction of measures starting from a content.

In this subsection we shall present a way of constructing a measure starting from a set function called a content. The origin of content dates back to [23, 81]. Ambrose defined regular contents in his unpublished article [81]. Motivated by his article, Halmos defined contents. This method covers most examples of interest and contains some beautiful applications in measure theory. We postulate a certain assumption on the underlying space  $X$ . In this subsection we assume that  $X$  is a locally compact space. That is, for all  $x \in X$  and all open sets  $U$  containing  $x$  we can find a compact set  $K$  such that  $x \in \text{Int}(K) \subset K \subset U$ .

*Exercise 12.* Show that  $\mathbb{R}^d$  is locally compact.

**Definitions.** Let us begin with presenting some definitions. The notation of content came about in the textbook of [23].

**Definition 3.13** (Content). Denote by  $\mathcal{K}_X$  the set of all compact sets in  $X$  and  $\mathcal{O}_X$  the set of all open sets in  $X$ . A set function  $\lambda : \mathcal{K}_X \rightarrow [0, \infty)$  is said to be a content, if

- (1)  $\lambda(\emptyset) = 0$ .
- (2)  $\lambda(K \cap L) = \lambda(K) + \lambda(L)$ , if  $K \in \mathcal{K}_X$  and  $L \in \mathcal{K}_X$  is disjoint.
- (3)  $\lambda(K \cap L) \leq \lambda(K) + \lambda(L)$  for all  $K, L \in \mathcal{K}_X$ .

Starting from a content, we are going to define a measure. This job is too heavy and we need to decompose the job into several steps. First, by using the topological structure that underlies  $X$ , we consider something close to the desired measure.

**Definition 3.14.** Given a content  $\lambda$ , write

$$(3.14) \quad \lambda_*(A) := \sup\{\lambda(K) : A \in \mathcal{K}_X, A \subset K\}, \quad \mu_e(A) := \inf\{\lambda_*(O) : A \subset O \in \mathcal{O}_X\}.$$

It will be understood that  $\emptyset \subset \emptyset$  and  $\emptyset \subset A$  for every  $A \in 2^X$ .

The outer measure associated with a content. In this paragraph, given a content  $\lambda$ , we shall show  $\mu_e$  is an outer measure. After proving this, we shall investigate some properties.

We note that  $\mu_e(O) = \lambda_*(O)$  for all  $O \in \mathcal{O}_X$ .

**Theorem 3.15.** *The set function  $\mu_e$  is an outer measure.*

*Proof.* We deduce easily that  $\mu(\emptyset) = 0$  from our agreement that  $\emptyset \subset \emptyset$ . Let  $E_1, E_2, \dots \in 2^X$  be arbitrary. We have to show the subadditive inequality

$$(3.15) \quad \mu_e \left( \bigcup_{j=1}^{\infty} E_j \right) \leq \sum_{j=1}^{\infty} \mu_e(E_j).$$

It can be assumed that the right-hand side is finite to prove (3.15).

For each  $j$  take  $V_j \in \mathcal{O}_X$  arbitrarily so that  $E_j \subset V_j$ . Then (3.15) is immediate, once we show

$$(3.16) \quad \mu_e \left( \bigcup_{j=1}^{\infty} V_j \right) \leq \sum_{j=1}^{\infty} \mu_e(V_j).$$

Taking into account the definition of  $\mu_e \left( \bigcup_{j=1}^{\infty} V_j \right)$ , we see (3.16) is reduced to showing

$$(3.17) \quad \lambda(K) \leq \sum_{j=1}^{\infty} \mu_e(V_j).$$

for any compact sets  $K$  contained in  $\bigcup_{j=1}^{\infty} V_j$ . Let  $J$  taken so large that  $K \subset \bigcup_{j=1}^J V_j$ .

Since we are assuming that  $X$  is locally compact, we can take  $W_j \in \mathcal{O}_X, j = 1, 2, \dots, J$  so that  $W_j \Subset V_j$ . Here and below we use  $A \Subset B$  to denote  $\bar{A} \subset \text{Int}(B)$  for subsets  $A$  and  $B$ . Then

$$(3.18) \quad \lambda(K) \leq \sum_{j=1}^J \lambda(K \cap \bar{W}_j) \leq \sum_{j=1}^J \mu_e(V_j) \leq \sum_{j=1}^{\infty} \mu_e(V_j).$$

Thus (3.17) is established. □

**Theorem 3.16.** *Let  $A \in 2^X$ . Then  $A$  is  $\mu_e$ -measurable if and only if*

$$(3.19) \quad \mu_e(O) = \mu_e(A \cap O) + \mu_e(A^c \cap O)$$

for all  $O \in \mathcal{O}_X$ .

Before we come to the proof, let us remark that  $A$  is  $\mu_e$ -measurable if and only if (3.19) holds for all  $O \in 2^X$ , which is just a definition of measurability.

*Proof.* Assume (3.19). Let  $E \in 2^X$ . We have to show that (3.19) holds even for  $E$  instead of  $O$ . Since we are assuming (3.19), we have

$$\mu_e(E) = \inf_{O: E \subset O \in \mathcal{O}_X} \mu_e(O) = \inf_{O: E \subset O \in \mathcal{O}_X} (\mu_e(O \cap A) + \mu_e(O \cap A^c)).$$

Let us use a trivial inequality

$$(3.20) \quad \inf_{\lambda \in \Lambda} (f(\lambda) + g(\lambda)) \geq \inf_{\lambda \in \Lambda} f(\lambda) + \inf_{\lambda \in \Lambda} g(\lambda)$$

for all real-valued functions  $f, g$  defined on a set  $\Lambda$ . We deduce from (3.20) that

$$\mu_e(E) \geq \inf_{E \subset O \in \mathcal{O}_X} \mu_e(O \cap A) + \inf_{E \subset O \in \mathcal{O}_X} \mu_e(O \cap A^c) \geq \mu_e(E \cap A) + \mu_e(E \cap A^c) \geq \mu_e(E).$$

Thus,  $A$  is measurable. □

**Theorem 3.17.** *Any open set is  $\mu_e$ -measurable.*

*Proof.* Let  $V \in \mathcal{O}_X$ . We have to show

$$\mu_e(O) = \mu_e(O \cap V) + \mu_e(O \cap V^c)$$

for any subset  $O$ . By the definition of  $\mu_e$  and subadditivity of  $\mu_e$  this amounts to showing

$$(3.21) \quad \mu_e(O) \geq \lambda(K) + \lambda(L)$$

for all  $K \in \mathcal{K}_X$  and  $L \in \mathcal{K}_X$  with  $K \subset O \cap V$  and  $L \subset O \cap V^c$ . Since  $K \subset O \cap V$  and  $L \subset O \cap V^c$ ,  $K$  and  $L$  are disjoint. Thus  $\lambda(K) + \lambda(L) = \lambda(K \cup L) \leq \mu_e(O)$ .  $\square$

Given a topological space  $X$ , we want to assign the volume to all open sets. So we start from the definition of a  $\sigma$ -algebra on  $X$ .

**Definition 3.18** (Borel  $\sigma$ -field). Let  $X$  be a topological space. Then  $\mathcal{B}(X)$  denotes the smallest  $\sigma$ -field and it is called the Borel  $\sigma$ -field of  $X$ .

Denote by  $\mu$  the restriction of  $\mu_e$  to the Borel sets  $\mathcal{B}(X)$ .

**Theorem 3.19.** *Let  $E$  be a measurable set.*

- (1)  $\mu(E) = \inf\{\mu(O) : O \in \mathcal{O}_X, E \subset O\}$ .
- (2)  $\mu(E) = \sup\{\mu(K) : K \in \mathcal{K}_X, K \subset E\}$ , if  $E \in \mathcal{O}_X$  or  $\mu(E) < \infty$ .
- (3) *Assume in addition that  $X$  is separable. Then  $X$  is  $\sigma$ -finite, that is,  $X$  can be written as the countable sum of sets of finite measure.*

*The property (2) is said to be outer regularity and (3) is said to be compact regularity.*

*Proof.* (3) is immediate: Indeed,  $X$  is locally compact and separable. Therefore it can be written as the sum of countable compact sets. As we have established, the  $\mu$ -measure of any compact set is finite. Hence, (3) is established.

(1) is also easy. By the definition we have

$$(3.22) \quad \mu(E) = \mu_e(E) = \inf\{\mu_e(O) : O \in \mathcal{O}_X, E \subset O\} = \inf\{\mu(O) : O \in \mathcal{O}_X, E \subset O\}.$$

It remains to prove (2) only.

2-a Any element in  $\mathcal{O}_X$  satisfies (2).

 This follows from the construction of  $\mu_e$ .

2-b Any finite measurable set  $E$  satisfies (2).

 Take  $O \in \mathcal{O}_X$  so that  $E \subset O$ ,  $\mu(O \setminus E) < 1$ .

Set

$$(3.23) \quad \mathcal{M} := \{E \in \mathcal{B}(X) : E \subset O \text{ and (2) holds}\}.$$

Our claim is that  $\mathcal{M}$  is a  $\sigma$ -algebra on  $O$  which contains  $\mathcal{O}_X \cap O$ . Here we have defined

$$(3.24) \quad \mathcal{O}_X \cap O := \{A \cap O : A \in \mathcal{O}_X\} = \{U : U \in \mathcal{O}_X, U \subset O\}.$$

Our present task is to prove the following.

- ( $\alpha$ )  $O \in \mathcal{M}$ .
- ( $\beta$ )  $A \in \mathcal{M}$  implies  $O \setminus A \in \mathcal{M}$ .
- ( $\gamma$ )  $A_1, A_2, \dots, A_k, \dots \in \mathcal{M}$  implies  $\bigcup_{k=1}^{\infty} A_k \in \mathcal{M}$ .
- ( $\delta$ ) If  $V \in \mathcal{O}_X$ , then  $V \cap O \in \mathcal{M}$ .

Conditions  $(\alpha)$  and  $(\gamma)$  are easy to verify. Also, the remaining assertion  $(\delta)$  is  $2 - a$  itself. Therefore it remains to show that  $(\beta)$  holds. For this purpose we take  $\varepsilon > 0$ . Let  $K$  be a compact set such that  $\mu(O \setminus K) < \varepsilon$ . Then

$$(3.25) \quad \mu(O \setminus A) \leq \mu((K \cap O) \setminus A) + \mu(O \setminus (K \cup A)) \leq \mu(K \setminus A) + \varepsilon.$$

As we have shown (1), there exists an open set  $U$  which engulfs  $A$  such that  $\mu(U \setminus A) < \varepsilon$ . Note that  $K \setminus U$  is compact and contained in  $O \setminus A$ . Observe also that

$$(3.26) \quad \mu(K \setminus A) = \mu(K \cap U \setminus A) + \mu(K \setminus U) < \varepsilon + \mu(K \setminus U).$$

Consequently

$$(3.27) \quad \mu(O \setminus A) \leq \sup\{\mu(L) : L \in \mathcal{K}_X, K \subset O \setminus A\} + 2\varepsilon.$$

Thus, it follows that  $(\beta)$  is established and (2) was completely proved.  $\square$

Before we finish this section, we give a sufficient condition of inner regularity.

**Corollary 3.20.** *Assume in addition that  $X$  is separable. Then  $\mu$  is inner regular, that is, for every  $E \in \mathcal{B}$ ,*

$$(3.28) \quad \mu(E) = \sup\{\mu(K) : K \subset E, K \in \mathcal{K}_X\}.$$

*Proof.* According to the proof of (3) of the previous theorem, we see that  $X$  is  $\sigma$ -finite. Therefore  $X$  can be partitioned into a countable collection of sets with finite  $\mu$ -measure. Let us write  $X = \sum_{j \in \mathbb{N}} X_j$ , where  $\mu(X_j) < \infty$  for every  $j \in \mathbb{N}$ .

Using this partition and (3) in the previous theorem, we obtain

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap X_j) = \sum_{j=1}^{\infty} \sup\{\mu(K \cap X_j) : K \subset E \cap X_j, K \in \mathcal{K}_X\}$$

Let us write out the sum as a limit of partial sum. The partial sum being made up of finite sums, we obtain

$$\mu(E) = \lim_{J \rightarrow \infty} \sum_{j=1}^J \left( \sup_{\substack{K \in \mathcal{K}_X \\ K \subset E \cap X_j}} \mu(K \cap X_j) \right) = \lim_{J \rightarrow \infty} \sup \left\{ \mu(K) : K \subset E \cap \bigcup_{j=1}^J X_j, K \in \mathcal{K}_X \right\}.$$

Therefore, we obtain

$$(3.29) \quad \mu(E) \leq \sup\{\mu(K) : K \subset E, K \in \mathcal{K}_X\}.$$

The reverse inequality being trivial, we obtain the desired result.  $\square$

It is important that we rephrase the above results for the Lebesgue measure. We take the liberty of repeating them.

**Theorem 3.21.** *Suppose that  $E \subset \mathbb{R}^d$  is  $dx$ -measurable and that  $E$  has finite  $dx$ -measure. Then, given  $\varepsilon > 0$ , we can find a compact set  $K$  and an open set  $U$  such that*

$$(3.30) \quad K \subset E \subset U, |U \setminus K| < \varepsilon.$$

*Proof.* This is a special case of Theorem 3.19 and Corollary 3.20.  $\square$

Having cleared up the definition of the measurable sets, we are now in the position of defining the functions which admit integration.

Throughout this section we assume that  $(X, \mathcal{B}, \mu)$  is a measure space. We adopt the following notation. Denote by  $\overline{\mathbb{R}}$  the two-sided compactification of  $\mathbb{R}$ :  $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ . It is also convenient to set  $\overline{\mathbb{K}} = \overline{\mathbb{R}}$  when  $\mathbb{K}$  denotes  $\mathbb{R}$  and  $\overline{\mathbb{K}} = \mathbb{C}$  when  $\mathbb{K}$  denotes  $\mathbb{C}$ . Let  $f$  be a function from  $X$  to  $\overline{\mathbb{R}}$  or from  $X$  to  $\mathbb{C}$ . Then we write

$$(3.31) \quad \{f \in A\} := \{x \in X : f(x) \in A\}$$

for  $A \in 2^{\overline{\mathbb{R}}}$  respectively.

### 3.3. Measurable functions.

Below we use the following convention.

**Definition 3.22** (Operations in  $\overline{\mathbb{R}}$ ). Let  $\alpha, \beta \in \overline{\mathbb{R}}$ . Suppose that  $\{a_j\}_{j \in \mathbb{N}}$  and  $\{b_j\}_{j \in \mathbb{N}}$  are sequences converging to  $\alpha$  and  $\beta$  respectively.

(1) Define

$$(3.32) \quad \alpha + \beta := \lim_{j \rightarrow \infty} (a_j + b_j), \quad \alpha - \beta := \lim_{j \rightarrow \infty} (a_j - b_j), \quad \alpha \cdot \beta := \lim_{j \rightarrow \infty} a_j b_j.$$

whenever each limit exists and does not depend on the particular choice of  $\{a_j\}_{j \in \mathbb{N}}$  and  $\{b_j\}_{j \in \mathbb{N}}$ .

(2) As an exception, define  $0 \cdot \infty = \infty \cdot 0 = 0$ .

According to our definition, we have the following.

#### Example 3.23.

- (1) The addition, subtraction and the multiplication are the usual ones for finite values.
- (2) If  $a \in \mathbb{R}$  and  $b > 0$ , then  $a + \infty = \infty + a = b \cdot \infty = \infty \cdot b = \infty \cdot \infty = \infty$
- (3)  $\infty - \infty$  does not make sense. Because on the one hand

$$(3.33) \quad \infty - \infty = \lim_{j \rightarrow \infty} j - \lim_{j \rightarrow \infty} j = \lim_{j \rightarrow \infty} (j - j) = 0$$

and on the other hand

$$(3.34) \quad \infty - \infty = \lim_{j \rightarrow \infty} (j + 1) - \lim_{j \rightarrow \infty} j = \lim_{j \rightarrow \infty} (j + 1 - j) = 1.$$

Therefore, the operation  $\infty - \infty$  does depend on the choice of the sequence.

*Exercise 13.*

- (1) Show that  $b \cdot \infty = -\infty$ , whenever  $b < 0$ .
- (2) Explain why  $2\infty - \infty$  does not make sense.

Measurable functions. We begin with the definition and a key property concerning the definition.

**Theorem 3.24.** *Let  $f : X \rightarrow \overline{\mathbb{R}}$  be a function. Then the following are equivalent.*

- (1)  $\{f > \lambda\} \in \mathcal{B}$  for all  $\lambda \in \mathbb{R}$ .
- (2)  $\{f < \lambda\} \in \mathcal{B}$  for all  $\lambda \in \mathbb{R}$ .
- (3)  $\{f \geq \lambda\} \in \mathcal{B}$  for all  $\lambda \in \mathbb{R}$ .
- (4)  $\{f \leq \lambda\} \in \mathcal{B}$  for all  $\lambda \in \mathbb{R}$ .

*Exercise 14.* Prove Theorem 3.24.



Keeping Theorem 3.24 in mind, we define measurable functions.

**Definition 3.25** (Real-valued measurable function). Let  $f : X \rightarrow \overline{\mathbb{R}}$  be a function. If all of the above conditions in Theorem 3.24 are fulfilled, then  $f$  is said to be  $\mu$ -measurable. In this case one writes  $f \in \mathcal{B}$ .

The notion of measurability of real functions can be readily extended to complex-valued functions.

**Definition 3.26** (Complex-valued measurable function). Let  $f : X \rightarrow \mathbb{C}$  be a function. The function  $f$  is said to be measurable, if both  $\operatorname{Re}(f)$  and  $\operatorname{Im}(f)$  are measurable. If this is the case, we write again  $f \in \mathcal{B}$ .

Having clarified the definition of measurability of functions, we prove Theorem 3.24.

*Proof of Theorem 3.24.* Now that  $\mathcal{B}$  is closed under complement, (1) and (4) are equivalent. Similarly (2) and (3) are equivalent. Let us prove that (1) implies (3) and that (2) implies (4).

(1)  $\implies$  (3) Note that  $\mathcal{B}$  is closed under countable intersections. Since we are assuming (1), we have  $\{f > q\} \in \mathcal{B}$  for all  $q \in \mathbb{Q}$ . Therefore, we claim that

$$(3.35) \quad \{f \geq \lambda\} = \bigcap_{q \in \mathbb{Q}, q < \lambda} \{f > q\}.$$

Once this is proved, then  $f$  is  $\mu$ -measurable, the right side belonging to  $\mathcal{B}$ . Since  $\{f \geq \lambda\} \subset \{f > q\}$ , we have

$$(3.36) \quad \{f \geq \lambda\} \subset \bigcap_{q \in \mathbb{Q}, q < \lambda} \{f > q\}.$$

Let  $x \in \bigcap_{q \in \mathbb{Q}, q < \lambda} \{f > q\}$ . Then we have  $f(x) > q$ , whenever  $q \in \mathbb{Q}$  and  $q < \lambda$ . Let  $[\cdot]$  be the Gauss symbol. That is,  $[x]$  denotes the largest integer not exceeding  $x$ . Define

$$(3.37) \quad \lambda_j := \frac{[qj] - 1}{j}$$

Then  $\lambda_j \in \mathbb{Q}$  and  $\lambda_j \uparrow \lambda$  as  $j \rightarrow \infty$ . Therefore, letting  $j \rightarrow \infty$  in  $f(x) > \lambda_j$  for all  $j \in \mathbb{N}$ , we obtain  $f(x) \geq \lambda$ . Therefore the reverse inclusion

$$(3.38) \quad \{f \geq \lambda\} \supset \bigcap_{q \in \mathbb{Q}, q < \lambda} \{f > q\}.$$

is established and hence we obtain

$$(3.39) \quad \{f \geq \lambda\} = \bigcap_{q \in \mathbb{Q}, q < \lambda} \{f > q\}.$$

belongs to  $\mathcal{B}$ .

(2)  $\implies$  (4) We pass to the complement of (3.39), which is always true regardless of measurability of  $f$ . The result is

$$(3.40) \quad \{f \leq \lambda\} = \bigcap_{q \in \mathbb{Q}, q > \lambda} \{f < q\}.$$

Therefore, assuming (2), we obtain (4).

In view of the above observations, (1)–(4) are equivalent. □

*Exercise 15.* Express the following sets without using  $\cup$  and  $\cap$ :

$$A_1 = \bigcup_{j=1}^{\infty} [0, j + 2^{-j}), \quad A_2 = \bigcap_{j=1}^{\infty} [-j, 2^{-j}], \quad A_3 = \bigcap_{j=1}^{\infty} (-j^{-2}, 2^{-j}], \quad A_4 = \bigcup_{j=1}^{\infty} (0, j], \quad A_5 = \bigcup_{j=1}^{\infty} (0, j).$$

Take care of the endpoints in each of your answers.

The following theorem is useful in considering the theory of integration on  $\mathbb{R}$ .

**Theorem 3.27.** *Any open set  $U \subset \mathbb{R}$  can be expressed as a disjoint union of countable open intervals.*

A minor modification of the proof below extends Theorem 3.27 to  $\overline{\mathbb{R}}$ .

*Proof.* Let  $U = \sum_{\lambda \in \Lambda} U_\lambda$  be the decomposition of connected components. Then each  $U_\lambda$  is open. Indeed, let  $x \in U_\lambda$ . Then there exists  $r > 0$  such that  $(x-r, x+r) \subset U$ . Since  $(x-r, x+r)$  is connected,  $(x-r, x+r) \subset U_\lambda$ .

Since any connected open set in  $\mathbb{R}$  is an open interval,  $U_\lambda$  is an open interval. Since  $\mathbb{Q}$  is countable,  $\Lambda$  is countable.  $\square$

Using Theorem 3.27, we can rephrase the measurability.

**Theorem 3.28.** *Let  $f : X \rightarrow \overline{\mathbb{R}}$  be a function on a measure space  $(X, \mathcal{B}, \mu)$ . Then the following are equivalent.*

- (1)  $f$  is measurable.
- (2)  $f^{-1}(E)$  is measurable for any measurable set  $E \subset \overline{\mathbb{R}}$ .

*Proof.*  $\boxed{(2) \implies (1)}$  Let  $\lambda \in \mathbb{R}$ . Then  $\{f > \lambda\} = f^{-1}((\lambda, \infty]) \in \mathcal{B}$ , proving  $f$  is measurable.

$\boxed{(1) \implies (2)}$  Denote by  $\mathcal{B}(\overline{\mathbb{R}})$  the Borel algebra in  $\overline{\mathbb{R}}$ . We define

$$(3.41) \quad \mathcal{C} := \{E \in \mathcal{B}(\overline{\mathbb{R}}) : f^{-1}(E) \in \mathcal{B}\}.$$

Then using elementary formulae of set theory

$$(3.42) \quad f^{-1}\left(\bigcap_{j=1}^{\infty} E_j\right) = \bigcap_{j=1}^{\infty} f^{-1}(E_j), \quad f^{-1}\left(\bigcup_{j=1}^{\infty} E_j\right) = \bigcup_{j=1}^{\infty} f^{-1}(E_j), \quad f^{-1}(E^c) = f^{-1}(E)^c,$$

for all  $E, E_1, E_2, \dots \in 2^{\overline{\mathbb{R}}}$ , we see that  $\mathcal{C}$  is a  $\sigma$ -algebra.

Let  $O \subset \mathbb{R}^d$  be an open set. Then  $O$  can be partitioned into a disjoint union of countable open intervals. The set  $O$  admits the following expression:

$$(3.43) \quad O = \sum_{j=1}^{\infty} (a_j, b_j)$$

Since  $f^{-1}\left(\bigcup_{j=1}^{\infty} (a_j, b_j)\right) = \bigcup_{j=1}^{\infty} f^{-1}((a_j, b_j))$ , it follows that  $O \in \mathcal{C}$ .

Thus,  $\mathcal{C}$  is a  $\sigma$ -algebra containing all the intervals in  $\mathbb{R}$ . Since  $\mathcal{B}(\overline{\mathbb{R}})$  is the smallest  $\sigma$ -algebra containing all the intervals in  $\mathbb{R}$ , we conclude  $\mathcal{B}(\overline{\mathbb{R}}) \subset \mathcal{C}$ . By definition of  $\mathcal{C}$  we have  $\mathcal{C} \subset \mathcal{B}(\overline{\mathbb{R}})$ . Thus,  $\mathcal{B}(\overline{\mathbb{R}}) = \mathcal{C}$ . As a result any element in  $\mathcal{B}(\overline{\mathbb{R}})$  belongs to  $\mathcal{C}$ , that is, any Borel set  $E$  satisfies  $f^{-1}(E) \in \mathcal{B}$ .  $\square$

*Exercise 16.* Prove (3.42).

Properties of measurable functions. Having set down the definition of measurability of functions, let us investigate its properties.

**Theorem 3.29** (Property of real measurable functions). *Suppose that we are given the functions  $f, f_1, f_2, \dots : X \rightarrow \overline{\mathbb{R}}$ .*

- (1)  $f_1, \dots, f_n \in \mathcal{B}$  implies  $\sup_{j \in \mathbb{N}} f_j \in \mathcal{B}$  and  $\inf_{j \in \mathbb{N}} f_j \in \mathcal{B}$ .
- (2) Both  $\limsup_{j \rightarrow \infty} f_j, \liminf_{j \rightarrow \infty} f_j$  are measurable. In particular, if the limit  $\lim_{j \rightarrow \infty} f_j$  exists, then we have  $\lim_{j \rightarrow \infty} f_j \in \mathcal{B}$ .
- (3) Let  $G : \overline{\mathbb{R}} \rightarrow \overline{\mathbb{R}}$  be a continuous function. Then  $G \circ f \in \mathcal{B}$ .
- (4) Assume that  $f, g \in \mathcal{B}$  and that the operation  $f + g$  makes sense. Then  $f + g \in \mathcal{B}$ .
- (5) Assume that  $f, g \in \mathcal{B}$  and that the operation  $f g$  makes sense. Then  $f g \in \mathcal{B}$ .
- (6) Any constant function is measurable.

(1). The fact that  $\sup_{j \in \mathbb{N}} f_j$  belongs to  $\mathcal{B}$  is easy to establish. The argument is straightforward.

To check that  $\sup_{j \in \mathbb{N}} f_j \in \mathcal{B}$ , we have to prove  $\left\{ \sup_{j \in \mathbb{N}} f_j > \lambda \right\} \in \mathcal{B}$ . Note that  $\left\{ \sup_{j \in \mathbb{N}} f_j > \lambda \right\} = \bigcup_{j \in \mathbb{N}} \{f_j > \lambda\}$ . Since each  $f_j$  is measurable, we see that  $\{f_j > \lambda\} \in \mathcal{B}$ . Thus, we conclude that  $\left\{ \sup_{j \in \mathbb{N}} f_j > \lambda \right\} \in \mathcal{B}$ .

In the same way, we establish  $\left\{ \inf_{j \in \mathbb{N}} f_j > \lambda \right\} \in \mathcal{B}$  when we prove  $\inf_{j \in \mathbb{N}} f_j \in \mathcal{B}$ . However, it is not the case that  $\left\{ \inf_{j \in \mathbb{N}} f_j > \lambda \right\} = \bigcap_{j \in \mathbb{N}} \{f_j > \lambda\}$ . To be sure, one inclusion

$$(3.44) \quad \left\{ \inf_{j \in \mathbb{N}} f_j > \lambda \right\} \subset \bigcap_{j \in \mathbb{N}} \{f_j > \lambda\}$$

is correct. However,  $\left\{ \inf_{j \in \mathbb{N}} f_j > \lambda \right\} \supset \bigcap_{j \in \mathbb{N}} \{f_j > \lambda\}$  fails. We need to change our point of view.

What is correct is the relation

$$(3.45) \quad \left\{ \inf_{j \in \mathbb{N}} f_j \geq \lambda \right\} = \bigcap_{j \in \mathbb{N}} \{f_j \geq \lambda\}.$$

Using (3.45), we see that  $\inf_{j \in \mathbb{N}} f_j \in \mathcal{B}$ . □

(2). Observe that  $\limsup_{j \rightarrow \infty} f_j = \inf_{j \in \mathbb{N}} \left( \sup_{k \geq j} f_k \right)$ ,  $\liminf_{j \rightarrow \infty} f_j = \sup_{j \in \mathbb{N}} \left( \inf_{k \geq j} f_k \right)$ . By (1) for each  $k$  we have  $\sup_{k \geq j} f_k, \inf_{k \geq j} f_k \in \mathcal{B}$ . Applying (1) to  $\sup_{k \geq j} f_k, \inf_{k \geq j} f_k \in \mathcal{B}, j \in \mathbb{N}$  again, we conclude

$$(3.46) \quad \limsup_{j \rightarrow \infty} f_j, \liminf_{j \rightarrow \infty} f_j \in \mathcal{B}.$$

Suppose the limit  $\lim_{j \rightarrow \infty} f_j(x)$  exists for all  $x \in X$ . Since we have  $\lim_{j \rightarrow \infty} f_j = \limsup_{j \rightarrow \infty} f_j$  in this case, it follows that  $\lim_{j \rightarrow \infty} f_j \in \mathcal{B}$ . □

(3). Since  $G$  is continuous,  $G^{-1}((\lambda, \infty])$  is an open set and it is partitioned into a disjoint union of open intervals (see Theorem 3.27):

$$(3.47) \quad G^{-1}((\lambda, \infty]) = \sum_{j \in J} I_j$$

where  $J$  is at most countable, yielding  $\{G \circ f > \lambda\} = \{f \in G^{-1}((\lambda, \infty])\} = \bigcup_{j \in J} \{f \in I_j\} \in \mathcal{B}$ .  $\square$

(4). Let  $f_j := \max(-j, \min(f, j))$  and  $g_j := \max(-j, \min(g, j))$ . Below we fix  $j \in \mathbb{N}$ . Instead of showing measurability of  $f + g$  directly, we prove that for  $f_j + g_j$ . According to the definition of the operation in  $\overline{\mathbb{R}}$ , it follows that  $f + g = \lim_{j \rightarrow \infty} (f_j + g_j)$ . Our strategy for the proof is to pass to the limit. Observe that

$$(3.48) \quad \{f_j + g_j > \lambda\} = \bigcup_{\substack{q, r \in \mathbb{Q} \\ q+r > \lambda}} (\{f_j > q\} \cap \{g_j > r\}).$$

Indeed, the inclusion

$$(3.49) \quad \{f_j + g_j > \lambda\} \supset \bigcup_{\substack{q, r \in \mathbb{Q} \\ q+r > \lambda}} (\{f_j > q\} \cap \{g_j > r\}).$$

is obvious. Let us prove the converse inclusion. Suppose that  $f_j(x) + g_j(x) > \lambda$ . Then there exists  $q_0 \in \mathbb{Q}$  such that  $f_j(x) + g_j(x) > q_0 > \lambda$ . Note that  $f_j(x) > q_0 - g_j(x)$ . Thus, there exists  $q_1 \in \mathbb{Q}$  such that  $f_j(x) > q_1 > q_0 - g_j(x)$ . Setting  $q = q_1$  and  $r = q_0 - q_1$ , we obtain  $f_j(x) > q$  and  $g_j(x) > q_1 - q_0 = r$ . Now that  $q + r = q_1 > q_0 > \lambda$ , we see that

$$(3.50) \quad \{f_j + g_j > \lambda\} \subset \bigcup_{\substack{q, r \in \mathbb{Q} \\ q+r > \lambda}} (\{f_j > q\} \cap \{g_j > r\}).$$

As a result (3.48) is established and we obtain

$$(3.51) \quad \{f_j + g_j > \lambda\} = \bigcup_{\substack{q, r \in \mathbb{Q} \\ q+r > \lambda}} (\{f_j > q\} \cap \{g_j > r\}) \in \mathcal{B}.$$

Thus,  $f_j + g_j$  is measurable. A passage to the limit along with (2) shows that  $f + g$  is measurable as well.  $\square$

(5). We may assume that  $f$  and  $g$  are finite by passing to the limit as we did in (4). We may even assume that  $f = g$  because we have proved (4) and

$$(3.52) \quad f \cdot g = \frac{1}{2} ((f + g)^2 - f^2 - g^2).$$

It is rather easy to prove  $f^2 \in \mathcal{B}$ . Indeed,

$$(3.53) \quad \{|f|^2 \geq \lambda\} = \{|f|^2 \geq \max(\lambda, 0)\} = \left\{f \geq \sqrt{\max(\lambda, 0)}\right\} \cup \left\{f \leq -\sqrt{\max(\lambda, 0)}\right\} \in \mathcal{B}.$$

As a consequence (5) is established.  $\square$

(6). Let  $f(x) = k$ , a constant function. Then  $\{f > \lambda\} = X$  if  $\lambda < k$  and  $\{f > \lambda\} = \emptyset$  if  $\lambda \geq k$ . Thus, whether  $\lambda$  is larger than  $k$  or not we have  $\{f > \lambda\} \in \mathcal{B}$ .  $\square$

*Exercise 17.* Suppose that  $f : X \rightarrow \overline{\mathbb{R}}$  is measurable. Use  $|a| = 2 \sup\{0, a\} - a$ ,  $a \in \mathbb{R}$  to give an alternative proof that  $|f| : X \rightarrow \overline{\mathbb{R}}$  is measurable.

We can transplant the assertions of the above theorem into the complex-valued functions.

**Theorem 3.30** (Property of complex measurable functions). *In this theorem by a function we mean a complex-valued function.*

- (1) Any constant function is measurable.
- (2) Let  $f, g \in \mathcal{B}$ . Then  $f + g \in \mathcal{B}$  and  $f \cdot g \in \mathcal{B}$ .
- (3) Let  $f_j : X \rightarrow \mathbb{C}$ ,  $j = 1, 2, \dots$  be measurable functions. If the limit  $\lim_{j \rightarrow \infty} f_j$  exists, then we have  $\lim_{j \rightarrow \infty} f_j \in \mathcal{B}$ .

*Proof.* The matters are reduced to the real-valued case because we have only to split the function into real and complex parts. □

**Example 3.31.** It is convenient for later considerations that we construct a sign function of a complex valued function  $f$ . We are going to construct a measurable function  $g$  such that  $\bar{g} \cdot f = |f|$ . Such a function can be described explicitly: We just put

$$(3.54) \quad g(x) := \lim_{j \rightarrow \infty} \frac{f(x)}{|f(x)| + j^{-1}} \quad (x \in X).$$

Then,  $g$  is a measurable function satisfying  $\bar{g} \cdot f = |f|$ .

Given a measurable function  $f$ , we want to consider its modulus  $|f|$ . So the following definition is of use.

**Definition 3.32** (sgn). Given a measurable function  $f : X \rightarrow \mathbb{C}$ , define  $\text{sgn}(f)$  as a function  $g$  in Example 3.31.

Simple functions. Now we prepare to define the integral for measurable functions. We intend to define the integral for general measurable functions via some approximation procedure. That is, we define the integral first for nice functions called simple functions. Let us begin by presenting the definition.

**Definition 3.33** (Simple functions). A measurable function is simple, if it assumes only a finite number of values.

We welcome positive functions because positive sequences are easy to handle when we consider doubly indexed series.

**Definition 3.34** ( $\mathcal{B}_+$ ). Write  $\mathcal{B}_+ := \{f \in \mathcal{B} : f \geq 0\}$ .

As the next theorem asserts, it is quite important to consider simple functions instead of measurable functions in some cases.

**Theorem 3.35.** *Let  $f \in \mathcal{B}_+$ . Then there exists a sequence of positive simple functions  $\{f_j\}_{j \in \mathbb{N}}$  such that  $0 \leq f_j \leq f$  and  $\lim_{j \rightarrow \infty} f_j = f$  pointwise.*

*Proof.* We set  $f_j := \min\left(j, \frac{[2^j f]}{2^j}\right)$  for each  $j \in \mathbb{N}$ . Then inequality  $[2a] \geq 2[a]$  for  $a \in [0, \infty)$  gives us that  $f_j$  is increasing. Since  $f_j(X) \subset \{0, 2^{-j}, 2 \cdot 2^{-j}, \dots, j2^j \cdot 2^{-j}\}$ , we conclude that  $f_j$  is simple. Since

$$(3.55) \quad \min\left(j, \frac{2^j f(x) - 1}{2^j}\right) \leq f_j(x) \leq f(x)$$

for each  $x \in X$  and  $\lim_{j \rightarrow \infty} \min\left(j, \frac{2^j f(x) - 1}{2^j}\right) = f(x)$  for all  $x \in X$ , we have  $\lim_{j \rightarrow \infty} f_j(x) = f(x)$ . Thus,  $\{f_j\}_{j \in \mathbb{N}}$  is an increasing sequence of simple functions converging to  $f$  pointwise. □

Support of a measurable function  $f$ . Finally we define support of a measurable function  $f$ , when  $X$  is a topological space and  $\mu : \mathcal{B}(X) \rightarrow [0, \infty]$  is a Borel measure.

**Definition 3.36** ( $\text{supp}(f)$ ). Suppose that  $f : X \rightarrow \overline{\mathbb{R}}$  is a measurable function. Then the support of  $f$  is the set of all points  $x$  for which the following property fails: There exists a neighborhood  $U$  of  $x$  such that  $\mu(U \cap \{f \neq 0\}) = 0$ . Below,  $\text{supp}(f)$  denotes the support of  $f$ .

*Exercise 18.* If  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  is a continuous function, then prove that  $\text{supp}(f) = \overline{\{f \neq 0\}}$ , the topological closure of  $\{f \neq 0\}$ . Here it will be understood that  $\text{supp}(f)$  is given in Definition 3.36 with respect to the Lebesgue measure.

### 3.4. Definition of the integral.

Integral of positive simple functions.

Having made clear what “simple” stands for, we now turn to the definition of integral for such functions. Recall that series consisting of positive numbers behave well. This is why we start with positive simple functions.

**Definition 3.37** (Integral for positive functions). Suppose that  $f \in \mathcal{B}_+$  is simple. Then define

$$(3.56) \quad \int_X f(x) d\mu(x) = \int_X f d\mu := \sum_{j=1}^n a_j \mu(E_j),$$

where  $f$  is represented as  $f = \sum_{j=1}^n a_j \chi_{E_j}$ . The sum  $f = \sum_{j=1}^n a_j \chi_{E_j}$  is said to be an admissible expression.

The point is that the definition does not depend on the choice of the representation.

**Lemma 3.38.** *Suppose that  $f \in \mathcal{B}_+$  is simple. Then the definition of  $\int_X f(x) d\mu(x)$  does not depend on the choice of the representation in (3.56).*

*Proof.* First, we shall verify the following.

**Claim 3.39.** *We can assume  $X = \sum_{j=1}^n E_j = \sum_{k=1}^m F_k$ .*

By symmetry, it suffices to treat  $\{E_j\}_{j=1}^n$ . Given an expression  $f = \sum_{j=1}^n a_j \chi_{E_j}$ , we shall construct an admissible representation  $f = \sum_{j=1}^n c_j \chi_{G_j}$  such that  $\{G_j\}_{j=1}^n$  is disjoint and that

$$\sum_{j=1}^n a_j \mu(E_j) = \sum_{j=1}^n c_j \mu(G_j).$$

To do this, setting  $E_0 := \emptyset$  for convenience, we define and  $G_j := E_j \setminus \bigcup_{l=0}^{j-1} E_l$ . Then we have

$$\bigcup_{j=1}^n G_j = \bigcup_{j=1}^n E_j \text{ and } \{G_j\}_{j=1}^n \text{ is disjoint.}$$

Observe that  $E_j = G_1 \cup G_2 \cup \dots \cup G_j$ . Inserting this formula and changing the order of the summations, we have

$$\sum_{j=1}^n a_j \chi_{E_j} = \sum_{j=1}^n \sum_{k=1}^j a_j \chi_{G_k} = \sum_{k=1}^n \left( \sum_{j=k}^n a_j \right) \chi_{G_k}.$$

Set

$$c_j := \sum_{l=j}^n a_l.$$

Then we have  $f = \sum_{j=1}^n c_j \chi_{G_j}$  and

$$\sum_{j=1}^n c_j \mu(G_j) = \sum_{k=1}^n \sum_{j=k}^n a_l \mu(G_k) = \sum_{k=1}^n \left( \sum_{j=k}^n a_j \right) \mu(G_k) = \sum_{j=1}^n \sum_{k=1}^j a_j \mu(G_k) = \sum_{j=1}^n a_j \mu(E_j).$$

If we set  $E_{n+1} = X \setminus (E_1 \cup E_2 \cup \dots \cup E_n)$  and  $c_{n+1} = 0$ , we obtain the desired expression.

Therefore, it was justified that we can assume  $X = \sum_{j=1}^n E_j = \sum_{k=1}^m F_k$ .

Suppose that  $f$  is represented as two different admissible representations :

$$(3.57) \quad f = \sum_{j=1}^n a_j \chi_{E_j} = \sum_{k=1}^m b_k \chi_{F_k}$$

with  $X = \sum_{j=1}^n E_j = \sum_{k=1}^m F_k$ . Then we have to show

$$(3.58) \quad \sum_{j=1}^n a_j \mu(E_j) = \sum_{k=1}^m b_k \mu(E_k).$$

Since  $\{E_j\}_{j=1}^n$  and  $\{F_k\}_{k=1}^m$  are disjoint, we have

$$\sum_{j=1}^n \sum_{k=1}^m b_k \chi_{E_j \cap F_k} = \sum_{k=1}^m b_k \chi_{F_k} = \sum_{j=1}^n a_j \chi_{E_j} = \sum_{j=1}^n \sum_{k=1}^m a_j \chi_{E_j \cap F_k}.$$

Thus, we have

$$(3.59) \quad a_j = b_k, \text{ provided } E_j \cap F_k \neq \emptyset.$$

By using once more the fact that both  $\{E_j\}_{j=1}^n$  and  $\{F_k\}_{k=1}^m$  partition  $X$ , we have

$$(3.60) \quad \sum_{j=1}^n a_j \mu(E_j) = \sum_{j=1}^n \sum_{k=1}^m a_j \mu(E_j \cap F_k) = \sum_{\substack{j=1,2,\dots,n \\ k=1,2,\dots,m \\ E_j \cap F_k \neq \emptyset}} a_j \mu(E_j \cap F_k).$$

By symmetry we have

$$(3.61) \quad \sum_{k=1}^m b_k \mu(E_k) = \sum_{j=1}^n \sum_{k=1}^m b_k \mu(E_j \cap F_k) = \sum_{\substack{j=1,2,\dots,n \\ k=1,2,\dots,m \\ E_j \cap F_k \neq \emptyset}} b_k \mu(E_j \cap F_k).$$

In view of (3.59), (3.60) and (3.61), we obtain (3.58).  $\square$

Integral of positive measurable functions. In the previous paragraph we have set down the definition of the integral for the nicest functions. The definition of simple positive functions are now complete. We now pass to the general function taking its value in  $[0, \infty]$ .

**Definition 3.40** (Definition of integral for positive functions). Let  $f \in \mathcal{B}_+$ . Define the integral

$$\int f d\mu = \int_X f d\mu = \int_X f(x) d\mu(x) \text{ by}$$

$$\int f(x) d\mu(x) = \sup \left\{ \int_X g(x) d\mu(x) : g \text{ is a simple function with } 0 \leq g \leq f \right\}.$$

Unless possible confusion can occur, we adopt the simplest expression.

The monotonicity of the integral is easy to see.

**Lemma 3.41.** *Let  $f, g \in \mathcal{B}_+$  and  $f \geq g$ . Then  $\int_X f(x) d\mu(x) \geq \int_X g(x) d\mu(x)$ .*

*Proof.* Taking into account that the set appearing in the supremum defining  $\int_X f(x) d\mu(x)$  is contained in that defining  $\int_X g(x) d\mu(x)$ , we obtain

$$\begin{aligned} \int_X f(x) d\mu(x) &= \sup \left\{ \int_X h(x) d\mu(x) : h \text{ is simple and satisfies } 0 \leq h \leq f \right\} \\ &\leq \sup \left\{ \int_X h(x) d\mu(x) : h \text{ is simple and satisfies } 0 \leq h \leq g \right\} \\ &= \int_X g(x) d\mu(x). \end{aligned}$$

This is the desired result. □

At first glance the integral operation is not linear. However, the linearity does hold. To prove this, we need the following theorem, which is important of its own right.

**Theorem 3.42** (Monotone convergence theorem). *Suppose that  $\{f_j\}_{j=1}^\infty \subset \mathcal{B}_+$  satisfies  $f_j \leq f_{j+1}$ . Then, if we write  $f = \sup_{j \in \mathbb{N}} f_j$ , we have  $\int_X f_j(x) d\mu(x) \rightarrow \int_X f(x) d\mu(x)$ , as  $j \rightarrow \infty$ .*

*Proof.* Let  $g$  be a simple function such that  $0 \leq g \leq f$ . It suffices to show that

$$(3.62) \quad \int_X \min(g(x), f_j(x)) d\mu(x) \rightarrow \int_X g(x) d\mu(x).$$

Indeed, once we prove (3.62), we obtain

$$\begin{aligned} \int_X f(x) d\mu(x) &= \sup \left\{ \int_X g(x) d\mu(x) : 0 \leq g \leq f \right\} \\ &= \sup \left\{ \left( \lim_{j \rightarrow \infty} \int_X \min(g, f_j) d\mu \right) : 0 \leq g \leq f \right\} \\ &\leq \lim_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) \\ &\leq \int_X f(x) d\mu(x). \end{aligned}$$

As a result we have only to show the theorem when  $f$  is simple.



The function  $f$  can be represented admissibly as  $f(x) = \sum_{l=1}^n a_l \chi_{E_l}(x)$ , where  $(E_l)_{l=1}^n$  is disjoint and  $a_l \geq 0$  for all  $l = 1, 2, \dots, n$ . Set

$$(3.63) \quad f_j^{(k)} := \sum_{l=1}^n a_l (1 - k^{-1}) \chi_{E_l \cap \{f_j > a_l(1 - k^{-1})\}} \quad j, k = 1, 2, \dots.$$

Then by Lemma 3.41 we have

$$(3.64) \quad \int_X f_j^{(k)}(x) d\mu(x) \leq \int_X f_j(x) d\mu(x) \leq \int_X f(x) d\mu(x).$$

As  $j \rightarrow \infty$  we have

$$\begin{aligned} \int_X f_j^{(k)}(x) d\mu(x) &= \sum_{l=1}^n a_l (1 - k^{-1}) \mu(E_l \cap \{f_j > a_l(1 - k^{-1})\}) \\ &\rightarrow \sum_{l=1}^n a_l (1 - k^{-1}) \mu(E_l) = (1 - k^{-1}) \int_X f(x) d\mu(x). \end{aligned}$$

Consequently, it follows from (3.64) that

$$(3.65) \quad (1 - k^{-1}) \int_X f(x) d\mu(x) \leq \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) \leq \int_X f(x) d\mu(x).$$

Since  $k$  is arbitrary as well, letting  $k \rightarrow \infty$ , we obtain  $\lim_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \int_X f(x) d\mu(x)$ .  $\square$

Now we prove the additivity of integral for positive functions.

**Lemma 3.43.** *Let  $f, g \in \mathcal{B}_+$ . Then  $\int_X (f + g)(x) d\mu(x) = \int_X f(x) d\mu(x) + \int_X g(x) d\mu(x)$ .*

*Proof.* Case 1: Assume that  $f$  and  $g$  are simple. Consider admissible representations of  $f$  and  $g$ .

$$(3.66) \quad f = \sum_{j=1}^n a_j \chi_{E_j}, \quad g = \sum_{k=1}^m b_k \chi_{F_k}.$$

As we did in Lemma 3.38, we may assume  $X = \sum_{j=1}^n E_j = \sum_{k=1}^m F_k$ . Since  $\{E_j\}_{j=1}^n$  and  $\{F_k\}_{k=1}^m$  are partitions of  $X$ , it follows that

$$(3.67) \quad f = \sum_{j=1}^n \sum_{k=1}^m a_j \chi_{E_j \cap F_k}, \quad g = \sum_{j=1}^n \sum_{k=1}^m b_k \chi_{E_j \cap F_k}.$$

Let  $a_{jk} := a_j$  and  $b_{jk} := b_k$  for  $j = 1, 2, \dots, n$  and  $k = 1, 2, \dots, m$ . Then we have

$$(3.68) \quad f = \sum_{j=1}^n \sum_{k=1}^m a_{jk} \chi_{E_j \cap F_k}, \quad g = \sum_{j=1}^n \sum_{k=1}^m b_{jk} \chi_{E_j \cap F_k}, \quad f + g = \sum_{j=1}^n \sum_{k=1}^m (a_{jk} + b_{jk}) \chi_{E_j \cap F_k}.$$

Using (3.68), we have

$$\int_X f(x) d\mu(x) + \int_X g(x) d\mu(x) = \sum_{j=1}^n \sum_{k=1}^m (a_{jk} + b_{jk}) \mu(E_j \cap F_k) = \int_X (f + g)(x) d\mu(x).$$

Case 2: General case Set  $f_j = \min\left(j, \frac{[2^j f(x)]}{2^j}\right)$  and  $g_j = \min\left(j, \frac{[2^j g(x)]}{2^j}\right)$ . Then we have  $f_j \uparrow f$  and  $g_j \uparrow g$ . Therefore, a repeated application of Theorem 3.42 gives us

$$\begin{aligned} \int_X (f+g)(x) d\mu(x) &= \lim_{j \rightarrow \infty} \int_X (f_j + g_j)(x) d\mu(x) \\ &= \lim_{j \rightarrow \infty} \left( \int_X f_j d\mu(x) + \int_X g_j(x) d\mu(x) \right) \\ &= \int_X f(x) d\mu(x) + \int_X g(x) d\mu(x). \end{aligned}$$

Here for the second inequality we have used the fact that  $f_j$  and  $g_j$  are simple. Additivity for simple functions is already established. This is the desired result.  $\square$

Positive homogeneity is much easier to prove than additivity.

**Lemma 3.44.** *Let  $f \in \mathcal{B}_+$  and  $a \geq 0$ . Then we have  $\int_X a \cdot f(x) d\mu(x) = a \int_X f(x) d\mu(x)$ .*

*Exercise 19.* Prove Lemma 3.44.

Integral of real-valued functions. Given an  $\mathbb{R}$ -valued function  $f$ , we can separate it into a difference of two positive functions. For example, a simple candidate for such a decomposition is

$$(3.69) \quad f = f_+ - f_-.$$

Therefore, it looks sensible to define

$$(3.70) \quad \int_X f(x) d\mu(x) := \int_X f_+(x) d\mu(x) - \int_X f_-(x) d\mu(x).$$

However, this definition (3.70) will not do as it stands. For example, we hope that the integral is linear. If we begin with this definition, then we will face a problem in proving linearity.

Actually, we overcome this difficulty by establishing Lemma 3.46.

**Definition 3.45** (Integrable functions). Let  $f \in \mathcal{B}$ . The function  $f$  is said to be  $(\mu)$ -integrable or  $(\mu)$ -summable, if  $\int_X |f(x)| d\mu(x) < \infty$ . If either  $f_+$  or  $f_-$  is integrable, one defines

$$(3.71) \quad \int_X f(x) d\mu(x) = \int_X f_+(x) d\mu(x) - \int_X f_-(x) d\mu(x).$$

The space  $L^1(\mu)$  denotes the set of all  $\mu$ -integrable functions and  $L^1(\mu)_+$  denotes the set of all positive  $\mu$ -integrable functions.

The crux of the proof of linearity lies in the following lemma.

**Lemma 3.46.** *Let  $f \in \mathcal{B}$  be integrable and assume that it is decomposed as  $f = g - h$  where  $g, h \in L^1(\mu)_+$ . Then*

$$\int_X f(x) d\mu(x) = \int_X g(x) d\mu(x) - \int_X h(x) d\mu(x).$$

*Proof.* Since  $g - h = f_+ - f_-$ , we have  $g + f_- = h + f_+$ . Therefore, it follows from Lemma 3.43 that

$$\begin{aligned} \int_X g(x) d\mu(x) + \int_X f_-(x) d\mu(x) &= \int_X (g + f_-)(x) d\mu(x) \\ &= \int_X (h + f_+)(x) d\mu(x) \\ &= \int_X h(x) d\mu(x) + \int_X f_+(x) d\mu(x). \end{aligned}$$

Since all the integrals above are finite, we are in the position of subtracting

$$\int_X f_-(x) d\mu(x) + \int_X h(x) d\mu(x) < \infty$$

from both sides. The result is

$$\int_X g(x) d\mu(x) - \int_X h(x) d\mu(x) = \int_X f_+(x) d\mu(x) - \int_X f_-(x) d\mu(x) = \int_X f(x) d\mu(x),$$

proving the lemma. □

Linearity of integral carries over to integrable functions, of course.

**Corollary 3.47.** *Let  $f, g \in L^1(\mu)$ . Then*

$$\int_X (f + g)(x) d\mu(x) = \int_X f(x) d\mu(x) + \int_X g(x) d\mu(x).$$

*Proof.* We use equality  $f(x) + g(x) = f_+(x) + g_+(x) - f_-(x) - g_-(x)$  and Lemma 3.43 to calculate  $\int_X f(x) d\mu(x) + \int_X g(x) d\mu(x)$ . By Definition 3.45 we have

$$\begin{aligned} \int_X f(x) d\mu(x) + \int_X g(x) d\mu(x) \\ = \int_X f_+(x) d\mu(x) - \int_X f_-(x) d\mu(x) + \int_X g_+(x) d\mu(x) - \int_X g_-(x) d\mu(x). \end{aligned}$$

Observe that all the integrals are finite and that the functions are positive. Thus, the order of the summation and the integration can be exchanged :

$$\begin{aligned} \int_X f_+(x) d\mu(x) - \int_X f_-(x) d\mu(x) + \int_X g_+(x) d\mu(x) - \int_X g_-(x) d\mu(x) \\ = \int_X f_+(x) d\mu(x) + \int_X g_+(x) d\mu(x) - \int_X f_-(x) d\mu(x) - \int_X g_-(x) d\mu(x) \\ = \int_X (f_+ + g_+)(x) d\mu(x) - \int_X (f_- + g_-)(x) d\mu(x). \end{aligned}$$

Now we invoke Lemma 3.46 with  $g = f_+ + g_+$  and  $h = f_- + g_-$  and we obtain

$$\int_X (f_+ + g_+)(x) d\mu(x) - \int_X (f_- + g_-) d\mu(x) = \int_X (f_+ + g_+ - f_- - g_-)(x) d\mu(x).$$

The integrand of the right-hand side being equal to  $f - g$ , we conclude the proof with all the observations above. □

Homogeneity of integral is immediate and we leave the proof for interested readers (see Exercise 20).

**Lemma 3.48.** *Let  $f \in L^1(\mu)$  and  $a \in \mathbb{R}$ . Then  $\int_X a \cdot f(x) d\mu(x) = a \int_X f(x) d\mu(x)$ .*

*Exercise 20.* Prove Lemma 3.48.

Our observation can be summarized as follows :

**Theorem 3.49.**

- (1) *The integral operation  $\int : L^1(\mu) \rightarrow \mathbb{R}$  is linear.*  
 (2) *Suppose that  $f, g \in L^1(\mu)$  satisfy  $f \geq g$ . Then  $\int_X f(x) d\mu(x) \geq \int_X g(x) d\mu(x)$ .*

Integral of complex-valued functions.

In Fourier analysis, for example, it is convenient to define the integrals taking value in  $\mathbb{C}$ . Having made an elaborate treatment, we are readily to extend our definition to such function. However, in this case it is important to rule out the possibility for the integral to be  $\infty$ . It is the case that in complex analysis  $\infty$  is added as a special point to  $\mathbb{C}$  to define the Riemann sphere. However, in integration theory we do not add  $\infty$  to  $\mathbb{C}$  to extend our operation.

**Definition 3.50** ( $L^1(\mu)$ ). Suppose that  $f \in \mathcal{B}$  is a complex valued function. Then  $f$  is integrable, if  $|f|$  is integrable. One still denotes  $L^1(\mu)$  by the set of complex valued integrable functions. By linearity one extends the integral.

We need to make what we obtained applicable on subsets. To this end, we use the characteristic functions.

**Definition 3.51** ( $L^1(E, \mu)$ ). Suppose that  $f \in \mathcal{B}$  and that  $E \in \mathcal{B}$  satisfy  $\chi_E f \in L^1(\mu)$ . (Write  $f \in L^1(E, \mu)$ .) Then define  $\int_E f d\mu = \int_E f(x) d\mu(x) := \int_X \chi_E(x) f(x) d\mu(x)$ .

Following this notation, we have

**Proposition 3.52.** *Suppose that  $E, F \in \mathcal{B}$  are disjoint. Then*

$$(3.72) \quad \int_{E \cup F} f(x) d\mu(x) = \int_E f(x) d\mu + \int_F f(x) d\mu(x)$$

for all  $f \in L^1(E \cup F, \mu)$ .

Needless to say, this corresponds to the formula

$$(3.73) \quad \int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

for Riemannian integrals.

*Proof.* Note that  $\chi_{E \cup F} = \chi_E + \chi_F$  characterizes the disjointness of  $E$  and  $F$ . Therefore, we obtain

$$\begin{aligned} \int_{E \cup F} f(x) d\mu(x) &= \int_X \chi_{E \cup F}(x) f(x) d\mu(x) \\ &= \int_X \chi_E(x) f(x) + \chi_F(x) f(x) d\mu(x) \\ &= \int_E f(x) d\mu(x) + \int_F f(x) d\mu(x). \end{aligned}$$

This is the desired result. □

*Exercise 21.* Let  $f : X \rightarrow \mathbb{C}$  be a complex-valued measurable function. Show that the following are equivalent.

- (1)  $f \in L^1(\mu)$ .

- (2)  $\operatorname{Re}(f) \in L^1(\mu)$  and  $\operatorname{Im}(f) \in L^1(\mu)$ .
- (3)  $\operatorname{Re}(f)_\pm \in L^1(\mu)$  and  $\operatorname{Im}(f)_\pm \in L^1(\mu)$ .

In considering the integral or something related to integral, there is no need to know the value of function at all points. Indeed, we just need to know them at almost all points. The next definition makes this more precise.

**Definition 3.53** (Almost all e.t.c.). Let  $(X, \mathcal{B}, \mu)$  be a measure space. A property holds for almost all / almost everywhere / almost every etc., if there exists a set  $A$  of  $\mu$ -measure 0 such that the property holds outside  $A$ . That is, the word “almost” means that the set of all points such that the property fails has  $\mu$ -measure 0. A null set means a set of measure 0.

**Example 3.54.** Below we exhibit examples of the usage of “almost”

- (1) Let  $(X, \mathcal{B}, \mu)$  be a measure space.  $f(x) \leq g(x)$   $\mu$ -a.e. means that there exists a  $\mu$ -null set  $A$  such that  $f(x) \leq g(x)$  for all  $x \in X \setminus A$ .
- (2) Let  $(\Omega, \mathcal{F}, \mu)$  be a probability space, that is,  $\mu(X) = 1$ . Then we use almost surely instead of almost every. We abbreviate this to a.s. in probability theory. For example  $\lim_{j \rightarrow \infty} X_j(\omega)$  exists almost surely means there exists a null set  $\Omega_0$  such that  $\lim_{j \rightarrow \infty} X_j(\omega)$  exists on  $\Omega \setminus \Omega_0$ .
- (3) The above two example are just rephrasing the definition. Let us see how this notion is used actually. For example,  $\lim_{j \rightarrow \infty} \sin(2\pi jx)$  diverges for  $dx$ -almost everywhere  $x \in \mathbb{R}$ . (Indeed, the limit does not exist if and only if  $\theta$  is irrational.)

Before proceeding, let us see another example of “almost”.

**Lemma 3.55.** Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $f : X \rightarrow [0, \infty]$  be a measurable function. Assume the integral is finite:  $\int_X f(x) d\mu(x) < \infty$ . Then  $f$  is finite a.e.

*Proof.* Suppose instead that  $f = \infty$  on a set  $A$  of positive  $\mu$ -measure. Then since  $\mu(A) > 0$ , we have

$$(3.74) \quad \int_X f(x) d\mu(x) \geq \int_A f(x) d\mu(x) = \mu(A) \cdot \infty = \infty.$$

This runs counter to the assumption. □

*Exercise 22.* Show that it is possible that  $f$  assumes  $\infty$  at some point even when

$$\int_X f(x) d\mu(x) < \infty$$

in Lemma 3.55.

Here we present a routine way with which to enlarge the class of sets for which we can consider its volume.

**Definition 3.56** (Completion). Let  $(X, \mathcal{B}, \mu)$  be a measure space.

- (1) Define

$$(3.75) \quad \mathcal{B}^* := \{B \in 2^X : A_0 \subset B \subset A_1 \text{ with } A_0, A_1 \in \mathcal{B} \text{ and } \mu(A_0) = \mu(A_1)\}$$

and

$$(3.76) \quad \mu^*(B) := \mu(A_0) = \mu(A_1), \text{ if } A_0 \subset B \subset A_1 \text{ with } A_0, A_1 \in \mathcal{B} \text{ and } \mu(A_0) = \mu(A_1).$$

The completion of  $(X, \mathcal{B}, \mu)$  is, by definition,  $(X, \mathcal{B}^*, \mu^*)$ .

- (2) The space  $X$  is said to be complete whenever a set which is contained in a  $\mu$ -null set is measurable.

*Exercise 23.* Show that  $(X, \mathcal{B}^*, \mu^*)$  is a measure space. If  $X$  is complete, then prove that  $\mathcal{B}^* = \mathcal{B}$ .

*Exercise 24.* Show that the set of all Lebesgue measurable sets is the completion of  $\mathcal{B}(\mathbb{R}^d)$ .

### 3.5. Convergence theorems.

In this section we deal with convergence theorems. Here and below for the sake of simplicity we use  $f \leq g$  to mean not only  $f(x) \leq g(x)$  for all  $x \in X$  but also  $f(x) \leq g(x)$  for  $\mu$ -almost all  $x \in X$ . Here we harvest the consequences of the complicated definitions we have made all the way.

Theorems dealing with the change of  $\lim$  and  $\int$ . This is one of the most important theorems in integration theory. Apart from their proofs, it is absolutely necessary that we utilize them throughout this book.

**Theorem 3.57.** *Suppose that  $(X, \mathcal{B}, \mu)$  is a measure space. Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence of  $\overline{\mathbb{R}}$ -valued  $\mu$ -measurable functions.*

- (1) (Monotone convergence theorem.) *Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a positive and increasing sequence: We have  $0 \leq f_j \leq f_{j+1}$ ,  $\mu$ -a.e.. Then*

$$(3.77) \quad \lim_{j \rightarrow \infty} \int_X (x) f_j d\mu(x) = \int_X \lim_{j \rightarrow \infty} f_j(x) d\mu(x).$$

- (2) (Fatou's lemma.) *Suppose  $\{f_j\}_{j \in \mathbb{N}}$  is a positive measurable sequence. Then we have*

$$(3.78) \quad \int_X \liminf_{j \rightarrow \infty} f_j(x) d\mu(x) \leq \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x).$$

- (3) (Dominated convergence theorem, Lebesgue's convergence theorem.) *Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  converges  $\mu$ -almost everywhere to  $f$ . Assume further that there exists  $g \in L^1(\mu)$  such that  $|f_j| \leq g$  a.e. for all  $j \in \mathbb{N}$ . Then*

$$(3.79) \quad \lim_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \int_X f(x) d\mu(x).$$

(1). (3.77) is taken up as Theorem 3.42, which is a starting point of (3.78) and (3.79). In Theorem 3.42 we have assumed that  $0 \leq f_j(x) \leq f_{j+1}(x)$  for all  $x \in X$  and  $j \in \mathbb{N}$ . However, if we dilate a set of  $\mu$ -measure zero, we can go through the same argument as before. Thus, (3.77) is complete.  $\square$

(2). Now we shall prove (3.78). If we set  $g_j(x) := \inf_{k \geq j} f_k(x)$  for  $x \in X$ , then  $\{g_j\}_{j \in \mathbb{N}}$  satisfies the assumption of Theorem 3.42 and  $\lim_{j \rightarrow \infty} g_j = \liminf_{j \rightarrow \infty} f_j$  for all  $x \in X$ . Therefore, (3.78) gives

$$\begin{aligned} \int_X \liminf_{j \rightarrow \infty} f_j(x) d\mu(x) &= \int_X \lim_{j \rightarrow \infty} g_j(x) d\mu(x) \\ &= \lim_{j \rightarrow \infty} \int_X g_j(x) d\mu(x) \\ &= \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) \\ &\leq \liminf_{j \rightarrow \infty} \int_X f_j d\mu(x). \end{aligned}$$

This is the desired result. □

(3). To prove (3.79), note that  $g - f_j$  is positive. Thus, we are in the position of using (3.78) to have

$$(3.80) \quad \int_X \liminf_{j \rightarrow \infty} (g - f_j)(x) d\mu(x) \leq \liminf_{j \rightarrow \infty} \int_X (g - f_j)(x) d\mu(x),$$

that is,

$$(3.81) \quad \int_X g(x) d\mu(x) - \int_X \limsup_{j \rightarrow \infty} f_j(x) d\mu(x) \leq \int_X g(x) d\mu(x) - \limsup_{j \rightarrow \infty} \int_X f_j(x) d\mu(x).$$

Equating (3.81), we obtain

$$(3.82) \quad \int_X \limsup_{j \rightarrow \infty} f_j(x) d\mu(x) \geq \limsup_{j \rightarrow \infty} \int_X f_j(x) d\mu(x).$$

Going through the same argument by using  $\{g + f_j\}_{j \in \mathbb{N}}$ , we obtain

$$(3.83) \quad \int_X \liminf_{j \rightarrow \infty} f_j(x) d\mu(x) \leq \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x).$$

By assumption that  $\{f_j\}_{j \in \mathbb{N}}$  converges to  $f$ , we have

$$(3.84) \quad \liminf_{j \rightarrow \infty} f_j = \limsup_{j \rightarrow \infty} f_j = \lim_{j \rightarrow \infty} f_j = f.$$

Also it is trivial that

$$(3.85) \quad \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) \leq \limsup_{j \rightarrow \infty} \int_X f_j(x) d\mu(x).$$

Putting together (3.82)–(3.85), we obtain

$$(3.86) \quad \liminf_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \limsup_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \int_X f(x) d\mu(x).$$

Thus, it follows that  $\lim_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \int_X f(x) d\mu(x)$ . □

Before we proceed further, two helpful remarks may be in order.

**Remark 3.58.** It is important that we extend the Lebesgue convergence theorem to complex-valued functions. Suppose that  $\{f_j\}_{j=1}^\infty$  is a sequence of complex-valued functions on a measure space  $(X, \mathcal{B}, \mu)$  and  $f$  is a complex-valued measurable function on  $X$ . Assume that the limit  $\lim_{j \rightarrow \infty} f_j(x)$  exists and coincides with  $f(x)$  for  $\mu$ -almost all  $x \in X$  and that there exists a function  $g \in L^1(\mu)$  such that

$$(3.87) \quad |f_j| \leq g \text{ for all } j \in \mathbb{N}$$

holds for  $\mu$ -almost all  $x \in X$ . Then we have

$$(3.88) \quad \lim_{j \rightarrow \infty} \int_X f_j(x) d\mu(x) = \int_X f(x) d\mu(x).$$

*Exercise 25.* Prove (3.88).

**Remark 3.59.** It is also important that we extend (3.79) to continuous variables. Let  $I$  be an open interval in  $\mathbb{R}$  and  $F : X \times I \rightarrow \overline{\mathbb{R}}$  measurable. Assume the following.

- (1)  $\lim_{t \rightarrow t_0, t \in I} F(x, t) = F(x, t_0)$  for almost every  $x \in X$ .
- (2) There exists an integrable function  $G : X \rightarrow [0, \infty]$  such that, for almost every  $x$ ,  $|F(x, t)| \leq G(x)$  for all  $t \in I$ .

Then we have

$$\lim_{t \rightarrow t_0} \int_X F(x, t) d\mu(x) = \int_X F(x, t_0) d\mu(x).$$

Below we call this assertion Lebesgue's convergence theorem as well.

*Proof.* Since the continuity of the function on  $I$  is equivalent to its sequential continuity, the assertion is immediate from (3.79).  $\square$

**Theorem 3.60** (Change of the order of integration and differentiation). *Let  $(X, \mathcal{B}, \mu)$  be a measure space. Assume that a function  $f : X \times (a, b) \rightarrow \mathbb{C}$  satisfies the following.*

- (1) For each  $t \in (a, b)$ ,  $f(\cdot, t)$  is a  $\mu$ -integrable function.
- (2) For  $\mu$ -almost all  $x \in X$  the function  $t \rightarrow f(x, t)$  is differentiable for all  $t \in (a, b)$ .
- (3) There exists a  $\mu$ -integrable function  $g$  such that  $\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x)$  for all  $t \in (a, b)$  for  $\mu$ -almost every  $x \in X$ .

Then we have

$$(3.89) \quad \frac{\partial}{\partial t} \int_X f(x, t) d\mu(x) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x).$$

*Proof.* Let  $h \neq 0$  be small enough. Then

$$(3.90) \quad \left| \frac{f(x, t+h) - f(x, t)}{h} \right| = \left| \int_0^1 \frac{\partial f}{\partial t}(x, t+hs) ds \right| \leq \int_0^1 \left| \frac{\partial f}{\partial t}(x, t+hs) \right| ds \leq g(x).$$

Thus, the key condition for the Lebesgue convergence theorem is satisfied. By the Lebesgue convergence theorem we obtain

$$\begin{aligned} \frac{\partial}{\partial t} \int_X f(x, t) d\mu(x) &= \lim_{h \rightarrow 0} \frac{1}{h} \left( \int_X f(x, t+h) d\mu - \int_X f(x, t) d\mu \right) \\ &= \lim_{h \rightarrow 0} \int_X \frac{1}{h} \left( f(x, t+h) - f(x, t) \right) d\mu(x) \\ &= \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x). \end{aligned}$$

This is the desired result.  $\square$

We conclude this paragraph with a cautionary example.

**Example 3.61.** Let  $\mathbb{Q} = \{r_1, r_2, \dots, r_j, \dots\}$  be a rearrangement of  $\mathbb{Q}$ . Define

$$(3.91) \quad \varphi(t) := \frac{\chi_{(0,1)}(t)}{t}, \quad f(t) := \sum_{j=1}^{\infty} \frac{1}{4^j} \varphi(t - r_j)$$

for  $t \in \mathbb{R}$ . Then we have the following.

- (1)  $f(t) < \infty$  for a.e.  $t \in \mathbb{R}$ .
- (2)  $\int_a^b f(t) dt = \infty$  for all  $a < b$ .



Indeed, note that  $\sqrt{f(t)} \leq \sum_{j=1}^{\infty} \frac{\chi_{(0,1)}(t-r_j)}{2^j \sqrt{t-r_j}}$ . Therefore,  $\sqrt{f}$  is integrable over  $\mathbb{R}$  and hence it is finite for a.e.  $t \in \mathbb{R}$  by Lemma 3.55, proving (1). For the proof of (2) we have only to note that  $\int_0^1 \frac{dt}{t} = \infty$ .

Riemann-integral and Lebesgue integral. After defining two types of integrals, the Riemann-integral and the Lebesgue integral, we are eager to connect them. We are going to investigate how they are related. As well as a condition for the two types of integrals to coincide, we obtain the necessary and sufficient conditions for a bounded function  $f$  defined on a compact rectangle  $R$  to be Riemann-integrable. We give an answer in terms of the Lebesgue integral.

**Theorem 3.62** (Darboux). *Let  $R$  be a compact rectangular in  $\mathbb{R}^d$ . Suppose that  $f : R \rightarrow \mathbb{R}$  is a bounded measurable function. Then  $f$  is Riemann-integrable, if and only if  $f$  is continuous at almost all points in  $R$ . If this is the case, then its Riemann-integral and its Lebesgue-integral coincide.*

*Proof.* We may assume that

$$R = \prod_{j=1}^d [0, R_j]$$

by translation. Bisect  $R$   $k$ -times to obtain rectangles  $R_j^k, j = 1, 2, \dots, 2^{kd}$ . Set

$$E := \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{kd}} \partial(R_j^k).$$

Given a point  $x \in R \setminus E$  and  $k \in \mathbb{N}$ , we define  $j(k, x) \in [1, 2^{kd}] \cap \mathbb{Z}$  uniquely so that  $x \in R_{j(k,x)}^k$ . Define

$$\begin{aligned} \bar{f}(x) &:= \lim_{k \rightarrow \infty} \left( \sup_{y \in R_{j(k,x)}^k} \chi_E(x) f(y) \right) \\ \underline{f}(x) &:= \lim_{k \rightarrow \infty} \left( \inf_{y \in R_{j(k,x)}^k} \chi_E(x) f(y) \right). \end{aligned}$$

Suppose that  $x \in R$  lies outside  $E$ . Then  $f$  is continuous precisely when  $\bar{f}(x) = \underline{f}(x)$ . Furthermore, by Darboux's theorem and Lebesgue's convergence theorem we have

$$(3.92) \quad \int_R \bar{f}(x) dx = \overline{\int_R f(x) dx}, \quad \int_R \underline{f}(x) dx = \underline{\int_R f(x) dx}.$$

Let us denote by  $\overline{\int_R f(x) dx}$  and  $\underline{\int_R f(x) dx}$  the upper and lower Riemann integrals over  $R$ . Then, if we put all observations together, it follows that

$$\begin{aligned} f \text{ is Riemann-integrable.} &\iff \overline{\int_R f(x) dx} = \underline{\int_R f(x) dx} \\ &\iff \int_R \bar{f}(x) dx = \int_R \underline{f}(x) dx \\ &\iff \bar{f}(x) = \underline{f}(x) \text{ a.e. } x \in R \\ &\iff f \text{ is continuous almost everywhere on } R. \end{aligned}$$

If this is the case, we see that two integrals coincide from the calculation above.  $\square$

*Exercise 26.* Calculate the Lebesgue integral  $\int_e^{e^2} \frac{\log x}{x} dx$ .

Now before we go further, let us solve some exercises.

*Exercise 27.*

- (1) Choose functions integrable on  $[0, \infty)$  with respect to the Lebesgue measure among the functions listed below.
- (2) Do the same thing replacing  $[0, \infty)$  with  $[1, \infty)$ .
- (3) Do the same thing replacing  $[0, \infty)$  with  $[0, 1)$ .

$$(1) e^{-x} \quad (2) x^\alpha (\alpha < -1) \quad (3) x^{-1} \quad (4) x^\alpha (-1 < \alpha < 0) \quad (5) 1 \quad (6) x^\alpha (\alpha > 0)$$

$$(7) xe^{-x} \quad (8) x^4 \sin x \quad (9) x^k e^{-\sqrt{x}} \quad (10) \frac{e^{-x}}{x-1} \quad (11) \frac{\log(x+1)^{\frac{1}{2}}}{x^{\frac{4}{3}}} \quad (12) \frac{e^{-x^2}}{\sqrt{|x-2|}}$$

*Exercise 28.* Find the necessary and sufficient condition for  $a, b \in \mathbb{R}$  to satisfy

$$(3.93) \quad \int_{\mathbb{R}^d} \frac{|x|^a}{1+|x|^b} dx < \infty.$$

### 3.6. Product measures and Fubini's theorem.

With the definitions of the integrations complete, our concern now goes to the change of the order of integrations. It is still important that our theory will work if the function is integrable or positive. We need to consider two measure spaces. Especially, we have to define their product. This will be a long process. As an intermediate step, we consider an outer measure  $(\mu \otimes \nu)^*$ .

**Definition 3.63**  $((\mu \otimes \nu)^*)$ . Let  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  be a couple of measure spaces. Then define an outer measure  $(\mu \otimes \nu)^*$  on  $X \times Y$  by

$$(\mu \otimes \nu)^*(\emptyset) = 0$$

and

$$(\mu \otimes \nu)^*(A) = \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : A \subset \bigcup_{j=1}^{\infty} E_j \times F_j \right\}, \quad A \in 2^{X \times Y}.$$

Note that the infimum runs over all the countable measurable coverings of  $A$  of the form  $E \times F$  with  $E \in \mathcal{M}$  and  $F \in \mathcal{N}$ . Let us say that a covering  $\{E_j^{(1)} \times F_j^{(1)}\}_{j \in \mathbb{N}}$  of  $A$  is finer than a covering  $\{E_j^{(2)} \times F_j^{(2)}\}_{j \in \mathbb{N}}$  of  $A$ , if  $\{E_j^{(1)} \times F_j^{(1)}\}_{j \in \mathbb{N}}$  is obtained by partitioning each elements in  $\{E_j^{(2)} \times F_j^{(2)}\}_{j \in \mathbb{N}}$ . The finer the covering is, the closer to the actual value of  $(\mu \otimes \nu)^*(A)$  the value  $\sum_{j=1}^{\infty} \mu(E_j) \nu(F_j)$  is.

**Theorem 3.64.** *The set of the form  $E \times F$  with  $E \in \mathcal{M}$  and  $F \in \mathcal{N}$  is  $(\mu \otimes \nu)^*$ -measurable and  $(\mu \otimes \nu)^*(E \times F) = \mu(E) \times \nu(F)$ .*

*Proof.* We go back to the definition of measurability. Fix  $G \in 2^{X \times Y}$ . Then we have

$$\begin{aligned} & (\mu \otimes \nu)^*((E \times F) \cap G) + (\mu \otimes \nu)^*((E \times F)^c \cap G) \\ &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : (E \times F) \cap G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \right\} \\ & \quad + \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : (E \times F)^c \cap G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \right\} \\ &:= \text{I} + \text{II}. \end{aligned}$$

Observe that  $\text{II} = \min(\text{II}_1, \text{II}_2, \text{II}_3)$ , where

$$\begin{aligned} \text{II}_1 &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : (E \times F^c) \cap G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \subset E \times F^c \right\} \\ \text{II}_2 &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : (E^c \times F) \cap G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \subset E^c \times F \right\} \\ \text{II}_3 &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : (E^c \times F^c) \cap G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \subset E^c \times F^c \right\} \end{aligned}$$

By the remark just below Definition 3.63 we obtain

$$\begin{aligned} (\mu \otimes \nu)^*((E \times F) \cap G) + (\mu \otimes \nu)^*((E \times F)^c \cap G) &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : G \subset \bigcup_{j=1}^{\infty} E_j \times F_j \right\} \\ &= (\mu \otimes \nu)^*(G). \end{aligned}$$

As a result we have established the measurability.

Now let us show the equality, assuming that  $E$  and  $F$  are both measurable. Taking into account the overlapping of the covering of  $E \times F$ , we obtain

$$\begin{aligned} & (\mu \otimes \nu)^*(E \times F) \\ &= \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j) : E \times F = \bigcup_{j=1}^{\infty} E_j \times F_j, (E_j)_{j=1}^{\infty} \subset \mathcal{M}, (F_j)_{j=1}^{\infty} \subset \mathcal{N} \right\} \\ &\leq \mu(E) \times \nu(F). \end{aligned}$$

Given covering  $E \times F = \bigcup_{j=1}^{\infty} E_j \times F_j$ , we can arrange that it is a disjoint covering. Consequently

we have only to prove that if  $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$  and  $\{F_j\}_{j=1}^{\infty} \subset \mathcal{N}$  satisfy  $E \times F = \bigcup_{j=1}^{\infty} E_j \times F_j$ , then

$$(3.94) \quad \mu(E) \nu(F) = \sum_{j=1}^{\infty} \mu(E_j) \nu(F_j).$$

This can be achieved as follows: Set

$$(3.95) \quad S_k(x, y) = \sum_{j=1}^k \chi_{E_j \times F_j}(x, y) \quad (x, y \in X \times Y).$$

Then by the definition of the integral we have

$$(3.96) \quad \int_X \left( \int_Y S_k(x, y) d\nu(y) \right) d\mu(x) = \sum_{j=1}^k \mu(E_j) \nu(F_j).$$

If  $k \rightarrow \infty$ , then

$$(3.97) \quad \int_Y S_k(x, y) d\nu(y) \uparrow \int_Y \chi_{E \times F}(x, y) d\nu(y)$$

by virtue of the monotone convergence theorem. Using the monotone convergence theorem once more, we have

$$(3.98) \quad \int_X \left( \int_Y S_k(x, y) d\nu(y) \right) d\mu(x) \uparrow \int_X \left( \int_Y \chi_{E \times F}(x, y) d\nu(y) \right) d\mu(x) = \mu(E) \nu(F).$$

If we let  $k$  tend to  $\infty$ , (3.94) is established.  $\square$

The product of the measure and the  $\sigma$ -field is readily defined now.

**Definition 3.65** (Tensor product of measure spaces). Denote by  $\mu \otimes \nu$  the restriction of  $(\mu \otimes \nu)^*$  to the measurable sets. Define  $\mathcal{M} \otimes \mathcal{N}$  as the smallest algebra generated by

$$(3.99) \quad \{E \times F : E \in \mathcal{M}, F \in \mathcal{N}\}.$$

It is worth mentioning that the following formula holds.

**Corollary 3.66.** *Suppose that  $A = E \times F$  with  $E \in \mathcal{M}$  and  $F \in \mathcal{N}$ . Then*

$$(3.100) \quad \int_X \left( \int_Y \chi_A(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X \chi_A(x, y) d\mu(x) \right) d\nu(y) = \mu \otimes \nu(A).$$

Let  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  be a couple of measure spaces.

Recall that a measure space  $(X, \mathcal{M}, \mu)$  is said to be  $\sigma$ -finite, if  $X$  is partitioned into a countable subsets of finite  $\mu$ -measure (see Definition 3.10).

Until the end of this section we assume that  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  are  $\sigma$ -finite.

**Proposition 3.67.** *Let  $A \in \mathcal{M} \otimes \mathcal{N}$ . Then  $A_x \in \mathcal{N}$  and  $A_y \in \mathcal{M}$ . Furthermore we have*

$$(3.101) \quad \mu \otimes \nu(A) = \int_X \nu(A_x) d\mu(x) = \int_Y \mu(A_y) d\nu(y).$$

Here we have defined

$$(3.102) \quad A_x := \{y \in Y : (x, y) \in A\}, \quad x \in X, \quad A_y := \{x \in X : (x, y) \in A\}, \quad y \in Y.$$

*Proof.* This is just a restatement of (3.100) when  $A = E \times F$  for some  $E \in \mathcal{M}$  and  $F \in \mathcal{N}$ . A passage to the general case can be achieved by the  $\pi$ - $\lambda$  system (Theorem 3.12).  $\square$

The next theorem will be a prototype of Fubini's theorem.

**Theorem 3.68** (Fubini's theorem). *Suppose that  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  are  $\sigma$ -finite. Then we have*

$$(3.103) \quad \iint_{X \times Y} f(x, y) d\mu \otimes \nu(x, y) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y),$$

provided  $f \in (\mathcal{M} \otimes \mathcal{N})_+$  or  $f \in L^1(\mathcal{M} \otimes \mathcal{N})$ .

*Proof.* Suppose that  $f \in L^1(\mathcal{M} \otimes \mathcal{N})$ . Then we can decompose

$$(3.104) \quad f = \operatorname{Re}(f)_+ - \operatorname{Re}(f)_- + i \cdot \operatorname{Im}(f)_+ - i \cdot \operatorname{Im}(f)_-.$$

Thus, by linearity we have only to prove (3.103) only when  $f \in (\mathcal{M} \otimes \mathcal{N})_+$ . If  $f \in (\mathcal{M} \otimes \mathcal{N})_+$ , then we can approximate it with a sequence of positive simple functions from above. Again by linearity we can assume  $f = \chi_A$ , where  $A$  is a  $\mu \otimes \nu$ -measurable set. If  $f = \chi_A$ , then (3.103) is (3.100) itself and the proof is now complete.  $\square$

*Exercise 29.* The author hit upon this exercise after the lecture of K. Yoneda [510].

- (1) Suppose that  $h : \mathbb{R} \rightarrow \mathbb{C}$  is a continuous even function such that  $\lim_{|x| \rightarrow \infty} h(x) = A$  does exists. Then show that

$$(3.105) \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T h(t) dt = A.$$

- (2) Under the same assumption, show that

$$(3.106) \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left( \log \frac{T}{|t|} \right) h(t) dt = A.$$

*Hint:* Justify that we can assume  $h$  even. Then use

$$\begin{aligned} \frac{1}{2T} \int_{-T}^T \left( \log \frac{T}{|t|} \right) h(t) dt &= \frac{1}{T} \int_0^T \left( \log \frac{T}{t} \right) h(t) dt \\ &= \frac{1}{T} \int_0^T \int_t^T \frac{ds}{s} h(t) dt \\ &= \frac{1}{T} \int_0^T \left( \frac{1}{s} \int_0^s h(t) dt \right) ds, \end{aligned}$$

where for the last equality we have used the Fubini theorem.

## 4. LEBESGUE SPACES

4.1.  $L^p(\mu)$ -spaces.

As is the case with the people who learned Riemannian integration theory, sometimes we are not interested in the exact value of integrals. Indeed, in many cases, it is again next to impossible to calculate precisely the value of integral. Instead, we are mainly interested in the size of functions. One of the simplest way to do so is to use Lebesgue spaces.

**Definition 4.1** (Lebesgue space  $L^p(\mu)$ ). Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $0 < p \leq \infty$ .

(1) The  $L^p(\mu)$ -(quasi-)norm of a measurable function  $f$  is given by

$$\|f\|_p := \|f\|_{L^p(\mu)} := \left( \int_X |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} = \left( \int_X |f(x)|^p d\mu(x) \right), \quad p < \infty$$

$$\|f\|_\infty := \|f\|_{L^\infty(\mu)} := \sup\{\lambda > 0 : |f(x)| \leq \lambda \text{ for } \mu\text{-a.e. } x \in X\}, \quad p = \infty.$$

(2) Define

$$L^p(\mu) := \{f : X \rightarrow \overline{\mathbb{K}} : f \text{ is measurable and } \|f\|_p < \infty\} / \sim.$$

Here the equivalence relation  $\sim$  is defined by

$$(4.1) \quad f \sim g \iff f = g \text{ a.e.}$$

and below omit this equivalence in defining function spaces.

For example,  $f$  stands for the representative of the class belonging to  $f$  as well as a measurable function  $f$ .

*Exercise 30.* The Lebesgue space  $L^p(\mu)$  with  $0 < p < 1$  is not a normed space. What is the property of the normed space that fails?

*Exercise 31.* Let  $(X, \mathcal{M}, \mu)$  be a finite measure space. Show that

$$(4.2) \quad \lim_{p \downarrow 0} \mu(X)^{-\frac{1}{p}} \|f\|_p = \exp\left(\frac{1}{\mu(X)} \int_X \log |f(x)| d\mu(x)\right), \quad \lim_{p \rightarrow \infty} \|f\|_{L^p(\mu)} = \|f\|_{L^\infty(\mu)}.$$

*Exercise 32.* Show the following scaling law in Lebesgue spaces  $L^p(dx)$ :

$$\|f(t \cdot)\|_p = t^{-n/p} \|f\|_p.$$

Now we state important integral inequalities.

**Theorem 4.2** (Minkowski's inequality). Let  $(X, \mathcal{B}, \mu)$  be a measure space. Let  $1 \leq p \leq \infty$ . Then

$$(4.3) \quad \|f + g\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)} + \|g\|_{L^p(\mu)}$$

for all measurable functions  $f, g : X \rightarrow \mathbb{C}$ .

Hence from this inequality we see that  $L^p(\mu)$  is a normed space.

*Proof.* If  $p = \infty$ , then this is just reduced to a triangle inequality for  $\mathbb{K}$ . Let us use the following convex inequality:

$$(4.4) \quad (\theta a + (1 - \theta)b)^p \leq \theta a^p + (1 - \theta)b^p$$

for  $0 < \theta < 1$  and  $a, b \geq 0$ , which can be readily obtained by an elementary calculation. Thus, it follows that

$$\begin{aligned} \left( \frac{\|f + g\|_p}{\|f\|_p + \|g\|_p} \right)^p &= \int_X \left( \frac{|f(x) + g(x)|}{\|f\|_p + \|g\|_p} \right)^p d\mu(x) \\ &\leq \frac{\|f\|_p}{\|f\|_p + \|g\|_p} \int_X \frac{|f(x)|^p}{\|f\|_p^p} d\mu(x) + \frac{\|g\|_p}{\|f\|_p + \|g\|_p} \int_X \frac{|g(x)|^p}{\|g\|_p^p} d\mu(x) \\ &= 1. \end{aligned}$$

As a result, we have  $\|f + g\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)} + \|g\|_{L^p(\mu)}$ . □

*Exercise 33.* Show that (4.4) holds with equality if and only if  $a = b$ .

We still have a substitute for (4.3) for  $0 < p \leq 1$ .

**Proposition 4.3** (*p*-convexity). *Let  $(X, \mathcal{B}, \mu)$  be a measure space. Let  $0 < p \leq 1$ . Then we have*

$$(4.5) \quad \|f + g\|_{L^p(\mu)}^p \leq \|f\|_{L^p(\mu)}^p + \|g\|_{L^p(\mu)}^p$$

for all measurable functions  $f, g : X \rightarrow \mathbb{C}$ .

*Exercise 34.* Let  $0 < p \leq 1$ .

- (1) Show that  $(a + b)^p \leq a^p + b^p$  for all  $a, b > 0$ .
- (2) Prove (4.5).

Let  $1 \leq p \leq \infty$ . Then define  $p' = \frac{p}{p-1}$ , if  $1 \leq p < \infty$ . If  $p = \infty$ , then define  $p' = 1$ . Note that

$$(4.6) \quad \frac{1}{p} + \frac{1}{p'} = 1,$$

whenever  $1 \leq p \leq \infty$ . Therefore, we have  $(p')' = p$ .

**Theorem 4.4** (Hölder's inequality). *Let  $(X, \mathcal{B}, \mu)$  be a measure space. Suppose that  $f$  and  $g$  are positive  $\mu$ -measurable functions on  $X$ . Then*

$$(4.7) \quad \int_X f(x)g(x) d\mu(x) \leq \left( \int_X f(x)^p d\mu(x) \right)^{\frac{1}{p}} \left( \int_X g(x)^{p'} d\mu(x) \right)^{\frac{1}{p'}}.$$

*Proof.* We make use of the following elementary inequality;  $ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$ . for all  $a, b > 0$  and  $p \in (1, \infty)$ . First, parameterize the above inequality with  $\theta$ ;  $ab \leq \frac{\theta^p}{p}a^p + \frac{\theta^{-p'}}{p'}b^{p'}$ . Let  $a = f(x)$  and  $b = g(x)$  and integrate the above inequality over  $X$ . Then we have

$$(4.8) \quad \int_X f(x)g(x) d\mu(x) \leq \frac{\theta^p}{p} \int_X f(x)^p d\mu + \frac{\theta^{-p'}}{p'} \int_X g(x)^{p'} d\mu.$$

Choose  $\theta$  so that it minimizes the right-hand side of the above inequality. Then we obtain the desired result. □

*Exercise 35.* Let  $f(\theta) := \frac{\theta^p}{p}A + \frac{\theta^{-p'}}{p'}B$  with  $A, B > 0$ . Find the minimum of  $f : [0, \infty) \rightarrow [0, \infty)$ .

*Exercise 36.* Let  $1 \leq p \leq \infty$ . Let  $f : X \rightarrow \mathbb{R}$  be a measurable function. Show that  $f \in L^p(\mu)$ , whenever  $f \cdot g \in L^1(\mu)$  for all  $g \in L^{p'}(\mu)$ .

In fact as the next theorem shows we have equality in (4.7).

**Theorem 4.5** (Duality  $L^p(\mu)$ - $L^{p'}(\mu)$ ). Assume that  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space. Let  $f$  be a function that is integrable on any set of finite measure. Then, for all  $1 \leq p \leq \infty$ , we have

$$\|f\|_{L^p(\mu)} = \sup \left\{ \left| \int_X f(x)g(x) d\mu(x) \right| : g \in L^\infty(\mu) \cap L^{p'}(\mu), \|g\|_{L^{p'}(\mu)} = 1, \mu\{g \neq 0\} < \infty \right\}.$$

*Proof.* Let  $g \in L^\infty(\mu) \cap L^{p'}(\mu)$  be such that  $\|g\|_{L^{p'}(\mu)} = 1$  and that  $\mu\{g \neq 0\} < \infty$ . It is easy to show that

$$(4.9) \quad \left| \int_X f(x) \cdot g(x) d\mu(x) \right| \leq \|f\|_{L^p(\mu)}$$

by using the Hölder inequality. Since  $X$  is assumed  $\sigma$ -finite, there exists  $\{f_j\}_{j \in \mathbb{N}} \subset L^\infty(\mu)$  such that

$$(4.10) \quad |f_j| \leq |f|, \lim_{j \rightarrow \infty} |f_j| = |f|, \mu\{f_j \neq 0\} < \infty \text{ for all } j \in \mathbb{N}.$$

Therefore, we may assume  $f \in L^\infty(\mu)$  with  $\mu\{f \neq 0\} < \infty$ . Let  $g := \overline{\text{sgn}(f)}|f|^{p-1}$ . Then we have

$$(4.11) \quad \|g\|_{L^{p'}(\mu)} = \|f\|_{L^p(\mu)}^{p-1}, \int_X f(x)g(x) d\mu(x) = \int_X |f(x)|^p d\mu(x) = \|f\|_{L^p(\mu)}^p.$$

Thus, if we set  $h = \frac{g}{\|f\|_{L^p(\mu)}^{p-1}}$ , then  $h$  attains the supremum in question.  $\square$

*Exercise 37.* Assume that  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space. Let  $f \in L^p(\mu)$  with  $1 \leq p < \infty$ . Then show that

$$(4.12) \quad \|f\|_{L^p(\mu)} = \max \left\{ \left| \int_X f(x)g(x) d\mu(x) \right| : g \in L^\infty(\mu), \|g\|_{L^{p'}(\mu)} = 1 \right\}.$$

**Theorem 4.6.** Let  $0 < p < \infty$ . Then the space  $L^p(\mu)$  is complete in the following sense: Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence in  $L^p(\mu)$  satisfying

$$(4.13) \quad \lim_{K \rightarrow \infty} \left( \sup_{j, k \geq K} \|f_j - f_k\|_{L^p(\mu)} \right) = 0.$$

Then there exists  $f \in L^p(\mu)$  so that

$$(4.14) \quad \lim_{j \rightarrow \infty} \|f - f_j\|_{L^p(\mu)} = 0.$$

*Proof.* We assume that  $0 < p \leq 1$ . The result for the case  $1 < p \leq \infty$  is well-known and left as Exercise 38 for the readers, the proof being similar to the case when  $0 < p \leq 1$ . The proof consists of two parts. We may also assume  $f$  is real-valued.

Construction of the limit  $f$  By assumption we can take a subsequence  $\{f_{j_k}\}_{k \in \mathbb{N}}$  so that

$$(4.15) \quad \|f_{j_{k+1}} - f_{j_k}\|_{L^p(\mu)} \leq \frac{1}{k!}.$$

Set  $g_1 := f_{j_1}$  and  $g_k := f_{j_{k+1}} - f_{j_k}$  for  $k \geq 2$ .

Then a repeated application of (4.5) gives

$$(4.16) \quad \| |g_1| + |g_2| + \dots + |g_k| \|_{L^p(\mu)}^p \leq \sum_{l=1}^k \|g_l\|_{L^p(\mu)}^p \leq \sum_{l=1}^k (k!)^{-p} < \infty.$$



By using the monotone convergence theorem we have

$$(4.17) \quad \| |g_1| + |g_2| + \dots \|_{L^p(\mu)} < \infty.$$

(4.17) gives us

$$(4.18) \quad |g_1(x)| + |g_2(x)| + \dots$$

is finite for  $\mu$ -a.e.  $x \in X$ . That is,

$$(4.19) \quad g_1(x) + g_2(x) + \dots$$

converges absolutely for  $\mu$ -a.e.  $x \in X$ . Set

$$(4.20) \quad f(x) := \limsup_{k \rightarrow \infty} \sum_{l=1}^k g_l(x).$$

Then proceeding in the same way as before, we have

$$(4.21) \quad \|f - g_1 - g_2 - \dots - g_k\|_{L^p(\mu)}^p \leq \sum_{l=k+1}^{\infty} (k!)^{-p} \rightarrow 0$$

as  $k \rightarrow \infty$ . In the next step we shall prove that  $f$  is the limit of the Cauchy sequence  $\{f_j\}_{j \in \mathbb{N}}$ .

$\{f_j\}_{j \in \mathbb{N}}$  tends to  $f$  Let  $\varepsilon > 0$  be fixed. Then there exists  $N$  so that  $m, n \geq N$  implies

$$(4.22) \quad \|f_m - f_n\|_{L^p(\mu)} \leq \varepsilon.$$

Let  $k \geq N$ . Then (4.22) gives us

$$(4.23) \quad \|f_m - g_1 - g_2 - \dots - g_k\|_{L^p(\mu)} = \|f_m - f_k\|_{L^p(\mu)} \leq \varepsilon.$$

By the Fatou lemma, we have

$$\|f_m - f\|_{L^p(\mu)} = \left\| \liminf_{k \rightarrow \infty} \left( f_m - \sum_{l=1}^k g_l \right) \right\|_{L^p(\mu)} \leq \liminf_{k \rightarrow \infty} \left\| f_m - \sum_{l=1}^k g_l \right\|_{L^p(\mu)}$$

Inserting the definition of the  $g_l$ , we obtain

$$\|f_m - f\|_{L^p(\mu)} \leq \liminf_{k \rightarrow \infty} \|f_m - f_k\|_{L^p(\mu)} \leq \varepsilon.$$

Therefore  $\{f_j\}_{j \in \mathbb{N}}$  converges to  $f$ . □

*Exercise 38.* Supply the proof of the above theorem with  $1 < p \leq \infty$ .

*Exercise 39.* What happens if we consider the  $L^p(dx)$  space in the frame of the Riemannian integral? Show that there exists a sequence of bounded functions  $\{f_j\}_{j \in \mathbb{N}}$  on  $[0, 1]$  with the following properties.

- (1)  $\int_0^1 |f_j(x) - f_k(x)| dx = 0$  for all  $j, k \in \mathbb{N}$ .
- (2)  $\lim_{j \rightarrow \infty} f_j(x) = f(x)$  exists for all  $x \in [0, 1]$ .
- (3)  $f$  is discontinuous everywhere on  $[0, 1]$ .

*Exercise 40.* Let  $\mathcal{B}$  denote the set of all measurable subsets in a measure space  $(X, \mathcal{B}, \mu)$ . If  $A$  and  $B$  have finite  $\mu$ -measure and  $\mu(A \ominus B) = 0$ , then let us say that  $A$  and  $B$  are equivalent. Denote by  $\mathcal{C}$  the equivalence class.  $[A]$  denotes the class of equivalence to which  $A \in \mathcal{B}$  of finite measure belongs.

- (1) Show that  $d([A], [B]) := \mu(A \ominus B)$  is a well-defined distance function.
- (2) Show that  $(\mathcal{C}, d)$  is complete.

## 4.2. Integration formulae.

In this section we prove some elementary calculus formulae. When we consider Lebesgue integral of the functions, we have to cut the space  $X$  according to the value of the functions. So it seems to count to consider  $\{x \in X : |f(x)| > \lambda\}$  instead of the precise structure of the functions themselves. This observation leads us to the following definition.

**Definition 4.7** (Distribution functions). Let  $(X, \mu, \mathcal{B})$  be a  $\sigma$ -finite measure space. Define, for a measurable function  $f$ ,

$$(4.24) \quad \varphi_f(\lambda)\mu\{|f| > \lambda\} = \mu\{x \in X : |f(x)| > \lambda\} \quad (\lambda > 0).$$

This function is said to be the distribution of  $f$ .

*Exercise 41.* In this exercise we consider a property and an example of distribution functions.

- (1) Show that  $\lambda \in [0, \infty) \mapsto \mu\{|f| > \lambda\} \in [0, \infty)$  is a right-continuous decreasing function.
- (2) Consider  $\log : (0, \infty) \rightarrow \mathbb{R}$ . Then prove that its distribution function is  $e^{-\lambda}$ .

The following inequality is elementary, which we frequently use without notice.

**Theorem 4.8** (Chebychev inequality). Let  $f \in L^p(\mu)$  and  $\lambda > 0$ . Then

$$(4.25) \quad \mu\{|f| > \lambda\} \leq \frac{1}{\lambda^p} \int_X |f(x)|^p d\mu(x).$$

*This inequality is known as the Chebychev inequality.*

*Proof.* We have only to re-examine the arguments in Chapter 1. The proof is obtained by expressing the left-hand side in terms of integral:

$$\mu\{|f| > \lambda\} = \int_X \chi_{\{|f| > \lambda\}}(x) d\mu(x) \leq \int_X \chi_{\{|f| > \lambda\}}(x) \cdot \frac{|f(x)|^p}{\lambda^p} d\mu(x) \leq \frac{1}{\lambda^p} \int_X |f(x)|^p d\mu(x).$$

This is the desired result. □

We use the following formula throughout this book.

**Theorem 4.9** (Distribution formula). Suppose that  $a$  is an increasing function on  $[0, \infty)$  with  $a(0) = 0$  and continuous derivative. Then for a  $\mathbb{C}$ -valued  $\mu$ -measurable function  $f$ , we have

$$(4.26) \quad \int_X a(|f(x)|) d\mu(x) = \int_0^\infty a'(\lambda)\mu\{|f| > \lambda\} d\lambda.$$

*Proof.* Our viewpoint is to insert the following trivial equality:

$$(4.27) \quad a(|f(x)|) = \int_0^{|f(x)|} a'(\lambda) d\lambda.$$

Thanks to (4.27), we can proceed further. By interchanging the order of integrations, we obtain

$$\begin{aligned}
\int_X a(|f(x)|) d\mu(x) &= \int_X \left( \int_0^{|f(x)|} a'(\lambda) d\lambda \right) d\mu(x) \\
&= \int_X \left( \int_0^\infty \chi_{\{|f(x)| > \lambda\}}(x, \lambda) a'(\lambda) d\lambda \right) d\mu(x) \\
&= \int_0^\infty \left( \int_X \chi_{\{|f(x)| > \lambda\}}(x, \lambda) a'(\lambda) d\mu(x) \right) d\lambda \\
&= \int_0^\infty a'(\lambda) \left( \int_X \chi_{\{|f(x)| > \lambda\}}(x, \lambda) d\mu(x) \right) d\lambda \\
&= \int_0^\infty a'(\lambda) \mu\{|f| > \lambda\} d\lambda \\
&= \text{R.H.S.}
\end{aligned}$$

Thus, the proof is now complete.  $\square$

To conclude this section we investigate the rearrangement of a function  $f$ . This is a tool with which to realize functions on  $X$  in the interval  $(0, \infty)$ .

**Definition 4.10** (Distribution functions). Let  $(X, \mathcal{B}, \mu)$  be a measure space and let  $f$  be a  $\mu$ -measurable function. Write  $\mu\{|f| > s\} = \mu\{x \in X : |f(x)| > s\}$ , the distribution function of  $f$ .

- (1)  $f^*(t) := \inf\{s > 0 : \mu\{|f| > s\} \leq t\}$  for  $t > 0$ .
- (2)  $f^{**}(t) := \frac{1}{t} \int_0^t f^*(s) ds$  for  $t > 0$ .

**Theorem 4.11.** Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $f$  a  $\mu$ -measurable function. Then  $f^*$  is a right-continuous decreasing function.

*Proof.* It is trivial that  $f^*$  is decreasing. To prove that  $f^*$  is right-continuous, we choose a sequence  $\{t_j\}_{j \in \mathbb{N}}$  decreasing to  $t$ . Then we have

$$(4.28) \quad \{s > 0 : \mu\{|f| > s\} \leq t\} = \bigcap_{j=1}^{\infty} \{s > 0 : \mu\{|f| > s\} \leq t_j\}.$$

Furthermore note that there exists  $s_0$  and an increasing sequence  $\{s_j\}_{j \in \mathbb{N}}$  such that

$$(4.29) \quad \{s > 0 : \mu\{|f| > s\} \leq t\} = [s_0, \infty), \quad \{s > 0 : \mu\{|f| > s\} \leq t_j\} = [s_j, \infty).$$

Therefore,

$$(4.30) \quad f^*(t) = s_0 = \lim_{j \rightarrow \infty} s_j = \lim_{j \rightarrow \infty} f^*(t_j),$$

which shows that  $f^*$  is right-continuous.  $\square$

**Theorem 4.12.** Suppose that  $f : X \rightarrow \mathbb{C}$  is a measurable function. Then we have

$$(4.31) \quad \|f\|_q^q = \int_0^\infty f^*(t)^q dt$$

for all  $0 < q < \infty$ .

*Proof.* We write out the right-hand side in terms of the distribution function of  $u^*$ .

$$\int_0^\infty f^*(t)^q dt = \int_0^\infty q \lambda^{q-1} |\{t \geq 0 : f^*(t) \geq \lambda\}| d\lambda = \int_0^\infty q \lambda^{q-1} |\{t \geq 0 : f^*(t) > \lambda\}| d\lambda.$$

Since  $\{t \geq 0 : f^*(t) > \lambda\} \subset \{t \geq 0 : \mu\{|f| > \lambda\} > t\} \subset \{t \geq 0 : f^*(t) \geq \lambda\}$ , we obtain

$$\int_0^\infty q \lambda^{q-1} |\{t \geq 0 : \mu\{|f| > \lambda\} > t\}| dt = \int_0^\infty q \lambda^{q-1} \mu\{|f| > \lambda\} d\lambda.$$

Therefore, it follows that

$$(4.32) \quad \int_0^\infty f^*(t)^q dt = \int_X |f(x)|^q d\mu(x).$$

This is the desired result.  $\square$

### 4.3. Convergence in measure.

One of the big advantages of Lebesgue integral is that measurability is closed under taking the countable supremum. This fact gives us a hope that we can do a lot about the operation of taking limit. Actually, in Riemannian integral, we could not change the order of the integration and the limit unless we check the uniform integrability. Even when we can check it, we have to place ourselves in the setting of bounded intervals. In Lebesgue integral, this strict condition can be relaxed to a large extent. Also, by using the measure from which we started, we can consider many types of convergences. Let  $(X, \mathcal{B}, \mu)$  be a measure space. Now we shall begin with convergence of functions in measure.

**Definition 4.13** (Convergence in measure). Let  $(X, \mathcal{B}, \mu)$  be a measure space. A sequence of measurable functions  $\{f_j\}_{j \in \mathbb{N}}$  converges to  $f$  in measure, if

$$(4.33) \quad \lim_{j \rightarrow \infty} \mu\{|f_j - f| > \varepsilon\} = 0$$

for all  $\varepsilon > 0$ .

Below we present an example of convergence in measure.

**Proposition 4.14.** Assume that  $\mu$  is finite. If a sequence of measurable functions  $\{f_j\}_{j \in \mathbb{N}}$  converges  $\mu$ -a.e., then it converges to  $f$  also in measure.

*Proof.* Let  $\varepsilon > 0$  be arbitrary. Then for all  $m \in \mathbb{N}$ , we have

$$(4.34) \quad \mu\{|f - f_j| > \varepsilon\} \leq \mu\left(\bigcup_{l=m}^{\infty} \bigcap_{k \geq l} \{|f_k - f_j| > \varepsilon\}\right) \leq \mu\left(\bigcup_{l=m}^{\infty} \{|f_l - f_j| > \varepsilon\}\right).$$

Assuming  $\mu$  finite, we see  $\lim_{m \rightarrow \infty} \left(\bigcup_{l=m}^{\infty} \{|f_l - f_j| > \varepsilon\}\right)$  differs from the empty set only by a set of measure zero. We obtain the desired assertion.  $\square$

**Proposition 4.15.** If  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence in  $L^p(\mu)$  that is convergent to  $f$  in the  $L^p(\mu)$ -topology, then  $\{f_j\}_{j \in \mathbb{N}}$  converges to  $f$  in measure.

*Proof.* This is immediate from Chebychev's inequality (Theorem 4.8).  $\square$

**Example 4.16.** It is important to keep in mind that the converse of the above proposition is false. We exhibit two examples on the measure space  $([0, 1], \mathcal{B}([0, 1]), dt|_{[0, 1]})$ .

- (1) Suppose that  $j \in \mathbb{N}$  satisfies  $\frac{k(k+1)}{2} \leq j < \frac{(k+1)(k+2)}{2}$  for some  $k \in \mathbb{N}$ . Let  $l(j) = j - \frac{k(k+1)}{2}$ . Then we define  $f_j = \chi_{[\frac{l(j)}{k+1}, \frac{l(j)+1}{k+1}]}$ . Then  $\{f_j\}_{j \in \mathbb{N}}$  converges to 0 in measure while  $\{f_j(t)\}_{j \in \mathbb{N}}$  never converges for any  $t \in [0, 1)$ .
- (2) Let  $f_j(t) = j! \chi_{(0, 2^{-j})}(t)$  for  $j \in \mathbb{N}$ . Then  $\{f_j\}_{j \in \mathbb{N}}$  converges in measure to 0 but never converges in the  $L^p(\mathbb{R})$ -topology for all  $0 < p \leq \infty$ .

To deal with the almost everywhere convergence, it is sometimes convenient to apply the Borel-Cantelli lemma. Later it will serve to extend the range of operators.

**Lemma 4.17** (Borel-Cantelli). *Assume  $\{A_j\}_{j \in \mathbb{N}}$  is a sequence of measurable sets satisfying*

$$(4.35) \quad \sum_{j=1}^{\infty} \mu(A_j) < \infty.$$

Then  $\mu \left( \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k \right) = 0$ .

*Proof.* Note that  $\mu \left( \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k \right) = \lim_{j \rightarrow \infty} \mu \left( \bigcup_{k=j}^{\infty} A_k \right)$ , since  $\mu \left( \bigcup_{k=1}^{\infty} A_k \right) \leq \sum_{k=1}^{\infty} \mu(A_k) < \infty$  by virtue of the subadditivity.

By the subadditivity of  $\mu$  together with (4.35) again we have

$$(4.36) \quad \mu \left( \bigcup_{k=j}^{\infty} A_k \right) \leq \sum_{k=j}^{\infty} \mu(A_k) \rightarrow 0$$

as  $j \rightarrow \infty$ . Thus, the result is immediate. □

The following theorem asserts that any Cauchy sequence has a limit in this topology.

**Theorem 4.18.** *Suppose that a sequence of measurable functions  $\{f_j\}_{j \in \mathbb{N}}$  satisfies*

$$(4.37) \quad \lim_{j, k \rightarrow \infty} \mu \{ |f_j - f_k| > \varepsilon \} = 0$$

for all  $\varepsilon > 0$ . Then  $\{f_j\}_{j \in \mathbb{N}}$  converges to some  $f$  in measure.

*Proof.* We may suppose that  $f$  is real-valued. Choose an increasing sequence  $N_1, N_2, \dots, N_l, \dots$  so that

$$(4.38) \quad \mu \{ |f_j - f_k| > 2^{-l} \} \leq 2^{-l},$$

whenever  $j$  and  $k$  are larger than  $N_l$ . Let

$$(4.39) \quad A := \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} \{ |f_j - f_k| > 2^{-l} \text{ for all } j, k \text{ with } j, k \geq N_l \}.$$

Then  $\mu(A) = 0$  and  $\lim_{l \rightarrow \infty} f_{N_l}(x)$  exists outside  $A$ . Set  $f(x) := \limsup_{l \rightarrow \infty} f_{N_l}(x)$  on  $X$ . Then on  $A$  we have

$$\{ |f_j - f| > 2^{-l} \} \subset \bigcup_{m=l}^{\infty} \left( \bigcap_{k=m}^{\infty} \{ |f_j - f_{N_k}| > 2^{-m} \} \right) \subset \bigcup_{m=l}^{\infty} \{ |f_j - f_{N_m}| > 2^{-m} \},$$

provided  $l \leq m$ . Therefore, it follows that  $\mu \{ |f_j - f| > 2^{-l} \} \leq 2^{1-l}$ . Thus,  $\{f_j\}_{j \in \mathbb{N}}$  converges to  $f$  in measure. □

Although the convergence in measure does not ensure almost everywhere convergence, it is important to keep in mind that this is the case, if we pass to a subsequence.

**Theorem 4.19.** *Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence of  $\mu$ -measurable functions converging in measure. Then one can choose a subsequence converging  $\mu$ -almost everywhere.*

*Proof.* Indeed,  $\lim_{l \rightarrow \infty} f_{N_l}(x)$  converges  $\mu$ -almost everywhere under the notation above.  $\square$

#### 4.4. Radon Nikodym derivative.

Next we turn to the density of measures. By the “density”, we can envisage the weight defined on underlying spaces. To formulate “weights” more precisely, let us again work on a measure space  $(X, \mathcal{B}, \mu)$ . First, it is easy to see that

$$(4.40) \quad \nu(E) = \int_E f(x) d\mu(x) \quad (E \in \mathcal{B})$$

defines another measure as long as  $f$  is positive and integral. When a measure  $\nu$  has such an expression, then  $\nu$  has a density (function)  $f$  with respect to  $\mu$ . The function  $f$  must be positive in order to guarantee that  $\nu(E) \geq 0$  for all  $E$ . However, as long as  $f$  is integrable,  $\nu$  satisfies the countable additivity defined below. This observation naturally leads us to the following definition.

**Definition 4.20** (Signed measure). A function  $\nu : \mathcal{B} \rightarrow \mathbb{C}$  is said to be countably additive, if it satisfies

$$(4.41) \quad \sum_{j=1}^{\infty} \nu(E_j) = \nu \left( \bigcup_{j=1}^{\infty} E_j \right),$$

whenever  $(E_j)_{j=1}^{\infty} \subset \mathcal{B}$  is disjoint.

An important remark concerning (4.41) is in order, which shows that (4.41) is much stronger than it looks.

**Remark 4.21.** The above series does not depend on the order of the sum. Speaking precisely, we take a bijective mapping  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ . Then

$$(4.42) \quad \sum_{j=1}^{\infty} \nu(E_{\sigma(j)}) = \nu \left( \bigcup_{j=1}^{\infty} E_{\sigma(j)} \right) = \nu \left( \bigcup_{j=1}^{\infty} E_j \right).$$

From this we conclude that the sum above converges absolutely.

For a measurable function  $f$  we can define its modulus  $|f|$ . The same can be said for countably additive functions.

**Definition 4.22** ( $|\nu|$ ). Suppose that  $\nu : \mathcal{B} \rightarrow \mathbb{C}$  is a countably additive function. Set

$$(4.43) \quad |\nu|(A) := \sup \left\{ \sum_{j=1}^{\infty} |\nu(A_j)| : \sum_{j=1}^{\infty} A_j = A, A_j \in \mathcal{B} \right\}, \quad |\nu(\emptyset)| = 0.$$

**Theorem 4.23.** *Let  $\nu : \mathcal{B} \rightarrow \mathbb{C}$  be a countably additive function. Then  $|\nu|$  is a finite measure.*

*Proof.* \$|\nu|\$ is a measure All that is not clear is the countable subadditivity. Suppose that  $A_j, j = 1, 2, \dots$  is a disjoint family in  $\mathcal{B}$ . Then there exists  $B_{j,k} \in \mathcal{B}, j, k \in \mathbb{N}$  such that

$$(4.44) \quad \sum_{k=1}^{\infty} |\nu(B_{j,k})| \geq |\nu|(A_j) - \frac{\varepsilon}{2^j}, \quad \sum_{k=1}^{\infty} B_{j,k} = A_j.$$

Summing this over  $j \in \mathbb{N}$ , we obtain

$$(4.45) \quad \sum_{j,k=1}^{\infty} |\nu(B_{j,k})| \geq \sum_{j=1}^{\infty} |\nu|(A_j) - \varepsilon, \quad \sum_{j,k=1}^{\infty} B_{j,k} = \sum_{j=1}^{\infty} A_j.$$

As a result it follows that

$$(4.46) \quad |\nu|\left(\sum_{j=1}^{\infty} A_j\right) \geq \sum_{j,k=1}^{\infty} |\nu(B_{j,k})| \geq \sum_{j=1}^{\infty} |\nu|(A_j) - \varepsilon.$$

Since  $\varepsilon > 0$  is taken arbitrarily, we obtain

$$(4.47) \quad |\nu|\left(\bigcup_{j=1}^{\infty} A_j\right) \geq \sum_{j=1}^{\infty} |\nu|(A_j).$$

To prove the converse inequality, we let

$$(4.48) \quad \sum_{k=1}^{\infty} B_k = \sum_{j=1}^{\infty} A_j, \quad B_k \in \mathcal{B} \quad k \in \mathbb{N}.$$

Then  $A_j$  can be partitioned into  $\{B_k \cap A_j\}_{k \in \mathbb{N}}$ . Thus, it follows that

$$(4.49) \quad \sum_{k=1}^{\infty} |\nu(B_k \cap A_j)| \leq |\nu|(A_j).$$

Taking this into account, we obtain

$$(4.50) \quad \sum_{k=1}^{\infty} |\nu(B_k)| \leq \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |\nu(B_k \cap A_j)| \leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\nu(B_k \cap A_j)| \leq \sum_{j=1}^{\infty} |\nu|(A_j).$$

Since the partition is arbitrary, we obtain

$$(4.51) \quad |\nu|\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} |\nu|(A_j).$$

Consequently we have proved that  $|\nu|$  is a measure.

\$|\nu|\$ is finite. It is easy to see that  $X$  is not partitioned into a countable sum  $\{A_j\}_{j \in \mathbb{N}}$  such that  $|\nu|(A_j) = \infty$ . Indeed, if there were such a partition, then each  $A_j$  would be partitioned into a subpartition  $\{A_j^{(k)}\}_{k \in \mathbb{N}}$  such that  $\sum_{k=1}^{\infty} |\nu(A_j^{(k)})| > 1$ . Thus,  $\{A_j^{(k)}\}_{j,k \in \mathbb{N}}$  is a partition such

that  $\sum_{j,k=1}^{\infty} |\nu(A_j^{(k)})| = \infty$ . This contradicts to Remark 4.21. Let us assume  $\nu$  is not zero.

From the definition of  $|\nu|(X)$  and the additivity of  $\nu$ , there exists a partition  $\{A_j\}_{j \in \mathbb{N}}$  of  $X$  such that each  $A_j$  is non-empty and that

$$(4.52) \quad \sum_{j=1}^{\infty} |\nu(A_j)| \geq \frac{1}{2} \min(1, |\nu|(X)).$$

Then from the above observation, there exists a non-empty measurable set  $A$  such that  $|\nu|(A)$  is finite.

Let us set

$$(4.53) \quad M := \sup\{|\nu|(A) : |\nu|(A) < \infty\}.$$

From the observation above, we see that the set defining  $M$  is not empty. From the definition of  $M$  there exists an increasing sequence  $\{A_j\}_{j \in \mathbb{N}}$  such that  $|\nu|(A_j) \rightarrow M$  as  $j \rightarrow \infty$ .

Assume that  $M$  is infinity. Then we can assume that  $|\nu|(A_{j+1} \setminus A_j) > 0$ . For each  $j \in \mathbb{N}$  we can take a countable partition  $\{B_j^{(k)}\}_{k \in \mathbb{N}}$  such that

$$(4.54) \quad \sum_{j=1}^{\infty} |\nu(B_j^{(k)})| > \frac{1}{2} |\nu|(A_{j+1} \setminus A_j).$$

Adding the above inequality over  $j \in \mathbb{N}$ , we would have a partition  $\{B_j^{(k)}\}_{j,k \in \mathbb{N}}$  satisfying

$$(4.55) \quad \sum_{j,k=1}^{\infty} |\nu(B_j^{(k)})| \geq \frac{1}{2} \sum_{j=1}^{\infty} |\nu|(A_{j+1} \setminus A_j) = \infty.$$

This is a contradiction.  $\square$

**Theorem 4.24** (Hahn-decomposition). *Let  $\nu$  be a real-valued countably additive function. Then there exists  $P \in \mathcal{B}$  such that  $\nu(\cdot \cap P)$ ,  $-\nu(\cdot \cap P^c)$  are measures, that is, positive countably additive functions.*

*Proof.* Set  $\alpha := \sup\{\nu(A) : A \in \mathcal{B}, \nu(A) = |\nu|(A)\} < \infty$ . Suppose that  $A, B \in \mathcal{B}$  satisfies  $\nu(A) = |\nu|(A)$  and  $\nu(B) = |\nu|(B)$ . Then we have

$$(4.56) \quad |\nu|(A) + |\nu|(B) = |\nu|(A \cup B) + |\nu|(A \cap B) \geq \nu(A \cup B) + \nu(A \cap B) = \nu(A) + \nu(B) = |\nu|(A) + |\nu|(B).$$

As a result  $\nu(A \cup B) = |\nu|(A \cup B)$ .

In view of this paragraph, we can find  $P$  such that  $\alpha = \nu(P)$ . If  $B \subset P^c$ , then  $\nu(B) \leq 0$ . As a result  $-\nu(\cdot \cap P^c)$  is a measure.

Suppose that  $B \subset P$ . Then

$$(4.57) \quad |\nu|(P) = |\nu|(P \setminus B) + |\nu|(B) \geq \nu(P \setminus B) + \nu(B) = \nu(P) = |\nu|(P).$$

As a result  $\nu(B) = |\nu|(B)$ . This implies that  $\nu(\cdot \cap P)$  is a measure.  $\square$

Based upon the above theorem, we can decompose  $\nu$ . This corresponds to the fact that  $f = f^+ - f^-$  for all real-valued measurable functions  $f$ .

**Definition 4.25** (Hahn-decomposition). Let  $\nu$  be a real-valued countably additive function. Then  $\nu = \nu(\cdot \cap P) - (-\nu(\cdot \cap P^c))$  is said to be Hahn-decomposition of  $\nu$ .

Absolute continuity and Radon-Nikodym derivative. We go back to the discussion about density above. If  $\nu$  is given by (4.40), then  $\nu$  is a countably additive function. A natural question arises: When countably additive functions can be expressed as above? To answer this question, we present a definition.

**Definition 4.26** (Absolute continuity). A countably additive function  $\nu$  is absolutely continuous with respect to a measure  $\mu$ , if  $\nu(A) = 0$  for all  $\mu$ -null sets. In this case write  $\nu \ll \mu$ .

The following proposition is easy to show.



**Lemma 4.27.** *The condition  $\nu \ll \mu$  implies  $|\nu| \ll \mu$ .*

**Theorem 4.28** (Radon-Nikodym derivative). *Assume  $X$  is  $\sigma$ -finite with respect to a measure  $\mu$ . Suppose that  $\nu$  are countably additive measures with  $\nu \ll \mu$ . Then there exists  $f \in L^1(\mu)$  such that*

$$(4.58) \quad \nu(A) := \int_A f(x) d\mu(x).$$

$f$  appearing in this theorem is called Radon-Nikodym derivative of  $\nu$  with respect to  $\mu$  and it is sometimes written as  $\frac{d\nu}{d\mu}$ . We may assume that  $\nu$  is positive by virtue of the Hahn-Decomposition.

*Proof.* It can be assumed that  $\mu$  is finite. Let us return to the proof of the theorem. We set

$$(4.59) \quad \mathcal{G} := \left\{ f \in L^1(\mu)_+ : \nu(E) \geq \int_E f(x) d\mu(x) \text{ for all } E \in \mathcal{B} \right\}.$$

Then we see that  $\mathcal{G}$  enjoys the following properties.

$$(4.60) \quad f \in \mathcal{G}, g \in \mathcal{G} \text{ implies } \max(f, g) \in \mathcal{G}$$

$$(4.61) \quad f_1, f_2, \dots \in \mathcal{G} \text{ and } f_j \leq f_{j+1} \text{ for all } j \in \mathbb{N} \text{ implies } \lim_{j \rightarrow \infty} f_j \in \mathcal{G}.$$

In view of these facts we see that there exists  $f \in \mathcal{G}$  such that

$$(4.62) \quad \int_X f(x) d\mu(x) = \sup_{g \in \mathcal{G}} \int_X g(x) d\mu(x).$$

We shall prove that  $f$  is the desired function.

To prove that  $f$  is the desired function, we note that it suffices to prove

$$(4.63) \quad \nu(X) = \int_X f(x) d\mu(x) \text{ or equivalently } \nu(X) \leq \int_X f(x) d\mu(x).$$

For this purpose we take  $\varepsilon \geq 0$  so that

$$(4.64) \quad \nu(X) \geq \int_X f(x) d\mu(x) + \varepsilon\mu(X).$$

We form the Hahn-decomposition of  $\nu - f\mu - \varepsilon\mu$ . Let  $P \in \mathcal{B}$  satisfy

$$(4.65) \quad (\nu - f\mu - \varepsilon\mu)(E) \geq 0, E \in \mathcal{B}|P, (\nu - f\mu - \varepsilon\mu)(E) \leq 0, E \in \mathcal{B}|P^c.$$

Set  $g(x) := \chi_{P^c}(x)f(x) + \chi_P(x)(f(x) + \varepsilon)$  for  $x \in X$ .

Suppose that  $E \subset P$ . Then, using  $(\nu - f\mu - \varepsilon\mu)(E) \geq 0, E \in \mathcal{B}|P$ , we obtain

$$(4.66) \quad \nu(E) \geq \int_E g(x) d\mu(x).$$

Suppose instead that  $E \subset P^c$ . Then, taking into account that  $g = f$  on  $P^c$ , we have

$$(4.67) \quad \nu(E) \geq \int_E f(x) d\mu(x) = \int_E g(x) d\mu(x).$$

From (4.66) and (4.67), we see that  $g \in \mathcal{G}$ . By the definition of  $f$  we have

$$(4.68) \quad \int_X f(x) d\mu(x) = \int_X g(x) d\mu(x).$$

This means that  $\varepsilon\mu(P) = 0$ . Suppose that  $\mu(P) = 0$ . Then  $\nu(R) = 0$  whenever  $R \subset P$  is  $\mathcal{B}$ -measurable. As a result, we see

$$(4.69) \quad (\nu - f\mu - \varepsilon\mu)|P \leq 0.$$

If we put (4.69) together with (4.65), we see that

$$(4.70) \quad \nu - f\mu - \varepsilon\mu \leq 0.$$

However, we are assuming

$$(4.71) \quad \int_X f(x) d\mu(x) + \varepsilon\mu(X) \leq \nu(X).$$

From (4.70) and (4.71) we conclude  $\varepsilon = 0$ .

Thus as a conclusion, we see that  $\int_X f(x) d\mu(x) + \varepsilon\mu(X) \leq \nu(X)$  implies  $\varepsilon = 0$ , which means that  $\int_X f(x) d\mu(x) = \nu(X)$ .

As a result we have obtained the desired function  $f$ . □

**Lemma 4.29.** *Let  $\mu$  be a signed measure. Then*

$$(4.72) \quad \left| \frac{d\mu}{d\|\mu\|} \right| = 1$$

for  $\|\mu\|$ -almost every  $x \in X$ .

*Proof.* Since  $|\mu(E)| \leq \|\mu\|(E)$ , we have  $\left| \frac{d\mu}{d\|\mu\|} \right| \leq 1$ . Indeed, for  $\varepsilon > 0$  and  $\theta \in \mathbb{R}$ , let

$$F := \left\{ x \in X : \left| \frac{d\mu}{d\|\mu\|} \right| > 1 \right\}$$

$$F_{\varepsilon, \theta} := \left\{ x \in X : \operatorname{Re} \left( e^{\theta} \frac{d\mu}{d\|\mu\|} \right) > 1 + \varepsilon \right\}.$$

Then since  $|\mu(F_{\varepsilon, \theta})| \leq \|\mu\|(F_{\varepsilon, \theta})$ , we have  $\|\mu\|(F_{\varepsilon, \theta}) = 0$ . Therefore, we obtain  $\|\mu\|(F) = 0$ .

Next, for  $\varepsilon > 0$ , we set

$$G := \left\{ \left| \frac{d\mu}{d\|\mu\|} \right| < 1 \right\}, \quad F_{\varepsilon} := \left\{ \left| \frac{d\mu}{d\|\mu\|} \right| < 1 - \varepsilon \right\}.$$

Then we have  $|\mu(G)| \leq (1 - \varepsilon)\|\mu\|(G)$  for all  $\|\mu\|$ -measurable set  $G$  contained in  $F_{\varepsilon}$ . Since

$$(4.73) \quad \|\mu\|(F_{\varepsilon}) = \sup \left\{ \sum_{j=1}^J |\mu(E_j)| : \{E_j\}_{j=1}^J \text{ is a finite partition of } F_{\varepsilon} \right\},$$

we conclude

$$(4.74) \quad \|\mu\|(F_{\varepsilon}) \leq (1 - \varepsilon)\|\mu\|(F_{\varepsilon}).$$

Therefore  $\|\mu\|(F_{\varepsilon}) = 0$ . Since  $\varepsilon > 0$  is arbitrary, we conclude  $\|\mu\|(G) = 0$ . □

**Lemma 4.30.** *Suppose that  $\mu$  and  $\nu$  are positive measures and  $\eta$  is a signed measure. Assume  $\eta \ll \nu \ll \mu$ . Then we have*

$$(4.75) \quad \frac{d\eta}{d\mu} = \frac{d\eta}{d\nu} \cdot \frac{d\nu}{d\mu}.$$

*Proof.* Let  $E$  be a measurable set. We may assume  $\eta$  is positive as well. Then we have

$$(4.76) \quad \eta(E) = \int_E \frac{d\eta}{d\nu}(x) d\nu(x).$$

Since  $\frac{d\eta}{d\nu} \geq 0$ , a passage to the limit gives us

$$\eta(E) = \lim_{N \rightarrow \infty} \sum_{j=0}^{\infty} \frac{j}{N} \nu \left( E \cap \left\{ \frac{j}{N} \leq \frac{d\eta}{d\nu} < \frac{j+1}{N} \right\} \right) = \lim_{N \rightarrow \infty} \sum_{j=0}^{\infty} \frac{j}{N} \int_{E \cap \left\{ \frac{j}{N} \leq \frac{d\eta}{d\nu} < \frac{j+1}{N} \right\}} \frac{d\nu}{d\mu}(x) d\mu(x),$$

where for the last equality we have used the definition of  $\frac{d\nu}{d\mu}$ . By the dominated convergence theorem we obtain

$$(4.77) \quad \eta(E) = \int_E \frac{d\eta}{d\nu}(x) \cdot \frac{d\nu}{d\mu}(x) d\mu(x).$$

From (4.77) we obtain

$$(4.78) \quad \frac{d\eta}{d\mu} = \frac{d\eta}{d\nu} \cdot \frac{d\nu}{d\mu},$$

which is the desired result. □

#### 4.5. Convolution and mollifier.

Now we place ourselves in the setting of  $\mathbb{R}^d$  coming with the Lebesgue measure  $dx$ . In this case, unless possible confusion happens, we write

$$(4.79) \quad \int_E f = \int_E f(x) dx$$

for a Lebesgue measurable set  $E$ .

Many mathematical transforms can be written in terms of convolutions, whose definition we present below.

**Definition 4.31.** Let  $f, g$  be a measurable functions. Then the convolution of  $f$  and  $g$  are given by

$$(4.80) \quad f * g(x) = \int_{\mathbb{R}^d} f(y)g(x - y) dy,$$

provided this definition makes sense.

*Exercise 42.* Let  $1 \leq p \leq \infty$ . Show that  $f * g$  makes sense in each case below. Give a detailed explanation of how  $f * g$  makes sense.

- (1)  $f \in L^p(\mathbb{R}^d), g \in L^{p'}(\mathbb{R}^d)$ .
- (2)  $f, g \in L^1(\mathbb{R}^d)$ . (Hint: see (4.83).)

Under these conditions, show that  $f * g = g * f$ .

Now we generalize the above exercise.

**Theorem 4.32** (Young inequality). *Let  $p, q, r$  satisfy*

$$(4.81) \quad 1 \leq p, q, r \leq \infty, \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1.$$

*Suppose that  $f \in L^p(\mathbb{R}^d)$  and  $g \in L^q(\mathbb{R}^d)$ . Then  $f * g$  makes sense and belongs to  $L^r(\mathbb{R}^d)$  with the estimate*

$$(4.82) \quad \|f * g\|_r \leq \|f\|_p \|g\|_q.$$

In particular,

$$(4.83) \quad \|f * g\|_p \leq \|f\|_p \|g\|_1.$$

*Proof.* Here we shall content ourselves with proving (4.83) for  $r = \infty$  and  $p = q = r = 1$ . Because the rest shall be proved via interpolation via interpolation in Chapter 16. If  $r = \infty$ , then we have only to use the Hölder inequality. We have a pointwise estimate

$$(4.84) \quad |f * g(x)| \leq \left( \int_{\mathbb{R}^d} |f(x-y)|^p dx \right)^{\frac{1}{p}} \left( \int_{\mathbb{R}^d} |g(y)|^q dx \right)^{\frac{1}{q}} = \|f\|_p \|g\|_q.$$

If  $p = q = r = 1$ , then we have

$$\int_{\mathbb{R}^d} |f * g(x)| dx \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |f(y)| \cdot |g(x-y)| dy \right) dx = \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |f(y)| \cdot |g(x-y)| dx \right) dy.$$

Factoring out  $|f(y)|$  and changing variables, we obtain

$$\int_{\mathbb{R}^d} |f * g(x)| dx = \int_{\mathbb{R}^d} |f(y)| \cdot \left( \int_{\mathbb{R}^d} |g(x-y)| dx \right) dy = \int_{\mathbb{R}^d} |f(y)| \cdot \|g\|_1 dy = \|f\|_1 \|g\|_1.$$

Thus, the proof is complete for  $p = q = r = 1$ .  $\square$

Now we take up the mollifiers. Fix a positive function  $\varphi \in C_c^\infty(\mathbb{R}^d)$  satisfying  $\int_{\mathbb{R}^d} \varphi(x) dx = 1$  and  $\text{supp}(\varphi) \subset B(1)$ . Then, for  $t > 0$ , let us write

$$(4.85) \quad \varphi_t(x) := \frac{1}{t^n} \varphi\left(\frac{x}{t}\right) \quad (x \in \mathbb{R}^n).$$

About the mollification, the following properties are fundamental.

**Theorem 4.33.** *Let  $1 \leq p \leq \infty$ .*

- (1) *Let  $f \in L_{\text{loc}}^1(\mathbb{R}^d)$ . Then  $\varphi_t * f \in C^\infty(\mathbb{R}^d)$ .*
- (2) *Let  $f \in L^p(\mathbb{R}^d)$ . Then*

$$(4.86) \quad \|\varphi_t * f\|_p \leq \|f\|_p.$$

- (3) *Suppose further that  $p < \infty$ . Then for all  $f \in L^p(\mathbb{R}^d)$ ,  $\varphi_t * f \rightarrow f$  in  $L^p(\mathbb{R}^d)$ .*
- (4) *Let  $f \in \text{BUC}(\mathbb{R}^d)$ . Then  $\varphi_t * f \rightarrow f$  in  $\text{BUC}(\mathbb{R}^d)$ .*
- (5) *Let  $f \in L_{\text{loc}}^1(\mathbb{R}^d)$ . Then for a.e.  $x \in \mathbb{R}^d$*

$$(4.87) \quad \varphi_t * f(x) \rightarrow f(x).$$

*Proof.* (1) is derived from Theorem 3.60, which we use without notice in what follows. (2) is just an application of (4.83). (5) is rather difficult and we postpone its proof. We now turn to the proof of (3). Suppose that  $f$  is a continuous function with compact support. Choose  $R > 0$  so that  $\text{supp}(f) \subset B(R)$ . Then we have

$$(4.88) \quad \varphi_t * f(x) - f(x) = \int_{\mathbb{R}^d} (f(x-ty) - f(x)) \varphi(y) dy.$$

Therefore, we have

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} |\varphi_t * f(x) - f(x)| &\leq \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} \left( \sup_{y \in B(1)} |f(x-ty) - f(x)| \varphi(y) \right) dy \\ &\leq \sup_{x \in \mathbb{R}^d, z \in B(t)} |f(x-z) - f(x)|. \end{aligned}$$

Since  $f$  is a compactly supported continuous function,  $f$  is uniformly continuous, that is

$$(4.89) \quad \sup_{x \in \mathbb{R}^d, z \in B(t)} |f(x-z) - f(x)| \rightarrow 0$$

as  $t \rightarrow 0$ . Furthermore we have  $\varphi_t * f$  is supported on  $B(R+1)$ . Now we pass to the general case. Let  $f \in L^p(\mathbb{R}^d)$ .

Take a compactly supported continuous function  $g$  arbitrarily. First, we decompose

$$\|\varphi_t * f - f\|_p \leq \|\varphi_t * (f - g)\|_p + \|\varphi_t * g - g\|_p + \|f - g\|_p$$

Now that (2) is established, we see that

$$(4.90) \quad \|\varphi_t * f - f\|_p \leq \|\varphi_t * g - g\|_p + 2\|f - g\|_p.$$

Due to a preparatory observation made above and the fact that  $g \in C_c$ , we see that

$$(4.91) \quad \limsup_{t \rightarrow 0} \|\varphi_t * f - f\|_p \leq \limsup_{t \rightarrow 0} \|\varphi_t * g - g\|_p + 2\|f - g\|_p = 2\|f - g\|_p.$$

Since  $g$  is arbitrary, we see

$$(4.92) \quad \limsup_{t \rightarrow 0} \|\varphi_t * f - f\|_p \leq 2 \inf_{g \in C_c} \|f - g\|_p = 0.$$

This is the desired result.

Finally we prove (4). Assume that  $f \in \text{BUC}(\mathbb{R}^d)$ . Note that

$$(4.93) \quad |\varphi_t * f(x) - f(x)| \leq \int_{\mathbb{R}^d} |f(x-ty) - f(x)| \varphi(y) dy \leq \sup_{z \in B(t)} |f(x-z) - f(x)|.$$

Therefore we have

$$(4.94) \quad \|\varphi_t * f - f\|_\infty \leq \sup_{\substack{x \in \mathbb{R}^d \\ z \in B(t)}} |f(x-z) - f(x)|.$$

Now that we are assuming  $f \in \text{BUC}(\mathbb{R}^d)$ , we have

$$(4.95) \quad \lim_{t \rightarrow 0} \left( \sup_{\substack{x \in \mathbb{R}^d \\ z \in B(t)}} |f(x-z) - f(x)| \right) = 0,$$

which yields, together with (4.94), the desired result.  $\square$

*Exercise 43.* Let  $\Gamma(t, x) := \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp(-t|x|^2)$ . Given a function  $f$  on  $\mathbb{R}^d$ , we write

$$(4.96) \quad e^{t\Delta} f(x) := \Gamma(t, *) * f(x)$$

as long as it makes sense. First, if  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p \leq \infty$ , then show that the integral defining  $e^{t\Delta} f(x)$  converges absolutely. More precisely, establish that  $\|e^{t\Delta} f\|_p \leq \|f\|_p$ . Below, we assume that  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p \leq \infty$ .

- (1) Show that  $e^{s\Delta}(e^{t\Delta} f) = e^{(t+s)\Delta} f$ . This property is called the semigroup property of  $e^{t\Delta}$ .
- (2) Let  $1 \leq p < \infty$ . Show that  $\lim_{t \rightarrow 0} e^{t\Delta} f = f$  in  $L^p(\mathbb{R}^d)$ .
- (3) Let  $f \in \text{BUC}(\mathbb{R}^d)$ . Then show that  $\lim_{t \rightarrow 0} e^{t\Delta} f = f$  in  $\text{BUC}(\mathbb{R}^d)$ .
- (4) Let  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p \leq \infty$ . Show that  $e^{t\Delta} f \in C^\infty(\mathbb{R}_+^{d+1}) = C^\infty(\mathbb{R}_x^d \times (0, \infty)_t)$ .
- (5) Let  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p \leq \infty$ . Then show that  $u(t, x) = e^{t\Delta} f(x)$  satisfies the following heat equation:  $\left( \frac{\partial}{\partial t} - \Delta \right) e^{t\Delta} f(x) = 0$  on  $\mathbb{R}_+^{d+1}$ .

#### 4.6. Density.

Now we turn to the density of  $L^p(\mathbb{R}^d)$  spaces.

**Theorem 4.34.** *Let  $1 \leq p < \infty$ . Then  $C_c$  is dense in  $L^p(\mathbb{R}^d)$ . That is, any  $f \in L^p(\mathbb{R}^d)$  can be approximated by a sequence  $\{\varphi_j\}_{j \in \mathbb{N}}$  of compactly supported continuous functions.*

*Proof.* Let  $f \in L^p(\mathbb{R}^d)$ . Then define  $f_j := \chi_{\{|f| \leq j\} \cap B(j)} \cdot f$ . By the monotone convergence theorem we have  $f_j \rightarrow f$  in the  $L^p(\mathbb{R}^d)$  topology. Thus, instead of approximating  $f \in L^p(\mathbb{R}^d)$ , we have only to approximate bounded function with bounded support. Using the same notation as before, we mollify such  $f$ . Note that we have established  $\varphi_t * f$  tends to  $f$  in the  $L^p(\mathbb{R}^d)$ -topology as  $t \downarrow 0$ . Furthermore  $\text{supp}(\varphi_t * f) \subset \text{supp}(f) + \text{supp}(\varphi_t)$  is a compact set for all  $t > 0$ . Therefore it follows that  $C_c$  is dense in  $L^p(\mathbb{R}^d)$ .  $\square$

*Exercise 44.* Let  $X$  be a topological space and  $A, B \subset X$ . If  $\overline{B} = X$  and  $B \subset \overline{A}$ , then show that  $\overline{A} = X$ .

Next we take up variation principle.

**Theorem 4.35** (Variation principle). *Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . If  $\int_{\mathbb{R}^d} f(x)g(x) dx = 0$  for all  $g \in C_c^\infty(\mathbb{R}^d)$ , then  $f(x) = 0$  for almost every  $x \in \mathbb{R}^d$ .*

*Proof.* We may assume  $f \in L^1(\mathbb{R}^d)$  by reducing the matter to the local one. Assume  $f$  is not zero almost everywhere.

We first choose  $h \in C_c^\infty(\mathbb{R}^d)$  so that  $\|f - h\|_1 \leq \frac{1}{4}\|f\|_1$ . In this case  $\|h\|_1 \geq \frac{3}{4}\|f\|_1$ . Let  $k$  be a measurable function given by  $k(x) = \overline{\text{sgn}(h(x))}$ . Then we have  $k(x)h(x) = |h(x)|$ .

We now construct  $g \in C_c^\infty(\mathbb{R}^d)$  so that  $\|k - g\|_1 \leq \frac{1}{4\|h\|_\infty + 4}\|f\|_1$  and  $\|g\|_\infty \leq 1$  by mollification.

Then

$$(4.97) \quad \int_{\mathbb{R}^d} f(x)g(x) dx = \int_{\mathbb{R}^d} (f(x) - h(x))g(x) dx + \int_{\mathbb{R}^d} (g(x) - k(x))h(x) dx + \int_{\mathbb{R}^d} k(x)h(x) dx.$$

Note that

$$(4.98) \quad \left| \int_{\mathbb{R}^d} (f(x) - h(x))g(x) dx \right| \leq \frac{1}{4}\|f\|_1, \quad \left| \int_{\mathbb{R}^d} (g(x) - k(x))h(x) dx \right| \leq \frac{1}{4}\|f\|_1,$$

while

$$(4.99) \quad \left| \int_{\mathbb{R}^d} k(x)h(x) dx \right| = \|h\|_1 \geq \frac{3}{4}\|f\|_1.$$

Therefore

$$(4.100) \quad \left| \int_{\mathbb{R}^d} f(x)g(x) dx \right| > \frac{1}{4}\|f\|_1 > 0.$$

This contradicts to the assumption. Therefore  $f = 0$  a.e. on  $\mathbb{R}^d$ .  $\square$

*Exercise 45.* Let  $E \subset \mathbb{R}^d$  be a measurable subset in  $\mathbb{R}^d$  with  $0 < |E| < \infty$ .

- (1) Show that  $f(x) := \int_{\mathbb{R}^d} \chi_E(x+y)\chi_E(y) dy$  is continuous.
- (2) Show that  $\{x - y : x, y \in E\}$  contains 0 as an interior point.

4.7. Dual spaces in connection with measure theory.

Riesz's representation theorem. Let  $X$  be a locally compact topological space. Denote by  $C(X)$  the set of bounded continuous functions. In this paragraph we set  $\mathcal{B}$  as a  $\sigma$ -algebra generated by open sets in  $X$ . In this paragraph we specify the dual of  $C(X)$ .

**Theorem 4.36** (Riesz). *Suppose that  $I$  is a positive  $\mathbb{R}$ -linear mapping from  $C_c(X)$  to  $\mathbb{R}$  in the sense that  $I(f) \geq 0$  for all  $f \in C_c(X)_+$ . Then there exists a unique positive Borel measure  $\mu$  such that  $\mu$  is compact and regular and that it satisfies*

$$(4.101) \quad I(f) = \int_X f(x) d\mu(x)$$

for all  $f \in C_c(X)$ .

*Proof.* Denote  $\mathcal{K}$  by the set of all compact sets. For  $K \in \mathcal{K}$  we define

$$(4.102) \quad \lambda(K) := \inf\{I(f) : f \geq \chi_K\}.$$

Then it is easy to see that  $\lambda$  is a content, that is,

$$(4.103) \quad \lambda\left(K \amalg L\right) = \lambda(K) + \lambda(L), \quad \lambda(K \cup L) \geq \lambda(K) + \lambda(L)$$

for all  $K, L \in \mathcal{K}$ . Let  $\mu$  be the measure induced by  $\lambda$ .

We have to show that

$$(4.104) \quad I(f) = \int_X f(x) d\mu(x)$$

for all  $f \in C_c(X)$ . We can decompose  $f = f_+ - f_-$ , so that it is assumed that  $f$  is positive.

Take  $\psi \in C_c(X)$  so that

$$(4.105) \quad \chi_{\text{supp}(f)} \leq \psi \leq \chi_{V_0},$$

where  $V_0$  is an open set engulfing  $\text{supp}(f)$ . Then  $|I(g)| \leq \|g\|_\infty I(\psi)$  for all  $g \in C_c(X)$  with  $\text{supp}(g) \subset \text{supp}(f)$ . In particular if  $\{g_n\}_{n \in \mathbb{N}}$  is a sequence of functions converging uniformly to  $g$  with  $\text{supp}(g_n) \subset \text{supp}(f)$ , then  $I(g_n) \rightarrow I(g)$ .

For  $k \in \mathbb{N}$  we set

$$(4.106) \quad M := \sup f, \quad V_{j,k} := \{f < jM/k\}.$$

We take  $\psi_{j,k}$  so that

$$(4.107) \quad 0 \leq \mu(V_{j,k}) - I(\psi_{j,k}) < 2^{-k}, \quad 0 \leq \psi_{j,k} \leq \chi_{V_{j,k}}.$$

We set  $g_k := \sum_{j=1}^k \frac{1}{k} \psi_{j,k}$ . Then  $I(f) = \lim_{k \rightarrow \infty} I(g_k) = \lim_{k \rightarrow \infty} \sum_{j=1}^k \frac{1}{k} \mu(V_{j,k}) = \int_X f(x) d\mu(x)$ .

As a result we have proved the theorem. □

*Exercise 46.* Suppose that  $A$  and  $B$  are open sets in  $\mathbb{R}^d$  with  $A \subset B \subset \mathbb{R}^d$  and that  $\varphi$  is a continuous function. Then what does the following inequality mean?

$$(4.108) \quad \chi_A \leq \varphi \leq \chi_B.$$

The following variants can be proved similarly.

**Theorem 4.37** (Riesz representation theorem). *Suppose that  $I : C(X) \rightarrow \mathbb{R}$  be a functional satisfying*

$$(4.109) \quad |I(f)| \leq C \|f\|_\infty, \quad I(f) \geq 0, \quad \text{for all } f \in C(X)_+.$$

*Then there exists a finite measure  $\mu : \mathcal{B} \rightarrow \mathbb{R}$  such that  $I(f) = \int_X f(x) d\mu(x)$  for all  $f \in C(X)$ .*

**Theorem 4.38** (Riesz representation theorem). *Suppose that  $I : C(X) \rightarrow \mathbb{R}$  be a functional satisfying*

$$(4.110) \quad |I(f)| \leq C \|f\|_\infty.$$

*Then there exists a finite signed measure  $\mu : \mathcal{B} \rightarrow \mathbb{R}$  such that  $I(f) = \int_X f(x) d\mu(x)$  for all  $f \in C(X)$ .*

Duality  $L^p(\mu)$ - $L^{p'}(\mu)$ . In the same way as  $C(X)$ , we now specify the dual of  $L^p(\mu)$  with  $1 \leq p < \infty$ .

**Theorem 4.39** (Duality  $L^p(\mu)$ - $L^{p'}(\mu)$ ). *Let  $1 \leq p < \infty$  and  $(X, \mathcal{B}, \mu)$  be a  $\sigma$ -finite measure space, that is,  $X$  is partitioned into countable sets of finite  $\mu$ -measure. Then  $L^p(\mu)^* = L^{p'}(\mu)$  with norm coincidence. Speaking precisely, let  $f \in L^{p'}(\mu)$  and define  $F_f \in L^p(\mu)^*$  by*

$$(4.111) \quad g \in L^{p'}(\mu) \mapsto \int_X f(x) g(x) d\mu(x).$$

*Then  $f \in L^{p'}(\mu) \mapsto F_f \in L^p(\mu)^*$  is well-defined and  $\|f\|_{L^{p'}(\mu)} = \|F_f\|_{L^p(\mu)}$ . If  $F : L^p(\mu) \rightarrow \mathbb{K}$  is a continuous linear functional, i.e., there exists  $M > 0$  such that*

$$(4.112) \quad |F(f)| \leq M \|f\|_{L^p(\mu)}$$

*for all  $f \in L^p(\mu)$ , then  $F$  is realized with some  $g \in L^{p'}(\mu)$ .*

*Proof.* By Hölder inequality it is easy to see  $f \mapsto F_f$  is well-defined and

$$(4.113) \quad \|f\|_{L^{p'}(\mu)} \geq \|F_f\|_{L^p(\mu)}.$$

The reverse inequality of this can be attained easily. Let  $f \neq 0$ . We can even write out the norm attainer  $g$ , that is, the non-zero element in  $g \in L^p(\mu)$  satisfying

$$(4.114) \quad |F_f(g)| = \|f\|_{L^{p'}(\mu)} \cdot \|g\|_{L^p(\mu)}.$$

Indeed, it suffices to take  $g := \lim_{\varepsilon \downarrow 0} \bar{f} \cdot (\varepsilon + |f|)^{\frac{p'}{p}-1}$ . Therefore, it remains to show the surjectivity of  $f \in L^{p'}(\mu) \rightarrow F_f \in L^p(\mu)^*$ .

Here we shall give an outline of the proof of the surjectivity, the Radon-Nikodym theorem plays a key part.

Assume first that  $X$  is finite. Let  $F \in L^p(\mu)^*$ . We define a (signed) measure  $\mu_F$  by

$$(4.115) \quad \mu_F(E) := F(\chi_E).$$

Then  $\mu(E) = 0$  implies  $\mu_F(E) = 0$ . Therefore, we are in the position of using the Radon-Nikodym theorem to obtain  $\mu_F = f(x) \cdot \mu$  for some  $f \in L^1(\mu)$ . An elaborate argument shows  $f \in L^{p'}(\mu)$ .

In general  $X$  can be expressed as  $X = \sum_{j=1}^{\infty} X_j$  with  $\mu(X_j) < \infty$  for all  $j = 1, 2, \dots$ . In this case we can find  $f_j \in L^{p'}(X_j)$ . Another elaborate argument shows the  $f_j$  defines an  $L^{p'}(\mu)$  function  $f$  that realizes  $F$ . Thus, the proof is complete.  $\square$

*Exercise 47.* Fill details in the above proof.



*Exercise 48.* By using the Riesz representation theorem, give an alternative proof of the above theorem for  $p = 2$ .

*Exercise 49.* Let  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  be a measure space. Show that any continuous linear mapping from  $L^1(\mu)$  to  $L^\infty(\nu)$  can be written as

$$(4.116) \quad Tf(y) = \int_X K(x, y) f(x) d\mu(x)$$

for some  $K \in L^\infty(X \times Y)$ .

**4.8. Some calculus formulas.**

To conclude this section, we take up some calculation facts concerning the unit ball  $B(1)$  and the unit sphere  $S^d$  and some calculation formulae.

Integration by parts and summation by parts. As for integration by parts, we content ourselves with viewing the following elementary one.

*Exercise 50.* Let  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  be  $C^1$ -functions and  $-\infty < a < b < \infty$ . Then show that  $f \cdot g \in C^1$  and that

$$(4.117) \quad \int_a^b f'(x)g(x) dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x) dx.$$

Now we turn to summation by parts.

*Exercise 51* (Summation by parts in  $\ell^p$ ). Suppose that  $0 < p < \infty$ . Let  $a_1, a_2, \dots, a_n \geq 0$  and  $x_1 \geq x_2 \geq \dots \geq x_n \geq 0$  be given. If  $p \geq 1$ , then

$$(4.118) \quad \left( \sum_{k=1}^n a_k x_k \right)^p \geq \sum_{k=1}^{n-1} A_k^p (x_k^p - x_{k+1}^p) + A_n^p x_n^p,$$

where  $A_k = a_1 + a_2 + \dots + a_k$ . The inequality reverses if  $0 < p \leq 1$ . Note that equality holds for  $p = 1$ .

*Exercise 52.* Assume that  $\{a_i\}_{i=1}^\infty$  is a bounded sequence such that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_i = y$ . Then show that

$$(4.119) \quad \lim_{r \uparrow 1} (1-r) \sum_{i=1}^\infty r^i a_i = y.$$

Change of variables formula. We supplement this section by proving the change of variables formula.

**Theorem 4.40** (Change of variables formula). *Let  $\Omega, D$  be domains on  $\mathbb{R}^d$ . Suppose that  $f : \Omega \rightarrow D$  is a  $C^1$ -diffeomorphism and  $g$  is a positive measurable mapping on  $D$ . Then*

$$(4.120) \quad \int_D g(x) dx = \int_\Omega g(f(x)) |\det(Df)(x)| dx.$$

We write  $Jf(x) = |\det(Df)(x)|$ .

**Lemma 4.41.** *Let  $t > 1$ . There exist an open covering  $\{O_k\}_{k \in O}$  and symmetric automorphism  $T_k$  such that*

$$(4.121) \quad t^{-1}|T_k(b-a)| \leq |f(b) - f(a)| \leq t|T_k(b-a)|$$

for  $a, b \in O_k$  and

$$(4.122) \quad t^{-d} |\det(T_k)| \leq Jf(x) \leq t^d |\det(T_k)|$$

for  $x \in O_k$ .

*Proof.* Let  $C$  be a countable dense set in  $\Omega$  and  $S$  be a countable subset of symmetric automorphisms in  $\mathbb{R}^d$ .

We denote

$$(4.123) \quad U(c, T, i) := \{b \in B(c, i^{-1}) \cap \Omega : b \text{ satisfies (4.124) and (4.125)}\}.$$

Here the conditions (4.124) and (4.125) are

$$(4.124) \quad t^{-\frac{1}{2}} |Tv| \leq |Df(b)v| \leq t^{\frac{1}{2}} |Tv| \quad \text{for all } v \in \mathbb{R}^d$$

and

$$(4.125) \quad |f(a) - f(b) - Df(b)(a - b)| \leq (t^{-1} - t^{-2})|a - b| \quad \text{for all } a \in B(b, 2i^{-1}).$$

From (4.124) we obtain

$$(4.126) \quad t^{-d} |\det(T)| \leq |\det(Df)(b)| \leq t^d |\det(T)|$$

If we combine (4.124) and (4.125), it follows that, taking into account  $t > 1$ ,

$$(4.127) \quad t^{-1} |T(b - a)| \leq |f(b) - f(a)| \leq t |T(b - a)|$$

for all  $b \in U(c, T, i)$  and  $a \in B(b, 2i^{-1})$ .

It remains to relabel  $U(c, T, i)_{x, T, i}$  as  $\{O_k\}$ . □

With this preparation in mind, we prove the theorem.

*Proof.* Here and below in this proof let  $A \sim B$  denote  $t^{-d} A \leq B \leq t^d A$ .

Denote by  $\mathcal{H}^0$  the counting measure. Set  $E_k = \{x \in \Omega : k = \min\{l \in \mathbb{N} : x \in O_l\}\}$ . Then, taking  $F \in \mathcal{B}(\Omega)$  arbitrarily, we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \mathcal{H}^0(F \cap f^{-1}(y)) dy &= \lim_{l \rightarrow \infty} \sum_{Q \in \mathcal{D}(l)} \sum_k \int_{\mathbb{R}^d} \chi_{f(Q \cap F \cap E_k)}(y) dy \\ &\sim \lim_{l \rightarrow \infty} \sum_{Q \in \mathcal{D}(l)} \sum_k \int_{\mathbb{R}^d} \chi_{T_k(Q \cap F \cap E_k)}(y) dy \\ &= \lim_{l \rightarrow \infty} \sum_{Q \in \mathcal{D}(l)} \sum_k \int_{\mathbb{R}^d} \chi_{(Q \cap F \cap E_k)}(y) |\det(T_k)| dy \\ &\sim \lim_{l \rightarrow \infty} \sum_{Q \in \mathcal{D}(l)} \sum_k \int_{\mathbb{R}^d} \chi_{(Q \cap F \cap E_k)}(y) Jf(y) dy \\ &= \int_F JF(y) dy. \end{aligned}$$

Consequently we obtain

$$(4.128) \quad \int_{\mathbb{R}^d} \mathcal{H}^0(F \cap f^{-1}(y)) dy = \int_F Jf(x) dx.$$

If we pass to the measurable functions, the above formula reads

$$(4.129) \quad \int_{\mathbb{R}^d} \left( \sum_{x \in f^{-1}(y)} h(x) \right) dy = \int_F h(x) Jf(x) dx$$

for all  $h \in \mathcal{B}(\mathbb{R}^d)_+$ .

In particular  $h = g \circ f$  we obtain

$$(4.130) \quad \int_{\mathbb{R}^d} g(x) dy = \int_F g(f(x)) Jf(x) dx.$$

Thus, the proof is complete. □

Gamma function and Beta function. Let us recall the property of the Gamma function.

**Definition 4.42** (Gamma function, Beta function). Set

$$(4.131) \quad \Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0$$

and

$$(4.132) \quad B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$$

**Proposition 4.43.** *Let  $x, \alpha, \beta > 0$ .*

- (1) *The integral defining  $\Gamma(x)$  converges absolutely.*
- (2) *The integral defining  $B(\alpha, \beta)$  converges absolutely.*
- (3) *We have  $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$ .*

*Proof.* The first two assertions are immediate. Therefore, we prove the third one. The proof is obtained by a series of changing variables. By the Fubini theorem we have

$$(4.133) \quad \Gamma(\alpha)\Gamma(\beta) = \int_0^\infty \int_0^\infty u^{\alpha-1} v^{\beta-1} e^{-u-v} du dv$$

If we set  $u = vt$ , then we obtain

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty \int_0^\infty (vt)^{\alpha-1} v^{\beta-1} e^{-vt-v} v dt dv = \int_0^\infty \int_0^\infty v^{\alpha+\beta-1} e^{-v(t+1)} t^{\alpha-1} dv dt.$$

Setting  $w = (1+t)v$ , we obtain

$$\Gamma(\alpha)\Gamma(\beta) = \Gamma(\alpha + \beta) \int_0^\infty \frac{t^{\alpha-1} dt}{(1+t)^{\alpha+\beta}}$$

Finally if we set  $s = \frac{1}{1+t}$ , then we obtain

$$(4.134) \quad \Gamma(\alpha)\Gamma(\beta) = \int_0^1 \left( \frac{1}{s} - 1 \right)^{\alpha-1} s^{\alpha+\beta-2} ds = \int_0^1 (1-s)^{\alpha-1} s^{\beta-1} ds = B(\alpha, \beta).$$

This is the desired result. □

*Exercise 53.* Verify the following equalities:

- (1)  $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ .
- (2)  $\int_0^1 x^8 (1-x)^7 dx = \frac{1}{102960}$ .

$$(3) \int_0^3 x^4(x-3)^4 dx = \frac{2187}{70}.$$

Volume of  $B(1)$  and area of  $S^{d-1}$ . Now let us calculate the volume of  $B(1)$  and the area of  $S^{d-1}$ . To do this we set

$$(4.135) \quad S_{d+1} := \{(x_1, x_2, \dots, x_{d+1}) : x_1, x_2, \dots, x_{d+1} \geq 0, x_1 + x_2 + \dots + x_{d+1} \leq 1\}.$$

**Theorem 4.44.** *Let  $\alpha_1, \alpha_2, \dots, \alpha_{d+1} > 0$ . Then*

$$(4.136) \quad \int_{S_d} x_1^{\alpha_1-1} \dots x_d^{\alpha_d-1} (1-x_1-\dots-x_d)^{\alpha_{d+1}-1} dx_1 \dots dx_d = \frac{\Gamma(\alpha_1) \dots \Gamma(\alpha_{d+1})}{\Gamma(\alpha_1 + \dots + \alpha_{d+1})}.$$

*Exercise 54.* Prove this theorem for  $d = 1$ .

*Proof.* We set

$$(4.137) \quad I(\alpha_1, \alpha_2, \dots, \alpha_{d+1}) = \int_{S_d} x_1^{\alpha_1-1} \dots x_d^{\alpha_d-1} (1-x_1-\dots-x_d)^{\alpha_{d+1}-1} dx_1 \dots dx_d.$$

If  $d = 1$ , then this integral is known as the  $\beta$ -function and the result is collect. Suppose that  $d \geq 2$ . Then

$$I(\alpha_1, \alpha_2, \dots, \alpha_{d+1}) = \int_{S_{d-1}} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \dots x_{d-1}^{\alpha_{d-1}-1} J(x_1, x_2, \dots, x_{d-1}) dx_1 \dots dx_{d-1}.$$

Here

$$(4.138) \quad J(x_1, x_2, \dots, x_{d-1}) := \int_{0 \leq x_d \leq 1-(x_1+x_2+\dots+x_{d-1})} x_d^{\alpha_d-1} (1-x_1-\dots-x_d)^{\alpha_{d+1}-1} dx_d.$$

Change variables to calculate  $J(x_1, \dots, x_{d-1})$ . Note that the  $\beta$ -function appears:

$$(4.139) \quad J(x_1, x_2, \dots, x_{d-1}) = (1-x_1-\dots-x_{d-1})^{\alpha_d+\alpha_{d+1}-1} \frac{\Gamma(\alpha_d)\Gamma(\alpha_{d+1})}{\Gamma(\alpha_d+\alpha_{d+1})}.$$

Inserting this, we obtain a recurrence formula.

$$(4.140) \quad I(\alpha_1, \alpha_2, \dots, \alpha_{d+1}) = I(\alpha_1, \alpha_2, \dots, \alpha_{d-1}, \alpha_d + \alpha_{d+1}) \frac{\Gamma(\alpha_d)\Gamma(\alpha_{d+1})}{\Gamma(\alpha_d + \alpha_{d+1})}.$$

With this recurrence formula (4.140), along with the initial condition  $I(\beta_1, \beta_2) = \frac{\Gamma(\beta_1)\Gamma(\beta_2)}{\Gamma(\beta_1 + \beta_2)}$ , we obtain the result.  $\square$

**Theorem 4.45.** *Let  $d \geq 2$ . Then  $v_d = |B(1)| = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d+2}{2}\right)}$  and  $\omega_d = |S^d| = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$ . In*

*particular, we have*

$$(4.141) \quad \frac{d}{dr} |B(r)| = |S^d(r)|.$$

*Exercise 55.* Verify (4.141) for  $d = 2, 3$ .

*Proof.* We may suppose that  $d \geq 3$ . Note that

$$\begin{aligned} |B(1)| &= |\{(x_1, x_2, \dots, x_d) \in \mathbb{R}^d : x_1^2 + x_2^2 + \dots + x_d^2 \leq 1\}| \\ &= 2^d \int_{\substack{x_1, x_2, \dots, x_d \geq 0 \\ x_1^2 + x_2^2 + \dots + x_d^2 \leq 1}} dx \\ &= \int_{(X_1, X_2, \dots, X_d) \in \Delta^{d-1}} X_1^{-\frac{1}{2}} X_2^{-\frac{1}{2}} \dots X_d^{-\frac{1}{2}} dX_1 dX_2 \dots dX_d \end{aligned}$$

Following the above notation, we obtain

$$(4.142) \quad |B(1)| = I\left(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}, 1\right) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d+2}{2}\right)}.$$

The exact value of  $\omega_d$  can be obtained from  $\omega_d = \int_{B(1)} \frac{dx'}{\sqrt{1-x_1^2-x_2^2-\dots-x_{d-1}^2}}$  similarly, where  $x' = (x_1, x_2, \dots, x_{d-1})$ . □

**Theorem 4.46.** *Let  $f$  be a continuous and integrable function. Then we have*

$$(4.143) \quad \int_{\mathbb{R}^d} f(x) dx = \int_0^\infty r^{d-1} \left( \int_{S^d} f(r\omega) d\omega \right) dr.$$

*In particular, if we assume in addition  $f$  is a radial function, that is,*

$$(4.144) \quad f(x) = F(|x|)$$

*for some function  $F$  on  $\mathbb{R}$ , then*

$$(4.145) \quad \int_{\mathbb{R}^d} f(x) dx = \omega_d \int_0^\infty r^{d-1} F(r) dr.$$

*Proof.* Assume first that  $f$  is radial. It is helpful to begin with the special case, that is,  $f$  is radial. A passage to the limit then allows us to assume that  $f$  is positive and compactly supported. Furthermore, it can be assumed that  $F$  given by (4.144) is smooth because we can approximate  $f$  with such functions. An elementary theorem on calculus gives us

$$(4.146) \quad f(x) = - \int_{|x|}^\infty F'(u) du = - \int_0^\infty \chi_{B(u)}(x) F'(u) du.$$

Therefore,

$$(4.147) \quad \int_{\mathbb{R}^d} f(x) dx = - \int_{\mathbb{R}^d} \left( \int_0^\infty \chi_{B(u)}(x) F'(u) du \right) dx.$$

If we use the Fubini theorem, then we obtain

$$(4.148) \quad \int_{\mathbb{R}^d} f(x) dx = -v_d \int_0^\infty u^d F'(u) du.$$

Now we integrate by parts:

$$(4.149) \quad \int_{\mathbb{R}^d} f(x) dx = d v_d \int_0^\infty u^{d-1} F(u) du.$$

Since  $d v_d = \omega_d$ , this is the desired result.

Passage to the general case Recall that  $O(d)$  is the set of all linear transforms preserving the Euclidean length in  $\mathbb{R}^d$ . Let  $\mu$  be the normalized Haar measure on  $O(d)$ , which we take up in Section 64. Here let us content ourselves with seeing that  $O(d)$  is a compact topological space equipped with a finite measure  $\mu$  with total mass 1 satisfying

$$(4.150) \quad \int_{O(d)} \varphi(AB) d\mu(B) = \int_{O(d)} \varphi(BA) d\mu(B) = \int_{O(d)} \varphi(B) d\mu(B)$$

for all  $B \in O(d)$  and for all continuous functions  $f : O(d) \rightarrow \mathbb{C}$ . Set

$$(4.151) \quad \bar{f}(x) = \int_{O(d)} f(Ax) d\mu(A).$$

As we see below  $\mu$  serves as an averaging tool.  $f$  and  $\bar{f}$  have the same integral over  $\mathbb{R}^d$ . Indeed, we have

$$\int_{\mathbb{R}^d} \bar{f}(x) dx = \int_{\mathbb{R}^d} \left( \int_{O(d)} f(Ax) d\mu(A) \right) dx = \int_{O(d)} \left( \int_{\mathbb{R}^d} f(Ax) dx \right) d\mu(A).$$

Note that  $O(A)$  leaves  $dx$  unchanged. Therefore, the integral of the most right-hand side does not depend on  $A \in O(d)$  and hence  $A$  can be replaced by  $\text{id}_{\mathbb{R}^d}$ . We deduce from this observation

$$\int_{\mathbb{R}^d} \bar{f}(x) dx = \int_{O(d)} \left( \int_{\mathbb{R}^d} f(x) dx \right) d\mu(A) = \mu(O(d)) \cdot \int_{\mathbb{R}^d} f(x) dx = \int_{\mathbb{R}^d} f(x) dx.$$

Note that  $\bar{f}$  is radial, that is,  $\bar{f}(Ax) = \bar{f}(x)$  for all  $A \in O(d)$ . Therefore, if we apply the special case, then we obtain

$$(4.152) \quad \int_{\mathbb{R}^d} f(x) dx = \int_{\mathbb{R}^d} \bar{f}(x) dx = \int_0^\infty r^{d-1} \left( \int_{S^d} \bar{f}(r\omega) d\omega \right) dr.$$

Note that, from (4.150), we have

$$(4.153) \quad \int_{S^d} \bar{f}(r\omega) d\omega = \int_{S^d} \left( \int_{O(d)} f(rA\omega) d\mu(A) \right) d\omega = \int_{O(d)} \left( \int_{S^d} f(rA\omega) d\omega \right) d\mu(A).$$

Since  $S^d$  is a homogeneous space of  $O(d)$ , that is, any rotation matrix  $A$  does not distort the measure  $d\omega$  as well as  $dx$ , we have

$$(4.154) \quad \int_{S^d} f(rA\omega) d\omega = \int_{S^d} f(r\omega) d\omega.$$

Inserting this, we have

$$(4.155) \quad \int_{S^d} \bar{f}(r\omega) d\omega = \int_{O(d)} \left( \int_{S^d} f(r\omega) d\omega \right) d\mu(A) = \mu(O(d)) \cdot \left( \int_{S^d} f(r\omega) d\omega \right) = \int_{S^d} f(r\omega) d\omega.$$

Therefore, putting our observations together, we obtain the desired result.  $\square$

Let us make a closer look of the property of  $O(d)$ .

*Exercise 56.* Denote by  $M(d)$  the set of all  $d \times d$  matrices. Identify  $M(d)$  naturally with  $\mathbb{R}^{d^2}$ . Then via this identification, we can regard  $M(d)$  as a topological space.

- (1) Let  $A = \{a_{ij}\}_{i,j=1,\dots,d}$ . Then write out in full the condition for which  $A \in O(d)$  in terms of the components.
- (2) Equip  $O(d)$  with a topology induced by  $M(d)$ . Show that  $O(d)$  is a compact space and the multiplication  $(A, B) \in O(d) \times O(d) \mapsto AB \in O(d)$  is continuous.

*Exercise 57.* Let  $S$  be the surface area of the unit ball in  $\mathbb{R}^d$ . The aim of this exercise is to obtain again the precise value of  $S$ .

$$(1) \text{ Show that } \int_{\mathbb{R}^d} e^{-|x|^2} dx = S \int_0^\infty r^{d-1} e^{-r^2} dr.$$

- (2) By calculating the precise value of the left-hand side of the above formula, prove that

$$(4.156) \quad S = 2\pi^{\frac{d}{2}} \Gamma\left(\frac{d}{2}\right)^{-1}.$$

The author was taught this exercise by Y. Giga [501].

Notes and references for Chapter 3.

Section 2.

Section 3. Lebesgue investigated his own theory for integration in [301, 302].

Theorem 3.6 Theorem 3.7 Theorem 3.12 Theorem 3.15 Theorem 3.16 Theorem 3.17 Theorem 3.19 Theorem 3.21 Theorem 3.24 Theorem 3.27 Theorem 3.28 Theorem 3.29 Theorem 3.30 Theorem 3.35 Theorem 3.42 Theorem 3.49 Theorem 3.57 Theorem 3.60 Theorem 3.62 Theorem 3.64

G. Fubini found out Theorem 3.68 in [197].

Theorem 4.2 was obtained by Minkowski in [334].

Theorem 4.4 was obtained by O. Hölder in [237]. Hölder was mainly concerned with number theory. He considered this inequality in view of the work by Herr. L. Rogers [402]. We remark that Rogers obtained the Hölder inequality in part. Hölder obtained this inequality from the convex inequality. Note that

$$(4.157) \quad \varphi((1-t)x + ty) \leq (1-t)\varphi(x) + t\varphi(y), \quad 0 \leq t \leq 1$$

for  $x, y \in \mathbb{R}$  if  $\varphi : (0, \infty) \rightarrow (0, \infty)$  is  $C^2$  and  $\varphi'' \geq 0$ . He used this inequality with  $\varphi(x) = x^m$  to obtain the Hölder inequality, which is not so difficult nowadays.

Theorem 4.5

The proof of Theorem 4.6 has a little history. In [406] F. Riesz proved it for  $p = 2$ . The passage to general  $p$  was made afterward also by him in [407].

Theorem 4.8 Theorem 4.9 Theorem 4.11 Theorem 4.12 Theorem 4.18 Theorem 4.19 Theorem 4.23 Theorem 4.24

Nikodym proved Theorem 4.28 in his paper [370]. I owe Professor Marcin Guest to his checking the presentation in English.

Section 4. Theorem 4.32

Theorem 4.33

Theorem 4.34

Theorem 4.35

Theorems 4.36–4.38 are due to Riesz, who found the theorem in 1909.

Theorem 4.39

Theorem 4.40

Theorem 4.44 Theorem 4.45 Theorem 4.46

## Part 4. The Fourier transform

We now turn to the definitions of the Fourier transform and Schwartz distributions. The Fourier transform is widely used in partial-differential equations. Here we present the definitions and investigate elementary properties.

### 5. $\mathcal{S}(\mathbb{R}^d)$ AND $\mathcal{S}'(\mathbb{R}^d)$

#### 5.1. Schwartz Space $\mathcal{S}(\mathbb{R}^d)$ .

Definition. First, we begin with the definitions.

**Definition 5.1.** The Schwartz function space  $\mathcal{S}(\mathbb{R}^d)$  is the set of functions given below :

$$(5.1) \quad \mathcal{S}(\mathbb{R}^d) = \{\varphi \in C^\infty(\mathbb{R}^d, \mathbb{C}) : x^\alpha D^\beta \varphi(x) \in L^\infty(\mathbb{R}^d) \text{ for all } \alpha, \beta \in \mathbb{N}_0^d\}.$$

Before we go further, let us exhibit some examples of elements in  $\mathcal{S}(\mathbb{R}^d)$ .

**Example 5.2.** Let  $\varphi(x) = e^{-x^2} \in \mathcal{S}(\mathbb{R}^d)$ . The function  $\varphi$  is called the Gaussian function.

*Exercise 58.* Prove that  $C_c^\infty \subset \mathcal{S}(\mathbb{R}^d)$ . In view of this and Example 5.2 we conclude that  $C_c^\infty$  is a proper subset of  $\mathcal{S}(\mathbb{R}^d)$ .

Topology of  $\mathcal{S}(\mathbb{R}^d)$ . Now we equip  $\mathcal{S}(\mathbb{R}^d)$  with a topology.

**Definition 5.3.** The topology of  $\mathcal{S}(\mathbb{R}^d)$  is defined as a topological vector space which are induced by the family of seminorms below :

$$(5.2) \quad p_{\alpha, \beta}(\varphi) := \sup_{x \in \mathbb{R}^d} |x^\alpha D^\beta \varphi(x)|,$$

where  $\alpha, \beta \in \mathbb{N}_0^d$  are multi-indices. According to Exercise 8, the topology of  $\mathcal{S}(\mathbb{R}^d)$  is the strongest topology such that  $p_{\alpha, \beta}$  is continuous for every  $\alpha, \beta \in \mathbb{N}_0^d$ . It is convenient to define

$$(5.3) \quad p_N(\varphi) := \sum_{\substack{\alpha, \beta \in \mathbb{N}_0^d \\ |\alpha|, |\beta| \leq N}} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N |\partial^\alpha \varphi(x)|$$

for  $N \in \mathbb{N}_0$ .

*Exercise 59.* Show that  $\mathcal{S}(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$  for all  $1 \leq p \leq \infty$ .

**Example 5.4.** Let  $U \subset \mathcal{S}(\mathbb{R}^d)$  be an open set and  $\varphi \in U$ . Then there exists  $N \in \mathbb{N}$  such that

$$(5.4) \quad \{\psi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi - \psi) < 1\} \subset U.$$

*Proof.* From the definition we can find finite collections  $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{N}$ ,  $\beta_1, \beta_2, \dots, \beta_k \in \mathbb{N}$  and  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k > 0$  such that

$$(5.5) \quad \{\psi \in \mathcal{S}(\mathbb{R}^d) : p_{\alpha_1, \beta_1}(\varphi - \psi) < \varepsilon_1, p_{\alpha_2, \beta_2}(\varphi - \psi) < \varepsilon_2, \dots, p_{\alpha_k, \beta_k}(\varphi - \psi) < \varepsilon_k\} \subset U.$$

If we set

$$(5.6) \quad N := \left[ \max(\varepsilon_1^{-1}(|\alpha_1| + |\beta_1|), \varepsilon_2^{-1}(|\alpha_2| + |\beta_2|), \dots, \varepsilon_k^{-1}(|\alpha_k| + |\beta_k|)) \right] + 1,$$

then  $N p_N(\varphi - \psi) < 1$ ,  $\psi \in \mathcal{S}(\mathbb{R}^d)$  implies  $p_{\alpha_j, \beta_j}(\varphi - \psi) < \varepsilon_j$  for all  $j = 1, 2, \dots, k$ .  $\square$



*Exercise 60.* Prove that the closure of  $\mathcal{S}(\mathbb{R}^d)$  in  $L^\infty(\mathbb{R}^d)$  is given by

$$\mathcal{S}(\mathbb{R}^d)^{\overline{L^\infty}} := \left\{ \varphi \in \text{BUC}(\mathbb{R}^d) : \lim_{|x| \rightarrow \infty} \varphi(x) = 0 \right\},$$

which is made up of uniformly continuous functions that decay at  $\infty$ .

To be accustomed with the definition, let us prove the following.

**Example 5.5.** Let  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then defining  $\varphi_t := \varphi(\cdot - t)$  for  $t \in \mathbb{R}^d$ , we have  $t \in \mathbb{R}^d \rightarrow \varphi_t \in \mathcal{S}(\mathbb{R}^d)$  is continuous.

*Proof.* We have only to show

$$(5.7) \quad \lim_{t \rightarrow 0} \varphi_t = \varphi$$

because of translation invariance of  $\mathcal{S}(\mathbb{R}^d)$ . Now we shall prove

$$(5.8) \quad \lim_{t \rightarrow 0} p_{\alpha, \beta}(\varphi - \varphi_t) = 0$$

for all  $\alpha, \beta \in \mathbb{N}_0^d$ . By replacing  $\partial^\beta \varphi$  with  $\varphi$ , we can assume  $\beta = 0$ . By the mean value theorem and the binomial expansion we have

$$(5.9) \quad \sup_{x \in \mathbb{R}^d} |x^\alpha(\varphi(x) - \varphi(x - t))| \leq t \sup_{x \in \mathbb{R}^d} \int_0^1 |x^\alpha \varphi(x - tu)| du \lesssim t p_{|\alpha|+1}(\varphi).$$

Thus, (5.8) is established. □

*Exercise 61.* Prove that

$$(5.10) \quad \sup_{t \in [0, 1]^d} p_N(\varphi_t) < \infty,$$

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

*Exercise 62.* Let  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then prove that

$$(5.11) \quad \lim_{t \rightarrow 0} \frac{1}{t}(\varphi - \varphi_{t e_j}) = \partial_j \varphi$$

in  $\mathcal{S}(\mathbb{R}^d)$ , where  $e_j = (\delta_{jk})_{k=1}^d$ , the  $j$ -th elementary vector.

**Definition 5.6.** A topological space  $(X, \mathcal{O}_X)$  is said to be metrizable, if there exists a distance function  $d : X \times X \rightarrow [0, \infty)$  so that the topology induced by  $d$  coincides with  $\mathcal{O}_X$ .

**Proposition 5.7.** *The space  $\mathcal{S}(\mathbb{R}^d)$  is metrizable.*

*Proof.* We set  $d(\varphi, \psi) := \sum_{j=1}^{\infty} \frac{1}{2^j} \min(p_j(\varphi - \psi), 1)$  for  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ .

$d$ -open sets are open sets in  $\mathcal{S}(\mathbb{R}^d)$  with respect to the original topology. Pick  $\psi \in B(\varphi, r)$  arbitrarily, where  $B(\varphi, r)$  denotes the  $d$ -open ball centered at  $\varphi$  of radius  $r > 0$ . Choose  $N$  large enough as to hold

$$(5.12) \quad \sum_{j=N+1}^{\infty} \frac{1}{2^j} < \frac{r - d(\varphi, \psi)}{2}.$$

Let  $\eta \in \mathcal{S}(\mathbb{R}^d)$  satisfy  $p_N(\eta - \psi) < \frac{r - d(\varphi, \psi)}{2}$ . Then by using

$$(5.13) \quad \min(a + b, c) \leq \min(a, c) + \min(b, c)$$

for all  $a, b, c > 0$ , we have

$$\begin{aligned}
d(\varphi, \eta) &= \sum_{j=1}^N \frac{1}{2^j} \min(p_j(\varphi - \eta), 1) + \sum_{j=N+1}^{\infty} \frac{1}{2^j} \min(p_j(\varphi - \eta), 1) \\
&\leq \sum_{j=1}^N \frac{1}{2^j} \min(p_j(\varphi - \psi) + p_j(\eta - \psi), 1) + \sum_{j=N+1}^{\infty} \frac{1}{2^j} \\
&\leq \sum_{j=1}^N \frac{1}{2^j} \min(p_j(\varphi - \psi), 1) + \sum_{j=1}^N \frac{1}{2^j} \min(p_j(\eta - \psi), 1) + \sum_{j=N+1}^{\infty} \frac{1}{2^j} \\
&< d(\varphi, \psi) + \frac{r - d(\varphi, \psi)}{2} + \frac{r - d(\varphi, \psi)}{2} = r.
\end{aligned}$$

Therefore,  $p_N(\eta - \psi) < \frac{r - d(\varphi, \psi)}{2}$  implies  $d(\varphi, \eta) < r$ . Therefore any  $d$ -ball is open in the original topology of  $r$ .

Any open set with respect to the original topology is  $d$ -open. In view of Example 5.4 it suffices to prove that  $A := \{\psi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi - \psi) < 1\}$  is  $d$ -open. Take an arbitrary point  $\psi \in A$ . We shall prove  $B(\psi, r) \subset A$  if  $r = 4^{-N-1}(1 - N p_N(\varphi - \psi))$ . Let  $\eta \in B(\psi, r)$ . Then

$$(5.14) \quad \frac{1}{2^N} p_N(\eta - \psi) < d(\psi, \eta) < r = 2^{-N-1}(1 - N p_N(\varphi - \psi)).$$

Therefore,  $p_N(\eta - \psi) \leq 2^{-1}(1 - N p_N(\varphi - \psi))$ . This implies  $\eta \in A$  because

$$(5.15) \quad N p_N(\varphi - \eta) < 2^{-2-N} N(1 - p_N(\varphi - \psi)) + N p_N(\varphi - \psi) < 2^{-1}(1 + N p_N(\varphi - \psi)) < 1.$$

Thus,  $A$  is  $d$ -open. □

*Exercise 63.* Prove (5.13).

We state a property of the distance space whose proof is left as an exercise for the readers.

*Exercise 64.* Let  $(X, d)$  be a metric space. Then  $F$  is closed precisely when every sequence in  $F$  has a limit in  $F$  whenever it converges.

## 5.2. Schwartz distribution $\mathcal{S}'$ .

We now turn to the topological dual of  $\mathcal{S}'(\mathbb{R}^d)$ . As we did in linear algebra, we use  $\mathbb{K}$  to denote  $\mathbb{R}$  or  $\mathbb{C}$ .

**Definition.**

We define  $\mathcal{S}'(\mathbb{R}^d)$  as the set of all continuous functionals on  $\mathcal{S}(\mathbb{R}^d)$ .

**Definition 5.8** (Schwartz distribution). Let  $\mathbb{K}$  be  $\mathbb{R}$  or  $\mathbb{C}$ . Denote by  $\text{Hom}_{\mathbb{K}}(\mathcal{S}(\mathbb{R}^d), \mathbb{K})$  the set of all linear mappings. The space  $\mathcal{S}'(\mathbb{R}^d)$  is the topological dual of  $\mathcal{S}(\mathbb{R}^d)$ , that is, the set of continuous linear mappings.

$$(5.16) \quad \mathcal{S}'(\mathbb{R}^d) := \{f \in \text{Hom}_{\mathbb{K}}(\mathcal{S}(\mathbb{R}^d), \mathbb{K}) : f \text{ is a continuous mapping} \}$$

or equivalently

$$\mathcal{S}'(\mathbb{R}^d) := \{f : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathbb{K} : \langle f, \varphi + \psi \rangle = \langle f, \varphi \rangle + \langle f, \psi \rangle, \langle f, a \cdot \varphi \rangle = a \cdot \langle f, \varphi \rangle, f \text{ is continuous} \}.$$

Later on it will turn out that this set is quite large.

**Example 5.9** (Dirac delta). Let  $a \in \mathbb{R}^d$ . It is easy to see that the evaluation mapping

$$(5.17) \quad \delta_a : \varphi \in \mathcal{S}(\mathbb{R}^d) \mapsto \varphi(a) \in \mathbb{K}$$

belongs to  $\mathcal{S}'(\mathbb{R}^d)$ . Indeed, we have

$$(5.18) \quad |\langle \delta_a, \varphi \rangle| \leq p_0(\varphi)$$

in words of the seminorm.

**Theorem 5.10.** A  $\mathbb{K}$ -linear mapping  $F$  from  $\mathcal{S}(\mathbb{R}^d)$  to  $\mathbb{K}$  is continuous if and only if there exists  $N \in \mathbb{N}$  so that

$$(5.19) \quad |F(\varphi)| \leq N p_N(\varphi)$$

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

*Proof.* “If” part is obvious, because  $F$  is continuous. To prove “only if” part, we observe

$$(5.20) \quad \{\varphi \in \mathcal{S}(\mathbb{R}^d) : |F(\varphi)| < 1\}$$

is an open set in  $\mathcal{S}(\mathbb{R}^d)$  containing 0. Therefore there exists  $N \in \mathbb{N}$  so that

$$(5.21) \quad \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi) < 1\} \subset \{\varphi \in \mathcal{S}(\mathbb{R}^d) : |F(\varphi)| < 1\}.$$

This implies  $|F(\varphi)| \leq 1$  whenever  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  satisfies  $N p_{2N}(\varphi) = \frac{1}{2}$ . Therefore homogeneity gives us

$$(5.22) \quad |F(\varphi)| \leq 2N p_{2N}(\varphi)$$

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

We give an example of elements in  $\mathcal{S}'(\mathbb{R}^d)$ .

**Definition 5.11.** A function  $h \in C^\infty(\mathbb{R}^d)$  is a tempered distribution, if for all  $\alpha \in \mathbb{N}_0^d$ , there exist constants  $c_\alpha, N_\alpha > 0$  such that

$$(5.23) \quad |\partial^\alpha h(x)| \leq c_\alpha \langle x \rangle^{N_\alpha}$$

for all  $x \in \mathbb{R}^d$ .

The next definition will give us a way to produce elements in  $\mathcal{S}'(\mathbb{R}^d)$ .

**Definition 5.12.** Given a tempered distribution  $h \in C^\infty(\mathbb{R}^d)$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$ , we define  $h \cdot f \in \mathcal{S}'(\mathbb{R}^d)$  by

$$(5.24) \quad \langle h \cdot f, \varphi \rangle := \langle f, h \cdot \varphi \rangle$$

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

**Example 5.13.** Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  and  $a \in \mathbb{R}$ . Then  $x^\alpha, \sin x \cdot f \in \mathcal{S}'(\mathbb{R}^d)$ ,  $\langle x \rangle^a f$  all make sense.

Topology of  $\mathcal{S}'(\mathbb{R}^d)$ . Now we equip  $\mathcal{S}'(\mathbb{R}^d)$  with a topology.

**Definition 5.14.** The topology of  $\mathcal{S}'(\mathbb{R}^d)$  is a topology generated by

$$(5.25) \quad \mathcal{U}_{f,\psi,r} := \{g \in \mathcal{S}'(\mathbb{R}^d) : |\langle f - g, \varphi \rangle| < r\}$$

is an open set for all  $f \in \mathcal{S}'(\mathbb{R}^d)$ ,  $r > 0$  and  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ .

Before we go further, let us familiarize ourselves with the topology above.

*Exercise 65.* Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence in  $\mathcal{S}'(\mathbb{R}^d)$ . Then prove that it is convergent to  $f \in \mathcal{S}'(\mathbb{R}^d)$ , precisely when

$$(5.26) \quad \lim_{j \rightarrow \infty} \langle f_j, \varphi \rangle = \langle f, \varphi \rangle$$

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

The following theorem is one of the most important theorems on Schwartz distributions.

**Theorem 5.15.** *Suppose that  $\{F_j\}_{j \in \mathbb{N}}$  is a sequence in  $\mathcal{S}'(\mathbb{R}^d)$  such that*

$$(5.27) \quad \lim_{j \rightarrow \infty} \langle F_j, \varphi \rangle$$

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

$$(5.28) \quad F(\varphi) := \lim_{j \rightarrow \infty} \langle F_j, \varphi \rangle,$$

*for  $\varphi$ . Then  $F \in \mathcal{S}'(\mathbb{R}^d)$  and  $F_j \rightarrow F$  as  $j \rightarrow \infty$ . Furthermore there exists  $N \in \mathbb{N}$  so that*

$$(5.29) \quad |\langle F_j, \varphi \rangle| \leq N p_N(\varphi)$$

*for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $j \in \mathbb{N}$ .*

*Proof.* We set

$$(5.30) \quad A_j := \{\varphi \in \mathcal{S}(\mathbb{R}^d) : |\langle F_j, \varphi \rangle - \langle F_k, \varphi \rangle| \leq 1 \text{ for all } k \geq j\}.$$

Then  $A_j$  is a closed set and  $\mathcal{S}(\mathbb{R}^d) = \bigcup_{j=1}^{\infty} A_j$ . We now apply the Baire category theorem (See Theorem 10.15 below) for  $\mathcal{S}(\mathbb{R}^d)$ . Since  $\mathcal{S}(\mathbb{R}^d)$  is metrizable and  $A_1 \subset A_2 \subset A_3 \subset \dots$ , it follows that  $A_j$  contains an interior point  $\varphi_0$  if  $j$  is large enough. Fix  $j$  such that  $A_j$  contains  $\varphi_0$  as an interior point.

Then there exists  $N \in \mathbb{N}$  so that

$$(5.31) \quad \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi - \varphi_0) < 1\} \subset A_j.$$

By symmetry we have

$$(5.32) \quad \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi + \varphi_0) < 1\} \subset A_j.$$

Since  $A_j$  is convex, it follows that

$$\begin{aligned} & \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi) < 1\} \\ & \subset \frac{1}{2} \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi - \varphi_0) < 1\} + \frac{1}{2} \{\varphi \in \mathcal{S}(\mathbb{R}^d) : N p_N(\varphi + \varphi_0) < 1\} \subset A_j. \end{aligned}$$

Thus, as before we have  $N p_N(\varphi) = \frac{1}{2}$  implies

$$(5.33) \quad |\langle F_j, \varphi \rangle - \langle F_k, \varphi \rangle| \leq 1$$

for all  $k \geq j$  and for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  with  $N p_N(\varphi) = \frac{1}{2}$ . A passage to the limit gives us that  $F - F_j$  and hence  $F$  itself is continuous. Taking into account the continuity of a finite set of functionals  $F_1, F_2, \dots, F_{j-1}$ , we can find  $N$  as in the theorem.  $\square$

In  $\mathcal{S}'(\mathbb{R}^d)$  we can readily interchange the order of the limit and differentiation, more precisely we have the following. Therefore, we can say that the differentiation in the sense of distribution differs totally from the usual differentiation.

**Theorem 5.16.** Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence in  $\mathcal{S}'(\mathbb{R}^d)$  convergent to  $f$ . Then

$$(5.34) \quad \lim_{j \rightarrow \infty} \partial^\alpha f_j = \partial^\alpha f$$

for all  $\alpha \in \mathbb{N}_0^d$ .

*Proof.* Despite its appearance this theorem is easy to prove. Indeed, pick a test function  $\varphi$ . Then we have

$$(5.35) \quad \lim_{j \rightarrow \infty} \langle \partial^\alpha f_j, \varphi \rangle = (-1)^{|\alpha|} \lim_{j \rightarrow \infty} \langle f_j, \partial^\alpha \varphi \rangle = (-1)^{|\alpha|} \langle f, \partial^\alpha \varphi \rangle = \langle \partial^\alpha f, \varphi \rangle,$$

which shows  $\lim_{j \rightarrow \infty} \partial^\alpha f_j = \partial^\alpha f$ .  $\square$

Regular distributions. First, we discuss a dense space in  $\mathcal{S}(\mathbb{R}^d)$ . For later consideration it is convenient to prove the following.

**Lemma 5.17** (Dyadic resolution). *There exists a family of compactly supported functions  $\{\varphi_j\}_{j \in \mathbb{N}_0}$  with the following properties.*

- (1)  $\text{supp}(\varphi_0) \subset B(4)$  and  $\text{supp}(\varphi_1) \subset B(8) \setminus B(1)$ .
- (2) There exists  $\eta \in C_c^\infty(\mathbb{R}^d)$  such that for every  $j \in \mathbb{N}$ , we have  $\varphi_j = \eta(2^{-j+1} \cdot)$ .
- (3)  $\sum_{j=0}^{\infty} \varphi_j \equiv 1$ .
- (4)  $\varphi_j \geq 0$  for all  $j \in \mathbb{N}_0$ .

*Proof.* Let us begin with a smooth function  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$(5.36) \quad c_0 \chi_{[-2,2]} \leq \eta \leq c_1 \chi_{[-3,3]}$$

for some  $c_0, c_1 > 0$ , whose explicit construction is left for readers as an exercise. Then define  $\varphi(x) := \eta(|x|)$  and

$$(5.37) \quad \varphi_j(x) = 2^{-jd} \varphi(2^{-j}x) - 2^{-(j-1)d} \varphi(2^{-(j-1)}x) \quad x \in \mathbb{R}^d.$$

Then  $\{\varphi_j\}_{j \in \mathbb{N}}$  satisfy the desired property.  $\square$

*Exercise 66.* Construct a smooth function  $\eta \in \mathcal{S}(\mathbb{R}^d)$  satisfying (5.36).

The following lemma is used frequently in our later consideration.

**Lemma 5.18.** *Suppose that  $\{\varphi_j\}_{j=0}^{\infty}$  is a family appearing in Lemma 5.17.*

- (1) For  $\psi \in \mathcal{S}(\mathbb{R}^d)$  we have  $\sum_{j=0}^J \varphi_j \cdot \psi \rightarrow \psi$  as  $J \rightarrow \infty$  in the topology of  $\mathcal{S}(\mathbb{R}^d)$ .
- (2) For  $f \in \mathcal{S}'(\mathbb{R}^d)$  we have  $\sum_{j=0}^J \varphi_j(D)f \rightarrow f$  as  $J \rightarrow \infty$  in the topology of  $\mathcal{S}'(\mathbb{R}^d)$ .

(1). Let  $m \in \mathbb{N}$ ,  $\beta \in \mathbb{N}_0^d$  taken arbitrarily. Then we have to show

$$(5.38) \quad \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \partial^\beta \left( \psi - \sum_{j=0}^J \varphi_j \cdot \psi \right) (x) \right| \rightarrow 0$$

as  $J \rightarrow \infty$ . However, since

$$(5.39) \quad \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \partial^\beta \left( \psi - \sum_{j=0}^J \varphi_j \cdot \psi \right) (x) \right| \leq 2^{-J} \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \partial^\beta \left( \psi - \sum_{j=0}^J \varphi_j \cdot \psi \right) (x) \right|,$$

it suffices to show that

$$(5.40) \quad \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \partial^\beta \left( \psi - \sum_{j=0}^J \varphi_j \cdot \psi \right) (x) \right| \leq c_{\alpha, \beta} < \infty.$$

This can be achieved easily by virtue of the inequality

$$(5.41) \quad \sup_{j \in \mathbb{N}_0} \sup_{x \in \mathbb{R}^d} \sup_{\substack{\gamma \in \mathbb{N}_0^d \\ \gamma \leq \beta}} |\partial^\gamma \varphi_j(x)| < \infty.$$

Thus, we have proved (1). □

(2). For the proof of (2) we take  $\psi \in \mathcal{S}(\mathbb{R}^d)$  and calculate

$$(5.42) \quad \lim_{J \rightarrow \infty} \left\langle \sum_{j=0}^J \varphi_j(D) f, \psi \right\rangle = \lim_{J \rightarrow \infty} \sum_{j=0}^J \langle \varphi_j \cdot \mathcal{F}f, \mathcal{F}^{-1}\psi \rangle = \langle \mathcal{F}f, \mathcal{F}^{-1}\psi \rangle = \langle f, \psi \rangle,$$

proving  $\sum_{j=0}^J \varphi_j(D) f \rightarrow f$  in  $\mathcal{S}'(\mathbb{R}^d)$  as  $J \rightarrow \infty$ . □

**Corollary 5.19.** *The space  $C_c^\infty(\mathbb{R}^d)$  is dense in  $\mathcal{S}(\mathbb{R}^d)$ .*

**Definition 5.20.** A locally integrable function  $f$  is said to belong  $\mathcal{S}'(\mathbb{R}^d)$ , if

$$(5.43) \quad \varphi \in C_c^\infty \mapsto \int_{\mathbb{R}^d} f(x) \varphi(x) dx \in \mathbb{C}$$

can be extended to a bounded linear functional on  $\mathcal{S}(\mathbb{R}^d)$ . If this is the case, identify  $f := F_f \in \mathcal{S}'(\mathbb{R}^d)$  by

$$(5.44) \quad \langle F_f, \varphi \rangle = \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} f(x) \varphi_j(x) dx.$$

Here  $\{\varphi_j\}_{j \in \mathbb{N}}$  is a family of  $C_c^\infty$  functions tending to  $\varphi$  as  $j \rightarrow \infty$ .  $f$  is sometimes identified with  $F_f$ .  $\mathcal{S}'(\mathbb{R}^d) \cap L_{\text{loc}}^1(\mathbb{R}^d) = L_{\text{loc}}^1(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d)$  is the set of all locally integrable functions satisfying the condition above. It is convenient to regard  $\mathcal{S}'(\mathbb{R}^d) \cap L_{\text{loc}}^1(\mathbb{R}^d)$  not only as a subset of  $L_{\text{loc}}^1(\mathbb{R}^d)$  but also as a subset of  $\mathcal{S}'(\mathbb{R}^d)$ .

**Example 5.21.** Let  $1 \leq p \leq \infty$ . Then  $L^p(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$  in the sense that every  $f \in L^p(\mathbb{R}^d)$  belongs to  $\mathcal{S}'(\mathbb{R}^d)$ . Furthermore  $f \in L^p(\mathbb{R}^d) \mapsto F_f \in \mathcal{S}'(\mathbb{R}^d)$  is continuous.

*Proof.* Pick a test function  $\varphi \in C_c^\infty(\mathbb{R}^d)$ . Then by Hölder's inequality we have

$$(5.45) \quad |\langle F_f, \varphi \rangle| = \left| \int_{\mathbb{R}^d} f(x) \varphi(x) dx \right| \leq \|f\|_p \cdot \|\varphi\|_{p'} \leq \|f\|_p \cdot pN(\varphi).$$

Here  $N$  is taken so that  $Np' > d + 1$ . Thus, the result is immediate. □

*Exercise 67.* Let  $1 \leq q \leq p < \infty$ . For an  $L_{\text{loc}}^q(\mathbb{R}^d)$ -function  $f$ , we define the quantity  $\|f\|_{\mathcal{M}_q^p}$  by

$$(5.46) \quad \|f\|_{\mathcal{M}_q^p} := \sup_{x \in \mathbb{R}^d, r > 0} |B(x, r)|^{\frac{1}{p} - \frac{1}{q}} \left( \int_{B(x, r)} |f|^q \right)^{\frac{1}{q}}.$$

The Morrey space  $\mathcal{M}_q^p(\mathbb{R}^d)$  is the set of all  $L^q(\mathbb{R}^d)$ -locally integrable functions  $f$  for which the norm  $\|f\|_{\mathcal{M}_q^p}$  is finite.

Prove that  $\mathcal{M}_q^p(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$  in the sense of Example 5.21.

**Definition 5.22.** Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then the support of  $f$  is the set of all points  $x \in \mathbb{R}^d$  failing the condition below: There exists  $r > 0$  such that

$$(5.47) \quad \langle f, \varphi \rangle = 0$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  with  $\text{supp}(\varphi) \subset B(x, r)$ .

**Proposition 5.23.** Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then for every  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  supported outside of  $\text{supp}(f)$  we have  $\langle f, \varphi \rangle = 0$ .

*Proof.* Let  $\eta$  be a function with  $\chi_{B(1)} \leq \eta \leq \chi_{B(2)}$ . Define  $\eta^{(j)} := \eta(j^{-1}\cdot)$  for  $j \in \mathbb{N}$ . Then we have

$$(5.48) \quad \lim_{j \rightarrow \infty} \eta^{(j)} \cdot \varphi = \varphi$$

in the topology of  $\mathcal{S}(\mathbb{R}^d)$ . Therefore, we can assume that  $\eta$  is compactly supported.

From the definition of  $\text{supp}(f)$  and the compactness of  $\text{supp}(\eta)$ , we can find a finite number of balls  $B_1, B_2, \dots, B_J$  such that

$$(5.49) \quad \text{supp}(\varphi) \subset B_1 \cup B_2 \cup \dots \cup B_J$$

and that  $\langle f, \psi \rangle = 0$  whenever  $\psi \in \mathcal{S}(\mathbb{R}^d)$  is supported on  $2B_j$  for some  $j = 1, 2, \dots, J$ . Let  $\tau_j$  be a function with  $\chi_{B_j} \leq \tau_j \leq \chi_{2B_j}$ . Set

$$(5.50) \quad \kappa_j := \frac{\varphi_j}{\sum_{k=1}^J \varphi_k + \prod_{k=1}^J (1 - \varphi_k)}.$$

Since the denominator of  $\kappa_j$  never vanishes,  $\kappa_j$  is a well-defined function. Furthermore, we have

$$(5.51) \quad \sum_{j=1}^J \kappa_j \equiv 1$$

on  $\text{supp}(\varphi)$ . Therefore, we obtain

$$(5.52) \quad \langle f, \varphi \rangle = \sum_{j=1}^J \langle f, \kappa_j \cdot \varphi \rangle = 0.$$

This is the desired result. □

**Theorem 5.24.** Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then  $\text{supp}(f) \subset \{0\}$  if and only if there exists a finite collection of coefficients  $\{a_\alpha\}_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq L}$  such that  $f = \sum_{|\alpha| \leq L} a_\alpha \partial^\alpha \delta_0$ .

*Proof.* The “if part” being obvious, we shall prove the “only if part”. From the assumption we can find a large integer  $N$  such that

$$(5.53) \quad |\langle f, \varphi \rangle| \leq N p_N(\varphi)$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Set  $E^\alpha(x) := x^\alpha e^{-|x|^2}$ ,  $\alpha \in \mathbb{N}_0^d$ . There exists a collection of coefficients  $\{k_\alpha\}_{|\alpha| \leq N}$  such that

$$(5.54) \quad \varphi(x) - \sum_{|\alpha| \leq N} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha(x) = O(|x|^{N+1})$$

and that  $k_\alpha = 0$  if  $|\alpha| \geq N$ . Let  $\kappa$  be a function satisfying  $\chi_{B(4) \setminus B(2)} \leq \kappa \leq \chi_{B(8) \setminus B(1)}$ . Then we have

$$(5.55) \quad \lim_{j \rightarrow \infty} p_N \left( \left( \varphi - \sum_{|\alpha| \leq N} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha \right) (1 - \kappa(2^j \cdot)) \right) = 0.$$

Therefore, it follows that

$$\begin{aligned} \langle f, \varphi \rangle &= \left\langle f, \sum_{\alpha \in \mathbb{N}_0^d} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha \right\rangle + \left\langle f, \varphi - \sum_{\alpha \in \mathbb{N}_0^d} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha \right\rangle \\ &= \left\langle f, \sum_{\alpha \in \mathbb{N}_0^d} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha \right\rangle + \lim_{j \rightarrow \infty} \left\langle f, \left( \varphi - \sum_{\alpha \in \mathbb{N}_0^d} k_\alpha \cdot \partial^\alpha \varphi(0) \cdot E^\alpha \right) \kappa(2^j \cdot) \right\rangle \\ &= \sum_{\alpha \in \mathbb{N}_0^d} \langle f, E^\alpha \rangle k_\alpha \partial^\alpha \varphi(0). \end{aligned}$$

Recall that the test function  $\varphi$  is selected arbitrarily. Therefore, we obtain

$$(5.56) \quad f = \sum_{\alpha \in \mathbb{N}_0^d} (-1)^{|\alpha|} \langle f, E^\alpha \rangle k_\alpha \partial^\alpha \delta_0.$$

This is the desired result.  $\square$

### 5.3. Convolution of $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ .

Now we take up the convolution operation  $f * g$  with  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $g \in \mathcal{S}'(\mathbb{R}^d)$ .

**Lemma 5.25** (Peetre's inequality). *For  $x, y \in \mathbb{R}^d$ ,  $\langle x + y \rangle \leq \sqrt{2} \langle x \rangle \cdot \langle y \rangle$ .*

*Proof.* The proof is simple and we leave it for readers as an exercise (Exercise 68).  $\square$

*Exercise 68.* Prove this inequality, using  $|x + y| \leq |x| + |y|$ .

**Theorem 5.26.**  *$f * g \in \mathcal{S}(\mathbb{R}^d)$  whenever  $f, g \in \mathcal{S}(\mathbb{R}^d)$ .*

*Proof.* Let  $\alpha \in \mathbb{N}_0^d$  and  $N \in \mathbb{N}$  with  $|\alpha| \leq N$ . Note that

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N |\partial^\alpha (f * g)(x)| &= \sup_{x \in \mathbb{R}^d} \langle x \rangle^N \int_{\mathbb{R}^d} |\partial^\alpha f(x - y)| \cdot |g(y)| dy \\ &\leq 2^N \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} \langle x - y \rangle^N |\partial^\alpha f(x - y)| \cdot \langle y \rangle^N |g(y)| dy \\ &\lesssim p_N(f) \cdot p_{N+d+1}(g). \end{aligned}$$

Thus, we conclude  $f * g \in \mathcal{S}(\mathbb{R}^d)$ .  $\square$

**Lemma 5.27.** *Let  $f, g \in \mathcal{S}(\mathbb{R}^d)$ . Then*

$$(5.57) \quad \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} f \left( * - \frac{j}{N} \right) g \left( \frac{j}{N} \right) = f * g,$$

where the convergence takes place in  $\mathcal{S}(\mathbb{R}^d)$ .

*Proof.* In view of the definition of the topology of  $\mathcal{S}$ , what we need to prove amounts to showing

$$(5.58) \quad \lim_{N \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} \partial^\alpha f \left( x - \frac{j}{N} \right) g \left( \frac{j}{N} \right) - \partial^\alpha f * g(x) \right| = 0$$



for all  $\alpha \in \mathbb{N}_0^d$  and  $N \in \mathbb{N}_0$ . Since  $\partial^\alpha f \in \mathcal{S}(\mathbb{R}^d)$ , the matters are reduced to showing

$$(5.59) \quad \lim_{N \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} f \left( x - \frac{j}{N} \right) g \left( \frac{j}{N} \right) - f * g(x) \right| = 0.$$

Note that

$$\begin{aligned} & f * g(x) - \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} f \left( x - \frac{j}{N} \right) g \left( \frac{j}{N} \right) \\ &= \sum_{j \in \mathbb{Z}^d} \int_{\frac{j}{N} + \frac{1}{N}[0,1]^d} \left( f(x-y)g(y) - f \left( x - \frac{j}{N} \right) g \left( \frac{j}{N} \right) \right) dy \\ &= \sum_{j \in \mathbb{Z}^d} \int_{\frac{j}{N} + \frac{1}{N}[0,1]^d} \left\{ \int_0^1 \frac{d}{dt} \left( f \left( x - t \left( y - \frac{j}{N} \right) - \frac{j}{N} \right) g \left( t \left( y - \frac{j}{N} \right) + \frac{j}{N} \right) \right) dt \right\} dy. \end{aligned}$$

By the chain rule and Peetre's inequality we have, whenever  $t \in [0, 1]$  and  $y \in \frac{j}{N} + \frac{1}{N}[0, 1]^d$ ,

$$(5.60) \quad \left| \langle x \rangle^m \frac{d}{dt} \left( f \left( x - t \left( y - \frac{j}{N} \right) - \frac{j}{N} \right) g \left( t \left( y - \frac{j}{N} \right) + \frac{j}{N} \right) \right) \right| \lesssim_{f,g} N^{-1} \langle y \rangle^{-d-1}.$$

Inserting this estimate, we have

$$(5.61) \quad \sup_{x \in \mathbb{R}^d} \langle x \rangle^m \left| f * g(x) - \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} f \left( x - \frac{j}{N} \right) g \left( \frac{j}{N} \right) \right| \lesssim N^{-1}.$$

Therefore, the lemma is proved.  $\square$

We want to define  $f * g$  for  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $g \in \mathcal{S}'(\mathbb{R}^d)$ . We defined

$$(5.62) \quad f * g(x) = \int_{\mathbb{R}^d} f(y)g(x-y) dy \quad (x \in \mathbb{R}^d)$$

if  $f, g \in \mathcal{S}(\mathbb{R}^d)$ . Note that (5.62) can be rephrased as

$$f * g(x) = \langle f, g(x - \cdot) \rangle \quad (x \in \mathbb{R}^d).$$

This justifies the definition below.

**Definition 5.28.** Let  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $g \in \mathcal{S}'(\mathbb{R}^d)$ . Then define a function  $f * g : \mathbb{R}^d \rightarrow \mathbb{C}$  by

$$(5.63) \quad f * g(x) := \langle g, f(x - \cdot) \rangle \quad (x \in \mathbb{R}^d).$$

Below we summarize the elementary properties of this operation.

**Theorem 5.29.** Let  $f, g \in \mathcal{S}(\mathbb{R}^d)$  and  $h \in \mathcal{S}'(\mathbb{R}^d)$ .

- (1)  $f * g$  defined in Definition 5.28 coincides with the one given for  $L^1(\mathbb{R}^d)$  functions.
- (2)  $f * h \in C^\infty(\mathbb{R}^d)$ . Furthermore, let  $\alpha \in \mathbb{N}_0^d$ . Then we have

$$(5.64) \quad \partial^\alpha (f * h) = (\partial^\alpha f) * h = f * \partial^\alpha h$$

and there exists  $c > 0$  and  $N = N_\alpha \in \mathbb{N}_0$  so that

$$(5.65) \quad \sup_{x \in \mathbb{R}^d} |\partial^\alpha f * h(x)| \lesssim \langle x \rangle^N$$

for all  $\alpha \in \mathbb{N}_0^d$ . In particular  $f * h \in \mathcal{S}'(\mathbb{R}^d)$ .

- (3)  $(f * g) * h = f * (g * h)$ .

*Proof.* (1) is trivial from the definition. For the proof of (2), it suffices to prove

$$(5.66) \quad \partial^\alpha(f * h) = f * \partial^\alpha h$$

when  $|\alpha| = 1$ . Once this is proved, by the definition of  $\partial^\alpha f$ , we obtain

$$(5.67) \quad \partial^\alpha(f * h) = (\partial^\alpha f) * h = f * \partial^\alpha h$$

when  $|\alpha| = 1$ . The passage to the general case can be achieved by induction. However, (5.66) is included in Exercise 62. To prove the remaining assertion of (2), we recall that there exists a constant  $N$  so that

$$(5.68) \quad |\langle h, \varphi \rangle| \lesssim p_N(\varphi)$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Let  $\varphi = f(x - \cdot)$ . Since  $p_N(\partial^\alpha f(x - \cdot)) \lesssim \langle x \rangle^N$ , it follows that

$$(5.69) \quad \sup_{x \in \mathbb{R}^d} |\partial^\alpha f * h(x)| \lesssim \langle x \rangle^N.$$

Finally we prove (3). To do this, we use Lemma 5.27. Thanks to this lemma, we have

$$\begin{aligned} (f * g) * h &= (g * f) * h = \left( \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} g \left( * - \frac{j}{N} \right) f \left( \frac{j}{N} \right) \right) * h \\ &= \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} g \left( * - \frac{j}{N} \right) * h \cdot f \left( \frac{j}{N} \right) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{j \in \mathbb{Z}^d} g * h \left( * - \frac{j}{N} \right) \cdot f \left( \frac{j}{N} \right) \end{aligned}$$

Now that we have verified (2), we are in the position of using the dominated convergence theorem to obtain

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

Thus, the proof is now complete.  $\square$

Convolution of  $\mathcal{S}'$  with compact support and  $\mathcal{S}'$ .

Here we deal with  $f * g$  with  $f \in \mathcal{S}'$  with compact support and  $g \in \mathcal{S}'$ . The idea is the same as before: We shall define the operation so that it is compatible with  $\mathcal{S} \leftrightarrow \mathcal{S}'$ .

**Lemma 5.30.** *Let  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$  with compact support. Then  $\varphi * f \in \mathcal{S}(\mathbb{R}^d)$ .*

*Proof.* Let  $\alpha, \beta \in \mathbb{N}_0^d$ . Choose  $R > 0$  so that  $\text{supp}(f) \subset B(R)$ . Then we have

$$(5.70) \quad |x^\alpha \partial^\beta \varphi * f(x)| = |y^\alpha \langle \partial^\beta f, (x - \cdot)^\alpha \varphi(x - \cdot) \rangle| \lesssim \sum_{\gamma \leq \alpha} |\langle y^\gamma \partial^\beta f, (x - y)^{\alpha - \gamma} \varphi(x - y) \rangle| \lesssim 1,$$

because  $\sup_{x \in \mathbb{R}^d} p_{\delta, \delta'}(y^{\alpha - \gamma} \varphi) = \sup_{y \in B(x, R)} |y^{\delta'} \partial^\delta (y^{\alpha - \gamma} \varphi(y))| < \infty$ . Therefore, the proof is complete.  $\square$

**Corollary 5.31.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  with compact support and  $N \in \mathbb{N}$ . Then there exists  $M \in \mathbb{N}$  such that*

$$(5.71) \quad p_N(\varphi * f) \lesssim p_M(\varphi)$$

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

Note that

$$(5.72) \quad f * g \in \mathcal{S}(\mathbb{R}^d)$$

as long as  $f \in \mathcal{S}'(\mathbb{R}^d)$  and  $g \in \mathcal{S}(\mathbb{R}^d)$ . Let  $f, g \in \mathcal{S}'(\mathbb{R}^d)$ . The formula (5.72) enables us to define  $f * g$  when  $f$  is compactly supported.

**Definition 5.32.** Given  $f, g \in \mathcal{S}'(\mathbb{R}^d)$  with  $f$  compactly supported, one defines  $f * g \in \mathcal{S}'(\mathbb{R}^d)$  by

$$(5.73) \quad \langle f * g, \varphi \rangle := \langle g, f * \varphi(\cdot) \rangle$$

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

In analogy with the operation  $\mathcal{S} * \mathcal{S}'$  that we have the following.

**Proposition 5.33.** Let  $f_1, f_2 \in \mathcal{S}'(\mathbb{R}^d)$  be compactly supported and  $g \in \mathcal{S}'(\mathbb{R}^d)$ . Suppose further that  $f \in C_c^\infty(\mathbb{R}^d)$ .

- (1)  $f_1 * f_2$  is compactly supported.
- (2)  $f_1 * (f_2 * g) = (f_1 * f_2) * g$
- (3) The operation  $f * g$  agrees with the one  $f * g$  defined for  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $g \in \mathcal{S}'(\mathbb{R}^d)$ .
- (4)  $\mathcal{F}[f_1 * g] = (2\pi)^{\frac{d}{2}} \mathcal{F}(f_1) \mathcal{F}(g)$ .

*Exercise 69.* Prove Proposition 5.33.

## 6. SCHWARTZ DISTRIBUTION ON THE TORUS $\mathbb{T}^d$

In this section we investigate the property of the Fourier series.

### 6.1. $2\pi$ -periodic functions.

A measurable function  $f$  on  $\mathbb{R}^d$  is  $2\pi$ -periodic, if it satisfies

$$(6.1) \quad f(x + 2\pi m) = f(x) \text{ dx-a.e. } x \in \mathbb{R}^d$$

for all  $m \in \mathbb{Z}^d$ . Hence it is natural to define that  $f \in \mathcal{S}'(\mathbb{R}^d)$  is  $2\pi$ -periodic when

$$(6.2) \quad \langle f, \varphi \rangle = \langle f, \varphi(\cdot - 2\pi m) \rangle$$

for all  $m \in \mathbb{Z}^d$  and  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . In this paragraph we investigate the Fourier series of such functions or distributions. It is convenient to start from a function spaces made up of nice functions.

**Definition 6.1.** The space  $\mathcal{D}(\mathbb{T}^d)$  is the set of all  $2\pi$ -periodic  $C^\infty(\mathbb{R}^d)$ -functions. Equip  $\mathcal{D}(\mathbb{T}^d)$  with a topology induced by  $\{p_\alpha\}_{\alpha \in \mathbb{N}_0^d}$ , where

$$(6.3) \quad p_\alpha(f) := \sup_{x \in \mathbb{T}^d} |\partial^\alpha f(x)|.$$

*Exercise 70.* Let  $f \in L^1(\mathbb{R}^d)$ . Set

$$(6.4) \quad F(x) := \sum_{m \in \mathbb{Z}^d} f(x - 2\pi m)$$

converges almost every  $x \in \mathbb{R}^d$  and  $F \in L^1(\mathbb{T}^d)$ .

*Exercise 71.* Show that  $\mathcal{D}(\mathbb{T}^d)$  is metrizable and complete with respect to its metric.

The topological dual can be defined in analogy with  $\mathcal{S}'(\mathbb{R}^d)$ .

**Definition 6.2.** The space  $\mathcal{D}'(\mathbb{T}^d)$  is the set of all continuous linear functionals on  $\mathcal{D}(\mathbb{T}^d)$ .

Next, we leave the set of periodic smooth functions for the set of functions having integrability.

**Definition 6.3.** Let  $1 \leq p \leq \infty$  and let  $\overline{\mathbb{K}}$  denote either  $[-\infty, \infty]$  or  $\mathbb{C}$ . Then define

$$(6.5) \quad L^p(\mathbb{T}^d) := \{f : \mathbb{R}^d \rightarrow \overline{\mathbb{K}} : f \text{ is } 2\pi\text{-periodic}, \|f\|_{L^p(\mathbb{T}^d)} < \infty\},$$

where the norm is defined by

$$(6.6) \quad \|f\|_{L^p(\mathbb{T}^d)} := \left( \int_{[0, 2\pi]^d} |f(x)|^p dx \right)^{\frac{1}{p}}.$$

If  $p = \infty$ , we modify the definition naturally. Define also

$$(6.7) \quad C(\mathbb{T}^d) := L^\infty(\mathbb{T}^d) \cap C(\mathbb{R}^d),$$

where the norm is defined by

$$(6.8) \quad \|f\|_{C(\mathbb{T}^d)} := \|f\|_{L^\infty(\mathbb{T}^d)}.$$

Let  $k \in \mathbb{N}$ . Then define

$$(6.9) \quad C^k(\mathbb{T}^d) := \{f \in C^k(\mathbb{R}^d) : \partial^\alpha f \in C(\mathbb{T}^d) \text{ for all } \alpha \in \mathbb{N}_0^d \text{ with } |\alpha| \leq k\}.$$

The norm is defined by

$$(6.10) \quad \|f\|_{C^k(\mathbb{T}^d)} := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq k}} \|\partial^\alpha f\|_{C(\mathbb{T}^d)}.$$

*Exercise 72.* Show that  $L^p(\mathbb{T}^d)$  with  $1 \leq p \leq \infty$  and  $C^k(\mathbb{T}^d)$  with  $k \in \mathbb{N}_0$  is a Banach space.

## 6.2. Fourier series.

Given  $f \in L^1(\mathbb{T}^d)$ , we intend to expand it into a series of the form

$$(6.11) \quad \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot x}.$$

If this is possible and the convergence takes place in  $L^1(\mathbb{T}^d)$ , then we have

$$(6.12) \quad a_k = \frac{1}{(2\pi)^d} \int_{[0, 2\pi]^d} f(x) e^{-ik \cdot x} dx.$$

With this motivation, below we write

$$(6.13) \quad f(x) \simeq \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot x},$$

where  $a_k = \frac{1}{(2\pi)^d} \int_{[0, 2\pi]^d} f(x) e^{-ik \cdot x} dx$ . Of course, the formal equality  $\simeq$  is actually  $=$  under some additional assumption.

We begin with presenting some negative and motivating results. To construct some (surprising) counterexamples, let us obtain information of the function

$$(6.14) \quad \sigma_N(x) := \frac{1 - \cos Nx}{4\pi N \sin^2 \frac{x}{2}} \quad (x \in \mathbb{R}).$$

The function  $\sigma_N$  is called the Fejer kernel.

**Lemma 6.4.** *The Fejer kernel  $\sigma_N$  enjoys the following properties.*

$$(1) \quad \int_0^{2\pi} \sigma_N(x) dx = 1.$$

(2)  $\sigma_N \geq 0$ .

(3) For each  $\delta \in (0, \pi)$ ,  $\sup_{\delta < x < 2\pi - \delta} \sigma_N(x) \leq \frac{1}{2\pi N} \operatorname{cosec}^2 \frac{\delta}{2}$ .

*Proof.* What is not clear is (1) only. The remaining assertion are readily deduced from the very definition of  $\sigma_N$ . To prove (1) we reverse the argument of the previous lemma. Then we notice

$$(6.15) \quad \sigma_N(x) = \frac{1}{2\pi N} \sum_{j=0}^{N-1} \sum_{k=-j}^j \exp(ik \cdot x).$$

Therefore, after integration over  $[0, 2\pi]$ , the terms for  $k = 0$  survive and the other terms vanish.

Hence it follows that  $\int_0^{2\pi} \sigma_N(x) dx = 1$ . □

**Theorem 6.5.** *There exists  $f \in L^1(\mathbb{T})$  whose Fourier series does not converge in  $L^1(\mathbb{T})$ .*

*Proof.* Observe that the partial sum  $\sum_{k=-j}^j a_k e^{ik \cdot x}$  can be written as  $S_j * f(x)$ , where the kernel is given by  $S_j(x) := \operatorname{cosec} \left( \frac{x}{2} \right) \sin \left\{ \left( j + \frac{1}{2} \right) x \right\}$ . Let us denote by  $S_j$  the convolution operator  $f \mapsto S_j * f$ , as well. Then we have

$$(6.16) \quad \|S_j\|_{B(L^1(\mathbb{T}))} \geq \frac{\|S_j * \sigma_N\|_{L^1(\mathbb{T})}}{\|\sigma_N\|_{L^1(\mathbb{T})}} = \|S_j * \sigma_N\|_{L^1(\mathbb{T})},$$

for all  $N \in \mathbb{N}$ . However, letting  $N \rightarrow \infty$ , we know  $S_j * \sigma_N$  converges to a function  $S_j$  in  $L^1(\mathbb{T})$ . Hence we obtain

$$(6.17) \quad \|S_j\|_{L^1(\mathbb{T})} \leq \|S_j\|_{B(L^1(\mathbb{T}))}.$$

Here  $B(L^1(\mathbb{T}))$  is the set of all linear transforms  $A$  on  $L^1(\mathbb{T})$  for which

$$\|A\|_{B(L^1(\mathbb{T}))} := \sup\{\|Af\|_{L^1(\mathbb{T})} : \|f\|_{L^1(\mathbb{T})} = 1\}.$$

In analogy, we can consider  $B(L^p(\mathbb{T}))$  with  $1 \leq p \leq \infty$ .

However,  $\|S_j\|_{L^1(\mathbb{T})}$  grows like  $\log j$ , as is easily verified by noting a bilateral estimate

$$(6.18) \quad \operatorname{cosec} \left( \frac{t}{2} \right) \sin \left\{ \left( j + \frac{1}{2} \right) t \right\} \simeq \frac{1}{t} \sin \left\{ \left( j + \frac{1}{2} \right) t \right\}$$

for  $\frac{1}{3j} \leq t \leq 1$ , where the implicit constant does not depend on  $j \in \mathbb{N}$ . Therefore, we conclude  $\{S_j\}_{j \in \mathbb{N}}$  is not a bounded family not only in  $L^1(\mathbb{T})$  but also in  $B(L^1(\mathbb{T}))$ .

Suppose that  $S_j f \rightarrow f$  in the  $L^1(\mathbb{T})$ -topology for all  $f \in L^1(\mathbb{T})$ . Then by the uniformly bounded principle (Theorem 10.19 below) the family  $\{S_j\}_{j \in \mathbb{N}}$  must be uniformly bounded in  $B(L^1(\mathbb{T}))$ . This is a contradiction to what we have obtained in the above paragraph. □

**Theorem 6.6** (du Bois-Reymond, 1876). *There exists  $f \in C(\mathbb{T})$  whose Fourier series does not converge in  $L^\infty(\mathbb{T})$ .*

*Proof.* As we have seen above, the partial sum is expressed as a convolution operator  $f \mapsto S_j * f$ . Let us consider

$$(6.19) \quad f \in C(\mathbb{T}) \mapsto S_j * f(0) = \int_0^{2\pi} S_j(t) f(t) dt.$$

The operator norm of the above operator is exactly  $\|S_j\|_{L^1(\mathbb{T})}$ . Therefore, we conclude that  $\{S_j\}_{j \in \mathbb{N}} \subset B(L^\infty(\mathbb{T}))$  is not uniformly bounded. As before, this fact ensures the existence of  $f \in C(\mathbb{T})$  such that  $S_j * f$  diverges.  $\square$

Cesaro means. In view of Theorems 6.5 and 6.6 it seems difficult to analyze the Fourier series directly. We adopt an indirect method to overcome the above difficulties. Let us assume that  $d = 1$  for the time being. Recall that

$$(6.20) \quad \lim_{j \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_j}{j} = \alpha$$

provided a sequence  $\{a_j\}_{j \in \mathbb{N}}$  satisfies  $\lim_{j \rightarrow \infty} a_j = \alpha$ .

Therefore, instead of dealing with the partial sum  $\sum_{|k| \leq N} a_k e^{-ik \cdot x}$  directly, we prefer to consider

$$(6.21) \quad \frac{1}{N} \sum_{j=0}^N \sum_{|k| \leq j} a_k \exp(ik \cdot x).$$

The next lemma gives us an expression of the above sum.

**Lemma 6.7.** *Given  $f \in L^1(\mathbb{T})$ , denote*

$$(6.22) \quad a_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) \exp(-ik \cdot x) dx.$$

*Then we have*

$$(6.23) \quad \frac{1}{N} \sum_{j=0}^{N-1} \sum_{|k| \leq j} a_k \exp(ik \cdot x) = \sigma_N * f(x)$$

*Proof.* First we write out the left-hand side in full.

$$(6.24) \quad \frac{1}{N} \sum_{j=0}^{N-1} \sum_{|k| \leq j} a_k \exp(ik \cdot x) = \frac{1}{2\pi N} \sum_{j=0}^N \sum_{|k| \leq j} \int_0^{2\pi} f(y) \exp(ik \cdot (x - y)) dy.$$

Calculating the geometric series, we obtain

$$(6.25) \quad \frac{1}{N} \sum_{j=0}^{N-1} \sum_{|k| \leq j} a_k \exp(ik \cdot x) = \frac{1}{2\pi N} \sum_{j=0}^{N-1} \int_0^{2\pi} \frac{\sin \left\{ \left( j + \frac{1}{2} \right) (x - y) \right\}}{\sin \frac{x - y}{2}} f(y) dy.$$

Observe that

$$\begin{aligned} \sin \frac{x - y}{2} \cdot \sum_{j=0}^{N-1} \sin \left\{ \left( j + \frac{1}{2} \right) (x - y) \right\} &= \frac{1}{2} \sum_{j=0}^{N-1} \cos[j(x - y)] - \cos[(j + 1)(x - y)] \\ &= \frac{1}{2} (1 - \cos N(x - y)). \end{aligned}$$

Therefore, we obtain

$$(6.26) \quad \frac{1}{N} \sum_{j=0}^{N-1} \sum_{|k| \leq j} a_k \exp(ik \cdot x) = \frac{1}{4\pi N} \int_0^{2\pi} \frac{1 - \cos N(x - y)}{\sin^2 \frac{x - y}{2}} f(y) dy.$$

This is the desired result.  $\square$

By “a polynomial” we mean a function of the form

$$(6.27) \quad f(X) = \sum_{\alpha} a_{\alpha} X^{\alpha},$$

where the sum is finite and the coefficients are complex numbers. The space  $\mathbb{C}[X]$  denotes the set of all polynomials. A trigonometric polynomial is a function such that it can be written as  $\varphi(x) = f(\exp(ix_1), \exp(ix_2), \dots, \exp(ix_d))$ .

In connection with the Fourier partial summation, let us prove the set of all trigonometric polynomials spans  $C(\mathbb{T}^d)$ .

**Theorem 6.8** (Weierstrass approximation theorem). *The finite linear span of  $\{\exp(ij \cdot x)\}_{j \in \mathbb{Z}^d}$  is a dense subspace of  $C(\mathbb{T}^d)$ .*

*Proof.* Let  $f \in C(\mathbb{T}^d)$ . Denote  $e_1 := (1, 0, 0, \dots, 0)$ . Then we have

$$(6.28) \quad |f(x) - \sigma_N * f(x)| \leq \int_0^{2\pi} \sigma_N(y_1) |f(x) - f(x - y_1 e_1)| dy_1,$$

where we defined  $\sigma_N * f(x) = \int_0^{2\pi} \sigma_N(y_1) f(x - y_1 e_1) dy_1$  in analogy with the convolution in  $\mathbb{R}^d$ . Fix  $\delta > 0$ . Then we have

$$\begin{aligned} & |f(x) - \sigma_N * f(x)| \\ & \leq \int_{-\delta}^{\delta} \sigma_N(y_1) |f(x) - f(x - y_1 e_1)| dy_1 + \int_{\delta}^{2\pi-\delta} \sigma_N(y_1) |f(x) - f(x - y_1 e_1)| dy_1 \\ & \leq \sup_{\substack{y, z \in \mathbb{R}^d \\ |y-z| \leq \delta}} |f(y) - f(z)| \cdot \int_{-\delta}^{\delta} \sigma_N(y_1) dy_1 + 2\|f\|_{\infty} \int_{\delta}^{2\pi-\delta} \sigma_N(y_1) dy_1 \\ & \leq \sup_{\substack{y, z \in \mathbb{R}^d \\ |y-z| \leq \delta}} |f(y) - f(z)| + \frac{\|f\|_{\infty}}{\pi N \sin^2 \frac{\delta}{2}}. \end{aligned}$$

Therefore, taking the supremum over  $x \in \mathbb{R}^d$ , we obtain

$$(6.29) \quad \|f - \sigma_N * f\|_{\infty} \leq \sup_{\substack{y, z \in \mathbb{R}^d \\ |y-z| \leq \delta}} |f(y) - f(z)| + \frac{\|f\|_{\infty}}{\pi N \sin^2 \frac{\delta}{2}}.$$

Letting  $N \rightarrow \infty$ , we are led to

$$(6.30) \quad \limsup_{N \rightarrow \infty} \|f - \sigma_N * f\|_{\infty} \leq \sup_{\substack{y, z \in \mathbb{R}^d \\ |y-z| \leq \delta}} |f(y) - f(z)|.$$

Since  $\delta > 0$  is still at our disposal, we conclude

$$(6.31) \quad \lim_{N \rightarrow \infty} \|f - \sigma_N * f\|_{\infty} = 0.$$

In view of this, we can approximate  $f$  with a function of the form

$$(6.32) \quad \sum_{j=-J}^J \exp(ij \cdot x_1) f_j(x'),$$

where we have written  $x' = (x_2, x_3, \dots, x_d)$ ,  $J \in \mathbb{N}$  and each  $f_j$  belongs  $C(\mathbb{T}^{d-1})$ . Iterating this approximation procedure for each variable, we see that the trigonometric polynomials form a dense subset in  $C(\mathbb{T}^d)$ .  $\square$

Convergence in  $L^2(\mathbb{T})$ . We collect some theorems on the Fourier series. Let us begin with the simplest case that  $p = 2$ .

**Theorem 6.9.** *Any function  $f \in L^2(\mathbb{T}^d)$  admits the following expansion.*

$$(6.33) \quad f(x) = \sum_{j \in \mathbb{Z}^d} a_j \exp(ij \cdot x),$$

where the convergence takes place in the topology of  $L^2(\mathbb{T}^d)$ .

*Proof.* It suffices to prove the finite linear span of  $\{\exp(ij \cdot x)\}_{j \in \mathbb{Z}^d}$  is a dense subspace of  $L^2(\mathbb{T})$ . The finite linear span of  $\{\exp(ij \cdot x)\}_{j \in \mathbb{Z}^d}$  is a dense subspace of  $C(\mathbb{T})$  by virtue of the Weierstrass theorem. Therefore, this result is immediate.  $\square$

Convergence in  $\mathcal{D}(\mathbb{T}^d)$ . With the result in  $L^2(\mathbb{T}^d)$  in mind, let us pass to the case when the functions are smooth.

**Theorem 6.10.** *Any function  $f \in \mathcal{D}(\mathbb{T}^d)$  admits the following expansion.*

$$(6.34) \quad f(x) = \sum_{j \in \mathbb{Z}^d} a_j \exp(ij \cdot x),$$

where the convergence takes place in the topology of  $\mathcal{D}(\mathbb{T}^d)$ .

*Proof.* We utilize identity

$$(6.35) \quad (1 - \Delta)^N \exp(-ij \cdot x) = \langle j \rangle^{2N} \exp(-ij \cdot x).$$

Then we have

$$(6.36) \quad a_j = \int_{\mathbb{T}^d} f(x) \exp(-ij \cdot x) dx = \langle j \rangle^{-2N} \int_{\mathbb{T}^d} f(x) \cdot (1 - \Delta)^N \exp(-ij \cdot x) dx.$$

If we carry out integration by parts, we obtain

$$(6.37) \quad a_j = \langle j \rangle^{-2N} \int_{\mathbb{T}^d} (1 - \Delta)^N f(x) \cdot \exp(-ij \cdot x) dx.$$

Therefore, we have  $|a_j| \leq c_N \langle j \rangle^{-2N}$ . Thus, we can say  $\{a_j\}_{j \in \mathbb{Z}^d}$  decays very rapidly. With this strong decay, we conclude that

$$(6.38) \quad \sum_{j \in \mathbb{Z}^d} a_j \exp(ij \cdot x)$$

converges in  $\mathcal{D}(\mathbb{T}^d)$ . Since we have shown in Theorem 6.9 that

$$(6.39) \quad \sum_{j \in \mathbb{Z}^d} a_j \exp(ij \cdot x)$$

converges to  $f \in L^2(\mathbb{T})$ ,  $f$  can be expanded by (6.34).  $\square$

The following formula is used in number theory and applied to harmonic analysis itself as well.

**Theorem 6.11 (Poisson).** *Let  $f \in \mathcal{S}(\mathbb{R}^d)$ . Then we have*

$$(6.40) \quad \sum_{j \in \mathbb{Z}^d} f(x - 2\pi j) = (2\pi)^{-\frac{d}{2}} \sum_{k \in \mathbb{Z}^d} \mathcal{F}f(k) \exp(ik \cdot x),$$

where the convergence takes place in  $\mathcal{D}(\mathbb{T}^d)$ .



*Proof.* We set

$$(6.41) \quad F(x) := \sum_{j \in \mathbb{Z}^d} f(x - 2\pi j) \quad (x \in \mathbb{R}^d).$$

Then by checking the uniform convergence of all partial derivatives, we conclude that  $F \in C^\infty(\mathbb{R}^d)$  and  $F$  is  $2\pi\mathbb{Z}^d$ -periodic. Therefore, we have

$$(6.42) \quad F(x) = \sum_{k \in \mathbb{Z}^d} a_k \exp(ik \cdot x),$$

where the coefficient  $a_k$  is given by

$$(6.43) \quad a_k = \frac{1}{(2\pi)^d} \int_{[0, 2\pi]^d} F(x) \exp(-ik \cdot x) dx = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(x) \exp(-ik \cdot x) dx = (2\pi)^{-\frac{d}{2}} \mathcal{F}f(k).$$

Inserting this formula, we obtain the desired result.  $\square$

Convergence in  $L^p(\mathbb{T})$  with  $1 < p < \infty$ . Despite the negative results above, if  $1 < p < \infty$ , then we have the following.

**Theorem 6.12.** *Let  $1 < p < \infty$ . Then the partial Fourier sum of  $f \in L^p(\mathbb{T})$  converges in  $L^p(\mathbb{T})$ .*

*Proof.* As we have verified, this is the case when  $p = 2$ . Since  $\mathcal{D}(\mathbb{T})$  is dense in  $L^p(\mathbb{T})$ , it suffices to prove

$$(6.44) \quad \|S_j f\|_p \lesssim \|f\|_p$$

independently of  $j$  and  $f$ . To see this we use the boundedness of conjugation. Let  $2 < p < \infty$  for the time being. Recall that we have obtained

$$(6.45) \quad f(t) = \sum_{j=1}^{\infty} a_j e^{ijt} \in L^p(\partial\mathbb{D}) \mapsto \tilde{f}(t) = \sum_{j=1}^{\infty} \operatorname{sgn}(j) a_j e^{ijt} \in L^p(\partial\mathbb{D})$$

is a bounded linear operator. Therefore, letting  $E_k : f \mapsto e^{ikt} f$ , we have

$$(6.46) \quad E_{-k}(E_k \tilde{f}) = \sum_{j=1}^{\infty} \operatorname{sgn}(j+k) a_j e^{ijt}.$$

is a bounded operator on  $L^p(\mathbb{T})$ . Since

$$(6.47) \quad T_j f := S_j f - \frac{E_{-k}(E_k \tilde{f}) + E_k(E_{-k} \tilde{f})}{2}$$

is a bounded operator whose norm is bounded by a constant independent of  $j$ , we conclude that  $\{S_j\}_{j \in \mathbb{N}}$  is a uniformly bounded operator on  $L^p(\mathbb{T})$  whenever  $2 \leq p < \infty$ . A passage to dual therefore gives us that this is the case when  $1 < p < 2$ .  $\square$

Although Theorem 6.6 presents a counterexample, under some more continuity condition, we can assert that the series converges absolutely. This is quantified by Theorem 6.13 below.

**Theorem 6.13** (Bernstein). *Suppose that  $F \in C(\mathbb{T})$  satisfies*

$$(6.48) \quad \|f\|_{C^\alpha(\mathbb{T})} := \|f\|_{C(\mathbb{T})} + \sup_{\substack{x, y \in \mathbb{R} \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

for  $\alpha > \frac{1}{2}$ . Then the Fourier series

$$(6.49) \quad F(x) := \sum_{j \in \mathbb{Z}} a_j \exp(2\pi i k \cdot x)$$

converges absolutely. Furthermore, we have

$$(6.50) \quad \sum_{j \in \mathbb{Z}} |a_j| \lesssim \|f\|_{C^\alpha(\mathbb{T})}.$$

*Proof.* First we decompose

$$(6.51) \quad \sum_{j \in \mathbb{Z}} |a_j| = |a_0| + \sum_{m \in \mathbb{N}_0} \sum_{2^m \leq |j| < 2^{m+1}} |a_j| \leq |a_0| + \sum_{m \in \mathbb{N}_0} 2^{\frac{m+2}{2}} \left( \sum_{2^m \leq |j| < 2^{m+1}} |a_j|^2 \right)^{\frac{1}{2}}.$$

Set  $h_m := \frac{2\pi}{3 \cdot 2^m}$  for  $m \in \mathbb{N}_0$ . Then, since  $|e^{i\theta} - 1| \sim \theta$  for all  $0 < \theta < 2$ , we have

$$\sum_{j \in \mathbb{Z}} |a_j| \lesssim |a_0| + \sum_{m \in \mathbb{N}_0} 2^{\frac{m}{2}} \left( \sum_{2^m \leq |j| < 2^{m+1}} |e^{ijh_m} - 1|^2 |a_j|^2 \right)^{\frac{1}{2}}.$$

Now we write  $e^{ijh_m} a_j - a_j$  out in full. By Theorem 6.10, we have

$$\begin{aligned} \sum_{j \in \mathbb{Z}} |a_j| &\lesssim \|f\|_{L^\infty(\mathbb{T})} + \sum_{m \in \mathbb{N}_0} 2^{\frac{m+2}{2}} \|f(\cdot + h_m) - f\|_2 \\ &\lesssim \|f\|_{C^\alpha(\mathbb{T})} + \sum_{m \in \mathbb{N}_0} 2^{\frac{m+2}{2}} h_m^{\frac{\alpha}{2}} \|f\|_{C^\alpha(\mathbb{T})} \\ &\lesssim \|f\|_{C^\alpha(\mathbb{T})}. \end{aligned}$$

This is the desired result.  $\square$

*Exercise 73.* Show that for all  $f \in \mathcal{D}'(\mathbb{T}^d)$  there exists  $N = N_f \in \mathbb{N}$  so that

$$(6.52) \quad |\langle f, \varphi \rangle| \leq \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq N} p_\alpha(\varphi)$$

for all  $\varphi \in D(\mathbb{T}^d)$ . Hint: Mimic the proof of Theorem 5.10.

Ogura-Shannon-Someya sampling theorem. As an application, we consider the expansion of the function in  $L^2(\mathbb{R}^d)$  whose Fourier transform is supported on  $Q(R)$ . Define a continuous function sinc so that  $\text{sinc}(t) = \frac{\sin t}{t}$   $t \in \mathbb{R} \setminus \{0\}$ .

**Theorem 6.14** (Ogura, Shannon-Someya). *Suppose that  $f \in L^2(\mathbb{R}^d)$  with  $\text{supp}(\mathcal{F}f) \subset Q(R)$ . Then*

$$(6.53) \quad f(x) = \sum_{m \in \mathbb{Z}^d} \mathcal{F}\left(\frac{\pi m}{R}\right) \text{sinc}(\pi(Rx - m))$$

in the sense of the  $L^2(\mathbb{R})$ -convergence.

*Proof.* We expand  $\mathcal{F}f$  into a Fourier series:

$$(6.54) \quad \mathcal{F}f(\xi) = \sum_{m \in \mathbb{Z}^d} a_m \exp\left(\frac{-\pi m \cdot \xi}{R}\right) \chi_{Q(R)}(\xi)$$

Here the coefficient is give by

$$(6.55) \quad a_m = \frac{1}{(2R)^d} \int_{Q(R)} \mathcal{F}(\xi) \exp\left(\frac{\pi m \cdot \xi}{R}\right) d\xi = \frac{1}{(2R)^{\frac{d}{2}}} f\left(\frac{\pi m}{R}\right).$$

We remark that, to obtain the second inequality, we used the fact that  $\mathcal{F}$  is supported on  $Q(R)$ .

We compute the inverse Fourier transform of (6.54). To do this we calculate

$$(6.56) \quad \mathcal{F}^{-1} \left[ \exp \left( \frac{-\pi m \cdot *}{R} i \right) \chi_{Q(R)} \right] = \frac{1}{(2R)^{\frac{d}{2}}} \int_{Q(R)} \exp \left( \frac{\pi(Rx - m) \cdot \xi}{R} i \right) d\xi.$$

Since  $\int_{-R}^R \cos \left( \frac{\pi A \cdot t}{R} \right) dt = 2R \frac{\sin(\pi A)}{\pi A} = 2R \operatorname{sinc}(\pi A)$ , it follows that

$$(6.57) \quad \mathcal{F}^{-1} \left[ \exp \left( \frac{-\pi m \cdot *}{R} i \right) \chi_{Q(R)} \right] = (2R)^{\frac{d}{2}} \prod_{j=1}^d \operatorname{sinc}(\pi(x_j - m_j)).$$

Inserting this formula, we obtain the desired result. □

*Exercise 74.* Show that the convergence in (6.53) takes place in  $L^\infty(\mathbb{R}^d)$ .

## 7. FOURIER TRANSFORM

On the torus  $\partial\mathbb{D}$ , we can expand functions into Fourier series. The Fourier transform is a tool with which to expand functions defined on  $\mathbb{R}^d$ .

**7.1. Definition and elementary properties.** The definition of the Fourier transform is made up of several steps. We begin with the simplest case. We can define the inverse Fourier transform simultaneously.

**Definition 7.1.** Define the Fourier transform and its inverse by

$$\begin{cases} \mathcal{F}f(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx \\ \mathcal{F}^{-1}f(x) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(\xi) e^{ix \cdot \xi} d\xi \end{cases}$$

if  $f$  is an integrable function.

**Remark 7.2.** For an integrable function  $f \in L^1(\mathbb{R}^d)$ , some prefer to define  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  by

$$\begin{cases} \mathcal{F}f(\xi) := \int_{\mathbb{R}^d} f(x) e^{-2\pi ix \cdot \xi} dx \\ \mathcal{F}^{-1}f(x) := \int_{\mathbb{R}^d} f(\xi) e^{2\pi ix \cdot \xi} d\xi \end{cases}$$

e.t.c. Although we have some plausible definitions, in this book we keep to Definition 7.1.

The Fourier transform turns out to be extremely important after we notice that it preserves  $L^2(\mathbb{R}^d)$ -functions. Indeed, the Fourier transform preserves  $L^2(\mathbb{R}^d)$ -norms. Namely, the Fourier transform is unitary. For this reason, it is important to define other unitary operators.

**Definition 7.3** (Translation, Modulation, Dilation). Define three types of unitary operators on  $L^2(\mathbb{R}^d)$  by

$$\begin{aligned} T_a f &:= f(\cdot - a) \\ M_\omega f &:= \exp(i \cdot \omega) f \\ D_\kappa f &:= \frac{1}{|\kappa|^{\frac{d}{2}}} f \left( \frac{\cdot}{\kappa} \right) \end{aligned}$$

for  $a \in \mathbb{R}^d$ ,  $\omega \in \mathbb{R}^d$  and  $\kappa \in \mathbb{R}$ .

*Exercise 75.* Show that  $T_y, M_\xi, D_\rho$  all preserve the  $L^2(\mathbb{R}^d)$ -norm and that

$$\begin{cases} T_y M_\xi = e^{-i\xi y} M_\xi T_y \\ M_\xi T_y = e^{i\xi y} T_y M_\xi \end{cases} \quad \begin{cases} T_y D_\rho = D_\rho T_y \\ D_\rho T_y = T_{\rho y} D_\rho \end{cases} \quad \begin{cases} M_\xi D_\rho = D_\rho M_{\rho\xi} \\ D_\rho M_\xi = M_{\xi/\rho} D_\rho \end{cases}$$

for all  $\xi, y \in \mathbb{R}^d$  and  $\rho > 0$ .

Having defined the Fourier transform, let us see what can be said for that. We begin with the following simple properties.

**Theorem 7.4.** For  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  we have

$$(7.1) \quad \mathcal{F} \left[ \frac{\partial \varphi}{\partial x_j} \right] (\xi) = i\xi_j \mathcal{F}\varphi(\xi)$$

$$(7.2) \quad \mathcal{F}[x_j \varphi](\xi) = i \frac{\partial \mathcal{F}\varphi}{\partial \xi_j}(\xi)$$

and, for all multiindices  $\alpha$  and  $\beta$ ,

$$(7.3) \quad |\xi^\alpha D^\beta \mathcal{F}\varphi(\xi)| \lesssim_{\alpha, \beta} \sup_{x \in \mathbb{R}^d} \left( \langle x \rangle^{|\beta|+d+1} \sum_{\gamma \in \mathbb{N}_0^d: \gamma \leq \alpha} |D^\gamma \varphi(x)| \right).$$

*Proof.* Equality (7.1) can be obtained if we carry out integration by parts. Meanwhile (7.2) is obtained by changing the order of integration and differentiation. Finally the inequality (7.3) follows from the first formula using formula for the differentiation of the product.  $\square$

Observe that (7.3) asserts the following:

**Corollary 7.5.** The Fourier transform and its inverse are continuous from  $\mathcal{S}(\mathbb{R}^d)$  to itself.

As for  $L^1$ -functions, we have the following results.

**Theorem 7.6** (Riemann-Lebesgue). Suppose that  $f \in L^1(\mathbb{R}^d)$ . Then we have the following.

$$(7.4) \quad \mathcal{F}f, \mathcal{F}^{-1}f \in \text{BUC}(\mathbb{R}^d).$$

$$(7.5) \quad \|\mathcal{F}f\|_\infty = \|\mathcal{F}^{-1}f\|_\infty \leq (2\pi)^{-\frac{d}{2}} \|f\|_1.$$

$$(7.6) \quad \lim_{|x| \rightarrow \infty} \mathcal{F}f(x) = \lim_{|x| \rightarrow \infty} \mathcal{F}^{-1}f(x) = 0.$$

*Proof.* First, (7.5) follows from the definition. Since  $\text{BUC}(\mathbb{R}^d)$  is closed and  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^1(\mathbb{R}^d)$ , (7.4) is immediate. Details are left as an exercise for readers.

As for (7.6) we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)$ , since  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^1(\mathbb{R}^d)$  and  $\text{BUC}(\mathbb{R}^d)$  is a closed subspace in  $L^\infty(\mathbb{R}^d)$ . (The subset of  $\text{BUC}(\mathbb{R}^d)$  with (7.6) forms a closed subspace.) If  $f \in \mathcal{S}(\mathbb{R}^d)$ , by the preceding theorem, we trivially have the assertion.  $\square$

*Exercise 76.* Show that  $\text{BUC}(\mathbb{R}^d)$  is closed in  $\text{BC}(\mathbb{R}^d)$ .

Gaussian kernel. Next we shall calculate the Fourier transform of the Gaussian kernel not only to exhibit an example of calculation but also to get a key formula of the proof of the inversion formula. Set

$$(7.7) \quad E(x) := \exp\left(-\frac{1}{2}|x|^2\right) \quad (x \in \mathbb{R}^d).$$

The function  $E$  is called the Gaussian function or the Gauss kernel.

**Theorem 7.7.** *Defining  $E$  by (7.7), we have  $\mathcal{F}[E] = E$ .*

*Proof.* We may assume by Fubini's theorem that  $d = 1$ . Thus we have to calculate

$$\mathcal{F}(E)(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} e^{-ix\xi} d\xi.$$

We have

$$(7.8) \quad (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} e^{-ix\xi} d\xi = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}(x-i\xi)^2} e^{-\frac{1}{2}\xi^2} d\xi.$$

By virtue of the line integral on the complex plane the integral in question can be rewritten as follows:

$$(7.9) \quad \int_{\mathbb{R}} e^{-\frac{1}{2}(x-i\xi)^2} e^{-\frac{1}{2}\xi^2} d\xi = \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}\xi^2} d\xi,$$

which can be expressed in terms of the Gamma function  $\Gamma$ . Therefore we obtain

$$(7.10) \quad (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} e^{-ix\xi} d\xi = E(\xi).$$

This is the desired result.  $\square$

Inverse formula. The following is a key formula to derive various important formulae from, whose proof is simple.

**Theorem 7.8** (The multiplication formula). *For  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ , we have*

$$(7.11) \quad \int_{\mathbb{R}^d} \varphi(x) \mathcal{F}\psi(x) dx = \int_{\mathbb{R}^d} \mathcal{F}\varphi(\xi) \psi(\xi) d\xi.$$

*Proof.* We write out both sides in full:

$$\begin{aligned} \text{L.H.S.} &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \varphi(x) \left( \int_{\mathbb{R}^d} \psi(\xi) e^{-ix \cdot \xi} d\xi \right) dx \\ \text{R.H.S.} &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \psi(\xi) \left( \int_{\mathbb{R}^d} \varphi(x) e^{-ix \cdot \xi} dx \right) d\xi. \end{aligned}$$

Since both  $\varphi$  and  $\psi$  are integrable, the result follows easily from Fubini's theorem.  $\square$

With this theorem in mind, we shall establish  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  are inverse to each other.

**Theorem 7.9.** *Suppose that  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then we have*

$$(7.12) \quad \mathcal{F}^{-1}(\mathcal{F}\varphi) = \mathcal{F}(\mathcal{F}^{-1}\varphi) = \varphi.$$

*Proof.* We prove  $\mathcal{F}^{-1}(\mathcal{F}\varphi) = \varphi$ ,  $\mathcal{F}(\mathcal{F}^{-1}\varphi) = \varphi$  being proved similarly. We set  $E_t(x) := E(tx)$ . Then we have by Theorem 7.8

$$(7.13) \quad \int_{\mathbb{R}^d} \mathcal{F}E_t(\xi) \varphi(\xi) d\xi = \int_{\mathbb{R}^d} \mathcal{F}\varphi(x) E_t(x) dx.$$

By Theorem 7.7 we have

$$(7.14) \quad \mathcal{F}E_t(\xi) = (2\pi)^{-\frac{d}{2}} t^{-d} E\left(\frac{\xi}{t}\right).$$

Inserting this identity we have

$$(7.15) \quad (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} t^{-d} E\left(\frac{\xi}{t}\right) \varphi(\xi) d\xi = \int_{\mathbb{R}^d} \mathcal{F}\varphi(x) E(tx) dx.$$

Note that

$$\begin{aligned} (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} t^{-d} E\left(\frac{\xi}{t}\right) \varphi(\xi) d\xi - \varphi(0) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} t^{-d} E\left(\frac{\xi}{t}\right) (\varphi(\xi) - \varphi(0)) d\xi \\ &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} E(\xi) (\varphi(t\xi) - \varphi(0)) d\xi. \end{aligned}$$

By virtue of the Lebesgue convergence theorem we see that this quantity tends to 0 as  $t \downarrow 0$ . Therefore, we obtain

$$(7.16) \quad \varphi(0) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}\varphi(\xi) d\xi,$$

if we let  $t \rightarrow 0$ .

A translation allows us to obtain

$$(7.17) \quad \varphi(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}[\varphi(\cdot + x)](\xi) d\xi = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}\varphi(\xi) e^{ix \cdot \xi} d\xi \quad (x \in \mathbb{R}^d),$$

which is the desired result.  $\square$

## 7.2. Fourier transform for Schwartz distributions.

Theorem 7.8 will be a good motivation for the definition of Fourier transform of Schwartz distribution.

**Definition 7.10.** For  $f \in \mathcal{S}'(\mathbb{R}^d)$  define  $\mathcal{F}f \in \mathcal{S}'(\mathbb{R}^d)$  by the formula

$$(7.18) \quad \langle \mathcal{F}f, \varphi \rangle := \langle f, \mathcal{F}\varphi \rangle, \quad \varphi \in \mathcal{S}(\mathbb{R}^d).$$

As we saw earlier,  $\mathcal{F}$  preserves  $\mathcal{S}(\mathbb{R}^d)$ . By duality the same can be said for the dual  $\mathcal{S}'(\mathbb{R}^d)$ .

**Theorem 7.11.** *The Fourier transform is a continuous linear mapping from  $\mathcal{S}'(\mathbb{R}^d)$  to itself.*

*Proof.* This follows from the fact that  $\mathcal{F}$  is continuous from  $\mathcal{S}(\mathbb{R}^d)$  to itself. In fact suppose that a sequence  $\{\varphi_k\}_{k \in \mathbb{N}}$  tends to  $\varphi$  in  $\mathcal{S}'(\mathbb{R}^d)$ . Then we have

$$(7.19) \quad \lim_{k \rightarrow \infty} \langle \mathcal{F}\varphi_k, \tau \rangle = \lim_{k \rightarrow \infty} \langle \varphi_k, \mathcal{F}\tau \rangle = \langle \varphi, \mathcal{F}\tau \rangle = \langle \mathcal{F}\varphi, \tau \rangle.$$

Thus we conclude  $\mathcal{F}$  is a continuous transform from  $\mathcal{S}'(\mathbb{R}^d)$  to itself.  $\square$

Plancherel's formula. The next formula is now just a corollary of the results in the previous section.

**Theorem 7.12** (Planchrel). *For  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ , we have*

$$(7.20) \quad \int_{\mathbb{R}^d} \mathcal{F}\varphi(\xi) \overline{\mathcal{F}\psi(\xi)} d\xi = \int_{\mathbb{R}^d} \varphi(x) \overline{\psi(x)} dx.$$

*Proof.* For the proof we apply  $\psi$  with  $\overline{\mathcal{F}\psi}$  to the right-hand-side of (7.20). Observe that the relation between taking conjugate and the Fourier transform is given by

$$(7.21) \quad \mathcal{F}[\overline{\mathcal{F}\psi}] = \mathcal{F}[\mathcal{F}^{-1}(\overline{\psi})] = \overline{\psi}$$

for  $\psi \in \mathcal{S}(\mathbb{R}^d)$ . If we insert this equality, then we obtain (7.20)  $\square$

Before we state the main theorem in this section, we recall a terminology from Hilbert space theory. A linear operator  $U$  in a Hilbert space  $\mathcal{H}$  is said to be a unitary, if  $U$  is bijective and

$$(7.22) \quad \|Ux\|_{\mathcal{H}} = \|x\|_{\mathcal{H}}$$

for all  $x \in \mathcal{H}$ .

The above theorem and Theorem 9.3, which is a well-known theorem in functional analysis, yields

**Theorem 7.13** (Planchrel). *We can extend  $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$  to an isometry in  $L^2(\mathbb{R}^d)$ .*

If the function is compactly supported, then its Fourier transform is not compactly supported. We quantify this fact.

**Definition 7.14.** Let  $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  and  $|x|f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ . Then define the center of  $f$  by

$$(7.23) \quad x^* := x_f^* = \frac{1}{\|f\|_2^2} \left( \int_{\mathbb{R}^d} x_j |f(x)|^2 dx \right)_{j=1}^d$$

and the width by

$$(7.24) \quad \Delta_x := \frac{1}{\|f\|_2} \left( \int_{\mathbb{R}^d} |x - x^*|^2 |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

Define  $\Delta_{\xi}$  analogously for  $\mathcal{F}f$ . One defines the time-frequency window of  $f$

$$(7.25) \quad \{x \in \mathbb{R}^d : \|x^* - x\|_{\infty} \leq \Delta_x\} \times \{\xi \in \mathbb{R}^d : \|\xi^* - \xi\|_{\infty} \leq \Delta_{\xi}\}.$$

Uncertainty principle asserts that  $\Delta_x \Delta_{\xi} \geq \frac{1}{2}$  and equality holds for

$$(7.26) \quad \varphi_c(x) = \exp\left(-\frac{|x|^2}{2c}\right)$$

with  $c > 0$ .

*Exercise 77.* Show that

$$(7.27) \quad x_{M_{\xi} T_b w}^* = x_w^* + b$$

for  $\xi, b \in \mathbb{R}^d$ .

*Exercise 78.* Show that

$$(7.28) \quad \langle f, M_{\xi} T_{\omega} w \rangle = \langle \mathcal{F}f, T_{\xi} M_{-\omega} \mathcal{F}w \rangle.$$

In particular  $\xi_{M_{\xi} T_b w}^* = \xi_w^* + b$  for  $\xi, b \in \mathbb{R}^d$ .

*Exercise 79.* Let  $\psi \in \mathcal{S}$ . Show that

$$(7.29) \quad x_{D_a \psi}^* = a x_{\psi}^*, \xi_{D_a \psi}^* = a^{-1} \xi_{\psi}^*, \Delta_{x, D_a \psi} = a \Delta_{x, \psi}, \Delta_{\xi, D_a \psi} = a^{-1} \Delta_{\xi, \psi}$$

for  $a > 0$  and  $\xi \in \mathbb{R}^d$ .

Convolution and Fourier transform. If  $f, g \in \mathcal{S}(\mathbb{R}^d)$ , then it is easy to show that

$$(7.30) \quad \mathcal{F}[f * g] = (2\pi)^{\frac{d}{2}} \mathcal{F}f \cdot \mathcal{F}g.$$

In this paragraph we prove that (7.30) is the case even when  $g \in \mathcal{S}'(\mathbb{R}^d)$ .

**Theorem 7.15.** *Let  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then*

$$(7.31) \quad \mathcal{F}[\varphi * f] = (2\pi)^{\frac{d}{2}} \mathcal{F}\varphi \cdot \mathcal{F}f.$$

*Proof.* Pick a test function  $\psi \in \mathcal{S}(\mathbb{R}^d)$ . Then we have

$$(7.32) \quad \langle \mathcal{F}[\varphi * f], \psi \rangle = \langle \varphi * f, \mathcal{F}\psi \rangle$$

by the definition of the Fourier transform for  $\mathcal{S}'(\mathbb{R}^d)$ . Set  $\eta(x) = \mathcal{F}\psi(-x)$ . Thus, we have

$$(7.33) \quad \langle \mathcal{F}[\varphi * f], \psi \rangle = \langle \varphi * f, \eta(0 - *) \rangle = \eta * (\varphi * f)(0) = (\eta * \varphi) * f(0) = \langle f, \eta * \varphi(0 - *) \rangle.$$

Now we write  $\eta * \varphi(0 - *)$  by using the change of variables.

$$\eta * \varphi(0 - x) = \int_{\mathbb{R}^d} \eta(-x - y)\varphi(y) dy = \int_{\mathbb{R}^d} \mathcal{F}\psi(x + y)\varphi(y) dy = \int_{\mathbb{R}^d} \mathcal{F}\psi(y)\varphi(y - x) dy$$

Furthermore, we write out the Fourier transform out in full and use the definition of the Fourier transform.

$$\eta * \varphi(0 - x) = \int_{\mathbb{R}^d} \psi(y)\mathcal{F}\varphi(y)e^{-2\pi i x \cdot y} dy = (2\pi)^{\frac{d}{2}} \mathcal{F}(\psi \cdot \mathcal{F}\varphi)(x).$$

If we insert this to the above formula, then we obtain

$$(7.34) \quad \langle \mathcal{F}[\varphi * f], \psi \rangle = (2\pi)^{\frac{d}{2}} \langle f, \mathcal{F}[\psi \cdot \mathcal{F}\varphi] \rangle = (2\pi)^{\frac{d}{2}} \langle \mathcal{F}f, \psi \cdot \mathcal{F}\varphi \rangle = (2\pi)^{\frac{d}{2}} \langle \mathcal{F}\varphi \cdot \mathcal{F}f, \psi \rangle.$$

Since  $\psi$  is a test function chosen arbitrarily, we have the desired result.  $\square$

*Exercise 80.* Show (7.30).

**Band-limited distributions.** A band-limited distribution is one such that the support of its Fourier transform is compact. Later it turns out that this class of distributions is of importance in connection with the Littlewood-Paley theory.

**Definition 7.16** (Band-limited distributions). The frequency support of a distribution is the support of its Fourier transform. An element is said to be band-limited, if its frequency support is compact.

**Theorem 7.17.** Any band-limited distribution can be represented by a smooth distribution.

*Proof.* Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  be an element whose Fourier transform is compactly supported. Then pick  $\psi \in \mathcal{S}(\mathbb{R}^d)$  that equals 1 on  $\text{supp}(\mathcal{F}f)$ . Then we have, for every test function  $\varphi$

$$(7.35) \quad \langle f, \varphi \rangle = \langle \mathcal{F}f, \mathcal{F}^{-1}(\varphi) \rangle = \langle \psi \cdot \mathcal{F}f, \mathcal{F}^{-1}(\varphi) \rangle = \langle \mathcal{F}^{-1}[\psi \cdot \mathcal{F}f], \varphi \rangle.$$

Therefore, we obtain

$$(7.36) \quad f = \mathcal{F}^{-1}[\psi \cdot \mathcal{F}f] = \mathcal{F}^{-1}\psi * f.$$

The most right-hand side being smooth, we conclude  $f$  can be represented by a smooth function.  $\square$

Suppose that  $f \in \mathcal{S}'(\mathbb{R}^d)$  is compactly supported. Then we see that its Fourier transform  $\mathcal{F}f$  agrees with a function

$$(7.37) \quad F(\xi) := \langle f, \eta \cdot e^{\xi \cdot *} \rangle,$$

where  $\eta$  is a function that equals 1 on a neighborhood of  $\text{supp}(f)$ . If we define

$$(7.38) \quad G(\xi + i\eta) := \langle f, \eta \cdot e^{(\xi + i\eta) \cdot *} \rangle,$$

then we see that  $G$  is holomorphic. Recall that the definition of holomorphy in a domain in  $\mathbb{C}^d$  is given as follows:

**Definition 7.18.** A  $\mathbb{C}$ -valued function  $f$  on an open set  $\Omega$  in  $\mathbb{C}^d$  is said to be holomorphic, if it is holomorphic with respect to each variable.



At first glance there are many plausible definitions of the holomorphy. However, according to Hartogs, all definitions agree. Below we present a key observation due to Hartogs. For the proof we refer to [36].

**Theorem 7.19** (Hartogs). *Keep to the notation above. Then if  $f$  is holomorphic, then  $f$  is continuous on  $\Omega$ .*

From this theorem we can deduce that  $f$  can be expanded into the Taylor series provided  $f$  is holomorphic in the sense above.

**Theorem 7.20** (Paley-Wiener-(1)). *Let  $f \in \mathcal{S}(\mathbb{R}^d)$  and set  $g = \mathcal{F}^{-1}f$ . Then  $\text{supp}(g) \subset \overline{B(R)}$ , if and only if  $f$  is a restriction of a holomorphic function  $F : \mathbb{C}^d \rightarrow \mathbb{C}$  for all  $N \in \mathbb{N}$  there exists  $c_N > 0$  such that*

$$(7.39) \quad |F(z)| \leq c_N (1 + |z|)^{-N} e^{R|\text{Im}(z)|}$$

for all  $z \in \mathbb{C}$ . Here  $\text{Im}(z)$  denotes  $(\text{Im}(z_1), \dots, \text{Im}(z_d))$ .

*Proof.* The proof of necessity being straightforward, let us prove the sufficiency. We observe that

$$(7.40) \quad g(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{ix \cdot \xi} F(x) dx.$$

Fix  $y \in \mathbb{R}^d$  arbitrarily. The Cauchy integral theorem gives us

$$\begin{aligned} g(\xi) &= (2\pi)^{-\frac{d}{2}} \lim_{R \rightarrow \infty} \int_{-R}^R \left( \int_{-R}^R \left( \dots \left( \int_{-R}^R e^{ix \cdot \xi} F(x) dx_d \right) dx_{d-1} \right) \dots \right) dx_1 \\ &= (2\pi)^{-\frac{d}{2}} \lim_{R \rightarrow \infty} \int_{-R}^R \left( \int_{-R}^R \left( \dots \left( \int_{-R}^R e^{i(x+iy) \cdot \xi} F(x+iy) dx_d \right) dx_{d-1} \right) \dots \right) dx_1 \\ &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{i(x+iy) \cdot \xi} F(x+iy) dx. \end{aligned}$$

Suppose that  $\xi \notin \overline{B(R)}$ . Let  $y = k\xi$  with  $k > 1$ . Then we obtain

$$(7.41) \quad |g(\xi)| \lesssim e^{kR|\xi| - k|\xi|^2}$$

from inequality (7.39). Note that  $k$  can be chosen as large as we wish. Since  $|\xi| > R$ , it follows that  $g(\xi) = 0$ . □

**Theorem 7.21** (Paley-Wiener-(2)). *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  and set  $g = \mathcal{F}^{-1}f$ . Then  $\text{supp}(g) \subset \overline{B(R)}$ , if and only if  $f$  is a restriction of a holomorphic function  $F : \mathbb{C}^d \rightarrow \mathbb{C}$  and the estimate*

$$(7.42) \quad |F(z)| \lesssim (1 + |z|)^N e^{R|\text{Im}(z)|}$$

holds for all  $z \in \mathbb{C}$ . Here  $\text{Im}(z) = (\text{Im}(z_1), \dots, \text{Im}(z_d))$ .

*Proof.* The necessity being proved straightforwardly, let us concentrate on the sufficiency. For the proof it suffices to prove  $\text{supp}(g) \subset B(R + \varepsilon_0)$  for all  $\varepsilon_0 > 0$ .

Pick a bump function  $\tau \in \mathcal{S}(\mathbb{R}^d)$  with integral 1 such that  $\chi_{B(1)} \leq \tau \leq \chi_{B(2)}$ . Let  $0 < \varepsilon < \frac{1}{2}\varepsilon_0$  and set  $\tau_\varepsilon(x) = \frac{1}{\varepsilon^d} \tau\left(\frac{x}{\varepsilon}\right)$ . Let  $g_\varepsilon = \tau_\varepsilon * g$ .

It is not so hard to see that  $g_\varepsilon$  clears the conditions of Theorem 7.20 with  $B$  replaced by  $B + \varepsilon$ . Therefore, a passage to the limit gives us  $\text{supp}(g) \subset \overline{B(R)}$ . □

The following proposition shows a distribution and its Fourier transform cannot have compact support simultaneously unless it is zero. This fact has a lot to do with the uncertainty principle.

**Theorem 7.22.** *Assume that a distribution  $f \in \mathcal{S}'(\mathbb{R}^d)$  and its Fourier transform are compactly supported. Then  $f = 0$ .*

*Proof.* Since  $\text{supp}(f)$  has a compact support,  $f$  can be extended into a holomorphic function  $F$  on  $\mathbb{C}^d$ . Fix  $z' \in \mathbb{C}^{d-1}$ . Since  $F(\cdot, z')$  is a holomorphic function with compactly support,  $F$  is constant with respect to the first variable. In the same way we conclude that  $F$  is constant with respect to each variable. Therefore  $F$  is constant. However,  $F$  is compactly supported,  $F$  must be 0. Hence  $f$  itself is zero.  $\square$

### 7.3. Examples.

In this section we take up some examples of the preceding sections.

Fourier transform of  $e^{-|x|} : \mathbb{R} \rightarrow \mathbb{R}$ . We begin with a simple function whose Fourier transform is computable.

**Example 7.23.** Let  $d = 1$ . Let us calculate  $\mathcal{F}[e^{-|x|}](\xi)$ .

$$\mathcal{F}[e^{-|x|}](\xi) = \sqrt{\frac{1}{2\pi}} \int_{\mathbb{R}} e^{-|x|-ix\xi} dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-x} \cos(x\xi) dx = \frac{\sqrt{2}}{\sqrt{\pi}(\xi^2 + 1)},$$

where we have used

$$(7.43) \quad \int e^{-x} \cos(\alpha x) dx = \frac{1}{\alpha^2 + 1} (e^{-x} \cos(\alpha x) + \alpha e^{-x} \sin(\alpha x)) + C.$$

From this example we can tell that non-smoothness of the function is transformed by the Fourier transform into non-integrability.

*Exercise 81.* Check (7.43).

**Corollary 7.24.** *The (inverse) Fourier transform of  $f(\xi) := \frac{1}{\xi^2 + 1}$  is  $\sqrt{\frac{\pi}{2}} e^{-|x|}$ .*

This example reveals us that nonintegrability is transformed into nonsmoothness.

*Exercise 82.* Calculate directly the Fourier transform of  $f(\xi) := \frac{1}{\xi^2 + 1}$  by way of line integral in the complex plane.

Fourier transform of Dirac delta.

**Proposition 7.25.** *We have  $\mathcal{F}(\delta_a)(\xi) = (2\pi)^{-\frac{d}{2}} e^{-ia \cdot \xi}$  for  $a \in \mathbb{R}^d$ .*

*Proof.* Pick a test function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then we have

$$(7.44) \quad \langle \mathcal{F}(\delta_a), \varphi \rangle = \langle \delta_a, \mathcal{F}\varphi \rangle = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \varphi(\xi) e^{-ia \cdot \xi} d\xi.$$

Since  $\varphi$  is arbitrary, we obtain  $\mathcal{F}(\delta_a)(\xi) = (2\pi)^{-\frac{d}{2}} e^{-ia \cdot \xi}$ .  $\square$

Principal value. The principal value which we learnt in complex analysis as well is of much importance in harmonic analysis. In harmonic analysis we define the principal value as a distribution. Later we will investigate it systematically.

**Definition 7.26.** One defines  $p.v.\frac{1}{x} \in \mathcal{S}'(\mathbb{R}^d)(\mathbb{R})$  as a distribution given by

$$(7.45) \quad \left\langle p.v.\frac{1}{x}, \varphi \right\rangle := \lim_{\varepsilon \rightarrow 0} \int_{\{|x|>\varepsilon\}} \frac{\varphi(x)}{x} dx.$$

**Proposition 7.27.** *The distribution  $p.v.\frac{1}{x}$  is well-defined as an element of  $\mathcal{S}'(\mathbb{R}^d)(\mathbb{R})$ .*

*Proof.* In fact we have

$$\int_{|x|>\varepsilon} \frac{\varphi(x)}{x} dx = \int_{|x|>1} \frac{\varphi(x)}{x} dx + \int_{\varepsilon<|x|<1} \frac{\varphi(x)}{x} dx = \int_{|x|>1} \frac{\varphi(x)}{x} dx + \int_{\varepsilon<|x|<1} \frac{\varphi(x) - \varphi(0)}{x} dx,$$

where we have used the fact  $\frac{1}{x}$  is odd for the last equality. We rewrite the last term by using the mean value theorem.

$$\int_{\varepsilon<|x|<1} \frac{\varphi(x) - \varphi(0)}{x} dx = \int_{\varepsilon<|x|<1} \frac{\varphi(x)}{x} dx + \int_{\varepsilon<|x|<1} \left( \int_0^1 \varphi'(tx) dt \right) dx.$$

Thus, letting  $\varepsilon \downarrow 0$ , we have

$$(7.46) \quad \left\langle p.v.\frac{1}{x}, \varphi \right\rangle = \int_{|x|>1} \frac{\varphi(x)}{x} dx + \int_{|x|<1} \left( \int_0^1 \varphi'(tx) dt \right) dx.$$

From this formula it follows that

$$\left| \left\langle p.v.\frac{1}{x}, \varphi \right\rangle \right| \lesssim \sup_{x \in \mathbb{R}} |x|^2 |\varphi(x)| + \sup_{x \in \mathbb{R}} |\nabla \varphi(x)| \lesssim p_2(\varphi),$$

which shows  $p.v.\frac{1}{x} \in \mathcal{S}'(\mathbb{R}^d)$ . □

Now let us calculate the Fourier transform of  $p.v.\frac{1}{x}$ .

**Proposition 7.28.** *We define  $\text{sgn} := \chi_{(0,\infty)} - \chi_{(-\infty,0)}$ . In the sense of Schwartz distributions we have*

$$(7.47) \quad \mathcal{F} \left[ p.v.\frac{1}{x} \right] (\xi) = -\pi i \cdot \text{sgn}(\xi),$$

or equivalently

$$\left\langle \mathcal{F} \left( p.v.\frac{1}{x} \right), \varphi \right\rangle = -\pi i \cdot \int_{\mathbb{R}^d} \text{sgn}(\xi) \varphi(\xi) d\xi$$

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

*Proof.* Firstly we claim that there exists a constant  $M > 0$  so that

$$(7.48) \quad \left| \int_{\varepsilon<|x|<\varepsilon^{-1}} \frac{e^{2\pi i x \xi}}{x} dx \right| \leq M.$$

This can be seen once we rewrite the integral.

$$(7.49) \quad \int_{\varepsilon<|x|<\varepsilon^{-1}} \frac{e^{2\pi i x \xi}}{x} dx = -i \cdot \text{sgn}(\xi) \int_{\varepsilon<|x|<\varepsilon^{-1}} \frac{\sin(2\pi x |\xi|)}{x} dx.$$

As is well-known, the integral of the right-hand side of (7.49) remains bounded. Therefore (7.48) was established. (7.48) gives us also that

$$(7.50) \quad \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon < |x| < \varepsilon^{-1}} \frac{e^{2\pi i x \xi}}{x} dx = -2\pi i \cdot \operatorname{sgn}(\xi).$$

Now we pick a test function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then we have

$$\left\langle \mathcal{F} \left( p.v. \frac{1}{x} \right), \varphi \right\rangle = \left\langle p.v. \frac{1}{x}, \mathcal{F}\varphi \right\rangle = \lim_{\varepsilon \downarrow 0} \left\langle \frac{\chi_{\{\varepsilon < |x| < \varepsilon^{-1}\}}}{x}, \mathcal{F}\varphi \right\rangle.$$

Change the order of the integration by using the Fubini theorem. Since  $x$  is now restricted to  $(\varepsilon, \varepsilon^{-1})$ , we are able to do so.

$$\begin{aligned} \left\langle \mathcal{F} \left( p.v. \frac{1}{x} \right), \varphi \right\rangle &= \lim_{\varepsilon \downarrow 0} \int_{\varepsilon}^{\varepsilon^{-1}} \left( \int_{\mathbb{R}} \varphi(\xi) e^{-ix \cdot \xi} d\xi \right) \frac{dx}{x} \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \left( \int_{\varepsilon}^{\varepsilon^{-1}} e^{-ix \cdot \xi} \frac{dx}{x} \right) \varphi(\xi) d\xi. \end{aligned}$$

Having established (7.48), we are now in the position of using Lebesgue's dominated convergence theorem to conclude

$$(7.51) \quad \left\langle \mathcal{F} \left( p.v. \frac{1}{x} \right), \varphi \right\rangle = \int_{\mathbb{R}} \left( \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{\varepsilon^{-1}} e^{-ix \cdot \xi} \frac{dx}{x} \right) \varphi(\xi) d\xi = -\pi i \cdot \langle \operatorname{sgn}, \varphi \rangle,$$

which is the desired result.  $\square$

*Exercise 83.* In the course of the proof we have used the Lebesgue convergence theorem. In this exercise we justify this. Let  $a, b, c > 0$ . Establish the following.

$$(1) \quad \int_0^{\infty} \frac{\sin ax}{x} dx = \frac{\pi}{2}.$$

$$(2) \quad \text{There exists } M > 0 \text{ independent of } b \text{ and } c \text{ such that } \left| \int_b^c \frac{\sin x}{x} dx \right| \leq M.$$

Various operations. As we have defined the dilation for the distributions, many other operations are defined so that it is commutative with the embedding mapping  $f \in L^1_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d) \mapsto F_f \in \mathcal{S}'(\mathbb{R}^d)$ . For example, we can define the translation of distributions. In order that the definition is commutative with the above embedding we need to define

$$(7.52) \quad \langle f(\cdot - a), \varphi \rangle = \langle f, \varphi(\cdot + a) \rangle.$$

Homogeneous distributions. Now we deal with distributions like monomials in algebra. Any monomial  $p(X)$  satisfies  $p(\alpha X) = \alpha^k p(X)$ ,  $\alpha \in \mathbb{K}$  for some  $k \in \mathbb{N}$ . Of course  $k$  appearing here is the degree of  $p(X)$ . Now we consider the counterpart.

**Definition 7.29** (Dilation in  $\mathcal{S}'(\mathbb{R}^d)$ ).

(1) Let  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $t > 0$ . Then define  $f(t \cdot) \in \mathcal{S}'(\mathbb{R}^d)$  by

$$(7.53) \quad \langle f(t \cdot), \varphi \rangle = \left\langle f, \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) \right\rangle.$$

(2) A distribution  $f \in \mathcal{S}'(\mathbb{R}^d)$  is said to be homogenous of degree  $k \in \mathbb{R}$ , if  $f(t \cdot) = t^k f$  for all  $t > 0$ .

**Example 7.30.** Here are example of homogeneous Schwartz distributions.

(1) The elements 1 and  $\delta_0$  are homogeneous of degree 0.

(2) Let  $k > -d$ . Then  $|x|^k$  is homogeneous of degree  $k$ .

**Theorem 7.31.** *Suppose that  $0 < a < d$ . Then*

$$(7.54) \quad \mathcal{F}(|x|^{-a})(\xi) = \frac{\pi^{a-\frac{d}{2}} \Gamma\left(\frac{d-a}{2}\right)}{\Gamma\left(\frac{a}{2}\right)} |\xi|^{a-d}.$$

*Proof.* A passage to the limit allows us to concentrate on the case  $a \neq \frac{d}{2}$ . First, we shall prove that  $\mathcal{F}(|x|^{-a})$  is a regular distribution whenever  $\frac{d}{2} < a < d$ . Indeed, we have

$$(7.55) \quad |x|^{-a} = \chi_{B(1)} |x|^{-a} + \chi_{\mathbb{R}^d \setminus B(1)} |x|^{-a}$$

and

$$(7.56) \quad \chi_{B(1)} |x|^{-a} \in L^1(\mathbb{R}^d), \chi_{\mathbb{R}^d \setminus B(1)} |x|^{-a} \in L^2(\mathbb{R}^d).$$

Since  $\mathcal{F}$  sends  $L^1(\mathbb{R}^d)$  to  $L^\infty(\mathbb{R}^d)$  and  $L^2(\mathbb{R}^d)$  to itself, we see that  $\mathcal{F}(|x|^{-a})$  is a regular distribution.

If  $a < \frac{d}{2}$ , we are still able to conclude that  $\mathcal{F}(|x|^{-a})$  is a regular distribution.

Indeed, we now have

$$(7.57) \quad |x|^{-a} = \chi_{B(1)} |x|^{-a} + \chi_{\mathbb{R}^d \setminus B(1)} |x|^{-a}$$

and

$$(7.58) \quad \chi_{B(1)} |x|^{-a} \in L^2(\mathbb{R}^d), \chi_{\mathbb{R}^d \setminus B(1)} |x|^{-a} \in L^1(\mathbb{R}^d).$$

The same argument as before works to conclude that  $\mathcal{F}(|x|^{-a})$  is a regular distribution.

Let us denote  $g(\xi) := \mathcal{F}(|x|^{-a})(\xi)$ . Denote by  $O(d)$  the set of all isometries preserving the origin as before. Then we have

$$(7.59) \quad g(t\xi) = t^{a-n} g(\xi), g(A\xi) = g(\xi)$$

for all  $t > 0$  and  $A \in O(d)$ . Therefore we have

$$(7.60) \quad g(\xi) = K |\xi|^{a-n}$$

for some  $K > 0$ .

Thus, we are left with the task of determining the value of above  $K$  exactly.

This can be achieved by using the Plancherel formula: We make use of

$$(7.61) \quad \int_{\mathbb{R}^d} |x|^{-a} e^{-\pi|x|^2} dx = \int_{\mathbb{R}^d} \mathcal{F}(|x|^{-a}) \mathcal{F}(e^{-\pi|x|^2}) dx = K \int_{\mathbb{R}^d} |x|^{a-n} e^{-\pi|x|^2} dx.$$

Observe that  $\mathcal{F}(e^{-\pi|\cdot|^2})(\xi) = e^{-\pi|\xi|^2}$ . We write the integral in the polar coordinate. The result is

$$(7.62) \quad |S^{n-1}| \int_0^\infty r^{n-a} e^{-\pi r^2} \frac{dr}{r} = K |S^{n-1}| \int_0^\infty r^a e^{-\pi r^2} \frac{dr}{r}.$$

Putting  $s = r^2$ , then we have

$$(7.63) \quad \int_0^\infty s^{\frac{d-a}{2}} e^{-\pi s} \frac{ds}{s} = K \int_0^\infty s^{\frac{a}{2}} e^{-\pi s} \frac{ds}{s}.$$

Finally we change variables  $t = \pi s$ . Then

$$(7.64) \quad \pi^{\frac{d-a}{2}} \int_0^\infty s^{\frac{d-a}{2}} e^{-\pi s} \frac{ds}{s} = K \pi^{\frac{a}{2}} \int_0^\infty s^{\frac{a}{2}} e^{-\pi s} \frac{ds}{s}.$$

It can be rephrased as

$$(7.65) \quad \pi^{\frac{d-a}{2}} \Gamma\left(\frac{d-a}{2}\right) = K \pi^{\frac{a}{2}} \Gamma\left(\frac{a}{2}\right).$$

Arranging (7.65), we see that  $K$  can be determined as we wish.  $\square$

*Exercise 84.* Let  $a > -d = 1$ . Then calculate  $\mathcal{F}(|x|^a)$ .

**Radial distribution.** A function is said to be radial if it depends on the absolute value. That is a function  $f$  on  $\mathbb{R}^d$  is said to be radial, if

$$(7.66) \quad f(x) = F(|x|), \quad x \in \mathbb{R}^d$$

for some function  $F$  on  $[0, \infty)$ . Note that this can be rephrased as

$$(7.67) \quad f(Ax) = f(x), \quad x \in \mathbb{R}^d$$

for all  $A \in O(d)$ . Here  $O(d)$  is the set of all linear isometries on  $\mathbb{R}^d$  as before. With this definition in mind, we formulate the definition of radial distributions.

**Definition 7.32** (Radial distribution). Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Define  $f(A \cdot) \in \mathcal{S}'(\mathbb{R}^d)$  by

$$(7.68) \quad \langle f(A \cdot), \varphi \rangle = \langle f, \varphi(A^{-1} \cdot) \rangle$$

for  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . The distribution  $f$  is said to be radial, if

$$(7.69) \quad f(A \cdot) = f$$

for all  $A \in O(d)$ .

The Fourier transform of radial functions has the following expression.

**Theorem 7.33.** *If  $f \in \mathcal{S}'(\mathbb{R}^d) \cap L^1_{\text{loc}}(\mathbb{R}^d)$  is a radial function, then*

$$(7.70) \quad \mathcal{F}f(\xi) = (2\pi)^{\frac{d}{2}} |\xi|^{-\frac{d-2}{2}} \int_0^\infty F(t) t^{\frac{d}{2}} J_{\frac{d-2}{2}}(|\xi|t) dt,$$

where  $J_\nu$  is the Bessel function given by

$$(7.71) \quad J_\nu(x) = \frac{2 \left(\frac{x}{2}\right)^\nu}{\Gamma\left(\frac{1}{2}\right) \cdot \Gamma\left(\nu + \frac{1}{2}\right)} \int_0^1 (1-t^2)^{\nu-\frac{1}{2}} \cos xt dt.$$

*Proof.* An approximation allows us to assume that  $f \in \mathcal{S}(\mathbb{R}^d)$ . Write  $F(t) = f(te_1)$  for  $t > 0$ , where  $e_1 := (1, 0, 0, \dots, 0)$ . First, we observe

$$(7.72) \quad \mathcal{F}f(\xi) = \mathcal{F}f(|\xi|e_1), \quad \xi \in \mathbb{R}^d \setminus \{0\}$$

by using a rotation matrix  $A$ , which sends  $e_1$  to  $\frac{\xi}{|\xi|}$ . Of course, (7.72) is trivial for  $\xi = 0$ .

We calculate  $\mathcal{F}f(|\xi|e_1)$ . Let us write  $x = (x_1, x') \in \mathbb{R} \times \mathbb{R}^{d-1}$ . Then we have

$$\mathcal{F}f(|\xi|e_1) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(x_1, x') e^{-i|\xi|x_1} dx = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{d-1}} f(x_1, x') dx' \right) e^{-i|\xi|x_1} dx_1.$$

Let  $e'_2 = (1, 0, \dots, 0) \in \mathbb{R}^{d-1}$ . Now we use the polar coordinate formula to obtain

$$(7.73) \quad \mathcal{F}f(|\xi|e_1) = (2\pi)^{-\frac{d}{2}} \omega_{d-2} \int_{\mathbb{R}} \left( \int_0^\infty f(x_1, r e'_2) r^{d-2} dr \right) e^{-i|\xi|x_1} dx_1.$$

Here  $\omega_{d-2} = \frac{2\pi^{\frac{d-2}{2}}}{\Gamma(\frac{d-2}{2})}$ . We use the polar coordinate formula in  $\mathbb{R}^2$  to obtain

$$\begin{aligned} \mathcal{F}f(\xi) &= (2\pi)^{-\frac{d}{2}} \omega_{d-2} \int_{\mathbb{R}} \left( \int_0^\pi f(t \cos \theta, t \sin \theta e'_2) t^{d-2} \sin^{d-2} \theta d\theta \right) e^{-i|\xi|t \cos \theta} dt \\ &= (2\pi)^{-\frac{d}{2}} \omega_{d-2} \int_{\mathbb{R}} \left( \int_0^\pi F(t) t^{d-1} \sin^{d-2} \theta d\theta \right) e^{-i|\xi|t \cos \theta} dt. \end{aligned}$$

If we change the order of integration, then we have

$$(7.74) \quad \mathcal{F}f(\xi) = (2\pi)^{\frac{d}{2}} |\xi|^{-\frac{d-2}{2}} \int_0^\infty F(t) t^{\frac{d}{2}} J_{\frac{d-2}{2}}(|\xi|t) dt.$$

This is the desired result. □

#### 7.4. Fourier transform of measures.

In probability theory, the Fourier transform is very important and it has a lot to do with convergence of many types. Here we do not stick to finite measures with mass 1 and we consider finite complex measures. Finite complex measures can be regarded naturally as elements in  $\mathcal{S}'(\mathbb{R}^d)$ . Let  $M(\mathbb{R}^d)$  be the set of all signed measures.

The aim of this subsection is to investigate the Fourier transform of the elements in  $M(\mathbb{R}^d)$  in more depth. In connection with probability theory, we adopt the notation slightly different from the one that we have been adopting.

**Definition 7.34.** Let  $\mu \in M$ . Then define

$$(7.75) \quad \varphi_\mu(\xi) := \int_{\mathbb{R}^d} e^{ix \cdot \xi} dx \quad (\xi \in \mathbb{R}^d).$$

We prefer not to add the multiplicative constant  $(2\pi)^{-d/2}$ . Also, it is customary not to add  $-$ .

First of all let us investigate the relation between the convolution and the characteristic function. Here, motivated by the convolution of functions, we define the convolution of measures as follows:

**Definition 7.35.** Given two signed measures  $\mu, \nu \in M(\mathbb{R}^d)$ , we set

$$(7.76) \quad \mu * \nu(E) = \int_{\mathbb{R}^d} \chi_E(x+y) d\mu(x) d\nu(y).$$

It is easy to see that  $\mu * \nu \in M$  again. Furthermore, if  $f$  and  $g$  are  $L^1(\mathbb{R}^d, dx)$  functions, then the definition  $f * g \in L^1(\mathbb{R}^d, dx)$  and  $f * g \in M$  coincides in the sense of above embedding.

**Proposition 7.36.** Let  $\mu, \nu \in M(\mathbb{R}^d)$ . Then

$$(7.77) \quad \varphi_{\mu * \nu} = \varphi_\mu \cdot \varphi_\nu.$$

The next proposition explicitly shows  $\mu$  can be expressed by terms of  $\varphi_\mu$ .

**Proposition 7.37.** Let  $\mu \in M(\mathbb{R}^d)$ . Write  $\varphi(\xi) := \varphi_\mu(\xi)$  for  $\xi \in \mathbb{R}^d$ , the Fourier transform of  $\mu$ . Assume that  $I = \prod_{k=1}^d (a_k, b_k)$  is an open rectangular whose boundary  $\mu$  does not charge.

Then we have

$$(7.78) \quad \mu(I) = \left(\frac{1}{2\pi}\right)^d \lim_{T \rightarrow \infty} \int_{[-T, T]^d} \prod_{k=1}^d \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} \varphi(t) dt.$$

*Proof.* By definition and change the order of the integration we have

$$(7.79) \quad \begin{aligned} & \int_{[-T, T]^d} \prod_{k=1}^d \frac{e^{-ix_k a_k} - e^{-ix_k b_k}}{ix_k} \varphi(x) dx \\ &= \int_{[-T, T]^d} \prod_{k=1}^d \frac{e^{-ix_k a_k} - e^{-ix_k b_k}}{ix_k} \left( \int_{\mathbb{R}^d} e^{ix \cdot y} d\mu(y) \right) dx \\ &= \int_{\mathbb{R}^d} \left( \int_{[-T, T]^d} \prod_{k=1}^d \left( \frac{e^{-ix_k a_k} - e^{-ix_k b_k}}{ix_k} \right) e^{ix \cdot y} dx \right) d\mu(y) \\ &= \int_{\mathbb{R}^d} \left( \int_{[-T, T]^d} \prod_{k=1}^d \left( \frac{e^{-ix_k a_k} - e^{-ix_k b_k}}{ix_k} e^{ix_i y_k} \right) dx \right) d\mu(y). \end{aligned}$$

By using the Fubini theorem, we separate the variables. To analyze each factor we calculate

$$\int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} e^{ity} dt = \int_{-T}^T \frac{e^{-it(a-y)} - e^{-it(b-y)}}{it} dt = \int_{-T}^T \frac{\sin(t(a-y)) - \sin(t(b-y))}{t} dt.$$

Since

$$\lim_{T \rightarrow \infty} \int_{-T}^T \frac{\sin(t(a-y))}{t} dt = - \lim_{T \rightarrow \infty} \int_{-T}^T \frac{\sin(t(y-a))}{t} dt = -\pi (\chi_{(a, \infty)}(y) - \chi_{(-\infty, a)}(y)),$$

we have

$$(7.80) \quad \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} e^{ity} dt = 2\pi \chi_{(a, b)}(y) + \pi \chi_{\{a, b\}}(y) \quad y \in \mathbb{R}.$$

Note that  $\left| \int_0^x \frac{\sin t}{t} dt \right| \leq M_0$ , where  $M_0$  is an absolute constant. From this, we obtain, for all  $T > 0$ ,

$$(7.81) \quad \left| \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} e^{itx} dt \right| \leq 8M_0.$$

Thus, in view of the assumption on  $\mu$ , (7.80) and (7.81), we are in the position of applying the Lebesgue convergence theorem to (7.79) to obtain

$$(7.82) \quad \lim_{T \rightarrow \infty} \int_{[-T, T]^d} \prod_{k=1}^d \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} \varphi(t) dt = (2\pi)^d \mu(I).$$

As a result we obtain the desired formula.  $\square$

*Exercise 85.* For  $t > 0$  we let  $f(t) := \left| \int_0^t \frac{\sin t}{t} dt \right|$ . Then show that  $f$  attains its maximum at  $\pi$ .

We verify here that  $\mu \rightarrow \varphi_\mu$  is invertible by using Proposition 7.37.

**Theorem 7.38.** Suppose that  $\mu, \nu \in M(\mathbb{R}^d)$  are finite measures. Then,  $\varphi_\mu \equiv \varphi_\nu$  implies  $\mu = \nu$ .



*Proof.* Since  $M(\mathbb{R}^d)$  is contained in  $\mathcal{S}'(\mathbb{R}^d)$ , this theorem is immediate from the usual version of the Fourier transform. More precisely, we proceed as follows: In view of the embedding  $M(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$ , we can at least show

$$(7.83) \quad \int_{\mathbb{R}^d} \varphi(x) d\mu(x) = \int_{\mathbb{R}^d} \varphi(x) d\nu(x)$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Let  $\mathcal{C}$  denote the set of all rectangles whose boundary is not charged by  $\mu$  and  $\nu$ . Then  $\mu$  and  $\nu$  coincide from Proposition 7.36. Since  $\mathcal{C}$  generates  $\mathcal{B}$ , the Borel algebra, we see that  $\mu$  and  $\nu$  are the same.  $\square$

Let  $f \in L^1(\mathbb{R}^d)$ . Then note that

$$(7.84) \quad \int_{\mathbb{R}^d} f(\xi) \varphi_\mu(\xi) d\xi = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}^{-1} f(x) d\mu(x)$$

for  $\mu \in M(\mathbb{R}^d)$ , which is immediate from Fubini's theorem as in the proof of Theorem 7.8.

Next, we seek to find a method of recovering  $\mu$  via the Cèsaro means.

**Theorem 7.39.** *Let  $\varphi \in C(\mathbb{R}^d)$ . Set*

$$(7.85) \quad \Phi_\lambda(x) := \frac{1}{(2\pi)^d} \int_{B(\lambda)} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) e^{ix\xi} d\xi \quad (x \in \mathbb{R}^d)$$

for each  $\lambda > 0$ . Then  $\{\Phi_\lambda\}_{\lambda>0}$  forms an  $L^1(\mathbb{R}^d)$ -bounded family, if and only if  $\varphi(\xi) = \varphi_\mu(\xi)$  for some  $\mu \in M(\mathbb{R}^d)$ .

*Proof.* If  $\varphi(\xi) = \varphi_\mu(\xi)$  for some  $\mu \in M(\mathbb{R}^d)$ , then

$$\frac{1}{(2\pi)^d} \int_{B(\lambda)} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) e^{ix\xi} d\xi = \frac{1}{(2\pi)^d} \iint_{B(\lambda) \times \mathbb{R}^d} \left(1 - \frac{|\xi|}{\lambda}\right) e^{i(x+y)\xi} d\xi \otimes \mu(\xi, y).$$

Since

$$\left| \int_{B(\lambda)} \left(1 - \frac{|\xi|}{\lambda}\right) e^{i(x+y)\xi} d\xi \right| \lesssim 1,$$

where the implicit constant in  $\lesssim$  is independent of  $x, y, \lambda$ ,  $\{\Phi_\lambda\}_{\lambda>0}$  forms an  $L^1(\mathbb{R}^d)$ -bounded family.

Suppose conversely that  $\{\Phi_\lambda\}_{\lambda>0}$  forms an  $L^1(\mathbb{R}^d)$ -bounded family. Then by the Banach Alaoglu theorem (see Theorem 9.27 to follow), there exists a subsequence  $\{\Phi_{\lambda_j}\}_{j \in \mathbb{N}}$  convergent to some  $\mu \in M$  weakly. It suffices to show that

$$(7.86) \quad \int_{\mathbb{R}^d} g(\xi) \varphi(\xi) d\xi = \int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi.$$

for all  $g \in \mathcal{S}(\mathbb{R}^d)$  by the variation principle (see Theorem 4.35).

By (7.84), we have

$$\int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}^{-1} g(x) d\mu(x).$$

Note that  $\mathcal{F}^{-1} g \in \mathcal{S}(\mathbb{R}^d)$ . By the weak convergence mentioned above, we have

$$\begin{aligned} \int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi &= (2\pi)^{\frac{d}{2}} \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} \mathcal{F}^{-1} g(x) \left( \frac{1}{(2\pi)^d} \int_{B(\lambda_j)} \left(1 - \frac{|\xi|}{\lambda_j}\right) \varphi(\xi) e^{ix\xi} d\xi \right) dx \\ &= (2\pi)^{\frac{d}{2}} \lim_{j \rightarrow \infty} \int_{B(\lambda_j)} \left(1 - \frac{|\xi|}{\lambda_j}\right) \varphi(\xi) \left( \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}^{-1} g(x) e^{ix\xi} dx \right) d\xi. \end{aligned}$$

By the inversion formula, we obtain

$$(7.87) \quad \int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi = \lim_{j \rightarrow \infty} \int_{B(\lambda_j)} \left(1 - \frac{|\xi|}{\lambda_j}\right) \varphi(\xi) g(\xi) d\xi.$$

Finally by the dominated convergence theorem, we have

$$(7.88) \quad \int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi = \int_{\mathbb{R}^d} g(\xi) \varphi(\xi) d\xi.$$

This is precisely what we want. □

Denote by  $\text{BC}_0(\mathbb{R}^d)$  the closure of  $\mathcal{S}(\mathbb{R}^d)$  in  $\text{BC}(\mathbb{R}^d)$ . Note that

$$\text{BC}_0(\mathbb{R}^d) = \left\{ f \in \text{BC}(\mathbb{R}^d) : \lim_{|x| \rightarrow \infty} f(x) = 0 \right\}.$$

We have other characterizations.

**Theorem 7.40.** *A bounded continuous function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$  is expressed as  $\varphi = \varphi_\mu$  with  $\mu \in M(\mathbb{R}^d)$  if and only if*

$$(7.89) \quad \left| \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}f(\xi) \varphi(-\xi) d\xi \right| \lesssim \|f\|_{L^\infty}$$

for all  $f \in L^1(\mathbb{R}^d) \cap \text{BC}_0(\mathbb{R}^d)$ .

*Proof.* If  $\varphi = \varphi_\mu$  for some  $\mu \in M(\mathbb{R}^d)$ , then

$$\left| \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}f(\xi) \varphi(-\xi) d\xi \right| = \left| \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(-\xi) d\mu \right| \lesssim \|f\|_{L^\infty}.$$

Here the equality is justified by taking an approximation sequence.

Assume (7.89) instead. For  $f \in L^1(\mathbb{R}^d) \cap \text{BC}_0(\mathbb{R}^d)$  we set

$$(7.90) \quad \Phi(f) := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}f(\xi) \varphi(-\xi) d\xi \in \mathbb{C}.$$

Since  $L^1(\mathbb{R}^d) \cap \text{BC}_0(\mathbb{R}^d)$  is dense in  $\text{BC}_0(\mathbb{R}^d)$  and we are assuming (7.89), we can extend  $\Phi$  continuously to whole  $\text{BC}_0(\mathbb{R}^d)$ . By Riesz's representation theorem, we can find  $\mu \in M(\mathbb{R}^d)$  such that

$$(7.91) \quad \Phi(f) = \int_{\mathbb{R}^d} f(x) d\mu(x)$$

for all  $f \in L^1(\mathbb{R}^d) \cap \text{BC}(\mathbb{R}^d)$ . By the Plancherel formula, we see

$$(7.92) \quad \int_{\mathbb{R}^d} \mathcal{F}f(x) \varphi_\mu(-\xi) d\xi = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} f(x) d\mu(x) = (2\pi)^{\frac{d}{2}} \Phi(f) = \int_{\mathbb{R}^d} \mathcal{F}f(x) \varphi(-\xi) d\xi$$

for all  $f \in \mathcal{S}(\mathbb{R}^d)$ . As a result we obtain  $\varphi(x) = \frac{1}{(2\pi)^d} \varphi_\mu(-x)$ . □

**Theorem 7.41.** *Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$  be a bounded continuous function. Then the necessary and sufficient condition for  $\varphi$  to be expressed as the Fourier transform of  $\mu \in M(\mathbb{R}^d)$  is that, for every  $\lambda > 0$ ,*

$$(7.93) \quad \varphi(\lambda n) = \int_{\mathbb{R}^d} e^{in \cdot x} d\mu_\lambda(x) \quad (n \in \mathbb{Z}^d)$$

for some measure  $\mu_\lambda \in M(\mathbb{R}^d)$  and that  $\{\mu_\lambda\}_{\lambda>0}$  forms a bounded subset in  $M(\mathbb{R}^d)$ .

*Proof.* Suppose that  $\varphi = \varphi_\mu$  with  $\mu \in M$ . There exists  $\mu_\lambda \in M(\mathbb{R}^d)$  so that for every  $f \in \mathcal{S}(\mathbb{R}^d)$ ,

$$(7.94) \quad \int_{\mathbb{R}^d} f(x) d\mu_\lambda(x) = \int_{\mathbb{R}^d} f(\lambda x) d\mu(x).$$

By the duality  $BC_0(\mathbb{R}^d)$ - $M(\mathbb{R}^d)$  (the Riesz-representation theorem), we see that  $\mu_\lambda$  has the same norm as  $\mu$ . Then

$$\varphi(\lambda n) = \varphi_\mu(\lambda n) = \int_{\mathbb{R}^d} e^{i\lambda n x} d\mu(x) = \int_{\mathbb{R}^d} e^{i n x} d\mu_\lambda(x).$$

Suppose instead that, for every  $\lambda > 0$ , (7.93) holds for some measure  $\mu_\lambda \in M$  and that  $\{\mu_\lambda\}_{\lambda>0}$  forms a bounded subset in  $M$ . We have to verify that

$$(7.95) \quad \left| \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi \right| \lesssim \|f\|_{L^\infty}$$

for  $f \in \mathcal{S}(\mathbb{R}^d)$  (or  $f \in L^1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ ). This can be achieved as follows: First, we write the integral in (7.95) by the limit of the Riemann sum. We have

$$(7.96) \quad \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi = \lim_{\lambda \rightarrow 0} \sum_{m \in \mathbb{Z}} \lambda^d \cdot \mathcal{F}f(\lambda m)\varphi(\lambda m).$$

Next, we insert  $\varphi = \varphi_\mu$ . The result is

$$\begin{aligned} \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi &= \lim_{\lambda \rightarrow 0} \sum_{m \in \mathbb{Z}} \lambda^d \mathcal{F}f(\lambda m) \int_{\mathbb{R}^d} e^{ix\lambda m} d\mu_\lambda(x) \\ &= \lim_{\lambda \rightarrow 0} \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z}} \lambda^d \mathcal{F}f(\lambda m) e^{ix\lambda m} d\mu_\lambda(x). \end{aligned}$$

Note that, by the Poisson summation, Theorem 6.11 with  $\lambda$  replaced by  $(2\pi)^d \lambda^{-1}$ , we obtain

$$\sum_{m \in \mathbb{Z}} \lambda^d \mathcal{F}f(\lambda m) e^{ix\lambda m} = (2\pi)^d \sum_{m \in \mathbb{Z}} f\left(x - \frac{2\pi m}{\lambda}\right)$$

Thus,

$$(7.97) \quad \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi = (2\pi)^d \lim_{\lambda \rightarrow 0} \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z}} f\left(x - \frac{2\pi m}{\lambda}\right) d\mu_\lambda(x).$$

Since  $f \in \mathcal{S}(\mathbb{R}^d)$ , for each compact set  $K \subset \mathbb{R}^d$ ,

$$(7.98) \quad \lim_{\lambda \rightarrow 0} \left\{ \sup_{x \in K} \left| \sum_{m \in \mathbb{Z} \setminus \{0\}} f\left(x - \frac{2\pi m}{\lambda}\right) \right| \right\} = 0.$$

Since  $\{\mu_\lambda\}_{\lambda>0}$  forms a bounded subset in  $M(\mathbb{R}^d)$ , for any given  $\varepsilon > 0$  there exist a compact set  $K$  and a sequence  $\{\lambda_j\}_{j \in \mathbb{N}}$  converging to 0 such that

$$(7.99) \quad |\mu_{\lambda_j}|(\mathbb{R}^d \setminus K) \leq \varepsilon.$$

We conclude from (7.98) and (7.99) that

$$(7.100) \quad \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z} \setminus \{0\}} f\left(x - \frac{2\pi m}{\lambda_j}\right) d\mu_{\lambda_j}(x) = 0$$

If we combine (7.97) and (7.100), then we have

$$(7.101) \quad \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi = (2\pi)^d \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} f(x) d\mu_{\lambda_j}(x).$$

Again we use the fact that  $\{\mu_\lambda\}_{\lambda>0}$  forms a bounded subset in  $M(\mathbb{R}^d)$ . From (7.101), we conclude

$$(7.102) \quad \left| \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi \right| = (2\pi)^d \lim_{j \rightarrow \infty} \left| \int_{\mathbb{R}^d} f(x) d\mu_{\lambda_j}(x) \right| \lesssim \|f\|_{L^\infty} \cdot \sup_\lambda \|\mu_\lambda\| = \|f\|_{L^\infty}.$$

Thus, we obtain (7.95).  $\square$

**Theorem 7.42.** *Let  $\varphi \in \text{BC}(\mathbb{R}^d)$ . Then  $\varphi$  is the Fourier transform of a positive finite measure  $\mu$ , if and only if*

$$(7.103) \quad \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(\xi) d\xi \geq 0$$

for all  $f \in C_{c,+}^\infty(\mathbb{R}^d)$ .

*Proof.* Necessity Necessity is clear. Given a function  $\tau$ , let us denote  $\check{\tau}(x) = \tau(-x)$ . Then we have

$$\begin{aligned} \int_{\mathbb{R}^d} \mathcal{F}f(x)\check{\varphi}(x) dx &= \int_{\mathbb{R}^d} \check{\mathcal{F}}f(x)\varphi_\mu(x) dx \\ &= (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} [\mathcal{F}^{-1}\check{\mathcal{F}}f](x)\varphi_\mu(x) dx \\ &= (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \check{f}(x) d\mu(x) \geq 0. \end{aligned}$$

Sufficiency First, we prove  $\varphi(0) \geq 0$ . For this purpose we take the Gauss function  $\Phi_t(\xi) := \exp(-t|\xi|^2)$ ,  $\xi \in \mathbb{R}^d$ . Since  $\mathcal{F}\Phi_t(\xi) = c_0 t^{-d/2} \exp(-t^{-1}|\xi|^2)$  with  $c_0 \geq 0$  (see (7.115 for the precise value of  $c_0$ ), we see that

$$\varphi(0) = \lim_{t \downarrow 0} \pi^{-d/2} \int_{\mathbb{R}^d} t^{-d/2} \exp(-t^{-1}|\xi|^2)\varphi(\xi) d\xi = \int_{\mathbb{R}^d} \mathcal{F}\Phi_t(\xi)\varphi(\xi) d\xi \geq 0.$$

A normalization allows us to assume that  $\varphi(0) = 1$ .

A passage to the limit shows that (7.103) is valid for all  $f \in C_{c,+}(\mathbb{R}^d)$  with  $\mathcal{F}f \in L^1(\mathbb{R}^d)$ . Keeping in mind that  $\varphi(0) \geq 0$ , we let  $f \in C_{c,+}(\mathbb{R}^d)$ . Set  $M_0 := \sup_{x \in \mathbb{R}^d} |f(x)|$ . We define a function  $K_\lambda$  on  $\mathbb{R}^d$  by

$$K_\lambda(\xi_1, \xi_2, \dots, \xi_d) := \frac{\lambda^d}{(2\pi)^d} \text{sinc}\left(\frac{\lambda\xi_1}{2}\right) \text{sinc}\left(\frac{\lambda\xi_2}{2}\right) \cdots \text{sinc}\left(\frac{\lambda\xi_d}{2}\right) \quad (\xi_1, \xi_2, \dots, \xi_d) \in \mathbb{R}^d,$$

where sinc is the sinc function defined by  $\text{sinc}(t) = \frac{\sin t}{t}$  for  $t \in \mathbb{R} \setminus \{0\}$ . If  $\lambda > 0$  is large enough, then

$$(7.104) \quad (2\pi)^d M_0 \cdot \mathcal{F}^{-1}K_\lambda - f \in C_{c,+}$$

By inserting (7.104) to (7.103) and the fact that  $\varphi(0) = 1$ , we have

$$(7.105) \quad \int_{\mathbb{R}^d} \{(2\pi)^d M_0 K_\lambda(\xi) - \mathcal{F}f(\xi)\} \varphi(-\xi) d\xi \geq 0.$$

As  $\lambda \rightarrow \infty$ , we obtain

$$(7.106) \quad \int_{\mathbb{R}^d} ((2\pi)^d M_0 K_\lambda(\xi) - \mathcal{F}f(\xi))\varphi(-\xi) d\xi \rightarrow \varphi(0)M_0 - \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(-\xi) d\xi.$$

Consequently, it follows from (7.105) and (7.106) that

$$(7.107) \quad 0 \leq \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(-\xi) d\xi \leq \varphi(0)M_0 = \varphi(0) \sup_{x \in \mathbb{R}^d} |f(x)| \quad (f \in C_{c,+}(\mathbb{R}^d)).$$

Now we are in the position of applying Riesz's representation theorem to the functional

$$f \in C_{c,+}(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} \mathcal{F}f(\xi)\varphi(-\xi) d\xi.$$

We can find a positive finite measure  $\mu \in M(\mathbb{R}^d)$  with

$$(7.108) \quad \int_{\mathbb{R}^d} \mathcal{F}f(\xi) \cdot \check{\varphi}(\xi) d\xi = \int_{\mathbb{R}^d} f(x) d\mu(x)$$

for all  $f \in BC(\mathbb{R}^d)$ . From (7.108), we deduce  $\varphi = \varphi_\mu$ . □

We have a good necessary and sufficient condition for bounded continuous functions to be expressed as the Fourier transform of measures. A natural question arises: When the measure is positive ?

To answer the question, we let  $M_+(\mathbb{R}^d)$  be the set of all positive measures and we present the following definition.

**Definition 7.43** (Positive definite functions). A function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$  is said to be positive definite, if

$$(7.109) \quad \{\varphi(\xi_j - \xi_k)\}_{j,k=1}^N$$

is positive definite for all finite sets  $\{\xi_1, \xi_2, \dots, \xi_N\}$ .

Positive definite functions can be characterized as the image of positive finite measures.

**Theorem 7.44** (Bochner). *A continuous function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$  is positive definite if and only if  $\varphi = \varphi_\mu$  with  $\mu \in M_+(\mathbb{R}^d)$ .*

*Proof.* The sufficiency is clear. Let us prove the necessity.

Let  $k \in \mathbb{N}$ . Set

$$\varphi_k(\xi) := \varphi(\xi) \exp(-k^{-1}|\xi|^2)$$

for  $\xi \in \mathbb{R}^d$  and

$$f_k(x) := \int_{\mathbb{R}^d} \varphi_k(\xi) e^{-ix\xi} d\xi$$

for  $x \in \mathbb{R}^d$ . Then let us establish that

$$(7.110) \quad f_k \geq 0$$

and that

$$(7.111) \quad \mathcal{F}(f_k dx)(\xi) = \int_{\mathbb{R}^d} f_k(x) e^{ix\xi} dx = (2\pi)^d \varphi_k(\xi)$$

Note that (7.111) follows from the inversion formula. For the proof of (7.110) we use positive definiteness of  $\varphi$  and the fact

$$\int_{\mathbb{R}^d} \exp\left(-\frac{2}{k}|x-y|^2 - ia(x-y)\right) \exp\left(-\frac{2}{k}|y|^2 - iay\right) dy \simeq_{k,d} \exp\left(-\frac{1}{k}|x|^2 - iax\right).$$

Keeping them in mind, we proceed as follows :

$$\begin{aligned}
f_k(a) &= \int_{\mathbb{R}^d} \varphi(\xi) \exp(-k^{-1}|\xi|^2 - ia\xi) d\xi \\
&\simeq \int_{\mathbb{R}^d} \varphi(\xi) \left( \int_{\mathbb{R}^d} \exp\left(-\frac{2}{k}|\xi - y|^2 - ia(\xi - y)\right) \exp\left(-\frac{2}{k}|y|^2 - iay\right) dy \right) d\xi \\
&\simeq \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\xi + \eta) \exp\left(-\frac{2}{k}|\xi|^2 - ia\xi\right) \exp\left(-\frac{2}{k}|\eta|^2 - ia\eta\right) d\eta d\xi \\
&\simeq \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\xi - \eta) \exp\left(-\frac{2}{k}|\xi|^2 - ia\xi\right) \overline{\exp\left(-\frac{2}{k}|\eta|^2 - ia\eta\right)} d\eta d\xi \\
&\simeq \lim_{N \rightarrow \infty} N^{2n} \sum_{k, j \in \mathbb{Z}^d} \varphi\left(\frac{j}{N} - \frac{k}{N}\right) \exp\left(-\frac{2}{k} \left|\frac{k}{N}\right|^2 - i\frac{a}{N}k\right) \overline{\exp\left(-\frac{2}{k} \left|\frac{j}{N}\right|^2 - i\frac{a}{N}j\right)},
\end{aligned}$$

which is positive since  $\varphi$  is positive definite.

Consequently (7.110) and (7.111) were established.

By the Banach Alaoglu theorem (see Theorem 9.27 again), we can take a subsequence  $f_{k(j)}, j = 1, 2, \dots$  such that it converges weak-\* to some positive  $\mu \in M(\mathbb{R}^d)$ . Let us check that  $\mu$  is a measure we are looking for.

Since each  $f_k$  is positive, so is  $\mu$ . What remains to show is that the Fourier transform of  $\mu$  is  $\varphi$ . Let  $g \in \mathcal{S}(\mathbb{R}^d)$ . Then

$$(7.112) \quad \int_{\mathbb{R}^d} g(\xi) \varphi(\xi) d\xi = \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} g(\xi) \varphi_{k(j)}(\xi) d\xi$$

by the dominated convergence theorem. Meanwhile, (7.111) gives us

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} g(\xi) \varphi_{k(j)}(\xi) d\xi = (2\pi)^{d/2} \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} g(\xi) \mathcal{F}(f_{k(j)} dx)(\xi) d\xi.$$

Now we use the Fubini formula to obtain

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} g(\xi) \varphi_{k(j)}(\xi) d\xi = (2\pi)^{d/2} \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} \mathcal{F}^{-1}g(x) f_{k(j)}(x) dx.$$

From the definition of the weak convergence and again by the Fubini theorem, we obtain

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} g(\xi) \varphi_{k(j)}(\xi) d\xi = (2\pi)^{d/2} \int_{\mathbb{R}^d} \mathcal{F}^{-1}g(x) d\mu(x) = \int_{\mathbb{R}^d} g(\xi) \varphi_\mu(\xi) d\xi.$$

Since  $g \in \mathcal{S}(\mathbb{R}^d)$  was taken arbitrarily, it follows that  $\varphi \equiv \varphi_\mu$ .  $\square$

The next theorem is used to prove the central limit theorem in probability theory.

**Theorem 7.45.** *Let  $\mu_j \in M, j = 1, 2, \dots$  be a sequence of Borel probability measures. Suppose that  $\varphi_{\mu_j}, j = 1, 2, \dots$  converges to  $\varphi$  pointwise and that  $\varphi$  is continuous at 0. Then  $\mu_j$  converges to some probability measure  $\mu$  weakly.*

*Proof.* We proceed in steps. First, we prove

**Claim 7.46.** *The set  $\{\mu_j\}_{j=1}^\infty$  is tight, that is,*

$$(7.113) \quad \lim_{R \rightarrow \infty} \left( \sup_{j \in \mathbb{N}} \mu_j(\mathbb{R}^d \setminus B(R)) \right) = 0.$$

*Proof of Claim 7.46.* We first compute the Fourier transform of  $\varphi_t(x) := \exp(-t^{-1}|x|^2)$ . It is not so difficult to see

$$(7.114) \quad \mathcal{F}\varphi_t(\xi) = C \exp(-t|\xi|^2).$$

To calculate the exact value of  $C = C_t$ , we note that  $C = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \varphi_t(x) = (2\pi)^{-\frac{d}{2}} \sqrt{(\pi t)^d}$ . Consequently we obtain

$$(7.115) \quad \mathcal{F}\varphi_t(\xi) = \mathcal{F}[\exp(-t^{-1}|x|^2)](\xi) = \left(\sqrt{\frac{t}{2}}\right)^d \exp(-t|\xi|^2) \quad (\xi \in \mathbb{R}^d).$$

We insert (7.115) to the formula  $\int_{\mathbb{R}^d} \mathcal{F}^{-1}\varphi_t(x) d\mu_j(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \varphi_t(\xi)\varphi_{\mu_j}(\xi) d\xi$  and obtain

$$(7.116) \quad \int_{\mathbb{R}^d} \exp(-t|x|^2) d\mu_j(x) = \frac{1}{\sqrt{(\pi t)^d}} \int_{\mathbb{R}^d} \exp(-t^{-1}|\xi|^2)\varphi_{\mu_j}(\xi) d\xi.$$

Let  $\varepsilon > 0$  be given. Then for large  $j$ , say  $j \geq J$ , we have

$$(7.117) \quad \left| \int_{\mathbb{R}^d} \exp(-t|x|^2) d\mu_j(x) - \frac{1}{\sqrt{(\pi t)^d}} \int_{\mathbb{R}^d} \exp(-t^{-1}|\xi|^2)\varphi(\xi) d\xi \right| < \varepsilon.$$

If  $t > 0$  is smaller than some  $t_0 > 0$ , then we have

$$(7.118) \quad \left| \frac{1}{(\sqrt{\pi t})^d} \int_{\mathbb{R}^d} \exp(-t^{-1}|\xi|^2)\varphi(\xi) d\xi - 1 \right| < \varepsilon.$$

Consequently we obtain

$$(7.119) \quad \int_{\mathbb{R}^d} \exp(-t|x|^2) d\mu_j(x) > 1 - 2\varepsilon.$$

for all  $j \geq J$  and  $0 < t \leq t_0$ . Since

$$(7.120) \quad \mu(\mathbb{R}^d \setminus B(R)) \leq \exp(tR^2) \int_{\mathbb{R}^d} \exp(-t|x|^2) d\mu_j(x) \leq 2\varepsilon \exp(tR^2),$$

for  $R_0 := t_0^{-\frac{1}{2}}$  we have  $\mu_j(\mathbb{R}^d \setminus B(R_0)) \leq 6\varepsilon, j \geq J$ . Take  $R_1$  so large that, for every  $j < J$ ,  $\mu_j(\mathbb{R}^d \setminus B(R_1)) \leq 6\varepsilon$ . Thus letting  $R := R_0 + R_1$ , we see that

$$(7.121) \quad \mu_j(\mathbb{R}^d \setminus B(R)) \leq 6\varepsilon$$

for all  $j \in \mathbb{N}$ . The claim is now therefore proved. □

Now we refer back to the proof of Theorem 7.45. From Claim 7.46, we can take a convergent subsequence for any subsequence. The uniqueness of the Fourier transform then shows that the original sequence converges to some probability measure  $\mu$ . □

To conclude this paragraph, we prove the Wiener theorem which parallels the Planchrel theorem. In the theorem, we use the symbol

$$(7.122) \quad \sum_{x \in \mathbb{R}^d} f(x)$$

for positive functions  $f : \mathbb{R}^d \rightarrow [0, \infty)$ . The summation  $\alpha$  in (7.122) means that  $f(x) \equiv 0$  on a cofinite subset  $E$  and that

$$\alpha = \sup \left\{ \sum_{x \in F} f(x) : F \text{ is a finite subset of } \mathbb{R}^d \right\}.$$

**Theorem 7.47** (Wiener). *Let  $\mu \in M$ . Then*

$$(7.123) \quad \sum_{x \in \mathbb{R}^d} |\mu(\{x\})|^2 = \lim_{\lambda \rightarrow \infty} \frac{1}{2\lambda} \int_{B(\lambda)} |\varphi_\mu(\xi)|^2 d\xi.$$

*Proof.* Let  $\nu \in M$  be defined by

$$(7.124) \quad \nu(E) := \int_{\mathbb{R}^d \times \mathbb{R}^d} \chi_E(x-y) d\mu(x) d\bar{\mu}(y).$$

Then by the definition of  $\nu$ , we have

$$(7.125) \quad \nu\{0\} = \int_{\mathbb{R}^d \times \mathbb{R}^d} \chi_{\{x=y\}}(x,y) d\mu(x) d\bar{\mu}(y) = \int_{\{\mu\{y\} \neq 0\}} \mu\{y\} d\bar{\mu}(y) = \sum_{x \in \mathbb{R}^d} |\mu(\{x\})|^2.$$

Furthermore, by using  $\varphi_\nu(\xi) = |\varphi_\mu(\xi)|^2$  we have

$$\int_{B(\lambda)} |\varphi_\mu(\xi)|^2 d\xi = \int_{B(\lambda)} \varphi_\nu(\xi) d\xi = \int_{\mathbb{R}^d} \left( \int_{B(\lambda)} e^{i\xi x} d\xi \right) d\nu(x) = 2\lambda \int_{\mathbb{R}^d} \text{sinc}(2\lambda x) d\nu(x).$$

By the dominated convergence theorem, we have

$$(7.126) \quad \lim_{\lambda \rightarrow \infty} \frac{1}{2\lambda} \int_{B(\lambda)} |\varphi_\mu(\xi)|^2 d\xi = \lim_{\lambda \rightarrow \infty} \int_{\mathbb{R}^d} \text{sinc}(2\lambda x) d\nu(x) = \nu\{0\}.$$

If we combine (7.125) and (7.126), we obtain the desired result.  $\square$

## 8. SOBOLEV SPACES

### 8.1. Weak derivative.

**Definition 8.1.** Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  and  $\alpha \in \mathbb{N}_0^d$ . A weak partial  $\alpha$ -derivative of  $f$ , if it exists, is a locally integrable function  $g$  such that

$$(8.1) \quad \int_{\mathbb{R}^d} f(x) \partial^\alpha \varphi(x) dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} g(x) \varphi(x) dx$$

for all  $\varphi \in C_c^\infty(\mathbb{R}^d)$ .

The following lemma is easy to prove.

**Lemma 8.2.** *Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  and  $\alpha \in \mathbb{N}_0^d$ .*

- (1) *The partial  $\alpha$ -derivative of  $f$ , if it exists, is unique.*
- (2) *If  $f \in C^{|\alpha|}$ , then  $\partial^\alpha f$  coincides with the usual partial  $\alpha$ -derivative.*

*Proof.* (1) follows from Theorem 4.35 and (2) can be obtained by carrying out integration by parts.  $\square$

**Example 8.3.** Let  $\varphi \in C_c^\infty(\mathbb{R}^d)$  and define  $\psi(x) := |x|\varphi(x)$  for  $x \in \mathbb{R}^d$ . Then we have

$$(8.2) \quad \partial_j \psi(x) = |x| \partial_j \varphi(x) + \frac{x_j}{|x|} \varphi(x)$$

in the sense of weak derivatives.



*Proof.* Let  $\eta \in C_c^\infty(\mathbb{R}^d)$ . For the sake of simplicity we can assume  $j = d$ . Then if we carry out integration by parts, we obtain

$$\begin{aligned} & \int_{|x_d| > \varepsilon} \psi(x) \cdot \partial_d \eta(x) \, dx \\ &= \int_{\mathbb{R}^{d-1}} (\psi(x' + \varepsilon e_d) \eta(x' + \varepsilon e_d) - \psi(x' - \varepsilon e_d) \eta(x' - \varepsilon e_d)) \, dx' - \int_{|x_d| > \varepsilon} \partial_d \psi(x) \cdot \eta(x) \, dx. \end{aligned}$$

If we let  $\varepsilon \downarrow 0$ , then the left-hand side tends to  $\int_{\mathbb{R}^d} \psi(x) \partial_d \eta(x) \, dx$  and the boundary terms of the right-hand side cancels. Therefore

$$\int_{\mathbb{R}^d} \psi(x) \partial_d \eta(x) \, dx = - \lim_{\varepsilon \downarrow 0} \int_{|x_d| > \varepsilon} \partial_d \psi(x) \eta(x) \, dx = - \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} \left( |x| \partial_j \varphi(x) + \frac{x_j}{|x|} \varphi(x) \right) \eta(x) \, dx.$$

Thus, we obtain (8.2), although  $\psi \notin C^1$  ! □

### 8.2. Sobolev space $W^{m,p}(\mathbb{R}^d)$ .

The aim of this section is very modest. We intend to define the Sobolev space  $W^{m,p}(\mathbb{R}^d)$  and investigate some elementary properties. We do not go into details, for example, the trace theorem, the extension theorem and so on are not taken up here. We content ourselves with making a brief view of the function spaces.

Sobolev space  $W^{m,p}(\mathbb{R}^d)$ . As we have seen before,  $C_c^\infty(\mathbb{R}^d)$  is not complete, if we endow it with the  $L^p(\mathbb{R}^d)$ -norm. If  $p$  is finite, then  $L^p(\mathbb{R}^d)$  is a completion of  $C_c^\infty(\mathbb{R}^d)$  with the  $L^p(\mathbb{R}^d)$ -norm. The space  $L^\infty(\mathbb{R}^d)$  is a natural extension of the  $L^p(\mathbb{R}^d)$ -norm.

The fall of completeness is the case, if we endow  $C_c^\infty(\mathbb{R}^d)$  with a norm given by

$$(8.3) \quad \|f\|_{W^{m,p}} := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq m}} \|\partial^\alpha f\|_p.$$

We can say that  $W^{m,p}(\mathbb{R}^d)$ , whose precise definition we are about to present, overcomes this defect, if  $p < \infty$ .

**Definition 8.4** (The Sobolev spaces in  $L^p(\mathbb{R}^d)$  with  $m$  derivatives). Let  $m \in \mathbb{N}_0$  and let  $1 \leq p \leq \infty$ . Then define

$$\begin{aligned} & W^{m,p}(\mathbb{R}^d) \\ & := \{f \in L^p(\mathbb{R}^d) : \partial^\alpha f \text{ exists in weak sense and belongs to } L^p(\mathbb{R}^d) \text{ for all } \alpha \text{ with } |\alpha| \leq m\}. \end{aligned}$$

The norm of  $f \in W^{m,p}(\mathbb{R}^d)$  is defined by  $\|f\|_{W^{m,p}} := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq m}} \|\partial^\alpha f\|_p$ .

Let  $1 \leq p \leq \infty$ . From the viewpoint of  $W^{m,p}(\mathbb{R}^d)$ , we see that  $f \in L^p(\mathbb{R}^d)$ , if and only if  $\partial^\alpha f \in L^p(\mathbb{R}^d)$  for all  $\alpha \in \mathbb{N}_0^d$  with  $|\alpha| \leq m$ .

**Theorem 8.5.** *The Sobolev space  $W^{m,p}(\mathbb{R}^d)$  is a Banach space.*

*Proof.* We leave the readers for proving that  $W^{m,p}(\mathbb{R}^d)$  is a normed space. Let us see the completeness. To do this, let  $\{f_j\}_{j \in \mathbb{N}}$  be a Cauchy sequence. Then, as we have established that  $L^p(\mathbb{R}^d)$  is complete, for each  $\alpha \in \mathbb{N}_0^d$  with length less than  $m$ , there exists  $g_\alpha$  so that  $\lim_{j \rightarrow \infty} \partial^\alpha f_j = g_\alpha$ . If  $\alpha = 0$ , then we write  $g := g_\alpha$ .

Choose a test function  $\varphi \in C_c^\infty(\mathbb{R}^d)$ . Then we have

$$(8.4) \quad \int_{\mathbb{R}^d} g(x) \partial^\alpha \varphi(x) dx = \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} f_j(x) \partial^\alpha \varphi(x) dx$$

and

$$(8.5) \quad (-1)^{|\alpha|} \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} \partial^\alpha f_j(x) \varphi(x) dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} g_\alpha(x) \varphi(x) dx.$$

Indeed, to verify (8.4), we have, by the Hölder inequality,

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} g(x) \cdot \partial^\alpha \varphi(x) dx - \int_{\mathbb{R}^d} f_j(x) \cdot \partial^\alpha \varphi(x) dx \right| \\ & \leq \int_{\mathbb{R}^d} |g(x) - f_j(x)| \cdot |\partial^\alpha \varphi(x)| dx \\ & \leq \left( \int_{\mathbb{R}^d} |g(x) - f_j(x)|^p dx \right)^{\frac{1}{p}} \cdot \left( \int_{\mathbb{R}^d} |\partial^\alpha \varphi(x)|^{p'} dx \right)^{\frac{1}{p'}} \\ & \leq \|g - f_j\|_p \cdot \|\partial^\alpha \varphi\|_{p'}. \end{aligned}$$

In the same way we can verify (8.5). It is worth noting that we do not use the Lebesgue convergence theorem. From the definition of the weak derivatives these four terms coincide, which means the weak partial  $\alpha$ -derivative of  $g$  exists and coincides with  $g_\alpha$ . Therefore, it follows that  $g \in W^{m,p}(\mathbb{R}^d)$ .

Once we obtain  $g \in W^{m,p}(\mathbb{R}^d)$ , it is easy to see  $\lim_{j \rightarrow \infty} f_j = g$ . Indeed,

$$(8.6) \quad \|f_j - g\|_{W^{m,p}} = \sum_{|\alpha| \leq m} \|\partial^\alpha (f_j - g)\|_p = \sum_{|\alpha| \leq m} \|\partial^\alpha f_j - g_\alpha\|_p \rightarrow 0$$

as  $j \rightarrow \infty$ .

Thus, it follows that  $W^{m,p}(\mathbb{R}^d)$  is complete.  $\square$

*Exercise 86.* Show that  $W^{m,p}(\mathbb{R}^d)$  is a normed space.

As a special case, when  $p = 2$ , the space  $W^{m,p}(\mathbb{R}^d)$  can be characterized in terms of the Fourier transform.

**Theorem 8.6.** *Let  $m \in \mathbb{N}_0$ . Then  $f \in L^2(\mathbb{R}^d)$  belongs to  $W^{m,2}(\mathbb{R}^d)$  if and only if  $\langle \cdot \rangle^m \mathcal{F} f \in L^2(\mathbb{R}^d)$ . Furthermore, we have*

$$(8.7) \quad \|\langle \cdot \rangle^m \mathcal{F} f\|_2 \sim \|f\|_{W^{m,2}}$$

for all  $f \in W^{m,2}(\mathbb{R}^d)$ .

*Exercise 87.* Prove Theorem 8.6 by using the Plancherel theorem.

Mollification and density. Now we are going to discuss the mollification and consider the dense subspace of  $W^{m,p}(\mathbb{R}^d)$ .

**Lemma 8.7.** *Let  $1 \leq p < \infty$ . Let  $\varphi \in C_c^\infty(\mathbb{R}^d)$  be a positive function with integral 1. Then for given  $f \in W^{m,p}(\mathbb{R}^d)$  and  $t > 0$ , we set  $f_t := \varphi_t * f$ , where  $\varphi_t = \frac{1}{t^d} \varphi\left(\frac{\cdot}{t}\right)$ . Then we have  $\lim_{t \downarrow 0} f_t = f$  in the topology of  $W^{m,p}(\mathbb{R}^d)$ .*

*Proof.* Note that  $\partial_i f_t = \varphi_t * \partial_i f$  because  $f \in W^{1,p}(\mathbb{R}^d)$  and the derivative in the sense of Schwartz distributions and that of usual derivative coincide. Therefore, this lemma is true for  $|\alpha| = 1$ . In general case we have only to prove the lemma inductively.  $\square$

**Corollary 8.8.** *Let  $1 \leq p < \infty$ . Then  $C_c^\infty(\mathbb{R}^d)$  is dense in  $W^{m,p}(\mathbb{R}^d)$ .*

*Proof.* By Lemma 8.7, we have only to approximate smooth elements in  $W^{m,p}(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$ . Let  $f \in W^{m,p}(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$ . Choose a truncation function  $\eta \in C_c^\infty(\mathbb{R}^d)$  that equals 1 on  $B(1)$ . Set  $f_m(x) := f(x)\eta\left(\frac{x}{m}\right)$ . Then by the Lebesgue convergence theorem  $f_m$  tends to  $f$  in  $W^{m,p}(\mathbb{R}^d)$ .  $\square$

Difference quotient and weak-partial derivative. Classically the differentiation was defined by the limit, that is, it was given by the limit of the difference quotient. However, the weak derivative we take up here seems to have nothing to do with the difference quotient. In this paragraph, we investigate the connection between the difference quotient and the weak-partial derivative.

**Definition 8.9.** Let  $h \in \mathbb{R} \setminus \{0\}$  and  $i = 1, 2, \dots, d$ . Then define, for  $f \in L^1_{loc}(\mathbb{R}^d)$ ,

$$(8.8) \quad D_i^h f(x) = \frac{f(x + he_i) - f(x)}{h},$$

where  $e_i = (\delta_{ij})_{j=1}^d$  is the  $j$ -th elementary vector.

**Theorem 8.10.** *Let  $1 < p < \infty$ . Then  $f \in L^p(\mathbb{R}^d)$  belongs to  $W^{1,p}(\mathbb{R}^d)$  if and only if*

$$(8.9) \quad \|D_i^h f\|_p \lesssim 1$$

for each  $h \in \mathbb{R} \setminus \{0\}$  and  $i = 1, 2, \dots, d$ .

*Proof.* Suppose that  $\|D_i^h f\|_p \leq c < \infty$  for each  $h \in \mathbb{R} \setminus \{0\}$  and  $i = 1, 2, \dots, d$ . Then by the Banach-Alaoglu theorem (see Theorem 9.27 again) along with the duality  $L^p(\mathbb{R}^d)$ - $L^{p'}(\mathbb{R}^d)$ , there exist a subsequence  $h_1 > h_2 > \dots > h_j > \dots \rightarrow 0$  and  $g_1, g_2, \dots, g_d \in L^p(\mathbb{R}^d)$  such that

$$(8.10) \quad \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} D_i^{h_j} f(x) \cdot h(x) dx = \int_{\mathbb{R}^d} g_i(x) \cdot h(x) dx$$

for all  $h \in C_c^\infty(\mathbb{R}^d)$  and  $i = 1, 2, \dots, d$ . Then by change of variables we obtain

$$(8.11) \quad \int_{\mathbb{R}^d} f(x) \partial_i h(x) dx = - \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} f(x) D_i^{-h_j} h(x) dx.$$

Therefore, letting  $j \rightarrow \infty$ , we obtain

$$\int_{\mathbb{R}^d} f(x) \partial_i h(x) dx = - \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} D_i^{h_j} f(x) h(x) dx = - \int_{\mathbb{R}^d} g_i(x) h(x) dx.$$

Therefore, it follows that  $f \in W^{1,p}(\mathbb{R}^d)$  and that its partial derivative  $\partial_i f$  coincides with  $g_i$ .

Suppose instead that  $f \in W^{1,p}(\mathbb{R}^d)$ . Let  $f_t$  be defined in Lemma 8.7. Then we have

$$(8.12) \quad D_i^h f_t(x) = \int_0^1 \partial_i f_t(x + hse_i) ds.$$

By the Minkowski inequality, we have

$$(8.13) \quad \|D_i^h f_t\|_p \leq \int_0^1 \|\partial_i f_t(* + hse_i)\|_p ds = \|\partial_i f_t\|_p \leq \|\partial_i f\|_p.$$

Since  $f_t \rightarrow f$  in  $L^p(\mathbb{R}^d)$ , we have  $\|D_i^h f\|_p \leq \|\partial_i f\|_p < \infty$ .

Therefore, the proof is now complete.  $\square$

Composition. Now we consider the chain rule. What is totally different from the classical analysis is that if  $f \in W^{1,p}(\mathbb{R}^d)$ , then we have  $|f| \in W^{1,p}(\mathbb{R}^d)$ . This property fails for the space  $C^1(\mathbb{R}^d)$  because the limit of difference quotient comes into play.

**Theorem 8.11.** *Let  $1 < p < \infty$ . Assume that  $\eta : \mathbb{C} \rightarrow \mathbb{R}$  is a  $C^1$  function with bounded derivative and that  $\eta(0) = 0$ . Let  $f \in W^{1,p}(\mathbb{R}^d)$ . Then  $\eta \circ f \in W^{1,p}(\mathbb{R}^d)$ .*

Here if the function  $f(x) = \pm\infty$ , then it will be understood that  $\eta \circ f(x) = 0$  and below we shall disregard such a point. We shall not allude to this point later.

*Proof.*  $\boxed{\eta \circ f \in L^p(\mathbb{R}^d)}$  To see this we note that  $|\eta(t)| \leq \sup_{s \in \mathbb{R}} |\eta'(s)|t$ . Using this inequality and  $f \in L^p(\mathbb{R}^d)$ , we readily obtain  $\eta \circ f \in L^p(\mathbb{R}^d)$ .

$\boxed{\eta \circ f \in W^{1,p}(\mathbb{R}^d)}$  Suppose that  $\{f_j\}_{j=1}^\infty$  is a sequence of compactly supported functions approximating  $f$  in  $W^{1,p}(\mathbb{R}^d)$ . A passage to a subsequence allows us to assume that  $\{f_j(x)\}_{j=1}^\infty$  converges to  $f(x)$  for almost all  $x \in \mathbb{R}^d$ . Let  $\tau \in C_c^\infty(\mathbb{R}^d)$  be a test function and fix  $k = 1, 2, \dots, d$ . Then the Lebesgue convergence theorem yields

$$(8.14) \quad \int_{\mathbb{R}^d} (\partial_k f(x) \cdot \eta'(f(x))\tau(x) dx = \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} (\partial_k f_j(x) \cdot \eta'(f_j(x))\tau(x) dx.$$

Now that  $f_j$  is smooth, we are in the position of carrying out the usual integration by parts. As a consequence we obtain

$$\int_{\mathbb{R}^d} \partial_k f(x) \eta'(f(x))\tau(x) dx = - \lim_{j \rightarrow \infty} \int_{\mathbb{R}^d} \eta(f_j(x)) \partial_k \tau(x) dx$$

Taking into account the inequality  $|\eta(s) - \eta(t)| \lesssim |s - t|$ , we see that  $\eta \circ f_j$  converges to  $\eta \circ f$  in the  $L^p(\mathbb{R}^d)$ -topology. Going through the same argument using the Hölder inequality as before, we obtain

$$\int_{\mathbb{R}^d} \partial_k f(x) \eta'(f(x))\tau(x) dx = - \int_{\mathbb{R}^d} \eta(f(x)) \partial_k \tau(x) dx.$$

Thus, we obtain  $\partial_k(\eta \circ f) = \partial_k f \cdot \eta' \circ f \in L^p(\mathbb{R}^d)$ . As a consequence  $\eta \circ f \in W^{1,p}(\mathbb{R}^d)$ .  $\square$

**Corollary 8.12.** *Let  $1 \leq p < \infty$ . Then, for every  $f \in W^{1,p}(\mathbb{R}^d)$ ,  $|f|$  belongs to  $W^{1,p}(\mathbb{R}^d)$  and*

$$(8.15) \quad \| |f| \|_{W^{1,p}} \leq \| f \|_{W^{1,p}}$$

for all  $f \in W^{1,p}(\mathbb{R}^d)$ .

*Proof.* We set

$$\eta_u(t) := \begin{cases} \frac{t^2}{2u} & |t| \leq u, \\ |t| - \frac{u}{2} & |t| \geq u \end{cases}$$

and define  $\text{sgn}(t) := \lim_{u \downarrow 0} \eta'_u(t)$ . Then for every test function  $\tau \in \mathcal{S}(\mathbb{R}^d)$ , we have

$$(8.16) \quad \int_{\mathbb{R}^d} \text{sgn}(f(x)) \partial_k f(x) \tau(x) dx = \lim_{u \downarrow 0} \int_{\mathbb{R}^d} \eta'_u(f(x)) \partial_k f(x) \tau(x) dx$$

by virtue of the dominated convergence theorem. Having approximated the integral with a smooth function, we are again in the position of carrying out integration by parts. Hence we obtain

$$(8.17) \quad \int_{\mathbb{R}^d} \text{sgn}(f) \cdot \partial_k f \cdot \tau = - \lim_{u \downarrow 0} \int_{\mathbb{R}^d} \eta_u \circ f \cdot \partial_k \tau.$$

Finally by using the Lebesgue convergence theorem, we obtain

$$\int_{\mathbb{R}^d} \operatorname{sgn}(f(x)) \partial_k f(x) \tau(x) dx = - \int_{\mathbb{R}^d} |f(x)| \partial_k \tau(x).$$

Thus, we have  $\partial_k |f| = \operatorname{sgn}(f) \partial_k f$ . As a result we obtain the desired estimate.  $\square$

Sobolev embedding theorem. Suppose that  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a smooth function with the first derivative integrable over  $\mathbb{R}$ . Then we have

$$(8.18) \quad f(x) = \int_0^x f(t) dt + f(0)$$

and hence  $f$  is bounded. In this way if we have some information on the partial derivatives, then we can say more about the function. The Sobolev embedding theorem quantifies such a situation.

We write

$$Du := (\partial_{x_1} u, \partial_{x_2} u, \dots, \partial_{x_d} u).$$

**Theorem 8.13** (Gagliardo-Nirenberg-Sobolev inequality). *Assume  $1 \leq p < d$ . Define  $q$  by*

$$(8.19) \quad \frac{1}{q} = \frac{1}{p} - \frac{1}{d}.$$

*Then*

$$(8.20) \quad \|u\|_q \lesssim_{p,d} \|Du\|_p$$

*for all  $u \in C_c^1$ .*

Before we come to the proof, let us see what happens if we replace  $q$  by some other values. Suppose that  $t > 0$  and set  $u_t(x) = u(tx)$ . Then we have

$$(8.21) \quad \|u_t\|_q = t^{-\frac{d}{q}} \|u\|_q, \quad \|Du_t\|_p = t^{1-\frac{d}{p}} \|Du\|_p.$$

Therefore, (8.20) can only possibly true for  $q$  defined above. Let us make a helpful remark. We really do need  $u$  to have compact support for (8.20) to hold, as the example  $u \equiv 1$  shows. However remarkably the constant here does not depend at all upon the size of the support of  $u$ .

Now let us prove Theorem 8.13.

*Proof.* First assume  $p = 1$ . Since  $u$  has compact support, for each  $j = 1, 2, \dots, d$  and  $x \in \mathbb{R}^d$  we have

$$(8.22) \quad u(x) = \int_{-\infty}^{x_j} u_{x_i}(x_1, x_2, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d) dy_j$$

and so

$$(8.23) \quad |u(x)| \leq \int_{\mathbb{R}} |Du(x_1, x_2, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| dy_j, \quad (i = 1, 2, \dots, d).$$

Consequently

$$(8.24) \quad |u(x)|^{\frac{d}{d-1}} \leq \left( \prod_{j=1}^d \int_{\mathbb{R}} |Du(x_1, x_2, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| dy_j \right)^{\frac{1}{d-1}}.$$

For each subset  $J \subset \{1, 2, \dots, d\}$ , we define  $A_J(x)$  uniquely so that

$$(8.25) \quad A_\emptyset(x) = Du(x), \quad A_{J \cup \{k\}} = \int_{\mathbb{R}} A_J(x_1, x_2, \dots, x_{k-1}, y_k, x_{k+1}, \dots, x_d) dy_k$$

whenever  $k \in \{1, 2, \dots, d\} \setminus J$ . In short,  $A_J(x)$  is obtained by integrating  $Du$  against all the variables in  $J$  over  $\mathbb{R}$ . Then (8.24) reads

$$(8.26) \quad |u(x)|^{\frac{d}{d-1}} \leq \prod_{j=1}^d A_{\{j\}}(x)^{\frac{1}{d-1}}.$$

We integrate this inequality with respect to  $x_1$ :

$$\begin{aligned} \int_{\mathbb{R}} |u(x_1, x_2, \dots, x_d)|^{\frac{d}{d-1}} dx_1 &\leq \int_{\mathbb{R}} \prod_{j=1}^d A_j(x) dx_1 \\ &= A_{\{1\}}(x)^{\frac{1}{d-1}} \cdot \int_{\mathbb{R}} \prod_{j=2}^d A_{\{j\}}(x)^{\frac{1}{d-1}} dx_1 \\ &\leq A_{\{1\}}(x)^{\frac{1}{d-1}} \cdot \prod_{j=2}^d A_{\{1,j\}}(x)^{\frac{1}{d-1}}. \end{aligned}$$

Note that in the product  $A_{\{1\}}(x)^{\frac{1}{d-1}} \cdot \prod_{j=2}^d A_{\{1,j\}}(x)^{\frac{1}{d-1}}$   $d-1$  factors contain the variable  $x_2$ . If we integrate this inequality with respect to  $x_2$ , then we obtain

$$(8.27) \quad \int_{\mathbb{R}} |u(x_1, x_2, \dots, x_d)|^{\frac{d}{d-1}} dx_1 dx_2 \leq A_{\{1,2\}}(x)^{\frac{2}{d-1}} \cdot \prod_{j=3}^d A_{\{1,2,j\}}(x)^{\frac{1}{d-1}}.$$

by virtue of the Hölder inequality.

Repeating this procedure, we obtain

$$(8.28) \quad \int_{\mathbb{R}} |u(x_1, x_2, \dots, x_d)|^{\frac{d}{d-1}} dx_1 dx_2 \dots dx_k \leq A_{\{1,2,\dots,k\}}(x)^{\frac{k}{d-1}} \cdot \prod_{j=k+1}^d A_{\{1,2,\dots,k,j\}}(x)^{\frac{1}{d-1}}$$

for all  $k = 1, 2, \dots, d$ . Therefore, it follows that

$$(8.29) \quad \int_{\mathbb{R}} |u(x_1, x_2, \dots, x_d)|^{\frac{d}{d-1}} dx_1 dx_2 \dots dx_d \leq A_{\{1,2,\dots,d\}}(x)^{\frac{d}{d-1}} = \|Du\|_1^{\frac{d}{d-1}}.$$

Thus the proof is complete when  $p = 1$ .

Consider now the case that  $1 < p < d$ . Then we set  $\gamma = \frac{p(d-1)}{d-p} > 1$ . Since  $\frac{\gamma d}{d-1} = q$ , we obtain

$$\left( \int_{\mathbb{R}^d} |u|^q \right)^{\frac{d-1}{d}} \leq \int_{\mathbb{R}^d} \gamma \cdot |u|^{\gamma-1} |Du| \leq \gamma \left( \int_{\mathbb{R}^d} |u|^q \right)^{\frac{p-1}{p}} \cdot \|Du\|_p.$$

Arranging this inequality, we obtain

$$(8.30) \quad \|u\|_q \leq \|Du\|_p.$$

This is the desired result.  $\square$

Function spaces are tools with which to describe the size and the smoothness of functions. Let  $p, q, m \in [1, \infty]$  and suppose that  $m$  is an integer. For example,  $L^q(\mathbb{R}^d)$  measures the size. Since the pointwise multiplication of  $L^\infty(\mathbb{R}^d)$ -functions is closed in  $L^q(\mathbb{R}^d)$ ,  $L^q(\mathbb{R}^d)$  never measures the smoothness of functions. However,  $W^{m,p}(\mathbb{R}^d)$  can control somehow the smoothness of functions. What counts about size and smoothness is that the control of smoothness can be transformed into that of size. This aspect of these two quantities can be illustrated by the next theorem.

**Theorem 8.14.** Let  $m \in \mathbb{N}$  and  $1 \leq p < d/m$ . Then define  $q$  by

$$(8.31) \quad m - \frac{d}{p} = -\frac{d}{q},$$

that is,

$$(8.32) \quad \frac{1}{p} - \frac{m}{d} = \frac{1}{q}.$$

Then,  $W^{m,p}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ .

*Proof.* It suffices to prove the case that  $m = 1$  because we can readily use the induction argument. In this case we have to prove that

$$(8.33) \quad \|u\|_q \lesssim \|u\|_{W^{1,p}}$$

for all  $u \in W^{1,p}(\mathbb{R}^d)$ . However,  $C_c^\infty(\mathbb{R}^d)$  being dense in  $W^{1,p}(\mathbb{R}^d)$ , we have only to prove (8.33) for  $u \in C_c^\infty(\mathbb{R}^d)$ . In this case Theorem 8.13 says more:

$$(8.34) \quad \|u\|_q \lesssim \|Du\|_p.$$

Therefore, the proof is now complete.  $\square$

**Definition 8.15.** In  $W^{m,p}(\mathbb{R}^d)$  the quantity  $m - \frac{d}{p}$  is referred to as the differential index of functions.

Related exercises. To conclude this section let us take up some calculus concerning to the differentiation.

*Exercise 88.* Let  $k \in \mathbb{N}$ . Show that, for each problem, there exists a polynomial  $p_k(x) \in \mathbb{Q}[x_1, \dots, x_d]$  whose degree is indicated below so that equality holds.

$$(1) \deg(p_k) = k.$$

$$(8.35) \quad \frac{d^k}{dx^k} e^{-x^2} = p_k(x) e^{-x^2}.$$

$$(2) \deg(p_k) = k.$$

$$(8.36) \quad \frac{d^k}{dx^k} \langle x \rangle = p_k(x) \langle x \rangle^{1-2k}.$$

$$(3) \deg(p_k) = mk - k.$$

$$(8.37) \quad \frac{d^k}{dx^k} (1 + |x|^2)^{-m} = p_k(x) (1 + |x|^2)^{-m(k+1)}.$$

Here  $m \in \mathbb{N}$ .

$$(4) \deg(p_k) = k - 1.$$

$$(8.38) \quad \frac{d^k}{dx^k} e^{-\frac{1}{x}} = \frac{p_k(x)}{x^{k+1}} e^{-\frac{1}{x}}.$$

for  $x > 0$ .

Notes and references for Chapter 4.

Section 5. Theorem 5.10 Theorem 5.15 Theorem 5.16 Theorem 5.24 Theorem 5.26 Theorem 5.29

Section 6.

## Section 7. Theorem 6.5

As for Theorem 6.6, Du Bois Reymond pointed out that there does exist a continuous function such that the Fourier series diverges in a point. After that Fejér and Lebesgue presented different proofs.

Fejér established the result for the convergence of the Fejér integral at a point having two-sided limit. Theorem 6.8 is just an application of his result. Kolmogorov considered this type of approximation by using the function space  $W^{r,\infty}(\mathbb{T})$  [272].

Theorem 6.9 Theorem 6.10 Theorem 6.11 Theorem 6.12 Theorem 6.13 is due to the work by Jackson in 1912 and to the work by Bernstein in 1913.

The Shanonn-Someya sampling theorem (Theorem 6.14) is discovered initially by Whittaker and Ogura. In 1915 Whittaker proposed a formulation of the theorem in terms of entire analytic functions. In 1920 K. Ogura pointed out a mistake of his formulation [377]. K. Ogura gave a correct proof of the Shanonn-Someya sampling theorem and Shanonn and Someya independently discovered the theorem in terms of the band-limited distributions. For more details we refer to [108].

Theorem 7.4

Theorem 7.6

Theorem 7.7

Theorem 7.8

Theorem 7.9

Theorem 7.11

Theorem 7.12

Theorem 7.13

Theorem 7.15

Theorem 7.17

Theorem 7.19

Theorem 7.20

Theorem 7.21

Theorem 7.22

Theorem 7.31

Theorem 7.33

Theorem 7.38

Theorem 7.39

Theorem 7.40

Theorem 7.41



Theorem 7.42

Theorem 7.44

Theorem 7.45

Theorem 7.47

Section 8. We refer to [13] for details on  $W^{m,p}(\mathbb{R}^d)$ .

Theorem 8.5 Theorem 8.6 Theorem 8.10 Theorem 8.11 Theorem 8.13 was obtained by Gagliard and Nirenberg in 1959.

In Theorem 8.14 when the domain is bounded and

$$(8.39) \quad m - \frac{d}{p} > -\frac{d}{q},$$

then the embedding is compact. This result is known as Rellich-Kondrachev results obtained in 1930 and 1945.

## Part 5. Elementary facts of functional analysis

The aim of this in this section is to review functional analysis. In Chapter 5 we build up theory on functional analysis. What is totally different from the usual book on functional analysis is that we have already kept abundant examples in mind. In Section 10 we make a view of the theory of Banach spaces. We develop theory of Hilbert spaces in Section 11. The aim of this section is just to make a quick review of Hilbert space theory.

### 9. NORMED SPACES

In this section we investigate normed spaces. The definition of normed spaces is given by the following.

**Definition 9.1** (Normed spaces).

- (1) A linear space  $X$  is said to be a normed space, if it comes with a function  $\|\cdot\|_X : X \rightarrow [0, \infty)$ , which satisfies the following property.  $\|\cdot\|_X : X \rightarrow [0, \infty)$  is said to be the norm on a linear space  $X$ .
  - (a) Let  $x \in X$ . Then  $\|x\|_X = 0$  implies  $x = 0$ .
  - (b)  $\|a \cdot x\|_X = |a| \cdot \|x\|_X$  for all  $a \in \mathbb{K}$  and  $x \in X$ .
  - (c) Let  $x, y \in X$ . Then  $\|x + y\|_X \leq \|x\|_X + \|y\|_X$ .
- (2) Equip a normed space  $X$  with a topology induced by the norm, that is, a subset  $B$  is an open set if for all  $x \in B$  there exists  $r = r_x > 0$  such that  $\{y \in X : \|x - y\|_X < r\} \subset B$ .
- (3) If any Cauchy sequence in  $X$  converges, then  $X$  is said to be a Banach space.

Sometimes we write  $\|\cdot\|_X$  for  $\|\cdot\|$  when there is no possibility of confusion.

*Exercise 89.*

- (1) Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $1 < p < \infty$ . Then show that  $L^p(\mu)$  is a normed space.
- (2) Let  $(X, d)$  be a distance space. Then show that  $BC(X)$  is a normed space.

#### 9.1. Linear operators on normed spaces.

In analysis, continuity is one of the key notion. When we deal with normed spaces, continuity often takes place of boundedness. So, the notion of boundedness is very important.

Bounded operators. Since  $X$  carries the structure of a normed space, it is of importance to consider linear mappings. What happens if a mapping is linear and continuous? To answer the question, we present a definition.

**Definition 9.2** (Bounded mapping). A linear mapping from a normed space  $X$  to a normed space  $Y$  is said to be bounded, if there exists a constant  $\kappa > 0$  such that

$$(9.1) \quad \|Ax\|_Y \leq \kappa \|x\|_X.$$

The infimum of such  $\kappa$  is called the operator norm of  $A$  and it is denoted by  $\|A\|_{B(X,Y)}$ . When  $X = Y$ , then abbreviate it to  $\|A\|_{B(X)}$ .

The following theorem is elementary in Banach space theory.

**Theorem 9.3.** *Let  $X$  and  $Y$  be normed spaces. A linear mapping from  $A : X \rightarrow Y$  is continuous if and only if  $A$  is bounded.*

*Proof.* Assume  $A$  is continuous. The set  $\{x \in X : Ax \in B_Y(1)\} = A^{-1}(B_Y(1))$  an open set containing  $0 \in X$  by virtue of the continuity of  $A$ . Therefore, there exists  $\eta > 0$  such that

$$(9.2) \quad B_X(\eta) \subset A^{-1}(B_Y(1)).$$

Thus, from linearity of  $A$  together with (9.2), we obtain

$$(9.3) \quad \|Ax\|_Y \leq \frac{1}{\eta} \|x\|_X \text{ for all } x \in X,$$

which implies  $A$  is bounded.

Assume  $A$  is bounded. Let  $\kappa$  be a constant in Definition 9.2. Since  $X$  is a metric space, we have only to show  $\{Ax_j\}_{j \in \mathbb{N}}$  is convergent to  $Ax$  whenever  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence converges to  $x$ . Since  $A$  is assumed linear and bounded, we obtain

$$(9.4) \quad \|Ax - Ax_j\|_Y = \|A(x - x_j)\|_Y \leq \kappa \|x - x_j\|_X.$$

Letting  $j \rightarrow \infty$ , we conclude  $\lim_{j \rightarrow \infty} Ax_j = Ax$ . Therefore  $A$  is continuous. □

**Example 9.4.** Let  $(X, \mathcal{B}, \mu)$  be a  $\sigma$ -finite measure space and  $1 \leq p \leq \infty$ . Suppose that  $f$  is a  $\mu$ -measurable function. Then

$$(9.5) \quad M_f : g \in L^p(\mu) \rightarrow f \cdot g \in L^p(\mu)$$

is a well-defined operator if and only if  $f \in L^\infty(\mu)$ .

*Proof.* If  $f \in L^\infty(\mu)$ , then it is clear that  $M_f(g) \in L^p(\mu)$  whenever  $g \in L^p(\mu)$ . Assume instead that  $f \notin L^\infty(\mu)$ . Then for all  $M \in \mathbb{N}$ , we can choose a measurable set  $A_M \in \mathcal{B}$  such that  $|f(x)| \geq M$  on  $A_M$  and that  $0 < \mu(A_M) < \infty$ . Set  $g_M(x) := \text{sgn}(\overline{f(x)})$  for  $x \in X$ . Then

$$(9.6) \quad \|M_g f\|_p^p = \int_X |g_M(x) \cdot f(x)|^p d\mu(x) = \int_{A_M} |f(x)|^p d\mu(x) \geq M^p \mu(A_M).$$

Thus,  $M_g$  cannot be bounded. □

*Exercise 90.* Let  $\{A_j\}_{j \in \mathbb{N}}$  and  $\{B_j\}_{j \in \mathbb{N}}$  be a sequence in  $B(X)$ . Assume  $\lim_{j \rightarrow \infty} A_j = A$  and  $\lim_{j \rightarrow \infty} B_j = B$ . Namely, assume that

$$\lim_{j \rightarrow \infty} \|A_j - A\|_{B(X)} = \lim_{j \rightarrow \infty} \|B_j - B\|_{B(X)} = 0.$$

Then show that  $\lim_{j \rightarrow \infty} A_j B_j = AB$ .

*Exercise 91.*

- (1) As long as the definition makes sense, we define the Laplace transform by

$$(9.7) \quad \mathcal{L}f(p) := \int_0^\infty e^{-tp} f(t) dt$$

for measurable functions  $f$  on  $(0, \infty)$ . Show that  $\mathcal{L}$  is not  $L^\infty((0, \infty))$ -bounded.

- (2) We define the operator

$$(9.8) \quad Lf(p) := p \int_0^\infty e^{-tp} f(t) dt = p \mathcal{L}f(p).$$

Show that  $L$  is  $L^\infty((0, \infty))$ -bounded.

**Example 9.5** (Schur's lemma). Suppose that  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{C}, \nu)$  are  $\sigma$ -finite measure spaces and assume that  $K : X \times Y \rightarrow \mathbb{C}$  is a measurable function satisfying

$$(9.9) \quad M_0 := \sup_{x \in X} \int_X |K(x, y)| d\mu(y), \quad M_1 := \sup_{y \in Y} \int_Y |K(x, y)| d\nu(x) < \infty.$$

Let  $1 \leq p \leq \infty$ . Then

$$(9.10) \quad T_K f(y) := \int_X K(x, y) f(x) d\mu(x)$$

is a bounded  $L^p(\mu)$ - $L^p(\nu)$  operator satisfying

$$(9.11) \quad \|T_K\|_{L^p(\mu) \rightarrow L^p(\nu)} \leq M_0^{1-\theta} M_1^\theta.$$

Indeed, if  $p = 1, \infty$ , this is clear by assumption. If  $1 < p < \infty$ , we use interpolation. The proof will be obtained in Chapter 16.

**Example 9.6** (Schmidt operator). Suppose that  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{C}, \nu)$  are  $\sigma$ -finite measure spaces and assume that  $K : X \times Y \rightarrow \mathbb{C}$  is a measurable function satisfying

$$(9.12) \quad \|K\|_{L^2(\mu \otimes \nu)} = \left( \iint_{X \times Y} |K(x, y)|^2 d\mu \otimes \nu(x, y) \right)^{\frac{1}{2}} < \infty.$$

Then

$$(9.13) \quad T_K f(y) := \int_X K(x, y) f(x) d\mu(x)$$

is a bounded  $L^2(\mu)$ - $L^2(\nu)$  linear operator with norm estimate

$$(9.14) \quad \|T_K\|_{B(L^2(\mu), L^2(\nu))} \leq \|K\|_{L^2(\mu \otimes \nu)}.$$

Indeed, by the Hölder inequality we have

$$\begin{aligned} \int_Y |T_K f(x)|^2 d\nu(x) &\leq \int_Y \left\{ \int_X |K(x, y)|^2 d\mu(x) \cdot \int_X |f(x)|^2 d\mu(x) \right\} d\nu(y) \\ &= \|K\|_{L^2(\mu \otimes \nu)}^2 \cdot \int_X |f(x)|^2 d\mu(x). \end{aligned}$$

This is the desired result.

Compact operators.

Let  $X, Y$  be Banach spaces. If the dimension is infinite, then the unit ball is not compact. Even when the set is bounded, the set is not pre-compact. Therefore, it makes sense to consider compactness in earnest.

**Definition 9.7** (Compact operator). Let  $A : X \rightarrow Y$  be a bounded linear operator. The operator  $A$  is said to be compact, if  $A$  maps  $X_1$ , the unit ball of  $X$  to a relatively compact set. Denote by  $K(X, Y)$  the set of all compact operators from  $X$  to  $Y$ . Also define  $K(X) := K(X, X)$ .

**Proposition 9.8.** *Let  $A$  be a finite rank operator, that is,  $\text{Range}(A)$  is a finite dimensional subspace of  $Y$ . Then  $A$  is compact.*

*Proof.* This is trivial because any bounded subset of a finite dimensional linear space is compact.  $\square$

The next proposition implicitly implies that some element in  $B(X, Y)$  is not approximated by the norm topology by  $K(X, Y)$ .

**Theorem 9.9.** *The space  $K(X, Y)$  is a closed subspace of  $B(X, Y)$  in the norm topology.*

*Proof.* Suppose that  $\{A_j\}_{j=1}^\infty$  is a sequence of compact operators converging to  $A \in B(X)$ . To prove that  $A$  is compact, it suffices to show that  $A(X_1)$  is relatively compact in  $Y$ . Let  $\varepsilon > 0$  be given. Then the convergence in norm topology gives us a large integer  $N$  with the following property: For all  $y \in A(X_1)$  we can find  $x \in X_1$  such that  $\|y - Ax\|_Y < \frac{\varepsilon}{2}$ . Since  $A_N$  is compact, we can find  $x_1, x_2, \dots, x_M$  such that

$$(9.15) \quad A_N(X_1) \subset B\left(x_1, \frac{\varepsilon}{2}\right) \cup B\left(x_2, \frac{\varepsilon}{2}\right) \cup \dots \cup B\left(x_M, \frac{\varepsilon}{2}\right).$$

From the property of  $N$  it follows that

$$(9.16) \quad A(X_1) \subset B(x_1, \varepsilon) \cup B(x_2, \varepsilon) \cup \dots \cup B(x_M, \varepsilon).$$

This is the desired result. □

*Exercise 92* (Approximation number). Let  $T : X \rightarrow Y$  be a bounded linear operator. For each  $n \in \mathbb{N}$ , we define

$$(9.17) \quad a_n(T) := \inf\{\|T - S\| : S \in B(X, Y), \dim(\text{Im}(S)) < n\}.$$

- (1) Show that  $\{a_n(T)\}_{n \in \mathbb{N}}$  is decreasing.
- (2) Show that  $T$  is compact if and only if  $\lim_{n \rightarrow \infty} a_n(T) = 0$ .

Here we content ourselves with seeing some examples of compact operators instead of going into more details.

**Example 9.10.** The integral operator  $T_K$ , presented in Example 9.6, is a compact operator in  $L^2(\mu \otimes \nu)$ . Indeed, we can approximate  $K$  in  $L^2(\mu \otimes \nu)$ -norm by the function of the form

$$(9.18) \quad \tilde{K} := \sum_{j=1}^k a_j \chi_{E_j \times F_j}, \quad k < \infty.$$

Since  $T_{\tilde{K}}$  is a finite rank operator, it is compact by Proposition 9.8. Therefore, by virtue of Theorem 9.9  $T_K$  is compact because it is a norm limit of such compact operators.

*Exercise 93.* Let  $X, Y, Z, W$  be Banach spaces. Show that  $ABC \in K(X, W)$ , whenever  $A \in B(X, Y)$ ,  $B \in K(Y, Z)$  and  $C \in B(Z, W)$ .

*Exercise 94.* Let  $a = \{a_j\}_{j \in \mathbb{N}} \in \ell^\infty(\mathbb{N})$ . Then for  $1 \leq p < \infty$ , the mapping

$$(9.19) \quad x = \{x_j\}_{j \in \mathbb{N}} \in \ell^p(\mathbb{N}) \mapsto \{a_j \cdot x_j\}_{j \in \mathbb{N}} \in \ell^p(\mathbb{N})$$

is compact if and only if  $\lim_{j \rightarrow \infty} a_j = 0$ .

**Closed operators.** In this paragraph we deal with the linear mapping which is not necessarily defined on  $X$ , that is, which is defined on a linear subspace of  $X$ .

In dealing with the partial differential equations it is indispensable that we use unbounded operators. Among such nasty operators, we can say that of importance is the class called closed. As an example, we shall see the pointwise multiplication operators and the differentiation operators are closed.

**Definition 9.11** (Closed operator). A mapping  $A$  is said to be a closed operator from  $X$  to  $Y$  with domain  $D(A)$ , if it satisfies the following conditions:

- (1)  $D(A)$  is a linear subspace.
- (2)  $A$  is a linear operator from  $D(A)$  to  $Y$ .  $D(A)$  is said to be the domain of  $A$ .
- (3) The graph of  $A$ , defined by  $\{(x, Ax) \in X \times Y : x \in D(A)\}$ , is closed.

If  $B$  is a linear operator that is an extension of  $A$ , then denote  $A \subset B$ .

We remark that  $A : X \rightarrow Y$  is a linear operator does not always mean  $D(A) = X$ .

**Example 9.12.** Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $F$  a  $\mu$ -measurable function. Let  $1 \leq p \leq \infty$ . Define

$$D(M_F) := \{f \in L^p(\mu) : F \cdot f \in L^p(\mu)\}, \quad M_F : D(M_F) \rightarrow L^p(\mu), \quad M_F(f) = F \cdot f.$$

Then  $M_F$  is a closed operator on  $L^p(\mu)$ .

*Proof.* To prove this, let us take a sequence  $f_1, f_2, \dots \in D(M_F)$  and assume that  $\lim_{j \rightarrow \infty} f_j = f$  and  $\lim_{j \rightarrow \infty} M_F f_j = g$  hold in  $L^p(\mu)$ . A passage to a subsequence allows us to assume that  $\lim_{j \rightarrow \infty} f_j(x) = f(x)$  and  $\lim_{j \rightarrow \infty} M_F f_j(x) = g(x)$  hold for  $\mu$ -almost all  $x \in X$ . Since  $\lim_{j \rightarrow \infty} M_F f_j = g$  takes place in  $L^p(\mu)$ , we have

$$(9.20) \quad \sup_{j \in \mathbb{N}} \|F \cdot f_j\|_{L^p(\mu)} = \sup_{j \in \mathbb{N}} \|M_F f_j\|_{L^p(\mu)} < \infty.$$

Thus, Fatou's lemma gives

$$(9.21) \quad \|F \cdot f\|_{L^p(\mu)} \leq \sup_{j \in \mathbb{N}} \|F \cdot f_j\|_{L^p(\mu)} < \infty$$

together with the fact  $\lim_{j \rightarrow \infty} f_j(x) = f(x)$  for  $\mu$ -almost all  $x \in X$ . Thus,  $f \in D(M_F)$ .

Another application of Fatou's lemma gives

$$(9.22) \quad \sup_{j \geq J} \|F \cdot f - F \cdot f_j\|_{L^p(\mu)} \leq \sup_{j \geq J} \liminf_{k \rightarrow \infty} \|F \cdot f_k - F \cdot f_j\|_{L^p(\mu)} \leq \sup_{j, k \geq J} \|F \cdot f_k - F \cdot f_j\|_{L^p(\mu)}.$$

Since the right-hand side of the above inequality tends to 0 as  $J \rightarrow \infty$ , we conclude  $F \cdot f_j \rightarrow F \cdot f$  in  $L^p(\mu)$  as  $j \rightarrow \infty$ . Since  $F \cdot f_j \rightarrow g$  in  $L^p(\mu)$  as  $j \rightarrow \infty$ , we conclude  $g = F \cdot f = M_F f$ .  $\square$

*Exercise 95.* Let  $X, Y, Z, W$  be Banach spaces. Suppose that we are given a closed operator  $A : X \rightarrow Y$  and isomorphisms  $B : W \rightarrow X$  and  $C : Y \rightarrow Z$ . Then  $CAB : W \rightarrow Z$  is a closed operator.

The following lemma characterizes closedness.

**Lemma 9.13.** Let  $A : X \rightarrow Y$  be a linear operator whose domain is  $D(A)$ , that is,  $A$  is a linear mapping from a linear subspace  $D(A)$  to  $Y$ . Endow  $D(A)$  with a graph norm defined by

$$(9.23) \quad \|x\|_{D(A)} := \|x\|_X + \|Ax\|_Y.$$

Then show that  $A$  is closed if and only if  $(D(A), \|\cdot\|_{D(A)})$  is a Banach space.

*Proof.* Assume that  $A$  is closed. Suppose we are given a Cauchy sequence  $\{x_j\}_{j \in \mathbb{N}}$  in  $D(A)$ . Then we deduce  $\{x_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence in  $X$  and that  $\{Ax_j\}_{j \in \mathbb{N}}$  is convergent from the definition of the graph norm of  $A$ . Therefore, letting  $x$  be a limit point of  $\{x_j\}_{j \in \mathbb{N}}$ , we see that  $x \in D(A)$  and  $\lim_{j \rightarrow \infty} Ax_j = Ax$ . This implies  $\{x_j\}_{j \in \mathbb{N}}$  converges to  $x$  in the graph norm. As a result  $(D(A), \|\cdot\|_{D(A)})$  is a Banach space.

Assume instead that  $(D(A), \|\cdot\|_{D(A)})$  is a Banach space. Suppose that  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence in  $D(A)$  such that it is convergent to  $x$  and that  $\{Ax_j\}_{j \in \mathbb{N}}$  is convergent. Then  $\{x_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence of  $(D(A), \|\cdot\|_{D(A)})$ . Therefore  $y := \lim_{j \rightarrow \infty} x_j$  exists in  $D(A)$ , which implies that  $Ay = \lim_{j \rightarrow \infty} Ax_j$ . Of course the convergence in  $D(A)$  is stronger than  $X$ , so that  $y = x$ . From the observation above we conclude  $x \in D(A)$  and  $Ax = \lim_{j \rightarrow \infty} Ax_j$ . Thus,  $A$  is closed.  $\square$

To conclude this section we introduce the notion of closable operators.

**Definition 9.14.** Let  $A : X \rightarrow Y$  and  $B : X \rightarrow Y$  be unbounded linear operators.

- (1)  $B$  is said to be an extension of  $A$ , if  $D(A) \subset D(B)$  and  $Ax = Bx$  for all  $x \in D(A)$ .
- (2)  $A$  is closable, if  $A$  admits an extension to a closed operator.

**9.2. Resolvent.**

Let  $X$  be a Banach space.

**Definition 9.15** (Resolvent). Let  $A$  be a closed operator on  $X$ . Then  $\rho(A)$  is the set of all complex numbers  $\lambda$  such that there exists a unique bounded linear operator  $B$  satisfying  $(A - \lambda)B = \text{id}_X$  and  $B(A - \lambda) = \text{id}_{D(A)}$ . Below write this  $B$  as  $R(\lambda) = R(\lambda; A) = (A - \lambda)^{-1}$ . The spectrum set  $\sigma(A)$  of  $A$  is the complement of  $\rho(A)$ .

**Lemma 9.16.** *The resolvent set  $\sigma(A)$  is not empty.*

*Proof.* The function  $R$  is a bounded holomorphic non-constant function. Therefore, by virtue of Lieuville's theorem its domain  $\rho(A)$  cannot coincide with  $\mathbb{C}$ . □

**Proposition 9.17** (Resolvent equation). *Let  $A$  be a closed operator. Then we have*

$$(9.24) \quad R(\lambda) - R(\mu) = (\lambda - \mu)R(\lambda)R(\mu)$$

for all  $\lambda, \mu \in \rho(A)$ .

*Proof.* If  $X = \mathbb{C}$ , then it could not be easier: Identify  $a$  with the multiplication operator and we obtain

$$(9.25) \quad \frac{1}{a - \lambda} - \frac{1}{a - \mu} = (\lambda - \mu) \frac{1}{a - \lambda} \cdot \frac{1}{a - \mu}.$$

The crux of the proof is to examine carefully the above calculation above. We use  $R(\mu)(A - \mu) = \text{id}_{D(A)}$  and  $(A - \lambda)R(\lambda) = \text{id}_X$  to obtain

$$(9.26) \quad R(\lambda) - R(\mu) = R(\mu)(A - \mu)R(\lambda) - R(\mu)(A - \lambda)R(\lambda) = (\lambda - \mu)R(\lambda)R(\mu).$$

This is the desired result. □

Keep to the same setting as Example 9.12. Let us calculate  $\sigma(M_F)$ . To formulate it, we need a definition.

**Definition 9.18.** Define

$$(9.27) \quad \text{Essrange}(F) := \{\lambda \in \mathbb{C} : \mu(F^{-1}(B(\lambda, r))) > 0 \text{ for all } r > 0\}.$$

**Proposition 9.19.** *If  $\mu$  is a measurable function, then  $\sigma(M_F) = \text{Essrange}(F)$ .*

*Proof.* Suppose that  $\lambda \notin \text{Essrange}(F)$ . Then

$$(9.28) \quad G := \frac{1}{a - F} \in L^\infty(\mu).$$

Therefore  $M_F - \lambda$  has inverse. Suppose instead that  $\lambda \in \rho(M_F)$ . Then there exists a linear operator  $R$  such that

$$(9.29) \quad (M_F - \lambda)Rg = g.$$

for  $g \in L^p(\mu)$ . From  $(M_F - \lambda)Rg = g$  for all  $g \in L^p(\mu)$ , we conclude  $\mu\{F = \lambda\} = 0$ . Therefore, by disregarding such a null set, we obtain

$$(9.30) \quad Rg = \frac{1}{F - \lambda} \cdot g.$$

In order that this multiplication operator is bounded, the multiplier must belong to  $L^\infty(\mu)$ . This implies  $\lambda \notin \text{Essrange}(F)$ .  $\square$

### 9.3. Quotient topology and quotient vector spaces.

General theory on quotient topology. To begin with let us investigate some elementary facts on general topology.

**Definition 9.20.** An equivalence relation of a set  $X$  is a subset  $R$  of  $X \times X$  satisfying the following. Below, for  $x, y \in X$ ,  $x \sim y$  means that  $(x, y) \in R$ .

- (1)  $x \sim x$  for all  $x \in X$  (Reflexivity).
- (2) Let  $x, y \in X$ . Then  $x \sim y$  implies  $y \sim x$  (Symmetry).
- (3) Let  $x, y, z \in X$ . Then  $x \sim y$  and  $y \sim z$  implies  $x \sim z$  (Transitivity).

In this case  $\sim$  is an equivalence relation of  $X$ . Given an equivalence relation of  $X$ , we write

$$(9.31) \quad [x] := \{y \in X : x \sim y\} \in 2^X$$

for  $x \in X$  and

$$(9.32) \quad X/\sim := \{[x] : x \in X\} \subset 2^X.$$

**Definition 9.21.** Let  $\sim$  be an equivalence relation of a topological space  $X$ . Then the quotient topology of  $X$  with respect to  $\sim$  is the weakest topology such that

$$(9.33) \quad p : X \rightarrow X/\sim, x \mapsto [x]$$

is continuous. The mapping  $p$  is said to be the (canonical/natural) projection.

According to the definition, we see that  $U \subset X/\sim$  is open, if and only if  $p^{-1}(U)$  is open.

As for this topology the following is elementary.

**Theorem 9.22.** Let  $X$  and  $Y$  be topological spaces and  $\sim$  an equivalence relation of  $X$ . A mapping  $f : X/\sim \rightarrow Y$  is continuous, if and only if  $f \circ p : X \rightarrow Y$  is continuous.

*Proof.* Suppose  $f$  is continuous. Then  $f \circ p$  is continuous since  $f$  and  $p$  are both continuous.

Conversely suppose that  $f \circ p$  is continuous. Let  $U$  be an open set in  $Y$ . Then we have to show  $f^{-1}(U)$  in order to prove that  $f$  is continuous. Set  $V = p^{-1}(f^{-1}(U))$ . Since  $p^{-1}(f^{-1}(U)) = (f \circ p)^{-1}(U)$  and  $f \circ p$  is continuous,  $p^{-1}(f^{-1}(U))$  is open. Therefore, from the remark just above, we see that  $f^{-1}(U)$  is open.  $\square$

Below we present an application of this theorem in order to be familiarized with the quotient topology.

**Example 9.23.** We define an equivalence relation  $\sim$  of  $\mathbb{C}^2 \setminus \{0\}$  by

$$(9.34) \quad (a, b) \sim (c, d) \iff a = kc, b = kd \text{ for some } k \in \mathbb{C} \setminus \{0\}.$$

Now it might be helpful to denote by  $[a : b]$  the equivalence class to which  $(a, b) \in \mathbb{C}^2 \setminus \{0\}$  belongs, which immediately reminds us that this equivalence class is defined by ratio of two complex numbers. Set  $\mathbb{C}P^1 := (\mathbb{C}^2 \setminus \{0\})/\sim$ . Define  $f : \mathbb{C}P^1 \rightarrow \mathbb{C}P^1$  by

$$(9.35) \quad [a, b] \mapsto [a^2 : b^2].$$

Note that this mapping is well-defined despite the ambiguity of the multiplicative constants. We are to show that  $f$  is continuous.



*Proof.* To prove this, we have only to show

$$(9.36) \quad f \circ p : \mathbb{C}^2 \setminus \{0\} \rightarrow \mathbb{C}P^1$$

is continuous. If we set

$$(9.37) \quad F(a, b) = (a^2, b^2), (a, b) \in \mathbb{C}^2 \setminus \{0\},$$

then  $f \circ p = p \circ F$ . Since  $p$  and  $F$  are continuous, so is  $f \circ p$ . Therefore  $f$  is continuous.  $\square$

**Example 9.24.** We equip  $X \times X$  with equivalence relation as follows:

$$(9.38) \quad (x_1, x_2) \sim (y_1, y_2) \iff x_1 \sim y_1, x_2 \sim y_2.$$

Then we have

$$(9.39) \quad X \times X / \sim \simeq X / \sim \times X / \sim .$$

*Exercise 96.* Prove (9.39).

*Exercise 97.* We define an equivalence relation  $\sim$  of  $\mathbb{C}^{d+1} \setminus \{0\}$  by

$$(9.40) \quad (a_0, a_1, \dots, a_d) \sim (b_0, b_1, \dots, b_d) \iff a_0 = k b_0, a_1 = k b_1, \dots, a_d = k b_d \text{ for some } k \in \mathbb{C} \setminus \{0\}.$$

We write  $[a_0, a_1, \dots, a_d]$  to denote the equivalence class  $(a_0, a_1, \dots, a_d)$  belongs to as before. Set  $\mathbb{C}P^d := (\mathbb{C}^{d+1} \setminus \{0\}) / \sim$ . Let  $m \in \mathbb{N}$ . Then  $f : \mathbb{C}P^d \rightarrow \mathbb{C}P^d$ , given by

$$(9.41) \quad [a_0, a_1, \dots, a_d] \mapsto [a_0^m, a_1^m, \dots, a_d^m],$$

is a continuous mapping.

The lift operator is an operator that will facilitate our discussion throughout : It will always reduce the argument to the one when  $s = 0$ . Having proved the multiplier theorem, we can prove boundedness of the lift operators.

Quotient normed space. Let  $X$  be a normed space and  $Y$  be a closed subspace. Set

$$(9.42) \quad X/Y := \{x + Y : x \in X\}.$$

We make  $X/Y$  into a normed space as follows :

**Definition 9.25.** The quotient space  $X/Y$  is a linear space whose operations are defined by

$$(9.43) \quad (a, x + Y) \in \mathbb{K} \times X/Y \mapsto a \cdot (x + Y) := a \cdot x + Y \in X/Y$$

and

$$(9.44) \quad (x + Y, y + Y) \in X/Y \times X/Y \mapsto (x + Y) + (y + Y) := (x + y) + Y \in X/Y.$$

Equip the linear space  $X/Y$  with the norm

$$(9.45) \quad \|x + Y\|_{X/Y} = \inf\{\|x'\|_X : x - x' \in Y\}.$$

**Proposition 9.26.** *The linear space  $X/Y$  defined above is a normed space whose topology agrees with the quotient topology.*

*Proof.* Here, we have to prove the following.

- (1) The scalar multiplication does not depend on the particular choice of the representative.
- (2) The definitions of the addition and the norm do not depend on the particular choice of the representative.
- (3) The set  $X/Y$  is a linear space.
- (4) The mapping  $x + Y \mapsto \|x + Y\|_{X/Y}$  is a norm.
- (5) The two topologies of  $X/Y$  coincide.

The proofs being all routine, we content ourselves with proving that the addition is well-defined. To do this, let  $x_1, x_2, x'_1, x'_2 \in X$  and assume that  $x + Y = x'_1 + Y$  and  $x_2 + Y = x'_2 + Y$ . Then the matter is to show that  $(x_1 + x_2) + Y = (x'_1 + x'_2) + Y$ . This is equivalent to showing  $x_1 + x_2 - x'_1 - x'_2 \in Y$  under the assumption that  $x_1 - x'_1, x_2 - x'_2 \in Y$ . This is the case because  $Y$  is a linear space. Therefore, the proof of our claim is complete.  $\square$

*Proof.* Complete the proof of Proposition 9.26 to fill the details of the assertions to check which is listed above.  $\square$

#### 9.4. Banach Alaoglu theorem.

The Banach Alaoglu theorem was taken up many times before. Let us formulate and prove it. The proof is obtained by the axiom choice. Namely our proof here largely depends on the Tychnov theorem, asserting that the product space of compact topological spaces is compact.

**Theorem 9.27** (Banach Alaoglu-(1)). *Let  $X$  be a normed space. Assume  $\{x_j^*\}_{j \in \mathbb{N}}$  belongs to the closed unit ball  $B$  in  $X^*$ . Then there exists a subsequence  $\{x_{j_k}^*\}_{k \in \mathbb{N}}$  convergent with respect to weak-\* topology. That is,*

$$(9.46) \quad \lim_{k \rightarrow \infty} x_{j_k}^*(x) = x^*(x)$$

for all  $x \in X$ .

*Proof.* We define a compact space  $\mathcal{E}$  by

$$(9.47) \quad \mathcal{E} := \prod_{x \in X} \{z \in \mathbb{C} : |z| \leq \|x\|_X\}.$$

Recall that the product topology, the topology of  $\mathcal{E}$ , is the weakest topology such that the projection  $\{z_x\}_{x \in B} \mapsto z_y$  is continuous for all  $y \in B$ .

Since  $x_j^*$  belongs to the unit ball  $B \subset X^*$ ,  $\{x_j^*(x)\}_{x \in B} \in \mathcal{E}$  for all  $j \in \mathbb{N}$ . Therefore we can choose a subsequence  $\{x_{j_k}^*\}_{k \in \mathbb{N}}$  so that  $\lim_{k \rightarrow \infty} x_{j_k}^*(x)$  exists for all  $x \in B$ . By homogeneity the limit  $\lim_{k \rightarrow \infty} x_{j_k}^*(x)$  exists for all  $x \in X$ . Therefore setting

$$(9.48) \quad x^*(x) := \lim_{k \rightarrow \infty} x_{j_k}^*(x),$$

then it follows that  $x_{j_k}^* \rightarrow x^*$  in the weak topology of  $X^*$ .  $\square$

#### 9.5. Hahn-Banach theorem.

Apart from the topological structure of normed space, we are now going to prove a theorem which has a strong flavor of algebra. The result does not involve any topological structure and hence it appears very frequently.

**Theorem 9.28** (Hahn-Banach theorem: real version). *Let  $V$  be an  $\mathbb{R}$ -vector space and  $p : V \rightarrow \mathbb{R}$  a function satisfying*

$$(9.49) \quad p(u + v) \leq p(u) + p(v), \quad p(\alpha u) = \alpha p(u)$$

for all  $\alpha \geq 0$  and  $u, v \in V$ . Assume that  $l_0 : W_0 \rightarrow \mathbb{R}$ , defined on a subspace  $W_0$ , is a linear mapping satisfying

$$(9.50) \quad l_0(u) \leq p(u)$$

for all  $u \in W_0$ . Then there exists a linear mapping  $L : V \rightarrow \mathbb{R}$  such that  $L|_{W_0} = l_0$  and

$$(9.51) \quad L(u) \leq p(u)$$

for all  $u \in V$ .

*Proof.* We define

$$(9.52) \quad \mathcal{Z} := \{(l, W) : W \text{ is a subspace of } V, W_0 \subset W, l|_W = W_0, l_0 \text{ satisfies (9.51)}\}.$$

Define a partial order of  $\mathcal{Z}$  by

$$(9.53) \quad (l_1, W_1) \ll (l_2, W_2) \iff W_1 \subset W_2 \text{ and } l_1 = l_2|_{W_1}.$$

Assume that  $W$  is a proper subset. Choose  $v \in V \setminus W$  and define

$$(9.54) \quad W' := \{av + w : a \in \mathbb{R}, w \in W\}.$$

Choose  $\alpha$  so that

$$(9.55) \quad \sup_{u \in V} (l(u) - p(u - v)) \leq \alpha \leq \sup_{u \in V} (p(u + v) - l(u)).$$

Then since any element  $w'$  in  $W'$  can be written uniquely as  $w' = av + w$  with some  $a \in \mathbb{R}$  and  $w \in W$ , the definition

$$(9.56) \quad l'(w') = a\alpha + l(w)$$

makes sense.

If  $a > 0$ , then we obtain

$$(9.57) \quad l(w') = a\alpha + l(w) \leq a \left( p \left( \frac{w}{a} + v \right) - l \left( \frac{w}{a} \right) \right) + l(w) = p(w + av) = p(w').$$

If  $a < 0$ , then we have

$$(9.58) \quad l(w') = a\alpha + l(w) \leq a \left( -p \left( -\frac{w}{a} - v \right) + l \left( -\frac{w}{a} \right) \right) + l(w) = p(w + av) = p(w').$$

Therefore,  $(l', W')$  is strictly larger than  $(l, W)$ .

Meanwhile  $\mathcal{Z}$  is an inductive partially ordered set. By Zorn's lemma  $\mathcal{Z}$  has a maximal element  $(L, W)$ . From the above observation and maximality of  $(L, W)$ , we conclude  $W = V$ .  $\square$

**Theorem 9.29** (Hahn-Banach theorem: complex version). *Let  $V$  be an  $\mathbb{C}$ -vector space and  $p : V \rightarrow \mathbb{C}$  a function satisfying*

$$(9.59) \quad p(u + v) \leq p(u) + p(v), \quad p(\alpha u) = |\alpha| p(u)$$

*for all  $\alpha \in \mathbb{C}$  and  $u, v \in V$ . Assume that  $l_0 : W_0 \rightarrow \mathbb{R}$ , defined on a subspace  $W_0$ , is a linear mapping satisfying*

$$(9.60) \quad l_0(u) \leq p(u)$$

*for all  $u \in W_0$ . Then there exists a linear mapping  $L : V \rightarrow \mathbb{R}$  such that  $L|_{W_0} = l_0$  and*

$$(9.61) \quad |L(u)| \leq p(u)$$

*for all  $u \in V$ .*

*Proof.* We set  $g(u) := \operatorname{Re} l_0(u)$  and  $h(u) := \operatorname{Im} l_0(u)$  for  $u \in V$ . Since  $l(iu) = il(u)$ , we obtain  $g(iu) = \operatorname{Re} il(u) = -\operatorname{Im} l_0(u) = -h(u)$ . Therefore,

$$(9.62) \quad l_0(u) = g(u) - ig(iu)$$

for all  $u \in W_0$ . Since  $g$  is  $\mathbb{R}$ -linear, we can extend  $g$  to a  $\mathbb{R}$  continuous linear functional  $G$  so that  $G(u) \leq p(u)$ . Since  $p(-u) = p(u)$  for all  $u \in V$ , we have  $|G(u)| \leq p(u)$ . Define

$$(9.63) \quad L(u) := G(u) - iG(iu) \quad u \in V.$$

Then since  $G$  is an extension of  $g$ , we see that  $L$  is an extension of  $l_0$ . Furthermore, since  $G$  is  $\mathbb{R}$ -linear, so is  $G$ . From the definition it follows that  $L(iu) = iL(u)$ . It remains to prove (9.61). To do this, choose a real number  $\theta$  so that  $e^{i\theta} L(u) \geq 0$ . Now that  $\mathbb{C}$ -linearity is already established, we obtain

$$(9.64) \quad |L(u)| = e^{i\theta} L(u) = \operatorname{Re} L(e^{i\theta} u) = G(e^{i\theta} u) \leq p(e^{i\theta} u) = p(u).$$

This is the desired result.  $\square$

Extension of linear mapping.

**Theorem 9.30.** *Let  $Y$  be a closed subspace of a normed space  $X$ . Then a bounded functional  $f : Y \rightarrow \mathbb{K}$  can be extended to a bounded functional  $F$  defined on  $X$  so that  $\|F\|_{X^*} = \|f\|_{Y^*}$ .*

*Proof.* It suffices to use 9.29 for

$$(9.65) \quad l_0 = f, p = \|f\|_{Y^*} \cdot \|\cdot\|_X.$$

$\square$

Norm attainer. Let  $X$  be a Banach space. Recall that we have defined

$$(9.66) \quad \|x^*\|_{X^*} = \sup_{x \in X_1} |x^*(x)|,$$

where  $X_1$  denotes the unit ball of  $X$ . Therefore, we have, denoting by  $X^*_1$  the unit ball of  $X^*$ ,

$$(9.67) \quad \|x\|_X \geq \sup_{x^* \in X^*_1} |x^*(x)|.$$

The following theorem asserts the supremum above can be attained by some  $x^* \in X^*_1$  as well as that the equality of the above formula holds.

**Theorem 9.31** (Existence of the norm attainer). *Let  $X$  be a normed space and  $x \in X$ . Then  $x$  has a norm attainer. Namely, there exists  $x^* \in X^*$  with unit norm such that*

$$(9.68) \quad x^*(x) = \|x\|_X.$$

*Proof.* We set

$$(9.69) \quad W_0 := \{k \cdot x : k \in \mathbb{K}\}, l(k \cdot x) := k \|x\|_X \text{ for } k \in \mathbb{K}, p(y) := \|y\|_X \text{ for } y \in X.$$

Then this triple satisfies the assumption of the Hahn-Banach theorem, whether  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{K} = \mathbb{C}$ . Therefore,  $l$  extends to a linear functional  $L$  on  $X$  satisfying

$$(9.70) \quad |L(y)| \leq \|y\|_X.$$

We remark that if  $\mathbb{K} = \mathbb{R}$ , we are still able to conclude that  $|L(y)| \leq p(y)$ , because  $p(-y) = p(y)$ . Set  $x^* := l$ . From (9.70), we see that  $x^* \leq 1$ . Since the equality is attained by  $x$ , we see that (9.68) holds.  $\square$

Complemental space. If  $X$  is a finite dimensional linear space, then for any subspace  $Y$  there exists a subspace  $Z$  such that  $X = Y \oplus Z$ . This is because, for example, the basis always exists. However, this is not the case when  $X$  is a normed space of infinite dimension.

**Definition 9.32.** Let  $X$  be a normed space and  $Y$  be a closed subspace. Then a closed subspace  $Z$  is said to be a complemental of  $Y$ , if  $X = Y \oplus Z$  as a direct sum.

**Lemma 9.33.** *If  $Y$  is a finite normed space, then any two norms are mutually equivalent.*

*Proof.* Let  $\{y_1, y_2, \dots, y_k\}$  be a basis of  $Y$ . We set

$$(9.71) \quad S := \left\{ \sum_{j=1}^k \alpha_j y_j : \sum_{j=1}^k |\alpha_j|^2 = 1 \right\}.$$

Then since  $S$  is compact, the function  $y \in S \mapsto \|y\|_Y$  has both maximum and minimum which are never 0. By homogeneity, we have

$$(9.72) \quad \left( \sum_{j=1}^k |\alpha_j|^2 \right)^{1/2} \sim \left\| \sum_{j=1}^k \alpha_j y_j \right\|_Y$$

for all  $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$ . If we set

$$(9.73) \quad \left\| \sum_{j=1}^k \alpha_j y_j \right\|_{Y_0} := \left( \sum_{j=1}^k |\alpha_j|^2 \right)^{1/2},$$

then this is a norm of  $Y$  equivalent to the original norm of  $Y$ . □

**Corollary 9.34.** *Any finite dimensional space of  $X$  is closed.*

*Proof.* Prove Corollary 9.34. □

**Theorem 9.35.** *If  $Y$  is a finite dimensional space of a normed space  $X$ , then  $Y$  has a complementary space.*

*Proof.* Let  $\{y_1, y_2, \dots, y_k\}$  be a basis of  $Y$ . Then define a continuous functional  $y_j^* : Y \rightarrow \mathbb{K}$  so that  $y_j^*(y_k) = \delta_{jk}$ .

By Lemma 9.33, if we keep to the same notation above, then we see there exists a constant  $c > 0$  such that

$$(9.74) \quad \|y^*(y)\|_* \lesssim \|y\|_{Y_0} \lesssim \|y\|_X$$

for all  $y \in Y$ . As a consequence  $y^*$  is continuous. If we set

$$(9.75) \quad Z = \bigcap_{j=1}^k \text{Ker}(y_j^*) = \{x \in X : y_j^*(x) = 0 \text{ for all } j = 1, 2, \dots, k\},$$

then a simple algebraic argument shows that  $Z$  is a complementary space of  $Y$ . □

*Exercise 98 (Herry's theorem).* The aim of this exercise is to consider extension of linear functionals. Let  $X$  be a normed space and  $f_1, f_2, \dots, f_d \in X^*$ . Suppose that  $f \in X^*$  satisfies

$$(9.76) \quad \bigcap_{j=1}^d \text{Ker} f_j \subset \text{Ker} f,$$

where we define

$$(9.77) \quad \text{Ker} g = g^{-1}(0)$$

for  $g \in X^*$ .

(1) Let us set

$$(9.78) \quad Y = \{(f_1(x), f_2(x), \dots, f_d(x)) : x \in X\} \subset \mathbb{R}^d.$$

Then show that the mapping

$$(9.79) \quad T : (f_1(x), f_2(x), \dots, f_d(x)) \in Y \mapsto f(x) \in \mathbb{R}$$

is well-defined.

(2) Let  $S$  be an arbitrary extension of  $T$ . Use this  $S$  to establish that  $f$  can be written as

$$(9.80) \quad f = \sum_{j=1}^d a_j f_j$$

for some constants  $a_1, a_2, \dots, a_d$ .

## 10. BANACH SPACES AND QUASI-BANACH SPACES

Classically, to solve the Dirichlet problem, for example, it is enough to consider Banach spaces or Hilbert spaces. However, a recent trend is that we go beyond the word of Banach spaces. One of the reasons may be non-linearity of functions. Suppose that  $f, g, h \in L^2(\mathbb{R}^d)$ . Then  $f \cdot g \cdot h \in L^{2/3}(\mathbb{R}^d)$  and  $L^{2/3}(\mathbb{R}^d)$  is not a Banach space. As this example shows, we are faced to the necessity of considering quasi-Banach spaces.

### 10.1. Elementary properties.

Recall that a normed space  $X$  is said to be a Banach space, if it is complete in the following sense: If  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence in  $X$  satisfying

$$(10.1) \quad \lim_{K \rightarrow \infty} \left( \sup_{j, k \in \mathbb{N}} \|x_j - x_k\|_X \right) = 0,$$

then  $\{x_j\}_{j \in \mathbb{N}}$  converges to  $x$ .

In Chapter 3 we have struggled to prove that  $L^p(\mu)$  with  $1 \leq p \leq \infty$  is a Banach space. For example  $\ell^\infty(\mathbb{N})$  is a Banach space. However, without learning the Lebesgue integral, we could not go beyond such a trivial example.

Banach Steinhaus principle. The Banach Steinhaus principle is a tool to reduce the matter to a dense subspace. Although the proof is simple, we frequently use this theorem for later consideration.

**Theorem 10.1** (Banach Steinhaus principle). *Suppose that we are given a family of operator  $\{T_t\}_{t>0}$  from a Banach space  $X$  to another Banach space  $Y$ . Assume that  $A := \sup_{t>0} \|T_t\|_{X \rightarrow Y} < \infty$  and that there exists a dense set  $D \subset X$  so that  $\lim_{t \downarrow 0} T_t x$  exists. Then the limit  $\lim_{t \downarrow 0} T_t x$  exists for all  $x \in X$ .*

*Proof.* It suffices to prove

$$(10.2) \quad \lim_{s, t \downarrow 0} \|T_t x - T_s x\|_Y \rightarrow 0$$

in view of the compactness of  $Y$ . Note that

$$\begin{aligned} \|T_t x - T_s x\|_Y &\leq \|T_t x - T_t x'\|_Y + \|T_t x' - T_s x'\|_Y + \|T_s x' - T_s x\|_Y \\ &\leq 2A \|x - x'\|_X + \|T_t x' - T_s x'\|_Y \end{aligned}$$

for all  $x' \in D$ . Thus, letting  $s, t \downarrow 0$ , we obtain

$$(10.3) \quad \limsup_{s, t \downarrow 0} \|T_t x - T_s x\|_Y \leq 2A \|x - x'\|_X.$$

Since  $x' \in D$  is arbitrary, and  $D$  is dense in  $X$ , we have

$$(10.4) \quad \limsup_{s, t \downarrow 0} \|T_t x - T_s x\|_Y = 0.$$

This is the desired result.  $\square$

Criterion of completeness. The following theorem is a nice criterion of the completeness. In this book we encounter many Banach spaces. If we prove the completeness directly from the definition, that is, if we intend to start with a Cauchy sequence and try to find the limit, then we always have to go through the tedious argument. This theorem will save us such a tedious discussion.

**Theorem 10.2.** *Let  $X$  be a normed space. Then  $X$  is complete if and only if  $\sum_{j=1}^{\infty} x_j$  is convergent whenever  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence in  $X$  satisfying*

$$(10.5) \quad \sum_{j=1}^{\infty} \|x_j\|_X < \infty.$$

*Proof.* Recall the proof of the fact that any absolutely convergent series in  $\mathbb{R}$  converges. If  $X$  is a Banach space, then the same argument as that for  $\mathbb{R}$  works. The proof is omitted.

To prove the converse we assume that  $\{x_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence. For all  $N \in \mathbb{N}$ , we can choose  $L_N$  so that

$$(10.6) \quad \|x_j - x_k\|_X \leq 2^{-N}$$

for all  $j, k \in \mathbb{N}$  with  $j, k \geq L_N$ . By replacing  $L_N$  with  $L_1 + L_2 + \dots + L_N$ , we can assume that  $L_1 < L_2 < \dots < L_N < \dots$

Consider a series

$$(10.7) \quad x_{L_1} + \sum_{N=1}^{\infty} (x_{L_{N+1}} - x_{L_N}).$$

By assumption we have posed, the series converges to some  $x \in X$ . This is equivalent to saying that  $\lim_{N \rightarrow \infty} x_{L_N} = x$ .

From (10.6) we deduce

$$(10.8) \quad \|x_j - x_{L_M}\|_X \leq 2^{-N}$$

for all  $M \geq N$  and  $j \geq L_N$ . Letting  $M \rightarrow \infty$ , we obtain

$$(10.9) \quad \|x_j - x\|_X \leq 2^{-N}$$

for all  $j \geq L_N$ . This implies  $\lim_{j \rightarrow \infty} x_j$  converges to  $x$ . Thus, we have established that any Cauchy sequence converges.  $\square$

*Exercise 99.* If  $\{x_j\}_{j=1}^{\infty}$  is a sequence in a Banach space  $X$  such that  $\sum_{j=1}^{\infty} \|x_j\|_X < \infty$ , then show that  $\sum_{j=1}^{\infty} x_j$  converges.

Dual spaces. As a special case of the linear mapping let us consider the case that the range is the underlying field  $\mathbb{K}$ .

**Definition 10.3.** Let  $X$  be a normed space. Then the dual space  $X^*$  is a set defined as

$$(10.10) \quad X^* := \{x^* : X \rightarrow \mathbb{K} : x^* \text{ is linear and continuous} \}.$$

That is, define  $X^* := B(X, \mathbb{K})$ .

*Exercise 100.* Let  $X$  be a normed space. Then show that  $X^*$  is a Banach space.

Reflexivity. Now we consider the dual of the dual. Reflexivity means that the dual of the dual coincide the original Banach space. This is the case when the dimension is finite. However, if the dimension is infinite, it can happen that the dual of the dual becomes strictly larger than the original Banach space.

**Definition 10.4.** Define a natural mapping  $Q : X \rightarrow X^{**}$  by  $Qx(x^*) = x^*(x)$ . The space  $X$  is said to be reflexive, if  $Q$  is an isomorphism from  $X$  to  $X^{**}$ .

**Remark 10.5.** It is important that  $Q$  is an isomorphism in the definition of reflexivity. James constructed a nonreflexive Banach space which is isomorphic to its bidual.

*Exercise 101.*

(1) Let  $Z \subset \ell^\infty(\mathbb{N})$  be a subspace given by

$$(10.11) \quad Z := \left\{ a \in \ell^\infty(\mathbb{N}) : \lim_{j \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_j}{j} \text{ exists} \right\}.$$

Let us define

$$(10.12) \quad \ell(a) := \lim_{j \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_j}{j}$$

for  $a \in Z$  and

$$(10.13) \quad q(a) := \limsup_{j \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_j}{j}, \quad r(a) := \liminf_{j \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_j}{j}$$

for  $a \in \ell^\infty$ . Show that  $Z$  is closed. Hint: Begin with proving that  $q$  and  $r$  are continuous. Use the Hahn-Banach theorem to construct a linear mapping to which  $\ell$  extends and which is dominated by  $q$ .

(2) Prove that  $\ell^1(\mathbb{N})$  and  $\ell^\infty(\mathbb{N})$  are not reflexive.

(3) Prove that a Banach space  $X$  is reflexive precisely when  $X^*$  is reflexive.

*Exercise 102.* Suppose that  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space. Use Theorem 4.39 to prove that  $L^p(\mu)$  is reflexive for all  $1 < p < \infty$ .

Weak-\* topology. We now return to the setting that we are given a Banach space  $X$ . The space  $X$  carries a topology induced by the norm. However,  $X$  is never locally compact with respect to this topology, unless it is finite dimensional. Indeed, we have shown that the unit ball of  $X$  is compact if and only if the dimension of  $X$  is finite. In this paragraph we consider the topology of  $X^*$ . Of course the unit ball of  $X^*$  is never compact as we have seen. However, there are several topologies of use. In this paragraph we investigate the weak-\* topology, under which the unit ball of  $X^*$  turns out compact.

**Definition 10.6.** The weak-\* topology of  $X^*$  is the topology generated by the following family.

$$(10.14) \quad \mathcal{U}_{x, x^*, \varepsilon} := \{y^* \in X^* : |x^*(x) - y^*(x)| < \varepsilon\},$$

where  $x, x^*$  and  $\varepsilon$  run through  $X, X^*$  and  $(0, \infty)$  respectively.

*Exercise 103.* Show that  $\{x_j^*\}_{j \in \mathbb{N}}$  is a sequence in  $X^*$ . Then it is convergent to  $x^* \in X^*$  in weak-\* topology, if and only if  $\lim_{j \rightarrow \infty} x_j^*(x) = x^*(x)$  for all  $x \in X$ .

Quasi-Banach spaces. We go into details of quasi-Banach spaces in order that we deal with  $L^p(\mu)$  for  $0 < p < 1$ . Here we place ourselves in the setting of a measure space  $(X, \mathcal{B}, \mu)$ .

**Definition 10.7.**

(1) A linear space  $X$  is said to be a quasi-normed space, if it comes with a function  $\|\cdot\|_X : X \rightarrow [0, \infty)$  with the following property.  $\|\cdot\|_X : X \rightarrow [0, \infty)$  is said to be quasi-norm of  $X$ .



- (a) Let  $x \in X$ . Then  $\|x\|_X = 0$  implies  $x = 0$ .
- (b)  $\|a \cdot x\|_X = |a| \cdot \|x\|_X$  for all  $a \in \mathbb{K}$  and  $x \in X$ .
- (c) There exists a constant  $c \geq 1$  such that  $\|x + y\|_X \lesssim \|x\|_X + \|y\|_X$ .
- (2) Equip  $X$  with the weakest topology such that the quasi-norm is continuous.
- (3) If  $X$  is complete in the following sense, then  $X$  is said to be a quasi-Banach space: If  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence in  $X$  satisfying

$$(10.15) \quad \lim_{K \rightarrow \infty} \left( \sup_{j, k \geq K} \|x_j - x_k\|_X \right) = 0,$$

then  $\{x_j\}_{j \in \mathbb{N}}$  is convergent.

Banach spaces are quasi-Banach spaces by definition. A typical example of the quasi-Banach space is, as we have seen,  $L^p(\mu)$  with  $0 < p < 1$ . The boundedness that we defined for Banach spaces can be extended readily.

*Exercise 104.* As for  $L^p(\mu)$  with  $0 < p \leq 1$ , show that we can take  $c = 2^{\frac{1}{p}-1}$  in (10.15).

**Definition 10.8.** Let  $X$  and  $Y$  be quasi-Banach spaces. A linear mapping from  $A : X \rightarrow Y$  is said to be bounded, if

$$(10.16) \quad \|Ax\|_Y \lesssim \|x\|_X$$

for all  $x \in X$ .

The same can be said for quasi-Banach spaces about the boundedness of linear operators.

**Theorem 10.9.** Let  $X$  and  $Y$  be quasi-Banach spaces. A linear mapping from  $A : X \rightarrow Y$  is continuous if and only if  $A$  is bounded.

*Proof.* We have only to re-examine the proof of Theorem 9.3. □

Now we shall show that any quasi-Banach space can be equipped with a distance compatible to the original topology.

**Theorem 10.10** (Aoki, Rolewicz). Let  $X$  be a quasi-normed spaces. Let  $\beta$  be a constant satisfying

$$(10.17) \quad \|x + y\|_X \leq \beta(\|x\|_X + \|y\|_X)$$

for all  $x, y \in X$ . Let  $\rho = \frac{\log 2}{\log 2\beta}$ , that is, define  $\rho \in (0, 1]$  by  $(2\beta)^\rho := 2$ . Then there exists a metric function  $d$  so that

$$(10.18) \quad \|x\|_X^\rho \leq d(x) \leq 2\|x\|_X^\rho.$$

*Proof.* We define

$$(10.19) \quad d(x) := \frac{1}{2} \inf \left\{ \sum_{j \in J} \|x_j\|_X^\rho : \{x_j\}_{j \in J} \subset X \text{ is finite and } x = \sum_{j \in J} x_j \right\}.$$

Then it is easy to see that  $d$  is metric and  $2\|x\|_X^\rho \geq d(x)$ . Before we prove the reverse inequality

$$(10.20) \quad \|x\|_X^\rho \leq d(x),$$

we claim the following:

**Claim 10.11.** *We have*

$$(10.21) \quad \left\| \sum_{j=1}^J x_j \right\|_X^\rho \leq \max_{1 \leq j \leq J} 2^{\nu_j} \|x_j\|_X^\rho$$

for any finite sequence  $x_1, x_2, \dots, x_J \in X$  and any multiindex  $\nu = (\nu_1, \nu_2, \dots, \nu_J) \in \mathbb{N}_0^J$  satisfying

$$(10.22) \quad \sum_{j=1}^J 2^{-\nu_j} \leq 1.$$

We prove this claim by induction on the length of  $\nu$ . Suppose that  $|\nu| = 1$ . Then the claim is true, since  $\nu_1 \geq 0$ .

Suppose that the claim is true for  $J_0 - 1$  with  $J_0 \geq 2$ . Suppose that we are given a finite sequence  $x_1, x_2, \dots, x_J \in X$  and a multiindex  $\nu = (\nu_1, \nu_2, \dots, \nu_J) \in \mathbb{N}_0^J$  satisfying

$$(10.23) \quad \sum_{j=1}^J 2^{-\nu_j} \leq 1.$$

By rearranging it in numerical order, it can be assumed that

$$(10.24) \quad \nu_1 \leq \nu_2 \leq \dots \leq \nu_J.$$

Then the key observation we have to make is that we can partition  $\{1, 2, \dots, J\}$  into a disjoint union of  $J_1$  and  $J_2$  so that

$$(10.25) \quad \sum_{j \in J_1} 2^{-\nu_j+1} \leq 1, \quad \sum_{j \in J_2} 2^{-\nu_j+1} \leq 1.$$

Indeed, since the matter is dyadical, we have only to set  $J_1 = \{1, 2, \dots, j_1\}$ , where  $j_1$  is the largest number  $\sum_{j=1}^{j_1} 2^{-\nu_j} \leq \frac{1}{2}$ .

Then by induction assumption, we have

$$(10.26) \quad \left\| \sum_{j \in J_1} x_j \right\|_X^\rho \leq \max_{j \in J_1} 2^{\nu_j-1} \|x_j\|_X^\rho, \quad \left\| \sum_{j \in J_2} x_j \right\|_X^\rho \leq \max_{j \in J_2} 2^{\nu_j-1} \|x_j\|_X^\rho.$$

Consequently we have

$$\begin{aligned} \left\| \sum_{j=1}^{J_0} x_j \right\|_X^\rho &\leq \beta^\rho \left( \left\| \sum_{j \in J_1} x_j \right\|_X + \left\| \sum_{j \in J_2} x_j \right\|_X \right)^\rho \\ &\leq (2\beta)^\rho \max \left( \left\| \sum_{j \in J_1} x_j \right\|_X^\rho, \left\| \sum_{j \in J_2} x_j \right\|_X^\rho \right) \\ &= 2 \max \left( \left\| \sum_{j \in J_1} x_j \right\|_X^\rho, \left\| \sum_{j \in J_2} x_j \right\|_X^\rho \right) \\ &\leq \max_{j=1,2,\dots,J_0} 2^{\nu_j} \|x_j\|_X^\rho. \end{aligned}$$

Thus, the claim is established.

Suppose again that we are given a finite number of collections  $x_1, x_2, \dots, x_k \in X$ . Then we set

$$(10.27) \quad M := \sum_{j=1}^k \|x_j\|_X^\rho.$$

Then we define  $\nu_j \in \mathbb{N}_0$  uniquely by

$$(10.28) \quad 2^{-\nu_j} < \frac{\|x_j\|_X^\rho}{M} \leq 2^{-\nu_j+1}$$

Then since  $\sum_{j=1}^k 2^{-\nu_j} \leq 1$ , we are in the position of applying the claim:

$$(10.29) \quad \left\| \sum_{j=1}^k x_j \right\|_X^\rho \leq \max_{1 \leq j \leq k} 2^{\nu_j} \|x_j\|_X^\rho \leq 2M = 2 \sum_{j=1}^k \|x_j\|_X^\rho.$$

Here for the second inequality we used (10.28). In view of (10.29), we can prove  $\|x\|_X^\rho \leq d(x)$  easily.  $\square$

*Exercise 105.* When  $X = L^p(\mu)$ , where  $(X, \mu)$  is a  $\sigma$ -finite space, what can we say about the constant  $\rho$  in Theorem 10.10?

A similar assertion to Theorem 10.2 still holds for quasi-Banach spaces.

**Theorem 10.12.** *Let  $X$  be a quasi-normed space. Suppose that  $\rho > 0$  and a metric function  $d$  satisfies*

$$(10.30) \quad \|x\|_X^\rho \leq d(x) \leq 2\|x\|_X^\rho.$$

*Then  $X$  is complete if and only if  $\sum_{j=1}^\infty x_j$  is convergent whenever  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence in  $X$  satisfying*

$$(10.31) \quad \sum_{j=1}^\infty \|x_j\|_X^\rho < \infty.$$

*Exercise 106.* Prove Theorem 10.12. Hint: We have only to mimic the proposition corresponding to  $L^p(\mu)$  with  $0 < p \leq 1$ .

**Theorem 10.13.** *Let  $0 < p \leq 1$ . Abbreviate  $\ell^q(\mathbb{N})$  to  $\ell^q$  for  $0 < q \leq \infty$ .*

- (1) *Let  $a \in \ell^\infty$ . Then  $b \in \ell^p \mapsto \sum_{j=1}^\infty a_j \cdot b_j$  defines a bounded linear functional on  $\ell^p$ .*
- (2) *Conversely any continuous linear functional on  $\ell^p$  can be realized with some  $a \in \ell^\infty$ .*

*Proof.* (1) is straightforward because  $\ell^p$  is continuously embedded into  $\ell^1$ . Let us prove (2). Pick a continuous linear functional  $\Phi$  on  $\ell^p$ . Let  $e_j = \{\delta_{jk}\}_{k \in \mathbb{N}}$  be the  $j$ -th elementary vector. We define  $a_j := \Phi(e_j)$  for each  $j \in \mathbb{N}$ . Then the boundedness of  $\Phi$  gives us that  $a = \{a_j\}_{j \in \mathbb{N}} \in \ell^\infty$ .

Given  $b = \{b_j\}_{j \in \mathbb{N}} \in \ell^p$ , we note that  $b = \sum_{j=1}^\infty b_j e_j$  converges in  $\ell^p$  by virtue of the monotone convergence theorem. Therefore,

$$(10.32) \quad \Phi(b) = \sum_{j=1}^\infty b_j \Phi(e_j) = \sum_{j=1}^\infty a_j b_j.$$

Thus, it follows that  $a$  realizes  $\Phi$ .  $\square$

**Definition 10.14.** Let  $X$  be quasi-Banach space. The space  $B(X)$  denotes the set of all bounded linear operators on  $X$ .

## 10.2. Baire category theorem and its applications.

Baire's category theorem. Here we list three theorems derived from the Baire category theorem. Examples of the notions and the theorems here are taken up later in this book. Baire's category theorem is one of the most important theorems in general topology.

**Theorem 10.15** (Baire's category theorem). *Let  $X$  be a complete metric space. Suppose that  $F_j, j = 1, 2, \dots$  are closed sets with  $\bigcup_{j=1}^{\infty} F_j = X$ . Then some  $F_j$  contains an interior point.*

*Proof.* We may assume  $\{F_j\}_{j=1}^{\infty}$  is increasing by replacing  $F_j$  with  $F_1 \cup F_2 \cup \dots \cup F_j$  if necessary. Assume that the interior of  $F_j$  is empty for all  $j \in \mathbb{N}$ . Let  $y_0 \in X$  and  $r_0 = 1$ . We define a sequence  $j_k$  and a sequence of balls as follows. Let

$$(10.33) \quad j_0 := \min \{j \in \mathbb{N} : F_j \cap B(y_0, r_0) \neq \emptyset\} < \infty.$$

We remark that the set appearing in the definition of  $j_0$  is not empty because we are assuming that  $\bigcup_{j=1}^{\infty} F_j = X$ . By assumption  $F_{j_0}$  contains no interior point, which allows us to select a ball  $B_1$  so that

$$(10.34) \quad \overline{B_1} \subset \frac{1}{2}B(y_0, r_0) \cap F_{j_0}^c.$$

Suppose that  $j_0, j_1, \dots, j_k$  and  $B_1, \dots, B_{k+1}$  is defined. Then we define

$$(10.35) \quad j_{k+1} := \min \{j \in \mathbb{N} : F_j \cap B_{k+1} \neq \emptyset\}.$$

We select  $B_{k+2}$  so that its closure is engulfed by an open set  $\frac{1}{2}B_{k+1} \cap F_{j_{k+1}}^c$ . This is possible because it cannot happen by assumption that  $\frac{1}{2}B_{k+1} \subset F_{j_{k+1}}$  and hence  $\frac{1}{2}B_{k+1} \cap F_{j_{k+1}}^c$  is not an empty set. This procedure does not terminate by assumption. Therefore,  $\bigcap_{k=1}^{\infty} \overline{B_k} \neq \emptyset$  is not

empty because  $X$  is complete. However,  $\bigcap_{k=1}^{\infty} \overline{B_k}$  never meets any  $F_j, j = 1, 2, \dots$ , since  $F_{j_k}$  does not intersect  $\overline{B_{k+1}}$ . This contradicts to the assumption  $\bigcup_{j=1}^{\infty} F_j = X$ .  $\square$

Open mapping theorem. Let us start with the definition. Let  $X$  and  $Y$  be Banach spaces, although the triangle inequality is not necessary just for the definition of the openness of the mapping.

**Definition 10.16** (Open mapping). A mapping from  $X$  to  $Y$  is said to be open, if it maps open sets in  $X$  to open sets in  $Y$ .

If  $A$  is an open and linear mapping from  $X$  to  $Y$ , then it is clear that  $A$  is surjective. Indeed the image of  $A$  contains the unit ball of  $Y$ . However, the open mapping theorem asserts that the converse is the case, which is far from trivial.

**Theorem 10.17** (Open mapping theorem). *Let  $A : X \rightarrow Y$  be a bounded surjective linear mapping. Then  $A$  is open.*

*Proof.* Let  $A : X \rightarrow Y$  be a bounded surjective from  $X$  to  $Y$ .

Since  $\overline{AB_X(k)}$  is closed subset in  $Y$  for  $k \in \mathbb{N}$  and  $\bigcup_{k \in \mathbb{N}} \overline{AB_X(k)} = X$ . By Baire's category theorem,  $\overline{AB_X(k)}$  has an interior point for some  $k$ . Multiplying  $k^{-1}$  to the above set if necessary, we conclude that the same holds for  $k = 1$ . Since  $A\{x \in X : \|x\|_X \leq 1\}$  is symmetric with respect to 0, there is  $\eta > 0$  such that  $B_Y(\eta) \subset \overline{AB_X(1)}$ . Let  $y \in Y$  with  $\|y\|_Y < \eta$ . Then there exists  $x_1 \in X$  with  $\|x_1\|_X \leq 1$  with  $\|y - Tx_1\| < \frac{\eta}{2}$ . Inductively we can take  $x_1, x_2, \dots$  such that

$$(10.36) \quad \|y - Tx_1 - Tx_2 - \dots - Tx_j\|_Y < 2^{-j}\eta, \quad x_j \in B_X\left(\frac{1}{2^j}\right)$$

for every  $j \in \mathbb{N}$ . Since  $x := \sum_{j=1}^{\infty} x_j$  is convergent, we have  $y = Tx$  with  $\|x\|_X \leq 2$ . Thus  $B_Y(\eta) \subset AB_X(2)$ , which implies  $A$  is open.  $\square$

*Exercise 107.* Given a set  $\Delta$ ,  $\ell^1(\Delta)$  denotes as usual the Banach space of complex valued functions  $\lambda$  on  $\Delta$  for which

$$\|\lambda\|_1 = \sum_{\delta \in \Delta} |\lambda(\delta)| < \infty.$$

Let  $X$  be a Banach space, let  $M_1, M_2$  be positive constants and let  $u$  be a mapping from a set  $\Delta$  to  $X$  such that

- (a)  $\|u(\delta)\| \leq M_1$  for all  $\delta \in \Delta$ ,
- (b)  $\sup\{|\psi^*(u(\delta))| : \delta \in \Delta\} \geq M_2\|\psi^*\|$  for all  $\psi^* \in X^*$ .

Then prove the following:

- (1) Show that the operator

$$T\lambda = \sum_{\delta \in \Delta} \lambda(\delta)u(\delta)$$

is bounded from  $\ell^1(\Delta)$  to  $X$ .

- (2) Show that

$$M_2\|\psi^*\| \leq \|T^*\psi^*\|$$

for all  $\psi^* \in X^*$ . Deduce from this that  $T^*$  is closed and injective.

- (3) By using the open mapping theorem, show that  $T^*$  and  $T$  are surjective.
- (4) Show that any  $f \in X$  is of the form

$$f = \sum_{\delta \in \Delta} \lambda(\delta)u(\delta)$$

for some  $\lambda \in \ell^1(\Delta)$  and that, if necessary, by replacing  $\lambda$  suitably, for any  $\varepsilon$ , we can arrange

$$\|\lambda\|_1 \leq \frac{1}{M_2}\|f\| + \varepsilon.$$

**Closed graph theorem.** Closed graph theorem is one of the basic tools in functional analysis, which is a by-product of the Baire category theorem.

**Theorem 10.18** (Closed graph theorem). *A closed operator from  $X$  to  $Y$  with domain  $X$  is bounded.*

*Proof.* Denote by  $Z$  a Banach space  $X$  endowed with the graph norm. Then  $Z \subset X$  in the sense of continuous embedding and  $X = Z$  as a set. Thus we are in the position of applying the open mapping theorem to the inclusion mapping  $\iota : Y \rightarrow X$ . This theorem tells us that  $\iota$  is an isomorphism. Therefore the norms of  $X$  and  $Y$  are mutually equivalent and in particular  $\|Ax\|_X + \|x\|_X \lesssim \|x\|_X$ . This shows that  $A$  is continuous.  $\square$

Uniformly bounded principle. Finally we take up the uniformly bounded principle. Once we obtain the open mapping theorem and the closed graph theorem, the proof is quite easy.

**Theorem 10.19** (Uniformly bounded principle). *Suppose that  $\{A_\lambda\}_{\lambda \in \Lambda}$  is a family of bounded linear mappings. Assume  $\sup_{\lambda \in \Lambda} \|A_\lambda x\|_Y < \infty$ . Then we have  $\sup_{\lambda \in \Lambda} \|A_\lambda\|_{X \rightarrow Y} < \infty$ .*

*Proof.* Define a closed subset  $X_j$  of  $X$  by  $X_j := \left\{ x \in X : \sup_{\lambda \in \Lambda} \|A_\lambda x\|_Y \leq j \right\}$  for  $j \in \mathbb{N}$ . Then  $X = \bigcup_{j=1}^{\infty} X_j$ . Then the Baire category tells us that  $X_j$  contains a open set, if  $X_j$  is large. Since  $X_j$  is symmetric,  $X_j$  contains 0 as an interior point. Therefore there exists  $\eta > 0$  such that  $\|x\|_X < \eta$  implies  $\|A_\lambda\|_Y \leq j$ . As a result we obtain  $\sup_{\lambda \in \Lambda} \|A_\lambda\|_{X \rightarrow Y} \leq j\eta^{-1}$ .  $\square$

## 11. HILBERT SPACES

Banach spaces can be considered as a natural extension of  $\mathbb{R}^d$ . However, they lose a rich structure: inner product of  $\mathbb{R}^d$  is used to measure angles of vectors. Hilbert spaces are Banach spaces equipped with inner product compatible with the norms. Although the unit ball is not compact even in Hilbert spaces, Hilbert spaces occur every field of mathematics.

### Example 11.1.

- (1) When we prove the Riemann-Roch theorem in Riemannian surfaces, a theorem in algebraic geometry, we need function spaces called  $A^2(\Omega)$ .
- (2) When we prove the Hodge-decomposition theorem in differential geometry, Sobolev spaces  $W^{2,k}(M)$  play an important role.
- (3) In stochastic analysis, to consider stochastic integrals, we need to consider  $L^2(P)$  spaces in connection with Brownian motions.

Here we concentrate on Hilbert spaces and consider their properties.

**11.1. Definitions and elementary properties.** Again we make the definition precise.

**Definition 11.2.** A Banach space  $(H, \|\cdot\|_H)$  is said to be a Hilbert space, if it comes with a mapping  $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{K}$ , with the following properties for all  $x, y, z \in H$  and  $a \in \mathbb{K}$ . And  $\langle \cdot, \cdot \rangle$  is said to be an inner product.

- (1)  $\langle x, x \rangle = \|x\|_H^2$  for all  $x \in H$ ,
- (2)  $\langle x, y \rangle = \overline{\langle y, x \rangle}$  for all  $x, y \in H$ .
- (3)  $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$  for all  $x, y, z \in H$ .
- (4)  $\langle a \cdot x, y \rangle = a \cdot \langle x, y \rangle$ .

**Theorem 11.3.** *Let  $H$  be a Hilbert space and  $x, y \in H$ . Then we have*

$$(11.1) \quad |\langle x, y \rangle| \leq \|x\|_H \cdot \|y\|_H$$

and

$$(11.2) \quad \|x + y\|_H^2 + \|x - y\|_H^2 = 2 (\|x\|_H^2 + \|y\|_H^2).$$

*Exercise 108.* Prove (11.1) and (11.2). Hint: To prove (11.1), we may assume  $\langle x, y \rangle \in \mathbb{R}$  by multiplying  $e^{i\theta}$ ,  $\theta \in \mathbb{R}$  if necessary. Consider the determinant of the function

$$(11.3) \quad t \in \mathbb{R} \mapsto \langle tx - y \rangle^2 \in [0, \infty).$$

**Theorem 11.4.** Assume that  $H$  is a Banach space whose norm obeys the parallelogram law:

$$(11.4) \quad \|x + y\|_H^2 + \|x - y\|_H^2 = 2 (\|x\|_H^2 + \|y\|_H^2)$$

for all  $x, y \in H$ . Then  $H$  carries the structure of a Hilbert space whose inner product is given as follows:

$$\text{If } \mathbb{K} = \mathbb{R}, \langle x, y \rangle = \frac{1}{2} (\|x + y\|_H^2 - \|x - y\|_H^2).$$

$$\text{If } \mathbb{K} = \mathbb{C}, \langle x, y \rangle = \frac{1}{4} (\|x + y\|_H^2 + i \|x + iy\|_H^2 - \|x - y\|_H^2 - i \|x - iy\|_H^2).$$

*Proof.* We concentrate on the case when  $\mathbb{K} = \mathbb{C}$ .

Additivity of  $\langle \cdot, \cdot \rangle$ .

 We shall prove

$$(11.5) \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

Since  $i \cdot \operatorname{Re}(\langle x, iy \rangle) = \operatorname{Im}(\langle x, y \rangle)$ , we have only to prove

$$(11.6) \quad \operatorname{Re}(\langle x + y, z \rangle) = \operatorname{Re}(\langle x, z \rangle) + \operatorname{Re}(\langle y, z \rangle).$$

By using the assumption we obtain

$$\begin{aligned} & 8 \operatorname{Re}(\langle x + y, z \rangle) - 8 \operatorname{Re}(\langle x, z \rangle) - 8 \operatorname{Re}(\langle y, z \rangle) \\ &= 2\|x + y + z\|_H^2 + 2\|x - z\|_H^2 + 2\|y - z\|_H^2 \\ &\quad - 2\|x + z\|_H^2 - 2\|y + z\|_H^2 - 2\|x + y - z\|_H^2 \\ &= 2\|x + y + z\|_H^2 + \|x + y - 2z\|_H^2 - \|x + y + 2z\|_H^2 - 2\|x + y - z\|_H^2. \end{aligned}$$

Thus, the function  $F(x, y, z) = 8 \operatorname{Re}(\langle x + y, z \rangle) - 8 \operatorname{Re}(\langle x, z \rangle) - 8 \operatorname{Re}(\langle y, z \rangle)$  depends only on  $x + y$  and  $z$ . Therefore, we obtain

$$(11.7) \quad F(x, y, z) = F(x + y, 0, z).$$

Since  $F(x + y, 0, z) = 0$ , we obtain  $F(x, y, z) = 0$ .

Remaining property of the inner product

 It is easy to show

$$(11.8) \quad \langle x, x \rangle = \|x\|_H^2, \langle i \cdot x, y \rangle = i \cdot \langle x, y \rangle, \langle x, y \rangle = \overline{\langle y, x \rangle}$$

for all  $x, y \in H$ . Therefore, it remains to show

$$(11.9) \quad \langle a \cdot x, y \rangle = a \langle x, y \rangle$$

for all  $x, y \in H$  with  $a \in \mathbb{R}$ . If  $a \in \mathbb{N}$ , this is just a consequence of additivity. If  $a = -1$ , this can be derived from the definition of  $\langle \cdot, \cdot \rangle$ . Suppose that  $a \in \mathbb{Q}$ . Let  $m$  be the denominator of  $a$ . Then since  $m \cdot a \in \mathbb{Z}$ , we have

$$(11.10) \quad m \langle a \cdot x, y \rangle = \langle ma \cdot x, y \rangle = m \cdot a \langle x, y \rangle.$$

Thus, (11.9) is valid for  $a \in \mathbb{Q}$ . It remains to pass to the case when  $a \in \mathbb{R}$  in general. To do this, we note that the inner product is continuous. By using the continuity we obtain

$$(11.11) \quad \langle a \cdot x, y \rangle = \lim_{j \rightarrow \infty} \left\langle \frac{[2^j a]x}{2^j}, y \right\rangle = \lim_{j \rightarrow \infty} \frac{[2^j a]}{2^j} \langle x, y \rangle = a \langle x, y \rangle.$$

Therefore, the remaining property that  $\langle \cdot, \cdot \rangle$  should satisfy is proved. □

*Exercise 109.* Let  $H$  be a Hilbert space over  $\mathbb{C}$ . The aim of this exercise is to generalize (11.2). Suppose that  $k$  is an integer larger than 2 and set  $\omega := \exp\left(\frac{2\pi i}{k}\right)$ . Then show that

$$(11.12) \quad \langle x, y \rangle_H = \frac{1}{k} \sum_{j=0}^{k-1} \omega^j \|x + \omega^j y\|_H^2.$$

*Exercise 110.* Show that  $L^1(0, \infty)$  does not carry the structure of a Hilbert space compatible with its original norm.

**Definition 11.5.** Let  $H$  be a Hilbert space and  $H_0$  and  $H_1$  its closed subspaces. Then  $H_0$  is perpendicular to  $H_1$  if  $\langle x, y \rangle = 0$  for all  $x \in H_0$  and  $y \in H_1$ . If in addition  $H = H_0 + H_1$  as a linear space, then denote  $H = H_0 \oplus H_1$ .

Until the end of this paragraph we assume that  $H$  is a Hilbert space.

Suppose that the subspaces  $H_0$  and  $H_1$  are perpendicular. Then  $H_0 \cap H_1 = \{0\}$ . Indeed, if  $x \in H_0 \cap H_1$ , then we have

$$(11.13) \quad \langle x, x \rangle = 0.$$

Thus,  $x = 0$ . From this we conclude that  $x \in H$  admits a unique decomposition  $x = x_1 + x_2$ , if  $H = H_0 \oplus H_1$ .

**Theorem 11.6.** Let  $H_0$  be a closed subspace of a Hilbert space  $H$ . Then there exists a unique closed subspace  $H_0^\perp$  such that

$$(11.14) \quad H = H_0 \oplus H_0^\perp.$$

The function space  $H_0^\perp$  is said to be the orthogonal complement of  $H_0$ .

*Proof.* Existence of  $H_0^\perp$  We set  $H_0^\perp := \{x \in H : \langle x, y \rangle = 0 \text{ for all } y \in H_0\}$ . We have to show  $H = H_0 + H_0^\perp$ . Let  $x \in H$ . Then define  $y \in H_0$  so that it minimizes the distance from  $x$  to  $H_0$ . Such a point does exist. Indeed, let  $\{y_j\}_{j \in \mathbb{N}}$  be a sequence in  $H_0$  such that  $\|y_j - x\|_H \downarrow \text{dist}(x, H_0)$ . Note that (11.2) gives

$$(11.15) \quad \|y_j - y_k\|_H^2 = 2\|x - y_j\|_H^2 + 2\|x - y_k\|_H^2 - \|2x - y_j - y_k\|_H^2.$$

Since  $\|2x - y_j - y_k\|_H^2 \geq \text{dist}(x, H_0)^2$ , we have

$$(11.16) \quad \|y_j - y_k\|_H^2 \leq 2\|x - y_j\|_H^2 + 2\|x - y_k\|_H^2 - 4 \text{dist}(x, H_0)^2.$$

Thus,  $\{y_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence. Since  $H_0$  is assumed closed, its limit still belongs to  $H_0$ . Therefore such a minimizer  $y \in H_0$  does exist. We have to show that  $z := x - y$  belongs to  $H_0^\perp$ . To do this let  $w \in H_0$  be chosen arbitrary. Then for every  $t \in \mathbb{R}$ , we have

$$(11.17) \quad \langle z + tw, z + tw \rangle = t^2 \langle w, w \rangle + 2t \text{Re}(\langle z, w \rangle) + \langle z, z \rangle$$

attains its minimum at  $t = 0$ , because  $y$  attains minimum the distance from  $x$  to  $H_0$ . Therefore, we obtain  $\text{Re}(\langle z, w \rangle) = 0$ . Replacing  $t$  with  $it$  and going through the same argument, we see that  $\text{Im}(\langle z, w \rangle) = 0$  as well. Therefore  $\langle z, w \rangle = 0$  for all  $w \in H_0$ . Thus,  $z \in H_0^\perp$  is established.

Uniqueness of  $H_0^\perp$  Assume that  $H = H_0 \oplus K$  for some closed space  $K$ . Let  $x \in K$ . Then for all  $y \in H_0$  we deduce  $\langle x, y \rangle = 0$  from the definition of  $\oplus$ . Thus  $K \subset H_0^\perp$ . Conversely let  $x \in H_0^\perp$ . Decompose  $x$  along the decomposition  $H = H_0 \oplus K$ , that is, split  $x$  by  $x = x_0 + x_1$  with  $x_0 \in H_0$  and  $x_1 \in K$ . Since  $\langle x_0, x_1 \rangle = 0$ , we conclude  $0 = \langle x, x_0 \rangle = \langle x_0, x_0 \rangle + \langle x_1, x_0 \rangle = \langle x_0, x_0 \rangle$ . Thus, it follows that  $x_0 = 0$  and  $x = x_1 \in K$ . As a consequence  $K = H_0^\perp$  is established.  $\square$



One of the reasons why Hilbert spaces have rich structures is that any bounded linear functional can be written by way of the inner product. Surprisingly enough, this can yield again the duality  $L^2(\mu)$  to  $L^2(\mu)$ .

**Theorem 11.7** (Riesz's representation theorem for Hilbert spaces). *Suppose that  $f \in H^*$ . Then there exists a unique  $x \in H$  such that*

$$(11.18) \quad f(y) = \langle y, x \rangle$$

for all  $y \in H$ .

*Proof.* Existence of such  $x \in H$  First of all we may assume  $f$  is nonzero. Otherwise  $0 \in H$  is an element we are looking for. Set  $H_0 := \text{Ker}(f) := f^{-1}(\{0\}) := \{y \in H : f(y) = 0\}$ . Then  $H_0$  is closed, since  $f$  is assumed continuous. Therefore  $H$  admits an orthonormal decomposition  $H = H_0 + H_0^\perp$ . Let  $z \in H_0^\perp \setminus \{0\}$ . If  $f(z)$  were 0, then we would have  $z \in H_0 \cap H_0^\perp = \{0\}$ . This is impossible. Therefore a normalization allows us to assume  $f(z) = 1$ . Then given  $y \in H$ , we have  $y - f(y)z \in \text{Ker}(f)$ . Therefore  $\langle y - f(y)z, z \rangle = 0$ . From this formula, we conclude

$$(11.19) \quad f(y) = \frac{\langle y, z \rangle}{\|z\|_H^2} \text{ for all } y \in H.$$

Thus, we have only to put  $x = \frac{z}{\|z\|_H^2} \in H$ .

Uniqueness of  $x$  Suppose that  $x_1, x_2 \in H$  realizes  $f \in H^*$ . Then

$$(11.20) \quad \langle y, x_1 \rangle = \langle y, x_2 \rangle$$

for all  $y \in H$ . Rearranging the above formula as  $\langle y, x_1 - x_2 \rangle = 0$  for all  $y$  and setting  $y = x_1 - x_2$ , we conclude  $x_1 - x_2 = 0$ . Therefore the element representing  $f$  turned out to be unique.  $\square$

Here is another reason why Hilbert spaces are attractive. The norm has the following expressions.

**Theorem 11.8.** *The operator norm of a linear operator  $A : H \rightarrow H$  is given by the following formula.*

$$(11.21) \quad \|A\|_H = \sup_{x, y \in H \setminus \{0\}} \frac{|\langle Ax, y \rangle|}{\|x\|_H \|y\|_H}.$$

*Proof.* The left-hand side is attained by the right-hand side, if we set  $y = Ax$ .

$$(11.22) \quad \text{R.H.S.} \geq \sup_{x \in H \setminus \{0\}} \frac{|\langle Ax, Ax \rangle|}{\|x\|_H \|Ax\|_H} = \sup_{x \in H \setminus \{0\}} \frac{\|Ax\|_H^2}{\|x\|_H^2} = \|A\|_H = \text{L.H.S.}$$

Meanwhile the right-hand side is less than or equal to the left-hand side by (11.1).

$$(11.23) \quad \text{R.H.S.} \leq \sup_{x \in H \setminus \{0\}} \frac{\|Ax\|_H \cdot \|y\|_H}{\|x\|_H \cdot \|y\|_H} = \sup_{x \in H \setminus \{0\}} \frac{\|Ax\|_H}{\|x\|_H} = \|A\|_H = \text{L.H.S.}$$

Thus, the both sides of (11.21) are identical.  $\square$

**Theorem 11.9** (Adjoint operator). *Let  $A \in B(H)$ . Then there uniquely exists  $B \in B(H)$  such that*

$$(11.24) \quad \langle Ax, y \rangle = \langle x, By \rangle$$

for all  $x, y \in H$ . Furthermore the operator norm of  $B$  equals that of  $A$ .

*Proof.* Existence of such  $B$  Let  $x$  be fixed. Then the functional

$$(11.25) \quad y \in H \mapsto \langle y, Ax \rangle \in \mathbb{C}$$

is bounded by virtue of Theorem 11.8. Therefore by the Riesz theorem there exists a unique  $B_x$  such that

$$(11.26) \quad \langle y, Ax \rangle = \langle y, B_x \rangle.$$

The mapping  $x \mapsto B_x$  is linear because of the uniqueness of the Riesz theorem. Indeed, if  $x \in H$  and  $a \in \mathbb{C}$ , we have

$$(11.27) \quad \langle y, B_{a \cdot x} \rangle = \langle y, A(a \cdot x) \rangle = \bar{a} \langle y, Ax \rangle = \bar{a} \langle y, B_x \rangle = \langle y, a \cdot B_x \rangle.$$

By the uniqueness of the Riesz representation theorem, we see that  $a \cdot B_x = B_{a \cdot x}$ . In the same way we can prove  $B_{x_1+x_2} = B_{x_1} + B_{x_2}$  for all  $x_1, x_2 \in H$ . Thus, the mapping  $x \mapsto B_x$  is linear. The norm of this linear mapping is given by

$$\begin{aligned} \sup_{x, y \in H \setminus \{0\}} \frac{|\langle B_x, y \rangle|}{\|x\|_H \cdot \|y\|_H} &= \sup_{x, y \in H \setminus \{0\}} \frac{|\langle y, B_x \rangle|}{\|x\|_H \cdot \|y\|_H} \\ &= \sup_{x, y \in H \setminus \{0\}} \frac{|\langle y, Ax \rangle|}{\|x\|_H \cdot \|y\|_H} \\ &= \sup_{x, y \in H \setminus \{0\}} \frac{|\langle Ax, y \rangle|}{\|x\|_H \cdot \|y\|_H} \\ &= \|A\|. \end{aligned}$$

Thus,  $B \in B(H)$  is the operator we wish to find.

Uniqueness of  $B$  If  $C \in B(H)$  is another bounded linear operator, then

$$(11.28) \quad \langle (B - C)x, y \rangle = \langle Bx, y \rangle - \langle Cx, y \rangle = \langle x, Ay \rangle - \langle x, Ay \rangle = 0.$$

Thus, the operator norm of  $B - C$  is 0 by Theorem 11.8. From this we conclude  $B = C$ .  $\square$

Theorem 11.9 justifies the following definition.

**Definition 11.10.** Let  $A \in B(H)$ .

- (1) The operator  $B$  obtained in Theorem 11.9 is said to be the adjoint of  $A$ .
- (2)  $A$  is said to be self-adjoint, if  $A = A^*$ . Below denote by  $S(H)$  the set of all self-adjoint operators in  $B(H)$ .

## 11.2. Complete orthonormal system.

In Chapter 4 we have shown that any  $f \in L^2(\mathbb{T})$  can be expanded as

$$(11.29) \quad f(x) = \sum_{j=-\infty}^{\infty} a_j e^{ijx},$$

where  $\{a_j\}_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ . Here the coefficient  $a_j$  is determined uniquely by  $f$ . In this sense  $\{e^{ijx}\}_{j \in \mathbb{Z}}$  plays a role of basis. Unlike algebra, in considering basis we take infinite sum provided the sum converges. Let us show that any separable Hilbert space enjoys the same property.

**Theorem 11.11.** Let  $\mathcal{H}$  be an infinite dimensional separable Hilbert space. Then there exists a countable set  $\{x_j\}_{j \in \mathbb{N}}$  such that any  $x \in \mathcal{H}$  admits an expansion

$$(11.30) \quad x = \sum_{j=1}^{\infty} \langle x, x_j \rangle \cdot x_j.$$

*Proof.* Let  $\{y_k\}_{k \in \mathbb{N}}$  be a countable dense subset. Set  $\{x_j\}_{j \in \mathbb{N}}$  as a sequence obtained by Schmidt's orthonormalization. Let  $x \in \mathcal{H}$  and  $J \in \mathbb{N}$ .

$$\begin{aligned} \left\| x - \sum_{j=1}^J \langle x, x_j \rangle \cdot x_j \right\|_{\mathcal{H}} &\leq \left\| (x - y_k) - \sum_{j=1}^J \langle x - y_k, x_j \rangle \cdot x_j \right\|_{\mathcal{H}} + \left\| y_k - \sum_{j=1}^J \langle y_k, x_j \rangle \cdot x_j \right\|_{\mathcal{H}} \\ &\leq \|x - y_k\|_{\mathcal{H}} + \left\| y_k - \sum_{j=1}^J \langle y_k, x_j \rangle \cdot x_j \right\|_{\mathcal{H}}. \end{aligned}$$

Note that  $y_k \in \text{Span}(x_1, x_2, \dots, x_J)$ , provided  $J$  is large enough. Letting  $J \rightarrow \infty$ , we obtain

$$(11.31) \quad \limsup_{J \rightarrow \infty} \left\| x - \sum_{j=1}^J \langle x, x_j \rangle \cdot x_j \right\|_{\mathcal{H}} \leq \|x - y_k\|_{\mathcal{H}}.$$

Since  $k$  is arbitrary and  $\{y_1, y_2, \dots\}$  is dense in  $\mathcal{H}$ , we see  $x = \sum_{j=1}^{\infty} \langle x, x_j \rangle \cdot x_j$ , which is the desired result.  $\square$

**Corollary 11.12.** *Any separable Hilbert space is isomorphic to  $\ell^2(\mathbb{Z})$ .*

**Definition 11.13.** Let  $\mathcal{H}$  be a separable Hilbert space. Then  $\{x_j\}_{j \in \mathbb{N}}$  is said to be a complete orthonormal system, or for short CONS, if it satisfies the following condition.

- (1)  $\langle x_j, x_k \rangle = \delta_{jk}$  for all  $j, k \in \mathbb{Z}$ .
- (2) Any  $x \in H$  admits the following expansion :

$$(11.32) \quad x = \sum_{j=1}^{\infty} \langle x, x_j \rangle \cdot x_j.$$

If  $\{x_j\}_{j \in \mathbb{N}}$  satisfies (1) only, then it is said to be orthonormal system or for short ONS.

In other words we can say that the completeness is a condition that we cannot add elements orthogonal to each element in the system any more.

**Theorem 11.14** (Parseval's formula). *Let  $\{x_j\}_{j \in \mathbb{N}}$  be an ONS. Then show that it is complete if and only if*

$$(11.33) \quad \|x\|_H^2 = \sum_{j=1}^{\infty} |\langle x, x_j \rangle|^2$$

for all  $x \in H$ .

*Proof.* It is straightforward to check that

$$(11.34) \quad \left\| x - \sum_{j=1}^N \langle x, x_j \rangle x_j \right\|_H^2 = \|x\|_H^2 - \sum_{j=1}^N |\langle x, x_j \rangle|^2.$$

Therefore, the assertion is immediate.  $\square$

Examples of CONS. Below we exhibit examples of CONS.

**Example 11.15** (Fourier series). Let  $t_j = e^{ijt} \in L^2(\mathbb{T}^d)$  for  $j \in \mathbb{Z}^d$ . Then  $\{t_j\}_{j \in \mathbb{Z}^d}$  is a CONS, which we have been struggled to prove in Chapter 4.

In view of this example, the expansion  $t = \sum_{j=1}^{\infty} \langle t, t_j \rangle \cdot t_j$  for a CONS is called generalized Fourier expansion.

*Exercise 111.* By using the Parseval formula, calculate the value of  $\zeta(2)$  and  $\zeta(4)$ , where  $\zeta$  is the zeta function given by

$$(11.35) \quad \zeta(s) = \sum_{j=1}^{\infty} \frac{1}{j^s}, \quad s > 0.$$

**Theorem 11.16** (Legendre polynomial). Define a polynomial  $P_j$  of degree  $j$  by

$$(11.36) \quad P_j(t) := \frac{1}{2^j j!} \frac{d^j}{dt^j} (t^2 - 1)^j \quad (t \in \mathbb{R})$$

for  $j \in \mathbb{N}$ . The family  $\left\{ \sqrt{j + \frac{1}{2}} P_j \right\}_{j \in \mathbb{N}}$  forms a CONS on  $L^2([0, 1])$ .

*Proof of Theorem 11.16.* Let  $j \leq k$ . Then since  $P_j$  is a polynomial of degree  $j$  and  $\frac{d^l}{dt^l} (t^2 - 1)^k$  vanishes at  $t = 0, 1$  whenever  $l < j$ , a repeated integration by parts gives us that

$$\int_{-1}^1 P_j(t) P_k(t) dt = \frac{(-1)^j}{4^j j!^2} \int_{-1}^1 \frac{d^j}{dt^j} P_j(t) \frac{d^{k-j}}{dt^{k-j}} (t^2 - 1)^j dt = \frac{(-1)^j \delta_{jk} (2j)!}{4^j j!^2} \int_{-1}^1 (1 - t^2)^j dt.$$

The integral above can be calculated by means of the Gamma function :

$$(11.37) \quad \int_0^1 (1 - t^2)^j dt = \frac{1}{2} \int_{-1}^1 (1 - t^2)^j dt = \frac{j!^2 2^{2j+1}}{(2j+1)!}.$$

Therefore, we obtain

$$(11.38) \quad \int_{-1}^1 P_j(t) P_k(t) dt = \delta_{jk}.$$

Therefore, we see that  $\{P_j\}_{j \in \mathbb{N}}$  is an ONS.

To see that this system is complete, we recall that  $\{e^{ijt}\}_{j \in \mathbb{N}}$  is dense in  $C([-1, 1])$ . Expand  $e^{ijt}$  into the Taylor series. Then since the series converges absolutely, we conclude that the set of all polynomials form a dense subset.  $\square$

*Exercise 112.* Prove (11.37).

**Theorem 11.17** (Hermite polynomial). Define  $H_j(t) = \exp(t^2) \frac{d^j}{dt^j} \exp(-t^2)$  for  $j \in \mathbb{N}$ . Then the family

$$\mathcal{A} = \left\{ \frac{1}{\sqrt{2^j j! \sqrt{\pi}}} H_j \cdot \exp\left(-\frac{t^2}{2}\right) \right\}_{j \in \mathbb{N}}$$

forms a CONS on  $L^2(\exp(-t^2))$ .

*Proof.* Let  $j$  and  $k$  be integers such that  $j \leq k$ . Then since  $H_j$  is a polynomial of degree  $j$  we have

$$\begin{aligned} \int_{\mathbb{R}} H_j(t)H_k(t) \exp(-t^2) dt &= \int_{\mathbb{R}} H_j(t) \cdot \frac{d^k}{dt^k} \exp(-t^2) dt \\ &= (-1)^j \delta_{jk} \int_{\mathbb{R}} \frac{d^j H_j}{dt^j}(t) \exp(-t^2) dt. \end{aligned}$$

Since the leading term of  $H_j$  is  $(-2)^j t^j$ , it follows that

$$(11.39) \quad \int_{\mathbb{R}} H_j(t)H_k(t) \exp(-t^2) dt = 2^j j! \int_{\mathbb{R}} \exp(-t^2) dt = \delta_{jk} 2^j j! \cdot \pi.$$

Therefore,  $\left\{ \frac{1}{\sqrt{2^j j! \sqrt{\pi}}} H_j \cdot \exp\left(-\frac{t^2}{2}\right) \right\}_{j \in \mathbb{N}}$  is an ONS.

To see that this system is complete, we take  $f \in L^2\left(\exp\left(-\frac{t^2}{2}\right)\right)$  and assume that

$$(11.40) \quad \int_{\mathbb{R}} f(t)H_j(t) \exp(-t^2) dt = 0$$

for all  $j \in \mathbb{N}$ . Then we have

$$(11.41) \quad \int_{\mathbb{R}} f(t)t^j \exp(-t^2) dt = 0$$

for all  $j \in \mathbb{N}$ . Define

$$(11.42) \quad G(\xi) := \int_{\mathbb{R}} f(t) \exp(-t^2) e^{it\xi} dt.$$

Then we have

$$(11.43) \quad \int_{\mathbb{R}} t^{2k} \exp\left(-\frac{t^2}{2}\right) dt = 2^{k-1} \int_{\mathbb{R}} t^{k-\frac{1}{2}} \exp(-t) dt = 2^{k-1} \Gamma\left(k + \frac{1}{2}\right).$$

Therefore,  $G(\xi)$  can be extended to a Taylor series convergent in a neighborhood of 0. Since

$$(11.44) \quad G^{(k)}(0) = i^k \int_{\mathbb{R}} f(t)t^k \exp(-t^2) dt = 0,$$

we see that  $G = 0$ . Therefore, the Fourier transform of  $F(t) := f(t) \exp(-t^2)$  is zero and hence  $f$  itself is zero. Therefore, it follows that  $\left\{ \frac{1}{\sqrt{2^j j! \sqrt{\pi}}} H_j \cdot \exp\left(-\frac{t^2}{2}\right) \right\}_{j \in \mathbb{N}}$  is a CONS.  $\square$

**Theorem 11.18** (Laguerre polynomial). *Let  $-1 < \alpha < \infty$ . Define*

$$(11.45) \quad L_j^\alpha(t) = \frac{\exp(t)t^{-\alpha}}{j!} \frac{d^j}{dt^j}(\exp(-t)t^{j+\alpha}).$$

*Then we have the family*

$$\mathcal{H} = \left\{ \sqrt{\frac{j!}{\Gamma(j + \alpha + 1)}} L_j^\alpha \right\}_{j \in \mathbb{N}}$$

*forms a CONS on the weighted space  $L^2([0, \infty), e^{-t}t^\alpha dt)$ .*

*Proof.* Let  $j \leq k$ . Then it is the same that

$$(11.46) \quad \int_0^\infty L_j^\alpha(t)L_k^\alpha(t) \exp(-t)t^\alpha dt = \delta_{jk} \int_0^\infty \frac{d^j}{dt^j} L_j^\alpha(t) \exp(-t)t^{j+\alpha} dt.$$

Observe that  $\frac{d^j}{dt^j} L_j^\alpha(t) = j! \cdot (j + \alpha)(j + \alpha - 1) \dots (1 + \alpha)$ . Therefore, we see  $\{L_j^\alpha\}_{j \in \mathbb{N}}$  is an ONS. To see that this is a CONS, take  $f \in L^2([0, \infty), e^{-t} t^\alpha dt)$  which is orthogonal to any  $L_j^\alpha$ . Extend  $f$  to  $\mathbb{R}$  by setting 0 on  $(-\infty, 0)$ . We set

$$(11.47) \quad F(s) := \int_0^\infty f(t) e^{-4t} t^\alpha e^{ist} dt = \mathbb{R} f(t) e^{-4t} t^\alpha e^{ist} dt.$$

Then  $F$  is a holomorphic function on  $\{s \in \mathbb{C} : \operatorname{Re}(s) > -1\}$ . Furthermore  $F^{(j)}(0) = 0$ . Hence  $F$  is zero. This means the Fourier transform of  $f(t) e^{-4t} t^\alpha$  is zero. Therefore, we conclude  $f = 0$ .  $\square$

**Theorem 11.19** (Chebychev polynomial). *Let  $j \in \mathbb{N}$ . Then define a polynomial  $T_j$  by*

$$(11.48) \quad T_j(\cos t) = \cos jt.$$

*Then the family  $\{T_j\}_{j \in \mathbb{N}}$  forms a CONS on the weighted space  $L^2\left((-1, 1), \frac{dt}{\sqrt{1-t^2}}\right)$ .*

*Proof.* By change of variables we see at least that  $\{T_j\}_{j \in \mathbb{N}}$  is an ONS. To see that it is complete, it suffices to establish that the system can approximate any function  $f \in L^2\left((-1, 1), \frac{dt}{\sqrt{1-t^2}}\right)$  such that  $\operatorname{supp}(f) \subset (-1 + \varepsilon, 1 - \varepsilon)$  for some  $\varepsilon > 0$ . By virtue of the fact that  $\{e^{ijt}\}_{j \in \mathbb{Z}}$  forms a dense subset in  $C([-\pi, \pi])$ , we can prove that  $\{\cos j\theta\}_{j \in \mathbb{N}}$  forms a dense subset in  $C([0, \pi])$ . Therefore, the system is complete.  $\square$

**11.3. Bounded linear operators defined on Hilbert spaces.** Due to the rich structure of Hilbert spaces, there are many good classes of bounded operators in Hilbert spaces. Operators in such classes can be regarded as atoms in the sense that many things are reduced to such operators.

Unitary operators. As we have seen before, the Fourier transform and its inverse preserve  $L^2(\mathbb{R}^d)$ -norms. Such operators are unitary in the following sense:

**Definition 11.20** (Unitary operators). Let  $H_1$  and  $H_2$  be Hilbert spaces. A unitary operator  $U : H_1 \rightarrow H_2$  is a linear bijection that is norm preserving. Denote by  $U(H)$  the set of all unitary operators in  $H$ .

A typical example of unitary operators is the Fourier transform  $\mathcal{F} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ .

*Exercise 113.* Let  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a distance preserving mapping in the sense that

$$(11.49) \quad \|T(x) - T(y)\|_{\mathbb{R}^d} = \|x - y\|_{\mathbb{R}^d}.$$

Assume in addition that  $T(0) = 0$ .

- (1) Show that  $T$  is an injection.
- (2) Show that  $T$  is linear and hence unitary.

Projection. In linear algebra, projection plays a central role. This can be applied in linear algebra in Hilbert spaces.

**Definition 11.21** (Projection). A self-adjoint bounded linear operator  $P \in B(H)$  is said to be a projection if  $P^2 = P$ . Denote by  $P(H)$  the set of all projections in  $H$ .

*Exercise 114.* Let  $U \in U(H)$  and  $P \in P(H)$ . Then show that  $U^{-1}PU \in P(H)$ .

**Theorem 11.22** (Projection operator). *Given a bounded operator  $A \in B(H)$ , we denote  $\operatorname{Range}(A) := \{Ax : x \in H\}$ .*

- (1) Let  $P \in P(H)$ . Then  $\text{Range}(P)$  is a closed subspace of  $H$  whose orthogonal complement is  $\text{Range}(1 - P)$ .
- (2) Suppose that  $H_0$  is a closed subspace. Then given  $x \in H$ , we define  $\text{proj}_{H \rightarrow H_0}(x)$  as follows :

$$(11.50) \quad \text{proj}_{H \rightarrow H_0}(x) = x_0,$$

where  $x = x_0 + x_1$  with  $x_0 \in H_0$  and  $x_1 \in H_0^\perp$ . Then  $\text{proj}(H \rightarrow H_0) \in P(H)$ .

*Proof.* (1) Let  $\{x_j\}_{j \in \mathbb{N}}$  be a sequence in  $\text{Range}(P)$  that is convergent to  $x$ . Then, we can express  $x_j$  as  $x_j = Py_j$  for some  $y_j \in H$ . Using  $P^2 = P$ , we obtain

$$(11.51) \quad Px = \lim_{j \rightarrow \infty} Px_j = \lim_{j \rightarrow \infty} P^2y_j = \lim_{j \rightarrow \infty} Py_j = \lim_{j \rightarrow \infty} x_j = x.$$

Using the self-adjointness of  $P$  and the fact  $P^2 = P$ , we obtain

$$(11.52) \quad \langle Px, (1 - P)y \rangle = \langle P(1 - P)x, y \rangle = \langle Px - P^2x, y \rangle = 0.$$

It is also easy to see  $H = \text{Range}(P) \oplus \text{Range}(1 - P)$ . Thus, we conclude  $\text{Range}(1 - P)$  is an orthonormal complement of  $\text{Range}(P)$ .

(2) We have to prove that  $P$  is self-adjoint because it is clear from the definition of  $P$  that  $P^2 = P$ . Let  $x, y \in H$ . Split  $x$  and  $y$  by  $x = x_0 + x_1, y = y_0 + y_1$  with  $x_0, y_0 \in H_0$  and  $x_1, y_1 \in H_0^\perp$ . Then we have

$$(11.53) \quad \langle Px, y \rangle = \langle x_0, y \rangle = \langle x_0, y_0 \rangle = \langle x, y_0 \rangle = \langle x, Py \rangle.$$

Thus,  $P$  is self-adjoint. □

**Definition 11.23.** Let  $P, Q \in P(H)$ . Then define an order  $P \ll Q$  by  $\text{Range}(P) \subset \text{Range}(Q)$ .

**Proposition 11.24.** Suppose  $\{H_j\}_{j \in \mathbb{N}}$  is a sequence of closed spaces.

- (1) If  $\{H_\lambda\}_{\lambda \in \Lambda}$  is decreasing, then we have

$$(11.54) \quad \lim_{j \rightarrow \infty} \text{proj}(H \rightarrow H_j)x = \text{proj} \left( H \rightarrow \bigcap_{j \in \mathbb{N}} H_j \right) x$$

for all  $x \in H$ .

- (2) If  $\{H_j\}_{j \in \mathbb{N}}$  is increasing, then we have

$$(11.55) \quad \lim_{j \rightarrow \infty} \text{proj}(H \rightarrow H_j)x = \text{proj} \left( H \rightarrow \overline{\bigcup_{j \in \mathbb{N}} H_j} \right) x$$

for all  $x \in H$ .

*Proof.* (1) If  $x \in \bigcap_{j \in \mathbb{N}} H_j$ , then it is trivial that the equality holds. Suppose that  $x \perp \bigcap_{j \in \mathbb{N}} H_j$ .

Then the equality trivially holds since both sides are zero. (2) The proof is similar. □

Hilbert-Schmidt class.

Let us now define the Hilbert-Schmidt class.

**Lemma 11.25.** Let  $\{x_j\}_{j \in \mathbb{N}}$  be a CONS of  $H$ . Assume that

$$(11.56) \quad M = \sum_{j=1}^{\infty} \langle T^*T x_j, x_j \rangle = \sum_{j=1}^{\infty} \|T x_j\|_H^2$$

is finite. Then for every CONS  $\{y_j\}_{j \in \mathbb{N}}$ , the quantity  $\sum_{j=1}^{\infty} \langle T^* T y_j, y_j \rangle$  converges to  $M$ .

*Proof.* By the Planchrel theorem, we have

$$M = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\langle T x_j, x_k \rangle|^2 = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\langle x_j, T^* x_k \rangle|^2 = \sum_{k=1}^{\infty} \|T^* x_k\|_H^2.$$

Using the Planchrel theorem once more, we obtain

$$M = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |\langle T^* x_k, y_j \rangle|^2 = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |\langle x_k, T y_j \rangle|^2 = \sum_{j=1}^{\infty} |\langle T y_j \rangle|^2.$$

This is the desired result.  $\square$

Earlier, we considered compact operators. However, it is not sufficient to consider compact operators. Sometimes, it may be useful to consider how different from the finite rank operators they are. One of such an attempt is to introduce the class of Hilbert-Schmidt operators.

**Definition 11.26.** Let  $\{x_j\}_{j \in \mathbb{N}}$  be a CONS in  $H$ . Then define

$$(11.57) \quad S(H) := \{T \in B(H) : \|T\|_{S(H)} < \infty\}.$$

where

$$(11.58) \quad \|T\|_{S(H)} = \left( \sum_{j=1}^{\infty} \langle T^* T x_j, x_j \rangle \right)^{\frac{1}{2}}.$$

**Proposition 11.27.** If  $T \in S(H)$ , then we have

$$(11.59) \quad \|T\|_{B(H)} \leq \|T\|_{S(H)}.$$

*Proof.* Let  $x \in H$ . Then we have

$$\begin{aligned} \langle T x, T x \rangle &= \sum_{j,k=1}^{\infty} \langle T x_j, T x_k \rangle \cdot \langle x_j, x \rangle \cdot \langle x, x_k \rangle \\ &\leq \sum_{j,k=1}^{\infty} |\langle T x_j, T x_k \rangle| \cdot |\langle x_j, x \rangle| \cdot |\langle x, x_k \rangle| \\ &\leq \sum_{j,k=1}^{\infty} \|T x_j\|_H \|T x_k\|_H |\langle x_j, x \rangle| \cdot |\langle x, x_k \rangle| \\ &= \left( \sum_{j=1}^{\infty} \|T x_j\|_H \cdot |\langle x_j, x \rangle| \right)^2 \\ &\leq \left( \sum_{j=1}^{\infty} \|T x_j\|_H^2 \right) \left( \sum_{j=1}^{\infty} |\langle x_j, x \rangle|^2 \right) \\ &\leq S(H)^2 \|x\|_H^2. \end{aligned}$$

This is the desired result.  $\square$

Let us denote by  $K(H)$  the set of all compact operators in  $B(H)$ .

**Corollary 11.28.** One has  $S(H) \subset K(H)$  for Hilbert spaces  $H$ .



*Proof.* Take  $T \in S(H)$  arbitrarily. Let  $P_N := \text{proj}(H \rightarrow \{x_1, x_2, \dots, x_N\})$ . Then we have by Proposition 11.27 that

$$(11.60) \quad \|T - TP_N\|_{B(H)} \leq \|T - TP_N\|_{S(H)} = \left( \sum_{j=N+1}^{\infty} \|Tx_j\|_{H^2}^2 \right)^{\frac{1}{2}} \rightarrow 0.$$

Therefore,  $TP_N$  being compact, so is  $T$ . □

**Proposition 11.29.** *The space  $S(H)$  is closed under taking adjoint and the adjoint operation preserves the norm.*

*Proof.* Let  $T \in S(H)$  and consider its polar decomposition:  $T = WS$ , where  $S = \sqrt{T^*T}$  is a positive self-adjoint operator and  $W$  is a partial isometry. Then we have

$$(11.61) \quad \|T^*\|_{S(H)}^2 = \sum_{j=1}^{\infty} \langle WSSWx_j, x_j \rangle = \sum_{j=1}^{\infty} \langle SWx_j, SWx_j \rangle \leq \sum_{j=1}^{\infty} \langle Sx_j, Sx_j \rangle,$$

because  $\{Wx_j\}_{j \in \mathbb{N}}$  forms an orthonormal system, although it can happen that it is not complete. Hence it follows that

$$(11.62) \quad \|T^*\|_{S(H)}^2 \leq \sum_{j=1}^{\infty} \langle Sx_j, Sx_j \rangle = \sum_{j=1}^{\infty} \langle S^2x_j, x_j \rangle = \sum_{j=1}^{\infty} \langle T^*Tx_j, x_j \rangle = \|T\|_{S(H)}^2 < \infty.$$

Hence  $T^*$  is in the Schmidt class. Changing the role of  $T$  and  $T^*$  it follows that

$$(11.63) \quad \|T^*\|_{S(H)}^2 \leq \|T\|_{S(H)}^2$$

and hence  $\|T\|_{S(H)} = \|T^*\|_{S(H)}$ . □

**Proposition 11.30.** *Let  $T_0 \in B(H)$  and  $T_1 \in S(H)$ . Then we have  $T_0T_1, T_0T_1 \in S(H)$ .*

*Proof.* It suffices to show that  $T_1T_0 \in S(H)$  because we have established that  $S(H)$  is closed under adjoint in Proposition 11.29. Indeed,

$$\|T_0T_1\|_{S(H)}^2 = \sum_{j=1}^{\infty} \langle (T_0T_1)^*T_0T_1x_j, x_j \rangle = \sum_{j=1}^{\infty} \|T_0T_1x_j\|_{H^2}^2 \leq \|T_0\|_{B(H)}^2 \sum_{j=1}^{\infty} \|T_1x_j\|_{H^2}^2 < \infty,$$

which proves the assertion. □

We conclude this section with an example.

**Example 11.31.** Let  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  be a  $\sigma$ -finite measure space. Assume that  $K : X \times Y \rightarrow \mathbb{C}$  is an  $\mathcal{M} \otimes \mathcal{N}$ -measurable function such that

$$(11.64) \quad M = \int_{X \times Y} |K(x, y)|^2 d\mu(x) d\nu(y) < \infty.$$

Then the operator

$$(11.65) \quad Tf(y) := \int_X K(x, y)f(x) d\mu(x) \quad (y \in Y)$$

is belongs to  $S(H)$  and it satisfies  $\|T\|_{S(H)} \leq M$ .

*Proof.* First we remark that the integral defining  $Tf(y)$  is finite for  $\nu$ -almost every  $y \in Y$ . Furthermore, we have

$$\begin{aligned} \int_Y |Tf(y)|^2 d\nu(y) &\leq \int_Y \left( \int_X |f(x)|^2 d\mu(x) \cdot \int_X |K(x, y)|^2 d\mu(x) \right) d\nu(y) \\ &\leq M^2 \int_X |f(x)|^2 d\mu(x). \end{aligned}$$

Hence  $T \in B(L^2(X), L^2(Y))$ . With this in mind, let  $\{e_j\}_{j \in \mathbb{N}}$  be a CONS. Then we have

$$\begin{aligned} \sum_{j=1}^{\infty} \langle T e_j, T e_j \rangle &= \sum_{j=1}^{\infty} \int_Y \left( \int_X K(x, y) e_j(x) d\mu(x) \cdot \int_X \overline{K(x', y)} e_j(x') d\mu(x') \right) d\mu(y) \\ &= \sum_{j=1}^{\infty} \int_{X \times X \times Y} K(x, y) e_j(x) \overline{K(x', y)} e_j(x') d\mu(x) d\mu(x') d\nu(y) \\ &= \sum_{j=1}^{\infty} \int_{X \times X} e_j(x) \overline{e_j(x')} \left( \int_Y K(x, y) \overline{K(x', y)} d\mu(y) \right) d\mu(x) d\mu(x'). \end{aligned}$$

Here our calculation is justified because the integral converges absolutely.  $\square$

Closed operators and adjoints (on Hilbert spaces).

Now let us investigate closed operators on Hilbert spaces. Recall that we have defined closed operators in Definition 9.11. Let  $H, H_1, H_2$  be Hilbert spaces throughout.

Let us begin with defining adjoint operators. Note that due to the sesquilinearity of inner product, we need to start again from the definition.

**Definition 11.32.** Assume that  $A : H_1 \rightarrow H_2$  is an unbounded operator such that  $D(A)$  is dense in  $H_1$ . Then define

$$\begin{aligned} D(A^*) &:= \{y \in H_2 : \text{there exists } v \in H_1 \text{ such that } \langle x, v \rangle_{H_1} = \langle Ax, y \rangle_{H_2} \text{ for all } x \in D(A)\} \\ A^*y &:= v, y \in D(A^*) \text{ if } \langle x, v \rangle_{H_1} = \langle Ax, y \rangle_{H_2} \text{ for all } x \in D(A). \end{aligned}$$

**Lemma 11.33.** Let  $A$  be an unbounded operator. Then we have

$$(11.66) \quad D(A^*) = \{y \in H_2 : |\langle Ax, y \rangle_{H_2}| \leq M_y \|x\|_{H_1} \text{ for all } x \in H_1\}.$$

*Proof.* Easy by using the Riesz representation theorem.  $\square$

**Proposition 11.34.** Let  $A : H_1 \rightarrow H_2$  be a closed operator. Then we have

$$H_2 = \overline{R(A)} \oplus N(A^*).$$

*Proof.* We have

$$\begin{aligned} y \in R(A)^\perp &\iff \langle y, Ax \rangle_{H_2} = 0 \text{ for all } x \in D(A) \\ &\iff y \in D(A^*) \text{ and } A^*y = 0 \\ &\iff A^*y = 0. \end{aligned}$$

Therefore, we conclude that  $R(A)^\perp = N(A^*)$ . Since  $H_2 = \overline{R(A)} \oplus R(A)^\perp$ , we have the desired result.  $\square$

**Corollary 11.35.** Let  $A : H_1 \rightarrow H_2$  be a densely defined closed operator. Then

$$(11.67) \quad H_1 \oplus H_2 = \{(x, Ax) : x \in D(A)\} \oplus \{(-A^*y, y) : y \in D(A^*)\}.$$

*Proof.* Let us define  $T : H_1 \rightarrow H_1 \oplus H_2$  by  $D(T) = D(A)$  and  $Tx = (x, Ax)$  for  $x \in D(A)$ . Then  $T$  is a closed operator. Note that

$$\begin{aligned} N(T^*) &= \{(u, v) \in D(T^*) : T^*(u, v) = 0\} \\ &= \{(u, v) \in D(T^*) : \langle T^*(u, v), x \rangle_{H_1} = 0 \text{ for all } x \in D(T)\} \\ &= \{(u, v) \in D(T^*) : \langle u, x \rangle_{H_1} + \langle v, Ax \rangle_{H_2} = 0\} \\ &= \{(-A^*v, v) : v \in D(A^*)\}. \end{aligned}$$

Thus, we have only to apply Proposition 11.34.  $\square$

**Theorem 11.36.** *Let  $A : H_1 \rightarrow H_2$  be a densely defined closed operator.  $A^*$  is densely defined and  $A^{**} = A$ .*

*Proof.* Let  $z \in D(A^*)^\perp$ . Then we have

$$(11.68) \quad (0, z) \in \{(-A^*y, y) : y \in D(A^*)\}^\perp.$$

By virtue of Corollary 11.35, we have

$$(11.69) \quad (0, z) \in \{(x, Ax) : x \in D(A)\},$$

which implies that  $z = 0$ .  $\square$

**Proposition 11.37.** *Let  $A : H_1 \rightarrow H_2$  be a densely closed operator. Assume in addition that  $0 \in \rho(A)$ . Then we have  $0 \in \rho(A^*)$  and  $(A^{-1})^* = (A^*)^{-1} : H_1 \rightarrow H_2$ .*

*Proof.* It is easy to see that  $A^* : H_2 \rightarrow H_1$  satisfies

$$(11.70) \quad N(A^*) = \{0\}, R(A^*) = H_1.$$

because by Corollary 11.35

$$(11.71) \quad N(A^*) = R(A)^\perp = H_2^\perp = \{0\}, R(A^*)^\perp = N(A^{**}) = N(A) = \{0\}.$$

Now that  $R(A) = H_2$  and  $R(A^*) = H_1$ , we have

$$(11.72) \quad \langle x, A^{-1}y \rangle_{H_1} = \langle (A^*)^{-1}x, y \rangle_{H_2},$$

which gives  $(A^{-1})^* = (A^*)^{-1}$ .  $\square$

Symmetric and self-adjoint operators.

**Definition 11.38.** Let  $A$  be an unbounded operator on  $H$ .

- (1)  $A$  is said to be symmetric if  $A \subset A^*$ .
- (2)  $A$  is said to be self-adjoint, if  $A = A^*$ .

**Proposition 11.39.** *Let  $S : H \rightarrow H$  be an injective self-adjoint operator. Then  $R(S) = D(S^{-1})$  is dense in  $H$  and  $S^{-1} : H \rightarrow H$  is a self-adjoint operator.*

*Proof.* First, let us prove that  $R(S) = D(S^{-1})$  is dense in  $H$ . To this end, we let  $y \in R(S)^\perp = D(S^{-1})^\perp$ . Then we have

$$(11.73) \quad \langle y, Sx \rangle_H = 0$$

for all  $x \in D(S)$ . This implies that

$$(11.74) \quad y \in D(S^*) \text{ and } S^*y = 0.$$

Since  $S = S^*$ , we conclude  $Sy = 0$ . By assumption  $S$  is injective. Thus, we have  $y = 0$ .

Let us calculate  $D((S^{-1})^*)$ .

$$\begin{aligned} D((S^{-1})^*) &= \{x \in H : \exists z \in H \text{ such that } \langle x, S^{-1}y \rangle_H = \langle z, y \rangle_H \text{ for all } y \in D(S^{-1})\} \\ &= \{x \in H : \exists z \in H \text{ such that } \langle x, y \rangle_H = \langle z, Sy \rangle_H \text{ for all } y \in D(S)\} \\ &= R(S^*) = R(S) = D(S^{-1}). \end{aligned}$$

Also, if  $x \in D(S^{-1})$ , then we deduce that  $x = Sz$ , that is,  $(S^{-1})^*z = z = S^{-1}x$ .  $\square$

The following theorem is due to von Neumann.

**Theorem 11.40** (von Neumann). *Let  $H_1$  and  $H_2$  be Hilbert spaces. Let  $A : H_1 \rightarrow H_2$  be a densely defined closed operator. Then  $A^*A$  is a self-adjoint operator.*

*Proof.* It is easy to see that  $A^*A$  is symmetric. Indeed, if  $x, y \in D(A^*A)$ , then we have

$$(11.75) \quad \langle A^*Ax, y \rangle_{H_1} = \langle Ax, Ay \rangle_{H_2}$$

because  $Ax \in D(A^*)$ ,  $y \in D(A)$ . Since  $Ay \in D(A^*)$ ,  $x \in D(A)$ , it follows that

$$(11.76) \quad \langle Ax, Ay \rangle_{H_2} = \langle x, A^*Ay \rangle_{H_1}.$$

Hence, we obtain  $A^*A$  is symmetric.

Let us set  $S = 1 + A^*A$ . Since we have established in Corollary 11.35 that

$$(11.77) \quad H_1 \oplus H_2 = \{(x, Ax) : x \in D(A)\} \oplus \{(-A^*y, y) : y \in D(A^*)\},$$

we see that  $S : D(A^*A) \rightarrow H_1$  is a bijection. We also have

$$(11.78) \quad \|x\|_{H_1} \leq \|Sx\|_{H_1}$$

for all  $x \in H_1$ .

Once we prove that  $S$  is self-adjoint, we will have been proved that  $A^*A$  is self-adjoint.

Observe that  $\langle S^{-1}x, y \rangle_{H_1} = \langle x, S^{-1}y \rangle_{H_1}$  for all  $x, y \in H_1$ . Indeed, to check this, we have only to write the above formula by using  $u, v \in D(S)$  such that  $x = Su$ ,  $y = Sv$  and use the fact that  $S \subset S^*$ . Since  $S^{-1}$  was proved to be self-adjoint, we see that  $S$  is self-adjoint.  $\square$

**Theorem 11.41.** *Let  $T : H_1 \rightarrow H_2$  be a densely defined operator. Then we have*

$$(11.79) \quad H_1 \oplus H_2 = \{(v, Tv) \in H_1 \oplus H_2 : v \in D(T)\} \oplus \{(-T^*w, w) \in H_1 \oplus H_2 : w \in D(T^*)\}.$$

*Proof.* Let  $(x, y) \in \{(v, Tv) : v \in D(T)\}^\perp \cap \{(-T^*w, w) : w \in D(T^*)\}^\perp$ . Then we have

$$(11.80) \quad \langle x, v \rangle + \langle y, Tv \rangle = 0, \quad v \in D(T).$$

Thus, we have  $y \in D(T^*)$  and  $x = -T^*y$ . Similarly,  $x \in D(T)$  and  $y = Tx$ . Hence we have  $(I + T^*T)x = 0$ . From this we conclude  $x = 0$  and  $y = 0$ . Hence (11.79) follows.  $\square$

Polar decomposition of bounded linear operators. We conclude this section with the polar decomposition of bounded linear operators. With this result, a lot of things are reduced to considering isometries and self-adjoint operators. Again we let  $H_0$  and  $H_1$  Hilbert spaces.

**Definition 11.42.** A bounded linear operator  $W \in B(H_0, H_1)$  is said to be a partial isometry, if  $W^*W$  is identity on  $\text{Ker}(T)^\perp$ .

*Exercise 115.* Let  $W \in B(H_0, H_1)$  be a partial isometry. Show the following.

- (1) There exist orthonormal systems  $\{x_j\}_{j \in J_1} \subset H_0$ ,  $\{\tilde{x}_j\}_{j \in J_2} \subset H_0$  and  $\{\tilde{y}_j\}_{j \in J_3} \subset H_1$  such that

$$(11.81) \quad \{x_j\}_{j \in J_1} \cup \{\tilde{x}_j\}_{j \in J_2}$$

is a CONS of  $H_0$  and

$$(11.82) \quad \{W x_j\}_{j \in J_1} \cup \{\tilde{y}_j\}_{j \in J_3}$$

is a CONS of  $H_1$ .

- (2) The adjoint  $W^* \in B(H_1, H_0)$  is a partial isometry as well.

**Theorem 11.43.** Any  $T \in B(H)$  admits the following decomposition:  $T$  is decomposed as  $T = W S$ , where  $S \in S(H)$  is a positive element and  $W$  is a partial isometry.

*Proof.* Let us define  $S := \sqrt{T^* T}$ . Then  $S$  is a positive element in  $S(H)$ . Let us claim that  $T$  is an injection if and only if  $S$  is. Indeed, if we let  $x, y \in H$ , then we have

$$\begin{aligned} \|S(x - y)\|_H^2 &= \langle Sx - Sy, Sx - Sy \rangle \\ &= \langle S(x - y), S(x - y) \rangle \\ &= \langle S^2(x - y), x - y \rangle \\ &= \langle T^* T(x - y), x - y \rangle \\ &= \|T(x - y)\|_H^2. \end{aligned}$$

Therefore, the operator

$$(11.83) \quad w : Tx \mapsto Sx$$

is a well-defined injection defined on  $\{Tx : x \in H\}$ . Furthermore, going through a similar argument as above, we see that it preserves the norm. Therefore, this mapping  $w$  extends to a partial isometry  $W_0$  such that  $W_0 T = S$ . Therefore,  $T = W_0^* S$  is our desired decomposition.  $\square$

#### 11.4. Reproducing kernel Hilbert spaces.

As examples of Hilbert spaces, we can list reproducing kernel Hilbert spaces. With reproducing kernel Hilbert spaces, we are convinced that Hilbert spaces do carry a rich structure.

Let us present the definition.

**Definition 11.44.** A complex-valued Hilbert space  $H$  is said to be a reproducing kernel Hilbert space, if there exists a function  $K : E \times E \rightarrow \mathbb{C}$  with the properties listed below, where  $E$  is a set.

- (1) For every  $s \in E$ ,  $K(\cdot, s) \in H$ .
- (2)  $\langle K(\cdot, s), K(\cdot, t) \rangle = K(t, s)$
- (3)  $\{K(\cdot, s) : s \in E\}$  spans a dense subspace in  $H$ .

The kernel  $K$  is said to be the reproducing kernel of  $H$ .

By the second condition, we have

$$(11.84) \quad K(t, s) = \langle K(\cdot, s), K(\cdot, t) \rangle = \overline{\langle K(\cdot, t), K(\cdot, s) \rangle} = \overline{K(s, t)} \quad (s, t \in E).$$

**Lemma 11.45.** The reproducing kernel  $K$  in a reproducing kernel Hilbert space  $H$ , if there exists, is unique.

*Proof.* A passage to the limit allows us to have

$$(11.85) \quad \langle f, K(\cdot, t) \rangle = f(t)$$

for all  $f \in H$ . Suppose that  $K^* : E \times E \rightarrow \mathbb{C}$  is another kernel. Then we have

$$(11.86) \quad K^*(t, s) = \langle K^*(\cdot, s), K(\cdot, t) \rangle = \overline{\langle K(\cdot, t), K^*(\cdot, s) \rangle} = \overline{K(s, t)} = K(t, s).$$

Therefore, the kernel is unique.  $\square$

**Lemma 11.46.** *Suppose that  $H$  is a reproducing kernel Hilbert space with reproducing kernel  $K$ . Then  $K$  is positive definite, that is,  $\{K(x_i, x_j)\}_{i,j=1,\dots,K}$  is positive definite for all finite subsets  $\{x_1, x_2, \dots, x_K\}$ .*

*Proof.* Let  $a_1, \dots, a_K \in \mathbb{C}$  be arbitrary scalars. Then we have

$$\begin{aligned} \sum_{i=1}^K \sum_{j=1}^K \bar{a}_i a_j K(x_i, x_j) &= \sum_{i=1}^K \sum_{j=1}^K \bar{a}_i a_j \langle K(\cdot, x_j), K(\cdot, x_i) \rangle \\ &= \left\langle \sum_{j=1}^K a_j K(\cdot, x_j), \sum_{i=1}^K a_i K(\cdot, x_i) \right\rangle_H \\ &\geq 0, \end{aligned}$$

which shows the positive definiteness.  $\square$

**Example 11.47.** Let  $w : (0, \infty) \rightarrow (0, \infty)$  be a continuous function such that  $w^{-1}$  is integrable over  $(0, R]$  for any  $R > 0$ . Then define

$$(11.87) \quad H_K(w) = \{f : (0, \infty) \rightarrow \mathbb{C} : f \text{ is absolutely continuous and } \|f\|_{H_K(w)} < \infty\},$$

where the norm is given by

$$(11.88) \quad \|f\|_{H_K(w)} = \left( \int_0^\infty |f'(t)|^2 w(x) dx \right)^{\frac{1}{2}}.$$

Then  $H_K(w)$  is a reproducing kernel Hilbert space whose reproducing kernel  $K(s, t)$  is given by

$$(11.89) \quad K(s, t) = \int_0^{\min(s,t)} w(\xi)^{-1} d\xi.$$

*Proof.* We can prove with ease that  $H_K(w)$  is a normed space whose norm obeys the parallelogram law. Therefore, for the proof that  $H_K(w)$  is a Hilbert space, it suffices to show that  $H_K(w)$  is complete. The key estimate is

$$(11.90) \quad |f(t)| \leq \left( \int_0^t w(s)^{-1} ds \right) \cdot \|f\|_{H_K(w)}.$$

By using this inequality we can prove in particular that any Cauchy sequence  $\{f_k\}_{k \in \mathbb{N}}$  converges locally uniformly to a continuous function  $f$ . It is not so hard to show that  $f$  is a limit of the Cauchy sequence.

Let us show that the kernel is given by (11.89). By the fundamental theorem of calculus we have  $\frac{\partial K}{\partial s}(s, t) \Big|_{s=\xi} = \chi_{(0,t)}(\xi) w(\xi)^{-1}$ . Let  $f \in H_K(w)$ . Then from the definition of the inner product of  $H_K(w)$ , we have

$$\langle f, K(\cdot, t) \rangle_{H_K(w)} = \int_0^\infty \frac{\partial K}{\partial s}(s, t) \Big|_{s=\xi} \cdot f'(\xi) w(\xi) d\xi = \int_0^t f'(\xi) d\xi = f(t),$$

because  $f(0) = 0$  by definition. Therefore, we conclude that the kernel is given by (11.89).  $\square$

Notes and references for Chapter 5.

We refer to [69, 71] for further facts. Below we describe how each theorem came about.

Section 9. Theorem 9.3

Theorem 9.9

Theorem 9.27

Banach proved Theorem 9.28 in [88].

Theorem 9.29

Theorem 9.30 was obtained by Hahn (1922) and Banach (1927) independently (see [226] and [88] respectively). Later, after the paper [88], Banach became aware of Hahn's paper [226] and acknowledged the priority of Hahn.

Theorem 9.31

Theorem 9.35

Theorem 10.1

Section 10. Theorem 10.2

Theorem 10.9

Theorem 10.10 is called the Aoki-Rolewicz theorem. We refer to [83, 411] for their independent proofs.

Theorem 10.12

Theorem 10.13

Theorem 10.15 was obtained by Baire [87].

Theorem 10.17 has a history. In the late 1920s S. Banach proved Theorem 10.17. In 1932 in his monograph, *Théorie des opérations linéaires*, he published this result. The proof we presented here is due to Schauder in 1930 [432].

Theorems 10.18 and 10.19 are also obtained in the monograph of S. Banach in 1932 [89].

Section 11. Theorem 11.3

Theorem 11.4 is due to Jordan and Von Neumann [262].

Theorem 11.6

Theorem 11.7

Theorem 11.8

Theorem 11.9

Theorem 11.11

Theorem 11.14

Theorem 11.16

Theorem 11.17

Theorem 11.18

Theorem 11.19

Theorem 11.22

Theorem 11.43

Theory of reproducing kernel dates back to [84]. An important theory connecting with the reproducing kernel Hilbert spaces is theory of Bergman spaces in complex analysis. For details we refer to [22].



**Part 6. Maximal operators and singular integral operators**

**Part 7. Maximal operators**

This part is the heart of this book. In this part we make a detailed look of Hardy-Littlewood maximal operators and singular integral operators. To explain Hardy-Littlewood maximal operators, we place ourselves in the setting of Lebesgue measure  $dx$  on  $\mathbb{R}^d$  for the time being. The history of the Hardy-Littlewood maximal operator dates back to

G. H. Hardy and J. E. Littlewood, A maximal theorem with function-theoretic applications, *Acta Mathematica* 54 (1930), 81–116

and the original definition is as follows. Denote by  $\mathcal{I}_t(\mathbb{R})$  the set of all open intervals containing  $t \in \mathbb{R}$ . The Hardy-Littlewood maximal operator  $M$  is defined by

$$Mf(t) = \sup_{t \in \mathcal{I}_t(\mathbb{R})} \frac{1}{|I|} \int_I |f(s)| ds$$

for measurable functions  $f : \mathbb{R} \rightarrow \mathbb{C}$ . The original aim was to apply it to the functions on the unit disk  $\Delta(1)$  on the complex plane.

Denote by  $\mathcal{B}_x(\mathbb{R}^d)$  the set of all open balls containing  $x \in \mathbb{R}^d$ . A natural passage to the higher dimension is

$$Mf(x) = \sup_{B \in \mathcal{B}_x(\mathbb{R}^d)} \frac{1}{|B|} \int_B |f(y)| dy$$

for measurable functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$ , which is made originally by

N. Wiener, The ergodic theorem, *Duke Mathematical Journal* 5 (1939), 1-18.

His purpose was to apply it to the ergodic theory.

However, recently, the Hardy-Littlewood maximal operator is used in many mathematical contexts.

The first example is closely related to PDE.

**Example 11.48.** Let  $I_\alpha$  be a fractional maximal operator given by

$$I_\alpha f(x) = \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}} dy \quad (x \in \mathbb{R}^d)$$

for positive measurable function  $f : \mathbb{R}^d \rightarrow [0, \infty]$ . Let  $1 \leq p < \frac{d}{\alpha}$  and define  $q$  so that  $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$ . Then Hedberg showed in 1972 that

$$I_\alpha f(x) \lesssim Mf(x)^{\frac{p}{q}} \|f\|_{L^p}^{1-\frac{p}{q}} \quad (x \in \mathbb{R}^d).$$

The next example is from Fourier analysis and theory of function spaces.

**Example 11.49.** Let  $R > 0$ . Denote by  $\mathcal{S}'(\mathbb{R}^d)_{B(R)}$  the set of all elements  $f \in \mathcal{S}'(\mathbb{R}^d)$  whose Fourier transform is supported in the open ball  $B(R) = \{y \in \mathbb{R}^d : |y| < R\}$ .

Then the Plancherel-Polya-Nikolskij theorem asserts that

$$\sup_{y \in \mathbb{R}^d} \frac{|f(x-y)|}{(1+R|y|)^{n/\eta}} \leq CM[|f|^\eta](x)^{1/\eta}$$

for  $\eta > 0$ . This theorem turns out to be important in conjunction with the theory of function spaces.

The next example is closely related to theory of probability.

**Example 11.50.** Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . Then consider

$$f_j \equiv \sum_{m \in \mathbb{Z}^d} \left( \int_{2^{-j}m + [0, 2^{-j}]^d} f(z) \frac{dz}{|2^{-j}m + [0, 2^{-j}]^d|} \right) \chi_{2^{-j}m + [0, 2^{-j}]^d}.$$

Note that  $f_j$  is obtained by the averaging procedure of  $f$  with respect to the dyadic grid  $\mathcal{D}_j = \{2^{-j}m + [0, 2^{-j}]^d\}_{m \in \mathbb{Z}^d}$ . Then it is easy to show that

$$|f_j| \leq C Mf$$

with constant independent of  $j$  and  $f$ .

This type of estimate is important when we consider the limit of  $f_j$  as  $j \rightarrow \infty$ . The topic taken up in this section is used and expanded in various directions. This part is devoted to making more precise what was dealt in Part 0. In Section 12 we shall deal with the Hardy-Littlewood maximal operators. Starting from the definitions and elementary properties, we collect some auxiliary but important results: Covering lemmas play a crucial role everywhere in harmonic analysis. Covering lemmas take many kind of forms and appear in the proofs. First we take up the  $5r$ -covering lemma. After building up a fundamental theory about Hardy-Littlewood maximal operators, we shall make a view of covering lemmas and of how it is used. In Section 13 we shall present non-trivial examples, where the Hardy-Littlewood maximal operator plays a powerful role.

## 12. MAXIMAL OPERATORS

In this section we shall make a systematic treatment of the Hardy-Littlewood maximal operators. As we have seen in Chapter 1, what lies behind the weak-(1, 1) boundedness is the covering lemma like Lemma 1.4.

### 12.1. Definition of Hardy-Littlewood maximal operators and elementary properties.

For a Lebesgue measurable function  $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\pm\infty\}$  or  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  define the centered Hardy-Littlewood maximal operator by

$$(12.1) \quad M'f(x) = \sup_{r>0} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y)| dy = \sup_{r>0} m_{B(x, r)}(|f|) \quad (x \in \mathbb{R}^d).$$

Denote by  $\mathcal{B}(x)$  the set of all open balls containing  $x$ . For a Lebesgue measurable function  $f$  define the uncentered Hardy-Littlewood maximal operator by

$$Mf(x) = \sup_{B \in \mathcal{B}(x)} \frac{1}{|B|} \int_B |f(z)| dz = \sup_{B \in \mathcal{B}(x)} m_B(|f|) \quad (x \in \mathbb{R}^d).$$

**Definition 12.1.** For a Lebesgue measurable function  $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\pm\infty\}$  or  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  define the centered Hardy-Littlewood maximal operator by

$$M'f(x) = \sup_{r>0} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y)| dy = \sup_{r>0} m_{B(x, r)}(|f|).$$

Denote by  $\mathcal{B}(x)$  the set of all open balls containing  $x$ . For a Lebesgue measurable function  $f$  define the uncentered Hardy-Littlewood maximal operator by

$$(12.2) \quad Mf(x) = \sup_{B \in \mathcal{B}(x)} \frac{1}{|B|} \int_B |f(z)| dz = \sup_{B \in \mathcal{B}(x)} m_B(|f|).$$

Here and below by “measurable function” we mean a measurable function taking its value in  $\mathbb{R} \cup \{\pm\infty\}$  or  $\mathbb{C}$ .

The following lemma is trivial but in more complicated situation we have to take into account the difference between these Hardy-Littlewood maximal operators. Detailed readers can solve exercises in 12.3.

**Lemma 12.2.** *For any measurable function  $f$  we have*

$$(12.3) \quad M'f(x) \leq Mf(x) \quad (x \in \mathbb{R}^d).$$

*Exercise 116.* Explain briefly why (12.3) holds.

We now prove the reverse estimate of (12.3).

**Lemma 12.3.** *For any measurable function  $f$  we have*

$$(12.4) \quad Mf(x) \leq 2^d M'f(x) \quad (x \in \mathbb{R}^d).$$

*Proof.* To prove (12.4) it is sufficient from the definition of sup defining  $Mf$  that we prove

$$(12.5) \quad m_B(|f|) \leq 2^d M'f(x)$$

for any ball  $B$  containing  $x$ . Let  $B^* = B(x, 2r(B))$ , a ball centered at  $x$  of radius  $2r(B)$ . Then a geometric observation shows  $B^*$  engulfs  $B$ . Since the volume of a ball  $B$  is proportional to  $r(B)^d$ , we conclude that

$$(12.6) \quad |B^*| = 2^d |B|.$$

Therefore, it follows that

$$m_B(|f|) = \frac{1}{|B|} \int_B |f(y)| dy \leq \frac{2^d}{|B^*|} \int_{B^*} |f(y)| dy \leq 2^d m_{B^*}(|f|) \leq 2^d M'f(x),$$

proving (12.5). □

In order to discuss the measurability of the Hardy-Littlewood maximal functions, we recall a terminology describing continuity of functions.

**Definition 12.4.** Let  $X$  be a topological space. A function  $f : X \rightarrow \mathbb{R} \cup \{\infty\}$  is said to be lower semicontinuous, if the set

$$(12.7) \quad \{x \in X : f(x) > \lambda\}$$

is open in  $X$  for every  $\lambda \in \mathbb{R}$ .

*Exercise 117.* Explain why any lower semicontinuous function on  $\mathbb{R}^d$  is Borel-measurable.

The following lemma ensures the measurability of the Hardy-Littlewood maximal operators.

**Lemma 12.5.** *The Hardy-Littlewood operators  $M'f, Mf$  are lower semicontinuous.*

*Exercise 118.* Prove Lemma 12.5. Hint: Use continuity of integral, if necessary.

In view of these lemmas we are tempted to confuse these Hardy-Littlewood maximal operators and we do so. Unless it is not specified, there is no harm in confusing these guys.

**Lemma 12.6** (Subadditivity of the Hardy-Littlewood maximal operators). *Let  $f, g$  be measurable functions and  $a \in \mathbb{R}$ . Then we have*

$$M[f + g](x) \leq Mf(x) + Mg(x), \quad M[af](x) = |a|Mf(x) \quad (x \in \mathbb{R}^d).$$

*The same can be said for  $M'$ .*

*Exercise 119.* Prove Lemma 12.6.

In functional analysis the boundedness of the operator is often of importance. A very simple result is given below.

**Theorem 12.7** ( $L^\infty(\mathbb{R}^d)$ -boundedness of the Hardy-Littlewood maximal operator). *For all  $f \in L^\infty$   $\|M'f\|_\infty \leq \|f\|_\infty$ .*

*Exercise 120.* Prove Theorem 12.7, returning to the definition.

*Exercise 121* (Translation and dilation invariance). Let  $a \in \mathbb{R}^d$  and  $k \in \mathbb{R} \setminus \{0\}$ . Then show that

$$(12.8) \quad M[f(\cdot - a)](x) \equiv Mf(x - a), \quad M[f(k\cdot)](x) \equiv Mf(kx) \quad (x \in \mathbb{R}^d).$$

We calculate an indicator function of the unit ball in order that the reader can be familiar with the definition.

**Example 12.8.** Here we consider  $d = 1$ . Define  $f := \chi_{[-1,1]}$ . Then a simple calculation gives us

$$M'f(t) = \sup_{r>0} \int_{[t-r, t+r]} \frac{\chi_{[-1,1]}(y) dy}{2r} = \sup_{r>0} \frac{|[t-r, t+r] \cap [-1, 1]|}{2r} = \min \left\{ 1, \frac{2}{1+|t|} \right\}$$

for  $t \in \mathbb{R}$ . Thus, in particular,  $M'f$  is not an  $L^1(\mathbb{R})$ -function:

$$\int_{-\infty}^{\infty} M'f(t) dt = \infty.$$

Thus we have to accept that  $M'$  does not send  $L^1(\mathbb{R})$  function to  $L^1(\mathbb{R})$ .

Furthermore we have the following.

**Proposition 12.9.** *Suppose that  $f$  is a nonzero measurable function. Then  $M'f$  is never an integrable function.*

*Proof.* Fix a ball  $B(x, r)$  so large that  $K := \int_{B(x,r)} |f(y)| dy > 0$ . By taking  $r$  sufficiently large we may assume that  $x = 0$ .

A simple calculation similar to the above case leads us to the inequality

$$(12.9) \quad \frac{Kr^d}{r^d + |x|^d} \lesssim M'f(x).$$

This disproves  $M'f \in L^1(\mathbb{R}^d)$ . □

## 12.2. $5r$ -covering lemma.

Motivated by the above example and Proposition 12.9, a question arises.

**Problem 12.10.** *Where do  $M$  and  $M'$  send the  $L^1(\mathbb{R}^d)$  functions?*

We try to answer this problem. Since (12.8) is almost the only example, where we can calculate the exact value of the maximal operators, we need some indirect approaches.

The covering lemma is one of the basic tools in harmonic analysis. Although the statement are referred to as a lemma, this lemma is one of the most important tools in harmonic analysis.

**Theorem 12.11** (5r-covering lemma). *Let  $(X, d)$  be a metric space. Suppose we are given a family of balls  $\{B(x_\lambda, r_\lambda)\}_{\lambda \in \Lambda}$  with bounded diameter: Assume that  $R := \sup_{\lambda \in \Lambda} r_\lambda < \infty$ . Then*

*we can take a disjoint subfamily of  $\{B(x_\lambda, r_\lambda)\}_{\lambda \in \Lambda_0}$  such that*

$$(12.10) \quad \bigcup_{\lambda \in \Lambda} B(x_\lambda, r_\lambda) \subset \bigcup_{\lambda \in \Lambda_0} B(x_\lambda, 5r_\lambda).$$

*Proof.* Special case Assume first that

$$(12.11) \quad \frac{R}{2} \leq \inf_{\lambda \in \Lambda} r_\lambda \leq R.$$

In this case the proof is easy and depends on the geometrical structure of  $\mathbb{R}^d$ . The proof is strict but very intuitive. We proceed in the following way. Take a maximal disjoint subfamily  $\{B(x_\lambda, r_\lambda)\}_{\lambda \in \Lambda_0}$ . This is the desired subfamily. In fact, if  $B(x_\lambda, r_\lambda)$  and  $B(x_\rho, r_\rho)$  meet ( $\lambda, \rho \in \Lambda$ ), then by assumption (12.11) we have  $B(x_\rho, r_\rho) \subset B(x_\lambda, 5r_\lambda)$  and  $B(x_\lambda, r_\lambda) \subset B(x_\rho, 5r_\rho)$ . Thus any maximal disjoint subfamily satisfies the requirement of the theorem.

General case Let us remove the additional assumption (12.11). This procedure to pass to the general case does not depend on the geometrical structure of  $\mathbb{R}^d$ .

Let us define subfamilies of the balls inductively.

We define

$$(12.12) \quad \mathcal{B}_1 := \left\{ B(x_\lambda, r_\lambda) : \frac{R}{2} \leq r_\lambda \leq R \right\}.$$

Suppose that  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$  and  $\Lambda_1, \Lambda_2, \dots, \Lambda_k$  are defined. We define  $\mathcal{B}_{k+1}$  as follows:

$$\mathcal{B}_{k+1} := \left\{ B(x_\lambda, r_\lambda) : \frac{R}{2^{k+1}} \leq r_\lambda \leq \frac{R}{2^k} \text{ and } B(x_\lambda, r_\lambda) \text{ is not contained in } \bigcup_{j=0}^k \bigcup_{B \in \Lambda_j} 5B \right\}.$$

It can happen that  $\mathcal{B}_j, j \in \mathbb{N}$  is empty. If  $\mathcal{B}_{k+1}$  is not empty, then we use the special case above to obtain  $\Lambda_{k+1}$  such that  $\{B_\lambda\}_{\lambda \in \Lambda_{k+1}}$  is disjoint and  $\bigcup_{B \in \mathcal{B}_{k+1}} B \subset \bigcup_{\lambda \in \Lambda_{k+1}} 5B_\lambda$ . Suppose that  $\lambda_0 \in \Lambda_{j_0}$  and  $\lambda_1 \in \Lambda_{j_1}$  with  $j_0 < j_1$ . Because  $B_{\lambda_1}$  is not contained in  $5B_{\lambda_0}$ ,  $B_{\lambda_1} \cap B_{\lambda_0} \neq \emptyset$ . Therefore, if we set  $\Lambda_0 = \bigcup_{j=1}^{\infty} \Lambda_j$ , then  $\{B_\lambda\}_{\lambda \in \Lambda_0}$  is disjoint.

Let  $\lambda \in \Lambda$ . Choose an integer  $j \in \mathbb{N}$  so that  $\frac{R}{2^j} \leq r(B_\lambda) \leq \frac{R}{2^{j-1}}$ . Assume  $B_\lambda$  is not contained in  $\bigcup_{k=1}^{j-1} \bigcup_{\rho \in \Lambda_k} 5B_\rho$  and  $j \geq 2$ , then from the definition of  $\Lambda_j$  and  $\mathcal{B}_j$ , we obtain  $B_\lambda \subset \bigcup_{\rho \in \Lambda_j} 5B_\rho$ . Therefore, we obtain

$$(12.13) \quad B_\lambda \subset \bigcup_{k=1}^j \bigcup_{\rho \in \Lambda_k} 5B_\rho,$$

if  $j \geq 2$ . The case when  $j = 1$  can readily be incorporated to (12.13) from the definition of  $\Lambda_1$  and  $\mathcal{B}_1$ . As a consequence, the condition

$$(12.14) \quad \bigcup_{\lambda \in \Lambda} B(x_\lambda, r_\lambda) \subset \bigcup_{\lambda \in \Lambda_0} B(x_\lambda, 5r_\lambda)$$

was achieved. □

### 12.3. Weak-(1, 1) boundedness of the Hardy-Littlewood maximal operators.

In this subsection we will answer the question posed in beginning of this section.

**Theorem 12.12** (Weak-(1,1) estimate). *The Hardy-Littlewood maximal operator  $M'$  is weak-(1, 1) bounded. More quantitatively, we have*

$$(12.15) \quad |\{M'f > \lambda\}| \leq \frac{5^d}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx$$

for all measurable functions  $f$  on  $\mathbb{R}^d$ .

*Proof.* Fix  $\lambda > 0$ . We define  $E_\lambda = \{Mf > \lambda\}$ . For all  $x \in E_\lambda$ , by its definition there exists  $r_x > 0$  such that

$$(12.16) \quad m_{B(x, r_x)}(|f|) > \lambda.$$

By Theorem 12.11 there exist  $x_1, x_2, \dots$  such that

$$(12.17) \quad \bigcup_{x \in E_\lambda} B(x, r_x) \subset \bigcup_j B(x_j, 5r_{x_j}) \text{ and } \{B(x_j, r_{x_j})\}_j \text{ is disjoint.}$$

Using this covering  $\{B(x_j, 5r_{x_j})\}_{j=1}^\infty$ , we have

$$(12.18) \quad |E_\lambda| \leq \left| \bigcup_j B(x_j, 5r_{x_j}) \right| \leq \sum_j |B(x_j, 5r_{x_j})|.$$

If we invoke (12.16), then we have

$$(12.19) \quad |E_\lambda| \leq \sum_j \frac{5^d}{\lambda} \int_{B(x_j, r_{x_j})} |f(x)| dx \leq \frac{5^d}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx.$$

This is the desired result.  $\square$

*Exercise 122.* Reexamine the proof of Theorem 12.12 to show that we are able to replace  $M'$  with  $M$  in the theorem.

*Exercise 123.* Let  $\mathcal{A} := \{A_\lambda\}_{\lambda \in \Lambda}$  be a family of subsets of a metric space  $(X, d)$ . Then  $\mathcal{A}$  is disjoint, if and only if,  $\sum_{\lambda \in \Lambda} \chi_{A_\lambda} \leq 1$ .

**Theorem 12.13** ( $L^p$ -boundedness of the Hardy-Littlewood maximal operator). *Let  $1 < p \leq \infty$ . Then*

$$(12.20) \quad \|Mf\|_p \lesssim_p \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

*Proof.* Let  $f \in L^p(\mathbb{R}^d)$ . Then we have

$$(12.21) \quad \int_{\mathbb{R}^d} Mf(x)^p dx = \int_0^\infty p \lambda^{p-1} |\{Mf > \lambda\}| d\lambda = 2^p \int_0^\infty p \lambda^{p-1} |\{Mf > 2\lambda\}| d\lambda.$$

Given  $\lambda > 0$ , we estimate the measure of the set  $\{Mf > 2\lambda\}$  at height  $\lambda$ . We now split  $f$  by  $f_1 + f_2$  with  $f_1 := \chi_{\{|f| \leq \lambda\}} \cdot f$  and  $f_2 := \chi_{\{|f| > \lambda\}} \cdot f$ .

Then taking into account  $Mf_1 \leq \lambda$ , we obtain  $|\{Mf > 2\lambda\}| \leq |\{Mf_2 > \lambda\}|$ . Apply the weak-(1, 1) inequality to obtain

$$(12.22) \quad |\{Mf_2 > \lambda\}| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f_2(x)| dx = \frac{1}{\lambda} \int_{\mathbb{R}^d} \chi_{\{(y, \rho) \in \mathbb{R}_+^{d+1} : |f(y)| > \rho\}}(x, \lambda) |f(x)| dx.$$

If we put these observations together, then we have

$$\begin{aligned} \int_{\mathbb{R}^d} Mf(x)^p dx &\lesssim \int_0^\infty \lambda^{p-2} \left( \int_{\mathbb{R}^d} \chi_{\{(x,\lambda) \in \mathbb{R}_+^{d+1} : |f(x)| > \lambda\}}(x, \lambda) |f(x)| dx \right) d\lambda \\ &= \iint_{\mathbb{R}_+^{d+1}} \chi_{\{(x,\lambda) \in \mathbb{R}_+^{d+1} : |f(x)| > \lambda\}}(x, \lambda) \lambda^{p-2} |f(x)| dx d\lambda \\ &\simeq_p \int_{\mathbb{R}^d} \left( \int_0^\infty \chi_{\{(x,\lambda) \in \mathbb{R}_+^{d+1} : |f(x)| > \lambda\}}(x, \lambda) \lambda^{p-2} |f(x)| d\lambda \right) dx \\ &\simeq_p \int_{\mathbb{R}^d} \left( |f(x)| \int_0^{|f(x)|} \lambda^{p-2} d\lambda \right) dx \\ &\simeq_p \int_{\mathbb{R}^d} |f(x)|^p dx. \end{aligned}$$

Therefore  $M$  is  $L^p(\mathbb{R}^d)$ -bounded. □

The following theorem supplements the weak-(1, 1) boundedness.

**Theorem 12.14** (Kolmogorov’s inequality). *Suppose that  $(X, \mu, \mathcal{B})$  is a finite measure space. Let  $S$  be a weak-(1, 1) bounded operator, that is, there exists a constant  $M > 0$  such that*

$$(12.23) \quad \mu \{ |Sf| > \lambda \} \leq \frac{M}{\lambda} \int_X |f(x)| d\mu(x)$$

for all  $f \in L^1(\mu)$ . Then we have

$$(12.24) \quad \int_X |Sf^\delta| d\mu \lesssim_M \mu(X)^{1-\delta} \|f\|_{L^1(\mu)}^\delta$$

for all  $0 < \delta < 1$ .

For the proof it is convenient to establish the following beforehand.

**Lemma 12.15.** *Let  $A, B > 0$  and  $\eta_1, \eta_2 > 0$ . Then we have*

$$(12.25) \quad \int_0^\infty \min(A\ell^{-\eta_1}, B\ell^{\eta_2}) \frac{d\ell}{\ell} \simeq A^{\frac{\eta_2}{\eta_1+\eta_2}} B^{\frac{\eta_1}{\eta_1+\eta_2}}.$$

*Proof of Lemma 12.15.* If we carry out the change of variables

$$(12.26) \quad \frac{A}{B} \ell = \lambda,$$

we obtain

$$\begin{aligned} \int_0^\infty \min(A\ell^{-\eta_1}, B\ell^{\eta_2}) \frac{d\ell}{\ell} &\simeq A^{\frac{\eta_2}{\eta_1+\eta_2}} B^{\frac{\eta_1}{\eta_1+\eta_2}} \int_0^\infty \min \left( \left( \frac{A}{B} \right)^{\frac{\eta_1}{\eta_1+\eta_2}} \ell^{-\eta_1}, \left( \frac{A}{B} \right)^{-\frac{\eta_2}{\eta_1+\eta_2}} \ell^{\eta_2} \right) \frac{d\ell}{\ell} \\ &\simeq A^{\frac{\eta_2}{\eta_1+\eta_2}} B^{\frac{\eta_1}{\eta_1+\eta_2}} \int_0^\infty \min(\lambda^{-\eta_1}, \lambda^{\eta_2}) \frac{d\lambda}{\lambda} \\ &\simeq A^{\frac{\eta_2}{\eta_1+\eta_2}} B^{\frac{\eta_1}{\eta_1+\eta_2}}. \end{aligned}$$

This is the desired result. □

*Proof of Theorem of 12.14.* The proof is simple. The key idea is to apply (12.24) of the left-hand side after writing the integral in question with the distribution formula. Our present assertion (12.24) can be summarized as follows, if we take into account the finiteness of  $X$ .

$$(12.27) \quad \mu \{ |Sf| > \lambda \} \lesssim \min \left\{ \frac{1}{\lambda} \int_X |f(x)| d\mu(x), \mu(X) \right\}.$$

Indeed, if we use (12.27) and Lemma 12.15, we obtain

$$\begin{aligned} \int |Sf(x)|^\delta d\mu(x) &= \int_0^\infty \delta \lambda^{\delta-1} \mu \{ |Sf| > \lambda \} d\lambda \\ &\lesssim_\delta \int_0^\infty \lambda^{\delta-1} \min \left\{ \frac{1}{\lambda} \int_X |f(x)| d\mu(x), \mu(X) \right\} d\lambda. \end{aligned}$$

A direct calculation of the above integral of the most right-hand side shows

$$\int_X |Sf(x)|^\delta d\mu(x) \lesssim_\delta \mu(X)^{1-\delta} \|f\|_{L^1(\mu)}^\delta.$$

This is the desired result.  $\square$

### 13. APPLICATIONS AND RELATED TOPICS

#### 13.1. Density argument.

Here we give an important principle of harmonic analysis. Density argument in functional analysis often means the Banach Steinhaus principle. But in harmonic analysis, the density argument means the following theorem.

**Theorem 13.1.** *Let  $I$  be an open interval containing 0 and let  $(X, \mathcal{B}, \mu)$  be a measure space. Suppose that  $\{T_t\}_{t \in I}$  is a family of the mapping from the set of measurable functions to measurable functions. Let*

$$(13.1) \quad \tilde{T}f(x) := \sup_{t \in I} |T_t f(x)| \quad (x \in X).$$

Assume that the estimate, called the weak- $L^p(\mu)$  estimate,

$$(13.2) \quad \left| \left\{ x \in X : \tilde{T}f(x) > \lambda \right\} \right| \lesssim \frac{1}{\lambda^p} \int_X |f(x)|^p dx$$

holds. Then

$$\begin{aligned} Y &:= \left\{ f \in L^p(\mu) : \lim_{\substack{t \downarrow 0 \\ t \in I}} T_t f(x) = f(x) \quad \mu\text{-a.e. } x \in X \right\} \\ Z &:= \left\{ f \in L^p(\mu) : \lim_{\substack{t \downarrow 0 \\ t \in I}} T_t f(x) \text{ exists } \mu\text{-a.e. } x \in X \right\} \end{aligned}$$

are closed in  $L^p(\mu)$ .

*Proof.* For the proof we take sequence  $\{f_j\}_{j \in \mathbb{N}}$  convergent to  $f$  in the  $L^p(\mu)$ -topology. Then we have to prove that

$$(13.3) \quad \hat{Y} := \left\{ x \in X : \limsup_{\substack{t \downarrow 0 \\ t \in I}} |T_t f(x) - f(x)| > 0 \right\}, \hat{Z} := \left\{ x \in X : \limsup_{\substack{t, s \downarrow 0 \\ t, s \in I}} |T_t f(x) - T_s f(x)| > 0 \right\}$$

have measure 0. By the monotonicity of the measure it suffices to prove

$$(13.4) \quad \hat{Y} := \left\{ x \in X : \limsup_{\substack{t \downarrow 0 \\ t \in I}} |T_t f(x) - f(x)| > \varepsilon \right\}, \hat{Z} := \left\{ x \in X : \limsup_{\substack{t, s \downarrow 0 \\ t, s \in I}} |T_t f(x) - T_s f(x)| > \varepsilon \right\}$$

have measure 0 for all  $\varepsilon > 0$  instead of proving (13.3) directly.



The proofs are similar, so that we will prove that  $\tilde{Y}$  is not charged by  $\mu$ .

Note that  $\left\{x \in X : \limsup_{t \downarrow 0} |T_t f_j(x) - f_j(x)| > \frac{\varepsilon}{2}\right\} = \emptyset$ , because  $f_j \in \tilde{Y}$ . As a consequence we have

$$\begin{aligned} & \left\{x \in X : \limsup_{t \downarrow 0} |T_t f(x) - f(x)| > \varepsilon\right\} \\ & \subset \left\{x \in X : \limsup_{t \downarrow 0} |T_t(f - f_j)(x) - (f - f_j)(x)| > \frac{\varepsilon}{2}\right\}. \end{aligned}$$

If we invoke our assumption (13.2), then we obtain

$$\begin{aligned} & \mu \left\{x \in X : \limsup_{t \rightarrow 0} |T_t f(x) - f(x)| > \varepsilon\right\} \\ & \leq \mu \left\{x \in X : \limsup_{t \rightarrow 0} |T_t(f - f_j)(x) - (f - f_j)(x)| > \frac{\varepsilon}{2}\right\} \\ & \lesssim \frac{1}{\lambda^p} \int_X |f(x) - f_j(x)|^p d\mu(x). \end{aligned}$$

Tending  $j \rightarrow \infty$  in the above inequality, we have

$$(13.5) \quad \mu \left\{x \in X : \limsup_{t \rightarrow 0} |T_t f(x) - f(x)| > \varepsilon\right\} = 0.$$

This is the result that we wish to prove. □

*Exercise 124.* Show that  $\mu(\tilde{Z}) = 0$ , where  $\tilde{Z}$  is given by (13.4).

**13.2. Application to the Lebesgue differentiation theorem.** We prove Lebesgue's differential theorem.

**Theorem 13.2.** *For a locally integrable function  $f$  we have*

$$(13.6) \quad \lim_{r \rightarrow 0} \frac{1}{|B(x, r)|} \int_{B(x, r)} f(y) dy = \lim_{r \rightarrow 0} m_{B(x, r)}(f) = f(x)$$

for almost every  $x \in \mathbb{R}^d$ .

*Exercise 125.* Prove that, if  $f \in C$ , then (13.6) is true.

*Proof.* For the proof we may assume that  $f \in L^1(\mathbb{R}^d)$  by using the truncation. Let

$$(13.7) \quad T_r f(x) = m_{B(x, r)}(f) \quad (r > 0).$$

Since  $M$  is weak-(1, 1) and strong-( $p, p$ ), so is the maximal operator  $\tilde{T}$ . Then we can use the above theorem to conclude that the functions for which (13.6) holds form a closed subspace in  $L^1(\mathbb{R}^d)$ .

Thus, noticing that  $C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$  is dense in  $L^1(\mathbb{R}^d)$ , we have the desired result. □

**Corollary 13.3.** *For a locally integrable function  $f$  we have*

$$(13.8) \quad |f(x)| \leq Mf(x) \quad dx\text{-almost everywhere.}$$

*Proof.* Indeed, we have, for  $x \in \mathbb{R}^d$  satisfying (13.6), that

$$(13.9) \quad |f(x)| = \left| \lim_{r \downarrow 0} m_{B(x, r)}(f) \right| = \lim_{r \downarrow 0} |m_{B(x, r)}(f)| \leq \limsup_{r \downarrow 0} m_{B(x, r)}(|f|) \leq Mf(x).$$

Thus, the proof is complete. □

*Exercise 126.*

- (1) Suppose that we are given a finite collection of cubes  $Q_1, Q_2, \dots, Q_N$ . Then if we relabel them, we can take  $L \leq N$  to arrange that  $Q_1, Q_2, \dots, Q_L$  be disjoint and that

$$\bigcup_{j=1}^N Q_j \subset \bigcup_{j=1}^L 3Q_j, \text{ where } 3Q_j \text{ denotes the triple of } Q_j.$$

- (2) Show that  $|\{Mf > \lambda\}| \leq \frac{3^d}{\lambda} \|f\|_1$  for all  $f \in L^1(\mathbb{R}^d)$ , where  $Mf$  is given by

$$(13.10) \quad Mf(x) = \sup_{x \in Q: Q: \text{cube}} \frac{1}{|Q|} \int_Q |f(y)| dy.$$

- (3) Let  $f \in L^1(\mathbb{R}^d)$ . Then there exists a set  $E$  of measure 0 such that

$$(13.11) \quad \lim_{Q \rightarrow x} \frac{1}{|Q|} \int_Q f(y) dy = f(x)$$

holds for all  $x \in E$ . Here (13.11) means that for all  $\varepsilon > 0$  there exists  $\delta > 0$  such that

$$(13.12) \quad \left| \frac{1}{|Q|} \int_Q f(y) dy - f(x) \right| < \varepsilon$$

whenever  $Q$  is a cube such that  $x \in Q$  and that  $\ell(Q) < \delta$ .

- (4) Extend this result to the case when  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  in general.

*Exercise 127.* Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a locally integrable function. Assume that there exists a constant  $L$  such that

$$(13.13) \quad \int_a^b f(x) dx = L(b-a)$$

for all  $-\infty < a < b < \infty$ . Then show that  $f(x) = L$  for a.e.  $x \in \mathbb{R}$ .

*Exercise 128.* Let  $1 \leq p < \infty$  and  $w, u : \mathbb{R}^d \rightarrow (0, \infty)$  be measurable functions. We define  $L^p(w)$  as follows:

$$(13.14) \quad L^p(w) := \{f : \mathbb{R}^d \rightarrow \mathbb{C} : \|f\|_{L^p(w)} < \infty\},$$

where

$$(13.15) \quad \|f\|_{L^p(w)} := \left( \int_{\mathbb{R}^d} |f(x)|^p w(x) dx \right)^{\frac{1}{p}}.$$

We say that  $L^p(w) \subset L^p(u)$  if

$$(13.16) \quad \|f\|_{L^p(u)} \leq c_0 \|f\|_{L^p(w)}.$$

Show that  $L^p(w) \subset L^p(u)$ , if and only if there exists a constant  $c_1 > 0$  such that  $u(x) \leq c_1 w(x)$  for a.e.  $x \in \mathbb{R}^d$ . Show as well that if this is the case, we can take  $c_1$  so that  $c_0 = c_1^{\frac{1}{p}}$ .

### 13.3. Application to the approximation to the unit.

We also consider an application for the approximation to the unit.

The notation below is valid only in this subsection. But it is frequently used in the book on harmonic analysis.

*Notation.* For a function  $g$  on  $\mathbb{R}^d$ , we set  $g_t(x) := \frac{1}{t^d} g\left(\frac{x}{t}\right)$  for  $t > 0$ .

**Proposition 13.4.** *Suppose that  $\tau$  is a radial positive and integrable function on  $\mathbb{R}^d$ . Assume that  $\rho$  is a function satisfying*

$$(13.17) \quad \rho(|x|) = \tau(x) \text{ for all } x \in \mathbb{R}^d, \rho \text{ is decreasing, } \tau \in L^1(\mathbb{R}^d).$$

*Then we have and for all  $t > 0$  that*

$$(13.18) \quad |\tau_t * f(x)| \leq \|\tau\|_1 Mf(x).$$

*for all  $f \in L^p(\mathbb{R}^d)$ ,  $1 \leq p \leq \infty$ .*

By the monotone convergence theorem we may assume that  $\rho$  is a step function with the same property as (13.17).

In this case we can write  $\rho = \sum_{j=1}^d a_j \cdot \chi_{B(R_j)}$ , where  $a_j \geq 0$  and  $0 < R_1 < \dots < R_n$ . Using this decomposition we have

$$|\tau_t * f(x)| \leq \sum_{j=1}^d \frac{a_j}{t^d} \int_{B(tR_j)} |f(x-y)| dy = \sum_{j=1}^d a_j |B(R_j)| \cdot m_{B(x,tR_j)}(|f|)$$

If we majorize the average by the maximal operator then we obtain

$$|\tau_t * f(x)| \leq \left( \sum_{j=1}^d a_j |B(R_j)| \right) \cdot Mf(x) = \left( \int_{\mathbb{R}^d} \tau(y) dy \right) \cdot Mf(x).$$

Thus we are done.

**Corollary 13.5.** *Under the assumption of the above proposition we have*

$$(13.19) \quad \lim_{t \rightarrow 0} \tau_t * f(x) = \int_{\mathbb{R}^d} \tau(y) dy \cdot f(x)$$

*for almost all  $x \in \mathbb{R}^d$ .*

*Proof.* Apply Theorem 13.1 and Proposition 13.4, keeping in mind that the assertion is trivial if  $f \in \mathcal{S}(\mathbb{R}^d)$ . □

We now reconsider the Fourier transform of  $L^1(\mathbb{R}^d)$  functions.

**Theorem 13.6.** *Assume that  $f \in L^1(\mathbb{R}^d)$  satisfies  $\mathcal{F}f \in L^1(\mathbb{R}^d)$ . Then we have*

$$(13.20) \quad \mathcal{F}^{-1}(\mathcal{F}f)(x) = \mathcal{F}(\mathcal{F}^{-1}f)(x) = f(x)$$

*for almost all  $x \in \mathbb{R}^d$ .*

*Proof.* If  $f \in \mathcal{S}(\mathbb{R}^d)$ , we have proved this assertion. The same proof holds. Reexamine the proof of Theorem 7.9. Notice that the above corollary shows that the operation  $t \rightarrow 0$  in (7.16) is available. □

*Exercise 129.* Let  $1 < p \leq \infty$ . Suppose that  $u : \mathbb{R}_+^{d+1} \rightarrow \mathbb{K}$  is a harmonic extension of  $f \in L^p(\mathbb{R}^d)$ . That is,

$$(13.21) \quad u(x, t) = c_d \int_{\mathbb{R}^d} \frac{t}{(|y|^2 + t^2)^{\frac{d+1}{2}}} f(x-y) dy.$$

Then show that  $u(x, t) \rightarrow f(x)$  for almost every  $x \in \mathbb{R}$ .

Lebesgue point of a function.

**Theorem 13.7.** *Suppose that  $f$  is a locally integrable function on  $\mathbb{R}^d$ . Then we have for almost all  $x \in \mathbb{R}^d$ .*

$$(13.22) \quad \lim_{r \rightarrow 0} \frac{1}{r^d} \int_{B(x,r)} |f(x) - f(y)| dy = 0.$$

*Proof.* It is the same as Theorem 13.2. Recall the proof of Theorem 13.2. The same reasoning can be used.  $\square$

**Definition 13.8.** A point for which (13.22) holds are called Lebesgue point of the function.

*Exercise 130.* Show that the assumption that  $f$  is locally integrable is absolutely necessary by using Example 3.61.

Finally having proved the differentiation theorem, we prove (4.87), which was left open in Chapter 3. Let us recall (4.87).

**Proposition 13.9.** *Suppose that  $\varphi$  is a positive smooth function satisfying*

$$(13.23) \quad \text{supp}(\varphi) \subset B(1), \quad \int \varphi(y) dy = 1.$$

*Then, for all locally integrable functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  we have*

$$(13.24) \quad \varphi_t * f(x) = \frac{1}{t^d} \int_{\mathbb{R}^d} \varphi\left(\frac{y}{t}\right) f(x-y) dy \rightarrow f(x) \text{ as } t \downarrow 0$$

*for any Lebesgue point in  $x \in \mathbb{R}^d$ . In particular, given a locally integrable function  $f$ , (13.24) holds for almost all  $x \in \mathbb{R}^d$ .*

*Proof.* Indeed,

$$|\varphi_t * f(x) - f(x)| \leq \frac{1}{t^d} \int_{\mathbb{R}^d} \varphi\left(\frac{y}{t}\right) |f(x) - f(x-y)| dy \leq \frac{\|\varphi\|_\infty}{t^d} \int_{B(x,t)} |f(x) - f(x-y)| dy \rightarrow 0$$

as  $t \downarrow 0$ .  $\square$

Harmonic extension and the non-tangential limit. Let us now consider the Poisson equation on the half space  $\mathbb{R}_+^d$ .

**Proposition 13.10.** *Let  $\Gamma(x, t) := \frac{c_d t}{(|x|^2 + t^2)^{\frac{n+1}{2}}}$ , where  $c_d$  is a normalization constant taken*

*so that*

$$(13.25) \quad \int_{\mathbb{R}^d} \Gamma(x, t) dx = 1$$

*for all  $t > 0$ . Given a function  $f : \mathbb{R}^d \rightarrow \mathbb{C}$ , we define*

$$(13.26) \quad u(x, t) = u_f(x, t) = \Gamma(*, t) * f(x) \quad x \in \mathbb{R}^d, \quad t > 0$$

*as long as it makes sense. Establish the following.*

- (1) *Let  $1 \leq p \leq \infty$ . Then  $\|u\|_{L^p(\mathbb{R}^{d+1})} \leq \|f\|_{L^p(\mathbb{R}^d)}$ .*
- (2) *Show that  $u_{u_f(*,t)}(x, s) = u_f(x, t+s)$   $t, s > 0$ ,  $x \in \mathbb{R}^d$ .*
- (3) *Let  $1 \leq p < \infty$ . Show that, for every  $f \in L^p(\mathbb{R}^d)$ ,  $\lim_{t \rightarrow 0} u(\cdot, t) = f$  in  $L^p(\mathbb{R}^d)$ .*
- (4) *Show that, for every  $f \in \text{BC}(\mathbb{R}^d)$ ,  $\lim_{t \rightarrow 0} u(*, t) = f$  in  $\text{BC}(\mathbb{R}^d)$ .*
- (5) *Let  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p \leq \infty$ . Show that  $u \in C^\infty(\mathbb{R}_+^{d+1})$ .*
- (6)  *$\Delta u(x, t) = 0$  on  $\mathbb{R}_+^{d+1}$ . Here  $\Delta$  is a Laplacian on  $\mathbb{R}_+^{d+1}$ .*

13.4. Dyadic maximal operator.

The CZ-decomposition is one of the key tools used in this book throughout.

Dyadic maximal operator.

In this subsection we define a dyadic maximal operator to investigate singular integral operators. To define a dyadic maximal operator, we define dyadic cubes. Dyadic cube plays an important role in harmonic analysis.

**Definition 13.11.** One defines the set of all dyadic cubes as

$$(13.27) \quad \mathcal{Q}^d = \left\{ Q \subset \mathbb{R}^d \mid Q = \left[ \frac{k_1}{2^j}, \frac{k_1+1}{2^j} \right) \times \dots \times \left[ \frac{k_n}{2^j}, \frac{k_n+1}{2^j} \right), k_1, \dots, k_n \in \mathbb{Z}, j \in \mathbb{Z} \right\}.$$

It will be helpful to define the  $j$ -th generation of them.

$$(13.28) \quad \mathcal{Q}_j^d = \left\{ Q \subset \mathbb{R}^d \mid Q = \left[ \frac{k_1}{2^j}, \frac{k_1+1}{2^j} \right) \times \dots \times \left[ \frac{k_n}{2^j}, \frac{k_n+1}{2^j} \right), k_1, \dots, k_n \in \mathbb{Z}, \right\}.$$

*Exercise 131.* Let  $d = 1$ . Then display the family  $\mathcal{Q}_0^d$  in the real line  $\mathbb{R}$ .

The following lemma is almost trivial. But it is important.

**Lemma 13.12.** *Suppose that  $Q, R \in \mathcal{Q}^d$  meet. Then we have either  $Q \subset R$  or  $R \subset Q$ .*

*Exercise 132.* Below we shall sketch the proof of Lemma 13.12. Supply its details for  $d = 1$ . Symmetry reduces the matter to two cases. Furthermore we can reduce one dimensional case.

Case 1 :  $l(Q) = l(R)$ . In this case we have  $Q = R$ .

Case 2 :  $l(Q) < l(R)$ . Then we have  $Q \subsetneq R$ .

We define a dyadic maximal operator.

**Definition 13.13.** For a locally integrable function  $f$ , we set

$$E_j f(x) := E_j[f](x) := \sum_{Q \in \mathcal{Q}_j^d} m_Q(f) \cdot \chi_Q(x)$$

$$M_{\text{dyadic}} f(x) := \sup_{j \in \mathbb{Z}} E_j[|f|](x)$$

for  $x \in \mathbb{R}^d$ . That is for a locally integrable function  $E_j f(x)$  is defined as a mean of  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  over  $Q \in \mathcal{Q}_j^d$  containing  $x$ .

**Lemma 13.14.** *Suppose that  $f$  is an integrable function. Then we have*

$$(13.29) \quad \lim_{j \rightarrow \infty} E_j f(x) = 0.$$

*Proof.* Indeed,

$$(13.30) \quad |E_j f(x)| \leq \frac{1}{2^{jd}} \int_{\mathbb{R}^d} |f(y)| dy$$

for all  $j \in \mathbb{Z}$  and  $x \in \mathbb{R}^d$ . □

**Theorem 13.15.** *Let  $f$  be a locally integrable function. Take  $\lambda > 0$ . Then there are disjoint dyadic cubes  $Q_j$  such that*

$$(13.31) \quad \{x \in \mathbb{R}^d : M_{\text{dyadic}} f(x) > \lambda\} = \bigcup_j Q_j, \quad \lambda \leq m_{Q_j}(|f|) \leq 2^d \lambda.$$

*Proof.* For the proof we set  $E_\lambda = \{M_{\text{dyadic}}f > \lambda\}$ . By definition of the dyadic maximal operator we have  $E_\lambda = \bigcup_{j \in \mathbb{Z}} \{E_j[|f|] > \lambda\}$ .

Note that there exists  $j^*$  so that  $E_j[|f|](x) \leq \lambda$  for all  $j \leq j^*$  by virtue of Lemma 13.14. Thus, we can partition  $E_\lambda$  into a disjoint union:

$$(13.32) \quad E_\lambda = \bigcup_{j \in \mathbb{Z}} \{x \in \mathbb{R}^d : E_j[|f|](x) > \lambda, E_k[|f|](x) \leq \lambda \text{ for all } k < j\}.$$

By Lemma 13.12 we can partition

$$(13.33) \quad \{x \in \mathbb{R}^d : E_j[|f|](x) > \lambda, E_k[|f|](x) \leq \lambda \text{ for all } k < j\}$$

into a union of  $Q_m^j \in \mathcal{Q}_j$  with  $m = 1, 2, \dots, j_M$ . If we rearrange  $Q_m^j$ 's, we can obtain a partition of  $E_\lambda$ .

$$E_\lambda = \sum_j Q_j.$$

If  $x \in Q_j = Q_m^{j_0}$ , then we have

$$\frac{1}{|Q_j|} \int_{Q_j} |f(y)| dy = E_{j_0}|f|(x) > \lambda.$$

and denoting  $Q_j^* \in \mathcal{Q}_{j_0-1}$  by the unique dyadic cube containing  $Q_j$

$$\frac{1}{|Q_j^*|} \int_{Q_j^*} |f(y)| dy = E_{j_0-1}|f|(x) \leq \lambda.$$

The last formula implies that

$$\frac{1}{2^d |Q_j|} \int_{Q_j} |f(y)| dy \leq m_{Q_j^*}(|f|) = E_{j_0-1}|f|(x) \leq \lambda.$$

Thus we have the desired results.  $\square$

Calderón-Zygmund decomposition.

Now we turn to the Calderón-Zygmund decomposition. We keep to the setting of  $\mathbb{R}^d$ .

**Theorem 13.16** ((dyadic) Calderón-Zygmund decomposition). *Given an  $L^1(\mathbb{R}^d)$ -function  $f$  and a positive constant  $\lambda$ , there exist dyadic cubes  $Q_j$  and  $L^1(\mathbb{R}^d)$ -functions  $g, \{b_j\}_j$  with the following properties.*

(1) [Decomposition]

$$(13.34) \quad f = g + \sum_j b_j.$$

(2) [Good part]

$$(13.35) \quad (L^\infty\text{-estimate}) \quad |g(x)| \leq 2^d \lambda$$

$$(13.36) \quad (L^1\text{-estimate}) \quad \int_{\mathbb{R}^d} |g(x)| dx \leq \int_{\mathbb{R}^d} |f(x)| dx.$$

(3) [Bad part]

$$(13.37) \quad \text{moment condition} \quad \int_{\mathbb{R}^d} b_j(y) dy = 0.$$

(4) [Support condition]

$$(13.38) \quad \text{supp}(b_j) \subset Q_j \text{ for all } j.$$

*Proof.* For the proof note that in the previous section we have taken a family of disjoint dyadic cubes  $\{Q_j\}_j$  with

$$(13.39) \quad \lambda \leq \frac{1}{|Q_j|} \int_{Q_j} |f(y)| dy \leq 2^d \lambda$$

and

$$\{x \in \mathbb{R}^d : M_{\text{dyadic}} f(x) > \lambda\} = \bigcup_j Q_j.$$

We shall construct  $b_j$  as follows: We put

$$(13.40) \quad b_j(x) := \chi_{Q_j}(x) \left( f(x) - \frac{1}{|Q_j|} \int_{Q_j} f(y) dy \right).$$

Once we have defined the  $b_j$ ,  $g$  is determined automatically by

$$g := f - \sum_j b_j.$$

By the definitions of  $g$  and  $b_j$ , (13.36) follows. All we have to verify is the formula (13.35).

Suppose that  $x \in Q_j$ . In this case we have, by disjointness of  $\{Q_j\}_j$ ,

$$(13.41) \quad g(x) = \frac{1}{|Q_j|} \int_{Q_j} f(y) dy.$$

By (13.39) we have  $|g(x)| \leq 2^d \lambda$ .

Suppose instead that  $x \notin \bigcup_j Q_j$ . In this case by Corollary 13.3 we have

$$(13.42) \quad |g(x)| = |f(x)| \leq M_{\text{dyadic}} f(x) \leq \lambda.$$

Thus all the requirements were verified. □

In the next corollary, we verify that the integrability property of  $f$  carries over to that of  $b_j$ .

**Corollary 13.17.** *Assume in addition that  $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ . Then each  $b_j$  belongs to  $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  and the convergence*

$$(13.43) \quad f = g + \sum_j b_j$$

*takes place in  $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ .*

*Proof.* The good part  $g$  belongs to  $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and hence  $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ . Once we assume  $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ , disjointness of the support of the  $b_j$  gives us  $b_j \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  for each  $j$  and the desired convergence. □

We also have the local version of the above theorem. The construction being identical, we leave the proof for the readers.

*Exercise 133.* Suppose that  $Q$  is a cube. Bisect  $Q$  to obtain  $2^d$  cubes. Let us name them  $Q_1^{(1)}, Q_2^{(1)}, \dots, Q_{2^d}^{(1)}$ . Set

$$(13.44) \quad \mathcal{Q}^{(1)} = \{Q_j^{(1)} : 1 \leq j \leq 2^d\}.$$

Now bisect each cube  $Q \in \mathcal{Q}^{(1)}$  to obtain  $Q_1^{(2)}, Q_2^{(2)}, \dots, Q_{2^{2d}}^{(2)}$ . Set

$$(13.45) \quad \mathcal{Q}^{(2)} := \{Q_j^{(2)} : 1 \leq j \leq 2^{2d}\}.$$

Repeat this procedure and then at the  $k$ -th stage we obtain  $2^{kd}$  cubes:

$$(13.46) \quad Q_1^{(k)}, Q_2^{(k)}, \dots, Q_{2^{kd}}^{(k)}$$

Set, as before,

$$(13.47) \quad \mathcal{Q}^{(j)} := \{Q_j^{(k)} : 1 \leq j \leq 2^{kd}\}.$$

(1) We define the corresponding maximal operator

$$(13.48) \quad M_{Q, \text{dyadic}} f(x) = \sup_{j \in \mathbb{N}} E_j[|f|](x),$$

where  $f$  is an integrable function on  $Q$ . Show that

$$(13.49) \quad |\{x \in Q : M_{Q, \text{dyadic}} f(x) > \lambda\}| \leq \frac{1}{\lambda} \int_Q |f(x)| dx.$$

(2) Mimic the proof of Theorem 13.16 to state and prove the analogue for  $Q$ .

(3) (After reading Section 23.) Given an integrable function  $f$  on  $Q$ , we define

$$(13.50) \quad E_j f(x) = \sum_{Q \in \mathcal{Q}^{(j)}} m_Q(f) \cdot \chi_Q.$$

Describe  $E_j$  in terms of the conditional expectation of the probability space  $\left(Q, \frac{dx}{|Q|}\right)$ .

**Remark 13.18.** Observe that the CZ-theory works if we replace  $dx$  with a Radon measure  $\mu$  satisfying the doubling condition

$$(13.51) \quad \mu(2Q) \lesssim \mu(Q)$$

for all cubes  $Q$  with positive  $\mu$ -measure.

### 13.5. Other covering lemmas and some related exercises.

When the underlying measure is the Lebesgue measure, then analysis of maximal operator is not so hard as we saw before. Several generalizations are also available, where the situations are close to the model case of the Lebesgue measure.

So what happens, for example,

- (1) on the graph ?
- (2) on the Gauss measure setting ?

Here the Gauss measure is the weighted measure given by  $d\gamma = \exp(-\pi|x|^2) dx$ .

In general, when we consider such a thing, we need a covering lemma or stopping time. Here, we employ a covering lemma.

A covering lemma is a tool with which to cover a set in an efficient way.

Here the word “efficient” varies according to the situation. In this section we state covering lemmas used in harmonic analysis.

Besicovitch’s covering lemma. The results needed later is the following, called Besicovitch’s covering lemma.

**Theorem 13.19** (Besicovitch). *Let  $\{B_\lambda\}_{\lambda \in \Lambda}$  be a family of balls. Suppose the diameters of balls  $\{B_\lambda\}_{\lambda \in \Lambda}$  are bounded.*

$$(13.52) \quad R = \sup_{\lambda \in \Lambda} r(B_\lambda)$$



Then there exists an integer  $N$  depending only on the dimension that has the following property:

We can pick  $N$  disjoint subfamilies

$$(13.53) \quad \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_N$$

so that,  $A$ , the set of centers of balls  $\{B_\lambda\}_{\lambda \in \Lambda}$  belong to a ball in some  $\mathcal{G}_j$ .

*Proof.* We may assume that  $A$  is bounded. We use the  $5r$ -covering lemma to the family

$$(13.54) \quad \Lambda_0^* := \left\{ \lambda : \lambda \in \Lambda, r(B_\lambda) > \frac{9R}{10} \right\}.$$

Then we obtain a countable family of balls  $\{B_\lambda\}_{\lambda \in \Lambda_0}$  satisfying two conditions below:

- (1)  $\bigcup_{\lambda \in \Lambda_0^*} \frac{1}{5}B_\lambda \subset \bigcup_{\lambda \in \Lambda_0} B_\lambda$ ,
- (2)  $\left\{ \frac{1}{5}B_\lambda \right\}_{\lambda \in \Lambda_0}$  is disjoint.

Since we assume  $A$  is bounded,  $\#\Lambda_0$  is finite. Label and arrange  $\{B_\lambda\}_{\lambda \in \Lambda_0}$  in numerical order:

Then we obtain a subfamily  $B_1, B_2, \dots, B_{N(1)}$  of  $\{B_\lambda\}_{\lambda \in \Lambda}$  such that  $\bigcup_{\lambda \in \Lambda_0^*} \frac{1}{5}B_\lambda \subset \bigcup_{j=1}^{N(1)} B_j$  and

that  $\left\{ \frac{1}{5}B_j \right\}_{j=1}^{N(1)}$  is disjoint. This is the first generation of the balls.

Let  $A_1 = A \setminus \bigcup_{j=1}^{N(1)} B_j$  and  $\Lambda_1^{**} := \{\lambda \in \Lambda : c(B_\lambda) \in A_1\}$ . Observe  $\sup_{\lambda \in \Lambda_1^{**}} r(B_\lambda) \leq \frac{9R}{10}$ . There-

fore, we are in the position of applying the above observation to  $\{B_\lambda\}_{\lambda \in \Lambda_1^{**}}$ . As a consequence we can find  $B_{N(1)+1}, B_{N(1)+2}, \dots, B_{N(2)}$  from  $\{B_\lambda\}_{\lambda \in \Lambda_1^{**}}$  with the following properties.

- (1) If  $\lambda \in \Lambda$  satisfy  $r(B_\lambda) \geq \frac{81R}{100}$ , then there exists  $\rho \in \Lambda_0 \cup \Lambda_1$  so that  $c(B_\lambda) \in B_\rho$ ,
- (2)  $\frac{1}{5}B_{N(1)+1}, \frac{1}{5}B_{N(1)+2}, \dots, \frac{1}{5}B_{N(2)}$  is disjoint.

Note that  $\frac{1}{5}B_j$  and  $\frac{1}{5}B_k$  are disjoint, if  $j \in \Lambda_0$  and  $k \in \Lambda_1$ . Indeed,  $c(B_k)$ , the center of  $B_k$ , does not belong to  $B_j$  and  $\frac{81}{100}r(B_j) \leq r(B_k) \leq \frac{100}{81}r(B_j)$ . This is the second generation of the balls.

Repeating this procedure we can find a sequence of balls  $\{B_j\}_{j \in J}$  from  $\{B_\lambda\}_{\lambda \in \Lambda}$  so that it satisfies the following properties.

- (1)  $J = \{1, 2, \dots, j_0\}$  or  $J = \mathbb{N}$ .
- (2) If  $i, j \in J$  satisfies  $i < j$ , then  $r(B_j) \leq \frac{10}{9}r(B_i)$ .
- (3)  $A \subset \bigcup_{j \in J} B_j$ .
- (4)  $\lim_{j \rightarrow \infty} r(B_j) = 0$ .
- (5)  $\left\{ \frac{1}{5}B_j \right\}_{j \in J}$  is disjoint.

Let  $k \in \mathbb{N}$ . Then we set

$$(13.55) \quad I_k := \{j \in \mathbb{N} : j < k, B_j \cap B_k \neq \emptyset\}, K_k := \{j \in I_j : r_j \leq 3r_k\}.$$

Let  $j \in K_k$ . Then we have  $\frac{9}{10}r_k \leq r_j \leq 3r_k$ . Since  $\left\{\frac{1}{5}B_j\right\}_{j \in J}$  is disjoint, it follows that  $K_k \leq c_d$ , where  $c_d$  is a constant depending only on the dimension  $d$ .

Let  $j_1, j_2 \in I_k \setminus J_k$ . If the generation of  $B_{j_1}$  and  $B_{j_2}$  are different, then  $B_{j_1}$  and  $B_{j_2}$  meet in a point and  $c(B_j)$  lies sufficiently close to the boundary of  $B_{j_1}$  and  $B_{j_2}$ . A geometric observation therefore shows the angle  $\angle c(B_{j_1})c(B_j)c(B_{j_2}) \geq \theta_0 > 0$ , where  $\theta_0$  is an absolute constant independent of the family and even of  $d$ !

For each  $j \in I_k \setminus J_k$  consider a point at which  $B_j$  and the segment  $\overline{c(B_j)c(B_k)}$  meet. Then the observation of the above paragraph says that such intersection points are torn apart. Even from this observation, we see the number of generations to which balls in  $I_k \setminus J_k$  are bounded by a number depending on  $d$ .

Of course for each generation the number of balls that can intersect  $B_k$  is also bounded by a number depending on  $d$ . Therefore  $I_k \setminus J_k$  is majorized in number by a constant  $c_d$ .

In view of this observation we conclude that the number of  $I_k$  is bounded by a constant independent of  $k$ , say  $M_d$ .

We define a mapping  $\sigma : \mathbb{N} \rightarrow \{1, 2, \dots, M_d + 1\}$  as follows: First we set  $\sigma(j) = j$  for  $j \leq M_d + 1$ . If  $j \geq M_d + 2$ , then we define inductively

$$(13.56) \quad \sigma(j) := \min\{l = 1, 2, \dots, M_d + 1 : B_j \cap B_m = \emptyset \text{ for all } m \in \sigma^{-1}(l) \text{ with } m \leq j - 1\}.$$

We remark that by the pigeon hole principle the set appearing in the above formula is not empty. Thus, we have  $\{\{B_j\}_{j \in \sigma^{-1}(l)}\}_{l=1}^{M_d+1}$  is the desired family.  $\square$

*Exercise 134.* Let  $\mu$  be a Radon measure. We set

$$(13.57) \quad M_\mu f(x) := \sup_{r>0} \frac{1}{\mu(B(x, r))} \int_{B(x, r)} |f(y)| d\mu(y).$$

Prove that

$$(13.58) \quad \mu\{M_\mu f > \lambda\} \leq \frac{N}{\lambda} \int_{\mathbb{R}^d} |f(x)| d\mu(x),$$

where  $N$  is a covering constant in Theorem 13.19.

*Exercise 135* (After solving Exercise 134). Generalize (13.6) to  $L^1_{\text{loc}}(\mu)$ .

Modified Hardy-Littlewood maximal operator. The covering lemma below can be found in [424], whose proof is omitted.

**Theorem 13.20.** *Let  $(X, d)$  be a metric space. For all  $k > 1$  there exists an integer  $N = N_k$ , depending only on the dimension and  $k$ , that satisfies the following:*

*Let  $\{B(x_\lambda, r_\lambda)\}_{\lambda \in L}$  be a family of balls in Euclidean space. Suppose that*

$$(13.59) \quad \sup_{\lambda \in L} r_\lambda < \infty.$$

*Then we can take disjoint subfamilies*

$$(13.60) \quad \{B(x_\rho, r_\rho)\}_{\rho \in L^1}, \{B(x_\rho, r_\rho)\}_{\rho \in L^2}, \dots, \{B(x_\rho, r_\rho)\}_{\rho \in L^N}$$

*such that  $\bigcup_{\lambda \in L} B(x_\lambda, r_\lambda) \subset \bigcup_{j=1, \dots, N} \bigcup_{\rho \in L^j} B(x_\rho, kr_\rho)$ .*

*Exercise 136.* Admit Theorem 13.20. Suppose that we are given a Radon measure  $\mu$ . Write  $\mathcal{B}_x$  for the set of all open balls containing  $x$ . Let  $k \geq 1$  and define

$$(13.61) \quad M'_{k,\mu} f(x) := \sup_{B=B(y,r) \in \mathcal{B}_x} \frac{1}{\mu(kB)} \int_{B(y,r)} |f(z)| d\mu(z).$$

Prove that, for  $k > 1$ ,

$$(13.62) \quad \mu\{M'_{k,\mu} f > \lambda\} \leq \frac{N_k}{\lambda} \int_{\mathbb{R}^d} |f(z)| d\mu(z).$$

In particular if  $\mu$  satisfies the doubling condition, that is, there exists a constant  $c_0 > 0$  so that

$$(13.63) \quad \mu(B(x, 2r)) \leq c_0 \mu(B(x, r)),$$

then we have

$$(13.64) \quad \mu\{M'_{1,\mu} f > \lambda\} \leq \frac{c_0 N}{\lambda} \int_{\mathbb{R}^d} |f(z)| d\mu(z).$$

Remark that the doubling condition was proposed initially by Coifman and Weiss [147].

1-dimensional case. In  $\mathbb{R}$  we have a following wonderful covering lemma. Denote by  $\mathcal{I}(\mathbb{R})$  the set of all closed intervals in  $\mathbb{R}$ .

**Theorem 13.21** (1-dimensional case). *Suppose that  $\{I_j\}_{j \in J}$  is a family in  $\mathcal{I}(\mathbb{R})$ . Then there exist  $J_1, J_2 \subset J$  such that*

$$(13.65) \quad \bigcup_{j \in J} I_j \subset \bigcup_{j \in J_1 \cup J_2} I_j$$

and that  $\{I_j\}_{j \in J_k}$  is disjoint ( $k = 1, 2$ ).

*Exercise 137.* Imagine that we are given three intervals meeting at one point. Observe that one of them is already covered by the remaining two intervals, so that we can delete it. With this in mind, prove Theorem 13.21.

*Exercise 138.* Now suppose that we are given a Radon measure  $\mu$  on  $\mathbb{R}$ . Write  $\mathcal{I}_t(\mathbb{R})$  for the set of all open intervals containing  $t \in \mathbb{R}$ . Define the uncentered Hardy-Littlewood maximal operator associated with  $\mu$  by

$$(13.66) \quad M_\mu f(t) := \sup_{I \in \mathcal{I}_t(\mathbb{R})} \frac{1}{\mu(I)} \int_I |f(s)| d\mu(s).$$

Keeping Theorem 13.21 in mind, prove that

$$(13.67) \quad \mu\{M_\mu f > \lambda\} \leq \frac{2}{\lambda} \int_{\mathbb{R}} |f(s)| d\mu(s).$$

*Exercise 139.* Assume further that  $\mu = dx$ . By considering the sequence  $\{\chi_{(0, k^{-1})}\}_{k \in \mathbb{N}}$  prove that we cannot replace 2 in (13.67) by any number strictly less than 2.

From this example, we see that  $d = 1$  is a special case in harmonic analysis.

*Exercise 140.* Returning to the setting of  $\mathbb{R}^d$  coming with a Radon measure  $\mu = e^{x^2} dx$ . Show that  $N_k$  in (13.62) blows up (tends to  $\infty$ ) as  $k \downarrow 1$ .

5r-covering lemma revisited. Next we place ourselves in the setting of a separable metric measure space  $(X, d)$ . Recall that a metric measure space  $(X, d)$  is said to be separable, if there exists a countable dense subset  $X_0$ .

**Lemma 13.22** (A variant of 5r-covering lemma). *Suppose we are given a family of  $N$  balls  $\{B(x_j, r_j)\}_{j=1}^N$  in a metric space  $(X, d)$ . Then we can extract a subfamily  $\{B(x_j, r_j)\}_{j \in A}$  such that*

- (1)  $\{B(x_j, r_j)\}_{j \in A}$  is disjoint.
- (2)  $\bigcup_{j=1}^N B(x_j, r_j) \subset \bigcup_{j \in A} B(x_j, 3r_j)$ .

*Exercise 141.* Prove Lemma 13.22 by induction on  $\#A$ .

*Exercise 142.* Envisage three situations in which we are given a metric measure space  $(X, d)$  coming with a Radon measure  $\mu$ .

Vitali's covering lemma. We now take up Vitali's covering lemma, which will be used to investigate that an increasing function on  $\mathbb{R}$  is differentiable almost everywhere.

**Definition 13.23** (Vitali covering). Let  $A$  be a set. A family of closed balls  $\{\overline{B}_\lambda\}_{\lambda \in \Lambda}$  is said to be a Vitali covering. For all  $x \in A$  and  $\varepsilon > 0$ , there exists  $\lambda \in \Lambda$  such that  $x \in \overline{B}_\lambda$  and  $\text{diam}(\overline{B}_\lambda) < \varepsilon$ .

**Theorem 13.24** (Vitali). *Let  $A$  be a measurable set and  $\{\overline{B}_\lambda\}_{\lambda \in \Lambda}$  be its Vitali covering. Then there exists a disjoint subfamily  $\{\overline{B}_\lambda\}_{\lambda \in \Lambda_0}$  such that*

$$(13.68) \quad \left| A \setminus \bigcup_{\lambda \in \Lambda_0} \overline{B}_\lambda \right| = 0.$$

*Proof.* We can assume that  $A$  is bounded and that  $\sup_{\lambda \in \Lambda} \text{diam}(\overline{B}_\lambda) \leq 1$ .

Let  $\lambda_1 \in \Lambda$  be taken so that

$$(13.69) \quad \text{diam}(\overline{B}_{\lambda_1}) > \frac{1}{2} \sup_{\lambda \in \Lambda} \text{diam}(\overline{B}_\lambda).$$

Suppose we have selected  $\lambda_1, \lambda_2, \dots, \lambda_k$  so that  $\overline{B}_{\lambda_1}, \overline{B}_{\lambda_2}, \dots, \overline{B}_{\lambda_k}$  are disjoint. Then choose  $\lambda_{k+1}$  so that it fulfills the following conditions.

- (1)  $\overline{B}_\lambda$  meets  $A$ .
- (2)  $\overline{B}_{\lambda_1}, \overline{B}_{\lambda_2}, \dots, \overline{B}_{\lambda_{k+1}}$  are disjoint.
- (3) If  $\overline{B}_\lambda$  intersects none of  $\overline{B}_{\lambda_1}, \overline{B}_{\lambda_2}, \dots, \overline{B}_{\lambda_k}$ , then

$$\text{diam}(\overline{B}_\lambda) \leq 2 \text{diam}(\overline{B}_{\lambda_{k+1}}).$$

Stop this procedure when the family  $\overline{B}_{\lambda_1}, \overline{B}_{\lambda_2}, \dots, \overline{B}_{\lambda_k}$  engulfs  $A$ . In this way inductively we have obtained a family of balls  $\overline{B}_{\lambda_1}, \overline{B}_{\lambda_2}, \dots, \overline{B}_{\lambda_J}$ .

The proof is complete once we prove

$$(13.70) \quad A \setminus \bigcup_{j=1}^N \overline{B}_{\lambda_j} \subset \bigcup_{j=N+1}^J 5\overline{B}_{\lambda_j}$$

for all  $N \in \mathbb{N}$ . Because we assume that  $A$  is bounded and each ball of the subfamily meets  $A$ . Therefore, we obtain

$$(13.71) \quad \lim_{N \rightarrow \infty} \left| \bigcup_{j=N+1}^J 5\bar{B}_{\lambda_j} \right| \rightarrow 0,$$

proving  $\left| A \setminus \bigcup_{j=1}^{\infty} \bar{B}_{\lambda_j} \right| = 0$ .

Suppose that  $x \in A \setminus \bigcup_{j=1}^N \bar{B}_{\lambda_j}$ . Then there exists  $\lambda \in \Lambda$  such that  $x \in \bar{B}_{\lambda}$  and

$$(13.72) \quad \text{diam}(\bar{B}_{\lambda}) < \min(\text{diam}(\bar{B}_{\lambda_1}), \text{diam}(\bar{B}_{\lambda_2}), \dots, \text{diam}(\bar{B}_{\lambda_N}))$$

Since the subfamily is disjoint, there exists  $j \geq N + 1$  so that

$$(13.73) \quad \text{diam}(\bar{B}_{\lambda_j}) < \frac{1}{4} \text{diam}(\bar{B}_{\lambda}).$$

This implies  $\bar{B}_{\lambda}$  intersects some ball  $\bar{B}_{\lambda_k}$  with  $\text{diam}(\bar{B}_{\lambda}) < \frac{1}{2} \text{diam}(\bar{B}_{\lambda_k})$ . A geometric observation shows that  $\bar{B}_{\lambda} \subset \bar{B}_{\lambda_k}$ . □

Linderöf’s covering lemma. Next we take up Linderöf’s covering lemma.

**Theorem 13.25** (Linderöf). *Suppose that  $\{B_{\lambda}\}_{\lambda \in \Lambda}$  is a family of balls in a separable metric space  $(X, d)$ . Then we have the union is already covered by a countable subfamily.*

*Proof.* Suppose that  $X_0$  is a countable dense subset. Let  $x \in \bigcup_{\lambda \in \Lambda} B_{\lambda}$ . Then we can choose  $r_x > 0$  so that

$$(13.74) \quad x \in B(x, 2r_x) \subset B_{\lambda}.$$

Since  $X_0$  is dense in  $X$ , we can pick  $p_x \in X_0$  so that  $d(x, p_x) < r_x$ . Choose a rational number  $q_x \in \mathbb{Q}$  between  $d(x, p_x)$  and  $r_x$ . Then we have

$$(13.75) \quad x \in B(x, q_x) \subset B(x, 2r_x) \subset B_{\lambda}.$$

Given  $x_0$  and  $q \in \mathbb{Q}$ , choose  $\lambda(x_0, q)$  so that  $B(x_0, q) \subset B_{\lambda(x_0, q)}$ , if it is possible. Collect  $B_{\lambda}$  such that  $\lambda = \lambda(x_0, q)$  for some  $x_0 \in X_0$  and  $q \in \mathbb{Q}$ . This is the desired family. □

Whitney covering lemma. Suppose that  $F$  is a non-empty closed set and  $O = F^c$  is its complement. We can “cover”  $O$  by a collection of balls that are essentially disjoint, and whose sizes are comparable to their distances from the set  $F$ .

In the general setting we consider here, things are not quite as elegant and we need to begin by fixing a pair of positive constants  $c^*$  and  $c^{**}$  (with  $1 < c^* < c^{**}$ ), which will not depend on the particular set  $F$  under consideration. Using them, for any ball  $B$  we define the balls  $B^* := c^* B$  and  $B^{**} = c^{**} B$ .

**Theorem 13.26** (Whitney covering lemma). *Given  $F$ , a closed non-empty set, there exists a collection of balls*

$$(13.76) \quad B_1, B_2, \dots, B_k, \dots$$

so that

$$(13.77) \quad \chi_O \leq \sum_{j=1}^{\infty} \chi_{B_j} \leq \sum_{j=1}^{\infty} \chi_{B_j^*} \leq N \chi_O, \quad B_j^{**} \cap F \neq \emptyset \text{ for all } j \in \mathbb{N}.$$

*Proof.* Let us consider  $B_x = B(x, \sqrt{c^* c^{**}} \text{dist}(x, F))$  for each  $x \in O$ . Set  $\mathcal{B} = \{B_x\}_{x \in O}$ . Choose a maximal collection  $\{B_j\}_{j \in \mathbb{N}} \subset \mathcal{B}$  satisfying

$$(13.78) \quad \frac{1}{3} B_j \cap \frac{1}{3} B_k = \emptyset.$$

We claim that

$$(13.79) \quad \chi_O \leq \sum_{j=1}^{\infty} \chi_{B_j}.$$

Indeed, let  $x \in O$ . Then there exists  $j \in \mathbb{N}$  such that  $\delta B_x \cap \delta B_j \neq \emptyset$  from the maximality. In this case, we have

$$(13.80) \quad \text{dist}(x, F) + \frac{1}{3} \text{dist}(x, F) + \frac{1}{3} \text{dist}(c(B_j), F) < \text{dist}(c(B_j), F),$$

which gives us

$$(13.81) \quad \text{dist}(x, F) < 2 \text{dist}(c(B_j), F).$$

We also have

$$(13.82) \quad |x - c(B_j)| < \frac{1}{3} (\text{dist}(x, F) + \text{dist}(c(B_j), F)) < \text{dist}(c(B_j), F).$$

Meanwhile the pointwise estimate

$$(13.83) \quad \sum_{j=1}^{\infty} \chi_{B_j^*} \leq N \chi_O$$

is a direct consequence that  $\text{dist}(B_j, F) \simeq c(B_j)$ .  $\square$

*Exercise 143.* Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . Assume that  $\Omega$  is a proper subset of  $\mathbb{R}^d$ . Show that there exists a smooth function  $D : \Omega \rightarrow \mathbb{R}$  such that

$$(13.84) \quad D(x) \simeq d(x), \quad \sup_{x \in \Omega} d(x)^{|\alpha|-1} |\partial^\alpha D(x)| < \infty$$

for all  $\alpha \in \mathbb{N}_0^d$ , where  $d(x) = \text{dist}(x, \mathbb{R}^d \setminus \Omega)$ .

### 13.6. Non-tangential maximal operator.

*Definition.*

Let  $F : \mathbb{R}_+^{d+1} \rightarrow \mathbb{C}$  be a function (which is not always measurable) and  $a > 0$ .

Set the cone of aperture  $a$  by

$$(13.85) \quad \Gamma_a(x) := \{(y, t) \in \mathbb{R}_+^{d+1} : |x - y| < at\}.$$

For simplicity we write  $\Gamma(x) := \Gamma_a(x)$ .

We set the non-tangential maximal function  $F_a^*$  by

$$(13.86) \quad F_a^*(x) := \sup_{(y, t) \in \Gamma_a(x)} |F(y, t)|.$$

We abbreviate  $F_1^*(x)$  to  $F^*(x)$ .

**Lemma 13.27.** *Let  $a > 0$ . Suppose that  $F$  is a complex valued function on  $\mathbb{R}^{d+1}$ . Then  $F_a^*$  is upper-semicontinuous. That is, for all  $\lambda > 0$  we have*

$$(13.87) \quad O_\lambda := \{F_a^* > \lambda\}$$

*is an open set. In particular  $F_a^*$  is measurable.*

*Proof.* Let  $x \in O_\lambda$ . Then there exists  $(y, t) \in \Gamma_a(x)$  such that

$$(13.88) \quad |F(y, t)| > \lambda.$$

We write down  $(y, t) \in \Gamma_a(x)$  explicitly:  $|x - y| < at$ . If  $x' \in \mathbb{R}^d$  satisfies

$$(13.89) \quad |x' - x| < at - |x - y|,$$

then we have  $|x' - y| < at$ . Thus we obtain  $B(y, at) \subset O_\lambda$ . □

$\gamma$ -density point.

**Definition 13.28.** Let  $A$  be a closed subset on  $\mathbb{R}^d$ . A point  $x \in \mathbb{R}^d$  has global  $\gamma$ -density with respect to  $A$ , if

$$(13.90) \quad \frac{|B(x, r) \cap A|}{|B(x, r)|} \geq \gamma$$

for all  $r > 0$ .

Let  $A^*$  be the points of global  $\gamma$ -density with respect to  $A$ ; then  $A^*$  is closed. Indeed, by Lebesgue's convergence theorem, we see  $\lim_{y \rightarrow x} |B(y, r) \cap A| = |B(x, r) \cap A|$ , which shows that  $A^*$  is closed.

We can say that  $A$  does not have global  $\gamma$ -density with respect to  $A$ , if

$$(13.91) \quad \frac{|B(x, r) \cap A^c|}{|B(x, r)|} > 1 - \gamma.$$

Thus, if we set  $O = A^c$  and  $O^* = (A^*)^c$ , then we have  $O^* := \{M[\chi_O] > 1 - \gamma\}$ .

By virtue of the weak-(1,1) property of the Hardy-Littlewood maximal operator  $M$ , we obtain

$$(13.92) \quad |O^*| = |\{M[\chi_O] > 1 - \gamma\}| \lesssim \frac{|O|}{1 - \gamma}.$$

The following lemma ensures us that it does not count what value we take for the aperture.

**Theorem 13.29.** *Let  $a \geq b > 0$ . Then*

$$(13.93) \quad |\{F_a^* > \alpha\}| \lesssim_d (1 + ab^{-1})^d |\{F_b^* > \alpha\}|$$

*for all  $\alpha > 0$ . In particular, we have*

$$(13.94) \quad \int_{\mathbb{R}^d} F_b^*(x)^p dx \leq \int_{\mathbb{R}^d} F_a^*(x)^p dx \lesssim_d (1 + ab^{-1})^d \int_{\mathbb{R}^d} F_b^*(x)^p dx.$$

*Proof.* We let  $O = \{F_b^* > \alpha\}$  and show that  $\{F_a^* > \alpha\}$  is contained in the complement of the set of points of global  $\gamma$ -density of  $A := O^c$ , provided we take  $\gamma$  slightly less than 1. Indeed, suppose  $F_a^*(x) > \alpha$  for a given  $x$ . Then there exists  $(y, t) \in \Gamma_a(x)$  with  $|F(y, t)| > \alpha$ . Now  $B(y, bt) \subset O$ , therefore  $O \cap B(x, a + b) \supset B(y, bt)$  and

$$(13.95) \quad \frac{|O \cap B(x, a + b)|}{|B(x, a + b)|} \geq \frac{b^d}{(b + a)^d}.$$

Thus

$$(13.96) \quad \frac{|A \cap B(x, a+b)|}{|B(x, a+b)|} \leq 1 - \frac{b^d}{(b+a)^d},$$

and  $x \notin A$  if  $\gamma = 1 - \frac{b^d}{2(b+a)^d}$ . Therefore by the weak-(1, 1) boundedness of the Hardy-Littlewood maximal operator

$$(13.97) \quad |\{F_a^* > \alpha\}| \leq |(A^*)^c| \lesssim |A^c| = |\{F_b^* > \alpha\}|.$$

The proof of the theorem is now complete.  $\square$

Carleson measure. In connection with the non-tangential maximal operator, we consider the Carleson measure.

**Definition 13.30.**

(1) Let  $B$  be a ball in  $\mathbb{R}^d$ . Then define

$$(13.98) \quad \tilde{B} := B \times (0, r(B)) \subset \mathbb{R}_+^{d+1}.$$

The set  $\tilde{B}$  is called the Carleson box of  $B$ .

(2) A measure  $\nu : \mathbb{R}_+^{d+1}$  is said to be a Carleson measure, if there exists a constant  $c > 0$  such that  $\nu(\tilde{B}) \leq c|B|$ . The smallest admissible value of  $c$  is called the Carleson constant and is denoted by  $\|\nu\|_{\text{Carleson}}$ .

**Theorem 13.31.** *Let  $O$  be an open set in  $\mathbb{R}^d$ . If we define*

$$(13.99) \quad \tilde{O} := \{(x, t) \in \mathbb{R}_+^{d+1} : B(x, t) \subset O\},$$

*then we have*

$$(13.100) \quad \nu(\tilde{O}) \leq 3^d \|\nu\|_{\text{Carleson}} |O|.$$

*Proof.* Let us consider

$$(13.101) \quad \tilde{O}_R := \{(x, t) \in \mathbb{R}^d \times (0, R) : B(x, t) \subset O\}$$

for  $R > 0$ . In view of separability of  $\mathbb{R}^d \times (0, R)$  we can find a countable sequence  $\{B(x_j, r_j)\}_{j=1}^\infty$  of balls such that

$$(13.102) \quad \tilde{O}_R = \bigcup_{j=1}^\infty \widetilde{B(x_j, r_j)}, \quad r_j \leq R.$$

Let  $N$  be fixed. Then by the  $5r$ -covering lemma, we have

$$(13.103) \quad \bigcup_{j=1}^N \widetilde{B(x_j, r_j)} \subset \bigcup_{k=1}^{M_N} \widetilde{B(x_{j_k^N}, 3r_{j_k^N})}, \quad \bigcup_{k=1}^{M_N} \chi_{B(x_{j_k^N}, r_{j_k^N})} \leq 1$$

for some sequence  $1 \leq j_1^N \leq j_2^N \leq \dots \leq j_{M_N}^N \leq N$ . With the help of this covering, we obtain

$$\begin{aligned} \nu \left( \bigcup_{j=1}^N \widetilde{B(x_j, r_j)} \right) &\leq \nu \left( \bigcup_{k=1}^{M_N} \widetilde{B(x_{j_k^N}, 3r_{j_k^N})} \right) \\ &\leq \sum_{k=1}^{M_N} \nu(\widetilde{B(x_{j_k^N}, 3r_{j_k^N})}) \\ &\leq \|\nu\|_{\text{Carleson}} \sum_{k=1}^{M_N} |B(x_{j_k^N}, 3r_{j_k^N})| \\ &\leq 3^d \|\nu\|_{\text{Carleson}} |O|. \end{aligned}$$



This is the desired result. □

Recall that  $\Gamma(x) = \Gamma_1(x) = \{(y, t) : |x - y| < t\} \subset \mathbb{R}_+^{d+1}$  for  $x \in \mathbb{R}^d$ .

**Proposition 13.32** (Carleson embedding). *Let  $\nu$  be a Carleson measure on  $\mathbb{R}_+^{d+1}$ . Then for every positive Borel measurable function  $F$  we have*

$$(13.104) \quad \int_{\mathbb{R}_+^{d+1}} F(x, t) d\nu(x, t) \leq 3^d \|\nu\|_{\text{Carleson}} \int_{\mathbb{R}^d} \left( \sup_{(y,t) \in \Gamma(x)} F(y, t) \right) dx.$$

*Proof.* We write  $F^*(x) := \sup_{(y,t) \in \Gamma(x)} F(y, t)$ . By the distribution formula we have

$$(13.105) \quad \int_{\mathbb{R}_+^{d+1}} F(x, t) d\nu(x, t) = \int_0^\infty \nu\{(x, t) \in \mathbb{R}_+^{d+1} : F(x, t) > \lambda\} d\lambda.$$

If we set  $O = \{F^* > \lambda\} \subset \mathbb{R}^d$ , then it follows immediately from the definition of  $F^*$  that  $O$  is an open set. Furthermore we have

$$(13.106) \quad \{(y, t) \in \mathbb{R}_+^{d+1} : F(y, t) > \lambda\} \subset \tilde{O}.$$

Indeed, let  $F(y, t) > \lambda$ . In order to show that  $(y, t) \in \tilde{O}$ , we need to prove  $B(y, t) \subset O$ . Let  $z \in B(y, t)$ , that is,  $(y, t) \in \Gamma(z)$ . Then we have

$$(13.107) \quad F^*(z) = \sup_{(w,s) \in \Gamma(z)} F(w, s) > F(y, t).$$

As a result, we have  $B(y, t) \subset O$ , or equivalently,  $(y, t) \in \tilde{O}$ , proving (13.106).

From (13.106), we deduce

$$(13.108) \quad \nu\{(y, t) \in \mathbb{R}_+^{d+1} : F(y, t) > \lambda\} \leq \left| \left\{ \sup_{(y,t) \in \Gamma(x)} F(y, t) > \lambda \right\} \right| \leq 3^d \|\nu\|_{\text{Carleson}} |O|.$$

If we insert (13.108) to (13.105), then we obtain the desired result. □

**Corollary 13.33.** *Suppose that  $\nu$  is a Carleson measure. Assume that  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  is an integrable function such that, for all  $x \in \mathbb{R}^d$ , the estimate  $|\varphi(x)| \leq (1 + |x|)^{-d-\varepsilon}$  holds for some  $\varepsilon > 0$ . Set  $\varphi_t := t^{-d}\varphi(\cdot/t)$  for  $t > 0$ . If  $1 < p < \infty$ , then*

$$(13.109) \quad \int_{\mathbb{R}_+^{d+1}} |\varphi_t * f(x)|^p d\nu(x, t) \lesssim_{d,p,\varepsilon} \|\nu\|_{\text{Carleson}} \int_{\mathbb{R}^d} |f(x)|^p dx.$$

*Proof.* We let  $F(x, t) = \sup_{(y,t) \in \Gamma(x)} |\varphi_t * f(y)|$ . We deduce a pointwise estimate  $F(x) \lesssim_\varepsilon Mf(x)$

from Proposition 13.32. Consequently, we obtain (13.109) by the  $L^p(\mathbb{R}^d)$ -boundedness of the Hardy-Littlewood maximal operator  $M$ ; more precisely,

$$\begin{aligned} \int_{\mathbb{R}_+^{d+1}} |\varphi_t * f(x)|^p d\nu(x, t) &\leq 3^d \|\nu\|_{\text{Carleson}} \int_{\mathbb{R}_+^{d+1}} \sup_{(y,t) \in \Gamma(x)} |\varphi_t * f(y)|^p d\nu(x, t) \\ &\lesssim \|\nu\|_{\text{Carleson}} \int_{\mathbb{R}_+^{d+1}} Mf(x)^p d\nu(x, t) \\ &\lesssim \|\nu\|_{\text{Carleson}} \int_{\mathbb{R}^d} |f(x)|^p dx, \end{aligned}$$

proving (13.109). □

Notes and references for Chapter 7. This chapter contains elementary theorems for maximal operators. The author has referred to [10, 16, 57, 58].

As for the interpolation technique used in this book, we refer to [113, 116] for more information. Caldeón-considered the interpolation functors in [111].

Section 12. Theorem 12.7

Theorem 12.11

The theory of the Hardy-Littlewood maximal operator in particular Theorems 12.12 and 12.13 goes back to the paper by G. Hardy and J. Littlewood [227] and by Wiener [494] in 1930, where they placed themselves in the periodic setting. G. Hardy and J. Littlewood considered the case when  $n = 1$  and N. Wiener generalized the result to the higher dimensional case. See [471] for a good account of the history of the Hardy-Littlewood maximal operator.

As we have seen in Theorem 12.13, the  $L^p(\mathbb{R}^d)$ -norm of the uncentered ball maximal operator was less than

$$(13.110) \quad \left( \frac{p \cdot 2^p \cdot 3^d}{p-1} \right)^{\frac{1}{p}}, \quad 1 < p < \infty.$$

As for the  $L^p(\mathbb{R}^d)$ -bounds of the centered maximal operator  $M'$ , given by (12.1), more can be said: In [460] Stein established that there exists a constant  $c_p$  independent of  $d$  such that

$$(13.111) \quad \|M'f\|_p \leq c_p \|f\|_p$$

where  $M$  now denotes the centered ball maximal operator.

Let us now consider the weak-(1,1) estimates. In general we denote by  $M_B$  the maximal operator with respect to an open, convex and symmetric set  $B$  in  $\mathbb{R}^d$ . The precise definition is

$$(13.112) \quad M_B f(x) = \sup_{r>0} \frac{1}{|rB|} \int_{rB} |f(x-y)| dy.$$

In [457] Stein showed that

$$(13.113) \quad \|\{M_B f > \lambda\}\| \lesssim \frac{d \log d}{\lambda} \|f\|_1,$$

where the implicit constant in  $\lesssim$  is independent of  $d$ .

There are a huge amount of results on the size of the constants in Theorems 12.12, 12.13. Let us set

$$\begin{aligned} \alpha_{M'} &:= \sup\{|\{M'f > 1\}| : \|f\|_1 = 1\} \\ \alpha_M &:= \sup\{|\{Mf > 1\}| : \|f\|_1 = 1\} \\ \beta_{p,M'} &:= \sup\{\|M'f\|_p : \|f\|_p = 1\} \\ \beta_{p,M} &:= \sup\{\|Mf\|_p : \|f\|_p = 1\} \end{aligned}$$

for  $1 < p < \infty$ . A naive estimate is that  $1 < \alpha_{M'} \leq \alpha_M \leq 2^d \alpha_{M'} < \infty$  and  $1 < \beta_{p,M'} \leq \beta_{p,M} \leq 2^d \beta_{p,M'} < \infty$  for  $1 < p < \infty$ . However, we can say even more.

Carleson proved that  $\alpha_{M'} \leq 2^d$  [132]. If  $d = 1$ , Melas proved that

$$(13.114) \quad \alpha_{M'} = \frac{11 + \sqrt{61}}{12}.$$

For related results we refer to [78]. The key to the proof of the weak-(1,1) boundedness is that we discretize the  $L^1(\mathbb{R}^d)$ -functions. This technique is firstly taken up in [340]. This technique

yields interesting inequalities. For example, Menarguez and Torrea proved

$$(13.115) \quad \left| \left\{ x \in \mathbb{R} : \left| \sum_{k=1}^N \sum_{j \in \mathbb{Z}} \frac{b_k^j e^{ijx}}{x - a_j} \right| > \lambda \right\} \right| \leq \frac{C}{\lambda} \sum_{k=1}^N \sum_{j \in \mathbb{Z}} |b_k^j|$$

for  $\lambda > 0$  and  $b_1, b_2, \dots, b_k \in \mathbb{C}$  [341]. For the related result we refer to [318], where Loomis proved that the inequality in (13.115) can be replaced by equality if  $j = 0$  and  $b_k^j = \delta_{j0}$  for all  $k = 1, 2, \dots, N$  and  $j \in \mathbb{Z}$ .

As for the non-homogeneous space, we refer to [471]. Grafakos and Smith established that  $\beta_{p,M}$  is a solution to the following equation

$$(13.116) \quad (p - 1)\beta_{p,M}^p - p\beta_{p,M}^{p-1} - 1 = 0,$$

if  $d = 1$ .

See [26, 211] for the one sided Hardy-Littlewood maximal operators given by

$$M^r f(t) := \sup_{r>0} \frac{1}{r} \int_t^{t+r} |f(s)| ds$$

$$M^l f(t) := \sup_{r>0} \frac{1}{r} \int_{t-r}^t |f(s)| ds$$

for  $t \in \mathbb{R}$ .

Theorem 12.14

Section 12.3. We refer to [374, 424, 471, 474] for discussions of the boundedness of the Hardy-Littlewood maximal operators on nonhomogeneous spaces.

Section 13. Theorem 13.1

Using the culmination of our work, we can prove the differential theorem with ease. For different approach of Theorem 13.2 we refer to [69], for example.

Proposition 13.4 is due to Stein.

Theorem 13.6

Theorem 13.7

Theorem 13.15

Theorem 13.16

Theorem 13.19

Theorem 13.20

Theorem 13.21

Theorem 13.24

Theorem 13.25

Exercise 143 was from

Whitney used Theorem 13.26 in [492]. He used it to investigate the differential properties of functions.

Theorem 13.29

Theorem 13.31

## Part 8. Singular integral operators

In this part, which is another heart of this book as well as Chapter 7, we take up singular integral operators. Suppose that  $f \in L^1(\mathbb{R}^d)$ . Then the operator  $Tg(x) := \int_{\mathbb{R}^d} f(x-y)g(y) dy$  is  $L^p(\mathbb{R}^d)$ -bounded for all  $1 \leq p \leq \infty$  by virtue of Hausdorff-Young inequality (4.82). However, what happens if we remove the assumption  $f \in L^1(\mathbb{R}^d)$ . In many aspects in harmonic analysis, it can often happen that the kernel is not an  $L^1(\mathbb{R}^d)$ -function.

Sections 14–17 are devoted to the singular integral operators. If we change our viewpoint, the Hausdorff Young inequality shows that the convolution operator  $f \rightarrow g * f$  is a bounded operator provided  $g$  enjoys nice integrability. However, what is amazing is that the boundedness is still available, if  $g$  is not sufficiently integrable. We shall take full advantage of the cancellation property of the integral kernel  $g$ . The Hilbert transform is a prototype of such “singular” integral operators which we shall take up in Section 14. After investigating the Hilbert transform, we intend to extend our result to the Riesz transform. The Riesz transform is one of the key singular integral operators used in partial differential equations such as Navier-Stokes equations. When we convert partial differential equations into the integral form, the Riesz transform often comes into play. In Section 16 we re-examine the proof of the boundedness of singular integral operators and we sharpen the assumption. This refinement of our results are essential for our later applications. The Calderón-Zygmund theory is a theory of (especially) singular integral operators and their maximal operators such as wavelet theory and so on. The crux is to decompose the functions into “good” part and “bad” part. The good part belongs to  $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and hence it could not be nicer, while the bad part is decomposed further into a countable sum of “not so bad” functions. In Section 17 we consider another type of singular integral operators called fractional integral operators and we discuss the difference between the Hilbert transform and the fractional integral operators.

### 14. HILBERT TRANSFORM

As we saw in Section 1, the Hilbert transform on the torus was a key tool for the theory of convergence of Fourier series. The aim of Section 14 is to prove the following theorem about the Hilbert transform. Therefore, throughout Section 14, we place ourselves in the real line  $\mathbb{R}$ .

**Theorem 14.1.** *Let  $f \in L^p(\mathbb{R})$  with  $1 \leq p < \infty$ .*

(1) *The limit*

$$(14.1) \quad Hf(t) := \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{|u| > \varepsilon} \frac{f(t-u)}{u} du$$

*exists for almost all  $t \in \mathbb{R}$ .*

(2) *If  $p = 2$ , then we have  $\|Hf\|_2 = \|f\|_2$ .*

(3) *If  $1 < p < \infty$ , then  $\|Hf\|_p \leq \|f\|_p$  for all  $f \in L^p(\mathbb{R})$ .*

(4) *If  $p = 1$ , then*

$$(14.2) \quad |\{ |Hf| > \lambda \}| \lesssim \frac{1}{\lambda} \|f\|_1$$

*for all  $\lambda \in \mathbb{R}$ .*

The transform (14.1) is called the Hilbert transform.

The mapping  $f \mapsto Hf$  is called Hilbert transform (on the real line) and it is a prototype of singular integral operators. The proof of this theorem is long and it has a long history. Here we content ourselves with giving the detailed proof apart from historical remarks. Before we investigate the properties of the operator  $H$ , let us make it clear what space  $H$  acts on. Denote by Meas the set of all measurable functions modulo the almost everywhere equivalence with respect to the Lebesgue measure.

In this chapter we always place ourselves in the setting of  $\mathbb{R}$  coming with the Lebesgue measure  $dt$ . Therefore, the measure is sometimes omitted.

Given a set of finite measure  $E$ , a measurable function  $f$  defined on  $E$  and  $\lambda > 0$ , we define

$$(14.3) \quad p_{\lambda, E}(f) := |E \cap \{|f| > \lambda\}|.$$

Recall that Meas is the set of all measurable functions on  $\mathbb{R}$ . We equip Meas with a topology generated by  $\{p_{\lambda, E}\}$ , where  $\lambda$  runs over all positive real numbers and  $E$  over the set of all subsets of finite measures. Then we can regard  $L^p(\mathbb{R})$ ,  $1 \leq p \leq \infty$  at least as a subset of Meas, which we do not allude to in what follows.

In view of the definition of the Hilbert transform, it is natural to define its truncation:

$$(14.4) \quad H_\varepsilon f(t) = \int_{|u| > \varepsilon} \frac{f(t-u)}{u} du \quad (\varepsilon > 0)$$

for  $f \in L^p(\mathbb{R})$  with  $1 \leq p < \infty$ . Of course this integral exists for all  $t \in \mathbb{R}$ .

**Example 14.2.** Let  $f = \chi_{[-1,1]}$  and  $\varepsilon < 1$ . Then we have

$$\begin{aligned} H_\varepsilon f(t) &= \int_{|u| > \varepsilon, |t-u| < 1} \frac{du}{u} \\ &= \int_{u > \varepsilon, t-1 < u < t+1} \frac{du}{u} + \int_{u < -\varepsilon, t-1 < u < t+1} \frac{du}{u} \\ &= \int_{\max(\varepsilon, t-1) < u < t+1} \frac{du}{u} + \int_{t-1 < u < \min(t+1, -\varepsilon)} \frac{du}{u}. \end{aligned}$$

Note that the first integral is not zero only when  $\varepsilon < t+1$ . Similarly, the second integral is not zero only when  $t-1 < -\varepsilon$ . Thus,

$$H_\varepsilon f(t) = \chi_{(\varepsilon-1, \infty)}(t) \log \frac{t+1}{\max(\varepsilon, t-1)} + \chi_{(-\infty, 1-\varepsilon)}(t) \log \frac{\min(t+1, -\varepsilon)}{t-1} \quad (t \in \mathbb{R}).$$

Letting  $\varepsilon \downarrow 0$ , then we have

$$Hf(t) = \chi_{(-1, \infty)}(t) \log \frac{t+1}{\max(0, t-1)} + \chi_{(-\infty, 1)}(t) \log \frac{\min(t+1, 0)}{t-1} \quad (t \in \mathbb{R}).$$

When  $t > 2$ , then the above calculation shows

$$Hf(t) = \log \frac{t+1}{t-1} \sim t^{-1},$$

Thus,  $Hf \notin L^1(\mathbb{R})$ .

Sometimes it is convenient to use the smooth truncation. That is, let  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  be an even function such that

$$(14.5) \quad \chi_{\mathbb{R} \setminus [-2,2]} \leq \eta \leq \chi_{\mathbb{R} \setminus [-1,1]}.$$

Then define the smooth truncation of the Hilbert transform by

$$(14.6) \quad \tilde{H}_\varepsilon f(t) = \int_{\mathbb{R}} \eta\left(\frac{t}{\varepsilon}\right) \frac{f(t-u)}{u} du.$$

As for these two truncated Hilbert transforms, we have the following.

**Lemma 14.3.** *Let  $f \in \bigcup_{1 \leq p < \infty} L^p(\mathbb{R})$  and  $\varepsilon > 0$ . Then*

$$(14.7) \quad |\tilde{H}_\varepsilon f(t) - H_\varepsilon f(t)| \lesssim Mf(t)$$

for all  $t \in \mathbb{R}$ . Here the implicit constant does not depend on  $\varepsilon$ .

*Proof.* Indeed, by the triangle inequality, we obtain

$$|\tilde{H}_\varepsilon f(t) - H_\varepsilon f(t)| = \int_{\varepsilon < |u| < 2\varepsilon} \left(1 - \eta\left(\frac{t}{\varepsilon}\right)\right) \frac{|f(t-u)|}{|u|} du \leq \frac{1}{\varepsilon} \int_{\varepsilon < |u| < 2\varepsilon} |f(t-u)| du \lesssim Mf(t),$$

proving the lemma.  $\square$

In view of Lemma 14.3, we see that the two operators  $H_\varepsilon$  and  $\tilde{H}_\varepsilon$  are equivalent: Suppose that we have proved

$$(14.8) \quad \|\tilde{H}_\varepsilon f\|_p \lesssim_p \|f\|_p, \quad f \in L^p(\mathbb{R})$$

for all  $1 < p < \infty$ . Then by virtue of the pointwise estimate (14.7) it follows that

$$(14.9) \quad \|H_\varepsilon f\|_p \lesssim_p \|f\|_p, \quad f \in L^p(\mathbb{R})$$

for all  $1 < p < \infty$ . Conversely (14.9) implies (14.8).

#### 14.1. Strong-(2, 2) estimate.

Keeping the above observations in mind, let us begin with the simplest case:

Assume  $p = 2$ .

Let  $a_\varepsilon(t) := \frac{1}{\pi t} \chi_{(-\infty, -\varepsilon) \cup (\varepsilon, \infty)}(t)$ . Since  $H_\varepsilon$  is a convolution operator generated by  $a_\varepsilon$ , it suffices to deal with its Fourier transform and investigate its  $L^\infty(\mathbb{R})$ -norm.

**Lemma 14.4.** *There exists  $M > 0$  such that*

$$(14.10) \quad |\mathcal{F}a_\varepsilon(\tau)| \leq M$$

for all  $\tau \in \mathbb{R}$  and  $\varepsilon > 0$ .

*Proof.* Observe that

$$(14.11) \quad a_\varepsilon \cdot \chi_{[-R, R]} \rightarrow a_\varepsilon$$

in the topology of  $\mathcal{S}'(\mathbb{R}^d)$ . Taking this into account, we write the Fourier transform out in full:

$$(14.12) \quad \mathcal{F}a_\varepsilon(\tau) = \sqrt{\frac{1}{2\pi}} \lim_{R \rightarrow \infty} \int_{(-R, -\varepsilon) \cup (\varepsilon, R)} \frac{e^{iu\tau}}{u} du = \sqrt{\frac{2}{\pi}} \lim_{R \rightarrow \infty} \int_\varepsilon^R \frac{\sin u\tau}{u} du = \sqrt{\frac{2}{\pi}} \int_\varepsilon^\infty \frac{\sin u\tau}{u} du.$$

Recall that  $a \in \mathbb{R} \mapsto \int_0^a \frac{\sin u\tau}{u} du$  is bounded. Here it will be understood that

$$(14.13) \quad \int_\varepsilon^\infty \frac{\sin u\tau}{u} du = \lim_{R \rightarrow \infty} \int_\varepsilon^R \frac{\sin u\tau}{u} du,$$

although the integral does not converge absolutely. Therefore, the assertion is immediate.  $\square$

**Theorem 14.5.** *Let  $\varepsilon > 0$ . Then*

$$(14.14) \quad \|H_\varepsilon f\|_2 \lesssim \|f\|_2$$

for all  $f \in L^2(\mathbb{R})$ . Furthermore, if  $f \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R})$ , then the limit  $\lim_{\varepsilon \downarrow 0} H_\varepsilon f(t) = Hf(t)$  exists for all  $t \in \mathbb{R}$  and we have

$$(14.15) \quad \|Hf\|_2 = \|f\|_2 \text{ for all } f \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R}).$$

*Proof.* To prove the first assertion, we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R})$ . Once we assume  $f \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R})$ , the assertion is immediate by the Plancherel theorem and Lemma 14.4. To obtain the second assertion, we remark that  $p.v. \frac{1}{t} \in \mathcal{S}'(\mathbb{R}^d)(\mathbb{R})$ . Therefore, the existence of the limit is clear. Reexamine the calculation of the proof of Lemma 14.4. Then we see that

$$(14.16) \quad Hf(t) = \lim_{\varepsilon \downarrow 0} H_\varepsilon f(t) = \lim_{\varepsilon \downarrow 0} \mathcal{F}_\tau^{-1} \left( \frac{2i}{\pi} \int_\varepsilon^\infty \frac{\sin u\tau}{u} du \cdot \mathcal{F}f(\tau) \right) (t) = i\mathcal{F}_\tau^{-1}(\text{sgn} \cdot \mathcal{F}f(\tau))(t),$$

where  $\mathcal{F}_\tau^{-1}$  denotes the Fourier transform with respect to the variable  $\tau$  and the convergence takes place in the  $L^2(\mathbb{R})$ -topology as well as we have the pointwise convergence. In view of this equality, we conclude (14.15).  $\square$

#### 14.2. Weak-(1,1) estimate.

The key tool for the proof of the boundedness of the singular integral operators is the CZ-decomposition.

Weak-(1,1) boundedness. With the Calderón-Zygmund decomposition in mind, let us state the weak-(1,1) boundedness.

**Theorem 14.6** (Kolmogorov). *The following inequality holds:*

$$(14.17) \quad |\{t \in \mathbb{R} : |H_\varepsilon f(t)| > \lambda\}| \lesssim \frac{1}{\lambda} \|f\|_1,$$

where the implicit constant in  $\lesssim$  are independent of  $f$  and  $\varepsilon$ .

*Proof.* Form the Calderón-Zygmund decomposition of  $f$  at height  $\lambda$ . Then we obtain a sequence of disjoint dyadic cubes  $\{Q_j\}_j$ ,  $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ -function  $g$  and a sequence of  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ -functions  $\{b_j\}_j$  satisfying the five conditions listed below:

$$(1) |g(t)| \leq \lambda, (2) \int_{\mathbb{R}} |g(t)| dt \leq 2 \int_{\mathbb{R}} |f(t)| dt, \\ (3) \text{supp}(b_j) \subset Q_j, (4) \int_{\mathbb{R}} b_j(t) dt = 0, (5) \sum_j |Q_j| \leq \frac{1}{\lambda}.$$

Here, for the sake of simplicity we set  $b := f - g$ . We first decompose

$$(14.18) \quad |\{|H_\varepsilon f| > \lambda\}| \leq |\{|H_\varepsilon g| > \lambda/2\}| + |\{|H_\varepsilon b| > \lambda/2\}|.$$

The good part is easy to estimate. Now that we established the  $L^2(\mathbb{R})$ -boundedness of  $H$ , we have

$$\left| \left\{ t \in \mathbb{R} : |H_\varepsilon g(t)| > \frac{\lambda}{2} \right\} \right| \leq \frac{4}{\lambda^2} \int_{\mathbb{R}} |H_\varepsilon g(t)|^2 dt \lesssim \frac{1}{\lambda^2} \int_{\mathbb{R}} |g(t)|^2 dt.$$

By virtue of (1), we obtain

$$\left| \left\{ t \in \mathbb{R} : |H_\varepsilon g(t)| > \frac{\lambda}{2} \right\} \right| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}} |g(t)| dt \lesssim \frac{1}{\lambda} \int_{\mathbb{R}} |f(t)| dt.$$



To obtain this estimate, we used the Chebychev inequality for the first inequality. The second inequality holds by virtue of the  $L^2(\mathbb{R})$ -boundedness of the Hilbert transform  $H$ . The remaining inequality can be obtained by the properties (1) and (2) of the Calderón-Zygmund decomposition. We turn to the “bad part”. First, we delete the influence of cubes. Speaking precisely, we proceed

$$(14.19) \quad |\{ |H_\varepsilon b| > \lambda/2 \}| \leq (2d)^d \sum_j |Q_j| + \left| \left( \mathbb{R} \setminus \bigcup_j 2dQ_j \right) \cap \left\{ |H_\varepsilon b| > \frac{\lambda}{2} \right\} \right|.$$

By (5) the first term of the right-hand side is dominated by  $\frac{1}{\lambda} \int_{\mathbb{R}} |f(t)| dt$ . Thus, we concentrate on the second term. By the Chebychev inequality again, we have

$$(14.20) \quad \left| \left( \mathbb{R} \setminus \bigcup_j 2dQ_j \right) \cap \left\{ |H_\varepsilon b| > \frac{\lambda}{2} \right\} \right| \leq \frac{2}{\lambda} \int_{\mathbb{R} \setminus \bigcup_j 2dQ_j} |H_\varepsilon b|.$$

Note that the convergence of the series

$$(14.21) \quad H_\varepsilon b = \sum_j H_\varepsilon b_j$$

takes place in  $L^2(\mathbb{R})$ , because we are assuming  $H_\varepsilon$  is  $L^2(\mathbb{R})$ -bounded. Therefore, a passage to a subsequence gives us the almost everywhere convergence as well. By the triangle inequality, we have

$$(14.22) \quad |H_\varepsilon b(t)| \leq \sum_j |H_\varepsilon b_j(t)|$$

holds for a.e.  $t \in \mathbb{R}$ . As a consequence by virtue of the monotone convergence theorem we obtain

$$(14.23) \quad \int_{\mathbb{R} \setminus \bigcup_j 2dQ_j} |H_\varepsilon b| \leq \sum_k \int_{\mathbb{R} \setminus \bigcup_j 2dQ_j} |H_\varepsilon b_k| \leq \sum_j \int_{\mathbb{R} \setminus 2dQ_j} |H_\varepsilon b_j|.$$

Now we use an explicit representation of  $H_\varepsilon b_j$ :

$$(14.24) \quad H_\varepsilon b_j(t) = \int_{\mathbb{R}} \eta \left( \frac{t-u}{\varepsilon} \right) \frac{b_j(u)}{t-u} du$$

for  $t \in \mathbb{R} \setminus 2dQ_j$ , taking into account (3). By virtue of (4), we obtain

$$H_\varepsilon b_j(t) = \int_{\mathbb{R}} \eta \left( \frac{t-u}{\varepsilon} \right) \frac{b_j(u)}{t-u} du = \int_{\mathbb{R}} \left\{ \eta \left( \frac{t-u}{\varepsilon} \right) \frac{1}{t-u} - \eta \left( \frac{t-c_{Q_j}}{\varepsilon} \right) \frac{1}{t-c_{Q_j}} \right\} b_j(u) du,$$

where  $c_{Q_j}$  denotes the center of  $Q_j$ .

Suppose  $t \in \mathbb{R}^d \setminus 2dQ_j$  and  $u \in Q_j$ . Denote by  $\ell(Q_j)$  the sidelength of  $Q_j$ . Then we have

$$(14.25) \quad \left| \eta \left( \frac{t-u}{\varepsilon} \right) \frac{1}{t-u} - \eta \left( \frac{t-c_{Q_j}}{\varepsilon} \right) \frac{1}{t-c_{Q_j}} \right| \lesssim \frac{\ell(Q_j)}{|t-c_{Q_j}|^2},$$

where the implicit constant in  $\lesssim$  does not depend on  $\varepsilon$ . Therefore,

$$(14.26) \quad |H_\varepsilon b_j(t)| \lesssim \frac{\ell(Q_j)}{|t-c_{Q_j}|^2} \int |b_j(s)| ds$$

for a.e.  $t \in \mathbb{R} \setminus 2dQ_j$ . Integrating this, we obtain

$$(14.27) \quad \sum_j \int_{\mathbb{R} \setminus 2dQ_j} |H_\varepsilon b_j(t)| dt \lesssim \sum_j \int_{\mathbb{R}} |b_j(t)| dt \lesssim \int_{\mathbb{R}} |f(t)| dt < \infty.$$

Putting together (14.20) and (14.27), we obtain the desired result. □

**Theorem 14.7** (Kolmogorov). *We have*

$$(14.28) \quad |\{t \in \mathbb{R} : |Hf| > \lambda\}| \lesssim \frac{1}{\lambda} \|f\|_1$$

for all  $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ .

*Proof.* Reexamine the proof of Theorem 14.6. If  $t \notin \text{supp}(f)$ , then we have

$$(14.29) \quad Hf(t) = \int_{\mathbb{R}} \frac{f(u)}{t-u} du$$

from the assumption  $\lim_{\varepsilon \downarrow 0} H_\varepsilon f = Hf$  in  $L^2(\mathbb{R})$ . Therefore, we obtain the desired result.  $\square$

### 14.3. Truncated Hilbert transform.

A part of Theorem 14.1 shall be proven in the next theorem.

**Theorem 14.8** (M. Riesz). *Suppose that  $1 < p < \infty$ . Then we have*

$$(14.30) \quad \|H_\varepsilon f\|_p \lesssim_p \|f\|_p$$

and

$$(14.31) \quad \|Hf\|_p \lesssim_p \|f\|_p$$

for all  $f \in \mathcal{S}(\mathbb{R})$ .

*Proof.* A passage to the limit gives us (14.30), once we prove (14.31). This is because we know that

$$(14.32) \quad Hf(t) = \lim_{\varepsilon \downarrow 0} H_\varepsilon f(t)$$

for  $f \in \mathcal{S}(\mathbb{R})$ . Suppose first that  $1 < p < 2$ . Then

$$\int |H_\varepsilon f(t)|^p dt = \int_0^\infty p \lambda^{p-1} |\{ |H_\varepsilon f| > \lambda \}| d\lambda = p \cdot 2^p \int_0^\infty \lambda^{p-1} |\{ |H_\varepsilon f| > 2\lambda \}| d\lambda.$$

Note that by the weak-(1, 1) estimate and the  $L^2(\mathbb{R})$ -estimate, we have

$$\begin{aligned} |\{ |H_\varepsilon f| > 2\lambda \}| &\leq |\{ |H_\varepsilon[\chi_{\{|f| < \lambda\}} \cdot f]| > \lambda \}| + |\{ |H_\varepsilon[\chi_{\{|f| \geq \lambda\}} \cdot f]| > \lambda \}| \\ &\lesssim \frac{1}{\lambda^2} \|H[\chi_{\{|f| < \lambda\}} \cdot f]\|_2^2 + \frac{1}{\lambda} \|\chi_{\{|f| \geq \lambda\}} \cdot f\|_1 \\ &\leq \frac{1}{\lambda^2} \|\chi_{\{|f| < \lambda\}} \cdot f\|_2^2 + \frac{1}{\lambda} \|\chi_{\{|f| \geq \lambda\}} \cdot f\|_1. \end{aligned}$$

Therefore, it follows that

$$\int |H_\varepsilon f(t)|^p dt \lesssim \int_0^\infty (\lambda^{p-3} \|\chi_{\{|f| < \lambda\}} \cdot f\|_2^2 + \lambda^{p-2} \|\chi_{\{|f| \geq \lambda\}} \cdot f\|_1) d\lambda = \int_{\mathbb{R}} |f(t)|^p dt.$$

This proves the theorem when  $1 < p < 2$ .

Therefore, what remains to be proved is the case  $2 < p < \infty$ . Note that

$$(14.33) \quad \int_{\mathbb{R}} H_\varepsilon f(t)g(t) dt = - \int_{\mathbb{R}} f(t)H_\varepsilon g(t) dt$$

for all  $g \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R})$ . Therefore by the duality  $L^p(\mathbb{R})$ - $L^{p'}(\mathbb{R})$  we have

$$\|Hf\|_p = \sup_{g \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R}) \setminus \{0\}} \frac{1}{\|g\|_{p'}} \left| \int_{\mathbb{R}} H_\varepsilon f(t)g(t) dt \right| = \sup_{g \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R}) \setminus \{0\}} \frac{1}{\|g\|_{p'}} \left| \int_{\mathbb{R}} f(t)H_\varepsilon g(t) dt \right|$$

Now that we have  $p' \in (1, 2)$ ,  $H$  is bounded on  $L^p$ . Therefore by the Hölder inequality we have

$$(14.34) \quad \|H_\varepsilon f\|_p \leq \sup_{g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}} \frac{\|f\|_p \cdot \|H_\varepsilon g\|_{p'}}{\|g\|_{p'}} \lesssim \|f\|_p.$$

This is the desired result. □

**14.4. Cotlar's inequality and almost everywhere convergence.** Now we are interested in the convergence of  $H_\varepsilon f(x)$  as  $\varepsilon \downarrow 0$ . As we did to prove the existence of the Lebesgue points, it is of use to consider the maximal operator. We set

$$(14.35) \quad H^* f(t) = \sup_{\varepsilon > 0} |H_\varepsilon f(t)| \quad (t \in \mathbb{R})$$

for  $f \in \bigcup_{1 \leq p < \infty} L^p(\mathbb{R})$ .

**Theorem 14.9** (Cotlar). *Let  $f \in \mathcal{S}(\mathbb{R}^d)$  and  $0 < \eta < 1$ . Then we have a pointwise estimate*

$$(14.36) \quad \sup_{\varepsilon > 0} |H_\varepsilon f(t)| \lesssim M^{(\eta)}[Hf](t) + Mf(t) \quad (t \in \mathbb{R}),$$

where  $M^{(\eta)}$  is a powered Hardy-Littlewood maximal operator given by

$$(14.37) \quad M^{(\eta)}g(t) = M^{(\eta)}[g](t) = \sup_{I \in \mathcal{I}_t(\mathbb{R})} \left( \frac{1}{|I|} \int_I |g(u)|^\eta du \right)^{\frac{1}{\eta}} \quad (t \in \mathbb{R}).$$

*Proof.* Needless to say, it suffices to prove that

$$(14.38) \quad |H_\varepsilon f(t)| \lesssim M^{(\eta)}[Hf](t) + Mf(t) \quad (t \in \mathbb{R}),$$

for fixed  $\varepsilon > 0$ , where the implicit constant in  $\lesssim$  does not depend on  $f$  and  $\varepsilon$ .

First observe that

$$(14.39) \quad H_\varepsilon f(t) = H[\chi_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)} f](t).$$

Let  $t - \frac{\varepsilon}{2} < z < t + \frac{\varepsilon}{2}$ . Then  $H[\chi_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)} f](z)$  differs essentially from  $H_\varepsilon f(t)$  only by  $Mf(t)$ . Indeed,

$$(14.40) \quad H_\varepsilon f(t) - H[\chi_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)} f](z) = \int_{\mathbb{R}} \chi_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)}(u) f(u) \frac{z-t}{(t-u)(z-u)} du.$$

Observe that the kernel is bounded by  $\frac{2\varepsilon}{(t-u)^2 + \varepsilon^2}$ . Therefore, it follows that

$$(14.41) \quad |H_\varepsilon f(t) - H[\chi_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)} f](z)| \lesssim Mf(t).$$

With (14.41), we are led to

$$(14.42) \quad |H_\varepsilon f(t)| \lesssim Mf(t) + |H[\chi_{(t-\varepsilon, t+\varepsilon)} f](z)| + |Hf(z)|.$$

Raise (14.42) to the power  $\eta$ . Then we get

$$(14.43) \quad |H_\varepsilon f(t)|^\eta \lesssim Mf(t)^\eta + |H[\chi_{(t-\varepsilon, t+\varepsilon)} f](z)|^\eta + |Hf(z)|^\eta.$$

Taking the average of (14.43) over  $\left(t - \frac{\varepsilon}{2}, t + \frac{\varepsilon}{2}\right)$  with respect to  $z$ . Then we obtain

$$(14.44) \quad |H_\varepsilon f(t)|^\eta \lesssim Mf(t)^\eta + M^{(\eta)}[Hf](t)^\eta + \frac{1}{\varepsilon} \int_{t-\frac{\varepsilon}{2}}^{t+\frac{\varepsilon}{2}} |H[\chi_{(t-\varepsilon, t+\varepsilon)} f](z)|^\eta dz.$$

By virtue of the Kolmogorov inequality, we obtain

$$(14.45) \quad \frac{1}{\varepsilon} \int_{t-\frac{\varepsilon}{2}}^{t+\frac{\varepsilon}{2}} |H[\chi_{(t-\varepsilon, t+\varepsilon)} f](z)|^\eta dz \lesssim \left( \frac{1}{\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} |f(s)| ds \right)^\eta \lesssim Mf(t)^\eta.$$

Therefore, inserting (14.45), then we obtain

$$(14.46) \quad |H_\varepsilon f(t)|^\eta \lesssim Mf(t)^\eta + M^{(\eta)}[Hf](t)^\eta.$$

If we raise (14.46) to the power  $\frac{1}{\eta}$ , we obtain the desired result.  $\square$

An immediate consequence of Theorem 14.9, we have a control of the maximal Hilbert transform  $H^*$ .

**Theorem 14.10.** *Let  $f \in L^p(\mathbb{R})$  with  $1 < p < \infty$ . Then there exists a constant  $c_p$  depending only on  $p$  such that*

$$(14.47) \quad \|H^* f\|_p \leq c_p \|f\|_p,$$

provided  $1 < p < \infty$  and

$$(14.48) \quad |\{H^* f > \lambda\}| \leq \frac{c_1}{\lambda} \|f\|_1,$$

if  $p = 1$ .

*Proof.* By continuity of  $H_\varepsilon f(t)$  with respect to  $\varepsilon > 0$ , we have

$$(14.49) \quad H^* f(t) = \sup_{\substack{\varepsilon \in \mathbb{Q} \\ \varepsilon > 0}} |H_\varepsilon f(t)|,$$

which shows the measurability of  $H^* f$  as well. Therefore, we can replace  $H^*$  with

$$(14.50) \quad T_J f(t) = \sup_{\varepsilon \in J} |H_\varepsilon f(t)|$$

where  $J \subset \mathbb{Q}$  is a finite set and it suffices to show

$$(14.51) \quad \|T_J^* f\|_p \leq c_p \|f\|_p, \quad |\{T_J^* f > \lambda\}| \leq \frac{c_1}{\lambda} \|f\|_1$$

with  $c_p$ ,  $1 \leq p < \infty$  dependent only on  $p$ . Assuming  $J$  finite, in view of Theorem 14.8, we may assume that  $f \in \mathcal{S}(\mathbb{R})$ .

Suppose that  $1 < p < \infty$ . Then by the maximal inequality, we have

$$(14.52) \quad \|T_J^* f\|_p \lesssim \|H^* f\|_p \lesssim \|M^{(\eta)}[Hf]\|_p + \|Mf\|_p \lesssim \|f\|_p.$$

Suppose instead that  $p = 1$ . Then we have

$$(14.53) \quad |\{T_J^* f > \lambda\}| \lesssim \frac{1}{\lambda} \|f\|_1 + \left| \left\{ M^{(\eta)}[Hf] > \frac{\lambda}{2} \right\} \right| \lesssim \frac{1}{\lambda} \|f\|_1.$$

Therefore, setting  $E = \left\{ M^{(\eta)}[Hf] > \frac{\lambda}{2} \right\}$ , we have only to show

$$(14.54) \quad |E| \lesssim \frac{1}{\lambda} \|f\|_1.$$

Since  $M$  and  $H$  are  $L^{2\eta-1}(\mathbb{R})$ -bounded, we know that  $E$  is of finite measure. By the Chebychev inequality and the distribution function, we have

$$(14.55) \quad |E| \leq \frac{1}{\lambda^\eta} \int_E |Hf(t)|^\eta dt = \frac{1}{\lambda^\eta} \int_0^\infty \eta \cdot \rho^{\eta-1} |E \cap \{|Hf| > \rho\}| d\rho$$

If we invoke the weak-(1, 1) boundedness of  $H$ , we obtain

$$(14.56) \quad |E| \leq \frac{\eta}{\lambda^\eta} \int_0^\infty \rho^{\eta-1} \min \left( |E|, \frac{1}{\rho} \int_{\mathbb{R}} |f(t)| dt \right) d\rho \simeq \frac{1}{\lambda^\eta} |E|^{1-\eta} \|f\|_1.$$

Keeping the finiteness of  $|E|$  in mind, we finally obtain

$$(14.57) \quad |E| \lesssim \frac{1}{\lambda} \|f\|_1.$$

This completes the proof. □

With this maximal inequality, we finally obtain Theorem 14.1.

*Proof of Theorem 14.1.* What has been intact up to now is the existence of the almost everywhere limit. However, having established the boundedness of the corresponding maximal operator and the limit does exist for  $f \in L^2(\mathbb{R})$ , we see that the limit does exist whenever  $f \in L^p(\mathbb{R})$  with  $1 \leq p < \infty$ . A passage to the limit gives us the desired operator estimate for  $1 \leq p < \infty$ . Having set down the almost everywhere convergence problem, the strong  $L^p(\mathbb{R})$ -estimate ( $1 < p < \infty$ ) is just a matter of using the Fatou lemma. What is less trivial is the weak-(1, 1) estimate of  $H$ . To prove this observe that

$$(14.58) \quad \chi_{\{|Hf|>\lambda\}} \leq \liminf_{j \rightarrow \infty} \chi_{\{|H_{2^{-j}}f|>\lambda\}}.$$

We remark that the inequality  $\leq$  in the above formula is strict because it can happen

$$(14.59) \quad 0 = \chi_{\{|Hf|>\lambda\}}(t) < \liminf_{j \rightarrow \infty} \chi_{\{|H_{2^{-j}}f|>\lambda\}}(t) = \chi_{\{|Hf|\geq\lambda\}}(t) = 1,$$

if  $Hf(t) = \lambda$ . Therefore, we are again in the position of using the Fatou lemma and the proof is now complete. □

#### 14.5. Cotlar's lemma and its application.

Cotlar's lemma. Before we investigate the  $L^2(\mathbb{R})$ -boundedness, we invoke a lemma on Hilbert space theory. Let  $\mathcal{H}$  be a Hilbert space.

We begin with some preliminary remarks about norms of operators on  $\mathcal{H}$ .

*Exercise 144.* Let  $T$  be a bounded operator on Hilbert space  $\mathcal{H}$ . Denote by  $\|\cdot\|$  the operator norm  $\|\cdot\|_{B(\mathcal{H})}$ .

- (1) Show that  $\|T^*\| = \|T\|$ .
- (2) Prove that  $\|T\|^2 = \|T^*T\|$ .
- (3) By using the spectral decomposition theorem, prove that  $\|S\|^m = \|S^m\|$ , if  $S$  is a bounded self-adjoint operator.
- (4) Finally conclude that

$$(14.60) \quad \|T\|^{2m} = \|(T^*T)^m\|.$$

**Theorem 14.11.** *Suppose that  $\{T_j\}_{j \in J}$  is a finite collection of bounded operators on  $\mathcal{H}$ . We assume that a sequence of positive constants  $\{\gamma_j\}_{j \in J}$  satisfies*

$$(14.61) \quad A := \sum_{j \in J} \gamma_j < \infty$$

and our hypothesis is that

$$(14.62) \quad \|T_i^*T_j\|_{B(\mathcal{H})} \leq \gamma_{i-j}^2, \quad \|T_iT_j^*\|_{B(\mathcal{H})} \leq \gamma_{i-j}^2.$$

Then the operator  $T_J := \sum_{j \in J} T_j$  satisfies

$$(14.63) \quad \|T_J\| \leq A.$$

*Proof.* In proving the theorem, we shall take full advantage of (14.60) because it allows us to most efficiently exploit our hypotheses. Written out in full,

$$(14.64) \quad (T^*T)^m = \sum_{i_1, i_2, \dots, i_{2m}} T_{i_1}^* T_{i_2} T_{i_3}^* \dots T_{i_{2m}}.$$

We shall estimate this sum by majorizing the norms of the individual summands.

First, associating the factors in each summand as

$$(14.65) \quad T_{i_1}^* T_{i_2} T_{i_3}^* \dots T_{i_{2m}} = (T_{i_1}^* T_{i_2}) (T_{i_3}^* T_{i_4}) \dots (T_{i_{2m-1}} T_{i_{2m}}),$$

and using the first inequality in (14.62), we get

$$(14.66) \quad \|T_{i_1}^* T_{i_2} T_{i_3}^* \dots T_{i_{2m}}\|_{B(\mathcal{H})} \leq \gamma(i_1 - i_2)^2 \gamma(i_3 - i_4)^2 \dots \gamma(i_{2m-1} - i_{2m})^2.$$

Alternatively, we can associate the factors as

$$(14.67) \quad T_{i_1}^* T_{i_2} T_{i_3}^* \dots T_{i_{2m}} = T_{i_1}^* (T_{i_2} T_{i_3}^*) \dots (T_{i_{2m-2}} T_{i_{2m-1}}^*) T_{i_{2m}}.$$

Then since  $\|T_{i_1}\| \leq \gamma(0) \leq A$ , and similarly  $\|T_{i_{2m}}\| \leq A$ , we get

$$(14.68) \quad \|T_{i_1}^* T_{i_2} T_{i_3}^* \dots T_{i_{2m}}\|_{B(\mathcal{H})} \leq A^2 \gamma(i_2 - i_3)^2 \gamma(i_4 - i_5)^2 \dots \gamma(i_{2m-2} - i_{2m-1})^2.$$

We take the geometric mean of (14.66) and (14.68) and insert this in (14.64). The result is

$$(14.69) \quad \|(T^* T)^m\|_{B(\mathcal{H})} \leq \sum_{i_1, \dots, i_{2m}} A \gamma(i_1 - i_2) \gamma(i_2 - i_3) \dots \gamma(i_{2m-1} - i_{2m}).$$

In the above, we first sum in  $i_1$  and use the fact that  $\sum_{i_1} \gamma(i_1 - i_2) \leq A$ . Next, we sum in  $i_2$ ,

using  $\sum_{i_2} \gamma(i_2 - i_3) \leq A$ . Continuing in this way for  $i_1, i_2, \dots, i_{2m-1}$  gives

$$(14.70) \quad \|(T^* T)^m\|_{B(\mathcal{H})} \leq A^{2m-1} \sum_{i_{2m}} 1.$$

We assumed that we had only finitely many non-zero  $T_i$ 's. Then by (14.60),

$$(14.71) \quad \|T\|_{B(\mathcal{H})} \leq A \cdot \sharp J^{\frac{1}{2m}}.$$

Finally, we let  $m \rightarrow \infty$ , proving the theorem.  $\square$

$L^2(\mathbb{R})$ -boundedness of the Hilbert transform  $H$ . There is another approach with which to prove the  $L^2(\mathbb{R})$ -boundedness of the Hilbert transform. We used the Fourier transform to this end but this is somehow limited: The Fourier transform is too heavy in some cases. In order that we apply the Cotlar inequality, we set

$$(14.72) \quad H_j(t) = \int_{2^j < |u| < 2^{j+1}} \frac{f(u)}{t-u} du$$

for  $j \in \mathbb{Z}$  and  $f \in L^2(\mathbb{R})$ .

**Lemma 14.12.** *Suppose  $\{H_j\}_{j \in \mathbb{Z}}$  is a family of uniformly bounded linear operators on  $L^2(\mathbb{R})$ . For all  $j, k$  we have*

$$(14.73) \quad \|H_j H_k\|_{B(L^2(\mathbb{R}))} \lesssim 2^{-|j-k|}.$$

*Proof.* By considering the adjoint it can be assumed that  $j \geq k$ . Since the estimate trivially holds near the diagonal, we may assume  $j > k + 4$ . Set

$$(14.74) \quad K_j(t, u) = \frac{\chi_{\Delta_j}(u)}{t-u}$$

for  $j \in \mathbb{Z}$ , where  $\Delta_j := \{t \in \mathbb{R} : 2^j < |t| < 2^{j+1}\}$ . Then we have

$$(14.75) \quad \|H_j H_k\|_{B(L^2(\mathbb{R}))} \leq \|K_j * K_k\|_{L^1(\mathbb{R})}$$

by virtue of the Hausdorff Young inequality. Write out the integral kernel out in full.

$$(14.76) \quad K_j * K_k(t) = \int_{2^j < |u| < 2^{j+1}, 2^k < |t-u| < 2^{k+1}} \frac{1}{(t-u)u} du.$$

Thus, we see that the kernel is even. For the purpose of estimating the size of the kernel we shall assume that  $t \geq 0$  for the time being. In order that the integral region

$$(14.77) \quad E := \{u \in \mathbb{R} : 2^j < |u| < 2^{j+1}, 2^k < |t - u| < 2^{k+1}\}$$

is non-empty, it is necessary and sufficient that  $t \in (2^j - 2^{k+1}, 2^{j+1} + 2^{k+1})$ . In this case the integral domain is contained in  $(0, \infty)$ . Set

$$(14.78) \quad E_1 = (2^{j+1} - 2^{k+1}, 2^j + 2^{k+1}), E_2 = (2^j - 2^{k+1}, 2^{j+1} + 2^{k+1}) \setminus (2^{j+1} - 2^{k+1}, 2^j + 2^{k+1}).$$

Suppose first that  $t \in E_1$ . In this case we have

$$\begin{aligned} K_j * K_k(t) &= \int_{t-2^{k+1} < u < t-2^k, t+2^k < u < t+2^{k+1}} \frac{1}{(t-u)t} dt \\ &= \int_{t-2^{k+1} < u < t-2^k, t+2^k < u < t+2^{k+1}} \left( \frac{1}{(t-u)t} - \frac{1}{(t-u)u} \right) dt \\ &= \int_{t-2^{k+1} < u < t-2^k, t+2^k < u < t+2^{k+1}} \frac{1}{tu} du. \end{aligned}$$

Therefore, taking into account that  $2^k$  is negligible in front of  $2^j$ , we have

$$(14.79) \quad |K_j * K_k(t)| \lesssim 2^{k-2j}.$$

Suppose instead that  $t \in E_2$ . Then we make use of  $t \simeq 2^j$  in the integral region. Using (14.79) we have

$$(14.80) \quad |K_j * K_k(t)| \lesssim 2^{-j}.$$

Taking into account the symmetry, we conclude

$$(14.81) \quad |K_j * K_k(t)| \lesssim 2^{k-2j} \chi_{E_1}(t) + 2^{-j} \chi_{E_2}(t) \quad (t \in \mathbb{R}).$$

Integrating (14.81) over  $\mathbb{R}$ , we conclude

$$(14.82) \quad \|H_j H_k\|_{L^2(\mathbb{R})} \leq \|K_j * K_k\|_1 \lesssim 2^{j-k}.$$

Thus, the proof is now complete. □

As a consequence of Lemma 14.12 and the Cotlar lemma, we can prove the boundedness of the Hilbert transform.

## 15. ROTATION METHOD AND THE RIESZ TRANSFORM

Here we go back to the setting of  $\mathbb{R}^d$  to consider counterparts of the Hilbert transform.

### 15.1. Rotation method.

In this section as an application of Hilbert transform we shall extend the class of singular integral operators. We say that a function  $K$  on  $\mathbb{R}^d \setminus \{0\}$  is homogeneous of degree  $-d$  if

$$(15.1) \quad K(rx) = r^{-d} K(x)$$

for all  $r > 0$  and  $x \in \mathbb{R}^d \setminus \{0\}$ .

**Definition 15.1.** Let  $K$  be a  $C^1$ -function on  $\mathbb{R}^d \setminus \{0\}$  of homogeneous of degree  $-d$ . Then define

$$(15.2) \quad T_K f(x) := \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y) f(x - y) dy,$$

provided the limit exists. Define also

$$(15.3) \quad T_{K,\varepsilon}f(x) := \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y)f(x-y) dy.$$

**Example 15.2.**

- (1) Identify  $\mathbb{C}$  naturally with  $\mathbb{R}^2$ . Equip the Lebesgue measure  $dA$  with  $\mathbb{C}$ . The Beurling transform

$$Bf(z) = -\frac{1}{\pi} \text{p.v.} \int_{\mathbb{C}} f(z-w) \frac{1}{w^2} dA(w)$$

is an example of singular integral with homogeneous kernel.

- (2) Let  $n \geq 2$  and  $1 \leq j, k \leq n$  with  $j \neq k$ . The integral operator

$$R_{j,k}f(x) = \text{p.v.} \int_{\mathbb{R}^d} f(x-y) \frac{y_j y_k}{|y|^{d+2}} dy$$

is called the Second order Riesz transform.

- (3) Analogously, we can consider

$$R_P f(x) = \text{p.v.} \int_{\mathbb{R}^d} f(x-y) \frac{P(y)}{|y|^{d+D}} dy$$

for homogeneous polynomials  $P(x)$  of order  $D \in \mathbb{N}$ .

The operators of the above type generalize the Hilbert transform.

**Theorem 15.3.** *Let  $K \in C^1(\mathbb{R}^d \setminus \{0\})$  be an odd function which is homogeneous of degree  $-d$ . Let  $1 < p < \infty$ . Then,*

$$(15.4) \quad \|T_{K,\varepsilon}f\|_p \lesssim_{K,p} \|f\|_p$$

for every  $f \in L^p(\mathbb{R}^d)$ .

We set  $\sigma$  be the Lebesgue measure on  $S^d$ .

*Proof.* It is not so hard to see that  $T_{K,\varepsilon}$  has a kernel  $K(\cdot) \cdot \chi_{B(\varepsilon)}$ . Thus, the point is to obtain a uniform estimate of the norm of  $T_{K,\varepsilon}$ . Therefore, to prove (15.4) we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)$ .

The operator  $T_{K,\varepsilon}$  factors through  $L^p(\mathbb{R}^d)$  as  $T_{K,\varepsilon} = S \circ T$ , where  $S$  is a nice integral operator and  $T$  is a variant of the Hilbert transform.

In this case we can write the integral by using the polar coordinate, since  $f$  and  $K$  are continuous. By using the polar coordinate we have

$$T_{K,\varepsilon}f(x) = \int_{S^{d-1}} \left( \int_{\varepsilon}^{\infty} K(r\sigma) f(x-r\sigma) r^{d-1} dr \right) d\sigma = \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^{\infty} f(x-r\sigma) \frac{dr}{r} \right) d\sigma.$$

By the change of variables  $\sigma \in S^{d-1} \mapsto -\sigma \in S^{d-1}$  we obtain

$$(15.5) \quad \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^{\infty} f(x-r\sigma) \frac{dr}{r} \right) d\sigma = \int_{S^{d-1}} K(-\sigma) \left( \int_{\varepsilon}^{\infty} f(x+r\sigma) \frac{dr}{r} \right) d\sigma.$$

Since we are assuming  $K$  odd, another change of variables yields

$$(15.6) \quad \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^{\infty} f(x-r\sigma) \frac{dr}{r} \right) d\sigma = \int_{S^{d-1}} K(\sigma) \left( \int_{-\infty}^{-\varepsilon} f(x-r\sigma) \frac{dr}{r} \right) d\sigma.$$

Therefore, taking average of the above two quantity, we obtain

$$(15.7) \quad T_{K,\varepsilon}f(x) = \frac{1}{2} \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f(x-r\sigma) \frac{dr}{r} \right) d\sigma.$$



By the Minkowski inequality, we obtain

$$\begin{aligned} & \left\{ \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y) f(x-y) dy \right)^p dx \right\}^{\frac{1}{p}} \\ & \leq \frac{1}{2} \int_{S^{d-1}} |K(\sigma)| \left\{ \int_{\mathbb{R}^d} \left( \left( \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f(x-r\sigma) \frac{dr}{r} \right) \right)^p dx \right\}^{\frac{1}{p}} d\sigma \\ & = \frac{1}{2} \int_{S^{d-1}} |K(\sigma_1)| \left\{ \int_{\mathbb{R}^d} \left( \left( \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f(x-r\sigma_1) \frac{dr}{r} \right) \right)^p dx \right\}^{\frac{1}{p}} d\sigma_1. \end{aligned}$$

Given  $\sigma_1 \in S^{d-1}$ , we pick an orthonormal system  $\sigma_2, \sigma_3, \dots, \sigma_d$  so that  $\{\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_d\}$  is an orthonormal basis. Then we have

$$\begin{aligned} & \int_{\mathbb{R}^d} \left( \left( \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f(x-r\sigma_1) \frac{dr}{r} \right) \right)^p dx \\ & = \int_{\mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R}} \left| \left( \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f((y_1-r)\sigma_1 + y_2\sigma_2 + \dots + y_d\sigma_d) \frac{dr}{r} \right) \right|^p dy_1 \right\} dy_2 dy_3 \dots dy_d \\ & = \pi \int_{\mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R}} |H_{\varepsilon}[f(*\sigma_1 + y_2\sigma_2 + \dots + y_d\sigma_d)](y_1)|^p dy_1 \right\} dy_2 dy_3 \dots dy_d. \end{aligned}$$

Therefore, the truncated integral can be bounded by the directional truncated Hilbert transform. Since  $H_{\varepsilon} : L^p(\mathbb{R}) \rightarrow L^p(\mathbb{R})$  is a bounded operator whose norm can be bounded by a quantity independent of  $\varepsilon$ , it follows that

$$(15.8) \quad \|T_{K,\varepsilon}f\|_p^p \lesssim \int_{\mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R}} |f(y_1\sigma_1 + y_2\sigma_2 + \dots + y_d\sigma_d)|^p dy_1 \right\} dy_2 dy_3 \dots dy_d = \|f\|_p^p.$$

This is the desired result. □

**Proposition 15.4.** *Keep to the same assumption as Theorem 15.3. If  $f \in \mathcal{S}(\mathbb{R}^d)$ , then the limit*

$$(15.9) \quad T_K f(x) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y) f(x-y) dy$$

*exists for all  $x \in \mathbb{R}^d$ .*

*Proof.* It suffices to show

$$(15.10) \quad \lim_{\varepsilon \downarrow 0} \int_{B(1) \setminus B(\varepsilon)} K(y) f(x-y) dy$$

exists for  $f \in \mathcal{S}(\mathbb{R}^d)$ . Assuming the kernel  $K$  odd, we obtain

$$(15.11) \quad \lim_{\varepsilon \downarrow 0} \int_{B(1) \setminus B(\varepsilon)} K(y) f(x-y) dy = \lim_{\varepsilon \downarrow 0} \int_{B(1) \setminus B(\varepsilon)} K(y) (f(x-y) - f(x)) dy.$$

The function satisfying  $|f(x) - f(x-y)| \lesssim |y|$ , the integral of the right-hand side converges absolutely. Therefore, we have the desired result. □

To prove that the limit  $T_{K,\varepsilon}f(x)$  exists for almost every  $x \in \mathbb{R}^d$ , we are led to the corresponding maximal operator

$$(15.12) \quad T_{K,\varepsilon}^* f(x) = \sup_{\varepsilon > 0} |T_{K,\varepsilon} f(x)|.$$

**Proposition 15.5.** *Keep to the same assumption as Theorem 15.3. Then we have*

$$(15.13) \quad \|T_K^* f\|_p \lesssim \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

*Proof.* If we re-examine the proof of Theorem 15.3, we obtain from (15.7)

$$(15.14) \quad |T_{K,\varepsilon} f(x)| = \frac{1}{2} \int_{S^{d-1}} |K(\sigma)| \sup_{\varepsilon > 0} \left| \int_{\varepsilon}^{\infty} + \int_{-\infty}^{-\varepsilon} f(x - r\sigma_1) \frac{dr}{r} \right| d\sigma_1.$$

Under the same notation by using this refined inequality, we are led to

$$(15.15) \quad \sup_{\varepsilon > 0} |T_{K,\varepsilon} f(x)| \lesssim \int_{\mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R}} |H^*[f(*\sigma_1 + y_2\sigma_2 + \dots + y_d\sigma_d)](y_1)|^p dy_1 \right\} dy_2 dy_3 \dots dy_d,$$

where  $H^*$  is the maximal operator of the truncated singular integral operators. Since we established that  $H^*$  is  $L^p(\mathbb{R}^d)$ -bounded as well in Section 11, we conclude that  $T_K^*$  is  $L^p(\mathbb{R}^d)$ -bounded.  $\square$

Taking into account the pointwise convergence when  $f \in \mathcal{S}(\mathbb{R}^d)$ , we see that the limit defining  $T_K f$  does exist for almost everywhere for each  $f \in L^p(\mathbb{R}^d)$ , whenever  $1 < p < \infty$ . By the dominated convergence theorem, we see that the limit takes place in  $L^p(\mathbb{R}^d)$  as well.

Now let us summarize our observation above.

**Theorem 15.6.** *Assume that  $K \in C^1(\mathbb{R}^d \setminus \{0\})$  is an odd function of homogeneous of degree  $-d$ . Let  $1 < p < \infty$ . Then, for each  $f \in L^p(\mathbb{R}^d)$ , the limit*

$$(15.16) \quad T_K f(x) := \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y) f(x - y) dy$$

exists for almost everywhere and in  $L^p(\mathbb{R}^d)$ . Furthermore it satisfies

$$(15.17) \quad \|T_K f\|_p \leq c_p \|f\|_p$$

for some constant  $c_p > 0$  independent of  $f$ .

Finally before we conclude this section, let us see that the operator  $T_K f$  is a Fourier multiplier.

**Theorem 15.7.** *Keep to the same assumption as Theorem 15.3. Then*

$$(15.18) \quad \mathcal{F}(T_K f) = m \cdot \mathcal{F}f$$

for all  $f \in L^2(\mathbb{R}^d)$ , where  $m$  is given by

$$(15.19) \quad m(\xi) = -\frac{\pi i}{2} \cdot (2\pi)^{-\frac{d}{2}} \int_{S^{d-1}} K(\sigma) \operatorname{sgn}(\sigma \xi) d\sigma.$$

*Proof.* Since  $T_K f = \lim_{\varepsilon \downarrow 0} T_{K,\varepsilon} f$  in the  $L^2(\mathbb{R}^d)$ -topology, it follows that

$$(15.20) \quad \mathcal{F}(T_K f) = (2\pi)^{\frac{d}{2}} \lim_{\varepsilon \downarrow 0} \mathcal{F}(\chi_{\mathbb{R}^d \setminus B(\varepsilon)} \cdot K) \cdot \mathcal{F}f.$$

Therefore, we have to calculate

$$(15.21) \quad m(\xi) := (2\pi)^{\frac{d}{2}} \lim_{\varepsilon \downarrow 0} \mathcal{F}(\chi_{\mathbb{R}^d \setminus B(\varepsilon)} \cdot K)(\xi).$$

Note that

$$\mathcal{F}(\chi_{\mathbb{R}^d \setminus B(\varepsilon)} \cdot K)(\xi) = \lim_{R \rightarrow \infty} \mathcal{F}(\chi_{B(R) \setminus B(\varepsilon)} \cdot K)(\xi) = (2\pi)^{-\frac{d}{2}} \lim_{R \rightarrow \infty} \int_{B(R) \setminus B(\varepsilon)} K(x) e^{-ix \cdot \xi} dx.$$

If we pass to the polar coordinate, then we obtain

$$\begin{aligned} \int_{B(R)\setminus B(\varepsilon)} K(x)e^{-ix\cdot\xi} dx &= -i \int_{B(R)\setminus B(\varepsilon)} K(x) \sin(x \cdot \xi) dx \\ &= -i \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^R \frac{\sin r\sigma \cdot \xi}{r} dr \right) d\sigma \end{aligned}$$

By the Lebesgue convergence theorem we obtain

$$(15.22) \quad \lim_{R \rightarrow \infty} \int_{B(R)\setminus B(\varepsilon)} K(x)e^{ix\cdot\xi} dx = -i \int_{S^{d-1}} K(\sigma) \left( \int_{\varepsilon}^{\infty} \frac{\sin r\sigma \cdot \xi}{r} dr \right) d\sigma,$$

under the understanding that  $\int_{\varepsilon}^{\infty} \frac{\sin r\sigma \cdot \xi}{r} dr = \lim_{R \rightarrow \infty} \int_{\varepsilon}^R \frac{\sin r\sigma \cdot \xi}{r} dr$ . Another application of this theorem gives us (15.19).  $\square$

**15.2. Riesz transform.**

We exhibit an example of singular integral operators to which the results in this section can be applied.

**Example 15.8.** The Riesz transform is a singular integral operators given by

$$(15.23) \quad R_j f(x) = c_d \text{p.v.} \frac{x_j}{|x|^{d+1}} * f(x),$$

where  $c_d := \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{d+1}{2}\right) 2^{2-\frac{d+1}{2}}$ .

The Riesz-transform  $R_j$  is an important singular integral operator. There are several reasons. One of the major reasons is that  $R_j$  takes a fundamental and simple form. It generalizes the Hilbert transform,

$$Hf(x) = \text{p.v.} \int_{\mathbb{R}} \frac{f(y)}{x-y} dy$$

The Hilbert transform was fundamental when we investigate the convergence of Fourier series. Another reason is that  $R_j$  appears in partial differential equations as an operator  $\partial_j(-\Delta)^{-1/2}$ .

In partial differential equations, singular integral operators play a fundamental role, for example, when we try to construct solutions. For example, in Navier Stokes equations, we need to consider the operator

$$\begin{pmatrix} 1 - R_1R_1 & -R_1R_2 & -R_1R_3 \\ -R_1R_2 & 1 - R_2R_2 & -R_2R_3 \\ -R_1R_3 & -R_2R_3 & 1 - R_3R_3 \end{pmatrix},$$

where

$$R_j f(x) = c_3 \int_{\mathbb{R}^3} \frac{x_j - y_j}{|x - y|^4} f(y) dy$$

denotes the  $j$ -th Riesz transform.

The following result is a starting point of the theory of the Riesz-transform.

**Theorem 15.9.** *Let  $d \geq 1$ . Then we have  $\mathcal{F}(R_j f)(\xi) = -i \frac{\xi_j}{|\xi|} \mathcal{F}(f)$  for all  $f \in L^2(\mathbb{R}^d)$ .*

*Proof.* We assume  $d \geq 2$ , for the case when  $d = 1$  is covered as the Hilbert transform. Observe that

$$(15.24) \quad \frac{\partial}{\partial x_j} |x|^{-d+1} = (1-d)p.v. \frac{x_j}{|x|^{d+1}}$$

in the sense of distributions. Indeed, to check (15.24), we have only to show

$$(15.25) \quad \int_{\mathbb{R}^d} |x|^{-d+1} \partial_j \varphi(x) dx = (d-1) \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} \frac{x_j}{|x|^{d+1}} \varphi(x) dx.$$

By the Stokes theorem we have

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^{-d+1} \partial_j \varphi(x) dx &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} |x|^{-d+1} \partial_j \varphi(x) dx \\ &= \lim_{\varepsilon \downarrow 0} \left\{ (d-1) \int_{\mathbb{R}^d \setminus B(\varepsilon)} \frac{x_j}{|x|^{d+1}} \varphi(x) dx - \int_{\partial B(\varepsilon)} \frac{x_j}{|x|^d} \varphi(x) d\sigma(x) \right\}. \end{aligned}$$

Here  $d\sigma$  is the area measure of the surface  $\partial B(\varepsilon)$ . We now utilize the cancellation condition and we obtain

$$(15.26) \quad \lim_{\varepsilon \downarrow 0} \int_{\partial B(\varepsilon)} \frac{x_j}{|x|^d} \varphi(x) d\sigma(x) = \lim_{\varepsilon \downarrow 0} \int_{\partial B(\varepsilon)} \frac{x_j}{|x|^d} (\varphi(x) - \varphi(0)) d\sigma(x) = 0,$$

because the integrand of the second term is of order  $|x|^{-d+2}$ . Therefore,

$$(15.27) \quad \frac{\partial}{\partial x_j} |x|^{-d+1} = (1-d)p.v. \frac{x_j}{|x|^{d+1}}.$$

By taking the Fourier transform, we obtain

$$\mathcal{F} \left( p.v. \frac{x_j}{|x|^{d+1}} \right) (\xi) = \frac{1}{1-d} \mathcal{F} \left( \frac{\partial}{\partial x_j} |x|^{-d+1} \right) (\xi) = \frac{i\xi_j}{1-d} \mathcal{F}(|x|^{-d+1})(\xi) = -i \frac{\Gamma(\frac{1}{2})}{2^{2-\frac{d}{2}} \Gamma(\frac{d+1}{2})}.$$

Therefore we have the desired result.  $\square$

Once Theorem 15.9 is proven, then we can resort to the Calderón-Zygmund theory and we have  $R_j$  is bounded on  $L^p(\mathbb{R}^d)$  whenever  $1 < p < \infty$ . It is important to restate the boundedness of the Riesz transform as a theorem. For it is of much importance.

**Theorem 15.10.** *Let  $1 < p < \infty$ . For  $f \in L^p(\mathbb{R}^d)$  we denote*

$$(15.28) \quad R_j f(x) = c_d \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} \frac{y_j}{|y|^{d+1}} f(x-y) dy.$$

*Then the limit defining  $R_j f$  converges almost everywhere on  $\mathbb{R}^d$  and the inequality*

$$(15.29) \quad \|R_j f\|_p \lesssim_p \|f\|_p.$$

*Furthermore, if  $p = 2$ , then we have*

$$(15.30) \quad \mathcal{F}(R_j f)(\xi) = -i \frac{\xi_j}{|\xi|} \mathcal{F}f(\xi).$$

Below is an example of the usage of the Riesz transform in connection with complex analysis and with partial differential equations.

**Example 15.11.** *Let  $1 < p < \infty$ . Then*

$$(15.31) \quad \|\partial_x f\|_p + \|\partial_y f\|_p \lesssim_p \|\partial_x f + i\partial_y f\|_p \leq \|\partial_x f\|_p + \|\partial_y f\|_p$$

for all  $f \in W^{1,p}(\mathbb{R}^d)$ .

Needless to say, Example 15.11 is significant only if  $f$  is  $\mathbb{C}$ -valued.

*Proof.* Let us write  $x = x_1$  and  $y = x_2$ . Since  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $W^{1,p}(\mathbb{R}^d)$ , we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)$ . Then use  $\partial_1 f = R_1(R_1 - iR_2)(\partial_1 + i\partial_2)f$ , because  $\partial_1 f, \partial_2 f \in L^2(\mathbb{R}^d)$ . If we use this, then we obtain the desired result.  $\square$

We shall exhibit another example. This example reveals us an important idea that the mixed derivatives can be absorbed by the pure derivatives. Let us again place ourselves in the setting of  $\mathbb{R}^d$ .

**Example 15.12.** Let  $1 < p < \infty$  and  $1 \leq j, k \leq d$ . Then

$$(15.32) \quad \|\partial_j \partial_k f\|_p \lesssim \|\Delta f\|_p$$

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

*Proof.* We have only to note that  $\partial_j \partial_k f = c R_j R_k \Delta f$  for  $f \in \mathcal{S}(\mathbb{R}^d)$  because the equality holds as an element in  $L^2(\mathbb{R}^d)$ .  $\square$

It is helpful to declare that the lower derivatives are absorbed by the highest pure derivatives and the original function. We present an example of this spirit as an exercise.

*Exercise 145.* Let  $1 < p < \infty$ .

- (1) Set  $m(\xi) = \frac{\xi_1}{1 + |\xi|^2}$ . Then show that  $|\partial^\alpha m(\xi)| \lesssim_\alpha \langle \xi \rangle^{-|\alpha|}$ .
- (2) Show that

$$(15.33) \quad \|\partial_1 f\|_p \lesssim \|(1 - \Delta)f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ . Hint: It is known that  $\|\mathcal{F}^{-1}(m \cdot \mathcal{F}f)\|_p \leq c_p \|f\|_p$  for all  $1 < p < \infty$  and  $f \in \mathcal{S}(\mathbb{R}^d)$ . The proof may be found in Chapter 24. An alternative approach is that we use Theorem 16.6.

## 16. GENERALIZED SINGULAR INTEGRAL OPERATORS

**Definition of Generalized CZ-operators.** The generalized Calderón-Zygmund operator and the theory that lies behind are key tools of this book. What we prove in this section seems just a general nonsense. However at least it has an advantage: In the rotation method we had to avoid the case when  $p = 1$ . This is because we used the Minkowski inequality.

In this book they take various forms and used throughout this book. The definition, presented below, is highly generalized in view of the singular operators we have been dealing. However, this generalization turns out to be essential in later applications.

**Definition 16.1.** An  $L^2(\mathbb{R}^d)$ -bounded linear operator is said to be a Calderón-Zygmund operator, if it satisfies the following conditions:

- (1) There is a measurable function  $K$  such that for all  $L^\infty(\mathbb{R}^d)$ -functions with compact support we have

$$(16.1) \quad Tf(x) = \int_{\mathbb{R}^d} K(x, y)f(y) dy \text{ for all } x \in \text{supp } f.$$

- (2) The kernel function satisfies the following estimates.

$$(16.2) \quad |K(x, y)| \lesssim \frac{1}{|x - y|^d},$$

if  $x \neq y$  and

$$(16.3) \quad |K(x, y) - K(x, z)| + |K(y, x) - K(z, x)| \lesssim \frac{|x - y|}{|x - z|^{n+1}},$$

if  $0 < 2|x - y| < |z - x|$ .

There are several conditions substituted for (16.3). For example, an typical one is the condition

$$(16.4) \quad |\nabla_x K(x, y)| + |\nabla_y K(x, y)| \lesssim |x - y|^{-d-1}$$

for all  $x \neq y$ , which is called gradient condition. The gradient condition is stronger than (16.3). Meanwhile (16.3) can be replaced by

$$(16.5) \quad \int_{|x-z|>2|x-y|} (|K(x, y) - K(x, z)| + |K(y, x) - K(z, x)|) dx \lesssim 1, \quad y \neq z,$$

which is referred to as the Hörmander condition. Before we investigate the generalized singular integral operators, helpful examples may be in order.

**Example 16.2.**

- (1) The Hilbert transform is a Calderón-Zygmund operator.
- (2) Strange to say, the identity operator  $I$  is a Calderón-Zygmund operator with kernel 0.
- (3) The above two examples are in the same category of the Calderón-Zygmund operator. But the former should be distinguished from the latter. To this end, we give another definition.

**Definition 16.3.** One says that  $T$  is a genuine Calderón-Zygmund operator, if the kernel satisfies in addition that for some fixed cone of the form

$$C = R\{x \in \mathbb{R}^d : |x'| < \theta|x_d|\}$$

with  $\theta > 0$  and  $R \in O(d)$ , the estimate from below

$$K(x, x + y) \gtrsim |y|^{-d}$$

holds for all  $x \in \mathbb{R}^d$  and  $y \in C$ .

This definition is due to Burenkov, Tararykova and Guliyev.

The generalized singular integral operators contain the Hilbert transform and the Riesz transform. Observe that the definition admits operators of the nonconvolution form as Example 16.4.

**Example 16.4.** The Cauchy transform is an important operator in complex analysis. For example, we take up the operator of the form

$$(16.6) \quad Tf(t) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus [t-s, t+s]} \frac{f(s)}{t - s + i(A(t) - A(s))} ds \quad (t \in \mathbb{R}),$$

where  $A : \mathbb{R} \rightarrow \mathbb{R}$  is a Lipschitz function. Note that this can be expanded formally as

$$(16.7) \quad Tf(t) = \lim_{\varepsilon \downarrow 0} \sum_{j=0}^{\infty} \int_{\mathbb{R}^d \setminus [t-s, t+s]} \frac{(A(t) - A(s))^j}{i^j (t - s)^{j+1}} f(s) ds.$$

We can investigate  $T$  via this expansion, about which we do not go into details. This is an example of the nonconvolution singular integral operators. In Chapter 17 we encounter an application of the Calderón-Zygmund theory of nonconvolution singular integral operators.

**Example 16.5.** Let  $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$  be a bounded linear operator. Assume (16.8), (16.9), (16.10) and (16.11).

(1) For all  $f \in L^2(\mathbb{R}^d)$ ,

$$(16.8) \quad \mathcal{F}(Tf) = m\mathcal{F}(f)$$

for some  $m \in L^\infty(\mathbb{R}^d)$ .

(2) There exists a kernel  $K$ , which is a complex valued measurable function on  $\mathbb{R}^d$ , such that

$$(16.9) \quad Tf(x) = \int_{\mathbb{R}^d} K(x, y)f(y) dy \text{ for a.e. } x \notin \text{supp}(f)$$

and the kernel  $K$  satisfies the Hölder condition

$$(16.10) \quad A := \sup_{y \in \mathbb{R}^d \setminus \{0\}} \int_{\mathbb{R}^d \setminus B(2|y|)} |K(x - y) - K(x)| dx < \infty$$

and the size condition

$$(16.11) \quad B := \sup_{x \in \mathbb{R}^d \setminus \{0\}} |x|^d |K(x)| < \infty.$$

This is an example of the singular integral operators. However, a natural question arises :

What on earth is the relation of  $m$  and  $K$  ?

To answer this question we shall prove the following.

**Theorem 16.6.** *Let  $m : \mathbb{R}^d \rightarrow \mathbb{C}$  be a bounded function that is smooth on  $\mathbb{R}^d \setminus \{0\}$ . Assume that the differential inequality*

$$(16.12) \quad |\partial^\alpha m(\xi)| \lesssim_\alpha |\xi|^{-|\alpha|}, \xi \in \mathbb{R}^d \setminus \{0\}$$

*holds for each  $\alpha \in \mathbb{N}_0^d$ . Then  $f \in L^2(\mathbb{R}^d) \mapsto \mathcal{F}^{-1}[m\mathcal{F}f] \in L^2(\mathbb{R}^d)$  is a Calderón-Zygmund operator.*

In what follows we abbreviate this type of operator to CZ-operator. The next proposition explains why we want to pose symmetric condition (16.3).

**Lemma 16.7.** *For all CZ-operator  $T$ , its dual  $T^*$  on the complex Hilbert space  $L^2(\mathbb{R}^d)$  satisfies*

$$(16.13) \quad T^*f(x) = \int_{\mathbb{R}^d} \overline{K(y, x)}f(y) dy \text{ for a.e. } x \notin \text{supp}(f).$$

*Proof.* For the proof we take a function  $g$  whose support does not meet that of  $f$ . Then we have by (16.1)

$$\begin{aligned} \int_{\mathbb{R}^d} f(x)\overline{T^*g(x)} dx &= \int_{\mathbb{R}^d} Tf(x)\overline{g(x)} dx \\ &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} K(x, y)f(y)\overline{g(x)} dx dy \\ &= \int_{\mathbb{R}^d} f(y) \overline{\left( \int_{\mathbb{R}^d} \overline{K(x, y)}g(x) dx \right)} dy. \end{aligned}$$

Thus since  $g$  is an arbitrary  $L^2(\mathbb{R}^d)$ -function with its support contained in  $\mathbb{R}^d \setminus \text{supp}(f)$ , we have

$$(16.14) \quad T^*f(x) = \int_{\mathbb{R}^d} \overline{K(y, x)}f(y) dy \text{ for almost everywhere } x \notin \text{supp}(f).$$

This is the desired result. □

Note that the condition on kernel is symmetric in the variables  $x$  and  $y$ .

**Corollary 16.8.** *For a CZ-operator  $T$  its adjoint  $T^*$  is a CZ-operator again.*

### 16.1. Weak-(1,1) boundedness.

We are going to prove two theorems for the CZ operators. Recall that given  $f \in L^1(\mathbb{R}^d)$  and  $\lambda > 0$ . There are a family of disjoint (dyadic) cubes  $\{Q_j\}_j$  such that

$$(16.15) \quad \lambda \leq \frac{1}{|Q_j|} \int_{Q_j} |f| \leq 2^d \lambda.$$

Using these cubes we can decompose  $f$  with the following properties:

$$(16.16) \quad (i) f = g + \sum_j b_j \quad (ii) \text{supp}(b_j) \subset Q_j \quad (iii) \int_{\mathbb{R}^d} b_j = 0.$$

From the definition of  $g$  we deduce  $\|g\|_{L^1} \leq 3\|f\|_{L^1}$ .

**Theorem 16.9.** *Suppose that  $T$  is a CZ-operator. Then  $T$  is weak-(1,1) bounded, that is*

$$(16.17) \quad |\{|Tf| > \lambda\}| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx$$

for all  $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ .

*Proof.* The proof can be obtained by re-examining the boundedness of the Hilbert transform  $H$ . Keeping to the above notation, we have

$$|\{|Tf| \geq \lambda\}| \leq \left| \left\{ |Tg| \geq \frac{\lambda}{2} \right\} \right| + \left| \left\{ \left| T \left( \sum_j b_j \right) \right| \geq \frac{\lambda}{2} \right\} \right|.$$

The first guy is easy to estimate. In fact, using the  $L^2(\mathbb{R}^d)$ -boundedness and (13.35) as usual, we have

$$(16.18) \quad \left| \left\{ |Tg| \geq \frac{\lambda}{2} \right\} \right| \lesssim \frac{1}{\lambda^2} \int_{\mathbb{R}^d} |Tg(x)|^2 dx$$

$$(16.19) \quad \lesssim \frac{1}{\lambda^2} \int_{\mathbb{R}^d} |g(x)|^2 dx$$

$$(16.20) \quad \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |g(x)| dx$$

$$(16.21) \quad \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx.$$

Hence

$$(16.22) \quad \left| \left\{ |Tg| \geq \frac{\lambda}{2} \right\} \right| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx.$$

To estimate the second term we need the following observation:

**Claim 16.10.** *We have the following estimate.*

$$(16.23) \quad \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} |Tb_j(x)| dx \lesssim \int_{\mathbb{R}^d} |b_j(x)| dx.$$



Admitting Claim 16.10, let us finish the proof of the theorem. Firstly we proceed as follows :

$$\begin{aligned} \left| \left\{ \left| T \left( \sum_j b_j \right) \right| \geq \frac{\lambda}{2} \right\} \right| &\leq \left| \left( \mathbb{R}^d \setminus \bigcup_j 2\sqrt{d} Q_j \right) \cap \left\{ \left| T \left( \sum_j b_j \right) \right| \geq \frac{\lambda}{2} \right\} \right| + \left| \bigcup_j 2\sqrt{d} Q_j \right| \\ &\lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d \setminus \bigcup_j 2\sqrt{d} Q_j} \left| T \left( \sum_j b_j \right) (x) \right| dx + \sum_j |Q_j| \end{aligned}$$

Recall that  $\sum_j |Q_j| = \left| \bigcup_j Q_j \right| = |\{M_{\text{dyadic}} f > \lambda\}| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| dx$ . By the triangle inequality we obtain

$$\begin{aligned} \int_{\mathbb{R}^d \setminus \bigcup_k 2\sqrt{d} Q_k} \left| T \left( \sum_j b_j \right) (x) \right| dx &\leq \sum_j \int_{\mathbb{R}^d \setminus \bigcup_k 2\sqrt{d} Q_k} |Tb_j(x)| dx \\ &\leq \sum_j \int_{\mathbb{R}^d \setminus 2\sqrt{d} Q_j} |Tb_j(x)| dx. \end{aligned}$$

If we invoke Claim 16.10, then we obtain

$$\begin{aligned} \int_{\mathbb{R}^d \setminus \bigcup_k 2\sqrt{d} Q_k} \left| T \left( \sum_j b_j \right) (x) \right| dx &\lesssim \sum_j \int_{Q_j} |b_j(x)| dx \\ &\lesssim \int_{\mathbb{R}^d} (|f(x)| + |g(x)|) dx \\ &\lesssim \int_{\mathbb{R}^d} |f(x)| dx. \end{aligned}$$

We can thus estimate  $\sum_j \int_{Q_j} |b_j(x)| dx$  because the  $b_j$  have support in  $Q_j$  and the  $Q_j$  are disjoint.

Thus the proof is over modulo Claim 16.10.

Let us prove Claim 16.10. For this purpose we can write out  $Tb_j(x)$  out in full :

$$(16.24) \quad Tb_j(x) = \int_{\mathbb{R}^d} K(x, y) b_j(y) dy \text{ for all a.e. } x \in \mathbb{R}^d \setminus 2\sqrt{d} Q_j.$$

Denoting  $c(Q_j)$  as a center of  $Q_j$  we can write from the moment condition and Hölmänder condition

$$|Tb_j(x)| = \left| \int_{\mathbb{R}^d} K(x, y) b_j(y) - K(c(Q_j), y) b_j(y) dy \right| \lesssim \int_{Q_j} \frac{|x - c(Q_j)|}{|x - y|^{n+1}} |b_j(y)| dy.$$

By integrating over  $\mathbb{R}^d \setminus 2\sqrt{d}Q_j$  we have with the aid of Fubini's theorem

$$\begin{aligned} \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} |Tb_j(x)| dx &\lesssim \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} \left( \int_{Q_j} \frac{|x - c(Q_j)|}{|x - y|^{d+1}} |b_j(y)| dy \right) dx \\ &= \int_{Q_j} \left( \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} \frac{|x - c(Q_j)|}{|x - y|^{d+1}} dx \right) |b_j(y)| dy \\ &\lesssim \int_{Q_j} \left( \int_{|x-y| > |x-c(Q_j)|} \frac{|x - c(Q_j)|}{|x - y|^{d+1}} dx \right) |b_j(y)| dy \\ &\lesssim \int_{Q_j} |b_j(y)| dy. \end{aligned}$$

Thus we have finished the proof of the claim hence the theorem.  $\square$

### 16.2. $L^p(\mathbb{R}^d)$ -boundedness.

**Theorem 16.11.** *Let  $1 < p < \infty$  and  $T$  be a generalized singular integral operator. Then there exists a constant  $c_p > 0$  such that*

$$(16.25) \quad \|Tf\|_p \leq c_p \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

*Proof.* We distinguish two cases.

Case 1 :  $1 < p < 2$  In this case we use interpolation of  $L^2(\mathbb{R}^d)$ -boundedness and weak- $(1, 1)$  boundedness, the same technique that we used for Hilbert transform.

Case 2 :  $p > 2$  In this case we use the duality  $L^p(\mathbb{R}^d)$ - $L^{p'}(\mathbb{R}^d)$ . Since  $C_c^\infty(\mathbb{R}^d)$  is dense in  $L^{p'}(\mathbb{R}^d)$  we have for all  $f \in L^2(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$  that

$$\begin{aligned} \|Tf\|_p &= \sup \left\{ \left\| \int_{\mathbb{R}^d} Tf(x) \overline{g(x)} dx \right\| : g \in C_c^\infty(\mathbb{R}^d), \|g\|_{p'} \leq 1 \right\} \\ &= \sup \left\{ \left\| \int_{\mathbb{R}^d} f(x) \overline{T^*g(x)} dx \right\| : g \in C_c^\infty(\mathbb{R}^d), \|g\|_{p'} \leq 1 \right\} \\ &\leq \|f\|_p \cdot \|T^*g\|_{p'}. \end{aligned}$$

Recall that  $T^*$  is still a CZ-operator and that  $1 < p' < 2$ . Thus we have

$$(16.26) \quad \|T^*g\|_{p'} \lesssim \|g\|_{p'}.$$

Putting together these estimates, we conclude  $T$  can be extended to a bounded operator on  $L^p(\mathbb{R}^d)$ .  $\square$

### 16.3. Truncation and pointwise convergence.

Having proved the boundedness property, we now turn to the maximal operator of the truncated singular integral.

Truncated singular integral operators and its maximal operator. Now we turn to the problem of pointwise convergence. To do this, it is very effective that we use the maximal operator, as this is the case when we investigated the Lebesgue point of measurable functions. The maximal operator associated with our present situation is the following one.

**Definition 16.12.** Let  $T$  be a singular integral operator. Then define

$$(16.27) \quad \tilde{T}f(x) := \sup_{\varepsilon > 0} \left| \int_{B(x, \varepsilon)} K(x, y)f(y) dy \right| \quad x \in \mathbb{R}^d.$$

The following inequality is called the Cotlar inequality.

**Theorem 16.13.** Suppose that  $T$  is a CZ operator and  $0 < \nu \leq 1$ . Then we have

$$(16.28) \quad \tilde{T}f(x) \lesssim_\nu M^{(\nu)}[Tf](x) + Mf(x) \quad (x \in \mathbb{R}^d)$$

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

*Proof.* For the proof, we fix  $\varepsilon > 0$ . We shall estimate

$$(16.29) \quad \int_{B(x, \varepsilon)} K(x, y)f(y) dy$$

independently on  $\varepsilon$ .

Let  $z \in B(x, \frac{1}{2}\varepsilon)$ . We shall decompose  $f$  according to  $B(x, 2\varepsilon)$ . Let  $f_0 := \chi_{B(x, \varepsilon)} \cdot f$  and  $f_1 := f - f_0$ .

As for the estimate of  $f_0$  we shall make use of  $|K(x, y)| \lesssim |x - y|^{-d}$ . It is easy to see that

$$\left| \int_{B(x, \varepsilon)} K(x, y)f(y) dy \right| \lesssim Mf(x).$$

Next we will tackle the estimate of  $f_1$ . We decompose the estimate further :

$$\int_{B(x, \varepsilon)} K(x, y)f_1(y) dy = \int_{B(x, \varepsilon)} (K(x, y) - K(z, x))f_1(y) dy + \int_{B(x, \varepsilon)} K(z, y)f_1(y) dy.$$

By the Hölder condition (16.3), the first term is readily estimated by a maximal function. Thus we have

$$\left| \int_{B(x, \varepsilon)} K(x, y)f_1(y) dy \right| \lesssim Mf(x) + |Tf_1(z)|.$$

Notice that by virtue of the inequality  $|K(x, y)| \lesssim \frac{1}{|x - y|^d}$ , we have

$$(16.30) \quad \left| \int_{B(x, \varepsilon)} K(x, y)f_0(y) dy \right| \lesssim Mf(x).$$

Thus, we have obtained

$$\left| \int_{B(x, \varepsilon)} K(x, y)f(y) dy \right| \lesssim Mf(x) + |Tf_1(z)| \lesssim Mf(x) + |Tf_0(z)| + |Tf(z)|$$

as long as  $x$  and  $z$  satisfies  $|x - y| \leq \frac{\varepsilon}{2}$ . Taking the  $\nu$ -power and averaging we have

$$(16.31) \quad \left| \int_{B(x, \varepsilon)} K(x, y)f(y) dy \right|^\nu \lesssim Mf(x)^\nu + \frac{1}{|Q|} \int_Q |Tf_0(y)|^\nu dy + \frac{1}{|Q|} \int_Q |Tf(y)|^\nu dy.$$

By Kolmogorov's inequality, proved in Theorem 12.14, we have

$$(16.32) \quad \frac{1}{|Q|} \int_Q |Tf_0(z)|^\nu dx \lesssim |Q|^{-\nu} \int_Q |f(x)|^\nu dx.$$

Putting together (16.30)–(16.32), we obtained the desired result. □

Now we prove a.e. convergence of CZ-operators, keeping the estimate of the maximal singular integral operators in mind.

**Theorem 16.14.** *Assume that the limit*

$$(16.33) \quad \lim_{\varepsilon \rightarrow 0} \int_{B(x,1) \setminus B(x,\varepsilon)} K(x,y) dy$$

*exists for almost all  $x \in \mathbb{R}^d$ . Then for all  $f \in C_c^\infty$  we have for almost all  $\varepsilon > 0$*

$$(16.34) \quad \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus B(x,\varepsilon)} K(x,y) f(y) dy$$

*exists.*

*Proof.* By the kernel condition the matter is reduced to the convergence of

$$(16.35) \quad \lim_{\varepsilon \rightarrow 0} \int_{B(x,1) \setminus B(x,\varepsilon)} K(x,y) f(y) dy.$$

However by assumption (16.33) (16.35) is reduced to showing

$$(16.36) \quad \lim_{\varepsilon \rightarrow 0} \int_{B(x,1) \setminus B(x,\varepsilon)} K(x,y) (f(x) - f(y)) dy.$$

By the mean value theorem we have  $|f(x) - f(y)| \lesssim |x - y|$ . Combining this with the size condition  $|K(x,y)| \lesssim |x - y|^{-d}$ , we see that the integrand of (16.36) is integrable. Therefore, the limit in question does exist.  $\square$

Combining Cotlar's inequality with Theorem 16.14, we have the following result.

**Theorem 16.15.** *Suppose that  $T$  is a CZ kernel and that  $K$  is an associated kernel. Then the limit*

$$(16.37) \quad Tf(x) := \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus B(x,\varepsilon)} K(x,y) f(y) dy$$

*exists for almost all  $x \in \mathbb{R}^d$ , if and only if*

$$\lim_{\varepsilon \rightarrow 0} \int_{B(x,1) \setminus B(x,\varepsilon)} K(x,y) dy$$

*exists for almost every  $x \in \mathbb{R}^d$ . If this is the case, then we have*

$$(16.38) \quad \|Tf\|_p \leq c_p \|f\|_p, \quad f \in L^p(\mathbb{R}^d)$$

*for all  $1 < p < \infty$  and*

$$(16.39) \quad |\{ |Tf| > \lambda \}| \leq \frac{c_1}{\lambda} \|f\|_1, \quad f \in L^1(\mathbb{R}^d) \text{ and } \lambda > 0.$$

We remark that the singular operator in Theorem 15.3 does satisfy the condition of the above theorem and hence we have the weak-(1, 1) estimate as well. Below let us say that a generalized singular integral operator is standard if the cancellation condition (16.33) is fulfilled. In this case we say that the kernel is standard as well. As an example of the standard singular integral operator, we can list the Riesz transform.

The structure of the generalized singular integral operator. It may be surprising that the identity operator is a singular integral operator, since the identity is far from singular. However, from the very definition of the singular integral operators, the kernel  $K \equiv 0$  corresponds to the operator.

In this paragraph we investigate the structure of a standard singular integral operator with a standard kernel  $K$ . As we have established above, the limit

$$(16.40) \quad T_K f(x) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(\varepsilon)} K(y) f(x - y) dy$$

exists for  $f \in L^2(\mathbb{R}^d)$ . Therefore, subtracting  $T_K$ , we can assume that  $T$  is a singular integral operator with kernel 0.

**Theorem 16.16.** *If  $T$  is a singular integral operator with a standard kernel  $K$ , then there exists  $m \in L^\infty(\mathbb{R}^d)$  such that*

$$(16.41) \quad T f(x) = m(x) f(x) + T_K f(x)$$

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

*Proof.* In view of the observation above, we may assume that the kernel is zero.

Let us define

$$(16.42) \quad m_j(x) := T \chi_{B(j)}(x).$$

Let  $j < k$ . Then since the kernel is zero, we have

$$(16.43) \quad m_j(x) - m_k(x) = -T \chi_{B(k) \setminus B(j)}(x) = 0$$

for  $x \in B(j)$ . Therefore, the limit

$$(16.44) \quad m(x) = \lim_{j \rightarrow \infty} m_j(x)$$

exists for almost everywhere  $x \in \mathbb{R}^d$ . Let  $R \subset \mathbb{R}$  be a compact rectangular, say it is contained in  $B(J)$ . If  $x \notin R$ , then we have

$$(16.45) \quad T \chi_R(x) = 0 = m(x) \cdot \chi_R(x).$$

If  $x \in R$ , then for all  $j \geq J$ , we have

$$(16.46) \quad T \chi_R(x) = T \chi_{B(j)}(x) - T \chi_{B(j) \setminus R}(x) = m_j(x) = m_j(x) \cdot \chi_R(x).$$

Letting  $j \rightarrow \infty$ , we have

$$(16.47) \quad T \chi_R(x) = m(x) \cdot \chi_R(x).$$

Therefore, it follows that if  $R$  is a compact rectangular, then we have

$$(16.48) \quad T \chi_R(x) = m(x) \cdot \chi_R(x).$$

By the  $\pi$ - $\lambda$  principle, it follows that  $R$  can be replaced by any measurable set of finite measure. A passage to the limit therefore gives us

$$(16.49) \quad T f(x) = m(x) \cdot f(x)$$

for all  $f \in L^2(\mathbb{R}^d)$ . In order that  $m \cdot f \in L^2(\mathbb{R}^d)$  for all  $f \in L^2(\mathbb{R}^d)$ , it is necessary and sufficient that  $m \in L^\infty(\mathbb{R}^d)$ , as we have verified in Chapter 5. Therefore, the proof is now complete.  $\square$

#### 16.4. Fourier multipliers.

Having clarified the properties of generalized singular integral operator, let us go into further examples.

**Definition 16.17.** Let  $1 \leq p \leq \infty$ . A bounded function  $m$  is said to be an  $L^p(\mathbb{R}^d)$ -multiplier, if the following operator, defined originally on  $L^2(\mathbb{R}^d)$ ,

$$f \in L^2(\mathbb{R}^d) \mapsto \mathcal{F}^{-1}(m \cdot \mathcal{F}f) \in L^2(\mathbb{R}^d).$$

can be extended to a bounded operator on  $L^p(\mathbb{R}^d)$ .

Before we go further, a helpful remarks may be in order.

- (1) If  $m$  is an  $L^p(\mathbb{R}^d)$ -multiplier, then by duality  $m$  is an  $L^{p'}$ -multiplier. An interpolation gives us that  $m$  is an  $L^q(\mathbb{R}^d)$ -multiplier for  $q$  between  $p$  and  $p'$ .
- (2) If  $m \in C_c^\infty(\mathbb{R}^d)$ , then  $m$  is an  $L^p(\mathbb{R}^d)$ -multiplier for  $1 \leq p \leq \infty$ , because in this case the operator is just an convolution operator generated by a Schwartz function  $(2\pi)^{\frac{d}{2}}\mathcal{F}^{-1}m$ .

The aim of this section is to present non-trivial examples of  $L^p(\mathbb{R}^d)$ -multipliers. Let us set  $\mathbb{R}_+^d = \{x \in \mathbb{R}^d : x_d > 0\}$ , the upper half plane.

**Proposition 16.18.** *The function  $\chi_{\mathbb{R}_+^d}$  is an  $L^p(\mathbb{R}^d)$ -Fourier multiplier for all  $1 < p < \infty$ .*

*Proof.* This is just a directional Fourier transform of  $f$  with respect to the  $x_d$ -axis. □

**Proposition 16.19.** *The characteristic function of a convex polygon is an  $L^p(\mathbb{R}^d)$ -multiplier for all  $1 < p < \infty$ .*

*Proof.* The translation of multipliers is transformed into the multiplication of  $e^{ix \cdot a}$  by the Fourier transform and the rotation of multipliers is preserved by Fourier transform. Therefore, the characteristic function of any half plane is an  $L^p(\mathbb{R}^d)$ -multiplier. The characteristic function of a convex polygon, obtained by a multiplication of a finite number of the indicators of half planes, is therefore an  $L^p(\mathbb{R}^d)$ -multiplier. □

In view of this proposition it seems natural to conjecture that the characteristic function of the unit ball is an  $L^p(\mathbb{R}^d)$ -multiplier. Because the unit ball is a limit of the convex polygon in some sense. However, this is not the case except when  $p = 2$ . We just cite the result due to Fefferman (See [178]).

**Theorem 16.20** (Non-multiplier). *The characteristic function of the unit ball is never an  $L^p(\mathbb{R}^d)$ -multiplier unless  $p = 2$ .*

### 17. FRACTIONAL INTEGRAL OPERATORS

Another class of singular integral operators of importance is the fractional integral operator (of order  $\alpha$ ) defined by

$$(17.1) \quad I_\alpha f(x) := \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}} dy \quad (x \in \mathbb{R}^d).$$

Here  $0 < \alpha < d$ . Let us see how this operator comes about.

By the Sobolev integral representation theorem, we have

$$(17.2) \quad \varphi(x) = \frac{1}{\omega_d} \int_{\mathbb{R}^d} \frac{x-y}{|x-y|^d} \cdot \text{grad}\varphi(y) dy, \quad \varphi \in C_c^\infty(\mathbb{R}^d).$$

This can be shown by using the Stokes theorem. However, it is more illustrating to provide a direct proof when  $d = 1$ . Indeed, the right-hand side equals to

$$(17.3) \quad \frac{1}{2} \int_{\mathbb{R}} \frac{x-y}{|x-y|} \text{grad}(\varphi(y)) dy = \frac{1}{2} \int_{-\infty}^x -\varphi'(y) dy + \frac{1}{2} \int_x^\infty \varphi'(y) dy = \varphi(x).$$

The following lemma is interesting and useful of its own right.

**Lemma 17.1** (Comparison lemma). *Let  $0 < \alpha < d$ .*

$$(17.4) \quad \int_0^\infty \chi_{B(x,l)}(y) \frac{dl}{l^{d+1-\alpha}} = \frac{|x-y|^{-d+\alpha}}{d-\alpha}.$$

*Proof.* The proof is simple. Indeed, we have

$$(17.5) \quad \int_0^\infty \chi_{B(x,l)}(y) \frac{dl}{l^{d+1-\alpha}} = \int_0^\infty \chi_{(|x-y|,\infty)}(l) \frac{dl}{l^{d+1-\alpha}} = \int_{|x-y|}^\infty \frac{dl}{l^{d+1-\alpha}} = \frac{|x-y|^{-d+\alpha}}{d-\alpha}$$

by using Fubini's theorem, which is the desired result. □

*Exercise 146.* Let  $0 < \alpha_1 < \alpha < \alpha_2 < d$ . Show that

$$I_\alpha f(x) \lesssim M_{\alpha_1} f(x)^{\frac{\alpha_2-\alpha}{\alpha_2-\alpha_1}} M_{\alpha_2} f(x)^{\frac{\alpha-\alpha_1}{\alpha_2-\alpha_1}}.$$

**Lemma 17.2.** *Let  $f \in L^p(\mathbb{R}^d)$  with  $1 \leq p < \frac{d}{\alpha}$ .*

- (1) *The integral defining  $I_\alpha f$  converges absolutely for almost everywhere  $x \in \mathbb{R}^d$ .*
- (2) *The pointwise estimate holds*

$$(17.6) \quad |I_\alpha f(x)| \lesssim Mf(x)^{\frac{p}{s}} \|f\|_p^{1-\frac{p}{s}}$$

*for a.e.  $x \in \mathbb{R}^d$ .*

*Proof.* It suffices to prove (2) and we can assume  $f$  is positive. We write

$$(17.7) \quad I_\alpha f(x) = (d-\alpha) \int_0^\infty \left( \frac{1}{l^{d+1-\alpha}} \int_{B(x,l)} f(y) dy \right) dl.$$

We have only to insert

$$(17.8) \quad \frac{1}{l^d} \int_{B(x,l)} f(y) dy \lesssim Mf(x), \quad \frac{1}{l^d} \int_{B(x,l)} f(y) dy \lesssim \left( \frac{1}{l^d} \int_{B(x,l)} f(y)^p dy \right)^{\frac{1}{p}}$$

to obtain (2). More precisely

$$\begin{aligned} |I_\alpha f(x)| &\leq \int_{\mathbb{R}^d} \frac{|f(y)|}{|x-y|^{d-\alpha}} dy \\ &\leq (d-\alpha) \int_0^\infty \left( \frac{1}{\ell^{d-\alpha+1}} \int_{B(x,\ell)} |f(y)| dy \right) d\ell \\ &\lesssim \int_0^\infty \min \left( Mf(x), \left( \frac{1}{l^d} \int_{B(x,l)} f(y)^p dy \right)^{\frac{1}{p}} \right) \frac{dl}{l^{-\alpha+1}} \\ &\lesssim Mf(x)^{\frac{p}{s}} \|f\|_p^{1-\frac{p}{s}}. \end{aligned}$$

Here we have used Lemma 12.15. □

Thanks to the maximal inequality we obtain

**Theorem 17.3** (Hardy-Littlewood-Sobolev). *Let  $0 < \alpha < d$ .*

(1) *We have*

$$(17.9) \quad \lambda^{\frac{d}{d-\alpha}} |\{ |I_\alpha f| > \lambda \}| \lesssim \|f\|_1$$

*for all  $f \in L^1(\mathbb{R}^d)$  and  $\lambda > 0$ .*

(2) *Assume that the parameters  $1 < p < \frac{d}{\alpha}$  and  $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$ . Then we have  $\|I_\alpha f\|_q \lesssim \|f\|_p$  for all  $f \in L^1(\mathbb{R}^d)$ .*

We complete this section by expressing  $I_\alpha$  in terms of the heat operator.

**Theorem 17.4** (A. Yoshikawa). *Let  $0 < \alpha < d$ . Then*

$$(17.10) \quad I_\alpha f(x) \simeq \int_0^\infty t^{\frac{\alpha}{2}-1} e^{t\Delta} f(x) dt \quad (x \in \mathbb{R}^d)$$

*for all positive measurable functions  $f$ .*

*Proof.* Let us set  $E(x, t) = \left(\frac{1}{4\pi t}\right)^{\frac{d}{2}} \exp\left(-\frac{|x|^2}{4t}\right)$ . we have

$$(17.11) \quad \int_0^\infty t^{\frac{\alpha}{2}-1} e^{t\Delta} f(x) dt \simeq \int_0^\infty t^{\frac{\alpha-d}{2}-1} E(\cdot, t) * f(x) dt$$

Note that

$$(17.12) \quad \int_0^\infty t^{\frac{\alpha-d}{2}-1} E(x, t) dt \simeq |x|^{\alpha-d}.$$

Thus if we insert this formula to the above expression, we obtain the desired result.  $\square$

A concluding remark of this section may be in order.

**Remark 17.5.** Let  $g(x) := |x|^{\alpha-d}$ . Compare the result in this section with (4.82). Summarizing what we have obtained, we see

$$(17.13) \quad \|f * g\|_r \lesssim \|f\|_p,$$

provided

$$(17.14) \quad 1 < p < r < \infty, \quad \frac{1}{r} = \frac{1}{p} - \frac{\alpha}{d}.$$

Meanwhile the Hausdorff-Young inequality (4.82) asserts

$$(17.15) \quad \|f * h\|_r \leq \|h\|_{d/\alpha} \cdot \|f\|_p.$$

Although  $g$  does not belong to  $L^{d/\alpha}$ , we still have (17.13).

**Example 17.6.** Here we give examples of lower bounds of  $I_\alpha f(x)$ .



(1) Let  $f(x) = \chi_{B(\frac{1}{4}) \setminus \{0\}}(x) |x|^{-\frac{d}{p}} \left(\log \frac{1}{|x|}\right)^{-q}$  with  $p > 1$  and  $q \in \mathbb{R}$ . Then we have

$$\begin{aligned} I_\alpha f(x) &\gtrsim \int_{B(\frac{1}{2}|x|)} |y|^{-\frac{d}{p}} |x-y|^{\alpha-d} \left(\log \frac{1}{|y|}\right)^{-q} dy \\ &\sim \int_{B(\frac{1}{2}|x|)} |y|^{-\frac{d}{p}} |x|^{\alpha-d} \left(\log \frac{1}{|y|}\right)^{-q} dy \\ &\gtrsim |x|^{\alpha-\frac{d}{p}} \left(\log \frac{1}{|x|}\right)^{-q}, \end{aligned}$$

if  $|x| \leq \frac{1}{8}$ .

(2) Let  $f(x) = \chi_{B(\frac{1}{4}) \setminus \{0\}}(x) |x|^{-\frac{d}{p}} \left(\log \frac{1}{|x|}\right)^{-1} \left(\log \log \frac{1}{|x|}\right)^{-q}$  with  $p > 0$  and  $q \in \mathbb{R}$ . Then we have

$$I_\alpha f(x) \geq |x|^{\alpha-d} \int_{B(\frac{1}{4}|x|)} f(y) dy$$

if  $|x| \leq \frac{1}{8}$ .

Notes and references for Chapter 8.

Section 14. Theorems 14.1, 14.5, 14.6 and 14.7, which deal with the Hilbert transform, are due to Kolmogorov [271] and M. Riesz. Later P. Stein, L. H. Loomis and A. P. Calderón gave different proofs [110, 318, 449]. Theorem 14.1 (2) is due to M. Riesz when  $p \in 2\mathbb{N}$ .

The proof of Theorem 14.8 first appeared in [409] in 1927. However, it is announced in 1924 by M. Riesz in [408].

Theorem 14.9

Theorem 14.10

Theorem 14.11 was obtained originally by M. Cotlar in [151]. Later it is generalized by Cotlar, Knapp and Stein (see [280]).

Section 16. Theorems 15.3 and 15.6 are due to A. P. Calderón and A. Zygmund [119].

We also refer to [117] for Theorems 15.3, 15.6 and 15.7. The paper [117] contains Theorems 15.3, 15.6 and 15.7 as a special case.

Theorem 15.9

Theorem 15.10

Theorem 16.6 admits a various extensions. For example, if  $\mathcal{F}H \in K_{11}^s$ , the Herz space, then we have  $H(D) = \mathcal{F}^{-1}[H \cdot \mathcal{F}[\cdot]] \in B(H^1)$ . We refer to [434] for more details. A generalization of Theorem 16.6 is made by Hörmander in 1960. Indeed, it suffices to assume that

$$m \in L^\infty, \sup_{0 < R < \infty} R^d \int_{R \leq |\xi| \leq 2R} |m'(\xi)|^2 d\xi \lesssim 1.$$

Theorem 16.9 and Theorem 16.11 are due to Calderón and Zygmund in 1952.

Theorem 16.13

Theorem 16.14

Theorem 16.15

In connection with Theorem 16.16, Calderón and Zygmund considered a special class of operators, namely, they assumed that the pointwise multiplier  $m$  is constant. In this case, the set of all such operators forms an algebra (see [119]).

Fefferman constructed a counterexample as in Theorem 16.20 in his paper [178].

The method of rotation appearing this book can be found originally in [168].

We can say the celebrated papers [117, 119] are the prototypes of the CZ-theory. The CZ-theory is taken up in great detail in [16, 10, 57, 58].

We refer to [453, 454] for more information.

If  $p = \infty$ , then Example 15.11 fails. This result is due to J. Boman [101]. To describe this result, we let  $X = \overline{C_c^\infty(\mathbb{R}^2)}^{L^\infty(\mathbb{R}^2)}$ . Then the norm estimate

$$(17.16) \quad \left\| \frac{\partial^2 f}{\partial x_1 \partial x_2} \right\|_X \lesssim \left\| \frac{\partial^2 f}{\partial x_1^2} \right\|_X + \left\| \frac{\partial^2 f}{\partial x_2^2} \right\|_X + \left\| \frac{\partial^2 f}{\partial x_1} \right\|_X + \left\| \frac{\partial^2 f}{\partial x_2} \right\|_X + \|f\|_X \quad (f \in X)$$

fails.

Section 17. The author borrowed its name of Lemma 17.1 from [373].

Theorem 17.3 is a fundamental theorem whose root lies especially in [228, 229, 447].

Theorem 17.4 is just an example of the idea due to A. Yoshikawa [490]. A. Yoshikawa used semi-groups to express the integral operators. Here as an example of his philosophy, we have connected the fractional integral operator with the heat group in Theorem 17.4. Exercise 146 is due to Welland.

There is a huge amount of works on generalization and refinement of Theorem 17.3, to which we shall allude in Chapter 23.

For example, many researchers investigated  $T_\rho$  given by

$$(17.17) \quad T_\rho f(x) = \int_{\mathbb{R}^d} \frac{\rho(x-y)}{|x-y|^d} f(y) dy.$$

Note that the case when  $\rho(x) = |x|^\alpha$  covers Theorem 17.3. We consider  $T_\rho$  in the context of elliptic differential operators such that  $1 - \Delta$ . We refer to [174, 222, 293, 365, 366, 367, 368, 465] for more details. Finally we remark that

$$(17.18) \quad \|M_\alpha f\|_p \simeq \|I_\alpha\|_p, \quad 1 < p < \infty$$

where  $M_\alpha$  denotes the fractional maximal integral operator (of order  $\alpha$ ) given by

$$(17.19) \quad M_\alpha f(x) = \sup_{r>0} \frac{1}{r^{n-\alpha}} \int_{B(x,r)} |f(y)| dy.$$

See [363]. In connection with this, we define the Hardy operator

$$H_\alpha f(x) = \frac{1}{|B(o,|x|)^{1-\alpha/d}} \int_{B(o,|x|)} |f(y)| dy.$$

Then note that

$$M_\alpha f(x) = \sup_{z \in \mathbb{R}^d} H_\alpha(|f(\cdot + x)|)(z).$$

The author has learnt the proof of Theorem 17.3 from Dr. H. Tanaka [508]. The pointwise estimate  $Mf(x) \lesssim \|f\|_p^{1-\frac{p}{s}} Mf(x)^{\frac{p}{s}}$  is due originally to Hedberg [233].

## Part 9. More about maximal and CZ-operators

From Section 18 to Section 21 we make further observations of Sections 12–Section 17. In Section 18 we are going to supplement the properties that the Hilbert transform, for example, fails. For example, in Section 14 we shall disprove that the Hilbert transform is  $L^1(\mathbb{R})$ -bounded. Instead of contenting ourselves with finding counterexamples, we modify our theory so that it works well. That is, we are going to find a function space contained in  $L^1(\mathbb{R})$  which is sent to  $L^1(\mathbb{R})$  by the Hilbert transform. We also consider its dual: What is the function space to which  $L^\infty(\mathbb{R}^d)$  is sent continuously by the Hilbert transform? The answer of the second question is the function space  $\text{BMO}(\mathbb{R}^d)$ . The crucial papers of Section 18 are the ones by C. Fefferman-E. M. Stein and John-Nirenberg [183, 261]. In Section 20 motivated by the definition of  $\text{BMO}(\mathbb{R}^d)$ , we are going to define another maximal operator called the sharp maximal operator. Finally in Section 21 we investigate the boundedness property after change of measures. That is, we are going to consider the weighted measures of the form  $\mu := w(x) dx$  instead of  $dx$ , where

$$(17.20) \quad \mu(E) := w(E) := \int_E w(x) dx$$

for a Lebesgue measurable set  $E$ .

### 18. THE HARDY SPACE $H^1(\mathbb{R}^d)$

To explain Hardy spaces, let us start with the Riesz transform. Recall that the ( $j$ -th) Riesz transform is a singular integral operators given by

$$(18.1) \quad R_j f(x) = c_d p.v. \frac{x_j}{|x|^{d+1}} * f(x),$$

where  $c_d := \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{d+1}{2}\right) 2^{2-\frac{d+1}{2}}$ .

Having investigated how singular integral operators behave, we hit upon some natural questions. Recall that the Hilbert transform does not send  $L^1(\mathbb{R}^d)$  to  $L^1(\mathbb{R}^d)$  nor  $L^\infty(\mathbb{R}^d)$  to  $L^\infty(\mathbb{R}^d)$ . Indeed, we have calculated the Hilbert transform of  $\chi_{[-1,1]}$  and concluded  $H\chi_{[-1,1]} \notin L^1(\mathbb{R}^d) \cup L^\infty(\mathbb{R}^d)$ . Therefore, it is natural to ask ourselves the following questions.

**Problem 18.1.** *What kind of subspaces of  $L^1(\mathbb{R}^d)$  is sent by singular integral operators to  $L^1(\mathbb{R}^d)$ ?*

**Problem 18.2.** *Where is  $L^\infty(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  sent by singular integral operators other than  $L^2(\mathbb{R}^d)$ ?*

One of the aims of this chapter is to answer these problems.

We begin with a definition.

**Definition 18.3.** An  $(L^2(\mathbb{R}^d)-)$ atom  $a$  is an  $L^2(\mathbb{R}^d)$ -function supported on a compact cube  $Q$  satisfying

$$(18.2) \quad \|a\|_2 \leq |Q|^{-\frac{1}{2}}, \quad \int_{\mathbb{R}^d} a(x) dx = 0.$$

First we state a property of atoms.

**Lemma 18.4.** *Any atom has the  $L^1(\mathbb{R}^d)$ -norm less than 1.*

The proof is immediate by using the Hölder inequality. With Lemma 18.4 in mind, we shall present a definition of Hardy spaces.

**Definition 18.5.** The function space  $H^1(\mathbb{R}^d)$  is the set of all  $L^1(\mathbb{R}^d)$ -functions  $f$  such that it can be represented as follows :

There exists  $\lambda = \{\lambda_j\}_{j \in \mathbb{N}} \in \ell^1$  and a countable family of atoms  $A = \{a_j\}_{j \in \mathbb{N}}$  such that

$$(18.3) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j.$$

The norm of  $H^1(\mathbb{R}^d)$  is defined by

$$(18.4) \quad \|f\|_{H^1} := \inf \left\{ \sum_{j=1}^{\infty} |\lambda_j| \lambda \in \ell^1 \text{ and } A = \{a_j\}_{j \in \mathbb{N}} \text{ satisfy (18.3)} \right\}.$$

*Exercise 147.* Show that the Hardy space  $H^1(\mathbb{R}^d)$  is a Banach space. Perhaps, this is one of the typical cases to which Theorem 10.12 is skillfully applied.

*Exercise 148.* Show that  $\chi_{[0,1]} - \chi_{[-1,0]} \in H^1(\mathbb{R})$ .

Singular integral operators. Having presented a candidate of the answer of Problem 18.1, we shall prove the following.

**Theorem 18.6.** Let  $T$  be a singular integral operator. Then  $T$ , which was originally defined in  $L^1(\mathbb{R}^d)$ , is  $H^1(\mathbb{R}^d)$ - $L^1(\mathbb{R}^d)$  bounded if it is restricted to  $H^1(\mathbb{R}^d)$ . Namely, we have

$$(18.5) \quad \|Tf\|_1 \lesssim \|f\|_{H^1}.$$

*Proof.* Instead of proving (18.5) directly, we have only to prove

$$(18.6) \quad \|Ta\|_1 \lesssim 1,$$

where  $a$  is an atom and  $c$  is a constant independent of  $a$ . Indeed, any  $f \in H^1(\mathbb{R}^d)$  can be approximated in the topology of  $H^1(\mathbb{R}^d)$  by a sequence  $\{f_j\}_{j=1}^{\infty}$  of finite linear combination of atoms. Note that the topology of  $H^1(\mathbb{R}^d)$  is stronger than that of  $L^1(\mathbb{R}^d)$ . Hence,  $\lim_{j \rightarrow \infty} Tf_j = Tf$  in the topology of weak- $L^1(\mathbb{R}^d)$ . So that, if we pass to a subsequence, we can assume

$$(18.7) \quad Tf(x) = \lim_{j \rightarrow \infty} Tf_j(x)$$

for a.e.  $x \in \mathbb{R}^d$ .

Since  $a$  is an atom, there exists a cube  $Q$  fulfilling  $\|a\|_2 \leq |Q|^{-\frac{1}{2}}$ ,  $\int_{\mathbb{R}^d} a(x) dx = 0$ . We separate the estimate (18.6) into (18.8) and (18.9), where (18.8) and (18.9) are given respectively by

$$(18.8) \quad \int_{2Q} |Ta(x)| dx \lesssim 1,$$

$$(18.9) \quad \int_{\mathbb{R}^d \setminus 2Q} |Ta(x)| dx \lesssim 1.$$

For the proof of (18.8) we use the Hölder inequality and the  $L^2(\mathbb{R}^d)$ -boundedness of  $T$  :

$$\int_{2Q} |Ta(x)| dx \lesssim |Q|^{\frac{1}{2}} \left( \int_{2Q} |Ta(x)|^2 dx \right)^{\frac{1}{2}} \lesssim |Q|^{\frac{1}{2}} \left( \int_{2Q} |a(x)|^2 dx \right)^{\frac{1}{2}} \lesssim 1.$$

Meanwhile for the proof of (18.9), we use the moment condition of  $a$ . Let us denote by  $c_Q$  the center of  $Q$ .

$$(18.10) \quad |Ta(x)| = \left| \int_{\mathbb{R}^d} K(x, y)a(y) dy \right| = \left| \int_{\mathbb{R}^d} (K(x, y) - K(c_Q, y))a(y) dy \right|.$$

Now we use the triangle inequality and the Hölder condition. We have

$$(18.11) \quad |Ta(x)| \leq \int_{\mathbb{R}^d} |K(x, y) - K(c_Q, y)| \cdot |a(y)| dy \lesssim \frac{\ell(Q)}{|x - c_Q|^{d+1}} \int_{\mathbb{R}^d} |a(y)| dy \lesssim \frac{\ell(Q)}{|x - c_Q|^{d+1}}.$$

For the last inequality we have used Lemma 18.4. If we integrate both sides over  $\mathbb{R}^d \setminus 2Q$ , then we obtain (18.9).

Thus, (18.8) and (18.9) are established.  $\square$

## 19. THE SPACE BMO

Having set down Problem 18.1, we are now oriented to the dual problem, Problem 18.2. To answer this question, we define and investigate  $\text{BMO}(\mathbb{R}^d)$ .

### 19.1. Definition.

We define

$$M^\sharp f := \sup_{x \in Q} m_Q(|f - m_Q(f)|) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y) - m_Q(f)| dy.$$

**Definition 19.1** ([183]). Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . Then define

$$(19.1) \quad \|f\|_* := \sup_{Q \in \mathcal{Q}} m_Q(|f - m_Q(f)|).$$

One says that  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  is said to be of bounded mean oscillation (abbreviated to  $f \in \text{BMO}(\mathbb{R}^d)$ ), if  $\|f\|_* < \infty$ .

Note that  $f \in \text{BMO}(\mathbb{R}^d)$  if and only if  $M^\sharp f \in L^\infty(\mathbb{R}^d)$ .

The next proposition reveals us how large  $\text{BMO}(\mathbb{R}^d)$  functions are. Here and in what follows we tacitly mean by  $f \in \text{BMO}(\mathbb{R}^d)$  that  $f$  belongs to the normed space  $\text{BMO}(\mathbb{R}^d)$  or that  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  with  $\|f\|_* < \infty$ . The confusion can rarely occur.

**Proposition 19.2.** Let  $f \in \text{BMO}(\mathbb{R}^d)$ . Then

$$(19.2) \quad \int_{\mathbb{R}^d} \frac{|f(x) - m_{Q(1)}(f)|}{1 + |x|^{d+1}} dx \lesssim \|f\|_*.$$

*Proof.* Let  $Q \in \mathcal{Q}$ . Then we have

$$(19.3) \quad |m_Q(f) - m_{2Q}(f)| = m_Q(|f - m_{2Q}(f)|) \leq 2^d m_{2Q}(|f - m_{2Q}(f)|) \leq 2^d \|f\|_*.$$

Therefore,

$$(19.4) \quad |m_Q(f) - m_{2^j Q}(f)| \leq 2^j \|f\|_*$$

for  $j \in \mathbb{N}$ . Therefore,

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{|f(x) - m_{Q(1)}(f)|}{1 + |x|^{d+1}} dx &\leq \int_{Q(1)} \frac{|f(x) - m_{Q(1)}(f)|}{1 + |x|^{d+1}} dx + \sum_{j=1}^{\infty} \int_{Q(2^j)} \frac{|f(x) - m_{Q(1)}(f)|}{1 + |x|^{d+1}} dx \\ &\leq \sum_{j=1}^{\infty} \frac{1}{2^{(j-1)(d+1)}} \int_{Q(2^j)} |f - m_{Q(1)}(f)| \\ &\lesssim \sum_{j=1}^{\infty} \frac{1}{2^j} \cdot \frac{1}{|Q(2^j)|} \int_{Q(2^j)} |f - m_{Q(1)}(f)|. \end{aligned}$$

By the definition of the  $BMO(\mathbb{R}^d)$ -norm, we have

$$\int_{\mathbb{R}^d} \frac{|f(x) - m_{Q(1)}(f)|}{1 + |x|^{d+1}} dx \lesssim \sum_{j=0}^{\infty} \frac{j}{2^j} \cdot \|f\|_* \lesssim \|f\|_*.$$

This is what we want. □

*Exercise 149.* Show that

$$(19.5) \quad |m_Q(f) - m_k Q(f)| \lesssim \log(k + 2) \|f\|_*$$

for  $k \geq 1$ .

**Remark 19.3.** In the next section we deal with more precise estimates.

Despite the fact that  $\|1\|_* = 0$ , we are still eager to regard  $BMO(\mathbb{R}^d)$  as a Banach space. To do this, we first observe the following, which is an obstacle when we regard  $BMO(\mathbb{R}^d)$  as a normed space.

**Lemma 19.4.** *Suppose that  $f \in L^1_{loc}(\mathbb{R}^d)$  satisfies  $\|f\|_* = 0$ . Then  $f$  is constant a.e. The converse is also true.*

*Proof.* It is straightforward from the definition of  $\|f\|_*$  that we prove  $\|f\|_* = 0$  if  $f$  is a constant function. Suppose that  $\|f\|_* = 0$ . Then letting  $Q_0 = [-1, 1]^d$ , we have

$$(19.6) \quad f(x) = m_{2^j Q_0}(f)$$

for almost everywhere on  $x \in 2^j Q_0$ . Thus, for almost everywhere  $x \in \mathbb{R}^d$ , we have

$$(19.7) \quad f(x) = \limsup_{j \rightarrow \infty} m_{2^j Q_0}(f).$$

The right-hand side being constant, we conclude that  $f$  is constant. □

Remark that  $\mathbb{C}$  is embedded into  $L^\infty(\mathbb{R}^d)$  naturally. As a consequence, if we define

$$(19.8) \quad BMO(\mathbb{R}^d) := \{f \in L^1_{loc}(\mathbb{R}^d) : \|f\|_* < \infty\} / \mathbb{C},$$

then  $BMO(\mathbb{R}^d)$  is a normed space. We write  $f \bmod \mathbb{C}$ , if we want to disregard the difference of additive constants of  $f$ .

We can use this estimate to show that  $BMO(\mathbb{R}^d)$  is a Banach space.

**Theorem 19.5.** *The space  $BMO(\mathbb{R}^d)$  is a Banach space.*

*Proof.* We use the criterion in Theorem 10.12. To prove this we fix  $Q = [0, 1]^d$ . Let  $f_j \in BMO(\mathbb{R}^d)$  for  $j = 1, 2, \dots$  and assume

$$(19.9) \quad \sum_{j=1}^{\infty} \|f_j\|_* < \infty.$$

We may assume that  $m_Q(f_j) \equiv 0$ . Then by Proposition 19.2 we have

$$(19.10) \quad \int_{\mathbb{R}^d} \frac{|f_j(x)|}{1+|x|^{d+1}} dx \leq 2^{d+2} \|f_j\|_*$$

for all  $j = 1, 2, \dots$ . Adding them over  $j \in \mathbb{N}$ , we obtain

$$(19.11) \quad \int_{\mathbb{R}^d} \left( \sum_{j=1}^{\infty} |f_j(x)| \right) \frac{dx}{1+|x|^{d+1}} \leq 2^{d+2} \sum_{j=1}^{\infty} \|f_j\|_* < \infty.$$

Consequently we see that

$$(19.12) \quad \sum_{j=1}^{\infty} |f_j(x)| < \infty$$

for  $dx$ -a.e.  $x \in \mathbb{R}^d$ . Set

$$(19.13) \quad F(x) := \sum_{j=1}^{\infty} f_j(x), G(x) := \sum_{j=1}^{\infty} |f_j(x)| \quad (x \in \mathbb{R}^d)$$

if the series converges. Extend  $F$  and  $G$  to whole  $\mathbb{R}^d$  by defining 0 where they do not defined.

We claim that  $F \in \text{BMO}(\mathbb{R}^d)$  and that  $\sum_{j=1}^{\infty} f_j = F$  in  $\text{BMO}(\mathbb{R}^d)$ .

Note first that  $F$  is an  $L^1_{\text{loc}}(\mathbb{R}^d)$ -function, since so is  $G$ . Next given  $R \in \mathcal{Q}$ , we have

$$(19.14) \quad m_R(|F - m_R(F)|) \leq \sum_{j=1}^{\infty} m_R(|f_j - m_R(f_j)|) \leq \sum_{j=1}^{\infty} \|f_j\|_*$$

by Lebesgue's convergence theorem. Taking sup over  $R \in \mathcal{Q}$ , we see that  $F \in \text{BMO}(\mathbb{R}^d)$ .

In the same way we can prove

$$(19.15) \quad m_R \left( \left| \left\{ F - \sum_{j=1}^J f_j \right\} - m_R \left( \left\{ F - \sum_{j=1}^J f_j \right\} \right) \right| \right) \leq \sum_{j=J+1}^{\infty} \|f_j\|_*.$$

Thus we have proved that  $\sum_{j=1}^{\infty} f_j = F$  in  $\text{BMO}(\mathbb{R}^d)$ .

As a result we have shown that  $\text{BMO}(\mathbb{R}^d)$  is a Banach space.  $\square$

Although we have remarked that we have to identify functions modulo additive constant, we are still able to embed  $\text{BMO}(\mathbb{R}^d)$  to  $\text{Meas}(dx)$ . To do this, we choose a special representative with integral over  $[-1, 1]^d$  zero. Nevertheless, we usually regard  $\text{BMO}(\mathbb{R}^d)$  as a function space modulo  $\mathbb{C}$ .

Singular integral operators.

Let us answer the second question. That is, we shall prove that  $L^\infty(\mathbb{R}^d)$  is mapped into  $\text{BMO}(\mathbb{R}^d)$  by the singular integral operators.

**Theorem 19.6.** *Suppose that  $T$  is a singular integral operator. Then there exists a continuous linear operator  $S : L^\infty(\mathbb{R}^d) \rightarrow \text{BMO}(\mathbb{R}^d)$  such that*

$$(19.16) \quad Sf = Tf \quad \text{modulo additive constants}$$

for all  $f \in L^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and that

$$(19.17) \quad \int_{\mathbb{R}^d} Sf(x) \cdot g(x) dx = \int_{\mathbb{R}^d} f(x) \cdot T^*(x) dx$$



for all  $f \in L^\infty(\mathbb{R}^d)$  and  $g \in H^1(\mathbb{R}^d)$ , where  $T^*$  is a formal adjoint singular integral operator whose kernel is given by  $(x, y) \mapsto K(y, x)$ .

*Proof.* Construction of a candidate for  $S$  Let  $f \in L^\infty(\mathbb{R}^d)$  and  $Q$  be a cube. We define a function  $\tilde{S}f(x)$  by

$$(19.18) \quad \tilde{S}_Q f(x) := T(\chi_Q \cdot f)(x) + \int_{\mathbb{R}^d \setminus Q} (K(x, y) - K(c_Q, y))f(y) dy.$$

Observe that thanks to the Hölmänder condition the integral defining  $\tilde{S}f(x)$  converges. Fix a cube  $Q$ . We define

$$(19.19) \quad Sf(x) := \tilde{S}_Q f(x) \text{ modulo } \mathbb{C} \text{ on } \text{Int}(Q).$$

We shall claim that  $S$  is defined unambiguously despite the choice of  $Q$  and hence  $S$  is the desired operator.

A property of  $S$  Suppose that  $R_1$  and  $R_2$  are cubes. We shall claim that  $\tilde{S}_{R_1}$  and  $\tilde{S}_{R_2}$  differ by a constant function on  $R_1 \cup R_2$ . Thus, the definition of  $S$  makes sense. To do this, by choosing a cube  $R$  larger than  $R_1$  and  $R_2$ , and replacing  $R_2$  with  $R$  if necessary, we may assume that  $R_1 \subset R_2$ . Indeed,

$$\begin{aligned} \tilde{S}_{R_1}(x) - \tilde{S}_{R_2}(x) &= T(\chi_{R_1 \setminus R_2} \cdot f)(x) - \int_{R_2 \setminus R_1} (K(x, y) - K(c_{R_1}, y))f(y) dy \\ &\quad + \int_{R_2 \setminus R_1} (K(c_{R_1}, y) - K(c_{R_2}, y))f(y) dy \\ &= \int_{R_2 \setminus R_1} (K(c_{R_1}, y) - K(c_{R_2}, y))f(y) dy. \end{aligned}$$

Thus, it follows that  $\tilde{S}_{R_1} - \tilde{S}_{R_2}$  is constant on  $R_1$ .

$S$  is an  $L^\infty(\mathbb{R}^d)$ -BMO( $\mathbb{R}^d$ ) bounded operator. Let  $R$  be a cube. Keeping the property of  $S$  in mind, we have

$$(19.20) \quad m_R(|Sf - m_R(Sf)|) = \frac{1}{|R|^2} \int_R \left| \int_R Sf(x) - Sf(y) dy \right| dx.$$

Thus, we have

$$\begin{aligned} m_R(|Sf - m_R(Sf)|) &\leq \frac{2}{|R|} \int_{\mathbb{R}^d} |T[\chi_{2R} \cdot f]| \\ &\quad + \frac{1}{|R|^2} \iiint_{R^2 \times \mathbb{R}^d \setminus 2R} |K(x, z) - K(y, z)| \cdot |f(z)| dx dy dz \\ &=: \text{I} + \text{II}. \end{aligned}$$

For the estimate of I we again use the  $L^2(\mathbb{R}^d)$ -boundedness of  $T$ .

$$(19.21) \quad \text{I} \lesssim \left( \frac{1}{|R|} \int_{\mathbb{R}^d} |T[\chi_{2R} \cdot f](y)|^2 dy \right)^{\frac{1}{2}} \lesssim \left( \frac{1}{|R|} \int_{2R} |f(y)|^2 dy \right)^{\frac{1}{2}} \lesssim \|f\|_\infty.$$

As for II we use the Hölmänder condition :

$$(19.22) \quad \text{II} \lesssim \ell(R) \int_{\mathbb{R}^d \setminus 2R} \frac{|f(z)|}{|z - c_R|^{d+1}} dz \lesssim \|f\|_\infty.$$

Consequently  $m_R(|Sf - m_R(Sf)|) \lesssim \|f\|_\infty$  is established.

Proof of (19.16) From the kernel condition of  $T$  we have

$$(19.23) \quad \tilde{S}f(x) = Tf(x) - \int_{\mathbb{R}^d \setminus Q} K(c_Q, y) f(y) dy$$

for  $f \in L^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . Since  $\int_{\mathbb{R}^d \setminus Q} K(c_Q, y) f(y) dy$  is a constant, (19.16) is established.

Proof of (19.17) From the definition of the  $H^1(\mathbb{R}^d)$ -norm, we may assume that  $g$  is an atom with associated cube  $R$ . As we have verified, the function

$$(19.24) \quad \tilde{S}f(x) - \left( T(\chi_{2R} \cdot f)(x) + \int_{\mathbb{R}^d \setminus 2R} (K(x, y) - K(c_{2R}, y)) f(y) dy \right)$$

is constant. Therefore, we can assume that  $Q = 2R$ . Once we assume this, it is easy to see that  $Sf$  is square integrable on  $R$ . By using the moment condition we calculate  $\int_{\mathbb{R}^d} Sf(x)g(x) dx$ .

$$\begin{aligned} \int_{\mathbb{R}^d} Sf(x)g(x) dx &= \int_{\mathbb{R}^d} T[\chi_{2R} \cdot f] \cdot g + \iint_{(\mathbb{R}^d \setminus 2R) \times R} (K(x, y) - K(c_Q, y)) f(y) g(x) dx dy \\ &= \int_{2R} f \cdot T^*g + \iint_{(\mathbb{R}^d \setminus 2R) \times R} K(x, y) f(y) g(x) dx dy \end{aligned}$$

Since  $T^*$  is a CZ-operator as well as  $T$ , we obtain

$$\int_{\mathbb{R}^d} Sf(x)g(x) dx = \int_{2R} f(y)T^*g(y) dy + \int_{\mathbb{R}^d \setminus 2R} f(y)T^*g(y) dy = \int_{\mathbb{R}^d} f(y)T^*g(y) dy.$$

Therefore,  $S$  satisfies all the properties required. □

The above theorem presents us with an element in  $\text{BMO} \setminus L^\infty(\mathbb{R}^d)$ .

**Example 19.7.** Let  $H$  be the Hilbert transform. Then

$$(19.25) \quad H\chi_{[-1,1]}(x) = \log \left| \frac{x-1}{x+1} \right| \in \text{BMO}(\mathbb{R}) \setminus L^\infty(\mathbb{R}).$$

## 19.2. The John-Nirenberg inequality.

We have seen the size of  $\text{BMO}(\mathbb{R}^d)$  functions very roughly in Proposition 19.2. The John-Nirenberg inequality shows us more precise information about how large  $\text{BMO}(\mathbb{R}^d)$  functions can be. The main inequalities in this section are all called the John-Nirenberg inequality.

**Theorem 19.8** (John-Nirenberg inequality-1). *There exist  $c_1, c_2 > 0$  so that*

$$(19.26) \quad |Q \cap \{|f - m_Q(f)| > \lambda\}| \leq c_1 |Q| \exp\left(-\frac{c_2 \lambda}{\|f\|_*}\right)$$

for all cubes  $Q$  and  $f \in \text{BMO} \setminus \{0\}$ .

*Proof.* We may also assume  $\lambda \geq 2^{d+1}$  by choosing  $c_1 \geq \exp(c_2 2^{d+1})$  because the left-hand side is always less than  $|Q|$ . Let us establish by induction that

$$(19.27) \quad |Q \cap \{|f - m_Q(f)| > k 2^{d+1} \|f\|_*\}| \leq 2^{1-k}.$$

If  $k = 1$ , then we have nothing to prove. Assuming

$$(19.28) \quad |R \cap \{|f - m_R(f)| > k 2^{d+1} \|f\|_*\}| \leq 2^{1-k} |R|$$

for all cubes  $R$ , let us establish

$$(19.29) \quad |Q \cap \{|f - m_Q(f)| > (k+1) 2^{d+1} \|f\|_*\}| \leq 2^{-k} |Q|.$$

Form the Carderón-Zygmund decomposition at height  $2\|f\|_*$ . Then we obtain a collection of cubes  $\{Q_j\}_{j \in J}$  such that

$$(19.30) \quad 2\|f\|_* \leq m_{Q_j}(|f - m_{Q_j}(f)|) \leq 2^{d+1}\|f\|_*, \sum_{j \in J} \chi_{Q_j} \leq 1, \text{ dx - a.e.}$$

and that

$$(19.31) \quad |f(x) - m_{Q_j}(f)| \leq 2\|f\|_*, \text{ dx - a.e. } x \in Q \setminus \bigcup_{j \in J} Q_j.$$

By induction assumption we have

$$(19.32) \quad |Q_j \cap \{|f - m_{Q_j}(f)| > k 2^{d+1}\|f\|_*\}| \leq 2^{1-k}|Q_j|$$

for all  $j \in J$ . Observe also that

$$(19.33) \quad \sum_{j \in J} |Q_j| \leq \frac{1}{2}|Q|$$

from the property of  $\{Q_j\}_{j \in J}$ .

Therefore it follows that

$$\begin{aligned} & |Q \cap \{|f - m_Q(f)| > (k+1)2^{d+1}\|f\|_*\}| \\ & \leq \sum_{j \in J} |Q_j \cap \{|f - m_{Q_j}(f)| > (k+1)2^{d+1}\|f\|_*\}| \\ & \leq \sum_{j \in J} |Q_j \cap \{|f - m_{Q_j}(f)| > 2^{d+1}\|f\|_*\}| \\ & \leq \sum_{j \in J} 2^{1-k}|Q_j| \\ & \leq 2^{-k}|Q|. \end{aligned}$$

Therefore, the desired result follows. □

*Exercise 150.* Let  $d = 1$  and  $H$  denote the Hilbert transform. Then calculate

$$(19.34) \quad |[-1, 1] \cap \{|f - m_{[-1,1]}(f)| > \lambda\}|$$

when  $f = H\chi_{[-1,1]}$ . Conclude from this that Theorem 19.8 is sharp.

**Theorem 19.9** (John-Nirenberg inequality-2). *Let  $1 < p < \infty$ . A locally integrable function  $f$  belongs to  $\text{BMO}(\mathbb{R}^d)$ , if and only if*

$$(19.35) \quad \|f\|_{*,p} := \sup_{Q \in \mathcal{Q}} m_Q(|f - m_Q(f)|^p)^{\frac{1}{p}} < \infty.$$

*In this case we have  $\|f\|_* \leq \|f\|_{*,p} \lesssim_p \|f\|_*$ .*

*Proof.* We have only to show

$$(19.36) \quad m_Q(|f - m_Q(f)|^p)^{\frac{1}{p}} \lesssim_p \|f\|_*,$$

because  $\|f\|_* \leq \|f\|_{*,p}$  is trivial by the Hölder inequality. To do this, we write the left-hand side by distribution formula.

$$(19.37) \quad m_Q(|f - m_Q(f)|^p)^{\frac{1}{p}} = \left( \frac{1}{|Q|} \int_0^\infty p \lambda^{p-1} |Q \cap \{|f - m_Q(f)| > \lambda\}| d\lambda \right)^{\frac{1}{p}}.$$

We insert the John-Nirenberg inequality to obtain

$$(19.38) \quad m_Q(|f - m_Q(f)|^p)^{\frac{1}{p}} \lesssim \left( \int_0^\infty \lambda^{p-1} \exp\left(-\frac{c'\lambda}{\|f\|_*}\right) d\lambda \right)^{\frac{1}{p}} = c\|f\|_*.$$

This is the desired result. □

As well as Theorem 19.8, Theorem 19.9 is referred to as the John-Nirenberg inequality.

The same idea can be used to characterize the  $BMO(\mathbb{R}^d)$ -norm.

**Theorem 19.10** (John-Nirenberg inequality 3). *A locally function belongs to  $BMO(\mathbb{R}^d)$  if and only if it satisfies the John-Nirenberg inequality, that is, there exists  $c_1, c_2 > 0$  so that*

$$(19.39) \quad |Q \cap \{|f - m_Q(f)| > \lambda\}| \leq c_1 |Q| \exp(-c_2 \lambda)$$

for all cubes  $Q$ .

*Exercise 151.* Prove Theorem 19.10. Hint: By the John-Nirenberg inequality the “only if” part is already established. To prove the “if” part, re-examine the proof of Theorem 19.9.

To conclude this section we give an example of the Carleson measure.

**Proposition 19.11.** *Suppose that  $\psi \in C_c^\infty(\mathbb{R}^d)$  has zero integral and that  $\psi$  is radial. Let*

$$\psi_t(x) = \frac{1}{t^d} \psi\left(\frac{x}{t}\right) \quad (t > 0, x \in \mathbb{R}^d).$$

If  $b \in BMO(\mathbb{R}^d)$ , then

$$\mu := |\psi_t * b(x)|^2 \frac{dx dt}{t}$$

is a Carleson measure.

*Proof.* Let  $Q \in \mathcal{Q}$  be a fixed cube. We are going to show that

$$(19.40) \quad \int_0^{\ell(Q)} \left( \int_Q |\psi_t * b(x)|^2 dx \right) \frac{dt}{t} \lesssim \|b\|_* |Q|.$$

We set

$$\begin{aligned} \text{I} &:= \int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\psi_t * [(b - m_{2Q}(b))\chi_{2Q}](x)|^2 dx \\ \text{II} &:= \int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\psi_t * [(b - m_{2Q}(b))\chi_{\mathbb{R}^d \setminus 2Q}](x)|^2 dx. \end{aligned}$$

We decompose the left-hand side of (19.40):

$$\int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\psi_t * b(x)|^2 dx = \text{I} + \text{II}.$$

By the John-Nirenberg inequality and the Fubini theorem, we have

$$\begin{aligned} \text{I} &= \int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\mathcal{F}\psi(t\xi)|^2 |\mathcal{F}[(b - m_{2Q}(b))\chi_{2Q}](\xi)|^2 d\xi \\ &\lesssim \int_0^\infty \left( \int_Q \max(|t\xi|, |t\xi|^{-1})^2 |\mathcal{F}[(b - m_{2Q}(b))\chi_{2Q}](\xi)|^2 d\xi \right) \frac{dt}{t} \\ &= \int_Q \left( \int_0^\infty \max(|t\xi|, |t\xi|^{-1})^2 \frac{dt}{t} \right) |\mathcal{F}[(b - m_{2Q}(b))\chi_{2Q}](\xi)|^2 d\xi. \end{aligned}$$

Since

$$\int_0^\infty \max(|t\xi|, |t\xi|^{-1})^2 \frac{dt}{t} = \int_0^\infty \max(|t|, |t|^{-1})^2 \frac{dt}{t}$$

is a finite constant independent of  $\xi \neq 0$ , we have,

$$\text{I} \lesssim \int_Q |\mathcal{F}[(b - m_{2Q}(b))\chi_{2Q}](\xi)|^2 d\xi \lesssim \int_{\mathbb{R}^d} |\mathcal{F}[(b - m_{2Q}(b))\chi_{2Q}](\xi)|^2 d\xi.$$

By the Planchrel inequality, we obtain  $I \lesssim \int_{2Q} |b(x) - m_{2Q}(b)|^2 d\xi \lesssim |Q| \cdot \|b\|_*$ . Meanwhile,  $\text{II}$  is estimated in a standard way.

$$\begin{aligned} \text{II}^{\frac{1}{2}} &= \sum_{j=1}^{\infty} \left( \int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\psi_t * [(b - m_{2Q}(b))\chi_{2^{j+1}Q \setminus 2^j Q}](x)|^2 dx \right) \\ &= \sum_{j=1}^{N_\psi} \left( \int_0^{\ell(Q)} \frac{dt}{t} \int_Q |\psi_t * [(b - m_{2Q}(b))\chi_{2^{j+1}Q \setminus 2^j Q}](x)|^2 dx \right) \\ &\lesssim \sqrt{|Q| \cdot \|b\|_*}. \end{aligned}$$

This proves (19.40). □

**19.3. Duality  $H^1$ -BMO( $\mathbb{R}^d$ ).**

Finally, we prove the relation between  $H^1(\mathbb{R}^d)$  and  $\text{BMO}(\mathbb{R}^d)$ . Here, we denote by  $\mathbb{K}$  either  $\mathbb{C}$  or  $\mathbb{R}$ .

**Theorem 19.12** ( $H^1$ -BMO duality). *We can say that  $\text{BMO}(\mathbb{R}^d)$  is a dual space of  $H^1(\mathbb{R}^d)$  in the following sense.*

- (1) *Suppose that  $f \in \text{BMO}(\mathbb{R}^d)$ . Then the functional  $g \in H^1(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} f(x) \cdot g(x) dx \in \mathbb{K}$  is unambiguously defined despite the difference of additive constants and it is continuous on  $H^1(\mathbb{R}^d)$ .*
- (2) *Conversely, any continuous functional on  $H^1(\mathbb{R}^d)$  can be realized with a certain  $f \in \text{BMO}(\mathbb{R}^d)$ .*

*Proof.* (1) is considerably easy to prove. Indeed, as before, we have only to deal with an atom  $a$  instead of  $g \in H^1(\mathbb{R}^d)$ . Suppose that  $Q$  is a cube associated to  $a$ . Then we have

$$\left| \int_{\mathbb{R}^d} f(x)a(x) dx \right| = \left| \int_Q (f(x) - m_Q(f))a(x) dx \right| \leq \left( \int_Q |f(x) - m_Q(f)|^2 dx \right)^{\frac{1}{2}} \cdot \|a\|_2.$$

By virtue of the John-Nirenberg inequality (Theorem 19.8) Suppose that  $x \in R \setminus \bigcup_j R_j$ . we have

$$\left( \int_Q |f(x) - m_Q(f)|^2 dx \right)^{\frac{1}{2}} \lesssim |Q|^{-\frac{1}{2}}.$$

Therefore, we conclude

$$\left| \int_{\mathbb{R}^d} f(x)a(x) dx \right| \lesssim \|f\|_*.$$

Now we turn to (2). Let  $\Phi$  be a continuous functional on  $H^1(\mathbb{R}^d)$ . Fix a cube  $Q$ . Then for all  $g \in L^2(\mathbb{R}^d)$  with  $\text{supp}(g) \subset Q$ , we have

$$\|g - m_Q(g)\|_{H^1} \lesssim |Q|^{\frac{1}{2}} \|g\|_2$$

and hence

$$|\Phi(g - m_Q(g))| \lesssim \|\Phi\|_* |Q|^{\frac{1}{2}} \|g\|_2.$$

By the Riesz representation theorem (see Theorem 11.7), there exists  $f_Q \in L^2(Q)$  so that, for every  $g \in L^2(\mathbb{R}^d)$ ,

$$(19.41) \quad \Phi(g - m_Q(g)) = \int_{\mathbb{R}^d} g(x)f_Q(x) dx, \quad \|f_Q\|_2 \lesssim |Q|^{\frac{1}{2}}.$$

By virtue of the uniqueness of Theorem 11.7, we have

$$f_Q(x) = f_R(x) - m_Q(f_R) \text{ on } Q,$$

for a.e.  $x \in \mathbb{R}^d$ . Let  $Q, R, S$  be cubes such that  $Q \subset R \subset S$ . Then we have

$$\begin{aligned} f_S(x) - m_Q(f_S) &= f_R(x) + m_R(f_S) - m_Q(f_S) \\ &= f_R(x) - m_Q(f_S - m_R(f_S)) \\ &= f_R(x) - m_Q(f_R) \end{aligned}$$

for a.e.  $x \in R$ .

Let  $Q$  be fixed again. We have to show

$$(19.42) \quad m_Q(|f - m_Q(f)|) \lesssim \|\Phi\|_*.$$

To do this, we let  $g = \overline{f_Q - m_Q(f_Q)}$  in (19.41). Then we obtain

$$m_Q(|f - m_Q(f)|) = m_Q(|f_Q - m_Q(f_Q)|) \leq m_Q(|f_Q - m_Q(f_Q)|^2)^{\frac{1}{2}}$$

by the Hölder inequality. Next, we express the right-hand side in terms of  $\Phi$ .

$$\begin{aligned} m_Q(|f - m_Q(f)|) &\leq \left( \frac{1}{|Q|} \int_Q g(x)(f_Q(x) - m_Q(f)) \right)^{\frac{1}{2}} \\ &= \left( \frac{1}{|Q|} \int_Q g(x)f_Q(x) dx \right)^{\frac{1}{2}} \\ &= \left( \frac{1}{|Q|} \Phi(g) \right)^{\frac{1}{2}}. \end{aligned}$$

Inserting the norm estimate of  $g$ , we finally obtain

$$m_Q(|f - m_Q(f)|) \lesssim \|\Phi\|_* |Q|^{\frac{1}{2}} m_Q(|f - m_Q(f)|)^{\frac{1}{2}}.$$

Therefore we have

$$(19.43) \quad m_Q(|f - m_Q(f)|) \lesssim \|\Phi\|_*.$$

(2) is therefore proved.  $\square$

The following exercise is similar to (2) of the above theorem in that it provides a way to obtain global information by patching local information.

*Exercise 152.* Given an open set  $O$ , we denote by  $\mathcal{O}(O)$  the set of all holomorphic functions on  $O$ . It will be understood that  $\mathcal{O}(\emptyset) = \{0\}$ . Prove the following.

- (1) Let  $U, V$  and  $W$  be open set satisfying  $V \subset U$ . Then we define the restriction mapping  $r_{VU}$  by

$$(19.44) \quad r_{VU} : f \in \mathcal{O}(U) \rightarrow f|_V \in \mathcal{O}(V).$$

If  $V$  is an open set, it will be understood that  $r_{VU}(f) = 0$  for all  $f \in \mathcal{O}(U)$ . We also define  $r_{WV}$  and  $r_{WU}$  analogously. Then prove that  $r_{WV} \circ r_{VU} = r_{WU}$ .

- (2) Let  $O$  be an open set and  $\{O_\lambda\}_{\lambda \in \Lambda}$  its open covering. Assume that  $f, g \in \mathcal{O}(O)$   $f|_{O_\lambda} = g|_{O_\lambda}$  for all  $\lambda \in \Lambda$ . Then  $f = g$  on  $O$ .
- (3) Let  $O$  be an open set and  $\{O_\lambda\}_{\lambda \in \Lambda}$  its open covering. Suppose that we are given a family of holomorphic functions  $\{f_\lambda\}_{\lambda \in \Lambda}$  satisfying

$$(19.45) \quad f_\lambda \in \mathcal{O}(O_\lambda), r_{O_\lambda \cap O_\mu} O_\lambda(f_\lambda) = r_{O_\lambda \cap O_\mu} O_\mu(f_\mu)$$

for all  $\lambda, \mu \in \Lambda$ . Then there exists a unique holomorphic function  $f$  on  $O$  so that

$$(19.46) \quad f|_{O_\lambda} = f_\lambda$$

for all  $\lambda \in \Lambda$ .

Let us generalize the property above.

**Definition 19.13.** Let  $X$  be a topological space and  $\mathcal{O}_X$  denote the set of all open sets.  $(\{\mathcal{S}(\mathbb{R}^d)(U)\}_{U \in \mathcal{O}_X}, \{r_{VU}\}_{V, U \in \mathcal{O}_X: V \subset U})$  is said to be a sheaf of rings, if it satisfies the following conditions.

- (1) For each  $U \in \mathcal{O}_X$ ,  $\mathcal{S}(\mathbb{R}^d)(U)$  is a ring.
- (2) For each  $U, V \in \mathcal{O}_X$  with  $V \subset U$ ,  $r_{VU}$  is a linear homomorphism and

$$(19.47) \quad r_{WU} = r_{WV} \circ r_{VU}$$

for all triples of open sets  $(U, V, W)$  with  $W \subset V \subset U$ .

- (3) Let  $U = \bigcup_{\lambda \in \Lambda} U_\lambda$  be an open covering of an open set  $U$ . Suppose that  $f \in \mathcal{S}(\mathbb{R}^d)(U)$ . If  $r_{U_\lambda U}(f) = 0$  for all  $\lambda \in \Lambda$ , then  $f = 0 \in \mathcal{S}(\mathbb{R}^d)(U)$ .

- (4) Let  $U = \bigcup_{\lambda \in \Lambda} U_\lambda$  be an open covering of an open set  $U$ . Suppose that we are given  $f_\lambda \in \mathcal{S}(\mathbb{R}^d)(U_\lambda)$  for each  $\lambda \in \Lambda$ . If  $r_{U_\lambda \cap U_\rho U_\lambda}(f_\lambda) = r_{U_\lambda \cap U_\rho U_\rho}(f_\rho)$  for all  $\lambda, \rho \in \Lambda$ , then there exists  $f \in \mathcal{S}(\mathbb{R}^d)(U)$  such that  $f_\lambda = r_{U_\lambda U}(f)$  for all  $\lambda \in \Lambda$ .

**Example 19.14** (Sheaf v.s. Banach space). Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . The following is a list of examples of sheaves and Banach spaces. For the definition of  $D(\Omega)$  and  $D'(\Omega)$  we refer to Subsection 32

	Sheaf	Banach space
$BC(\Omega)$	No	Yes
$C(\Omega)$	Yes	No
$C^m(\Omega)$	No	Yes
$C^\infty(\Omega)$	Yes	No
$L^p(\Omega)$	No	Yes
$BMO(\mathbb{R}^d)$	No	Yes
$H^1(\mathbb{R}^d)$	No	Yes
$L^1_{loc}(\Omega)$	Yes	No
$D(\Omega)$	Yes	No
$D'(\Omega)$	Yes	No

From this chart we can say sheaves reflect local structures of things while Banach space deals with global structures. We remark that for the sake of convenience of this book we regard  $C^m(\Omega)$  as a Banach space while  $C^\infty(\Omega)$  is a sheaf.

## 20. SHARP-MAXIMAL OPERATORS

### 20.1. Definition.

Motivated by the definition of  $BMO(\mathbb{R}^d)$ , we define the sharp-maximal operator. The definition arose from the paper [183] which deals systematically with Hardy spaces and the  $BMO(\mathbb{R}^d)$  space.

**Definition 20.1.** [183] For a locally integrable function  $f$ , define

$$(20.1) \quad M^\sharp f(x) := \sup_{x \in Q} m_Q(|f - m_Q(f)|) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y) - m_Q(f)| dy,$$

where  $Q$  runs over all cubes containing  $x$ .

As is easily seen, we have

- (1) The operator  $M^\sharp$  annihilates the constant functions.
- (2)  $M^\sharp f \leq 2Mf$  for all measurable functions  $f$ .
- (3) Let  $1 < p \leq \infty$ . Then  $M^\sharp$  is  $L^p(\mathbb{R}^d)$ -bounded.
- (4) Let  $f \in \text{BMO}(\mathbb{R}^d)$ .  $M^\sharp f$  makes sense despite its ambiguity of additive constants.

To be familiarize the definition let us prove the following.

**Lemma 20.2.** *We define an auxiliary maximal operator  $M^{\sharp,*}$  by*

$$(20.2) \quad M^{\sharp,*}f(x) = \sup_{x \in Q \in \mathcal{Q}} \inf_{a \in \mathbb{C}} m_Q(|f - a|).$$

*Then we have*

$$(20.3) \quad M^{\sharp,*}f(x) \leq M^\sharp f(x) \leq 2M^{\sharp,*}f(x)$$

*for all measurable functions  $f$ .*

*Proof.* The left inequality is obvious: Indeed, let  $Q$  be a fixed cube containing  $x$ . If we take  $a = m_Q(f)$ , we see

$$(20.4) \quad \inf_{b \in \mathbb{C}} m_Q(|f - b|) \leq m_Q(|f - a|) \leq M^\sharp f(x).$$

Since  $Q$  is arbitrary, we see  $M^{\sharp,*}f(x) \leq M^\sharp f(x)$ . Let us prove the right inequality. To do this, we take  $a \in \mathbb{C}$  arbitrarily and a cube containing  $x$ . Then we have

$$(20.5) \quad m_Q(|f - m_Q(f)|) = m_Q(|f - a - m_Q(f - a)|) \leq m_Q(|f - a|) + m_Q(m_Q(|f - a|)) = 2m_Q(|f - a|).$$

Therefore, taking the infimum over  $a \in \mathbb{C}$ , we obtain

$$(20.6) \quad m_Q(|f - m_Q(f)|) \leq 2 \inf_{a \in \mathbb{C}} m_Q(|f - a|).$$

Now we take the supremum over  $Q$  containing  $x$ , we have

$$(20.7) \quad M^\sharp f(x) \leq 2M^{\sharp,*}f(x).$$

This is the desired result.  $\square$

*Exercise 153.* Show that  $M^\sharp[|f|](x) \lesssim M^{\sharp,*}f(x)$  for every locally integrable function  $f$ . Hint: Notice that

$$(20.8) \quad M^{\sharp,*}|f|(x) = \sup_{x \in Q \in \mathcal{Q}} \inf_{a \in \mathbb{C}} m_Q(|f| - a) = \sup_{x \in Q \in \mathcal{Q}} \inf_{a > 0} m_Q(|f| - a).$$

## 20.2. Good $\lambda$ -inequality.

To obtain an inequality, we use so called “good  $\lambda$ ”-inequality. Here we place ourselves in the setting of a measure space  $(X, \mathcal{B}, \mu)$ . Nowadays it is used not only in harmonic analysis but also in probability theory, as we see Chapter 20.

**Theorem 20.3** (Good- $\lambda$  inequality). *Let  $0 < p < \infty$ . Suppose that  $f$  and  $g$  are measurable functions. Assume that  $\min(1, |f|) \in L^p(\mu)$  and that there exists  $\beta > 0$ ,  $0 < \eta < 2^{-p}$  so that*

$$(20.9) \quad \mu\{|f| > \lambda, |g| \leq \beta\lambda\} \leq \eta\mu\{|f| > 2\lambda\}$$

*for all  $\lambda > 0$ . Then we have*

$$(20.10) \quad \|f\|_{L^p(\mu)} \lesssim_{\beta, \eta, p} \|g\|_{L^p(\mu)}.$$



Now we really need  $f$  to have some speed of decay at infinity, as the example  $f \equiv 1$  shows. We remark that the integrability assumption is often formulated as follows:

**Corollary 20.4.** *Let  $0 < p < \infty$ . Suppose that  $f$  and  $g$  are measurable functions with  $f \in L^{p_0}$  for some  $p_0 < p$ . If the inequality (20.9) holds, (20.10) is still available.*

Since  $\min(1, |f|) \in L^p(\mu)$  is weaker than the assumption that  $f \in L^{p_0}(\mu)$  for some  $0 < p_0 < p$ , it follows that the corollary is an immediate consequence of the theorem.

*Proof of Theorem 20.3.* We use the distribution formula. Let  $R > 0$  be fixed.

$$\begin{aligned} \int_0^{2R} p \lambda^{p-1} \mu\{|f| > \lambda\} d\lambda &= 2^p \int_0^R p \lambda^{p-1} \mu\{|f| > 2\lambda\} d\lambda \\ &\leq 2^p \int_0^R p \lambda^{p-1} (\mu\{|f| > 2\lambda, |g| \leq \beta\lambda\} + \mu\{|g| > \beta\lambda\}) d\lambda \\ &\lesssim \int_X |g(x)|^p d\mu(x) + 2^p \eta \int_0^{2R} p \lambda^{p-1} \mu\{|f| > \lambda\} d\lambda. \end{aligned}$$

Assuming  $\min(1, |f|) \in L^p(\mu)$ , we have

$$2^p \eta \int_0^{2R} p \lambda^{p-1} \mu\{|f| > \lambda\} d\lambda = \int_X \min(2R, |f(x)|)^p d\mu(x) < \infty.$$

Thus, we are now in the position of bringing the second term of the right-hand side to the left-hand side. The result is

$$(20.11) \quad \int_0^{2R} p \lambda^{p-1} \mu\{|f| > \lambda\} d\lambda \lesssim \int_X |g(x)|^p d\mu(x).$$

Here the implicit constant in  $\lesssim$  does not depend on  $R$ . Letting  $R \rightarrow \infty$ , we obtain the desired result.  $\square$

**20.3. Sharp-maximal inequality.**

We return to the setting of the usual  $\mathbb{R}^d$  coming with the Lebesgue measure  $dx$ .

**Theorem 20.5** (Sharp maximal inequality). *Let  $1 < p < \infty$ . Then*

$$(20.12) \quad \|f\|_p \lesssim_p \|M^\sharp f\|_p$$

for every measurable function  $f$  with  $\min(Mf, 1) \in L^p(\mathbb{R}^d)$ .

We really have to assume that  $\min(1, Mf) \in L^p(\mathbb{R}^d)$  as the example  $f \equiv 1$  again shows.

*Proof.* To prove this theorem, we have only to prove the corresponding good  $\lambda$ -inequality

$$(20.13) \quad |\{Mf > 2\lambda : M^\sharp f \leq \beta\lambda\}| \lesssim \beta |\{Mf > \lambda\}|.$$

To do this we set

$$E_\lambda := \{Mf > 2\lambda : M^\sharp f \leq \beta\lambda\}, F_\lambda := \{Mf > \lambda\}.$$

Let  $x \in E_\lambda$ . Then by definition, there exists a cube  $Q = Q_x$  so that  $m_Q(|f|) > \frac{3}{2}\lambda$ . If we replace  $Q$  with larger one, then we can assume

$$(20.14) \quad m_S(|f|) \leq \frac{3}{2}\lambda$$

for all cubes  $S$  that engulfs  $2Q$ . The crux of the proof is *not* to choose  $Q$  so that  $m_Q(|f|) > 2\lambda$ . Indeed, suppose that  $S_1, S_2, \dots, S_j, \dots$  are cubes such that  $2S_j \subset S_{j+1}$ , then  $\lim_{j \rightarrow \infty} m_{S_j}(|f|) = 0$ . Therefore, such a cube  $Q$  exists.

Use the Besicovitch covering lemma, and we obtain the countable family of cubes  $\{S_j\}_{j \in J}$  so that

$$(20.15) \quad m_{S_j}(|f|) > \frac{3}{2}\lambda, \quad \sum_{j \in J} \chi_{S_j} \lesssim 1.$$

We claim

$$(20.16) \quad |S_j \cap E_\lambda| \lesssim \beta |S_j|$$

assuming that  $0 < \beta \ll 1$ . Once we prove (20.16), then we have

$$(20.17) \quad |E_\lambda| \lesssim \beta \sum_{j \in J} |S_j| \lesssim \beta \left| \bigcup_{j \in J} S_j \right| \lesssim \beta |F_\lambda|.$$

Thus, the theorem is proved modulo the claim. Let  $x \in S_j \cap E_\lambda$ . Then there exists a cube  $R$  such that  $m_R(|f|) > 2\lambda$ . If  $R$  were not contained in  $\frac{5}{4}S_j$ , then  $100R$  engulfs  $2S_j$  and hence

$$(20.18) \quad m_{100R}(|f|) \geq m_R(|f|) - |m_R(|f|) - m_{100R}(|f|)| \geq 2\lambda - cM^\sharp f(x) \geq 2\lambda - c\beta\lambda \geq \frac{7}{4}\lambda,$$

contradicts (20.14). Therefore, it follows that

$$(20.19) \quad M\left(\chi_{\frac{5}{4}S_j} \cdot f\right)(x) > 2\lambda.$$

If we use this observation, we obtain

$$(20.20) \quad M\left(\chi_{\frac{5}{4}S_j} \cdot (f - m_{S_j}(f))\right)(x) \geq M\left(\chi_{\frac{5}{4}S_j} \cdot f\right)(x) - |m_{S_j}(f)| \geq 2\lambda - \frac{3}{2}\lambda = \frac{1}{2}\lambda.$$

If we had replace  $\frac{3}{2}$  in (20.14) by 2, then our calculation would not have worked. Therefore, if we use the weak-(1, 1) boundedness of  $M$ , then we obtain

$$(20.21) \quad |S_j \cap E_\lambda| \lesssim \frac{1}{\lambda} \int_{\frac{5}{4}S_j} |f(x) - m_{S_j}(f)| dx \lesssim \beta |S_j|.$$

Therefore the claim is proved.  $\square$

Now we obtain an alternative proof of Theorem 20.5.

**Theorem 20.6.** *Let  $f, g$  be locally integrable functions. Assume that  $\{|f| > \lambda\} < \infty$  for all  $\lambda > 0$ . Then we have*

$$(20.22) \quad \int_{\mathbb{R}^d} |f(x)w(x)| dx \leq \int_{\mathbb{R}^d} M^\sharp f(x) M_{\text{dyadic}} w(x) dx.$$

Once this theorem is established, it is easy to prove Theorem 20.5.

*Proof.* Let  $k \in \mathbb{Z}$  be fixed. We can assume that  $f$  is bounded and that  $w$  is positive, compactly supported and bounded. We partition  $\{M_{\text{dyadic}} w > 2^k\}$  into a collection of maximal dyadic cubes such that  $m_{Q_j^k}(w) > 2^k$ . We write  $\{M_{\text{dyadic}} w > 2^k\} = \prod_{j \in J^k} Q_j^k$ . Define

$$(20.23) \quad G^k(x) = \chi_{\{M_{\text{dyadic}} w \leq 2^k\}}(x)w(x) + \sum_{j \in J^k} m_{Q_j^k}(w)\chi_{Q_j^k}(x), \quad B^k(x) = w(x) - G^k(x).$$

Then for each  $j' \in J^{k+1}$ , there exists  $j \in J^k$  such that  $Q_{j'}^{k+1} \subset Q_j^k$ . Therefore, we have  $\int_{Q_j^k} (B^{k+1}(x) - B^k(x)) dx = 0$ . In view of the property of the good part of the Calderón-Zygmund decomposition, we have  $|B^k(x) - B^{k+1}(x)| = |G^k(x) - G^{k+1}(x)| \leq 3 \cdot 2^d$ . Since  $w \in L^\infty(\mathbb{R}^d)$ , there exists  $K \in \mathbb{N}$  such that  $B^k \equiv 0$  whenever  $k \geq K$ . Therefore,

$$\begin{aligned} \int_{\mathbb{R}^d} f(x)g(x) dx &= \int_{\mathbb{R}^d} f(x)G^k(x) dx + \int_{\mathbb{R}^d} f(x)B^k(x) dx \\ &= \int_{\mathbb{R}^d} f(x)G^k(x) dx + \sum_{l=k}^{\infty} \int_{\mathbb{R}^d} f(x)(B^l(x) - B^{l+1}(x)) dx \\ &=: \mathbf{I}_k + \mathbf{II}_k. \end{aligned}$$

Since  $\mathbf{I}_k \leq \|f \cdot g\|_\infty \{ \int_{\{|f| > \lambda\}} + 2^k \varepsilon \int_{\{|f| \leq \varepsilon\}} \}$ , it follows that  $\lim_{k \rightarrow -\infty} \mathbf{I}_k = 0$ . As for  $\mathbf{II}_k$ , we have

$$\begin{aligned} |\mathbf{II}_k| &\leq \sum_{l=k}^{\infty} \sum_{j \in J^l} \int_{Q_j^l} |f(x) - m_{Q_j^l}(f)| \cdot |B^l(x) - B^{l+1}(x)| dx \\ &\leq 3 \sum_{l=k}^{\infty} 2^l \sum_{j \in J^l} \int_{Q_j^l} |f(x) - m_{Q_j^l}(f)| dx. \end{aligned}$$

A trivial but interesting observation is that

$$(20.24) \quad \int_{Q_j^l} |f(x) - m_{Q_j^l}(f)| w(x) dx \leq \int_{Q_j^l} M^\sharp f(x) dx.$$

If we invoke (20.24), then we have

$$|\mathbf{II}_k| \leq 3 \sum_{\substack{j \in J^l \\ l \in \mathbb{Z}}} 2^l \int_{Q_j^l} M^\sharp f(x) dx = 3 \sum_{l \in \mathbb{Z}} 2^l \int_{\{M^\sharp f > 2^l\}} M^\sharp f(x) dx \leq 6 \int_{\mathbb{R}^d} M^\sharp f(x) M_{\text{dyadic}} w(x) dx.$$

The proof is therefore complete. □

## 21. WEIGHTED NORM ESTIMATES

In this section we consider the maximal inequalities with the Lebesgue measure replaced by a Borel measure. Leaving the definition of maximal operator unchanged, we consider the weighted measure  $w dx$ . We consider the following problems ?

For what weights  $w$  is it the case that

$$\int_{\mathbb{R}^d} Mf(x)^p w(x) dx \lesssim \int_{\mathbb{R}^d} |f(x)|^p w(x) dx$$

for  $p \in (1, \infty)$  and that

$$\int_{\{Mf > \lambda\}} w(x) dx \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| w(x) dx \quad ?$$

The weight such that  $M$  is bounded on  $L^p(w dx)$  will be called the Muckenhoupt class. This definition dates back to the pioneering work [361].

21.1.  $A_1$ -weights.

This theorem will be a motivation of definition of various classes of weights for the Hardy-Littlewood maximal operator. In [182] for any locally integrable functions  $f, g$  and  $\lambda > 0$ , Fefferman and Stein proved

$$\int_{\{Mf > \lambda\}} |g(x)| dx \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)| Mg(x) dx.$$

The next result refines the result in [182].

**Theorem 21.1** (Dual inequality of Stein type). *For any locally integrable functions  $f, g$  and  $\lambda > 0$ , we have*

$$(21.1) \quad \int_{\{Mf > \lambda\}} |g(x)| dx \lesssim \frac{1}{\lambda} \int_{\{|f| > 5^{-d}\lambda\}} |f(x)| Mg(x) dx.$$

*Proof.* Let  $E_\lambda = \{Mf > \lambda\}$ . As before by the definition of  $E_\lambda$ , for all  $x \in E$  there exists  $r_x$  such that

$$m_{B(x, r_x)}(|f|) > \lambda.$$

By Theorem 12.11, there exists a disjoint subfamily of balls  $\{B_j\}_{j \in \Lambda}$  with  $\Lambda \subset E_\lambda$  that satisfies the following properties.

$$(21.2) \quad E_\lambda \subset \bigcup_{x \in E_\lambda} B(x, r_x) \subset \bigcup_{j \in \Lambda} 5B_j.$$

Let  $x \in 5B_j$ . Then the average of  $|g|$  over  $5B_j$  is bounded by the maximal operator :

$$(21.3) \quad m_{5B_j}(|g|) \leq Mg(x).$$

Using (21.2) and (21.3), we obtain

$$\int_{E_\lambda} |g(x)| dx \leq \int_{\bigcup_{j \in \Lambda} 5B_j} |g(x)| dx \leq \sum_{j \in \Lambda} \int_{5B_j} |g(x)| dx \leq \frac{5^d}{\lambda} \sum_{j \in \Lambda} \int_{B_j} |f(x)| dx \cdot m_{5B_j}(|g|).$$

Note that from the definition of the (uncentered) maximal operator

$$m_{5B_j}(|g|) \leq \inf_{z \in B_j} Mg(z).$$

Inserting this, we obtain

$$\int_{E_\lambda} |g(x)| dx \leq \frac{5^d}{\lambda} \sum_{j \in \Lambda} \int_{B_j} |f(x)| Mg(x) dx \leq \frac{5^d}{\lambda} \int_{\{|f| > 5^{-d}\lambda\}} |f(x)| Mg(x) dx,$$

which is the desired result. □

This corollary is an example of the usage of interpolation for the measure other than the Lebesgue measure  $dx$ .

**Corollary 21.2.** *Suppose that  $1 < p < \infty$ . Then one has*

$$(21.4) \quad \int_{\mathbb{R}^d} Mf(x)^p |g(x)| dx \lesssim_p \int_{\mathbb{R}^d} |f(x)|^p Mg(x) dx.$$

*Proof.* We write out the right-hand side out in full by using the distribution:

$$\int_{\mathbb{R}^d} Mf(x)^p |g(x)| dx = \int_{\mathbb{R}^d} \left( \int_0^\infty \lambda^{p-1} \chi_{\{x \in \mathbb{R}^d : Mf(x) > \lambda\}}(\lambda) \cdot |g(x)| d\lambda \right) dx.$$

If we change the order of integration by using the Fubini theorem and invoke Theorem 21.1, we obtain

$$(21.5) \quad \int_{\mathbb{R}^d} Mf(x)^p |g(x)| dx \lesssim_p \int_0^\infty \lambda^{p-2} \left( \int_{\{x \in \mathbb{R}^d : |f(x)| > 5^{-d}\lambda\}} |f(x)| Mg(x) dx \right) d\lambda.$$

Another application of the Fubini theorem yields

$$(21.6) \quad \int_0^\infty \lambda^{p-2} \left( \int_{\{x \in \mathbb{R}^d : |f(x)| > 5^{-d}\lambda\}} |f(x)| Mg(x) dx \right) d\lambda \simeq_p \int_{\mathbb{R}^d} |f(x)|^p Mg(x) dx.$$

Putting together (21.5) and (21.6), we obtain the desired result. □

Definition of the class  $A_1$ . Let us consider the following class of weights.

**Definition 21.3.** A locally integrable function  $w$  is said to be an  $A_1$ -weight, if  $w(x)$  is strictly positive and finite and there exists  $C_0$  such that

$$(21.7) \quad Mw(x) \leq C_0 w(x)$$

for a.e.  $x \in \mathbb{R}^d$ . The  $A_1$  constant, which is denoted by  $A_1(w)$  below, is the minimum of  $C_0$  satisfying (21.7).

*Exercise 154.* Let  $w$  be a weight and  $a > 0$ . Then show that  $A_1(w) = A_1(aw)$ . Be careful for the usage of the term “norm” : It is an abuse of language. Because  $A_1(w)$  is not a linear space over  $\mathbb{R}$ .

Here we present standard notations about weighted function spaces.

**Definition 21.4.** Denote by  $\text{Meas}(\mathbb{R}^d)$  the set of all Lebesgue measurable functions on  $\mathbb{R}^d$ .

- (1) Write  $L^p(w) := L^p(w dx)$  for the sake of simplicity.
- (2) A sublinear operator  $S$  from  $L^p(w)$  to  $\text{Meas}$  is said to be weak  $L^p(w)$ -bounded, if

$$\int_{\{x \in \mathbb{R}^d : Sf > \lambda\}} w(x) dx \lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^d} |f(x)|^p w(x) dx$$

for all  $f \in L^p(w)$ .

- (3) Given a weight  $w$  and a measurable set  $E$ , we denote

$$w(E) := \int_E w(x) dx.$$

When  $E = \{\dots\}$ , then we write  $w(E) := w\{\dots\}$ .

Key properties of the class  $A_1(w)$ . In this paragraph we relate  $A_1(w)$  with the weak- $L^1(w)$  boundedness of the Hardy-Littlewood maximal operator  $M$ .

**Theorem 21.5.** A weight  $w$  belongs to  $A_1$  if and only if  $M$  is weak  $L^1(w)$ -bounded, that is,

$$w\{x \in \mathbb{R}^d : Mf(x) > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx \quad (f \in L^1(w), \lambda > 0).$$

*Proof.* Assume that  $w$  is an  $A_1$ -weight. We have  $Mw(x) \lesssim w(x)$  by assumption. If we invoke Theorem 21.1, then we obtain

$$(21.8) \quad w\{x \in \mathbb{R}^d : Mf(x) > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|Mw(x) dx \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx,$$

showing the weak- $L^1(w)$  boundedness of  $M$ .

Assume instead that  $M$  is weak- $L^1(w)$  bounded. We shall show that interval testing suffices. More precisely, we proceed as follows: The assumption reads

$$(21.9) \quad w\{x \in \mathbb{R}^d : Mf(x) > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx.$$

We shall show that, if  $x$  is a Lebesgue point of  $w$ , that is, if  $x$  is a point satisfying

$$\lim_{\substack{R \in \mathcal{Q} \\ R \downarrow x}} m_R(w) = w(x),$$

then  $m_Q(w) \lesssim w(x)$  for any cube  $Q$  containing  $x$ . Let  $R$  be another cube containing  $x$  and contained by  $Q$ . Observe that  $M\chi_R > \frac{|R|}{2|Q|}$  on  $Q$ . Indeed, we have  $M\chi_R(x) = \sup_{x \in S \in \mathcal{Q}} \frac{|R \cap S|}{|S|} > \frac{|R|}{2|Q|}$ . Therefore we have  $w(Q) \leq w \left\{ M\chi_R > \frac{|R|}{2|Q|} \right\} \lesssim \frac{w(R)|Q|}{|R|}$ . Arranging this inequality, we obtain  $m_Q(w) \lesssim m_R(w)$ . Shrinking  $R$  to a point  $x$ , we obtain  $m_Q(w) \lesssim w(x)$ , which is a desired result.  $\square$

Example of  $A_1$ -weights. Now we present some examples of  $A_1$ -weights.

**Lemma 21.6.** *Let  $\mu$  be a Radon measure and set (temporarily)*

$$(21.10) \quad M\mu(x) := \sup \left\{ \frac{\mu(Q)}{|Q|} : Q \in \mathcal{Q}, c_Q = x \right\}.$$

Then we have  $|\{M\mu > \lambda\}| \leq \frac{N}{\lambda} \mu(\mathbb{R}^d)$ , where  $N$  is a covering constant of the Besicovitch covering lemma.

*Exercise 155.* Using the Besicovitch covering lemma, prove Lemma 21.6.

**Theorem 21.7.** *Suppose that  $0 < \delta < 1$ . Assume  $\mu$  is a Radon measure such that  $M\mu$  is finite for a.e.  $x \in \mathbb{R}^d$ . Then  $w = M\mu^\delta$  is an  $A_1$ -weight, where  $M\mu$  is given by (21.10). Furthermore,  $A_1(w) \lesssim_{\delta,d} 1$ .*

*Proof.* We have to establish  $m_Q(w) \lesssim w(x)$  whenever  $Q$  is a cube containing  $x$ .

We decompose  $\mu$  according to  $2Q$ . Denote  $(\mu|A)(E) = \mu(A \cap E)$ , the restriction of  $\mu$  to a measurable subset  $A$ . Split  $f$  by  $\mu = \mu_1 + \mu_2$  with  $\mu_1 = \mu|_{2Q}$  and  $\mu_2 = \mu|_{(\mathbb{R}^d \setminus 2Q)}$ . We shall prove

$$m_Q(Mf_1^\delta) \lesssim w(x), \quad m_Q(Mf_2^\delta) \lesssim w(x)$$

for all  $x \in Q$ . To deal with  $f_1$  we use the idea used in the proof of the Kolmogorov inequality (Theorem 12.14). Going through a similar argument in that proof, we obtain

$$m_Q(M\mu_1^\delta) \lesssim m_{2Q}(M\mu_1^\delta) \lesssim \left( \frac{\mu(2Q)}{|2Q|} \right)^\delta \lesssim w(x).$$

For the treatment of  $f_2$ , we shall make use of a pointwise estimate, which we encounter frequently. Let  $y \in Q$ . First, we write out  $Mf_2(y)$  in full:

$$M\mu_2(y) = \sup \left\{ \frac{\mu(R \setminus 2Q)}{|R|} : R \in \mathcal{Q}, y \in R \right\}.$$

A geometric observation readily shows  $\ell(R) \geq \frac{1}{2}\ell(Q)$  whenever  $R$  is a cube intersecting both  $Q$  and  $\mathbb{R}^d \setminus 2Q$ . Therefore such  $R$  engulfs  $Q$ , if we triple it. In view of this fact, we deduce

$$M\mu_2(y) \leq 3^d \sup \left\{ \frac{\mu(S)}{|S|} : S \in \mathcal{Q}, Q \subset S \right\} \leq 3^d M\mu(x).$$

This pointwise estimate readily gives  $m_Q(Mf_2^\delta) \lesssim w(x)$ . Therefore the theorem is now completely proved.  $\square$

The following proposition gives us a concrete example of  $A_1$ -weights.

**Proposition 21.8.** *Let  $0 < a < d$ . Then  $w(x) = |x|^{-a}$  belongs to  $A_1$ .*

*Proof.* It suffices to take  $\mu = \delta_0$ , the point measure massed on the origin, which is a special case of Theorem 21.7. Then  $w$  equals  $M\mu$  modulo multiplicative constants.  $\square$

Now we prove the converse of Theorem 21.7. For the proof, we need a lemma.

**Lemma 21.9.** *Denote by  $M_{\text{dyadic}}^Q$  the dyadic maximal operator with respect to a cube  $Q$ . The following inequality holds*

$$w \left\{ x \in Q : M_{\text{dyadic}}^Q w(x) > \lambda \right\} \leq 2^d \lambda \left| \left\{ x \in Q : M_{\text{dyadic}}^Q w(x) > \lambda \right\} \right|,$$

whenever  $\lambda > m_Q(w)$ .

*Proof.* We can take the set of dyadic cubes  $\mathcal{D}_{\lambda,Q}$  with respect to  $Q$  so that  $\lambda < m_R(w) \leq 2^d \lambda$  and that

$$\left\{ x \in Q : M_{\text{dyadic}}^Q w(x) > \lambda \right\} = \prod_{R \in \mathcal{D}_{\lambda,Q}} R \subsetneq Q.$$

Using this partition, we obtain

$$\text{R.H.S.} = \sum_{R \in \mathcal{D}_{\lambda,Q}} w(R) \leq 2^d \lambda \sum_{R \in \mathcal{D}_{\lambda,Q}} |R| = 2^d \lambda \left| \left\{ x \in Q : M_{\text{dyadic}}^Q w(x) > \lambda \right\} \right|.$$

The proof is therefore complete.  $\square$

The next proposition is called the reverse Hölder inequality, whose proof is due to Lerner.

**Theorem 21.10.** *Let  $w \in A_1$ . Define*

$$M_{Q,\text{dyadic}} w(x) := \sup_{R \in \mathcal{D}(Q)_x} \frac{1}{|R|} \int_R |f(y)| dy.$$

*If we set  $\delta := \frac{1}{2^{n+1}[w]_{A_1}}$ , then we have*

$$\left( \frac{1}{|Q|} \int_Q M_{Q,\text{dyadic}} w(x)^\delta w(x) dx \right)^{\frac{1}{1+\delta}} \leq 2 \frac{1}{|Q|} \int_Q w(x) dx$$

for all cubes  $Q$ .

*Proof.* By replacing  $w$  with  $\min(w, R)$  with  $R > 0$ , we can and do assume that  $w \in L^\infty(\mathbb{R}^n)$ . Abbreviate  $m_Q(w)$  to  $\mu$ . Then we have

$$\begin{aligned} & \frac{1}{|Q|} \int_Q M_{Q,\text{dyadic}} w(x)^\delta w(x) dx \\ &:= \frac{1}{|Q|} \int_0^\infty \delta \lambda^{\delta-1} w \{x \in Q : M_{Q,\text{dyadic}} w(x) > \lambda\} d\lambda \\ &= \frac{1}{|Q|} \left( \int_0^\mu + \int_\mu^\infty \right) \delta \lambda^{\delta-1} w \{x \in Q : M_{Q,\text{dyadic}} w(x) > \lambda\} d\lambda \\ &\leq \mu^{\delta+1} + \frac{1}{|Q|} \int_\mu^\infty \delta \lambda^{\delta-1} w \{x \in Q : M_{Q,\text{dyadic}} w(x) > \lambda\} d\lambda. \end{aligned}$$

Let  $\lambda > \mu$ . Then we can decompose

$$\{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\} = \bigcup_j Q_j$$

into a union of dyadic cubes  $\{Q_j\}_j$  such that

$$\frac{1}{|Q_j|} \int_{Q_j} w(x) dx > \lambda \geq \frac{1}{2^n |Q_j|} \int_{Q_j} w(x) dx = \frac{1}{2^n |Q_j|} w(Q_j)$$

and that

$$|Q_j \cap Q_{j'}| = 0 \quad (j \neq j').$$

Hence

$$\begin{aligned} & w \{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\} \\ &= \sum_j w(Q_j) \\ &\leq 2^n \sum_j |Q_j| \lambda \\ &= 2^n \lambda |\{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\}|. \end{aligned}$$

Inserting this estimate, we obtain

$$\begin{aligned} & \frac{1}{|Q|} \int_\mu^\infty \delta \lambda^{\delta-1} w \{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\} d\lambda \\ &\leq \frac{2^n}{|Q|} \int_\mu^\infty \delta \lambda^\delta |\{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\}| d\lambda \\ &\leq \frac{2^n}{|Q|} \int_0^\infty \delta \lambda^\delta |\{x \in Q : M_{Q,\text{dyadic}}w(x) > \lambda\}| d\lambda \\ &= \frac{2^n \delta}{1 + \delta} \frac{1}{|Q|} \int_Q M_{Q,\text{dyadic}}w(x)^{1+\delta} dx. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} & \frac{1}{|Q|} \int_Q M_{Q,\text{dyadic}}w(x)^\delta w(x) dx \\ &\leq \mu^{\delta+1} + \frac{2^n \delta}{\delta + 1} \frac{1}{|Q|} \int_Q M_{\text{dyadic},Q}w(x)^{1+\delta} dx \\ &\leq \mu^{\delta+1} + \frac{2^n \delta [w]_{A_1}}{\delta + 1} \frac{1}{|Q|} \int_Q M_{\text{dyadic},Q}w(x)^\delta w(x) dx \\ &\leq \mu^{\delta+1} + \frac{1}{2} \frac{1}{|Q|} \int_Q M_{\text{dyadic},Q}w(x)^\delta w(x) dx. \end{aligned}$$

Now that we are assuming that  $w \in L^\infty(\mathbb{R}^n)$ , it follows from the absorbing argument that

$$\frac{1}{|Q|} \int_Q M_{\text{dyadic},Q}w(x)^\delta w(x) dx \leq 2\mu^{\delta+1} = 2 \left( \frac{1}{|Q|} \int_Q w(x) dx \right)^{1+\delta}.$$

The proof is therefore complete.  $\square$

In Theorem 21.7, we have seen an example of  $A_1$ -weights. A function  $a : \mathbb{R}^d \rightarrow \mathbb{R}$  is said to be logarithmic bounded, if it is positive and  $\log a \in L^\infty(\mathbb{R}^d)$ . From Theorem 21.7, we deduce

$$w = K \cdot Mf^\delta \in A_1,$$

whenever  $K$  is a logarithmic bounded function and  $0 < \delta < 1$ . The following theorem asserts the converse is also true.



**Theorem 21.11.** *A weight  $w$  is an  $A_1$ -weight precisely when it can be expressed as*

$$w = K \cdot Mf^\delta,$$

where  $K$  is a positive logarithmic bounded function and  $0 < \delta < 1$ .

*Proof.* We have only to decompose  $w \in A_1$  into the product described above. Since  $w$  satisfies the  $A_1$ -condition, it satisfies the reverse Hölder inequality. Hence there exists  $0 < \delta < 1$  such that

$$m_Q(w^{\frac{1}{\delta}}) \lesssim m_Q(w)^{\frac{1}{\delta}}$$

for all  $Q \in \mathcal{Q}$ . In addition we set  $c_0 := A_1(w)$ .

Let  $x \in \mathbb{R}^d$  be a point satisfying  $Mw(x) \leq c_0 w(x)$ . Fix a  $Q$  is a cube containing  $x$ . Then  $m_Q(w) \leq c_0 w(x)$  yields

$$m_Q(w^{\frac{1}{\delta}}) \leq c_0 w(x)^{\frac{1}{\delta}}.$$

Since  $Q$  is a cube taken arbitrarily, we see that

$M[w^{\frac{1}{\delta}}](x) \lesssim w(x)^{\frac{1}{\delta}}$ . The reverse inequality of the above estimate being trivial by virtue of the Lebesgue differentiation theorem, we conclude that

$$w(x)^{\frac{1}{\delta}} \leq M[w^{\frac{1}{\delta}}](x) \lesssim w(x)^{\frac{1}{\delta}} \text{ a.e. } x \in \mathbb{R}^d,$$

from which we deduce that  $K(x) := \frac{w(x)}{M[w^{\frac{1}{\delta}}](x)^\delta}$  is logarithmic bounded.

Therefore, if we set  $f := w^{\frac{1}{\delta}}$ , then we have the desired decomposition  $w = K \cdot Mf^\delta$ . □

### 21.2. $A_p$ -weights.

Definition of the class  $A_p$ . Let us view the definition due to Muckenhoupt [361].

**Definition 21.12.** Let  $1 < p < \infty$ . A weight  $w$  is said to be an  $A_p$ -weight, if it satisfies

$$A_p(w) := \sup_{Q \in \mathcal{Q}} m_Q(w) \cdot m_Q^{(1/(p-1))}(w^{-1}) < \infty.$$

The quantity  $A_p(w)$  is said to be the  $A_p$ -norm of the weight  $w$ .

The class  $A_1$  is the smallest class among  $\{A_p\}_{1 \leq p < \infty}$  as the following proposition shows.

**Proposition 21.13.** *Let  $1 < p < \infty$ . Then we have  $A_1 \subset A_p$  and  $A_p(w) \leq A_1(w)$ .*

*Proof.* Let  $w \in A_1$  and  $Q$  be a fixed cube. Then we have

$$m_Q(w) \lesssim w(x)$$

for almost every  $x \in Q$  from the definition of  $A_1$ -class and the Hardy-Littlewood maximal operator  $M$ . Therefore, it follows that

$$m_Q(w^{-\frac{1}{p-1}}) \lesssim m_Q(w)^{-\frac{1}{p-1}},$$

proving that  $w \in A_p$  and that  $A_p(w) \leq A_1(w)$ . □

From the proposition above, we see that  $A_p$  contains a lot of nonconstant functions. Now we summarize several fundamental properties.

**Lemma 21.14.** *Let  $1 < q < p < \infty$ .*

- (1) *We have  $A_q \subset A_p$ . Speaking precisely,  $A_p(w) \leq A_q(w)$  for all  $w \in A_q$ .*

- (2) If  $w \in A_p$ , then so does  $w^{-\frac{1}{p-1}}$ . Speaking precisely, we have  $A_p(w^{-\frac{1}{p-1}}) = A_p(w)^{\frac{1}{p-1}}$ .
- (3) Let  $w_0, w_1 \in A_p$ . Then we have  $w_0 w_1^{-p+1} \in A_p$ .

The case when  $p = 2$  is symmetric:  $w \in A_2$  precisely when  $w^{-1} \in A_2$ .

*Proof.* (1) is a direct consequence of the Hölder inequality. (2) is also immediate. Let us prove (3). To do this, we fix a cube  $Q$ . By the Hölder inequality we have

$$m_Q(w_0) \lesssim w_0(x), m_Q(w_1) \lesssim w_1(x)$$

for all  $x \in Q$ . Inserting the estimate, we have

$$\begin{aligned} A_p(w_0 w_1^{-p+1}) &= m_Q(w_0 w_1^{-p+1}) \cdot m_Q(w_0^{1-p'} w_1)^{\frac{1}{p'-1}} \\ &\lesssim m_Q(w_0) m_Q(w_1)^{-p+1} \cdot m_Q(w_0)^{-1} m_Q(w_1)^{\frac{1}{p'-1}} \\ &\lesssim 1. \end{aligned}$$

This is a desired result. □

Lemma 21.14 gives us an important example of the class  $A_p$ .

**Corollary 21.15.** *Let  $1 < p < \infty$  and  $a \in \mathbb{R}$ . Then  $w(x) = |x|^a$  belongs to  $A_p$  if and only if  $-d < a < d(p - 1)$ .*

*Proof.* By Proposition 21.8, we have  $|x|^A \in A_1$  for all  $-d < A < 0$ . Therefore, we have  $|x|^{A-B(p-1)} \in A_p$  for all  $-d < A, B < 0$ . Therefore the condition is sufficient. The condition is also necessary. Indeed if  $a \leq -d$ , the the function cannot be locally integrable. If  $a \geq (p - 1)d$ , then  $w^{-\frac{1}{p-1}}$  in turn cannot be locally integrable. Therefore the condition is also necessary. □

A slight variant may be in order as well.

**Remark 21.16.** Let  $1 < p < \infty$  and  $a \in \mathbb{R}$ . Then  $w(x) = (1 + |x|)^a$  belongs to  $A_p$  if and only if  $-d < a < d(p - 1)$ .

Here we shall characterize the class  $A_p$ .

**Theorem 21.17.** *Let  $1 < p < \infty$ . Then the following conditions on a weight  $w$  are equivalent.*

- (1)  $w \in A_p$ .
  - (2)  $M$  is weak- $L^p(w)$  bounded, that is, for every  $f \in L^p(w)$ ,
- $$(21.11) \quad w\{Mf > \lambda\} \lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^d} |f(x)|^p w(x) dx.$$
- (3)  $M$  is strong  $L^p(w)$ -bounded, that is, for every  $f \in L^p(w)$ ,
- $$(21.12) \quad \int_{\mathbb{R}^d} Mf(x)^p w(x) dx \lesssim \int_{\mathbb{R}^d} |f(x)|^p w(x) dx.$$

Needless to say, the implication  $w \in A_p \implies (21.12)$  is significant. Indeed, (21.12) is stronger than (21.11) by virtue of the Chebychev inequality. Meanwhile once we assume (21.11), then taking

$$(21.13) \quad f := w^{-\frac{1}{p-1}} \chi_Q, \quad \lambda =: m_Q(w)$$

for a given cube  $Q$ , we can easily deduce that  $w \in A_p$ .

$w \in A_p \implies (21.12)$ . Denote by  $M_w^c$  the centered weighted maximal operator given by

$$M_w^c f(x) := \sup_{Q \in \mathcal{Q}, c_Q=x} \frac{1}{w(Q)} \int_Q |f(y)| w(y) dy.$$

By virtue of the 5r-covering lemma, it is easy to see that  $M_w$  is  $L^p(w)$ -bounded. Write  $\sigma := w^{-\frac{1}{p-1}}$  and define  $M_\sigma^c$  analogously. By the definition of the  $A_p$ -weights we have

$$\frac{w(Q)}{|Q|} \left( \frac{\sigma(3Q)}{|Q|} \right)^{p-1} \leq 3^{dp} \frac{w(3Q)}{|3Q|} \left( \frac{\sigma(3Q)}{|3Q|} \right)^{p-1} \leq 3^{dp} A_p(w).$$

Hence it follows that

$$\begin{aligned} m_Q(|f|) &\leq 3^{dp} A_p(w)^{\frac{1}{p-1}} \left( \frac{|Q|}{w(Q)} \left( \frac{1}{w^{-\frac{1}{p-1}}(3Q)} \int_Q |f(y)| dy \right)^{p-1} \right)^{\frac{1}{p-1}} \\ &\lesssim_d A_p(w)^{\frac{1}{p-1}} M_w^c (M_\sigma^c [f \sigma^{-1}]^{p-1} w^{-1})(x)^{\frac{1}{p-1}} \end{aligned}$$

for all  $x \in Q$ . Or equivalently,

$$Mf(x) \lesssim_d A_p(w)^{\frac{1}{p-1}} M_w^c (M_\sigma^c [|f| w^{\frac{1}{p-1}}] w^{-1})(x)^{\frac{1}{p-1}}.$$

Inserting this pointwise estimate, we obtain the desired  $L^p(w)$ -estimate. □

**Remark 21.18.** The condition above is sometimes referred to as the testing conditions.

*Exercise 156.* Let  $1 < p < \infty$ . Then show that a weight  $w$  belongs to  $A_p$  if and only if  $\frac{1}{|Q|} \int_Q |f(x)| dx \lesssim \left( \frac{1}{w(Q)} \int_Q |f(x)|^p w(x) dx \right)^{1/p}$  for all measurable functions  $f$ .

**21.3.  $A_\infty$ -weights.**

If we use the Hölder inequality, then we have

$$(21.14) \quad m_Q^{(1+\varepsilon)}(w) \geq m_Q(w)$$

for all weights. However, since  $w \in L^1_{\text{loc}}(\mathbb{R}^d)$  does not imply  $w \in L^{1+\varepsilon}_{\text{loc}}(\mathbb{R}^d)$  for any  $\varepsilon > 0$ , the reverse inequality

$$(21.15) \quad m_Q^{(1+\varepsilon)}(w) \leq m_Q(w)$$

and its weaker version

$$(21.16) \quad m_Q^{(1+\varepsilon)}(w) \leq m_Q(w)$$

fail. However, (21.16) is attractive and deserves a special name. Here we investigate a class of weights satisfying (21.16).

**Lemma 21.19.** *Suppose that  $w$  is an  $A_p$  weight with  $1 \leq p < \infty$ . For all  $\alpha \in (0, 1)$  there exists  $\beta = \beta(\alpha) \in (0, 1)$  so that*

$$(21.17) \quad |E| \leq \alpha |Q| \implies w(E) \leq \beta w(Q)$$

whenever  $E$  is a measurable subset of a cube  $Q$ .

*Proof.* We may assume  $1 < p < \infty$  because  $A_1$  is a subset of  $A_2$ . There exists  $c_0 > 0$  such that

$$(21.18) \quad w\{Mf > \lambda\} \leq \frac{c_0}{\lambda^p} \int |f(x)|^p \cdot w(x) dx$$

for all  $f \in L^p(\mathbb{R}^d)$ . Let  $\lambda := \frac{|Q| - |E|}{|Q|}$  and  $f := \chi_{Q \setminus E}$ . Then we have

$$(21.19) \quad w(Q) \leq w\{Mf > \lambda\} \leq c_1 \frac{1}{\lambda^p} \int |f(x)|^p \cdot w(x) dx \leq c_0 c_1 w(Q \setminus E) \frac{|Q|^p}{(|Q| - |E|)^p}$$

for some  $c_1 > 0$ . Arranging this, we have

$$(21.20) \quad w(E) \leq w(Q) \left( 1 - \frac{1}{c_0 c_1} \left( \frac{|Q| - |E|}{|Q|} \right)^p \right).$$

Therefore we have only to set  $\beta := 1 - \frac{(1 - \alpha)^p}{c_0 c_1}$ . □

We are going to characterize weights which satisfy (21.17). With this purpose, the next lemma seems too generalized. However, we need it in full generality for our purpose.

**Lemma 21.20** (Reverse Hölder inequality). *Suppose that  $\mu_1, \mu_2$  are  $\sigma$ -finite Radon measures on  $\mathbb{R}^d$  that satisfies, for some  $0 < \alpha, \beta < 1$ ,*

$$(21.21) \quad \mu_1(E) \leq \alpha \mu_2(Q) \implies \mu_2(E) \leq \beta \mu_2(Q).$$

Then;

- (1) *The measure  $\mu_2$  is absolutely continuous with respect to  $\mu_1$ .*
- (2) *Assume in addition that  $\mu_1$  is doubling. Then there exists  $\varepsilon > 0$  so that, for every cube  $Q$ , the reverse Hölder inequality below holds:*

$$(21.22) \quad \left( \frac{1}{\mu_1(Q)} \int_Q w(x)^{1+\varepsilon} d\mu_1(x) \right)^{\frac{1}{1+\varepsilon}} \lesssim \frac{1}{\mu_1(Q)} \int_Q w(x) d\mu_1(x).$$

where  $w$  denotes the density of  $\mu_2$  with respect to  $\mu_1$ .

In particular, if  $w$  is an  $A_\infty$ -weight, then, for some  $\varepsilon > 0$ ,

$$(21.23) \quad m_Q^{(1+\varepsilon)}(w) \lesssim m_Q(w).$$

By virtue of the classical Hölder inequality it is trivial that

$$(21.24) \quad m_Q(w) \leq m_Q^{(1+\varepsilon)}(w).$$

This is why (21.22) deserves its name.

*Proof.*

- (1) Suppose that  $E$  is a bounded open set such that  $\mu_1(E) = 0$ . Decompose  $E$  into the sum of disjoint dyadic cubes  $\{R_j\}_{j=1}^\infty$ . Then we have

$$(21.25) \quad 0 = \mu_1(E \cap R_j) \leq \alpha \mu_1(R_j)$$

for each  $j \in \mathbb{N}$ . By assumption (21.21), we have

$$(21.26) \quad \mu_2(E \cap R_j) \leq \beta \mu_2(R_j).$$

Adding (21.26) over  $j$ , we obtain

$$(21.27) \quad \mu_2(E) \leq \beta \mu_2(E) < \infty$$

and hence  $\mu_2(E) = 0$ . Since this is the case for all bounded set  $E$  with  $\mu_1(E) = 0$ , we see that  $\mu_2$  is absolutely continuous with respect to  $\mu_1$ .

(2) Having proved the absolute continuity of  $\mu_2$  with respect to  $\mu_1$ , we denote by  $w$  its Radon-Nikodym derivative. Denote by  $D$  the doubling constant of  $\mu_1$ :

$$\mu_1(2Q) \leq D\mu_1(Q)$$

for all cubes  $Q$ .

To prove (21.22), fix a cube  $Q$  and set  $\lambda_k := (\alpha^{-1}D)^k \lambda_0$  for each  $k \in \mathbb{N}_0$ , where

$$(21.28) \quad \lambda_0 := \frac{1}{\mu_1(Q)} \int_Q w(x) d\mu_1(x).$$

Let us decompose  $\{x \in \mathbb{R}^d : M_{\text{dyadic}} w(x) > \lambda_k\}$  into a disjoint union of dyadic cubes  $Q_{k,j}$  such that

$$(21.29) \quad \lambda_k < \frac{1}{\mu_1(Q_{k,j})} \int_{Q_{k,j}} w(x) d\mu_1(x) = \frac{\mu_2(Q_{k,j})}{\mu_1(Q_{k,j})} \leq D\lambda_k = \alpha\lambda_{k+1}$$

and that  $w(x) \leq \lambda_k$  for  $\mu_1$ -a.e.  $x \in Q \setminus \bigcup_j Q_{k,j}$ . We write  $\Omega_k := \bigcup_j Q_{k,j}$ .

Recall that two dyadic cubes are disjoint unless one is contained in the other. Therefore we have

$$\mu_1(\Omega_{k+1} \cap Q_{k,j_0}) \leq \sum_{\substack{Q_{k+1,j} \\ Q_{k+1,j} \subset Q_{k,j_0}}} \mu_1(Q_{k+1,j}) \leq \sum_{\substack{Q_{k+1,j} \\ Q_{k+1,j} \subset Q_{k,j_0}}} \frac{\mu_2(Q_{k+1,j})}{\lambda_{k+1}} \leq \frac{\mu_2(Q_{k,j_0})}{\lambda_{k+1}}.$$

If we use (21.29) once again, then we have

$$\mu_1(\Omega_{k+1} \cap Q_{k,j_0}) \leq \frac{\mu_2(Q_{k,j_0})}{\lambda_{k+1}} \leq \alpha\mu_1(Q_{k,j_0}).$$

We are now assuming (21.17) and hence we can pass the above inequality to the weighted one:

$$(21.30) \quad \mu_2(\Omega_{k+1} \cap Q_{k,j_0}) \leq \beta\mu_2(Q_{k,j_0}),$$

where we have used (21.21) with  $\alpha$ . Adding the above inequality, we obtain

$$(21.31) \quad \mu_2(\Omega_{k+1}) \leq \beta\mu_2(\Omega_k).$$

This implies that  $\mu_2\left(\bigcap_{k=0}^{\infty} \Omega_k\right) = 0$ . Similarly by using  $\mu_1(\Omega_{k+1} \cap Q_{k,j_0}) \leq \alpha\mu_1(Q_{k,j_0})$  again, we also have

$$(21.32) \quad \mu_1\left(\bigcap_{k=0}^{\infty} \Omega_k\right) = 0.$$

Denote  $\Omega_{-1} := Q_0$  for the sake of simplicity of notations. Then by (21.32), we have

$$\int_Q w(x)^{1+\varepsilon} d\mu_1(x) = \frac{1}{\mu_1(Q)} \sum_{k=-1}^{\infty} \int_{\Omega_k \setminus \Omega_{k+1}} w(x)^{1+\varepsilon} d\mu_1(x).$$

If we invoke (21.29)–(21.31), then we obtain

$$\begin{aligned} \int_Q w(x)^{1+\varepsilon} d\mu_1(x) &\leq \frac{1}{\mu_1(Q)} \sum_{k=-1}^{\infty} \lambda_{k+1}^\varepsilon \mu_2(\Omega_k) \\ &\leq \frac{\lambda_0^\varepsilon}{\mu_1(Q)} \sum_{k=-1}^{\infty} (\alpha^{-1}D)^{(k+1)\varepsilon} \beta^k \mu_2(\Omega_0) \\ &\leq \lambda_0^\varepsilon \frac{1}{\mu_1(Q)} \int_Q w(x) d\mu_1(x) \cdot \sum_{k=-1}^{\infty} (\alpha^{-1}D)^{(k+1)\varepsilon} \beta^k. \end{aligned}$$

Note that  $\beta$  is given by (21.21). However,  $\varepsilon$  is still at our disposal. Therefore, by choosing  $\varepsilon > 0$  sufficiently small the series of the above inequality converges.

Thus, we obtain

$$(21.33) \quad \left( \frac{1}{\mu_1(Q)} \int_Q w(x)^{1+\varepsilon} d\mu_1(x) \right)^{\frac{1}{1+\varepsilon}} \lesssim \frac{1}{\mu_1(Q)} \int_Q w(x) d\mu_1(x)$$

and the proof of (21.23) is therefore complete. □

We summarize the above observations.

**Theorem 21.21.** *Let  $w$  be a weight. Then the following are equivalent.*

- (1)  $w \in A_p$  for some  $1 \leq p < \infty$ .
- (2) For every  $0 < \alpha < 1$  there exists  $0 < \beta = \beta(\alpha) < 1$  such that

$$(21.34) \quad |E| \leq \alpha|Q| \implies w(E) \leq \beta w(Q).$$

- (3)  $w$  satisfies the reverse Hölder inequality: there exists  $\varepsilon > 0$  such that, for any cubes  $Q$ ,

$$(21.35) \quad m_Q^{(1+\varepsilon)}(w) \lesssim m_Q(w).$$

**Definition 21.22.** A weight  $w$  is said to be an  $A_\infty$ -weight, if it satisfies the conditions in Theorem 21.21.

*Proof of Theorem 21.21.* Note first that (1)  $\implies$  (2)  $\implies$  (3) is proved in Lemmas 21.19 and 21.20. Here we used Lemma 21.20 with  $\mu_1 = dx$  and  $\mu_2 = w dx$ . We shall prove (3) implies (1) by establishing the equivalence to the following additional assertions.

- (4) There exists  $\delta > 0$  such that

$$(21.36) \quad \frac{w(A)}{w(Q)} \lesssim \left( \frac{|A|}{|Q|} \right)^\delta,$$

whenever  $A$  is a subset of a cube  $Q$ .

- (5) There exists  $0 < \alpha, \beta < 1$  such that

$$(21.37) \quad |E| \leq \alpha|Q| \implies w(E) \leq \beta w(Q)$$

for all measurable sets  $E$  and  $Q$  such that  $Q$  is a cube containing  $E$ .

- (6) There exists  $0 < \alpha', \beta' < 1$  such that

$$(21.38) \quad w(E) < \alpha' w(Q) \implies |E| < \beta' |Q|$$

for all measurable sets  $E$  and  $Q$  such that  $Q$  is a cube containing  $E$ .

- (7) There exists  $\varepsilon > 0$  such that

$$(21.39) \quad \left( \frac{1}{|Q|} \int_Q w(x)^{-(1+\varepsilon)} dx \right)^{\frac{1}{1+\varepsilon}} \lesssim \frac{w(Q)}{|Q|}$$

for all measurable sets  $E$  and  $Q$  such that  $Q$  is a cube containing  $E$ .

We plan to conclude the proof of Theorem 21.21 by proving

$$(3) \implies (4) \implies (5) \implies (6) \implies (7) \implies (1).$$

$(3) \implies (4).$

By the Hölder inequality and the reverse Hölder inequality we have

$$(21.40) \quad \frac{w(A)}{w(Q)} \leq \frac{|A|^{\frac{\varepsilon}{1+\varepsilon}}}{w(Q)} \left( \int_A w(x)^{1+\varepsilon} dx \right)^{\frac{1}{1+\varepsilon}} \leq \frac{|A|^{\frac{\varepsilon}{1+\varepsilon}}}{w(Q)} \left( \int_Q w(x)^{1+\varepsilon} dx \right)^{\frac{1}{1+\varepsilon}} \lesssim \left( \frac{|A|}{|Q|} \right)^{\frac{\varepsilon}{1+\varepsilon}}.$$

$(4) \implies (5)$ . Denote by  $C_0$  the implicit constant in (21.36):

$$\frac{w(A)}{w(Q)} \leq C_0 \left( \frac{|A|}{|Q|} \right)^\delta,$$

Choose  $\alpha > 0$  so small that  $C_0 \alpha^\delta < 1$ . Then we have only to set  $\beta = C_0 \alpha^\delta$  to obtain (5).

$(5) \implies (6)$ . We have only to set  $\alpha' = 1 - \beta$  and  $\beta' = 1 - \alpha$ .

$(6) \implies (7)$ . First, we contrapose (21.38):

$$|E| \geq \beta'|Q| \implies w(E) \geq \alpha'w(Q).$$

Therefore, by substituting  $E = \sqrt[p]{\beta'}Q$ , we obtain

$$(21.41) \quad \alpha^{-1}w(\sqrt[p]{\beta'}Q) \geq w(Q).$$

Since  $\beta < 1$ , (21.41) implies the doubling condition. Note that  $\mu_2 = dx$  has a density  $w(x)^{-1}$  with respect to the doubling measure  $\mu_1 = w(x) dx$ . Thus, we can invoke Lemma 21.20 again.

$(7) \implies (1)$ . Now that we have

$$(21.42) \quad \left( \frac{1}{|Q|} \int_Q w(x)^{-(1+\varepsilon)} dx \right)^{\frac{1}{1+\varepsilon}} \lesssim \frac{w(Q)}{|Q|}$$

We have only to take  $p := 1 + \frac{1}{1+\varepsilon}$ . □

Properties of  $A_\infty$ -weights. Having cleared up some elementary properties concerning the reverse Hölder inequality, we now turn to characterize  $A_p$ -weights.

Now we observe the openness property of the weight.

**Theorem 21.23** (Openness property of  $A_p$ -weights). *Suppose that  $w \in A_p$  with  $1 < p < \infty$ . Then there exists  $1 < q < p$  so that  $w \in A_q$ .*

Before we prove this theorem, let us see why this theorem deserves its name. From this theorem, we conclude that the set of all  $p \geq 1$  for which  $w \in A_p$  form an open set in  $[1, \infty)$ .

*Proof.* Note that  $w^{-(p'-1)}$  satisfies the reverse Hölder inequality:

$$(21.43) \quad m_Q(w^{-(1+\varepsilon)(p'-1)})^{\frac{1}{1+\varepsilon}} \lesssim m_Q(w^{-(p'-1)}).$$

Define  $q$  by  $(q' - 1) = (1 + \varepsilon)(p' - 1)$ . Then the reverse Hölder inequality reads

$$(21.44) \quad m_Q(w^{-(q'-1)})^{\frac{1}{(q'-1)}} \lesssim m_Q(w^{-(p'-1)})^{\frac{1}{(p'-1)}}.$$

Inserting this estimate to the definition of  $A_q(w)$ , we obtain

$$(21.45) \quad A_q(w) = \sup_{Q \in \mathcal{Q}} m_Q(w) \cdot m_Q(w^{-(q'-1)})^{\frac{1}{(q'-1)}} \lesssim m_Q(w) \cdot m_Q(w^{-(p'-1)})^{\frac{1}{(p'-1)}} \lesssim A_p(w).$$

Therefore, we conclude that  $w \in A_q$ . □

Finally we shall prove a structure theorem of the class  $A_p$  with  $1 < p < \infty$ .

**Theorem 21.24.** *A weight  $w$  belongs to  $A_p$  precisely when  $w$  can be factorized into  $w_0 w_1^{1-p}$  with  $w_0, w_1 \in A_1$ .*

*Proof.* The heart of the matters is the factorization of a given  $A_1$ -weight because the converse is already proved in Lemma 21.14.

First, we may assume  $1 < p \leq 2$ . Because  $w \in A_p$  if and only if  $w^{1-p} \in A_{p'}$ .

We define a sublinear operator  $S$  on  $L^p(w)$  by

$$(21.46) \quad Su := w^{-1}M(uw) + M(|u|^{\frac{1}{p-1}})^{p-1}.$$

Then  $S$  is  $L^p(w)$ -bounded. Indeed,

$$(21.47) \quad \|Su\|_{L^{p'}(w)} \leq \|M(uw)\|_{L^{p'}(w^{1-p'})} + \|M(|u|^{\frac{1}{p-1}})\|_{L^p(w)}^{p-1} \lesssim \|u\|_{L^{p'}(w)}.$$

Let us put the operator norm by  $\beta$ . Let  $u$  be a positive non-zero function in  $L^{p'}(w)$  and set

$$(21.48) \quad v = \sum_{j=0}^{\infty} \frac{1}{(2\beta)^j} S^j u,$$

where  $S^j$  denotes the  $j$ -times composition of  $S$ . Then since  $S$  is sublinear, we have

$$(21.49) \quad Sv \leq 2\beta v.$$

Therefore, in this case from the definition of  $S$  we have

$$(21.50) \quad M(uw) \leq 2\beta uw, \quad M(u^{\frac{1}{p-1}}) \leq 2\beta u^{\frac{1}{p-1}}.$$

We set  $w_0 = uw$  and  $w_1 = u^{\frac{1}{p-1}}$ . Then the above inequality reads that  $w_0$  and  $w_1$  are  $A_1$ -weights. Therefore, it follows that  $w = w_0 w_1^{1-p}$  is the desired decomposition.  $\square$

The proof of this theorem provides us another method to generate  $A_1$ -weights, which we outline in the next exercise.

*Exercise 157.* Let  $w$  be a measurable function belonging to some  $L^p(w)$  with  $1 < p \leq \infty$ . Denote by  $\beta$  the  $L^p(w)$ -norm of the maximal operator  $M$ . Then show that

$$(21.51) \quad v := \sum_{j=0}^{\infty} \frac{1}{(2\beta)^j} M^j w$$

is an  $A_1$ -weight satisfying

$$(21.52) \quad Mv \leq 2\beta v.$$

#### 21.4. $A_p$ -weights and singular integral operators.

Recall that, in the course of the proof of Theorem 21.21, we have proved

$$(21.53) \quad \frac{w(A)}{w(Q)} \lesssim \left( \frac{|A|}{|Q|} \right)^\delta$$

for any couple  $(A, Q)$  of measurable sets such that  $Q$  is a cubes that contains  $A$ .

**Theorem 21.25.** *Suppose that  $w \in A_\infty$ . Then,*

$$(21.54) \quad w\{Mf > 2\lambda, M^\sharp f < \beta\lambda\} \lesssim \beta^\delta w\{Mf > \lambda\}$$

for all  $\lambda > 0, \beta > 0$ .



*Proof.* We keep to the notation of Theorem 20.5. Under the same notation as the unweighted version, we have

$$(21.55) \quad |E_\lambda \cap S_j| \lesssim \beta |S_j|.$$

Therefore, as we have remarked above, we obtain

$$(21.56) \quad w(E_\lambda \cap S_j) \lesssim \beta^\delta w(S_j)$$

Going through the same argument as the unweighted case, we obtain our assertion.  $\square$

Hence by the good- $\lambda$  inequality we obtain

**Corollary 21.26.** *Let  $w \in A_\infty$ . Assume that a measurable function  $f$  satisfies  $Mf \in L^p(w)$ . Then*

$$(21.57) \quad \|Mf\|_{L^p(w)} \lesssim \|M^\sharp f\|_{L^p(w)}.$$

**Lemma 21.27.** *Let  $T$  be a singular integral operator. Assume that  $w \in A_\infty$ . Then  $Tf \in L^p(w)$  for every  $f \in L_c^\infty(\mathbb{R}^d)$  and  $1 < p < \infty$ .*

*Proof.* Let us take  $\varepsilon > 0$  so that the reverse Hölder inequality

$$(21.58) \quad m_Q^{(1+\varepsilon)}(w) \lesssim m_Q(w) \text{ for all } Q \in \mathcal{Q}$$

holds. Suppose that  $B(R)$  engulfs the support of  $f$ . Then we have

$$\begin{aligned} \int_{B(2R)} |Tf(x)|^p w(x) dx &\leq \left( \int_{B(2R)} |Tf(x)|^{p(1+\varepsilon^{-1})} dx \right)^{\frac{\varepsilon}{1+\varepsilon}} \left( \int_{B(2R)} w(x)^{1+\varepsilon} dx \right)^{\frac{1}{1+\varepsilon}} \\ &\lesssim \left( \int_{\mathbb{R}^d} |f(x)|^{p(1+\varepsilon^{-1})} dx \right)^{\frac{\varepsilon}{1+\varepsilon}} \cdot w(B(2R)) \\ &< \infty. \end{aligned}$$

Outside  $B(2R)$ , we have  $|Tf(x)| \lesssim |x|^{-d}$  by the size condition of the kernel. Invoking this estimate, we have

$$(21.59) \quad \int_{\mathbb{R}^d \setminus B(2R)} |Tf(x)|^p w(x) dx \lesssim \sum_{j=1}^{\infty} \frac{1}{2^{jp}} w(B(2^j R)).$$

By the  $A_\infty$ -condition  $w(B(2^j R)) \lesssim 2^{j\delta} w(B(R))$  we obtain

$$(21.60) \quad \int_{\mathbb{R}^d \setminus B(2R)} |Tf(x)|^p w(x) dx \lesssim \sum_{j=1}^{\infty} \frac{1}{2^{j(p-\delta)}} w(B(R)) < \infty.$$

This is the desired result.  $\square$

**Lemma 21.28.** *Let  $r > 1$  and  $T$  be a singular integral operator. Then for all  $f \in L_c^\infty(\mathbb{R}^d)$ , we have*

$$(21.61) \quad M^\sharp[Tf](x) \lesssim M^{(r)}f(x)$$

for all  $x \in \mathbb{R}^d$ .

We remark that Lemmas 21.27 and 21.28 with  $r = 2$  will give us a new proof of the  $L^p(\mathbb{R}^d)$ -boundedness of singular integral operators. Namely, from Lemmas 21.27 and 21.28 we obtain

$$(21.62) \quad \|Tf\|_p \lesssim \|f\|_p$$

for all  $f \in L_c^\infty(\mathbb{R}^d)$ .

*Proof.* Let  $Q$  be a cube containing  $x$ . Let us prove

$$(21.63) \quad m_Q(|Tf - m_Q(Tf)|) \lesssim M^{(r)}f(x).$$

Let us decompose  $f$  according to  $2Q$ . By the  $L^r$ -boundedness we obtain

$$(21.64) \quad m_Q(|T[\chi_{2Q} \cdot f] - m_Q(T[\chi_{2Q} \cdot f])|) \leq 2m_Q^{(r)}(|T[\chi_{2Q} \cdot f]|) \lesssim m_{2Q}^{(r)}(|f|) \lesssim M^{(r)}f(x).$$

Meanwhile the estimate outside  $2Q$  is obtained by the Hölmänder condition. Let  $y \in Q$ .

$$(21.65) \quad T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f](y) - m_Q(T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f]) = \frac{1}{|Q|} \iint_{Q \times \mathbb{R}^d \setminus 2Q} (K(y, z) - K(w, z))f(z) dw dz.$$

By using  $|K(y, z) - K(x, z)| \lesssim \frac{\ell(Q)}{|x - z|^{d+1}}$ , we obtain

$$(21.66) \quad |T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f](y) - m_Q(T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f])| \lesssim Mf(x).$$

Hence it follows that

$$(21.67) \quad m_Q(|T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f] - m_Q(T[\chi_{\mathbb{R}^d \setminus 2Q} \cdot f])|) \lesssim Mf(x).$$

This is the desired result. □

**Theorem 21.29.** *Assume that  $w \in A_p$  with  $1 < p < \infty$ . Let  $T$  be a singular integral operator. Then  $\|Tf\|_{L^p(w)} \lesssim_p \|f\|_{L^p(w)}$  for all  $f \in L_c^\infty(\mathbb{R}^d)$ .*

*Proof.* As we have verified in Lemma 21.27, we have  $Tf \in L^p(w)$ . Then we have

$$(21.68) \quad \|Tf\|_{L^p(w)} \leq \|MTf\|_{L^p(w)} \lesssim \|M^\sharp[Tf]\|_{L^p(w)}.$$

by virtue of Corollary 21.26. By the openness property we can choose  $r > 1$  so that  $w \in A_{p/r}$ . Using this  $r > 1$ , we have

$$(21.69) \quad \|M^\sharp[Tf]\|_{L^p(w)} \lesssim \|M^{(r)}f\|_{L^p(w)} = \|M[|f|^r]\|_{L^{p/r}(w)}^{\frac{1}{r}}$$

Now that  $w \in A_{p/r}$  we obtain

$$(21.70) \quad \|Tf\|_{L^p(w)} \lesssim \| |f|^r \|_{L^{p/r}(w)}^{\frac{1}{r}} \lesssim \|f\|_{L^p(w)}.$$

This is the desired result. □

**Theorem 21.30.** *Let  $T$  be a singular integral operator and  $w \in A_1$ . Then we have that  $\sup_{\lambda > 0} \lambda w\{|Tf| > \lambda\} \lesssim_w \int_{\mathbb{R}^d} |f(x)|w(x) dx$  for all  $f \in L_c^\infty(\mathbb{R}^d)$  and  $\lambda > 0$ .*

*Proof.* Form the Calderón-Zygmund decomposition of  $f$ . Under the same notation as usual, we have to show

$$(21.71) \quad w\{|Tg| > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx, \quad w\{|Tb| > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx.$$

Observe that the set  $\{Mf > \lambda\}$  is a compact set and hence in view of the construction we see that  $g$  and the  $b_j$ 's are compactly supported. As for the good part, now that we have established  $\|Tg\|_{L^2(w)} \lesssim \|g\|_{L^2(w)}$ , the same argument as the unweighted case works and we obtain

$$(21.72) \quad w\{|Tg| > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |g(x)|w(x) dx.$$

On the cube  $Q_j$ , we have  $g = m_{Q_j}(f)$ . Therefore, it follows from the  $A_1$ -condition that

$$(21.73) \quad \int_{Q_j} |g(x)|w(x) dx = \frac{w(Q_j)}{|Q_j|} \int_{Q_j} |f(x)| dx \leq \int_{Q_j} |f(x)|Mw(x) dx \lesssim \int_{Q_j} |f(x)|w(x) dx.$$

Meanwhile outside  $\bigcup_j Q_j$ ,  $g$  and  $f$  agree. Therefore, we obtain

$$(21.74) \quad \int_{Q_j} |g(x)|w(x) dx \lesssim \int_{Q_j} |f(x)|w(x) dx.$$

Thus, the estimate for the “good ” part is valid.

By the condition  $\frac{1}{|Q_j|} \int_{Q_j} |f(x)| dx \geq \lambda$ , we obtain

$$w(2\sqrt{d}Q_j) \leq m_{2\sqrt{d}Q_j}(w) \cdot \frac{2^d \sqrt{d}^d}{\lambda} \int_{Q_j} |f(x)| dx \lesssim \frac{1}{\lambda} \int_{Q_j} |f(x)|Mw(x) dx \lesssim \frac{1}{\lambda} \int_{Q_j} |f(x)|w(x) dx.$$

Thus, for the treatment of the “bad ” part, it remains to prove

$$(21.75) \quad \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} |Tb_j(x)|w(x) dx \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |b_j(x)|w(x) dx.$$

Once we prove (21.75), we obtain

$$\begin{aligned} w\{|Tb| > \lambda\} &\lesssim \frac{1}{\lambda} \left( \int_{\mathbb{R}^d} |f(x)|w(x) dx + \sum_j \int_{\mathbb{R}^d} |b_j(x)|w(x) dx \right) \\ &\simeq \frac{1}{\lambda} \left( \int_{\mathbb{R}^d} |f(x)|w(x) dx + \int_{\mathbb{R}^d} |g(x)|w(x) dx \right) \\ &\lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |f(x)|w(x) dx, \end{aligned}$$

in view of the disjointness of the  $Q_j$ 's.

To see (21.75) we obtain

$$(21.76) \quad \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} |Tb_j(x)|w(x) dx \leq \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} \left( \int_{Q_j} \frac{\ell(Q_j)}{|x - c(Q_j)|^{d+1}} |b_j(y)| \right) w(x) dx$$

by the Hölmänder condition. Since  $w \in A_1$ , we obtain

$$(21.77) \quad \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_j} \frac{w(x) dx}{|x - c(Q_j)|^{d+1}} \lesssim \int_{\ell(Q_j)}^\infty \frac{w(Q(c(Q_j), \ell))}{\ell^{d+2}} d\ell \lesssim \inf_{z \in Q} Mw(z) \int_{\ell(Q_j)}^\infty \frac{|Q(c(Q_j), \ell)|}{\ell^{d+2}} d\ell.$$

If we insert this estimate, we obtain (21.75). □

Finally to conclude this section, let us prove the boundedness of the maximal operator of singular integral operators. Recall that, if  $T$  is a generalized singular integral operator with kernel  $K$ , then we defined

$$(21.78) \quad T^* f(x) := \sup_{\varepsilon > 0} \left| \int_{\mathbb{R}^d \setminus B(x, \varepsilon)} K(x, y) f(y) dy \right|$$

for  $f \in L_c^\infty(\mathbb{R}^d)$ . Recall that we proved the Cotlar inequality

$$(21.79) \quad T^* f(x) \lesssim_\eta M^{(\eta)}[Tf](x) + Mf(x) \quad (x \in \mathbb{R}^d)$$

for  $0 < \eta \leq 1$ . The proof is similar to the unweighted case and once this is proved, the almost everywhere pointwise limit of the truncated singular integral operators are easy to deduce.

**Theorem 21.31.** *Suppose that  $T$  is a generalized Calderón-Zygmund operator. Let  $w \in A_p$  with  $1 \leq p < \infty$ .*

- (1) *Suppose that  $1 < p < \infty$ . Then  $\|T^* f\|_{L^p(w)} \lesssim_{p,w} \|f\|_{L^p(w)}$  for all  $f \in L_c^\infty(\mathbb{R}^d)$ .*

(2) Let  $\lambda > 0$ . Then

$$\lambda w\{T^*f > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}} |f(x)|w(x) dx$$

for all  $f \in L_c^\infty(\mathbb{R}^d)$ .

Notes and references for Chapter 9.

Section 18. As in [183], the classical theory of  $H^p(\mathbb{R}^d)$  spaces could be considered as a chapter of complex function theory.

Theorem 18.6 was obtained in [183], where the definition of the norm is different. The definition dealt in [183] will be taken up in Chapter 23. The proof of Theorem 18.6 is very subtle, since we have taken into account the weak- $L^1(\mathbb{R}^d)$  boundedness of the singular integral operators. This aspect is well-described in the papers [102, 335].

Section 19. The origin of the space  $BMO(\mathbb{R}^d)$  is the paper by John and Nirenberg [261], where they investigated the space  $BMO(\mathbb{R}^d)$  on a fixed cube. Before the study by C. Fefferman and E. M. Stein, the John-Nirenberg inequality was obtained in [261]. In the original paper the John-Nirenberg inequality was stated without using the BMO norm: It was formulated as an inequality just in connection with a locally integrable function and an average over cube. In [183] C. Fefferman and E. M. Stein began to use the terminology of BMO, who investigated in connection with Hardy spaces and the singular integral operators. Proposition 19.2 is investigated as well. Nowadays, however, the definition of (global) BMO being prevailing, Theorem 19.8 was stated in the form in the modern fashion. Spanne and Stein obtained Theorem 19.6 in [444] and [452] respectively. In [183] the result is somehow generalized. As we have seen in Theorem 19.6, the class BMO arises as the image of  $L^\infty(\mathbb{R}^d)$  under singular integral transformations. However, the converse is also true. In [177] C. Fefferman presented a characterization: In summary we have  $BMO(\mathbb{R}^d) = L^\infty(\mathbb{R}^d) + \sum_{j=1}^d R_j(L^\infty(\mathbb{R}^d))$ . We have shown  $\supset$  in this book. However, equality does hold. For more details we refer to [16, 58]. We remark that the converse is known: Let  $K$  be a kernel satisfying the estimates in this section. We do not assume that the operator

$$(21.80) \quad T : f \mapsto \int_{\mathbb{R}^d} K(x, y)f(y) dy$$

is  $L^2(\mathbb{R}^d)$ -bounded. All we assume is that  $T : \mathcal{S} \rightarrow \mathcal{S}'$  is a continuous operator such that

$$(21.81) \quad Tf(x) = \int_{\mathbb{R}^d} K(x, y)f(y) dy, \quad x \notin \text{supp}(f)$$

for all  $f \in C_c^\infty(\mathbb{R}^d)$ . In order that  $T$  is  $L^2(\mathbb{R}^d)$ -bounded it is necessary and sufficient for  $T1$  to be defined as an element in  $BMO(\mathbb{R}^d)$  in some sense. This theorem is known as the T1-theorem (see [158]). The definition of  $H^1(\mathbb{R})$  has a lot of variants. For example, Stein and Weiss defined

$$\|f\|_{H^1} = \|f\|_{L^1} + \|Hf\|_{L^1}$$

for  $f \in H^1(\mathbb{R})$ .

Theorem 19.5

Theorem 19.9

Theorem 19.10

Theorem 19.12 is due to Fefferman in 1972.

Section 20. The sharp maximal operator came about in the paper [183], where Fefferman and Stein used  $f^\sharp$  instead of  $M^\sharp f$ . It appeared implicitly in [259]. In the subsequent textbook [58] Stein took up the sharp maximal operator. Theorem 20.3, which was called the distributional inequality in the age of [183], began to appear in [183]. Nowadays this stuff can be found in [10].

Theorem 20.5 was proved by C. Fefferman and E. M. Stein in the celebrated paper [183]. It is true that we really need  $f$  to have some decay property at infinity. However, the condition  $\min(1, |f|) \in L^p(\mathbb{R}^d)$ , or  $f \in L^q(\mathbb{R}^d)$ ,  $0 < q < p$  is somehow strong. It deserves attention to loosen this assumption. In [198], Fujii showed that it suffices to assume

$$(21.82) \quad \lim_{j \rightarrow \infty} m_{2^j I}(f) = 0$$

for some fixed cube  $I$ . For a generalization of this result to a non-homogeneous space we refer to [427, 428]. In 2010 Lerner revisited this idea [309]. In [352] A. Miyachi and K. Yabuta found another condition that can be used as a substitute to  $f \in L^q(\mathbb{R}^d)$ . K. Yabuta and A. Miyachi investigated the condition of the (weighted) sharp maximal inequality and concluded that the integrability condition is replaced by the following condition.

$$(21.83) \quad \sup_{\lambda > 0} \lambda^\alpha w \{x \in \mathbb{R}^d : |f(x)| > \lambda\} < \infty$$

for some  $\alpha > 0$ .

Section 21. The theory of  $A_p$ -weights dates back to Muckenhoupt [361]. In [16] the technique of vector-valued inequalities are dealt in great detail. We refer to [257] as well.

Theorem 21.1 is obtained by Fefferman and E. M. Stein [182].

Theorem 21.5

Theorem 21.7 is due to Córdoba and Fefferman (see [150]).

Stein firstly considered the weight of the type  $|x|^\alpha$  as in Proposition 21.8. Muckenhoupt proved Theorem 21.21 in [361, 362] and Coifman and Fefferman proved it independently in [142].

Theorem 21.11 is obtained by Coifman and Rochberg (see [145]).

Theorem 21.23

Theorem 21.17 was originally due to Muckenhoupt [362]. An alternative proof using Theorem 21.23 was obtained by Coifman-Fefferman [142]. Later Hunt, Kurtz and Neugebauer gave a new proof avoiding the usage of Theorem 21.23 (see [248]). Later Christ and Fefferman gave another new proof without using Theorem 21.23 but depends on the Calderón-Zygmund decomposition. The proof given here depends on [306]. A. K. Lerner noticed that the proof can be made without using such a classical and heavy tools. The proof was surprisingly simple, as we have seen.

Theorem 21.24 was obtained by Peter W. Jones [385]. Later B. Garnett and P. Jones obtained a simpler proof in [208]. We refer to [403, 404] as well.

We refer to [307] for the extended version of Theorem 21.25.

Hunt, Muckenhoupt and Wheeden obtained Theorems 21.29, 21.30 for the Hilbert transform in [249].

Theorem 21.31

We remark that  $w \in A_p$  is a necessary condition for singular integral operators to be  $L^p(w)$ -bounded, as well. Speaking precisely, we have the following result.

**Theorem 21.32.** *Suppose that  $w$  is a weight.*

- (1) *Let  $1 < p < \infty$ . Assume  $w$  satisfies the following weighted norm estimates of the Riesz transforms  $R_1, R_2, \dots, R_d$ : For every  $j = 1, 2, \dots, d$ ,  $\|R_j f\|_{L^p(w)} \lesssim \|f\|_{L^p(w)}$  for all  $f \in L_c^\infty(\mathbb{R}^d)$ . Then  $w \in A_p$ .*
- (2) *Assume  $w$  satisfies the following weighted norm estimates of the Riesz transforms: For every  $j = 1, 2, \dots, d$ ,  $f \in L_c^\infty(\mathbb{R}^d)$  and  $\lambda > 0$ , we have  $\lambda |\{ |R_j f| > \lambda \}| \lesssim \|f\|_{L^1(w)}$ . Then  $w \in A_1$ .*

For the proof of this theorem we refer to the textbook [16] for example.

The reverse Hölder inequality is an equivalent condition to the  $A_\infty$  condition, as we have seen. There are several attempts to quantify the reverse Hölder inequality. Let  $1 < p \leq \infty$ . We say that a weight  $V$  belongs to the class  $(RH)_p$ , if

$$(21.84) \quad \frac{1}{|B(x, r)|^{\frac{1}{p}}} \|\chi_{B(x,r)} V\|_p \lesssim m_{B(x,r)}(V)$$

for all  $x \in \mathbb{R}^d$  and  $r > 0$ . Let  $k \in \mathbb{Z}$  and  $\alpha > 0$ . In [180] Fefferman established that

$$(21.85) \quad \frac{1}{|B(x, r)|^{\frac{1}{p}}} \|\chi_{B(x,r)} V\|_p \lesssim_{k,\alpha} m_{B(x,r)}(V)$$

holds for all functions  $V$  such that  $V(x) = |P(x)|^\alpha$  for some polynomial  $P$  of degree  $k$ .

Finally we remark that the weighted inequality for the potential operator  $I_\alpha$  such as

$$(21.86) \quad \int_{\mathbb{R}^d} |I_\alpha f(x)|^p v(x) dx \lesssim \int_{\mathbb{R}^d} |f(x)|^p dx$$

was investigated by Kokilashvili and Meskhi [283]. We refer to [11] for a huge amount of this field, where we can find so many about this type of weighted inequality.

Proposition 13.4 immediately gives us

$$(21.87) \quad \left\| \sup_{t>0} \left| \varphi \left( \frac{\cdot}{t} \right) * f \right| \right\|_p \lesssim \|\phi\|_1 \|f\|_p$$

for all  $1 < p < \infty$  whenever  $\varphi$  is a radial decreasing function. The aforementioned proposition allows us to extend it to the weighted inequality.

$$(21.88) \quad \left\| \sup_{t>0} \left| \varphi \left( \frac{\cdot}{t} \right) * f \right| \right\|_{L^p(w)} \lesssim \|\phi\|_1 \|f\|_{L^p(w)}$$

for  $w \in A_p$ . However, more can be said about this weighted inequality. H. Gunawan established

$$(21.89) \quad \left\| \sup_{t>0} \left| \varphi \left( \frac{\cdot}{t} \right) * f \right| \right\|_{L^p(w^\delta)} \lesssim \|\phi\|_1 \|f\|_{L^p(w^\delta)}$$

if  $w \in A_p$  and  $0 \leq \delta < \frac{p(n-1)-n}{n(p-1)}$ . H. Gunawan proved this by means of the Mellin transform [223].

We refer to [252] for an example of doubling weights that are not  $A_\infty$ -weights.

**Part 10. Probability theory, martingale and ergodicity****Part 11. Probability theory**

In this part we take up two topics which are very close to the ones in Part 2.

The aim of the first half of this part, Chapter 11, is to make an introductory view of probability theory and to compare martingale theory with the ones of maximal and of singular integral operators. In Section 22 we build up elementary notions. Section 23 we are going to define martingale with discrete time. Section 24 is the heart of the martingale theory, where we will see a lot of similarities of maximal operators and martingales: They are very close to each other in mathematical structure. The book is not oriented to a thorough introduction of martingale theories. For example, we do not take up theory of martingale with continuous time, which is important in finance. However, it is benefiting to compare probability theory with harmonic analysis. In this part we discuss martingales with discrete time  $\mathbb{N}_0$ . It will be arranged that preliminary facts for stochastic integral be ready after reading this part.

## 22. SOME ELEMENTARY NOTIONS

## 22.1. Probability spaces.

We begin with the definition of probability space on which probability theory is staged.

**Definition 22.1.** A probability space is a measure space  $(\Omega, \mathcal{F}, P)$  such that  $P(\Omega) = 1$ .

At first glance it seems to have nothing with the probability appearing in daily life. To get feeling of probability let us make a brief view of examples of the probability spaces.

**Example 22.2** (Coin tossing). Now we present a model of coin-tossing. Let

$$(22.1) \quad \Omega = \{ \text{head}, \text{tail} \}, \mathcal{F} = 2^\Omega.$$

Define a measure  $P$  uniquely so that

$$(22.2) \quad P(\{ \text{head} \}) = P(\{ \text{tail} \}) = \frac{1}{2}.$$

If we are to toss a coin twice, the model is given as follows:

$$(22.3) \quad \Omega := \{ \text{head head}, \text{head tail}, \text{tail head}, \text{tail tail} \}$$

and  $\mathcal{F} := 2^\Omega$ .

$$(22.4) \quad P(\{ \text{head head} \}) = P(\{ \text{head tail} \}) = P(\{ \text{tail head} \}) = P(\{ \text{tail tail} \}) = \frac{1}{4}$$

**Example 22.3.** Let  $\Omega = [0, 1]$ ,  $\mathcal{F}$  the Borel  $\sigma$ -field on  $[0, 1]$  and  $P$  the restriction of the Lebesgue measure on  $\Omega$ . Then  $(\Omega, \mathcal{F}, P)$  is a probability space.

*Exercise 158.* Present a mathematical model of the following situations.

- (1) Tossing a coin three times.
- (2) Tossing a die once.
- (3) Tossing two dice.
- (4) Two persons are playing a game of "Rock Paper Scissors".
- (5) Tossing a die and a coin.

Events. In probability theory we often consider events, that is, what can happen under the situation we envisage. For example, if we toss four dice, then we can expect that all the numbers are different.

Let us see how this is expressed mathematically.

**Definition 22.4.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space. Elements in  $\mathcal{F}$  is called events.

**Example 22.5.** Let  $\Omega = \{1, 2, 3, 4, 5, 6\}^4$  and  $\mathcal{F} = 2^\Omega$ . Define a probability measure  $P$  by

$$(22.5) \quad P\{(a, b, c, d)\} = \frac{1}{1296}$$

for all  $(a, b, c, d) \in \Omega$ . Then we can express the event that all the numbers are different as follows:

$$(22.6) \quad \{(a, b, c, d) : a \neq b, a \neq c, a \neq d, b \neq c, b \neq d, c \neq d\}.$$

*Exercise 159.* In Example 22.2, how many events are there ?

Random variables. Having set down the probability space, we are going to give the definition of random variables. In the discrete setting it is sufficient to define them as the mapping from  $\Omega$ . However, the measure coming into play, we need take into account the measurability. Thus, we are led to the following definition.

**Definition 22.6.** A random variable on a probability space  $(\Omega, \mathcal{F}, P)$  is a measurable mapping from  $\Omega$  to  $\mathbb{R} \cup \{\pm\infty\}$  or  $\mathbb{C}$ .

**Example 22.7.** In Example 22.5, the functions

$$(22.7) \quad f((a, b, c, d)) = \frac{a + b + c + d}{6}, \quad g((a, b, c, d)) = abcd, \quad h((a, b, c, d)) = \frac{a}{b}$$

are all random variables.

**Definition 22.8.** A random variable  $X$  is said to be integrable or to belong to  $L^1(P)$ , if

$$(22.8) \quad \int_{\Omega} |X(\omega)| dP(\omega) < \infty.$$

In this case the expectation of  $X$  is defined as

$$(22.9) \quad E[X] := \int_{\Omega} X(\omega) dP(\omega).$$

**Definition 22.9.** Suppose that  $X$  and  $Y$  are  $L^2(P)$ -random variables. Then define

$$(22.10) \quad \text{Var}(X) := E[(X - E[X])^2], \quad \text{Cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Example 22.10.** We toss a die once and let  $X$  be the number. Then  $E[X] = \frac{7}{6}$ .

**Example 22.11.** Imagine that you arrive at the bus stop and catch one. The bus will come every 20-minutes. If you reach the bus stop without knowing what time it is, then how long will you be kept waited ? To model this setting, we set

$$(22.11) \quad \Omega := [0, 60), \quad \mathcal{F} := \mathcal{B}([0, 60)), \quad P := \frac{1}{60} dx|_{[0, 60)}.$$

Define  $f : \Omega \rightarrow \mathbb{R}$  by

$$f(t) = \begin{cases} 20 - t & 0 \leq t < 20 \\ 40 - t & 20 \leq t < 40 \\ 60 - t & 40 \leq t < 60. \end{cases}$$

Then the expected waiting time is given by

$$(22.12) \quad E[f] = \frac{1}{60} \int_0^{20} (20 - t) dt + \frac{1}{60} \int_{20}^{40} (40 - t) dt + \frac{1}{60} \int_{40}^{60} (60 - t) dt = \frac{600}{60} = 10.$$



*Exercise 160.* In the above example the diagram has changed and in every hour the bus will come

$$(22.13) \quad 0, 12, 20, 32, 40, 52 \text{ min.}$$

Then how long is the expected waiting time ?

**Definition 22.12** (Independent random variables). Let  $(\Omega, \mathcal{F}, P)$  be a probability space.

(1) A family of  $\mathbb{R}$ -valued random variables  $\{X_\lambda\}_{\lambda \in \Lambda}$  is said to be independent, if, for every

$$(22.14) \quad \lambda_1, \lambda_2, \dots, \lambda_n \in \Lambda \text{ and } A_1, A_2, \dots, A_n \in \mathcal{B}(\mathbb{R} \cup \{\pm\infty\}),$$

we have

$$(22.15) \quad P(X_{\lambda_1} \in A_1, X_{\lambda_2} \in A_2, \dots, X_{\lambda_n} \in A_n) = \prod_{j=1}^n P(X_{\lambda_j}).$$

(2) Let  $\mathcal{G}_0$  and  $\mathcal{G}_1$  be  $\sigma$ -subalgebras of  $\mathcal{F}$ .  $\mathcal{G}_0$  and  $\mathcal{G}_1$  are said to be independent if  $\chi_{G_0}$  and  $\chi_{G_1}$  are independent for all  $G_0 \in \mathcal{G}_0$  and  $G_1 \in \mathcal{G}_1$ . In this case write  $\mathcal{G}_0 \amalg \mathcal{G}_1$ .

**Example 22.13.** In Example 22.2 we define

$$(22.16) \quad f(\text{head head}) = 1, f(\text{head tail}) = 1, f(\text{tail head}) = 0, f(\text{tail tail}) = 0$$

and

$$(22.17) \quad g(\text{head head}) = 1, g(\text{head tail}) = 0, g(\text{tail head}) = 1, g(\text{tail tail}) = 0.$$

Then  $f$  and  $g$  are independent.

We shall postpone the proof of the following theorem till Chapter 27.

**Theorem 22.14.** Let  $(\Omega_j, \mathcal{F}_j, P_j)_{j \in \mathbb{N}}$  be a sequence of probability spaces. Suppose that for each  $j$  we are given a random variable  $X_j$ . Then there exist a probability space  $(\Omega, \mathcal{F}, P)$  and a sequence of random variables  $Y_j$  on  $\Omega$  such that  $Y_j$  and  $X_j$  are identically distributed.

**Proposition 22.15.** Let  $\{X_\lambda\}_{\lambda \in \Lambda}$  be a family of independent  $\mathbb{R}$ -valued random variables. Suppose that each  $f_\lambda : \mathbb{R} \rightarrow \mathbb{R}$  is measurable. Then  $\{f_\lambda(X_\lambda)\}_{\lambda \in \Lambda}$  is independent.

**Theorem 22.16.** Suppose that  $X, Y \in L^1(P)$  are independent. Then  $XY \in L^1(P)$  and

$$(22.18) \quad E[XY] = E[X]E[Y].$$

*Proof.* We may assume that  $X$  and  $Y$  are positive. By considering the composition with  $f_n$ , given by  $f_n(t) = \min\left(n, 2^n \left\lfloor \frac{t}{2^n} \right\rfloor\right)$ , we may further assume that  $X$  and  $Y$  are step functions. In this case the proof is not so difficult from Definition 22.12.  $\square$

Without using Theorem 22.14, we shall construct a countable sequence of independent random variables.

**Example 22.17** (Rademacher sequence). In this example we let

$$(22.19) \quad (\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), dx|_{[0, 1]}).$$

We define a sequence of random variables  $r_n, n = 1, 2, \dots$  by

$$(22.20) \quad r_n(t) := \sum_{j=1}^{2^n} (-1)^{j-1} \mathbf{1}_{[(j-1)2^{-n}, j2^{-n})}(t).$$

*Exercise 161.* Show that  $r_n(t) = \text{sgn}(\sin(2^n \pi t))$  a.e.  $t \in [0, 1]$ .

Later, this sequence will turn out to be powerful in extending the boundedness of operators.

**Lemma 22.18.** *The functions  $\{r_n(t)\}_{n=1}^\infty$  form a family of independent random variables.*

*Proof.* We have to show for  $A_1, A_2, \dots, A_k \in \mathcal{B}(\mathbb{R})$

$$(22.21) \quad P(r_{n_1}(t) \in A_1, r_{n_2}(t) \in A_2, \dots, r_{n_k}(t) \in A_k) = \prod_{j=1}^k P(r_{n_j}(t) \in A_j),$$

where  $n_1 < n_2 < \dots < n_k$ . By setting  $A_j := \mathbb{R}$ , if  $j \in \{1, 2, \dots, n_k\} \setminus \{n_1, n_2, \dots, n_k\}$ , we may assume  $n_j = j$  for all  $j = 1, 2, \dots, k$ . Taking into account what value  $r_n$ 's take, we may assume that  $A_j \subset \{-1, 1\}$ . Once we show (22.21) for  $A_j$  such that  $\#A_j = 1$ , the passage for general  $A_j$ 's can be achieved by summing up (22.21) with  $\#A_j = 1, j = 1, 2, \dots, k$ . Thus we have only to prove

$$(22.22) \quad P(r_1(t) = (-1)^{n_1}, r_2(t) = (-1)^{n_2}, \dots, r_k(t) = (-1)^{n_k}) = 2^{-k}.$$

for  $n_1, n_2, \dots, n_k$  with  $\{n_1, n_2, \dots, n_k\} \in \{0, 1\}$ . Suppose  $\sum_{j=1}^k \frac{m_j}{2^j} \leq t < \sum_{j=1}^k \frac{m_j}{2^j} + 2^{-k}$ , where

$m_1, m_2, \dots, m_k \in \{0, 1\}$ . Then  $r_j(t) = r_j \left( t - \sum_{l=1}^j \frac{m_l}{2^l} \right) = (-1)^{m_j}$  Consequently

$$(22.23) \quad \{t \in \mathbb{R} : r_j(t) = (-1)^{n_j}, j = 1, 2, \dots, k\} = \left\{ t \in \mathbb{R} : \sum_{j=1}^k \frac{n_j}{2^j} \leq t < \sum_{j=1}^k \frac{n_j}{2^j} + 2^{-k} \right\}.$$

Thus (22.22) is justified. □

**Example 22.19.** Suppose that  $(X, \Omega, P) = ([0, 1), \mathcal{B}([0, 1)), dx|[0, 1))$ . Consider a sequence  $\{X_n\}_{n \in \mathbb{N}}$  of random variables given by

$$(22.24) \quad X_n = r_1 + r_2 + \dots + r_n$$

for  $n \in \mathbb{N}$ .  $\{X_n\}_{n \in \mathbb{N}}$  is used as a model of random walking: It grasps mathematically how likely is a drunken man stay in the coordinate of the real line  $\mathbb{R}$  after the  $n$ -th step.

### 22.2. The Characteristic functions.

Distribution functions. In considering something uncertain it is not what the exact value of a random variables that counts. Let us toss a die twice. If we do not have to consider the relation between two dice, then it is sufficient to consider the number of each die independently.

We can say the distribution reflects the structure of the random variable in the sense that how distributed the random variable is and that the distribution kills the information of the underlying probability space.

**Definition 22.20.** Given a random variable  $X : \Omega \rightarrow \mathbb{R}^d$ , one defines the distribution  $\mu_X$  of  $X$  as the measure on  $\mathbb{R}^d$ , which satisfies

$$(22.25) \quad \mu_X(E) = P(X^{-1}(E)) \quad (E \in \mathcal{B}(\mathbb{R}^d)).$$

**Lemma 22.21.** *Let  $f$  be a positive measurable function on  $\mathbb{R}^d$ . Then we have*

$$(22.26) \quad \int_{\Omega} f(X(\omega)) dP(\omega) = \int_{\mathbb{R}^d} f(x) d\mu_X(x)$$

for every  $\mathbb{R}^d$ -valued random variable  $X$ .

*Proof.* The case when  $f = \mathbf{1}_E$  being trivial, we have only to go through a limiting argument as usual.  $\square$

**Example 22.22.** Let  $X = (X_1, X_2, \dots, X_d)$  be a random variable whose distribution is

$$(22.27) \quad \mu_{m,A} = \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx,$$

where  $A$  is a  $d \times d$  strictly positive definite matrix and  $m \in \mathbb{R}^d$  is a fixed vector. First of all let us verify that the total probability of  $\mu_{m,A}$  is 1. That is, let us check

$$(22.28) \quad \mu_{m,A}(\mathbb{R}^d) = 1.$$

Before we calculate let us make a review of the square root of positive matrices. A matrix is said to be positive, if  ${}^t x A x \geq 0$  for all  $x \in \mathbb{R}^d$ . Here  ${}^t x$  denotes the transpose of  $x$ . The square root of  $A$  is a unique positive matrix  $B$  such that  $B^2 = A$ .

Returning to the proof of (22.28), let us set  $B$  as the square root of  $A$ . Then a series of changing variables gives us

$$\begin{aligned} \mu_{m,A}(\mathbb{R}^d) &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}z \cdot A^{-1}z\right) dz \\ &= \frac{\det(B)}{(2\pi)^{\frac{d}{2}}(\det(B^2))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}Bw \cdot A^{-1}Bw\right) dw \end{aligned}$$

Since  $BA^{-1}B = \text{id}_{\mathbb{R}^d}$ , we obtain

$$\mu_{m,A}(\mathbb{R}^d) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}w^2\right) dw = \frac{1}{(2\pi)^{\frac{d}{2}}} \cdot \Gamma\left(\frac{1}{2}\right)^d = 1.$$

Next, let us calculate the average and covariance. As for the average, we have

$$\begin{aligned} E[X_j] &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} x_j \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} (x_j - m_j) \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &\quad + \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} m_j \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &= \frac{m_j}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &= m_j. \end{aligned}$$

Let  $j \neq k$ . Denote by  $(Bz)_j$  the  $j$ -th component of  $Bz \in \mathbb{R}^d$ . Then another series of changing variables gives us

$$\begin{aligned} \text{Cov}(X_j, X_k) &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} (x_j - m_j)(x_k - m_k) \exp\left(-\frac{1}{2}(x-m) \cdot A^{-1}(x-m)\right) dx \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}(\det(A))^{\frac{1}{2}}} \int_{\mathbb{R}^d} y_j y_k \exp\left(-\frac{1}{2}y \cdot A^{-1}y\right) dy \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} (Bz)_j \cdot (Bz)_k \exp\left(-\frac{1}{2}Bz \cdot A^{-1}Bz\right) dz \end{aligned}$$

Expanding  $(Bz)_j \cdot (Bz)_k$ , we obtain

$$\text{Cov}(X_j, X_k) = \sum_{l,m=1}^d \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} b_{jl} b_{km} z_l z_m \exp\left(-\frac{1}{2}z^2\right) dz = \sum_{l=1}^d b_{jl} b_{kl} = a_{jk}.$$

Therefore,  $N(m; A)$ , whose distribution is  $\mu_{m,A}$ , has average  $m$  and covariance matrix  $A$ .

**Characteristic functions.** The characteristic function of measures, roughly speaking, is the Fourier transform of the distribution of the measures. In this book, due to the notation adopted in the previous chapters, the characteristic function is the constant multiple of the Fourier transform of the measures, whose definition we gave in Chapter 4. The characteristic function and the distribution are connected to each other. Furthermore, the convergence of the distributions can be characterized by the convergence of the measure. In this present paragraph as an application we consider the central limit theorem as well.

**Definition 22.23.** Let  $X$  be an  $\mathbb{R}^d$ -valued random variable. Then denoting by  $\mu_X$  the distribution of  $X$ , one defines

$$(22.29) \quad \varphi_X(\xi) := \int_{\mathbb{R}^d} e^{ix \cdot \xi} d\mu_X(\xi).$$

**Example 22.24.** Let  $A$  be strictly positive definite matrix. Let us calculate the distribution of  $N(m; A)$ . Recall that the density is given by

$$(22.30) \quad \mu_{m,A} = \frac{1}{(2\pi)^{\frac{d}{2}} (\det(V))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-m) \cdot V^{-1}(x-m)\right) dx.$$

Let  $B$  be the square root of  $A$ . Therefore, the characteristic function is give by

$$\begin{aligned} \varphi_{m,V}(\xi) &= \frac{1}{(2\pi)^{\frac{d}{2}} (\det(V))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}(x-m) \cdot V^{-1}(x-m) + ix \cdot \xi\right) dx \\ &= \exp(im \cdot \xi) \frac{1}{(2\pi)^{\frac{d}{2}} (\det(V))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}y \cdot V^{-1}y + iy \cdot \xi\right) dy \\ &= \exp(im \cdot \xi) \frac{\det(B)}{(2\pi)^{\frac{d}{2}} (\det(B^2))^{\frac{1}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}(Bz) \cdot V^{-1}Bz + iBz \cdot \xi\right) dz. \end{aligned}$$

If we complete the square, then we obtain

$$\begin{aligned} \varphi_{m,V}(\xi) &= \exp(im \cdot \xi) \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}(z - iB\xi)^2 - \frac{1}{2}B\xi \cdot B\xi\right) dy \\ &= \exp(im \cdot \xi) \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}z^2 - \frac{1}{2}B\xi \cdot B\xi\right) dy \\ &= \exp\left(im \cdot \xi - \frac{1}{2}\xi \cdot A\xi\right), \end{aligned}$$

where for the fifth equality we have used the complex line integral.

From the results we obtain in Subsection 7.4, we have the following.

**Theorem 22.25.** Let  $X$  and  $Y$  be  $\mathbb{R}^d$ -valued random variables. If their characteristic functions agree, then so do their distributions.

**Definition 22.26.** A sequence of random variables  $X_1, X_2, \dots, X_n, \dots$  is said to converge in law, if

$$(22.31) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f(x) d\mu_{X_n}(x) = \int_{\mathbb{R}^d} f(x) d\mu_X(x)$$

for all bounded continuous functions in  $\mathbb{R}^d$ .

Theorem 7.45 can be rephrased as follows :

**Theorem 22.27.** *The necessary and sufficient condition for the sequence of random variables  $X_1, X_2, \dots, X_n, \dots$  to converges to  $X$  in law is that  $\lim_{n \rightarrow \infty} \varphi_{X_n}(\xi) = \varphi_X(\xi)$  for all  $\xi \in \mathbb{R}^d$ .*

**Example 22.28** (Generalized normal distribution). Let  $A$  be a positive definite matrix whose eigenvalues can take 0. For  $\varepsilon > 0$  we define  $\mu_\varepsilon = N(m; A + \varepsilon I)$ . Then the characteristic function  $\varphi_\varepsilon$  of  $\mu_\varepsilon$  is given by

$$(22.32) \quad \varphi_\varepsilon(\xi) = \exp \left( im \cdot \xi - \frac{1}{2} \xi \cdot (A + \varepsilon I) \xi \right).$$

Therefore,  $\varphi_\varepsilon$  tends to a limit  $\varphi$  uniformly over a neighborhood of 0. As a consequence  $\varphi$  is a characteristic function of a distribution  $\mu$ . We define  $N(m; A) = \mu$ , if  $A$  is not strictly positive definite. This is the natural extension of normal distributions.

Central limit theorem. As an application of characteristic functions let us prove the central limit theorem.

**Theorem 22.29** (The de Moivre and Laplace central theorem). *Suppose that we are given  $X_1, X_2, \dots, X_n, \dots$  are sequence of i.i.d. square integrable random variables. Set*

$$(22.33) \quad Y_n := \sqrt{n} \left( \frac{X_1 + X_2 + \dots + X_n}{n} - E[X_1] \right)$$

*Then  $\{Y_n\}_{n \in \mathbb{N}}$  converges to the normal distribution  $N(0, \text{Var}[X])$ .*

*Proof.* To prove this, we consider the characteristic function of  $Y_n$ .

$$\begin{aligned} \varphi_{Y_n}(\xi) &= \int_{\Omega} \exp \left( i\xi \frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}} - i\xi \sqrt{n} E[X_1] \right) dP \\ &= \prod_{j=1}^n \int_{\Omega} \exp \left( i\xi \frac{X_j}{\sqrt{n}} - i\xi \sqrt{n} E[X_1] \right) dP \\ &= \left( \int_{\Omega} \exp \left( i\xi \frac{X_1 - E[X_1]}{\sqrt{n}} \right) dP \right)^n. \end{aligned}$$

Here for the last equality, we have used the assumption that  $\{X_j\}_{j \in \mathbb{N}}$  is an i.i.d. sequence. Observe that

$$\begin{aligned} &n \left\{ 1 - \frac{\xi^2}{2} \text{Var}(X_1) - \left( \int_{\Omega} \exp \left( i\xi \frac{X_1 - E[X_1]}{\sqrt{n}} \right) dP \right) \right\} \\ &= \int_{\Omega} n \left( 1 + i\xi \frac{X_1 - E[X_1]}{\sqrt{n}} - \frac{\xi^2}{2} (X_1 - E[X_1])^2 - \exp \left( i\xi \frac{X_1 - E[X_1]}{\sqrt{n}} \right) \right) dP. \end{aligned}$$

Now we observe that the integrand is bounded by

$$(22.34) \quad c |X_1|^2,$$

where  $c$  is the absolute constant independent of  $\omega$ . Therefore, we are in the position of using the Lebesgue convergence theorem to obtain

$$(22.35) \quad \lim_{n \rightarrow \infty} n \left\{ 1 - \frac{\xi^2}{2} \text{Var}(X_1) - \left( \int_{\Omega} \exp \left( i\xi \frac{X_1 - E[X_1]}{\sqrt{n}} \right) dP \right) \right\} = 0.$$

Consequently letting  $n \rightarrow \infty$ , we finally obtain

$$(22.36) \quad \lim_{n \rightarrow \infty} \varphi_{Y_n}(\xi) = \lim_{n \rightarrow \infty} \left( 1 - \frac{\xi^2}{2n} \text{Var}(X_1) \right)^n = \exp(-\xi^2 \text{Var}(X_1)).$$

Since the characteristic function of the normal distribution  $N(0; \text{Var}(X_1))$  is the right-hand side itself, we obtain the desired result.  $\square$

### 22.3. Conditional expectation.

In this section we keep to the notation before. We let  $(\Omega, \mathcal{F}, P)$  be a probability space.

Suppose that there are five white balls and three red balls. We put each of them in each box. The balls are in the boxes so that we cannot tell their color from outside the boxes.

Choose one of the boxes and guess the color of the ball inside of it. Then it is more likely that we say “white”.

After opening three boxes, it turned out that all the balls are white. Then we are going to open the fourth one. What color is more likely? Of course the answer is red. Because there are three red balls remaining in the boxes, while there are only two white balls.

As the above situation suggests, the strategy differs from the situations. Thus, the probability according to the present situation counts a lot.

**Definition 22.30.** Let  $A$  be an event with  $P(A) > 0$ . Then the conditional probability of  $B$  under  $A$  is given by

$$(22.37) \quad P_A(B) = \frac{P(A \cap B)}{P(A)}.$$

Let us describe the above situation guessing the color of the balls mathematically.

**Example 22.31.** Let  $A = \{(w, 1), (w, 2), (w, 3), (w, 4), (w, 5), (r, 1), (r, 2), (r, 3)\}$ .

$$(22.38) \quad \Omega = \{(a_1, a_2, \dots, a_8) : \{a_1, a_2, \dots, a_8\} = A\}.$$

We set  $\mathcal{F} = 2^\Omega$  as usual. Note that  $\#\Omega = 8!$ . Define a measure  $P$  by  $P(\{\omega\}) = \frac{1}{8!} = \frac{1}{40320}$  for each  $\omega \in \Omega$ . Define  $p_i : \Omega \rightarrow \{w, r\}$  by

$$(22.39) \quad p_i((a_1, a_2, \dots, a_8)) := \text{“the first component of the } a_i\text{”}$$

for  $i = 1, 2, \dots, 8$ . Then we have

$$\text{“event that the color of the ball of the first box is white”} = p_1^{-1}(w)$$

$$\text{“event that the color of the ball of the fourth box is red”} = p_4^{-1}(r).$$

Then the probability that the color of the ball of the first box is white is

$$(22.40) \quad P(p_1^{-1}(w)) = \frac{\#p_1^{-1}(w)}{8!} = \frac{5 \cdot 7!}{8!} = \frac{5}{8}.$$

Meanwhile, after knowing that the first three boxes contain white balls, the probability that the color of the ball of the fourth box is red is

$$\begin{aligned} P_{p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w)}(p_4^{-1}(r)) &= \frac{P(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w) \cap p_4^{-1}(r))}{P(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w))} \\ &= \frac{\#(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w) \cap p_4^{-1}(r))/8!}{\#(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w))/8!} \\ &= \frac{\#(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w) \cap p_4^{-1}(r))}{\#(p_1^{-1}(w) \cap p_2^{-1}(w) \cap p_3^{-1}(w))} \\ &= \frac{5 \cdot 4 \cdot 3 \cdot 3 \cdot 4!}{5 \cdot 4 \cdot 3 \cdot 5!} \\ &= \frac{3}{5}. \end{aligned}$$

This does not run counter to our intuition.

*Exercise 162.* Suppose that we have to select two numbers from 1, 2, 3, 3 randomly.

- (1) Suppose that the first number is 2. Then how likely do you choose 3 ?
- (2) Suppose that the first number is 3. Then how likely do you choose 3 ?

**Theorem 22.32.** *Suppose that  $(\Omega, \mathcal{F}, P)$  is a probability space. Let  $\mathcal{G} \subset \mathcal{F}$  be a sub- $\sigma$  field. We write  $\tilde{P} = P|_{\mathcal{G}}$ . Take  $X \in L^1(\Omega, \mathcal{F}, P)$  arbitrarily. Then we can find  $Z \in L^1(\Omega, \mathcal{G}, P)$  such that*

$$(22.41) \quad \int_A Z(\omega) d\tilde{P}(\omega) = \int_A X(\omega) dP(\omega)$$

for all  $A \in \mathcal{G}$ .

*Proof.* We define a signed measure  $Q$  by

$$(22.42) \quad Q(A) = \int_A X(\omega) dP(\omega), \quad A \in \mathcal{G}.$$

Then  $Q$  is absolutely continuous with respect to  $P$ , that is,  $A \in \mathcal{G}$  and  $P(A) = 0$  implies  $Q(A) = 0$ . This implies  $Q$  has a density  $Z \in L^1(P)$  with respect to  $P$  by the Radon-Nikodym theorem. Therefore, we obtain

$$(22.43) \quad \int_A Z(\omega) dP(\omega) = Q(A) = \int_A X(\omega) dP(\omega)$$

for all  $A \in \mathcal{G}$ .

The uniqueness can be proved by noting the fact that in the measure space  $(\Omega, \mathcal{G}, P)$  0 is the only  $\mathcal{G}$ -measurable function of which the integral on every event is zero.  $\square$

**Proposition 22.33.** *Let  $(\Omega, \mathcal{F}, P)$  be a probability space. Suppose that  $\Omega$  is partitioned into a finite sum:*

$$(22.44) \quad \Omega = \Omega_1 \amalg \Omega_2 \amalg \dots \amalg \Omega_k.$$

Set

$$(22.45) \quad \mathcal{B} := \left\{ \prod_{j \in \Lambda} \Omega_j : \Lambda \subset \{1, 2, \dots, k\} \right\}.$$

Let  $P$  be a probability measure. Then we have

$$(22.46) \quad E[\mathbf{1}_A : \mathcal{B}] = \sum_{j=1}^k \frac{P(A \cap \Omega_k)}{P(\Omega_k)} \mathbf{1}_{\Omega_k} = \sum_{j=1}^k P_{\Omega_k}(A) \mathbf{1}_{\Omega_k}$$

for all  $A \in \mathcal{F}$ .

*Proof.* First, since we have

$$(22.47) \quad \sum_{j=1}^k P_{\Omega_k}(A) \mathbf{1}_{\Omega_k} \in \mathcal{B}$$

by definition, what remains to be proved is

$$(22.48) \quad \int_B \left( \sum_{j=1}^k P_{\Omega_j}(A) \mathbf{1}_{\Omega_j}(\omega) \right) dP(\omega) = \int_B \mathbf{1}_A(\omega) dP(\omega)$$

for all  $B \in \mathcal{B}$ . By definition of  $\mathcal{B}$ , we can write  $B = \prod_{l \in \Lambda} \Omega_l$  with some  $\Lambda \subset \{1, 2, \dots, k\}$ . Using this expression, we have

$$\begin{aligned} \int_B \left( \sum_{j=1}^k P_{\Omega_j}(A) \mathbf{1}_{\Omega_j}(\omega) \right) dP(\omega) &= \sum_{l \in \Lambda} \int_{\Omega_l} \left( \sum_{j=1}^k P_{\Omega_j}(A) \mathbf{1}_{\Omega_j}(\omega) \right) dP(\omega) \\ &= \sum_{l \in \Lambda} P(A \cap \Omega_l) \\ &= P(A \cap B) \\ &= \int_B \mathbf{1}_A(\omega) dP(\omega). \end{aligned}$$

This is the desired result.  $\square$

**Definition 22.34.** Write  $E[X : \mathcal{G}] := Z$  and call it the conditional expectation of  $X$  with respect to  $\mathcal{G}$ .

Here we collect fundamental properties of the conditional expectation operator.

**Theorem 22.35.** *Suppose that  $(\Omega, \mathcal{F}, P)$  is a probability space,  $X, Y \in L^1(\Omega, \mathcal{F}, P)$ ,  $c \in \mathbb{R}$  and  $\mathcal{G} \subset \mathcal{F}$  is a sub- $\sigma$ -field.*

(1) *Suppose further  $\mathcal{G} \perp \mathcal{H}$ . Then*

$$(22.49) \quad E[X : \mathcal{G}] = E[X] \text{ a.s.}$$

*for all  $X \in \mathcal{H}$ . In particular, if  $\mathcal{G} = \{\varphi, X\}$ . Then  $E[X : \mathcal{G}] = E[X]$  a.s.*

(2)  *$E[X + Y : \mathcal{G}] = E[X : \mathcal{G}] + E[Y : \mathcal{G}]$  a.s.*

(3) *Let  $a \in \mathbb{K}$ . Then  $E[aX : \mathcal{G}] = a \cdot E[X : \mathcal{G}]$ .*

(4) *If  $Z \in L^\infty(\Omega, \mathcal{G}, P)$ , then  $E[ZX : \mathcal{G}] = ZE[X : \mathcal{G}]$  a.s.*

(5)  *$X \geq 0$  implies  $E[X : \mathcal{G}] \geq 0$ .*

(6) *Suppose that  $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}$  are sub- $\sigma$ -fields and  $\mathcal{G}_1 \subset \mathcal{G}_2$ . Then*

$$(22.50) \quad E[E[X : \mathcal{G}_2] : \mathcal{G}_1] = E[E[X : \mathcal{G}_1] : \mathcal{G}_2] = E[X : \mathcal{G}_1] \text{ a.s.}$$

In view of (22.50) we can say that the smaller  $\sigma$ -field always wins.

(1). Let  $A \in \mathcal{G}$ . Then we have

$$(22.51) \quad \int_A X(\omega) dP(\omega) = P(A)E[X].$$

Indeed, if  $X = \mathbf{1}_Y$  for some  $Y \in \mathcal{H}$ , this is trivial. In general a simple limiting argument gives us this equality. Therefore, since

$$(22.52) \quad P(A)E[X] = \int_A E[X] dP(\omega),$$

we obtain

$$(22.53) \quad \int_A X(\omega) dP(\omega) = \int_A E[X] dP(\omega).$$

From this we conclude  $E[X : \mathcal{G}] = E[X]$ .  $\square$

(2). We have to show

$$\int_A (E[X : \mathcal{G}] + E[Y : \mathcal{G}])(\omega) dP(\omega) = \int_A (X(\omega) + Y(\omega)) dP(\omega)$$



for all  $A \in \mathcal{G}$ . However, by linearity we have

$$\begin{aligned} \int_A (E[X : \mathcal{G}] + E[Y : \mathcal{G}])(\omega) dP(\omega) &= \int_A E[X : \mathcal{G}](\omega) dP(\omega) + \int_A E[Y : \mathcal{G}](\omega) dP(\omega) \\ &= \int_A X(\omega) dP(\omega) + \int_A Y(\omega) dP(\omega) \\ &= \int_A (X(\omega) + Y(\omega)) dP(\omega). \end{aligned}$$

Thus, the equality was proved. □

(3). A similar argument to (2) works and we omit the proof. □

(4). Let  $A \in \mathcal{G}$ . Then we have to show

$$(22.54) \quad \int_A Z(\omega)E[X : \mathcal{G}](\omega) dP(\omega) = \int_A Z(\omega)X(\omega) dP(\omega).$$

A passage to limit allows us to assume that  $Z$  is a simple function, that is,

$$(22.55) \quad Z = \sum_{j=1}^k a_j \mathbf{1}_{A_j}$$

for some  $a_1, a_2, \dots, a_k \in \mathbb{K}$  and  $A_1, A_2, \dots, A_k \in \mathcal{G}$ . Inserting this expression, we have

$$\begin{aligned} \int_A Z(\omega)X(\omega) dP(\omega) &= \sum_{j=1}^k a_j \int_{A \cap A_j} X(\omega) dP(\omega) \\ &= \sum_{j=1}^k a_j \int_{A \cap A_j} E[X : \mathcal{G}](\omega) dP(\omega) \\ &= \int_A Z(\omega)E[X : \mathcal{G}](\omega) dP(\omega). \end{aligned}$$

Thus, (4) is proved. □

5. We observe

$$\begin{aligned} X \geq 0 &\iff \int_A X(\omega) dP(\omega) \geq 0 \text{ for all } A \in \mathcal{F} \\ &\implies \int_A X(\omega) dP(\omega) \geq 0 \text{ for all } A \in \mathcal{G} \\ &\iff \int_A E[X : \mathcal{G}](\omega) dP(\omega) \geq 0 \text{ for all } A \in \mathcal{G} \\ &\iff E[X : \mathcal{G}] \geq 0. \end{aligned}$$

Thus, 5 is proved. □

6. Let  $A \in \mathcal{G}_1$ . Then we have to show

$$(22.56) \quad \int_A E[E[X : \mathcal{G}_2] : \mathcal{G}_1](\omega) dP(\omega) = \int_A E[E[X : \mathcal{G}_1] : \mathcal{G}_2](\omega) dP(\omega) = \int_A X(\omega) dP(\omega).$$

First, we have

$$(22.57) \quad \int_A E[E[X : \mathcal{G}_2] : \mathcal{G}_1](\omega) dP(\omega) = \int_A E[X : \mathcal{G}_2](\omega) dP(\omega)$$

from the definition of  $E[E[X : \mathcal{G}_2] : \mathcal{G}_1]$ . Since  $A \in \mathcal{G}_2$ , we have

$$(22.58) \quad \int_A E[X : \mathcal{G}_2](\omega) dP(\omega) = \int_A X(\omega) dP(\omega).$$

Similarly from the definition of  $E[E[X : \mathcal{G}_1] : \mathcal{G}_2]$  and the fact that  $A \in \mathcal{G}_2$ , we have

$$(22.59) \quad \int_A E[E[X : \mathcal{G}_1] : \mathcal{G}_2](\omega) dP(\omega) = \int_A E[X : \mathcal{G}_1](\omega) dP(\omega).$$

Therefore, we also obtain

$$(22.60) \quad \int_A E[E[X : \mathcal{G}_1] : \mathcal{G}_2](\omega) dP(\omega) = \int_A X(\omega) dP(\omega).$$

This is the desired result.  $\square$

**Theorem 22.36** (Integration theorems for conditional expectation). *Suppose that  $(\Omega, \mathcal{F}, P)$  is a probability space,  $X_1, X_2, \dots, X, Y \in L^1(\Omega, \mathcal{F}, P)$ .*

(1) *Suppose that  $0 \leq X_1 \leq X_2 \leq \dots \rightarrow X$ . Then*

$$(22.61) \quad \lim_{n \rightarrow \infty} E[X_n : \mathcal{G}] = E[X : \mathcal{G}] \text{ a.s.}$$

(2) *Suppose that  $|X| \leq Y$  a.s. and  $\lim_{n \rightarrow \infty} X_n = Y$ . Then*

$$(22.62) \quad \lim_{n \rightarrow \infty} E[X_n : \mathcal{G}] = E[X : \mathcal{G}] \text{ a.s.}$$

*Exercise 163.* Prove Theorem 22.36.

This proposition is a special case of the separation theorem of convex sets, which we deal in Part 13. Here we give a proof without the choice of axiom.

**Theorem 22.37.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is convex, then we have*

$$(22.63) \quad f(x) = \sup\{ax + b : a, b \in \mathbb{R}, f(t) \geq at + b \text{ for all } t \in \mathbb{R}\}.$$

*Proof.* The right-hand side is trivially less than or equal to the left-hand side. To prove the converse inequality, let  $x_0 \in \mathbb{R}$  be fixed. Choose  $a \in \mathbb{R}$  so that

$$(22.64) \quad \limsup_{\varepsilon \downarrow 0} \frac{f(x_0) - f(x_0 - \varepsilon)}{\varepsilon} \leq a \leq \liminf_{\varepsilon \downarrow 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon}.$$

Assume that  $a(x' - x_0) + f(x_0) > f(x')$  for some  $x' \in \mathbb{R}$ . By symmetry we may suppose  $x' > x_0$ . Then we have

$$(22.65) \quad \liminf_{\varepsilon \downarrow 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} > \frac{f(x') - f(x_0)}{x' - x_0}.$$

This is a contradiction. Therefore

$$(22.66) \quad a(x' - x_0) + f(x_0) \leq f(x')$$

for all  $x' \in \mathbb{R}$ .  $\square$

**Theorem 22.38** (Jensen's inequality). *Suppose that  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is convex. Then*

$$(22.67) \quad \varphi(|E[X : \mathcal{G}]|) \leq E[\varphi(|X|) : \mathcal{G}].$$

*In particular, if  $1 \leq p < \infty$ , then*

$$(22.68) \quad |E[X : \mathcal{G}]|^p \leq E[|X|^p : \mathcal{G}].$$

*and, if  $1 \leq p \leq \infty$  then we have*

$$(22.69) \quad \|E[X : \mathcal{G}]\|_p \leq \|X\|_p$$

*for all  $X \in L^p(\Omega, \mathcal{F}, P)$*

*Proof.* We have only to show (22.67), the rest being immediate from this. Utilize Theorem 22.37:

$$(22.70) \quad \varphi(t) = \sup_{a,b \in \mathbb{R}} \{at + b : ax + b \leq \varphi(x)\}.$$

Let  $t = E[X : \mathcal{G}]$ . Then we have

$$\begin{aligned} \varphi(E[X : \mathcal{G}]) &= \sup_{a,b \in \mathbb{R}} \{a E[X : \mathcal{G}] + b : ax + b \leq \varphi(x)\} \\ &= \sup_{a,b \in \mathbb{R}} \{E[aX + b : \mathcal{G}] : ax + b \leq \varphi(x)\} \\ &\leq E[\varphi(X) : \mathcal{G}]. \end{aligned}$$

This is the desired result. □

**Theorem 22.39.** *Suppose that  $1 < p < \infty$ . Assume  $X \in L^p(\Omega, \mathcal{F}, P)$  and  $Y \in L^{p'}(\Omega, \mathcal{F}, P)$ . Then*

$$(22.71) \quad E[|XY| : \mathcal{G}] \leq E[|X|^p : \mathcal{G}]^{\frac{1}{p}} E[|Y|^{p'} : \mathcal{G}]^{\frac{1}{p'}} \text{ a.s.}$$

*Exercise 164.* Prove Theorem 22.39.

*Exercise 165.* Let  $\mathcal{G}$  be a  $\sigma$ -field smaller than  $\mathcal{F}$ . Then show that

$$(22.72) \quad E[\cdot : \mathcal{G}]L^2(\mathcal{F}, P) \rightarrow L^2(\mathcal{F}, P)$$

is a projection in  $L^2(\mathcal{F}, P)$ .

### 23. MARTINGALES WITH DISCRETE TIME

#### 23.1. Martingales.

**Definition 23.1.** A stochastic process  $X = (X_n)_{n \in \mathbb{N}_0}$  is said to be a martingale, if it satisfies the following conditions.

- (1)  $X_n \in L^1(\Omega, \mathcal{F}, P)$  for all  $n \in \mathbb{N}_0$ .
- (2)  $0 \leq m < n$  implies

$$(23.1) \quad E[X_n : \mathcal{F}_m] = X_m \text{ a.s.}$$

**Example 23.2.** Let  $\xi \in L^1(\mathcal{F})$ . Set  $M_n := E[\xi : \mathcal{F}_n]$  for  $n \in \mathbb{N}_0$ . Then

$$(23.2) \quad M := (M_n)_{n \in \mathbb{N}}$$

is a martingale.

*Proof.* Since  $\xi \in L^1(\mathcal{F})$ , each  $M_n$  is integrable. Let  $m > n$ . Recall that the smaller  $\sigma$ -field  $\mathcal{F}_n$  always wins the larger  $\sigma$ -field  $\mathcal{F}_m$ . Thus,

$$(23.3) \quad E[M_m : \mathcal{F}_n] = E[E[\xi : \mathcal{F}_m] : \mathcal{F}_n] = E[\xi : \mathcal{F}_n] = M_n \text{ a.s.},$$

which shows  $M$  is a martingale. □

**Example 23.3.** Let  $\xi_n \in L^1(\mathcal{F}, P), n = 1, 2, \dots$  be a sequence of random variables such that  $\xi_n \perp \mathcal{F}_{n-1}$  and  $E[\xi_n] = 0$  for all  $n \in \mathbb{N}$ . Then

$$(23.4) \quad M := (M_n)_{n \in \mathbb{N}}, M_n := \sum_{j=1}^n \xi_j$$

is a martingale.

**Definition 23.4.** Let  $X = (X_n)_{n \in \mathbb{N}}$  be a stochastic process.

- (1)  $X$  is said to be a submartingale, if
- (a)  $X_n \in L^1(\Omega, \mathcal{F}, P)$  for all  $n \in \mathbb{N}$ .
  - (b)  $m < n$  implies

$$(23.5) \quad E[X_n : \mathcal{F}_m] \geq X_m \text{ a.s.}$$

- (2)  $X$  is said to be a super-martingale, if  $-X = (-X_n)_{n \in \mathbb{N}_0}$  is a submartingale.

Starting from a submartingale, we can easily generate submartingales.

**Theorem 23.5.** *Suppose that  $\varphi$  is an  $N$ -function and  $M := \{M_n\}_{n \in \mathbb{N}_0}$  is a submartingale. Assume that  $\varphi(M_n) \in L^1(P)$  for each  $n$ . Then  $X = (\varphi(M_n))_{n \in \mathbb{N}_0}$  is a submartingale.*

*Proof.* Let  $m > n$ . Jensen's inequality gives us that

$$(23.6) \quad [\varphi(M_m) : \mathcal{F}_n] \leq \varphi([M_m : \mathcal{F}_n]) \leq \varphi(M_n) \text{ a.s.},$$

which shows  $X = (\varphi(M_n))_{n \in \mathbb{N}_0}$  is a submartingale.  $\square$

**Difference of martingales.** Given a process  $X = (X_n)_{n \in \mathbb{N}_0}$ , the difference process is the difference as is given just below. In this paragraph we characterize the (discrete) martingale in terms of the difference.

**Definition 23.6.** Given a stochastic process  $X = (X_n)_{n \in \mathbb{N}_0}$ , define

$$(23.7) \quad d_0 = X_0, \quad d_n = X_n - X_{n-1}, \quad n \geq 1.$$

$d = (d_n)_{n \in \mathbb{N}_0}$  is called difference of  $X$ .

Martingale is characterized in terms of their difference.

**Theorem 23.7.** *Suppose that  $X = (X_n)_{n \in \mathbb{N}_0}$  is a stochastic process and denote by  $d = (d_n)_{n \in \mathbb{N}_0}$  its difference. Then  $X$  is martingale if and only if  $d_n \in L^1(P)$  and*

$$(23.8) \quad E[d_{n+1} : \mathcal{F}_n] = 0$$

*a.s. for all  $n \in \mathbb{N}_0$ .*

*Proof.* The proof being simple, we leave it for readers as Exercise 166.  $\square$

*Exercise 166.* Prove Theorem 23.7.

**Integrability of martingales.** In this paragraph we give some definitions concerning to the integrability of martingales.

**Definition 23.8.** Let  $1 \leq p < \infty$  and  $X = \{X_n\}_{n \in \mathbb{N}_0}$  a martingale. The process  $X$  is said to be an  $L^p$ -martingale, if  $X_n \in L^p(\Omega, \mathcal{F}, P)$  for each  $n \in \mathbb{N}_0$ , and an  $L^p$ -bounded martingale, if  $\sup_{n \in \mathbb{N}_0} E[|X_n|^p] < \infty$ .

The next theorem exhibits the way of generating submartingales.

**Example 23.9.** Let  $1 \leq p < \infty$  and  $X$  be an  $L^p$ -martingale. Then Theorem 23.5 with  $\varphi(t) = |t|^p$  gives us that  $(|X_n|^p)_{n \in \mathbb{N}_0}$  is a submartingale.

Martingale transform.

Martingale transform is a prototype of the stochastic integral due to Seizo Ito, which provides us a method of generating new martingale starting from a given martingale.

**Definition 23.10** (Predictable process). Let  $H = (H_n)_{n \in \mathbb{N}_0}$  be a stochastic process. Then  $H$  is said to be predictable, if  $H_0 \in \mathcal{F}_0$  and  $H_n \in \mathcal{F}_{n-1}$ .

**Definition 23.11.** Let  $M$  be a stochastic process and  $H$  a predictable process. Then we denote

$$(23.9) \quad (H \cdot M)_n := \sum_{j=1}^n H_j(M_j - M_{j-1}), \quad (H \cdot M)_0 := 0.$$

In the first half of the last century it had been investigated in connection with gambling, exclusively when  $H_j$  takes only 0, 1. In [104] Burkholder generalized to the form which we shall present below.

**Theorem 23.12** (Martingale transform). *Suppose that  $H$  is a positive bounded predictable process. and  $M$  is a submartingale. Then  $H \cdot M$  is a submartingale.*

*Proof.* The fact that  $(H \cdot M)_n \in L^1(P)$  is easy to prove since the sum is made up of a finite number of elements. We have to show that

$$(23.10) \quad E[(H \cdot M)_n : \mathcal{F}_{n-1}] \geq (H \cdot M)_{n-1}$$

for all  $n \in \mathbb{N}$ . Suppose that  $n = 1$ . Then

$$(23.11) \quad E[(H \cdot M)_1 : \mathcal{F}_0] = E[H_1(M_1 - M_0) : \mathcal{F}_0] = H_1(E[M_1 : \mathcal{F}_0] - M_0) \geq 0.$$

Suppose that  $n \geq 2$ . In this case by induction to prove (23.10) we have only to show that

$$(23.12) \quad E[H_n(M_n - M_{n-1}) : \mathcal{F}_0] \geq 0.$$

However, this is essentially the same as the case when  $n = 1$ . □

**Corollary 23.13.** *Suppose that  $H$  is a martingale and  $M$  is a martingale. Then  $H \cdot M$  is a martingale.*

### 23.2. Decomposition of martingales.

Here we make a view of two decompositions of martingales. The first decomposition, which is called the Doob decomposition, clarifies the structure of submartingales. The second one corresponds to the decomposition of  $\mathbb{R}$ -valued measurable functions into the difference of two positive measurable functions, as we have been doing in measure theory.

**Theorem 23.14** (Doob decomposition). *Any submartingale  $Y = (Y_n)_{n \in \mathbb{N}_0}$  admits uniquely the following decomposition :*

$$(23.13) \quad Y_n = X_n + A_n \quad n \in \mathbb{N}_0,$$

where  $X = (X_n)_{n \in \mathbb{N}_0}$  is a martingale and  $A = (A_n)_{n \in \mathbb{N}_0}$  is a predictable increasing process with  $A_0 = 0$ .

*Proof.* It might be helpful to begin with the uniqueness, which suggests how to decompose  $Y$ .

To do this, let us assume that  $Y$  admits a decomposition described above. Then we have

$$(23.14) \quad Y_{n+1} = X_{n+1} + A_{n+1}, \quad A_{n+1} \in \mathcal{F}_n, \quad E[X_{n+1} : \mathcal{F}_n] = X_n,$$

In view of this, we have

$$(23.15) \quad A_{n+1} = E[A_{n+1} : \mathcal{F}_n] = E[Y_{n+1} : \mathcal{F}_n] - E[X_{n+1} : \mathcal{F}_n] = E[Y_{n+1} : \mathcal{F}_n] - X_n.$$

Therefore once we are given  $Y_{n+1}$  and  $X_n$ ,  $A_{n+1}$  and  $X_{n+1}$  are determined uniquely by the following recurrence formula:

$$\begin{aligned} A_{n+1} &= E[Y_{n+1} : \mathcal{F}_n] - X_n \\ X_{n+1} &= X_n + Y_{n+1} - E[Y_{n+1} : \mathcal{F}_n]. \end{aligned}$$

Of course, the initial data  $A_0$  and  $X_0$  are given by

$$(23.16) \quad A_0 = 0, X_0 = Y_0.$$

Therefore, the decomposition is unique.

To deal with the uniqueness, we have only to reverse the argument above. Set  $A_0 = 0$  and  $X_0 = 0$ . Set, for  $n \geq 1$ ,

$$(23.17) \quad X_n = Y_0 + \sum_{j=1}^{n-1} (Y_j - E[Y_j : \mathcal{F}_{j-1}]), \quad A_n = \sum_{j=1}^n (E[Y_j : \mathcal{F}_{j-1}] - Y_{j-1}).$$

Then, since  $Y$  is a submartingale, each summand defining  $A_n$  is positive. Therefore,  $A$  is increasing. Since  $E[Y_n : \mathcal{F}_{n-1}] \in \mathcal{F}_{n-1}$ ,  $A$  is predictable.  $E[Y_j - E[Y_j : \mathcal{F}_{j-1}] : \mathcal{F}_{j-1}] = 0$  implies  $X$  is a martingale. Therefore, the desired decomposition was obtained.  $\square$

This is a concrete example of the above theorem.

**Example 23.15.** Let  $X = (X_n)_{n \in \mathbb{N}}$  be an  $L^2(P)$ -martingale. Then by Theorem 23.14  $X_n^2$  can be decomposed into

$$(23.18) \quad X_n^2 = M_n + A_n.$$

Here,  $X = (X_n)_{n \in \mathbb{N}_0}$  is a martingale and  $A = (A_n)_{n \in \mathbb{N}_0}$  is an increasing process with  $A_0 = 0$ . Following the construction of the above proof,  $M_n$  and  $A_n$  are given by

$$\begin{aligned} M_n &= X_0^2 + \sum_{j=1}^{n-1} (X_j^2 - E[X_j^2 : \mathcal{F}_{j-1}]) \\ A_n &= \sum_{j=1}^n (E[X_j^2 : \mathcal{F}_{j-1}] - X_{j-1}^2). \end{aligned}$$

$A_n$  can be written in terms of the martingale difference  $d$  of  $X$ . In fact,

$$\begin{aligned} E[X_j^2 : \mathcal{F}_{j-1}] - X_{j-1}^2 &= E[(X_j - X_{j-1})^2 + 2X_{j-1}(X_j - X_{j-1}) + X_{j-1}^2 : \mathcal{F}_{j-1}] - X_{j-1}^2 \\ &= E[(X_j - X_{j-1})^2 + 2X_{j-1}(X_j - X_{j-1}) : \mathcal{F}_{j-1}] \\ &= E[(X_j - X_{j-1})^2 : \mathcal{F}_{j-1}] + 2X_{j-1}E[X_j - X_{j-1} : \mathcal{F}_{j-1}] \\ &= E[d_j^2 : \mathcal{F}_{j-1}]. \end{aligned}$$

Therefore we obtain

$$(23.19) \quad A_n = \sum_{j=1}^n E[d_j^2 : \mathcal{F}_{j-1}].$$

**Definition 23.16.** A martingale  $M$  is said to be  $L^1(P)$ -bounded, if  $\sup_{n \in \mathbb{N}} \|M_n\|_1 < \infty$ .

**Theorem 23.17** (Krickeberg's decomposition theorem). *Any  $L^1(P)$ -bounded martingale  $X$  admits the following decomposition.*

$$(23.20) \quad X = Y - Z,$$

where  $Y$  and  $Z$  are non-negative martingales.

*Proof.* Let  $j, n \in \mathbb{N}_0$ . Set

$$(23.21) \quad Y_n^{(j)} := E[(X_{n+j})_+ : \mathcal{F}_n], \quad Z_n^{(j)} := E[(X_{n+j})_- : \mathcal{F}_n].$$

Suppose that  $k \geq j$ . Then by the convexity of  $x \mapsto x_+$  we have

$$(23.22) \quad (X_{n+j})_+ = (E[X_{m+k} : \mathcal{F}_{n+j}])_+ \leq E[(X_{m+k})_+ : \mathcal{F}_{n+j}].$$

Taking the conditional expectation with respect to  $\mathcal{F}_n$  we obtain

$$(23.23) \quad Y_n^{(j)} = E[(X_{n+j})_+ : \mathcal{F}_n] \leq E[(X_{n+k})_+ : \mathcal{F}_n] = Y_n^{(k)}.$$

Taking into account the  $L^1(P)$ -boundedness of  $X$  and the fact that  $\{Y_n\}_{n \in \mathbb{N}}$  is increasing, we see that  $\lim_{j \rightarrow \infty} Y_n^{(j)} = Y_n$  exists in  $L^1(P)$  and almost surely. It is easy to verify that  $(Y_n)_{n \in \mathbb{N}}$  is a non-negative martingale. Since  $\{X_n\}_{n \in \mathbb{N}}$  is a martingale, we have  $Y_n^{(j)} - Z_n^{(j)} = X_n$ . Therefore  $\lim_{j \rightarrow \infty} Z_n^{(j)} = Z_n$  exists and is positive. Furthermore,  $Z = X - Y$  is martingale because so are  $X$  and  $Y$ . As a result we have the desired conclusion.  $\square$

### 23.3. Stopping time.

Stopping time is a tool for stochastic analysis.

**Definition 23.18** (Filtration). A filtration  $\mathbb{F}$  is an increasing sequence of sub- $\sigma$  fields  $\mathbb{F} = \{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$ , that is,

$$(23.24) \quad \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$$

for all  $n \in \mathbb{N}_0$ , where  $\mathcal{F}$  is a fixed  $\sigma$ -field.

In the sequel it will be tacitly understood that a filtration  $\mathbb{F}$  is given.

**Definition 23.19** (Stopping time). A stopping time  $T : \Omega \rightarrow \mathbb{N}_0$  is said to be a random variable, if

$$(23.25) \quad \{T \leq n\} \in \mathcal{F}_n$$

for all  $n \in \mathbb{N}_0$ .  $\mathcal{T}$  denotes the set of all stopping times.

**Proposition 23.20.** *Let  $T : \Omega \rightarrow \mathbb{N}_0$  be a random variable. Then the following are equivalent.*

- (1)  $T$  is a stopping time.
- (2)  $\{T = n\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}_0$ .

*Exercise 167.* Prove Proposition 23.20.

### 23.4. Elementary properties.

**Theorem 23.21.**  $\mathcal{T}$  has the following properties.

- (1) Let  $S, T \in \mathcal{T}$ . Then  $S \wedge T, S \vee T \in \mathcal{T}$ .
- (2) Let  $\{S_n\}_{n \in \mathbb{N}} \subset \mathcal{T}$  be a sequence of stopping times. Then

$$(23.26) \quad \bigwedge_{n \in \mathbb{N}_0} S_n := \inf_{n \in \mathbb{N}_0} S_n, \quad \bigvee_{n \in \mathbb{N}_0} S_n := \sup_{n \in \mathbb{N}_0} S_n \in \mathcal{T}.$$

- (3)  $n \in \mathbb{N}_0$  can be embedded into  $\mathcal{T}$  in the following sense :

$$(23.27) \quad n \in \mathbb{N}_0 \mapsto \tilde{n} := [\omega \in \Omega \mapsto n \in \mathbb{N}_0] \in \mathcal{T}.$$

Here and below in view of this  $n \in \mathbb{N}_0$  will be identified with a constant function  $\tilde{n}$ .

(1). Let  $n \in \mathbb{N}_0$ . Then

$$(23.28) \quad \{S \wedge T \leq n\} = \{S \leq n\} \cup \{T \leq n\} \in \mathcal{F}_n, \{S \vee T \leq n\} = \{S \leq n\} \cap \{T \leq n\} \in \mathcal{F}_n.$$

Therefore it follows that  $S \wedge T, S \vee T \in \mathcal{T}$ .  $\square$

(2). The proof is the same as (1), although it is a bit more complicated. Let  $k \in \mathbb{N}_0$  as before. Then, taking into account that all the stopping times take their value in  $\mathbb{N}$ , we have

$$(23.29) \quad \left\{ \bigvee_{n \in \mathbb{N}_0} S_n \leq k \right\} = \bigcup_{n \in \mathbb{N}_0} \{S_n \leq k\} \in \mathcal{F}_k, \left\{ \bigwedge_{n \in \mathbb{N}_0} S_n \leq k \right\} = \bigcap_{n \in \mathbb{N}_0} \{S_n \leq k\} \in \mathcal{F}_k.$$

Therefore, it follows that  $\bigwedge_{n \in \mathbb{N}_0} S_n, \bigvee_{n \in \mathbb{N}_0} S_n \in \mathcal{T}$ .  $\square$

(3). Let  $k \in \mathbb{N}$ . Then  $\{\tilde{n} \leq k\} = \Omega$ , provided  $n > k$  and  $\{\tilde{n} \leq k\} = \emptyset$ , if  $n \leq k$ . Therefore  $\tilde{n} \in \mathcal{T}$ .  $\square$

**Definition 23.22.** Let  $T \in \mathcal{T}$ . Then define

$$(23.30) \quad \mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n \in \mathbb{N}_0\}.$$

$\mathcal{F}_T$ , as is expected, is a  $\sigma$ -field.

**Proposition 23.23.**  $\mathcal{F}_T$  is a  $\sigma$ -field for every  $T \in \mathcal{T}$ . Furthermore,  $\mathcal{F}_{\tilde{n}} = \mathcal{F}_n$ , where  $\tilde{n}$  is defined by (23.27).

*Exercise 168.* Prove Proposition 23.23.

**Theorem 23.24.** Let  $\tau, \mu \in \mathcal{T}$ . Then the following are true.

- (1) Let  $A \in \mathcal{F}_\tau$ . Then  $A \cap \{\tau \leq \mu\} \in \mathcal{F}_\mu$ .
- (2)  $\tau \leq \mu$  implies  $\mathcal{F}_\tau \subset \mathcal{F}_\mu$ .
- (3)  $\{\tau = \mu\}, \{\tau \geq \mu\} \in \mathcal{F}_\tau \cap \mathcal{F}_\mu$ .
- (4)  $\mathcal{F}_{\min(\tau, \mu)} = \mathcal{F}_\tau \cap \mathcal{F}_\mu$ .

(1). We have to show

$$(23.31) \quad (A \cap \{\tau \leq \mu\}) \cap \{\mu \leq n\} \in \mathcal{F}_n.$$

Note that this can be expressed as follows:

$$(23.32) \quad (A \cap \{\tau \leq \mu\}) \cap \{\mu \leq n\} = \bigcup_{\substack{k, l \in \mathbb{N}_0 \\ 0 \leq k \leq l \leq n}} A \cap \{\tau = k\} \cap \{\mu = l\}.$$

Since  $A \cap \{\tau = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$  and  $\{\mu = l\} \in \mathcal{F}_l \subset \mathcal{F}_n$ , (23.31) is established.  $\square$

(2). Since  $\tau \leq \mu$ , it follows that  $A \cap \{\tau \leq \mu\} = A$  for all  $A \in \mathcal{F}_\tau$ . Therefore, we see  $\mathcal{F}_\tau \subset \mathcal{F}_\mu$  from (1).  $\square$

(3). Let  $n \in \mathbb{N}$ . Then  $\{\tau = \mu\} \cap \{\mu \leq n\} = \bigcup_{0 \leq k \leq n} \{\tau = \mu = k\} \in \mathcal{F}_n$ . Therefore,  $\{\tau = \mu\} \in \mathcal{F}_\mu$ .

By symmetry, we have  $\{\tau = \mu\} \in \mathcal{F}_\tau$ . Thus,  $\{\tau = \mu\} \in \mathcal{F}_\mu \cap \mathcal{F}_\tau$  is proved.

Since  $\Omega \in \mathcal{F}_\mu$ , we have  $\{\tau \geq \mu\} \in \mathcal{F}_\tau$  by (1), from which we deduce  $\{\tau \leq \mu\} \in \mathcal{F}_\mu$  and, taking its complement, we further obtain  $\{\tau > \mu\} \in \mathcal{F}_\mu$ . Since  $\{\tau \geq \mu\} = \{\tau > \mu\} \cup \{\tau = \mu\} \in \mathcal{F}_\mu$  by (1), it follows that  $\{\tau \geq \mu\} \in \mathcal{F}_\tau \cap \mathcal{F}_\mu$ .  $\square$



(4). By (2) we have  $\mathcal{F}_{\tau \wedge \mu} \subset \mathcal{F}_\tau \cap \mathcal{F}_\mu$ . Let  $A \in \mathcal{F}_\tau \cap \mathcal{F}_\mu$ . Then we have

$$(23.33) \quad A \cap \{\tau \wedge \mu \leq n\} = (A \cap \{\tau \leq n\}) \cup (A \cap \{\mu \leq n\}) \in \mathcal{F}_n.$$

Therefore, the reverse inclusion  $\mathcal{F}_{\tau \wedge \mu} \supset \mathcal{F}_\tau \cap \mathcal{F}_\mu$  is also established. □

## 24. PROPERTIES OF MARTINGALES

### 24.1. The optional sampling theorem.

*Notation* . For a stochastic process  $X = (X_n)_{n \in \mathbb{N}_0}$  and  $\tau \in \mathcal{T}$  we denote  $X^\tau := (X_{\min(\tau, n)})_{n \in \mathbb{N}}$ , where we have defined  $X_\mu(\omega) := X_{\mu(\omega)}(\omega)$  for  $\mu \in \mathcal{T}$ .

**Theorem 24.1** (Doob's optimal sampling theorem). *Let  $\tau \in \mathcal{F}$  and  $X$  be a martingale. Then  $X^\tau$  is a  $\mathcal{F}$ -martingale. Furthermore,  $X^\tau$  is a  $(\mathcal{F}_{\tau \vee n})_{n \in \mathbb{N}}$ -martingale.*

*Proof.*  $X^\tau$  is a  $\mathcal{F}$ -martingale. To verify this, let  $m \geq n$  be two integers and we have to show

$$(24.1) \quad E[X_m^\tau : \mathcal{F}_n] = X_n^\tau$$

for almost surely. Let  $A \in \mathcal{F}_n$ . Then

$$\begin{aligned} E[X_m^\tau : A] &= \sum_{j=1}^n E[X_j : A \cap \{\tau = j\}] + \sum_{j=n+1}^m E[X_j : A \cap \{\tau = j\}] \\ &\quad + \sum_{j=m+1}^{\infty} E[X_m : A \cap \{\tau = j\}]. \end{aligned}$$

Note that if  $\tau = j \leq n$ , then  $X_j = X_j^\tau$ . Therefore the first term can be written as

$$(24.2) \quad \sum_{j=1}^n E[X_j : A \cap \{\tau = j\}] = \sum_{j=1}^n E[X_n^\tau : A \cap \{\tau = j\}].$$

The second term is equal to

$$\begin{aligned} \sum_{j=n+1}^m E[X_j : A \cap \{\tau = j\}] &= \sum_{j=n+1}^m E[E[X_m : \mathcal{F}_j] : A \cap \{\tau = j\}] \\ &= \sum_{j=n+1}^m E[X_m : A \cap \{\tau = j\}], \end{aligned}$$

where we used  $A \cap \{\tau = j\} \in \mathcal{F}_m$ . As a result we are led to

$$(24.3) \quad \sum_{j=n+1}^m E[X_j : A \cap \{\tau = j\}] + \sum_{j=m+1}^{\infty} E[X_m : A \cap \{\tau = j\}] = \sum_{j=n+1}^{\infty} E[X_m : A \cap \{\tau = j\}].$$

Since  $B := \bigcup_{j=n+1}^{\infty} (A \cap \{\tau = j\}) = \Omega \setminus \bigcup_{j=1}^n (A \cap \{\tau = j\}) \in \mathcal{F}_n$ , we obtain

$$\sum_{j=n+1}^{\infty} E[X_m : A \cap \{\tau = j\}] = E[E[X_m \mathbf{1}_B : \mathcal{F}_n]] = E[X_n \mathbf{1}_B : \mathcal{F}_n].$$

Inserting this estimate we obtain

$$(24.4) \quad \sum_{j=n+1}^{\infty} E[X_m : A \cap \{\tau = j\}] = \sum_{j=n+1}^{\infty} E[X_n : A \cap \{\tau = j\}].$$

This is the desired result.

$X^\tau$  is a  $\mathcal{F}_{\tau \wedge n}$ -martingale. To see this, we have to show

$$(24.5) \quad E[X_m^\tau : \mathcal{F}_{\tau \wedge n}] = X_n^\tau.$$

To prove this, we proceed as before. First we take  $A \in \mathcal{F}_{\tau \wedge n}$ . We shall calculate

$$E[X_m^\tau : A] = \sum_{j=1}^n E[X_j : A \cap \{\tau = j\}] + \sum_{j=n+1}^{\infty} E[X_m : A \cap \{\tau = j\}].$$

The treatment of the first term is complete.

$$(24.6) \quad \sum_{j=1}^n E[X_j : A \cap \{\tau = j\}] = \sum_{j=1}^n E[X_n^\tau : A \cap \{\tau = j\}].$$

Note that

$$(24.7) \quad A \cap \{\tau = j\} \in \mathcal{F}_{\tau \wedge n \wedge j} \subset \mathcal{F}_n.$$

Taking into account that  $X_n^\tau \in \mathcal{F}_n$ , we obtain

$$(24.8) \quad E[X_j : A \cap \{\tau = j\}] = E[X_n^\tau : A \cap \{\tau = j\}].$$

The rest is the same and we omit some details.  $\square$

## 24.2. Doob's maximal inequality.

Our elementary inequality is as follows: Note that this is similar to the weak-(1, 1) inequality of the Hardy-Littlewood maximal operators.

**Theorem 24.2** (Doob's maximal inequality). *Suppose that  $M$  is a submartingale. Then for all  $\lambda > 0$*

$$(24.9) \quad P\left(\max_{1 \leq k \leq n} M_k > \lambda\right) \leq \frac{1}{\lambda} E\left[M_n : \max_{1 \leq k \leq n} M_k > \lambda\right].$$

*Proof.* Let  $\tau = \min\{k \in \mathbb{N} : |M_k| > \lambda\}$ . Then  $\tau \in \mathcal{T}$ .

$$P\left(\max_{1 \leq k \leq n} M_k > \lambda\right) = \sum_{k=1}^n P(\tau = k) \leq \frac{1}{\lambda} \sum_{k=1}^n E[M_k : \tau = k].$$

Since  $M$  is a submartingale, we have  $M_k \geq E[M_n : \mathcal{F}_k]$ . Thus

$$(24.10) \quad E[M_k : \tau = k] \leq E[E[M_n : \mathcal{F}_k] \mathbf{1}_{\{\tau=k\}}] = E[E[M_n \mathbf{1}_{\{\tau=k\}} : \mathcal{F}_k]] = E[M_n \mathbf{1}_{\{\tau=k\}}].$$

Summing up the above estimate for  $k = 1, 2, \dots, n$ , we obtain

$$(24.11) \quad P\left(\max_{1 \leq k \leq n} M_k > \lambda\right) \leq \frac{1}{\lambda} \sum_{k=1}^n E[M_n : \tau = k] = \frac{1}{\lambda} E\left[M_n : \max_{1 \leq j \leq n} M_j > \lambda\right].$$

This is the desired result.  $\square$

**Corollary 24.3** (Doob). *Suppose that  $M = (M_n)_{n \in \mathbb{N}_0}$  is a submartingale. Then*

$$(1) \quad P\left(\max_{1 \leq k \leq n} |M_k| > \lambda\right) \leq \frac{1}{\lambda} E[|M_n|].$$

(2) *Let  $1 < p < \infty$ . Then*

$$(24.12) \quad E\left[\max_{1 \leq k \leq n} |M_k|^p\right] \leq \left(\frac{p}{p-1}\right)^p E[|M_n|^p].$$

*In particular, letting  $p = 2$ , we have*

$$(24.13) \quad E\left[\max_{1 \leq k \leq n} |M_k|^2\right] \leq 4E[|M_n|^2].$$

*Exercise 169.* Prove Corollary 24.3 by means of interpolation. Reexamine the proof of the interpolation to obtain the size of constants.

**Remark 24.4.** Compare Corollary 24.3 with the boundedness of the maximal operator.

*Exercise 170.* Set  $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), dx|_{[0, 1]})$ . Let  $M_n = 2^n \chi_{[0, 2^{-n}]}$  for  $n \in \mathbb{N}$ . Then show that by defining a suitable filtration 24.3 is not available for  $p = 1$ .

**24.3. Convergence theorems of martingales.**

We begin with our criterion of almost sure convergence. The following theorem gives us a general principle for the almost sure convergence.

**Theorem 24.5.** *Suppose that  $X = \{X_n\}_{n=1}^\infty$  is a stochastic process such that*

$$(24.14) \quad \lim_{n \rightarrow \infty} P \left( \sup_{n \leq k, l} |X_k - X_l| > \varepsilon \right) = 0$$

for all  $\varepsilon > 0$ . Then  $X = \{X_n\}_{n \in \mathbb{N}}$  converges almost surely.

*Proof.* We have to show that

$$(24.15) \quad P \left( \limsup_{m, k \rightarrow \infty} |X_k - X_m| \neq 0 \right) = 0.$$

To this end it suffices to show that

$$(24.16) \quad P \left( \limsup_{m, k \rightarrow \infty} |X_k - X_m| > \varepsilon \right) = 0$$

for all  $\varepsilon > 0$ . However, this is just a matter of taking limit of the assumption, we omit the details. □

**Lemma 24.6.** *Let  $m \leq n$  and  $M = \{M_n\}_{n=1}^\infty$  be a submartingale. Then we have*

$$E[M_n^2], E[M_n M_m] \geq E[M_m^2].$$

*Proof.* The first inequality is a direct consequence of Jensen's inequality and the fact that  $M$  is a non-negative submartingale.

$$(24.17) \quad M_m^2 \leq E[M_n : \mathcal{F}_m]^2 \leq E[M_n^2 : \mathcal{F}_m].$$

The second inequality is similar.

$$(24.18) \quad M_m^2 \leq M_m E[M_n : \mathcal{F}_m] = E[M_m M_n : \mathcal{F}_m].$$

It remains to take the expectation in both estimates. □

**Theorem 24.7.** *Suppose that  $M = (M_n)_{n \in \mathbb{N}_0}$  is an  $L^2(P)$ -bounded non-negative submartingale. Then*

$$(24.19) \quad M_\infty := \lim_{n \rightarrow \infty} M_n$$

exists in  $L^2(P)$ .

*Proof.* From the lemma above and the assumption that  $X$  is  $L^2(P)$ -bounded, we deduce that  $\lim_{n \rightarrow \infty} E[M_n^2]$  exists and finite. We also have

$$(24.20) \quad E[(X_n - X_m)^2] \leq E[X_{\max(m, n)}^2] - E[X_{\min(m, n)}^2]$$

by virtue of the second inequality. As a consequence  $(X_n)_{n \in \mathbb{N}}$  converges in  $L^2(P)$ . □

**Theorem 24.8.** *Suppose that  $M = (M_n)_{n \in \mathbb{N}}$  is a positive martingale. Then  $\lim_{n \rightarrow \infty} M_n(\omega)$  exists for almost sure  $\omega \in \Omega$ .*

*Proof.* Let  $X_n = \exp(-M_n)$  for  $n \in \mathbb{N}$ . Since  $\exp : \mathbb{R} \rightarrow (0, \infty)$  is a homeomorphism, it suffices to prove

$$(24.21) \quad \lim_{n \rightarrow \infty} P \left( \max_{n \leq k, l} |X_k - X_l| > \varepsilon \right) = 0$$

for all  $\varepsilon > 0$ .

First of all, we note that  $\{X_n\}_{n=1}^\infty$  is a submartingale. Indeed,  $\{-M_n\}_{n=1}^\infty$  is a martingale and hence as a composition of a convex function,  $\{X_n\}_{n=1}^\infty$  is a submartingale. By positivity of  $M$ ,  $X_n$  satisfies  $0 \leq X_n \leq 1$ .

Thus we are in the position of using Theorem 24.7, to have  $X_n, n = 1, 2, \dots$  converges in  $L^2(P)$ . Keeping the  $L^2(P)$ -convergence of  $X$  in mind, we prove (24.21). Let  $X_\infty$  be the  $L^2(P)$ -limit of  $(X_n)_{n \in \mathbb{N}}$ . Then we have

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left( \max_{n \leq k, l} |X_k - X_l| > \varepsilon \right) &= \lim_{n \rightarrow \infty} \left( \lim_{m \rightarrow \infty} P \left( \max_{n \leq k, m} |X_k - X_l| > \varepsilon \right) \right) \\ &\leq \frac{4}{\varepsilon^2} \lim_{n \rightarrow \infty} \left( \lim_{m \rightarrow \infty} E[|X_m - X_n|^2] \right) \\ &\leq \frac{4}{\varepsilon^2} \lim_{n \rightarrow \infty} E[|X_\infty - X_n|^2] \\ &= 0. \end{aligned}$$

Consequently we have proved the theorem.  $\square$

#### 24.4. Applications of convergence theorems.

Recall that

$$(24.22) \quad \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

is divergent while

$$(24.23) \quad \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

is convergent. If we choose the sign  $\pm$  randomly in the above series, what happens? The following example is very interesting in that it gives a definitive answer in terms of mathematics.

**Example 24.9.** Let  $\xi_j, j = 1, 2, \dots$  be a sequence of independent and identically distributed random variables with  $P(\xi_1 = -1) = P(\xi_1 = 1) = \frac{1}{2}$ . Then

$$(24.24) \quad Y(\omega) = \sum_{j=1}^{\infty} \frac{\xi_j(\omega)}{j}$$

converges for almost sure  $\omega \in \Omega$ .

*Proof.* First, we set

$$(24.25) \quad X_n(\omega) = \sum_{j=1}^n \frac{\xi_j(\omega)}{j}.$$

Then  $(X_n)_{n \in \mathbb{N}}$  is a martingale. Since  $E[\xi_1^2] = 1$ ,  $E[\xi_1] = 0$ ,  $\xi_1, \xi_2, \xi_3, \dots, \xi_n$  are i.i.d. and it follows that

$$(24.26) \quad E[X_n^2] = \sum_{j=1}^n \frac{1}{j^2} < \frac{\pi^2}{6}.$$

Therefore,  $X$  is an  $L^2(P)$ -bounded martingale and hence admits its almost sure limit  $\lim_{n \rightarrow \infty} X_n$ .

Consequently  $Y(\omega)$  is convergent for almost sure  $\omega \in \Omega$ . □

Example 24.9 can be explained from the next theorem as well.

**Theorem 24.10** (Kolmogorov's 0-1 principle). *Suppose that  $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n, \dots$  are independent  $\sigma$ -field. Define the tail  $\sigma$ -field by  $\mathcal{T}^* := \bigcap_{k \in \mathbb{N}} \left( \bigcup_{n \geq k} \mathcal{G}_n \right)$ . Then any event in  $\mathcal{T}^*$  happens almost surely or never happens almost surely.*

*Proof.* Let  $\mathcal{F}_n = \bigvee_{j=1}^n \mathcal{G}_j$  for  $n = 1, 2, \dots$  and  $\mathcal{F} = \bigvee_{j \in \mathbb{N}} \mathcal{G}_j$ . Let  $A \in \mathcal{T}^*$ . Define  $X_n := E[\mathbf{1}_A : \mathcal{F}_n]$ .

Then  $(X_n)_{n \in \mathbb{N}}$  is a martingale converging to  $\mathbf{1}_A$  almost surely. Therefore

$$(24.27) \quad \mathbf{1}_A = \lim_{n \rightarrow \infty} X_n.$$

However,  $A$  is independent of  $\mathcal{F}_n$  because  $A \in \bigvee_{j=n+1}^{\infty} \mathcal{G}_j$ . As a result we have  $X_n = P(A)$ . Therefore  $\mathbf{1}_A$  is constant except on a null set. This means  $P(A) = 0$  or  $1$ . □

*Exercise 171.* Suppose that  $X_1, X_2, \dots, X_n, \dots$  is a sequence of real valued independent random variables. Then show that the probability for which

$$(24.28) \quad Y := \lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n}$$

converges is 0 or 1.

### 24.5. The strong law of large numbers.

As a further application of the convergence theorem we prove the strong law of large numbers.

**Theorem 24.11** (Kronecker's theorem). *Suppose that  $\{a_j\}_{j \in \mathbb{N}}$  is an increasing sequence tending to  $\infty$  and  $\{x_j\}_{j \in \mathbb{N}}$  is another sequence. Assume that  $\sum_{j=1}^{\infty} \frac{x_j}{a_j}$  converges. Then*

$$(24.29) \quad \lim_{j \rightarrow \infty} \frac{x_1 + x_2 + \dots + x_j}{a_j} = 0.$$

*Proof.* Let us set  $y_j := \sum_{k=1}^j \frac{x_k}{a_k}$  and  $y_0 := 0$ . Then we have  $x_j = a_j(y_j - y_{j-1})$  for all  $j \in \mathbb{N}$ .

Therefore, we obtain

$$(24.30) \quad \sum_{k=1}^j x_k = \sum_{k=1}^j a_k y_k - \sum_{k=1}^j a_k y_{k-1} = \sum_{k=1}^{j-1} y_k (a_k - a_{k+1}) + a_j y_j.$$

The assumption reads

$$(24.31) \quad y_{\infty} := \lim_{j \rightarrow \infty} y_j$$

exists. Therefore, it suffices to prove that

$$(24.32) \quad \lim_{j \rightarrow \infty} \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) y_k = y_\infty.$$

It is clear that

$$(24.33) \quad \lim_{j \rightarrow \infty} \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) y_\infty = y_\infty,$$

since  $\{a_j\}_{j \in \mathbb{N}}$  is increasing to  $\infty$ . In view of this, let us estimate

$$(24.34) \quad \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) (y_k - y_\infty)$$

Let  $N \in \mathbb{N}$  be fixed. In view of the fact that  $\{a_j\}_{j \in \mathbb{N}}$  is increasing, we obtain

$$(24.35) \quad \left| \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) (y_k - y_\infty) \right| \leq \frac{1}{a_j} \sum_{k=1}^N (a_{k+1} - a_k) |y_k - y_\infty| + \sup_{l \geq N} |y_l - y_\infty| \cdot \frac{a_j - a_N}{a_j}$$

for all  $j > N$ . Letting  $j \rightarrow \infty$ , we obtain

$$(24.36) \quad \limsup_{j \rightarrow \infty} \left| \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) (y_k - y_\infty) \right| \leq \sup_{l \geq N} |y_l - y_\infty|.$$

Since  $N$  is still at our disposal, we obtain

$$(24.37) \quad \lim_{j \rightarrow \infty} \frac{1}{a_j} \sum_{k=1}^{j-1} (a_{k+1} - a_k) y_k = y_\infty.$$

Thus, the proof is now complete.  $\square$

**Theorem 24.12** (Laws of large numbers, LLN). *Suppose that  $(X_j)_{j \in \mathbb{N}}$  is a sequence of i.i.d. random variables. Then*

$$(24.38) \quad \lim_{j \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = E[X_1]$$

for almost surely.

*Proof.* Let us set  $Y_j = \mathbf{1}_{\{|X_j| < j\}} \cdot X_j$ . Denote by  $\mu$  the distribution of  $X_1$ . Then we have

$$\sum_{j=1}^{\infty} P(X_j \neq Y_j) = \sum_{j=1}^{\infty} \mu((-\infty, -j] \cup [j, \infty)) \leq \int_{\mathbb{R}} |t| d\mu(t) \leq E[|X_1|] < \infty.$$

Therefore, we have only to prove  $\lim_{n \rightarrow \infty} \frac{Y_1 + Y_2 + \dots + Y_n}{n} = E[X_1]$  for almost surely.

In the same way we obtain

$$(24.39) \quad \sum_{j=1}^{\infty} \frac{E[|Y_j|^2]}{j^2} = \sum_{j=1}^{\infty} \int_0^{\infty} \lambda^2 \cdot \chi_{\{|\lambda| < j\}}(\lambda) d\mu(\lambda) = \int_0^{\infty} \lambda^2 \left( \sum_{j=|\lambda|+1}^{\infty} \frac{1}{j^2} \right) d\mu(\lambda).$$

Since  $\sum_{j=|\lambda|+1}^{\infty} \frac{1}{j^2} \lesssim \frac{1}{|\lambda|+1}$ , we obtain  $\sum_{j=1}^{\infty} \frac{E[|Y_j|^2]}{j^2} \lesssim \int_0^{\infty} |\lambda| d\mu(\lambda) \lesssim E[|X_1|] < \infty$ .

Let  $j \in \mathbb{N}$ . Then we set  $M_n = \sum_{k=1}^n \frac{Y_k - E[Y_k]}{k}$ . Then  $(M_n)_{n \in \mathbb{N}}$  is a martingale with

$$(24.40) \quad \sup_{n \in \mathbb{N}} E[|M_n|^2] \leq \sum_{j=1}^{\infty} \frac{E[|Y_j|^2]}{k^2} < \infty$$

in view of the paragraph above. Therefore, the convergence theorem above gives us

$$(24.41) \quad \lim_{j \rightarrow \infty} M_n(\omega)$$

exists for almost every  $\omega \in \Omega$ . If we apply Theorem 24.11 with  $a_n = n$  and  $x_n = Y_n(\omega) - E[Y_n]$ , we obtain

$$(24.42) \quad \lim_{n \rightarrow \infty} \frac{Y_1 + Y_2 + \dots + Y_n - E[Y_1 + Y_2 + \dots + Y_n]}{n}$$

exists. Since

$$(24.43) \quad \frac{E[Y_1 + Y_2 + \dots + Y_n]}{n} = \int_{\mathbb{R}} \frac{\chi_{(-1,1)}(t) + \chi_{(-2,2)}(t) + \dots + \chi_{(-n,n)}(t)}{n} d\mu(t),$$

we are in the position of using the Lebesgue convergence theorem to obtain

$$(24.44) \quad \lim_{n \rightarrow \infty} \frac{Y_1 + Y_2 + \dots + Y_n}{n} = \lim_{n \rightarrow \infty} \frac{E[Y_1 + Y_2 + \dots + Y_n]}{n} = E[X_1].$$

This is the desired result. □

**24.6. Uniform integrability.**

Suppose that  $(X_n)_{n \in \mathbb{N}}$  converges to  $X$  almost surely and that  $E[X_n | \mathcal{F}_m] = X_{\min(m,n)}$  almost surely. Can we say that there exists  $X \in L^1(P)$  such that  $X_n = E[X | \mathcal{F}_n]$  almost surely? The answer is no. Indeed, let  $(Y_n)_{n=1}^{\infty}$  be a sequence of i.i.d. random variables such that  $P(Y_n = \pm 1) = 1/2$ . Then put

$$X_n = \sum_{k=1}^n Y_k$$

for each  $n \in \mathbb{N}$ . The sequence  $(X_n)_{n \in \mathbb{N}}$  satisfies the condition above but there is no limit  $X$ . We attempt to give a necessary and sufficient condition for this problem.

**Definition 24.13.** A set of random variables  $(X_\lambda)_{\lambda \in \Lambda}$  is said to be uniformly integrable, if

$$(24.45) \quad \lim_{\rho \rightarrow \infty} \left( \sup_{\lambda \in \Lambda} E[|X_\lambda| : |X_\lambda| > \rho] \right) = 0.$$

Let us see the family of uniformly integrable random variables.

**Example 24.14.** Suppose that  $(\mathcal{G}_\lambda)_{\lambda \in \Lambda}$  be a family of sub- $\sigma$  fields. Let  $X \in L^1(\Omega, \mathcal{F}, P)$ . Then  $(E[X : \mathcal{G}_\lambda])_{\lambda \in \Lambda}$  forms a uniformly integrable family.

*Proof.* We have to show

$$(24.46) \quad \lim_{\rho \rightarrow \infty} \sup_{\lambda \in \Lambda} E[|E[X : \mathcal{G}_\lambda]| : |E[X : \mathcal{G}_\lambda]| > \rho] = 0.$$

Taking into account that  $E[X : \mathcal{G}_\lambda] \in \mathcal{G}_\lambda$ , we have

$$(24.47) \quad E[|E[X : \mathcal{G}_\lambda]| : |E[X : \mathcal{G}_\lambda]| > \rho] \leq E[|X| : |E[X : \mathcal{G}_\lambda]| > \rho].$$

By absolute continuity of measures, (24.46) can be obtained. □

Another important example is

**Example 24.15.** Let  $p > 1$  and  $(X_\lambda)_{\lambda \in \Lambda}$  is  $L^p$ -bounded. Then  $(X_\lambda)_{\lambda \in \Lambda}$  is uniformly integrable.

*Proof.* Immediate from  $E[|X_\lambda| : |X_\lambda| > \lambda] \leq \frac{1}{\lambda} \sup_{\lambda \in \Lambda} E[|X_\lambda|^p]$ .  $\square$

Having made a brief look of examples of the uniformly integrable family, let us characterize the uniform integrable family.

**Definition 24.16.** A family of  $L^1(P)$ -random variables  $\mathcal{X} = \{X_\lambda\}_{\lambda \in \Lambda}$  is said to be uniformly absolutely continuous, if

$$(24.48) \quad \lim_{\delta \rightarrow 0} \sup_{\substack{A \in \mathcal{F} \\ P(A) < \delta}} E[|X| : A] = 0.$$

**Theorem 24.17.** Let  $\mathcal{X} = (X_\lambda)_{\lambda \in \Lambda}$  be a family of  $L^1(P)$ -random variables. Then  $\mathcal{X}$  is uniformly integrable if and only if it is  $L^1(P)$ -bounded and uniformly absolutely continuous.

**Theorem 24.18.** Suppose that  $(X_n)_{n \in \mathbb{N}}$  converges to  $X$  almost surely and that  $E[X_n | \mathcal{F}_m] = X_{\min(m,n)}$  almost surely. Then  $(X_n)_{n \in \mathbb{N}}$  is uniformly integrable if and only if  $X_n$  converges to  $X$  in the  $L^1(P)$ -topology.

*Proof.* Suppose that  $X_n$  converges to  $X$  in the  $L^1(P)$ -topology.

We have to prove

$$(24.49) \quad \lim_{\lambda \rightarrow \infty} \sup_{n \in \mathbb{Z}} E[|X_n| : |X_n| > \lambda] = 0.$$

For this purpose it suffices to prove

$$(24.50) \quad \lim_{\lambda \rightarrow \infty} \sup_{n \in \mathbb{Z}} E[|X| : |X_n| > \lambda] = 0,$$

since the estimate

$$(24.51) \quad \sup_{n \in \mathbb{Z}} E[|X_n - X| : |X_n| > \lambda] \leq \sup_{n \geq m} E[|X - X_n|] + \sum_{k=1}^m E[|X - X_k| : |X_k| > \lambda]$$

and  $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$  for almost sure  $\omega$  implies

$$(24.52) \quad \overline{\lim}_{\lambda \rightarrow \infty} \sup_{n \in \mathbb{Z}} E[|X_n - X| : |X_n| > \lambda] \leq \sup_{n \geq m} E[|X - X_n|].$$

(24.50) is a consequence of absolute continuity of the measure.

Suppose that  $(X_n)_{n \in \mathbb{N}}$  is uniformly integrable. Then

$$\begin{aligned} E[|X_n - X|] &\leq \sup_{n \in \mathbb{N}} E[|X_n| : |X_n| > \lambda] + \sup_{n \in \mathbb{N}} E[|X| : |X_n| > \lambda] \\ &\quad + E[|X| + \lambda : |X_n| \leq \lambda, |X| > \lambda] + E[|X_n - X| : |X_n| \leq \lambda, |X| \leq \lambda]. \end{aligned}$$

Let  $\varepsilon > 0$  be arbitrary. Then there exists  $\delta > 0$  such that for each  $A \in \mathcal{F}$  with  $P(A) < \delta$  we have  $E[|X| : A] < \varepsilon$ .

If we take  $\lambda > 0$  large enough, then

$$(24.53) \quad \lambda P(|X| > \lambda) \leq E[|X| : |X| > \lambda] < \varepsilon, \sup_{n \in \mathbb{N}} E[|X_n| : |X_n| > \lambda] < \varepsilon, \sup_{n \in \mathbb{N}} P(|X_n| > \lambda) < \delta.$$

We fix  $\lambda$  so large that the above condition holds. Since  $|X_n - X| \leq \lambda$ , there exists  $N$  such that, for every  $n \geq N$ , we have

$$(24.54) \quad E[|X_n - X| : |X_n| \leq \lambda, |X| \leq 2\lambda] < \varepsilon.$$



Putting these observations together,  $\lim_{n \rightarrow \infty} E[|X_n - X|] = 0$ .

As a consequence the proof is finished. □

**24.7. Upcrossing time and almost sure convergence.**

Upcrossing number and upcrossing time are useful for almost sure convergence.

**Definition 24.19** (Upcrossing number). Let  $X = (X_n)_{n \in \mathbb{N}_0}$  be a submartingale and let  $-\infty < a < b < \infty$ .

(1) One first defines

$$(24.55) \quad \sigma_1 := \min\{n \geq 1 : X_n \leq a\}, \quad \sigma_2 := \min\{n > \sigma_1 : X_n \leq b\}.$$

Suppose that  $\sigma_1, \sigma_2, \dots, \sigma_{2k-2}$  are defined. Then set

$$\begin{aligned} \sigma_{2k-1} &:= \min\{n > \sigma_{2k-2} : X_n \leq a\} \\ \sigma_{2k} &:= \min\{n > \sigma_{2k-1} : X_n \geq b\}. \end{aligned}$$

(2) Set the upcrossing number by

$$(24.56) \quad U_n := \max\{k \in \mathbb{N}_0 : \sigma_{2k} \leq n\} (\leq n).$$

Here it will be understood that  $\min \emptyset = \infty$ .

Note that  $\sigma_{k+1} \geq \sigma_k + 1$  for all  $k \in \mathbb{N}$  in the above definition.

*Exercise 172.* Show that each  $U_n$  is a random variable.

The following inequality, called upcrossing inequality, plays an important role for the proof of almost everywhere convergence.

**Theorem 24.20** (Upcrossing inequality). *Let  $-\infty < a < b < \infty$ . Then we have  $(b-a)E[U_n] \leq E[(X_n - a)_+]$ .*

*Proof.* We set  $Y_n := (X_n - a)_+$ . Suppose that  $U_n = k$ . Then, taking into account that  $Y_{\sigma_{2j}} - Y_{\sigma_{2j-1}} \geq b - a$ , we obtain

$$(24.57) \quad (b-a)U_n = (b-a)k \leq \sum_{j=1}^k (Y_{\sigma_{2j}} - Y_{\sigma_{2j-1}}).$$

Set  $H_n = \sum_{k=1}^{\infty} \mathbf{1}_{\{\sigma_{2k-1} < n \leq \sigma_{2k}\}}$  for  $n \geq 1$  and  $H_0 = 0$ . Then  $H_n$  is predictable. Indeed,

$$(24.58) \quad \mathbf{1}_{\{\sigma_{2k-1} < n \leq \sigma_{2k}\}} = \mathbf{1}_{\{\sigma_{2k-1} \leq n-1\}} - \mathbf{1}_{\{\sigma_{2k} \leq n-1\}} \in \mathcal{F}_{n-1}.$$

Then  $\sum_{j=1}^n Y_{\{\sigma_{2j}} - Y_{\sigma_{2j-1}}\}$  can be written as

$$(24.59) \quad \sum_{j=1}^n (Y_{\sigma_{2j}} - Y_{\sigma_{2j-1}}) = (H \cdot X)_n.$$

And  $H \cdot X$  is a submartingale. Consequently

$$(24.60) \quad E[(b-a)U_n] \leq E[(H \cdot Y)_n] = E[Y_n] - E[((1-H) \cdot Y)_n] \leq E[Y_n].$$

This is the desired result. □

Finally to conclude this part, we refine the almost sure convergence theorem.

**Theorem 24.21.** *If  $X$  is a submartingale such that  $\sup_{n \in \mathbb{N}} E[X_n^+] < \infty$ . Then the almost sure limit  $\lim_{n \rightarrow \infty} X_n =: X_\infty$  exists.*

*Proof.* Set

$$(24.61) \quad \bar{X} := \limsup_{n \rightarrow \infty} X_n, \quad \underline{X} := \liminf_{n \rightarrow \infty} X_n.$$

We shall show this theorem by proving

$$(24.62) \quad P(\bar{X} > \underline{X}) = 0$$

or equivalently

$$(24.63) \quad P(\bar{X} > b > a > \underline{X}) = 0$$

for all  $a, b \in \mathbb{Q}$  with  $a < b$ . Now that  $\{X_n(\omega)\}_{n \in \mathbb{N}_0}$  upcrosses  $(a, b)$  infinitely many times, if  $\bar{X}(\omega) > b > a > \underline{X}(\omega)$ . Thus,  $(b - a)E[U_n] \leq E[(X_n - a)_+]$  for  $n \in \mathbb{N}$  implies, along with the assumption,  $(b - a)E[U_\infty] < \infty$ . Thus,  $U_\infty < \infty$  almost surely. This implies that  $P(\bar{X} > b > a > \underline{X}) = 0$  and the proof is now complete.  $\square$

**Theorem 24.22.** *Suppose that  $M = (M_n)_{n \in \mathbb{N}}$  is a uniformly integrable martingale. Then*

$$(24.64) \quad M_\infty := \lim_{n \rightarrow \infty} M_n$$

*exists almost surely and in  $L^1(P)$ . Furthermore we have*

$$(24.65) \quad M_n = E[M_\infty : \mathcal{F}_n]$$

*holds.*

*Proof.* Since  $M$  is uniformly integrable,  $M$  is  $L^1(P)$ -bounded. Therefore, by Theorem 24.21 (24.64) holds almost surely. Again by the uniform integrability and the almost sure convergence, we see that (24.64) takes place in  $L^1(P)$ . Since  $M_n = E[M_{n+j} : \mathcal{F}_n]$  for  $j \in \mathbb{N}$ , the  $L^1(P)$ -convergence gives us (24.65).  $\square$

*Exercise 173.* Use the random walk to show that we really need to assume that  $M$  is uniformly integrable.

Notes and references for Chapter 11.

Section 22. Theorem 22.14

Theorem 22.16

Theorem 22.25

Theorem 22.27

Theorem 22.29 is due to de Moivre and Laplace.

Section 23. Theorem 22.32

Theorem 22.35

Theorem 22.36

Theorem 22.37

Theorem 22.38

Theorem 22.39

Theorem 23.5

Theorem 23.7

Burkholder considered Theorem 23.12 in [104, 104].

Theorem 23.14

Theorem 23.17

Theorem 23.24

Theorem 24.1

Theorems 23.21 and 24.2 are due to Doob [8]. For Corollary 24.3 we refer to [349].

Theorem 24.5

Section 24. Theorem 24.7

Theorem 24.8

Theorem 24.10

Theorem 24.11

Theorem 24.12

Theorem 24.17

Theorem 24.18

Theorem 24.20

Theorem 24.21

Theorem 24.22

## Part 12. Ergodic theory

In this chapter we consider ergodic theory.

In Section 25 we introduce the notion of ergodicity. In Section 26 we deal with the maximal inequalities.

Throughout this chapter we assume that  $(\Omega, \mathcal{F}, P)$  is a probability space.

### 25. ERGODICITY

Let us begin with stating fundamental concepts.

**Definition 25.1.** A measurable bijection  $T : \Omega \rightarrow \Omega$  is measure preserving, if  $T(\mathcal{F}) = \mathcal{F}$  and  $P(TA) = P(A)$  for all  $A \in \mathcal{F}$ .

Here and below, by convention of ergodic theory, we write  $TA$  instead of  $T(A)$ .

**Definition 25.2.** A measure preserving mapping  $T : \Omega \rightarrow \Omega$  is ergodic, if  $TA = A$  implies  $P(A) = 0, 1$  for all  $A \in \mathcal{F}$ .

**Example 25.3.** Let us denote by  $S^1$  the unit circle on  $\mathbb{C}$ . Then for every  $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ , the mapping  $T : z \in S^1 \rightarrow \exp(2\pi\alpha)z \in S^1$  is ergodic.

*Proof.* Let  $A$  be an invariant set. Then consider its indicator and expand it into Fourier series.

$$(25.1) \quad \chi_A(z) = \sum_{j=-\infty}^{\infty} c_j z^j.$$

Since  $A$  is invariant, we have

$$(25.2) \quad \sum_{j=-\infty}^{\infty} c_j z^j = \chi_A(z) = \chi_A(\exp(2\pi\alpha)z) = \sum_{j=-\infty}^{\infty} c_j \exp(2\pi\alpha j) z^j$$

Hence it follows that

$$(25.3) \quad c_j = c_j \exp(2\pi\alpha j)$$

for all  $j \in \mathbb{Z}$ , which yields  $\chi_A(z)$  is a constant function a.e.. As a result  $A$  or  $S^1 \setminus A$  has measure zero.  $\square$

*Exercise 174.* If  $\alpha \in \mathbb{Q}$ , show that  $T$ , defined in Example 25.3, is not ergodic.

### 26. ERGODIC MAXIMAL FUNCTION

Here we will view the similarity between harmonic analysis and ergodic theory.

$L^2(P)$ -theory.

**Theorem 26.1.** Let  $f \in L^2(P)$  and  $T$  be a measure preserving mapping. Then

$$(26.1) \quad \lim_{j \rightarrow \infty} \frac{1}{j} \sum_{k=1}^j f \circ T^k(\omega)$$

exists for a.s.  $\omega \in \Omega$  and in the topology of  $L^2(P)$ .

*Proof.* It is easy to see that the assertion is true if  $f$  can be written as  $f = g - g \circ T$  for some  $g \in L^2(P)$  or if  $f$  is invariant under  $T$ . Therefore, by the Banach-Steinhaus principle it remains to show that the sum of such functions spans a dense space.

Suppose that  $f$  is a function that is perpendicular to such functions. Since  $f \perp (f - f \circ T)$ , we see that

$$(26.2) \quad \langle f, f - f \circ T \rangle_{L^2(P)} = 0$$

As a consequence, since  $T$  is a measure preserving, we obtain

$$(26.3) \quad \langle f - f \circ T, f - f \circ T \rangle_{L^2(P)} = \langle -f \circ T, f \circ T^{-1} \circ T - f \circ T \rangle_{L^2(P)}$$

$$(26.4) \quad = \langle -f, f \circ T^{-1} - f \rangle_{L^2(P)}$$

$$(26.5) \quad = \langle -f, (f \circ T^{-1}) - (f \circ T^{-1}) \circ T \rangle_{L^2(P)}$$

$$(26.6) \quad = 0$$

Hence  $f$  is invariant under  $T$ . However,  $f$  is assumed to be perpendicular to any function invariant under  $T$ . As a result, we see that  $f = 0$ .  $\square$

Maximal inequality. The following theorem, called the maximal ergodic theorem, is a heart of the matters in ergodic theory.

**Theorem 26.2.** *Let  $T : \Omega \rightarrow \Omega$  be a measure preserving mapping. Then we have*

$$(26.7) \quad \int_G f(\omega) dP(\omega) \geq 0,$$

where

$$(26.8) \quad G = \left\{ \sup_{j \in \mathbb{N}_0} \sum_{l=0}^j f \circ T^l > 0 \right\}.$$

*Proof.* Let us set

$$(26.9) \quad G_J = \left\{ \sup_{0 \leq j \leq J} \sum_{l=1}^j f \circ T^l > 0 \right\}, \quad J = 0, 1, 2, \dots$$

Notice that  $G_J$  converges monotonically to  $G$ . Thus, we have

$$(26.10) \quad \int_G f(\omega) dP(\omega) = \lim_{J \rightarrow \infty} \int_{G_J} f(\omega) dP(\omega) = \lim_{J \rightarrow \infty} \frac{1}{J} \sum_{k=0}^J \int_{G_k} f(\omega) dP(\omega).$$

With this in mind, let us prove that

$$(26.11) \quad \sum_{k=0}^J \int_{G_k} f(\omega) dP(\omega) \geq 0.$$

Let us now consider the case  $J = J_0$ . Now that  $T$  preserves measure, we have

$$(26.12) \quad \sum_{k=0}^J \int_{G_k} f(\omega) dP(\omega) = \sum_{k=0}^J \int_{\left\{ \sup_{0 \leq j \leq k} \sum_{l=j-k}^{j-1} f \circ T^l > 0 \right\}} f \circ T^{J-k}(\omega) dP(\omega).$$

Let  $f_k = f \circ T^k$ . Then we have

$$(26.13) \quad \sum_{k=0}^J \int_{G_k} f(\omega) dP(\omega) = \int_{\Omega} \sum_{k=0}^J \chi_{\left\{ \sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} f_l > 0 \right\}} f_k(\omega) dP(\omega).$$

Therefore, we have only to prove that

$$(26.14) \quad \sum_{k=0}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} f_l > 0\}} f_k \geq 0.$$

More generally, let us establish that

$$(26.15) \quad \sum_{k=0}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} a_l > 0\}} a_k \geq 0$$

for any sequence  $\{a_l\}_{l \in \mathbb{N}}$ .

If  $J = 0$ , then there is nothing to prove. Assume that this is the case for  $J_0$ . If

$$(26.16) \quad \sup_{0 \leq j \leq J} \sum_{l=0}^j a_l \leq 0$$

then we can readily obtain

$$(26.17) \quad \sum_{k=0}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} a_l > 0\}} a_k = \sum_{k=1}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} a_l > 0\}} a_k \geq 0$$

by induction assumption. Assume  $\sup_{0 \leq j \leq J} \sum_{l=0}^j a_l > 0$  instead. Let  $k_0 \geq 0$  be the smallest integer

such that  $\sum_{l=0}^{k_0} a_l > 0$ . Then we have that  $\sum_{l=k+1}^{k_0} a_l > 0$  for all  $k \leq k_0 - 1$ . because  $\sum_{l=0}^k a_l < 0$ . As a result, we obtain

$$(26.18) \quad \sum_{k=0}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} a_l > 0\}} a_k = \sum_{k=0}^{k_0} a_k + \sum_{k=k_0+1}^J \chi_{\{\sup_{0 \leq j \leq J-k} \sum_{l=k}^{k+j} a_l > 0\}} a_k > 0.$$

This is the desired result. □

*L<sup>p</sup>-theory.* Once we obtain Theorem 26.2, we can prove fundamental theorems in ergodic theory in a satisfactory manner.

**Theorem 26.3.** *Let  $T$  be a measure preserving.*

- (1) *Let  $\lambda > 0$ . Then we have  $P \left( \sup_{J \in \mathbb{N}} \frac{1}{J} \sum_{j=0}^{J-1} |f \circ T^j| > \lambda \right) \leq \frac{1}{\lambda} E[|f|]$*
- (2) *Let  $1 < p < \infty$ . Then we have*

$$(26.19) \quad E \left[ \sup_{J \in \mathbb{N}} \left| \frac{1}{J} \sum_{j=0}^{J-1} f \circ T^j \right|^p \right] \leq \frac{p \cdot 2^p}{p-1} E[|f|^p].$$

- (3) *If  $f \in L^1(P)$ , then the limit*

$$(26.20) \quad \lim_{j \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N f \circ T^j$$

*exists for a.s. (Birkhoff's ergodic theorem).*

(4) Assume in addition that  $T$  is ergodic. Then

$$(26.21) \quad E[f] = \lim_{j \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N f \circ T^j$$

for a.s. (von Neumann's ergodic theorem).

*Exercise 175.* Prove Theorem 26.3 by mimicking the proof of the boundedness of the maximal operators.

Notes and references for Chapter 12.

Section 25.

Section 26. Theorem 26.1 stems from the work of von Neumann (1932).

Theorem 26.2 dates back to Riesz (1945).

Theorem 26.3 (3), which is called the pointwise ergodic theorem, is due to Birkhoff (1931). The proof given here depends upon Riesz (1945) who considered and applied Theorem 26.2.

**Part 13. Functional analysis and harmonic analysis****Part 14. More about functional analysis**

The aim of this in this part is to review functional analysis, keeping rich examples obtained in this book in mind. For example, what we have proved in Chapter 8 provides rich examples.

**Example 26.4.** Singular integral operators are bounded on  $L^p$  for all  $1 < p < \infty$ . Let  $0 < \alpha < d$ . Then the fractional integral operator  $I_\alpha$  is bounded from  $L^p(\mathbb{R}^d)$  to  $L^q(\mathbb{R}^d)$  whenever  $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$ .

In Chapter 14 we build up theory on functional analysis. What is totally different from the usual book on functional analysis is that we have already kept abundant examples in mind. Section 27 is a combination of Banach space theory and theory of integration. In this section as an application of the Bochner integral we consider the semigroup of operators. In Section 29 we are going to prove the spectral decomposition as well as to develop theory of Banach algebras.

Having set down the elementary notions in functional analysis, the remaining three chapters are devoted to deeper consideration of functional analysis.

## 27. BOCHNER INTEGRAL

The Bochner integral is an advanced topic for beginners. However, we shall need it in this book. We assume that  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space.

## 27.1. Measurable functions.

Our present aim, as we have been referring to, is the construction of the theory of integration for  $B$ -valued functions. The theory of the underlying measure space  $(X, \mathcal{B}, \mu)$  is already set up, so that it seems appropriate that we start from the definition of the measurability of  $B$ -valued functions,

**Definition 27.1.** Let  $B$  be a Banach space and  $(X, \mathcal{B}, \mu)$  a measure space. Suppose  $\varphi : X \rightarrow B$  is a function.

- (1)  $\varphi$  is said to be weakly measurable, if  $b^* \circ \varphi : X \rightarrow \mathbb{K}$  is measurable for all  $b^* \in B^*$ .
- (2)  $\varphi$  is said to be simple, if there exists  $x_1, x_2, \dots$  and  $A_1, A_2, \dots \in \mathcal{B}$  such that

$$\varphi(x) = \sum_{j=1}^{\infty} \chi_{A_j}(x) \cdot x_j$$

for  $\mu$ -almost all  $x \in X$

- (3)  $\varphi$  is said to be strongly measurable, if there exists a sequence of simple functions  $\{\varphi_j\}_{j \in \mathbb{N}}$  such that  $\lim_{j \rightarrow \infty} \varphi_j(x) = \varphi(x)$  for  $\mu$ -a.e.  $x \in X$ .
- (4)  $\varphi$  is said to be separably valued, if there exists a measurable set  $X_0$  such that  $\varphi(X_0)$  is separable and  $\mu(X \setminus X_0) = 0$ .

The next lemma is useful when we consider separably valued functions.



**Lemma 27.2.** *If  $Y$  is a separable Banach space, then there exists a countable set  $Z^* \subset (Y^*)_1$  such that*

$$(27.1) \quad \|y\|_Y = \sup_{z^* \in Z^*} |z^*(y)|$$

for all  $y \in Y$ .

*Proof.* Pick a countable dense set  $Y_0 = \{y_j\}_{j \in \mathbb{N}}$ . For each  $j \in \mathbb{N}$  by Hahn-Banach theorem we can find a norm attainer  $z_j^* \in (Y^*)_1$  of  $y_j$ . We have only to set  $Z^* := \{z_j^* : j \in \mathbb{N}\}$ .  $\square$

Below given a  $B$ -valued function  $\varphi : X \rightarrow B$ , let us write

$$(27.2) \quad \|\varphi\|_B(x) := \|\varphi(x)\|_B,$$

that is,  $\|\varphi\|_B$  is a non-negative function.

**Theorem 27.3.** *Let  $\varphi : X \rightarrow B$  be a function. Then  $\varphi$  is strongly measurable, if and only if  $\varphi$  is separably valued and weakly measurable.*

*Proof.* It is straightforward to prove “only if” part and we leave this for the readers. Assume that  $\varphi$  is separably valued and weakly measurable. We may assume by disregarding a set of measure zero that  $\varphi(X)$  itself is separable. Let  $Y = \overline{\varphi(X)}$  and  $\{y_k\}_{k \in \mathbb{Z}}$  a countable dense set in  $Y$ . Then by Lemma 27.2 we obtain  $z_1^*, z_2^*, \dots, z_k^*, \dots \in (Y^*)_1$  such that

$$(27.3) \quad \|y\|_Y = \sup_{j \in \mathbb{N}} |z_j^*(y)|.$$

From this we conclude that  $\|\varphi(\cdot) - z\|_B$  is a measurable function.

Let  $j \in \mathbb{N}$ . Then

$$(27.4) \quad A_{k,j} := \{\|\varphi - y_k\|_X < j^{-1}\}.$$

Set

$$(27.5) \quad B_{1,j} = A_{1,j}, B_{k,j} = A_{k,j} \setminus (A_{1,j} \cup A_{2,j} \cup \dots \cup A_{k-1,j}), k \geq 2.$$

Then we have only to set

$$(27.6) \quad \varphi_j(x) := \sum_{k=1}^{\infty} \chi_{B_{k,j}}(x) \cdot y_k.$$

From the property of  $A_{k,j}$ ,  $k, j \in \mathbb{N}$  we conclude that  $\varphi_j$  tends to  $\varphi$  and hence  $\varphi$  is separably valued.  $\square$

*Exercise 176.* Prove that  $\varphi$  is separably valued and weakly measurable, if  $\varphi$  is strongly measurable.

### 27.2. Definition of the Bochner integral.

Having clarified the definition of measurability, we turn to the definition of the Bochner integral.

Bochner integral for ( countably ) simple functions. Now we are going to define integrals for countably simple functions. As we did for the usual measurable measurable functions, it is convenient to start from the definition of the countably simple function. But we need be careful: It will not work if we define the countably simple functions as the one taking finitely many values. What is quite different is that countably simple functions are allowed to take countably many values.

**Definition 27.4** (Integrable countably simple functions). Let  $(X, \mathcal{B}, \mu)$  be a measure space and  $B$  a Banach space. A countably simple function  $\varphi: X \rightarrow B$  is said to be integrable, if it admits a representation

$$(27.7) \quad \varphi(x) = \sum_{j=1}^{\infty} a_j \chi_{E_j}(x)$$

$\mu$ -a.e.  $x \in X$  with

$$(27.8) \quad \sum_{j=1}^{\infty} \|a_j\|_B \cdot \mu(E_j) < \infty.$$

The above representation is called integrable representation and write

$$(27.9) \quad \varphi \simeq \sum_{j=1}^{\infty} a_j \chi_{E_j}.$$

**Lemma 27.5.** *Suppose that  $\varphi$  is a countably simple integrable functions. Then there exists a unique element  $\Phi$  such that*

$$(27.10) \quad b^*(\Phi) = \int_X b^*(\varphi(x)) d\mu(x)$$

for all  $b^* \in B^*$ . Below we shall write

$$(27.11) \quad \Phi = \int_X \varphi(x) d\mu(x).$$

*Proof.* Existence Choose an integrable representation  $\varphi(x) \simeq \sum_{j=1}^{\infty} a_j \chi_{E_j}(x)$ . Then define an element  $\Phi := \sum_{j=1}^{\infty} a_j \mu(E_j) \in B$ . Note that the convergence is absolute. Let  $y \in B^*$ . Then we have

$$(27.12) \quad b^*(\Phi) = \sum_{j=1}^{\infty} b^*(a_j) \mu(E_j) = \int_X b^* \circ \varphi(x) d\mu(x),$$

where we have used the Lebesgue convergence theorem for the second equality.

Uniqueness Uniqueness is almost clear from the Hahn-Banach extension theorem. □

**Lemma 27.6.** *Let  $\varphi: X \rightarrow B$  be an integrable countably simple function. Then we have*

$$(27.13) \quad \left\| \int_X \varphi(x) d\mu(x) \right\|_B \leq \int_X \|\varphi\|_B(x) d\mu(x).$$

*In particular if  $\varphi$  has an integrable representation  $\varphi \simeq \sum_{j=1}^{\infty} a_j \chi_{E_j}$ , then*

$$(27.14) \quad \left\| \int_X \varphi(x) d\mu(x) \right\|_B \leq \sum_{j=1}^{\infty} \|a_j\|_B \mu(E_j).$$

*Proof.* By the Hahn-Banach theorem we can choose  $y^*$  from the closed unit ball of  $B^*$ . Then we have

$$\begin{aligned} \left\| \int_X \varphi(x) d\mu(x) \right\|_B &= y^* \left( \int_X \varphi(x) d\mu(x) \right) = \int_X y^* \circ \varphi(x) d\mu(x) \\ &\leq \int_X \|\varphi(x)\|_B d\mu(x), \end{aligned}$$

where the second inequality follows from the definition of  $\int_X \varphi(x) d\mu(x)$ . The second inequality is a direct consequence of the inequality

$$(27.15) \quad \|\varphi(x)\|_B \leq \sum_{j=1}^{\infty} \|a_j\|_B \chi_{E_j}(x)$$

for  $\mu$ -almost every  $x \in X$ . Therefore, the proof is now complete.  $\square$

**Lemma 27.7.** *Suppose that  $\varphi$  is a countably simple measurable function and  $\varepsilon > 0$ . Then  $\varphi$  has an integrable representation  $\varphi \simeq \sum_{j=1}^{\infty} a_j \chi_{E_j}$  such that*

$$(27.16) \quad \sum_{j=1}^{\infty} \|a_j\|_X \mu(E_j) < \int_X \|\varphi\|_B(x) d\mu(x) + \varepsilon.$$

*Proof.* Suppose that  $\varphi \simeq \sum_{j=1}^{\infty} b_j \chi_{F_j}$  is an integrable representation. Then there exists an increasing sequence of integers  $N_1 < N_2 < \dots$  such that

$$(27.17) \quad \sum_{j=N_k}^{\infty} \|b_j\|_X \mu(F_j) < \frac{\varepsilon}{2^{k+1}}.$$

for each  $k$ . Let  $N_0 = M_0 = 1$ . By partitioning  $F_{N_{k-1}}, F_{N_{k-1}+1}, \dots, F_{N_k}$  we can find a collection of disjoint measurable sets  $E_{M_{k-1}}, E_{M_{k-1}+1}, \dots, E_{M_k}$  and  $a_{M_{k-1}}, a_{M_{k-1}+1}, \dots, a_{M_k} \in B$

$$(27.18) \quad \sum_{l=M_{k-1}}^{M_k-1} a_l \chi_{E_l} = \sum_{j=N_{k-1}}^{N_k-1} b_j \chi_{F_j}.$$

Then disjointness of  $\{E_j\}_{j=M_{k-1}}^{M_k-1}$  yields

$$(27.19) \quad \int_X \left\| \sum_{j=N_{k-1}}^{N_k-1} b_j \chi_{F_j}(x) \right\|_B d\mu(x) = \int_X \left\| \sum_{l=M_{k-1}}^{M_k-1} a_l \chi_{E_l}(x) \right\|_B d\mu(x) = \sum_{l=M_{k-1}}^{M_k-1} \|a_l\|_B \mu(E_l).$$

Therefore, we obtain

$$\begin{aligned}
\sum_{l=1}^{\infty} \|a_l\|_B \mu(E_l) &= \sum_{k=1}^{\infty} \int_X \left\| \sum_{j=N_{k-1}}^{N_k-1} b_j \chi_{F_j}(x) \right\|_B d\mu(x) \\
&\leq \int_X \left\| \sum_{j=1}^{N_1-1} b_j \chi_{F_j}(x) \right\|_B d\mu(x) + \sum_{k=1}^{\infty} \sum_{j=N_k}^{\infty} \|b_j\|_X \mu(F_j) \\
&\leq \int_X \left\| f - \sum_{j=N_1}^{\infty} b_j \chi_{F_j}(x) \right\|_B d\mu(x) + \sum_{k=1}^{\infty} \sum_{j=N_k}^{\infty} \|b_j\|_X \mu(F_j) \\
&\leq \int_X \|f(x)\|_B d\mu(x) + 2 \sum_{k=1}^{\infty} \sum_{j=N_k}^{\infty} \|b_j\|_X \mu(F_j) \\
&\leq \int_X \|f(x)\|_B d\mu(x) + \varepsilon.
\end{aligned}$$

The number  $\varepsilon > 0$  being, arbitrary, this is the desired result.  $\square$

Bochner integral for strongly measurable functions. Keeping our observations above, we pass to the general case. We are going to define the integrable functions. In the  $\mathbb{R}$ -valued case we have allowed the integral to take the value  $\pm\infty$ . However, in the Bochner integral we rule out the possibility for the integral to diverge.

**Definition 27.8** (Bochner integrable functions). A function  $\varphi : X \rightarrow B$  is said to be Bochner integrable, if there exists a sequence of countably simple functions  $\{\varphi_j\}_{j \in \mathbb{N}}$  such that

$$(27.20) \quad \lim_{j \rightarrow \infty} \varphi_j(x) = \varphi(x)$$

for  $\mu$ -almost every  $x \in X$  and that

$$(27.21) \quad \lim_{j \rightarrow \infty} \int_X \|\varphi(x) - \varphi_j(x)\|_B d\mu(x) = 0.$$

If this is the case, define the integral of  $\varphi$  on  $X$  by

$$(27.22) \quad \int_X \varphi(x) d\mu(x) = \int_X \varphi d\mu := \lim_{j \rightarrow \infty} \int_X \varphi_j(x) d\mu(x).$$

Below denote by  $L^1(X; B)$  the set of all Bochner integrable functions.

**Lemma 27.9.** *The definition of the integral  $\int_X \varphi(x) d\mu(x)$  is makes sense, that is, in that the limit defining  $\int_X \varphi(x) d\mu(x)$  does exist and the element  $\int_X \varphi(x) d\mu(x) \in B$  does not depend on the choice of  $\{\varphi_j\}_{j \in \mathbb{N}}$ .*

*Proof.* The limit in (27.22) exists. To see this, we observe, if  $j, k \geq J$

$$\left\| \int_X \varphi_j(x) d\mu(x) - \int_X \varphi_k(x) d\mu(x) \right\|_B \leq 2 \sup_{l \geq J} \int_X \|\varphi(x) - \varphi_l(x)\|_B d\mu(x).$$

Therefore  $\left\{ \int_X \varphi_j(x) d\mu(x) \right\}_{j \in \mathbb{N}}$  is a Cauchy sequence in  $B$  and the limit therefore does exist.

The integral does not depend on the admissible representation of  $\varphi$ . Suppose  $\{\psi_j\}_{j \in \mathbb{N}}$  is another sequence of countably simple integrable functions such that

$$(27.23) \quad \lim_{j \rightarrow \infty} \psi_j = \varphi, \quad \lim_{j \rightarrow \infty} \int_X \|\varphi(x) - \psi_j(x)\|_B d\mu(x) = 0.$$

Then by the triangle inequality we obtain

$$\left\| \int_X \varphi_j(x) d\mu(x) - \int_X \psi_j(x) d\mu(x) \right\|_B \leq \int_X (\|\varphi(x) - \varphi_j(x)\|_B + \|\varphi(x) - \psi_j(x)\|_B) d\mu(x),$$

which tends to 0 as  $j \rightarrow \infty$ . Thus, both  $\left\{ \int_X \varphi_j(x) d\mu(x) \right\}_{j \in \mathbb{N}}$  and  $\left\{ \int_X \psi_j(x) d\mu(x) \right\}_{j \in \mathbb{N}}$  tend to the same limit. Therefore, the integral does not depend on the admissible representations of  $\varphi$ .  $\square$

**Lemma 27.10.** *Keep to the same setting as above. Let  $\varphi : X \rightarrow B$  be a Bochner integrable function. Then we have*

$$(27.24) \quad \left\| \int_X \varphi(x) d\mu(x) \right\|_B \leq \int_X \|\varphi(x)\|_B d\mu(x).$$

*Proof.* It is just a matter of passage to the limit of the countably simple function case. However, it is worth considering the case when  $B = \mathbb{C}$ . Writing out the statement in full with  $B = \mathbb{C}$ , we can easily imagine what is the crux of the proof of Lemma 27.10.  $\square$

**Theorem 27.11.** *The function  $\varphi : X \rightarrow B$  is Bochner integrable, if and only if  $\varphi$  is an integrable and countably simple function.*

*Proof.* Suppose that  $\varphi$  is an integrable countably simple function. Then we can take an integrable representation

$$(27.25) \quad \varphi \simeq \sum_{j=1}^{\infty} a_j \chi_{E_j}.$$

Set  $\varphi_j = \sum_{k=1}^j a_k \chi_{E_k}$ . Then each  $\varphi_j$  is an integrable countably simple function. By the Fatou lemma we have

$$(27.26) \quad \int_X \|\varphi(x) - \varphi_j(x)\|_B d\mu(x) \leq \int_X \left\| \sum_{k=j+1}^{\infty} a_k \chi_{E_k} \right\|_B d\mu(x) \leq \sum_{k=j+1}^{\infty} \|a_k\|_B \mu(E_k).$$

The most right-hand side tending to 0 as  $k \rightarrow \infty$ , so does the left-hand side. Therefore,  $\varphi$  is Bochner integrable.

Suppose that  $\varphi$  is Bochner integrable. Then there exists a sequence of countably simple integrable functions  $\{\varphi_j\}_{j \in \mathbb{N}}$  such that

$$(27.27) \quad \lim_{j \rightarrow \infty} \int_X \|\varphi(x) - \varphi_j(x)\|_B d\mu(x) = 0.$$

A passage to the subsequence, we can assume

$$(27.28) \quad \int_X \|\varphi(x) - \varphi_j(x)\|_B d\mu(x) < 2^{-j-1}.$$

In this case we have

$$(27.29) \quad \int_X \|\varphi_{j+1}(x) - \varphi_j(x)\|_B d\mu(x) < \frac{3}{4} \cdot 2^{-j}.$$

If we invoke Lemma 27.7, then we can find an integrable representation

$$(27.30) \quad \varphi_{j+1} - \varphi_j \simeq \sum_{k=1}^{\infty} a_{j+1,k} \chi_{E_{j+1,k}} \quad \text{with} \quad \sum_{k=1}^{\infty} \|a_{j,k}\|_B \mu(E_{j,k}) < 2^{-j}$$

for each  $j \in \mathbb{N}$ . We also have an integrable representation of  $\varphi_1$ :  $\varphi_1 \simeq \sum_{k=1}^{\infty} a_{1,k} \chi_{E_{1,k}}$ . Using these integrable representations, we see that  $\varphi$  has a following integrable representation

$$(27.31) \quad \varphi \simeq \sum_{j,k=1}^{\infty} a_{j,k} \chi_{E_{j,k}}.$$

Therefore  $\varphi$  is an integrable countably simple function as well.  $\square$

### 27.3. Convergence theorems.

Having set down the definition of the integrals, we now turn to the convergence theorems such as the dominated convergence theorem.

**Theorem 27.12.** *Suppose that  $\{\varphi_j\}_{j \in \mathbb{N}}$  is a sequence of strongly integrable functions such that  $\{\varphi_j(x)\}_{j \in \mathbb{N}}$  is a Cauchy sequence for almost every  $x \in X$ . Let  $\varphi$  be a weakly measurable function such that  $\varphi(x) = \lim_{j \rightarrow \infty} \varphi_j(x)$  for  $\mu$ -almost every  $x \in X$ . Assume in addition there exists  $\psi \in L^1(\mu; \mathbb{R})$  such that*

$$(27.32) \quad \|\varphi_j(x)\|_B \leq \psi(x)$$

for  $\mu$ -a.e.  $x \in X$ . Then  $\varphi \in L^1(X; B)$

$$(27.33) \quad \lim_{j \rightarrow \infty} \int_X \varphi_j d\mu = \int_X \varphi d\mu.$$

*Proof.*  $\varphi$  is strongly measurable. This follows from the fact that  $Y := \overline{\bigcup_{j=1}^{\infty} Y_j}$  is separable, whenever each  $Y_j$  is separable.

$\varphi \in L^1(X, B)$  and (27.33) holds. Since  $\varphi_j \in L^1(X, B)$ , by Theorem 27.11 we conclude that  $\varphi_j$  is an integrable countably simple function. Furthermore,  $\|\varphi(x) - \varphi_j(x)\|_B \leq 2\psi(x)$  for  $\mu$ -almost everywhere  $x \in X$ . Thus, if we invoke Lebesgue's convergence theorem, we obtain

$$(27.34) \quad \lim_{j \rightarrow \infty} \int_X \|\varphi(x) - \varphi_j(x)\|_B d\mu(x) = 0.$$

As a result  $\varphi \in L^1(X; B)$  and (27.33) holds.  $\square$

*Exercise 177.* Suppose that  $\{\varphi_j\}_{j \in \mathbb{N}} \in L^1(X; B)$  satisfies

$$(27.35) \quad \sum_{j=1}^{\infty} \int_X \|\varphi_j\|_B < \infty.$$

Then we have  $\sum_{j=1}^{\infty} \varphi_j(x)$  converges  $\mu$ -almost everywhere  $x \in X$ . Modifying a set of measure zero, we have  $\sum_{j=1}^{\infty} \varphi_j \in L^1(X; B)$  and

$$(27.36) \quad \int_X \left( \sum_{j=1}^{\infty} \varphi_j \right) d\mu = \sum_{j=1}^{\infty} \int_X \varphi_j d\mu.$$

**Theorem 27.13.** *The space  $L^1(X; B)$  is complete, once we identify the function having the same integrable representation.*

*Proof.* By Theorem 10.12, we have only to prove  $\sum_{j=1}^{\infty} \|\varphi_j\|_{L^1(B; X)}$  converges then so does  $\sum_{j=1}^{\infty} \varphi_j$ .

To do this we pick an integrable representation

$$(27.37) \quad \varphi_j \simeq \sum_{k=1}^{\infty} a_{j,k} \chi_{E_{j,k}} \text{ with } \sum_{k=1}^{\infty} \|a_{j,k}\|_B \mu(E_{j,k}) < \|\varphi_j\|_{L^1(B; X)} + 2^{-j}.$$

Then we define

$$(27.38) \quad \psi(x) := \sum_{j,k=1}^{\infty} a_{j,k} \chi_{E_{j,k}}(x),$$

provided the series converges and  $\psi(x) = 0$  otherwise. Then  $\psi$  has an integrable representation

$$(27.39) \quad \psi \simeq \sum_{j,k=1}^{\infty} a_{j,k} \chi_{E_{j,k}}.$$

Therefore  $\psi \in L^1(X; B)$ . Since  $\psi - \sum_{j=1}^l \varphi_j$  has an integrable representation

$$(27.40) \quad \psi - \sum_{j=1}^l \varphi_j \simeq \sum_{j=l+1}^{\infty} \sum_{k=1}^{\infty} a_{j,k} \chi_{E_{j,k}}(x),$$

we conclude  $\psi = \sum_{j=1}^{\infty} \varphi_j$ . □

Finally let us characterize the Bochner integral in terms of the dual spaces.

**Proposition 27.14.** *Let  $\varphi$  be a Bochner integrable function. Then  $b := \int_X \varphi(x) d\mu(x)$  is a unique element in  $B$  satisfying*

$$(27.41) \quad b^*(b) = \int_X b^*(\varphi(x)) d\mu(x)$$

for all  $b^* \in X^*$ .

*Proof.* Now that “Bochner integrable function” and “integrable countably simple function” are synonymous, this proposition is Lemma 27.5 itself. □

#### 27.4. Fubini's theorem for Bochner integral.

In this subsection to discuss Fubini's theorem let  $(X, \mathcal{M}, \mu)$  and  $(Y, \mathcal{N}, \nu)$  be  $\sigma$ -finite measure spaces.

**Theorem 27.15.** *Suppose that  $\varphi \in L^1(X \times Y; B)$ . Then we have the following.*

- (1) *For almost every  $x \in X$ , the function  $\varphi(x, *)$  belongs to  $L^1(Y; B)$ . For almost every  $y \in Y$ , the function  $\varphi(*, y)$  belongs to  $L^1(X; B)$ . Below it will be understood tacitly that  $\int_Y \varphi(x, y) d\nu(y) = 0$  and that  $\int_X \varphi(x, y) d\mu(x) = 0$ , if the integrand is not integrable.*
- (2) *We have*

$$\begin{aligned} \iint_{X \times Y} \varphi(x, y) d\mu \otimes \nu(x, y) &= \int_X \left( \int_Y \varphi(x, y) d\nu(y) \right) d\mu(x) \\ &= \int_Y \left( \int_X \varphi(x, y) d\mu(x) \right) d\nu(y). \end{aligned}$$

*Proof.* (1) is immediate because

$$(27.42) \quad \iint_{X \times Y} \|\varphi(x, y)\|_B d\mu \otimes \nu(x, y) < \infty.$$

Let us prove (2). Pick an integrable representation of  $\varphi$ .

$$(27.43) \quad \varphi \simeq \sum_{j=1}^{\infty} a_j \chi_{G_j},$$

where each  $G_j$  is a  $\mu \otimes \nu$ -measurable set. By disregarding a set of measure zero we may assume

$$(27.44) \quad \sum_{j=1}^{\infty} \|a_j\|_B \cdot \int_Y \chi_{G_j}(x, y) d\nu(y) < \infty$$

for all  $x \in X$ . Then for all  $x \in X$ ,

$$(27.45) \quad \int_Y \varphi(x, y) d\mu(y) = \sum_{j=1}^{\infty} a_j \int_Y \chi_{G_j}(x, y) d\nu(y).$$

Since

$$(27.46) \quad \sum_{j=1}^{\infty} \int_X \left\| a_j \int_Y \chi_{G_j}(x, y) d\nu(y) \right\|_B d\mu(x) = \sum_{j=1}^{\infty} \|a_j\|_B \cdot \int_Y \chi_{G_j}(x, y) d\nu(y) < \infty,$$

we are in the position of using the Lebesgue convergence theorem (Theorem 27.12) for The Bochner integral to obtain

$$\int_X \left( \int_Y \varphi(x, y) d\mu(y) \right) d\nu(x) = \sum_{j=1}^{\infty} a_j \int_{X \times Y} \chi_{G_j}(x, y) d\mu(x) d\nu(y) = \iint_{X \times Y} \varphi d\mu \otimes \nu.$$

This is the desired result. □

27.5.  $L^p(X; B)$ -spaces. Now we pass our theory of Lebesgue integral to the  $L^p(X; B)$ -space.

**Definition 27.16.** Let  $1 \leq p < \infty$ . Define  $L^p(X; B)$  as the set of all simple functions  $f : X \rightarrow B$  for which the norm

$$(27.47) \quad \|f\|_{L^p(X; B)} := \left( \int_X \|f\|_B^p d\mu \right)^{\frac{1}{p}} < \infty.$$



The space  $L^\infty(X; B)$  is the set of all simple functions satisfying

$$(27.48) \quad \|f\|_{L^\infty(X; B)} := \| \|f\|_B \|_{L^\infty(X)} < \infty.$$

*Exercise 178.* Let  $1 \leq p \leq \infty$ . Then show that  $L^p(X; B)$  is a Banach space.

Some duality results. Now we shall collect some duality result that is needed in this book.

**Theorem 27.17.** *Let  $B$  be a Banach space. Then the dual of  $L^1(\mathbb{R}; B)$  is canonically identified with  $L^\infty(\mathbb{R}; B^*)$ .*

(1) *If  $g \in L^\infty(\mathbb{R}; B^*)$ , then there exists a linear functional  $L_g$  such that*

$$L_g(\chi_A \cdot b) = \int_A \langle b, g(x) \rangle dx$$

*for all  $b \in B$  and measurable sets  $A$ .*

(2) *Let  $\varphi \in L^1(\mathbb{R}; B^*)$ . Then there exists  $g \in L^\infty(\mathbb{R}; B^*)$  such that  $L_g = \varphi$ .*

*Proof.* It is not so hard to see that  $L^\infty(\mathbb{R}; B^*)$  defines a continuous functional on  $L^1(\mathbb{R}; B)$ . Suppose that  $\varphi \in L^1(\mathbb{R}; B^*)$ . Then define  $\Psi : \mathbb{R} \rightarrow B^*$  by

$$\langle \Psi(t), b \rangle := \begin{cases} \varphi(\chi_{[0,t]} \cdot b) & t \geq 0 \\ -\varphi(\chi_{[t,0]} \cdot b) & t \leq 0. \end{cases} \quad (b \in B).$$

For  $-\infty < s < t < \infty$ , we note that

$$\langle \Psi(t), b \rangle - \langle \Psi(s), b \rangle = \varphi(\chi_{[s,t]} \cdot b).$$

Hence, denoting  $B_1$  by the unit ball of  $B$ , we obtain

$$(27.49) \quad \|\Psi(t) - \Psi(s)\|_* = \sup_{b \in B_1} |\langle \Psi(t), b \rangle - \langle \Psi(s), b \rangle| = \sup_{b \in B_1} |\varphi(\chi_{[s,t]} \cdot b)| \leq \|\varphi\|_* \cdot |t - s|.$$

Let us define

$$(27.50) \quad \Phi_j(t) := 2^j(\Psi(t + 2^{-j}) - \Psi(t)) \quad (t \in \mathbb{R}).$$

Then we have

$$(27.51) \quad \|\Phi_j(t)\|_{B^*} \leq \|\varphi\|_* \quad (t \in \mathbb{R})$$

by virtue of (27.49). Therefore, by the Banach Alaoglu theorem there exists a subsequence  $\{\Phi_{j_k}\}_{k \in \mathbb{N}}$  such that the weak-\* limit

$$(27.52) \quad \Phi(t) := \lim_{k \rightarrow \infty} \Phi_{j_k}(t)$$

exists for all  $t \in \mathbb{Q}$ . If  $t, s \in \mathbb{Q}$ , then we have

$$(27.53) \quad \|\Phi(t) - \Phi(s)\|_{B^*} = \sup_{b \in B} \frac{|\langle \Phi(t) - \Phi(s), b \rangle|}{\|b\|_B} \leq \|\varphi\|_* |t - s|.$$

Therefore,  $\Phi$  can be extended to a continuous function from  $\mathbb{R}$  to  $B^*$  satisfying

$$(27.54) \quad \|\Phi(t) - \Phi(s)\|_{B^*} \leq \|\varphi\|_* |t - s|.$$

This implies

$$\begin{aligned} |\langle \Phi(t) - \Phi_{j_k}(t), b \rangle| &\leq |\langle \Phi(t) - \Phi(s) - \Phi_{j_k}(t) + \Phi_{j_k}(s), b \rangle| + |\langle \Phi(s) - \Phi_{j_k}(s), b \rangle| \\ &\leq 2\|\varphi\|_* |t - s| + |\langle \Phi(s) - \Phi_{j_k}(s), b \rangle| \end{aligned}$$

for all  $t \in \mathbb{R}$  and  $s \in \mathbb{Q}$ . Letting  $j$  to  $\infty$ , then we obtain

$$(27.55) \quad \limsup_{j \rightarrow \infty} \|\Phi(t) - \Phi_{j_k}(t)\|_{B^*} = \limsup_{j \rightarrow \infty} \sup_{b \in B_1} |\langle \Phi(t) - \Phi_{j_k}(t), b \rangle| \leq 2\|\varphi\|_* |t - s|.$$

Since  $s \in \mathbb{Q}$  is arbitrary, this implies  $\lim_{k \rightarrow \infty} \Phi_{j_k}(t) = \Phi(t)$  for all  $t \in \mathbb{R}$ . As a result, for each  $g \in L_c^\infty(\mathbb{R}^d)(\mathbb{R}; B)$ ,

$$(27.56) \quad \int_{\mathbb{R}} \langle \Phi(t), g(t) \rangle dt = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} \langle \Phi_{j_k}(t), g(t) \rangle dt = \varphi(g).$$

Since  $L_c^\infty(\mathbb{R}^d)(\mathbb{R}; B)$  is dense in  $L^1(\mathbb{R}; B)$ , we see that  $\varphi$  is realized by  $\Phi$ .  $\square$

We conclude this section with another duality result. The proof is similar to Theorem 27.17 and we omit the proof.

**Theorem 27.18.** *Suppose that  $B$  is reflexive and that  $1 \leq p < \infty$ . Then*

$$(27.57) \quad L^p(X; B)^* = L^{p'}(X; B^*).$$

## 28. SEMIGROUPS

### 28.1. Continuous semigroups.

In this book, in order that we concentrate on concrete examples we do not go into the details of the semigroup. Our aim is modest. Our final aim is to consider the heat semigroup. This semigroup enjoys nice properties such as continuity and so on. Thus, we content ourselves with presenting a definition of the nice class of semigroups.

**Definition 28.1** (Semigroup). Let  $X$  be a Banach space. A family of operators  $\{T(t)\}_{t \geq 0} \subset B(X)$  is said to be a (continuous) semigroup on  $X$  if it satisfies the following.

- (1)  $T(0) = \text{id}_X$ .
- (2) The family  $\{T(t)\}_{t \geq 0}$  is strongly continuous, that is,  $t \in [0, \infty) \mapsto T(t)x \in X$  is continuous for all  $x \in X$ .
- (3) The family  $\{T(t)\}_{t \geq 0}$  satisfies the semigroup property:

$$(28.1) \quad T(t+s) = T(t)T(s) \quad (t, s \geq 0).$$

*Exercise 179.* Use the uniformly bounded principle to show that there exists  $M > 0$  such that  $\|T(t)\|_{B(X)} \lesssim e^{Mt}$ .

**Definition 28.2** (Generator of semigroup). Let  $X$  be a Banach space and  $\{T(t)\}_{t \geq 0}$  a semigroup on  $X$ . The generator  $A$  of  $\{T(t)\}_{t \geq 0}$  is a linear operator given by

$$(28.2) \quad D(A) := \left\{ x \in X : \lim_{t \downarrow 0} \frac{T(t)x - x}{t} \text{ exists} \right\}, \quad Ax := \lim_{t \downarrow 0} \frac{T(t)x - x}{t} \text{ for } x \in D(A).$$

**Proposition 28.3.** *The generator  $A$  of a continuous semigroup is closed.*

*Proof.* If  $x \in D(A)$ , then the function

$$(28.3) \quad t \in [0, \infty) \mapsto \frac{d}{dt} T(t)x = \lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{T(t+h) - T(t)}{h} x = T(t) \lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{T(h) - 1}{h} x = T(t)Ax$$

is continuous. Here it will be understood that  $\left. \frac{d}{dt} T(t)x \right|_{t=0} = \lim_{h \downarrow 0} \frac{T(h) - 1}{h} x$ . Therefore, we obtain

$$(28.4) \quad (T(h) - 1)x = \int_0^h T(s)Ax ds, \quad x \in D(A).$$

Suppose that  $x_1, x_2, \dots \in D(A)$  satisfy  $\lim_{j \rightarrow \infty} x_j =: x$  and  $\lim_{j \rightarrow \infty} Ax_j = y$ . Then from (28.4), we see

$$(28.5) \quad (T(h) - 1)x_j = \int_0^h T(s)Ax_j ds.$$

A passage to the limit therefore gives

$$(28.6) \quad (T(h) - 1)x = \int_0^h T(s)y ds.$$

Thus, it follows that

$$(28.7) \quad y = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h T(s)y ds = \lim_{h \rightarrow 0} \frac{T(h) - 1}{h}x.$$

This equation proves  $x \in D(A)$  and  $y = Ax$ . □

### 28.2. Sectorial operators and semigroups.

Now we turn to the class of semigroup whose generator is nice. That is, the class of the sectorial operators contains important closed operators such as Laplacian. Therefore, we are still able to obtain results of interest if we restrict to sectorial operators.

Semigroup generated by a sectorial operator. As we mentioned before, our aim is to view a theory of heat semigroups, which has a nice property. To formulate this property, we give a definition of the sectorial operators.

**Definition 28.4** (Sectorial operator). A densely defined closed operator  $A$  is said to be sectorial, if there exist  $M \in \mathbb{R}$ ,  $\omega \in \mathbb{R}$  and  $\frac{\pi}{2} < \eta < \pi$  with the following properties.

- (1)  $\rho(A) \supset S_{\omega, \eta} := \{z \in \mathbb{C} : z \neq \omega, |\arg(z - \omega)| < \eta\}$ . Here  $\arg$  is a branch taking  $(-\pi, \pi)$ .
- (2)  $\|R(\lambda)\|_X \leq \frac{M}{|\lambda - \omega|}$  for all  $\lambda \notin S_{\omega, \eta}$ .

**Example 28.5.** As it will turn out, we see that the Laplacian is a sectorial operator.

*Exercise 180.* Show that any element in  $B(X)$  is sectorial operator.

**Definition 28.6** (Exponential of the sectorial operator). Given a sectorial operator  $A$  as in Definition 28.4, the exponential  $e^{tA}$  is defined by

$$(28.8) \quad e^{tA} := \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r}} e^{tz}(z - A)^{-1} dz.$$

Here,  $\gamma_{\omega, \theta, r}$  is a path given by

$$\gamma_{\omega, \theta, r}(t) = \begin{cases} \omega + (t + r)e^{-i\theta} & t \leq -r \\ \omega + re^{i\frac{t\theta}{r}} & -r \leq t \leq r \\ \omega + (t - r)e^{-i\theta} & t \geq r \end{cases}$$

and  $\frac{\pi}{2} < \theta < \eta$  and  $r > 0$ . Furthermore  $z$  lies in the right side of the domain defined by  $\gamma_{\omega, \theta, r}$ .

**Proposition 28.7.** In Definition 28.6, the integral defining  $e^{tA}$  converges absolutely and the definition of  $e^{tA}$  does not depend on admissible choices of  $\theta$  and  $r$ .

*Proof.* The integral converges absolutely. Set

$$\begin{aligned}\gamma_{1,\omega,\theta,r}(t) &:= \omega + t e^{-i\theta}, \quad t \leq -r \\ \gamma_{2,\omega,\theta,r}(t) &:= \omega + r e^{i\frac{\theta}{r}}, \quad -r \leq t \leq r \\ \gamma_{3,\omega,\theta,r}(t) &:= \omega + t e^{-i\theta}, \quad t \geq r\end{aligned}$$

From the definition of integral, we have

$$\int_{\gamma_3} e^{tz}(z-A)^{-1} dz = \int_r^\infty e^{t(\omega+s e^{-i\theta})} (A-\omega-s e^{-i\theta})^{-1} e^{i\theta} ds.$$

Thus, the matter is to prove

$$(28.9) \quad \int_r^\infty \left\| e^{t(\omega+s e^{-i\theta})} (A-\omega-s e^{-i\theta})^{-1} \right\|_{B(X)} ds < \infty.$$

Since  $A$  is sectorial, we have

$$(28.10) \quad \left\| e^{t(\omega+s e^{-i\theta})} (A-\omega-s e^{-i\theta})^{-1} \right\|_{B(X)} \leq \frac{M e^{t\omega+t s \cos \theta}}{|s|}.$$

Since  $\frac{\pi}{2} < \theta < \eta < \pi$ , we conclude (28.9) holds. A similar argument works for  $\gamma_2, \gamma_3$  and therefore the integral converges absolutely.

$e^{tA}$  does not depend on the path.

It is clear that the definition does not depend on  $r > 0$  by virtue of the Cauchy integral theorem. Let  $\frac{\pi}{2} < \theta < \theta' < \eta$ . Then let  $\ell_{1,R}$  be an oriented line segment starting from  $\gamma_{1,\omega,\theta',r}(R)$  and ending at  $\gamma_{1,\omega,\theta,r}(R)$ . Then let  $\ell_{3,R}$  be an oriented line segment starting from  $\gamma_{3,\omega,\theta',r}(R)$  and ending at  $\gamma_{3,\omega,\theta,r}(R)$ . Then

$$(28.11) \quad \int_{\gamma_{\omega,\theta,r}[-R,R]} - \int_{\gamma_{\omega,\theta',r}[-R,R]} e^{tz}(z-A)^{-1} dz = \int_{\ell_{1,R}} + \int_{\ell_{3,R}} e^{tz}(z-A)^{-1} dz.$$

On  $\ell_{1,R} \cup \ell_{3,R}$  the integrand is bounded by  $c R^{-1} e^{tR \cos \theta'}$ . Therefore, we see

$$(28.12) \quad \lim_{R \rightarrow \infty} \int_{\ell_{1,R}} + \int_{\ell_{3,R}} e^{tz}(z-A)^{-1} dz = 0.$$

As a result, we see for all  $\frac{\pi}{2} < \theta' < \pi$ ,

$$\begin{aligned}\int_{\gamma_{\omega,\theta,r}} e^{tz}(z-A)^{-1} dz &= \frac{1}{2\pi i} \lim_{R \rightarrow \infty} \int_{\gamma_{\omega,\theta,r}[-R,R]} e^{tz}(z-A)^{-1} dz \\ &= \frac{1}{2\pi i} \lim_{R \rightarrow \infty} \int_{\gamma_{\omega,\theta',r}[-R,R]} e^{tz}(z-A)^{-1} dz \\ &= \int_{\gamma_{\omega,\theta',r}} e^{tz}(z-A)^{-1} dz.\end{aligned}$$

Therefore, the integral defining  $e^{tA}$  does not depend on the path. □

*Exercise 181.* Suppose that  $A$  is a sectorial operator. Use the resolvent equation to show that  $e^{tA}$  and  $(A-\lambda)^{-1}$  commute whenever  $\lambda \in \rho(A)$ .

The case when  $A$  is bounded. If  $A$  is bounded, we have now two plausible definitions. One is to regard  $A$  as a sectorial operator. The other is to expand the exponential, as we did in defining the exponential of matrices.

**Proposition 28.8.** *If  $A \in B(X)$ , then  $e^{tA} \in B(X)$  and it can be expanded into the Taylor series:  $e^{tA} := \sum_{j=0}^{\infty} \frac{1}{j!} A^j$ .*

*Proof.* Proceeding in the same way as before, we see

$$(28.13) \quad e^{tA} := \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r}[-R, R]} e^{tz}(z - A)^{-1} dz := \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r, R}} e^{tz}(z - A)^{-1} dz,$$

where  $\gamma_{\omega, \theta, r, R}$  is a curve obtained by connecting end points of  $\gamma_{\omega, \theta, r}[-R, R]$ . Since  $A$  is assumed bounded, we see that

$$(28.14) \quad e^{tA} := \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r, R}} e^{tz}(z - A)^{-1} dz,$$

provided  $R$  is large enough. It remains to expand  $e^{tz}$  into a Taylor series. □

If  $A$  is a sectorial operator, the boundedness of  $A$  is equivalent to the compactness of  $\sigma(A)$ .

**Proposition 28.9.** *Suppose that  $A$  is a sectorial operator. Then  $A$  is bounded, if and only if  $\sigma(A)$  is compact.*

*Proof.* By the Neumann expansion it is trivial that  $\sigma(A)$  is compact whenever  $A$  is bounded.

Suppose that  $\sigma(A)$  is compact. Then under the same notation as before we have

$$(28.15) \quad e^{tA} = \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r, R}} e^{tz}(A - z)^{-1} dz,$$

provided  $R$  is sufficiently large. Differentiation yields

$$(28.16) \quad A e^{tA} = \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r, R}} z e^{tz}(A - z)^{-1} dz.$$

Let  $x \in X$ . Then there exists a sequence  $\{x_j\}_{j \in \mathbb{N}} \subset D(A)$  such that  $\lim_{j \rightarrow \infty} x_j = x$ . Observe that

$$(28.17) \quad \|A(x_j - x_k) : X\| = \left\| \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r, R}} z e^{tz}(A - z)^{-1}(x_j - x_k) dz : X \right\| \lesssim \|x_j - x_k\|_X.$$

Therefore  $\lim_{j \rightarrow \infty} A x_j$  exists. As a result  $x \in D(A)$ . Since  $x \in X$  is chosen arbitrarily, we conclude that  $D(A) = X$  and hence  $A$  is a bounded operator. □

Regularizing effect. Let us return to the setting that  $A$  is a sectorial operator. In this paragraph we consider regularizing effect of the semigroup  $\{e^{tA}\}_{t \geq 0}$ . We considered the heat semigroup before. Let  $t > 0$ . Then define  $E(x, t) := \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{4t}\right)$ . Then the semigroup

$$(28.18) \quad f \in \mathcal{S}'(\mathbb{R}^d) \rightarrow E(\cdot, t) * f \in \mathcal{S}'(\mathbb{R}^d)$$

has some regularizing effect. That is the solution of the heat equation becomes smooth as soon as the time passes even only a little. This applies for the semigroup  $e^{tA}$ . Speaking precisely we have the following. To formulate the result, let us define inductively

$$(28.19) \quad D(A^{k+1}) := \{x \in D(A^k) : Ax \in D(A^k)\}, \quad A^{k+1}x = A^k(Ax) \quad x \in D(A^{k+1})$$

for  $k \in \mathbb{N}$ .

**Proposition 28.10.** *If  $x \in X$ , then  $e^{tA}x \in D(A^k)$  for each  $k \in \mathbb{N}$  and  $t > 0$ . Furthermore*

$$(28.20) \quad Ae^{tA} = \frac{1}{2\pi i} \int_{\gamma_{\omega, \theta, r}} e^{tz} A(z - A)^{-1} dz.$$

*In particular, if  $x \in D(A)$ , then  $Ae^{tA}x = e^{tA}Ax$ .*

*Proof.* We begin with proving  $e^{tA}x \in D(A^k)$  for each  $k \in \mathbb{N}$ . The semigroup property  $e^{tA} = (e^{\frac{t}{k}A})^k$  and an induction argument reduces to the matter to showing the case that  $k = 1$ . It suffices to prove

$$(28.21) \quad \int_{\gamma_{\omega, \theta, r}} e^{tz} (z - A)^{-1} dz \in D(A) \text{ and } A \left( \int_{\gamma_{\omega, \theta, r}} e^{tz} (z - A)^{-1} dz \right) = \int_{\gamma_{\omega, \theta, r}} e^{tz} A(z - A)^{-1} dz.$$

Indeed,

$$\begin{aligned} & \left( \int_{\gamma_{\omega, \theta, r}} e^{tz} (z - A)^{-1} dz \right) \\ &= \lim_{j \rightarrow \infty} \sum_{k=-2^j}^{2^j-1} e^{t\gamma_{\omega, \theta, r}(2^{-j}k)} \gamma_{\omega, \theta, r}(2^{-j}k) - A)^{-1} (\gamma_{\omega, \theta, r}(2^{-j}(k+1)) - \gamma_{\omega, \theta, r}(2^{-j}k)). \end{aligned}$$

Since  $A(z - A)^{-1} = 1 - z(z - A)^{-1}$  is continuous,

$$\begin{aligned} & \lim_{j \rightarrow \infty} A \left( \sum_{k=-2^j}^{2^j-1} e^{t\gamma_{\omega, \theta, r}(2^{-j}k)} \gamma_{\omega, \theta, r}(2^{-j}k) - A)^{-1} (\gamma_{\omega, \theta, r}(2^{-j}(k+1)) - \gamma_{\omega, \theta, r}(2^{-j}k)) \right) \\ &= \lim_{j \rightarrow \infty} \sum_{k=-2^j}^{2^j-1} e^{t\gamma_{\omega, \theta, r}(2^{-j}k)} \gamma_{\omega, \theta, r} A(2^{-j}k) - A)^{-1} (\gamma_{\omega, \theta, r}(2^{-j}(k+1)) - \gamma_{\omega, \theta, r}(2^{-j}k)) \\ &= \int_{\gamma_{\omega, \theta, r}} e^{tz} A(z - A)^{-1} dz \end{aligned}$$

converges. Therefore, (28.21) holds.  $\square$

**Theorem 28.11.** *If  $A$  is a sectorial operator described in Definition 28.4, then, for each  $k \in \mathbb{N}_0$ ,*

$$(28.22) \quad \|t^k(A - \omega)^k e^{tA}; B(X)\| \lesssim_k e^{\omega t}$$

*for all  $t > 0$ .*

*Proof.* Suppose  $k = 0$ . Recall that

$$(28.23) \quad \|e^A\|_X \lesssim M_0,$$

in Proposition 28.10, where  $M_0$  is a constant in Definition 28.4. Then if we set  $B = A - \omega$ , then  $B$  is a sectorial operator and

$$(28.24) \quad e^{tB} = \frac{1}{2\pi i} \int_{\gamma_{0, \theta, r}} e^{tz} (z - B)^{-1} dz = \frac{1}{2\pi i} \int_{\gamma_{0, \theta, r}} e^{tz} (z + \omega - A)^{-1} dz = e^{-t\omega} e^{tA}.$$

Therefore we can reduce the matter to the case when  $\omega = 0$ . In this case we have

$$(28.25) \quad e^{tA} = \frac{1}{2\pi i} \int_{\gamma_{0, \theta, tr}} e^{tz} (z - A)^{-1} dz = \frac{1}{2\pi i t} \int_{\gamma_{0, \theta, r}} e^w (w - t^{-1}A)^{-1} dw$$

By (28.23) with  $A$  replaced by  $t^{-1}A$  and  $M_0$  by  $M_0/t$ , we obtain

$$(28.26) \quad \|e^{tA}\|_X \lesssim M_0.$$

Therefore, the case when  $k = 0$  is proved.

Assume  $k \in \mathbb{N}$ . Similarly we can prove the case when  $k = 1$ . Assume  $k \geq 2$ . Then we have

$$\|t^k(A - \omega)^k e^{tA}\|_{B(X)} \leq k^k \left\| \frac{t}{k}(A - \omega)e^{\frac{tA}{k}} \right\|_{B(X)}^k \leq (kM_1)^k.$$

Therefore, the case when  $k \geq 2$  was established, too. Here we used the fact that  $k!e^k = \sum_{j=0}^{\infty} \frac{k!}{j!} k^j \geq k^k$ . □

Semigroup property. Let us see that  $\{e^{tA}\}_{t \geq 0}$  defined by means of the complex line integral is a semigroup defined in the beginning of this section.

**Theorem 28.12.** *If  $A$  is a sectorial operator, then  $\{e^{tA}\}_{t \geq 0}$  is a semigroup generated by  $A$ .*

*Proof.*  $e^{(t+s)A} = e^{tA}e^{sA}$ . Suppose that  $\gamma_{\omega, \theta_0, r_0}$  lies in the left side of  $\gamma_{\omega, \theta_1, r_1}$ . We write

$$\begin{aligned} e^{tA}e^{sA} &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_0, r_0}} e^{tz}(z - A)^{-1} dz \int_{\gamma_{\omega, \theta_1, r_1}} e^{sw}(w - A)^{-1} dw \\ &= \frac{1}{(2\pi i)^2} \iint_{\gamma_{\omega, \theta_0, r_0} \times \gamma_{\omega, \theta_1, r_1}} e^{sw+tz}(z - A)^{-1}(w - A)^{-1} dz dw \\ &= \frac{1}{(2\pi i)^2} \iint_{\gamma_{\omega, \theta_0, r_0} \times \gamma_{\omega, \theta_1, r_1}} \frac{e^{sw+tz}}{w - z} ((z - A)^{-1} - (w - A)^{-1}) dz dw. \end{aligned}$$

Here we used the Fubini theorem for the second inequality and the resolvent equation for the third inequality. In view of this we set

$$\begin{aligned} I_1 &:= \frac{1}{(2\pi i)^2} \iint_{\gamma_{\omega, \theta_0, r_0} \times \gamma_{\omega, \theta_1, r_1}} \frac{e^{sw+tz}}{w - z} (z - A)^{-1} dz dw \\ I_2 &:= \iint_{\gamma_{\omega, \theta_0, r_0} \times \gamma_{\omega, \theta_1, r_1}} \frac{e^{sw+tz}}{w - z} (w - A)^{-1} dz dw. \end{aligned}$$

Then, by virtue of the Cauchy integral theorem, we obtain

$$\begin{aligned} I_1 &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_0, r_0}} \left( \int_{\gamma_{\omega, \theta_1, r_1}} \frac{e^{sw}}{w - z} dw \right) e^{tz}(z - A)^{-1} dz \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_0, r_0}} \lim_{R \rightarrow \infty} \left( \int_{\gamma_{\omega, \theta_1, r_1, R}} \frac{e^{sw}}{w - z} dw \right) e^{tz}(z - A)^{-1} dz \\ &= \frac{1}{(2\pi i)} \int_{\gamma_{\omega, \theta_0, r_0}} e^{(t+s)z}(z - A)^{-1} dz, \\ I_2 &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_1, r_1}} \left( \int_{\gamma_{\omega, \theta_0, r_0}} \frac{e^{tz}}{w - z} dz \right) e^{sw}(w - A)^{-1} dw \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_1, r_1}} \lim_{R \rightarrow \infty} \left( \int_{\gamma_{\omega, \theta_0, r_0, R}} \frac{e^{tz}}{w - z} dz \right) e^{sw}(w - A)^{-1} dw \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\omega, \theta_1, r_1}} 0 \cdot e^{sw}(w - A)^{-1} dw = 0. \end{aligned}$$

This is the desired result.

$\lim_{t \rightarrow 0} e^{tA}x = x$  for all  $x \in X$ . By the Banach Steinhaus principle and (28.22) we can assume  $x \in D(A)$ . Then, going through the same argument as (28.4), we obtain

$$(28.27) \quad e^{tA}x - x = \int_0^t e^{sA}Ax \, ds.$$

Since if  $t \leq 1$ , then the integrand  $e^{sA}Ax$  remains bounded. Therefore, if we pass to the limit, then we see  $x = \lim_{t \downarrow 0} e^{tA}x$ .

The generator of  $e^{tA}$  is  $A$ . Denote by  $B$  the generator of  $e^{tA}$ .

Suppose that  $x \in D(A)$ . Then again by (28.4) we have

$$(28.28) \quad Ax = \lim_{t \downarrow 0} \frac{e^{tA} - 1}{t}x.$$

Therefore  $x \in D(B)$  and hence  $A \subset B$ .

Suppose that  $x \in D(B)$ . Then

$$(28.29) \quad y = \lim_{t \downarrow 0} \frac{e^{tA} - 1}{t}x$$

exists. Let  $z \in \rho(A) \setminus \{1\}$ . By multiplying  $(z - A)^{-1}$ , we obtain

$$(28.30) \quad (z - A)^{-1}y = \lim_{t \downarrow 0} \frac{e^{tA} - 1}{t}(z - A)^{-1}x.$$

Since  $(z - A)^{-1}x \in D(A)$ , we obtain

$$(28.31) \quad (z - A)^{-1}y = \lim_{t \downarrow 0} \frac{e^{tA} - 1}{t}(z - A)^{-1}x = A(z - A)^{-1}x.$$

Therefore,

$$(28.32) \quad x = z(z - A)^{-1}x - (z - A)^{-1}y \in D(A).$$

This implies along with  $A \subset B$  that  $A = B$ . □

## 29. BANACH AND $C^*$ -ALGEBRA

In this section we assume that the coefficient field  $\mathbb{K}$  is  $\mathbb{C}$ .

### 29.1. Banach algebras.

**Definition.** Banach algebras carry a structure similar to  $B(X)$ . Recalling the property of  $B(X)$ , let us present the definition.

#### **Definition 29.1.**

- (1) A Banach algebra  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  is a  $\mathbb{C}$ -Banach space  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  which comes with an additional operation  $\cdot : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$  satisfying

$$(29.1) \quad \|x \cdot y\|_{\mathcal{A}} \leq \|x\|_{\mathcal{A}} \cdot \|y\|_{\mathcal{A}}$$

for all  $x, y \in \mathcal{A}$ . The operation  $\cdot$  is said to be multiplication.

- (2) A Banach algebra  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  is said to be commutative

$$(29.2) \quad a \cdot b = b \cdot a$$

for all  $a, b \in \mathcal{A}$ .



- (3) A Banach algebra  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  is said to be unital, if there exists an element  $e$  called neutral such that

$$(29.3) \quad e \cdot a = a \cdot e = a$$

for all  $a \in \mathcal{A}$ .

**Example 29.2.** Below we present some examples of commutative and noncommutative Banach algebras.

- (1)  $\mathbb{C}$  is a commutative Banach algebra.
- (2)  $B(X)$  is a noncommutative Banach algebra, where  $X$  is a Banach space with dimension larger than 2.
- (3)  $L^1(\mathbb{R}^d)$  is a commutative Banach algebra whose multiplication is given by convolution.
- (4)  $L^\infty(\mathbb{R}^d)$  is a commutative Banach algebra whose multiplication is given by pointwise multiplication.

*Exercise 182.* Concerning to Example 29.2, prove the following.

- (1) Let  $X$  be a Banach space whose dimension is greater than 2. Then  $B(X)$  is unital but not commutative.
- (2)  $L^1(\mathbb{R}^d)$  is commutative but not unital.
- (3)  $L^\infty(\mathbb{R}^d)$  is commutative and unital.

*Exercise 183.* Let  $(X, \mathcal{B}, \mu)$  be a  $\sigma$ -finite space. For a measurable function  $K : X \times X \rightarrow \mathbb{C}$ , we define

$$(29.4) \quad \|K\|_{\mathcal{A}} = \max \left\{ \operatorname{esssup}_{x \in X} \int_X |K(x, y)| d\mu(y), \operatorname{esssup}_{y \in X} \int_X |K(x, y)| d\mu(x) \right\}.$$

The linear space  $\mathcal{A}$  is the set of all measurable functions  $K : X \times X \rightarrow \mathbb{C}$  such that  $\|K\|_{\mathcal{A}}$  is finite.

- (1) Show that  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  is a Banach algebra. Here we define the multiplication by

$$(K_1 \cdot K_2)(x, y) = \int_X K_1(x, z) K_2(z, y) d\mu(z)$$

for  $K_1, K_2 \in \mathcal{A}$ .

- (2) Show that, for each  $K \in \mathcal{A}$ ,

$$(29.5) \quad I_K : f \in L^p(\mu) \mapsto \int_X K(x, y) f(y) d\mu(y)$$

is a bounded linear mapping. Show also that  $\|I_K\|_{B(L^p(\mu))} \leq \|K\|_{\mathcal{A}}$  by means of complex interpolation (if necessary).

- (3) Let  $p = 2$ . Show that the adjoint of  $I_K \in B(L^2(\mu))$  is given by  $I_{K^*}$ , where we defined  $K^*(x, y) = \overline{K(y, x)}$  for  $x, y \in X$ .

*Exercise 184.* Keep to the same notation as Exercise 183. Suppose that we are given a positive measurable function  $m : X \times X \rightarrow [1, \infty)$  satisfying

$$(29.6) \quad m(x, y) \leq m(x, z)m(z, y), \quad m(x, y) = m(y, x), \quad m(x, x) \leq C$$

for all  $x, y, z \in X$ . Define

$$\|K\|_{\mathcal{A}_m} = \|Km\|_{\mathcal{A}}$$

for a measurable function  $K : X \times X \rightarrow \mathbb{C}$ .

- (1) Show that  $\mathcal{A}_m$  is a Banach algebra, where we define the multiplication by

$$(K_1 \cdot K_2)(x, y) = \int_X K_1(x, z) K_2(z, y) d\mu(z)$$

for  $K_1, K_2 \in \mathcal{A}$ .

(2) Show that  $\mathcal{A}_m$  is a Banach- $*$ -algebra by defining the adjoint suitably.

(3) Let  $v : X \rightarrow (0, \infty)$  be a measurable function. Define

$$(29.7) \quad m(x, y) := \max \left\{ \frac{v(x)}{v(y)}, \frac{v(y)}{v(x)} \right\}$$

for  $x, y \in X$ . Then show that  $m$  satisfies (29.6) above.

(4) If  $m$  is given by (29.7), then  $I_K$  maps  $L^p(w^p)$  into itself continuously.

**Definition 29.3.** Let  $\mathcal{A}$  and  $\mathcal{B}$  be Banach algebras. Then define  $\text{Hom}(\mathcal{A}, \mathcal{B})$  as the set of all  $\mathbb{C}$ -linear mappings  $\varphi$  satisfying  $\varphi(a_0 a_1) = \varphi(a_0) \varphi(a_1)$  for all  $a_0, a_1 \in \mathcal{A}$ .

Unitalization of nonunital Banach algebra. If  $\mathcal{A}$  is not unital, we can unitalize  $\mathcal{A}$  as follows:

**Definition 29.4.** Assume that  $\mathcal{A}$  is not unital. Then define

$$(29.8) \quad \mathcal{A}_0 := \mathcal{A} \times \mathbb{C}.$$

Equip  $\mathcal{A}_0$  with operations as follows:

$$\begin{aligned} (a_0, k_0) + (a_1, k_1) &:= (a_0 + a_1, k_0 + k_1) \\ (a_0, k_0) \cdot (a_1, k_1) &:= (a_0 a_1 + k_0 \cdot a_1 + k_1 \cdot a_0, k_0 k_1). \end{aligned}$$

$\mathcal{A}_0$  is called unitalization of  $\mathcal{A}$ . The norm of  $\mathcal{A}_0$  is defined as

$$(29.9) \quad \|(a_0, k_0)\|_{\mathcal{A}_0} := \|a_0\|_{\mathcal{A}} + |k_0|.$$

$\mathcal{A}_0$  enjoys the following properties. The proofs are left as an exercise because it is all straightforward.

**Theorem 29.5.** Suppose that  $\mathcal{A}$  is a Banach algebra which is not unital.

- (1) The Banach algebra  $\mathcal{A}_0$  is a unital Banach algebra whose neutral element is  $1 := (0, 1)$ .
- (2) Define  $\varphi(a) = (a, 0)$  for  $a \in \mathcal{A}$ . Then we have  $\varphi \in \text{Hom}(\mathcal{A}, \mathcal{A}_0)$ .
- (3)  $\mathcal{A}_0$  is commutative, if  $\mathcal{A}$  is commutative.

*Exercise 185.* Prove Theorem 29.5.

Ideal and quotient. As we did for normed spaces, it is possible that we define the quotient of Banach algebras.

**Definition 29.6** (Ideal). Let  $\mathcal{A}$  be a Banach algebra. A subspace  $I$  of  $\mathcal{A}$  is said to be a two-sided ideal or for short an ideal, if  $a \cdot b, b \cdot a \in I$  whenever  $a \in \mathcal{A}$  and  $b \in I$ .

**Example 29.7.** Equip  $\text{BC}(\mathbb{R}^d)$  with the pointwise multiplication. Then the space  $\text{BC}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$  is an ideal of  $\text{BC}(\mathbb{R}^d)$ , which can be verified by using the Lebesgue convergence theorem.

*Exercise 186.* Let  $X$  be a Banach space. Then prove that  $K(X)$ , the set of all compact operators, is an ideal of  $B(X)$ .

**Definition 29.8.** Let  $\mathcal{A}$  be a Banach algebra and  $I$  an ideal.

- (1)  $I$  is said to be maximal, if there is no ideal  $J$  such that  $I \subsetneq J \subsetneq \mathcal{A}$ .
- (2) Let  $a_0 \in \mathcal{A}$ . The ideal  $[I, a_0]$  denotes the smallest ideal containing  $a_0$  and  $I$ .

*Exercise 187.* Keep to the same notation above. Show the following:

- (1)  $[I, a] = \{b + b'a_0 : b \in I, b' \in \mathcal{A}\}$ .
- (2)  $I$  is maximal precisely when  $[I, a] = \mathcal{A}$  for all  $a \notin I$ .

**Definition 29.9.** Let  $I$  be a closed ideal. Then  $\mathcal{A}/I$  is a Banach algebra equipped with a structure of the Banach algebra. That is, define

$$(29.10) \quad \mathcal{A}/I := \{a + I : a \in \mathcal{A}\}$$

as a set and define addition and multiplication

$$(29.11) \quad (a + I) + (b + I) := (a + b) + I, (a + I)(b + I) := ab + I$$

The norm is the quotient norm as a normed space.

As we have seen before, the multiplication does not depend on the choice of the representative of  $(a + I)$ .

*Exercise 188.* As for the definition of  $\mathcal{A}/I$  prove the following.

- (1) The multiplication is independent of the choice of the representative of  $(a + I)$  and  $(b + I)$ .
- (2)  $\|(a + I)(b + I)\|_{\mathcal{A}/I} \leq \|a + I\|_{\mathcal{A}/I} \cdot \|b + I\|_{\mathcal{A}/I}$ .
- (3)  $\mathcal{A}/I$  is unital, if  $\mathcal{A}$  is unital.  $\mathcal{A}/I$  is commutative, if  $\mathcal{A}$  is commutative.
- (4) If  $I$  is maximal, then for all  $a \in \mathcal{A} \setminus I$  there exists  $b \in \mathcal{A} \setminus I$  such that  $ab - 1 \in I$ . Therefore, if  $I$  is maximal, then  $\mathcal{A}/I$  is a field, that is,  $(\mathcal{A}/I)^* = \mathcal{A} \setminus \{0\}$ .

Invertible elements and spectrum.

Having seen how to define invertible elements for  $B(X)$ , we are readily to generalize the definitions.

**Definition 29.10.** Assume that  $\mathcal{A}$  is a unital Banach algebra with unit  $e_{\mathcal{A}}$ . Let  $a \in \mathcal{A}$ .

- (1)  $a$  is said to be invertible, if there exists  $b \in \mathcal{A}$  such that  $ba = ab = e_{\mathcal{A}}$ . Below define  $\mathcal{A}^{\times}$  as the set of all invertible elements in  $\mathcal{A}$ .
- (2) For  $\lambda \in \mathbb{C}$ , identify  $\lambda$  with  $\lambda e_{\mathcal{A}} \in \mathcal{A}$ .

Here and below we abbreviate  $a e_{\mathcal{A}}$  simply to  $a$ .

**Definition 29.11** (Resolvent, Spectrum). Let  $\mathcal{A}$  be a unital Banach algebra. Define

$$\begin{aligned} \rho(a) &:= \{\lambda \in \mathbb{C} : \lambda - a \in \mathcal{A}^{\times}\} \\ \sigma(a) &:= \{\lambda \in \mathbb{C} : \lambda - a \notin \mathcal{A}^{\times}\}. \end{aligned}$$

$\rho(a)$  and  $\sigma(a)$  are called the resolvent set and the spectrum set of  $a$  respectively.

**Theorem 29.12.** Let  $\mathcal{A}$  be a unital Banach algebra. Then  $\sigma(a)$  is not empty.

*Proof.* Let  $f(\lambda) = (\lambda - a)^{-1}$  for  $\lambda \in \rho(a)$ . Then  $f$  is a nonconstant holomorphic function which is bounded near  $\infty$ . Thus, if  $f$  were defined everywhere, that is,  $\rho(a) = \mathbb{C}$ ,  $f$  would become a nonconstant bounded holomorphic function defined everywhere on  $\mathbb{C}$ . As the Liouville theorem asserts, there is no such a holomorphic function. From this we conclude  $\rho(a) \subsetneq \mathbb{C}$ .  $\square$

**Theorem 29.13.** If  $\mathcal{A}$  is a unital Banach algebra which is a field as well, then  $\mathcal{A}$  is isomorphic to  $\mathbb{C}$ .

*Proof.* Let  $a \in \mathcal{A}$ . Then  $\sigma(a)$  consists of a point  $\lambda$ . Since the only non-invertible element in  $\mathcal{A}$  is 0, we must have

$$(29.12) \quad a - \lambda \in \mathcal{A} \setminus \mathcal{A}^{\times} = \{0\}.$$

Thus,  $a = \lambda$  and this is the desired result.  $\square$

**Proposition 29.14.** *Let  $a, b \in \mathcal{A}$ . Then  $\sigma(ab) \setminus \{0\} = \sigma(ba) \setminus \{0\}$ .*

*Proof.* By normalizing  $a$ , we have only to prove that  $1 - ab$  is invertible if  $1 - ba$  is invertible. Denote  $c = (1 - ba)^{-1}$ . Then we have  $c - cba = c - bac = 1$ . Therefore a simple calculation gives us

$$(29.13) \quad (1 - ab)(acb + 1) = 1 + acb - abacb - ab = 1 + acb - a(c - 1)b - ab = 1$$

and

$$(29.14) \quad (acb + 1)(1 - ab) = acb + 1 - acbab - ab = acb + 1 - a(c - 1)b - ab = 1.$$

This shows that  $(1 - ab)^{-1} = acb + 1$ . □

**Remark 29.15.** This example will explain why we excluded 0 in Proposition 29.14. Let  $\mathcal{A} = B(\ell^2(\mathbb{N}))$ . Then define

$$(29.15) \quad A(\{a_j\}_{j \in \mathbb{N}}) := \{a_{j+1}\}_{j \in \mathbb{N}}, \quad B(\{a_j\}_{j \in \mathbb{N}}) := \{a_{j-1}\}_{j \in \mathbb{N}}.$$

Here it will be understood that  $a_0 = 0$  for  $a = \{a_j\}_{j \in \mathbb{N}}$ . Then

$$(29.16) \quad AB = \text{id}_{\ell^2(\mathbb{N})}, \quad BA(\{a_j\}_{j \in \mathbb{N}}) = (0, a_2, a_3, \dots).$$

Therefore,

$$(29.17) \quad 0 \in \sigma(BA) \setminus \sigma(AB).$$

**Lemma 29.16.** *Let  $\mathcal{A}$  be a Banach algebra. Then  $r(a) := \lim_{j \rightarrow \infty} \|a^j\|_{\mathcal{A}}^{\frac{1}{j}}$  exists for all  $a \in \mathcal{A}$ . Furthermore we have  $r(a) \leq \|a\|_{\mathcal{A}}$ .*

*Proof.* Let  $j$  be fixed. For  $k > j$  let us divide it by  $j$ : We write  $k = j \cdot m(k) + l(k)$  with  $m \in \mathbb{N}$  and  $0 \leq l(k) < j(k)$ . Then we have

$$(29.18) \quad \|a^k\|_{\mathcal{A}}^{\frac{1}{k}} \leq \|a^j\|_{\mathcal{A}}^{\frac{m(k)}{k}} \cdot \|a\|_{\mathcal{A}}^{\frac{l(k)}{k}}.$$

Then, taking into account that  $\lim_{k \rightarrow \infty} \frac{m(k)}{k} = \frac{1}{j}$ , we have

$$(29.19) \quad \limsup_{k \rightarrow \infty} \|a^k\|_{\mathcal{A}}^{\frac{1}{k}} \leq \|a^j\|_{\mathcal{A}}^{\frac{1}{j}}.$$

Since  $j$  is arbitrary, we see

$$(29.20) \quad \limsup_{k \rightarrow \infty} \|a^k\|_{\mathcal{A}}^{\frac{1}{k}} \leq \liminf_{j \rightarrow \infty} \|a^j\|_{\mathcal{A}}^{\frac{1}{j}}.$$

Hence the limit in question exists.

The remaining assertion is easy from  $\|a^j\|_{\mathcal{A}}^{\frac{1}{j}} \leq \|a\|_{\mathcal{A}}$ . □

**Proposition 29.17.**  *$r(ab) = r(ba)$  for  $a, b \in \mathcal{A}$ .*

*Proof.* Let  $j \in \mathbb{N}$ .

$$\begin{aligned} \|(ab)^{j+1}\|_{\mathcal{A}}^{\frac{1}{j+1}} &\leq \|(ba)^j\|_{\mathcal{A}}^{\frac{1}{j+1}} \|a\|_{\mathcal{A}}^{\frac{1}{j+1}} \|b\|_{\mathcal{A}}^{\frac{1}{j+1}} \\ &\leq \|(ba)^j\|_{\mathcal{A}}^{\frac{1}{j}} \|a\|_{\mathcal{A}}^{\frac{1}{j+1}} \|b\|_{\mathcal{A}}^{\frac{1}{j+1}} \|(ba)^j\|_{\mathcal{A}}^{\frac{1}{j(j+1)}} \end{aligned}$$

Taking  $\limsup_{j \rightarrow \infty}$ , we obtain  $r(ab) \leq r(ba)$ . By symmetry we also have  $r(ba) \leq r(ab)$ . Thus, it follows that  $r(ba) = r(ab)$ . □

What can be said for complex numbers can be said for Banach algebras.

**Theorem 29.18** (Neumann expansion). *Let  $a \in \mathcal{A}$  and  $\lambda \in \mathbb{C}$ . If  $\lambda > r(a)$ , then  $\lambda \in \rho(a)$  and we have*

$$(29.21) \quad (\lambda - a)^{-1} = \sum_{j=1}^{\infty} \frac{a^j}{\lambda^{j+1}}.$$

*Proof.* We have dealt with a similar assertion for  $B(X)$ . Just mimic its proof. □

**Example 29.19.** Let  $X$  be a Banach space and  $A \in B(X)$  with  $\|A\|_X < 1$ . Then  $(1 - A)$  is invertible.

We list two corollaries that are special cases of Theorem 29.18.

**Corollary 29.20.** *Let  $a \in \mathcal{A}$ . Then if  $|\lambda| > \|a\|_{\mathcal{A}}$ , then  $\lambda \in \rho(\mathcal{A})$  and*

$$(29.22) \quad \|(\lambda - a)^{-1}\|_{\mathcal{A}} \leq \frac{1}{|\lambda| - \|a\|_{\mathcal{A}}}.$$

**Corollary 29.21.** *Let  $a \in \mathcal{A}^{\times}$ . Assume that  $b \in \mathcal{A}$  satisfies*

$$(29.23) \quad \|a - b\|_{\mathcal{A}} < \frac{1}{\|a^{-1}\|_{\mathcal{A}}}.$$

*Then  $b \in \rho(\mathcal{A})$ . In particular  $\rho(\mathcal{A})$  is an open set in  $\mathbb{C}$ .*

*Exercise 189.* Prove Corollaries 29.20 and 29.21.

Concerning Theorem 29.18, we note the following.

**Theorem 29.22** (Spectral radius theorem). *Let  $a \in \mathcal{A}$ . Then*

$$(29.24) \quad r(a) = \lim_{j \rightarrow \infty} \|a^j\|_{\mathcal{A}} = \sup_{\lambda \in \sigma(a)} |\lambda|.$$

*Proof.* We have seen in Theorem 29.18 that

$$(29.25) \quad r(a) \geq \sup_{\lambda \in \sigma(a)} |\lambda|.$$

To prove the converse, we note the following on the convergence radius of a Taylor series. Set  $\mathbb{D}(r) = \{z \in \mathbb{C} : |z| < r\}$ .

**Claim 29.23.** *If  $f : \mathbb{D}(r) \rightarrow \mathcal{A}$  is a holomorphic function whose Taylor expansion is given by*

$$(29.26) \quad f(z) = \sum_{j=0}^{\infty} a_j z^j, \quad a_0, a_1, \dots \in \mathcal{A},$$

*then  $\limsup_{j \rightarrow \infty} \|a_j\|_{\mathcal{A}}^{-\frac{1}{j}} \geq r$ .*

Claim 29.23 is proved by expressing each coefficient by means of complex line integral. Although  $f$  takes its value in  $\mathcal{A}$ , we can go through the same argument when  $\mathcal{A} = \mathbb{C}$ .

In view of this, if the inequality in (29.25) were strict, then we would have

$$(29.27) \quad r(a) = \limsup_{j \rightarrow \infty} \|a^{j-1}\|_{\mathcal{A}}^{\frac{1}{j}} > \sup_{\lambda \in \sigma(a)} |\lambda|$$

for the expansion  $(a - \lambda)^{-1} = \sum_{j=0}^{\infty} \frac{a^j}{\lambda^{j+1}}$ . (29.27) is a contradiction. Thus, (29.25) is actually an identity. □

**Theorem 29.24.** *Any maximal ideal of a unital Banach algebra  $\mathcal{A}$  is closed.*

*Proof.* Let  $M$  be a maximal ideal of  $\mathcal{A}$ . Then its closure  $\overline{M}$  is also closed. By Theorem 29.18, we see that  $B(e, 1) \subset \mathcal{A}^\times$ . Since  $M$  is a maximal ideal,  $M$  contains no invertible element. Therefore,  $M \cap B(e, 1) \neq \emptyset$ . Since  $B(e, 1)$  is open, we conclude  $\overline{M} \cap B(e, 1) = \emptyset$ . Therefore  $\overline{M}$  is a proper subset of  $\mathcal{A}$ . Since  $\overline{M}$  is an ideal satisfying  $M \subset \overline{M} \subsetneq \mathcal{A}$ , it follows from the maximality of  $M$  that  $M = \overline{M}$ . Therefore  $M$  is closed.  $\square$

**Theorem 29.25.** *Let  $M$  be a maximal ideal of a unital Banach algebra  $\mathcal{A}$ . Then  $\mathcal{A}/M$  is isomorphic to  $\mathbb{C}$ .*

*Proof.* Since  $\mathcal{A}/M$  is a field and a Banach algebra, it follows that  $\mathcal{A}/M$  is isomorphic to  $\mathbb{C}$ .  $\square$

Maximal ideal space. In this paragraph we assume that  $\mathcal{A}$  is a commutative Banach algebra.

**Definition 29.26** (Homomorphism).

(1)  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  is said to be a homomorphism, if

$$\varphi(a + b) = \varphi(a) + \varphi(b), \varphi(ab) = \varphi(a)\varphi(b)$$

for all  $a, b \in \mathcal{A}$ .

(2) The maximal ideal space  $M_{\mathcal{A}}$  is given by

$$(29.28) \quad M_{\mathcal{A}} := \{ \varphi : \varphi \text{ is a nonzero homomorphism} \}.$$

Assume for the time being that  $\mathcal{A}$  is unital. Let us see why  $M_{\mathcal{A}}$  deserves its name.

**Theorem 29.27.** *Assume that  $\mathcal{A}$  is unital.*

(1) *Let  $\varphi \in M_{\mathcal{A}}$ . Then  $\text{Ker}(\varphi)$  is a maximal ideal.*

(2) *If  $M$  is a maximal ideal, then there exists a unique  $\varphi$  whose kernel is  $M$ .*

*Proof.* (1) Since  $\varphi$  is not zero,  $\text{Ker}(\varphi)$  is not empty. Pick an element  $a_0 \notin \text{Ker}(\varphi)$ . Let  $a \in \mathcal{A}$ . Then we have  $a - \frac{\varphi(a)}{\varphi(a_0)}a_0 \in \text{Ker}(\varphi)$ . Thus,  $a \in [\text{Ker}(\varphi), a_0]$ . This implies  $[\text{Ker}(\varphi), a_0] = \mathcal{A}$ .  $\square$

*Proof.* (2) Consider a composition of natural mappings.

$$(29.29) \quad \mathcal{A} \mapsto \mathcal{A}/M \simeq \mathbb{C}.$$

If we set this mapping as  $\varphi$ , then it is easy to see that  $\varphi \in M_{\mathcal{A}}$  and  $\text{Ker}(\varphi) = M$ .  $\square$

Before we proceed further, let us make a little deeper look at the isomorphism

$$(29.30) \quad \mathcal{A}/M \simeq \mathbb{C}.$$

Note that  $z \in \mathbb{C}$  is mapped to  $[a] \in \mathcal{A}/M$  such that  $a - z \in M$ .

Given a polynomial  $f(x) = \sum_{j=0}^d a_j x^{d-j}$  and  $a \in \mathcal{A}$ , we define  $f(a) := \sum_{j=0}^d a_j a^{d-j}$

**Proposition 29.28.** *Suppose that  $f(x)$  is a polynomial with complex coefficients. Then for every  $a \in \mathcal{A}$ , we have*

$$(29.31) \quad \sigma(f(a)) = \{f(\lambda) : \lambda \in \sigma(a)\}.$$

*Proof.* Suppose that  $\lambda \in \sigma(a)$ . We are going to show that  $f(\lambda) \in \sigma(f(a))$ . In this case by virtue of the remainder theorem we can write

$$(29.32) \quad f(a) - f(\lambda) = (a - \lambda)g(a)$$

with some polynomial  $g(x)$ . If  $f(\lambda) \notin \sigma(f(a))$ , then we would have  $a - \lambda$  would have inverse, which is a contradiction.

Suppose instead that  $\lambda \in \sigma(f(a))$ . Then there exist  $\gamma_1, \gamma_2, \dots, \gamma_k \in \mathbb{C}$  such that

$$(29.33) \quad f(z) - \lambda = (z - \gamma_1)(z - \gamma_2) \dots (z - \gamma_k),$$

by virtue of the fundamental theorem of algebra. There exists  $j_0 = 1, 2, \dots, k$  such that  $a - \gamma_{j_0}$  is not invertible: Otherwise  $f(a) - \lambda$  is invertible and this contradicts to  $\lambda \in \sigma(f(a))$ . Therefore, it follows that

$$(29.34) \quad \lambda = f(\gamma_{j_0}) \in \{f(\mu) : \mu \in \sigma(a)\}.$$

This is the desired result. □

**Theorem 29.29.** *Assume that  $\mathcal{A}$  is unital.  $\mathcal{A}^*$  denotes the set of all bounded continuous functional. Then  $M_{\mathcal{A}}$  is a weak-\* compact subset of the closed unit ball  $\mathcal{A}^*$ .*

*Proof.*   $M_{\mathcal{A}}$  is a subset of the closed unit ball of  $\mathcal{A}^*$ .  Let  $\varphi \in M_{\mathcal{A}}$ . Suppose that  $z \in \mathbb{C}$  satisfies  $|z| > \|a\|_{\mathcal{A}}$ . Then  $z - a \in \mathcal{A}^*$ . Therefore, there exists  $b_z \in \mathcal{A}$  such that  $(z - a)b_z = e_{\mathcal{A}}$ . Using this  $b_z$ , we obtain  $\varphi(z - a)\varphi(b_z) = 1$ . Thus  $z \neq \varphi(a)$ . Since  $z$  is an arbitrary point outside the closed disk  $\{|z| \leq \|a\|_{\mathcal{A}}\}$ , we conclude  $|\varphi(a)| \leq \|a\|_{\mathcal{A}}$ . Thus  $\varphi$  lies in the closed unit ball in  $\mathcal{A}^*$ .

$M_{\mathcal{A}}$  is weak-\* compact in  $\mathcal{A}^*$ .  Let us write out  $M_{\mathcal{A}}$  in terms of  $\mathcal{A}^*$ . Since

$$(29.35) \quad M_{\mathcal{A}} = \bigcap_{a, b \in \mathcal{A}} \{\varphi \in \mathcal{A}^* : \varphi(ab) = \varphi(a)\varphi(b)\},$$

and each  $\{\varphi \in \mathcal{A}^* : \varphi(ab) = \varphi(a)\varphi(b)\}$  is weak-\* closed in  $\mathcal{A}^*$ , we see that  $M_{\mathcal{A}}$  is weak-\* closed. By the Banach Alaoglu theorem and the fact that weak-\* topology is Hausdorff, we see that  $M_{\mathcal{A}}$  is weak-\* compact. □

**Examples.** Now let us exhibit some examples of the above notions.

**Example 29.30.** Let  $X$  be a compact Hausdorff space. For a point  $x \in X$ , define

$$(29.36) \quad \text{ev}_x(f) = f(x), f \in C(X).$$

Then it follows from the definition of  $M_{C(X)}$  that  $\text{ev}_x \in M_{C(X)}$ .

**Theorem 29.31.** *Keep to the same setting as Example 29.30. Then the evaluation mapping  $\text{ev} : X \mapsto M_{C(X)}$ ,  $\text{ev}(x) = \text{ev}_x$  is a homeomorphism.*

*Proof.*  **Injectivity**  Let  $x, y \in X$  be distinct points. Then Uryzohn's lemma yields  $f \in C(X)$  such that  $f(x) = 1$  and  $f(y) = 0$ . Then  $[\text{ev}(x)](f) = 1$  and  $[\text{ev}(y)](f) = 0$ . Thus  $\text{ev}(x) \neq \text{ev}(y)$ .

**Surjectivity**  Let  $\varphi \in M_{C(X)}$ . Then we have

$$(29.37) \quad \{x \in X : f_1(x) = f_2(x) = \dots = f_n(x) = 0\} \neq \emptyset$$

for any finite collection  $f_1, f_2, \dots, f_n \in \text{Ker}(\varphi)$ . Indeed, if we set  $g := \sum_{j=1}^n |f_j|^2$ , then we have

$g \in \text{Ker}(\varphi)$ . Therefore we obtain  $g \notin C(X)^\times$ , since  $\text{Ker}(\varphi) \cap C(X)^\times = \emptyset$ . Let us recall the condition for  $h \in C(X)$  to be invertible:  $h$  is invertible if and only if  $\inf_{x \in X} |h(x)| > 0$ . In view

of this observation  $g$  vanishes at some point  $\tilde{x} \in X$ . Thus, (29.37) is established. Since  $X$  is compact, we are in the position of applying the finite intersection property. According to this property,  $\text{Ker}(\varphi)$  has a common vanishing point  $x_0$ . This can be rephrased as

$$(29.38) \quad \text{Ker}(\varphi) \subset \text{Ker}(\text{ev}_{x_0}).$$

Finally let us prove that  $\text{ev}_{x_0} = \varphi$ , which shows the surjectivity of  $\text{ev}$ . If we pick  $f \in C(X)$  arbitrarily, then we have  $f - \varphi(f) \in \text{Ker}(\varphi) \subset \text{Ker}(\text{ev}_{x_0})$ . Therefore,

$$(29.39) \quad \varphi(f) = \varphi(f - \varphi(f)) + \varphi(f) = f(x_0) - \varphi(f) + \varphi(f) = f(x_0).$$

Thus,  $\varphi = \text{ev}_{x_0}$ .

Continuity Since we have proved the surjectivity of  $\text{ev}$ ,

$$(29.40) \quad \{\mathcal{U}_{x,\varepsilon,f} : x \in X, \varepsilon > 0, f \in C(X)\}$$

is an open basis of  $M_{C(X)}$ , where

$$(29.41) \quad \mathcal{U}_{x,\varepsilon,f} := \{\tau \in M_{C(X)} : |\tau(f) - f(x)| > \varepsilon\}.$$

Since  $\text{ev}^{-1}(\mathcal{U}_{x,\varepsilon,f}) = \{y \in X : |f(y) - f(x)| > \varepsilon\}$  is an open set of  $X$ , it follows that  $\text{ev}$  is continuous.

Openness Since  $X$  is compact,  $M_{C(X)}$  is Hausdorff and  $\text{ev}$  is continuous, we see that  $\text{ev}$  is open. □

*Exercise 190.* Let  $X$  be a compact topological space and  $Y$  a Hausdorff topological space. Then any continuous bijection  $f : X \rightarrow Y$  is homeomorphism.

**Example 29.32.** Let  $\lambda \in \partial\mathbb{D} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ . Then define

$$(29.42) \quad \tau_\lambda(a) = \sum_{j \in \mathbb{Z}} a_j \lambda^j$$

for  $a = \{a_j\}_{j \in \mathbb{Z}} \in \ell^1$ . Define  $\tau : \lambda \in \partial\mathbb{D} \mapsto M_{\ell^1}$ .

**Theorem 29.33.** *Keep to the same setting above. Then  $\tau$  is a homeomorphism.*

*Proof.* It is the same as Example 29.30 that openness of  $\tau$  follows automatically once we prove  $\tau$  is a continuous bijection. However, unlike Example 29.30 it is easy to see that  $\tau$  is a continuous injection. Thus, it remains to prove the bijectivity. The same argument as Example 29.30 works for this case. In particular we can show the following: Let  $\varphi \in M_{\ell^1}$ . There exists  $\lambda \in \partial\mathbb{D}$  with  $\text{Ker}(\varphi) \subset \text{Ker}(\tau_\lambda)$ . Thus, going through the same argument as Example 29.30, we see  $\tau$  is surjective. □

*Exercise 191.* Fill in the details of Example 29.32.

Gelfand transform for unital Banach algebra. In this paragraph we investigate the structure of unital and commutative Banach algebras.

**Definition 29.34** (Gelfand transform). Let  $\mathcal{A}$  be a unital and commutative Banach algebra. Given  $f \in \mathcal{A}$ , define  $\hat{f} \in C(M_{\mathcal{A}})$  by  $\hat{f}(\varphi) = \varphi(f)$ .

**Theorem 29.35** (Gelfand).  $\text{ev} : \mathcal{A} \mapsto C(M_{\mathcal{A}})$ ,  $f \mapsto \hat{f}$  is a homomorphism. Furthermore,

$$(29.43) \quad \|\hat{f}\|_{C(M_{\mathcal{A}})} \leq \|f\|_{\mathcal{A}}.$$



*Proof.* We content ourselves with proving well-definedness of  $ev$ , that is, we are going to show that  $ev(f) = \hat{f} \in C(M_{\mathcal{A}})$ . To do this, we pick an open set  $U \subset \mathbb{C}$ . Then

$$(29.44) \quad \{\varphi \in M_{\mathcal{A}} : \hat{f}(\varphi) \in U\} = \{\varphi \in M_{\mathcal{A}} : \varphi(f) \in U\}$$

is an open set in  $M_{\mathcal{A}}$ . Therefore,  $ev$  is well-defined. □

*Exercise 192.* Complete the proof of Theorem 29.35.

**Theorem 29.36.** For every  $f \in \mathcal{A}$ , we have  $\hat{f}(M_{\mathcal{A}}) = \sigma(f) \subset \mathbb{C}$ .

*Proof.*  $\hat{f}(M_{\mathcal{A}}) \subset \sigma(f)$  Let  $\lambda := \hat{f}(\varphi) = \varphi(f)$  for some  $\varphi \in M_{\mathcal{A}}$ . Then  $f - \lambda = f - \varphi(f)$  cannot be invertible because it belongs to  $\text{Ker}(\varphi)$ . Therefore  $\lambda \in \sigma(f)$ .

$\hat{f}(M_{\mathcal{A}}) \supset \sigma(f)$  Conversely let  $\lambda \in \sigma(f)$ . Then there exists a maximal ideal  $I$  such that  $f - \lambda \in I$ . Given  $g \in \mathcal{A}$ , we define  $\varphi(g)$  as a unique complex number such that  $g - \varphi(g) \in I$ . It follows from the definition that  $\varphi(f) = \lambda$  and  $\varphi$  preserves operations. If  $g \notin I$ , then  $\varphi(g) \neq 0$  by definition. Therefore, and  $\varphi \in M_{\mathcal{A}}$ . □

### 29.2. $C^*$ -algebras.

$C^*$ -algebras. As we have seen, if  $X$  is a Banach space, then  $B(X)$  is a Banach algebra. Suppose in addition that  $X$  is a Hilbert space. Then  $B(X)$  has another operation. That is, the adjoint operation  $A \mapsto A^*$ . Keeping the property of this operation in mind, let us make a view of the definition of  $C^*$ -algebras.

**Definition 29.37** ( $C^*$ -algebra). A  $C^*$ -algebra is a triple  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}}, *)$  of a Banach algebra  $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$  and an operation called involution  $*$ :  $\mathcal{A} \rightarrow \mathcal{A}$  with the following properties.

- (1)  $(a^*)^* = a$  for all  $a \in \mathcal{A}$ .
- (2)  $(k \cdot a)^* = \bar{k} \cdot a^*$  for all  $k \in \mathbb{K}$  and  $a \in \mathcal{A}$ .
- (3)  $(a + b)^* = a^* + b^*$  for all  $a, b \in \mathcal{A}$ .
- (4)  $\|a^*\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}$  for all  $a \in \mathcal{A}$ .
- (5)  $\|a^*a\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}^2$  for all  $a \in \mathcal{A}$ .

**Example 29.38.** All the examples in Example 29.2 are  $C^*$ -algebras, if we assume  $X$  is a Hilbert space in (2). In particular  $M_{\text{dyadic}}(\mathbb{C})$ , the set of all  $d \times d$  matrices, is an example of  $C^*$ -algebra with  $X = \mathbb{R}^d$ .

**Definition 29.39.** Let  $\mathcal{A}$  be a  $C^*$ -algebra and  $a \in \mathcal{A}$ .

- (1)  $a$  is said to be normal, if  $a^*a = aa^*$ .
- (2)  $a$  is said to be self-adjoint if  $a^* = a$ .  $S(\mathcal{A})$  denotes the set of all self-adjoint elements in  $\mathcal{A}$ .
- (3)  $a$  is said to be projection, if  $a^2 = a$  and  $a \in S(\mathcal{A})$ .  $P(\mathcal{A})$  denotes the set of all projections in  $\mathcal{A}$ .
- (4)  $a$  is said to be unitary, if  $a^*a = aa^* = e_{\mathcal{A}}$ .  $U(\mathcal{A})$  denotes the set of all unitary elements in  $\mathcal{A}$ .

*Exercise 193.* Let  $f \in L^\infty(X)$ , where  $(X, \mathcal{B}, \mu)$  is a measure space. Establish:

- (1)  $f \in S(L^\infty(X))$  if and only if  $\text{Essrange}(f) \subset \mathbb{R}$ .
- (2)  $f \in P(L^\infty(X))$  if and only if  $\text{Essrange}(f) \subset \{0, 1\}$ .
- (3)  $f \in U(L^\infty(X))$  if and only if  $\text{Essrange}(f) \subset \partial\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ .

*Exercise 194.* Show that all these notions are natural extension of the ones corresponding to  $B(H)$  which we took up in Section 11, where  $H$  is a Hilbert space,

*Exercise 195.* Let  $A$  be a  $2 \times 2$  matrix which is not a constant multiple of the unit matrix  $I$ .

- (1) Let  $B$  be a  $2 \times 2$  matrix. Then characterize the condition for which  $AB = BA$  in terms of  $A$ .
- (2) In  $M_2(\mathbb{C})$ , the  $C^*$ -algebra consisting of  $2 \times 2$  matrix, exhibit a normal element which is not self-adjoint.

*Exercise 196.* Let  $A$  be a non-unital  $C^*$ -algebra. Denote by  $A \times \mathbb{C}$  the unitalization of  $A$ . Show that

$$(29.45) \quad \|(a, \lambda) : A \times \mathbb{C}\| := \sup\{\|ab + \lambda b : A\| : \|b : A\| \leq 1\}$$

defines a norm that makes  $A \times \mathbb{C}$  into a unital  $C^*$ -algebra.

In this paragraph we assume that  $A$  is a unital  $C^*$ -algebra.

**Theorem 29.40.** *Let  $\mathcal{A}$  be a  $C^*$ -algebra and  $a \in \mathcal{A}$ .*

- (1) *If  $a$  is normal, then  $r(a) = \|a\|_{\mathcal{A}}$ .*
- (2) *If  $a \in S(\mathcal{A})$ , then  $\sigma(a) \subset \mathbb{R}$ .*
- (3) *If  $a \in U(\mathcal{A})$ , then  $\sigma(a) \subset \mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ .*

(1). Let  $j \in \mathbb{N}$ .

$$(29.46) \quad \|a^{2^j}\|_{\mathcal{A}} = \|a^{(2^{j-1})}a^{*(2^{j-1})}\|_{\mathcal{A}} = \|(aa^*)^{(2^{j-1})}\|_{\mathcal{A}} = \|aa^*\|_{\mathcal{A}}^{(2^{j-1})} = \|a\|_{\mathcal{A}}^{2^j}.$$

Therefore, it follows that  $r(a) = \lim_{j \rightarrow \infty} \|a^{2^j}\|_{\mathcal{A}}^{\frac{1}{2^j}} = \|a\|_{\mathcal{A}}$ . □

(2). Let  $\beta + i\gamma \in \sigma(a)$ , where  $\beta, \gamma \in \mathbb{R}$ . Observe that, for each  $n \in \mathbb{N}$ ,  $i(n+1)\gamma \in \sigma(a - \beta + in\gamma)$ . Theorem 29.18 gives us

$$(29.47) \quad (n+1)^2\gamma^2 \leq r(a - \beta + in\gamma)^2 \leq \|a - b + ic\|_{\mathcal{A}}^2.$$

Therefore, it follows that

$$(29.48) \quad (n+1)^2\gamma^2 \leq \|(a - \beta)^2 + n^2\gamma^2\|_{\mathcal{A}} \leq \|(a - \beta)^2\|_{\mathcal{A}} + n^2\gamma^2.$$

From this formula we see that  $\gamma = 0$  because  $n \in \mathbb{N}$  is arbitrary. □

(3). It is easy to see

$$(29.49) \quad \sigma(b^{-1}) = \{\lambda \in \mathbb{C} : \lambda^{-1} \in \sigma(a)\}, \quad \sigma(b^*) = \{\lambda \in \mathbb{C} : \bar{\lambda} \in \sigma(a)\},$$

provided  $b^{-1}$  exists. Thus, we obtain, using  $a^* = a^{-1}$ ,

$$\sigma(a) = \{\lambda \in \mathbb{C} : \bar{\lambda} \in \sigma(a^*)\} = \{\lambda \in \mathbb{C} : \bar{\lambda} \in \sigma(a^{-1})\} = \{\lambda \in \mathbb{C} : \overline{\lambda^{-1}} \in \sigma(a)\}.$$

As a result we obtain  $\lambda \in \sigma(a) \iff \overline{\lambda^{-1}} \in \sigma(a)$ . However, again by Theorem 29.18  $\sigma(a)$  is contained in  $\{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$ . Therefore,  $\sigma(a)$  is contained in  $\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ . □

**Theorem 29.41** (Functional calculus). *Let  $a$  be a self-adjoint element in  $\mathcal{A}$ . Then there exists a unique continuous linear mapping*

$$(29.50) \quad f \in C(\sigma(a)) \mapsto f(a) \in \mathcal{A}$$

such that  $f(a)$  has its elementary meaning when  $f$  is a polynomial. Moreover, this mapping enjoys the following property for all  $f, g \in C(\sigma(a))$ .

- (1)  $\|f(a)\|_{\mathcal{A}} = \|f : C(\sigma(A))\|$  ;
- (2)  $(f \cdot g)(a) = f(a)g(a)$  ;
- (3)  $\overline{f(a)} = f(a)^*$  ;

- (4)  $f(a)$  is normal ;
- (5) If  $b$  commutes with  $A$ , then  $bf(a) = f(a)b$ .

*Outline.* Since the set of all polynomial is dense in  $C(\sigma(A))$  by virtue of Weierstrass's approximation theorem, we have only to show the first assertion. The remaining assertion follows immediately from this. Furthermore by approximation, it is enough even to prove it only for polynomial.

This can be achieved easily by virtue of Theorem 29.27,:

$$(29.51) \quad \|f(a)\|_{\mathcal{A}} = \sup_{\varphi \in M_{\mathcal{A}}} |\varphi(f(a))| = \sup_{\varphi \in M_{\mathcal{A}}} |f(\varphi(a))| = \|f : C(\sigma(a))\|.$$

This is the desired result. □

*Exercise 197.* Fill in the details of the proof of this theorem.

**Theorem 29.42** (Spectral mapping theorem). *Let  $a \in \mathcal{A}$ . Let  $f \in C(\sigma(a))$ . Then we have*

$$(29.52) \quad \sigma(f(a)) = \{f(\lambda) : \lambda \in \sigma(a)\}.$$

*Proof.* To prove this theorem, we begin with a setup. Choose a sequence of polynomials  $\{f_j\}_{j \in \mathbb{N}}$  such that  $r_j := \|f - f_j : C(\sigma(a))\| < e^{-j}$  for each  $j \in \mathbb{N}$ .

First let us establish

$$(29.53) \quad \{f(\lambda) : \lambda \in \sigma(a)\} \subset \sigma(f(a)).$$

or equivalently

$$(29.54) \quad \mathbb{C} \setminus \sigma(f(a)) \subset \mathbb{C} \setminus \{f(\lambda) : \lambda \in \sigma(a)\}.$$

To prove (29.54) we let  $\mu \notin \sigma(f(a))$ . Since  $\mu \notin \sigma(f(a))$ , i.e.,  $\mu \in \rho(f(a))$ ,  $(\mu - f(a))^{-1}$  makes sense and is not a zero element. With this in mind, let us set

$$(29.55) \quad R := \frac{1}{2\|(f(a) - \mu)^{-1}\|_{\mathcal{A}}}.$$

We claim

$$(29.56) \quad B(\mu, R) \cap \left( \bigcup_{j \in \mathbb{N} \cap (1 + |\log R|, \infty)} \{f_j(\lambda) : \lambda \in \sigma(a)\} \right) = \emptyset.$$

To prove (29.56), we suppose that  $j > 1 + |\log R|$  is an integer.

We take  $\tau \in B(\mu, R)$ . Then we have

$$(29.57) \quad \|(f(a) - \mu) - (f_j(a) - \tau)\|_{\mathcal{A}} \leq \|f(a) - f_j(a)\|_{\mathcal{A}} + |\mu - \tau|.$$

By the fact that  $j > 1 + |\log R|$  and the functional calculus, we have

$$(29.58) \quad \|f(a) - f_j(a)\|_{\mathcal{A}} + |\mu - \rho| \leq 2r_j < \frac{1}{\|(f(a) - \mu)^{-1}\|_{\mathcal{A}}}.$$

If we combine (29.57) and (29.58), we obtain

$$(29.59) \quad \|(f(a) - \mu) - (f_j(a) - \tau)\|_{\mathcal{A}} < \frac{1}{\|(f(a) - \mu)^{-1}\|_{\mathcal{A}}}.$$

(29.59) allows us to use the Neumann expansion. The neumann expansion shows that  $f_j(a) - \tau$  is invertible and hence

$$(29.60) \quad \tau \in \rho(f_j(a)) = \sigma(f_j(a))^c, j \in \mathbb{N}$$

Since each  $f_j$  is a polynomial, we have

$$(29.61) \quad \sigma(f_j(a)) = \{f_j(\lambda) : \lambda \in \sigma(a)\}, j \in \mathbb{N}.$$

Thus, we finally obtain (29.56) from (29.60) and (29.61).

Since  $f(\lambda) = \lim_{j \rightarrow \infty} f_j(\lambda)$ , it follows that  $\mu \notin \{f(\lambda) : \lambda \in \sigma(a)\}$ . As a consequence (29.53) was justified.

Suppose that

$$(29.62) \quad \mu \notin \{f(\lambda) : \lambda \in \sigma(a)\}.$$

Then, there exists  $R > 0$  such that

$$(29.63) \quad \left( \bigcup_{\lambda \in f^{-1}(\mu)} B(\lambda, 3R) \right) \cap \sigma(a) = \emptyset.$$

By uniform convergence of the sequence  $\{f_j\}_{j \in \mathbb{N}}$  we have

$$(29.64) \quad \left( \bigcup_{\lambda \in f_j^{-1}(\mu)} B(\lambda, 2R) \right) \cap \sigma(a) = \emptyset.$$

Therefore, it follows that  $|\mu - f_j(\lambda)| \geq 2R$ ,  $\lambda \in \sigma(a)$ , if  $j > \log R^{-1}$ . This implies

$$(29.65) \quad g_j(\lambda) = \frac{1}{\mu - f_j(\lambda)}$$

is a continuous function on  $\sigma(a)$ . By the estimate of  $\mu - f_j(\lambda)$  from below we see that  $\{g_j\}_{j \in \mathbb{N}}$  converges uniformly to  $g \in C(\sigma(a))$ . Since

$$(29.66) \quad g_j(a)(\mu - f_j(a)) = (\mu - f_j(a))g_j(a) = 1,$$

a passage to the limit gives us  $(\mu - f(a))^{-1} = g(a)$ . Therefore, we obtain  $\mu \in \rho(a) = \mathbb{C} \setminus \sigma(a)$ .  $\square$

Positive elements. In the  $C^*$ -algebra world we are able to define positivity.

**Definition 29.43.**  $a \in \mathcal{A}$  is said to be positive, if  $a \in S(\mathcal{A})$  and  $\sigma(a) \subset [0, \infty)$ .  $S(\mathcal{A})_+$  denotes the set of all positive elements in  $\mathcal{A}$ .

**Lemma 29.44.** *If  $a \in S(\mathcal{A})_+$  satisfies  $-a \in S(\mathcal{A})_+$ , then  $a = 0$ .*

*Proof.* By assumption  $\sigma(a) = 0$ . Since  $a \in S(\mathcal{A})$ , it follows that  $\|a\|_{\mathcal{A}} = r(a) = \sup_{\lambda \in \mathcal{A}} |\lambda| = 0$ . Therefore we conclude  $a = 0$ .  $\square$

The following lemma is a key to our observations.

**Lemma 29.45.** *Let  $a \in S(\mathcal{A})$ . Then  $a$  is positive if and only if*

$$(29.67) \quad \|a - \|a\|_{\mathcal{A}}\|_{\mathcal{A}} \leq \|a\|_{\mathcal{A}}.$$

*In particular,  $S(\mathcal{A})_+$  is closed.*

*Proof.* Assume that  $a$  is positive. Then we have

$$(29.68) \quad \sigma(a - \|a\|_{\mathcal{A}}) \subset [-\|a\|_{\mathcal{A}}, 0]$$

by assumption. Therefore,

$$(29.69) \quad \|a - \|a\|_{\mathcal{A}}\|_{\mathcal{A}} = \sup_{\lambda \in \sigma(a - \|a\|_{\mathcal{A}})} |\lambda| \leq \|a\|_{\mathcal{A}}.$$

Assume  $\|a - \|a\|_{\mathcal{A}}\|_{\mathcal{A}} \leq \|a\|_{\mathcal{A}}$ . Then we have

$$(29.70) \quad \sigma(a - \|a\|_{\mathcal{A}}) \subset [-\|a\|_{\mathcal{A}}, \|a\|_{\mathcal{A}}].$$

Therefore, by the spectral mapping theorem we obtain

$$(29.71) \quad \sigma(a) \subset [0, 2\|a\|_{\mathcal{A}}],$$

which shows that  $a$  is positive. □

**Theorem 29.46.** *If  $a, b \in S(\mathcal{A})_+$ , then  $a + b \in S(\mathcal{A})_+$ .*

*Proof.* From Lemma 29.45 we obtain

$$(29.72) \quad \|a - \|a\|_{\mathcal{A}}\| \leq \|a\|_{\mathcal{A}}, \|b - \|b\|_{\mathcal{A}}\| \leq \|b\|_{\mathcal{A}}.$$

As a consequence, we see

$$(29.73) \quad \|a + b - \|a\|_{\mathcal{A}} - \|b\|_{\mathcal{A}}\| \leq \|a\|_{\mathcal{A}} + \|b\|_{\mathcal{A}}.$$

Therefore, we obtain

$$(29.74) \quad \sigma(a + b) \subset [0, 2(\|a\|_{\mathcal{A}} + \|b\|_{\mathcal{A}})],$$

proving  $a + b \in S(\mathcal{A})_+$ . □

Positive and negative parts. Having defined positivity, we are now oriented to decomposition of the elements into positive and negative parts.

**Theorem 29.47.** *Let  $a \in S(\mathcal{A})$ . Then there exists a unique element  $b_0, c_0 \in S(\mathcal{A})_+$  such that  $a = b_0 - c_0$ ,  $b_0 c_0 = c_0 b_0 = 0$ .*

*Proof.* Existence We set  $\alpha(t) = \max(|t|, 0)$  and  $\beta(t) = t - \alpha(t)$ . If we define  $b_0 := \alpha(a)$  and  $c_0 := \beta(a)$ , then we have  $a = b_0 - c_0$ ,  $b_0 c_0 = c_0 b_0 = 0$ .

Uniqueness Suppose that  $b, c \in S(\mathcal{A})_+$  such that  $a = b - c$ ,  $b c = c b = 0$ . Then we have  $a^k = b^k + (-c)^k$  for each  $k \in \mathbb{N}_0$ . Therefore, by linearity this relation extends to any polynomial  $p(x)$ .

$$(29.75) \quad p(a) = p(b) + p(-c)$$

A passage to the limit therefore gives

$$(29.76) \quad \alpha(a) = \alpha(b) + \alpha(-c).$$

However,  $\alpha(b) = b$  and  $\alpha(-c) = 0$  because on  $[0, \infty)$  we have  $\alpha(t) = t$  and on  $(-\infty, 0]$  we have  $\alpha(t) = 0$ . Therefore,  $b_0 = \alpha(b) = b$  and  $c_0 = b_0 - a = b - a = c$ . □

**Definition 29.48.**  $b_0$  in Theorem 29.47 is said to be positive part of  $a$ , while  $c_0$  is said to be the negative part of  $a$ .

Characterization of positivity and square root of positive elements. Starting from the positivity, we shall define the square root of positive elements which is of importance. A real number is positive if and only if it is a square of some real numbers. We can even say that a real number is positive precisely when it is a product of a complex number and its conjugate.

**Lemma 29.49.** *If  $a \in \mathcal{A}$  satisfies  $-a^*a \in S(\mathcal{A})_+$ , then  $a = 0$ .*

*Proof.* Since  $\mathcal{A}$  is a  $C^*$ -algebra, we have only to prove  $a^*a = 0$ . Furthermore, Lemma 29.44 and assumption reduces the matter to showing

$$(29.77) \quad a^*a \in \mathcal{A}_+.$$

Let  $a = b + ic$  with  $b, c \in S(\mathcal{A})$ . Then we have  $a^*a + aa^* = 2b^2 + 2c^2$ . Hence  $a^*a = 2b^2 + 2c^2 - aa^*$ . Since  $\sigma(-aa^*) \subset \sigma(a^*a) \cup \{0\}$ , we conclude that  $-aa^* \in S(\mathcal{A})_+$ . By spectral mapping theorem, we see that  $b^2, c^2 \in \mathcal{A}_+$  as well. Since  $S(\mathcal{A})_+$  is closed under addition, (29.77) is established. □

**Theorem 29.50.** *Let  $a \in S(\mathcal{A})$ . Then the following are equivalent.*

- (1)  $a \in S(\mathcal{A})_+$ .
- (2)  $a = b^2$  for some  $b \in S(\mathcal{A})_+$ .
- (3)  $a = c^*c$  for some  $c \in \mathcal{A}$ .

Furthermore  $b$  in (2) is unique and below we write  $b = \sqrt{a}$ .

*Proof.* Let  $f(t) = \sqrt{t}$  for  $t \geq 0$ . If  $a \in S(\mathcal{A})_+$ , then  $b = f(a)$  satisfies the condition of (2). Let  $b' \in S(\mathcal{A})_+$  satisfies  $b'^2 = a$ . Let  $g(t) = t$ . Then pick a sequence of polynomials  $\{p_j(t)\}_{j \in \mathbb{N}}$  such that  $p_j(t) = p_j(-t)$  and  $\|g - p_j : \sigma(a) \cup \sigma(b')\| \rightarrow 0$ . The fact that  $p_j(t) = p_j(-t)$  and  $a = b^2 = b'^2$  gives  $p_j(b') = p_j(b)$ . Thus,  $b' = g(b') = \lim_{j \rightarrow \infty} p_j(b) = b$ . Thus,  $b$  in (2) is unique. Therefore, since (2) is clearly stronger than (3), we are left with the task of proving (3) implies (1).

Set  $d = ca_-$  assuming (3). Then we have  $d^*d = a_-c^*ca_- = a_-aa_- = -a_-^3$ . Since  $a_-^3 \in S(\mathcal{A})_+$ , we see that  $d = 0$  from Lemma 29.49. As a consequence  $a_-^3 = 0$ . Therefore by the spectral mapping theorem,  $\sigma(a_-)$  consists only of 0. Since  $a_- \in S(\mathcal{A})$ , we see that  $a_- = 0$ .  $\square$

**Corollary 29.51.**  $aba^* \in S(\mathcal{A})_+$ , whenever  $a \in \mathcal{A}$  and  $b \in S(\mathcal{A})_+$ .

*Proof.*  $aba^* = (a\sqrt{b})(a\sqrt{b})^* \in S(\mathcal{A})_+$ .  $\square$

**Corollary 29.52.** *Let  $a, b \in S(\mathcal{A})_+$ . If  $a, b$  commutes, then  $ab \in S(\mathcal{A})_+$ .*

*Proof.* Since  $b$  commutes with  $\sqrt{a}$ , it follows that  $ab = \sqrt{ab}\sqrt{a} \in S(\mathcal{A})$ .  $\square$

Order structure. Having set down the properties of the positive elements, we now turn to the order structure of a  $C^*$ -algebra.

**Definition 29.53.** Let  $\mathcal{A}$  be a  $C^*$ -algebra. Then define  $a \leq b$  by  $b - a \in S(\mathcal{A})_+$ .

We see that  $S(\mathcal{A})$  is partially ordered.

**Theorem 29.54.** *Suppose that  $a, b \in S(\mathcal{A})$ .*

- (1) *If  $-b \leq a \leq b$ , then  $\|a\|_{\mathcal{A}} \leq \|b\|_{\mathcal{A}}$ .*
- (2) *If  $0 \leq a \leq b$ , then  $\sqrt{a} \leq \sqrt{b}$ .*
- (3) *If  $0 \leq a \leq b$  and  $a$  is invertible, then  $b$  is invertible and  $b^{-1} \leq a^{-1}$ .*

*Proof.* (1) Since  $-\|b\|_{\mathcal{A}} \leq -b \leq a \leq b \leq \|b\|_{\mathcal{A}}$ , we conclude

$$(29.78) \quad \sigma(a) \subset [-\|b\|_{\mathcal{A}}, \|b\|_{\mathcal{A}}].$$

Therefore, we have  $\|a\|_{\mathcal{A}} \leq \|b\|_{\mathcal{A}}$ .  $\square$

*Proof.* (2) Let  $\varepsilon > 0$  and define  $f_\varepsilon(t) = \sqrt{t + \varepsilon}$ . Then observe that  $f_\varepsilon \rightarrow f$  uniformly on every compact set in  $[0, \infty)$ . Therefore, we may assume that  $a$  and  $b$  are positive.

Since  $a \leq b$ , we have  $\sqrt{b}^{-1}a\sqrt{b}^{-1} \leq 1$ . As a consequence we obtain

$$(29.79) \quad \|\sqrt{b}^{-1}a\|_{\mathcal{A}} = \|\sqrt{b}^{-1}a\sqrt{b}^{-1}\|_{\mathcal{A}}^{\frac{1}{2}} \leq 1.$$

Thus, by Proposition 29.17 we obtain

$$(29.80) \quad \|({}^4\sqrt{b})^{-1}\sqrt{a}({}^4\sqrt{b})^{-1}\|_{\mathcal{A}} = r(({}^4\sqrt{b})^{-1}\sqrt{a}({}^4\sqrt{b})^{-1}) = r(\sqrt{b}^{-1}\sqrt{a}) \leq 1.$$

Hence we obtain

$$(29.81) \quad ({}^4\sqrt{b})^{-1}\sqrt{a}({}^4\sqrt{b})^{-1} \leq 1.$$

Arranging this inequality, we see  $\sqrt{a} \leq \sqrt{b}$ . □

*Proof.* Since  $a$  is invertible by assumption and  $\sigma(a)$  is a closed set that does not contain 0, we have  $\sigma(a) \subset [r, \infty)$  for some  $r > 0$ . Thus,  $a \geq r$ . Since  $S(\mathcal{A})_+$  is partially ordered, by transitivity we see that  $b \geq r$ . Thus we conclude  $b \in \mathcal{A}^*$  as well. It is the same as (2) that we have  $\|\sqrt{b}^{-1}\sqrt{a}\|_{\mathcal{A}} \leq 1$ . Since  $\sqrt{a}$  and  $\sqrt{b}$  is self-adjoint, we have  $\|\sqrt{a}\sqrt{b}^{-1}\|_{\mathcal{A}} \leq 1$ . Therefore, we have  $\|\sqrt{ab}^{-1}\sqrt{a}\|_{\mathcal{A}} \leq 1$ . This means  $\sqrt{ab}^{-1}\sqrt{a} \leq 1$ . As a consequence  $b^{-1} \leq a^{-1}$ . □

### 30. SPECTRAL DECOMPOSITION OF BOUNDED SELF-ADJOINT OPERATORS

#### 30.1. Spectral decomposition.

In this subsection we consider the spectral decomposition of self-adjoint operators on a Hilbert space  $H$ .

**Definition 30.1.** A spectral family  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  is a family in  $P(H)$  that fulfills the following requirements.

- (1) The family  $\{E_\lambda H\}_{\lambda \in \mathbb{R}}$  is increasing.
- (2) For all  $x \in H$ ,  $\lim_{\varepsilon \downarrow 0} E_{\lambda+\varepsilon}x = E_\lambda x$ .
- (3) If  $\lambda > 0$  is sufficiently large, then  $E_\lambda = 1$ .
- (4) If  $\lambda > 0$  is negative enough, then  $E_\lambda = 0$ .

**Theorem 30.2** (Spectral decomposition). *Suppose that  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  is a spectral family. Then*

$$(30.1) \quad \int_{\mathbb{R}} \lambda dE_\lambda := \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} k 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}})$$

*defines a self-adjoint operator.*

*Proof.* Define

$$(30.2) \quad A_j := \sum_{k=-\infty}^{\infty} k 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}}).$$

The, the sum defining  $A_j$  is finite and hence  $A_j \in S(H)$ . Observe that

$$\begin{aligned} A_{j+1} - A_j &= \sum_{k=-\infty}^{\infty} k 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}}) - \sum_{k=-\infty}^{\infty} k 2^{-j-1} (E_{(k+1)2^{-j-1}} - E_{k2^{-j-1}}) \\ &= - \sum_{k=-\infty}^{\infty} 2^{-j-1} (E_{(2k+2)2^{-j-1}} - E_{(2k+1)2^{-j-1}}). \end{aligned}$$

Therefore, we obtain

$$(30.3) \quad \|A_{j+1} - A_j\|_{B(H)} \leq 2^{-j-1}.$$

From this we conclude that  $\{A_j\}_{j \in \mathbb{N}}$  converges. □

**Theorem 30.3** (Spectral decomposition). *Let  $A \in S(H)$  be a self-adjoint operator. Then there exists a spectral family  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  is a family in  $P(H)$  that fulfills the following.*

- (1)  $E_\lambda = 0$  for  $\lambda < -\|A\|_H$  and  $E_\lambda = 1$  for  $\lambda \geq \|A\|_H$ .

(2) For all  $x, y \in H$ ,

$$(30.4) \quad \int_{\mathbb{R}} d\langle E_{\lambda}x, y \rangle = \langle Ax, y \rangle.$$

*Proof.* Uniqueness of the spectral family. From the continuity it suffices to prove that  $E_{k2^{-j}}$  is determined uniquely from the expression

$$(30.5) \quad A := \lim_{m \rightarrow \infty} \sum_{k=-\infty}^{\infty} l 2^{-m} (E_{(l+1)2^{-m}} - E_{l2^{-m}})$$

for each  $k, j \in \mathbb{Z}$ . Suppose that

$$(30.6) \quad A = \lim_{m \rightarrow \infty} \sum_{k=-\infty}^{\infty} l 2^{-m} (E_{(l+1)2^{-m}} - E_{l2^{-m}})$$

is another expression. Then

$$(A - k 2^{-j})_+ = \lim_{j \rightarrow \infty} \sum_{l=-\infty}^{\infty} (l 2^{-j} - k 2^{-j})_+ (F_{(l+1)2^{-j}} - F_{l2^{-j}})$$

$$(A - k 2^{-j})_- = \lim_{j \rightarrow \infty} \sum_{l=-\infty}^{\infty} (l 2^{-j} - k 2^{-j})_- (F_{(l+1)2^{-j}} - F_{l2^{-j}}).$$

because  $(A - k 2^{-j})_+$  and  $(A - k 2^{-j})_-$  are positive operators characterized by

$$(30.7) \quad (A - k 2^{-j})_+ (A - k 2^{-j})_- = 0, \quad A - k 2^{-j} = (A - k 2^{-j})_+ - (A - k 2^{-j})_-.$$

As a result,  $F_{k2^{-j}} = (A - k 2^{-j})_+$ . This shows the uniqueness of the spectral family.

The construction of the spectral family. Define

$$(30.8) \quad E_{\lambda} = \text{proj} (H \rightarrow \text{Ker} (A - \lambda)_+).$$

We claim that  $\{E_{\lambda}\}_{\lambda \in \mathbb{R}}$  is monotone. Indeed, let  $\lambda > \rho$ . Then there exists a continuous function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  such that  $(x - \lambda)_+ = \varphi(x)(x - \rho)_+$ . Therefore, we have  $(A - \lambda)_+ = \varphi(A)(A - \rho)_+$  and hence  $\text{Ker} (A - \rho)_+ \subset \text{Ker} (A - \lambda)_+$ .

$\lim_{\varepsilon \downarrow 0} E_{\lambda+\varepsilon}x = E_{\lambda}x$  for  $\lambda \in \mathbb{R}$  It suffices to check that

$$(30.9) \quad \bigcap_{\varepsilon > 0} \text{Ker} (A - \lambda - \varepsilon)_+ = \text{Ker} (A - \lambda)_+.$$

However, it is not so hard to see

$$(30.10) \quad \|(A - \lambda - \varepsilon)_+ - (A - \lambda)_+\|_{B(X)} \leq \varepsilon.$$

Therefore, the claim is immediate.

$A$  and  $E_{\lambda}$  commute. We have  $A(A - \lambda)_+ = (A - \lambda)_+A$  by the functional calculus. Therefore, if  $x \in \text{Ker} (A - \lambda)_+$ , then  $(A - \lambda)_+Ax = A(A - \lambda)_+x = 0$ . Thus,  $Ax \in \text{Ker} (A - \lambda)_+$ . As a consequence  $E_{\lambda}Ax = Ax$  for all  $x \in \text{Ker} (A - \lambda)$ . Note that this can be rephrased as  $E_{\lambda}AE_{\lambda} = AE_{\lambda}$ . By passing to the adjoint, we obtain

$$(30.11) \quad E_{\lambda}A = (AE_{\lambda})^* = (E_{\lambda}AE_{\lambda})^* = E_{\lambda}AE_{\lambda} = AE_{\lambda}.$$

Therefore  $A$  and  $E_{\lambda}$  commute.

We claim  $\lambda E_{\lambda} \geq AE_{\lambda}$  and  $\lambda(1 - E_{\lambda}) \leq A(1 - E_{\lambda})$  Indeed,

$$(30.12) \quad (A - \lambda)E_{\lambda} = E_{\lambda}(A - \lambda)E_{\lambda} \leq E_{\lambda}(A - \lambda)_+E_{\lambda} = 0.$$



Note  $E_\lambda(A-\lambda)_- = (A-\lambda)_-$  and hence  $(A-\lambda)_-E_\lambda = (A-\lambda)_-$ , because  $(A-\lambda)_+(A-\lambda)_- = 0$ . Therefore,

$$(30.13) \quad (A-\lambda)(1-E_\lambda) = (A-\lambda)_+(1-E_\lambda) \geq 0.$$

This is the desired result.

$$\boxed{A = \int_{\mathbb{R}} \lambda dE_\lambda.} \quad \text{Suppose that } \rho < \lambda. \text{ Then we have}$$

$$(30.14) \quad \rho(E_\lambda - E_\rho) = \rho(1 - E_\rho)E_\lambda \leq A(1 - E_\rho)E_\lambda = AE_\lambda(1 - E_\rho) \leq \lambda E_\lambda(1 - E_\rho) = \lambda(E_\lambda - E_\rho).$$

Therefore, it follows that

$$\int_{\mathbb{R}} \lambda dE_\lambda = \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} k 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}}) \leq \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} A (E_{(k+1)2^{-j}} - E_{k2^{-j}}) = A$$

and

$$A \leq \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} (k+1) 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}}) = \int_{\mathbb{R}} \lambda dE_\lambda.$$

Putting together these inequalities, we obtain the desired  $A = \int_{\mathbb{R}} \lambda dE_\lambda$ . □

**Definition 30.4.** Under the notation in Theorem 30.3, given a continuous  $\mathbb{R}$ -valued function  $f$  define a bounded operator  $\int_0^\infty f(\lambda) dE_\lambda \in S(H)$  by

$$(30.15) \quad \int_0^\infty f(\lambda) dE_\lambda = \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} f(k 2^{-j}) (E_{(k+1)2^{-j}} - E_{k2^{-j}}).$$

**Proposition 30.5.** Suppose that  $f$  is a continuous function on  $\mathbb{R}$ . Then

$$(30.16) \quad \left\| \int_0^\infty f(\lambda) dE_\lambda \right\|_X = \|f\|_\infty.$$

*Proof.* Let  $\alpha = \|f\|_\infty$  and  $\varepsilon > 0$ . Then there exists  $j \in \mathbb{N}$  and  $k \in \mathbb{Z}$  such that

$$(30.17) \quad \inf_{k 2^{-j} \leq \lambda \leq (k+1)2^{-j}} |f(\lambda)| \geq \alpha - \varepsilon.$$

Therefore, it follows that

$$(30.18) \quad \left\| \int_0^\infty f(\lambda) dE_\lambda \right\|_X \geq \left\| \int_0^\infty f(\lambda) dE_\lambda \cdot E_{(k+1)2^{-j}} - E_{k2^{-j}} \right\|_X \geq \inf_{k 2^{-j} \leq \lambda \leq (k+1)2^{-j}} |f(\lambda)| \geq \alpha - \varepsilon.$$

As for the reverse inequality we use

$$(30.19) \quad \left\| \sum_{k=-\infty}^{\infty} (k+1) 2^{-j} (E_{(k+1)2^{-j}} - E_{k2^{-j}}) \right\|_X \leq \|f\|_\infty.$$

A passage to the limit therefore gives the desired result. □

**Theorem 30.6.** Keep to the same notation and the assumption as Theorem 30.3. Then

$$(30.20) \quad f \in C(\mathbb{R}) \mapsto f(A) \in S(H)$$

is a  $\mathbb{K}$ -algebra isomorphism, that is, for all  $f, g \in C(\mathbb{R})$  and  $a, b \in \mathbb{K}$ ,

$$(af + bg)(A) = af(A) + bg(A) \\ (fg)(A) = f(A)g(A).$$

Furthermore, if  $f$  is a polynomial, then  $f(A)$  coincides with the usual definition.

*Proof.* Suppose that  $\{A_j\}_{j \in \mathbb{N}}$  and  $\{B_j\}_{j \in \mathbb{N}}$  are sequences of bounded operators convergent in the norm topology. Then so is  $\{A_j B_j\}_{j \in \mathbb{N}}$ . Therefore, we obtain

$$\begin{aligned} & \int_{\mathbb{R}} f(\lambda) dE_\lambda \cdot \int_{\mathbb{R}} g(\lambda) dE_\lambda \\ &= \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} f(k 2^{-j}) (E_{(k+1)2^{-j}} - E_{k2^{-j}}) \sum_{l=-\infty}^{\infty} g(l 2^{-j}) (E_{(l+1)2^{-j}} - E_{l2^{-j}}) \\ &= \lim_{j \rightarrow \infty} \sum_{k=-\infty}^{\infty} f(k 2^{-j}) g(k 2^{-j}) (E_{(k+1)2^{-j}} - E_{k2^{-j}}) \\ &= \int_{\mathbb{R}} f \cdot g(\lambda) dE_\lambda \end{aligned}$$

The proof of the linearity is simpler.

Thus, it remains to prove that

$$(30.21) \quad f(A) = \int_{\mathbb{R}} f(\lambda) dE_\lambda.$$

Proposition 30.5 along with linearity reduces the matter to the case when  $f$  is a monomial. That is, we have only to show that

$$(30.22) \quad A^k = \int_{\mathbb{R}} \lambda^k dE_\lambda.$$

If  $k = 1$ , this is just a definition. For general  $k \in \mathbb{N}$ , now that we have proved that the mapping  $f \in C(\mathbb{R}) \mapsto \int_{\mathbb{R}} f(\lambda) dE_\lambda$  is an algebraic homomorphism, a simple induction argument works.  $\square$

**Proposition 30.7.** *Let  $\lambda \in \mathbb{R}$ . Then*

$$(30.23) \quad \lim_{\varepsilon \downarrow 0} (\text{proj}(H \rightarrow \text{Ker}(A - \lambda)_+) - \text{proj}(H \rightarrow \text{Ker}(A - \lambda - \varepsilon)_+)) = \text{proj}(H \rightarrow \text{Ker}(A - \lambda)).$$

*Proof.* Note that

$$\begin{aligned} & \lim_{\varepsilon \downarrow 0} (\text{proj}(H \rightarrow \text{Ker}(A - \lambda)_+) - \text{proj}(H \rightarrow \text{Ker}(A - \lambda - \varepsilon)_+)) \\ &= \text{proj} \left( H \rightarrow \bigcap_{\varepsilon > 0} \text{Ker}(A - \lambda)_+ \cap \text{Ker}(A - \lambda - \varepsilon)_+^\perp \right). \end{aligned}$$

Suppose that  $x \in \text{Ker}(A - \lambda)$ . Then we have  $x \in \text{Ker}(A - \lambda)_+ \cap \text{Ker}(A - \lambda - \varepsilon)_+^\perp$ , for each  $\varepsilon > 0$ . Suppose instead that  $x \in \bigcap_{\varepsilon > 0} (\text{Ker}(A - \lambda)_+ \cap \text{Ker}(A - \lambda - \varepsilon)_+^\perp)$ . Then we have

$$(30.24) \quad (A - \lambda)_+ x = \lim_{j \rightarrow \infty} \sum_{k=0}^{\infty} k 2^{-j} (E_{(k+1)2^{-j} + \lambda} - E_{k 2^{-j} + \lambda}).$$

Therefore  $\text{proj}(H \rightarrow \text{Ker}(A - \lambda)_+) = 1 - E_\lambda$ . As a result,

$$(30.25) \quad (A - \lambda)x = \lim_{j \rightarrow \infty} \sum_{k=0}^{\infty} k 2^{-j} (E_{(k+1)2^{-j} + \lambda} - E_{k 2^{-j} + \lambda})x = 0.$$

This is the desired result.  $\square$

**30.2. Compact self-adjoint operators.** As a special case let us see what happens if  $A \in B(H)$  is a compact and self-adjoint.

**Theorem 30.8.** *Let  $A \in S(H)$  be compact. Then there exists a sequence  $\{r_j\}_{j \in \mathbb{N}} \subset \mathbb{R} \setminus \{0\}$  converging to 0 and a sequence of projections to finite dimensional subspace such that*

$$(30.26) \quad A = \sum_{j=1}^{\infty} r_j E_j.$$

*Proof.* We may assume  $A$  is positive by decomposing  $A$  into  $A_+$  and  $A_-$ . First we observe

$$(30.27) \quad A = \lim_{\varepsilon \rightarrow 0} A(1 - E_\varepsilon).$$

Note that  $\text{Ker}(A - \lambda)_+^\perp = \bigcap_{\rho > \lambda} \text{Ker}(A - \rho)$  is finite dimensional, if  $\lambda$ . Using this, we see that there exists a sequence  $r_1 > r_2 > \dots > r_j > \dots \rightarrow 0$  such that  $E_{r_j} - \lim_{\rho \uparrow a_j} E_r$  is a nonzero projections and  $E_r - \lim_{\rho \uparrow r} E_r$  if  $r > 0$  and  $r \notin \{r_1, r_2, \dots, r_j, \dots\}$ . Using this projection, we see that

$$(30.28) \quad A = \sum_{j=1}^{\infty} r_j E_j.$$

This is the desired result. □

*Exercise 198.* Let  $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  be a compact linear operator from a Hilbert space  $\mathcal{H}_1$  to a Hilbert space  $\mathcal{H}_2$ . By using the fact that  $T^*T$  is a compact operator on  $\mathcal{H}_1$ , establish that there exist a positive sequence  $\{\lambda_j\}_{j \in \mathbb{N}}$ , an orthonormal sequence  $\{u_j\}_{j=1}^\infty$  in  $\text{Ker}(T)^\perp$  and an orthonormal sequence  $\{v_j\}_{j=1}^\infty$  in  $\text{Ker}(T^*)^\perp$  such that

$$(30.29) \quad Tu_j = \lambda_j v_j, Tv_j = \lambda_j u_j$$

for all  $j \in \mathbb{N}$ .

*Exercise 199.* Let  $T$  be a compact linear operator on a Hilbert space  $H$ . Let  $\{x_n\}_{n=1}^\infty$  a disjoint sequence which converges weakly to 0. Show that  $\{Tx_n\}_{n=1}^\infty$  converges strongly to 0. In this sense compact operators are Dunford-Pettis operators.

*Exercise 200.* Let  $S, T$  be bounded operators on a Hilbert space  $H$ . Assume in addition that  $0 \leq S \leq T$  and that  $T$  is compact.

- (1) Let  $\{P_n\}_{n=1}^\infty$  be a sequence of projections such that  $P_n T P_n \rightarrow T$  as  $n \rightarrow \infty$ . Then show that  $(1 - P_n)S(1 - P_n) \rightarrow 0$  as  $n \rightarrow \infty$ .
- (2) Show that  $S$  is compact.

Notes and references for Chapter 14.

Section 27. Theorem 27.3 is due to Pettis.

Theorem 27.11

Theorem 27.12

Theorem 27.13

Theorem 27.15

Theorem 27.17

Theorem 27.18

Let  $T$  be a bounded linear operator from a Banach space  $A$  to a Banach space  $B$ . Then Virot showed that  $T$  can be extended to a bounded operator from  $L^p(A)$  to  $L^p(B)$  (see [487]).

We refer to [27, Chapter] for a description of the theory of Bochner integral.

We list [71] as a textbook of the theory of Bochner integral.

Section 28. We refer to [38, 291] for more details about the analytic semigroup as well as its applications.

There are several ways to define  $e^{tA}$  in Theorem 28.11. Here we followed the method due to Lunardi [38]. Alternative way to define  $e^{tA}$  can be found in [157, 332].

Theorem 28.12

Section 29. Theorem 29.5

Theorem 29.12

Theorem 29.13

Theorem 29.18

Theorem 29.22

Theorem 29.24

Theorem 29.25

Theorem 29.27

Theorem 29.29

Theorem 29.31

Theorem 29.33

Theorem 29.35 and Theorem 29.36 are due to Gelfand.

Theorem 29.40

Theorem 29.41

Theorem 29.42

Theorem 29.46

Theorem 29.47

Theorem 29.50

Theorem 29.54

Section 30. Theorem 30.2

Theorem 30.3

Theorem 30.6

Theorem 30.8

## Part 15. Topological vector spaces

We begin with investigating Banach spaces having the structure such as  $D(\Omega)$  in Chapter 15. In Section 31 we shall characterize some notions on topological spaces by means of nets. Nets will turn out useful in later applications. In Section 32 we consider topological vector spaces, which are generalization of normed spaces and Schwartz distribution spaces  $\mathcal{S}'(\mathbb{R}^d)$ . That, is we are going to discuss the function spaces which we can not endow with norms.

### 31. NETS AND TOPOLOGY

When a topological space is a first countable space, then a subset  $A$  is closed if and only if the limit belongs to  $A$  whenever the sequence in  $A$  is convergent. This is a very convenient result. However, there are many topological spaces which do not satisfy first axiom of countability. For such spaces, we do not have such a criterion. To overcome this difficulty, we need a notion of nets.

**Definition 31.1** (Partially ordered set, Net and so on). Let  $X$  be a topological space and  $A$  a set.

- (1) The subset  $A$  is said to be a partially ordered set if  $A$  comes with a subset  $R$  of  $A \times A$  with the following conditions fulfilled.
  - (a)  $(\alpha, \alpha) \in R$
  - (b)  $(\alpha, \beta) \in R$  and  $(\beta, \alpha) \in R$  implies  $\alpha = \beta$ .
  - (c)  $(\alpha, \beta) \in R$  and  $(\beta, \gamma) \in R$  implies  $(\alpha, \gamma) \in R$ .
 Let  $A$  is a partially ordered set as above. Then denote  $(\alpha, \beta) \in R$  by  $\alpha \gg \beta$  or  $\beta \ll \alpha$  and  $A$  is said to be a partially ordered set with order  $\gg$ .
- (2)  $A$  is said to be a directed set if  $A$  is a partially ordered set with order  $\gg$  and for all  $\alpha, \beta \in A$ , we can find  $\gamma$  such that  $\gamma \gg \alpha, \beta$ .
- (3) A subset  $A_0$  of a directed set  $A$  is said to be cofinal, if for all  $\alpha \in A$  we can find  $\beta \in A_0$  with  $\alpha \ll \beta$ .
- (4) A net is a collection of elements  $\{x_\alpha\}_{\alpha \in A} \subset X$  in  $X$  indexed by a directed set  $A$ .
- (5) Let  $\{x_\alpha\}_{\alpha \in A} \subset X$  be a net. A subnet of  $\{x_\alpha\}_{\alpha \in A}$  is a net  $\{x_\alpha\}_{\alpha \in A_0}$  indexed by a cofinal subset  $A_0$  of  $A$ .

We exhibit a series of examples that will be helpful for later considerations.

**Example 31.2.** Let  $X$  be a topological space and denote by  $\mathcal{O}_X$  the set of all open sets in  $X$ .

- (1) Any subset of  $\mathbb{R}$  is a directed set.
- (2) Let  $x \in X$ . Then set

$$(31.1) \quad \mathcal{N}_x := \{U \in \mathcal{O}_X : x \in U\}.$$

Define  $U_1 \gg U_2$  to be  $U_1 \subset U_2$  for  $U_1, U_2 \in \mathcal{N}_x$ . Then  $\mathcal{N}_x$  is a net. Suppose that we are given a map  $f : \mathcal{N}_x \rightarrow X$  satisfying  $f(U) \in U$  for all  $U$ . Then  $\{f(U)\}_{U \in \mathcal{N}_x}$  is a net converging to  $x$ .

- (3) Define an order of  $2^{\mathbb{N}}$  by  $A \gg B \iff A \supset B$ . Then  $2^{\mathbb{N}}$  is a directed set.

**Definition 31.3.** Let  $X$  be a topological space and  $\{x_\alpha\}_{\alpha \in A} \subset X$  be a net. Then  $\{x_\alpha\}_{\alpha \in A}$  converges to  $y$  (we write  $\lim_{\alpha \in A} x_\alpha = y$ ), if for any neighborhood  $U$  of  $y$ , there exists  $\alpha_0 \in A$  such that  $x_\alpha \in U$  for all  $\alpha \gg \alpha_0$ .

Characterization of closed sets in terms of nets.

If  $(X, d)$  is a metric space, then a set  $A$  is closed if and only if the limit of a sequence in  $A$  always lies in  $A$  whenever it is convergent to some point  $X$ . However, this is not the case for topological spaces in general. However, we have the following substitute.

**Theorem 31.4.** *A subset  $A$  of the topological space is closed if and only if  $x = \lim_{\mu \in M} x_\mu \in A$  whenever  $\{x_\mu\}_{\mu \in M}$  is a converging net in  $A$ .*

*Proof.* Assume that  $A$  is closed

Suppose  $\{x_\mu\}_{\mu \in M}$  is a converging net in  $A$ . If  $x$  were not in  $A$ , we would have some  $x_\mu$  belongs to an open set  $X \setminus A$ . This runs counter to the assumption that  $x_\mu \in A$ .

Assume that for any converging  $\{x_\mu\}_{\mu \in M}$  net in  $A$  the limit, if there exists, lies in  $A$ .

We shall show that  $A = \overline{A}$ , which implies that  $A$  is closed.

Let  $x \in \overline{A}$ . Denote

$$(31.2) \quad \mathcal{U}_x := \{U \in \mathcal{O}_X : x \in U\}.$$

By the axiom of choice we can choose  $(x_U)_{U \in \mathcal{U}}$  such that  $x_U \in U \cap A$ .

We endow  $\mathcal{U}$  with its order by defining  $U_1 \gg U_2 \Leftrightarrow U_1 \subset U_2$ .

Then by definition  $(x_U)_{U \in \mathcal{U}}$  converges to  $x$ . Thus by assumption we have  $x \in U$ . □

Characterization of compact sets in terms of nets.

We have a characterization of compactness in terms of sequences when  $(X, d)$  is a metric space. The counterpart of the general topological space is the following.

**Theorem 31.5.** *A topological space  $X$  is compact, if and only if any net in  $X$  has a converging subnet.*

*Proof.* Suppose that any net in  $X$  has a converging subnet.

We are going to show that a family of closed sets always has a point in common, whenever it enjoys the finite intersection property. It is known to be equivalent to the compactness of  $X$ . Given a family of closed sets  $\mathcal{F} := \{F_\lambda\}_{\lambda \in \Lambda}$  with the finite intersection property, we shall find a common element of  $\mathcal{F}$ .

To do this, for all  $A \subset \Lambda$  we can find  $x_A \in \bigcap_{\lambda \in A} F_\lambda$  thanks to the finite intersection property.

Denote by  $\mathcal{N}(\Lambda)$  the set of all finite subsets.  $(x_A)_{A \in \mathcal{N}(\Lambda)}$  is a net, if we induce the order by defining  $A \gg B \Leftrightarrow A \supset B$ . By assumption  $(x_A)_{A \in \mathcal{N}(\Lambda)}$  has a converging subnet  $(x_A)_{A \in \mathcal{N}(\Lambda)_0}$ . Let us denote its limit point by  $x$ .

Then the common element for which we are looking is  $x$ . Indeed, let  $\mu \in \Lambda$ . Then we can find  $A_0 \in \mathcal{N}(\Lambda)_0$  such that  $\mu \in A_0$ . Since  $\bigcap_{\lambda \in A_0} F_\lambda$  is closed, we have  $x = \lim_{A_0 \ll A \in \mathcal{N}(\Lambda)_0} x_A \in$

$\bigcap_{\lambda \in A_0} F_\lambda \subset F_\mu$ . Thus, it follows that  $X$  is compact.

Assume  $X$  is compact.

Let  $(x_\lambda)_{\lambda \in \Lambda}$  be a net in  $X$ . We observe

There is no subnet converging to  $x$ .

- $\Leftrightarrow$  For any subnet there exists an open set  $U \ni x$   
 such that there exists a sufficiently large  $\beta$  with  $x_\beta \notin U$ .  
 $\Leftrightarrow$  For any subnet there exists an open set  $U \ni x$   
 such that there exists a further subnet outside  $U$ .

Thus, if there is no subnet converging to  $x$ , it follows that there exists  $U_x \in \mathcal{O}_X$  containing  $x$  and  $\lambda_x \in \Lambda$  such that  $\lambda \gg \lambda_x$  implies  $x_\lambda \notin U_x$ .

Keeping this observation in mind, assume that for all  $x$  there is no subnet converging to  $x$ . Then we can take  $\lambda_x$  and  $U_x$  described in the above paragraph.

By compactness we can cover  $X$  with finite open sets  $U_{x_1}, \dots, U_{x_k}$ . Let  $\mu$  be taken so that  $\mu \gg \lambda_{x_1}, \dots, \lambda_{x_k}$ .

Then if  $\lambda \gg \mu$ , we have  $x_\lambda \notin U_{x_j}$  for all  $j = 1, 2, \dots$ . This is a contradiction because  $U_{x_1}, \dots, U_{x_k}$  forms an open cover in  $X$  and  $x_\lambda \in X$  for all  $\lambda \in \Lambda$ .

From this we deduce that  $(x_\lambda)_{\lambda \in \Lambda}$  has a converging subnet. □

Characterization of continuous mappings in terms of nets. In the same way we now characterize the continuity of the mapping.

**Theorem 31.6.** *Let  $X$  and  $Y$  be topological spaces. A mapping  $f : X \rightarrow Y$  is continuous if and only if  $\{f(x_\alpha)\}_{\alpha \in A}$  to a net in  $Y$  convergent to  $f(x)$  whenever  $\{x_\alpha\}_{\alpha \in A}$  is a net in  $X$  convergent to  $x$ .*

*Proof.* Assume that  $f$  is continuous.

Suppose that we are given a net  $\{x_\alpha\}_{\alpha \in A}$  convergent to  $x$ . Let  $V$  be an open set of  $Y$  containing  $f(x)$ . Then  $f^{-1}(V)$  is an open neighborhood of  $x$ . Therefore there exists  $\alpha \in A$  such that  $x_\beta \in f^{-1}(V)$  whenever  $\beta \gg \alpha$ . This implies  $f(x_\beta) \in V$  for all  $\beta \gg \alpha$ . Therefore,  $f(x_\beta)$  is convergent to  $f(x)$ .

Assume instead that  $f$  is discontinuous.

Then there exists an open set  $V$  of  $Y$  such that  $f^{-1}(V)$  is not open. Let  $x \in f^{-1}(V)$  be a point that is not an interior point of  $f^{-1}(V)$ . We define an ordered set  $\mathcal{N}_x$  by (31.1). From the definition of  $\mathcal{N}_x$  and the fact that  $f^{-1}(V)$  is not open, given  $U \in \mathcal{N}_x$ , we can take  $x_U \in U \setminus f^{-1}(V)$ . Then  $\{x_U\}_{U \in \mathcal{N}_x}$  is net converging to  $x$ . However  $\{f(x_U)\}_{U \in \mathcal{N}_x}$  never converges to  $f(x)$  because  $V$  is a neighborhood of  $f(x)$  which contains no element in  $\{f(x_U)\}_{U \in \mathcal{N}_x}$ . Thus, the converse is established as well. □

## 32. TOPOLOGICAL VECTOR SPACES

### 32.1. Definition.

As we shall see,  $\mathcal{S}'$  can not carry a structure of a Banach space compatible to its original topology. The aim of this section is to discuss a linear space like  $\mathcal{S}'$ . It is not so hard to see that  $\mathcal{S}'$  is a topological vector space, whose definition is given below.



**Definition 32.1.** A  $\mathbb{K}$ -linear space  $X$  is said to be a topological linear space, if the operations

$$(32.1) \quad (x, y) \in X \mapsto x + y \in X, (a, x) \in \mathbb{K} \times X \rightarrow ax \in X$$

are both continuous.

**Definition 32.2.** Let  $X$  be a  $\mathbb{K}$ -linear space. A seminorm on  $X$  is a function from  $X$  to  $[0, \infty)$  satisfying the following conditions.

- (1)  $p(x + y) \leq p(x) + p(y)$  for all  $x, y \in X$
- (2)  $p(\alpha x) = |\alpha| \cdot p(x)$  for all  $x \in X$  and  $\alpha \in \mathbb{K}$ .

From the definition all the notions are clearly generalizations of Banach spaces.

**Definition 32.3.** Let  $X$  be a topological vector space. Then equip  $X^*$  with the weak-\* topology. That is,  $X^*$  is endowed with the weakest topology such that  $x^* \in X^* \mapsto x^*(x) \in \mathbb{K}$  is continuous for all  $x \in X$ . Below it will be understood tacitly that  $X^*$  carries the weak-\* topology. Instead of adding “weak-\*”, add  $\sigma(X^*, X)$  to refer to this topology. In the same way equip  $X$  with another topology called weak topology, that is,  $x^* \in X^* \mapsto x^*(x) \in \mathbb{K}$  is continuous for all  $x \in X$ . To refer to this topology, add  $\sigma(X, X^*)$ .

Polar subset. Having set down the elementary facts, we now turn to the polar subsets. Here we will encounter an analogy with Galois theory.

**Definition 32.4.** Let  $X$  be a topological vector space.

- (1) Let  $A \subset X$ . Then define  $A^\circ := \{x^* \in X^* : |x^*(x)| \leq 1\}$ .
- (2) Let  $F \subset X^*$ . Then define  ${}^\circ F := \{x \in X : |x^*(x)| \leq 1\}$ .

**Theorem 32.5.** Let  $A$  and  $B$  be subsets in a topological vector space  $X$ . Then the following are true.

- (1)  $A \subset B$  implies  $B^\circ \subset A^\circ$ .
- (2)  $A \cup B^\circ = A^\circ \cap B^\circ$ .
- (3) If  $\lambda \in \mathbb{K} \setminus \{0\}$ , then  $(\lambda A)^\circ = \lambda^{-1} A^\circ$ .
- (4)  $A \subset {}^\circ(A^\circ)$ .
- (5)  $({}^\circ(A^\circ))^\circ = A^\circ$ .

*Proof.* Here we shall prove the last assertion, accepting the remaining ones. The first four assertions are left as an exercise. By (4) with  $A$  replaced by  $A^\circ$  we have  $({}^\circ(A^\circ))^\circ \subset A^\circ$ , while (4), combined with (1), gives  $({}^\circ(A^\circ))^\circ \supset A^\circ$ . □

*Exercise 201.* Prove the remaining assertions in Theorem 32.5.

**Definition 32.6.** A subset  $A$  in a topological vector space  $X$  is said to be absolutely convex, if  $ax + by \in X$  for each  $x, y \in X$  and  $a, b \in \mathbb{K}$  with  $|a| + |b| \leq 1$ .

**Theorem 32.7.** Let  $X$  be a topological vector space.

- (1) For each  $A \subset X$ ,  $A^\circ$  is an absolutely convex  $\sigma(X, X^*)$ -closed set.
- (2) For each  $F \subset X^*$ ,  ${}^\circ F$  is an absolutely convex  $\sigma(X^*, X)$ -closed set.

*Exercise 202.* Following the definitions, prove Theorem 32.7.

**Definition 32.8.** If  $A$  is a linear subspace of  $X$ , then denote  $A^\perp := A^\circ$ . Similarly if  $F$  is a linear subspace of  $X^*$ , then denote  ${}^\perp F := {}^\circ F$ .

**Theorem 32.9.** *Suppose that  $A$  is a linear subspace of  $X$  and  $F$  is a linear subspace of  $X^*$ . Then we have*

$$(32.2) \quad A^\perp = \{x^* \in X^* : x^*(x) = 0 \text{ for all } x \in A\}$$

$$(32.3) \quad {}^\perp F = \{x \in X : x^*(x) = 0 \text{ for all } x^* \in F\}.$$

*Proof.* We content ourselves with proving (32.2), (32.3) being proved in the same way. It follows from the definition of  $A^\circ = A^\perp$  that

$$(32.4) \quad A^\perp \supset \{x^* \in X^* : x^*(x) = 0 \text{ for all } x \in A\}.$$

To prove the reverse inclusion, let  $x^* \in A^\perp$ . Then for all  $y \in A$ , we have  $|x^*(y)| \leq 1$ . Note that  $A$  is a linear subspace of  $X$ . Hence  $y \in A$  implies  $n \cdot y \in A$  for all  $n \in \mathbb{N}$ . With  $y$  replaced by  $n \cdot y$ , we obtain  $|x^*(x)| \leq n^{-1}$  for all  $n \in \mathbb{N}$ . Therefore, since  $n \in \mathbb{N}$  is arbitrary, we conclude that  $x^*(x) = 0$ .  $\square$

*Exercise 203.* Prove (32.3).

Minkowski functional. In  $\mathbb{R}^d$ , given a convex set  $K$  with non-empty interior, we can define many norms. Given such a set  $K$ , we define

$$(32.5) \quad d(x, y) = 2 \inf\{\alpha \in [0, \infty) : x - y \in \alpha \cdot K + z \text{ for some } z \in \mathbb{R}^d\}.$$

For example, if  $K$  is a unit square, then the corresponding metric is the  $\ell^\infty$ -norm. If  $K$  is a unit ball, then we obtain the  $\ell^2$ -norm. Some other metrics of interest may be obtained starting from a convex polygon  $K$ . In this paragraph we generalize this procedure to the case when  $X$  is a topological vector space.

**Definition 32.10.** A subset  $K \in X$  is circular, if  $\alpha \cdot x \in K$  for all  $\alpha$  with absolute value 1 and  $x \in K$ .

**Definition 32.11.** A subset  $K \in X$  is said to be absorptive, if for all  $x \in X$  there exists  $\alpha_x > 0$  such that  $\alpha_x \cdot x \in K$ .

This is equivalent to saying that

$$(32.6) \quad p_K(x) := \inf\{\alpha > 0 : \alpha^{-1} \cdot x \in K\}$$

is finite for all  $x \in X$ .  $p_K$  is called the Minkowski functional of  $K$ .

**Theorem 32.12.** *Suppose that  $K$  is a convex and absorptive subset of  $X$ . Then we have the following for all  $x, y \in X$  and  $\lambda \geq 0$ .*

- (1)  $0 \in K$  and hence  $p_K(0) = 0$ .
- (2)  $p_K(\lambda \cdot x) = \lambda p_K(x)$ .
- (3)  $p_K(x + y) \leq p_K(x) + p_K(y)$ .

Furthermore if  $K$  is assumed circular in addition, then  $p_K(\rho \cdot x) = |\rho| p_K(x)$  for all  $\rho \in \mathbb{K}$ .

*Proof.* (1) is immediate. Indeed, since  $K$  is absorptive,  $0 \in \alpha K$  for some  $\alpha > 0$ . Therefore,  $0 = \alpha^{-1} \cdot 0 \in K$ . (2) is straightforward as well. Let us prove (3). Suppose that  $\varepsilon > 0$ . Then by definition of  $p_K(x)$  and  $p_K(y)$ , there exists  $\alpha, \beta > 0$  such that  $x \in \alpha \cdot K, y \in \beta \cdot K$  with  $\alpha < p_K(x) + \varepsilon$  and  $\beta < p_K(y) + \varepsilon$ . Since  $K$  is assumed convex, we see that  $\alpha \cdot K + \beta \cdot K = (\alpha + \beta)K$ . Therefore,  $x + y \in (\alpha + \beta)K$ . Thus,  $p_K(x + y) \leq \alpha + \beta \leq p_K(x) + p_K(y) + 2\varepsilon$ . Since  $\varepsilon > 0$  is arbitrary, (3) was proved. Finally it is again straightforward that we prove the remaining assertion under the additional assumption that  $K$  is circular.  $\square$

**Theorem 32.13** (Gauge functional). *Let  $K$  be a convex and absorptive set. Define*

$$K_1 := \{x \in X : p_K(x) < 1\}$$

$$K_2 := \{x \in X : p_K(x) \leq 1\}.$$

- (1)  $K_1 \subset K \subset K_2$ .
- (2)  $K$  is a neighborhood of 0 if and only if  $p_K$  is continuous.
- (3) If (2) is the case, then  $K_1 = \text{Int}(K)$  and  $K_2 = \overline{K}$ .

*Proof.* (1) It is trivial that  $K \subset K_2$ . Let  $k \in K_1$ . Then there exists  $\alpha < 1$  such that  $x \in \alpha \cdot K$ . Therefore,

$$(32.7) \quad x = \alpha \cdot \frac{1}{\alpha}x + (1 - \alpha) \cdot 0 \in K.$$

This is the desired result.

(2) : Suppose that  $K$  is a neighborhood of 0. Given  $\varepsilon > 0$ , we have

$$(32.8) \quad |p(y) - p(x)| \leq p(x - y) \leq \varepsilon$$

for  $y \in x + \varepsilon \cdot K$ . Therefore  $p_K$  is continuous.

(2) : Suppose that  $p_K$  is continuous. Then  $K_1 = p_K^{-1}((-1, 1))$  is an open neighborhood of 0. Since  $K$  is larger than  $K_1$ , so is  $K$ .

(3) Part 1 : Proof of  $K_1 = \text{Int}(K)$  Since we have shown that  $K_1$  is an open set, we see  $K_1 \subset \text{Int}(K)$ . Let  $x \in \text{Int}(K)$ . Then by the continuity of the multiplication, there exists  $\alpha > 1$  such that  $\alpha \cdot x \in \text{Int}(K)$ . Therefore  $p_K(x) \leq \alpha^{-1} < 1$  and the reverse inclusion is proved as well.

(3) Part 2 : Proof of  $\overline{K} = K_2$  As before, since  $K_2$  is closed,  $\overline{K} \subset K_2$  is trivial. If  $x \in K_2$ , we have  $\beta \cdot x \in K_1$  then for all  $0 < \beta < 1$ . Therefore,  $x \in \overline{K_1} \subset \overline{K}$ . As a result,  $\overline{K} = K_2$  is proved as well. □

Separation theorems. Suppose that  $A$  and  $B$  are convex open sets. Then  $A$  and  $B$  are separated by a line. That is there exists a linear mapping  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that

$$(32.9) \quad A \subset \{f > c\}, B \subset \{f < c\}.$$

The aim of this paragraph is to generalize this present situation.

**Theorem 32.14** (Hahn-Banach theorem / separation version). *Suppose that  $A$  is a closed convex set in a topological vector space  $X$ . If  $b$  is a point outside  $A$ , then there exists a continuous functional  $x^* : X \rightarrow \mathbb{K}$  such that*

$$(32.10) \quad \sup_{a \in A} \text{Re } x^*(a) < x^*(b).$$

Bounded set. A bounded set of a Banach space is by definition a set which is contained by  $\{b \in B : \|b\|_B < M\}$  for some  $M > 0$ . Motivated by this, we are led to the following definition.

**Definition 32.15.** A subset  $A$  in a topological vector space is bounded, if for all open set  $U$  containing 0 there exists  $\alpha_U > 0$  such that  $\alpha_U \cdot A \subset U$ .

**Example 32.16.** Let  $X$  be a normed space and  $A$  a bounded set in the sense of Definition 32.15. Then, since  $\{x \in X : \|x\|_X < 1\}$  is a neighborhood, there exists  $\alpha > 0$  such that  $A \subset \{x \in X : \|x\|_X < \alpha\}$ . Conversely, Suppose  $A$  is a bounded set in the sense that there exists  $\alpha > 0$  such that  $A \subset \{x \in X : \|x\|_X < \alpha\}$ . If  $U$  is a neighborhood of 0, then there exists

$\beta > 0$  such that  $U \supset \{x \in X : \|x\|_X < \beta\}$ . Therefore we conclude  $\alpha^{-1}\beta A \subset U$  and hence  $A$  is bounded in the sense of Definition 32.15. Thus, the notion of boundedness is a natural extension of the boundedness in normed spaces.

Equicontinuous family of linear functional. Suppose that  $F = \{f_\lambda\}_{\lambda \in \Lambda}$  is a family of continuous functions on a domain  $\Omega \subset \mathbb{C}$ . Then recall that  $F$  is said to be equicontinuous, if for all  $\varepsilon > 0$  and  $z \in \Omega$  there exists  $\delta > 0$  such that

$$(32.11) \quad |f_\lambda(z) - f_\lambda(w)| < \varepsilon$$

whenever  $w \in B(z, \delta)$  and  $\lambda \in \Lambda$ . What counts in the definition that the constant  $\delta$  can be taken uniformly over  $\Lambda$ . With this in mind, let us consider the equicontinuous family of the linear functionals on a topological vector space  $X$ .

**Definition 32.17.** A subset  $F$  of  $X^*$  is equicontinuous, if there exists a neighborhood  $U$  of 0 such that  $|x^*(x)| \leq 1$  for all  $x \in U$  and  $x^* \in F$ .

**Example 32.18.** If  $X$  is a normed space, then the unit ball of  $X^*$  is equicontinuous. Indeed, the corresponding neighborhood  $U$  can be taken as the unit ball of  $X$ .

*Exercise 204.* If  $F$  is a finite collection of  $X^*$ , then show that  $F$  is equicontinuous.

**Theorem 32.19** (Minkowski functional). *Let  $X$  be a topological vector space and  $F \subset X^*$  an equicontinuous family. Define  $p := p_{\circ F}$ , the Minkowski functional of  $\circ F$ . Then  $p$  is continuous and*

$$(32.12) \quad (\circ F)^\circ = \{x^* \in \text{Hom}(X, \mathbb{K}) : |x^*(x)| \leq p(x) \text{ for all } x \in X\}.$$

*Proof.* Let  $U$  be an open set as in Definition 32.17. Then from the definition of  $\circ F$  we have  $U \subset \circ F$ . Therefore,  $\circ F$  is a neighborhood of 0. We deduce from Theorem 32.12 that this is equivalent to  $p$  is continuous.

Let us prove (32.12). Suppose that  $x^* \in (\circ F)^\circ$ . That is, we assume

$$(32.13) \quad |x^*(x)| \leq 1 \text{ whenever } x \text{ satisfies } |y^*(x)| \leq 1 \text{ for all } y^* \in F.$$

Since  $\circ F$  is closed and hence  $p^{-1}([0, 1]) = \circ F$ . Furthermore, we have  $p^{-1}(0) = \{0\}$  because  $\circ F$  is a neighborhood of  $0 \in X$ . Therefore, we have  $\frac{x}{p(x)} \in \circ F$  for every  $x \in X \setminus \{0\} = p^{-1}((0, \infty))$ .

This implies  $\left| y^* \left( \frac{x}{p(x)} \right) \right| \leq 1$  again from the fact that  $p^{-1}([0, 1]) = \circ F$ . By our assumption we have

$$(32.14) \quad \left| x^* \left( \frac{x}{p(x)} \right) \right| \leq 1.$$

Arranging this inequality, we obtain  $|x^*(x)| \leq p(x)$  for all  $x \in X \setminus \{0\}$ . This being trivial for  $x = 0$ , we conclude  $x^*$  belongs to the right-hand side of (32.12).

Suppose instead that  $x^*$  belongs to the right-hand side of (32.12). Then we have to show  $|x^*(x)| \leq 1$  whenever  $x \in \circ F$ . Note that  $x \in \circ F$  is equivalent to  $p(x) \leq 1$ . Therefore, combining  $|x^*(x)| \leq p(x)$ , we obtain  $|x^*(x)| \leq 1$ . As a result (32.12) is established.  $\square$

**Theorem 32.20** (Banach Alaoglu-(2)). *Let  $F \subset X^*$  be a closed set. Assume that  $F$  is equicontinuous. Then  $F$  is  $\sigma(X^*, X)$ -compact.*

*Proof.* Since  $F$  is a subset of  $(\circ F)^\circ$ , by (32.12) it suffices to show

$$(32.15) \quad (\circ F)^\circ := \{x^* \in \text{Hom}(X, \mathbb{K}) : |x^*(x)| \leq p(x) \text{ for all } x \in X\}$$

is compact.

Given  $x \in X$ , we define  $D_x := \{z \in \mathbb{K} : |z| \leq p(x)\}$ . To prove the compactness of  $({}^\circ F)^\circ$ , we use Theorem 31.5. Suppose that we are given a net  $\{x_\alpha^*\}_{\alpha \in A}$ . Our present task is to find a subnet  $\{x_\alpha^*\}_{\alpha \in A_0}$  converging in weak-\* topology. That is, we have to find a subnet  $\{x_\alpha^*\}_{\alpha \in A_0}$  such that  $\{x_\alpha^*(x)\}_{\alpha \in A_0}$  converges for all  $x \in X$ .

Since each  $x_\alpha$  belongs to  $({}^\circ F)^\circ$ , we have  $|x_\alpha^*(x)| \leq p(x)$ . Therefore, we conclude

$$(32.16) \quad \{x_\alpha^*(x)\}_{x \in X} \in \prod_{x \in X} D_x$$

for each  $\alpha \in A$ . Note that  $\prod_{x \in X} D_x$  is equipped with a topology such that the projection to each component is continuous. Furthermore by the Tychnov theorem we see that  $\prod_{x \in X} D_x$  is a compact space. Therefore, we can find a subnet

$$(32.17) \quad \{x_\alpha^*\}_{\alpha \in A_0} \subset \{x_\alpha^*\}_{\alpha \in A}$$

such that  $\{x_\alpha^*(x)\}_{\alpha \in A_0}$  converges for all  $x \in X$ . As a result  $({}^\circ F)^\circ$  is compact. □

**Theorem 32.21** (Alaoglu-Bourbaki-Kakutani-(3)). *The closed ball  $B$  in  $X^*$  is weak-\* compact.*

*Proof of Theorems 32.20 and 32.21.* Since  $B$  is equicontinuous, Theorem 32.21 is a special case when  $F = B$ . It is well-known that if a topological space is compact, then any sequence has a convergent subsequence. Therefore, Theorem 9.27 is a further special case of Theorem 32.21. □

Extremal sets. First we present the definition to formulate the Klein-Milman theorem.

Let  $X$  be a topological vector space. Denote by  $X^*$  the dual space.

**Definition 32.22** (extremal set). Let  $\emptyset \neq K \subset X$ .  $E \neq \emptyset$  is said to be an extremal set of  $K$ , if

$$(32.18) \quad x, y \in K, t \in (0, 1), tx + (1 - t)y \in E$$

implies  $x, y \in E$ .  $x \in K$  is said to be extremal, if  $\{x\}$  is an extremal set.

The following lemma gives us an example of the extremal set.

**Lemma 32.23.** *Let  $f \in X^*$  and  $K$  be a non-empty compact set. Then*

$$(32.19) \quad \left\{ y \in K : \operatorname{Re} f(y) = \sup_{x \in K} \operatorname{Re} f(x) \right\}$$

*is an extremal set of  $K$ .*

*Proof.* Let  $x, y \in K$  and  $0 < t < 1$ . Assume that

$$(32.20) \quad \operatorname{Re} f(tx + (1 - t)y) = \sup_{z \in K} \operatorname{Re} f(z).$$

Then  $\sup_{z \in K} \operatorname{Re} f(z) = \operatorname{Re} f(tx + (1 - t)y) \leq t \operatorname{Re} f(x) + (1 - t) \operatorname{Re} f(y) \leq \sup_{z \in K} \operatorname{Re} f(z)$ . This implies that  $\operatorname{Re} f(x) = \operatorname{Re} f(y) = \sup_{z \in K} \operatorname{Re} f(z)$ . As a result the set in question is an extremal set of  $K$ . □

**Lemma 32.24.** *Let  $K \supset E \neq \emptyset$ . Suppose that  $E$  is an extremal set of  $K$  and that  $F$  is an extremal set of  $E$ . Then  $F$  is an extremal set of  $K$ .*

*Proof.* Assume that

$$(32.21) \quad x, y \in K, t \in (0, 1), tx + (1 - t)y \in F.$$

Since  $E$  is an extremal set of  $K$  and

$$(32.22) \quad x, y \in K, t \in (0, 1), tx + (1 - t)y \in E,$$

it follows that  $x, y \in E$ . Thus we obtain

$$(32.23) \quad x, y \in E, t \in (0, 1), tx + (1 - t)y \in F.$$

Since  $F$  is an extremal set of  $E$ , this implies  $x, y \in F$ . □

**Definition 32.25.** Let  $A \subset X$ . Then the set of extremal points of  $A$  is denoted by  $ex(A)$ .

The Klein-Milman theorem.

Having cleared up the definitions, we are now in the position of formulating the Klein-Milman's theorem.

**Theorem 32.26** (Klein-Milman). *Suppose that  $K$  is an compact set of  $X$ . Then*

- (1)  $ex(K)$  is non-empty.
- (2)  $K \subset \overline{co}(ex(K))$ .

*Proof.* The idea to prove the first statement is to find the smallest extremal set of  $K$ . Set

$$(32.24) \quad \mathcal{S}(\mathbb{R}^d) := \{E \subset K : E \text{ is closed and an extremal set of } K\}.$$

As we have constructed in Lemma 32.24,  $\mathcal{S}(\mathbb{R}^d)$  contains at least one element. We define the order of  $\mathcal{S}(\mathbb{R}^d)$  by

$$(32.25) \quad E_1 \gg E_2 \text{ if and only if } E_1 \subset E_2.$$

Assume that  $E \in \mathcal{S}(\mathbb{R}^d)$  contains at least two elements. Then it is easy to see that there exists  $F$  strictly greater than  $E$ . That is, we can find  $F \in \mathcal{S}(\mathbb{R}^d)$  such that

$$(32.26) \quad E \supset F, E \neq F.$$

Indeed, let  $x, y \in E$  be two different element. Then we can take a continuous linear functional  $f : X \rightarrow \mathbb{C}$  such that  $\operatorname{Re} f(x) > \operatorname{Re} f(y)$ . We have only to put

$$(32.27) \quad F := \left\{ y \in E : \operatorname{Re} f(y) = \sup_{x \in K} \operatorname{Re} f(x) \right\}.$$

Since the intersection of any collection of non-empty compact sets is compact and non-empty again,  $\mathcal{S}(\mathbb{R}^d)$  has a maximal element by Zorn's lemma. Let  $H$  be the maximal element. Then, from the preceding paragraph  $H$  cannot contain two different elements. Thus, it follows that  $H := \{x_0\}$  and  $x_0 \in ex(K)$ .

To prove the second statement, suppose that

$$(32.28) \quad a \in K \setminus \overline{co}(ex(K)).$$

Then there exists a continuous functional  $f : X \rightarrow \mathbb{C}$  such that

$$(32.29) \quad \operatorname{Re} f(a) > \sup_{y \in \overline{co}(ex(K))} \operatorname{Re} f(y).$$

If we set

$$(32.30) \quad K_0 = \left\{ y \in K : \operatorname{Re} f(y) = \sup_{x \in K} \operatorname{Re} f(x) \right\},$$

then  $K_0 \cap \overline{co}(ex(K)) \neq \emptyset$  by (32.29) and  $K_0$  is an extremal set of  $K$ . As we have shown  $K_0$  contains an extremal point  $z$  and  $z \in ex(K_0) \subset ex(K)$ ,  $K_0 \cap \overline{co}(ex(K)) \neq \emptyset$  is a contradiction.  $\square$

The Choquet-Singer theorem. Let  $Y$  be a closed subspace of a normed space  $X$ . Now characterize the condition for  $b \in Y$  to be the distance attainer of  $a \in X$ .

**Theorem 32.27** (Choquet-Singer). *Let  $X$  be a normed space and  $Y$  is a closed proper subspace. Suppose  $a \in X$  and  $b \in Y$ . Then the following are equivalent.*

- (1)  $b$  realizes the distance of  $a$  and  $Y$ .
- (2) For every  $y \in Y$  there exists  $x^* \in \text{ex}(X^*_1)$  such that

$$(32.31) \quad \text{Re } x^*(y) \leq \text{Re } x^*(b), \quad x^*(a - b) = \|a - b\|_X.$$

*Proof.* (1)  $\implies$  (2) Let  $y \in Y$ . Then we have

$$(32.32) \quad \|a - b\|_X = \text{Re } x^*(a - b) \leq \text{Re } x^*(a - y) \leq \|a - y\|_X.$$

Therefore,  $b$  realizes the distance from  $a$  to  $Y$ . In this proof we remark that  $x^* \in \text{ex}(X^*_1)$  is not used at all.

$$\boxed{(2) \implies (1)}$$

Define

$$(32.33) \quad F := \{x^* \in X^*_1 : \text{Re } x^*(y) \leq \text{Re } x^*(b), x^*(a - b) = \|a - b\|_X\}.$$

Then since  $F$  is compact, we have only to prove that  $F$  is not empty.

If  $a \in A$ . Then  $x^* = 0 \in F$ . If  $a \notin A$  and  $y = b$ , then any norm attainer of  $a - b$  belongs to  $F$ . Suppose that  $a \notin A$  and  $y \neq b$ . Then  $W_0 := \mathbb{K}(a - b) + \mathbb{K}(y - b)$  is 2-dimensional subspace. Define the linear functional  $l$  by

$$(32.34) \quad l(k_0 \cdot (a - b) + k_1 \cdot (y - b)) := k_0 \|a - b\|_X.$$

Then we have

$$(32.35) \quad |l(k_0 \cdot (a - b) + k_1 \cdot (y - b))| = \|k_0(a - b)\|_X \leq \|k_0(a - b) + k_1(y - b)\|_X,$$

since  $k_0 b$  is a norm attainer of  $k_0 a$ . By using the Hahn Banach theorem  $l$  can be extended to a bounded linear functional  $L$  with norm 1. Since  $L(y - b) = 0$ , we see that  $L \in F$ .  $\square$

The de Brange theorem and application to Weierstrass theorem. In this paragraph we let  $T$  be a compact Hausdorff space consisting of more than 1 points.

**Theorem 32.28.** *Let  $A$  be a subspace of  $C(T)$  that is closed under multiplication. Denote by  $M(T)$  the set of all signed Borel measures. Set*

$$(32.36) \quad K := \left\{ \mu \in M(T) : \|\mu : M(T)\| \leq 1, \int_T f d\mu = 0 \text{ for all } f \in A \right\}.$$

*Then if  $\mu \in \text{ex}(K)$  and  $f \in A$  satisfies  $0 < \inf f \leq \sup f < 1$ , then  $f$  is constant on  $\text{supp}(\mu)$ .*

*Proof.* If  $\mu$  is zero, then  $\text{supp}(\mu) = \emptyset$  and hence there is nothing to prove. Below we assume that  $\mu \neq 0$ . Since  $\mu$  is an extremal point of  $K$ , we see that  $\|\mu : M(T)\| = 1$ .

Define two positive measures absolutely continuous with respect to  $\mu$  by

$$(32.37) \quad \nu(E) := \int_E f d\mu, \quad \lambda(E) := \int_E (1 - f) d\mu.$$

Then  $\nu$  and  $\lambda$  are not zero by assumption. From the definition, we have

$$(32.38) \quad \int_T g d\nu = \int_E f \cdot g d\mu, \quad \int_T g d\lambda = \int_E f \cdot g d\lambda. \quad \int_F g d\nu = \int_E f \cdot g d\mu,$$

Note that

$$(32.39) \quad \mu = \|\nu : M(T)\| \cdot \frac{\nu}{\|\nu : M(T)\|} + \|\lambda : M(T)\| \cdot \frac{\lambda}{\|\lambda : M(T)\|}$$

and

$$(32.40) \quad \|\nu : M(T)\| + \|\lambda : M(T)\| = \int_E f d|\mu| + \int_E (1-f) d|\mu| = |\mu|(E) = 1.$$

Since  $\mu$  is an extremal point of  $K$ , we have  $\mu = \nu$ . By the uniqueness of the density, we have  $f$  is constant on  $\text{supp}(\mu)$ .  $\square$

Finally to conclude this section we present a proof of the Weierstrass theorem due to de Brange.

**Theorem 32.29.** *Suppose that  $\Theta \subset C(T)$  is a closed subspace that is closed under pointwise multiplication and under complex conjugate ( if  $\mathbb{K} = \mathbb{C}$  ) as well. Assume that  $1 \in \Theta$  and  $\Theta$  separates  $T$ , that is, for two distinct points  $x, y$  there exists  $f \in \Theta$  such that  $f(x) \neq f(y)$ . Then  $\Theta$  is dense in  $C(T)$ .*

*Proof.* By the Hahn Banach theorem, our present task is to show

$$\begin{aligned} K &:= \left\{ \mu \in M(T) : \|\mu : M(T)\| \leq 1, \int_T f d\mu = 0 \text{ for all } f \in \Theta \right\} \\ &= \left\{ \mu \in M(T) : \|\mu : M(T)\| \leq 1, \int_T f d\mu = 0 \text{ for all } f \in \overline{\Theta} \right\} \end{aligned}$$

consists only of  $\{0\}$ . Let  $\mu \in \text{ex}(K)$ . If  $\text{supp}(\mu)$  were made up of more than 1 point, then it would run counter to the fact that  $\Theta$  separates  $T$ . Therefore,  $\mu = \alpha \delta_p$  for some  $p \in T$  and  $\alpha \in \mathbb{K}$ . Since  $\mu$  annihilates 1, we have

$$(32.41) \quad \alpha = \int_T 1 d\mu = 0.$$

Therefore,  $\text{ex}(K) = \{0\}$ . By the Klein-Milman theorem, we have  $K \subset \overline{\text{co}}(\text{ex}(K)) = \{0\}$ . This is the desired result.  $\square$

## 32.2. Locally convex spaces.

After learning topological vector spaces, we go on to investigate objects carrying a little more structures. Locally convex spaces are defined in 1930 by Von. Neumann.

**Definition 32.30.** A topological linear space  $X$  over  $\mathbb{K}$  is locally convex, if there exists a collection of seminorms  $\{p_\lambda\}_{\lambda \in \Lambda}$  which topologizes  $X$ .

**Example 32.31.** Banach spaces are typical examples of the locally convex spaces. We shall take up non-trivial examples later along with their characteristic properties.

**Example 32.32.** We define

$$(32.42) \quad D_{L^1}(\mathbb{R}^d) := \{\varphi \in C^\infty(\mathbb{R}^d) : \partial^\alpha \varphi \in L^1(\mathbb{R}^d)\}.$$

For a multiindex  $\alpha \in \mathbb{N}_0^d$ , we set

$$(32.43) \quad p_\alpha(\varphi) := \left( \int_{\mathbb{R}^d} |\partial^\alpha \varphi(x)| dx \right).$$

Then  $D_{L^1}(\mathbb{R}^d)$  is a locally convex space with respect to these seminorms.



Characterization of Hausdorff locally convex spaces. Let us see how the Hausdorff property is characterized.

**Proposition 32.33.** *Let  $X$  be a locally convex space with topologizing seminorms  $\{p_\lambda\}_{\lambda \in \Lambda}$ . Then  $X$  is Hausdorff, if and only if for all  $x \in X$ , there exists  $\lambda \in \Lambda$  such that  $p_\lambda(x) \neq 0$ .*

*Proof.* If  $X$  is Hausdorff, then there exists a neighborhood  $U$  of 0 such that

$$(32.44) \quad (x + U) \cap U \neq \emptyset.$$

By replacing  $U$  with smaller neighborhood, we may assume that  $U = \bigcap_{j=1}^k p_{\lambda(j)}^{-1}([0, \varepsilon_j])$  for some  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k > 0$  and  $\lambda(1), \lambda(2), \dots, \lambda(k) \in \Lambda$ . Since  $x \notin U$ , we see that  $p_{\lambda(j)}(x) \neq 0$  for some  $j$ .

If for each  $x \in X$  there exists  $\lambda \in \Lambda$  such that  $p_\lambda(x) \neq 0$ . Then

$$(32.45) \quad U_1 := \left\{ y \in X : |p(y)| < \frac{p(x)}{2} \right\}, U_2 := \left\{ y \in X : |p(y) - p(x)| < \frac{p(x)}{2} \right\}$$

are open sets that separate 0 and  $x$ . Therefore,  $X$  is Hausdorff. □

*Exercise 205.* Show that  $D_{L^1}(\mathbb{R}^d)$  is a Hausdorff space.

Bounded set in a locally compact space. Now let us characterize bounded sets in a locally convex space.

**Theorem 32.34.** *Assume that  $X$  is a locally convex space topologized by  $\{p_\lambda\}_{\lambda \in \Lambda}$ . A subset  $A$  in  $X$  is bounded if and only if*

$$(32.46) \quad \sup_{a \in A} p_\lambda(a) < \infty$$

for each  $\lambda \in \Lambda$ .

*Proof.* Note that for any open set  $U$  there exists  $\lambda \in \Lambda$  and  $\varepsilon > 0$  such that

$$(32.47) \quad \{\varphi \in X : p_\lambda(\varphi) < \varepsilon\} \subset U.$$

Therefore, we can say  $A$  is bounded if and only if for each  $\varepsilon > 0$  and  $N$  there exists  $\alpha > 0$  such that

$$(32.48) \quad A \subset \alpha \{\varphi \in X : p_\lambda(\varphi) < \varepsilon\}.$$

By positive homogeneity of  $p_\lambda$  this is equivalent to saying for each  $\lambda \in \Lambda$  there exists  $\alpha > 0$  such that

$$(32.49) \quad A \subset \{\varphi \in X : p_\lambda(\varphi) < \alpha\},$$

which is equivalent to

$$(32.50) \quad \sup_{a \in A} p_\lambda(a) < \infty$$

for each  $\lambda \in \Lambda$ . □

Finally, in this paragraph, we shall prove the Mackey theorem on bounded subsets in topological vector spaces. In Definition 32.15 we have defined the boundedness for subsets in topological vector spaces. Here we shall define another notion of boundedness.

**Definition 32.35.** Let  $X$  be a topological vector space and  $X^*$  its dual. A  $\sigma(X^*, X)$ -bounded subset in  $X$  is said to be weakly bounded.

**Theorem 32.36** (Mackey). *Any weakly bounded set in locally convex spaces is bounded.*

*Proof.* Let  $A \subset X$  be a weakly bounded set in a locally convex space  $X$ . Let  $p$  be a continuous semi-norm. We define

$$(32.51) \quad \|x^*\|_p = \sup\{|x^*(x)| : x \in p^{-1}(\{|z| \leq 1\})\}$$

for  $x^* \in X^*$ . Denote by  $X_p^*$  the set of all functionals  $x^* \in X^*$  such that  $\|x^*\|_p < \infty$ . Then  $X_p^*$  carries the structure of a Banach space with norm  $\|\cdot\|_p$ .

Let  $q(x^*) := \sup\{|x^*(x)| : x \in A\}$ . Then  $q < \infty$  because  $A$  is weakly-bounded. Furthermore, as is easily seen  $q^{-1}((j, \infty)) \subset X_p^*$  is open in the norm topology of  $X_p^*$ . Since

$$(32.52) \quad X_p^* = \bigcup_{j=1}^{\infty} q^{-1}([0, j]),$$

by the Baire category theorem  $q^{-1}([0, 1])$  is an open set. As a result, we see that  $q$  is continuous, which implies that  $p(A)$  is bounded. □

**Fixed point theorems.** In this paragraph we consider a fixed point  $T : X \rightarrow X$ , where  $X$  is a locally convex space. A fixed point  $x$  of  $T$  is a point satisfying  $Tx = x$ . Fixed point theorems are helpful to guarantee the existence of the partial differential / integral equations.

**Theorem 32.37** (Brouwer). *Let  $K$  be a convex compact set in  $\mathbb{R}^d$ . If  $T$  is a continuous mapping from  $K$  to  $K$ , then  $T$  has a fixed point.*

*Proof.* An argument using homology shows that the assertion is the case if  $K$  is a unit disc. For the proof we refer to [60], for example. Since we have to go into the details of topological argument, we accept that this case is true.

In general case it can be assumed that  $\text{Span}(K) = \mathbb{R}^d$  by replacing  $\mathbb{R}^d$  with  $\text{Span}(K)$ . Translation allows us to assume that  $0 \in K$  as well. Assuming  $\text{Span}(K) = \mathbb{R}^d$  and  $0 \in K$ , there exists  $d$ -independent vectors  $v_1, v_2, \dots, v_d$ . Consider a polygon  $P$  whose vertices are  $0, v_1, v_2, \dots, v_d$ . Then  $P$  contains an interior point. Another translation therefore allows us to assume that  $0$  is an interior point of  $K$ . Let  $p$  be the Minkowski functional. Then the mapping  $\varphi : x \in K \setminus \{0\} \mapsto \frac{p(x)x}{|x|} \in \overline{B(1)}$  extends to a homeomorphism from  $K$  to the closed unit ball  $\overline{B(1)}$ . Therefore, the matters are reduced to the case when  $K = \overline{B(1)}$ . □

Our intention here is to extend it to a large extent. Define

$$(32.53) \quad C := \{a = \{a_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N}) : |a_j| \leq 2^{-j} \text{ for all } j \in \mathbb{N}\}.$$

**Lemma 32.38.**  *$C \subset \ell^2(\mathbb{N})$  is compact.*

*Proof.* Let  $\{a^{(k)}\}_{k \in \mathbb{N}}$  be a sequence in  $C$ . Let us write  $a^{(k)} = \{a_j^{(k)}\}_{j \in \mathbb{N}}$ . Then a passage to a subsequence allows us to assume that  $\{a_j^{(k)}\}_{k \in \mathbb{N}}$  converges for all  $j \in \mathbb{N}$ . Set  $\lim_{k \rightarrow \infty} a_j^{(k)} = a_j$ .

Since  $\{a^{(k)}\}_{k \in \mathbb{N}} \in C$ , we have

$$(32.54) \quad \sum_{j=1}^{\infty} |a_j^{(k)} - a_l^{(k)}|^2 \leq \sum_{j=1}^N |a_j^{(k)} - a_l^{(k)}|^2 + 4 \sum_{j=N+1}^{\infty} \frac{1}{4j}$$

for all  $N \in \mathbb{N}$ . By the Fatou lemma, we obtain

$$(32.55) \quad \sum_{j=1}^{\infty} |a_j^{(k)} - a^{(k)}|^2 \leq \sum_{j=1}^N |a_j^{(k)} - a^{(k)}|^2 + 4 \sum_{j=N+1}^{\infty} \frac{1}{4j},$$

letting  $l \rightarrow \infty$ . Now if we let  $k \rightarrow \infty$ , it follows that

$$(32.56) \quad \limsup_{k \rightarrow \infty} \sum_{j=1}^{\infty} |a_j^{(k)} - a^{(k)}|^2 \leq 4 \sum_{j=N+1}^{\infty} \frac{1}{4^j}.$$

Since  $N$  is at our disposal, we finally obtain

$$(32.57) \quad \limsup_{k \rightarrow \infty} \sum_{j=1}^{\infty} |a_j^{(k)} - a^{(k)}|^2 \leq 0.$$

Thus,  $\{a^{(k)}\}_{k \in \mathbb{N}}$  tends to a limit. □

**Lemma 32.39.** *Any continuous mapping from  $C$  to itself has a fixed point.*

*Proof.* Let  $T$  be a continuous mapping from  $C$  to itself. We define

$$(32.58) \quad P_k((a_1, a_2, \dots, a_k, a_{k+1}, a_{k+2}, \dots)) := (a_1, a_2, \dots, a_k, 0, 0, \dots).$$

Then  $P_k T|_{P_k(C)} : P_k(C) \rightarrow P_k(C)$  is a continuous mapping. Therefore, the Brouwer fixed point theorem gives us a point  $y_k \in P_k(C)$  with  $P_k T(y_k) = y_k$ . If we pass to a subsequence  $\{y_{k_j}\}_{j \in \mathbb{N}}$ , we may assume it is convergent. Note that  $|P_k(x) - x| \leq 2^{-k}$  for all  $x \in C$  because of the definition of  $C$ . Since

$$(32.59) \quad |y_{k_j} - T(y_{k_j})| = |y_{k_j} - P_{k_j} T(y_{k_j}) + P_{k_j} T(y_{k_j}) - T(y_{k_j})| \leq \frac{1}{2^{k_j-1}},$$

a passage to the limit  $j \rightarrow \infty$  gives us  $y = Ty$ . Therefore,  $C$  has a fixed point. □

**Lemma 32.40.** *Let  $K$  be a convex compact set in  $C$ . Then any continuous mapping  $T$  from  $K$  to itself has a fixed point.*

*Proof.* Let  $T$  be a continuous mapping from  $K$  to  $K$ . From the the parallelogram law for all  $p \in C$ , there exists unique point  $N(p) \in K$  such that  $\text{dist}(p, K) = |p - N(p)|$ . Here we use the the parallelogram law to guarantee the uniqueness of  $N(p)$ . Since  $p \mapsto N(p)$  is continuous, speaking precisely a contraction:  $|N(p) - N(q)| \leq |p - q|$  for all  $p, q \in C$ , we have  $T \circ N : C \rightarrow C$  is a continuous mapping. Therefore, from Lemma 32.39 we have a fixed point  $k$  of  $T \circ N$ .  $k$  satisfies

$$(32.60) \quad k = T(N(k)) \in T(K) \subset K.$$

Therefore  $k = N(k) \in K$ . Substituting this, we obtain  $Tk = TN(k) = k \in K$ . Therefore,  $k$  is a fixed point of  $T$  as well. □

**Theorem 32.41** (Schauder-Tykonov). *Let  $K$  be a compact convex subset in a locally convex space  $X$ . Then any continuous mapping  $f : K \rightarrow K$  has a fixed point.*

*Proof.* We claim that a net  $\{x_\alpha\}_{\alpha \in A}$  in  $K$  converges to  $x$  in the topology, if and only if  $\{x^*(x_\alpha)\}_{\alpha \in A}$  in  $K$  converges to  $x^*(x)$  for all  $x^* \in X^*$ . Indeed, let  $(K, \mathcal{O}_X|K)$  be the topological space equipped with the induced topology of  $X$  and  $(K, \mathcal{O}^*)$  be a topological space topologized by  $X^*$ . Then the identity mapping

$$(32.61) \quad \iota : (K, \mathcal{O}_X|K) \rightarrow (K, \mathcal{O}^*)$$

is a continuous bijection from a compact space  $(K, \mathcal{O}_X|K)$  to the Hausdorff space  $(K, \mathcal{O}^*)$ . Therefore, it follows that  $\iota$  is a bijection, which shows our claim. Hence, we may assume  $X$  is equipped with the weak topology from the beginning.

Next, let us say that a collection of  $\{f_j\}_{j \in J}$  is determined by  $\{g_k\}_{k \in K}$ , if and only if the following conditions are fulfilled. For all  $j \in J$  and  $\varepsilon > 0$  there exists  $\delta > 0$  and a finite collection of index  $k_1, k_2, \dots, k_m$  in  $K$  such that

$$(32.62) \quad |f(Tp) - f(Tq)| < \varepsilon,$$

whenever  $p, q \in K$  satisfy

$$(32.63) \quad |g_1(p) - g_1(q)|, \dots, |g_m(p) - g_m(q)| < \delta.$$

In this case we write

$$(32.64) \quad \{f_j\}_{j \in J} \ll \{g_k\}_{k \in K}.$$

First, we claim that for any  $f \in X^*$ , we can pick a countable set  $G \subset X^*$  determining a singleton  $\{f\}$ .

Let  $n \in \mathbb{N}$  be fixed. For all  $p \in K$ , there exists a convex and symmetric neighborhood  $U_p$  of 0 such that

$$(32.65) \quad |f(Tp) - f(Tq)| < 2^{-n-1}$$

for all  $q \in p + U_p$ . Since  $K$  is assumed compact, we can find finite subset  $\{p_1, p_2, \dots, p_m\}$  such that

$$(32.66) \quad K \subset \bigcup_{j=1}^m \left( p_j + \frac{1}{2}U_{p_j} \right).$$

Define  $U = \bigcap_{j=1}^m \frac{1}{2}U_{p_j}$ . Suppose that  $p, q \in K$  with  $p - q \in U$ . Then  $p - p_j \in \frac{1}{2}U_{p_j}$  for some  $j = 1, 2, \dots, m$ . Therefore,

$$(32.67) \quad |f(Tp) - f(Tp_j)| < 2^{-n-1}.$$

Note that  $q - p_j = q - p + p - p_j \in U + \frac{1}{2}U_j \subset U_j$ . Therefore, there exists a neighborhood  $U$  such that

$$(32.68) \quad |f(Tp) - f(Tq)| < 2^{-n}$$

for all  $p, q$  with  $p - q \in U$ . Since  $X$  is assumed locally convex, we conclude that there exists a finite subset  $G_n \subset X^*$  such that

$$(32.69) \quad |f(Tp) - f(Tq)| < 2^{-n}$$

whenever  $\sup_{g \in G_n} |g(p - q)| \leq 1$ . If we set  $G = \bigcup_{n=1}^{\infty} G_n$ , then we obtain  $\{f\} \ll G$ .

From the previous paragraph, given a countable set  $F \subset X^*$ , we can find another countable set  $G$  such that  $F \ll G$ . Suppose that  $F = F_0$  is a countable set. Then we can take inductively a sequence of countable subsets  $F_1, F_2, \dots$  such that  $F_{j-1} \ll F_j$  for all  $j \in \mathbb{N}$ . Define

$$(32.70) \quad F := \bigcup_{j=0}^{\infty} F_j.$$

Then we obtain  $F \ll F$ . Therefore, starting from a countable set  $F_0$ , we can construct a countable set such that  $F_0 \ll F \ll F$  and that  $F_0 \subset F$ .

Since we can assume that  $K$  contains more than 1 point, there exists  $f \in X^*$  and  $F \subset X^*$  such that  $f|_K$  is not constant,  $\{f\} \subset F$  and  $\{f\} \ll F \ll F$ . Let us fix this  $F = \{f_j\}_{j \in \mathbb{N}}$  until the end of the proof.

We define

$$(32.71) \quad H : K \rightarrow C, H(k) = \left( \frac{f_j(k)}{2^j \sup_{k' \in K} |f_j(k')| + 2^j} \right)_{j \in \mathbb{N}}.$$

Then since each  $f_j$  is continuous, we see that  $H$  is continuous as well. Set  $K_0 := H(K)$ . Then  $K_0$  is a convex and compact set. Suppose that  $k_1, k_2 \in K$  satisfies  $H(k_1) = H(k_2)$ . Then we have  $H(Tk_1) = H(Tk_2)$ , since  $F \ll F$ .

In view of the above paragraph, the mapping

$$(32.72) \quad T_0 : a \in K_0 \mapsto H(Tk), a = H(k)$$

is well-defined. We shall prove its continuity. Let  $\varepsilon > 0$ . Then there exists  $N$  such that

$$(32.73) \quad \sum_{j=N+1}^{\infty} \frac{1}{4^j} < \frac{\varepsilon}{2}.$$

Then since  $F \ll F$ , there exists  $\delta$  such that

$$(32.74) \quad |f_j(Tk) - f_j(Tl)| < \frac{\varepsilon}{2},$$

if  $|f_j(k - l)| < \delta$ . Assume that  $a, b \in K_0$  with

$$(32.75) \quad \|a - b\|_{\ell^\infty(\mathbb{N})} < \frac{\delta}{4^N \sum_{j=1}^N \sup_{k' \in K} |f_j(k')| + 4^N}.$$

If  $a = H(k)$  and  $b = H(l)$ , then

$$(32.76) \quad \|T_0(a) - T_0(b)\|_{\ell^\infty(\mathbb{N})} \leq \left\{ \sum_{j=1}^N \left( \frac{|f_j(Tk) - f_j(Tl)|^2}{2^j \sup_{k' \in K} |f_j(k')| + 2^j} \right)^2 \right\}^{\frac{1}{2}} + \frac{\varepsilon}{2} \leq \varepsilon.$$

Therefore,  $T_0$  is continuous.

In view of the previous lemma,  $T_0$  has a fixed point. Let  $a = H(k)$  be such a fixed point. Then

$$(32.77) \quad T(H^{-1}(k)) \subset H^{-1}(k).$$

Since  $H$  is not an injection, we see  $H^{-1}(k)$  is a proper subset of  $K$ .

Up to this point we have found a proper subset  $K_1$  of  $K$  which is preserved by  $T$ . By Zorn's lemma, we see that  $T$  has a fixed point.  $\square$

**Theorem 32.42.** *Let  $X$  be a locally convex space and  $K$  be a convex compact set. Then any continuous mapping  $T : K \rightarrow K$  has a fixed point.*

**Theorem 32.43.** *Let  $X$  be a locally convex space and  $K$  be a convex compact set. Suppose that we are given a commuting family of affine mapping  $\mathcal{T} = \{T_\lambda\}_{\lambda \in \Lambda}$  on  $K$ . That is, for each  $\lambda$ ,  $T_\lambda$  is a continuous mapping from  $K$  to itself that satisfies  $T_\lambda(\alpha x + (1 - \alpha)y) = \alpha T_\lambda x + (1 - \alpha)T_\lambda y$ . Then there exists a common fixed point of  $\mathcal{T}$ .*

*Proof.* Assume that  $\Lambda$  is a finite set We prove the theorem by induction. If  $\#\Lambda = 1$ , then Theorem 32.42 asserts more: We do not have to assume that  $T$  is affine. Anyway the case when  $\#\Lambda = 1$  is immediate from Theorem 32.42. Assume that Theorem 32.43 is true when  $\Lambda$  consists of  $k$  elements. Let  $\Lambda$  be a set consisting of  $k + 1$  elements. Let  $\lambda \in \Lambda$ . Then again by Theorem 32.43 with  $\Lambda = k$ , the set  $\tilde{K} = \{x \in K : T_\lambda x = x\}$  is not empty. Since  $T_\lambda$  is affine,  $\tilde{K}$  is convex as well. Since  $\mathcal{T}$  is commutative,  $T_\rho$  preserves  $\tilde{K}$  for any  $\rho \in \Lambda$ . In view of the fact that  $\#(\Lambda \setminus \{\lambda\}) = k$  we are in the position of the inductive hypothesis to  $\Lambda \setminus \{\lambda\}$  and  $\{T_\lambda|_{\tilde{K}}\}_{\lambda \in \Lambda}$ .

As a consequence there exists  $y \in \tilde{K}$  such that  $T_\rho y = y$  for all  $\rho \in \Lambda \setminus \{\lambda\}$ . Since  $y \in \tilde{K}$ ,  $T_\lambda$  fixes  $y$ , too. Therefore,  $y$  is the desired common fixed point.

Assume that  $\Lambda$  is an infinite set We define

$$(32.78) \quad \mathcal{L} := \{L \in 2^\Lambda : \#L < \infty\}.$$

Let  $L_1, L_2 \in \mathcal{L}$ . Then define  $L_1 \ll L_2$  by  $L_1 \subset L_2$ . Note that  $\ll$  defines a partial order which makes  $\mathcal{L}$  a direct set. Our strategy of the proof is that we take full advantage of the net argument.

By virtue of the previous paragraph, for each  $L \in \mathcal{L}$  we can find  $x_L$  such that  $T_\lambda x_L = x_L$  for all  $\lambda \in L$ . Since  $K$  is assumed compact, we can find a convergent subnet  $\{x_L\}_{L \in \mathcal{L}_0}$ . Define  $x := \lim_{L \in \mathcal{L}_0} x_L$ .

Pick  $\lambda \in \Lambda$  arbitrarily. Then, since  $\mathcal{L}_0$  is cofinal, there exists  $L_0 \in \mathcal{L}_0$  such that  $\{\lambda\} \ll L_0$ , that is,  $\lambda \in L_0$ . Therefore, if  $L \in \mathcal{L}_0$  contains  $L_0$ , then  $T_\lambda x_L = x_L$ . This implies  $\lim_{L \in \mathcal{L}_0} T_\lambda x_L = \lim_{L \in \mathcal{L}_0} x_L = x$ . Since  $T_\lambda$  is assumed continuous, we obtain  $T_\lambda x = \lim_{L \in \mathcal{L}_0} T_\lambda x_L = x$ , proving that  $x$  is a common fixed point. □

*Exercise 206.* Suppose that  $\{A_\lambda\}_{\lambda \in \Lambda}$  is a family of  $d \times d$ -normal matrices:  $A_\lambda^* A_\lambda = A_\lambda A_\lambda^*$  for all  $\lambda \in \Lambda$ . Assume that  $A_\lambda \cdot A_\rho = A_\rho \cdot A_\lambda$  for  $\lambda, \rho \in \Lambda$ . Then  $\{A_\lambda\}_{\lambda \in \Lambda}$  can be diagonalized by a common unitary matrix.

Direct limit of the locally convex spaces. Finally we consider the “sum” of the “increasing” locally convex spaces. We begin with a preparatory lemma.

**Lemma 32.44.** *Let  $X$  be a locally convex space and  $M$  a sublinear space. Equip  $M$  with the relative topology of  $X$ . Assume that  $V$  is a convex and circular open set in  $M$ . Then there exists a convex circular neighborhood  $U$  of  $0 \in X$  such that  $V = U \cap M$ .*

*Proof.* Let  $W_0$  be an open set in  $X$  such that  $0 \in W_0 \cap M \subset V$ . Then the smallest convex and circular set  $W$  containing  $0$  is again open. Now that  $W_0 \cap M \subset V$  and  $V$  is convex and circular, we have  $W \cap M \subset V$ . Let  $U = \text{co}(V \cap W)$ , which is open, convex and circular in  $X$ . We claim that  $U$  is the desired set, that is, we prove that  $U \cap M = V$ . Let  $x \in U \cap M$ . Then we can write

$$(32.79) \quad x = t x_0 + (1 - t) x_1$$

for some  $x_0 \in V, x_1 \in W$  and  $0 \leq t \leq 1$ . If  $t = 0, 1$ , then it is easy to deduce that  $x \in V$ . Suppose otherwise. Since  $x_0 \in V \subset M$  and  $x \in M$ , we see that  $x_1 \in M$ . Since  $x_1 \in V$ , we obtain  $x_1 \in V$ . Now that  $V$  is convex, we conclude  $x \in V$ . □

**Definition 32.45.** Suppose that  $\Lambda$  is an oriented set and  $\{X_\lambda\}_{\lambda \in \Lambda}$  is an increasing family of locally convex spaces. Let  $X = \bigcup_{\lambda \in \Lambda} X_\lambda$ . Assume  $X_\tau$  coincides the relative topology of  $X_\lambda$

whenever  $\tau < \lambda$ . Define

$$(32.80) \quad \mathcal{U} := \{U \in 2^X : U \cap X_\lambda \text{ is an open set containing } 0\}$$

and  $\mathcal{U}_x := \{x + U : U \in \mathcal{U}\}$ . Equip  $X$  with a topology generated by  $\bigcup_{x \in X} \mathcal{U}_x$ .

**Theorem 32.46.** *Keep to the same setting above. Then  $X$ , equipped with the topology in Definition 32.45, is a locally convex space. Furthermore it satisfies the following.*

- (1)  $X_\lambda \subset X$  is a continuous embedding for all  $\lambda \in \Lambda$ .

(2) Assume in addition that each  $X_\lambda$  is Hausdorff. Then so is  $X$ .

*Proof.* It is straightforward to prove that  $X$  is a locally convex space. Let us prove (1). It follows from the definition of  $X$  that the topology of  $X_\lambda$  induced by  $X$  is weaker than the original topology of  $X_\lambda$ . Indeed, the generator of  $X$  is open set, if restricted to  $X_\lambda$ . Conversely pick an open, convex and circular set  $V_\lambda$  of  $0$ . Then by Lemma 32.44 for each  $\rho > \lambda$ , we can choose a set  $V_\rho \subset X_\rho$  which is open in the topology of  $X_\rho$  such that  $V_\rho \cap X_\lambda = V_\lambda$ . Let

$$(32.81) \quad V := \bigcup_{\rho > \lambda} V_\rho \subset X.$$

Since  $\{V_\rho\}_{\rho \gg \lambda}$  is an increasing family, we conclude that  $V$  is convex and circular and that it contains  $X$ . Observe that  $V \cap X_\rho = V_\rho$ . Therefore,  $V$  is open in  $X$ . Hence,  $V_\lambda$  is open with respect to the relative topology of  $X$ .

Next, let us prove (2). Let  $u \in X \setminus \{0\}$ . Choose  $\lambda$  so large that  $u \in X_\lambda$ . Then there exists a convex and circular neighborhood  $U$  of  $0$  such that  $u \notin U + U$ . Now that we have established that the topology of  $X_\lambda$  is induced by that of  $X$ , we are in the position of using Lemma 32.44 and we obtain a convex and circular neighborhood  $V$  of  $0 \in X$  such that  $X_\lambda \cap V = U$ . Then we see that  $V \cap (V + u) = \emptyset$  from the fact that  $u \notin U + U$  and that  $X_\lambda \cap V = U$ . Therefore,  $X$  is a Hausdorff space.  $\square$

### 33. EXAMPLES OF LOCALLY CONVEX SPACES: $D(\Omega)$ AND $E(\Omega)$

We now begin to be oriented to concrete settings. We deal with  $D(\Omega)$ , the set of all  $C^\infty(\mathbb{R}^d)$ -functions supported in  $\Omega$  and  $E(\Omega)$ , the set of all  $C^\infty(\mathbb{R}^d)$ -functions in  $\Omega$ .

#### 33.1. $D(\Omega)$ and $D'(\Omega)$ .

$C^\infty(\Omega; K)$ . First we set up the theory of  $C_c^\infty(\Omega)$  functions supported on a fixed compact set  $K$ .

**Definition 33.1.** Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . Denote by  $C_c^\infty(\Omega)$  the set of all compactly supported functions on  $\Omega$ .

- (1) Define  $\mathcal{K}(\Omega)$  as the set of all compact sets in  $\Omega$ .
- (2) Let  $K \in \mathcal{K}(\Omega)$ . Then denote

$$(33.1) \quad C_c^\infty(\Omega; K) := \{\varphi \in C_c^\infty(\Omega) : \text{supp}(\varphi) \subset K\}.$$

- (3) Let  $\alpha \in \mathbb{N}_0^d$ . Then  $p_\alpha$  is a functional on  $C_c^\infty(\Omega)$  given by

$$(33.2) \quad p_\alpha(\varphi) := \sup_{x \in \Omega} |\partial^\alpha \varphi(x)|$$

for  $\varphi \in C_c^\infty(\Omega)$ .

- (4) Equip  $C_c^\infty(\Omega; K)$  with the topology generated by a family of functional  $\{p_\alpha\}_{\alpha \in \mathbb{N}_0^d}$ .

**Lemma 33.2.**  $C_c^\infty(\Omega; K)$  is metrizable.

*Proof.* We content ourselves in presenting a metric function for which the topology coincide with the original topology. We set

$$(33.3) \quad d(f, g) := \sum_{\alpha \in \mathbb{N}_0^d} \frac{1}{\alpha!} \min(1, p_\alpha(f - g)).$$

It is straightforward to verify that  $d$  defines the same topology as the original one.  $\square$

**Corollary 33.3.**  $C_c^\infty(\Omega; K)$  is complete in the sense that  $d$ , appearing in the above proof, defines a complete topology.

$D(\Omega)$ . With this definition in mind, we shall endow  $C_c^\infty(\Omega)$  with a suitable topology.

**Definition 33.4.** Equip  $C_c^\infty(\Omega)$  with the topology indicated in Definition 32.45.

Characterization of convergence of sequence. Having presented the definition of the topology, let us see how the convergence of sequence is characterized.

**Lemma 33.5.** Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence. Assume that  $\{x_j\}_{j \in \mathbb{N}}$  is a sequence that has no accumulation point in  $\Omega$  and  $f_j(x_j) \neq 0$ . Then  $\{f_j\}_{j \in \mathbb{N}}$  is never convergent.

*Proof.* If we pass to a subsequence, we may assume  $x_j \neq x_k$  for all  $j \neq k$ . In this case by assumption, for each  $j$  we can choose  $r_j > 0$  so that  $B(x_j, 2r_j)$  do not contain  $x_k$  if  $k \neq j$ . Let  $\varphi_j$  be defined so that

$$(33.4) \quad \chi_{B(x_j, r_j)} \leq \varphi_j \leq \chi_{B(x_j, 2r_j)}.$$

Then define

$$(33.5) \quad p(f) := \sum_{j=1}^{\infty} \frac{\|\varphi_j \cdot f\|_\infty}{|f_j(x_j)|}$$

Note that  $p : C_c^\infty(\Omega) \rightarrow [0, \infty)$  is continuous. Indeed, to verify this, we need to check that the subset in  $C_c^\infty(\Omega)$ , given by

$$(33.6) \quad \{f \in C_c^\infty(\Omega) : p(f) < R\}$$

is open for all  $R > 0$ . However, by the definition of the norm, this is reduced to proving

$$(33.7) \quad \{f \in C_c^\infty(\Omega; K) : p(f) < R\}$$

for all  $K \in \mathcal{K}(\Omega)$ . Since  $p|_K$  consists of finite sum, it is clear that the above subset in  $C_c^\infty(\Omega; K)$  is open. Therefore,  $p$  is continuous.

Note that  $p(f_j - f_k) \geq 2$  if  $j$  and  $k$  are not the same. Therefore,  $\{f_j\}_{j \in \mathbb{N}}$  is not convergent.  $\square$

**Proposition 33.6.** Let  $\{\varphi_k\}_{k \in \mathbb{N}}$  be a sequence in  $C_c^\infty(\Omega; K)$ . Then  $\{\Phi_k\}_{k \in \mathbb{N}}$  is convergent to  $\varphi \in C_c^\infty(\Omega; K)$ , precisely when the following conditions are fulfilled.

(1) There exists  $K \in \mathcal{K}(\Omega)$  so that

$$(33.8) \quad \text{supp}(\varphi_k) \subset K$$

for all  $K$ .

(2) For all  $\alpha \in \mathbb{N}_0^d$ , we have

$$(33.9) \quad p_\alpha(\varphi_k - \varphi) \rightarrow 0$$

as  $k \rightarrow \infty$ .

We should keep in mind that (33.8) is a necessary condition for convergence.

*Proof.* It is easy to see  $\varphi_k \rightarrow \varphi$  as  $k \rightarrow \infty$  once we assume (33.8) and (33.9).

Suppose that  $\varphi_k \rightarrow \varphi$  as  $k \rightarrow \infty$ . Then  $\bigcup_{k=1}^{\infty} \varphi_k$  is relatively compact by Lemma 33.5, which means (33.8). Therefore, (33.9) is now immediate from the definition of the convergence.  $\square$



Bounded set of  $D(\Omega)$ . Following the definition, let us characterize bounded sets in  $D(\Omega)$ . Let  $A$  be a bounded set.

**Theorem 33.7.** *A subset  $A$  in  $D(\Omega)$  is bounded, if there exists a compact set  $K$  such that*

$$(33.10) \quad A \subset C_c^\infty(\Omega; K)$$

and

$$(33.11) \quad \sup_{a \in A} p_N(a) < \infty$$

for each  $N \in \mathbb{N}$ .

*Proof.* Assume that  $A$  satisfies the condition above. Then  $A$  is a subset of  $C_c^\infty(\Omega; K)$ . Let  $U$  be an open set containing  $0$ . Then  $U \cap C_c^\infty(\Omega; K)$  is an open set in  $C_c^\infty(\Omega; K)$  containing  $0$ . Therefore, there exists  $\alpha \in \mathbb{N}$  such that  $A \subset \alpha(U \cap C_c^\infty(\Omega; K)) \subset \alpha U$ . Therefore, it follows that  $A$  is bounded in  $D(\Omega)$ .

Assume first that  $A$  is bounded. Let us deny the first inclusion. Then there exists a sequence of points  $\{x_j\}_{j \in \mathbb{N}}$  with no culmination points and  $\{\varphi_j\}_{j \in \mathbb{N}} \subset A$  such that

$$(33.12) \quad \varphi_j(x_j) \neq 0.$$

By discarding some points and functions, we may assume the points  $\{x_j\}_{j \in \mathbb{N}}$  are all distinct. In this case there exists a disjoint sequence of balls  $\{B_j\}$  centered at  $x_j$ . Let  $\{\psi_j\}_{j \in \mathbb{N}}$  be a sequence of cut-off smooth functions such that  $\psi_j$  equals 1 near  $x_j$  for each  $j$ . Now define

$$(33.13) \quad p(\varphi) = \sum_{j=1}^{\infty} j \frac{\|\psi_j \varphi\|_{\infty}}{|\varphi_j(x_j)|}.$$

As before,  $p$  is a continuous functional. Therefore, there exists  $M > 0$  such that  $A \subset p^{-1}([0, M])$ . However, this is a contradiction. Indeed,  $\varphi_j \in A \cap p^{-1}([j, \infty))$  for all  $j$ . This is impossible, if  $j > M$ . Having proved this, we see that  $A$  is a bounded subset of  $D(\Omega; K)$ . Therefore, the second condition is trivial. □

Let us prove functions of the tensored type form a dense subset in  $D(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$ .

**Theorem 33.8.** *Let  $d_1, d_2 \in \mathbb{N}$ . The set of all the functions of tensored type, that is, the set of the functions of the form*

$$(33.14) \quad \Phi(x, y) = \sum_{j=1}^k f_j(x)g_j(y), \quad k \in \mathbb{N}, \quad f_1, f_2, \dots, f_k \in D(\mathbb{R}^{d_1}), \quad g_1, g_2, \dots, g_k \in D(\mathbb{R}^{d_2})$$

is a dense subset of  $D(\mathbb{R}^{d_1, d_2})$ .

*Proof.* Set up Let  $F \in D(\mathbb{R}^{d_1+d_2})$ . Choose  $R$  so large that

$$(33.15) \quad \text{supp}(F) \subset \{(x_1, x_2, \dots, x_{d_1}, y_1, y_2, \dots, y_{d_2}) \in \mathbb{R}^{d_1+d_2} : \|(x, y)\|_{\infty} < R\},$$

where

$$(33.16) \quad \|(x, y)\|_{\infty} := \max(|x_1|, |x_2|, \dots, |x_{d_1}|, |y_1|, |y_2|, \dots, |y_{d_2}|).$$

Pick a cut-off function  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  so that

$$(33.17) \quad \chi_{[-R, R]} \leq \eta \leq \chi_{[-2R, 2R]}.$$

$$\text{Set } \zeta(x, y) = \prod_{j=1}^{d_1} \eta(x_j) \cdot \prod_{j=1}^{d_2} \eta(y_j).$$

Approximation procedure Define

$$(33.18) \quad G(x, y) = \sum_{j \in \mathbb{Z}^{d_1}} \sum_{k \in \mathbb{Z}^{d_2}} F(x - 2Rj, y - 2Rk).$$

Then  $G$  is  $2R$ -periodic and smooth. Therefore, we can expand it into a Fourier series

$$(33.19) \quad G(x, y) = \sum_{j \in \mathbb{Z}^{d_1}} \sum_{k \in \mathbb{Z}^{d_2}} a_{jk} \cdot \exp\left(i \frac{\pi j \cdot x + \pi k \cdot y}{2R}\right).$$

Recall that any partial derivative converges uniformly. Since  $F(x, y) = \zeta(x, y)G(x, y)$ , we have

$$(33.20) \quad F(x, y) = \sum_{j \in \mathbb{Z}^{d_1}} \sum_{k \in \mathbb{Z}^{d_2}} a_{jk} \cdot \zeta(x, y) \exp\left(\pi i \frac{j \cdot x + k \cdot y}{2R}\right).$$

Since  $\zeta(x, y) \exp\left(\pi i \frac{j \cdot x + k \cdot y}{2R}\right)$  is a tensor product of a  $D(\mathbb{R}^{d_1})$  function and a  $D(\mathbb{R}^{d_2})$  function, we see that  $F$  can be approximated in the desired way. □

$D'(\Omega)$ . Let us now turn to the topological dual.  $D'(\Omega)$  is defined as the dual of  $D(\Omega)$  in the way described in the previous section.

**Definition 33.9.** The distribution space  $D'(\Omega)$  is the set of all continuous functional on  $D(\Omega)$ . Its topology is the weakest one such that

$$(33.21) \quad f \in D'(\Omega) \rightarrow \langle f, \varphi \rangle \in \mathbb{C}$$

is continuous for all  $\varphi \in D(\Omega)$ .

**Theorem 33.10.** *Suppose that  $H$  is a subset of  $D'(\Omega)$ . Then the following is equivalent.*

- (1)  $\sup_{F \in H} |\langle F, \varphi \rangle| < \infty$  for all  $\varphi \in D(\Omega)$ .
- (2) *There exists a neighborhood  $U$  of  $0 \in D(\Omega)$  such that  $\sup_{F \in H} \sup_{\varphi \in U} |\langle F, \varphi \rangle| \leq 1$ .*

*Proof.* Needless to say (2) is much stronger than (1). Thus, we have to show (2) assuming (1). Define

$$(33.22) \quad A_j := \{\varphi \in D(\Omega) : \sup_{F \in H} |\langle F, \varphi \rangle| \leq j\}, \quad B_j := \{\varphi \in D(\Omega) : \sup_{F \in H} |\langle F, \varphi \rangle| < j\}.$$

It suffices to show that  $B_1$  is open in  $D(\Omega)$ . Then the assumption reads  $D(\Omega) = \bigcup_{j \in \mathbb{N}} A_j$ .

Observe also that  $A_j = j \cdot A_1$ .

Let  $K$  be a compact set chosen arbitrarily. Then

$$(33.23) \quad C_c^\infty(\Omega; K) = \bigcup_{j=1}^\infty (A_j \cap C_c^\infty(\Omega; K)).$$

Since  $C_c^\infty(\Omega; K)$  can be topologized by a complete metric, we are in the position to apply the Baire category theorem to conclude that  $A_j \cap C_c^\infty(\Omega; K)$  contains an interior point for some  $j$ . Since  $A_j \cap C_c^\infty(\Omega; K) = j \cdot (A_1 \cap C_c^\infty(\Omega; K))$  and  $A_1 \cap C_c^\infty(\Omega; K)$  is symmetric, that is,  $\varphi \in A_1 \cap C_c^\infty(\Omega; K)$  implies  $-\varphi \in A_1 \cap C_c^\infty(\Omega; K)$ , we see that  $0 \in A_1 \cap C_c^\infty(\Omega; K)$ . From the fact that  $0 \in A_1 \cap C_c^\infty(\Omega; K)$ , we deduce that  $B_1 \cap C_c^\infty(\Omega; K)$  is open for each compact set  $K$ . As a result  $B_1$  is an open set we are looking for. □

Finally we conclude this section by presenting a condition for  $f \in D'(\mathbb{R}^d)$  to be regarded as an element in  $D'_{L^1}(\mathbb{R}^d)$ , the dual of  $D_{L^1}(\mathbb{R}^d)$ .

**Proposition 33.11.** *In order that  $f \in D'(\mathbb{R}^d)$  can be regarded as an element in  $D'_{L^1}(\mathbb{R}^d)$ , it is necessary and sufficient that  $f$  can be expressed as follows :*

$$(33.24) \quad f = \sum_{|\alpha| \leq N} \partial^\alpha f_j,$$

where the  $f_j \in L^\infty(\mathbb{R}^d)$  and  $N \in \mathbb{N}$ .

*Proof.* We prove the necessity, the sufficiency being trivial. Let  $f \in D'_{L^1}(\mathbb{R}^d)$ . Then there exist  $c, N \in \mathbb{N}$  such that

$$(33.25) \quad |\langle f, \varphi \rangle| \lesssim \sum_{|\alpha| \leq N} \int_{\mathbb{R}^d} |\partial^\alpha \varphi|$$

for all  $\varphi \in D_{L^1}(\mathbb{R})$ . Let  $X$  be the closure of the subspace

$$(33.26) \quad \left\{ \{ \partial^\alpha \varphi \}_{|\alpha| \leq N} \in \prod_{|\alpha| \leq N} L^1(\mathbb{R}^d) : \varphi \in D_{L^1}(\mathbb{R}^d) \right\}$$

in  $\prod_{|\alpha| \leq N} L^1(\mathbb{R}^d)$ . Then  $f$  can be regarded as a continuous linear functional  $F$  on  $X$ . By Hahn-Banach theorem and the duality  $L^1(\mathbb{R}^d)$ - $L^\infty(\mathbb{R}^d)$ , there exists a collection  $\{f_\alpha\}_{|\alpha| \leq N}$  of  $D_{L^1}(\mathbb{R}^d)$  functions

$$(33.27) \quad \langle f, \varphi \rangle = \sum_{|\alpha| \leq N} \int_{\mathbb{R}^d} f_j \partial^\alpha \varphi$$

for all  $\varphi \in D_{L^1}(\mathbb{R}^d)$ . As a result, we conclude that

$$(33.28) \quad f = \sum_{|\alpha| \leq N} \partial^\alpha f_j,$$

which is the desired result. □

This assertion contains a surprising corollary, which gives us a characterization of elements in  $\mathcal{S}'(\mathbb{R}^d)$ .

**Theorem 33.12.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then it can be expressed as*

$$(33.29) \quad f = \sum_{j=1}^N \partial^{\alpha_j} f_j,$$

where the  $f_j$  are regular distributions that satisfy

$$(33.30) \quad |f_j(x)| \lesssim \langle x \rangle^{M_j}$$

and  $\alpha_1, \dots, \alpha_N \in \mathbb{N}_0^d$ .

*Proof.* Let  $m \in \mathbb{N}$  be an integer such that  $(1+|x|^2)^{-m} f \in D'_{L^1}(\mathbb{R}^d)$ . Then the above proposition shows us

$$(33.31) \quad (1+|x|^2)^{-m} f = \sum_{|\alpha| \leq N} \partial^\alpha f_\alpha$$

for some  $\{f_\alpha\}_{|\alpha| \leq N}$  with  $N$  large. It remains to arrange the above inequality by using the Leibnitz rule. □

33.2.  $E(\Omega)$  and  $E'(\Omega)$ .

We now make  $C^\infty(\Omega)$ , the set of all smooth functions not always compactly supported, into a topological vector space. Let us investigate the structure of  $C^\infty(\Omega)$  as a topological vector space.

**Definition 33.13.**  $E(\Omega)$  is a topological vector space that equals to  $C^\infty(\Omega)$  as a set. The topology of  $E(\Omega)$  is the weakest topology such that

$$(33.32) \quad \varphi \in E(\Omega) \rightarrow p_{K,\alpha}(\varphi) := \sup_{x \in K} |\partial^\alpha \varphi(x)|$$

is continuous for all  $K \in \mathcal{K}(\Omega)$ .  $E'(\Omega)$  is the topological dual of  $E(\Omega)$ . For the sake of simplicity we set

$$(33.33) \quad p_{K,N}(\varphi) := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq N}} p_{K,\alpha}(\varphi).$$

Below we regard  $D(\Omega)$  as a subset of  $E(\Omega)$  naturally.

**Lemma 33.14.** *Let  $U$  be an open set in  $E(\Omega)$ . Then  $U \cap D(\Omega; K)$  is open in  $D(\Omega; K)$  for each compact set  $K \in \mathcal{K}(\Omega)$ .*

*Proof.* By translation we may assume  $0 \in U \cap D(\Omega; K)$  and  $0$  is an interior of  $U \cap D(\Omega; K)$ . From the definition of the topology we have

$$(33.34) \quad \{\varphi \in E(\Omega) : N p_{K,N}(\varphi) < 1\} \subset U$$

for some  $K \in \mathcal{K}(\Omega)$  and  $N \in \mathbb{N}$ . Since

$$(33.35) \quad \{\varphi \in D(\Omega; K) : N p_{K,N}(\varphi) < 1\}$$

is open in  $D(\Omega; K)$ , we conclude that  $U \cap D(\Omega; K)$  is open.  $\square$

**Definition 33.15.** Let  $F \in E'(\Omega)$  or  $D'(\Omega)$ . Then the support of  $F$  is the set of all points  $x$  such that the following property *fails*: There exists a neighborhood  $U(\subset \Omega)$  of  $x$  such that

$$(33.36) \quad \langle F, \varphi \rangle = 0$$

whenever  $\varphi \in C_c^\infty(\Omega)$  is supported on  $U$ .

**Proposition 33.16.** *Let  $\{\eta^{(j)}\}_{j \in \mathbb{N}} \subset D(\Omega)$  be an increasing sequence of cut-off functions satisfying*

$$(33.37) \quad \Omega = \bigcup_{j=1}^{\infty} \text{supp}(\eta^{(j)}), \quad \eta_{j+1} \equiv 1 \text{ on } \text{supp}(\eta^{(j)})$$

for each  $j$ . Then

$$(33.38) \quad \lim_{j \rightarrow \infty} \eta^{(j)} \cdot f = f$$

for every  $f \in E(\Omega)$ .

*Proof.* For the proof, we have to prove

$$(33.39) \quad \lim_{j \rightarrow \infty} p_{K,N}(f - f_j) = 0$$

for each  $K \in \mathcal{K}(\Omega)$  and  $N \in \mathbb{N}$ . Because every neighborhood of  $0$  contains a finite intersection of sets of the form

$$(33.40) \quad \{\varphi \in E(\Omega) : p_{K,N}(\varphi) < \varepsilon\},$$

where  $K \in \mathcal{K}(\Omega)$ ,  $N \in \mathbb{N}$  and  $\varepsilon > 0$ . Having clarified what to prove, it is not hard to prove (33.39). Indeed,  $p_{K,N}(f - f_j) = 0$  if  $j$  is sufficiently large. Thus, the proof is therefore complete.  $\square$

**Theorem 33.17.** *Let  $\Omega$  be an open set in  $\mathbb{R}^d$ .*

(1) *The restriction mapping*

$$(33.41) \quad F \in E'(\Omega) \mapsto F|D(\Omega) \in D'(\Omega)$$

*is well-defined, injective and continuous.*

(2) *Conversely  $f \in D'(\Omega)$  is realized by some  $F \in E'(\Omega)$  in the manner above, if and only if  $f$  is compactly supported.*

*Proof.* (1) : well-definedness  $\square$  We are going to show that  $F|D(\Omega)$  belongs to  $D'(\Omega)$ . It suffices to show

$$(33.42) \quad \{\varphi \in D(\Omega) : |\langle F, \varphi \rangle| < 1\}$$

is an open set in  $D(\Omega)$ . To see this, we have only to check that

$$(33.43) \quad \{\varphi \in D(\Omega; K) : |\langle F, \varphi \rangle| < 1\}$$

is open in  $D(\Omega; K)$  for each compact set  $K$ . However, we can write

$$(33.44) \quad \{\varphi \in D(\Omega; K) : |\langle F, \varphi \rangle| < 1\} = \{\varphi \in E(\Omega) : |\langle F, \varphi \rangle| < 1\} \cap D(\Omega; K).$$

Therefore, the set in question is open in  $D(\Omega; K)$ .

(1) : injectivity  $\square$  Let  $f, g \in E'(\Omega)$ . Then keeping to the same notation as Proposition 33.16 we have

$$(33.45) \quad \langle f, \varphi \rangle = \lim_{j \rightarrow \infty} \langle f, \eta^{(j)} \cdot \varphi \rangle = \lim_{j \rightarrow \infty} \langle g, \eta^{(j)} \cdot \varphi \rangle = \langle g, \varphi \rangle.$$

Here for the second inequality we used the assumption that  $f$  and  $g$  agree on  $D(\Omega)$ .

(1) : continuity  $\square$  Continuity is almost trivial. Because if  $\{f_\alpha\}_{\alpha \in A} \subset E'(\Omega)$  is a net convergent to  $f \in E'(\Omega)$ , then

$$(33.46) \quad \lim_{\alpha \in A} \langle f_\alpha, \varphi \rangle = \langle f, \varphi \rangle$$

for each  $\varphi \in D(\Omega) (\subset E(\Omega))$ . Therefore,  $\{f_\alpha|D(\Omega)\}_{\alpha \in A} \subset D'(\Omega)$  is a net convergent to  $f|D(\Omega) \in D'(\Omega)$ . Thus, the mapping in question is continuous.

(2) : "Only if" part  $\square$  Let  $F \in E'(\Omega)$ . Then there exists  $K \in \mathcal{K}(\Omega)$  and  $N \in \mathbb{N}$  such that

$$(33.47) \quad \{\varphi \in E(\Omega) : N p_{K,N}(\varphi) \leq 1\} \subset \{\varphi \in E(\Omega) : |\langle F, \varphi \rangle| \leq 1\}.$$

Therefore

$$(33.48) \quad |\langle F, \varphi \rangle| \leq N p_{K,N}(\varphi).$$

Thus, if  $\varphi$  vanishes in  $K$ , then  $\langle F, \varphi \rangle = 0$ . Therefore any point outside  $K$  is not a support of  $F$ . Thus,  $F$  is compactly supported. Therefore,  $f \in D'(\Omega)$  can be expressed as a restriction of some  $F \in E'(\Omega)$  only if it is compactly supported.

(2) : "If" part  $\square$  "If part" is easier to prove. Indeed, assume  $f \in D'(\Omega)$  is compactly supported. Then choose a cut-off function  $\zeta \in D(\Omega)$  that equals 1 on and near the support of  $f$ . Define  $F \in E'(\Omega)$  by

$$(33.49) \quad \langle F, \varphi \rangle = \langle f, \zeta \cdot \varphi \rangle.$$

Then  $f$  is a restriction of  $F$  to  $D(\Omega)$ . This is the desired result.  $\square$

Notes and references for Chapter 15.

Section 31. Moore and Smith considered a generalization of the limit of sequences in [358].

Theorems 31.4, 31.5 and 31.6 can be regarded as a triumph of their theory in that they generalize theorems corresponding to  $\mathbb{R}^d$  and countable sequences.

Section 32. Theorem 32.5

Theorem 32.7

Theorem 32.9

Theorem 32.12 and Theorem 32.13 were investigated by S. Nakano.

Theorem 32.14

Theorem 32.19

Theorems 32.20 and 32.21 are called the Banach-Alaoglu theorem. In [77] Alaoglu found the theorem.

M. G. Klein and D. P. Milman obtained Theorem 32.26 in [270].

Section 33. Theorem 32.27

Theorem 32.28

Theorem 32.29

Theorem 32.34

Theorem 32.37

The author owed the idea for the proof of Theorem 32.41 to the textbook [9].

Theorem 32.42

Theorem 32.43

Theorem 32.46

Theorem 33.7

Theorem 33.8

Theorem 33.10

Theorem 33.12

Theorem 33.17

We refer to [230, 149] for other examples of locally convex spaces.

The author has referred to [4, 61, 71] in writing this section.

## Part 16. Interpolation

Here we consider intermediate quasi-Banach spaces of given two quasi-Banach spaces. For example, for  $k \in \mathbb{N}$ , we want to take up the intermediate spaces between  $C^k$  and  $C^0$ . To formulate the results, we need to clarify what pair of (quasi)-Banach spaces can be interpolated. Section 34 is devoted to this formulation. That is, we are going to define pairs of Banach spaces that we can interpolate. In Section 35 we actually interpolate such Banach spaces. We deal with real interpolation and complex interpolation. Each method has advantage and shortcomings. The real interpolation method is applicable to sub-linear operators. The target need not be a Banach space; we can consider quasi-Banach spaces, or more general topological linear spaces. Another merit of real interpolation is to improve function spaces: For example, the starting point can be a little bad function spaces. We have seen in the proof of the boundedness of the Hardy-Littlewood maximal operator  $M$ : Recall that  $M$  is not linear, nor is it  $L^1(\mathbb{R}^d)$ -bounded. We interpolated boundedness of weak- $L^1(\mathbb{R}^d)$  type and of  $L^\infty(\mathbb{R}^d)$  type. As we have seen the weak- $L^1$  space is not a Banach space; it fails the triangle inequality. Meanwhile, the complex interpolation came from a natural context of complex analysis. It can be also applied to multi-linear operators and a family of operators containing the complex parameters. Section 36 intends a series of examples in Section 35: We interpolate  $L^p$ -spaces.

### 34. INTERPOLATION

Interpolation is one of the elementary tools in harmonic analysis.

#### 34.1. Compatible couple.

Interpolation, as the name suggests, gives us information from some given pieces of information. Envisage the Fourier transform  $\mathcal{F}$ . We have shown before that  $\mathcal{F} : L^1(\mathbb{R}^d) \rightarrow L^\infty(\mathbb{R}^d)$  and  $\mathcal{F} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ . Although we did not referred to the  $L^p(\mathbb{R}^d)$ -boundedness property ( $1 < p < 2$ ), an interpolation tells us that  $\mathcal{F} : L^p(\mathbb{R}^d) \rightarrow L^{p'}(\mathbb{R}^d)$  is a continuous mapping. This is an interpolation between two Lebesgue spaces. However, in general, if two Banach spaces are similar in some sense, we can interpolate them. First of all, we shall formulate what “a pair of two quasi-Banach spaces is similar ” means. We begin with presenting the definitions and then exhibit some examples of their notions.

**Definition 34.1.** A pair of quasi-Banach spaces  $(X_0, X_1)$  is a compatible couple if there exists a topological space  $X$  into which both  $X_0$  and  $X_1$  are continuously embedded. In this case  $(X_0, X_1)$  is said to be a compatible quasi-Banach couple. If  $X_0$  and  $X_1$  are Banach spaces, (even if  $X$  is not) then one says that  $(X_0, X_1)$  is a compatible Banach couple.

Given a compatible couple  $(X_0, X_1)$  embedded into  $X$ , the most fundamental spaces are  $X_0 \cap X_1$  and  $X_0 + X_1$ , which can be given as follows:

**Definition 34.2.** Suppose that  $(X_0, X_1)$  is a compatible couple embedded into a topological vector space  $X$ . Below one regards  $X_0$  and  $X_1$  as subsets equipped with different topology. Define

$$\begin{aligned} X_0 \cap X_1 &:= \{x \in X : x \in X_0, \text{ and } x \in X_1\} \\ X_0 + X_1 &:= \{x \in X : x = x_0 + x_1 \text{ for some } x_0 \in X_0, x_1 \in X_1\}. \end{aligned}$$

Equip  $X_0 \cap X_1$  and  $X_0 + X_1$  with the following norms respectively.

$$\begin{aligned} \|x\|_{X_0 \cap X_1} &:= \max(\|x\|_{X_0}, \|x\|_{X_1}) \\ \|x\|_{X_0 + X_1} &:= \inf \{\|x_0\|_{X_0} + \|x_1\|_{X_1} : x_0 \in X_0, x_1 \in X_1, x = x_0 + x_1\}. \end{aligned}$$

**Example 34.3.**

- (1) A pair  $(\mathcal{S}(\mathbb{R}^d), \mathcal{S}'(\mathbb{R}^d))$  is a compatible couple, since  $\mathcal{S}(\mathbb{R}^d)$  is embedded into  $\mathcal{S}'(\mathbb{R}^d)$ .
- (2) Let  $(X, \mu)$  be a  $\sigma$ -finite space. Then  $(L^{p_1}(\mu), L^{p_2}(\mu))$  is a compatible couple for all  $0 < p_1 \leq p_2 \leq \infty$ . Indeed, they are both continuously embedded into the set of all measurable functions. The set of all measurable functions has a linear topology generated by the set of functional given by

$$(34.1) \quad \{[f \mapsto \mu\{|f| > \lambda\}] : \lambda > 0\}.$$

Before we investigate some more examples, let us remark the following elementary fact.

**Theorem 34.4.** *Suppose that  $(X_0, X_1)$  is a compatible couple of (quasi-)Banach spaces embedded into a topological space  $X$ . Then  $X_0 \cap X_1$  and  $X_0 + X_1$  are (quasi-)Banach spaces.*

*Exercise 207.* Prove Theorem 34.4. Especially show their completeness by using Theorem 10.12.

Dual space. Let us investigate the dual space of a compatible couple  $(X_0, X_1)$ .

**Proposition 34.5.** *Suppose that  $(X_0, X_1)$  is a compatible couple. Then so is  $(X_0^*, X_1^*)$ .*

*Proof.* Just notice that  $(X_0 \cap X_1)^*$  is one of the topological spaces into which  $X_0^*$  and  $X_1^*$  are embedded. □

**Theorem 34.6** (Duality theorem). *Suppose that  $X_0$  and  $X_1$  are Banach spaces. Assume in addition that  $X_0 \cap X_1$  is dense in  $X_0$  and  $X_1$ . Then*

$$(34.2) \quad (X_0 \cap X_1)^* = X_0^* + X_1^*, \quad (X_0 + X_1)^* = X_0^* \cap X_1^*$$

with

$$\|x^*\|_{X_0^* + X_1^*} = \sup_{x \in X_0 \cap X_1} \frac{|\langle x^*, x \rangle|}{\|x\|_{X_0 \cap X_1}} \text{ for all } x^* \in X_0^* + X_1^*$$

$$\|x^*\|_{X_0^* \cap X_1^*} = \sup_{x \in X_0 + X_1} \frac{|\langle x^*, x \rangle|}{\|x\|_{X_0 + X_1}} \text{ for all } x^* \in X_0^* \cap X_1^*.$$

*Proof.* It is straightforward to prove that

$$(34.3) \quad (X_0 \cap X_1)^* \leftrightarrow X_0^* + X_1^*, \quad (X_0 + X_1)^* = X_0^* \cap X_1^*$$

and that

$$(34.4) \quad \|x^*\|_{X_0^* + X_1^*} \geq \sup_{x \in X_0 \cap X_1} \frac{|\langle x^*, x \rangle|}{\|x\|_{X_0 \cap X_1}}$$

$$(34.5) \quad \|x^*\|_{X_0^* \cap X_1^*} = \sup_{x \in X_0 + X_1} \frac{|\langle x^*, x \rangle|}{\|x\|_{X_0 + X_1}}.$$

We leave the proof of (34.3), (34.4) and (34.5) as an exercise (Exercise 208.)

Let us prove the reverse inclusion of (34.3). To this end, we take  $x^* \in (X_0 \cap X_1)^*$ . Define

$$(34.6) \quad E := \{(x_0, x_1) \in X_0 \oplus X_1 : x_0 = x_1 \in X_0 \cap X_1\}.$$

Recall that the norm of  $X_0 \oplus X_1$  is given by

$$(34.7) \quad \|(x_0, x_1) : X_0 \oplus X_1\| = \max(\|x_0\|_{X_0}, \|x_1\|_{X_1}),$$

which immediately yields that  $E$  is a closed subspace. Furthermore, the dual of  $X_0 \oplus X_1$  is canonically identified with  $X_0^* \oplus X_1^*$ , whose norm is given by

$$(34.8) \quad \|(x_0^*, x_1^*)\|_{X_0^* \oplus X_1^*} = \|x_0^*\|_{X_0} + \|x_1^*\|_{X_1}.$$



Then  $l : (x_0, x_1) \in E \mapsto \frac{1}{2}x^*(x_0 + x_1)$  is a continuous functional which is dominated by the norm of  $X_0 \oplus X_1$ . Therefore,  $l$  extends to a continuous linear functional  $L$  on  $X_0 \oplus X_1$ . As a result we obtain  $x_0^*$  and  $x_1^*$  such that

$$(34.9) \quad \|x_0^*\|_{X_0^*} + \|x_1^*\|_{X_1^*} \leq \|L\|_{(X_0 \oplus X_1)^*} = \|x^*\|_{(X_0 \cap X_1)^*}$$

and that

$$(34.10) \quad L(x_0, x_1) = \langle x_0^*, x_0 \rangle + \langle x_1^*, x_1 \rangle$$

for all  $(x_0, x_1) \in X_0 + X_1$ . Letting  $x_0 = x_1 = x$ , we obtain

$$(34.11) \quad \langle x^*, x \rangle = L(x, x) = \langle x_0^*, x \rangle + \langle x_1^*, x \rangle.$$

Thus,  $x^* := x_0^* + x_1^*$  and we deduce from 34.9 that

$$\|x^*\|_{X_0^* + X_1^*} \leq \|x_0^*\|_{X_0^*} + \|x_1^*\|_{X_1^*} \leq \|x^*\|_{(X_0 \cap X_1)^*}.$$

This is the desired converse inequality. □

*Exercise 208.* Prove (34.3)–(34.5).

*Exercise 209.*

- (1) If  $X$  and  $Y$  are Banach spaces such that  $X$  is continuously embedded into  $Y$ , then show that  $(X, Y)$  is a compatible couple.
- (2) What is the space into which both  $X$  and  $Y$  are continuously embedded ?

*Exercise 210.* We define two functions  $\Phi, \Psi$  by

$$\Phi(t) := \min(t, t^2), \quad \Psi(t) := \begin{cases} t^2 & 0 \leq t \leq 1, \\ 2t - 1 & 1 \leq t < \infty \end{cases}$$

for  $t \geq 0$ . Then prove that  $L^\Phi \simeq L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  and  $L^\Psi \simeq L^1(\mathbb{R}^d) + L^2(\mathbb{R}^d)$  with norm equivalence.

*Exercise 211.* Prove that  $L^1(\mathbb{R}^d)$  and  $L^\infty(\mathbb{R}^d)$  are Banach spaces which can be embedded into  $\mathcal{S}'(\mathbb{R}^d)$ .

### 34.2. Weak-type function spaces.

Here we suppose that  $(X, \mathcal{B}, \mu)$  is a measure space.

**Definition 34.7.** Let  $1 \leq p < \infty$ . A measurable function  $f$  is a weak- $L^p(\mu)$  function, if it satisfies

$$(34.12) \quad \|f\|_{L^{p,\infty}(\mu)} := \sup_{\lambda > 0} \lambda^p \mu\{|f| > \lambda\} < \infty.$$

Denote the totality of weak- $L^p(\mu)$  functions by  $L^{p,\infty}(\mu)$ .

A typical advantage of dealing with  $L^{p,\infty}(\mu)$  functions can be seen in the following example.

**Example 34.8.** Let  $a > 0$  and  $0 < p < \infty$ . Then  $|x|^{\frac{d}{a}} \notin L^p((0, \infty))$ . However, we still have  $|x|^{\frac{d}{a}} \in L^{a^{-1}, \infty}((0, \infty))$ .

*Exercise 212.* Let  $X = [0, 1]$ ,  $\mu = dx|_X$ . Define  $f(\rho) = \rho$ ,  $g(\rho) = 1 - \rho$  for  $\rho \in X$ . Then show that

$$(34.13) \quad \|f\|_{L^{1,\infty}} = \|g\|_{L^{1,\infty}} = \frac{1}{4}, \quad \|f + g\|_{L^{1,\infty}} = 1.$$

In particular the triangle inequality fails.

*Exercise 213.* This is a good contrast of the fact that we can deduce from the Fubini theorem that  $h \in L^1(\mathbb{R}^2)$  if we assume  $f, g \in L^1(\mathbb{R})$ . Let  $f \in L^{1,\infty}(\mathbb{R})$  and  $g \in L^{1,\infty}(\mathbb{R})$ . Then define  $h(x, y) = (f \otimes g)(x, y) = f(x)g(y)$ . Can we say that  $h \in L^{1,\infty}(\mathbb{R}^2)$ ? Hint: Consider  $f(x) = |x|^{-1}$ ,  $g(y) = |y|^{-1}$ .

**Remark 34.9.** Let  $1 \leq p < \infty$ .

- (1) The function space  $L^{1,\infty}(\mu)$  is not a normed space.
- (2) The function  $f(x) = \min \left\{ 1, \frac{2}{1+|x|} \right\}$ , appearing in the previous part, belongs to  $L^{1,\infty}$ .
- (3) For  $0 < p < \infty$ , we have  $L^p(\mu) \hookrightarrow L^{p,\infty}(\mu)$ , which is immediate from the Chebychev inequality.

*Exercise 214.* Fill in the details of Remark 34.9 and prove each assertion.

*Exercise 215.* What is the axiom of the normed space for which  $L^{1,\infty}(\mu)$  fails?

**Definition 34.10.** Denote by  $\text{Meas}(\mu)$  the linear space of all measurable complex valued functions modulo  $\mu$ -null sets. Topologize  $\text{Meas}(\mu)$  with the following family of functionals.

$$(34.14) \quad M(E; \varepsilon, g)(f) := \mu\{x \in E : |f(x) - g(x)| < \varepsilon\} \quad (E \in \mathcal{B}, \varepsilon > 0, g \in \text{Meas}(\mu)).$$

If  $\mu = dx$ , the Lebesgue measure on  $\mathbb{R}^d$ , then omit  $(\mu)$  to write  $\text{Meas}$ .

**Example 34.11.** Let  $(X, \mathcal{B}, \mu)$  be a measure space. Then establish that  $(L^1(\mu), L^\infty(\mu))$  is a compatible couple such that  $\text{Meas}(\mu)$  is a topological vector space into which both  $L^1(\mu)$  and  $L^\infty(\mu)$  are continuously embedded.

### 34.3. Interpolation techniques.

Having set down the preliminary facts, we are now going to state the elementary theorem. First, we state the interpolation of two types of weak boundedness.

**Definition 34.12** (Sublinear mappings). A mapping  $T$  from a Banach space  $X$  to  $\text{Meas}(\mu)$  is said to be sublinear, if

$$\begin{aligned} |T(f+g)(x)| &\leq |Tf(x)| + |Tg(x)| \\ |T(a \cdot f)(x)| &= |a| \cdot |Tf(x)| \end{aligned}$$

for all  $f \in L^\infty(\mu) + L^1(\mu)$ ,  $a \in \mathbb{C}$ .

**Definition 34.13** (Weak- $(p, q)$  bounded operator). A sublinear operator  $T$  is said to be a weak- $(p, q)$  bounded operator from  $L^p(\mu)$  to  $\text{Meas}(\mu)$ , if

$$(34.15) \quad \|T\|_{L^p(\mu) \rightarrow L^{q,\infty}(\mu)} := \sup_{f \in L^p(\mu)} \frac{\|Tf\|_{L^{q,\infty}(\mu)}}{\|f\|_{L^p(\mu)}}.$$

*Exercise 216.* Let  $T$  be a bounded linear operator from  $L^p(\mu)$  to  $L^q(\mu)$ . Show that  $T$  is weak- $(p, q)$  bounded.

Needless to say, a typical example of sublinear weak- $(1, 1)$  bounded operators is the Hardy-Littlewood maximal operator  $M : L^1 + L^\infty(\mathbb{R}^d) \rightarrow \text{Meas}$ , where  $\text{Meas}$  denotes the set of all measurable functions in  $\mathbb{R}^d$ .

The next theorem concerns interpolation of weak- $L^{p_0}$  and  $L^\infty(\mu)$

**Theorem 34.14** (Marcinkiewicz's interpolation theorem I). Let  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{F}, \nu)$  be measure spaces and  $1 \leq p_0 < \infty$ . Let  $T$  be a mapping sending  $L^{p_0}(\mu) + L^\infty(\mu)$  functions to

measurable functions. Assume that  $T$  is sublinear and that there exist  $M_{p_0}$  and  $M_\infty$  such that

$$(34.16) \quad \|Tf\|_{L^{p_0, \infty}(\nu)} \leq M_{p_0} \|f\|_{L^{p_0}(\mu)}$$

$$(34.17) \quad \|Tf\|_{L^\infty(\nu)} \leq M_\infty \|f\|_{L^\infty(\mu)}$$

for all  $f \in L^{p_0}(\mu)$  and  $f \in L^\infty(\mu)$  respectively. Then for all  $p_0 < p < \infty$  we have

$$(34.18) \quad \|Tf\|_{L^p(\nu)} \lesssim_{p, M_{p_0}, M_\infty} \|f\|_{L^p(\mu)}.$$

*Proof.* If  $p_0 = \infty$ , then mimic the proof of the  $L^p(\mathbb{R}^d)$ -boundedness of the Hardy-Littlewood maximal operator  $M$ . If we replace  $M$  with  $T$ , we will obtain the proof. Therefore assume that  $p_0 < \infty$  below.

We begin with changing variables.

$$\|Tf\|_{L^p(\nu)}^p = \int_0^\infty p \lambda^{p-1} \nu\{|Tf| > \lambda\} d\lambda = (1 + M_\infty)^p \int_0^\infty p \lambda^{p-1} \nu\{|Tf| > (1 + M_\infty)\lambda\} d\lambda.$$

We decompose  $f$  at height  $\lambda$ : Set  $f_1 = \chi_{\{|f| > \lambda\}} \cdot f$  and  $f_2 = \chi_{\{|f| \leq \lambda\}} \cdot f$ . Observe that (34.17) yields  $|Tf_2(x)| \lesssim \lambda$   $\nu$ -almost everywhere. The triangle inequality gives

$$(34.19) \quad \{|Tf| > (1 + M_\infty)\lambda\} \leq \{|Tf_1| + |Tf_2| > (1 + M_\infty)\lambda\} \leq \{|Tf_1| > \lambda\}.$$

Therefore, by virtue of (34.16) we have

$$(34.20) \quad \nu\{|Tf| > (1 + M_\infty)\lambda\} \lesssim \frac{1}{\lambda^{p_0}} \int_X |f_1(x)|^{p_0} d\mu(x) = \frac{1}{\lambda^{p_0}} \int_X \chi_{\{|f| > \lambda\}}(x) |f(x)|^{p_0} d\mu(x).$$

Inserting this inequality and using the Fubini theorem, we obtain

$$(34.21) \quad \|Tf\|_{L^p(\nu)}^p \lesssim \int_0^\infty \left( \int_X \lambda^{p-p_0-1} \chi_{\{|f| > \lambda\}}(x) |f(x)|^{p_0} d\mu(x) \right) d\lambda \simeq \int_X |f(x)|^p d\mu(x).$$

This is the desired result. □

Accordingly, we have the result on the interpolation of weak- $L^{p_0}$  and weak- $L^{p_1}$  for  $1 \leq p_0 < p_1 < \infty$ .

**Theorem 34.15** (Marcinkiewicz's interpolation II). *Suppose that  $1 \leq p_0 < p_1 < \infty$ . Let  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{F}, \nu)$  be measure spaces. Suppose that  $T$  is a sublinear mapping from  $L^{p_0}(\mu) + L^{p_1}(\mu)$  to the set of  $\nu$ -measurable functions. Assume*

$$(34.22) \quad \nu\{|Tf| > \lambda\} \leq \frac{M_{p_0}}{\lambda^{p_0}} \int_X |f(x)|^{p_0} d\mu(x)$$

$$(34.23) \quad \nu\{|Tf| > \lambda\} \leq \frac{M_{p_1}}{\lambda^{p_1}} \int_X |f(x)|^{p_1} d\mu(x)$$

for all  $f \in L^{p_0}(\mu) \cap L^{p_1}(\mu)$ . Then we have for all  $p_0 < p < p_1$  that

$$(34.24) \quad \|Tf\|_{L^p(\nu)} \lesssim_{p, p_0, p_1, M_{p_0}, M_{p_1}} \|f\|_{L^p(\mu)}.$$

*Proof.* The proof depends upon the distribution formula again:

$$\begin{aligned} \|Tf\|_{L^p(\nu)}^p &= p \int_0^\infty \rho^{p-1} \nu\{y \in Y : |Tf(y)| > \rho\} d\rho \\ &= p 2^p \int_0^\infty \lambda^{p-1} \nu\{|Tf| > 2\lambda\} d\lambda. \end{aligned}$$

As before, we have changed variables in order to transform the formula to the form of our disposal.

In order to utilize the assumptions we decompose  $f$  at height  $\lambda$ : Split  $f = f_0 + f_1$  with  $f_0 = \chi_{\{|f|>\lambda\}} \cdot f$  and  $f_1 = f - f_0$ . Observe that

$$(34.25) \quad \nu\{|Tf| > 2\lambda\} \leq \nu\{y \in Y : |Tf_0(y)| > \lambda\} + \nu\{y \in Y : |Tf_1(y)| > \lambda\}.$$

Along this decomposition, first of all, we shall estimate

$$(34.26) \quad \int_0^\infty p\lambda^{p-1} \nu\{y \in Y : |Tf_0(y)| > \lambda\} d\lambda.$$

Using assumption (34.22), we have

$$\int_0^\infty p\lambda^{p-1} \nu\{y \in Y : |Tf_0(y)| > \lambda\} d\lambda \lesssim \int_0^\infty \lambda^{p-1} \cdot \left( \frac{1}{\lambda^{p_0}} \int_X |f_0(x)|^{p_0} d\mu(x) \right) d\lambda.$$

Thus, if we make use of the Fubini theorem, it follows that

$$\begin{aligned} \int_0^\infty p\lambda^{p-1} \nu\{y \in Y : |Tf_0(y)| > \lambda\} d\lambda &\lesssim \int_X \left( \int_0^{|f(x)|} \lambda^{p-p_0-1} d\lambda \right) |f(x)|^{p_0} d\mu(x) \\ &\simeq_{p,p_0} \int_X |f(x)|^p d\mu(x). \end{aligned}$$

The treatment of  $f_1$  is the same as that of  $f_0$ . Using assumption (34.23), we have

$$\int_0^\infty p\lambda^{p-1} \nu\{y \in Y : |Tf_1(y)| > \lambda\} d\lambda \lesssim \int_0^\infty \lambda^{p-p_1-1} \cdot \left( \int_{\{|f|\leq\lambda\}} |f(x)|^{p_1} d\mu(x) \right) d\lambda.$$

By the Fubini theorem for positive functions, we obtain

$$\begin{aligned} \int_0^\infty p\lambda^{p-1} \nu\{y \in Y : |Tf_1(y)| > \lambda\} d\lambda &\lesssim \int_X \left( \int_{|f(x)|}^\infty \lambda^{p-p_1-1} d\lambda \right) |f(x)|^{p_1} d\mu(x) \\ &\simeq_{p,p_1} \int_X |f(x)|^p d\mu(x). \end{aligned}$$

Putting together these estimates, we have for all  $p_0 < p < p_1$

$$\|Tf\|_{L^p(\nu)}^p = \int_Y |Tf(y)|^p d\nu(y) \lesssim_{p,p_0,p_1} \int_X |f(x)|^p d\mu(x).$$

This is the desired estimate. □

The next theorem is just a rephrasing of Theorems 34.14 and 34.15. However, it has been playing a key role to obtain the boundedness of the Hardy-Littlewood maximal operator and the singular integral operators.

**Theorem 34.16.**

- (1) *Let  $M$  denote the Hardy-Littlewood maximal operator. Suppose that  $1 < p \leq \infty$ . Then we have*

$$(34.27) \quad \|Mf\|_p \lesssim_p \|f\|_p$$

*for all  $f \in L^p(\mathbb{R}^d)$ .*

- (2) *Suppose that  $T$  is a singular integral operator. Then for  $1 < p < 2$  we have*

$$(34.28) \quad \|Tf\|_p \lesssim_p \|f\|_p$$

*for all  $f \in L^p(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ .*

**Interpolation technique of  $L^p(\mathbb{R}^d)$  and  $BMO(\mathbb{R}^d)$**

Before we conclude this paragraph, let us place ourselves in the setting of  $\mathbb{R}^d$  with the Lebesgue measure and reconsider the function space  $BMO(\mathbb{R}^d)$ . As we have seen,  $BMO(\mathbb{R}^d)$  plays a substitute role of  $L^\infty(\mathbb{R}^d)$ . Therefore, it is natural to guess the following holds.

**Theorem 34.17.** *Suppose that  $S : L^\infty(\mathbb{R}^d) \rightarrow BMO(\mathbb{R}^d)$  and  $T : L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)$  with  $1 < p < \infty$  are bounded linear operators. If*

$$(34.29) \quad Sf = Tf \text{ modulo } \mathbb{C} \text{ for all } f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d),$$

*then  $T$  extends to a bounded linear operator on  $L^q(\mathbb{R}^d)$  with  $p < q < \infty$ .*

*Proof.* Observe that  $M^\sharp$  acts on  $Tf \in L^p(\mathbb{R}^d)$  as well as on  $Sf \in BMO(\mathbb{R}^d)$ , when  $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . In this case the assumption that  $Sf = Tf$  modulo additive constants reads

$$(34.30) \quad M^\sharp[Tf] = M^\sharp[Sf].$$

Meanwhile the boundedness of  $T$  and  $S$  can be rephrased as

$$(34.31) \quad \|M^\sharp \circ T(f)\|_p = \|M^\sharp[Tf]\|_p \lesssim \|f\|_p, \quad \|M^\sharp \circ T(f)\|_\infty = \|M^\sharp[Tf]\|_\infty = \|M^\sharp[Sf]\|_\infty \lesssim \|f\|_\infty$$

for all  $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . If we interpolate this, then we have

$$(34.32) \quad \|M^\sharp[Tf]\|_q \lesssim \|f\|_q$$

for all  $p \leq q < \infty$  with  $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ .

Let  $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . Since  $Tf \in L^p(\mathbb{R}^d)$ , we are in the position of using Theorem 20.5 to obtain

$$(34.33) \quad \|Tf\|_q \lesssim \|M^\sharp[Tf]\|_q.$$

Combining (34.32) and (34.33), we obtain

$$(34.34) \quad \|Tf\|_q \lesssim \|f\|_q.$$

Now we have  $L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  is dense in  $L^q(\mathbb{R}^d)$ , assuming  $q$  is finite. Therefore,  $T$  can be extended to a bounded linear operator on  $L^q(\mathbb{R}^d)$ .  $\square$

35. INTERPOLATION FUNCTORS

In this section we suppose that  $(X_0, X_1)$  is a compatible couple such that  $X_0$  and  $X_1$  are embedded into a topological linear space  $X$ .

We introduce to two methods with which to generate Banach spaces starting from  $(X_0, X_1)$ . In general such a method is called interpolation functor. To be rigorous, this is an abuse of language from algebra. However, we do not allude to this subtle point and we shall be naive to use this terminology in this book.

Here is a general principle: Suppose that  $(X_0, X_1)$  is a compatible couple. The interpolation functors are method with which to create Banach spaces between  $X_0 \cap X_1$  and  $X_0 + X_1$ .

35.1. Real interpolation functors.

**K-method.**

Let us investigate the K-method, which is a standard way to define the real interpolation functor.

**Definition 35.1** (quasi-Banach space  $tX$ ). Let  $t > 0$  and  $X$  a quasi-Banach space. Then denote by  $tX$  a quasi-Banach space which coincides with  $X$  as a set and which is normed by

$$(35.1) \quad \|x\|_{tX} := t\|x\|_X$$

for  $x \in tX (= X)$ .

Roughly speaking, the point of  $K$ -functional is to give a “strict norm” and consider the elements with finite norms.

**Definition 35.2** ( $K$ -functional). The  $K$ -functional of  $(X_0, X_1)$  is a mapping

$$(35.2) \quad K : (X_0 + X_1) \times [0, \infty) \rightarrow [0, \infty)$$

which is given by

$$(35.3) \quad K(x, t) := \inf \{ (\|x_0\|_{X_0} + t\|x_1\|_{X_1}) : x_0 \in X_0, x_1 \in X_1, x = x_0 + x_1 \} = \|x\|_{X_0 + tX_1}.$$

The method using the  $K$ -functional is called the  $K$ -method.

**Definition 35.3** (Real interpolation functor). Let  $0 < \theta < 1$  and  $0 < q \leq \infty$ . Then define  $K_{\theta, q}(X_0, X_1) := (X_0, X_1)_{\theta, q}$  as a subspace of  $X_0 + X_1$  seminormed by

$$(35.4) \quad \|x\|_{(X_0, X_1)_{\theta, q}} = \left( \int_0^\infty (t^{-\theta} K(x, t))^q \frac{dt}{t} \right)^{\frac{1}{q}}.$$

It is elementary to prove the following inequalities.

**Lemma 35.4.** Let  $t, s > 0$  and  $x \in X_0 + X_1$ . Then we have  $K(t, x) \leq \min\left(1, \frac{t}{s}\right) K(s, x)$ .

**Corollary 35.5.** Let  $0 < \theta < 1$  and  $0 < q \leq \infty$ . Then

$$(35.5) \quad \sup_{t>0} t^{-\theta} K(t, x) \lesssim \|x\|_{[X_0, X_1]_{\theta, q}}.$$

*Exercise 217.* Prove Lemma 35.4 and Corollary 35.5.

*Exercise 218.* Show that  $(X_0, X_1)_{\theta, q}$  is complete. Hint: Perhaps, the Fatou lemma will help you.

It is sometimes helpful to use the discrete version of the norms.

**Lemma 35.6.** Let  $0 < q \leq \infty$  and  $0 < \theta < 1$ . Then there exists a constant  $c > 0$  such that

$$(35.6) \quad \left( \sum_{j \in \mathbb{Z}} 2^{-j\theta} K(2^j, x; X_0, X_1)^q \right)^{\frac{1}{q}} \simeq \|x\|_{(X_0, X_1)_{\theta, q}}$$

for all  $x \in X_0 + X_1$ .

*Proof.* Let  $2^j \leq s \leq 2^{j+1}$  with  $j \in \mathbb{Z}$ . Then we have

$$(35.7) \quad K(2^j, x; X_0, X_1) \leq K(s, x; X_0, X_1) \leq 2K(2^j, x; X_0, X_1)$$

for all  $x \in X$ . Therefore, inserting this bilateral estimate to equality

$$\|x\|_{(X_0, X_1)_{\theta, q}} = \left( \sum_{j \in \mathbb{Z}} \int_{2^j}^{2^{j+1}} \left( \frac{K(t, x; X_0, X_1)}{t^\theta} \right)^q \frac{dt}{t} \right)^{\frac{1}{q}},$$

we obtain

$$\|x\|_{(X_0, X_1)_{\theta, q}} \simeq \left( \sum_{j \in \mathbb{Z}} \int_{2^j}^{2^{j+1}} \left( \frac{K(2^j, x; X_0, X_1)}{2^{j\theta}} \right)^q \frac{dt}{t} \right)^{\frac{1}{q}} = \left( \log 2 \sum_{j \in \mathbb{Z}} \frac{K(2^j, x; X_0, X_1)^q}{2^{jq\theta}} \right)^{\frac{1}{q}}.$$

This is the desired result. □

**Lemma 35.7.** *Let  $0 < p \leq q \leq \infty$  and  $0 < \theta < 1$ . Then we have  $(X_0, X_1)_{\theta, p} \hookrightarrow (X_0, X_1)_{\theta, q}$ .*

*Proof.* Instead of using the original norm directly, we use the equivalent norm above. Then the matters are reduced to using  $\ell^p \subset \ell^q$ . □

**Theorem 35.8.** *Let  $0 < q \leq \infty$  and  $0 < \theta < 1$ . Suppose that  $(X_0, X_1)$  and  $(Y_0, Y_1)$  are compatible couples respectively. Let  $T : X_0 + X_1 \rightarrow Y_0 + Y_1$  be a linear operator such that there exists  $M_i > 0$*

$$\|Tx\|_{Y_i} \leq M_i \|x\|_{X_i} \quad (x \in X_i)$$

for each  $i = 0, 1$ . Then we have

$$(35.8) \quad \|Tx\|_{(Y_0, Y_1)_{\theta, q}} \leq M_0^{1-\theta} M_1^\theta \|x\|_{(X_0, X_1)_{\theta, q}}.$$

*Proof.* First, we obtain an estimate of the  $K$ -functional. Note that

$$\begin{aligned} K(Tx, t) &= \inf \{ \|y_0\|_{Y_0} + t \|y_1\|_{Y_1} : y_0 \in Y_0, y_1 \in Y_1, Tx = y_0 + y_1 \} \\ &\leq \inf \{ \|Tx_0\|_{Y_0} + t \|Tx_1\|_{Y_1} ; x_0 \in X_0, x_1 \in X_1, x = x_0 + x_1 \} \end{aligned}$$

because the decomposition  $Tx = Tx_0 + Tx_1$  as above is a special case of the one  $Tx = y_0 + y_1$  with  $y_0 \in Y_0$  and  $y_1 \in Y_1$ . Therefore

$$K(Tx, t) \leq \inf_{\substack{x_0 \in X_0, x_1 \in X_1 \\ x = x_0 + x_1}} M_0 \|x_0\|_{X_0} + t M_1 \|x_1\|_{X_1} = M_0 K \left( x, t \frac{M_1}{M_0} \right).$$

Therefore, inserting this pointwise estimate, we obtain

$$\|Tx\|_{(Y_0, Y_1)_{\theta, q}} = \left( \int_0^\infty (t^{-\theta} K(Tx, t))^q \frac{dt}{t} \right)^{\frac{1}{q}} \leq M_0 \left( \int_0^\infty \left( t^{-\theta} K \left( x, t \frac{M_1}{M_0} \right) \right)^q \frac{dt}{t} \right)^{\frac{1}{q}}.$$

To take full advantage of the Haar measure  $\frac{dt}{t}$ , we change variables. Then we have

$$\|Tx\|_{(Y_0, Y_1)_{\theta, q}} \leq M_0^{1-\theta} M_1^\theta \left( \int_0^\infty (s^{-\theta} K(x, s))^q \frac{ds}{s} \right)^{\frac{1}{q}} \leq M_0^{1-\theta} M_1^\theta \|x\|_{(X_0, X_1)_{\theta, q}}.$$

This implies  $\|Tx\|_{(Y_0, Y_1)_{\theta, q}} \leq M_0^{1-\theta} M_1^\theta \|x\|_{(X_0, X_1)_{\theta, q}}$ . Therefore the proof is complete. □

**J-method.**

There is another way to define a real interpolation functor. This method is called the J-method and turns out to be equivalent to the K-method. With two equivalent methods at our disposal, we can investigate the real interpolation functor in more depth. Unlike the K-method, we give a “light norm” to  $X_0 \cap X_1$  and consider the closure.

**Definition 35.9** (J-functional). Let  $(X_0, X_1)$  be a compatible couple. Then define

$$(35.9) \quad J(t, x) := J(t, x; X_0, X_1) := \max(\|x\|_{X_0}, t\|x\|_{X_1}) = \|x\|_{X_0 \cap_t X_1}$$

for  $x \in X_0 \cap X_1$ .

**Definition 35.10.** Let  $(X_0, X_1)$  be a compatible couple. Suppose that the parameters  $\theta, q$  satisfy  $0 < \theta < 1$  and  $0 < q \leq \infty$ . Define

$$(35.10) \quad J_{\theta,q}(X_0, X_1) := \{x \in X_0 + X_1 : x \text{ satisfies (35.11), (35.12)}\}$$

where the conditions (35.11) and (35.12) are

$$(35.11) \quad x = \sum_{j=-\infty}^{\infty} x_j \text{ in } X_0 + X_1,$$

$$(35.12) \quad \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j))^q < \infty.$$

The quasi-norm of  $J_{q,\theta}(X_0, X_1)$  is given by

$$J_{q,\theta}(X_0, X_1) := \inf \{ \{2^{-j\theta} J(2^j, x_j)\}_{\ell^q(\mathbb{Z})} : \{x_j\}_{j=-\infty}^{\infty} \subset X_0 \cap X_1, \{x_j\}_{j=-\infty}^{\infty} \text{ fulfills (35.11) and (35.12)} \}.$$

Equivalence of K-method and J-method. As is announced in the previous paragraph, we prove that two functors are equivalent. The following lemma is a key to our argument.

**Lemma 35.11.** Assume that  $x \in X_0 + X_1$  satisfies

$$(35.13) \quad \lim_{t \rightarrow \infty} (\min(1, t^{-1})K(t, x)) = \lim_{t \rightarrow 0} (\min(1, t^{-1})K(t, x)) = 0.$$

Then  $x$  admits the following decomposition.

$$(35.14) \quad x = \sum_{j=-\infty}^{\infty} x_j \text{ in } X_0 + X_1,$$

$$(35.15) \quad J(2^j, x_j) \lesssim_{X_0, X_1} K(2^j, x_j).$$

*Proof.* Let  $j \in \mathbb{Z}$  be fixed. Then there exists a decomposition  $x = x_{0,j} + x_{1,j}$  such that

$$(35.16) \quad x_{0,j} \in X_0, x_{1,j} \in X_1, \|x_{0,j}\|_{X_0} + 2^j \|x_{1,j}\|_{X_1} \leq 2K(2^j, x).$$

Set  $x_j = x_{0,j} - x_{0,j-1} \in X_0$ . Since  $x = x_{1,j-1} - x_{1,j}$ ,  $x$  also belongs to  $X_1$ . By assumption we have

$$(35.17) \quad \|x_{0,j}\|_{X_0} \leq 2K(2^j, x) \rightarrow 0, \|x_{1,j}\|_{X_1} \leq 2^{1-j}K(2^j, x) \rightarrow 0$$

as  $j \rightarrow -\infty$  and  $j \rightarrow \infty$  respectively.

The above observation implies  $\lim_{j \rightarrow -\infty} x_{0,j} = 0$  and  $\lim_{j \rightarrow \infty} x_{0,j} = x$  in the topology of  $X_0 + X_1$ . Therefore, it follows that

$$(35.18) \quad \sum_{j=-J}^J x_j = x_{0,J} - x_{0,-J-1} \rightarrow x$$

as  $J \rightarrow \infty$ . Thus, the condition (35.14) is verified.

It remains to check (35.15). By the quasi-triangle inequality we have

$$(35.19) \quad J(2^j, x_j) \lesssim \max(\|x_{0,j}\|_{X_0} + \|x_{0,j-1}\|_{X_0}, 2^j \|x_{1,j}\|_{X_1} + 2^j \|x_{1,j-1}\|_{X_1}) \lesssim K(2^j, x).$$

Thus, the proof is complete. □

**Theorem 35.12.** Let  $0 < \theta < 1$  and  $0 < q \leq \infty$ . Then

$$(35.20) \quad K_{\theta,q}(X_0, X_1) \simeq J_{\theta,q}(X_0, X_1)$$

with norm equivalence.



*Proof.* Let  $x \in J_{\theta,q}(X_0, X_1)$ . Then from the definition of J-method,  $x$  admits a decomposition

$$x = \sum_{j=-\infty}^{\infty} x_j,$$

where  $\{x_j\}_{j=-\infty}^{\infty} \subset X_0 \cap X_1$  and

$$(35.21) \quad \left( \sum_{j=-\infty}^{\infty} 2^{-jq\theta} J(2^j, x_j)^q \right)^{\frac{1}{q}} \leq 2 \|x\|_{J_{\theta,q}(X_0, X_1)}.$$

We have to show

$$(35.22) \quad \left( \sum_{j=-\infty}^{\infty} 2^{-jq\theta} K(2^j, x)^q \right)^{\frac{1}{q}} \lesssim \|x\|_{J_{\theta,q}(X_0, X_1)}.$$

Let  $\rho > 0$  be so small that  $\|\cdot\|_{X_0}^\rho$  and  $\|\cdot\|_{X_1}^\rho$  satisfy the condition of Theorem 10.10. Then we have

$$(35.23) \quad 2^{-j\rho\theta} K(2^j, x)^\rho \lesssim \sum_{k=-\infty}^{\infty} 2^{-j\rho\theta} K(2^j, x_k)^\rho.$$

From the definition of the J-functional we deduce

$$(35.24) \quad K(2^j, x_k) \leq \min(\|x_k\|_{X_0}, 2^j \|x_k\|_{X_1}) = \min(1, 2^{j-k}) J(2^k, x_k).$$

As a consequence we obtain

$$2^{-j\rho\theta} K(2^j, x)^\rho \leq \sum_{k=-\infty}^{\infty} \min(2^{-j\rho\theta}, 2^{-j\rho\theta + \rho(j-k)}) J(2^k, x_k)^\rho.$$

Inserting this estimate, we obtain

$$\begin{aligned} & \left( \sum_{j=-\infty}^{\infty} 2^{-\theta q} K(2^j, x)^q \right)^{\frac{1}{q}} \\ & \lesssim \left( \sum_{j=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} \min(2^{-j\rho\theta}, 2^{-j\rho\theta + \rho(j-k)}) J(2^k, x_k)^\rho \right)^{\frac{q}{\rho}} \right)^{\frac{1}{q}} \\ & = \left( \sum_{j=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} \min(2^{\rho\theta(k-j)}, 2^{\rho(\theta-1)(k-j)}) (2^{-k\theta} J(2^k, x_k))^\rho \right)^{\frac{q}{\rho}} \right)^{\frac{1}{q}}. \end{aligned}$$

Since we are assuming that  $\rho$  is sufficiently small, we are in the position of using the Hölder inequality. Therefore, assuming that  $0 < \theta < 1$ , we have

$$\begin{aligned} & \left( \sum_{k=-\infty}^{\infty} \min(2^{\rho\theta(k-j)}, 2^{\rho(\theta-1)(k-j)}) (2^{-k\theta} J(2^k, x_k))^\rho \right)^{\frac{q}{\rho}} \\ & \leq \left( \sum_{k=-\infty}^{\infty} \min(2^{\frac{\theta}{2}}, 2^{\frac{\theta-1}{2}q(k-j)}) (2^{-k\theta} J(2^k, x_k))^q \right) \cdot \sum_{k=-\infty}^{\infty} \min(2^{\frac{\theta}{2}}, 2^{\frac{\theta-1}{2}})^{\rho(\frac{q}{\rho})'(k-j)} \\ & \lesssim \left( \sum_{k=-\infty}^{\infty} \min(2^{\frac{\theta}{2}}, 2^{\frac{\theta-1}{2}})^{q(k-j)} (2^{-k\theta} J(2^k, x_k))^q \right). \end{aligned}$$

Inserting this inequality, we obtain

$$\begin{aligned} \left( \sum_{j=-\infty}^{\infty} 2^{-\theta q} K(2^j, x)^q \right)^{\frac{1}{q}} &\lesssim \left( \sum_{k,j=-\infty}^{\infty} \min(2^{\frac{\theta}{2}}, 2^{\frac{\theta-1}{2}})^{q(k-j)} (2^{-k\theta} J(2^k, x_k))^q \right)^{\frac{1}{q}} \\ &= \left( \sum_{j=-\infty}^{\infty} 2^{-\theta q} J(2^j, x_j)^q \right)^{\frac{1}{q}} \\ &\lesssim \|x\|_{J_{\theta,q}(X_0, X_1)}. \end{aligned}$$

Therefore, we obtain  $J_{\theta,q}(X_0, X_1) \hookrightarrow K_{\theta,q}(X_0, X_1)$ .

Suppose that  $x \in K_{\theta,q}(X_0, X_1)$ . Then

$$(35.25) \quad \min(1, t^{-1})K(t, x) \lesssim \min(t^\theta, t^{\theta-1}) \left( \int_t^{2t} (s^{-\theta} K(s, x))^q \frac{ds}{s} \right)^{\frac{1}{q}} \rightarrow 0$$

as  $t \rightarrow 0$  and as  $t \rightarrow \infty$ . Thus we can apply Lemma 35.11 to this  $x$ . Note that  $x$  admits a decomposition  $x = \sum_{j=-\infty}^{\infty} x_j$  as in Lemma 35.11.

Therefore, it follows that

$$(35.26) \quad \left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j))^q \right)^{\frac{1}{q}} \lesssim \left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} K(2^j, x))^q \right)^{\frac{1}{q}} \lesssim \|x\|_{(X_0, X_1)_{\theta,q}}.$$

Thus,  $J_{\theta,q}(X_0, X_1) \hookrightarrow K_{\theta,q}(X_0, X_1)$  is established. □

**Corollary 35.13** (Density result). *Let  $0 < \theta < 1$  and  $0 < q < \infty$ . Then  $X_0 \cap X_1$  is dense in  $(X_0, X_1)_{\theta,q}$ .*

Dual space. Let us see the relation between the dual space and the real interpolation.

**Theorem 35.14** (Duality). *Suppose that  $1 \leq q \leq \infty$  and  $0 < \theta < 1$ . Let  $(X_0, X_1)$  be a compatible couple of Banach space such that  $X_0 \cap X_1$  is dense in  $X_0$  and  $X_1$ . Then*

$$(35.27) \quad (X_0, X_1)_{\theta,q}^* = (X_0^*, X_1^*)_{\theta,q'}.$$

*Proof.*  $(X_0, X_1)_{\theta,q}^* \hookrightarrow (X_0^*, X_1^*)_{\theta,q'}$  Let  $x' \in (X_0, X_1)_{\theta,q}^*$ . Then given  $\varepsilon > 0$ , we can find  $x_j \in X_0 \cap X_1$  such that  $\langle x', x_j \rangle \geq 0$  and

$$(35.28) \quad J(2^j, x_j; X_0, X_1) > 0, K(2^{-j}, x'; X_0^*, X_1^*) < \frac{\langle x', x_j \rangle}{J(2^j, x_j; X_0, X_1)} + \varepsilon \min(1, 2^{-j})$$

by virtue of Theorem 34.6. Therefore, we have

$$\begin{aligned} & \left( \sum_{j \in \mathbb{Z}} (2^{j\theta} K(2^{-j}, x'; X_0^*, X_1^*))^{q'} \right)^{\frac{1}{q'}} \\ &= \sup \left\{ \sum_{j \in \mathbb{Z}} 2^{j\theta} a_j K(2^{-j}, x'; X_0^*, X_1^*) : a = \{a_j\}_{j \in \mathbb{N}} \in \ell^q, \|a\|_{\ell^q} = 1 \right\} \\ &\leq C \varepsilon + \sup \left\{ \sum_{j \in \mathbb{Z}} \frac{2^{j\theta} a_j \langle x', x_j \rangle}{J(2^j, x_j; X_0, X_1)} : a = \{a_j\}_{j \in \mathbb{N}} \in \ell^q, \|a\|_{\ell^q} = 1 \right\} \\ &\leq C \varepsilon + \sup \left\{ \left\langle x', \sum_{j \in \mathbb{Z}} \frac{2^{j\theta} a_j x_j}{J(2^j, x_j; X_0, X_1)} \right\rangle : a = \{a_j\}_{j \in \mathbb{N}} \in \ell^q, \|a\|_{\ell^q} = 1 \right\}. \end{aligned}$$

Since

$$(35.29) \quad \left( \sum_{j \in \mathbb{Z}} 2^{-j\theta} J \left( 2^j, \frac{2^{j\theta} a_j x_j}{J(2^j, x_j; X_0, X_1)}; X_0, X_1 \right)^q \right)^{\frac{1}{q}} = \|a\|_{\ell^q} = 1,$$

we obtain

$$(35.30) \quad \left( \sum_{j \in \mathbb{Z}} (2^{j\theta} K(2^{-j}, x'; X_0^*, X_1^*))^{q'} \right)^{\frac{1}{q'}} \leq \|x'\|_{(X_0, X_1)_{\theta, q}} + C \varepsilon.$$

The positive number  $\varepsilon > 0$  being arbitrarily, we obtain

$$(35.31) \quad \left( \sum_{j \in \mathbb{Z}} (2^{j\theta} K(2^{-j}, x'; X_0^*, X_1^*))^{q'} \right)^{\frac{1}{q'}} \leq \|x'\|_{(X_0, X_1)_{\theta, q}}.$$

Thus,  $(X_0, X_1)_{\theta, q^*} \leftrightarrow (X_0^*, X_1^*)_{\theta, q'}$  is proved.

$(X_0, X_1)_{\theta, q^*} \leftrightarrow (X_0^*, X_1^*)_{\theta, q'}$  Suppose that  $x' \in (X_0^*, X_1^*)_{\theta, q'}$ . Then we have

$$(35.32) \quad x' = \sum_{j=-\infty}^{\infty} x'_j \text{ with } \left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x'_j; X_0, X_1))^q \right)^{\frac{1}{q}} \lesssim \|x'\|_{(X_0^*, X_1^*)_{\theta, q'}}.$$

Therefore,

$$|\langle x', x \rangle| \leq \sum_{j=-\infty}^{\infty} |\langle x'_j, x \rangle| \leq \sum_{j=-\infty}^{\infty} J(2^j, x'_j; X_0, X_1) K(2^{-j}, x; X_0, X_1).$$

Using the Hölder inequality, we obtain

$$\begin{aligned} |\langle x', x \rangle| &\leq \left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x'_j; X_0, X_1))^q \right)^{\frac{1}{q}} \left( \sum_{j=-\infty}^{\infty} (2^{j\theta} K(2^{-j}, x; X_0, X_1))^{q'} \right)^{\frac{1}{q'}} \\ &\lesssim \|x' : (X_0^*, X_1^*)_{\theta, q'}\| \left( \sum_{j=-\infty}^{\infty} (2^{j\theta} K(2^{-j}, x; X_0, X_1))^{q'} \right)^{\frac{1}{q'}}. \end{aligned}$$

Therefore, we obtain  $(X_0, X_1)_{\theta, q^*} \leftrightarrow (X_0^*, X_1^*)_{\theta, q'}$ . □

**Iteration theorem.**

Before we start to formulate this theorem, let us start with a heuristic observation. Let  $p_0 := 0 \in \mathbb{R}$  and  $p_1 := 1 \in \mathbb{R}$  and let  $p_\theta := \theta \in [0, 1]$ . From the definition of these points  $p_\theta$  separates  $p_0$  and  $p_1$  by the ratio of  $\theta$  and  $(1 - \theta)$ . Let  $\theta_0, \theta_1, \eta \in [0, 1]$ . Then the point  $q := p_{(1-\eta)\theta_0 + \eta\theta_1}$  separates the intermediate points  $p_{\theta_0}$  and  $p_{\theta_1}$  by the ratio of  $\eta$  and  $1 - \eta$ .

Let us now return to the setting of a compatible couple  $(X_0, X_1)$ . We can regard the intermediate space  $(X_0, X_1)_{\theta_0, q}$  as an intermediate space of order  $\theta_0$ , if we ignore the value  $q$ . Therefore, as a point in  $\mathbb{R}$ , we can identify  $(X_0, X_1)_{\theta_0, q}$  with  $p_{\theta_0}$ . Similarly  $(X_0, X_1)_{\theta_1, r}$  corresponds to  $p_{\theta_1}$ . Let us interpolate  $(X_0, X_1)_{\theta_0, q}$  and  $(X_0, X_1)_{\theta_1, r}$ . This corresponds to the interpolation of  $p_{\theta_0}$  and  $p_{\theta_1}$  according to the interpolation of the points in  $\mathbb{R}$ .

**Theorem 35.15** (Iteration theorem for real interpolation). *Let  $(X_0, X_1)$  be a compatible couple. Suppose that  $0 < \theta_0 < \theta_1 < 1$  and that  $\theta = (1 - \eta)\theta_0 + \eta\theta_1$  with  $0 < \eta < 1$ . Then, for  $0 < q_0, q_1, q \leq \infty$ ,*

$$(35.33) \quad ((X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})_{\eta, q} = (X_0, X_1)_{\theta, q}.$$

*Proof.* First, let us show that

$$(35.34) \quad ((X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})_{\eta, q} \hookrightarrow (X_0, X_1)_{\theta, q}.$$

For this purpose we may assume that  $q_0 = q_1 = \infty$  by virtue of Lemma 35.7. Let

$$(35.35) \quad x \in ((X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty})_{\eta, q}.$$

Consider a decomposition  $x = x_0 + x_1$ , where  $x_0 \in (X_0, X_1)_{\theta_0, \infty}$  and  $x_1 \in (X_0, X_1)_{\theta_1, \infty}$ . Then

$$(35.36) \quad \begin{aligned} K(t, x; X_0, X_1) &\lesssim K(t, x_0; X_0, X_1) + K(t, x_1; X_0, X_1) \\ &\lesssim t^{\theta_0} \|x_0\|_{(X_0, X_1)_{\theta_0, \infty}} + t^{\theta_1} \|x_1\|_{(X_0, X_1)_{\theta_1, \infty}}. \end{aligned}$$

Therefore, since the above decomposition is arbitrary, we have

$$(35.37) \quad K(t, x; X_0, X_1) \lesssim t^{\theta_0} K(t^{\theta_1 - \theta_0}, x; (X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty})$$

Inserting this inequality we have

$$\begin{aligned} \|x\|_{(X_0, X_1)_{\theta, q}} &= \left( \int_0^\infty (t^{-\theta} K(t, x; X_0, X_1))^q \frac{dt}{t} \right)^{\frac{1}{q}} \\ &\lesssim \left( \int_0^\infty (t^{\theta_0 - \theta} K(t^{\theta_1 - \theta_0}, x; (X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty}))^q \frac{dt}{t} \right)^{\frac{1}{q}} \\ &= \left( \int_0^\infty (t^{-\eta(\theta_1 - \theta_0)} K(t^{\theta_1 - \theta_0}, x; (X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty}))^q \frac{dt}{t} \right)^{\frac{1}{q}} \\ &= \left( \int_0^\infty (s^{-\eta} K(s, x; (X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty}))^q \frac{ds}{s} \right)^{\frac{1}{q}} \\ &\simeq \|x\|_{((X_0, X_1)_{\theta_0, \infty}, (X_0, X_1)_{\theta_1, \infty})_{\eta, q}}. \end{aligned}$$

Thus, we obtain  $((X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})_{\eta, q} \hookrightarrow (X_0, X_1)_{\theta, q}$ .

Now let us prove the reverse inclusion. To this end let  $x \in (X_0, X_1)_{\theta, q}$ . Then  $x$  admits a decomposition  $x = \sum_{j=-\infty}^{\infty} x_j$ , where

$$(35.38) \quad \left( \sum_{j=-\infty}^{\infty} 2^{-j\theta q} J(2^j, x_j)^q \right)^{\frac{1}{q}} \lesssim \|x\|_{(X_0, X_1)_{\theta, q}}.$$

Now we estimate  $K(2^k, x_j; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})$ . To do this, we obtain an estimate of  $\|x_j : (X_0, X_1)_{\theta, q_0}\|$ . First of all, we have

$$\begin{aligned} K(t, x_j; X_0, X_1) &\leq \min(\|x_j\|_{X_0}, t\|x_j\|_{X_1}) \\ &\leq \min(J(2^j, x_j; X_0, X_1), 2^{-j}t J(2^j, x_j; X_0, X_1)) \\ &= \min(1, 2^{-j}t) J(2^j, x_j; X_0, X_1). \end{aligned}$$

Therefore we have, for every  $0 < q < \infty$ ,

$$\begin{aligned} \|x_j\|_{(X_0, X_1)_{\theta, q}} &= \left( \int_0^{\infty} (t^{-\theta} K(t, x_j; X_0, X_1))^q \frac{dt}{t} \right)^{\frac{1}{q}} \\ &\leq J(2^j, x_j; X_0, X_1) \left( \int_0^{\infty} (t^{-\theta} \min(1, 2^{-j}t))^q \frac{dt}{t} \right)^{\frac{1}{q}} \\ &\simeq 2^{-j\theta} J(2^j, x_j; X_0, X_1). \end{aligned}$$

Now we turn to the  $K$ -functional with respect to  $(X_0, X_1)_{\theta_0, q_0}$  and  $(X_0, X_1)_{\theta_1, q_1}$ . Inserting the estimate above, we have

$$(35.39) \quad \begin{aligned} K(2^k, x_j; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1}) &\leq \min(\|x_j\|_{(X_0, X_1)_{\theta_0, q_0}}, 2^k \|x_j\|_{(X_0, X_1)_{\theta_1, q_1}}) \\ &\lesssim J(2^j, x_j; X_0, X_1) \min(2^{-j\theta_0}, 2^{k-j\theta_1}). \end{aligned}$$

Therefore, with  $\rho$  sufficiently small, we have

$$\begin{aligned} &(2^{-k\eta} K(2^k, x; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1}))^\rho \\ &\leq 2^{-k\rho\eta} \sum_{j=-\infty}^{\infty} K(2^k, x_j; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})^\rho \\ &\leq \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j; X_0, X_1))^\rho \min(2^{j\theta-k\eta-j\theta_0}, 2^{j\theta+k-k\eta-j\theta_1})^\rho \\ &= \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j; X_0, X_1))^\rho \min(2^{-\eta(k-j(\theta_1-\theta_0))}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))})^\rho. \end{aligned}$$

Assuming  $\rho$  small enough, we obtain by the Hölder inequality

$$\begin{aligned} &\left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j; X_0, X_1))^\rho \min(2^{-\eta(k-j(\theta_1-\theta_0))}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))})^\rho \right)^{\frac{q}{\rho}} \\ &\lesssim \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j; X_0, X_1))^q \min(2^{-\eta(k-j(\theta_1-\theta_0))/2}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))/2})^q. \end{aligned}$$

Here we decomposed the factor as follows:

$$\begin{aligned} & (2^{-j\theta} J(2^j, x_j; X_0, X_1))^\rho \min(2^{-\eta(k-j(\theta_1-\theta_0))}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))})^\rho \\ &= (2^{-j\theta} J(2^j, x_j; X_0, X_1))^\rho \min(2^{-\eta(k-j(\theta_1-\theta_0))/2}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))/2})^\rho \\ & \quad \times \min(2^{-\eta(k-j(\theta_1-\theta_0))/2}, 2^{-(1-\eta)(k-j(\theta_1-\theta_0))/2})^\rho. \end{aligned}$$

Adding this inequality over  $k$ , we obtain

$$\left( \sum_{k=-\infty}^{\infty} (2^{-k\theta} K(2^k, x; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})^q) \right)^{\frac{1}{q}} \lesssim \left( \sum_{j=-\infty}^{\infty} (2^{-j\theta} J(2^j, x_j; X_0, X_1)^q) \right)^{\frac{1}{q}}.$$

By (35.38) we see

$$(35.40) \quad \left( \sum_{k=-\infty}^{\infty} (2^{-k\theta} K(2^k, x; (X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})^q) \right)^{\frac{1}{q}} \lesssim \|x\|_{(X_0, X_1)_{\theta, q}}.$$

Hence the reverse inclusion  $((X_0, X_1)_{\theta_0, q_0}, (X_0, X_1)_{\theta_1, q_1})_{\eta, q} \leftrightarrow (X_0, X_1)_{\theta, q}$  is proved as well.  $\square$

**Remark 35.16.** Let us consider the limiting case. If we use

$$(35.41) \quad K(t, x; X_0, X_1) \lesssim t^{\theta_0} \|x_0\|_{(X_0, X_1)_{\theta_0, \infty}} + t^{\theta_1} \|x_1\|_{X_1}.$$

instead of (35.36) and

$$(35.42) \quad K(2^k, x_j; [X_0, X_1], X_1) \lesssim J(2^j, x_j; X_0, X_1) \max(2^{-j\theta_0}, 2^{k-j})$$

instead of (35.39), then we obtain a result analogous to Theorem 35.15:

$$(35.43) \quad ((X_0, X_1)_{\theta_0, q_0}, X_1)_{\eta, q} = (X_0, X_1)_{(1-\eta)\theta_0 + \eta, q}.$$

*Exercise 219.* Let  $X$  and  $Y$  be Banach spaces such that  $X$  is continuously embedded into  $Y : X \hookrightarrow Y$ . Then

$$(X, Y)_{q_0, \theta} \hookrightarrow (X, Y)_{q_1, \theta'}$$

whenever  $0 < q_0, q_1 \leq \infty$  and  $0 < \theta < \theta' \leq 1$ .

### 35.2. Complex interpolation functors.

In this subsection we assume that  $X_0$  and  $X_1$  are complex Banach spaces.

**Definition 35.17.** Let  $\Omega \subset \mathbb{C}$  be an open set and  $A$  a topological vector space. A continuous function  $f : \Omega \rightarrow A$  is said to be holomorphic, if

$$(35.44) \quad \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists for all  $z_0 \in \Omega$ . In this case one writes  $f'(z_0) := \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$ .

Let  $E$  be a subset on  $\mathbb{C}$  and  $X$  a Banach space. Then define

$$(35.45) \quad \text{BC}(E, X) := \left\{ f : E \rightarrow X : f \text{ is continuous and satisfies } \sup_{z \in E} \|f(z)\|_X < \infty \right\}.$$

In the theory of complex interpolation, the set  $S$  has a special meaning as the following definition shows:

**Definition 35.18.** Set  $S := \{z \in \mathbb{C} : 0 < \operatorname{Re}(z) < 1\}$  and  $\bar{S} := \{z \in \mathbb{C} : 0 \leq \operatorname{Re}(z) \leq 1\}$ . Define

$$(35.46) \quad \mathcal{F}(X_0; X_1) := \{f \in \operatorname{BC}(\bar{S}, X_0 + X_1) : f \text{ satisfies (35.47), (35.48) and (35.49)}\}.$$

Here, conditions (35.47), (35.48) and (35.49) are given by

$$(35.47) \quad f \in \operatorname{BC}(\bar{S}, X_0 + X_1),$$

$$(35.48) \quad f|_{\{z \in S : \operatorname{Re}(z) = i\}} \in C(\{z \in S : \operatorname{Re}(z) = i\})$$

for  $i = 1, 2$  and

$$(35.49) \quad f \in \mathcal{O}(S, X_0 + X_1).$$

Let  $\{f_j\}_{j \in \mathbb{N}}$  be a sequence of  $\mathcal{F}(\mathbb{C}, \mathbb{C})$ , that is, suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of continuous functions on  $\bar{S}$  holomorphic on  $S$ . If  $\{f_j\}_{j \in \mathbb{N}}$  converges in  $\mathcal{F}(\mathbb{C}, \mathbb{C})$  or equivalently in  $\operatorname{BC}(\bar{S})$ , then the limit belongs to  $\mathcal{F}(\mathbb{C}, \mathbb{C})$ . Since  $X_0$  and  $X_1$  are Banach spaces, so is  $X_0 + X_1$ . Hence, we are in the position of developing the theory of the Bochner integral and we conclude that  $\mathcal{F}(X_0, X_1)$  is a Banach space.

**Definition 35.19** (Calderón's first complex interpolation). Define  $[X_0, X_1]_\theta$  by

$$(35.50) \quad [X_0, X_1]_\theta = \{f(\theta) \in X_0 + X_1 : f \in \mathcal{F}(X_0; X_1)\},$$

where its norm is given by

$$(35.51) \quad \|a\|_{[X_0, X_1]_\theta} := \inf \{\|f\|_{\mathcal{F}(X_0; X_1)} : f \in \mathcal{F}(X_0; X_1), a = f(\theta)\}.$$

**Theorem 35.20.** Let  $0 < \theta < 1$ . Suppose that  $(X_0, X_1)$  and  $(Y_0, Y_1)$  are compatible couples respectively. Let  $T : X_0 + X_1 \rightarrow Y_0 + Y_1$  be a linear operator satisfying

$$\|Tx\|_{Y_i} \leq M_i \|x\|_{X_i}$$

for  $i = 0, 1$ . Then we have

$$(35.52) \quad \|Tx\|_{[Y_0, Y_1]_\theta} \leq M_0^{1-\theta} M_1^\theta \|x\|_{[X_0, X_1]_\theta}$$

for all  $x \in [X_0, X_1]_\theta$ .

*Proof.* Let  $x \in [X_0, X_1]_\theta$  and take  $f \in \mathcal{F}(X_0; X_1)$  so that  $f(\theta) = x$ . Note that the mapping  $g$ , given by

$$(35.53) \quad g(z) := \exp(A(z - \theta)) \cdot T(f(z)) \in Y_0 + Y_1,$$

belongs to  $\mathcal{F}(Y_0; Y_1)$ , where  $A$  is a constant fixed later. By assumption we have

$$(35.54) \quad \|Tx\|_{[Y_0, Y_1]_\theta} \leq \sup_{t \in \mathbb{R}, j=0,1} \|g(j + it)\|_{Y_j} \leq \sup_{t \in \mathbb{R}, j=0,1} M_j \exp(A(j - \theta)) \|f(j + it)\|_{X_j}.$$

Choose  $A$  so that  $M_0 \exp(-A\theta) = M_1 \exp(A(1 - \theta))$ , that is,  $A = \log(M_0/M_1)$ . Inserting this, we obtain

$$(35.55) \quad \|Tx\|_{[Y_0, Y_1]_\theta} \leq M_0^{1-\theta} M_1^\theta \max_{j=0,1} \|f(j + it)\|_{X_j}.$$

Since  $f$  is taken arbitrarily, we finally have

$$(35.56) \quad \|Tx\|_{[Y_0, Y_1]_\theta} \leq M_0^{1-\theta} M_1^\theta \|x\|_{[X_0, X_1]_\theta}.$$

This is the desired result. □

**Lemma 35.21** (Three line lemma, Doetsch). Let  $f$  be a  $\mathbb{C}$ -valued bounded continuous function on  $\bar{S}$  whose restriction to  $S$  is holomorphic. Then we have

$$(35.57) \quad |f(\theta)| \leq \sup_{t \in \mathbb{R}} |f(it)|^{1-\theta} \cdot \sup_{t \in \mathbb{R}} |f(1 + it)|^\theta, \theta \in (0, 1).$$

*Proof.* For the sake of simplicity we write  $M_0 := \sup_{t \in \mathbb{R}} |f(it)|$ ,  $M_1 := \sup_{t \in \mathbb{R}} |f(1 + it)|$ . Assume first that  $M_0 = M_1$ . Let  $\varepsilon > 0$  be taken arbitrarily. Then the function  $g(z) = \exp(\varepsilon z^2)f(z)$  satisfies

$$(35.58) \quad \lim_{z \rightarrow \infty, z \in \bar{S}} g(z) = 0,$$

which shows the value of  $|g|$  near  $\infty$  is small relative to the ones near 1. Therefore, we are in the position of using the maximum principle to disprove that  $f$  takes its maximum in  $S$ . For details of the facts on complex analysis of one variable, we refer to [2]. Therefore, we have

$$(35.59) \quad |g(\theta)| \leq \max(e^{-\varepsilon t^2} |f(it)|, e^{\varepsilon(1-t^2)} |f(1 + it)|).$$

A passage to the limit  $\varepsilon \rightarrow 0$  then gives us  $|f(z)| \leq 1$ .

Let us turn to the general case. First if we replace  $f$  by  $f + \delta$  with  $\delta > 0$ , we may assume  $M_0, M_1 \neq 0$ . Set

$$(35.60) \quad h(z) = \exp(-\log M_0 \cdot (1 - z) - \log M_1 \cdot z)f(z).$$

Then  $|h(z)| \leq 1$  on the boundary of  $S$ . Therefore we can apply the above special case to  $h$  and we obtain  $|h(\theta)| \leq 1$ . Equating this inequality, we finally have  $|f(\theta)| \leq M_0^{1-\theta} M_1^\theta$ .  $\square$

Poisson kernel of the strip  $S$ . To deal with complex interpolation it turns out indispensable to obtain some complex analytic information on the functions on  $\bar{S}$ .

**Definition 35.22.** Let  $0 \leq s \leq 1$  and  $t, \tau \in \mathbb{R}$ . Then define

$$(35.61) \quad P_j(s + it, \tau) = \frac{e^{-\pi(\tau-t)} \sin \pi s}{\sin^2 \pi s + (\cos \pi s - e^{ij\pi - \pi(\tau-t)})^2}$$

for  $j = 0, 1$ .

**Lemma 35.23.** *We have*

$$(35.62) \quad \int_{\mathbb{R}} P_0(\theta + it, \tau) d\tau = 1 - \theta, \quad \int_{\mathbb{R}} P_1(\theta + it, \tau) d\tau = \theta$$

for all  $0 \leq s \leq 1$  and  $t \in \mathbb{R}$ .

*Proof.* We deal with  $P_0$ , the treatment of  $P_1$  being the same. Write out the integral in full:

$$(35.63) \quad \int_{\mathbb{R}} P_0(\theta + it, \tau) d\tau = \int_{\mathbb{R}} \frac{e^{-\pi(\tau-t)} \sin \pi \theta}{\sin^2 \pi \theta + (\cos \pi \theta - e^{-\pi(\tau-t)})^2} dt = \int_{\mathbb{R}} \frac{e^{-\pi t} \sin \pi \theta}{\sin^2 \pi \theta + (\cos \pi \theta - e^{-\pi t})^2} dt.$$

Let  $u = e^{\pi t}$ . Then we have

$$(35.64) \quad \int_{\mathbb{R}} P_0(\theta + it, \tau) d\tau = \frac{1}{\pi} \int_0^\infty \frac{\sin \pi \theta}{\sin^2 \pi \theta + (u - \cos \pi \theta)^2} du.$$

Choose a branch of  $\tan^{-1}$  so that  $\tan^{-1}(\mathbb{R}) \subset \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ . Then we have

$$(35.65) \quad \int_{\mathbb{R}} P_0(\theta + it, \tau) d\tau = \frac{1}{\pi} \left[ \tan^{-1} \left( \frac{u - \cos \pi \theta}{\sin \pi \theta} \right) \right]_0^\infty = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left( \tan \left( \frac{\pi}{2} - \pi \theta \right) \right) = 1 - \theta.$$

This is the desired result.  $\square$

As is easily verified, we have the following:



**Lemma 35.24.** *Then the function  $(s, t) \in (0, 1) \times \mathbb{R} \mapsto P_j(s + it, \tau)$  is harmonic with  $\tau \in \mathbb{R}$  fixed. Namely,*

$$(35.66) \quad \left( \frac{\partial^2}{\partial s^2} + \frac{\partial^2}{\partial t^2} \right) P_j(s + it, \tau) = 0$$

for  $j = 0, 1, 0 < s < 1$  and  $t, \tau \in \mathbb{R}$ .

*Exercise 220.* Prove Lemma 35.24.

**Lemma 35.25.** *Let  $f \in \mathcal{F}(X_0; X_1)$ . Then we have the following.*

$$\log \|f(\theta)\|_{(X_0, X_1)_\theta} \leq \sum_{j=0}^1 \int_{\mathbb{R}} \log \|f(j + i\tau)\|_{X_j} P_j(\theta, \tau) d\tau.$$

*Proof.* Set  $f_\varepsilon(z) = e^{-\varepsilon z^2} f(z)$  for  $z \in \bar{S}$ . Then by replacing  $f$  with  $f_\varepsilon$  we may assume that

$$(35.67) \quad \lim_{z \in \bar{S}, z \rightarrow \infty} f(z) = 0.$$

Let  $l \in \mathbb{N}$ . Pick  $\varphi_{0,l}, \varphi_{1,l} \in C^\infty(\mathbb{R}) \cap L^\infty(\mathbb{R})$  such that

$$(35.68) \quad \log(l^{-1} + \|f(j + it)\|_{X_j}) \leq \varphi_{j,l}(t) \leq \log(2l^{-1} + \|f(j + it)\|_{X_j}), \varphi'_j \in L^\infty(\mathbb{R})$$

for  $j = 0, 1$  and  $t \in \mathbb{R}$ . Let  $\Phi(z)$  be an analytic function on  $S$  satisfying

$$(35.69) \quad \operatorname{Re} \Phi(z) = \int_{\mathbb{R}} \varphi_{0,l}(\tau) P_0(z, \tau) d\tau + \int_{\mathbb{R}} \varphi_{1,l}(\tau) P_1(z, \tau) d\tau \text{ for } z \in S.$$

Note that  $\operatorname{Re} \Phi$  is a bounded function with bounded first derivative. Furthermore, we have

$$(35.70) \quad \|\exp(-\Phi(j + it))f(j + it)\|_{X_j} \leq \exp(-\varphi_{j,l}(t)) \cdot \|f(j + it)\|_{X_j} \leq 1.$$

Therefore, applying the three line theorem to  $z \in \bar{S} \mapsto \exp(\eta z^2 - \Phi(z))f(z)$ ,  $\eta > 0$ , we conclude that

$$(35.71) \quad \log \|e^{\eta z^2} f(\theta)\|_{(X_0, X_1)_\theta} \leq \int_{\mathbb{R}} \varphi_0(\tau) P_0(\theta, \tau) d\tau + \int_{\mathbb{R}} \varphi_1(\tau) P_1(\theta, \tau) d\tau.$$

A passage to the limit  $\eta \downarrow 0$  gives us the first inequality. □

We refer to [2] for precise statement of the three line theorem.

Lemma 35.25 is transformed into the following form. The first inequality is a combination of the first inequality and the Jensen inequality. The second one is obtained by using

$$(35.72) \quad ab \leq (1 - \theta)a^{1-\theta} + \theta b^\theta.$$

**Corollary 35.26.** *Keep to the same notation above. Then*

$$\begin{aligned} \|f(\theta)\|_{[X_0, X_1]_\theta} &\leq \left( \int_{\mathbb{R}} \|f(i\tau)\|_{X_0} \frac{P_0(\theta, \tau)}{1 - \theta} d\tau \right)^{1-\theta} \left( \int_{\mathbb{R}} \|f(1 + i\tau)\|_{X_1} \frac{P_1(\theta, \tau)}{\theta} d\tau \right)^\theta \\ \|f(\theta)\|_{[X_0, X_1]_\theta} &\leq \int_{\mathbb{R}} \|f(i\tau)\|_{X_0} P_0(\theta, \tau) d\tau + \int_{\mathbb{R}} \|f(1 + i\tau)\|_{X_1} P_1(\theta, \tau) d\tau. \end{aligned}$$

Second complex interpolation. As we did in the real interpolation method, we seek to find a functor equivalent to the complex interpolation functor. However, unlike the real interpolation we need to postulate another assumption on Banach spaces.

**Definition 35.27** (Calderón’s second complex interpolation). Let  $(X_0, X_1)$  be compatible couples. Then define

$$\mathcal{G}(X_0; X_1) := \{g \in \mathcal{O}(S; X_0 + X_1) \cap \text{BC}(S; X_0 + X_1) : g(j + i^*) \in \text{BC}(X_j) \text{ for } j = 0, 1, \|g\|_{\mathcal{G}(X_0; X_1)} < \infty\},$$

where

$$\|g\|_{\mathcal{G}(X_0; X_1)} := \max \left( \sup_{-\infty < t_1 < t_2 < \infty} \left\| \frac{g(it_1) - g(it_2)}{t_1 - t_2} \right\|_{X_0}, \sup_{-\infty < t_1 < t_2 < \infty} \left\| \frac{g(it_1) - g(it_2)}{t_1 - t_2} \right\|_{X_1} \right).$$

**Definition 35.28.** Define

$$(35.73) \quad [X_0, X_1]^\theta = \{g(\theta) : g \in \mathcal{G}(X_0; X_1)\}$$

and the norm is given by

$$(35.74) \quad \|a\|_{[X_0, X_1]^\theta} = \inf \{ \|g\|_{\mathcal{G}(X_0; X_1)} : g \in \mathcal{G}(X_0; X_1), a = g(\theta) \}.$$

In this paragraph under additional assumption we see that the two methods of complex interpolation coincide.

The following observation made by a skillful usage of theory of integration paves the way to the equivalence theorem.

**Lemma 35.29.** Let  $f \in \mathcal{G}(X_0; X_1)$ . Assume that there exists a measurable set  $E$  such that

$$(35.75) \quad \lim_{h \rightarrow 0} \frac{1}{h} (f(it + ih) - f(it)), t \in E$$

exists in the topology of  $X_0$ . Then we have  $f'(\theta) \in [X_0, X_1]_\theta$  for  $0 < \theta < 1$ .

*Proof.* We set

$$(35.76) \quad f_j(z) = -ij \left( f \left( z + \frac{i}{j} \right) - f(z) \right).$$

Then we have

$$(35.77) \quad \|f_n(it) - f_m(it)\|_{X_0} \rightarrow 0.$$

for all  $t \in E$ . Therefore, we have

$$\log \|f_k(\theta) - f_j(\theta)\|_{[X_0, X_1]^\theta} \leq \sum_{l=0}^1 \int_{\mathbb{R}} \log \|f_k(l + i\tau) - f_j(l + i\tau)\|_{X_j} \cdot P_l(\theta, \tau) d\tau.$$

Note that the integrand of the right-hand side is bounded by an integrable function. Therefore, we are in the position of using the Fatou lemma we obtain

$$(35.78) \quad \lim_{j, k \rightarrow \infty} \log (\|f_k(\theta) - f_j(\theta)\|_{[X_0, X_1]^\theta}) = 0.$$

This implies that  $f'(\theta) = \lim_{j \rightarrow \infty} f_j(\theta)$  belongs to  $[X_0, X_1]_\theta$ . □

**Theorem 35.30.** Let  $X_0, X_1$  be compatible couple. Then

$$(35.79) \quad \|x\|_{[X_0, X_1]^\theta} \leq \|x\|_{[X_0, X_1]_\theta}$$

for all  $x \in [X_0, X_1]_\theta$ . Assume in addition that at least one of  $X_0$  and  $X_1$  is reflexive. Then

$$(35.80) \quad [X_0, X_1]_\theta = [X_0, X_1]^\theta$$

with norm coincidence.

*Proof.*  $\|a\|_{[X_0, X_1]^\theta} \leq \|a\|_{[X_0, X_1]^\theta}$ . Let  $a \in [X_0, X_1]^\theta$  and  $\varepsilon > 0$  be given. Then there exists  $f \in \mathcal{F}(X_0; X_1)$  such that

$$(35.81) \quad a = f(\theta), \|f\|_{\mathcal{F}(X_0; X_1)} \leq \|a\|_{[X_0, X_1]^\theta} + \varepsilon.$$

Define

$$(35.82) \quad g(z) := \int_0^z f(\zeta) d\zeta.$$

Then  $g \in \mathcal{G}(X_0; X_1)$  with  $g'(\theta) = f(\theta) = a$ . Furthermore, we have

$$(35.83) \quad \|g\|_{\mathcal{G}(X_0; X_1)} \leq \|f\|_{\mathcal{F}(X_0; X_1)}.$$

As a result, we obtain

$$(35.84) \quad \|a\|_{[X_0, X_1]^\theta} \leq \|a\|_{[X_0, X_1]^\theta}.$$

The half of the inclusion is therefore proved.

$\|a\|_{[X_0, X_1]^\theta} \geq \|a\|_{[X_0, X_1]^\theta}$ . By symmetry we may assume that  $X_0$  is reflexive. Suppose instead that  $a \in [X_0, X_1]^\theta$ . Define

$$(35.85) \quad f_j(z) = -i \cdot j \left( f \left( z + \frac{i}{j} \right) - f(z) \right)$$

Since  $X_0$  is assumed reflexive, by a diagonal argument and continuity there exists a subsequence  $f_{j_k}$  such that  $f_{j_k}(it)$  is weakly convergent to  $g(t)$  for all  $t \in \mathbb{R}$ . We remark that  $g$  is weakly measurable and separably valued, since each  $f_j$  enjoys the same property. Therefore, the Bochner integral of  $g$  makes sense. A passage to the limit gives us

$$(35.86) \quad f(it) = f(0) + i \cdot t \int_0^t g(\tau) d\tau,$$

since we are in the position of applying the Lebesgue convergence theorem. An argument using the  $X_0$ -valued maximal operator works to obtain

$$(35.87) \quad g(\tau) = \lim_{h \rightarrow 0} \frac{f(it + ih) - f(it)}{h}$$

for almost every  $t \in \mathbb{R}$ . Thus, we are in the position of invoking Lemma 35.29 to conclude that  $f \in \mathcal{F}(X_0, X_1)$ . Since

$$(35.88) \quad \|f\|_{\mathcal{F}(X_0; X_1)} \leq \|f\|_{\mathcal{G}(X_0; X_1)}$$

from the definition, it follows that  $\|a\|_{[X_0, X_1]^\theta} \leq \|a\|_{[X_0, X_1]^\theta}$ . This is the desired result.  $\square$

Density. Let us turn to the density. The lemma below is a key to our observation, whose crux of the proof is the line integral on a complex plane.

**Lemma 35.31.** *The set given below is a dense subset of  $\mathcal{F}(X_0, X_1)_0$ , where  $\mathcal{F}(X_0, X_1)_0$  denotes the closure of the set  $\{\exp(\delta z^2)f : \delta > 0\}$ .*

$$(35.89) \quad \left\{ \exp(\delta z^2) \sum_{j=1}^N \exp(\lambda_j z) x_j : N \in \mathbb{N}, x_1, x_2, \dots, x_N \in X_0 \cap X_1, \lambda_1, \lambda_2, \dots, \lambda_N \in \mathbb{R}, \delta > 0 \right\}.$$

*Proof.* Let  $f \in \mathcal{F}(X_0, X_1)$ . We have to approximate the function  $g(z) = \exp(\delta z^2)f(z)$  of  $z \in S$ . Set

$$(35.90) \quad g_j(z) = \sum_k g(z + 2\pi i k j)$$

for  $j \geq 1$  and  $z \in S$ . Then we have

$$(35.91) \quad \lim_{j \rightarrow \infty} \|g_j - g\|_{\mathcal{F}(X_0, X_1)} = 0.$$

As a consequence the matters are reduced to approximating  $g_j$ . We expand  $g_j$  into a Fourier series;

$$(35.92) \quad g_j(s + it) = \sum_k \exp\left(\frac{ks + ikt}{j}\right) x_{kj}(s),$$

where

$$(35.93) \quad x_{kj}(s + it) = \frac{1}{2\pi jm} \int_{-\pi jm}^{\pi jm} g_j(s + it) e^{-k(s+it)j} dt.$$

Note that

$$(35.94) \quad \frac{1}{2\pi jm} \int_0^1 g_j(s + it) e^{-k(s+it)j} ds \rightarrow 0.$$

Therefore

$$(35.95) \quad x_{kj}(s) = \lim_{m \rightarrow \infty} \frac{1}{2\pi jm} \int_{-\pi jm}^{\pi jm} g_j(s + it) e^{-k(s+it)j} dt$$

is independent of  $s \in [0, 1]$ . This means that  $a_{kj} \in X_0 \cap X_1$ . Therefore, a passage to a limit once more proves the lemma. □

An immediate consequence of this lemma is the following result on density.

**Corollary 35.32** (Density). *Let  $0 < \theta < 1$ . Then  $X_0 \cap X_1$  is dense in  $[X_0, X_1]_\theta$ .*

Duality. Having set down the density property, let us investigate the dual space.

**Theorem 35.33.** *Suppose that  $(X_0, X_1)$  is a compatible couple such that  $X_0 \cap X_1$  are dense in  $X_0$  and  $X_1$  respectively. Then we have*

$$(35.96) \quad (X_0, X_1)_{\theta}^* = (X_0^*, X_1^*)^{\theta}$$

with norm coincidence. If in addition, one of  $X_0$  and  $X_1$  is reflexive, then we have

$$(35.97) \quad (X_0, X_1)_{\theta}^* = (X_0^*, X_1^*)_{\theta}$$

with norm coincidence.

*Proof.*  $\boxed{(X_0, X_1)_{\theta}^* \supset (X_0^*, X_1^*)^{\theta}}$ . Let  $x^* \in (X_0^*, X_1^*)^{\theta}$  and  $\varepsilon > 0$ . Then there exists  $f^* \in \mathcal{G}(X_0^*; X_1^*)$  such that  $f^{*'}(\theta) = x^*$  with

$$(35.98) \quad \|f^*\|_{\mathcal{G}(X_0^*; X_1^*)} \leq (1 + \varepsilon) \|x^*\|_{(X_0^*, X_1^*)_{\theta}}.$$

Let  $x \in (X_0, X_1)_{\theta}$ . Then  $x$  can be realized as  $x = g(\theta)$  with some  $g \in \mathcal{F}(X_0; X_1)$  such that  $\|g\|_{\mathcal{F}(X_0; X_1)} \leq (1 + \varepsilon) \|x\|_{(X_0, X_1)_{\theta}}$ . Therefore,

$$(35.99) \quad x^*(x) = \langle f^{*'}(\theta), g(\theta) \rangle.$$

Observe that  $F(z) = f^{*'}(z)(g(z))$  is holomorphic. Indeed,

$$(35.100) \quad \sup\{\|g(z + h)\|_{X_0 + X_1} : z, h, z + h \in S\} < \infty$$

by virtue of the three line theorem. Therefore, we have

$$\begin{aligned} & \left| \frac{\langle f^{*'}(z+h), g(z+h) \rangle - \langle f^{*'}(z), g(z+h) \rangle - h \langle f^{*''}(z), g(z+h) \rangle}{h} \right| \\ &= \left| \left\langle \frac{f^{*'}(z+h) - f^{*'}(z) - h f^{*''}(z)}{h}, g(z+h) \right\rangle \right| \\ &\leq \left\| \frac{f^{*'}(z+h) - f^{*'}(z) - h f^{*''}(z)}{h} \right\|_{X_0^* + X_1^*} \end{aligned}$$

which tends to 0 as  $h \rightarrow 0$ . In the same way we have

$$(35.101) \quad \frac{f^{*'}(z)(g(z+h)) - f^{*'}(z)(g(z)) - h f^{*'}(z)(g'(z))}{h}$$

tends to 0 as  $h \rightarrow 0$ . Therefore,  $F(z) = \langle f^*(z), g(z) \rangle$  is holomorphic.

By the three line theorem again we have

$$(35.102) \quad |x^*(x)| = |F(\theta)| \leq \max_{z \in \partial S} |F(z)| \leq (1 + \varepsilon)^2 \|x^*\|_{(X_0^*, X_1^*)_\theta} \cdot \|x\|_{(X_0, X_1)_\theta}.$$

Consequently we have  $x^* \in (X_0, X_1)_\theta^*$ .

$$(35.103) \quad \boxed{(X_0, X_1)_\theta^* \hookrightarrow (X_0^*, X_1^*)_\theta.} \quad \text{Let } x^* \in (X_0, X_1)_\theta^*. \text{ Then the mapping } f \in \mathcal{F}(X_0; X_1) \mapsto x^*(f(\theta)) \in \mathbb{C}$$

is continuous. Let us write

$$E := \{(f_0, f_1) \in L^1(\mathbb{R}; A_0) \oplus L^1(\mathbb{R}; A_1) : f_i(\tau) = f(j + i\tau)P_j(0, \tau) \text{ for some } f \in \mathcal{F}(X_0; X_1)\}.$$

Then

$$(35.104) \quad (f_0, f_1) \in E \mapsto x^*(f)$$

where  $f_i(\tau) = f(j + i\tau)P_j(0, \tau)$  is a continuous mapping with norm less than

$$(35.105) \quad \|x^*\|_{(X_0, X_1)_\theta^*} \cdot (\|f_0\|_{L^1(\mathbb{R}; X_0)} + \|f_1\|_{L^1(\mathbb{R}; X_1)}).$$

By the Hahn-Banach theorem, there exists  $g_0 \in L^\infty(X_0^*)$  and  $g_1 \in L^\infty(X_1^*)$  such that

$$(35.106) \quad x^*(f) = \int_{\mathbb{R}} g_0(\tau)(f(\tau))P_0(0, \tau) + g_1(\tau)(f(\tau))P_1(0, \tau) d\tau.$$

Let  $a \in X_0 \cap X_1$ . Then

$$h(\theta)l(a) = l(h(\theta)a) = \int_{\mathbb{R}} (h(i\tau)P_0(\theta, \tau)\langle g_0(\tau), a \rangle + h(1 + i\tau)P_1(\theta, \tau)\langle g_1(\tau), a \rangle) d\tau.$$

We prefer to work on the unit disk instead of dealing with the strip  $S$ . Because we have been investigating a lot on conformal mappings on the unit disk in complex analysis with one variable.

Define

$$(35.107) \quad \mu(z) := \frac{\exp(i\pi z) - \exp(i\pi\theta)}{\exp(i\pi z) - \exp(-i\pi\theta)}.$$

We define  $\tilde{k}_a : \partial\mathbb{D} \rightarrow X_0 + X_1$  by

$$(35.108) \quad \tilde{k}_a(\mu(j + i\tau)) = \langle g_j(\tau), a \rangle.$$

Denote by  $\Delta(1)$  the open unit disk in  $\mathbb{C}$ . Then by change of variables, we obtain

$$(35.109) \quad \int_{\Delta(1)} \tilde{k}_a(z)z^k dz = 0,$$

which yields  $\tilde{k}_a$  is a boundary value of an analytic function  $k_a$  defined on  $\Delta(1)$ . Furthermore, we have

$$(35.110) \quad |k_a(z)| \leq \max(|\langle g_0(\tau), a \rangle|, |\langle g_1(\tau), a \rangle|) \leq \max(\|g_0\|_{L^\infty(X_0)}, \|g_1\|_{L^\infty(X_1)}) \cdot \|a\|_{X_0 \cap X_1}.$$

Thus,  $|k_a(z)| \leq \|l\|_{[X_0, X_1]_\theta}$ . We define

$$(35.111) \quad \langle k(z), a \rangle = k_a(z) \text{ for } z \in S.$$

Then  $k(z) \in X_0^* \cap X_1^* = (X_0 + X_1)^*$ . We define

$$(35.112) \quad g(z) = \int_{\frac{1}{2}}^z k(\zeta) d\zeta$$

for  $z \in \bar{S}$ . Then a passage to the non-tangential limit gives us that

$$(35.113) \quad \langle g(j + it + ih) - g(j + it), a \rangle = i\langle g_j(\tau + h) - g_j(\tau), a \rangle, \quad j = 0, 1.$$

By the density assumption,

$$(35.114) \quad g(j + it + ih) - g(j + it) = ig_j(\tau + h) - g_j(\tau) \in X_j^*.$$

As a result, we obtain  $\|g\|_{\mathcal{G}(X_0; X_1)} = \max(\|g_0\|_{L^\infty(X_0^*)}, \|g_1\|_{L^\infty(X_1^*)})$ . Finally,

$$(35.115) \quad l(f(\theta)) = \int_{\mathbb{R}} (g_0(\tau)P_0(\theta, \tau)f(i\tau) + g_1(\tau)P_1(\theta, \tau)f(1 + i\tau)) d\tau = \langle ig'(\theta), f(\theta) \rangle$$

for all  $f \in \mathcal{F}(X_0; X_1)$ . Thus,  $l = g'(\theta)$  and belongs to  $[X_0^*, X_1^*]^\theta = [X_0^*, X_1^*]_\theta$ . Therefore the proof is complete.  $\square$

Iteration theorem. With our culmination in mind, in this paragraph we prove an iteration theorem. The situation is now more transparent than the real interpolation because we have only to deal with one parameter  $\theta$ .

**Theorem 35.34.** *Suppose that the parameters  $\theta_0, \theta_1, \theta$  and  $\eta$  satisfy*

$$(35.116) \quad 0 < \theta_0, \theta_1, \theta, \eta < 1, \quad \theta = (1 - \eta)\theta_0 + \eta\theta_1.$$

*Assume in addition that  $X_0 \cap X_1$  is dense in  $X_0$  and  $X_1$ . Then*

$$(35.117) \quad [[X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1}]_\eta = [X_0, X_1]_\theta$$

*with norm coincidence.*

*Proof.* It is not so hard to see that

$$(35.118) \quad [[X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1}]_\eta \supset [X_0, X_1]_\theta, \quad [[X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1}]^\eta \supset [X_0, X_1]^\theta.$$

Indeed, let  $f \in \mathcal{F}(X_0; X_1)$ . Then

$$(35.119) \quad g(\eta) := f((1 - \eta)\theta_0 + \eta\theta_1).$$

Then by virtue of the three line theorem we have

$$(35.120) \quad g \in \mathcal{F}([X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1}) \text{ with } \|g\|_{\mathcal{F}([X_0, X_1]_{\theta_0}; [X_0, X_1]_{\theta_1})} \leq \|f\|_{\mathcal{F}(X_0; X_1)}.$$

This implies the first inclusion. The second one being proved similarly.

A passage to the dual gives us

$$(35.121) \quad \|l\|_{([X_0, X_1]_\theta)^*} = \|l\|_{[X_0^*, X_1^*]^\theta} \leq \|l\|_{[[X_0^*, X_1^*]^{\theta_0}, [X_0^*, X_1^*]^{\theta_1}]^\eta} \leq \|l\|_{[[X_0^*, X_1^*]^{\theta_0}, [X_0^*, X_1^*]^{\theta_1}]_\eta}.$$

This is the desired result.  $\square$

Real interpolation and complex interpolation. Finally let us consider the mixture of the real interpolation and the complex interpolation.

**Theorem 35.35.** *Let  $0 < \theta_0 < \theta_1 < 1$  and  $0 < \eta < 1$ . We set  $\theta = (1 - \eta)\theta_0 + \eta\theta_1$ . Suppose that  $(X_0, X_1)$  is a compatible Banach couple. Then we have the following.*

- (1)  $(X_0, X_1)_{\theta,1} \hookrightarrow [X_0, X_1]_{\theta} \hookrightarrow (X_0, X_1)_{\theta,\infty}$
- (2) We have

$$(35.122) \quad ([X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1})_{\eta,p} = (X_0, X_1)_{\theta,p}$$

for all  $0 < p \leq \infty$  with norm equivalence.

- (3) Suppose that  $1 \leq p_0, p_1 \leq \infty$  and  $\frac{1}{p} = \frac{1-\eta}{p_0} + \frac{\eta}{p_1}$ . Then we have

$$(35.123) \quad [(X_0, X_1)_{\theta_0,p_0}, (X_0, X_1)_{\theta_1,p_1}]_{\eta} = (X_0, X_1)_{\theta,p}.$$

*Proof.* (1): Step 1  $(X_0, X_1)_{\theta,1} \hookrightarrow (X_0, X_1)_{\theta}$ . If  $x \in (X_0, X_1)_{\theta,1}$ , then  $x$  can be expressed as

$$(35.124) \quad x = \sum_{j \in \mathbb{Z}} x_j, \text{ with } \sum_{j \in \mathbb{Z}} 2^{-j\theta} \max(\|x_j\|_{X_0}, 2^j \|x_j\|_{X_1}) \leq \|x\|_{(X_0, X_1)_{\theta,1}}$$

by the J-method. Note that  $T_j : a \in \mathbb{C} \mapsto a \cdot x_j \in X_k$  has norm  $\|x_j\|_{X_k}$  for  $k = 0, 1$ . Therefore, by complex interpolation we have

$$\|x_j\|_{[X_0, X_1]_{\theta}} = \|T_j\|_{B(\mathbb{C}, [X_0, X_1]_{\theta})} \leq \|T_j\|_{B(\mathbb{C}, X_0)}^{1-\theta} \cdot \|T_j\|_{B(\mathbb{C}, X_1)}^{\theta} \leq \|x_j\|_{X_0}^{1-\theta} \cdot \|x_j\|_{X_1}^{\theta}$$

for all  $x \in X_0 \cap X_1$ . As a result we obtain

$$(35.125) \quad \|x_j\|_{[X_0, X_1]_{\theta}} \leq 2^{-j\theta} \max(\|x_j\|_{X_0}, 2^j \|x_j\|_{X_1}).$$

Adding them therefore gives us the first inclusion.

(1): Step 2  $(X_0, X_1)_{\theta} \hookrightarrow (X_0, X_1)_{\theta,\infty}$ . Let  $x \in (X_0, X_1)_{\theta}$  be fixed. Then there exists  $f \in \mathcal{F}(X_0; X_1)$  such that

$$(35.126) \quad x = f(\theta), \|f\|_{\mathcal{F}(X_0, X_1)} \leq 2 \|x\|_{[X_0, X_1]_{\theta}}.$$

We define  $g(z) = e^{\alpha(z-\theta)} f(z)$  for  $z \in S$ . Here  $\alpha$  is chosen so that

$$(35.127) \quad e^{\alpha t} \sup_{\tau \in \mathbb{R}} \|g(1+i\tau)\|_{X_1} = \sup_{\tau \in \mathbb{R}} \|g(i\tau)\|_{X_0}.$$

For such  $\alpha$ , we have

$$\begin{aligned} K(t, f(\theta); X_0, X_1) &= K(t, g(\theta); X_0, X_1) \\ &\leq \max(\|g(i\tau)\|_{X_0}, t \|g(1+i\tau)\|_{X_1}) \\ &= t^{\theta} \left( \sup_{\tau \in \mathbb{R}} \|f(i\tau)\|_{X_0} \right)^{1-\theta} \left( \sup_{\tau \in \mathbb{R}} \|f(1+i\tau)\|_{X_1} \right)^{\theta} \\ &= t^{\theta} \|f\|_{\mathcal{F}(X_0; X_1)}. \end{aligned}$$

This is the desired result.

(2) :  $([X_0, X_1]_{\theta_0}, [X_0, X_1]_{\theta_1})_{\eta,p} = (X_0, X_1)_{\theta,p}$ . This is clear because we know that

$$(35.128) \quad ((X_0, X_1)_{\theta_0,1}, (X_0, X_1)_{\theta_1,1})_{\eta,p} = ((X_0, X_1)_{\theta_0,\infty}, (X_0, X_1)_{\theta_1,\infty})_{\eta,p} = (X_0, X_1)_{\theta,p}.$$

This relation and (1) give us the desired result.

(3) Step 1 :  $[(X_0, X_1)_{\theta_0, p_0}, (X_0, X_1)_{\theta_1, p_1}]_\eta \supset (X_0, X_1)_{\theta, p}$ . Let  $x \in (X_0, X_1)_{\theta, p}$ . Then we can decompose  $x$  as follows :

$$(35.129) \quad x = \sum_{j \in \mathbb{Z}} x_j, \left( \sum_{j \in \mathbb{Z}} 2^{-j\theta} J(2^j, x_j; X_0, X_1)^p \right)^{\frac{1}{p}} \lesssim \|x\|_{(X_0, X_1)_{\theta, p}}.$$

Define

$$(35.130) \quad f_j(z) := \left\{ 2^{(\theta_1 - \theta_0)j} \left( \frac{2^{-j\theta} J(2^j, x_j; X_0, X_1)}{\|x\|_{(X_0, X_1)_{\theta, p}}} \right)^{\frac{p}{p_1} - \frac{p}{p_0}} \right\}^{z - \eta} x_j.$$

Define

$$(35.131) \quad f := \sum_{j \in \mathbb{Z}} f_j.$$

Then we have

$$\begin{aligned} & \|f(it)\|_{(X_0, X_1)_{\theta_0, p_0}} \\ & \lesssim \left( \sum_{j \in \mathbb{Z}} (2^{-\theta_0 j} J(2^j, f_j(it); X_0, X_1))^{p_0} \right)^{\frac{1}{p_0}} \\ & \lesssim \left( \sum_{j \in \mathbb{Z}} (2^{-\theta_0 j} J(2^j, x_j; X_0, X_1))^{p_0} \left\{ 2^{(\theta_1 - \theta_0)j} \left( \frac{2^{-j\theta} J(2^j, x_j; X_0, X_1)}{\|x\|_{(X_0, X_1)_{\theta, p}}} \right)^{\frac{p}{p_1} - \frac{p}{p_0}} \right\}^{-p_0 \eta} \right)^{\frac{1}{p_0}} \\ & = \|x\|_{(X_0, X_1)_{\theta, p}}^{(1 - \frac{p_0}{p})} \left( \sum_{j \in \mathbb{Z}} 2^{-j\theta} J(2^j, x_j; X_0, X_1)^p (2^{-j\theta}) \right)^{\frac{1}{p_0}} \\ & = \|x\|_{(X_0, X_1)_{\theta, p}}. \end{aligned}$$

Similarly we can prove that

$$(35.132) \quad \|f(1 + it)\|_{(X_0, X_1)_{\theta_1, p_1}} \lesssim \|x\|_{(X_0, X_1)_{\theta, p}}.$$

Therefore, it follows that  $[(X_0, X_1)_{\theta_0, p_0}, (X_0, X_1)_{\theta_1, p_1}]_\eta \leftrightarrow (X_0, X_1)_{\theta, p}$ .

(3) Step 2 :  $[(X_0, X_1)_{\theta_0, p_0}, (X_0, X_1)_{\theta_1, p_1}]_\eta \leftrightarrow (X_0, X_1)_{\theta, p}$ .

Let  $x \in [(X_0, X_1)_{\theta_0, p_0}, (X_0, X_1)_{\theta_1, p_1}]_\eta$ . Let  $f \in \mathcal{F}((X_0, X_1)_{\theta_0, p_0}; (X_0, X_1)_{\theta_1, p_1})$  with

$$(35.133) \quad \|f\|_{\mathcal{F}((X_0, X_1)_{\theta_0, p_0}; (X_0, X_1)_{\theta_1, p_1})} \leq 2\|x\|_{[(X_0, X_1)_{\theta_0, p_0}, (X_0, X_1)_{\theta_1, p_1}]_\eta}.$$

Let  $\gamma \in \mathbb{R}$  be chosen later and define

$$(35.134) \quad g_j(z) = 2^{(z - \eta)((\theta_0 - \theta_1)j + \gamma)} f_j(z).$$

Then

$$(35.135) \quad x = \int_{\mathbb{R}} P_0(\eta, \tau) g_j(i\tau) d\tau + \int_{\mathbb{R}} P_1(\eta, \tau) g_j(1 + i\tau) d\tau$$



by the Cauchy integral formula. Therefore, choosing  $\gamma$  appropriately, we have

$$\begin{aligned} 2^{-j\theta} K(2^j, x; X_0, X_1) &\leq 2^{-j\theta - j\eta(\theta_0 - \theta_1) - \eta\gamma} \int_{\mathbb{R}} P_0(\eta, \tau) K(2^j, f(i\tau); X_0, X_1) d\tau \\ &\quad + 2^{-j\theta + j(1-\eta)(\theta_0 - \theta_1) + (1-\eta)\gamma} \int_{\mathbb{R}} P_1(\eta, \tau) K(2^j, f(1+i\tau); X_0, X_1) d\tau \\ &\lesssim \left( \int_{\mathbb{R}} 2^{-j\theta_0} P_0(\eta, \tau) K(2^j, f(i\tau); X_0, X_1) d\tau \right)^{1-\eta} \\ &\quad \cdot \left( \int_{\mathbb{R}} 2^{-j\theta_1} P_1(\eta, \tau) K(2^j, f(1+i\tau); X_0, X_1) d\tau \right)^\eta. \end{aligned}$$

It follows from the definition of the norm that

$$\begin{aligned} &2^{-j\theta} K(2^j, x; X_0, X_1) \\ &\lesssim \left( \int_{\mathbb{R}} P_0(\eta, \tau) \|f(i\tau)\|_{(X_0, X_1)_{\theta_0, p_0}} d\tau \right)^{1-\eta} \left( \int_{\mathbb{R}} P_1(\eta, \tau) \|f(1+i\tau)\|_{(X_0, X_1)_{\theta_0, p_1}} d\tau \right)^\eta \\ &\lesssim \|f\|_{\mathcal{F}((X_0, X_1)_{\theta_0, p_0}; (X_0, X_1)_{\theta_1, p_1})} \\ &\lesssim \|x\|_{[(X_0, X_1)_{\theta_0, p_0}; (X_0, X_1)_{\theta_1, p_1}]_\theta}. \end{aligned}$$

This is the desired inequality. □

### 36. INTERPOLATION OF $L^p(\mu)$ -SPACES

In this section, let us see how Banach spaces interpolate.

#### 36.1. Real interpolation of $L^p(\mu)$ -spaces.

A fundamental result. As an example of real interpolation, we shall interpolate  $L^p(\mu)$ -spaces. Throughout this paragraph, we assume that  $(X, \mathcal{B}, \mu)$  is a measure space. We shall give an example of calculation of the  $K$ -functionals.

**Theorem 36.1.** *Let  $0 < p < \infty$ . Then,*

$$(36.1) \quad K(t, f; L^p(\mu), L^\infty(\mu)) \simeq_p \left( \int_0^{t^p} f^*(s)^p ds \right)^{\frac{1}{p}}.$$

If  $p = 1$ , the equivalence  $\simeq_p$  can be replaced with the equality  $=$ .

In the proof, it will be understood tacitly that  $\simeq_1$  stands for  $=$ .

*Proof.* We begin with the following problem to calculate the  $K$ -functional:

$$\begin{aligned} &[\text{Problem}] \text{ Minimize } \|f - g\|_{L^p(\mu)} \\ &\text{under the condition that } g \in L^\infty(\mu), f - g \in L^p(\mu) \text{ and } \|g\|_{L^\infty(\mu)} \leq s \end{aligned}$$

Among all functions of the given  $L^\infty(\mu)$ -norm, we need to look for the function  $g$ . This is achieved as follows: Split  $f$  by  $f = f_0^s + f_1^s$  with  $f_0^s := \max(|f| - s, 0)\text{sgn}f$  and  $f_1^s := \min(|f|, s)\text{sgn}f$ .

We claim that  $f_s^1$  is the desired minimizer. Indeed, if  $g \in L^\infty(\mu)$  has its norm less than  $s$ , then we have

$$\begin{aligned} \left( \int_X |f(x) - g(x)|^p d\mu(x) \right)^{\frac{1}{p}} &\geq \left( \int_X \max(|f(x)| - s, 0)^p d\mu(x) \right)^{\frac{1}{p}} \\ &= \left( \int_X |f(x) - f_s^1(x)|^p d\mu(x) \right)^{\frac{1}{p}}. \end{aligned}$$

Therefore,  $f_s^1$  is the desired minimizer.

Computation of the K-functional. By means of the distribution kernel we have

$$\|f_0^s\|_p = \left( \int_0^\infty \mu\{\max(|f| - s, 0)^p > \lambda\} d\lambda \right)^{\frac{1}{p}} = \left( \int_0^\infty \mu\{|f| > \sqrt[p]{\lambda} + s\} d\lambda \right)^{\frac{1}{p}}.$$

Note that there exists a constant  $c_p \geq 1$  depending only on  $p > 0$  such that

$$(36.2) \quad c_1 = 1, \quad c_p^{-1} (\lambda + s^p)^{\frac{1}{p}} \leq \sqrt[p]{\lambda} + s \leq c_p (\lambda + s^p)^{\frac{1}{p}}.$$

Using this constant  $c_p$ , we see

$$(36.3) \quad \left( \int_{(c_p s)^p}^\infty \mu\{|f| > \sqrt[p]{\lambda}\} d\lambda \right)^{\frac{1}{p}} \leq \|f_0^s\|_p \leq \left( \int_{(c_p^{-1} s)^p}^\infty \mu\{|f| > \sqrt[p]{\lambda}\} d\lambda \right)^{\frac{1}{p}}.$$

Therefore, we have

$$(36.4) \quad \left( \int_0^\infty F(t, c_p \cdot s, \lambda) d\lambda \right)^{\frac{1}{p}} \simeq_p \|f_0^s\|_p + t \min(s, \|f\|_\infty) \leq c'_p \left( \int_0^\infty F(t, c_p^{-1} \cdot s, \lambda) d\lambda \right)^{\frac{1}{p}},$$

where we have written

$$(36.5) \quad F(t, s, \lambda) := t^p \chi_{[0, \min(s, \|f\|_\infty)^p]}(\lambda) + \mu\{|f| > \sqrt[p]{\lambda}\} \chi_{[s^p, \infty)}(\lambda).$$

Inserting (36.5), we obtain

$$\begin{aligned} &\inf\{\|f - g\|_p + t\|g\|_\infty : g \in L^\infty(\mu), f - g \in L^p(\mu)\} \\ &= \inf\{\|f_0^s\|_p + t \min(s, \|f\|_\infty) : s \geq 0\} \\ &\simeq_p \inf\left\{ \left( \int_0^\infty F(t, s, \lambda) d\lambda \right)^{\frac{1}{p}} : s \geq 0 \right\} \\ &= \inf\left\{ \left( \int_0^\infty \left( t^p \chi_{[0, \min(s, \|f\|_\infty)^p]}(\lambda) + \mu\{|f| > \sqrt[p]{\lambda}\} \chi_{[s^p, \infty)}(\lambda) \right) d\lambda \right)^{\frac{1}{p}} : s \geq 0 \right\}. \end{aligned}$$

Note that

$$\begin{aligned} &\int_0^\infty \left( t^p \chi_{[0, \min(s, \|f\|_\infty)^p]}(\lambda) + \mu\{|f| > \sqrt[p]{\lambda}\} \chi_{[s^p, \infty)}(\lambda) \right) d\lambda \\ &\geq \int_0^\infty \left( t^p \chi_{[0, \min(s, \|f\|_\infty)^p]}(\lambda) + \mu\{|f| > \sqrt[p]{\lambda}\} \chi_{[\min(s, \|f\|_\infty)^p, \infty)}(\lambda) \right) d\lambda \\ &\geq \int_0^\infty \min(t^p, \mu\{|f| > \sqrt[p]{\lambda}\}) d\lambda. \end{aligned}$$

and that all the equalities hold for  $s = f^*(t) \leq \|f\|_\infty$ . Furthermore, in this case,

$$\begin{aligned} \int_0^\infty \min(t^p, \mu\{|f| > \sqrt[p]{\lambda}\}) d\lambda &= \int_0^\infty \left( \int_0^{t^p} \chi_{[0, \mu\{|f| > \sqrt[p]{\lambda}\})}(s) ds \right) d\lambda \\ &= \int_0^{t^p} \left( \int_0^\infty \chi_{[0, \mu\{|f| > \sqrt[p]{\lambda}\})}(s) d\lambda \right) ds \\ &= \int_0^{t^p} f^*(s)^p ds. \end{aligned}$$

Here, for the last equality, we have used

$$\begin{aligned} \left( \int_0^\infty \chi_{[0, \mu\{|f| > \sqrt[p]{\lambda}\})}(s) d\lambda \right) &= \sup \{ \lambda \geq 0 : \mu\{|f| > \sqrt[p]{\lambda}\} > s \} \\ &= (\inf \{ \rho \geq 0 : \mu\{|f| > \rho\} \leq s \})^p \\ &= f^*(s)^p. \end{aligned}$$

Inserting this, we obtain the desired result. □

**Corollary 36.2.** *Suppose that  $0 < p_0 < \infty$  and  $0 < \theta < 1$ . Set  $p = \frac{p_0}{1-\theta}$ . Then one has*

$$(36.6) \quad \|f\|_{(L^{p_0}(\mu), L^\infty(\mu))_{\theta, q}} \simeq \left( \int_0^\infty t^{\frac{q}{p}-1} f^*(t)^q dt \right)^{\frac{1}{q}}$$

for  $p_0 \leq q \leq \infty$ .

*Proof.* By Theorem 36.1 we obtain

$$\begin{aligned} \|f\|_{(L^{p_0}(\mu), L^\infty(\mu))_{\theta, q}} &\simeq \left( \int_0^\infty t^{-\theta q} \left( \int_0^{t^{p_0}} f^*(s)^{p_0} ds \right)^{\frac{q}{p_0}} \frac{dt}{t} \right)^{\frac{1}{q}} \\ &\simeq \left( \int_0^\infty t^{-\frac{\theta q}{p_0}} \left( \int_0^t f^*(s)^{p_0} ds \right)^{\frac{q}{p_0}} \frac{dt}{t} \right)^{\frac{1}{q}} \\ &= \left( \int_0^\infty t^{\frac{q}{p}-1} \left( \int_0^1 f^*(ts)^{p_0} ds \right)^{\frac{q}{p_0}} dt \right)^{\frac{1}{q}}. \end{aligned}$$

Since  $f^*$  is decreasing, we obtain

$$\|f\|_{(L^{p_0}(\mu), L^\infty(\mu))_{\theta, q}} \gtrsim \left( \int_0^\infty t^{\frac{q}{p}-1} f^*(t)^q dt \right)^{\frac{1}{q}}.$$

Meanwhile, assuming  $q \geq p_0$ , we are in the position of applying the Minkowski inequality to obtain

$$\begin{aligned} \left( \int_0^\infty t^{\frac{q}{p}-1} \left( \int_0^1 f^*(ts)^{p_0} ds \right)^{\frac{q}{p_0}} dt \right)^{\frac{1}{q}} &\leq \left( \int_0^1 \left( \int_0^\infty f^*(ts)^q t^{\frac{q}{p}-1} dt \right)^{\frac{p_0}{q}} ds \right)^{\frac{1}{p_0}} \\ &\simeq \left( \int_0^1 s^{-\frac{p_0}{p}} \left( \int_0^\infty f^*(ts)^q t^{\frac{q}{p}-1} dt \right)^{\frac{p_0}{q}} ds \right)^{\frac{1}{p_0}} \\ &\simeq \left( \int_0^\infty t^{\frac{q}{p}-1} f^*(t)^q dt \right)^{\frac{1}{q}}. \end{aligned}$$

Therefore, we obtain (36.6). □

**Theorem 36.3.** *Let  $0 < p_0 < p_1 \leq \infty$  and  $0 < \theta < 1$ . Define  $p$  by*

$$(36.7) \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

*Then we have*

$$(36.8) \quad (L^{p_0}(\mu), L^{p_1}(\mu))_{\theta,p} = L^p(\mu)$$

*with norm equivalence.*

*Proof.* Step 1 : Assume  $p_1 = \infty$ . Then by Corollary 36.2 and Theorem 4.12 we have

$$(36.9) \quad \|f\|_{[L^{p_0}(\mu), L^\infty(\mu)]_{\theta,p}} \simeq \left( \int_0^\infty f^*(t)^p dt \right)^{\frac{1}{p}} \simeq \|f\|_p.$$

Therefore, (36.8) holds.

Step 2 : Assume  $p_1 < \infty$ .

It is convenient to use the iteration procedure. Let  $0 < r < p_0$ . Then we have

$$(L^{p_0}(\mu), L^{p_1}(\mu))_{\theta,p} = \left[ (L^r(\mu), L^\infty(\mu))_{1-\frac{r}{p_0}, p_0}, (L^r(\mu), L^\infty(\mu))_{1-\frac{r}{p_1}, p_1} \right]_{\theta,p}.$$

Now that we know the interpolation result for  $L^r(\mu)$  and  $L^\infty(\mu)$ , we obtain

$$(L^{p_0}(\mu), L^{p_1}(\mu))_{\theta,p} \approx (L^r(\mu), L^\infty(\mu))_{(1-\frac{r}{p_0})(1-\theta) + (1-\frac{r}{p_1})\theta, p} = (L^r(\mu), L^\infty(\mu))_{1-\frac{r}{p}, p} \approx L^p(\mu).$$

This is the desired result. □

In view of Corollary 36.2 it is quite meaningful to define the following space:

**Definition 36.4** (Lorentz norm). Let  $0 < p, q \leq \infty$ . Then define the Lorentz space by the norm given by

$$(36.10) \quad \|f\|_{\mathcal{L}^{p,q}(\mu)} := \left( \int_0^\infty t^{\frac{q}{p}-1} f^*(t)^q dt \right)^{\frac{1}{q}}.$$

Here one modifies the definition naturally when  $q = \infty$ .

In terms of Lorentz spaces Corollary 36.2 can be translated as follows :

**Theorem 36.5.** *Suppose that  $0 < p_0 < p < \infty$  and  $0 < q \leq \infty$ . Define  $\theta$  by  $\theta = 1 - \frac{p_0}{p}$ . Then,*

$$(36.11) \quad (L^{p_0}(\mu), L^\infty(\mu))_{\theta,q} = \mathcal{L}^{p,q}(\mu).$$

**Theorem 36.6.** *Suppose that the parameters  $p_0, p_1, q_0, q_1, q$  satisfy*

$$(36.12) \quad 0 < p_0 < p_1 \leq \infty, 0 < q_0 \leq \infty, 0 < q_1 \leq \infty, 0 < q \leq \infty.$$

*Assume in addition that the parameters  $\theta$  and  $p$  fulfill*

$$(36.13) \quad 0 < \theta < 1, \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

*Then we have*

$$(36.14) \quad [\mathcal{L}^{p_0,q_0}(\mu), \mathcal{L}^{p_1,q_1}(\mu)]_{\theta,q} = \mathcal{L}^{p,q}(\mu)$$

*with norm equivalence.*

*Exercise 221.* Prove Theorem 36.6 by mimicking the proof of Theorem 36.3.

*Exercise 222.* Explain Theorems 34.14 and 34.15 in terms of Lorentz spaces.

*Exercise 223.* Show that  $\mathcal{L}^{p,q}(\mu)$  is a Banach space when  $1 < p < \infty$  and  $1 \leq q \leq \infty$ .

*Exercise 224.* Let  $f$  be a  $\mu$ -measurable function. Recall that  $f^*$  is given by

$$(36.15) \quad f^*(t) = \sup \{ \operatorname{ess\,inf}_{x \in E} |f(x)| : E : \text{measurable and } \mu(E) \leq t \}.$$

Then show that the following are equivalent.

- (1)  $f \in L^{p,\infty}(\mu)$ .
- (2)  $\sup_{s>0} s \mu\{|f| > s\}^{1/p} < \infty$ .
- (3)  $\sup_{t>0} t^{1/p} f^*(t) < \infty$ .

*Exercise 225.* Show that the dual of  $L^{1,q}(\mathbb{R})$  is  $\{0\}$ .

Powered space and interpolation. Although it is possible to calculate the exact value of the K-functional  $K(f, t; L^p(\mu), L^\infty(\mu))$ , it was very difficult. In this section we shall present an alternative method of interpolation of  $L^p(\mu)$ -spaces. Here we keep to the assumption that  $(X, \mathcal{B}, \mu)$  is a measure space.

**Definition 36.7** (Powered quasi-Banach space). Given  $0 < p < \infty$  and a quasi-Banach space  $X$ , define  $X^p$  as follows:  $X^p = X$  as a set and the norm is given by

$$(36.16) \quad \|x\|_{X^p} := \|x\|_X^p.$$

**Lemma 36.8.** Let  $0 < p_0 \leq p_1 < \infty$ . We define

$$(36.17) \quad F(z, t; p_0, p_1) := \min_{z_0 \in \mathbb{C}} (|z_0|^{p_0} + t|z - z_0|^{p_1})$$

for  $z \in \mathbb{C} \setminus \{0\}$  and  $t \geq 0$ . Then we have

$$(36.18) \quad F(z, t; p_0, p_1) = |z|^{p_0} F(1, t|z|^{p_1-p_0}; p_0, p_1), \quad F(1, t; p_0, p_1) \leq \min(1, |t|).$$

The proof is straightforward and simple. Therefore, the proof is left for the readers as an exercise.

*Exercise 226.* Prove Lemma 36.8.

Note that in the next theorem the definition of  $p$  is different from those in other theorems.

**Theorem 36.9.** Suppose that  $0 < p_0 < p_1 < \infty$  and  $0 < \eta < 1$ . Define  $p := (1 - \eta)p_0 + \eta p_1$ . Then we have

$$(36.19) \quad (L^{p_0}(\mu)^{p_0}, L^{p_1}(\mu)^{p_1})_{\eta,p} = L^p(\mu)^p.$$

More precisely, there exists  $\alpha(p, p_0, p_1) > 0$  independent of the measure space  $X$  such that

$$(36.20) \quad \|f\|_{(L^{p_0}(\mu)^{p_0}, L^{p_1}(\mu)^{p_1})_{\eta,p}} = \alpha(p, p_0, p_1) \|f\|_{L^p(\mu)^p}.$$

*Proof.* We calculate the K-functional. Under the same notation as Lemma 36.8, we have

$$\begin{aligned} K(t, f; L^{p_0}(\mu)^{p_0}, L^{p_1}(\mu)^{p_1}) &= \inf_{h \in L^{p_0}(\mu)} \int_X (|h(x)|^{p_0} + t|f(x) - h(x)|^{p_1}) d\mu(x) \\ &= \int_{\{f \neq 0\}} F(f(x), t; p_0, p_1) d\mu(x) \\ &= \int_{\{f \neq 0\}} |f(x)|^{p_0} F(1, t|f(x)|^{p_1-p_0}; p_0, p_1) d\mu(x). \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \|f\|_{[L^p(\mu)^p, L^\infty(\mu)^p]_{\eta,1}} &= \int_0^\infty \left( \int_{\{f \neq 0\}} t^{-\eta} |f(x)|^{p_0} F(1, t|f(x)|^{p_1-p_0}; p_0, p_1) d\mu(x) \right) \frac{dt}{t} \\ &= \int_{\{f \neq 0\}} \left( \int_0^\infty t^{-\eta} |f|^{p_0} F(1, t|f|^{p_1-p_0}; p_0, p_1) \frac{dt}{t} \right) d\mu \\ &= \int_{\{f \neq 0\}} \left( \int_0^\infty s^{-\eta} |f|^p F(1, s; p_0, p_1) \frac{ds}{s} \right) d\mu \\ &\simeq_{p,p_0,p_1} \int_{\{f \neq 0\}} |f(x)|^p d\mu(x) \\ &= \|f\|_{L^p(\mu)}^p. \end{aligned}$$

This is the desired result. □

**Theorem 36.10.** *Suppose that  $0 < p_0, p_1, q_0, q_1 \leq \infty$  and  $0 < \theta < 1$ . Define  $p, q$  by*

$$(36.21) \quad \frac{1}{p} := \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} := \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$$

*Let  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{C}, \nu)$  be measure spaces. Assume that*

$$(36.22) \quad T : L^{p_0}(X, \mu) + L^{p_1}(X, \mu) \rightarrow L^{q_0}(Y, \nu) + L^{q_1}(Y, \nu)$$

*is a sublinear operator satisfying*

$$\begin{aligned} \|Tf\|_{L^{q_0}(Y, \nu)} &\leq M_0 \|f\|_{L^{p_0}(X, \mu)} \\ \|Tf\|_{L^{q_1}(Y, \nu)} &\leq M_1 \|f\|_{L^{p_1}(X, \mu)}. \end{aligned}$$

*Then there exists a constant  $\alpha(p, p_0, p_1, q, q_0, q_1) \geq 1$  such that*

$$(36.23) \quad \alpha(p, p_0, p_1, p, p_0, p_1) = 1 \text{ whenever } p_0, p_1 \geq 1$$

$$(36.24) \quad \|Tf\|_{L^q(Y, \nu)} \leq \alpha(p, p_0, p_1, q, q_0, q_1) M_0^{1-\theta} M_1^\theta \|f\|_{L^p(X, \mu)}.$$

*Proof.* As for (36.24) we use Theorems 35.8 and 36.6, while to obtain the precise estimate (36.23) we combine Theorems 35.8 and 36.9 instead. □

### 36.2. Complex interpolation of $L^p(\mu)$ -spaces.

Below let us investigate complex interpolations.

**Theorem 36.11.** *Let  $1 \leq p_0 < p_1 \leq \infty$  and  $0 < \theta < 1$ . Define  $p$  by  $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ . Then we have*

$$(36.25) \quad [L^{p_0}(\mu), L^{p_1}(\mu)]_\theta = L^p(\mu)$$

*with norm coincidence. In particular if  $p = (1-\theta)^{-1} \in (1, \infty)$ , then*

$$(36.26) \quad [L^1(\mu), L^\infty(\mu)]_\theta = L^p(\mu).$$

*Proof.* It is sufficient to prove that

$$(36.27) \quad \|f\|_{[L^{p_0}(\mu), L^{p_1}(\mu)]_\theta} = \|f\|_{L^p(\mu)}$$

for all bounded measurable functions  $f$  such that  $\{f \neq 0\}$  has finite  $\mu$ -measure. Put

$$(36.28) \quad F(z) := \exp(\varepsilon z^2 - \varepsilon \theta^2) |f(x)|^{\frac{p}{p(z)}} \operatorname{sgn}(f)(x),$$

where  $\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}$ . Since  $f \in L^\infty(\mu)$  and  $\mu\{f \neq 0\} < \infty$ , we have  $F \in \mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))$ .

$$(36.29) \quad |F(it)| = \exp(\varepsilon \operatorname{Re}(z^2) - \varepsilon \theta^2) |a(x)|^{\frac{p}{p_0}}$$

implies  $\|F(it)\|_{p_0} \leq e^\varepsilon \|f\|_p$ . Similarly  $\|F(1+it)\|_{p_1} \leq e^\varepsilon \|f\|_p$ . Therefore

$$(36.30) \quad \|F\|_{\mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))} \leq e^\varepsilon \|f\|_p$$

for all  $\varepsilon > 0$ . Hence  $\|f\|_{[L^{p_0}(\mu), L^{p_1}(\mu)]_\theta} \leq \|f\|_p$  is established. To prove the converse inequality we write

$$(36.31) \quad \|f\|_{L^p(\mu)} = \sup_g \left| \int_X f(x) \cdot g(x) d\mu(x) \right|,$$

where  $g$  runs over all elements in  $L^{p'}(\mu) \cap L^\infty(\mu)$  such that  $\{g \neq 0\}$  has finite  $\mu$ -measure and that  $\|g\|_{p'} = 1$ . Let  $\varepsilon > 0$  be arbitrary and choose  $F \in \mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))$  with

$$(36.32) \quad \|f\|_{[L^{p_0}(\mu), L^{p_1}(\mu)]_\theta} \leq (1 + \varepsilon) \|F\|_{\mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))}.$$

Let  $\frac{1}{p'(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}$ . Set

$$(36.33) \quad G(z; x) = \exp(\varepsilon z^2 - \varepsilon \theta^2) |b(x)|^{\frac{p'}{p'(z)}} \operatorname{sgn}(b)(x)$$

Then  $H(z) := \int_X f(z)(x)G(z; x) d\mu(x)$  defines a holomorphic function on  $S$  that satisfies

$$(36.34) \quad |H(it)|, |H(1+it)| \leq e^{2\varepsilon} \|F\|_{\mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))}.$$

Therefore, the three line theorem yields

$$(36.35) \quad \|F\|_p = |H(\theta)| \leq e^{2\varepsilon} \|F\|_{\mathcal{F}(L^{p_0}(\mu), L^{p_1}(\mu))}.$$

Hence the converse inequality was proved. □

As an application, we shall complete the proof of Theorem 4.82 in Chapter 3. Recall that the Young inequality concerns the inequality of convolutions of Lebesgue measurable functions on  $\mathbb{R}^d$ .

**Theorem 36.12.** *Suppose that  $1 \leq p, q, r \leq \infty$  satisfy  $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$ . Then we have*

$$(36.36) \quad \|f * g\|_r \leq \|f\|_p \cdot \|g\|_q$$

for all  $f \in L^p(\mathbb{R}^d)$  and  $g \in L^q(\mathbb{R}^d)$ .

*Proof.* Note that  $\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1$ , provided  $f \in L^1(\mathbb{R}^d)$  and  $g \in L^1(\mathbb{R}^d)$  and that  $\|f * g\|_\infty \leq \|f\|_\infty \cdot \|g\|_1$ , provided  $f \in L^\infty(\mathbb{R}^d)$  and  $g \in L^1(\mathbb{R}^d)$ .

Let  $g \in L^1(\mathbb{R}^d)$ . Then we define  $T : L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d) \rightarrow L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$  by  $Tf = f * g$ . Then the above inequality means  $\|T\|_{L^1} \leq \|g\|_1$  and  $\|T\|_{L^\infty} \leq \|g\|_1$ . By interpolation we have

$$(36.37) \quad \|T\|_{L^p} \leq \|g\|_p, \text{ that is, } \|f * g\|_p \leq \|f\|_p \cdot \|g\|_1.$$

Note that

$$(36.38) \quad \|f * g\|_\infty \leq \|f\|_p \cdot \|g\|_{p'}$$

by virtue of the Schwarz inequality and we now define  $S : L^p(\mathbb{R}^d) + L^\infty(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$  by  $Sg = f * g$ . Then we have

$$(36.39) \quad \|S\|_{L^1 \rightarrow L^p} \leq \|f\|_p, \quad \|S\|_{L^{p'} \rightarrow L^\infty} \leq \|f\|_p.$$

Interpolate this once more, we have

$$(36.40) \quad \|S\|_{L^q \rightarrow L^r} \leq \|f\|_p, \text{ that is, } \|f * g\|_r \leq \|f\|_p \cdot \|g\|_q.$$

This is the desired result.  $\square$

Notes and references for Chapter 16.

We refer to [6] for fundamental facts on interpolation technique. The author has learnt a lot from Professor T. Sobukawa [507].

Gagliardo investigated interpolation of Banach spaces in [200, 201, 202].

Section 34. Theorem 34.4

Theorem 34.6

Marcinkiewicz proved Theorems 34.14, 34.15 in [326]. Theorems 34.14, 34.15 are proved by Marcinkiewicz but the results are not published. Actually Zygmund published the proof due to Marcinkiewicz in 1956. We also refer to [498] as well for these theorems. The term “weak-boundedness” was given by Wolf who reviewed [498].

Theorem 34.16

Theorem 34.17

Section 35. The  $K$ -method was paved by J. Peetre in [386]. Thorin proved the three line theorem [472].

Theorem 35.8 is a typical result in [386], for example.

Theorem 35.12

Theorem 35.14 was obtained by Lions and Peetre. For details we refer to [311, 313, 314, 315, 316].

Lemma 35.25 can be found in [236].

Theorem 35.15 was due to Holmstedt and Peetre [243].

The complex interpolation was obtained independently by Peetre and Lions. The reader is referred to [386, 312].

Theorem 35.20 is a direct consequence of the aforementioned papers such as [386]. However, the origin of this theorem seems to stem from the work by Stein [450], where Stein used the Thorin theorem on the strip  $S$ .

Theorem 35.30

Theorem 35.33

Theorem 35.34

Theorem 35.35 (1) was obtained by Peetre [392]. Theorem 35.35 2,3 were investigated by Karadžov in [266].

Kreuzer considered the complex method for operator norms in [275].

Other interpolation functors were considered in [391, 372].



To conclude this paragraph, let us describe one of the interpolations in [391]. Let  $(X_0, X_1)$  be the compatible couple of Banach spaces. We define  $l_p(A)$  to be the set of sequences  $\{a_j\}_{j \in \mathbb{Z}}$  satisfying

$$(36.41) \quad \|\{a_j\}_{j \in \mathbb{Z}}\|_{\ell^p(A)} := \left( \sum_{j \in \mathbb{Z}} \|a_j\|_{A^p} \right)^{\frac{1}{p}} < \infty.$$

Of course, a natural modification is made when  $p = \infty$ .

Let  $1 \leq p_0, p_1 \leq \infty$ . We say that  $a \in X_0 + X_1$  belongs to  $(A_0, A_1)_{\theta, p_0, p_1}$  if there exists  $\{u_j\}_{j \in \mathbb{Z}} \in X_0 \cap X_1$

$$(36.42) \quad a = \sum_{j \in \mathbb{Z}} u_j, \{2^{-j\theta} u_j\}_{j \in \mathbb{Z}} \in \ell^{p_0}(X_0), \{2^{j(1-\theta)} u_j\}_{j \in \mathbb{Z}} \in \ell^{p_1}(X_1)$$

and the norm is given by

$$(36.43) \quad \inf_{(36.42)} \left[ \max \left( \|\{2^{-j\theta} u_j\}_{j \in \mathbb{Z}}\|_{\ell^{p_0}(X_0)}, \|\{2^{j(1-\theta)} u_j\}_{j \in \mathbb{Z}}\|_{\ell^{p_1}(X_1)} \right) \right].$$

Section 36. As early as 1927, Marcel Markov found that if  $T$  is a linear operator from  $\ell_0^p$  to  $\ell_0^p$  and from  $\ell_1^p$  to  $\ell_1^p$  then for all  $p_0 < p < p_1$   $T$  has the operator norm from  $\ell^p$  to itself.

It is in 1938 that Thorin found the general results on interpolation by using the complex method.

Lorentz spaces, which were defined originally by G. Lorentz, were investigated by W. Luxenberg as well, which were just a part of Luxenberg's paper [39]. Calderón investigated Lorentz spaces in connection with interpolation and established interpolation results such as Theorem 36.6 (See [113]). Hunt synthesised the results of Lorentz spaces in [250]. We remark that Stein-Weiss and M. Riesz established the following interpolation result, which we use in Chapter 19. In [410, 458] the interpolation with change of measures were considered. If a linear operator  $T : \text{Meas}(X, \mu) \rightarrow \text{Meas}(Y, \nu)$  satisfies

$$(36.44) \quad \|Tf\|_{L^{q_i}(\nu)} \leq M_i \|f\|_{L^{p_i}(\mu)}, \quad i = 0, 1,$$

then we have

$$(36.45) \quad \|Tf\|_{L^q(\nu)} \lesssim_{\theta} M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p(\mu)},$$

where the parameters satisfy

$$(36.46) \quad 1 \leq p_i, q_i \leq \infty, \quad 0 \leq \theta \leq 1,$$

and

$$(36.47) \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

For details we refer to [410, 458].

Theorems 36.1 and 36.3 can be found in [386]. Theorem 36.1 dates back to Holmstedt's result in 1970.

Theorem 36.5

Theorem 36.9

Theorem 36.10

Theorem 36.11 was obtained in [113].

## Theorem 36.12

Nilsson obtained interpolation results for Banach couples such that the couple carries the structure of lattice. For details we refer to [372].

**Part 17. Wavelets**

In Chapter 17 we exhibit an application of theory of singular integral operators. Wavelet is one of the mathematical subjects which grew up so rapidly based on mathematical preparations of other field. We place ourselves in the setting of the Euclidean space  $\mathbb{R}^d$  as usual. Wavelet is applied to analysis of image processing. But the mathematical background largely depends on harmonic analysis. In Section 37 we make a brief review of wavelet theory. Section 38 is devoted to developing theory of unconditional basis, where we develop our theory in a general Banach space  $B$ . As an application of wavelet in Section 39 we shall construct “basis” of  $L^p(\mathbb{R})$  spaces with  $1 < p < \infty$ .

37. WAVELETS AND SCALING FUNCTIONS

37.1. Definition.

Below, Definition 37.1 is a definition in terms of Grafakos.

**Definition 37.1.** An  $L^2(\mathbb{R}^d)$ -function  $\psi : \mathbb{R}^d \rightarrow \mathbb{C}$  is said to be a wavelet (function) if the system  $\{\psi_{j,k}\}_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} = \{2^{\frac{jd}{2}} \psi(2^j x - k)\}_{j \in \mathbb{Z}, k \in \mathbb{Z}^d}$  forms an orthonormal basis in  $L^2(\mathbb{R}^d)$ .

**Definition 37.2.** A multi-resolution analysis, which is abbreviated to MRA, is a sequence  $\{V_j\}_{j \in \mathbb{Z}}$  of closed subspaces of  $L^2(\mathbb{R}^d)$  that fulfills the following conditions listed below :

- (1) (Nested property)  $V_j \subset V_{j+1}$  for all  $j \in \mathbb{Z}$ .
- (2)  $\bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R}^d)$ .
- (3)  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ .
- (4)  $V_0$  is spanned by the set of the integer translates of a function  $\varphi$ . Also,  $\{\varphi(* - m)\}_{m \in \mathbb{Z}^d}$  is an orthonormal sequence.
- (5) (Scaling property) Let  $j \in \mathbb{Z}$ . Then  $f(2^j *) \in V_j$  precisely when  $f \in V_0$ .

The function  $\varphi$  is said to be a scaling function of the multi-resolution analysis  $\{V_j\}_{j \in \mathbb{Z}}$ .

Recall also that the Fourier transform is defined by  $\mathcal{F}f(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx$  for  $f \in L^1(\mathbb{R}^d)$ .

**Lemma 37.3.** Let  $g \in L^2(\mathbb{R}^d)$ . In order that  $\{g(* - m)\}_{m \in \mathbb{Z}^d}$  forms a CONS in  $L^2(\mathbb{R}^d)$  it is necessary and sufficient that

$$(37.1) \quad \sum_{k \in \mathbb{Z}^d} |\mathcal{F}g(\xi + 2\pi k)|^2 = \frac{1}{(2\pi)^d}$$

for a.e.  $\xi \in \mathbb{R}^d$ .

*Proof.* Let  $k \in \mathbb{Z}^d$ . First, we observe

$$\int_{\mathbb{R}^d} g(x) \overline{g(x - k)} dx = \int_{\mathbb{R}^d} |\mathcal{F}g(\xi)|^2 e^{ik \cdot \xi} d\xi = \sum_{l \in \mathbb{Z}^d} \int_{[0, 2\pi]^d} |\mathcal{F}g(\xi + 2\pi l)|^2 e^{ik \cdot \xi} d\xi.$$

Here, for the first equality, we used the Planchrel theorem and, for the second equality, we employed the partition  $\mathbb{R}^d = \bigcup_{l \in \mathbb{Z}^d} (2\pi l + [0, 2\pi]^d)$  and the fact that  $e^{ik \cdot \xi}$  is  $2\pi$ - $\mathbb{Z}$  periodic.

Therefore, since  $\sum_{l \in \mathbb{Z}^d} |\mathcal{F}g(* + 2\pi l)|^2 \in L^1([0, 2\pi]^d)$ , if we expand this function by the Fourier series,

$$(37.2) \quad \sum_{l \in \mathbb{Z}^d} |\mathcal{F}g(* + 2\pi l)|^2 = \frac{1}{(2\pi)^d} \sum_{k \in \mathbb{Z}^d} \left( \int_{\mathbb{R}^d} g(x) \overline{g(x-k)} dx \right) e^{-ik \cdot \xi}.$$

From this expansion, the lemma is easy to prove.  $\square$

Recall that  $V_0$  is a linear subspace of  $V_1$ . We let  $W_0$  be the orthogonal complement of  $V_0$  in  $V_1$ , that is,  $V_1 = V_0 \oplus W_0$ . We define  $W_j := \{f \in L^2(\mathbb{R}^d) : f(2^j *) \in L^2(\mathbb{R}^d)\}$ . Then, from the scaling property, we have  $V_{j+1} = V_j \oplus W_j$ . Assuming that  $V_j \rightarrow \{0\}$  as  $j \rightarrow -\infty$ , that is,  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ , we have

$$(37.3) \quad V_{j+1} = V_j \oplus W_j = V_{j-1} \oplus W_{j-1} \oplus W_j = V_{j-2} \oplus W_{j-2} \oplus W_{j-1} \oplus W_j = \dots = \bigoplus_{k=-\infty}^j W_k.$$

Since  $\overline{\bigcup_{j=-\infty}^{\infty} V_j} = L^2(\mathbb{R}^d)$ , we have  $L^2(\mathbb{R}^d) = \bigoplus_{k=-\infty}^{\infty} W_k$ .

The goal for the time being is to find  $\psi \in L^2(\mathbb{R}^d)$  with norm 1 such that  $W_0 = \bigoplus_{m \in \mathbb{Z}^d} \mathbb{C}\psi(* - m)$ . Occasionally, we write  $\psi_m = \psi(* - m)$ .

Given an MRA  $\{V_j\}_{j=-\infty}^{\infty}$ , we remark that the scaling function  $\varphi$  is not unique. Here, keeping in mind that the scaling function  $\varphi$  is not unique, we still choose and fix a scaling function  $\varphi$ . Since  $V_0 = W_{-1} \oplus V_{-1}$ , we have  $\frac{1}{2^d} \varphi\left(\frac{*}{2}\right) \in V_{-1}$ . Since  $V_{-1} \subset V_0 = \overline{\text{Span}(\{\varphi(* - m)\}_{m \in \mathbb{Z}^d})}$ , we have

$$(37.4) \quad \frac{1}{2^d} \varphi\left(\frac{x}{2}\right) = \sum_{k \in \mathbb{Z}^d} a_k \varphi(x + k).$$

Here the coefficient is given by  $a_k := \left\langle \frac{1}{2^d} \varphi\left(\frac{*}{2}\right), \varphi(* + k) \right\rangle_{L^2(\mathbb{R}^d)}$ . Taking the Fourier transform of the both sides of (37.4), we have

$$(37.5) \quad \mathcal{F}\varphi(2\xi) = \left( \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot \xi} \right) \mathcal{F}\varphi(\xi).$$

The function  $m_0(\xi) = \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot \xi}$  has a special name.

**Definition 37.4** (Low pass filter). The function  $m_0 : \mathbb{R}^d \rightarrow \mathbb{C}$ , which is given by

$$(37.6) \quad m_0(\xi) = \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot \xi},$$

is called the low pass filter of a scaling function  $\varphi$ .

**Lemma 37.5.** The equality  $\sum_{l \in \{0,1\}^d} |m_0(\xi + \pi l)|^2 = 2^d$  holds for a.e.  $\xi \in \mathbb{R}^d$ .

*Proof.* By Lemma 37.3 we have

$$(37.7) \quad \sum_{k \in \mathbb{Z}^d} |\mathcal{F}\varphi(\xi + 2\pi k)|^2 = \frac{1}{(2\pi)^d}$$

for a.e.  $\xi \in \mathbb{R}^d$ . From (37.5) and (37.7), we have

$$(37.8) \quad \frac{1}{(2\pi)^d} = \sum_{k \in \mathbb{Z}^d} |\mathcal{F}\varphi(2\xi + 2\pi k)|^2 = \sum_{k \in \mathbb{Z}^d} |\mathcal{F}\varphi(\xi + \pi k)|^2 \cdot |m_0(\xi + \pi k)|^2.$$

Since  $\xi$  is arbitrary, we have

$$(37.9) \quad \frac{1}{(2\pi)^d} = \sum_{k \in \mathbb{Z}^d} |\mathcal{F}\varphi(\xi + \pi l + \pi k)|^2 \cdot |m_0(\xi + \pi l + \pi k)|^2.$$

Add (37.9) over  $l \in \{0, 1\}^d$ . Then we have

$$\frac{1}{\pi^d} = \sum_{k \in \mathbb{Z}^d} \sum_{l \in \{0, 1\}^d} |\mathcal{F}\varphi(\xi + 2\pi k + \pi l)|^2 \cdot |m_0(\xi + 2\pi k + \pi l)|^2.$$

A change of the order of integration then gives us

$$\begin{aligned} \frac{1}{\pi^d} &= \sum_{k \in \mathbb{Z}^d} \sum_{l \in \{0, 1\}^d} |\mathcal{F}\varphi(\xi + 2\pi k + \pi l)|^2 \cdot |m_0(\xi + \pi l)|^2 \\ &= \sum_{l \in \{0, 1\}^d} \left( \sum_{k \in \mathbb{Z}^d} |\mathcal{F}\varphi(\xi + 2\pi k + \pi l)|^2 \right) \cdot |m_0(\xi + \pi l)|^2. \end{aligned}$$

If we invoke Lemma 37.3 once again, then we have

$$\frac{1}{\pi^d} = \frac{1}{(2\pi)^d} \sum_{l \in \{0, 1\}^d} |m_0(\xi + \pi l)|^2.$$

Therefore, Lemma 37.5 is now proved. □

### 37.2. Construction of wavelets for $d = 1$ .

Let us construct an orthonormal wavelets starting from MRA.

**Lemma 37.6.** *Let  $\{V_j\}_{j \in \mathbb{Z}}$  be an MRA,  $\varphi$  a scaling function and  $m_0$  its low pass filter. Then we have*

$$\begin{aligned} V_{-1} &= \{f \in L^2(\mathbb{R}) : \mathcal{F}f(\xi) = m(2\xi)m_0(\xi)\mathcal{F}\varphi(\xi), m \in L^2(\mathbb{T})\} \\ V_0 &= \{f \in L^2(\mathbb{R}) : \mathcal{F}f(\xi) = \ell(\xi)m_0(\xi)\mathcal{F}\varphi(\xi), \ell \in L^2(\mathbb{T})\}, \end{aligned}$$

where  $L^2(\mathbb{T}) := \left\{ f \in L^1_{\text{loc}}(\mathbb{R}) : f(* + 2\pi k) = f \text{ for all } k \in \mathbb{Z}, \int_{[0, 2\pi]} |f(x)|^2 dx < \infty \right\}$ .

We remark that the notation of  $L^2(\mathbb{T})$  is slightly different from that in Chapter 1.

*Proof.* Let us denote by  $Z_{-1}$  and  $Z_0$ , respectively the spaces appearing in the right-hand side.

Let  $f \in V_{-1}$  and let us prove that  $f$  has an expression described in the right-hand side. Then it follows from the scaling property and the fact that  $V_0 = \overline{\text{Span}}(\{\varphi(* - m)\}_{m \in \mathbb{Z}^d})$  that there exists  $c = \{c_k\}_{k \in \mathbb{Z}} \in \ell^2$  such that

$$f = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} c_k \varphi\left(\frac{1}{2} * - k\right),$$

where the convergence takes place in  $L^2(\mathbb{R})$ . Therefore, taking the Fourier transform, we are led to  $\mathcal{F}f(\xi) = \sqrt{2}\mathcal{F}\varphi(2\xi) \left( \sum_{k \in \mathbb{Z}} c_k e^{-2k \cdot \xi} \right)$ . We claim that the desired  $L^2(\mathbb{T})$ -function  $m$  is given by  $m(\xi) = \sqrt{2} \sum_{k \in \mathbb{Z}} c_k e^{-k \cdot \xi}$ . Indeed, from the definition we have

$$(37.10) \quad \mathcal{F}f(\xi) = m(2\xi)\mathcal{F}\varphi(2\xi) = m(2\xi)m_0(\xi)\mathcal{F}\varphi(\xi).$$

Consequently  $V_{-1} \subset Z_{-1}$  is established.

To prove the reverse inclusion, we have to make a few preparatory observations. First, if  $h \in L^2(\mathbb{T})$ , then  $h \cdot \mathcal{F}\varphi \in L^2(\mathbb{R})$ .

Indeed, taking into account the periodicity of  $h$ , we have

$$\begin{aligned} \int_{\mathbb{R}} |h(\xi)\mathcal{F}\varphi(\xi)|^2 d\xi &= \sum_{k \in \mathbb{Z}} \int_{[0, 2\pi]} |h(\xi + 2\pi k)\mathcal{F}\varphi(\xi + 2\pi k)|^2 d\xi \\ &= \sum_{k \in \mathbb{Z}} \int_{[0, 2\pi]} |h(\xi)\mathcal{F}\varphi(\xi + 2\pi k)|^2 d\xi \\ &= \sum_{k \in \mathbb{Z}} \int_{[0, 2\pi]} |h(\xi)\mathcal{F}\varphi(\xi)|^2 d\xi \end{aligned}$$

By Lemma 37.3, we have

$$\int_{\mathbb{R}} |h(\xi)\mathcal{F}\varphi(\xi)|^2 d\xi = \frac{1}{(2\pi)} \sum_{k \in \mathbb{Z}} \int_{[0, 2\pi]} |h(\xi)|^2 d\xi < \infty.$$

Since  $m_0$  is also essentially bounded from Lemma 37.5, it follows that  $m(2^*)m_0\mathcal{F}\varphi \in L^2(\mathbb{R})$ .

With this in mind, we define  $f \in L^2(\mathbb{R})$  by  $f := \mathcal{F}^{-1}(m(2^*)m_0\mathcal{F}\varphi)$ . Then since  $m \in L^2(\mathbb{T})$ , as before there exists  $c = \{c_k\}_{k \in \mathbb{Z}} \in \ell^2$  such that  $m = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} c_k e^{-ik \cdot \xi}$ . Inserting this expression along with the formula  $\mathcal{F}\varphi(2\xi) = m_0(\xi)\mathcal{F}\varphi(\xi)$ , we obtain

$$(37.11) \quad f = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} c_k \varphi \left( \frac{1}{2} * -k \right) \in L^2(\mathbb{R}).$$

The characterization for  $V_0$  being the same, the proof is now complete. □

In the course of the proof, we have obtained the following: A direct consequence of Lemma 37.5 is that  $m_0 \in L^\infty(\mathbb{R})$ . Thus, we can rewrite the definition of  $V_{-1}$  and  $V_0$  as follows:

$$\begin{aligned} V_{-1} &= \{ \mathcal{F}^{-1}[m(2^*)m_0\mathcal{F}\varphi] : m \in L^2(\mathbb{T}) \}. \\ V_0 &= \{ \mathcal{F}^{-1}[\ell m_0\mathcal{F}\varphi] : \ell \in L^2(\mathbb{T}) \}. \end{aligned}$$

**Lemma 37.7.** *We have*

$$(37.12) \quad W_{-1} = \{ f \in L^2(\mathbb{R}) : \mathcal{F}f(\xi) = e^{i\xi} s(2\xi) \overline{m_0(\xi + \pi)} \mathcal{F}\varphi(\xi), s \in L^2(\mathbb{T}) \}.$$

*Proof.* We define an operator  $U : V_0 \rightarrow L^2(\mathbb{T})$  by

$$Uf(\xi) := \begin{cases} \frac{\mathcal{F}f(\xi)}{\mathcal{F}\varphi(\xi)} & \text{if } \mathcal{F}\varphi(\xi) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then using the expansion  $f = \sum_{j \in \mathbb{Z}} d_j \varphi(* - j)$ , we have  $Uf = \chi_{\text{supp}(\varphi)} \cdot \sum_{j \in \mathbb{Z}} d_j e^{ij \cdot *}$ . Therefore,

$$(37.13) \quad \|Uf\|_{L^2(\mathbb{T})} = \|\{d_j\}_{j \in \mathbb{Z}}\|_{\ell^2(\mathbb{Z})} = \|f\|_{L^2(\mathbb{R})},$$

which implies  $U$  is an isometry. By polarization, this immediately leads us

$$(37.14) \quad \langle f, g \rangle_{L^2(\mathbb{R})} = \langle Uf, Ug \rangle_{L^2(\mathbb{T})}$$

for all  $f, g \in V_0$ .

From (37.14) we deduce that  $f \in W_{-1}$  if and only if  $Uf$  is perpendicular to  $m(2*)m_0$  for all  $m \in L^2(\mathbb{T})$ . Therefore

$$(37.15) \quad \ell(\xi) \overline{m_0(\xi)} + \ell(\xi + \pi) \overline{m_0(\xi + \pi)} = 0.$$

Since  $(\overline{m_0(\xi)}, \overline{m_0(\xi + \pi)})$  is a unit vector, there exists a  $2\pi$ -periodic function  $\lambda$  such that

$$(37.16) \quad \ell(\xi) = -\lambda(\xi + \pi) \overline{m_0(\xi + \pi)}, \ell(\xi + \pi) = \lambda(\xi) \overline{m_0(\xi)}.$$

If we substitute  $\xi + \pi$  instead of  $\xi$ , then

$$(37.17) \quad \ell(\xi + \pi) = -\lambda(\xi + 2\pi) \overline{m_0(\xi + 2\pi)}, \ell(\xi + 2\pi) = \lambda(\xi + \pi) \overline{m_0(\xi + \pi)}.$$

Taking into account the periodicity again, we conclude

$$(37.18) \quad \ell(\xi + \pi) = -\lambda(\xi) \overline{m_0(\xi)}, \ell(\xi) = \lambda(\xi + \pi) \overline{m_0(\xi + \pi)}.$$

In view of (37.16) and (37.18), it follows that  $\lambda(\xi + \pi) = -\lambda(\xi)$ . Finally if we set

$$(37.19) \quad s(\xi) := e^{-i\frac{\xi}{2}} \lambda\left(\frac{\xi}{2}\right),$$

we obtain

$$(37.20) \quad \ell(\xi) = e^{i\xi} s(2\xi) \overline{m_0(\xi + \pi)}.$$

Note that  $\ell$  is  $\pi$ -periodic. For the reverse inclusion, we have to reverse the argument above.  $\square$

The next theorem characterizes a condition for  $\psi \in W_0 = V_1 \cap V_0^\perp$  to be an orthonormal wavelet.

**Theorem 37.8.** *Let  $\{V_j\}_{j \in \mathbb{Z}}$  be MRA with scaling function  $\varphi$ . Denote by  $m_0$  its low-pass filter. Then  $\psi \in W_0 = V_1 \cap V_0^\perp$  is an orthonormal wavelet if and only if  $\psi$  can be expressed as follows:*

$$(37.21) \quad \mathcal{F}\psi(2\xi) = e^{i\xi} \nu(2\xi) \overline{m_0(\xi + \pi)} \mathcal{F}\varphi(\xi),$$

where  $\nu \in L^2(\mathbb{T})$  with  $|\nu(\xi)| = 1$  for almost everywhere.

*Proof.* Sufficiency Assume that  $\psi$  can be expressed as above.

Let  $g \in W_0$ . Then there exists  $s \in L^2(\mathbb{T})$  such that  $\mathcal{F}g(\xi) = e^{i\xi} s(2\xi) \overline{m_0(\xi + \pi)} \mathcal{F}\varphi(\xi)$ . Using the expression of  $\mathcal{F}\psi$ , we obtain  $\mathcal{F}g(\xi) = s(\xi) \overline{\nu(\xi)} \mathcal{F}\psi(\xi)$ . Since  $s \cdot \overline{\nu} \in L^2(\mathbb{T})$ , we can expand it into the Fourier series:

$$(37.22) \quad s(\xi) \overline{\nu(\xi)} = \sum_{k \in \mathbb{Z}} c_k e^{-ik\xi},$$

where  $\{c_k\}_{k \in \mathbb{Z}} \in \ell^2$ . Inserting this expression and taking the inverse Fourier transform, we obtain

$$(37.23) \quad g = \sum_{k \in \mathbb{Z}} c_k \psi(* - k).$$

This shows that  $\{\psi(* - k)\}_{k \in \mathbb{Z}}$  spans  $W_0$ . To prove that this is an orthonormal system, we have only to show  $\sum_{k \in \mathbb{Z}} |\mathcal{F}\psi(\xi + 2k\pi)|^2 = \frac{1}{2\pi}$ .

Since  $\{\varphi(* - k)\}_{k \in \mathbb{Z}}$  forms an orthonormal system, we have

$$(37.24) \quad \sum_{l \in \mathbb{Z}} |\mathcal{F}\varphi(\xi + 2l\pi)|^2 = \frac{1}{2\pi}.$$

Keeping this in mind, we calculate

$$\sum_{k \in \mathbb{Z}} |\mathcal{F}\psi(\xi + 2k\pi)|^2 = \sum_{k \in \mathbb{Z}} \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + k\pi\right) \right|^2 \cdot \left| m_0\left(\frac{\xi}{2} + k\pi + \pi\right) \right|^2.$$

Since by periodicity

$$\begin{aligned} \sum_{k \in \mathbb{Z}} \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + 2k\pi\right) \right|^2 \cdot \left| m_0\left(\frac{\xi}{2} + 2k\pi + \pi\right) \right|^2 &= \frac{1}{2\pi} \left| m_0\left(\frac{\xi}{2}\right) \right|^2 \\ \sum_{k \in \mathbb{Z}} \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + 2k\pi + \pi\right) \right|^2 \cdot \left| m_0\left(\frac{\xi}{2} + 2k\pi\right) \right|^2 &= \frac{1}{2\pi} \left| m_0\left(\frac{\xi}{2} + \pi\right) \right|^2, \end{aligned}$$

we have the desired result.

Necessity Now we turn to the necessity. Assume  $\psi$  is an orthonormal wavelet and that  $\psi \in W_0$ . Then, there exists  $\nu \in L^2(\mathbb{T})$  such that

$$(37.25) \quad \mathcal{F}\psi(\xi) = e^{i\frac{\xi}{2}} \nu(\xi) \overline{m_0(\xi + \pi)} \mathcal{F}\varphi\left(\frac{\xi}{2}\right).$$

Since  $\psi$  is an orthonormal wavelet, we have

$$1 = \sum_{k \in \mathbb{Z}} |\mathcal{F}\psi(\xi + 2k\pi)|^2 = \sum_{k \in \mathbb{Z}} |\nu(\xi)|^2 \cdot \left| m_0\left(\frac{\xi}{2} + k\pi + \pi\right) \right|^2 \cdot \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + k\pi\right) \right|^2.$$

It remains to separate  $k$  according to its parity.

$$\begin{aligned} 1 &= |\nu(\xi)|^2 \sum_{k \in \mathbb{Z}} \left\{ \left| m_0\left(\frac{\xi}{2} + \pi\right) \right|^2 \cdot \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + 2k\pi\right) \right|^2 + \left| m_0\left(\frac{\xi}{2}\right) \right|^2 \cdot \left| \mathcal{F}\varphi\left(\frac{\xi}{2} + 2k\pi + \pi\right) \right|^2 \right\} \\ &= |\nu(\xi)|^2 \left( \left| m_0\left(\frac{\xi}{2} + \pi\right) \right|^2 + \left| m_0\left(\frac{\xi}{2}\right) \right|^2 \right) \\ &= |\nu(\xi)|^2. \end{aligned}$$

Thus, the converse is proved as well. □

### 37.3. Examples when $d = 1$ .

In this section we shall present some examples for  $d = 1$ . To do this it is convenient to give a criterion for MRA.

**Theorem 37.9.** *Assume that we are given a function  $\Phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  with the following properties:*

(A) *For all complex-valued sequences  $a = \{a_l\}_{l \in \mathbb{Z}} \in \ell^2$ ,*

$$\left\| \sum_{l \in \mathbb{Z}} a_l \Phi(* - l) \right\|_{L^2(\mathbb{R})} = \left( \sum_{l \in \mathbb{Z}} |a_l|^2 \right)^{\frac{1}{2}}.$$



(B) There exists  $a = \{a_l\}_{l \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$  such that  $\Phi\left(\frac{x}{2}\right) = \sum_{l \in \mathbb{Z}} a_l \Phi(x-l)$ , where the convergence takes place in  $L^2(\mathbb{R})$ .

Set  $V_j := \text{Span}\{\Phi(2^j * -k)\}_{k \in \mathbb{Z}}$  for each  $j \in \mathbb{Z}$ . Then  $\{V_j\}_{j \in \mathbb{Z}}$  forms an MRA.

*Proof.* Observe first that (A) implies  $\{\Phi(* - k)\}_{k \in \mathbb{Z}}$  is an orthonormal system. Indeed, by polarization, we have

$$\int_{\mathbb{R}} \left( \sum_{l \in \mathbb{Z}} a_l \Phi(x-l) \right) \overline{\left( \sum_{l \in \mathbb{Z}} b_l \Phi(x-l) \right)} dx = \sum_{l \in \mathbb{Z}} a_l \bar{b}_l$$

for all  $a = \{a_l\}_{l \in \mathbb{Z}}, b = \{b_l\}_{l \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ . Thus,

$$\int_{\mathbb{R}} \Phi(x-k) \overline{\Phi(x-m)} dx = \delta_{k,m} \quad (k, m \in \mathbb{Z})$$

by letting  $a_l = \delta_{l,k}$  and  $b_l = \delta_{l,m}$  for each  $l \in \mathbb{Z}$ . Once we prove that

$$(37.26) \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\}$$

and that

$$(37.27) \quad \bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R}),$$

then we see immediately that  $\{V_j\}_{j \in \mathbb{Z}}$  is the desired MRA with scaling function  $\Phi$ . Indeed, the scaling property is from the definition of  $V_j$  itself and nested property follows from (B). Finally,  $V_0$  is certainly spanned by a single function from its definition. So let us prove that (37.26) and (37.27)

Proof of (37.26) Setting  $P_j = \text{proj}(L^2(\mathbb{R}) \rightarrow V_j)$ , the projection from  $L^2(\mathbb{R})$  to  $V_j$ , we have to show  $\lim_{j \rightarrow \infty} P_j f = 0$  for all  $f \in L^2(\mathbb{R})$ . However, the operator norm of these projections is 1 and  $C_c^\infty(\mathbb{R})$  is dense in  $L^2(\mathbb{R})$ . Therefore, we can assume  $f \in C_c^\infty(\mathbb{R})$ . Suppose that  $\text{supp}(f) \subset [-R, R]$ .

As we did in Definition 37.1, we define  $\Phi_{j,k}(x) := 2^{\frac{j}{2}} \Phi(2^j x - k)$  for  $j, k \in \mathbb{Z}$ . Then we have, taking into account that  $P_j : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$  is a self adjoint operator satisfying  $P_j \Phi_{j,k} = \Phi_{j,k}$  and that  $\{\Phi_{j,k}\}_{k \in \mathbb{Z}}$  is an orthonormal system,

$$\begin{aligned} \|P_j f\|_{L^2(\mathbb{R})}^2 &= \sum_{k \in \mathbb{Z}} |\langle f, \Phi_{j,k} \rangle_{L^2(\mathbb{R})}|^2 \\ &= \sum_{k \in \mathbb{Z}} \left| \int_{-R}^R f(x) \Phi_{j,k}(x) dx \right|^2 \\ &\leq \|f\|_{L^2(\mathbb{R})}^2 \cdot \sum_{k \in \mathbb{Z}} \int_{-R}^R |\Phi_{j,k}(x)|^2 dx. \end{aligned}$$

Keeping in mind that  $R$  is a fixed number that depends on the fixed function  $f$ , we observe

$$(37.28) \quad \lim_{j \rightarrow \infty} \sum_{k \in \mathbb{Z}} \int_{-R}^R |\Phi_{j,k}(x)|^2 dx = \lim_{j \rightarrow \infty} \sum_{k \in \mathbb{Z}} \int_{k-2^j R}^{k+2^j R} |\Phi(x)|^2 dx = 0,$$

where, for the last equality, we used the Lebesgue convergence theorem. Therefore we obtain (37.26).

**Proof of (37.27)** Suppose that  $f \in L^2(\mathbb{R})$  is perpendicular to each  $V_j$ . That is  $P_j f = 0$  for all  $j \in \mathbb{Z}$ . Set  $f_N = \mathcal{F}^{-1}(\chi_{[-2^N\pi+1, 2^N\pi-1]}\mathcal{F}f)$  for  $N \in \mathbb{N}$ . Let  $\varepsilon > 0$  be given.  $\|f - f_N\|_2 < \varepsilon$  provided  $N$  is large enough. Observe that

$$(37.29) \quad \|P_j[f_N]\|_2 = \|P_j[f - f_N] - P_j f\|_2 = \|P_j[f - f_N]\|_2 \leq \|f - f_N\|_2 \leq \varepsilon.$$

Note that  $P_j[f_N] = \text{proj}(L^2(\mathbb{R}) \rightarrow V_j)[f_N] \in V_j$  and that  $\{\Phi_{j,k}\}_{k=1}^\infty$  is an orthonormal system from (A) and (B). Hence

$$\|P_j[f_N]\|_2^2 = \sum_{k \in \mathbb{Z}} |\langle P_j f_N, \Phi_{j,k} \rangle_{L^2(\mathbb{R})}|^2.$$

Keeping to the same notation above, we calculate

$$\|P_j[f_N]\|_2^2 = \sum_{k \in \mathbb{Z}} |\langle P_j f_N, \Phi_{j,k} \rangle_{L^2(\mathbb{R})}|^2 = \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} f_N(\xi) \overline{\Phi_{j,k}(\xi)} d\xi \right|^2$$

from another side. We use the Plancherel theorem.

$$\|P_j[f_N]\|_{L^2(\mathbb{R})}^2 = \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \mathcal{F}f_N(\xi) \overline{\mathcal{F}\Phi_{j,k}(\xi)} d\xi \right|^2 = \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \mathcal{F}f_N(\xi) 2^{-\frac{j}{2}} e^{ik2^{-j}\xi} \overline{\mathcal{F}\Phi(2^{-j}\xi)} d\xi \right|^2.$$

We interpret the integral  $\int_{\mathbb{R}} \mathcal{F}f_N(\xi) 2^{-\frac{j}{2}} e^{ik2^{-j}\xi} \overline{\mathcal{F}\Phi(2^{-j}\xi)} d\xi$  as the  $(-k)$ -th Fourier coefficient of  $\sqrt{2\pi} \mathcal{F}f_N \cdot \overline{\mathcal{F}\Phi(2^{-j}\cdot)} \in L^2([-2^j\pi, 2^j\pi])$ . By the Plancherel formula we obtain

$$(37.30) \quad \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \mathcal{F}f_N(\xi) 2^{-\frac{j}{2}} e^{ik2^{-j}\xi} \overline{\mathcal{F}\Phi(2^{-j}\xi)} d\xi \right|^2 = 2\pi \int_{\mathbb{R}} |\mathcal{F}f_N(\xi) \mathcal{F}\Phi(2^{-j}\xi)|^2 d\xi.$$

Since  $\mathcal{F}\Phi$  is continuous and bounded, we conclude

$$(37.31) \quad \varepsilon^2 \gtrsim |\mathcal{F}\varphi(0)| \cdot \|f_N\|_{L^2(\mathbb{R})}^2$$

provided  $N$  is large enough. Letting  $N \rightarrow \infty$ , we conclude

$$(37.32) \quad \varepsilon^2 \gtrsim |\mathcal{F}\varphi(0)| \cdot \|f\|_{L^2(\mathbb{R})}^2.$$

Since  $\varepsilon > 0$  is arbitrary, we conclude  $f = 0$ .  $\square$

*Exercise 227.* Let  $H$  be a Hilbert space and  $H_0$  its subspace. Then show that  $H_0$  is dense in  $H$  if and only if there is no nonzero vector in  $H$  perpendicular to  $H_0$ .

Haar wavelet. Now start with a function  $\varphi = \chi_{[0,1]} \in L^2(\mathbb{R})$ . We intend to construct an MRA whose scaling function is  $\varphi$ . To do this we set  $\varphi_{j,k} := 2^{\frac{j}{2}}\varphi(2^j * -k)$ ,  $j, k \in \mathbb{Z}$  and  $V_j := \text{Span}(\{\varphi_{j,k}\}_{k \in \mathbb{Z}})$ ,  $j \in \mathbb{Z}$ .

To see that  $\{V_j\}_{j \in \mathbb{Z}}$  is an MRA whose scaling function is  $\varphi$ , we have to verify the following.

$$(37.33) \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}).$$

$$(37.34) \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\}.$$

$$(37.35) \quad \dots \subset V_j \subset V_{j+1} \subset \dots$$

$$(37.36) \quad \{\varphi(* - k)\}_{k \in \mathbb{Z}} \text{ forms an orthonormal basis in } V_0.$$

$$(37.37) \quad f \in V_j \iff f(2*) \in V_{j+1}.$$

Conditions (37.35)–(37.37) are immediate from the definition. What is less trivial is to check (37.33) and (37.34). Now we present a skillful usage of the dyadic maximal operator. Let

us recall the definition. Set  $E_j[f] = \sum_{k \in \mathbb{Z}} m_{[2^{-j}k, 2^{-j}(k+1)]}(f) \cdot \chi_{[2^{-j}k, 2^{-j}(k+1)]}$  for  $j \in \mathbb{Z}$  and  $M_{\text{dyadic}}f(x) = \sup_{j \in \mathbb{Z}} E_j[|f|](x)$ . To prove (37.33) it suffices to show that  $\lim_{j \rightarrow \infty} E_j[f] = f$  in  $L^2(\mathbb{R})$ . However if  $f \in C_c^\infty(\mathbb{R})$  this is trivial. Because  $E_j[f], j = 1, 2, \dots$  have a compact support in common and  $\{E_j[f]\}_{j \in \mathbb{N}}$  converges to  $f$ . To pass to the general case we note that  $\|E_j[f]\|_2 \leq \|M_{\text{dyadic}}f\|_2 \lesssim \|f\|_2$  for all  $f \in L^2(\mathbb{R})$ . Let  $f \in L^2(\mathbb{R})$  and  $g \in C_c^\infty(\mathbb{R})$ . Then we have

$$(37.38) \quad \|E_j[f] - f\|_2 \leq \|E_j[f - g] - (f - g)\|_2 + \|E_j[g] - g\|_2 \lesssim \|f - g\|_2 + \|E_j[g] - g\|_2$$

for all  $j \in \mathbb{N}$ . Letting  $j \rightarrow \infty$ , we conclude

$$(37.39) \quad \limsup_{j \rightarrow \infty} \|E_j[f] - f\|_2 \lesssim \|f - g\|_2.$$

Given  $f \in L^2(\mathbb{R})$ , we can make the right-hand side as small as we wish by selecting  $g \in C_c^\infty(\mathbb{R})$ . Therefore it follows that  $E_j[f] \rightarrow f$  as  $j \rightarrow \infty$ .

The proof of (37.34) is similar. Let  $f \in \bigcap_{j \in \mathbb{Z}} V_j$ . In this case we have  $E_j[f] = f$ . We observe  $E_j[f] \rightarrow 0$  almost everywhere in  $\mathbb{R}$ , as we proved in Chapter 7. In Chapter 7, we proved this assuming  $f \in L^1(\mathbb{R})$ . But a similar proof still works for  $f \in L^2(\mathbb{R})$ . Since  $|E_j[f]| \leq M_{\text{dyadic}}f$  and  $M_{\text{dyadic}}f$  is an  $L^2(\mathbb{R})$ -function, we are in the position of using the dominated convergence theorem to conclude that  $f = E_j[f] \rightarrow 0$  as  $j \rightarrow \infty$ . Thus it follows that  $f = 0$ .

Finally we note that  $\chi_{[0, \frac{1}{2}]} - \chi_{[\frac{1}{2}, 1]}$  is a orthonormal wavelet of this MRA. Indeed, a direct calculation shows

$$\left\| \sum_{k \in \mathbb{Z}} a_k \chi_{[k2^{-j}, (k+1)2^{-j}]} \right\|_{L^2(\mathbb{R})} = \left( \sum_{k \in \mathbb{Z}} |a_k|^2 \right)^{1/2}.$$

Also, if we expand  $f \in L^2(\mathbb{R})$  into the series  $f = \sum_{k \in \mathbb{Z}} a_k \chi_{[k2^{-j}, (k+1)2^{-j}]}$ , then we have  $f \perp V_0$  if and only if  $a_{2k-1} + a_{2k} = 0$  for all  $k \in \mathbb{Z}$ .

Summarizing the above observations above we obtain

**Theorem 37.10.** *Let  $\varphi = \chi_{[0,1]}$ . We set*

$$(37.40) \quad \varphi_{jk}(x) = 2^{\frac{j}{2}} \varphi(2^j x - k), j, k \in \mathbb{Z}, V_j = \text{Span}(\{\varphi_{jk}\}_{k \in \mathbb{Z}}), j \in \mathbb{Z}.$$

*Then  $\{V_j\}_{j \in \mathbb{Z}}$  is an MRA whose scaling function is  $\varphi$ . Furthermore  $\psi := \chi_{[0, \frac{1}{2}]} - \chi_{[\frac{1}{2}, 1]}$  is an orthonormal wavelet of this MRA.*

**Definition 37.11** (Haar wavelet). The function  $\psi = \chi_{[0, \frac{1}{2}]} - \chi_{[\frac{1}{2}, 1]}$  is called Haar wavelet.

Meyer wavelet. We start with an even  $C^\infty(\mathbb{R})$ -real valued function  $\Theta$  satisfying the following conditions:

$$(37.41) \quad \chi_{[-\frac{2}{3}\pi, \frac{2}{3}\pi]} \leq \sqrt{2\pi}\Theta \leq \chi_{[-\frac{4}{3}\pi, \frac{4}{3}\pi]}, \Theta(\xi)^2 + \Theta(\xi - 2\pi)^2 = 1 \text{ for } 0 \leq \xi \leq 2\pi.$$

Set  $\Phi := \mathcal{F}^{-1}\Theta \in \mathcal{S}(\mathbb{R}^d)$ . Recall that in (37.5), we have considered  $m_0(\xi) = \sum_{k \in \mathbb{Z}^d} a_k e^{ik \cdot \xi}$ .

**Theorem 37.12.** *Under the notation above,  $\Phi$  is a scaling function of an MRA. The low-pass filter  $m_0$  is a  $2\pi$ -periodic function which on  $(-\pi, \pi)$  equals  $\sqrt{2\pi}\Theta(2\xi)$ .*

*Proof.* Note by the support condition and the periodicity that

$$(37.42) \quad \sum_{l \in \mathbb{Z}} \Theta(\xi + 2\pi l)^2 = \frac{1}{2\pi} \quad (\xi \in \mathbb{R}).$$

Since we are assuming that  $\Theta \in C_c^\infty(\mathbb{R})$ ,  $\Phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  with an estimate  $|\Phi(x)| \lesssim \frac{1}{1+x^2}$ . Thus, we are in the position of applying Theorem 37.9 to conclude that  $\Phi$  is a scaling function of an MRA. As for the assertion for low-pass filter, we note that

$$(37.43) \quad \Theta(2\xi) = \sqrt{2\pi}\Theta(2\xi)\Theta(\xi) = m_0(\xi)\Theta(\xi) \quad (\xi \in \mathbb{R})$$

because  $\chi_{[-\frac{2}{3}\pi, \frac{2}{3}\pi]} \leq \sqrt{2\pi}\Theta \leq \chi_{[-\frac{4}{3}\pi, \frac{4}{3}\pi]}$ . Since  $\mathcal{F}\Phi = \Theta$ , we obtain the desired conclusion.  $\square$

**Remark 37.13.** In view of Theorem 37.12, we have obtained a wavelet  $\psi \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R})$ .

To conclude this section, once again we give a definition of the Haar function more precise through Exercise 228 below.

*Exercise 228.* We define  $h := \chi_{[0,1/2)} - \chi_{[1/2,1)}$ . The function  $h$  is called the Haar function, the Haar wavelet, or the Haar (scaling) function.

- (1) Show that  $\{h_{j,k}\}_{j,k \in \mathbb{Z}}$  forms a complete orthonormal basis in  $L^2(\mathbb{R})$ .
- (2) By using the dyadic Calderón-Zygmund decomposition (see Theorem 44.3), establish that

$$(37.44) \quad \left| \left\{ x \in \mathbb{R} : \left( \sum_{j,k \in \mathbb{Z}} |\langle f, h_{j,k} \rangle_{L^2(\mathbb{R})}|^2 \chi_{2^{-j}[k,k+1)}(x) \right)^{\frac{1}{2}} > \lambda \right\} \right| \leq \frac{16}{\lambda} \|f\|_1.$$

- (3) Let  $1 < p < \infty$ . Then show that

$$(37.45) \quad \|f\|_p \sim \left\| \left( \sum_{j,k \in \mathbb{Z}} |\langle f, h_{j,k} \rangle_{L^2(\mathbb{R})}|^2 \chi_{2^{-j}[k,k+1)} \right)^{\frac{1}{2}} \right\|_p.$$

If necessary, use the technique presented in the proof of Theorem 44.3.

### 38. UNCONDITIONAL BASIS

Throughout this section we assume that  $(B, \|\cdot\|_B)$  is a Banach space. The aim of this section is to investigate properties of wavelet basis. Let us begin with some generality about the basis in Banach spaces over  $\mathbb{K}$ , where, as usual,  $\mathbb{K}$  denotes either  $\mathbb{R}$  or  $\mathbb{C}$ .

**38.1. Unconditional convergence.** The following is the most fundamental in the theory of basis in Banach spaces.

**Definition 38.1.** Let  $\{x_j\}_{j \in \mathbb{N}}$  be a sequence in  $B$  and  $x \in B$ . A series  $\sum_{j=1}^{\infty} x_j$  is said to converge to  $x$  unconditionally if, for all  $\varepsilon > 0$ , there exists a finite set  $J_0 \subset \mathbb{N}$  so that

$$\left\| x - \sum_{j \in J} x_j \right\|_B \leq \varepsilon$$

for every finite set  $J$  larger than  $J_0$ .

To be accustomed with the definition above, let us rephrase this convergence in terms of nets.

*Exercise 229.* Induce a suitable order of  $2^{\mathbb{N}}$  so that the following holds: A sequence  $\{x_j\}_{j \in \mathbb{N}}$  in  $B$  is such that  $\lim_{A \in 2^{\mathbb{N}}} \sum_{j \in A} x_j$  converges if and only if  $\sum_{j=1}^{\infty} y_j$  converges unconditionally.

Definition 38.1 can be rephrased as follows:

**Lemma 38.2.** Let  $\{x_j\}_{j \in \mathbb{N}}$  be a sequence in  $B$  and  $x \in B$ . A series  $\sum_{j=1}^{\infty} x_j$  in  $B$  converges to  $x$  unconditionally, if and only if  $\sum_{j=1}^{\infty} x_{\sigma(j)}$  converges to  $x$  for all bijections  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ .

*Proof.* Assume that  $\sum_{j=1}^{\infty} x_j$  converges unconditionally to  $x$ .

Choose a bijection  $\sigma$  on  $\mathbb{N}$ . Let  $\varepsilon > 0$ . Since  $\sum_{j=1}^{\infty} x_j$  converges unconditionally, we can take a subset  $\mathcal{N}_0 \in 2^{\mathbb{N}}$  so that  $\left\| x - \sum_{j \in \mathcal{N}} x_j \right\|_B \leq \varepsilon$ , whenever  $\mathcal{N}$  is a finite set containing  $\mathcal{N}_0$ . Set  $N_0 := \max\{n \in \mathbb{N} : \sigma(n) \in \mathcal{N}_0\}$ . Then if  $J \geq N_0$ , then we have  $\mathcal{N}_0 \subset \{1, 2, \dots, J\}$ . Hence, it follows that  $\left\| x - \sum_{j=1}^J x_j \right\|_B \leq \varepsilon$  for all  $J \geq N_0$ .

Suppose instead that  $\sum_{j=1}^{\infty} x_j$  does not converge unconditionally to  $x$ .

Then there exists  $\varepsilon > 0$  with the following property: For any subset  $\mathcal{N}$  we can find a subset  $\mathcal{N}'$  larger than  $\mathcal{N}$  so that  $\left\| x - \sum_{j \in \mathcal{N}'} x_j \right\|_B \geq \varepsilon$ .

For this  $\varepsilon > 0$ , we shall construct an increasing sequence of subsets  $\{\mathcal{N}_j\}_{j \in \mathbb{N}}$  so that, for every  $j \in \mathbb{N}$ ,  $\left\| x - \sum_{k \in \mathcal{N}_j} x_k \right\|_B \geq \varepsilon$  and  $\{1, 2, \dots, j\} \subset \mathcal{N}_j$ . To do this, first of all we pick  $\mathcal{N}_1$  so that  $\left\| x - \sum_{j \in \mathcal{N}_1} x_j \right\|_B \geq \varepsilon$  and  $1 \in \mathcal{N}_1$ . Suppose that we have constructed  $\mathcal{N}_1, \dots, \mathcal{N}_k$ . Then we choose  $\mathcal{N}_{k+1}$  so that it engulfs  $\mathcal{N}_k \cup \{k+1\}$  and that  $\left\| x - \sum_{j \in \mathcal{N}_{k+1}} x_j \right\|_B \geq \varepsilon$ . Then  $\{\mathcal{N}_j\}_{j \in \mathbb{N}}$  obtained in this way clearly satisfies the desired property.

Take a bijection  $\sigma$  so that  $k_1 \geq k_2$  whenever  $j_1 \geq j_2$ ,  $\sigma(j_1) \in \mathcal{N}_{k_1}$  and  $\sigma(j_2) \in \mathcal{N}_{k_2}$ . Then it is easy to see  $\sum_{j=1}^{\infty} x_{\sigma(j)}$  does not converges to  $x$ .  $\square$

The next lemma is a crucial property that can be deduced from the unconditional convergence.

**Lemma 38.3.** *If the series  $\sum_{j=1}^{\infty} y_j$  converges unconditionally, then we have*

$$(38.1) \quad \lim_{N \rightarrow \infty} \sup \left\{ \sum_{j=N}^{\infty} |f(y_j)| : f \in B^*, \|f\|_{B^*} \leq 1 \right\} = 0.$$

*Proof.* Choose  $\varepsilon > 0$  arbitrarily. Then there exists  $\mathcal{N}_0$  so that  $\left\| \sum_{j \in \mathcal{N}} y_j - y \right\|_B < \varepsilon$  for any finite subset  $\mathcal{N}$  larger than  $\mathcal{N}_0$ . By replacing  $\mathcal{N}_0$  with  $[1, \#\mathcal{N}_0] \cap \mathbb{N}$ , if necessary, we can assume  $\mathcal{N}_0 = \{1, 2, \dots, N_0\}$  for some  $N_0 \in \mathbb{N}$ . Suppose that we are given  $f \in B^*$  with  $\|f\|_* \leq 1$ . Denote

$$(38.2) \quad \sum_1 := \sum_{\substack{j: j \geq N_0+1 \\ \operatorname{Re}(f(y_j)) \geq 0}}, \sum_2 := \sum_{\substack{j: j \geq N_0+1 \\ \operatorname{Re}(f(y_j)) < 0}}, \sum_3 := \sum_{\substack{j: j \geq N_0+1 \\ \operatorname{Im}(f(y_j)) \geq 0}}, \sum_4 := \sum_{\substack{j: j \geq N_0+1 \\ \operatorname{Im}(f(y_j)) < 0}}$$

for the sake of simplicity.

Then we have

$$\begin{aligned} \sum_{j=N_0+1}^{\infty} |f(y_j)| &\leq \operatorname{Re} \left[ f \left( \sum_1 y_j \right) \right] - \operatorname{Re} \left[ f \left( \sum_2 y_j \right) \right] + \operatorname{Im} \left[ f \left( \sum_3 y_j \right) \right] - \operatorname{Im} \left[ f \left( \sum_4 y_j \right) \right] \\ &\leq \sum_{l=1}^4 \left\| \sum_l y_j \right\|_B \\ &\leq 4\varepsilon. \end{aligned}$$

This is the desired result.  $\square$

With the help of Lemma 38.3, we obtain an important characterization of the unconditional convergence.

**Theorem 38.4.** *Let  $\{x_j\}_{j \in \mathbb{N}}$  be a sequence in  $B$ . The following are equivalent.*

- (1)  $\sum_{j=1}^{\infty} x_j$  converges unconditionally.
- (2) Whenever  $\{\beta_j\}_{j \in \mathbb{N}}$  is a sequence in a unit ball in  $\mathbb{K}$ , the series  $\sum_{j=1}^{\infty} \beta_j x_j$  converges.
- (3) Let  $\{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N}$  be a strictly increasing sequence. Then  $\sum_{j=1}^{\infty} x_{n_j}$  converges.
- (4) If  $\{\varepsilon_j\}_{j \in \mathbb{N}}$  is a sequence consisting of  $\{\pm 1\}$ , then  $\sum_{j=1}^{\infty} \beta_j x_j$  converges.

*Proof.*  $\boxed{(1) \implies (2)}$  Set  $S_\beta^J(x) := \sum_{j=1}^J \beta_j \alpha_j(x) x_j$ . Then we have, for  $J_2 > J_1 > J_0$ ,

$$\begin{aligned} \left\| S_\beta^{J_1}(x) - S_\beta^{J_2}(x) \right\|_B &= \sup_{\substack{f \in B^* \\ \|f\|_{B^*} \leq 1}} |f(S_\beta^{J_1}(x) - S_\beta^{J_2}(x))| \\ &\leq \sum_{j=J_1+1}^{J_2} |\beta_j| \cdot |\alpha_j(x)| \cdot |f(x_j)| \\ &\leq \sum_{j=J_0}^{\infty} |\alpha_j(x)| \cdot |f(x_j)|. \end{aligned}$$

By virtue of Lemma 38.3 it follows that  $\lim_{N \rightarrow \infty} \sum_{j=N}^{\infty} |\alpha_j(x)| \cdot |f(x_j)| = 0$ , from which we conclude  $\{S_\beta^J(x)\}_{J \in \mathbb{N}}$  is a Cauchy sequence and hence it is convergent.

$\boxed{(2) \implies (3) \iff (4)}$  It is clear because  $\sum_{j=1}^{\infty} x_j$  converges.

$\boxed{(3) \implies (1)}$  Suppose that  $\sum_{j=1}^{\infty} x_j$  is convergent but the convergence is not unconditional.

Then we can take  $\varepsilon > 0$  with the following property. For all  $j \in \mathbb{N}$  there exists  $\mathcal{N}_j$  that contains  $1, 2, \dots, j$  and satisfies

$$(38.3) \quad \left\| \left( \sum_{k \in \mathcal{N}_j} x_k \right) - x \right\|_B \geq \varepsilon.$$

Since we are assuming that  $\sum_{j=1}^{\infty} x_j$  is convergent to  $x$ , we can take  $J_0 \in \mathbb{N}$  so that

$$(38.4) \quad \left\| \left( \sum_{k=1}^j x_k \right) - x \right\|_B < \frac{1}{2} \varepsilon$$

for all  $j \geq J_0$ .

We define an increasing sequence  $\{J_l\}_{l=1}^{\infty}$  by the following recurrence formula:

$$(38.5) \quad J_l := \max \mathcal{N}_{J_{l-1}}, \quad l = 1, 2, \dots$$

Then from the construction of the  $\mathcal{N}_j$ , we have

$$(38.6) \quad \left\| \sum_{k \in \mathcal{N}_{J_l} \setminus \{1, 2, \dots, J_l\}} x_k \right\|_B \geq \left\| \left( \sum_{k \in \mathcal{N}_{J_l}} x_k \right) - x \right\|_B - \left\| \left( \sum_{k=1}^{J_l} x_k \right) - x \right\|_B > \frac{1}{2} \varepsilon.$$

We define  $\mathcal{N} := \bigcup_{l=1}^{\infty} (\mathcal{N}_{J_l} \setminus \{1, 2, \dots, J_l\})$ . Arrange the number in  $\mathcal{N}$  in numerical order to

obtain  $j_1 < j_2 < \dots < j_m < \dots$ . Then from (38.6) we conclude  $\sum_{m=1}^{\infty} x_{j_m}$  diverges.  $\square$

**38.2. Bases for Banach spaces.** Now we define basis in Banach spaces and investigate their properties.

**Definition 38.5.** A countable subset  $\{x_j\}_{j \in \mathbb{N}}$  of  $B$  is said to be a basis of  $B$ , if for every  $x \in B$

there exists a unique sequence  $\{\alpha_j\}_{j \in \mathbb{N}}$  such that  $\lim_{J \rightarrow \infty} \left\| x - \sum_{j=1}^J \alpha_j x_j \right\|_B = 0$ .

Assume that  $\{x_j\}_{j \in \mathbb{N}}$  is a basis of  $B$ . We remark that  $x \mapsto \alpha_j$  is a linear mapping for each  $j \in \mathbb{N}$ , where  $\{\alpha_j\}_{j \in \mathbb{N}}$  is a sequence corresponding to  $x$ . Let us denote by  $\alpha_j$  the above mapping. This mapping is called coefficient mapping.

We obtain an equivalent norm in the lemma below:

**Lemma 38.6.** Let  $B$  be a Banach space and  $\{x_j\}_{j \in \mathbb{N}}$  its basis. Denote by  $\{\alpha_j\}$  the coefficient mapping. Then  $\| \|x\|_B := \sup_{n \in \mathbb{N}} \left\| \sum_{j=1}^n \alpha_j(x) x_j \right\|_B$  is a norm equivalent to the original norm of  $B$ .

*Proof.* Thanks to the continuity of the original norm of  $B$  we obtain

$$(38.7) \quad \|x\|_B \leq \| \|x\|_B$$

and that  $\| \|x\|_B < \infty$  for all  $x \in B$ . Let  $(B, \|\cdot\|_B)$  denote the original Banach space and  $(B, \| \cdot \|_B)$  the normed space equipped with a norm  $\| \cdot \|_B$ . Then the inclusion mapping

$$(38.8) \quad \iota : (B, \| \cdot \|_B) \rightarrow (B, \|\cdot\|_B)$$

is a continuous bijection. Therefore, once we prove  $(B, \| \cdot \|_B)$  is complete, then it follows by the closed graph theorem that  $\iota$  is an isomorphism.

To establish  $(B, \| \cdot \|_B)$  is complete, we pick a Cauchy sequence  $\{x_j\}_{j \in \mathbb{N}}$ . Therefore once we prove  $(B, \| \cdot \|_B)$  is complete, then the open mapping theorem shows that two norms are equivalent.

To do this we take a Cauchy sequence  $\{x_j\}_{j \in \mathbb{N}}$  in  $(B, \| \cdot \|_B)$ . Then we have

$$(38.9) \quad \left\| \sum_{k=1}^K \alpha_k(x_{j_1}) x_k - \sum_{k=1}^K \alpha_k(x_{j_2}) x_k \right\|_B \leq \|x_{j_1} - x_{j_2}\|_B, \quad K \in \mathbb{N}$$

by definition of the norm. Therefore

$$(38.10) \quad \|(\alpha_K(x_{j_1}) - \alpha_K(x_{j_2})) x_K\|_B \leq 2 \|x_{j_1} - x_{j_2}\|_B$$

for all  $K \in \mathbb{N}$ , which implies  $\{\alpha_K(x_j)\}_{j \in \mathbb{N}}$  is a Cauchy sequence for all  $K \in \mathbb{N}$ . Set

$$(38.11) \quad \alpha_K = \lim_{j \rightarrow \infty} \alpha_K(x_j).$$

Fix  $j$  for the time being. Then we obtain

$$(38.12) \quad \left\| \sum_{k=1}^K \alpha_k x_k - \sum_{k=1}^K \alpha_k(x_j) x_k \right\|_B \leq \sup_{l \geq j} \|x_l - x_j\|_B$$

for all  $K$ . This implies  $\left\| \sum_{k=K+1}^M \alpha_k x_k \right\|_B \leq 2 \sup_{l \geq j} \|x_l - x_j\|_B + \left\| \sum_{k=K+1}^M \alpha_k(x_j) x_k \right\|_B$ . Since  $K, M$  are still at our disposal, we have

$$(38.13) \quad \sup_{K, M \geq L} \left\| \sum_{k=K+1}^M \alpha_k x_k \right\|_B \leq 2 \sup_{l \geq j} \|x_l - x_j\|_B + \sup_{K, M \geq L} \left\| \sum_{k=K+1}^M \alpha_k(x_j) x_k \right\|_B.$$



Letting  $L \rightarrow \infty$ , we have  $\limsup_{K, M \rightarrow \infty} \left\| \sum_{k=K+1}^M \alpha_k x_k \right\|_B \leq 2 \sup_{l \geq j} \|x_l - x_j\|_B$ . Note that  $j$  is a number fixed arbitrary. Therefore if we let  $j \rightarrow \infty$ , we obtain  $\limsup_{K, M \rightarrow \infty} \left\| \sum_{k=K+1}^M \alpha_k x_k \right\|_B = 0$ . Since  $(B, \|\cdot\|_B)$  is complete, we conclude  $x = \sum_{k=1}^{\infty} \alpha_k x_k$  exists. By virtue of the uniqueness of the coefficient  $\alpha_k$ , we have  $\alpha_k = \alpha_k(x)$  for all  $k$ .

Returning to (38.12), we insert  $\alpha_k = \alpha_k(x)$ . Then we finally have

$$(38.14) \quad \|x - x_j\|_B = \sup_{K \in \mathbb{N}} \left\| \sum_{k=1}^K \alpha_k(x) x_k - \sum_{k=1}^K \alpha_k(x_j) x_k \right\|_B \leq \sup_{l \geq j} \|x_l - x_j\|_B.$$

Finally letting  $j \rightarrow \infty$ , we obtain  $\lim_{j \rightarrow \infty} \|x - x_j\|_B = 0$ . □

**Theorem 38.7.** *Let  $B$  be a Banach space and  $\{x_j\}_{j \in \mathbb{N}}$  its basis. Denote by  $\{\alpha_j\}_{j \in \mathbb{N}}$  the coefficient mapping. Then each  $\alpha_j$  is continuous.*

*Proof.* By Lemma 38.6, we have

$$(38.15) \quad |\alpha_j(x)| = \left\| \frac{\alpha_j(x)}{\|x_j\|_B} x_j \right\|_B \leq \frac{2\|x\|_B}{\|x_j\|_B} \lesssim \frac{\|x\|_B}{\|x_j\|_B}.$$

Therefore,  $\alpha_j$  is continuous for all  $j$ . □

**38.3. Unconditional bases and applications to wavelets.** One of the advantages of wavelet bases is that they can often be unconditional. We actually show this aspect in Section 39. The precise definition of unconditional basis is as follows:

**Definition 38.8.** Let  $B$  be a Banach space. A basis  $\{x_j\}_{j \in \mathbb{N}}$  is said to be unconditional if the expression  $\sum_j \alpha_j(x) x_j$ , with the coefficient mapping  $\{\alpha_j\}_{j \in \mathbb{N}}$ , converges unconditionally for all  $x \in B$ .

Suppose that  $B$  is a Banach space and  $\{x_j\}_{j \in \mathbb{N}}$  is a basis with a family of coefficient functionals  $\{\alpha_j\}_{j \in \mathbb{N}}$ . Then we write

$$(38.16) \quad S_\beta f(x) := \sum_{j=1}^{\infty} \beta_j \alpha_j(x) x_j$$

for  $\beta = \{\beta_j\}_{j \in \mathbb{N}} \in \ell^\infty(\mathbb{N})$ .

**Theorem 38.9.** *Let  $B$  be a Banach space and  $\{x_j\}_{j \in \mathbb{N}}$  its basis. The following are equivalent.*

- (1)  $\{x_j\}_{j \in \mathbb{N}}$  is unconditional.
- (2)  $\{x_{\sigma(j)}\}_{j \in \mathbb{N}}$  is a basis for any bijection  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ .
- (3) Denote by  $\{\alpha_j\}_{j \in \mathbb{N}}$  the coefficient mapping. Then  $S_\beta(x)$  converges for all  $x \in X$ , whenever  $\beta = \{\beta_j\}_{j \in \mathbb{N}}$  is a sequence in the unit ball.

*Proof.* All of these conditions are equivalent to the fact that  $\sum_{j=1}^{\infty} \alpha_j(x) x_j$  converges unconditionally. Therefore, this is trivial. □

**Lemma 38.10.** *Suppose that  $\{x_j\}_{j \in \mathbb{N}}$  is an unconditional basis on a Banach space  $B$ . Then*

$$(38.17) \quad \|S_\beta(x)\|_B \lesssim \|x\|_B$$

for all  $x$  and a sequence  $\beta = \{\beta_j\}_{j \in \mathbb{N}}$  in the unit ball  $(\ell^\infty)_1$  in  $\ell^\infty(\mathbb{N})$ .

*Proof.* Before proving this lemma, we make a brief remark about how to use unconditional convergence of  $\sum_{j=1}^{\infty} x_j$ . The assumption enables us to define  $S_\beta(x)$  and to assume

$$(38.18) \quad \sup_{\beta \in (\ell^\infty)_1} \|\alpha_j(x)x_j\|_B < \infty.$$

First we observe that  $S_\beta$  is a bounded operator. To see this, we have only to show that  $S_\beta$  is closed by the closed graph theorem.

Let  $\{x_j\}_{j \in \mathbb{N}}$  be a sequence such that  $\{x_j\}_{j \in \mathbb{N}}$  and  $\{S_\beta(x_j)\}_{j \in \mathbb{N}}$  converge to  $x$  and  $y$  respectively. Since  $\alpha_k$  is continuous for all  $k \in \mathbb{N}$ , we have

$$(38.19) \quad \alpha_k(y) = \lim_{j \rightarrow \infty} \alpha_k(S_\beta(x_j)) = \lim_{j \rightarrow \infty} \beta_k \alpha_k(x_j) = \beta_k \alpha_k(x),$$

which implies  $S_\beta(x) = y$ . Therefore  $S_\beta$  is closed and hence  $S_\beta$  is bounded.

The uniformly bounded principal reduces the matter to showing

$$(38.20) \quad \sup_{\beta \in (\ell^\infty)_1} \|S_\beta(x)\|_B < \infty.$$

However, this is easy to prove once we clarify what we are assuming and what to establish. Indeed,

$$(38.21) \quad \sup_{\beta \in (\ell^\infty)_1} \|S_\beta(x)\|_B = \sup_{\beta \in (\ell^\infty)_1} \left( \sup_{f \in B^*} |f(S_\beta(x))| \right) \leq \sup_{f \in B^*} \sum_{j=1}^{\infty} \|\alpha_j(x)x_j\|_B < \infty.$$

This is the desired result.  $\square$

**Lemma 38.11.** *Let  $\{x_j\}_{j \in \mathbb{N}}$  be a basis on a Banach space  $B$ . Assume that there exists a constant  $C > 0$  so that*

$$(38.22) \quad \|S_\beta(x)\|_B \leq C \|x\|_B$$

for all  $x \in B$ . Then  $\{x_j\}_{j \in \mathbb{N}}$  is an unconditional basis.

*Proof.* It suffices to show that the series defining  $S_\beta(x)$  converges whenever  $\beta \in (\ell^\infty(\mathbb{N}))_1$ , the unit ball of  $\ell^\infty(\mathbb{N})$ . To do this, we cut off the sum and define  $S_\beta^J(x) := \sum_{j=1}^J \beta_j \alpha_j(x)x_j$ . Let

$J_2 > J_1 \geq J_0$ . Then we set  $\beta_j^{J_1, J_2} := \beta_j$  for  $J_1 < j \leq J_2$  and  $\beta_j^{J_1, J_2} := 0$  otherwise. Then we have

$$(38.23) \quad S_\beta^{J_2}(x) - S_\beta^{J_1}(x) = S_\beta^{J_1, J_2} \left( \sum_{j=J_1+1}^{J_2} \alpha_j(x)x_j \right)$$

and  $\beta^{J_1, J_2}$  has finitely many non-zero entry. Therefore

$$(38.24) \quad \sup_{J_2 > J_1 \geq J_0} \|S_\beta^{J_2}(x) - S_\beta^{J_1}(x)\|_B \lesssim \sup_{J_2 > J_1 \geq J_0} \left\| \sum_{j=J_1+1}^{J_2} \alpha_j(x)x_j \right\|_B.$$

Letting  $J_0 \rightarrow \infty$ , we conclude that  $S_\beta^J(x)$  is a Cauchy sequence. Therefore, it follows that  $S_\beta^J(x)$  converges.  $\square$

**Theorem 38.12.** *Let  $\{x_j\}_{j \in \mathbb{N}}$  be a basis in  $B$ . Then the following are equivalent.*

- (1)  $\{x_j\}_{j \in \mathbb{N}}$  is unconditional.
- (2) For every sequence  $\beta = \{\beta_j\}_{j \in \mathbb{N}}$ ,  $\|S_\beta\|_B \lesssim 1$ .
- (3) For every sequence  $\varepsilon = \{\varepsilon_j\}_{j \in \mathbb{N}}$  consisting of  $\pm 1$ ,  $\|S_\varepsilon\|_B \lesssim 1$ .
- (4) For every sequence  $\varepsilon = \{\varepsilon_j\}_{j \in \mathbb{N}}$  consisting of 0 and finitely many 1,  $\|S_\varepsilon\|_B \lesssim 1$ .

*Proof.* This is clear from the lemmas above. □

### 39. EXISTENCE OF UNCONDITIONAL BASIS

Now we shall construct an unconditional basis of  $L^p(\mathbb{R})$  for  $1 < p < \infty$ . Suppose that  $\{V_j\}_{j \in \mathbb{Z}}$  is an MRA of  $L^2(\mathbb{R})$  with the scaling function  $\varphi$ . Set  $W_j := V_{j+1} \cap V_j^\perp$ . Hence, we have the following:

- (1) There exists an orthonormal wavelet function  $\psi \in W_0$ .
- (2) There exists a positive decreasing function  $W : [0, \infty) \rightarrow \mathbb{R}$  satisfying

$$(39.1) \quad \int_0^\infty (1+s)W(s)ds < \infty.$$

- (3) The estimate  $|\psi(x)| + |\psi'(x)| \leq W(|x|)$  holds for all  $x \in \mathbb{R}$ .

Remark that the Haar wavelet does not fall under the scope of the theory we shall develop from now on.

As for this function  $W$  we have the following quantitative information.

**Lemma 39.1.** *The estimate  $\sum_{k \in \mathbb{Z}} W(|x-k|)W(|y-k|) \lesssim W\left(\frac{|x-y|}{2}\right)$  holds for all  $x, y \in \mathbb{R}$  with  $x \neq y$ .*

*Proof.* We split the sum according as  $|x-k| \geq \frac{1}{2}|x-y|$  or not.

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} W(|x-k|)W(|y-k|) \\ & \leq \sum_{\substack{k \in \mathbb{Z} \\ |x-k| \geq \frac{1}{2}|x-y|}} W(|x-k|)W(|y-k|) + \sum_{\substack{k \in \mathbb{Z} \\ |y-k| \geq \frac{1}{2}|x-y|}} W(|x-k|)W(|y-k|). \end{aligned}$$

Since  $W$  is decreasing, we have

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} W(|x-k|)W(|y-k|) \\ & \leq W\left(\frac{|x-y|}{2}\right) \left\{ \sum_{\substack{k \in \mathbb{Z} \\ |x-k| \geq \frac{1}{2}|x-y|}} W(|y-k|) + \sum_{\substack{k \in \mathbb{Z} \\ |y-k| \geq \frac{1}{2}|x-y|}} W(|x-k|) \right\} \\ & \lesssim W\left(\frac{|x-y|}{2}\right) \left( \int_0^\infty W(s)ds + W(0) \right) \\ & \lesssim W\left(\frac{|x-y|}{2}\right). \end{aligned}$$

Thus, we obtain the desired estimate.  $\square$

Recall that we defined  $\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^jx - k)$  in Definition 37.2. Given a bounded sequence  $\beta = \{\beta_{j,k}\}_{j,k \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z} \times \mathbb{Z})$ , we define

$$(39.2) \quad T_\beta f := \sum_{j,k \in \mathbb{Z}} \beta_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}.$$

Note that this can be written as  $T_\beta f = \int_{\mathbb{R}} K_\beta(x, y) f(y) dy$ , in terms of integral kernel, where  $K_\beta(x, y)$  is defined by

$$(39.3) \quad K_\beta(x, y) := \sum_{j,k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}(x) \overline{\psi_{j,k}(y)}.$$

**Theorem 39.2.** *Let  $1 < p < \infty$  and  $\beta \in \{0, 1\}^{\mathbb{Z} \times \mathbb{Z}}$ . Then we have*

$$(39.4) \quad \|T_\beta\|_{L^p \rightarrow L^p} \lesssim_p 1.$$

*Proof.* It is trivial that  $T_\beta$  is an  $L^2(\mathbb{R})$ -bounded operator with norm less than 1. Now we estimate the kernel.

$$|K_\beta(x, y)| \leq \sum_{j,k \in \mathbb{Z}} 2^j |\psi(2^jx - k)\psi(2^jy - k)| \leq \sum_{j,k \in \mathbb{Z}} 2^j W(2^jx - k)W(2^jy - k).$$

We pass to a continuous variables and we obtain

$$(39.5) \quad |K_\beta(x, y)| \lesssim \sum_{j \in \mathbb{Z}} 2^j W(2^{j-1}|x - y|) \lesssim \int_0^\infty W(t|x - y|) dt = \frac{1}{|x - y|} \int_0^\infty W(t) dt$$

for all  $x, y \in \mathbb{R}^d$  with  $x \neq y$ .

A similar calculation works for the partial derivatives:

$$|\partial_x K_\beta(x, y)| \leq \sum_{j,k \in \mathbb{Z}} 2^{2j} |\psi'(2^jx - k)\psi(2^jy - k)| \leq C_0 \sum_{j,k \in \mathbb{Z}} 2^{2j} W(2^jx - k)W(2^jy - k),$$

where  $C_0 := \int_0^\infty (1 + s)W(s) ds$ . It is the same as before that we pass to a continuous variable.

$$(39.6) \quad |\partial_x K_\beta(x, y)| \lesssim \sum_{j \in \mathbb{Z}} 2^{2j} W(2^{j-1}|x - y|) \lesssim \int_0^\infty t W(t|x - y|) dt = \frac{1}{|x - y|^2}.$$

By symmetry we also have

$$(39.7) \quad |\partial_y K_\beta(x, y)| \lesssim \frac{1}{|x - y|^2}.$$

What counts about (39.5), (39.6) and (39.7) is that the constants in  $\lesssim$  are not dependent on  $\beta \in \{0, 1\}^{\mathbb{Z} \times \mathbb{Z}}$ . Therefore we are in the position of using the CZ-theory uniformly over  $\beta$ . Thus the operator norm does not depend on  $\beta$ .  $\square$

**Theorem 39.3.** *Keep to the same assumption as above. Then  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  is an unconditional basis of  $L^p(\mathbb{R})$  with  $1 < p < \infty$ .*

*Proof.* This theorem is also trivial for  $p = 2$ . Suppose otherwise.

We set

$$(39.8) \quad S_{m,n} f := \sum_{|j| \leq m} \sum_{|k| \leq n} \langle f, \psi_{j,k} \rangle \psi_{j,k}.$$

Then by Theorem 39.2 we see that the operator norm of  $S_{m,n}$  is majorized by a constant independent of  $m$  and  $n$ .

We are to prove

$$(39.9) \quad \lim_{m,n \rightarrow \infty} S_{m,n}f = f$$

for each  $f \in L^p(\mathbb{R})$ . To do this, the uniform boundedness of  $\{S_{m,n}\}$  allows us to assume that  $f \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$ . Now we shall make use of the interpolating inequality

$$(39.10) \quad \|g\|_p \leq \|g\|_2^\theta \cdot \|g\|_q^{1-\theta},$$

where  $q$  and  $\theta \in (0, 1)$  satisfies

$$(39.11) \quad (q-2)(q-p) < 0, \quad \frac{1}{p} = \frac{\theta}{2} + \frac{1-\theta}{q}$$

In this case we have

$$(39.12) \quad \|f - S_{m,n}f\|_p \leq \|f - S_{m,n}f\|_2^\theta \cdot \|f - S_{m,n}f\|_q^{1-\theta} \lesssim \|f - S_{m,n}f\|_2^\theta \cdot \|f\|_q^{1-\theta}.$$

Letting  $m, n \rightarrow \infty$ , we have the desired convergence.

From the observation above, we see that  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  is a basis. In particular the coefficient functional can be obtained by using the orthogonality  $\langle \psi_{j_1,k_1}, \psi_{j_2,k_2} \rangle = \delta_{j_1,j_2} \delta_{k_1,k_2}$ . Once we show that  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  is a basis, then it follows immediately from Theorem 39.3 that the basis is unconditional.  $\square$

**Theorem 39.4.** *Assume that  $1 < p < \infty$  and that  $\varphi \in C^\infty(\mathbb{R})$  is an even real-valued function satisfying*

$$(39.13) \quad \chi_{[-\frac{2}{3}\pi, \frac{2}{3}\pi]} \leq \varphi \leq \chi_{[-\frac{4}{3}\pi, \frac{4}{3}\pi]}, \quad \varphi(\xi)^2 + \varphi(\xi - 2\pi)^2 = 1 \text{ on } [0, 2\pi].$$

Define

$$(39.14) \quad \psi(x) := \mathcal{F}^{-1} \left[ \mathcal{F}\varphi \left( \frac{\xi}{2} \right) (\mathcal{F}\varphi(\xi + 2\pi) + \mathcal{F}\varphi(\xi - 2\pi)) \right] (x)$$

and  $\psi_{j,k}(x) := 2^{\frac{j}{2}} \psi(2^j x - k)$  for  $j, k \in \mathbb{Z}$ . Then  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  is an unconditional basis in  $L^p(\mathbb{R})$ .

*Proof.* Since  $\psi \in \mathcal{S}(\mathbb{R}^d)$ , the desired majorant function  $W$  does exist. Thus, we are in the position of using Theorem 39.2 and the proof is complete.  $\square$

Notes and references for Chapter 17.

Section 37. Before we make a review at each theorem, let us describe a history of wavelet. Haar found out the Haar wavelet in [225]. Exercise 228 when  $p = 2$  is a starting point. Meyer introduced the notion of wavelet [342]. Mallat introduced and investigated MRA in [321]. This paper practically contains all the contents in Section 37 such as Theorems 37.8, 37.9, 37.10, 37.12. Furthermore, the notion of  $r$ -regularity was investigated implicitly. A smooth function  $f \in C^r$  is said to be  $r$ -regular if

$$\sup_{x \in \mathbb{R}^n} (1 + |x|)^k \left| \frac{\partial^\alpha}{\partial x^\alpha} f(x) \right| < \infty$$

for all  $k = 0, 1, 2, \dots$  and  $|\alpha| \leq \max(r, 0)$ . Here  $C^0 = C$  and  $C^{-1}$  denotes the set of all measurable functions. We remark that Exercise 228 when  $1 < p < \infty, p \neq 2$  is due to Paley (see [384]).

The theory of wavelets branches according as we place ourselves in  $\mathbb{R}$  or  $\mathbb{R}^d$ . First, we consider wavelets which do not depend on MRAs. In terms of Daubechis they are pathological. For

example, we refer to Journé and Meyer. Some Korean party constructed unimodular wavelets in 1994. We refer to [20] for more details.

Ashino constructed the multi-wavelet. Other systems which are not orthonormal occur in many branches of mathematics.

Wavelets based on MRA of one variable have a long history. Around 1910 Haar considered a wavelet  $\chi_{[0,1/2)} - \chi_{[1/2,1)}$ , which is well-known as the Haar wavelet. Around 1940 and 1950 Shannon constructed band-limited wavelets whose frequency support are compact. Such functions are called band-limited. In 1983 Strömberg considered wavelets. Meyer considered functions in  $\mathcal{S}(\mathbb{R})$  which are band-limited. Daubechies considered wavelets belonging to  $C_c^\infty(\mathbb{R})$ .

Section 38. The idea of bases in normed spaces was introduced by a young Polish mathematician J. Schauder.

The notion of unconditional basis dates back to [210], which is based on the definition of basis by Karlin [265]. The definition of unconditional convergence and Theorem 38.4 are due to James [255].

Theorem 38.7

Theorem 38.9

There are many attempts of characterizing function spaces concerning with Theorem 38.12.

Section 39. We refer to the paper by Kelly-Kon-Raphael for a statement similar to Lemma 39.1. Theorems 39.2, 39.3, 39.4, which are dealing with constructions of MRA, examples of wavelet based on MRA and examples of wavelets, can be found in [24, 43, 68, 159] for 1-variables. The paper by Daubechies [159] is the first one which contains compact wavelets on  $\mathbb{R}$ . In [304] Lemarié-Rieusset and Malgouyres specified the size of the support. [24] covers the result on  $\mathbb{R}$ . The books [43, 68] contain the results for  $d$ -variables as well. These attempts are made in the weighted setting as well. In [76, 303].  $L^p(w)$  spaces with  $w \in A_p$  are characterized in terms of wavelet. The textbook of Meyer [43] is a fundamental source of the characterization of function spaces in terms of wavelet. The characterization of Hardy spaces (see Section 50) goes back to Strömberg work in 1983. The space  $L^p(w)$  characterized by Lemarié in 1994, by García-Cuerva and Kazarian in 1995, by García-Cuerva and Martell in 2001, and by Aimar, Bernardis, Martín and Reyes. For rearrangement invariant spaces such an attempt had been done in 1997 by Soardi.

The weighted Hardy space  $H^p(w)$  (see Section 50) is characterized by García-Cuerva and Martell in 2001. As for weighted Triebel-Lizorkin spaces and weighted Besov spaces, we refer to the works by Deng, Xu and Yang in 2002 and by Izuki and Sawano in 2009.

In 1986 Sharapudinov characterized the space  $L^{p(\cdot)}([0,1])$ . In 2008 Izuki and Kopaliani independently characterized the space  $L^{p(\cdot)}(\mathbb{R}^d)$ . Here, for a function  $p$  on a measure space  $X$  which assumes the value in  $[1, \infty)$ , the variable Lebesgue space  $L^{p(\cdot)}(X)$  is the set of all functions  $f$  such that

$$\int_X \left( \frac{|f(x)|}{\lambda} \right)^{p(x)} dx < \infty$$

for some  $\lambda > 0$  and the norm is given by

$$\|f\|_{L^{p(\cdot)}(X)} = \inf \left\{ \lambda > 0 : \int_X \left( \frac{|f(x)|}{\lambda} \right)^{p(x)} dx \leq 1 \right\}.$$

To write this section the author has referred to [24]. I have learnt a lot from M. Izuki [503].

We refer to [463] for the unconditional basis of the Hardy spaces.

**Part 18. Vector-valued norm inequalities****Part 19. Vector-valued norm inequalities on  $\mathbb{R}^d$** 

This part is a culmination of Chapters 7, 8, 9 and 11. As an application of the results in these sections, we shall present a very strong tool called Littlewood-Paley theory.

## 40. VECTOR-VALUED INEQUALITIES

By vector-valued inequalities we mean the inequality of the form

$$(40.1) \quad \|T_j f_j\|_{L^p(\ell^q)} \lesssim \|f_j\|_{L^p(\ell^q)},$$

where for a sequence of measurable functions  $\{f_j\}_{j \in \mathbb{N}_0}$  we set

$$(40.2) \quad \|f_j\|_{L^p(\ell^q)} := \left\| \left( \sum_{j=0}^{\infty} |f_j|^q \right)^{\frac{1}{q}} \right\|_p$$

and  $\{T_j\}_{j \in \mathbb{N}}$  is a sequence of (sub)linear operators. It can happen that  $\mathbb{N}_0$  and  $\mathbb{Z}$  take the place of  $\mathbb{N}$ .

A natural question will arise: There seems to be a routine procedure to extend known inequalities to the vector-valued version. For example, suppose that we are given a family of operators  $\{T_j\}_{j \in \mathbb{N}}$  on  $L^1(\mathbb{R})$  such that

$$(40.3) \quad \|T_j f\|_1 \leq \|f\|_1 \text{ for all } j \in \mathbb{Z} \text{ and } f \in L^1(\mathbb{R}).$$

Here for the sake of brevity to produce a counterexample, we replaced  $\mathbb{N}_0$  by  $\mathbb{Z}$ . It seems easy to have, for example,

$$(40.4) \quad \|T_j f_j\|_{L^1(\ell^\infty)} \lesssim \|f_j\|_{L^1(\ell^\infty)} \text{ for all } \{f_j\}_{j \in \mathbb{Z}} \in L^1(\ell^\infty).$$

However, this is not the case. Assume to the contrary that (40.4) is true. Let us define  $T_j f := \frac{1}{|B(2^j)|} \chi_{B(2^j)} * f$ . Then the assumption is cleared. Meanwhile, letting  $f_j = f$  for  $j \in \mathbb{Z}$ , we see

$$(40.5) \quad \|T_j f\|_{L^1(\ell^\infty)} \lesssim \|f\|_{L^1(\ell^\infty)} = \|f_j\|_{L^1(\ell^\infty)}$$

cannot hold, since  $T_j f$  is essentially the maximal operator which does not enjoy the  $L^1(\mathbb{R})$ -boundedness property. This inequality looks too heavy and it seems to be meaningless in analyzing this type of inequality. However, what is surprising about this inequality is that we can rewrite the  $L^p(\mathbb{R}^d)$ -norm with  $1 < p < \infty$ . Let us fix another notation. Pick a smooth function  $\varphi$  satisfying  $\chi_{B(4) \setminus B(2)} \leq \varphi \leq \chi_{B(8) \setminus B(1)}$ . Denote  $\varphi_j(\xi) = \varphi(2^{-j}\xi)$  for  $j \in \mathbb{Z}$ . Then we define

$$(40.6) \quad \varphi_j(D)f(x) := \mathcal{F}^{-1}(\varphi_j \cdot \mathcal{F}f)(x) := (2\pi)^{-\frac{d}{2}} \mathcal{F}^{-1} \varphi_j * f(x).$$

Then Littlewood and Paley showed the following: Suppose that  $1 < p < \infty$ . Then

$$(40.7) \quad \|\varphi_j(D)f\|_{L^p(\ell^2)} \simeq_p \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ . Its probability space version is also known as the Burkholder-Gundy-Davis inequality. They are widely used and support harmonic and stochastic analysis nowadays.

To develop a theory of vector-valued inequality we begin with extending the duality formula to the vector-valued version.



**Proposition 40.1.** *Let  $1 \leq p, q \leq \infty$ . Suppose  $\{f_j\} \subset L^p(\mathbb{R}^d)$  be measurable functions. Then we have*

$$(40.8) \quad \|f_j\|_{L^p(\ell^q)} = \sup \left| \int_{\mathbb{R}^d} \sum_j f_j(x) g_j(x) dx \right|,$$

where  $\{g_j\}$  runs over the system of functions such that

$$\|g_j\|_{L^{p'}(\ell^{q'})} = 1, g_N = g_{N+1} = g_{N+2} = \dots = 0$$

for some large  $N$ .

*Proof.* We may assume  $f_N = f_{N+1} = f_{N+2} = \dots = 0$  for some large  $N$  by virtue of the monotone convergence theorem. The proof

$$(40.9) \quad \|f_j\|_{L^p(\ell^q)} \geq \sup \left| \sum_j \int_{\mathbb{R}^d} f_j(x) g_j(x) dx \right|$$

is easy by a repeated application of Hölder's inequality.

We write out  $\|f_j\|_{L^p(\ell^q)}$  in full :

$$(40.10) \quad \|f_j\|_{L^p(\ell^q)} = \left( \int_{\mathbb{R}^d} \left( \sum_j |f_j(x)|^q dx \right)^{\frac{p}{q}} \right)^{\frac{1}{p}}.$$

By the duality  $\ell^q$ - $\ell^{q'}$  we have

$$(40.11) \quad \left( \sum_j |f_j(x)|^q \right)^{\frac{1}{q}} = \sum_j h_j(x) f_j(x), \|h_j(x)\|_{\ell^q} = 1, h_N = h_{N+1} = h_{N+2} = \dots = 0.$$

for some sequence of measurable functions  $\{h_j\}_{j \in \mathbb{N}}$ .

By the duality  $L^p(\mathbb{R}^d)$ - $L^{p'}(\mathbb{R}^d)$  we have

$$(40.12) \quad \left( \int_{\mathbb{R}^d} \left( \left| \sum_j h_j(x) \cdot f_j(x) \right|^p dx \right)^{\frac{1}{p}} = \int_{\mathbb{R}^d} \sum_j g(x) \cdot h_j(x) \cdot f_j(x) dx, \|g\|_p = 1$$

for some  $L^{p'}$  function  $g$ . Putting  $g_j(x) = g(x)h_j(x)$ , we thus obtain the attainer of sup in the formula. □

**Remark 40.2.** The proof shows that we can take  $\{g_j\}_{j \in \mathbb{N}}$  so that  $g_N = g_{N+1} = g_{N+2} = \dots = 0$ , if  $f_N = f_{N+1} = f_{N+2} = \dots = 0$  for some large  $N$ .

Before we proceed further, some helpful remark may be in order.

**Remark 40.3.** Consider the vector-valued inequality

$$(40.13) \quad \|T_j f_j\|_{L^p(\ell^q)} \leq c_0 \|f_j\|_{L^p(\ell^q)}$$

where  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of  $L^p(\mathbb{R}^d)$ -functions and  $c_0$  is independent of  $\{f_j\}_{j \in \mathbb{N}}$ .

- (1) If (13.3) holds, then each  $T_j$  is  $L^p(\mathbb{R}^d)$ -bounded with norm  $c_0$ . Conversely if each  $T_j$  is  $L^p(\mathbb{R}^d)$ -bounded with norm  $c_0$ , then (13.3) holds for  $p = q$ .

- (2) By the monotone convergence theorem (40.13) holds, once we prove (40.13) for every sequence of  $L^p(\mathbb{R}^d)$ -functions  $\{f_j\}_{j \in \mathbb{N}}$  such that  $f_N = f_{N+1} = \dots = 0$  for some large  $N$ . Of course  $c_0$  must not depend on  $N$ .

#### 41. VECTOR-VALUED INEQUALITIES FOR MAXIMAL OPERATORS

Now let us extend Theorem 12.13 to the vector-valued inequality.

**Theorem 41.1.** *Suppose that  $1 < p < \infty$  and  $1 < q \leq \infty$ . Then we have*

$$(41.1) \quad \|M f_j\|_{L^p(\ell^q)} \lesssim \|f_j\|_{L^p(\ell^q)}.$$

**Remark 41.2.** The proof will be long. By interpolation we can shorten the proof slightly. However we prefer to depend on an elementary method. Here we are going to prove with our own method. The proof will be a simplified version of [424].

*Proof.* Case 1 :  $q = \infty$ . In this case the theorem is clear because

$$(41.2) \quad \sup_{j \in \mathbb{N}_0} M f_j(x) \leq M \left[ \sup_{j \in \mathbb{N}_0} M f_j \right] (x) \text{ for all } x \in \mathbb{R}^d$$

and we can resort to the  $L^p(\mathbb{R}^d)$ -boundedness of the Hardy-Littlewood maximal operator  $M$ .

Case 2 :  $1 < p = q < \infty$

As we have referred to in Remark 40.3, the theorem is clear.

Case 3 :  $1 < q < \infty$  Put  $r = \frac{p}{q} > 1$ . By using Proposition 40.1, we can choose  $g \in L^{r'}$  with unit norm so that

$$(41.3) \quad \|M f_j\|_{L^p(\ell^q)}^q = \int_{\mathbb{R}^d} \left( \sum_{j \in \mathbb{N}_0} M f_j(x)^q \right) \cdot g(x) dx = \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} M f_j(x)^q \cdot g(x) dx$$

Theorem 21.1 gives us

$$(41.4) \quad \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} M f_j^q(x) \cdot g(x) \lesssim \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} |f_j(x)|^q \cdot M g(x)$$

By Hölder's inequality and  $L^{r'}$ -boundedness of  $M$  we obtain

$$\begin{aligned} \|M f_j\|_{L^p(\ell^q)}^q &\lesssim \left\{ \int_{\mathbb{R}^d} \left( \sum_{j \in \mathbb{N}_0} |f_j(x)|^q \right)^r \right\}^{\frac{1}{r}} \left\{ \int_{\mathbb{R}^d} M g^{r'}(x) \right\}^{\frac{1}{r'}} dx \\ &\lesssim \|f_j\|_{L^p(\ell^q)}^q \cdot \sup_{\|g\|_{r'}=1} \left( \int_{\mathbb{R}^d} |g(x)|^{r'} dx \right)^{\frac{1}{r'}} \\ &= \|f_j\|_{L^p(\ell^q)}^q. \end{aligned}$$

Raising to the  $\frac{1}{q}$ -th power both sides, we obtain

$$(41.5) \quad \|M f_j\|_{L^p(\ell^q)} \lesssim \|f_j\|_{L^p(\ell^q)}.$$

This is the desired result.

Case 4 :  $1 < p < q < \infty$  We do not keep to the same notation as Case 3. Take another  $r < q$  so close to  $q$  that  $\frac{pr}{q} > 1$ . As in Remark 40.3, we can assume  $f_N = f_{N+1} = \dots = f_{-N} =$

$f_{-N-1} = \dots = 0$  for some large  $N$  as long as the constants do not depend on  $N$ . According to Proposition 40.1 we have

$$(41.6) \quad \|Mf_j\|_{L^p(\ell^q)}^{\frac{q}{r}} = \left\| Mf_j^{\frac{q}{r}} \right\|_{L^{\frac{pr}{q}}(\ell^r)} = \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} Mf_j(x)^{\frac{q}{r}} \cdot g_j(x)$$

for some sequence  $\{g_j\}_{j \in \mathbb{N}_0}$  of  $L^{(\frac{pr}{q})'(\ell^r)}$ -functions with norm 1. Now that we have gone through Case 3, we have

$$\begin{aligned} \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} Mf_j(x)^{\frac{q}{r}} \cdot g_j(x) dx &\lesssim \sum_{j \in \mathbb{N}_0} \int_{\mathbb{R}^d} |f_j(x)|^{\frac{q}{r}} \cdot Mg_j(x) dx \\ &\leq \int_{\mathbb{R}^d} \left( \sum_{j \in \mathbb{N}_0} |f_j(x)|^q \right)^{\frac{1}{r}} \cdot \left( \sum_{j \in \mathbb{N}_0} Mg_j^{r'} \right)^{\frac{1}{r'}} \\ &\lesssim \|f_j\|_{L^p(\ell^q)}^{\frac{1}{r}} \|Mg_j\|_{L^{(\frac{pr}{q})'(\ell^r)}} \\ &\lesssim \|f_j\|_{L^p(\ell^q)}^{\frac{1}{r}} \|g_j\|_{L^{(\frac{pr}{q})'(\ell^r)}} \\ &= \|f_j\|_{L^p(\ell^q)}^{\frac{1}{r}}. \end{aligned}$$

Putting together this and first observations, we finish the proof. □

Finally to conclude this section we present an application of the Marcinkiewicz integral. Before we give the definition of the Marcinkiewicz integral operator, let us begin with a simple observation.

**Theorem 41.3.** *Let  $1 < p < \infty, 1 < q < \infty$ . Suppose that we are given a disjoint collection*

*of cubes  $\{Q_j\}_{j \in \mathbb{N}}$ . Set  $J_q(x) := \sum_{j=1}^{\infty} \frac{\ell(Q_j)^{dq}}{\ell(Q_j)^{dq} + |x - c(Q_j)|^{dq}}$ . Then  $\|J_q^{\frac{1}{q}}\|_p \lesssim \left( \sum_{j=1}^{\infty} |Q_j| \right)^{\frac{1}{p}}$ .*

*Proof.* Theorem 41.3 follows from the fact that  $J_q(x)$  is pointwise equivalent to  $\sum_{j=1}^{\infty} M[\chi_{Q_j}](x)^q$ . □

**Definition 41.4** (Marcinkiewicz integral). Let  $\Omega$  be an open set and define

$$(41.7) \quad d_{\Omega}(x) = \text{dist}(x, \Omega^c)$$

for  $x \in \Omega$ . Then define

$$(41.8) \quad I_q(x) := \int_{\Omega} \frac{\delta(x)^{n(q-1)}}{|x - y|^{nq} + \delta(x)^{nq}} dy$$

for  $1 < q < \infty$ .

Let us prove the following as an application of Theorem 41.1.

**Theorem 41.5.** *Suppose that  $1 < p < \infty, 1 < q < \infty$  and that  $\Omega$  is an open set. Then we have*

$$(41.9) \quad \left\| J_q^{\frac{1}{q}} \right\|_p \simeq |\Omega|^{\frac{1}{p}},$$

where the implicit constant in  $\simeq$  does not depend on  $\Omega$ .

*Proof.* We use Theorem 13.26 to find a collection of disjoint collection cubes  $\{Q_j\}_{j \in \mathbb{N}}$

$$(41.10) \quad \chi_\Omega \leq \sum_{j=1}^{\infty} \chi_{2Q_j} \leq \sum_{j=1}^{\infty} \chi_{4Q_j} \leq N_d \chi_\Omega, \quad 8Q_j \cap \Omega^c \neq \emptyset,$$

where  $N_d$  is a constant depending only on  $d$ . By using this covering, it is easy to see that

$$(41.11) \quad I_q(x) \simeq \sum_{j=1}^{\infty} \frac{\ell(Q_j)^{dq}}{\ell(Q_j)^{dq} + |x - c(Q_j)|^{dq}}.$$

Therefore, we are in the position of applying Theorem 41.3 to obtain

$$(41.12) \quad \left\| J_q^{\frac{1}{q}} \right\|_p \simeq \left( \sum_{j=1}^{\infty} |Q_j| \right)^{\frac{1}{p}} \simeq |\Omega|^{\frac{1}{p}}.$$

Thus, we obtain the desired result.  $\square$

#### 42. VECTOR-VALUED INEQUALITIES FOR CZ-OPERATORS

In this section we are going to prove this inequality.

Calderón-Zygmund decomposition for vector-valued functions. Using the same idea as in the  $\mathbb{R}$ -valued case we can obtain the following decomposition.

**Lemma 42.1.** *Let  $1 \leq p \leq \infty$ . Assume that  $F : \mathbb{R}^d \rightarrow \ell^p$  is a function with following properties.*

(A)  *$F$  is componentwise measurable, i.e.*

$$(42.1) \quad F = (f_n)_{n=1}^{\infty} \text{ with } f_n \text{ measurable for all } n \in \mathbb{N}.$$

(B)  *$F$  is integrable, i.e.*

$$(42.2) \quad \int_{\mathbb{R}^d} \|F(x)\|_{\ell^2} dx < \infty.$$

Then  $F$  admits decomposition  $F = G + \sum B_j$  with the following properties.

- (1)  $G, B_j (j \in \mathbb{N})$  is componentwise measurable.
- (2)  $G, B_j (j \in \mathbb{N})$  is integrable.
- (3) There is a disjoint family of cubes  $Q_l (l \in \mathbb{N})$ . And  $G$ , the  $B_j$  and the  $Q_l$  satisfy

$$G(x) = \left\{ \chi_{\mathbb{R}^d \setminus \Omega} \cdot f_n(x) + \sum_{l \in \mathbb{N}} m_{Q_l}(f_n) \cdot \chi_{Q_l} \right\}_{n \in \mathbb{N}}$$

$$B_j(x) = \left\{ \chi_{Q_j}(x) \left( f_n(x) - \frac{1}{|Q_j|} \int_{Q_j} f_n \right) \right\}_{n \in \mathbb{N}}.$$

- (4)  $\|G_j(x)\|_{\ell^q} \leq 2^d t$  for a.e.  $x \in \mathbb{R}^d$ .
- (5)  $\sum_j |Q_j| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} \|F_j(x)\|_{\ell^q} dx.$

The proof of this lemma can be achieved by re-examining the proof of the usual version.

*Sketch of the proof.* For the proof we proceed in the same way as in the  $\mathbb{R}$ -valued case. Set

$$(42.3) \quad M_{\text{dyadic}}[\{F_j\}_{j \in \mathbb{Z}}](x) := M_{\text{dyadic}}[\|F_j\|_{\ell^q}](x).$$

Then we will obtain by using  $\mathbb{R}_+$  valued version of the CZ-decomposition to obtain disjoint cubes  $\{Q_j\}$ . The properties listed in the statement can be verified easily once we define  $G$  and  $B_j$ .  $\square$

*Exercise 230.* Supply the details of the proof of the above theorem.

Vector-valued inequality for singular integral operators. We also have the same type of the results for singular integral operators.

**Theorem 42.2.** *Suppose that  $1 < p, q < \infty$  and that  $T$  is a CZ-operator. Then we have*

$$(42.4) \quad \|Tf_j\|_{L^p(\ell^q)} \lesssim \|f_j\|_{L^p(\ell^q)}.$$

*Proof.* We have only to go through a vector-valued version of the preceding proof along with Lemma 42.1.  $\square$

A different approach of Theorem 41.1.

We have proved Theorem 41.1 by means of the weight technique. The maximal operator controls somehow singular integral operators, as we have seen in Cotlar's inequality. When we formed the CZ-decomposition, the maximal operator was indispensable. Therefore, the proof of singular integral operators heavily depends on the maximal operator.

However, once we develop the CZ-theory, we can change our viewpoint.

Here, only in this subsection, it is more convenient to deal with the centered maximal operator. Recall that

$$(42.5) \quad M^*f(x) := \sup_{j \in \mathbb{Z}} m_{B(x, 2^j)}(|f|).$$

A trivial but important point is to use an equivalent definition of the above operator. Let  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  be a radial function taken so that

$$(42.6) \quad \chi_{B(1)} \leq \varphi \leq \chi_{B(2)}.$$

Set  $\varphi_k(x) := \frac{1}{2^{dk}} \varphi\left(\frac{x}{2^j}\right)$ . Then we have

$$(42.7) \quad Tf(x) := \sup_{k \in \mathbb{Z}} \varphi_k * |f|(x) \simeq M^*f(x).$$

We define a vector-valued kernel function  $K : \mathbb{R}^d \rightarrow \ell^\infty$  by

$$(42.8) \quad K(x) := \{\varphi_k(x)\}_{k \in \mathbb{Z}}.$$

Denote by  $M_A$  the multiplication operator  $\{a_k\}_{k \in \mathbb{Z}} \mapsto \{A_k \cdot a_k\}_{k \in \mathbb{Z}}$ , where  $A \in \ell^\infty$ . Then it is easy to see that it satisfies the following vector-valued norm estimates.

$$\begin{aligned} \|M_K(x)\|_{\ell^q} &= \|K(x)\|_{\ell^\infty} \lesssim \frac{1}{|x|^d} \\ \|M_{K'}(x)\|_{\ell^q} &= \|K'(x)\|_{\ell^\infty} \lesssim \frac{1}{|x|^{d+1}}. \end{aligned}$$

Therefore, we are in the position of using the CZ-theory and its vector-valued extension to obtain

$$(42.9) \quad \|Tf_j\|_{L^p(\ell^q)} \leq \|f_j\|_{L^p(\ell^q)}$$

for all  $1 < p, q \leq \infty$ . The case when  $p < \infty$  and  $q = \infty$  can be readily supplemented since

$$(42.10) \quad \sup_{j \in \mathbb{N}} |Tf_j(x)| \leq T \left( \sup_{j \in \mathbb{N}} |f_j| \right) (x).$$

Therefore, putting the above observations together, we obtain

$$(42.11) \quad \|Mf_j\|_{L^p(\ell^q)} \leq \|f_j\|_{L^p(\ell^q)},$$

for all  $1 < p, q \leq \infty$ . Together with the weak estimate of vector-valued version, we also have the following.

**Theorem 42.3.** *Let  $1 < q \leq \infty$ . Then*

$$(42.12) \quad \left| \left\{ x \in \mathbb{R}^d : \left( \sum_{j=1}^{\infty} Mf_j(x)^q \right)^{\frac{1}{q}} > \lambda \right\} \right| \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} \left( \sum_{j=1}^{\infty} Mf_j(x)^q \right)^{\frac{1}{q}} dx$$

for every  $\{f_j\}_{j=1}^{\infty} \in L^1(\ell^q)$  and  $\lambda > 0$ .

*Proof.* Supply the details of the proof of Theorem 42.3. □

*Exercise 231.* Formulate and prove counterparts of Theorems 41.3 and 41.5 by using Theorem 42.3.

*Exercise 232.* Let  $1 < p < \infty$  and  $1 < q \leq \infty$ . Then by mimicking the proof of Theorem 21.17, establish Theorem 21.17 that

$$(42.13) \quad \int_{\mathbb{R}^d} \left( \sum_{j=1}^{\infty} Mf_j(x)^q \right)^{\frac{p}{q}} w(x) dx \lesssim \int_{\mathbb{R}^d} \left( \sum_{j=1}^{\infty} |f_j(x)|^q \right)^{\frac{p}{q}} w(x) dx$$

for  $w \in A_p$ .

### 43. KINTCHINE'S INEQUALITY AND ITS APPLICATIONS

In this section we encounter theorems to which probability theory is applied skillfully. We can say that probability theory and harmonic analysis are closely connected.

Rademacher sequence. The Rademacher sequence, which appeared in Part 10, is a powerful tool in harmonic analysis. As an application of the Rademacher sequence, let us see how it is used for the vector-valued extension. As a preparatory step, we prove an equivalence inequality.

**Definition 43.1** (Rademacher sequence). Let  $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), dx|_{[0, 1]})$ . Define a sequence of random variables  $r_n, n = 1, 2, \dots$  by

$$(43.1) \quad r_n(t) := \sum_{j=1}^{2^n} (-1)^{j-1} \chi_{[(j-1)2^{-n}, j2^{-n})}(t).$$

Recall that the functions  $\{r_n(t)\}_{n=1}^{\infty}$  form a family of independent random variables.

**Theorem 43.2** (Kinchine inequality). *Let  $\{a_j\}_{j \in \mathbb{N}} \in \ell^2$ . Then, for  $0 < p < \infty$ ,*

$$(43.2) \quad \left\| \sum_{j=1}^{\infty} a_j r_j \right\|_p \sim \|a_j\|_{\ell^2}.$$

*Proof.* Homogeneity allows us to normalize  $\|a_j\|_{\ell^2} = 1$ . We may also assume there exists  $j_0$  such that  $a_j = 0$  for any  $j \geq j_0$ . We write

$$(43.3) \quad A(t) = \sum_{j=1}^{\infty} a_j r_j(t) \quad (t \in \mathbb{R}).$$

Then, in view of the fact that

$$(43.4) \quad |t|^p \leq \max(1, t^{[p]+1}) \leq ([p] + 1)!(e^t + e^{-t}),$$

we obtain

$$\begin{aligned} \|A\|_p^p &= \int_0^1 |A(t)|^p dt \\ &\lesssim \left\{ \int_0^1 \exp(A(t)) dt + \int_0^1 \exp(-A(t)) dt \right\} \\ &= \left\{ \prod_j \int_0^1 \exp(a_j r_j(t)) dt + \prod_j \int_0^1 \exp(-a_j r_j(t)) dt \right\} \\ &= 2 \prod_j \frac{\exp(a_j) + \exp(-a_j)}{2}. \end{aligned}$$

We now exploit the inequality  $\frac{e^t + e^{-t}}{2} \leq \exp(t^2)$  ( $t \in \mathbb{R}$ ).

$$\|A\|_p^p \lesssim \prod_j \frac{e^{a_j} + e^{-a_j}}{2} \lesssim \prod_j \exp(a_j^2) = e.$$

Thus, the right inequality was proved.

To prove the converse inequality, we may assume that  $p < 2$  by virtue of the monotonicity with respect to  $p$ . If  $p = 2$ , then the equivalence formula in question is actually the inequality because of the orthogonality of  $r_j$ 's. Thus

$$(43.5) \quad \|a_j\|_{\ell^2} = \|A\|_2 \lesssim \|A\|_p^{1-\theta} \cdot \|A\|_4^\theta \lesssim \|A\|_p^{1-\theta} \cdot \|a_j\|_{\ell^2}^\theta,$$

where  $\frac{1}{2} = \frac{1-\theta}{p} + \frac{\theta}{4}$ . If we arrange the above inequality, then we obtain

$$(43.6) \quad \|a_j\|_{\ell^2} \lesssim \|A\|_p$$

for  $p < 2$ .

If we put the above observation together, then we obtain the desired result. □

Application to vector-valued inequalities. By the Rademacher sequence method we see that bounded linear operators immediately admit  $\ell^2$ -valued extension.

**Theorem 43.3.** *Let  $(X, \mathcal{B}, \mu)$  be a measure space. Suppose that  $T : L^p(\mu) \rightarrow L^p(\mu)$  is a bounded linear operator. Then we have*

$$(43.7) \quad \|Tf_j\|_{L^p(\ell^2, \mu)} \lesssim \|f_j\|_{L^p(\ell^2, \mu)}.$$

*Proof.* Let  $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), dx|_{[0, 1]})$  as above.  $(r_n)_{n \in \mathbb{N}}$  is a Rademacher sequence as above. The monotone convergence theorem allows us to assume  $f_j$  is zero for  $j$  larger than some  $J$ .

Now we use the Kintchine inequality.

$$\begin{aligned} (\text{L.H.S.})^p &\lesssim \int_X \int_\Omega \left| \sum_{j=1}^\infty r_j(\omega) T f_j(x) \right|^p dP(\omega) d\mu(x) \\ &= \int_X \int_\Omega \left| T \left( \sum_{j=1}^\infty r_j(\omega) f_j(x) \right) \right|^p dP(\omega) d\mu(x). \end{aligned}$$

By the  $L^p(\mathbb{R}^d)$ -boundedness of  $T$  and the Fubini theorem, we obtain

$$\begin{aligned} (\text{L.H.S.})^p &\lesssim \iint_{\Omega \times X} \left| T \left( \sum_{j=1}^\infty r_j(\omega) f_j(x) \right) \right|^p d\mu(x) dP(\omega) \\ &\lesssim \iint_{\Omega \times X} \left| \sum_{j=1}^\infty r_j(\omega) f_j(x) \right|^p d\mu(x) dP(\omega). \end{aligned}$$

Changing the order of the integration once more, we finally obtain

$$(\text{L.H.S.})^p \lesssim \int_X \int_\Omega \left| \sum_{j=1}^\infty r_j(\omega) f_j(x) \right|^p dP(\omega) d\mu(x) \lesssim \|f_j\|_{L^p(\ell^2, \mu)}^p.$$

This is the desired result. □

Application to unconditional basis. As an another application of the equivalence inequality of Rademacher sequence, we exhibit a counterexample showing that it is a basis but is not an unconditional one.

**Theorem 43.4.** *Let  $1 < p < 2$  or  $2 < p < \infty$ . Then  $\{e^{ik \cdot x}\}_{k \in \mathbb{Z}^d}$  is a basis but is not an unconditional basis of  $L^p(\mathbb{T}^d)$ .*

*Proof.* By Theorem 6.12 it is true that  $\{e^{ik \cdot x}\}_{k \in \mathbb{Z}^d}$  is a basis. Let us disprove that it is unconditional. Set

$$(43.8) \quad S_\beta f(x) = \sum_{j \in \mathbb{Z}^d} \beta_j \left( \int_{[0, 2\pi]^d} f(x) e^{-ijy} dy \right) e^{ijx},$$

for  $\beta \in \ell^\infty$  with finite non-zero entry. Then by dual formula

$$(43.9) \quad \int_{\mathbb{T}^d} S_\beta f \cdot g = \int_{\mathbb{T}^d} g \cdot S_\beta f$$

for  $f, g \in C(\mathbb{T})$ , we have  $\|S_\beta\|_{L^p(\mathbb{T}^d)} = \|S_\beta\|_{L^{p'}(\mathbb{T}^d)}$ . Therefore, we may assume that  $1 < p < 2$ . Assume that  $\{e^{ik \cdot x}\}_{k \in \mathbb{Z}^d}$  is an unconditional basis of  $L^p(\mathbb{T})$ . Denote by  $\{r_j\}_{j \in \mathbb{Z}^d}$  the Rademacher sequence. Namely, rearrange the Rademacher sequence in numerical order. Then

$$(43.10) \quad \left( \sum_{j \in \mathbb{Z}^d} |a_j|^2 \right) \simeq \left( \int_0^1 \left| \sum_{j \in \mathbb{Z}^d} a_j r_j(t) e^{-ijx} \right|^p dt \right)^{\frac{1}{p}}.$$

Therefore, it follows that

$$(43.11) \quad \left( \sum_{j \in \mathbb{Z}^d} |a_j|^2 \right) \simeq \left( \int_0^1 \left( \int_{[0, 2\pi]^d} \left| \sum_{j \in \mathbb{Z}^d} a_j r_j(t) e^{-ijx} \right|^p dx \right) dt \right)^{\frac{1}{p}}.$$



Since  $\{e^{ik \cdot x}\}_{k \in \mathbb{Z}^d}$  is assumed unconditional,

$$\int_{[0,2\pi]^d} \left| \sum_{j \in \mathbb{Z}^d} a_j r_j(t) e^{-ijx} \right|^p dx = \left\| S_{\{r_j(t)\}_{j \in \mathbb{Z}^d}} \left( \sum_{j \in \mathbb{Z}^d} a_j e^{ij \cdot * } \right) \right\|_{L^p(\mathbb{T}^d)}^p \lesssim \left\| \sum_{j \in \mathbb{Z}^d} a_j e^{ij \cdot * } \right\|_{L^p(\mathbb{T}^d)}^p .$$

This is impossible because

$$(43.12) \quad \left\| \sum_{j \in \mathbb{Z}^d} a_j e^{ij \cdot * } \right\|_{L^2(\mathbb{T}^d)} \simeq \left( \sum_{j \in \mathbb{Z}^d} |a_j|^2 \right)^{1/2} \leq \left\| \sum_{j \in \mathbb{Z}^d} a_j e^{ij \cdot * } \right\|_{L^p(\mathbb{T}^d)}$$

and hence  $L^2(\mathbb{T}^d) \hookrightarrow L^p(\mathbb{T}^d)$ . This is a contradiction. □

Notes and references for Chapter 19.

Section 40. In this part the author has referred to [16]. for the vector-valued inequalities, which are explained in great detail.

Vector-valued inequalities are taken up in [433].

Section 41. Theorem 41.1 is originally proved in [182]. The proof of the vector valued inequality is based on [424].

Nowadays we often use this inequality to develop the theory of the Triebel-Lizorkin space. For the example of this usage we refer to Chapter 24 of this book and to [64].

Marcinkiewicz integral has a long history. Marcinkiewicz investigated the Marcinkiewicz integral in [323, 324, 325]. In fact there are many crucial papers using the Marcinkiewicz integral. For example we refer to [119, 131, 182, 499] as well as a textbook [73] for more details. Theorem 41.3, 41.5 and Exercise 231 are due to the paper by C. Fefferman and E. M. Stein [182].

Section 42. The spirit of the proof of Theorem 21.17 can be found in the proof of [117, Theorem 6].

J. Garcia and J. Martel extended Theorems 42.2 and 42.3 to the non-homogeneous space (See [207]).

Section 43. Theorem 43.2 is found by Kinchine in 1923.

Theorem 43.3

Theorem 43.4

## Part 20. Littlewood-Paley theory

In this section we develop the Littlewood-Paley theory, where we will see the square sum is very powerful.

### 44. LITTLEWOOD-PALEY THEORY

Littlewood-Paley theory is a theory to decompose the function and obtain the equivalent norm.

44.1. **G-functionals.** As a starting point of Littlewood-Paley theory, we define  $G$ -functional.

**Lemma 44.1.** *We can take  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  such that*

$$(44.1) \quad \text{supp}(\varphi) \subset \overline{B(2) \setminus B\left(\frac{1}{2}\right)}, \quad \int_0^\infty |\mathcal{F}\varphi(tx)|^2 \frac{dt}{t} = 1, \quad (x \in \mathbb{R}^d \setminus \{0\}).$$

*Proof.* Take a nonzero radial function  $\varphi_0 \in \mathcal{S}(\mathbb{R}^d)$  supported on an annulus  $\overline{B(2) \setminus B\left(\frac{1}{2}\right)}$ . Then  $\varphi_1 = \Delta\varphi_0$  is radial because  $\mathcal{F}\varphi_1$  is radial. As a consequence we have

$$(44.2) \quad \int_0^\infty |\mathcal{F}\varphi_1(tx)|^2 \frac{dt}{t} = \int_0^\infty |\mathcal{F}\varphi_1(t|x|e_1)|^2 \frac{dt}{t} = \int_0^\infty |\mathcal{F}\varphi_1(te_1)|^2 \frac{dt}{t},$$

where  $e_1 = (1, 0, 0, \dots, 0)$ . Note that

$$(44.3) \quad |\mathcal{F}\varphi_1(te_1)| \leq \frac{t}{t^2 + 1}, \quad (t > 0).$$

Thus, the above integral is finite. In order to obtain the desired  $\varphi$ , it remains to normalize  $\varphi_1$  by multiplying a constant.  $\square$

*Notation .* Throughout this section, we fix  $\varphi$  from (44.1). We define a functional  $s_\varphi$  by the formula

$$(44.4) \quad s_\varphi(f)(x) := \left( \int_0^\infty |f * \varphi_t(x)|^2 \frac{dt}{t} \right)^{\frac{1}{2}},$$

where  $\varphi_t(x) := \frac{1}{t^d} \varphi\left(\frac{x}{t}\right)$ . Given a measurable function  $F : (0, \infty) \rightarrow \mathbb{C}$ , we define

$$(44.5) \quad \|F(t)\|_{L_t^2\left(\frac{dt}{t}\right)} := \left( \int_0^\infty |F(t)|^2 \frac{dt}{t} \right)^{\frac{1}{2}}.$$

$L_t^2\left(\frac{dt}{t}\right)$  is a set of all measurable functions  $F : (0, \infty) \rightarrow \mathbb{C}$  for which  $\|F\|_{L_t^2\left(\frac{dt}{t}\right)}$  is finite.

Given a measurable function  $F : \mathbb{R}_+^{d+1} \rightarrow \mathbb{C}$ , we define

$$(44.6) \quad \|F(x, t)\|_{L_x^2\left(L_t^2\left(\frac{dt}{t}\right)\right)} := \left( \int_0^\infty \left( \int_0^\infty |F(x, t)|^2 \frac{dt}{t} \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}.$$

We have the following proposition.

**Proposition 44.2.** *If  $f \in L^2(\mathbb{R}^d)$ , then we have*

$$(44.7) \quad \|s_\varphi(f)\|_2 = \|f\|_2.$$

*Proof.* First, we change the order of integrations.

$$(44.8) \quad \|s_\varphi(f)\|_2^2 = \int_{\mathbb{R}^d} \left( \int_0^\infty |f * \varphi_t(x)|^2 \frac{dt}{t} \right) dx = \int_0^\infty \left( \int_{\mathbb{R}^d} |f * \varphi_t(x)|^2 dx \right) \frac{dt}{t}.$$

Then we use the Plancherel theorem.

$$(44.9) \quad \|s_\varphi(f)\|_2^2 = \int_0^\infty \left( \int_{\mathbb{R}^d} |\mathcal{F}f(\xi)\mathcal{F}\varphi_t(\xi)|^2 d\xi \right) \frac{dt}{t}.$$

Change the order of integrations once more.

$$(44.10) \quad \|s_\varphi(f)\|_2^2 = \int_{\mathbb{R}^d} \left( \int_0^\infty |\mathcal{F}f(\xi)\mathcal{F}\varphi_t(\xi)|^2 \frac{dt}{t} \right) d\xi = \int_{\mathbb{R}^d} |\mathcal{F}f(\xi)|^2 \left( \int_0^\infty |\mathcal{F}\varphi_t(\xi)|^2 \frac{dt}{t} \right) d\xi.$$

Invoke the assumption and we obtain

$$(44.11) \quad \|s_\varphi(f)\|_2^2 = \int_{\mathbb{R}^d} |\mathcal{F}f(\xi)|^2 d\xi = \int_{\mathbb{R}^d} |f|^2.$$

This is the desired result. □

The following theorem is a part of the Littlewood-Paley theory.

**Theorem 44.3.** *Suppose that  $1 < p < \infty$ . Then for every  $f \in L^p(\mathbb{R}^d)$  we have  $\|s_\varphi(f)\|_p \simeq \|f\|_p$ .*

The proof below shows that we can actually prove that

$$(44.12) \quad c_p^{-1} \|f\|_p \leq \|s_\varphi(f)\|_p \leq c_p \|f\|_p$$

for a collection of constants  $\{c_p\}_{p>1}$  with  $c_p = c_{p'}$ .

*Proof.* Proof of  $\|s_\varphi(f)\|_p \leq c_p \|f\|_p$  By definition it follows that

$$(44.13) \quad s_\varphi(f)(x) = \|\varphi_t * f(x)\|_{L_t^2(\frac{dt}{t})}$$

Thus we can consider the mapping defined by convolution

$$(44.14) \quad \Phi : f \in L^2(\mathbb{R}^d) \mapsto [t \mapsto f * \varphi_t(\cdot)] \in L_t^2\left(\frac{dt}{t}\right).$$

And our job is to prove  $\|\Phi(f)\|_p \leq c_p \|f\|_p$ .

Notice that

$$(44.15) \quad \|\varphi_t(x)\|_{L_t^2(\frac{dt}{t})} \lesssim |x|^{-d}, \quad \|\partial_j \varphi_t(x)\|_{L_t^2(\frac{dt}{t})} \lesssim |x|^{-d-1}$$

for all  $j = 1, 2, \dots, d$ .

This implies the linear operator

$$(44.16) \quad K(x) : \mathbb{R} \rightarrow B, \quad [K(x)](t) := [t \rightarrow \varphi_t(x)]$$

satisfies Hörmander's condition. If we invoke the vector-valued CZ-theory,  $\|s_\varphi(f)\|_p \leq c_p \|f\|_p$  is established. □

*Proof.* Proof of  $\|s_\varphi(f)\|_p \geq c_{p'} \|f\|_p$  By polarization formula (polarize (44.7)) we have

$$\int_{\mathbb{R}^d} f(x) \cdot \bar{g}(x) dx = \int_{\mathbb{R}^d} \left( \int_0^\infty f * \varphi_t(x) \overline{g * \varphi_t(x)} \frac{dt}{t} \right) dx$$

for all  $g \in \mathcal{S}(\mathbb{R}^d)$ . Using Hölder's inequality twice, we have

$$\begin{aligned} \int_{\mathbb{R}^d} f(x) \cdot \bar{g}(x) dx &\leq \int_{\mathbb{R}^d} \|f * \varphi_t(x)\|_{L_t^2(\frac{dx}{t})} \cdot \|g * \varphi_t(x)\|_{L_t^2(\frac{dx}{t})} dx \\ &\leq \|f * \varphi_t(x)\|_{L_x^p(L_t^2(\frac{dx}{t}))} \cdot \|g * \varphi_t(x)\|_{L_x^{p'}(L_t^2(\frac{dx}{t}))}. \end{aligned}$$

However in the last step we have proved that

$$(44.17) \quad \|g * \varphi_t(x)\|_{L_x^{p'}(L_t^2(\frac{dx}{t}))} \leq c_{p'} \|g\|_{p'}.$$

Thus it follows that for all  $g \in L^{p'} \setminus 0$ ,

$$(44.18) \quad \frac{1}{\|g\|_{p'}} \left| \int_{\mathbb{R}^d} f(x) \cdot \bar{g}(x) dx \right| \leq c_{p'} \|f * \varphi_t(x)\|_{L_x^p(L_t^2(\frac{dx}{t}))}.$$

If we take sup of the left-hand-side of the last formula over  $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$ , we obtain

$$(44.19) \quad \|f\|_p \leq c_{p'} \|f * \varphi_t(x)\|_{L_x^p(L_t^2(\frac{dx}{t}))}.$$

This is precisely what we want.  $\square$

#### 44.2. Discrete Littlewood-Paley theory.

Now we take up the discrete version, which is of importance as well. To do this, now we choose a smooth function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  so that

$$(44.20) \quad A^{-1} \chi_{B(4) \setminus B(2)} \leq \varphi \leq A \chi_{B(8) \setminus B(1)}$$

for some  $A > 0$ . Let  $\varphi_j(x) := \varphi(2^{-j}x)$ . We define

$$(44.21) \quad \|f\|_{\dot{F}_{p2}^0} := \|\{\varphi_j(D)f\}_{j \in \mathbb{Z}}\|_{L^p(\ell^2)} := \|\{\mathcal{F}^{-1}(\varphi_j \mathcal{F}f)\}_{j \in \mathbb{Z}}\|_{L^p(\ell^2)}.$$

Later on it will be made clear why this norm is denoted by  $\dot{F}_{p2}^0(\mathbb{R}^d)$ .

**Theorem 44.4.** *Let  $1 < p < \infty$ . Then*

$$(44.22) \quad \|f\|_{\dot{F}_{p2}^0} \sim \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

*Proof.* Step 1 : The case when  $p = 2$  In this case the theorem follows from the Plancherel theorem.

Step 2 : The left-hand inequality

It suffices to show an estimate

$$(44.23) \quad \left\| \left( \sum_{j=0}^N |\varphi_j(D)f|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \|f\|_p,$$

whose implicit constant is independent of  $N$ .

To do this we take the Rademacher sequence  $\{r_j\}_{j \in \mathbb{N}_0}$ . Then we have only to prove

$$(44.24) \quad \left\| \sum_{j=0}^N r_j(t) \varphi_j(D)f \right\|_p \lesssim \|f\|_p$$

for every  $t \in [0, 1]$  in view of equivalence

$$(44.25) \quad \|\{a_j\}_{j \in \mathbb{N}_0}\|_{\ell^2} \sim \left\| \sum_{j \in \mathbb{N}_0} a_j r_j \right\|_{L^p([0,1])}.$$

Let us denote

$$(44.26) \quad \sum_{j=0}^N r_j(t) \varphi_j(D) f = A_t * f,$$

where  $A_t$  is a smooth function for every fixed  $t$ . Note that the family  $A_t$  satisfies

$$(44.27) \quad |A_t(x)| \lesssim |x|^{-d}, \quad |\partial_j A_t(x)| \lesssim |x|^{-d-1}.$$

Furthermore, by the Plancherel theorem  $\|A_t * f\|_2 \leq \|f\|_2$ .

Step 3 : The right-hand inequality

To prove this, we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)$ . Then observe that

$$(44.28) \quad \int_{\mathbb{R}^d} f(x) \cdot g(x) dx = \sum_{j=0}^{\infty} \int_{\mathbb{R}^d} \varphi_j(D) f \cdot \varphi_j(D) g$$

for all  $g \in \mathcal{S}(\mathbb{R}^d)$ . Therefore, it follows that

$$(44.29) \quad \left| \int_{\mathbb{R}^d} f(x) \cdot g(x) dx \right| \leq \|f\|_{F_{p,2}^0} \cdot \|g\|_{F_{p',2}^0}$$

by virtue of the Hölder inequality. Now that we have established in Step 2 that

$$(44.30) \quad \|g\|_{F_{p',2}^0} \lesssim \|g\|_{p'},$$

it follows that

$$(44.31) \quad \|f\|_p \lesssim \|f\|_{F_{p,2}^0}$$

by duality. □

### 45. BURKHOLDER-GUNDY-DAVIS INEQUALITY

Now we discuss the analogy in probability theory.

The Burkholder-Gundy-Davis inequality. We adopt the following notation.

*Notation* . Let  $X = (X_n)_{n \in \mathbb{N}_0}$  be a random variable.

- (1) Set  $X_n^*(\omega) := \sup_{0 \leq k \leq n} |X_k(\omega)|$  and  $X^*(\omega) := \lim_{n \rightarrow \infty} X_n^*(\omega)$ .
- (2) Let  $d_n = X_n - X_{n-1}$ ,  $n \in \mathbb{N}$ ,  $d_0 = X_0$  be the martingale difference.
- (3) Set  $S(X)_n := \sqrt{\sum_{k=1}^n d_k^2}$  for  $n \in \mathbb{N}_0$  and  $S(X) := \lim_{n \rightarrow \infty} S(X)_n$ .

Below we assume that  $X$  is a martingale.

**Theorem 45.1** (Burkholder-Gundy-Davis). *Suppose that  $p \in [1, \infty)$ . Then*

$$(45.1) \quad \|S(X)\|_p \sim_p \|X^*\|_p$$

for every  $L^p$ -martingale  $X$ .

Key-lemma. The following good- $\lambda$  inequality is a key to our observation. Recall that we have taken up the same type of inequality in Chapter in 9.

**Theorem 45.2.** *Let  $\beta > \delta + 1$  and  $0 < \delta \leq \frac{1}{2}$ . Suppose that  $v_n \in \mathcal{F}_{n-1}$ ,  $n \in \mathbb{N}$  such that  $|d_n| \leq v_n$  for all  $n \in \mathbb{N}$ . Then*

$$(45.2) \quad P(X^* > \beta\lambda, S(X) \vee v^* \leq \delta\lambda) \leq \frac{2\delta^2}{(\beta - \delta - 1)^2} P(X^* > \lambda)$$

$$(45.3) \quad P(S(X) > \beta\lambda, X^* \vee v^* \leq \delta\lambda) \leq \frac{4\delta^2}{(\beta - \delta - 1)^2} P(X^* > \lambda).$$

*Proof of (45.2).* We define three stopping times.

$$\begin{aligned} \tau &:= \min\{n : |X_n| > \lambda\} \\ \mu &:= \min\{n : |X_n| > \beta\lambda\} \\ \sigma &:= \min\{n : S(X)_n > \delta\lambda \text{ or } v_n > \delta\lambda\}. \end{aligned}$$

Here it will be understood that  $\min(\emptyset) = \infty$ . Note that  $\tau \leq \mu$  by their definitions. By definition of these stopping times we can write

$$(45.4) \quad \{X^* > \beta\lambda, S(X) \vee v^* \leq \beta\lambda\} = \{\mu < \infty, \sigma = \infty\}.$$

We set  $G_n = \sum_{k=0}^n \mathbf{1}_{(\tau, \mu \wedge \sigma]} d_k$ . On  $\{\mu < \infty, \sigma = \infty\}$  we have  $G_n = X_{n \wedge \mu} - X_{n \wedge \tau}$ . Note that

$$(45.5) \quad G^* \geq |G_\mu| = |X_\mu - X_\tau| \geq |X_\mu| - |X_\tau|.$$

It cannot happen that  $\tau = 0$ , if  $\sigma = \infty$ . This is because  $|X_0| = |d_0| \leq |v_0| \leq \delta$ . Consequently  $\tau \geq 1$ . Then  $|X_\tau| \leq |X_{\tau-1}| + |d_\tau|$ . Since  $\tau - 1 < \min\{n : |X_n| > \lambda\} = \tau$ , it follows that  $\tau - 1 \notin \{n : |X_n| > \lambda\}$ , that is,  $|X_{\tau-1}| \leq \lambda$ . Furthermore, since  $v_n \leq \delta\lambda$  on  $\{\sigma = \infty\}$  for all  $n$ , it follows that  $|d_\tau| \leq \delta\lambda$ . Finally, since  $\mu = \min\{n : |X_n| > \beta\lambda\}$ , it follows that  $|X_\mu| > \beta\lambda$ . If we combine these observation, we obtain  $G^* \geq (\beta - 1 - \delta)\lambda$ .

As a result we obtain  $\{\mu < \infty, \sigma = \infty\} \subset \{G^* > (\beta - 1 - \delta)\lambda\}$ . The Chebychev inequality then gives us

$$(45.6) \quad P\{\mu < \infty, \sigma = \infty\} \leq \frac{1}{(\beta - 1 - \delta)^2 \lambda^2} \sup_n \|G_n\|_2.$$

Note that  $\{G_n^2 - S_n(G)^2\}_{n \in \mathbb{N}}$  is a martingale. If  $X^* \leq \lambda$ , then  $\tau = \infty$  consequently  $G_n$  is 0. Thus, we obtain

$$(45.7) \quad \|S(G)^2\|_2^2 = \int_{\{X^* > \lambda\}} S(G)^2 dP.$$

Now we have  $|d_\sigma| \leq v_\sigma \leq \delta\lambda$  and  $S_{\sigma-1}(X) \leq \delta\lambda$ , if  $\sigma < \infty$ . Consequently we obtain

$$(45.8) \quad S(G)^2 = S_\sigma(G)^2 = (S_{\sigma-1}(G)^2 + d_\sigma^2) I_{\{\sigma < \infty\}} + S_\sigma(X)^2 I_{\{\sigma = \infty\}} \leq 2\delta^2 \lambda^2.$$

As a result we obtain (45.2). □

*Proof of (45.3).* We reset the definition of  $\tau, \sigma, \mu$ . We define three stopping times.

$$\begin{aligned} \tau &:= \min\{n : S(X)_n > \lambda\} \\ \mu &:= \min\{n : S(X)_n > \beta\lambda\} \\ \sigma &:= \min\{n : |X_n| > \delta\lambda\} \cup \{n : v_{n+1} > \delta\lambda\}. \end{aligned}$$

It is the same that  $\tau \leq \mu$  as before. We have

$$\{S(X) > \beta\lambda, X^* \vee v^* \leq \delta\lambda\} = \{\tau < \infty, \sigma = \infty\}.$$

We set

$$H_n = \left( \sum_{k=1}^{n \wedge \mu \wedge \sigma} d_k^2 - \sum_{k=1}^{n \wedge \tau \wedge \sigma} d_k^2 \right)^{\frac{1}{2}} = \left( \sum_{k=0}^n [(\mathbf{1}_{[0, \mu \wedge \sigma]} - \mathbf{1}_{[0, \tau \wedge \sigma]}) d_k]^2 \right)^{\frac{1}{2}}.$$

Then on  $\{\tau < \infty, \sigma = \infty\}$  we have

$$(45.9) \quad H_n \geq \left( \sum_{k=0}^n [\mathbf{1}_{[0, \mu \wedge \sigma]} d_k]^2 \right)^{\frac{1}{2}} - \left( \sum_{k=0}^n [\mathbf{1}_{[0, \tau \wedge \sigma]} d_k]^2 \right)^{\frac{1}{2}}.$$

In the same way as before, on the event  $\{\tau < \infty, \sigma = \infty\}$  we have

$$(45.10) \quad \tau \geq 1, H^* \geq (\beta - 1 - \delta)\lambda.$$

Consequently we obtain

$$(45.11) \quad P\{S(X) > \beta\lambda, X^* \vee v^* \leq \delta\lambda\} \leq \frac{1}{(\beta - 1 - \delta)^2 \lambda^2} E[H^{*2}].$$

Note that

$$(45.12) \quad E[G^{*2}] = \lim_{n \rightarrow \infty} \sum_{k=0}^n E[\mathbf{1}_{\tau \wedge \sigma < k \leq \mu \wedge \sigma} d_k^2] = \lim_{n \rightarrow \infty} E[(\mathbf{1}_{[0, \mu \wedge \sigma]} - \mathbf{1}_{[0, \tau \wedge \sigma]}) X_n^2].$$

We observe that  $\tau = \infty$  on  $\{S(X) \leq \lambda\}$ . As a result the identity

$$(45.13) \quad (\mathbf{1}_{[0, \mu \wedge \sigma]} - \mathbf{1}_{[0, \tau \wedge \sigma]}) X_n^2 = \mathbf{1}_{\{S(X) > \lambda\}} (\mathbf{1}_{[0, \mu \wedge \sigma]} - \mathbf{1}_{[0, \tau \wedge \sigma]}) X_n^2$$

holds. We are to obtain the upper bound of this term.

Suppose that  $\tau \wedge \sigma < n \leq \mu \wedge \sigma$ . First, we notice that  $\tau < \sigma < \mu$ . As a result we obtain  $\tau < n \leq \sigma < \mu$ . If  $n < \sigma$ , then  $|X_n| \leq \delta\lambda$ . If  $n = \sigma$ , then  $|X_n - X_{n-1}| \leq v_n \leq \delta\lambda$  and  $|X_{n-1}| \leq \delta\lambda$ . Thus, if we assume  $\tau \wedge \sigma < n \leq \mu \wedge \sigma$ , then  $|X_n| \leq 2\delta\lambda$ .

Consequently we obtain

$$(45.14) \quad (\mathbf{1}_{[0, \mu \wedge \sigma]} - \mathbf{1}_{[0, \tau \wedge \sigma]}) X_n^2 \leq 4\delta^2 \lambda^2.$$

From these observations (45.3) is completely proved. □

**Corollary 45.3.** *Suppose that  $X$  is an  $L^p$ -martingale, that is,  $X$  is a martingale such that  $X_n$  is  $L^p$ -integrable for each  $n \in \mathbb{N}$ . Assume in addition that  $(v_n)_{n \in \mathbb{N}}$  is a predictable process dominating  $(d_n)_{n \in \mathbb{N}}$ . Then one has*

$$(45.15) \quad E[S(X)^p] \lesssim E[\min(X^*, v^*)^p], \quad E[X^{*p}] \lesssim E[\min(S(X), v^*)^p],$$

where  $c$  does not depend on  $v$  nor  $X$ .

*Proof.* These estimates follow from the good  $\lambda$ -inequality argument. To apply the good  $\lambda$ -inequality, we have verify some integrability condition. This can be achieved by the stopping time argument. Speaking precisely, let  $T_k \equiv k$  be a constant stopping time with  $k \in \mathbb{N}$ . Then we are in the position of using the good  $\lambda$ -inequality with  $X$  replaced by  $X^{T_k}$ . By the monotone convergence theorem, the passage of the general case can be obtained. □

Davis decomposition. To proceed further we consider the Davis decomposition.

**Theorem 45.4.** *Any martingale  $X := (X_n)_{n \in \mathbb{N}}$  admits the decomposition into a sum of two martingales  $G = (G_n)_{n \in \mathbb{N}}$  and  $H = (H_n)_{n \in \mathbb{N}}$  with the following properties. Below denote by  $d = (d_n)_{n \in \mathbb{N}}, e = (e_n)_{n \in \mathbb{N}}, u = (u_n)_{n \in \mathbb{N}}$  the martingale differences of  $X, G$  and  $H$  respectively.*

$$(1) \quad |e_n| \leq 4d_{n-1}.$$

(2) Then there exists a sequence of Random variables  $(\beta_n)_{n \in \mathbb{N}}$  such that

$$(45.16) \quad |\beta_n| \leq 2(d_n^* - d_{n-1}^*)$$

$$(45.17) \quad \sum_{n=0}^{\infty} |u_n| \leq \sum_{n=0}^{\infty} |\beta_n| + \sum_{n=0}^{\infty} E[|\beta_n| : \mathcal{F}_{n-1}].$$

Here, we have set  $\mathcal{F}_{-1} := \{\emptyset, \Omega\}$ .

*Proof.* The proof begins with constructing auxiliary processes  $\alpha = (\alpha_n)_{n \in \mathbb{N}}$  and  $\beta = (\beta_n)_{n \in \mathbb{N}}$ . First we set  $\alpha_n := d_n \mathbf{1}_{\{|d_n| \leq 2d_{n-1}^*\}}$ ,  $\beta_n := d_n \mathbf{1}_{\{|d_n| > 2d_{n-1}^*\}}$ . Then  $\alpha$  and  $\beta$  are  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -adapted. Next, define the martingale differences of  $G$  and  $H$ :  $e_n := \alpha_n - E[\alpha_n : \mathcal{F}_{n-1}]$ ,  $u_n := \beta_n - E[\beta_n : \mathcal{F}_{n-1}]$ .  $G$  and  $H$  are defined so that their martingale differences are  $e$  and  $u$ :

$$(45.18) \quad G_n := \sum_{k=0}^n e_k, \quad H_n := \sum_{k=0}^n u_k.$$

Now we shall verify that  $G$  and  $H$  satisfy the requirements of the theorem. To begin with (45.17) is clear by definition of  $\beta$ . If  $\beta_n \neq 0$ , then we have  $|d_n| > 2d_{n-1}^*$ . It follows from this that

$$(45.19) \quad 2(|d_n|^2 - |d_{n-1}|^2) > |d_n|^2 + 4(d_{n-1}^*)^2 - 2|d_{n-1}|^2 > |d_n|^2.$$

Consequently (45.16) is also verified. It remains to check  $|e_n| \leq 4|d_{n-1}^*|$ . Since  $|\alpha_n| \leq 2d_{n-1}^*$  by definition of  $\alpha_n$ , we have  $|e_n| \leq |\alpha_n| + E[|\alpha_n| : \mathcal{F}_{n-1}] \leq 4d_{n-1}^*$ . As a result we obtain the desired decomposition.  $\square$

Proof of the Burkholder-Gundy-Davis inequality. First, we form the Davis decomposition. We split  $X$  by  $X = G + H$ , where  $G$  and  $H$  satisfy

- (1)  $|e_n| \leq 4d_{n-1}$ ,  $e_n = G_n - G_{n-1}$ ,  $n \geq 2$  and  $e_1 = G_1$ .
- (2) Let  $(u_n)_{n \in \mathbb{N}}$  be the martingale difference of  $(H_n)_{n \in \mathbb{N}}$  and  $(d_n)_{n \in \mathbb{N}}$  that of  $(X_n)_{n \in \mathbb{N}}$ . Then there exists a sequence of Random variables  $(\beta_n)_{n \in \mathbb{N}}$  such that

$$|\beta_n| \leq 2(d_n^* - d_{n-1}^*)$$

$$\sum_{n=0}^{\infty} |u_n| \leq \sum_{n=0}^{\infty} |\beta_n| + \sum_{n=0}^{\infty} E[|\beta_n| : \mathcal{F}_{n-1}].$$

Then, since  $|H_n| \leq \sum_{k=0}^n |u_k|$  and  $S(H) \leq \left( \sum_{n=0}^{\infty} |u_n|^2 \right)^{\frac{1}{2}} \leq \sum_{n=0}^{\infty} |u_n|$ , we obtain

$$(45.20) \quad H^* \vee S(H) \leq \sum_{n=0}^{\infty} |u_n|.$$

A direct consequence of this inequality is

$$(45.21) \quad E[H^* \vee S(H)] \leq 2 \sum_{k=0}^{\infty} E[|u_k|] \leq 4E[d^*] \leq 8E[X^* \wedge S(X)].$$

As a result we obtained a key estimate that immediately completes the proof of the Burkholder-Gundy-Davis inequality.

$$(45.22) \quad E[H^* \vee S(H)] \leq 8E[X^* \wedge S(X)].$$

With (45.22) in mind, we prove  $E[X^*] \lesssim E[S(X)]$ . Speaking precisely, we shall prove

$$(45.23) \quad E[G^*] \lesssim E[S(X)], \quad E[H^*] \lesssim E[S(X)].$$



From (45.22) we obtain, taking into account  $S(G) \leq S(H) + S(X)$

$$(45.24) \quad E[H^*], E[S(G)], E[S(H)] \lesssim E[S(X)].$$

Note that by applying the key lemma with  $X$  replaced by  $G$  and  $v_n = 4d_{n-1}, n = 1, 2, \dots$  we obtain

$$(45.25) \quad E[G^*] \lesssim E[S(G) \vee v^*].$$

Since  $v^* \leq 2S(X)^*$ , it follows that

$$(45.26) \quad E[v^*] \lesssim E[S(X)].$$

Finally from (45.22) we obtain

$$(45.27) \quad E[H^*] \leq 8E[X^*].$$

If we combine (45.24)–(45.27), then we obtain

$$(45.28) \quad E[X^*] \lesssim E[S(X)]$$

The proof of  $E[S(X)] \lesssim E[X^*]$  is similar. We have only to reverse the role of the square function  $S$  and the maximal operator  $*$ .

Finally to conclude this section we present an application of this inequality.

**Definition 45.5.** Denote by  $\mathcal{M}^p$  the set of all  $L^p$ -bounded martingales. Equip  $\mathcal{M}^p$  with a norm

$$(45.29) \quad \|M\|_{\mathcal{M}^p} := \sup_{n \in \mathbb{N}} E[|M_n|^p]^{\frac{1}{p}}.$$

*Exercise 233.* Show that  $\mathcal{M}^p$  is a Banach space with its norm.

**Theorem 45.6.** Let  $H = (H_n)_{n \in \mathbb{N}_0}$  be a bounded predictable process and  $1 < p < \infty$ . Then

$$(45.30) \quad M \in \mathcal{M}^p \mapsto H \cdot M \in \mathcal{M}^p$$

is a bounded linear operator.

*Proof.* Note that

$$(45.31) \quad H \cdot M_n - H \cdot M_{n-1} = H_{n-1}d_n,$$

where  $d = (d_n)$  is the martingale difference of  $M$ . Therefore by the Burkholder-Gundy-Davis inequality we have

$$\begin{aligned} \|H \cdot M\|_{\mathcal{M}^p} &\lesssim \left\| \left( \sum_{n=1}^{\infty} |H_{n-1}d_n|^2 \right)^{\frac{1}{2}} \right\|_p \\ &\lesssim \left( \sup_n \|H\|_{\infty} \right) \cdot \left\| \left( \sum_{n=1}^{\infty} |H_{n-1}d_n|^2 \right)^{\frac{1}{2}} \right\|_p \\ &\lesssim \left( \sup_n \|H\|_{\infty} \right) \cdot \|M\|_{\mathcal{M}^p}. \end{aligned}$$

Therefore the proof is now complete. □

Notes and references for Chapter 20.

Section 44. Littlewood and Paley introduced the  $g$ -functional, which was published in Proc. London Math. Soc. in 1936. The idea that the vector-valued singular integral operators are applied to the Littlewood theory dates back to [413].

Littlewood-Paley theory is a widely spread theory and many function spaces are characterized in terms of this theory.

For example, in [254] M. Izuki characterized  $L^p(w)$  with  $w$  belonging to a class wider than  $A_p$  in connection with wavelet.

Theorem 44.3

Theorem 44.4

We refer to [435] for the application of the Rudin function.

An example of the application of the Littlewood-Paley theory can be found in [452].

Section 45. The connection between the square sum operator and the maximal operator is described in [105]. Theorem 45.1 reflects this aspect.

The proof of Theorem 45.1 using Theorem 45.2 depends on the work [106].

The proof of Theorem 45.1 is very ingenious because we have to depend on the good- $\lambda$  inequality. However, the result is already known [104, 8] when  $1 < p < \infty$ .

Theorem 45.4

Theorem 45.6

Finally, there are many other “Littlewood Paley”-type operators such as

$$(45.32) \quad L(f) = \left( E[M : \mathcal{F}_1]^2 + \frac{1}{k} \sum_{j=1}^k \left| E[M : \mathcal{F}_k] - \frac{1}{k} \sum_{l=1}^k E[M : \mathcal{F}_l] \right|^2 \right)^{\frac{1}{2}} .$$

For details of this type of Littlewood Paley operators we refer to [456, 479].

## Part 21. Function spaces appearing in harmonic analysis

One of the big branches in analysis is theory of function spaces. Function spaces are tools used not only in analysis but also in algebra and in geometry.

- (1) When we prove the Riemann-Roch theorem in Riemannian surfaces, a theorem in algebraic geometry we need function spaces called  $A^2(\Omega)$ .
- (2) When we prove the Hodge-decomposition theorem in differential geometry, Sobolev spaces  $W^{2,k}(M)$  play an important role. Indeed, one has to start from function spaces with lower smoothness order instead of considering  $C^\infty$ -smoothness.
- (3) In stochastic analysis, to consider the Ito stochastic integrals, we need to consider  $L^2(P)$  spaces in connection with Brownian motions.

Therefore, function spaces occur in every field of mathematics.

The simplest function spaces appearing in analysis seems  $BC(X)$ , the set of all complex-valued continuous functions defined on a metric space  $(X, d)$ . This function spaces appears as an example or an exercise of complete metric spaces.

The next example we encounter in the course of analysis is the Lebesgue space  $L^p(\mathbb{R}^d)$ . Unlike the space  $BC(X)$ , it had been very difficult to prove its completeness. One of the reasons why theory of function spaces are prevailant is that the triangle inequality is available in addition to the completeness.

Everything is fine as long as we have only to work within the framework of  $BC(X)$  or  $L^p(\mathbb{R}^d)$ .

However, things are not so nice once we consider many other mathematical problems. For example, these spaces are not sufficient when we consider the Laplace equation  $-\Delta u = f$ . For the purpose of tackling this equation, we are led to thinking of Sobolev spaces. Indeed, in this case, it is convenient to introduce

$$\|f\|_{H^1} = \sqrt{\|f\|_{L^2}^2 + \sum_{j=1}^d \|\partial_j f\|_{L^2}^2}$$

because we can use the integration by parts. Since there are many problems which can be described in terms of partial differential equations, it could not be better if everything would be solved within the framework of Sobolev spaces.

However, many elliptic problems (appearing in many branches of mathematics) require some other spaces of fractional smoothness order. Spaces of fractional smoothness order may not enough to describe other partial differential equations. So we need more delicate spaces such as Besov spaces which we are going to define in this chapter.

Another problem concering  $BC(X)$  and  $L^p(\mathbb{R}^d)$  is that neither of them contains  $|x|^{-\alpha}$  ( $\alpha > 0$ ).

This simplest function appears everywhere but to handle this function, we need to restrict its domain. However, by the Morrey space  $\mathcal{M}_q^{n/\alpha}(\mathbb{R}^d)$  with  $1 < q < n/\alpha$  we have  $|x|^{-\alpha} \in \mathcal{M}_q^{n/\alpha}(\mathbb{R}^d)$ . Note that the Morrey norm  $\|f\|_{\mathcal{M}_q}$  is given by

$$\|f\|_{\mathcal{M}_q} = \sup_B |B|^{\frac{1}{p} - \frac{1}{q}} \|f\|_{L^q(B)},$$

where  $B$  runs over all balls.

There are many function spaces appearing in harmonic analysis. Let us describe some of them together with some backgrounds. Suppose that we are given a function  $\Phi : (0, \infty) \rightarrow (0, \infty)$ .

Then define the Orlicz norm by

$$\|f\|_{L^\Phi} = \inf \left\{ \lambda > 0 : \Phi \left( \frac{|f(x)|}{\lambda} \right) \leq 1 \right\}.$$

Seemingly, the definition is hard to understand. But when we consider the case  $\Phi(t) = t^p$ , then we have  $\|f\|_{L^\Phi} = \|f\|_p$ .

One of the reasons why we are fascinated with this space is that it can describe the behaviour of the Hardy-Littlewood maximal operator  $M$ . Here  $M$  is the maximal operator given by

$$Mf(x) = \sup_{x \in Q: \text{cubes}} \frac{1}{|Q|} \int_Q |f(y)| dy$$

and the trouble is that

$$\|Mf\|_1 \leq C \|f\|_1 \quad (f \in L^1(\mathbb{R}^d))$$

fails for any positive constant  $C$ . Orlicz spaces can recover such a difficulty.

Another remedy to recover the failure of the inequality  $\|Mf\|_1 \lesssim \|f\|_1$  is to use Lorentz spaces. Let us define

$$f^*(t) = \sup \{ \text{ess. inf}_E |f(x)| : |E| = t \}$$

and

$$\|f\|_{p,q} = \left( \int_0^\infty (t^{1/p} f^*(t))^q \frac{dt}{t} \right)^{1/q}.$$

The Lorentz space  $L^{p,q}(\mathbb{R}^d)$  is given by this norm.

In this chapter, with these problems in mind, we consider many function spaces.

## Part 22. Functions on $\mathbb{R}$

### OVERVIEW

Orlicz spaces measures integrability more precisely than the  $L^p(\mathbb{R}^d)$ -spaces. In this chapter we are going to present further examples of applications of the results in Part 6. In Section 46 we apply the results in Section 12. Sections 47 and 48 are devoted to an introduction to Orlicz spaces. For the sake of simplicity we are not going into the details here.

#### 46. 1-VARIABLE FUNCTIONS AND THEIR DIFFERENTIABILITY A.E.

Having cleared up, the maximal theory and so on, we can now go into the investigation of 1-variable functions.

##### 46.1. Monotone functions.

First, we investigate monotone functions.

**Theorem 46.1.** *Any increasing function  $\mathbb{R}$  is continuous except for a countable number of points.*

*Proof.* We have only to prove this by assuming  $f$  is constant on  $[1, \infty)$  and  $(-\infty, -1]$ . Denote by  $D$  the point at which  $f$  is not continuous. Then we have

$$(46.1) \quad D = \bigcup_{k=1}^{\infty} \left\{ x \in [-1, 1] : f_+(x) - f_-(x) > \frac{1}{k} \right\},$$

where  $f_+(x)$  and  $f_-(x)$  are given respectively by

$$(46.2) \quad f_+(x) = \lim_{\varepsilon \downarrow 0} f(x + \varepsilon), \quad f_-(x) = \lim_{\varepsilon \downarrow 0} f(x - \varepsilon).$$

It is easy to see that

$$(46.3) \quad \left\{ x \in [-1, 1] : f_+(x) - f_-(x) > \frac{1}{k} \right\}$$

is finite. Therefore  $D$  is countable. □

**Theorem 46.2.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is an increasing function, then it is differentiable almost everywhere.*

*Proof.* As before, we write  $f_+(x) = \lim_{\varepsilon \downarrow 0} f(x + \varepsilon)$ . and let

$$(46.4) \quad A := \{x \in \mathbb{R} : f_+(x) = f(x)\}.$$

Then the complement of  $A$  is countable and hence  $A$  is dense in  $\mathbb{R}$ . If  $f_+(x)$  is differentiable at  $x \in A$ , then so is  $f$ . Indeed, Let  $\varepsilon > 0$ . Then, letting  $a$  the differential of  $f_+$  at  $x$ , we can find  $\delta > 0$  such that

$$(46.5) \quad \left| \frac{f_+(x+h) - f_+(x)}{h} - a \right| \leq \varepsilon$$

for all  $0 < |h| < \delta$ . We claim that the same inequality holds for  $f$ . Let  $0 < |h| < \delta$ . If  $x+h \in A$ , then the matter is trivial because  $f_+(x+h) = f(x+h)$  and  $f_+(x) = f(x)$ . Let  $x+h \notin A$ . Then, we have

$$(46.6) \quad \frac{f(x+h) - f(x)}{h} - a \leq \lim_{\substack{\theta \downarrow 0 \\ h+x+\theta \in A}} \frac{f(x+h+\theta) - f(x)}{h+\theta} - a \leq \varepsilon$$

and similarly

$$(46.7) \quad -\varepsilon \leq \lim_{\substack{\theta \downarrow 0 \\ h+x-\theta \in A}} \frac{f(x+h-\theta) - f(x)}{h-\theta} - a \leq \frac{f(x+h) - f(x)}{h} - a.$$

Therefore the same inequality as (46.5) holds for  $f$ . This implies that  $f$  is differentiable at  $x$ . Therefore it is justified that we can replace  $f$  with  $f_+$ . Thus, it follows that  $f$  can be assumed right-continuous from the beginning. Therefore

$$(46.8) \quad S_{\pm}f(x) := \limsup_{t \downarrow 0} \frac{f(x \pm t) - f(x)}{\pm t}, \quad I_{\pm}f(x) := \liminf_{t \downarrow 0} \frac{f(x \pm t) - f(x)}{\pm t}$$

are all measurable functions of  $x$ ,

We shall prove  $\{x \in \mathbb{R}^d : S_+(x) > I_-(x)\}$  has measure zero. Since  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , this is reduced to proving  $E = \{x \in \mathbb{R}^d : S_+(x) > b > a > I_-(x)\}$  for all  $a, b \in \mathbb{Q}$  with  $a < b$ . Let  $\varepsilon > 0$  be given arbitrarily and pick an open set  $O$  engulfing  $E$  so that  $|O| \leq |E| + \varepsilon$ . We set

$$(46.9) \quad \mathcal{I} := \{[x-h, x] : x \in E, h > 0, f(x) - f(x-h) < ha, [x-h, x] \subset O\}$$

Then  $\mathcal{I}$  is a Vitali covering. Therefore, we can select a disjoint family  $\{I_j\}_{j \in J}$  of  $\mathcal{I}$  so that  $E \setminus \bigcup_{j \in J} I_j$  has measure 0. We write  $I_j = [x_j - h_j, x_j]$ . Set

$$(46.10) \quad B := \bigcup_{j \in J} E \cap I_j.$$

Then  $|B| = |E|$ . Furthermore

$$(46.11) \quad \sum_{j \in J} f(x_j - h_j) - f(x_j) \leq \sum_{j \in J} v h_j \leq a|O| \leq a(|E| + \varepsilon).$$

Next, we set

$$(46.12) \quad \mathcal{J} := \{[x, x+h] : x \in B, h > 0, f(x+h) - f(x) < ha, [x, x+h] \subset I_k \text{ for some } k\}$$

Then  $\mathcal{J}$  is a Vitali covering of  $B$ . Therefore, we can select a disjoint family  $\{L_k\}_{k \in K}$  of  $\mathcal{J}$  so that  $B \setminus \bigcup_{k \in K} L_k$  has measure 0. We write  $L_k = [y_k, y_k + \delta_k]$ . Then we have

$$(46.13) \quad b|E| \leq \sum_{k \in K} f(y_k + \delta_k) - f(y_k) \leq \sum_{j \in J} f(x_j - h_j) - f(x_j) \leq a(|E| + \varepsilon).$$

Therefore, we obtain  $|E| \leq \frac{\varepsilon}{b-a}$ , implying  $|E| = 0$ .

Analogously we can prove  $\{x \in \mathbb{R}^d : S_-(x) > I_+(x)\}$  has measure zero. Hence

$$(46.14) \quad I_+(x) \leq I_-(x) \leq S_+(x) \leq S_-(x) \leq I_+(x)$$

holds for almost everywhere  $x \in \mathbb{R}$ . If this chain of inequality holds at some  $x$ , then it follows that  $f$  is differentiable at  $x$ .  $\square$

## 46.2. Functions of bounded variations.

**Definition 46.3.** An  $\mathbb{R}$ -valued function  $f$  defined on an interval  $I$  and  $[a, b]$  be an interval contained in  $I$ .

- (1) A partition  $\Delta$  of  $[a, b]$  is a finite sequence  $a = t_0 < t_1 < \dots < t_k = b$ . We write  $\Delta : a = t_0 < t_1 < \dots < t_k = b$ .
- (2) A partition  $\Delta$  of  $[a, b]$  is said to be finer than  $\Delta'$ , if  $\Delta'$  is obtained by adding some partition points to  $\Delta$ .
- (3) Given two partitions  $\Delta_1$  and  $\Delta_2$  of  $[a, b]$ , we define  $\Delta_1 \cup \Delta_2$  as the least finest partition finer than  $\Delta_1$  and  $\Delta_2$ .
- (4) Given a partition  $\Delta : a = t_0 < t_1 < \dots < t_k = b$ , define

$$(46.15) \quad P_\Delta[a, b] := \sum_{j=1}^k (f(t_j) - f(t_{j-1}))_+, \quad N_\Delta[a, b] := \sum_{j=1}^k (f(t_j) - f(t_{j-1}))_-,$$

$$(46.16) \quad T_\Delta[a, b] := P_\Delta[a, b] + N_\Delta[a, b].$$

and

$$(46.17) \quad P[a, b] = \sup_{\Delta} P_\Delta[a, b], \quad N[a, b] = \sup_{\Delta} N_\Delta[a, b], \quad T[a, b] = \sup_{\Delta} T_\Delta[a, b],$$

where  $\Delta$  runs over all partitions of  $[a, b]$ .

- (5) A function  $f$  is said to be of bounded variation on  $[a, b]$ , if

$$(46.18) \quad \text{Var}(f) = \sup\{T_\Delta[a, b] : \Delta : \text{is a partition of } [a, b]\} < \infty$$

**Example 46.4.** A set of all functions of bounded variation  $[a, b]$  forms a linear space. Any monotone function on  $[a, b]$  is of bounded variation. Therefore, if a function on  $[a, b]$  can be expressed as a difference of increasing functions, then it is of bounded variation.

We intend to prove the converse of this example.

**Lemma 46.5.** *Suppose that  $f$  is of bounded variation. Then we have the following.*

$$\begin{aligned} T[a, b] &= P[a, b] + N[a, b] \\ f(b) - f(a) &= P[a, b] - N[a, b]. \end{aligned}$$

*Proof.* Noting that  $f(b) - f(a) + N_{\Delta}[a, b] = P_{\Delta}[a, b]$ , we see  $f(b) - f(a) + N[a, b] = P[a, b]$ .

By definition we have

$$(46.19) \quad P[a, b] + N[a, b] = \sup_{\Delta_1} P_{\Delta_1}[a, b] + \sup_{\Delta_2} N_{\Delta_2}[a, b] = \sup_{\Delta_1, \Delta_2} (P_{\Delta_1}[a, b] + N_{\Delta_2}[a, b]),$$

where  $\Delta_1$  and  $\Delta_2$  run over all the partitions of  $[a, b]$ . Since

$$(46.20) \quad P_{\Delta_1}[a, b] + N_{\Delta_2}[a, b] \leq P_{\Delta_1 \cup \Delta_2}[a, b] + N_{\Delta_1 \cup \Delta_2}[a, b] = T_{\Delta_1 \cup \Delta_2}[a, b],$$

it follows that

$$(46.21) \quad P[a, b] + N[a, b] \leq \sup_{\Delta} T_{\Delta}[a, b] \leq T[a, b].$$

The converse inequality being immediate,  $T[a, b] = P[a, b] + N[a, b]$  is proved.  $\square$

The formula  $f(b) - f(a) + N[a, b] = P[a, b]$  gives us information of differentiability.

**Theorem 46.6.** *Any function of bounded variation on  $[a, b]$  can be represented as a difference of monotone function. In particular it is differentiable for almost everywhere on  $(a, b)$ .*

### 46.3. Absolutely continuous functions.

Now we encounter a skillful usage of the maximal function.

**Theorem 46.7** (Lebesgue). *Let  $f \in L^1_{\text{loc}}(\mathbb{R})$ . Then the function*

$$(46.22) \quad F(t) := \int_0^t f(u) \, du,$$

*is differentiable  $dt$ -almost every  $t \in \mathbb{R}$  and we have*

$$(46.23) \quad F'(t) = f(t)$$

*$dt$ -almost everywhere on  $\mathbb{R}$ .*

*Proof.* The matter being local, we can assume that  $f \in L^1(\mathbb{R})$ . We have only to show

$$(46.24) \quad \left\{ x \in \mathbb{R} : \limsup_{h \rightarrow 0} \left| \frac{1}{h} \int_x^{x+h} f(u) \, du - f(x) \right| > \varepsilon \right\}$$

has measure zero. However, if  $g \in C_c(\mathbb{R})$  then

$$(46.25) \quad \lim_{h \rightarrow 0} \left( \frac{1}{h} \int_x^{x+h} g(u) \, du - g(x) \right) = 0$$

for all  $x \in \mathbb{R}$ . Keeping this in mind, we pick  $g \in C_c$  arbitrarily. Then by virtue of the maximal inequality

$$\begin{aligned} & \left| \left\{ x \in \mathbb{R} : \limsup_{h \rightarrow 0} \left| \frac{1}{h} \int_x^{x+h} f(u) du - f(x) \right| > \varepsilon \right\} \right| \\ &= \left| \left\{ x \in \mathbb{R} : \limsup_{h \rightarrow 0} \left| \frac{1}{h} \int_x^{x+h} (f(u) - g(u)) du - (f(x) - g(x)) \right| > \varepsilon \right\} \right| \\ &\leq |\{M[f - g] > \varepsilon\}| \\ &\lesssim \frac{1}{\varepsilon} \|f - g\|_1. \end{aligned}$$

Since  $g$  is arbitrary, the measure in question is zero and the theorem is proved.  $\square$

**Definition 46.8.** A function  $f : [a, b] \rightarrow \mathbb{R}$  is said to be absolutely continuous, if for any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$(46.26) \quad \sum_{j=1}^k |f(b_j) - f(a_j)| < \varepsilon$$

whenever  $a \leq a_1 < b_1 < a_2 < \dots < a_{k-1} < b_{k-1} < a_k < b_k \leq b$  satisfies  $\sum_{j=1}^k (b_j - a_j) < \delta$ .

**Theorem 46.9.** A continuous function  $f : [a, b] \rightarrow \mathbb{R}$  is absolutely continuous, if and only if there exists a function  $g \in L^1([a, b])$  such that

$$(46.27) \quad f(t) = \int_{\frac{a+b}{2}}^t g(u) du + f\left(\frac{a+b}{2}\right).$$

*Proof.* “If” part is immediate. To prove “only if” part, we define a measure  $\mu$

$$(46.28) \quad \mu([c, d]) = f(d) - f(c).$$

Then  $\mu$  is a signed measure absolutely continuous with respect to  $dx$ . The density of  $\mu$  is the desired function  $g$ :  $g$  is a derivative of  $f$ .

Indeed, if  $t > c = \frac{a+b}{2}$ , then we have

$$f(t) - f(c) = \mu(c, t) = \int_c^t g(u) du + f\left(\frac{a+b}{2}\right).$$

If  $t \leq \frac{a+b}{2}$ , a similar argument works and the proof is therefore complete.  $\square$

Integration-by-parts formula revisited. As an application we can generalize the integration-by-parts formula.

**Theorem 46.10.**  $f \cdot g$  is absolutely continuous, if  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  are absolutely continuous functions. Furthermore,

$$(46.29) \quad \int_a^b f'(x)g(x) dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x) dx$$

for all  $-\infty < a < b < \infty$ .



*Proof.* Just notice that the function  $f \cdot g$  is again absolutely continuous and use the equality

$$(46.30) \quad \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx = \int_a^b (f(x)g(x))' dx = f(b)g(b) - f(a)g(a).$$

□

Let us take this advantage to prove the Bari-Stechkin lemma as an application of integration-by-parts formula.

**Theorem 46.11.** *Let  $\varphi : (0, \infty) \rightarrow (0, \infty)$  be a continuous function such that*

$$(46.31) \quad \int_r^\infty \varphi(t)t^{-1} dt \leq A\varphi(r) \quad (r > 0).$$

*Then for all  $\varepsilon < A^{-1}$  we have*

$$(46.32) \quad \int_r^\infty \varphi(t)t^{\varepsilon-1} dt \leq \frac{A}{1-A\varepsilon} \varphi(r)r^\varepsilon \quad (r > 0).$$

*Proof.* Let us write

$$(46.33) \quad \Phi(t) = \int_t^R \varphi(v)v^{-1} \quad (t > 0).$$

Then we have

$$\begin{aligned} \int_r^R \varphi(t)t^{\varepsilon-1} dt &= [-\Phi(t) \cdot t^\varepsilon]_r^R + \varepsilon \int_r^R \Phi(t) \cdot t^{\varepsilon-1} dt \\ &\leq \int_r^R \varphi(v)v^{-1} dv \cdot r^\varepsilon + A\varepsilon \int_r^R \varphi(u) \cdot t^{\varepsilon-1} dt \\ &\leq A\varphi(r)r^\varepsilon + A\varepsilon \int_r^R \varphi(u) \cdot t^{\varepsilon-1} dt. \end{aligned}$$

As a result, we have

$$(46.34) \quad \int_r^R \varphi(u) \cdot t^{\varepsilon-1} dt \leq \frac{A}{1-A\varepsilon} \varphi(r)r^\varepsilon.$$

Letting  $R \rightarrow \infty$ , we obtain the desired result. □

Lipschitz functions.

**Definition 46.12.** Let  $I$  be an interval on  $\mathbb{R}$ . A function  $f : I \rightarrow \mathbb{R}$  is said to be Lipschitz, if

$$(46.35) \quad \|f\|_{\text{Lip}(I)} := \sup \left\{ \frac{|f(x) - f(u)|}{|x - u|} : x, u \in I, x \neq u \right\} < \infty.$$

*Exercise 234.* Let  $I = [a, b]$  be a closed interval on  $\mathbb{R}$ . If a continuous function  $f : I \rightarrow \mathbb{R}$  is differentiable on  $(a, b)$  and its derivative is bounded, then prove that  $f$  is Lipschitz.

*Exercise 235.* Let  $I = (a, b)$  be an open finite interval on  $\mathbb{R}$  and  $f \in \text{Lip}(I)$ . Then  $f$  is extendible to a Lipschitz function on  $\mathbb{R}$ .

**Corollary 46.13.** *If  $f : I \rightarrow \mathbb{R}$  is a Lipschitz function, then it is differentiable a.e. on  $\mathbb{R}$ , and the derivative is bounded.*

#### 46.4. Convex functions.

Let  $I$  be an interval on  $\mathbb{R}$ . A function  $f : I \rightarrow \mathbb{R}$  is said to be convex, if

$$(46.36) \quad f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y),$$

while  $f$  is said to be concave, if  $-f$  is convex. In this section we investigate convex functions.

*Exercise 236.* Let  $I$  be an open interval on  $\mathbb{R}$ . Use the mean value theorem to prove that  $f : I \rightarrow \mathbb{R}$  is convex, if  $f \in C^2$  and  $f'' \geq 0$ .

*Exercise 237.* Suppose that  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function. We are going to prove

$$(46.37) \quad f\left(\frac{t_1 + t_2 + \dots + t_n}{n}\right) \leq \frac{1}{n}(f(t_1) + f(t_2) + \dots + f(t_n)).$$

We proceed in several steps to prove (46.37).

- (1)  $n = 4$ .
- (2)  $n$  is a power of 2 in general.
- (3)  $n \in \mathbb{N}$ . Hint: Pick an integer  $m$  so that  $n < 2^m$ . Then consider (2) for the special case:

$$(46.38) \quad t_{n+1} = t_{n+2} = \dots = t_{2^m} = \frac{t_1 + t_2 + \dots + t_n}{n}.$$

*Exercise 238* (Jensen's inequality). Suppose that  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function. Generalize the above inequality to the following form:

$$(46.39) \quad f\left(\sum_{k=1}^N \alpha_k x_k\right) \leq \sum_{k=1}^N \alpha_k f(x_k)$$

whenever  $N \in \mathbb{N}$ ,

$$(46.40) \quad x_1, x_2, \dots, x_N \in \mathbb{R}, \alpha_1, \alpha_2, \dots, \alpha_N \geq 0$$

and  $\sum_{k=1}^N \alpha_k = 1$ .

**Theorem 46.14.** Let  $I \in \mathcal{I}(\mathbb{R})$  and  $f : I \rightarrow \mathbb{R}$  a convex function. Then it is differentiable almost everywhere. If we let  $g \in L^1(I)$  be the derivative of  $f$ , then

$$(46.41) \quad f(t) = f(a) + \int_a^t f(u) du$$

for  $a \in I$ .

*Proof.* We may assume  $I$  is open. We shall prove that  $f$  is Lipschitz on any proper closed interval  $[a, b]$  of  $I$ . Choose auxiliary  $c$  and  $d$  so that  $[a, b] \subset (c, d) \Subset I$ . Since  $f$  is convex, we have

$$(46.42) \quad \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leq \frac{f(\gamma) - f(\beta)}{\gamma - \beta}.$$

Thus, for any  $x, y \in [a, b]$  with  $x > y$ , we have

$$(46.43) \quad M_1 := \frac{f(a) - f(c)}{a - c} \leq \frac{f(x) - f(a)}{x - a} \leq \frac{f(x) - f(y)}{x - y} \leq \frac{f(b) - f(y)}{b - y} \leq \frac{f(d) - f(b)}{d - b} := M_2$$

Therefore  $\frac{f(x) - f(y)}{x - y}$  remains bounded:

$$(46.44) \quad M_1 \leq \frac{f(x) - f(y)}{x - y} \leq M_2.$$

Therefore  $f$  is Lipschitz and hence  $f$  is differentiable a.e. on  $I$ . The assertion concerning inequality follows because  $f$  is absolutely continuous.  $\square$

Finally we take up Theorem 22.37 without proof for the sake of convenience.

**Theorem 46.15.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is convex, then we have*

$$(46.45) \quad f(x) = \sup\{ax + b : a, b \in \mathbb{R}, f(t) \geq at + b \text{ for all } t \in \mathbb{R}\}.$$

## 47. N-FUNCTIONS

### 47.1. Right-inverse.

**Lemma 47.1.** *Let  $\varphi : [0, \infty) \rightarrow [0, \infty)$  be a right-continuous increasing function satisfying*

$$(47.1) \quad \varphi(0) = 0, \lim_{t \rightarrow \infty} \varphi(t) = \infty.$$

*We define its right-inverse by*

$$(47.2) \quad \varphi^*(t) := \sup\{s \in [0, \infty) : \varphi(s) \leq t\}.$$

*Here it will be understood that  $\sup \emptyset = 0$ . Then  $\varphi^* : [0, \infty) \rightarrow [0, \infty]$  is a finite-valued, right-continuous and increasing function satisfying*

$$(47.3) \quad \varphi^*(0) = 0, \lim_{t \rightarrow \infty} \varphi^*(t) = \infty.$$

*Proof.* Since  $\varphi$  is increasing and  $\varphi(0) = 0$ ,  $\varphi^*(0) = 0$ . Inequality  $\varphi^*(\varphi(t)) \geq t$  gives  $\lim_{t \rightarrow \infty} \varphi^*(t) = \infty$ . Finiteness of  $\varphi^*(t)$  follows from  $\lim_{t \rightarrow \infty} \varphi(t) = \infty$  and the property of the set appearing in the definition of  $\varphi^*$ . From the definition of  $\varphi^*$ , it is increasing. To see that it is right-continuous, let  $S > \varphi^*(t)$ . Then  $S > \varphi^*(t)$  implies  $\varphi(S) > t$ . Therefore there exists  $\varepsilon > 0$  such that  $\varphi(S) > t + \varepsilon$ . This implies  $\varphi^*(t + \varepsilon) \leq S$ . Therefore  $\varphi^*$  is right-continuous.  $\square$

**Lemma 47.2.** *Let  $\varphi : [0, \infty) \rightarrow [0, \infty)$  be a right-continuous increasing function. Then*

$$(47.4) \quad (\varphi^*)^* = \varphi.$$

*Proof.* Let  $t > 0$  be a common continuous point of  $\varphi$  and  $(\varphi^*)^*$ . Then we have

$$\begin{aligned} (\varphi^*)^*(t) &= \lim_{\varepsilon \downarrow 0} (\varphi^*)^*(t - \varepsilon) \\ &= \lim_{\varepsilon \downarrow 0} \sup\{u \geq 0 : (\varphi^*)(u) \leq t - \varepsilon\} \\ &\leq \lim_{\varepsilon \downarrow 0} \sup\left\{u \geq 0 : (\varphi^*)(u) < t - \frac{\varepsilon}{2}\right\} \\ &\leq \lim_{\varepsilon \downarrow 0} \sup\left\{u \geq 0 : u < \varphi\left(t - \frac{\varepsilon}{2}\right)\right\} \\ &\leq \lim_{\varepsilon \downarrow 0} \varphi\left(t - \frac{\varepsilon}{2}\right) = \varphi(t). \end{aligned}$$

Meanwhile we have

$$(\varphi^*)^*(t) = \lim_{\varepsilon \downarrow 0} \sup\{(\varphi^*)(u) \leq t + \varepsilon\} \geq \lim_{\varepsilon \downarrow 0} \sup\{u \geq 0 : u < \varphi(t + \varepsilon)\} = \lim_{\varepsilon \downarrow 0} \varphi(t + \varepsilon) = \varphi(t).$$

Therefore, since the set of all common continuous points form a dense subset, it follows that  $\varphi = (\varphi^*)^*$ .  $\square$

Lemmas 47.1 and 47.2, along with the theorem below, explain why  $\varphi^*$  deserves its name.

**Theorem 47.3.** *Suppose that  $a$  is a non-negative function that is right-continuous and increasing. Assume in addition that  $\lim_{t \rightarrow \infty} a(t) = \infty$ . Define*

$$(47.5) \quad c(t) = \inf \{ a > t \}$$

for  $t > 0$ . Then we have the following.

(1)  $c$  is a non-negative, right-continuous and increasing function that satisfies  $\lim_{t \rightarrow \infty} c(t) = \infty$ .

(2) Define  $a_-(t) = \lim_{\varepsilon \downarrow 0} a(t - \varepsilon)$  and  $c_-(t) = \lim_{\varepsilon \downarrow 0} c(t - \varepsilon)$ . Then we have

$$(47.6) \quad a(c(t)) \geq a(c_-(t)) \geq t, \quad a_-(c_-(t)) \leq a_-(c(t)) \leq t.$$

(3)  $a(s) = \inf \{ c > s \}$  for all  $t > 0$ .

(1). It is clear that  $c$  is non-negative because there does exist a positive element in the set appearing in the infimum defining  $c(t)$ . It is also clear that  $c$  is increasing as well. To prove that  $c$  is right-continuous, let us choose a decreasing sequence  $\{t_j\}_{j \in \mathbb{N}}$  converging to  $t$ . Then

$$\lim_{j \rightarrow \infty} c(t_j) = \lim_{j \rightarrow \infty} \inf \{ a > t_j \} = \inf \left( \bigcup_{j=1}^{\infty} \{ a > t_j \} \right) = \inf \{ a > t \} = c(t).$$

Since  $c(a(s) + 1) \geq s$  for all  $s > 0$ , it is not so hard to see  $\lim_{t \rightarrow \infty} c(t) = \infty$ .  $\square$

(2). Assume that  $a(c_-(t)) < t$ . Then there exists  $\varepsilon > 0$  and  $\delta > 0$  such that  $a(c_-(t) + \delta) < t - \varepsilon$  by virtue of the right-continuity of  $a$ . As a consequence we have  $c_-(t) + \delta \notin \{s > 0 : a(s) > t - \varepsilon\}$ . Since  $\{s > 0 : a(s) > t - \varepsilon\}$  is open or closed interval, it follows that  $c_-(t) + \delta \leq \inf \{s > 0 : a(s) > t - \varepsilon\} = c(t - \varepsilon) \leq c_-(t)$ . This is a contradiction. As a result  $a(c_-(t)) \geq t$ .

Assume now that  $a_-(c(t)) > t$ . Then there exists another  $\delta > 0$  such that  $a(c(t) - \delta) > t$ . As a consequence there exists  $\varepsilon > 0$  such that  $a(c(t) - \delta) > t + \varepsilon$ . This implies  $c(t) - \delta \geq c(t + \varepsilon)$ . Since  $\delta, \varepsilon > 0$ , this contradicts to the fact that  $c$  is increasing.  $\square$

(3). First, we are going to prove  $a(s) \leq \inf \{t > 0 : c(t) > s\}$ . To prove this, for all  $\varepsilon > 0$  we shall prove  $a(s) - \varepsilon \notin \inf \{t > 0 : c(t) > s\}$ . Indeed, we have  $c(a(s) - \varepsilon) = \inf \{u : a(u) > a(s) - \varepsilon\} \leq s$ . It remains to show  $a(s) \geq \inf \{t > 0 : c(t) > s\}$ . Let  $\varepsilon > 0$ . Then it suffices to show  $a(s) + \varepsilon \geq \inf \{t > 0 : c(t) > s\}$ . To do this, we calculate  $c(a(s) + \varepsilon) \geq \inf \{u > 0 : a(u) > a(s) + \varepsilon\} > s$ . For the last inequality we have used the fact that  $a$  is increasing and right continuous.  $a(s) + \varepsilon \in \{t > 0 : c(t) > s\}$ . Thus,  $a(s) + \varepsilon \geq \inf \{t > 0 : c(t) > s\}$  and letting  $\varepsilon \downarrow 0$ , we obtain  $a(s) \geq \inf \{t > 0 : c(t) > s\}$ . Thus, (3) is proved completely.  $\square$

## 47.2. Definition and fundamental properties of N-functions.

N-functions. N-functions are good convex functions. Some literature allows the functions to take  $\infty$ . However, here for the sake of simplicity, we content ourselves with investigating the properties of functions taking finite values.

**Definition 47.4.** Denote by  $\mathbb{H}$  the set of all non-decreasing functions  $\varphi : (0, \infty) \rightarrow (0, \infty)$  such that  $\lim_{t \downarrow 0} \varphi(t) = 0$ .

**Definition 47.5.** A function  $\Phi \in \mathbb{H}$  is said to be a Young function, if it satisfies the following conditions.

(1)  $\Phi$  is continuous.

(2)  $\Phi$  is convex, that is,

$$(47.7) \quad \Phi((1 - \theta)t_1 + \theta t_2) \leq (1 - \theta)\Phi(t_1) + \theta\Phi(t_2)$$

for all  $t_1, t_2 \in (0, \infty)$  and  $0 < \theta < 1$ .

By convention define  $\phi(0) = 0$  and  $\phi(\infty) = \infty$ .

**Example 47.6.** The functions  $e^t - 1$  and  $t \log(t + 1)$  are N-functions.

*Exercise 239.* Show that  $t^\alpha$  is an N-function, where  $1 \leq \alpha < \infty$ .

**Theorem 47.7.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be a continuous function. Then  $\Phi$  is an N-function, if and only if there exists an increasing right-continuous function  $\varphi : [0, \infty) \rightarrow [0, \infty)$  satisfying

$$(47.8) \quad \Phi(t) = \int_0^t \varphi(s) ds > 0$$

for all  $t > 0$ . Furthermore, if this is the case, then  $\varphi$  is unique.

Here and below we call  $\Phi(t) = \int_0^t \varphi(s) ds$  canonical representation of an N-function  $\Phi$ .

**Lemma 47.8.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be an N-function.

- (1)  $\Phi(\alpha t) \leq \alpha\Phi(t)$  for all  $0 < \alpha < 1$  and  $0 \leq t < \infty$ .
- (2)  $t \mapsto \frac{\Phi(t)}{t}$  is increasing.
- (3)  $\Phi$  itself is strictly increasing.

*Proof.* We content ourselves with showing (1) only. Because the rest will be derived easily from (1). By the convexity we have

$$(47.9) \quad \Phi(\alpha t) = \Phi(\alpha t + (1 - \alpha)0) \leq \alpha\Phi(t) + (1 - \alpha)\Phi(0) = \alpha\Phi(t).$$

Therefore the lemma is proved. □

*Exercise 240.* Use (1) and prove the remaining assertions of the above lemma.

Next, we shall treat the information of  $\varphi$  appearing in the canonical representation.

**Lemma 47.9.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be a Young function with canonical representation

$$(47.10) \quad \Phi(t) = \int_0^t \varphi(s) ds.$$

Let  $t > 0$ .

- (1)  $\frac{\Phi(t)}{t} \leq \varphi(t) \leq \frac{\Phi(2t)}{t}$ .
- (2)  $\varphi\left(\frac{t}{2}\right) \leq \int_0^t \frac{\varphi(s)}{s} ds \leq \varphi(t)$ .

*Proof.* All the estimates are easy to derive.

$$(47.11) \quad \frac{\Phi(t)}{t} = \frac{1}{t} \int_0^t \varphi(s) ds \leq \frac{1}{t} \int_0^t \varphi(t) ds = \varphi(t).$$

$$(47.12) \quad \varphi(t) = \frac{1}{t} \int_t^{2t} \varphi(t) dt \leq \frac{1}{t} \int_t^{2t} \varphi(s) ds \leq \frac{1}{t} \int_0^{2t} \varphi(s) ds = \frac{\Phi(2t)}{t}.$$

To obtain the remaining estimate, we use the monotonicity of  $t \in [0, \infty) \mapsto \frac{\varphi(t)}{t}$ .

$$(47.13) \quad \varphi\left(\frac{t}{2}\right) \leq \int_{t/2}^t \frac{\varphi(t/2)}{t/2} ds \leq \int_{t/2}^t \frac{\varphi(s)}{s} ds \leq \int_0^t \frac{\varphi(s)}{s} ds.$$

$$(47.14) \quad \int_0^t \frac{\varphi(s)}{s} ds \leq \int_0^t \frac{\varphi(t)}{t} ds = \varphi(t).$$

Thus, the proof is therefore complete.  $\square$

*Exercise 241.* Let  $f : \mathbb{R} \rightarrow [0, \infty)$  be a convex function.

$$(1) \text{ Let } a < b. \text{ Show that } f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(t) dt \leq \frac{f(a)+f(b)}{2}.$$

(2) Let  $a \in \mathbb{R}$  and  $h > 0$ . Show that

$$(47.15) \quad \sum_{j=1}^k f\left(a + \left(j - \frac{1}{2}\right)h\right) \leq \int_a^{a+kh} f(t) dt \leq \frac{f(a)+f(a+kh)}{2} + \sum_{j=1}^{k-1} f(a + (j-1)h).$$

$$(3) \text{ Show that } \frac{2}{3} < \log 2 = \int_1^2 \frac{dx}{x} < \frac{3}{4}.$$

Conjugate function. It is not so hard to prove

$$(47.16) \quad ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'},$$

where  $1 < p < \infty$  and  $a, b > 0$ . This inequality is a special case of the Young inequality. Changing the viewpoint, this inequality can be seen as the one for the function  $\varphi(t) = \frac{1}{p}t^p$ .

We are led to consider the following generalization:

$$(47.17) \quad ab \leq \varphi(a) + \psi(b),$$

where  $\psi$  is a convex functions. Our present problem is to obtain a function  $\psi$  for a given N-function  $\varphi$ .

**Definition 47.10.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be an N-function with canonical representation

$$(47.18) \quad \Phi(t) = \int_0^t \varphi(s) ds.$$

$\Phi$  is said to be an N-function if  $\varphi(0) = 0$  and  $\lim_{t \rightarrow \infty} \varphi(t) = \infty$ .

According to the definition above,  $\Phi(t) = t$  is not an N-function.

**Definition 47.11.** Let  $\Phi$  be an N-function. Then the conjugate function of  $\Phi$  is defined by

$$(47.19) \quad \Phi^*(t) := \int_0^t \varphi^*(s) ds.$$

**Proposition 47.12.** *If  $\Phi$  is an N-function, then so is  $\Phi^*$ .*

*Proof.* Note that  $\varphi^*$  is finite because  $\lim_{t \rightarrow \infty} \varphi(t) = \infty$ . As we have verified in Lemma 47.1,  $\varphi^*$  is right-continuous and increasing. Since  $\varphi^*(0) = 0$  and  $\lim_{t \rightarrow \infty} \varphi^*(t) = \infty$  from the definition of  $\varphi^*$ , the canonical representation of  $\Phi^*$  satisfies the condition in Definition 47.11.  $\square$

**Theorem 47.13.** *Let  $\Phi$  be an  $N$ -function and  $\Phi^*$  its conjugate. Then*

$$(47.20) \quad \Phi^*(t) = \sup\{st - \Phi(s) : s \in [0, \infty)\}.$$

Given a function  $\Psi : (0, \infty) \rightarrow \mathbb{R}$ , its Fenchel-Legendre transform is given by

$$(47.21) \quad t \in (0, \infty) \mapsto \sup\{st - \Psi(s) : s \in [0, \infty)\}.$$

Hence, Theorem 47.13 asserts that  $\Phi^*$  is the Fenchel-Legendre transform of  $\Phi$ .

*Proof.* Fix  $s, t \geq 0$ . We are going to show

$$(47.22) \quad \Phi(t) + \Phi^*(s) \geq st$$

for all  $s, t \geq 0$  and that for some  $t$  we have the equality. If  $st = 0$ , then the inequality is trivial. Therefore, we may assume  $s, t > 0$ . Let  $v > 0$  and  $w = \varphi^*(v)$ . Since  $\varphi(0) = 0$  and  $\varphi$  is increasing, we have

$$(47.23) \quad \varphi^{-1}([0, v]) = [0, w] \text{ or } \varphi^{-1}([0, v]) = [0, w].$$

Therefore  $w = |\varphi^{-1}([0, v])| = \int_0^\infty \chi_{\varphi^{-1}([0, v])}(x) dx$ , that is,

$$(47.24) \quad \varphi^*(v) = \int_0^\infty \chi_{\varphi^{-1}([0, v])}(x) dx$$

for every  $v > 0$ . Inserting this equality we have

$$(47.25) \quad \Phi^*(s) = \int_0^s \left( \int_0^\infty \chi_{\varphi^{-1}([0, v])}(x) dx \right) dv = \iint_{[0, \infty)^2} \chi_{\{(x, v) : 0 \leq v < s, \varphi(x) \leq v\}}(x, v) dv dx.$$

Meanwhile

$$(47.26) \quad \Phi(t) = \int_0^t \varphi(x) dx = \iint_{[0, \infty)^2} \chi_{\{(x, v) : 0 \leq x < t, 0 \leq v < \varphi(x)\}}(x, v) dv dx.$$

Finally observe that

$$(47.27) \quad [0, t] \times [0, s] \subset \{(x, v) : 0 \leq v < s, \varphi(x) \leq v\} \cup \{(x, v) : 0 \leq x < t, 0 \leq v < \varphi(x)\}.$$

Therefore we conclude

$$(47.28) \quad \Phi^*(s) + \Phi(t) \geq \int_0^\infty \int_0^\infty \chi_{[0, t] \times [0, s]}(x, v) dx dv = st.$$

To show that equality holds for some  $t$ , we set  $t = \varphi^*(s)$ . We shall claim the two sets in (47.27) coincide. Suppose that

$$(47.29) \quad (x, v) \in \{(x, v) : 0 \leq v < s, \varphi(x) \leq v\} \cup \{(x, v) : 0 \leq x < t, 0 \leq v < \varphi(x)\}.$$

Let  $0 < v < s$  and  $\varphi(x) \leq v$ . Then  $t = \varphi^*(s) \geq \varphi^*(v) \geq x$ . Let  $0 \leq x < t$  and  $0 \leq v < \varphi(x)$ . Since  $x < t = \varphi^*(s)$ , we have  $\varphi(x) \leq s$ . Therefore  $v \leq s$ . This implies (47.27) is actually an equality. In this case carrying out the same calculation as before, we conclude that  $\Phi^*(s) + \Phi(t) = st$ .  $\square$

**Proposition 47.14.** *Let  $\varphi$  be an  $N$ -function and  $a, b > 0$ . Set  $\psi(t) = a\varphi(bt)$ . Then  $\psi$  is an  $N$ -function as well and the conjugate is given by  $\psi^*(t) = a\varphi^*(t/ab)$ .*

**Theorem 47.15.** *Let  $\Phi$  be an  $N$ -function. Then  $(\Phi^*)^* = \Phi$ .*

*Proof.* This follows from the fact that  $(\varphi^*)^* = \varphi$ .  $\square$

**Theorem 47.16** (Young's inequality). *Let  $\Phi$  be an  $N$ -function and denote by  $\Phi^*$  its conjugate. Then*

$$(47.30) \quad st \leq \Phi(t) + \Phi^*(s)$$

for all  $s, t \geq 0$ . If we denote by  $\varphi$  the right derivative of  $\Phi$ , then

$$(47.31) \quad t\varphi(t) = \Phi(t) + \Phi^*(\varphi(t)).$$

*Proof.* We have only to re-examine the proof of Theorem 47.13. □

**Lemma 47.17.** *Let  $\Phi$  be an  $N$ -function and  $\Phi^*$  its conjugate. Then we have the following:*

- (1) *The inequality  $\Phi^*\left(\frac{\Phi(t)}{t}\right) \leq \Phi(t) \leq \Phi^*\left(\frac{2\Phi(t)}{t}\right)$  holds for all  $t > 0$ .*  
 (2) *For all  $t, s > 0$ , we have*

$$s \leq \frac{\varphi(t)}{t} \implies \frac{\varphi^*(s)}{s} \leq t$$

$$\frac{\varphi(t)}{t} \leq s \implies t \leq \frac{\varphi^*(2s)}{s}.$$

- (3) *Let  $s, t, \lambda > 0$ . If  $\varphi(t) = \varphi^*(s) = \lambda$ , then we have  $\lambda \leq ts \leq 2\lambda$ .*

- (1). Recall that  $\Phi(t)/t$  is increasing with respect to  $t$ .

$$\Phi^*\left(\frac{\Phi(t)}{t}\right) = \sup_{0 < s < t} s \left(\frac{\Phi(t)}{t} - \frac{\Phi(s)}{s}\right) \leq t \sup_{0 < s < t} \left(\frac{\Phi(t)}{t} - \frac{\Phi(s)}{s}\right) = \Phi(t).$$

The right inequality is easier to prove. We have only to use Theorem 47.13.

$$\Phi^*\left(\frac{2\Phi(t)}{t}\right) = \sup_{s > 0} s \left(\frac{2\Phi(t)}{t} - \frac{\Phi(s)}{s}\right) \leq \Phi(t).$$

Thus, (1) is proved. □

- (2). We use Theorem 47.13. Let  $s \leq \frac{\Phi(t)}{t}$ . Since  $\Phi(t)/t$  is increasing with respect to  $t$ , we have

$$(47.32) \quad \frac{\Phi^*(s)}{s} = \sup_{u > 0} \left(u - \frac{\Phi(u)}{s}\right) \leq t \sup_{u > 0} \left(\frac{u}{t} - \frac{\Phi(u)}{\Phi(t)}\right) \leq t \sup_{0 < u < t} \left(\frac{u}{t} - \frac{\Phi(u)}{\Phi(t)}\right) \leq t.$$

Suppose instead that  $s \geq \frac{\Phi(t)}{t}$ . Then

$$(47.33) \quad \frac{\Phi^*(2s)}{s} = \sup_{u > 0} \left(2u - \frac{\Phi(u)}{s}\right) \geq 2t - \frac{\Phi(t)}{s} \geq t.$$

Thus, the proof of (2) is complete. □

- (3). It follows from the Young inequality that  $st \leq \Phi(s) + \Phi(t) = 2\lambda$ . If  $\varphi(t) \leq s$ , then we have

$$(47.34) \quad \lambda = \Phi(t) = \int_0^t \varphi(u) du \leq st.$$

If  $\varphi(t) > s$ , then  $\varphi^*(s) \leq t$ . Therefore

$$(47.35) \quad \lambda = \Phi^*(s) = \int_0^s \varphi^*(u) du \leq st.$$

Thus, (3) is established. □



$\Delta_2$ -condition and  $\nabla_2$ -condition. In the next section we are going to generalize the  $L^p(\mathbb{R}^d)$ -spaces. Our generalization here is oriented to the parameter  $p$ . We can say that  $L^p(\mathbb{R}^d)$  is a function space based on the function  $\Phi(t) = t^p$ . The Orlicz space  $L^\Phi$  is a generalization such the function  $\Phi$  is an N-function. In this preparatory paragraph, we are going to introduce some condition under which the maximal operator  $M$  and singular integral operators, such as the Hilbert transform  $H$ , are bounded on  $L^\Phi$ . The function  $\Phi(t) = t, t \in \mathbb{R}$  is an N-function. However, in this case,  $L^\Phi(\mathbb{R})$  being the  $L^1(\mathbb{R})$ -space, the maximal operator and the singular operators cannot be  $L^\Phi(\mathbb{R})$ -bounded. Thus, we need some conditions to exclude such functions. Now we are going to formulate the conditions of the N-functions  $\Phi$  under which the maximal operator and the singular integral operators are  $L^\Phi$ -bounded.

**Definition 47.18.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be an N-function.

- (1) The function  $\Phi$  is said to satisfy  $\Delta_2$ -condition, if there exists a constant  $\mu > 1$  such that

$$(47.36) \quad \Phi(2t) \leq \mu \Phi(t), t > 0.$$

In this case one writes  $\Phi \in \Delta_2$ .

- (2) The function  $\Phi$  is said to satisfy  $\nabla_2$ -condition, if  $\Phi^* \in \Delta_2$ . In this case one writes  $\Phi \in \nabla_2$ .

**Example 47.19.** It is important to note that  $t \notin \nabla_2$ .

**Proposition 47.20.** Let  $\Phi$  be an N-function.

- (1) The function  $\Phi \in \nabla_2$  if and only if there exists a constant  $A > 1$  such that  $\Phi(At) \geq 2A\Phi(t)$ .
- (2) If  $\Phi \in \nabla_2$ , then there exist  $\mu > 1$  and  $\varepsilon > 0$  such that

$$(47.37) \quad \Phi(ut) \geq \mu^{-1}u^{1+\varepsilon}\Phi(t), \Phi(vt) \leq \mu v^{1+\varepsilon}\Phi(t)$$

whenever  $0 < v \leq 1 \leq u$  and  $t > 0$ .

*Proof.* As for the first assertion we note that

$$(47.38) \quad \Phi(A\cdot)^* = \Phi^*(\cdot/A), 2A\Phi(\cdot)^* = 2A\Phi(\cdot/2A).$$

Therefore it follows that

$$(47.39) \quad \Phi(2t) \geq 2A\Phi(t) \iff \Phi^*(s/A) \leq 2A\Phi^*(s/2A) \iff \Phi^*(2s) \leq 2A\Phi^*(s).$$

The second assertion can be obtained by induction. Indeed, we have

$$(47.40) \quad \varphi(A^k t) \geq (2A)^k \varphi(t)$$

for all  $k$ . Therefore if we set  $\varepsilon = \log_2 A$ , then

$$(47.41) \quad \varphi(A^k t) \geq A^{k(1+\varepsilon)} \varphi(t)$$

follows. It remains to pass to the continuous valuable from the above formula. □

**Proposition 47.21.** Let  $\Phi$  be an N-function with canonical representation

$$(47.42) \quad \Phi(t) = \int_0^t \varphi(s) ds$$

for  $t \geq 0$ .

- (1) Assume that  $\Phi \in \Delta_2$ . More precisely  $\Phi(2t) \leq A\Phi(t)$  for some  $A \geq 2$ . Set  $\beta = \log_2 A$ . If  $p > \beta + 1$ , then

$$(47.43) \quad \int_t^\infty \frac{\varphi(s)}{s^p} ds \lesssim \frac{\Phi(t)}{t^p}$$

for all  $t > 0$ .

- (2) Assume that  $\Phi \in \nabla_2$ . Then

$$(47.44) \quad \int_0^t \frac{\varphi(s)}{s} ds \lesssim \frac{\Phi(t)}{t}.$$

*Proof.* (1) If we carry out integration by parts we obtain

$$\int_t^\infty \frac{\varphi(s)}{s^p} ds = \int_t^\infty \frac{\Phi'(s)}{s^p} ds = \lim_{R \rightarrow \infty} \left( \left[ \frac{\Phi(s)}{s^p} \right]_t^R + p \int_t^R \frac{\Phi(s)}{s^p} ds \right).$$

Now we invoke an estimate  $\Phi(s) \lesssim \left(\frac{s}{t}\right)^\beta \Phi(t)$  for  $s \geq t$ . From this estimate we obtain

$$(47.45) \quad \lim_{R \rightarrow \infty} \Phi(R)R^p = 0, \quad \int_t^\infty \frac{\Phi(s)}{s^p} ds \lesssim \frac{\Phi(t)}{t^p}.$$

Therefore, (1) follows.

The estimate of (2) is similar. First, we carry out integration by parts.

$$(47.46) \quad \int_0^t \frac{\varphi(s)}{s} ds = \left[ \frac{\Phi(s)}{s} \right]_0^t + \int_0^t \frac{\Phi(s)}{s^2} ds.$$

Now we use  $\Phi(A s) \geq 2A\Phi(s)$ , which yields

$$\int_0^t \frac{\Phi(s)}{s^2} ds = \sum_{j=0}^{\infty} \int_{A^{-j-1}t}^{A^{-j}t} \frac{\Phi(s)}{s^2} ds \lesssim \frac{1}{t} \sum_{j=0}^{\infty} A^j \Phi(A^{-j}t) \lesssim \frac{1}{t} \sum_{j=0}^{\infty} A^j (2A)^{-j} \Phi(t) \lesssim \frac{\Phi(t)}{t}.$$

Inserting this estimate, we obtain

$$(47.47) \quad \int_0^t \frac{\varphi(s)}{s} ds \leq \frac{\Phi(t)}{t} + \int_0^t \frac{\Phi(s)}{s^2} ds \lesssim \frac{\Phi(t)}{t}.$$

Therefore (2) is established.  $\square$

## 48. ORLICZ SPACES

We now turn to Orlicz spaces, which are function spaces we encounter first in this book. As a review of the previous section, in particular Theorem 10.12, let us take up Orlicz spaces and prove its completeness.

### 48.1. Definition.

As an application of Theorem 10.12, we investigate Orlicz spaces. Let  $(X, \mu)$  be a measure space.

**Definition 48.1.** Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be an N-function. Then define

$$(48.1) \quad \|f\|_{L^\Phi(X)} := \inf \left\{ \lambda > 0 : \int_X \Phi \left( \frac{|f(x)|}{\lambda} \right) d\mu(x) \leq 1 \right\}.$$

The space  $L^\Phi(X)$  is the set of  $\mu$ -measurable functions normed by  $\|\cdot\|_{L^\Phi(X)}$ .

*Exercise 242.* What is  $L^\Phi(X)$  when  $\Phi(t) = \infty \chi_{(1, \infty)}$ ?

The definition of the Orlicz norm was, of course, in connection with Orlicz. However, he did not give the definition in the form as above. It seems that Nakano and Luxemburg originally considered such a norm in connection with the Minkowski functional (see Theorem 32.12). The references are [39, 45, 46]. Orlicz gave a different definition of Orlicz norms and later it turned out that the definitions are essentially identical.

**Example 48.2.** We are interested, in particular, in the case when  $\Phi(x) = \exp(x^\alpha) - 1$  for  $1 \leq \alpha < \infty$ . In this case we write  $L^\Phi = \exp(L^\alpha)$ .

Here we shall prove that  $L^\Phi(X)$  is a Banach space. In this section we do not go into the details. In Chapter 22, we make a deeper observation.

**Theorem 48.3.** *Let  $\Phi : [0, \infty) \rightarrow [0, \infty)$  be an  $N$ -function. Then  $L^\Phi(X)$  is a Banach space with its norm.*

*Proof.* Proof of the triangle inequality Here we content ourselves with showing the triangle inequality which is least trivial. We are now going to show that

$$(48.2) \quad \|f + g\|_{L^\Phi(X)} \leq \|f\|_{L^\Phi(X)} + \|g\|_{L^\Phi(X)}.$$

Note that (4.3) is a special case of (48.2) with  $\Phi(t) = t^p$ .

Let  $\varepsilon > 0$  be taken arbitrarily. Then (48.2) can be reduced to showing

$$(48.3) \quad \|f + g\|_{L^\Phi(X)} \leq \|f\|_{L^\Phi(X)} + \|g\|_{L^\Phi(X)} + 2\varepsilon.$$

From the definition of the set defining the norm  $\|f\|_{L^\Phi(X)}$ , we deduce

$$(48.4) \quad \|f\|_{L^\Phi(X)} + \varepsilon \in \left\{ \lambda > 0 : \int_X \Phi\left(\frac{|f|}{\lambda}\right) d\mu \leq 1 \right\},$$

that is, we have

$$(48.5) \quad \int_X \Phi\left(\frac{|f|}{\|f\|_{L^\Phi(X)} + \varepsilon}\right) d\mu \leq 1.$$

The same can be said about  $g$ ,

$$(48.6) \quad \int_X \Phi\left(\frac{|g|}{\|g\|_{L^\Phi(X)} + \varepsilon}\right) d\mu \leq 1.$$

Let us set  $\alpha = \|f\|_{L^\Phi(X)} + \varepsilon$ ,  $\beta = \|g\|_{L^\Phi(X)} + \varepsilon$  and  $\theta = \frac{\beta}{\alpha + \beta}$ .

Since  $\Phi$  is convex, we have

$$\Phi\left(\frac{|f(x) + g(x)|}{\alpha + \beta}\right) = \Phi\left((1 - \theta)\frac{|f(x)|}{\alpha} + \theta\frac{|g(x)|}{\beta}\right) \leq (1 - \theta)\Phi\left(\frac{|f(x)|}{\alpha}\right) + \theta\Phi\left(\frac{|g(x)|}{\beta}\right).$$

Integrating this inequality over  $X$ , we obtain

$$(48.7) \quad \int_X \Phi\left(\frac{|f + g|}{\|f\|_{L^\Phi(X)} + \|g\|_{L^\Phi(X)} + 2\varepsilon}\right) d\mu = \int_X \Phi\left(\frac{|f + g|}{\alpha + \beta}\right) d\mu \leq 1,$$

which is equivalent to (48.3). □

*Exercise 243.* Show that  $L^\Phi(X)$  is a Banach space. It might be helpful to use Theorem 10.12 to prove that any Cauchy sequence is convergent.

*Exercise 244.* Use Theorem 32.12 to give an alternative proof.

*Exercise 245.* Let  $\Psi : [0, \infty) \rightarrow [0, \infty)$  be an N-function. If  $Q$  is a cube and  $f$  is a measurable function, then define the  $\psi$ -average

$$(48.8) \quad \|f\|_{\psi, Q} := \inf \left\{ \frac{1}{|Q|} \int_Q \psi \left( \frac{|f(x)|}{\lambda} \right) dx \leq 1 \right\}.$$

Correspondingly, define the maximal operator  $M_\psi$  by  $M_\psi f(x) = \sup_{x \in Q} \|f\|_{\psi, Q}$ . Show that the mapping  $f \mapsto M_\psi f$  is sublinear.

*Exercise 246.* Suppose that we are given a measurable set  $\Omega$  with  $|\Omega| = 1$ . Let  $\Phi_a(t) =: t \exp(t^a)$  with  $a > 0$  and set  $L_{\text{exp}, a}(\Omega) := L^\Phi(\Omega)$ . Also, let  $\Psi_{p, a}(t) := \{(t \log(C+t))^a\}^p$  and write  $L^p(\log L)_a(\Omega) := L^{\Psi_{p, a}}(\Omega)$ .

(1) Show that

$$\|f\|_{L_{\text{exp}, a}(\Omega)} \sim \sup_{0 < t < 1} \frac{f^*(t)}{(1 + |\log t|^a)}.$$

(2) Show that

$$\|f\|_{L_{\text{exp}, a}(\Omega)} \sim \sup_{q > 1} \frac{\|f\|_{L^q(\Omega)}}{q^a}$$

(3) Show that

$$\|f\|_{L^p(\log L)_a} \sim \left\| |f| \log^a \left( 3 + \frac{|f|}{\|f\|_{L^p(\Omega)}} \right) \right\|_{L^p(\Omega)}.$$

(4) Show that

$$\|f\|_{L^p(\log L)_a} \sim \|f^* \log^a(\cdot^{-1})\|_{L^p(0,1)}.$$

See the paper by Iwaniec and Verde in 1999 for the case when  $a = 1/p$  and the one by Krbeć and Schmeisser in 2011.

**48.2. Convergence theorems.** Like  $L^p(\mu)$ , we have the following convergence theorems.

**Theorem 48.4** (Monotone convergence theorem for Orlicz spaces). *Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of functions satisfying*

$$(48.9) \quad 0 \leq f_j(x) \leq f_{j+1}(x) \text{ for almost every } x \in X$$

for all  $j$ . Set  $\limsup_{j \rightarrow \infty} f_j(x) = f(x)$ . Then we have

$$(48.10) \quad \lim_{j \rightarrow \infty} \|f_j\|_{L^\Phi(X)} = \|f\|_{L^\Phi(X)}.$$

*Proof.* By the monotonicity of the norm we have

$$(48.11) \quad \|f_j\|_{L^\Phi(X)} \leq \|f\|_{L^\Phi(X)}$$

for all  $j \in \mathbb{N}$ . For the proof of the theorem, we may assume that  $f \neq 0$ . Pick a constant  $M$  smaller than  $\|f\|_{L^\Phi(X)}$  arbitrarily. Then we have

$$(48.12) \quad \int_X \Phi \left( \frac{|f|}{M} \right) d\mu > 1$$

from the definition of the norm. By virtue of the usual monotone convergence theorem, we have

$$(48.13) \quad \int_X \Phi \left( \frac{|f_j|}{M} \right) d\mu > 1,$$

provided  $j$  is sufficiently large. Therefore, it follows that

$$(48.14) \quad \|f_j\|_{L^\Phi(X)} \geq M$$

for such  $j$ . Since  $M$  is a constant taken arbitrarily from  $(0, \|f\|_{L^\Phi(X)})$ , we obtain

$$(48.15) \quad \lim_{j \rightarrow \infty} \|f_j\|_{L^\Phi(X)} = \|f\|_{L^\Phi(X)},$$

which is the desired result. □

*Exercise 247.* Mimic the proof of the usual Fatou lemma, to prove that

$$(48.16) \quad \left\| \liminf_{j \rightarrow \infty} f_j \right\|_{L^\Phi(X)} \leq \liminf_{j \rightarrow \infty} \|f_j\|_{L^\Phi(X)},$$

whenever  $f_j \geq 0$  and  $\{f_j\}_{j \in \mathbb{N}} \subset L^\Phi(X)$ .

**Theorem 48.5.** *Let  $\Phi$  be an  $N$ -function. Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of  $L^\Phi(X)$  converging to  $f$  almost everywhere. Assume that there exists  $g \in L^\Phi(X)$  such that  $|f_j| \leq g$  a.e. for all  $j \in \mathbb{N}$ . Then we have*

$$(48.17) \quad \lim_{j \rightarrow \infty} \|f - f_j\|_{L^\Phi(X)} = 0.$$

*Proof.* Let  $\varepsilon \in (0, 1)$  be arbitrary. Then it suffices to show

$$(48.18) \quad \int_X \Phi \left( \frac{|f - f_j|}{\varepsilon} \right) d\mu \leq 1,$$

whenever  $j$  is sufficiently large. However, it is trivial. Indeed, the integrand is majorized by  $\frac{2}{\varepsilon} \Phi(|g|)$ . Therefore, we are in the position of using the Lebesgue convergence theorem to conclude (48.18) holds. □

As an application we can prove that  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^\Phi$ .

**Theorem 48.6.** *The space  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^\Phi(\mathbb{R}^d)$ .*

*Exercise 248.* Prove Theorem 48.6. Hint: First, by using Theorem 48.5 prove that any  $f \in L^\Phi$  can be approximated by bounded and compactly supported functions. Thus, we have only to approximate such functions. And then we use the mollification procedure.

### 48.3. Hölder inequality for Orlicz spaces.

In this section we consider the Hölder inequality for Orlicz spaces.

**Lemma 48.7.** *Let  $\Phi_1, \Phi_2, \Phi_3$  be  $N$ -functions such that*

$$(48.19) \quad \Phi_1^{-1}(x)\Phi_2^{-1}(x) \leq \Phi_3^{-1}(x).$$

*Then*

$$(48.20) \quad \Phi_3(xy) \leq \max(\Phi_1(x), \Phi_2(y)).$$

*In particular*

$$(48.21) \quad \Phi_3(xy) \leq \Phi_1(x) + \Phi_2(y).$$

*Proof.* By symmetry we may assume  $\Phi_1(x) \leq \Phi_2(y)$ . Then by (48.19) we obtain

$$(48.22) \quad xy = \Phi_1^{-1}(\Phi_1(x))\Phi_2^{-1}(\Phi_2(y)) \leq \Phi_1^{-1}(\Phi_2(y))\Phi_2^{-1}(\Phi_2(y)) \leq \Phi_3^{-1}(\Phi_2(y)).$$

This chain of inequality yields  $\Phi_3(xy) \leq \Phi_2(y)$ , proving the theorem. □

**Theorem 48.8** (Hölder inequality for Orlicz spaces). *Let  $\Phi_1, \Phi_2, \Phi_3$  be  $N$ -functions such that*

$$(48.23) \quad \Phi_1^{-1}(x)\Phi_2^{-1}(x) \leq \Phi_3^{-1}(x).$$

*Then*

$$(48.24) \quad \|fg\|_{L^{\Phi_3}} \leq 2\|f\|_{L^{\Phi_1}} \cdot \|g\|_{L^{\Phi_2}}, \quad f \in L^{\Phi_1}(\mu), \quad g \in L^{\Phi_2}(\mu).$$

*Proof.* For the proof we may assume that  $f, g \neq 0$ . Having verified that  $L^{\Phi_i}, i = 1, 2, 3$  is homogeneous, we may assume that  $\|f\|_{L^{\Phi_1}} = \|g\|_{L^{\Phi_2}} = 1$ . By

$$(48.25) \quad \Phi_3(xy) \leq \Phi_1(x) + \Phi_2(y),$$

we have

$$\int_X \Phi_3\left(\frac{|f_1 f_2|}{2}\right) d\mu \leq \frac{1}{2} \int_X \Phi_3(|f_1| \cdot |f_2|) d\mu \leq \frac{1}{2} \left( \int_X \Phi_1(|f_1|) d\mu + \int_X \Phi_2(|f_2|) d\mu \right) \leq 1,$$

proving  $\|f_1 f_2\|_{L^{\Phi_3}} \leq 2$ . □

**Theorem 48.9.** *Let  $\Phi$  be an  $N$ -function and  $\Phi^*$  its conjugate.*

(1) *Let  $f \in L^\Phi(X)$  and  $g \in L^{\Phi^*}(X)$ . Then*

$$(48.26) \quad \int_X |f(x)g(x)| dx \leq 2\|f\|_{L^\Phi(X)} \cdot \|g\|_{L^{\Phi^*}(X)}.$$

(2) *Assume that  $X$  is  $\sigma$ -finite and  $f \in L^\Phi(X)$ . Then*

$$(48.27) \quad \|f\|_{L^\Phi(X)} \leq \sup \left\{ \int_X |fg| d\mu : \|g\|_{L^{\Phi^*}(X)} \leq 1 \right\} \leq 2\|f\|_{L^\Phi(X)}.$$

*Proof.* For the proof of (1) and the right inequality of (2), we use the Young inequality with  $a = \frac{|f(x)|}{\|f\|_{L^\Phi(X)}}$  and  $b = \frac{|g(x)|}{\|g\|_{L^{\Phi^*}(X)}}$ .

$$(48.28) \quad \frac{|f(x)g(x)|}{\|f\|_{L^\Phi(X)} \cdot \|g\|_{L^{\Phi^*}(X)}} \leq \Phi\left(\frac{|f(x)|}{\|f\|_{L^\Phi(X)}}\right) + \Phi^*\left(\frac{|g(x)|}{\|g\|_{L^{\Phi^*}(X)}}\right).$$

Integrating this inequality over  $X$ , we obtain

$$(48.29) \quad \int_X \frac{|fg| d\mu}{\|f\|_{L^\Phi(X)} \cdot \|g\|_{L^{\Phi^*}(X)}} \leq \int_X \left\{ \Phi\left(\frac{|f|}{\|f\|_{L^\Phi(X)}}\right) + \Phi^*\left(\frac{|g|}{\|g\|_{L^{\Phi^*}(X)}}\right) \right\} d\mu \leq 2.$$

Arranging this inequality, we obtain the desired result.

For the proof of the left inequality of (2), by Theorem 48.5 we may assume that  $f \in L^\Phi(X)$ . Of course there is nothing to prove, if  $f = 0$ . Therefore, it can be assumed that  $f \in L^\Phi(X) \setminus \{0\}$ . Let  $\lambda$  be a number larger than  $\|f\|_{L^\Phi(X)}$ . Then define

$$(48.30) \quad g = \chi_{\{\lambda < |f| < \infty\}} \Phi\left(\frac{|f|}{\lambda}\right) \frac{\lambda}{|f|}.$$

Then by virtue of the inequality  $\Phi^*\left(\frac{\varphi(t)}{t}\right) \leq \varphi(t)$ , we obtain

$$(48.31) \quad \int_X \Phi^*(g) d\mu = \int_X \chi_{\{\lambda < |f| < \infty\}} \varphi^*\left(\varphi\left(\frac{|f|}{\lambda}\right) \frac{\lambda}{|f|}\right) d\mu \leq \int_X \varphi\left(\frac{|f|}{\lambda}\right) d\mu \leq 1.$$

As a consequence it follows that  $\|g\|_{L^{\Phi^*}} \leq 1$ . Furthermore, we have

$$(48.32) \quad \int_X |f \cdot g| d\mu = \lambda \int_X \Phi\left(\frac{|f|}{\lambda}\right) d\mu.$$

Thus,

$$(48.33) \quad \sup \left\{ \int_X |fg| d\mu : \|g\|_{L^{\Phi^*}(X)} \leq 1 \right\} \geq \lambda \int_X \Phi \left( \frac{|f|}{\lambda} \right) d\mu.$$

for every  $\lambda > \|f\|_{L^\Phi(X)}$ . Letting  $\lambda \rightarrow \|f\|_{L^\Phi(X)}$ , we obtain the desired inequality.  $\square$

**48.4. Maximal operators and singular integral operators.**

In this section we shall obtain a boundedness property of the maximal operators. Until the end of this chapter, we place ourselves in the setting of  $\mathbb{R}^d$  with the Lebesgue measure.

Refinement of the weak-(1, 1) boundedness of maximal operators.

This section is devoted to reminding the readers of the boundedness of the maximal operator by presenting a different proof. The result is somehow strengthened.

Here and below we denote by  $M$  the uncentered maximal operator given by

$$(48.34) \quad Mf(x) = \sup_{x \in Q \in \mathcal{Q}} m_Q(|f|).$$

**Theorem 48.10.** *For all measurable functions  $f$  we have*

$$(48.35) \quad |\{x \in \mathbb{R}^d : Mf(x) > \lambda\}| \leq \frac{5^d}{\lambda} \int_{\{x \in \mathbb{R}^d : Mf(x) > \lambda\}} |f(x)| dx$$

*Proof.* In fact  $B_j$  satisfies that  $B_j \subset \{x \in \mathbb{R}^d : Mf(x) > \lambda\}$ . We only have to modify the proof slightly.  $\square$

We also have the following variant.

**Proposition 48.11.** *Let  $f \in L^1(\mathbb{R}^d)$  and  $\lambda > 0$ . Then we have*

$$(48.36) \quad |\{x \in \mathbb{R}^d : Mf(x) > 2\lambda\}| \lesssim \frac{1}{\lambda} \int_{\{x \in \mathbb{R}^d : |f| > \lambda\}} |f|.$$

*Proof.* By using the weak-(1, 1) boundedness of  $M$  we have

$$\begin{aligned} |\{x \in \mathbb{R}^d : Mf(x) > 2\lambda\}| &\leq |\{x \in \mathbb{R}^d : M[\chi_{\{y \in \mathbb{R}^d : |f(y)| > \lambda\}} \cdot f](x) > \lambda\}| \\ &\lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} \chi_{\{x \in \mathbb{R}^d : |f| > \lambda\}} |f(x)| dx = \frac{1}{\lambda} \int_{\{x \in \mathbb{R}^d : |f| > \lambda\}} |f|. \end{aligned}$$

This is the desired result.  $\square$

We now prove the converse of the inequality above called the sunrise lemma.

**Theorem 48.12.** *Suppose that  $f$  is a locally integrable function. Then we have*

$$(48.37) \quad |\{x \in \mathbb{R}^d : Mf(x) > \lambda\}| \geq \frac{1}{\lambda} \int_{\{x \in \mathbb{R}^d : |f(x)| > \lambda\}} |f(x)| dx.$$

*Proof.* For the proof we may assume that  $f \in L^1(\mathbb{R}^d)$  by using monotone convergence theorem.

We form the CZ-decomposition of  $f$  at level  $\lambda$ . Then we will obtain a disjoint family of cube  $Q_j$  such that

$$(48.38) \quad \lambda \leq m_{Q_j}(|f|) \leq 2^d \lambda.$$

Then as we have observed, it follows that

$$(48.39) \quad \{x \in \mathbb{R}^d : |f| > \lambda\} \subset \{x \in \mathbb{R}^d : Mf > \lambda\} = \bigcup Q_j.$$

Then we have

$$\frac{1}{\lambda} \int_{\{y \in \mathbb{R}^d : |f(y)| > \lambda\}} |f(x)| dx \leq \frac{1}{\lambda} \int_{\bigcup Q_j} |f(x)| dx \leq \sum 2^d |Q_j| = |\{x \in \mathbb{R}^d : Mf > \lambda\}|.$$

Thus we have the desired result. □

Orlicz-boundedness of maximal operator. With this preparatory observations in mind, let us prove the boundedness of the maximal operator on Orlicz spaces.

**Theorem 48.13.** *The maximal operator  $M$  is bounded on  $L^\Phi$ , provided  $\Phi \in \nabla_2$ .*

*Proof.* Let  $\Lambda > 0$  and  $f \in L^\Phi \setminus \{0\}$ . Then we have

$$\int_{\mathbb{R}^d} \Phi\left(\frac{Mf}{\Lambda}\right) = \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{\lambda}{\Lambda}\right) |\{Mf > \lambda\}| d\lambda = \frac{2}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) |\{Mf > 2\lambda\}| d\lambda$$

Now we invoke our new maximal inequality and change the order of integration.

$$\begin{aligned} \int_{\mathbb{R}^d} \Phi\left(\frac{Mf}{\Lambda}\right) &\lesssim \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) \left(\int_{\{|f|>\lambda\}} |f|\right) \frac{d\lambda}{\lambda} \\ &\lesssim \frac{1}{\Lambda} \int_{\mathbb{R}^d} |f(x)| \left(\int_0^{|f(x)|} \varphi\left(\frac{2\lambda}{\Lambda}\right) \frac{d\lambda}{\lambda}\right) dx \\ &\lesssim \frac{1}{\Lambda} \int_{\mathbb{R}^d} |f(x)| \left(\int_0^{2\lambda^{-1}|f(x)|} \varphi(\lambda) \frac{d\lambda}{\lambda}\right) dx. \end{aligned}$$

Now we invoke Proposition 47.21. From Proposition 47.21 we obtain

$$(48.40) \quad \left(\int_0^{2\lambda^{-1}|f(x)|} \varphi(\lambda) \frac{d\lambda}{\lambda}\right) \lesssim |f(x)|^{-1} \Lambda \Phi\left(\frac{2|f(x)|}{\Lambda}\right),$$

if  $f(x) \neq 0$ . Recall that  $k\Phi(t) \leq \Phi(kt)$  for  $k \geq 1$  and  $t > 0$ , assuming  $\Phi$  convex. Therefore, it follows that

$$(48.41) \quad \int_{\mathbb{R}^d} \Phi\left(\frac{Mf}{\Lambda}\right) \leq c_0 \int_{\mathbb{R}^d} \Phi\left(\frac{2|f|}{\Lambda}\right) \leq \int_{\mathbb{R}^d} \Phi\left(\frac{c_0|f|}{\Lambda}\right).$$

Here  $c_0$  is a constant we would like to shed light on. Choosing  $\Lambda = c_0 \|f\|_{L^\Phi}$ , we obtain

$$(48.42) \quad \int_{\mathbb{R}^d} \Phi\left(\frac{Mf}{\Lambda}\right) \leq 1.$$

This means

$$(48.43) \quad \|Mf\|_{L^\Phi} \leq \Lambda = c_0 \|f\|_{L^\Phi}$$

from the definition of the norm. □

Finally we prove the boundedness of singular integral operators.

**Theorem 48.14.** *Let  $T$  be a singular integral operator and assume  $\Phi \in \Delta_2 \cap \nabla_2$ . Then  $T$  is bounded on  $L^\Phi$ .*



*Proof.* It is the same as before that we use the distribution functions.

$$\int_{\mathbb{R}^d} \Phi\left(\frac{|Tf|}{\Lambda}\right) = \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{\lambda}{\Lambda}\right) |\{|Tf| > \lambda\}| \, d\lambda = \frac{2}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) |\{|Tf| > 2\lambda\}| \, d\lambda.$$

What is different from the estimate for the maximal operator is the point that  $T$  is not  $L^\infty$ -bounded. Let  $p > 1$  be sufficiently large. Then

$$|\{|Tf| > 2\lambda\}| \leq |\{|T(\chi_{\{|f|>\lambda\}} \cdot f)| > \lambda\}| + |\{|T(\chi_{\{|f|\leq\lambda\}} \cdot f)| > \lambda\}|$$

By the weak-(1, 1) boundedness and the  $L^p(\mathbb{R}^d)$ -boundedness of  $T$  gives us

$$\begin{aligned} |\{|T(\chi_{\{|f|>\lambda\}} \cdot f)| > \lambda\}| &\lesssim \frac{1}{\lambda} \int_{\{|f|>\lambda\}} |f| \\ |\{|T(\chi_{\{|f|\leq\lambda\}} \cdot f)| > \lambda\}| &\lesssim \frac{1}{\lambda^p} \int_{\{|f|\leq\lambda\}} |f|^p. \end{aligned}$$

The same calculation as we used for the maximal operator works for the first term to obtain

$$(48.44) \quad \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) |\{|T(\chi_{\{|f|>\lambda\}} \cdot f)| > \lambda\}| \, d\lambda \leq \int_{\mathbb{R}^d} \Phi\left(\frac{c|f|}{\Lambda}\right).$$

As for the second term a similar computation still works but we use the fact that  $\Phi \in \Delta_2$ .

$$\begin{aligned} \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) |\{|T(\chi_{\{|f|\leq\lambda\}} \cdot f)| > \lambda\}| \, d\lambda &\lesssim \frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) \left(\int_{\{|f|\leq\lambda\}} |f|^p\right) \frac{d\lambda}{\lambda^p} \\ &\lesssim \frac{1}{\Lambda} \int_{\mathbb{R}^d} |f(x)|^p \left(\int_{|f(x)|}^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) \frac{d\lambda}{\lambda^p}\right) dx \end{aligned}$$

Using Proposition 47.21 (1), we have

$$(48.45) \quad \begin{aligned} &\frac{1}{\Lambda} \int_0^\infty \varphi\left(\frac{2\lambda}{\Lambda}\right) |\{|T(\chi_{\{|f|\leq\lambda\}} \cdot f)| > \lambda\}| \, d\lambda \\ &\lesssim \int_{\mathbb{R}^d} \Phi\left(\frac{2|f|}{\Lambda}\right) \leq \int_{\mathbb{R}^d} \Phi\left(\frac{c|f|}{\Lambda}\right). \end{aligned}$$

Thus, putting together (48.44) and (48.45), we obtain

$$(48.46) \quad \int_{\mathbb{R}^d} \Phi\left(\frac{|Tf|}{\Lambda}\right) \leq \int_{\mathbb{R}^d} \Phi\left(\frac{c_0|f|}{\Lambda}\right).$$

Again we shall label the constant we want to distinguish from other less important constants. As before, if we set  $\Lambda = c_2 \|f\|_{L^\Phi}$ , then we obtain

$$(48.47) \quad \int_{\mathbb{R}^d} \Phi\left(\frac{|Tf|}{\Lambda}\right) \leq 1.$$

Hence the operator norm of  $T$  is less than  $c_2$ . □

## 49. REFERENCE FOR CHAPTER 22

### Section 46. Theorem 46.1

Theorem 46.2

Theorem 46.6

Lebegue proved Theorem 46.7 in [300].

Theorem 46.9

Exercise 237 was already obtained by J. L. W. V. Jensen and O. Hölder [258, 237]. Hölder proved “his” Hölder inequality as a corollary. He used a natural mathematical induction argument: It is trivial that the inequality holds for  $n = 1$ . Assuming that the inequality is true for  $n$ , he established that the inequality for  $n + 1$ .

Theorem 46.14

Section 47. Lemma 47.1 stems from the paper by Young [488]. A detailed description of  $N$ -functions can be found in [51, 35]. In [395], one can find inequalities in Lemma 47.17, which asserts that  $\frac{\Phi(t)}{t}$  and  $\frac{\Phi^*(t)}{t}$  are almost mutually inverse to each other. We refer to [330] for more information of  $N$ -functions.

Theorem 47.3

Theorem 47.7

Theorem 47.13

Theorem 47.15

Theorem 47.16

As for the  $\Delta_2$ -condition we refer to [273], for example.

Section 48. The root of Orlicz spaces, together with the properties such as Theorem 48.3, dates back to [379]. However, this was just a special case of the theory by Nakano and Luxemburg. Afterward Matuszewska investigated Orlicz spaces in more generalized setting [329].

Theorem 48.4

Theorem 48.5

Theorem 48.6

Theorem 48.8

Theorem 48.9

Theorem 48.10 was originally obtained by Wiener in 1939.

Theorem 48.12 is due to the work by Stein in 1969.

Kita obtained Theorem 48.13 in [268] while the Orlicz-boundedness of the fractional integral operator  $I_\alpha$  was investigated in [355, 356, 438, 439]. In particular in [355, 438] Y. Mizuta and T. Shimomura investigated the Taylor expansion of the integral kernel  $|x - y|^{\alpha-d}$ .

Bloom and Kerman investigated Orlicz spaces without the  $\Delta_2$  condition (see [100]).

The weighted version of Theorem 48.13 can be found in [267].

Theorem 48.14

We encounter Orlicz spaces as a limit case. For example we have disproved in Chapter 7 that  $Mf$  never integrable over  $\mathbb{R}^d$  unless  $f$  vanishes identically. However, we still have the local integrability if  $f \in L^2(\mathbb{R}^d)$  because  $Mf \in L^2(\mathbb{R}^d)$ . Let us consider the limit case of this. Stein proved the following theorem [455].

**Theorem 49.1.** *Suppose that  $f \in L^1(\mathbb{R}^d)$  is supported on a ball  $B$ . Then  $Mf$  is integrable over  $B$ , if and only if  $f \log |f| \in L^1(\mathbb{R}^d)$ .*

The proof is not so hard, if we use the weak-(1, 1) inequality. Stein proved the analogous for the Riesz transform as well in [455].

Another example of using Orlicz spaces to describe the boundedness of the operator is the work due to O’Neil (see [378]).

Let  $G$  be a bounded domain and  $\frac{1}{p} = 1 - \frac{\alpha}{d} > 0$ . Then

$$(49.1) \quad \left( \int_G I_\alpha f(x)^p \{\log(1 + I_\alpha f(x))\}^{qp-1} dx \right)^{\frac{1}{p}} \lesssim_q \int_G f(y) \{\log(e + f(y))\}^q dy.$$

Interpolation of Orlicz spaces are investigated in [224, 274, 401].

The idea of defining Orlicz spaces immediately leads us to defining the variable exponent Lebesgue spaces. Recall that we can define the  $L^p(\mathbb{R}^d)$ -norm by

$$(49.2) \quad \inf \left\{ \lambda > 0 : \int_{\mathbb{R}^d} \left( \frac{|f(x)|}{\lambda} \right)^p dx \leq 1 \right\}.$$

By variable exponent we mean that the power  $p$  is a function of  $x \in \mathbb{R}^d$ . That is, we are going to consider the function space such as

$$(49.3) \quad \int_{\mathbb{R}^d} |f(x)|^{p(x)} dx < \infty.$$

However, this is not a norm. Indeed the homogeneity fails. To overcome this shortcoming, we use (49.2).

$$(49.4) \quad \inf \left\{ \lambda > 0 : \int_{\mathbb{R}^d} \left( \frac{|f(x)|}{\lambda} \right)^{p(x)} dx \leq 1 \right\}.$$

L. Diening obtained the maximal inequality for this variable exponent space. For details we refer to [163]. T. Futamura, Y. Mizuta and T. Shimomura obtained the Sobolev embedding theorem for this variable exponent space [199]. As early as 1986, Sharapudinov shed light on variable Lebesgue spaces and obtained the Haar wavelet characterization [419]. We refer to [369] for variable Lebesgue sequence spaces.

Finally let us make a remark of Theorem 46.11. This theorem is obtained originally by Bari and Stechkin (see [90]). We refer to [367] as well.

The author is grateful to Professors A. Miyachi and E. Nakai for their helpful comments [505, 504].

**Part 23. Function spaces with one or two parameters**

In this section we consider function spaces appearing in harmonic analysis and partial differential equations. We have considered the  $L^p(\mathbb{R}^d)$ -spaces and Sobolev spaces. Furthermore, in the context of application we have dealt with Lorentz spaces. However, there are many other function spaces that measure integrability. In this chapter we are going to investigate function spaces reflecting such properties. As for the smoothness, we need more precise information. That is, we are going to obtain information of smoothness of fractional order. The aim of this chapter is to investigate function spaces reflecting such subtle information. Before we go into this vast chapter, let us describe what we will investigate here.

In Section 50 we define Hardy spaces  $H^p(\mathbb{R}^d)$  in particular with  $0 < p \leq 1$ , which is a natural extension of what we did in Chapter 9. There are several equivalent ways to define Hardy spaces. One is to use the atomic decomposition as we have done in Chapter 9. Let  $\psi \in \mathcal{S}(\mathbb{R}^d)$  such that  $\int_{\mathbb{R}^d} \psi(x) dx \neq 0$ . The other, which is new in this book, is to use the grand maximal function. Our new definition is

$$(49.5) \quad \|f\|_{H^p} := \left\| \sup_{t>0} \frac{1}{t^d} \left| \psi\left(\frac{\cdot}{t}\right) * f \right| \right\|_p.$$

Our first concern is that the definition of our new norm is independent of the choice of admissible  $\psi$ . As we shall show in Theorems 50.12, this norm is equivalent to

$$(49.6) \quad \left\| \sup_{\varphi \in \mathcal{F}_N, t>0} \frac{1}{t^d} \left| \varphi\left(\frac{\cdot}{t}\right) * f \right| \right\|_p.$$

Here we have set  $\mathcal{F}_N$  as the closed unit ball with respect to the seminorm  $p_N$  given by

$$(49.7) \quad p_N(\varphi) := \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^d} (1 + |x|^2)^{\frac{N}{2}} |\partial^\alpha \varphi(x)|, \varphi \in \mathcal{S}(\mathbb{R}^d).$$

Subsections 50.3 and 50.4 are devoted to the proof that our two definitions, the definition by atomic decomposition and by the grand maximal operator, are equivalent. Subsection 50.4 occupies vast part of this section. So, we summarize our results in Subsection 50.5.

50. HARDY SPACES  $H^p(\mathbb{R}^d)$  WITH  $0 < p \leq 1$

In this section we follow closely to [183] again.

Recall that  $H^1(\mathbb{R}^d)$  is the set of functions in  $L^1(\mathbb{R}^d)$  that can be written as

$$(50.1) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j,$$

where each  $a_j$  is an atom and  $\{\lambda_j\}_{j=1}^{\infty}$ . The definition of atoms will be reviewed just below in the parameterized setting. In this section we define the Hardy space  $H^p(\mathbb{R}^d)$  with  $0 < p \leq 1$  or even with  $0 < p < \infty$ , although we allude to Hardy spaces with  $1 < p < \infty$  only a little.

The most important observation about the decomposition or the definition by means of decomposition is that we separate the quantity and the quality. First, if we are given a function  $f$ , the consider

$$f = \sum_{j=1}^{\infty} \lambda_j a_j.$$

Of course, each  $a_j$  enjoys good property. This is the quality. The quality plays an important role when we consider the boundedness of operators, for example. Meanwhile to measure the norm, instead of  $a_j$ , we just have to consider the indicator functions similar to  $a_j$ . Since the indicator functions do not have any information other than size, measuring the norms eventually amount to the quantity.

**50.1. Definition by means of atomic decomposition.**

Now we present a definition of  $H^p(\mathbb{R}^d)$  by means of atomic decomposition, which is simpler than the characterization of the grand maximal operator.

**Definition 50.1.** Let  $0 < p \leq 1$  and  $L \geq L_0 := [d(p^{-1} - 1)]$ . An  $H^p(\mathbb{R}^d)$ -atom is a measurable function  $A$  so that

- (1)  $A$  is supported in a cube  $Q$ ,
- (2)  $\|A\|_2 \leq |Q|^{-\frac{1}{p} + \frac{1}{2}}$ ,
- (3)  $A$  satisfies the moment condition of order  $L$ , that is,  $\int_{\mathbb{R}^d} x^\alpha A(x) dx = 0$  for all  $\alpha$  with  $|\alpha| \leq L$ .

Here and below we write  $A \perp \mathcal{P}_L(\mathbb{R}^d)$ , if  $A$  satisfies the moment condition of order  $L$ .

Recall that when  $p = 1$ , we had been called it atom for short in Part 6.

**Definition 50.2.** The Hardy space  $H^p(\mathbb{R}^d)$  is the set of all Schwartz distributions  $f \in L^p(\mathbb{R}^d)$  that can be expressed as

$$(50.2) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j,$$

where the convergence takes place in  $\mathcal{S}'(\mathbb{R}^d)$ , each  $a_j$  is an  $H^p(\mathbb{R}^d)$ -atom and

$$(50.3) \quad \sum_{j=1}^{\infty} |\lambda_j|^p < \infty$$

For  $f \in H^p(\mathbb{R}^d)$ , define

$$(50.4) \quad \|f\|_{H^p} := \inf_{(50.2)} \left( \sum_{j=1}^{\infty} |\lambda_j|^p \right)^{\frac{1}{p}},$$

where inf runs over all the admissible expressions in (50.2).

Condition  $\|a\|_2 \leq |Q|^{-\frac{1}{p} + \frac{1}{2}}$  yields  $\|a\|_p \leq 1$  for any  $H^p(\mathbb{R}^d)$ -atom  $a$ . An immediate consequence of this together with (50.3) is the following.

**Proposition 50.3.** Let  $0 < p \leq 1$ .  $H^p(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$  in the sense of continuous embedding.

Before we investigate more, let us verify the convergence in  $\mathcal{S}'(\mathbb{R}^d)$ .

**Proposition 50.4.** Suppose that we are given a sequence of  $H^p(\mathbb{R}^d)$ -atoms  $\{a_j\}_{j \in \mathbb{N}}$ . Assume additionally that  $\{\lambda_j\}_{j \in \mathbb{N}}$  satisfies (50.3). Then  $\sum_{j=1}^{\infty} \lambda_j a_j$  converges in  $\mathcal{S}'(\mathbb{R}^d)$ .

*Proof.* We begin with a quantitative observation of atoms. Suppose that  $a \in L^2(\mathbb{R}^d)$  such that

$$(50.5) \quad \text{supp}(a) \subset Q, \|a\|_2 \leq |Q|^{\frac{1}{2} - \frac{1}{p}}$$

for some cube  $Q$  and that  $a$  satisfies the moment condition of order  $L$ .

Pick a test function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then the moment condition leads us to subtract the Taylor polynomial of order  $L$ :

$$(50.6) \quad \int_{\mathbb{R}^d} a \cdot \varphi = \int_{\mathbb{R}^d} a(x) \left( \varphi(x) - \sum_{|\alpha| \leq L} \frac{\partial^\alpha \varphi(c(Q))}{\alpha!} (x - c(Q))^\alpha \right) dx.$$

Now that  $\varphi(x) - \sum_{|\alpha| \leq L} \frac{\partial^\alpha \varphi(c(Q))}{\alpha!} (x - c(Q))^\alpha$  is a remainder term of the Taylor expansion that can be bounded by a constant multiple of  $p_{L+1}(\varphi)\ell(Q)^{L+1}$ , we obtain

$$(50.7) \quad \left| \int_{\mathbb{R}^d} a \cdot \varphi \right| \lesssim \ell(Q)^{L+1+d-\frac{d}{p}} p_{L+1}(\varphi).$$

Meanwhile, it is easy to see

$$(50.8) \quad \left| \int_{\mathbb{R}^d} a \cdot \varphi \right| \lesssim \ell(Q)^{\frac{d}{2} - \frac{d}{p}} p_d(\varphi)$$

from the fact that  $\|a\|_2 \leq |Q|^{\frac{1}{2} - \frac{1}{p}}$ . Now that  $L \geq L_0 = [d(1/p - 1)]$  and  $0 < p \leq 1$ , we can combine (50.7) and (50.8) and obtain a key estimate

$$(50.9) \quad \left| \int_{\mathbb{R}^d} a \cdot \varphi \right| \lesssim p_{d+L+1}(\varphi).$$

Now suppose that we are given a sequence of atoms  $\{a_j\}_{j \in \mathbb{N}}$  and  $\{\lambda_j\}_{j \in \mathbb{N}} \in \ell^p$ . Taking into account assumption  $0 < p \leq 1$  once more, we obtain

$$\sum_{j=1}^{\infty} \left| \lambda_j \int_{\mathbb{R}^d} a_j \cdot \varphi \right| \lesssim p_{d+L+1}(\varphi) \sum_{j=1}^{\infty} |\lambda_j| \lesssim p_{d+L+1}(\varphi) \left( \sum_{j=1}^{\infty} |\lambda_j|^p \right)^{\frac{1}{p}}$$

by (50.9), which shows that  $\sum_{j=1}^{\infty} \lambda_j a_j$  converges in  $\mathcal{S}'(\mathbb{R}^d)$ . □

### 50.2. Maximal characterization of the Hardy space $H^p(\mathbb{R}^d)$ .

Throughout this subsection, we use the following notation. Let  $j \in \mathbb{Z}$  and  $\tau \in \mathcal{S}(\mathbb{R}^d)$ . Then define

$$(50.10) \quad \tau_j(x) := 2^{jd} \tau(2^j x).$$

Furthermore, let  $0 < p < \infty$ . Here we do not pose a condition that  $p > 1$  because what we will obtain in this subsection is available even for  $p \geq 1$ .

**Definition 50.5.** Suppose that  $\psi \in \mathcal{S}(\mathbb{R}^d)$  is non-degenerate in the sense that

$$(50.11) \quad \int_{\mathbb{R}^d} \psi(x) dx \neq 0.$$

Let  $0 < p < \infty$ . Then define

$$\|f\|_{H^p}^* := \left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p$$

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

Our first task is to show that  $\|f\|_{H^p}^*$  does not essentially depend on  $\psi$ . Speaking precisely, we have to show the following.

**Theorem 50.6.** *Let  $0 < p < \infty$ . Assume that  $\psi, \psi' \in \mathcal{S}(\mathbb{R}^d)$  satisfy (50.11). Then*

$$(50.12) \quad \left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p \simeq \left\| \sup_{j \in \mathbb{Z}} |\psi'_j * f| \right\|_p$$

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

Reproducing formula. To investigate Hardy spaces, which was defined by means of the grand maximal operator, we need an approximation procedure called reproducing formula.

Define

$$\varphi(x) := \psi(x) - 2^{-n}\psi(2^{-1}x) = \psi(x) - \psi_{-1}(x)$$

for  $x \in \mathbb{R}^d$ .

**Lemma 50.7.** *Let  $\psi \in \mathcal{S}(\mathbb{R}^d)$  and  $L \in \mathbb{N}_0$  be given. Then there exists  $\tilde{\psi}, \tilde{\varphi} \in \mathcal{S}(\mathbb{R}^d)$  such that*

$$(50.13) \quad \tilde{\psi} * \psi + \sum_{j=1}^{\infty} \tilde{\varphi}_j * \varphi_j = \delta$$

in the topology of  $\mathcal{S}'(\mathbb{R}^d)$  and that  $\tilde{\varphi}$  has vanishing moment up to order  $L$ . If  $\psi \in C_c^\infty(\mathbb{R}^d)$ , then we can arrange even that both  $\tilde{\psi}$  and  $\tilde{\varphi}$  have compact support.

*Proof.* Let us start from the identity

$$(50.14) \quad \lim_{j \rightarrow \infty} \psi_j * \psi_j = \delta,$$

as is verified easily by the Fourier transform.

Define  $g := \psi * \psi$ ,  $G = g - g_{-1}$ . Then (50.14) can be rephrased as

$$(50.15) \quad g + \sum_{j=1}^{\infty} G_j = \lim_{j \rightarrow \infty} \psi_j * \psi_j = \delta.$$

Denote by  $*^L h$  the  $L$ -fold convolution of  $h$ . Then (50.15) yields

$$(50.16) \quad *^L \left( g + \sum_{j=1}^{\infty} G_j \right) = \lim_{N \rightarrow \infty} *^L \left( g + \sum_{j=1}^N G_j \right) = \delta.$$

Below it will be understood that

$$(50.17) \quad *^L \left( g + \sum_{j=1}^{\infty} G_j \right) = \lim_{N \rightarrow \infty} *^L \left( g + \sum_{j=1}^N G_j \right), *^L \left( \sum_{j=k}^{\infty} G_j \right) = \lim_{N \rightarrow \infty} *^L \left( \sum_{j=k}^N G_j \right)$$

for  $k \in \mathbb{N}$ .

Let us expand (50.16). Denote by  ${}_n C_k$  the binomial coefficient.

$$\begin{aligned}
 *^L \left( g + \sum_{j=1}^{\infty} G_j \right) &= \sum_{l=0}^L {}_L C_l *^l g *^{L-l} \left( \sum_{j=1}^{\infty} G_j \right) \\
 &= g * \left( \sum_{l=1}^L {}_L C_l *^l g *^{L-l-1} \left( \sum_{j=1}^{\infty} G_j \right) \right) + *^L \left( \sum_{j=1}^{\infty} G_j \right) \\
 &= g * h + G_1 * \left( \sum_{l=1}^L {}_L C_l *^{l-1} G_1 * \left( *^{L-l} \left( \sum_{j=2}^{\infty} G_j \right) \right) \right) + *^L \left( \sum_{j=2}^{\infty} G_j \right) \\
 &= g * h + G_1 * H_1 + *^L \left( \sum_{j=2}^{\infty} G_j \right),
 \end{aligned}$$

where we have defined  $h$  and  $H$  so that

$$\begin{aligned}
 h &= \sum_{l=1}^L {}_L C_l *^l g *^{L-l-1} \left( \sum_{j=1}^{\infty} G_j \right) \\
 H_1 &= \sum_{l=1}^L {}_L C_l *^{l-1} G_1 * \left( *^{L-l} \left( \sum_{j=2}^{\infty} G_j \right) \right).
 \end{aligned}$$

Due to the self-similarity, we can express  $*^L \left( \sum_{j=2}^{\infty} G_j \right)$  in terms of  $G$  and  $H$ :

$$(50.18) \quad *^L \left( \sum_{j=2}^{\infty} G_j \right) = G_2 * \left( \sum_{l=1}^L {}_L C_l *^{l-1} G_2 * \left( *^{L-l} \left( \sum_{j=3}^{\infty} G_j \right) \right) \right) = G_2 * H_2.$$

Repeating this procedure, we obtain

$$(50.19) \quad g * h + \sum_{j=1}^{\infty} G_j * H_j = \delta.$$

Now let us factorize  $G = \varphi * (g + g_{-1})$  and set  $\tilde{\psi} := \tilde{\varphi} * h$ ,  $\tilde{\varphi} := (g + g_{-1}) * H$ . Then (50.19) gives us that  $\tilde{\psi} * \psi + \sum_{j=1}^{\infty} \tilde{\varphi}_j * \varphi_j = \delta$ .

The moment condition posed on  $\tilde{\varphi}$  is verified easily (c.f. Exercise 250). We had not alluded to the compactness of the supports of  $\tilde{\psi}$  and  $\tilde{\varphi}$ . However, in view of our actual construction, it is clear. Hence, the proof is complete.  $\square$

For the proof of Theorem 50.6 we also use the following lemma.

**Lemma 50.8.** *Let  $N \in \mathbb{N}_0$  and suppose that  $f$  is a measurable function such that*

$$(50.20) \quad \int_{\mathbb{R}^d} |x|^{L+1} |f(x)| dx < \infty.$$

*Assume that  $f \in L^1(\mathbb{R}^d)$  satisfies the moment condition of order  $L \in \mathbb{N}_0$ :*

$$(50.21) \quad f \perp \mathcal{P}_L(\mathbb{R}^d)$$



for all  $\alpha \in \mathbb{N}_0^d$  with  $|\alpha| \leq L$ . Suppose further that  $g \in \mathcal{S}(\mathbb{R}^d)$ . Then

$$(50.22) \quad \sup_{x \in \mathbb{R}^d} |x^\alpha f * g(x)| \lesssim_\alpha \| |\cdot|^{L+1} f \|_1 \sup_{\beta \leq \alpha} \int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} \mathcal{F}g(\xi)| d\xi$$

where the implicit constant in  $\lesssim$  depends only on  $\alpha$  with length less than  $L$ .

*Proof.* By (50.20) we see that the partial derivatives of  $\mathcal{F}f$  up to order  $L + 1$  are all bounded and continuous. Meanwhile (50.21) yields that  $\partial^\alpha \mathcal{F}f(0) = 0$  for all  $\alpha \in \mathbb{N}_0^d$  with  $|\alpha| \leq L$ . As a consequence, we obtain

$$(50.23) \quad |\partial^\alpha \mathcal{F}f(\xi)| \lesssim |\xi|^{L+1-|\alpha|} \sup_{|\beta|=L+1} \|\partial^\beta \mathcal{F}f\|_\infty \lesssim |\xi|^{L+1-|\alpha|} \| |\cdot|^{L+1} f \|_1$$

for such  $\alpha$ .

By the Riemann-Lebesgue theorem, we obtain

$$(50.24) \quad \sup_{x \in \mathbb{R}^d} |x^\alpha f * g(x)| \lesssim \int_{\mathbb{R}^d} |\partial_\xi^\alpha [\mathcal{F}f(\xi) \mathcal{F}g(\xi)]| d\xi.$$

If we use the Leibnitz rule and insert (50.23), then we are led to

$$|\partial_\xi^\alpha [\mathcal{F}f(\xi) \mathcal{F}g(\xi)]| \lesssim \sup_{\beta \leq \alpha} |\partial_\xi^\beta \mathcal{F}f(\xi) \partial_\xi^{\alpha-\beta} \mathcal{F}g(\xi)| \lesssim \| |\cdot|^{L+1} f \|_1 \sup_{\beta \leq \alpha} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} \mathcal{F}g(\xi)|.$$

Integrating this over  $\mathbb{R}^d$ , then we obtain the desired result. □

**Corollary 50.9.** *Let  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$  and assume that  $\varphi$  satisfies the moment condition of order  $L$ . Then for every  $N \in \mathbb{N}$  we have*

$$(50.25) \quad |\varphi_j * \psi_k(x)| \lesssim 2^{(k-j)(L+1)+kd} \langle 2^k x \rangle^{-N}$$

for all  $x \in \mathbb{R}^d$  with  $c$  independent of  $j$  and  $k$ .

*Proof.* Let  $L$  be fixed. Then

$$\varphi_j * \psi_k(x) = (\varphi * \psi_{k-j})_j(x).$$

Therefore, the dilation allows us to assume that  $j = 0$ . In this case Lemma 50.8 yields

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} |x^\alpha \varphi * \psi_k(x)| &\lesssim \sup_{\beta \leq \alpha} \int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} [\mathcal{F}\psi(2^{-k}\xi)]| d\xi \\ &\lesssim \sup_{\beta \leq \alpha} 2^{-k|\alpha-\beta|} \int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} [\mathcal{F}\psi](2^{-k}\xi)| d\xi \\ &\lesssim 2^{k(L+1-|\alpha|+d)}. \end{aligned}$$

This gives us the desired result,  $\alpha$  being arbitrary. □

To formulate the grand maximal operator, we define a seminorm by

$$(50.26) \quad p_N(\zeta) := \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N |\partial^\alpha \zeta(x)|, \zeta \in \mathcal{S}(\mathbb{R}^d)$$

for  $N \in \mathbb{N}$ . Here  $\langle x \rangle = \sqrt{1 + |x|^2}$  for  $x \in \mathbb{R}^d$ . Note that  $\{p_N\}_{N \in \mathbb{N}}$  topologizes  $\mathcal{S}(\mathbb{R}^d)$ .

Let us define an auxiliary maximal operator  $\psi_j^* f$  by

$$\psi_j^* f(x) := \sup_{y \in \mathbb{R}^d} \langle 2^j y \rangle^{-Ar} |\psi_j * f(x - y)|, f \in \mathcal{S}'(\mathbb{R}^d)$$

where  $A$  and  $r$  are fixed constants  $r < \frac{1}{\max(1,p)}$ ,  $Ar > d$ .

As a preparatory step for the proof of Theorem 50.6 we shall establish the following, which is of importance of its own right.

**Theorem 50.10.** *Let  $0 < p \leq 1$ . Then*

$$(50.27) \quad \left\| \sup_{j \in \mathbb{Z}} \psi_j^* f \right\|_p \lesssim_p \left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p$$

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

Before the proof we remark that the reverse estimate  $\left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p \lesssim \left\| \sup_{j \in \mathbb{Z}} \psi_j^* f \right\|_p$  is trivial because of the pointwise estimate of two integrands.

*Proof.* First of all, we fix  $y$  and  $j$ . By (50.38) we obtain

$$\begin{aligned} & \langle 2^j y \rangle^{-A} |\varphi_j * f(x - y)| \\ & \leq \langle 2^j y \rangle^{-A} \left( |\varphi_j * \psi_j * \tilde{\psi}_j * f(x - y)| + \sum_{l=1}^{\infty} |\varphi_j * \tilde{\varphi}_{j+l} * \varphi_{j+l} * f(x - y)| \right) \\ & \leq \langle 2^j y \rangle^{-A} \left( |(\varphi * \tilde{\psi})_j * \psi_j * f(x - y)| + \sum_{l=1}^{\infty} |(\varphi * \tilde{\varphi}_l)_j * \varphi_{j+l} * f(x - y)| \right). \end{aligned}$$

Here  $\tilde{\varphi}_l$  is assumed to satisfy the moment condition of order  $L$ . Now we shall make use of Corollary 50.9. Assuming  $1 \ll A, L$ , we obtain

$$(50.28) \quad |\varphi_j * \tilde{\varphi}_{l+j}(x)| \leq 2^{-l(L+1)+jd} \langle 2^j y \rangle^{-A}.$$

If we invoke Proposition 13.4, we obtain

$$\begin{aligned} \langle 2^j y \rangle^{-A} |(\varphi * \tilde{\varphi}_l)_j * \varphi_{j+l} * f(x - y)| & \leq \int_{\mathbb{R}^d} \langle 2^j y \rangle^{-A} |(\varphi * \tilde{\varphi}_l)_j(z) \varphi_{j+l} * f(x - y - z)| dz \\ & \lesssim 2^{-l(L+1)+jd} \int_{\mathbb{R}^d} \langle 2^j y \rangle^{-A} \langle 2^j z \rangle^{-A} |\varphi_{j+l} * f(x - y - z)| dz \\ & \lesssim 2^{-l(L+1)+jd} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\varphi_{j+l} * f(x - z)| dz \\ & \lesssim 2^{-l(L+1)+(j+l)d} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\varphi_{j+l} * f(x - z)| dz \end{aligned}$$

A similar estimate for  $(\varphi * \tilde{\psi})_j(x)$  is also available. As a result, we obtain

$$(50.29) \quad \langle 2^j y \rangle^{-A} |\psi_j * f(x - y)| \lesssim \sum_{l=0}^{\infty} 2^{-l(L+1)+(j+l)d} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\psi_{j+l} * f(x - z)| dz.$$

Now let us set

$$(50.30) \quad \mathcal{N}f(x, j; A) := \sup_{k \geq j} 2^{-(j-k)(L+1)} \langle 2^j y \rangle^{-A} |\psi_k * f(x - y)|.$$

Then we obtain

$$\begin{aligned}
 \mathcal{N}f(x, j; A) &\lesssim \sup_{k \geq j} 2^{-(j-k)(L+1)} \sum_{l=0}^{\infty} 2^{-l(L+1)+(k+l)d} \int_{\mathbb{R}^d} \langle 2^k z \rangle^{-A} \langle 2^j y \rangle^{-A} |\psi_{k+l} * f(x-y-z)| dz \\
 &\lesssim \sup_{k \geq j} \sum_{l=0}^{\infty} 2^{-(j-k-l)(L+1)+(k+l)d} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\psi_{k+l} * f(x-z)| dz \\
 &\simeq \sup_{k \geq j} \sum_{l=k}^{\infty} 2^{-(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\psi_l * f(x-z)| dz \\
 &\simeq \sum_{l=j}^{\infty} 2^{-(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A} |\psi_l * f(x-z)| dz \\
 &\lesssim \mathcal{N}f(x, j; A)^{1-r} \left( \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz \right).
 \end{aligned}$$

As a result we obtain

(50.31)

$$\mathcal{N}f(x, j; A) \lesssim \mathcal{N}f(x, j; A)^{1-r} \left( \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz \right).$$

We would be very happy, if we were fortunate enough to have  $\mathcal{N}f(x, j; A) < \infty$ . If this happens, then we can readily divide  $\mathcal{N}f(x, j; A)$  and the proof is complete. However, it is, of course, not always the case. To overcome this difficulty, we have to be more tricky.

Since  $f \in \mathcal{S}'(\mathbb{R}^d)$ , there exists  $A_f$  depending on  $f$  such that  $\mathcal{N}f(x, j; A_f) < \infty$  and hence

$$(50.32) \quad \mathcal{N}f(x, j; A_f)^r \lesssim_f \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A_f r} |\psi_l * f(x-z)|^r dz$$

holds. Assume that

$$(50.33) \quad \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz < \infty$$

with  $A < A_f$ . Then we have

$$\begin{aligned}
 \mathcal{N}f(x, j; A_f)^r &\lesssim_f \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-A_f r} |\psi_l * f(x-z)|^r dz \\
 &\lesssim_f \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz < \infty
 \end{aligned}$$

by virtue of monotonicity of the right-hand side with respect to  $A$ . As a result

$$\begin{aligned}
 \mathcal{N}f(x, j; A)^r &\lesssim_f \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} \langle 2^j y \rangle^{-Ar} |\psi_l * f(x-y-z)|^r dz \\
 &\lesssim_f \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz < \infty.
 \end{aligned}$$

Going back to (50.31), we obtain

$$(50.34) \quad \mathcal{N}f(x, j; A)^r \lesssim \sum_{l=j}^{\infty} 2^{-r(j-l)(L+1)+ld} \int_{\mathbb{R}^d} \langle 2^j z \rangle^{-Ar} |\psi_l * f(x-z)|^r dz.$$

Using the powered maximal operator  $M^{(r)}f(x) = M[|f|^r](x)^{\frac{1}{r}}$  and (50.34), we have

$$(50.35) \quad \psi_j^* f(x) \leq \mathcal{N}f(x, j; A)^r \lesssim M^{(r)} \left[ \sup_{l \in \mathbb{Z}} |\psi_l * f| \right] (x).$$

If (50.33) fails with  $A < A_f$ , then we have (50.34) and (50.35) trivially.

If  $A \geq A_f$ , then we still have  $\mathcal{N}(x, j, A) < \infty$  and we have an estimate independent of  $f$  which readily yields (50.34) and (50.35).

Now, with (50.35) established, the proof of (50.27) is just a matter of applying the  $L^{\frac{p}{r}}$ -maximal inequality. □

The following theorem asserts more than Theorem 50.6, whose proof will be finally obtained. To formulate our result we define

$$(50.36) \quad \mathcal{F}_N := \{\varphi \in \mathcal{S}(\mathbb{R}^d) : p_N(\varphi) \leq 1\},$$

where  $p_N$  is given by (49.7).

**Theorem 50.11.** *Let  $0 < p < \infty$ . Assume  $\psi \in \mathcal{S}(\mathbb{R}^d)$  is non-degenerate in the sense that  $\int_{\mathbb{R}^d} \psi(x) dx \neq 0$ . Then there exist  $N \gg 1$  and  $c > 0$  such that*

$$(50.37) \quad \left\| \sup_{\zeta \in \mathcal{F}_N, j \in \mathbb{Z}} |\zeta_j * f| \right\|_p \lesssim \left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p$$

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

*Proof.* Let  $L \in \mathbb{N}$  be chosen sufficiently large. Then there exists  $\tilde{\psi}, \tilde{\varphi} \in \mathcal{S}(\mathbb{R}^d)$  such that  $\tilde{\varphi}$  has vanishing moment up to order  $L$  and that

$$(50.38) \quad \tilde{\psi} * \psi + \sum_{l=0}^{\infty} \tilde{\varphi}_l * \varphi_l = \delta$$

by virtue of Lemma 50.7.

In view of Theorem 50.10 we have only to show

$$(50.39) \quad \left\| \sup_{\zeta \in \mathcal{F}_N, j \in \mathbb{Z}} |\zeta_j * f| \right\|_p \lesssim \left\| \sup_{j \in \mathbb{Z}} \psi_j^* f \right\|_p.$$

Fix  $\zeta \in \mathcal{F}_N$  and  $j \in \mathbb{Z}$ . Using (50.39), we obtain

$$(50.40) \quad |\zeta_j * f(x)| \leq |\zeta_j * \tilde{\psi}_j * \psi_j * f(x)| + \sum_{l=0}^{\infty} |\zeta_j * \tilde{\varphi}_{j+l} * \varphi_{j+l} * f(x)|.$$

Another application of Lemma 50.8 yields

$$(50.41) \quad |\zeta_j * \tilde{\varphi}_{j+l} * \varphi_{j+l} * f(x)| \lesssim 2^{-(L+1-\frac{d}{r})l} (\psi_{j+l-1}^* f(x) + \psi_{j+l}^* f(x))$$

A similar estimate is valid for  $|\zeta_j * \tilde{\psi}_j * \psi_j * f(x)|$ . If  $L$  is sufficiently large, say  $L + 1 - \frac{d}{r} > 0$ , (50.41) is summable over  $l \in \mathbb{N}_0$  and we have

$$(50.42) \quad |\zeta_j * f(x)| \lesssim \sup_{k \in \mathbb{Z}} \psi_k^* f(x).$$

From (50.42), we obtain

$$(50.43) \quad \left\| \sup_{\zeta \in \mathcal{F}_N, j \in \mathbb{Z}} |\zeta_j * f| \right\|_p \lesssim \left\| \sup_{j \in \mathbb{Z}} \psi_j^* f \right\|_p,$$

which is the desired result. □

Given  $f \in \mathcal{S}'(\mathbb{R}^d)$ , we define

$$(50.44) \quad \begin{aligned} \mathcal{M}_0 f(x) &:= \sup_{\varphi \in \mathcal{F}_N, j \in \mathbb{Z}} |\varphi_j * f(x)| \\ \mathcal{M} f(x) &:= \sup_{\varphi \in \mathcal{F}_N, t > 0} \left| \frac{1}{t^d} \varphi\left(\frac{\cdot}{t}\right) * f(x) \right|. \end{aligned}$$

These maximal operators are called the grand maximal operators. It is easy to see that

$$(50.45) \quad \mathcal{M} f(x) \simeq_{d,N} \mathcal{M}_0 f(x).$$

Let us summarize what we have proved.

**Theorem 50.12.** *Let  $\psi \in \mathcal{S}(\mathbb{R}^d)$  be a non-degenerate function in the sense that  $\int_{\mathbb{R}^d} \psi(x) dx \neq 0$ . Then the following equivalence holds*

$$(50.46) \quad \|\mathcal{M} f\|_p \simeq \left\| \sup_{j \in \mathbb{Z}} |\psi_j * f| \right\|_p, \quad f \in \mathcal{S}'(\mathbb{R}^d)$$

for every  $0 < p \leq 1$ .

*Exercise 249.* Let  $j, k \in \mathbb{Z}$  and  $\tau, \eta \in \mathcal{S}(\mathbb{R}^d)$ . Then show that

$$(50.47) \quad (\tau_j)_k = \tau_{j+k}, \quad \tau_j * \eta_j = (\tau * \eta)_j.$$

*Exercise 250.* Assume that  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$  satisfy

$$\varphi \perp \mathcal{P}_L(\mathbb{R}^d), \quad \psi \perp \mathcal{P}_M(\mathbb{R}^d)$$

for some  $L, M \in \mathbb{N}_0$ . Then show that

$$(50.48) \quad \varphi * \psi \perp \mathcal{P}_{L+M+1}.$$

### 50.3. Atomic decomposition vs. the grand maximal operator - I.

Let  $0 < p \leq 1$  throughout this subsection. Now we are to obtain

$$(50.49) \quad \|\mathcal{M} f\|_p \lesssim \|f\|_{H^p}$$

for all  $f \in \mathcal{S}'$ , where  $\mathcal{M}$  is the grand maximal operator defined by (50.44). Here let us assume  $f \in H^p(\mathbb{R}^d)$  throughout this subsection.

**Lemma 50.13.** *Suppose that a function  $A$  is an  $H^p(\mathbb{R}^d)$ -atom. Then  $\|\mathcal{M} A\|_p \leq c$ .*

*Proof.* Let  $Q$  be a cube such that

$$(50.50) \quad \text{supp}(A) \subset Q, \quad \|A\|_2 \leq |Q|^{\frac{1}{2} - \frac{1}{p}}.$$

A translation allows us to assume that  $Q$  is centered at the origin. Pick  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Then

$$\left| x^\alpha \frac{1}{t^d} \varphi\left(\frac{\cdot}{t}\right) * A(x) \right| \lesssim \| |\cdot|^{L+1} A \|_1 \cdot \sup_{\beta \leq \alpha} \int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} \mathcal{F}\varphi(t\xi)| d\xi \lesssim \frac{\ell(Q)^{L+1+d-\frac{d}{p}}}{t^{L+1+d-|\alpha|}},$$

because  $\int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} \mathcal{F}\varphi(t\xi)| d\xi = \int_{\mathbb{R}^d} |\xi|^{L+1-|\beta|} |\partial_\xi^{\alpha-\beta} [\mathcal{F}\varphi(t\xi)]| d\xi \simeq t^{-(L+1+d-|\alpha|)}$ . As a result, choosing  $\alpha$  so that  $|\alpha| = L + d + 1$ , we obtain

$$(50.51) \quad \mathcal{M} A(x) \lesssim \frac{\ell(Q)^{L+1+d-\frac{d}{p}}}{|x|^{L+1+d}}.$$

Meanwhile Proposition 13.4 yields  $\mathcal{M}A(x) \lesssim MA(x)$  for all  $x \in \mathbb{R}^d$ . Therefore,

$$(50.52) \quad \mathcal{M}A(x) \lesssim \chi_Q(x)MA(x) + \frac{\ell(Q)^{L+1+d-\frac{d}{p}}}{|x|^{L+1+d}}\chi_{\mathbb{R}^d \setminus Q}(x)$$

for all  $x \in \mathbb{R}^d$ . Integrating this over  $\mathbb{R}^d$ , we obtain

$$\|\mathcal{M}A\|_p \lesssim \|\chi_Q \cdot MA\|_p + 1 \lesssim |Q|^{\frac{1}{2}-\frac{1}{p}}\|\chi_Q \cdot MA\|_2 + 1 \lesssim |Q|^{\frac{1}{2}-\frac{1}{p}}\|A\|_2 + 1 \leq 1.$$

Thus, the proof is now complete. □

**Proposition 50.14.** *Let  $f \in H^p(\mathbb{R}^d)$ . Then*

$$(50.53) \quad \|\mathcal{M}f\|_p \lesssim_p \|f\|_{H^p}.$$

*Proof.* Since  $f \in H^p(\mathbb{R}^d)$ ,  $f$  admits the following decomposition :

$$(50.54) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j,$$

where  $a_j$  is an  $H^p(\mathbb{R}^d)$ -atom and

$$(50.55) \quad \sum_{j=1}^{\infty} |\lambda_j|^p \leq 2 \|f\|_{H^p}^p$$

Let us denote  $f_j = \sum_{k=1}^j \lambda_k a_k$ . Then we have

$$(50.56) \quad |\langle \check{\varphi}, f_{j_1} - f_{j_0} \rangle| \lesssim \inf_{|y| \leq 1} \mathcal{M}[f_{j_1} - f_{j_0}](y) \lesssim \|\mathcal{M}[f_{j_1} - f_{j_0}]\|_p \lesssim \|f_{j_1} - f_{j_0}\|_{H^p}.$$

Here we have set  $\check{\varphi}(x) := \varphi(-x)$ . Therefore, it follows that the limit  $\lim_{j \rightarrow \infty} \langle \check{\varphi}, f_j \rangle$  exists. The function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  being arbitrary, it follows that  $f$  converges in  $\mathcal{S}'(\mathbb{R}^d)$  as well. Once we prove this, we obtain

$$(50.57) \quad \|\mathcal{M}f\|_p \leq \liminf_{J \rightarrow \infty} \left\| \mathcal{M} \left[ \sum_{j=1}^J \lambda_j a_j \right] \right\|_p \lesssim \left( \sum_{j=1}^{\infty} |\lambda_j|^p \right)^{\frac{1}{p}} \lesssim \|f\|_{H^p}$$

by virtue of Lemma 50.13. This is the desired result. □

Let us summarize our observations up to now.

**Theorem 50.15.** *Let  $0 < p \leq 1$ . Then  $H^p(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$  in the sense of continuous embedding. Speaking precisely, we have*

$$(50.58) \quad |\langle f, \varphi \rangle| \lesssim \|f\|_{H^p} \cdot p_N(\varphi)$$

for some large  $N$ , which depends only on  $p$  and  $d$ . Furthermore, we have

$$(50.59) \quad \|\mathcal{M}f\|_p \lesssim \|f\|_{H^p}$$

for all  $f \in H^p(\mathbb{R}^d)$ .

*Proof.* To obtain a quantitative estimate (50.58), we have only to re-examine the proofs above. □

50.4. Atomic decomposition vs. the grand maximal operator - II.

Now we are going to establish the converse of Theorem 50.15. That is, if we assume that

$$(50.60) \quad \|\mathcal{M}f\|_p < \infty,$$

then we are to show  $f$  admits the decomposition as above.

Throughout this subsection we assume that  $f \in \mathcal{S}'(\mathbb{R}^d)$  satisfies (50.60) with  $0 < p \leq 1$ .

A set up. Let  $\lambda > 0$  be fixed. Then set

$$(50.61) \quad O := \{\mathcal{M}f > \lambda\},$$

which is open in  $\mathbb{R}^d$ . Form the Whitney decomposition of  $O$ . That is, we choose a disjoint collection  $\{Q_j\}_{j \in \mathbb{N}}$  so that  $\text{dist}(O^c, Q_j) \leq 10000\ell(Q_j)$  for every  $j \in \mathbb{N}$  and that

$$(50.62) \quad \chi_O \leq \sum_{j=1}^{\infty} \chi_{10Q_j} \leq \sum_{j=1}^{\infty} \chi_{1000Q_j} \leq M \chi_O$$

for some fixed constant  $M \in \mathbb{N}$ . Furthermore, let us assume

$$(50.63) \quad \sup_{j \in \mathbb{N}} \{l \in \mathbb{N} : 1000Q_j \cap 1000Q_l \neq \emptyset\} < \infty.$$

Let  $\zeta \in \mathcal{S}(\mathbb{R}^d)$  be a bump function satisfying

$$(50.64) \quad \chi_{[-10,10]^d} \leq \zeta \leq \chi_{[-100,100]^d}, \zeta(x) > 0, x \in (-100, 100)^d.$$

With this preparation in mind, we define

$$(50.65) \quad K(x) = \chi_{\mathbb{R}^d \setminus O}(x) + \sum_{k=1}^{\infty} \zeta_k(x)$$

and

$$(50.66) \quad \zeta_j(x) := \zeta\left(\frac{x - c(Q_j)}{\ell(Q_j)}\right), \eta_j(x) := \frac{\zeta_j(x)}{K(x)}$$

for  $j \in \mathbb{N}$ . Note that  $\text{supp}(\zeta_j) \subset O$ , which implies  $\eta_j \in C_c^\infty(\mathbb{R}^d)$ .

**Lemma 50.16.** *Let  $\alpha \in \mathbb{N}_0^d$  and assume that each function  $\zeta_j \in C_c^\infty(\mathbb{R}^d)$  satisfies the natural differential inequalities:*

$$(50.67) \quad |\partial^\alpha \eta_j(x)| \lesssim_\alpha \ell(Q_j)^{-|\alpha|} \chi_{10Q_j}(x).$$

Lemma 50.16 says that  $\eta$  behaves almost in the same manner as  $\zeta$ .

*Proof.* A simple induction argument suffices because

$$(50.68) \quad \sum_{k=1}^{\infty} \zeta_k(x) \gtrsim \chi_O(x), \sum_{k=1}^{\infty} |\partial^\alpha \zeta_k(x)| \lesssim_\alpha \text{dist}(x, \mathbb{R}^d \setminus O)^{-|\alpha|}$$

for all  $x \in O$ . □

Now let  $j \in \mathbb{N}$  be fixed. We also take  $L \in \mathbb{N}$  so that  $L \geq L_0 = [d(1/p - 1)]$ .

**Lemma 50.17.** *There exists uniquely a polynomial  $c_j$  of degree  $L$  such that*

$$(50.69) \quad \langle f - c_j, \eta_j \cdot q \rangle = 0$$

for all polynomials  $q$  of degree  $L$ .

*Proof.* This is just a matter of simple linear algebra. Let us set

$$(50.70) \quad c_j(x) := \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq L} \lambda_\alpha x^\alpha$$

and write condition (50.69) out in full. Then we obtain

$$(50.71) \quad \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq L} \lambda_\alpha \int_{\mathbb{R}^d} x^{\alpha+\beta} \eta_j(x) dx = \langle f, \eta_j \cdot x^\beta \rangle$$

for all  $\beta$  with length less than  $L$ . For the purpose of finding a unique solution of (50.71) it suffices to show  $A = \{a_{\alpha, \beta}\}_{|\alpha|, |\beta| \leq L}$  is non-degenerate, which is given by

$$(50.72) \quad a_{\alpha, \beta} := \int_{\mathbb{R}^d} x^\alpha x^\beta \eta_j(x) dx.$$

Let  $\{k_\alpha\}_{|\alpha| \leq L} \subset \mathbb{C}$  be a vector. Then a simple calculation shows

$$(50.73) \quad \sum_{|\alpha|, |\beta| \leq L} a_{\alpha, \beta} k_\alpha \overline{k_\beta} = \int_{\mathbb{R}^d} \left| \sum_{|\alpha| \leq L} k_\alpha x^\alpha \right|^2 \eta_j(x) dx.$$

Therefore, this quadratic form is not zero precisely when  $k_\alpha = 0$  for all  $\alpha$  with  $|\alpha| \leq L$ , which shows that  $A$  is non-degenerate.  $\square$

Let us define a Hilbert space  $\mathcal{H}_j$  of functions on  $100Q_j$  by

$$(50.74) \quad \|g\|_{\mathcal{H}_j} := \sqrt{\frac{1}{\|\eta_j\|_{100Q_j}} \int_{100Q_j} |g(x)|^2 \eta_j(x) dx}.$$

Let  $\mathcal{H}_{j,L}$  be the subspace of  $\mathcal{H}_j$  consisting of all polynomials of degree  $L$ . Note that the inner product is given by

$$(50.75) \quad \langle g_1, g_2 \rangle_{\mathcal{H}_j} = \frac{1}{\|\eta_j\|_{L^1(100Q_j)}} \int_{100Q_j} g_1(x) \overline{g_2(x)} \eta_j(x) dx.$$

**Lemma 50.18.** *Let  $\alpha \in \mathbb{N}_0^d$  be fixed. Let  $q \in \mathcal{H}_{j,L}$ . Then*

$$(50.76) \quad \|\partial^\alpha q\|_{\text{BC}(100Q_j)} \lesssim_{\alpha, L} \ell(Q_j)^{-|\alpha|} \|q\|_{\mathcal{H}_j},$$

where the implicit constant in  $\lesssim$  is independent of  $q$ .

*Proof.* The function  $\eta_j$  appearing implicitly in the right-hand side can be readily replaced by  $\zeta_j$  for the proof of (50.76). If  $\ell(Q) = 1$ , then this is clearly the case because  $\mathcal{H}_{j,L}$  is of finite dimensional. The passage to the general case is just a matter of scaling.  $\square$

**Proposition 50.19.** *Let  $\{e_{j,k}\}_{k \in K_j}$  be a CONS of  $\mathcal{H}_{j,L}$ , where  $K$  is a finite set. Then we have*

$$(50.77) \quad c_j(x) = \sum_{k \in K_j} \frac{\langle f, e_k \cdot \eta_j \rangle}{\|\eta_j\|_{L^1(100Q_j)}} \cdot e_{j,k}(x).$$

Below, given a polynomial  $P$ , we omit  $|_{100Q_j}$ , which shows the restriction to  $100Q_j$ , and we write  $P$  instead of  $P|_{100Q_j}$ .

*Proof.* Let us denote the right-hand side by  $d_j$ . Since  $c_j$  is characterized by the cancellation property, let us pick a polynomial  $q$  of degree  $L$ .



Noting that  $\langle \cdot, \cdot \rangle_{\mathcal{H}} = \frac{\langle \cdot, \cdot \rangle}{\int_{100Q_j} \eta_j}$ , we have

$$\begin{aligned} \langle f - d_j, \eta_j \cdot q \rangle &= \langle f, \eta_j \cdot q \rangle - \sum_{k \in K_j} \frac{\langle f, e_{j,k} \cdot \eta_j \rangle \cdot \langle e_{j,k}, \eta_j \cdot q \rangle}{\int_{100Q_j} \eta_j} \\ &= \langle f, \eta_j \cdot q \rangle - \left\langle f, \sum_{k \in K_j} \frac{\langle e_{j,k}, \eta_j \cdot q \rangle}{\int_{100Q_j} \eta_j} e_{j,k} \cdot \eta_j \right\rangle \\ &= \langle f, \eta_j \cdot q \rangle - \left\langle f, \left( \sum_{k \in K_j} \langle e_{j,k}, q \rangle_{\mathcal{H}_j} \cdot e_{j,k} \right) \cdot \eta_j \right\rangle \\ &= \langle f, \eta_j \cdot q \rangle - \langle f, \eta_j \cdot q \rangle = 0. \end{aligned}$$

This is exactly a condition that  $c_j(x)$  must satisfy and the proof is complete.  $\square$

Now we claim the following.

**Claim 50.20.** *We have an estimate*

$$(50.78) \quad |c_j(x) \eta_j(x)| \lesssim \lambda \chi_{100Q_j}(x)$$

uniformly over  $j$ .

*Proof.* Let  $x \in 100Q_j$ . It suffices to show that

$$(50.79) \quad \left| \frac{\langle f, e_{j,k} \cdot \eta_j \rangle}{\|\eta_j\|_{L^1(100Q_j)}} \cdot e_{j,k}(x) \eta_j(x) \right| \lesssim_{\lambda} 1$$

for each  $k \in K_j$  because  $K$  is a finite set. By Lemma 50.18 we have

$$(50.80) \quad |\partial^\alpha c_j(x)| \leq c_\alpha \ell(Q_j)^{-|\alpha|}$$

for  $\alpha \in \mathbb{N}_0^d$ . A similar inequality is also the case for  $\eta_j$ , as was shown in Lemma 50.16:

$$(50.81) \quad |\partial^\alpha \eta_j(x)| \leq c_\alpha \ell(Q_j)^{-|\alpha|}.$$

As a result, we obtain a valid bound for  $|e_{j,k}(x) \eta_j(x)|$ :

$$(50.82) \quad |e_{j,k}(x) \eta_j(x)| \leq c < \infty.$$

Now we are to estimate  $\langle f, c_j \cdot \eta_j \rangle$ . To do this we pick a point  $z_j \in 1000Q_j \setminus O$ . Then we have

$$(50.83) \quad \langle f, c_j \cdot \eta_j \rangle = f * [c_j(z_j - \cdot) \eta_j(z_j - \cdot)](z_j).$$

By virtue of (50.80) and (50.81) we obtain

$$\begin{aligned} & p_N (\ell(Q_j))^d c_j(z_j - \ell(Q_j) \cdot) \eta_j(z_j - \ell(Q_j) \cdot) \\ &= \ell(Q_j)^d \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N |\partial_x^\alpha [c_j(z_j - \ell(Q_j)x) \eta_j(z_j - \ell(Q_j)x)]| \\ &= \ell(Q_j)^d \sum_{|\alpha| \leq N} \sup_{|z_j - \ell(Q_j)x - c(Q_j)| \leq \ell(Q_j)} \langle x \rangle^N |\partial_x^\alpha [c_j(z_j - \ell(Q_j)x) \eta_j(z_j - \ell(Q_j)x)]| \\ &= \ell(Q_j)^d \sum_{|\alpha| \leq N} \sup_{1 \leq |x| \leq 2000} \langle x \rangle^N |\partial_x^\alpha [c_j(z_j - \ell(Q_j)x) \eta_j(z_j - \ell(Q_j)x)]| \lesssim \ell(Q_j)^d. \end{aligned}$$

As a result, we obtain

$$(50.84) \quad |\langle f, c_j \cdot \eta_j \rangle| \lesssim \ell(Q_j)^d \mathcal{M}f(z_j) \lesssim \lambda \ell(Q_j)^d.$$

Now that  $\int_{\mathbb{R}^d} \eta_j(x) dx$  differs from  $\ell(Q_j)^d$  essentially by a multiplicative constant independent of  $j$  and  $f$ , we have the desired result.  $\square$

A variant of the Calderón-Zygmund decomposition. Now we consider a variant of the Calderón-Zygmund decomposition. Let us set

$$(50.85) \quad b_j := (f - c_j) \cdot \eta_j$$

for  $j \in \mathbb{N}$  and

$$(50.86) \quad g := f - \sum_{j=1}^{\infty} b_j.$$

$b_j$  corresponds to the “bad” part, while  $g$  is a counterpart for the “good” part.

Our next goal is to control  $\mathcal{M}b_j(x)$ , the grand maximal operator of the “bad” part. What we are going to establish are pointwise estimates. So, let us freeze  $x$  and write it as  $x_0$ .

**Lemma 50.21.** *Let  $x_0 \in 100Q_j$  be fixed. Then we have*

$$(50.87) \quad \mathcal{M}[\eta_j \cdot f](x_0), \mathcal{M}[\eta_j \cdot c_j](x_0) \lesssim \mathcal{M}f(x_0)$$

uniformly over  $j, f$  and  $x_0$ .

*Proof.* By the definition of the grand maximal operator  $\mathcal{M}$  we have

$$\mathcal{M}[\eta_j \cdot f](x_0) = \sup_{\varphi \in \mathcal{F}_N, t > 0} \left| \left\langle \eta_j \cdot f, \frac{1}{t^d} \varphi \left( \frac{x_0 - \cdot}{t} \right) \right\rangle \right| = \sup_{\varphi \in \mathcal{F}_N, t > 0} \left| \left\langle f, \eta_j \cdot \frac{1}{t^d} \varphi \left( \frac{x_0 - \cdot}{t} \right) \right\rangle \right|.$$

Let  $t \leq \ell(Q_j)$ . Then we consider  $\Phi_{j,t}(x) = \eta_j(x_0 - tx) \cdot \varphi(x)$ . Then it follows from (50.76) that  $\sup_{j \in \mathbb{N}, t > 0} p_N(\Phi_{j,t}) < \infty$ , which yields

$$(50.88) \quad \begin{aligned} \left| \left\langle f, \eta_j \cdot \frac{1}{t^d} \varphi \left( \frac{x_0 - \cdot}{t} \right) \right\rangle \right| &= \left| \left\langle f, \frac{1}{t^d} \Phi_{j,t} \left( \frac{x_0 - \cdot}{t} \right) \right\rangle \right| \\ &= \left| f * \left( \frac{1}{t^d} \Phi_{j,t} \left( \frac{\cdot}{t} \right) \right) (x_0) \right| \\ &\lesssim \mathcal{M}f(x_0). \end{aligned}$$

Now let  $t > \ell(Q_j)$ . Then we consider  $\Psi_{j,t}(x) = \frac{\ell(Q_j)^d}{t^d} \eta_j(x_0 - \ell(Q_j)x) \Phi \left( \frac{\ell(Q_j)}{t} x \right)$ . Note that  $\Psi_{j,t}(x)$  is supported on a cube  $Q(100)$ . Hence it follows that  $\sup_{j \in \mathbb{N}} p_N(\Psi_{j,t}) < \infty$ . Going through a similar argument as above, we obtain

$$(50.89) \quad \begin{aligned} \left| \left\langle f, \eta_j \cdot \frac{1}{t^d} \varphi \left( \frac{x_0 - \cdot}{t} \right) \right\rangle \right| &= \left| \left\langle f, \frac{1}{\ell(Q_j)^d} \Psi_{j,t} \left( \frac{x_0 - \cdot}{\ell(Q_j)} \right) \right\rangle \right| \\ &= \left| f * \left( \frac{1}{\ell(Q_j)^d} \Psi_{j,t} \left( \frac{\cdot}{\ell(Q_j)} \right) \right) (x_0) \right| \\ &\lesssim \mathcal{M}f(x_0). \end{aligned}$$

Thus, the treatment of the first inequality is complete.

Let us turn to the second inequality. To do this, we insert (50.77), which yields

$$\mathcal{M}[c_j \cdot \eta_j](x_0) \leq \sum_{k \in K_j} \frac{|\langle f, e_{j,k} \eta_j \rangle|}{\int_{100Q_j} \eta_j} \mathcal{M}[e_{j,k} \cdot \eta_j](x_0).$$

As we have seen,  $|e_{j,k}(x)\eta_j(x)|$  is bounded by a constant independent of  $j$  and  $k$ . It follows from Lemma 50.18 that

$$(50.90) \quad \sup_{j \in \mathbb{N}} p_N(e_{j,k}(x_0 - \ell(Q_j)\cdot)\eta_j(x_0 - \ell(Q_j)\cdot)) < \infty.$$

As a consequence, we obtain

$$(50.91) \quad \mathcal{M}[c_k \cdot \eta_j](x_0) \leq c < \infty, \quad |\langle f, e_{j,k}\eta_j \rangle| \lesssim \mathcal{M}f(x_0)$$

uniformly over  $j$  and  $k$ . and hence

$$(50.92) \quad \mathcal{M}[c_j\eta_j](x_0) \lesssim \mathcal{M}f(x_0)$$

and the proof is complete. □

**Proposition 50.22.** *We have an estimate*

$$(50.93) \quad \mathcal{M}b_j(x_0) \lesssim_{\zeta,d} \mathcal{M}f(x_0)\chi_{1000Q_j}(x_0) + \frac{\lambda \ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x_0 - c(Q_j)|^{d+L+1}}$$

uniformly over  $x_0 \in \mathbb{R}^d$ .

*Proof.* In view of Lemma 50.21 we are left with the task of proving (50.93) for  $x_0 \notin 1000Q_j$ . Let us pick a reference point  $z_0 \in 10000Q_j \setminus O$ . We shall claim

$$\mathcal{M}b_j(x_0) \lesssim \frac{\mathcal{M}f(z_0) \ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x_0 - c(Q_j)|^{d+L+1}}$$

or equivalently

$$(50.94) \quad \left| \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) * b_j(x_0) \right| \leq \frac{c \mathcal{M}f(z_0) \ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x_0 - c(Q_j)|^{d+L+1}}$$

for all  $\varphi \in \mathcal{F}_N$  and  $t > 0$ . Once this is established, the proposition will have been proved because  $\mathcal{M}f(z_0) \leq \lambda$  by the definition of  $O$ . We also remark that

$$(50.95) \quad \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) * b_j(x_0) = \langle f - c_j, \Phi_{j,t}(z_0 - \cdot) \rangle = \langle f - c_j, \tilde{\Phi}_{j,t}(z_0 - \cdot) \rangle,$$

where

$$\begin{aligned} \Phi_{j,t}(x) &:= \eta_j(z_0 - x) \cdot \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + x}{t} \right) \\ \tilde{\Phi}_{j,t}(x) &:= \eta_j(z_0 - x) \cdot \left( \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + x}{t} \right) - \sum_{|\alpha| \leq L} \frac{1}{\alpha! t^d} \partial^\alpha \varphi \left( \frac{x_0 - z_0}{t} \right) \left( \frac{x}{t} \right)^\alpha \right). \end{aligned}$$

Assume first that  $x_0 \in 100000Q_j$ . If  $t > \ell(Q_j)$ , then we have

$$(50.96) \quad p_N(\ell(Q_j)^d \Phi_{j,t}(\ell(Q_j)\cdot)) \leq c < \infty,$$

where  $c$  does not depend on  $j, z_0, x_0, t$ . If  $t < \ell(Q_j)$ , then

$$(50.97) \quad \langle x \rangle^N \left| \partial^\alpha \varphi \left( \frac{x_0 - z_0 + tx}{t} \right) \right| \chi_{100Q_j}(z_0 - tx) \lesssim \left\langle \frac{z_0 - c(Q_j)}{t} \right\rangle^N \left\langle \frac{x_0 - c(Q_j)}{t} \right\rangle^{-N} \leq c < \infty,$$

which implies

$$(50.98) \quad p_N(t^d \Phi_{j,t}(t\cdot)) \leq c < \infty,$$

where  $c$  does not depend on  $j, z_0, x_0, t$ . As a result, we obtain

$$(50.99) \quad \left| \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) * b_j(x_0) \right| \lesssim \mathcal{M}f(z_0).$$

Assume instead that  $x_0 \notin 100000Q_j$ . We make use of the following basic estimate:

$$\left| \partial^\beta \left( \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + x}{t} \right) - \sum_{|\alpha| \leq L} \frac{1}{\alpha! t^d} \partial^\alpha \varphi \left( \frac{x_0 - z_0}{t} \right) \left( \frac{x}{t} \right)^\alpha \right) \right| \lesssim \frac{\ell(Q_j)^{L+1-|\beta|}}{t^{d+L+1} \left\langle \frac{x_0 - c(Q_j)}{t} \right\rangle^M},$$

if  $z_0 - x \in \text{supp}(\eta_j)$  and  $|\beta| \leq L + 1$

$$\left| \partial^\beta \left( \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + x}{t} \right) - \sum_{|\alpha| \leq L} \frac{1}{\alpha! t^d} \partial^\alpha \varphi \left( \frac{x_0 - z_0}{t} \right) \left( \frac{x}{t} \right)^\alpha \right) \right| \lesssim t^{-d-|\beta|} \left\langle \frac{x_0 - c(Q_j)}{t} \right\rangle^{-M},$$

if  $z_0 - x \in \text{supp}(\eta_j)$  and  $|\beta| > L + 1$ .

In summary

(50.100)

$$I := \left| \partial^\beta \ell(Q_j)^d \left( \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + \ell(Q_j)x}{t} \right) - \sum_{|\alpha| \leq L} \frac{1}{\alpha! t^d} \partial^\alpha \varphi \left( \frac{x_0 - z_0}{t} \right) \left( \frac{\ell(Q_j)x}{t} \right)^\alpha \right) \right|$$

has an estimate

$$I \leq \frac{c \ell(Q_j)^{d+\max(|\beta|, L+1)}}{(t + |x_0 - c(Q_j)|)^{d+\max(|\beta|, L+1)}} \leq \frac{c \ell(Q_j)^{d+L+1}}{(t + |x_0 - c(Q_j)|)^{d+L+1}},$$

which yields

$$(50.101) \quad \left| \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) * b_j(x_0) \right| \lesssim \mathcal{M}f(z_0).$$

Thus, the proof is now complete. □

**Proposition 50.23.** *We have*

$$(50.102) \quad \mathcal{M}g(x_0) \lesssim_{d,\zeta} \chi_{O^c}(x_0) \mathcal{M}f(x_0) + \sum_{j=1}^{\infty} \frac{\lambda \ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x_0 - c(Q_j)|^{d+L+1}}$$

for all  $x_0 \in \mathbb{R}^d$ .

*Proof.* If  $x_0 \notin O$ , then the inequality is clear from the previous estimate. Suppose that  $x \in O$ , more specifically,  $x \in 10Q_j$ . Let us set

$$\mathcal{J}_{j,0} := \{l \in \mathbb{N} : 100Q_l \cap 100Q_j \neq \emptyset\}, \mathcal{J}_{j,1} := \mathbb{N} \setminus \mathcal{J}_{j,0}$$

for such  $j$ .

Let us write

$$(50.103) \quad g(x_0) = \left( f - \sum_{l \in \mathcal{J}_{j,0}} b_l \right) (x_0) - \sum_{l \in \mathcal{J}_{j,1}} b_l(x_0).$$

It is easy to deal with the second summation. Indeed, a crude estimate using the triangle inequality  $\mathcal{M}[F + G](x) \leq \mathcal{M}F(x) + \mathcal{M}G(x)$  suffices:

$$\mathcal{M} \left[ \sum_{l \in \mathcal{J}_{j,1}} b_l \right] (x_0) \leq \sum_{l \in \mathcal{J}_{j,1}} \mathcal{M}b_l(x_0) \lesssim \sum_{l \in \mathcal{J}_{j,1}} \frac{\lambda \ell(Q_l)^{d+L+1}}{\ell(Q_l)^{d+L+1} + |x_0 - c(Q_l)|^{d+L+1}}.$$

Now consider the first summation. To do this we decompose the summation further:

$$(50.104) \quad \mathcal{M} \left[ f - \sum_{l \in \mathcal{J}_{j,0}} b_l \right] (x_0) \leq \mathcal{M} \left[ f - \sum_{l \in \mathcal{J}_{j,0}} \eta_l \cdot f \right] (x_0) + \sum_{l \in \mathcal{J}_{j,0}} \mathcal{M}[c_l \cdot \eta_l](x_0).$$

The second term is less complicated to estimate

$$(50.105) \quad \mathcal{M}(c_l \cdot \eta_l)(x_0) \lesssim \frac{\lambda \ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x_0 - c(Q_j)|^{d+L+1}}$$

as we did in Lemma 50.21.

Let us turn to the first term. The support property of  $f - \sum_{l \in \mathcal{J}_{j,0}} \eta_l \cdot f$ , that is, the fact that  $f - \sum_{l \in \mathcal{J}_{j,0}} \eta_l \cdot f$  vanishes outside  $100Q_j$ , gives

$$(50.106) \quad \mathcal{M} \left[ f - \sum_{l \in \mathcal{J}_{j,0}} \eta_l \cdot f \right] (x_0) \leq c \lambda.$$

Indeed, let  $z_0 \in 10000Q_j \setminus O$  be a fixed reference point. We define

$$\begin{aligned} \tilde{\Psi}_{j,t}(x) &:= \frac{1}{t^d} \varphi \left( \frac{x_0 - x}{t} \right) \left( 1 - \sum_{l \in \mathcal{J}_{j,0}} \eta_l(x) \right) \\ \Psi_{j,t}(x) &:= \tilde{\Psi}_{j,t}(z_0 - x) = \frac{1}{t^d} \varphi \left( \frac{x_0 - z_0 + x}{t} \right) \left( 1 - \sum_{l \in \mathcal{J}_{j,0}} \eta_l(z_0 - x) \right) \end{aligned}$$

given a function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . We have to show that

$$(50.107) \quad \left| \left\langle f - \sum_{l \in \mathcal{J}_{j,0}} \eta_l \cdot f, \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) \right\rangle \right| \leq |\Psi_{j,t} * f(z_0)| \lesssim \mathcal{M}f(z_0) \lesssim \lambda$$

with  $c$  independent of  $t > 0$ . Note that  $\Psi_{j,t}(x) = 0$  for all  $x$  with  $|z_0 - x - c(Q_j)| \leq 100\ell(Q_j)$ , hence for all  $x$  with  $|z_0 - x - x_0| \leq \frac{1}{100}|z_0 - x_0|$ . As a result, it follows that

$$\begin{aligned} &p_N(\ell(Q_j)^d \Psi_{j,t}(\ell(Q_j) \cdot)) \\ &= \frac{\ell(Q_j)^d}{t^d} \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N \left| \partial^\alpha \left[ \varphi \left( \frac{x_0 - z_0 + \ell(Q_j)x}{t} \right) \left( 1 - \sum_{l \in \mathcal{J}_{j,0}} \eta_l(z_0 - \ell(Q_j)x) \right) \right] \right| \\ &\leq \frac{c \ell(Q_j)^d}{t^d} \leq c < \infty, \end{aligned}$$

if  $t > \ell(Q_j)$ .

A geometric observation shows  $|x + x_0 - z_0| \geq \frac{1}{100}|x|$ , if  $|z_0 - x - x_0| \leq \frac{1}{100}|z_0 - x_0|$ . If  $t \leq \ell(Q_j)$ , then we obtain

$$\begin{aligned} p_N(t^d \Psi_{j,t}(t \cdot)) &= \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^d} \langle x \rangle^N \left| \partial^\alpha \left[ \varphi \left( \frac{x_0 - z_0 + tx}{t} \right) \left( 1 - \sum_{l \in \mathcal{J}_{j,0}} \eta_l(z_0 - tx) \right) \right] \right| \\ &\lesssim 1. \end{aligned}$$

Consequently (50.107) was justified. □

**Proposition 50.24.** *Let  $\lambda$  be fixed. Keep to the same notation as above. Then*

$$\int_{\mathbb{R}^d} \mathcal{M}g(x) dx \lesssim \lambda^{1-p} \int_{\mathbb{R}^d} \mathcal{M}f(x)^p dx, \quad \int_{\mathbb{R}^d} \mathcal{M}b(x)^p dx \lesssim \int_{\{\mathcal{M}f > \lambda\}} \mathcal{M}f^p.$$

*Proof.* If we integrate (50.102) over  $\mathbb{R}^d$ , then we obtain

$$(50.108) \quad \int_{\mathbb{R}^d} \mathcal{M}g(x) \lesssim \int_{\{\mathcal{M}f \leq \lambda\}} \mathcal{M}f(x) dx + \lambda \sum_{j=1}^{\infty} \ell(Q_j)^d \lesssim \lambda^{1-p} \int_{\{\mathcal{M}f \leq \lambda\}} (\mathcal{M}f)^p + \lambda \sum_{j=1}^{\infty} \ell(Q_j)^d,$$

while  $\sum_{j=1}^{\infty} \ell(Q_j)^d = \sum_{j=1}^{\infty} |Q_j| \leq |\{\mathcal{M}f > \lambda\}| < \infty$ . As a result we have

$$(50.109) \quad \int_{\mathbb{R}^d} \mathcal{M}g(x) dx \lesssim \lambda^{1-p} \int_{\{\mathcal{M}f \leq \lambda\}} \mathcal{M}f(x)^p dx.$$

Therefore, we conclude  $\int_{\mathbb{R}^d} \mathcal{M}g(x) dx < \infty$ , that is,  $g \in H^1(\mathbb{R}^d)$ .

Let us deal with the “bad” part. If we sum (50.93), which we proved in Proposition 50.22, over  $j \in \mathbb{N}$ , then we obtain

$$(50.110) \quad \mathcal{M}b(x) \leq \sum_{j=1}^{\infty} \mathcal{M}b_j(x) \lesssim \chi_{O}(x)\mathcal{M}f(x) + c\lambda \sum_{j=1}^{\infty} \frac{\ell(Q_j)^{d+L+1}}{\ell(Q_j)^{d+L+1} + |x - c(Q_j)|^{d+L+1}}.$$

Let us recall that  $0 < p \leq 1$ . In view of the assumption that  $0 < p \leq 1$  we have

$$(50.111) \quad \mathcal{M}b(x)^p \lesssim \chi_{O}(x)\mathcal{M}f(x)^p + c\lambda^p \sum_{j=1}^{\infty} \frac{\ell(Q_j)^{p(d+L+1)}}{\ell(Q_j)^{p(d+L+1)} + |x - c(Q_j)|^{p(d+L+1)}}.$$

Thus, if we integrate this inequality over  $\mathbb{R}^d$ , then we obtain

$$(50.112) \quad \int_{\mathbb{R}^d} \mathcal{M}b(x)^p dx \lesssim \int_{\{\mathcal{M}f > \lambda\}} \mathcal{M}f(x)^p dx + \lambda^p \left( \sum_{j=1}^{\infty} \ell(Q_j)^d \right)^p \leq \int_{\{\mathcal{M}f > \lambda\}} \mathcal{M}f^p + \lambda^p \sum_{j=1}^{\infty} \ell(Q_j)^d.$$

In the same way as before, we can prove

$$(50.113) \quad \lambda^p \sum_{j=1}^{\infty} \ell(Q_j)^d \leq \int_{\{\mathcal{M}f > \lambda\}} \lambda^p \cdot \chi_{\{\mathcal{M}f > \lambda\}} \lesssim \int_{\{\mathcal{M}f > \lambda\}} \mathcal{M}f^p$$

and the proof is now complete. □

Conclusion of the proof of the converse of Theorem 50.15. To stress the dependence of  $g$  on  $\lambda$ , let us write it as  $g_\lambda$ . We shall do the same for  $b_j$  and  $Q_j$  and so on. Under our new notation, we have a collection of Hilbert spaces  $\mathcal{H}_{k,\lambda}$ ,  $k \in \mathbb{N}$ , pointwise estimates

$$\begin{aligned} \mathcal{M}g_\lambda(x) &\lesssim \chi_{\mathbb{R}^d \setminus O_\lambda}(x)\mathcal{M}f(x) + \sum_{k=1}^{\infty} \frac{\lambda \ell(Q_{k,\lambda})^{d+L+1}}{\ell(Q_{k,\lambda})^{d+L+1} + |x - c(Q_{k,\lambda})|^{d+L+1}} \\ \mathcal{M}b_{j,\lambda}(x) &\lesssim \chi_{100Q_{j,\lambda}}(x)\mathcal{M}f(x) + \frac{\lambda \ell(Q_{j,\lambda})^{d+L+1}}{\ell(Q_{j,\lambda})^{d+L+1} + |x - c(Q_{j,\lambda})|^{d+L+1}}, \quad j \in \mathbb{N} \end{aligned}$$

and a Whitney covering  $\{Q_{j,\lambda}\}_{j \in \mathbb{N}}$  of  $\{\mathcal{M}f > \lambda\}$ . Let us set  $\sum_{j=1}^{\infty} b_{j,\lambda} =: b_\lambda$ .

Proposition 50.24 readily gives us

**Proposition 50.25.** *Let  $0 < p \leq 1$ . Then*

$$(50.114) \quad \int_{\mathbb{R}^d} \mathcal{M}g_\lambda(x) < \infty,$$

$$(50.115) \quad \lim_{\lambda \rightarrow \infty} \int_{\mathbb{R}^d} \mathcal{M}[f - g_\lambda](x)^p dx = 0.$$

Now we claim the following.

**Claim 50.26.**  $|g_\lambda(x)| \lesssim \lambda$  for all  $\lambda > 0$ .

*Proof.* Let  $x \notin O_\lambda$ . Then we have

$$(50.116) \quad |g_\lambda(x)| = |f(x)| \lesssim \mathcal{M}f(x) \lesssim \lambda$$

from the definition of  $g_\lambda$ : Recall that we defined

$$(50.117) \quad b_{j,\lambda}(x) = \eta_j(x) \cdot (f(x) - c_{j,\lambda}(x)), \quad g_\lambda(x) = f(x) - \sum_{j=1}^{\infty} b_{j,\lambda}(x).$$

Suppose instead that  $x \in 10Q_{j,\lambda}$  for some  $j \in \mathbb{N}$ . Then we can use  $|c_{j,\lambda}(x)\eta_{j,\lambda}(x)| \lesssim \lambda$ . □

Now we make a little reduction to proceed further.

**Lemma 50.27.** *Let  $X$  be a set. Let  $D$  and  $Y$  be subsets of  $X$  such that  $D \subset Y \subset X$ . Suppose that  $X$  and  $Y$  come with quasi-norms  $\|\cdot\|_X$  and  $\|\cdot\|_Y$ , respectively, which make  $X$  and  $Y$  into quasi-normed spaces. Assume the following.*

- (1)  $D$  is dense in  $X$  with respect to  $\|\cdot\|_X$ .
- (2)  $(Y, \|\cdot\|_Y)$  is continuously embedded into  $(X, \|\cdot\|_X)$ .
- (3) We have

$$(50.118) \quad \|y\|_Y \simeq \|y\|_X$$

for all  $y \in D$ .

- (4)  $(Y, \|\cdot\|_Y)$  is a quasi-Banach space.

Then  $X = Y$  as a set and their quasi-norms are mutually equivalent.

*Proof.* Let  $x \in X$ . Then there exists a sequence  $\{d_j\}_{j \in \mathbb{N}} \subset D$  that approximates  $x$  in the topology of  $X$ . Now that the norms are assumed equivalent if we restrict them to  $D$ , we have

$$(50.119) \quad \lim_{j,k \rightarrow \infty} \|d_j - d_k\|_Y = 0.$$

Since  $(Y, \|\cdot\|_Y)$  is a quasi-Banach space,  $\lim_{j \rightarrow \infty} d_j =: y \in Y$  exists with respect to  $\|\cdot\|_Y$ . By virtue of the continuous embedding  $Y \subset X$  we conclude that  $\lim_{j \rightarrow \infty} d_j = y$  takes place with respect to  $\|\cdot\|_X$  as well.  $X$  being assumed a quasi-normed space, we see that  $x = y \in Y$ . As a consequence we see that  $X = Y$ . Now that the fact that any element in  $X$  can be approximated with respect to both  $\|\cdot\|_X$  and  $\|\cdot\|_Y$ , we have only to pass to the limit of (50.118) to see that the norms are mutually equivalent. □

We apply Lemma 50.27 as follows:

$$X := X^p(\mathbb{R}^d) = \{f \in \mathcal{S}'(\mathbb{R}^d) : \|\mathcal{M}f\|_p < \infty\}$$

$$Y := H^p(\mathbb{R}^d)$$

$$D := \{f \in X^p : \|\mathcal{M}f\|_1, \|f\|_\infty < \infty\}.$$

We have been verified some of the assumptions of the lemma:

- (1)  $D$  is dense in  $X^p$  (Proposition 50.25 and Claim 50.26).
- (2)  $H^p(\mathbb{R}^d) \hookrightarrow X^p(\mathbb{R}^d)$  is a continuous embedding (Subsection 50.3).

(3) As is easily verified by using Theorem 10.12,  $H^p(\mathbb{R}^d)$  is a quasi-Banach space.

Therefore, it remains to show that  $D \subset H^p(\mathbb{R}^d)$  and that

$$(50.120) \quad \|f\|_{H^p} \lesssim \|f\|_{X^p}$$

for all  $f \in D$ .

This allows us to assume that  $f \in X^p \cap L^1(\mathbb{R}^d)$ . If we assume that  $f \in D$ , then the key estimate

$$(50.121) \quad \mathcal{M}g_\lambda(x) \lesssim \chi_{\mathbb{R}^d \setminus O_\lambda}(x) \mathcal{M}f(x) + \lambda \sum_{j=1}^{\infty} \frac{\ell(Q_{j,\lambda})^{d+L+1}}{\ell(Q_{j,\lambda})^{d+L+1} + |x - c(Q_{j,\lambda})|^{d+L+1}},$$

which we have strived to show in this subsection, yields that

$$\begin{aligned} \int_{\mathbb{R}^d} \mathcal{M}g_\lambda(x)^p dx &\lesssim \int_{\{\mathcal{M}f \leq \lambda\}} \mathcal{M}f(x)^p dx + \lambda^p \sum_{j=1}^{\infty} |Q_{j,\lambda}| \\ &\lesssim \int_{\mathbb{R}^d} \min(\mathcal{M}f(x), \lambda)^p dx. \end{aligned}$$

In particular we have  $g_\lambda \rightarrow 0$  as  $\lambda \rightarrow 0$  in  $X^p$ . As a consequence we have

$$(50.122) \quad f = \lim_{J \rightarrow \infty} g_{2^J} - g_{2^{J-1}}$$

in the topology of  $X^p$ .

Let  $j \in \mathbb{Z}$  and  $k \in \mathbb{N}$ . Denote by  $P_k^j$  the projection to the Hilbert space  $\mathcal{H}_{k,2^j}$ , which we have defined earlier. By definition given  $f \in \mathcal{S}'(\mathbb{R}^d)$  with  $\|\mathcal{M}f\|_p < \infty$ ,  $P_k^j(f)$  is a unique polynomial satisfying

$$(50.123) \quad \langle f, q \cdot \eta_{k,2^j} \rangle = \langle P_k^j(f), q \cdot \eta_{k,2^j} \rangle$$

for all polynomials  $q$  with degree  $L$ . Let us set

$$(50.124) \quad c_{j,k,l} = P_l^j((f - c_{l,2^j})\varphi_{k,2^{j-1}}).$$

If  $100Q_{l,2^j} \cap 100Q_{k,2^{j-1}} = \emptyset$ , then  $c_{j,k,l} = 0$ , as is easily seen from (50.123). Since  $O_{2^j} \subset O_{2^{j-1}}$ , we have an estimate  $\ell(Q_{l,2^j}) \lesssim \ell(Q_{k,2^{j-1}})$  uniformly over  $j$  and  $k$  as long as they meet. In view of this geometric observation we conclude that

$$(50.125) \quad |c_{j,k,l}(x)\varphi_{l,2^j}(x)| \lesssim 2^j \chi_{c_{l,2^j} Q_k^{2^{j-1}}}(x)$$

from Proposition 50.19. Below equalities will be understood as the one in  $\mathcal{S}'(\mathbb{R}^d)$ . Let us show that

$$\sum_{k,l=1}^{\infty} c_{j,k,l}\varphi_{l,2^j} = \sum_{l=1}^{\infty} \left( \sum_{k=1}^{\infty} P_l^j((f - c_{l,2^j})\varphi_{k,2^{j-1}}) \right) \varphi_{l,2^j}$$

equals identically 0. Now that  $\sum_{k=1}^{\infty} (f - c_{l,2^j})\varphi_{k,2^{j-1}} = (f - c_{l,2^j})\chi_{O_{2^{j-1}}}$  takes place in  $L^2(\mathbb{R}^d)$ ,

we have

$$(50.126) \quad \sum_{k,l=1}^{\infty} c_{j,k,l}\varphi_{l,2^j} = \sum_{l=1}^{\infty} P_l^j((f - c_{l,2^j})\chi_{O_{2^{j-1}}})\varphi_{l,2^j} = \sum_{l=1}^{\infty} (c_{l,2^j} - c_{l,2^j})\varphi_{l,2^j} = 0$$

from the definition of the projection  $P_l^j$ .

Let us set

$$A_{j,k} := (f - c_{k,2^{j-1}})\varphi_{k,2^{j-1}} - \sum_{l=1}^{\infty} (f - c_{l,2^j})\varphi_{l,2^j}\varphi_{k,2^{j-1}} + \sum_{l=1}^{\infty} c_{j,k,l}\varphi_{l,2^j}.$$



As is seen from the definition and the geometric observation above,  $A_{j,k}$  is supported on  $cQ_{k,2^{j-1}}$ , where  $c$  is independent of  $j$  and  $k$ .

First, recompose of the terms defining  $A_{j,k}$

$$(50.127) \quad A_{j,k} = \left(1 - \sum_{l=1}^{\infty} \varphi_{l,2^j} \varphi_{k,2^{j-1}}\right) f - c_{k,2^{j-1}} \varphi_{k,2^{j-1}} + \sum_{l=1}^{\infty} c_{l,2^j} \varphi_{l,2^j} \varphi_{k,2^{j-1}} + \sum_{l=1}^{\infty},$$

which gives us

$$(50.128) \quad |A_{j,k}(x)| \lesssim \chi_{c_0 Q_{k,2^j}} \left( \left(1 - \sum_{l=1}^{\infty} \varphi_{l,2^j} \varphi_{k,2^{j-1}}\right) |f(x)| + 2^j \right) \lesssim 2^j \chi_{c_0 Q_{k,2^{j-1}}}(x)$$

together with Claim 50.20.

Furthermore, since  $\{Q_{k,2^{j-1}}\}_{k \in \mathbb{N}}$  is a Whitney covering of  $\{\mathcal{M}f > 2^{j-1}\}$ , we obtain

$$(50.129) \quad \sum_{j,k=1}^{\infty} 2^{jp} |Q_{k,2^{j-1}}| \lesssim \sum_j 2^{jp} |\{\mathcal{M}f > 2^{j-1}\}| \lesssim \int_{\mathbb{R}^d} \mathcal{M}f(x)^p dx < \infty.$$

Finally, for all polynomials  $q$  of degree  $L$ , we have

$$\int_{\mathbb{R}^d} A_{j,k}(x)q(x) dx = \left\langle (f - c_{k,2^{j-1}}) - \sum_{l=1}^{\infty} (f - c_{l,2^j}) \varphi_{l,2^j}, q \cdot \varphi_{k,2^{j-1}} \right\rangle + \sum_{l=1}^{\infty} \langle c_{j,k,l}, q \cdot \varphi_{l,2^j} \rangle.$$

The bracket involving the part of  $f - c_{k,2^{j-1}}$  vanishes from the definition of  $c_{k,2^{j-1}}$ . Furthermore, with  $k$  fixed, we are in the position of using the Lebesgue convergence theorem to have

$$(50.130) \quad \int_{\mathbb{R}^d} A_{j,k}(x)q(x) dx = - \sum_{l=1}^{\infty} \langle (f - c_{l,2^j}) \varphi_{l,2^j}, q \cdot \varphi_{k,2^{j-1}} \rangle + \sum_{l=1}^{\infty} \langle c_{j,k,l}, q \cdot \varphi_{l,2^j} \rangle$$

By virtue of the fact that  $P_l^j$  is a projection we see

$$(50.131) \quad A_{j,k} \perp \mathcal{P}_L(\mathbb{R}^d)$$

where for the second equality we have used the fact that  $c_{k,2^{j-1}} = P_k^{j-1}(f)$ .

With (50.128)–(50.131) established, we can change the order of the summation for  $j$  and  $k$  freely. As a consequence we have

$$\begin{aligned} & \sum_{j=-\infty}^{\infty} \sum_{k=1}^{\infty} A_{j,k} \\ &= \lim_{J \rightarrow \infty} \sum_{j=-J}^J \sum_{k=1}^{\infty} \left( (f - c_{k,2^{j-1}}) \varphi_{k,2^{j-1}} - \sum_{l=1}^{\infty} (f - c_{l,2^j}) \varphi_{l,2^j} \varphi_{k,2^{j-1}} + \sum_{l=1}^{\infty} c_{j,k,l} \varphi_{l,2^j} \right) \\ &= \lim_{J \rightarrow \infty} \sum_{j=-J}^J \left( \sum_{k=1}^{\infty} (f - c_{k,2^{j-1}}) \varphi_{k,2^{j-1}} - \sum_{l=1}^{\infty} (f - c_{l,2^j}) \varphi_{l,2^j} \right) \\ &= \lim_{J \rightarrow \infty} \sum_{j=-J}^J b_{2^j-1} - b_{2^j} = \lim_{J \rightarrow \infty} \sum_{j=-J}^J g_{2^j} - g_{2^{j-1}} = f \end{aligned}$$

Here we have used (50.126) and (50.122) for the second and the last equalities respectively and equalities will be understood as the one in  $\mathcal{S}'(\mathbb{R}^d)$ .

In view of (50.128)–(50.131) as well as the last equality we see that  $\sum_{j,k=1}^{\infty} A_{j,k} = f$  is the desired atomic decomposition.

- Exercise 251.* (1) Use the polarization to deduce (50.75).  
 (2) The norm of  $\mathcal{H}_{j,\lambda}$  remains equivalent if we replace  $\eta_j$  with  $\zeta_j$ .

**50.5. Summary of Sections 50.3 and 50.4.**

Here let us summarize what we have obtained before we conclude this section.

**Theorem 50.28** (Coifman, 1974). *Let  $0 < p \leq 1$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Set  $L_0 = [d(1/p - 1)]$ . Assume that  $L \geq L_0$ .*

(1) *The following are equivalent.*

[H1]  $\|\mathcal{M}f\|_p < \infty$ , where we have defined

$$(50.132) \quad \mathcal{M}f(x) = \sup_{\varphi \in \mathcal{F}, t > 0} \left| \frac{1}{t^d} \varphi \left( \frac{\cdot}{t} \right) * f(x) \right|.$$

[H2] *Let  $\psi \in \mathcal{S}(\mathbb{R}^d)$  be a non-degenerate function, that is,  $\int_{\mathbb{R}^d} \psi(x) dx \neq 0$ . Then*

$$(50.133) \quad \left\| \sup_{j \in \mathbb{Z}} 2^{jn} |\psi(2^j \cdot) * f| \right\|_p < \infty.$$

[H3] *f admits a decomposition*

$$(50.134) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j,$$

where the functions  $\{a_j\}_{j=1}^{\infty}$  and the coefficient  $\{\lambda_j\}_{j=1}^{\infty}$  satisfy

- (i)  $\text{supp}(a_j)$  is contained in a cube  $Q_j$  for each  $j \in \mathbb{N}$ .
- (ii)  $\|a_j\|_2 \leq |Q_j|^{-\frac{1}{p} + \frac{1}{2}}$ .
- (iii)  $a \perp \mathcal{P}_L(\mathbb{R}^d)$ .
- (iv)  $\{\lambda_j\}_{j=1}^{\infty} \in \ell^p$ .

[H4] *f admits a decomposition*

$$(50.135) \quad f = \sum_{j=1}^{\infty} \lambda_j a_j,$$

where the functions  $\{a_j\}_{j=1}^{\infty}$  and the coefficient  $\{\lambda_j\}_{j=1}^{\infty}$  satisfy

- (i)  $\text{supp}(a_j)$  is contained in a cube  $Q_j$  for each  $j \in \mathbb{N}$ .
- (ii)  $\|a_j\|_{\infty} \leq |Q_j|^{-\frac{1}{p}}$ .
- (iii)  $A \perp \mathcal{P}_L(\mathbb{R}^d)$
- (iv)  $\{\lambda_j\}_{j=1}^{\infty} \in \ell^p(\mathbb{N})$ .

Furthermore, we have the following norm equivalence:

$$(50.136) \quad \|\mathcal{M}f\|_p \simeq \left\| \sup_{j \in \mathbb{Z}} 2^{jn} |\psi(2^j \cdot) * f| \right\|_p \simeq \inf_{(50.134)} \|\lambda\|_p \simeq \inf_{(50.135)} \|\lambda\|_p,$$

where the infimums are taken over all sequences satisfying (50.134) and (50.135) respectively.

(2) *The space  $H^1(\mathbb{R}^d) \cap H^p(\mathbb{R}^d)$  is dense in  $H^p(\mathbb{R}^d)$ .*

*Proof.* Indeed, it is trivial that [H4] implies [H3]. Furthermore, we have shown that [H1] and [H2] are equivalent in Subsection 50.2. As we have seen in Subsection 50.3 [H3] implies [H2]. In Subsection 50.4 we have struggled to show that [H1] implies [H4] as well as that  $H^1(\mathbb{R}^d) \cap H^p(\mathbb{R}^d)$  is dense in  $H^p(\mathbb{R}^d)$ . □

As for the Hardy space  $H^p(\mathbb{R}^d)$  with  $1 < p < \infty$  it is easy to define and investigate it. If we set

$$(50.137) \quad \|f\|_{H^p} := \|\mathcal{M}f\|_p,$$

then it is easy to prove  $H^p(\mathbb{R}^d) \approx L^p(\mathbb{R}^d)$  with norm equivalence by using the maximal operator.

**Remark 50.29.** The proof of [H1]  $\implies$  [H4] shows that the decomposition of  $f$  is highly non-linear.

### 51. LIPSCHITZ SPACE

Recall that  $\text{Lip}(\mathbb{R}^d)$  is a Banach space of functions from  $\mathbb{R}^d$  to  $\mathbb{C}$  modulo additive constants normed by

$$(51.1) \quad \|f\|_{\text{Lip}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

The aim of this section is to make a deeper investigation of the differentiability of the Lipschitz functions.

**Definition 51.1.** A function  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  is said to be differentiable at  $x \in \mathbb{R}^d$ , if there exist a linear mapping  $A$  and a function  $F_x(y)$  so that

$$(51.2) \quad f(y) = f(x) + A(y - x) + F_x(y)$$

with

$$(51.3) \quad \lim_{y \rightarrow x} \frac{|F_x(y)|}{|y - x|} = 0.$$

By differentiability we mean the total differentiability. Now that the partial derivative exists almost everywhere, our main concern goes to the total differentiability.

**Theorem 51.2.** *Any Lipschitz function is differentiable for almost everywhere.*

*Proof.* Existence of directional differential Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a Lipschitz function. Fix any  $v \in \mathbb{R}^d$  with  $|v| = 1$ , and define

$$\begin{aligned} \overline{D}_v f(x) &= \limsup_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} \\ \underline{D}_v f(x) &= \liminf_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}. \end{aligned}$$

**Claim 51.3.** *For a.e.  $x \in \mathbb{R}^d$ ,*

$$(51.4) \quad \overline{D}_v f(x) = \underline{D}_v f(x).$$

*Proof of the claim.* Since  $f$  is continuous,

$$\begin{aligned} \overline{D}_v f(x) &= \limsup_{\substack{t \in \mathbb{Q} \\ t \rightarrow 0}} \frac{f(x + tv) - f(x)}{t} \\ \underline{D}_v f(x) &= \liminf_{\substack{t \in \mathbb{Q} \\ t \rightarrow 0}} \frac{f(x + tv) - f(x)}{t} \end{aligned}$$

are measurable. Thus,

$$(51.5) \quad A_v := \{x \in \mathbb{R}^d : \overline{D}_v f(x) = \underline{D}_v f(x)\}$$

is measurable. Now, for each  $x, v \in \mathbb{R}^d$ , with  $|v| = 1$ , define

$$(51.6) \quad \varphi(t) := f(x + tv) \quad t \in \mathbb{R}.$$

Then  $\varphi$  is Lipschitz, thus absolutely continuous, and thus differentiable a.e.  $t \in \mathbb{R}$ . Hence

$$(51.7) \quad |A_v \cap L| = 0$$

for each line  $L$  parallel to  $v$ . Fubini's theorem then implies

$$(51.8) \quad |A_v| = 0.$$

Thus, the claim is proved.  $\square$

Total differential As a consequence of the claim above, we see

$$(51.9) \quad \text{grad}f(x) := (\partial_1 f(x), \partial_2 f(x), \dots, \partial_d f(x))$$

exists for almost everywhere.

**Claim 51.4.**  $D_v f(x) = v \cdot \text{grad}f(x)$  for almost everywhere.

*Proof.* Let  $\xi \in C_c^\infty(\mathbb{R}^d)$ . Then

$$(51.10) \quad \int_{\mathbb{R}^d} \frac{f(x + tv) - f(x)}{t} \xi(x) dx = - \int_{\mathbb{R}^d} f(x) \frac{\xi(x) - \xi(x - tv)}{t} dx.$$

Let  $t = \frac{1}{k}$  for  $k = 1, 2, \dots$  in the above inequality and note

$$(51.11) \quad k \left| f\left(x + \frac{v}{k}\right) - f(x) \right| \leq \|f\|_{\text{Lip}} |v| = \|f\|_{\text{Lip}}.$$

Thus, the dominated convergence theorem implies

$$\begin{aligned} \int_{\mathbb{R}^d} \overline{D}_v f(x) \xi(x) dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} k \left( f\left(x + \frac{v}{k}\right) - f(x) \right) \xi(x) dx \\ &= - \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} f(x) k \left( \xi\left(x - \frac{v}{k}\right) - \xi(x) \right) dx \\ &= - \sum_{j=1}^d v_j \int_{\mathbb{R}^d} f(x) \partial_j \xi(x) dx \\ &= \sum_{j=1}^d v_j \int_{\mathbb{R}^d} \partial_j f(x) \xi(x) dx. \end{aligned}$$

Therefore

$$(51.12) \quad \overline{D}_v f(x) = \sum_{j=1}^d v_j \partial_j f(x) = v \cdot \text{grad}f(x)$$

for almost everywhere.  $\square$

Construction of the set of points at which  $f$  is differentiable. Now choose  $\{v_k\}_{k=1}^\infty$  to be a countable, dense subset of  $S^{d-1}$ . Set

$$(51.13) \quad A_k := \{x \in \mathbb{R}^d : \overline{D}_{v_k} f(x) = \underline{D}_{v_k} f(x), \overline{D}_{v_k} f(x) = v_k \cdot \text{grad}f(x)\}$$

and define

$$(51.14) \quad A := \bigcap_{j=1}^\infty A_j.$$

Observe  $A = \mathbb{R}^d$  for almost everywhere.

To conclude the proof of the theorem we have only to show the following.

**Claim 51.5.** *f is differentiable at each point x ∈ A.*

*Proof.* Fix any x ∈ A. Choose v ∈ S<sup>d-1</sup>, t ∈ ℝ \ {0} and write

$$(51.15) \quad Q(x, v, t) = \frac{f(x + tv) - f(x)}{t} - v \cdot \text{grad}f(x).$$

Then if v' ∈ S<sup>d-1</sup>, we have

$$(51.16) \quad |Q(x, v, t) - Q(x, v', t)| \leq \left| \frac{f(x + tv) - f(x + tv')}{t} \right| + |(v - v') \cdot \text{grad}f(x)| \leq (d + 1)|v - v'|.$$

Now fix ε > 0 and choose N so large that if v ∈ S<sup>d-1</sup>, then

$$(51.17) \quad |v - v_k| \leq \frac{\varepsilon}{2(d + 1)\|f\|_{\text{Lip}} + 1}$$

for some k = 1, 2, ..., N. That is {v<sub>k</sub>}<sub>k=1</sub><sup>N</sup> forms an  $\frac{\varepsilon}{2(d + 1)\|f\|_{\text{Lip}} + 1}$ -net of S<sup>d-1</sup>, that is, for each point v in S<sup>d-1</sup> we can find k = 1, 2, ..., N such that  $|v - v_k| \leq \frac{\varepsilon}{2(d + 1)\|f\|_{\text{Lip}} + 1}$ .

Now

$$(51.18) \quad \lim_{t \rightarrow 0} Q(x, v_k, t) = 0, \quad k = 1, 2, \dots, N$$

and thus there exists δ > 0 so that

$$(51.19) \quad |Q(x, v_k, t)| < \frac{\varepsilon}{2} \text{ for all } 0 < |t| < \delta, \quad k = 1, 2, \dots, N.$$

Consequently, for each v ∈ S<sup>d-1</sup>, there exists k = 1, 2, ..., N such that

$$(51.20) \quad |Q(x, v, t)| \leq |Q(x, v_k, t)| + |Q(x, v, t) - Q(x, v_k, t)| < \varepsilon,$$

if 0 < |t| < δ according to (51.17) and (51.19). Note the same δ > 0 works for all v ∈ S<sup>d-1</sup>.

Now choose any y ∈ ℝ<sup>d</sup> y ≠ x. Write v =  $\frac{y - x}{|y - x|}$ , so that y = x + tv, t = |x - y|. Then

$$f(y) - f(x) - \text{grad}f(x) \cdot (y - x) = f(x + tv) - f(x) - tv \cdot \text{grad}f(x) = o(t) = o(|x - y|).$$

Hence f is differentiable at x with Df(x) = gradf(x). □

In view of this claim the proof is now complete. □

## 52. HÖLDER-ZYGMUND SPACES

In this section, we deal with Hölder-Zygmund spaces, which describes the differentiability of functions.

**52.1. Hölder-Zygmund space C<sup>θ</sup>(ℝ<sup>d</sup>).** We define the Hölder-Zygmund space C<sup>θ</sup>(ℝ<sup>d</sup>) with 0 < θ ≤ 1. The case when 0 < θ < 1 is used in PDEs. For example, we frequently encounter the a-priori estimates of elliptic differential equations. Indeed, it is not effective to use C<sup>2</sup>(ℝ<sup>d</sup>) or C<sup>m</sup>(ℝ<sup>d</sup>) for any m ∈ ℕ.

**Definition 52.1.** Let  $0 < \theta \leq 1$ . The function space  $\text{Lip}_\theta(\mathbb{R}^d)$  is the set of all continuous functions for which the seminorm

$$(52.1) \quad \|f\|_{\text{Lip}_\theta} := \sup \left\{ \frac{|f(x) - f(y)|}{|x - y|^\theta} : x, y \in \mathbb{R}^d, x \neq y \right\}$$

is finite.

It is surprising that this norm can be expressed in terms of the average.

**Theorem 52.2.** Let  $0 < \theta \leq 1$ . Suppose that  $f$  is a locally integrable function. Then  $f$  belongs to  $\mathcal{C}^\theta(\mathbb{R}^d)$  after redefining  $f$  on a set of zero measure, if and only if

$$(52.2) \quad \|f\|_{\Lambda^\theta} := \sup_{Q \in \mathcal{Q}} \frac{m_Q(|f - m_Q(f)|)}{|Q|^{\frac{\theta}{d}}}$$

is finite.

*Proof.* It is easy to show

$$(52.3) \quad \|f\|_{\Lambda^\theta} \leq \|f\|_{\mathcal{C}^\theta}$$

for all  $f \in \mathcal{C}^\theta(\mathbb{R}^d)$ . Therefore we concentrate on the reverse inequality.

For the proof of the reverse inequality, we may assume that  $f$  is real-valued. Suppose that  $\|f\|_{\mathcal{C}^\theta} < \infty$ .

Note that, for every cube  $Q$ ,

$$(52.4) \quad |m_Q(f) - m_{2Q}(f)| \leq m_Q(|f - m_{2Q}(f)|) \lesssim m_{2Q}(|f - m_{2Q}(f)|) \lesssim \ell(Q)^\theta.$$

In the same way a geometric observation shows

$$(52.5) \quad |m_Q(f) - m_S(f)| \lesssim \ell(Q)^\theta,$$

whenever  $Q$  and  $S$  are cubes of comparable size with non-empty intersection.

Let  $E$  be the set of all Lebesgue points of  $f$ . Fix  $x, y \in E$  and write

$$(52.6) \quad Q_j := Q(x, 2^{-j+1}|x - y|), \quad R_j := Q(y, 2^{-j+1}|x - y|).$$

Then we have

$$(52.7) \quad f(x) = \lim_{j \rightarrow \infty} m_{Q_j}(f), \quad f(y) = \lim_{j \rightarrow \infty} m_{R_j}(f)$$

from the definition of the Lebesgue point. We deduce from the preceding paragraph

$$|m_{Q_j}(f) - m_{Q_{j+1}}(f)|, |m_{R_j}(f) - m_{R_{j+1}}(f)| \lesssim 2^{-j\theta}|x - y|^\theta, \quad |m_{Q_0}(f) - m_{R_0}(f)| \lesssim |x - y|^\theta.$$

Therefore, we have

$$\begin{aligned} |f(x) - f(y)| &\leq |m_{Q_0}(f) - m_{R_0}(f)| + \sum_{j=1}^{\infty} (|m_{Q_j}(f) - m_{Q_{j+1}}(f)| + |m_{R_j}(f) - m_{R_{j+1}}(f)|) \\ &\lesssim |x - y|^\theta. \end{aligned}$$

Thus, if we define

$$(52.8) \quad g(x) := \limsup_{\substack{z \in E \\ z \rightarrow x}} f(z),$$

we see that  $g$  belongs to  $\mathcal{C}^\theta(\mathbb{R}^d)$ . □

52.2. Hölder-Zygmund spaces with higher regularity.

Now we consider Hölder-Zygmund spaces with higher regularity.

**Definition 52.3.** Let  $m \in \mathbb{N}$  and  $\alpha \in (0, \infty) \setminus \mathbb{N}$ .

- (1) The function space  $C^m$  is the set of all functions whose partial differentials exist and are bounded up to order  $m$ . The norm is given by

$$(52.9) \quad \|f\|_{C^m} := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq m}} \|\partial^\alpha f\|_{BC}$$

for  $f \in C^m$ .

- (2) The function space  $C^\alpha(\mathbb{R}^d)$  is the set of all functions  $f \in C^{[\alpha]}(\mathbb{R}^d)$  such that and  $\partial^\beta f \in \text{Lip}_{\alpha-[\beta]}(\mathbb{R}^d)$  for all multiindices  $\beta$  with length  $m$ . The norm is given by

$$(52.10) \quad \|f\|_{C^\alpha} := \|f\|_{C^m} + \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta|=m}} \|\partial^\beta f\|_{\text{Lip}_{\alpha-[\beta]}}$$

for  $f \in C^\alpha(\mathbb{R}^d)$ .

**Theorem 52.4** (Interpolation inequality). *Let  $m \in \mathbb{N}$ . Then we have*

$$(52.11) \quad \|\partial^\alpha f\|_{BC} \lesssim \|f\|_{BC}^{1-\frac{|\alpha|}{m}} \cdot \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta|=m}} \|\partial^\beta f\|_{BC}^{\frac{|\alpha|}{m}}$$

for all  $f \in C^m(\mathbb{R}^d)$ .

*Proof.* It suffices to prove

$$(52.12) \quad \|\partial^\alpha f\|_{BC} \lesssim \|f\|_{BC} + \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta|=m}} \|\partial^\beta f\|_{BC}.$$

Indeed, once we prove (52.12), we can use the dilation argument: We can replace  $f$  with  $f(t \cdot)$  for  $t > 0$ . Choosing  $t > 0$  suitably, we obtain the desired inequality.

Let  $i = 1, 2, \dots, d$  and  $h > 0$ . Then define

$$(52.13) \quad D_i^h f(x) := \frac{f(x + h e_i) - f(x)}{h}.$$

A repeated application of the mean value theorem yields

$$(52.14) \quad \partial^\alpha f(x + y) = D_{\alpha_1}^1 D_{\alpha_2}^1 \dots D_{\alpha_d}^1 f(x)$$

for some  $y \in Q(1)$ , the unit cube. Therefore, another application of the mean value theorem gives us

$$(52.15) \quad \partial^\alpha f(x) = \partial^\alpha f(x) - \partial^\alpha f(x + y) + \partial^\alpha f(x + y) = \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta|=|\alpha|+1}} c_\beta \partial^\beta f(x + y_\beta) + D_{\alpha_1}^1 D_{\alpha_2}^1 \dots D_{\alpha_d}^1 f(x)$$

for some  $|c_\beta| \leq 1$  and  $y_\beta \in Q(1)$ . In view of this inequality, we obtain

$$(52.16) \quad \|\partial^\alpha f\|_{BC} \lesssim \|f\|_{BC} + \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta|=|\alpha|+1}} \|\partial^\beta f\|_{BC}.$$

This is a bridge that connects the lower derivatives, the highest ones and the original function. A repeated usage of this key inequality gives us the desired result.  $\square$

**Theorem 52.5.** *Let  $0 < \alpha < 1$  or  $1 < \alpha < 2$  and  $f \in \text{BC}(\mathbb{R}^d)$ . Then  $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$ , if and only if*

$$(52.17) \quad \|f\|_{\tilde{\mathcal{C}}^\alpha} := \|f\|_{\text{BC}} + \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{1}{|x - y|^\alpha} \left| f(x) + f(y) - 2f\left(\frac{x + y}{2}\right) \right| < \infty.$$

An axiom of choice, or more precisely the Hamel basis on  $\mathbb{Q}$ , yields an example showing that the assumption  $f \in \text{BC}(\mathbb{R}^d)$  is absolutely necessary in order that  $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$  for all  $f$  with  $\|f\|_{\tilde{\mathcal{C}}^\alpha} < \infty$ . See Exercise 252.

*Proof.* Let  $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$ . It is not so hard to see that

$$(52.18) \quad \|f\|_{\tilde{\mathcal{C}}^\alpha} \lesssim \|f\|_{\mathcal{C}^\alpha}$$

directly from the definition if  $0 < \alpha < 1$  and by using the mean value theorem if  $1 < \alpha < 2$ .

The crux is, of course, to prove the reverse inequality. To do this, assuming  $f \in \tilde{\mathcal{C}}^\alpha(\mathbb{R}^d)$ , let us obtain a key estimate. We write

$$(52.19) \quad g(h) = g(x; h) = f(x + h) - f(x)$$

for a fixed  $x \in \mathbb{R}^d$ . Then from the definition of the norm  $\|\cdot\|_{\tilde{\mathcal{C}}^\alpha}$  we have

$$(52.20) \quad \left| g(h) - 2g\left(\frac{h}{2}\right) \right| = \left| f(x + h) - 2f\left(\frac{h}{2}\right) + f(x) \right| \lesssim |h|^\alpha \|f\|_{\tilde{\mathcal{C}}^\alpha}.$$

Therefore, we obtain a key inequality

$$(52.21) \quad \left| 2^j g\left(\frac{h}{2^j}\right) - 2^{j-1} g\left(\frac{h}{2^{j-1}}\right) \right| \lesssim 2^{j(1-\alpha)} |h|^\alpha \|f\|_{\tilde{\mathcal{C}}^\alpha}, \quad j = 1, 2, \dots$$

by letting  $h$  replaced by  $\frac{h}{2^j}$  in (52.20).

Let  $0 < \alpha < 1$ . In this case we add (52.21) over  $j = 1, 2, \dots, k$ . Then we obtain

$$(52.22) \quad \left| g(h) - 2^k g\left(\frac{h}{2^k}\right) \right| \lesssim 2^{k(1-\alpha)} |h|^\alpha \|f\|_{\tilde{\mathcal{C}}^\alpha}.$$

Suppose that  $y$  is a point in  $B(x, 1) \setminus \{x\}$ . Then we can choose  $k \in \mathbb{N}_0$  so that  $\frac{1}{2} \leq 2^k |x - y| < 1$ .

Let  $h = 2^k(y - x)$ . Then we have

$$(52.23) \quad \left| f(2^k(y - x) + x) - f(x) - 2^k(f(y) - f(x)) \right| \lesssim 2^k |x - y|^\alpha \|f\|_{\tilde{\mathcal{C}}^\alpha}$$

and hence

$$(52.24) \quad |f(y) - f(x)| \lesssim |x - y|^\alpha \|f\|_{\tilde{\mathcal{C}}^\alpha} + 2^{-k} |f(2^k(y - x) + x) - f(x)|.$$

As a consequence, if we arrange this inequality and use the fact that  $k \in \mathbb{N}_0$ , we obtain

$$(52.25) \quad |f(y) - f(x)| \lesssim |x - y|^\alpha (\|f\|_\infty + \|f\|_{\tilde{\mathcal{C}}^\alpha}).$$

If  $|x - y| > 1$ , then this inequality is trivial. Therefore, it follows that  $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$  with the desired norm estimate.

Let  $1 < \alpha < 2$ . The first job to do is proving the existence of all partial derivatives of 1-st order. Choose an even function  $\rho$  with integral 1. Notice that we can even arrange that  $\rho(x) - 2^d \rho(2x) = \Delta \psi(x)$  for some  $\psi \in \mathcal{S}$ . Define a mollifying sequence  $\{f_j\}_{j \in \mathbb{N}}$  of  $f$  by

$$(52.26) \quad f_j(x) := 2^{jn} \rho(2^j \cdot) * f(x) = 2^{jn} \int_{\mathbb{R}^d} \rho(2^j(x - y)) f(y) dy.$$



Then, since  $\rho$  is smooth, we have

$$(52.27) \quad \partial_k f_j(x) = 2^{j(n+1)} \int_{\mathbb{R}^d} \partial_k \rho(2^j(x-y)) f(y) dy = 2^j \int_{\mathbb{R}^d} \partial_k \rho(z) f(x-2^{-j}z) dz.$$

Therefore, we have

$$\begin{aligned} & \partial_k f_j(x) - \partial_k f_{j+1}(x) \\ &= \int_{\mathbb{R}^d} \partial_k \rho(z) \cdot 2^j (f(x-2^{-j}z) - f(x)) dz - \int_{\mathbb{R}^d} \partial_k \rho(z) \cdot 2^{j+1} (f(x-2^{-j-1}z) - f(x)) dz \\ &= \int_{\mathbb{R}^d} \partial_k \rho(z) \cdot (2^j g(x; 2^{-j}z) - 2^{j+1} g(x; 2^{-j-1}z)) dz. \end{aligned}$$

Therefore, the key estimate (52.21) gives

$$(52.28) \quad \|\partial_k f_j - \partial_k f_{j+1}\|_{BC} \lesssim 2^{j(1-\alpha)} \|f\|_{\tilde{C}^\alpha} \cdot \int_{\mathbb{R}^d} |\partial_k \rho(z)| dz = c 2^{j(1-\alpha)} \|f\|_{\tilde{C}^\alpha}.$$

Since  $\alpha > 1$ , this formula shows that  $\partial_k f_j$  is a Cauchy sequence in  $BC(\mathbb{R}^d)$ . Therefore,  $\{f_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence convergent to  $f$  in the topology  $BC^1(\mathbb{R}^d)$ . As a consequence we conclude at least that  $f \in BC^1(\mathbb{R}^d)$ .

Since  $\partial_l \partial_k \rho$  is even, we have

$$(52.29) \quad \partial_l \partial_k f_j(x) = 2^{2j} \int_{\mathbb{R}^d} \partial_l \partial_k \rho(z) f(x-2^{-j}z) dz = 2^{2j} \int_{\mathbb{R}^d} \partial_l \partial_k \rho(z) f(x+2^{-j}z) dz.$$

Therefore, it follows that

$$(52.30) \quad \partial_l \partial_k f_j(x) = 2^{2j} \int_{\mathbb{R}^d} \partial_l \partial_k \rho(z) (f(x+2^{-j}z) + f(x-2^{-j}z) - 2f(x)) dz.$$

As a result we have

$$(52.31) \quad \|\partial_l \partial_k f_j\|_{BC} \lesssim 2^{2j(2-\alpha)} \|f\|_{\tilde{C}^\alpha}.$$

By (52.28) and (52.31), we have

$$(52.32) \quad |\partial_k f_j(x) - \partial_k f_{j+1}(x) - \partial_k f_j(y) + \partial_k f_{j+1}(y)| \lesssim \min(2^{j(2-\alpha)}|x-y|, 2^{j(1-\alpha)}) \|f\|_{\tilde{C}^\alpha}.$$

Since  $\rho$  is smooth, it is easy to see

$$(52.33) \quad |\partial_k f_0(x) - \partial_k f_0(y)| \lesssim \min(|x-y|, 1) \|f\|_{\tilde{C}^\alpha}.$$

If we add (52.32) over  $j \in \mathbb{N}$  and use (52.33), then we obtain the desired estimate. □

*Exercise 252.* Show that by using a Hamel basis, that is, a collection  $\{x_\lambda\}_{\lambda \in \Lambda}$  which is  $\mathbb{Q}$ -linearly independent and  $\mathbb{Q}$ -spans  $\mathbb{R}$  to show that there exists a non-continuous function such that  $f\left(\frac{x+y}{2}\right) = \frac{1}{2}(f(x) + f(y))$ .

**Definition 52.6** ( $\mathcal{C}^\alpha(\mathbb{R}^d)$  for  $\alpha > 0$ ). Let  $\alpha > 0$ . Then define  $\mathcal{C}^\alpha(\mathbb{R}^d)$  to be the set of all functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  such that  $f \in C^{\max([\alpha]-1, 0)}$  and that  $\|f\|_{\mathcal{C}^\alpha}$  redefined in (52.34) is finite.

$$(52.34) \quad \|f\|_{\mathcal{C}^\alpha} := \|f\|_{C^{\max([\alpha]-1, 0)}} + \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta| \leq \max(\alpha-1, 0)}} \|\partial^\beta f\|_{\tilde{C}^{\alpha-[\alpha]+1}}.$$

*Exercise 253.* Let  $\tau \in C^\infty(\mathbb{R}^d)$  be a cutoff function that equals 1 near the origin and 0 sufficiently away from the origin. Set  $F(x) = |x|\tau(x)$ . Then show that  $F \in \tilde{C}^1 \setminus C^1$ .

Note that from Theorem 52.5 the norms (52.17) and (52.34) are equivalent.

52.3. Interpolation of Hölder spaces.

**Theorem 52.7.** *Let  $l, m \in \mathbb{N}$  and  $0 < \theta < 1$ . If  $l < m$ , then we have*

$$(52.35) \quad (C^l(\mathbb{R}^d), C^m(\mathbb{R}^d))_{\theta, \infty} = \mathcal{C}^{l(1-\theta)+m\theta}(\mathbb{R}^d).$$

*Proof.* Some inequalities on convolution. Denote by  $D_h$  the set translation operator  $f \in \text{BC}(\mathbb{R}^d) \mapsto f(\cdot - h) \in \text{BC}(\mathbb{R}^d)$ . Given  $f \in \text{BC}(\mathbb{R}^d)$ , we define

$$(52.36) \quad f_t(x) = \int_{\mathbb{R}^d} \tau(y)[(D_{ty} - 1)^M + 1]f(x) dy,$$

where  $\tau \in C_{c,+}^\infty(\mathbb{R}^d)$  is an even function with integral 1 and  $M$  is sufficiently large. By a change of variables we obtain

$$(52.37) \quad f_t(x) = \int_{\mathbb{R}^d} \rho(y)f(x - ty) dy,$$

where  $\rho \in C_{c,+}^\infty(\mathbb{R}^d)$  is an even function with integral 1.

Assume that  $[l(1 - \theta) + m\theta]$  is even. Suppose that  $f \in \tilde{\mathcal{C}}^{l(1-\theta)+m\theta}(\mathbb{R}^d)$ . Let  $\beta \in \mathbb{N}_0$  such that  $|\beta|$  is an odd integer greater than  $[l(1 - \theta) + m\theta]$ . Let  $\gamma$  be an multiindex less than  $\beta$  whose length is  $[l(1 - \theta) + m\theta] - 1$ . Then we have

$$\begin{aligned} \partial^\beta f_t(x) &= -\partial^{\beta-\gamma} \int_{\mathbb{R}^d} \rho(y)\partial^\gamma f(x - ty) dy \\ &= -t^{-|\beta|+|\gamma|} \int_{\mathbb{R}^d} \partial^{\beta-\gamma} \rho(y)\partial^\gamma f(x - ty) dy \\ &= -\frac{1}{2}t^{-|\beta|+|\gamma|} \int_{\mathbb{R}^d} \partial^{\beta-\gamma} \rho(y) (\partial^\gamma f(x - ty) + \partial^\gamma f(x + ty) - 2\partial^\gamma f(x)) dy \end{aligned}$$

by symmetry of  $\rho$ . As a result we obtain

$$(52.38) \quad \|\partial^\beta f\|_{\text{BC}} \lesssim t^{l(1-\theta)+m\theta-|\beta|} \|f\|_{\tilde{\mathcal{C}}_{m\theta}}$$

for all  $\beta$  such that it is sufficiently large and odd. If  $[l(1 - \theta) + m\theta]$  is odd, then reverse odd and even in the above calculation. The result is

$$(52.39) \quad \|\partial^\beta f\|_{\text{BC}} \lesssim t^{l(1-\theta)+m\theta-|\beta|} \|f\|_{\tilde{\mathcal{C}}_{m\theta}}$$

for all  $\beta$  such that it is sufficiently large and even. In the same way we obtain

$$(52.40) \quad \|f_t - f\|_{C^l} \lesssim t^{l(1-\theta)+m\theta} \|f\|_{\tilde{\mathcal{C}}_{2\theta}},$$

using  $f_t(x) - f(x) = \int_{\mathbb{R}^d} \tau(y)(D_{ty} - 1)^M f(x) dy$ . By the interpolating inequality we obtain

$$(52.41) \quad \|\partial^\beta f\|_{\text{BC}} \lesssim t^{l(1-\theta)+m\theta-|\beta|} \|f\|_{\tilde{\mathcal{C}}_{2\theta}},$$

whether  $[m\theta]$  is even or not. Therefore,

$$(52.42) \quad \|f_t\|_{C^M} \lesssim_M t^{l(1-\theta)+m\theta-M} \|f\|_{\tilde{\mathcal{C}}_{2\theta}}$$

for  $0 < t \leq 1$ .

As a result, we obtain

$$(52.43) \quad K(t, f; C^l, C^m) \leq \|f - f_{m-\sqrt{t}}\|_{C^l} + t\|f_{m-\sqrt{t}}\|_{C^m} \lesssim t^\theta \|f\|_{\tilde{\mathcal{C}}^{l(1-\theta)+m\theta}},$$

if  $0 < t < 1$  and

$$(52.44) \quad K(t, f; C^l, C^m) \leq \|f\|_{C^l} \leq t^\theta \left\| f : \tilde{\mathcal{C}}^{l(1-\theta)+m\theta} \right\|$$

if  $t \geq 1$ . Thus, we conclude  $f \in (C^l, C^m)_{\theta, \infty}$ .

Let us prove the reverse inequality. Let  $f \in (C^l, C^m)_{\theta, \infty}$ . Then from the definition of the  $K$ -functional  $f$  can be split into the sum of a bounded continuous function  $f_0$  and a  $C^m$ -function  $f_1$  satisfying

$$(52.45) \quad \|f_0\|_{C^l} + |x - y|^{m-l} \|f_1\|_{C^m} \leq 2|x - y|^{(m-l)\theta} \|f\|_{(C^l, C^m)_{\theta, \infty}}.$$

With this preparation in mind, we shall prove the reverse inequality by induction on  $[m\theta]$ . Assume first that  $m\theta < 2$ . Then, if we set  $z = \frac{x+y}{2}$ , then we obtain

$$\begin{aligned} & |\partial^\gamma f(x) + \partial^\gamma f(y) - 2\partial^\gamma f(z)| \\ & \leq |\partial^\gamma f_0(x) + \partial^\gamma f_0(y) - 2\partial^\gamma f_0(z)| + |\partial^\gamma f_1(x) + \partial^\gamma f_1(y) - 2\partial^\gamma f_1(z)| \\ & \leq 4\|f\|_{C^l} + |x - y|^{m-l} \|f\|_{C^m} \\ & \leq 8\|f\|_{(C^l, C^m)_{\theta, \infty}}. \end{aligned}$$

Assume that  $k \geq 2$  and

$$(52.46) \quad (C^l(\mathbb{R}^d), C^m(\mathbb{R}^d))_{\theta, \infty} = C^{m\theta}(\mathbb{R}^d)$$

holds for  $\theta$  with  $(m-l)\theta < k$ . Assume that  $\theta$  satisfies  $k \leq (m-l)\theta < k+1$ . Let  $\theta_0 = \frac{k-1}{m-l}$ . Define  $\eta$  by  $\theta = (1-\eta)\theta_0 + \eta$ . Since

$$(52.47) \quad (m-l-k+1)\eta = (m-l)\eta(1-\theta_0) = (m-l)(\theta - \theta_0) < k+1 - (k-1) = 2,$$

we can use the induction assumption to obtain

$$\begin{aligned} (C^l(\mathbb{R}^d), C^m(\mathbb{R}^d))_{\theta, \infty} &= ([C^l(\mathbb{R}^d), C^m(\mathbb{R}^d)]_{\eta_0, \infty}, C^m(\mathbb{R}^d)(\mathbb{R}^d))_{\theta, \infty} \\ &= (C^{m\theta_0}(\mathbb{R}^d), C^m(\mathbb{R}^d))_{\eta, \infty} \\ &\subset C^{m\theta_0(1-\eta)+m\eta}(\mathbb{R}^d) \\ &= C^{m\theta}(\mathbb{R}^d). \end{aligned}$$

This is the desired result. □

### 53. MORREY SPACES

Now we turn to Morrey spaces, which we took up earlier in Exercise 67. Let us recall the definition of the norm.

**Definition 53.1.** Let  $1 \leq q \leq p < \infty$ . For an  $L^q_{\text{loc}}(\mathbb{R}^d)$ -function  $f$  the Morrey norm is defined by

$$(53.1) \quad \|f\|_{\mathcal{M}_q^p} := \sup_{x \in \mathbb{R}^d, r > 0} |B(x, r)|^{\frac{1}{p} - \frac{1}{q}} \left( \int_{B(x, r)} |f(y)|^q dy \right)^{\frac{1}{q}}.$$

The Morrey space  $\mathcal{M}_q^p(\mathbb{R}^d)$  is the set of all  $L^q(\mathbb{R}^d)$ -locally integrable functions  $f$  for which the norm  $\|f\|_{\mathcal{M}_q^p}$  is finite.

*Exercise 254.* Let  $f \in \mathcal{M}_q^p(\mathbb{R}^{d_1})$  and  $g \in \mathcal{M}_q^p(\mathbb{R}^{d_2})$ . Show that  $h = f \otimes g \in \mathcal{M}_q^p(\mathbb{R}^{d_1+d_2})$ , where we have defined  $h(x, y) := f(x)g(y)$ .

*Exercise 255.* Let  $1 \leq q \leq p < \infty$ .

- (1) Let  $B$  be an open ball. Then show that  $\|\chi_B\|_{\mathcal{M}_q^p} = \|\chi_B\|_p$ .
- (2) Let  $g \in L^\infty(\mathbb{R}^d)$ . Show that the operator norm  $f \in \mathcal{M}_q^p(\mathbb{R}^d) \mapsto g \cdot f \in \mathcal{M}_q^p(\mathbb{R}^d)$  is  $\|g\|_\infty$ .

### 53.1. Boundedness of maximal operators.

Now we present a typical argument proving the Morrey-boundedness.

**Theorem 53.2.** *Let  $1 < q \leq p < \infty$ . Then*

$$(53.2) \quad \|Mf\|_{\mathcal{M}_q^p} \lesssim_{p,q} \|f\|_{\mathcal{M}_q^p}$$

for all  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ .

*Proof.* For the proof we have only to show, from the definition, that

$$(53.3) \quad |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf(y)^q dy \right)^{\frac{1}{q}} \lesssim \|f\|_{\mathcal{M}_q^p}, Q \in \mathcal{Q}.$$

Write  $f = f_1 + f_2$ , where  $f_1 = f$  on  $5Q$  and  $f_2 = f$  outside  $5Q$ . The estimate of (53.3) can be decomposed into

$$(53.4) \quad |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf_1(y)^q dy \right)^{\frac{1}{q}} \lesssim \|f\|_{\mathcal{M}_q^p}$$

$$(53.5) \quad |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf_2(y)^q dy \right)^{\frac{1}{q}} \lesssim \|f\|_{\mathcal{M}_q^p}.$$

As we have shown,  $M$  is  $L^q(\mathbb{R}^d)$ -bounded. Thus (53.4) can be shown easily ;

$$\begin{aligned} |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf_1(y)^q dy \right)^{\frac{1}{q}} &\leq |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_{\mathbb{R}^d} Mf_1(y)^q dy \right)^{\frac{1}{q}} \\ &\lesssim |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_{5Q} |f(y)|^q dy \right)^{\frac{1}{q}} \\ &\lesssim |5Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_{5Q} |f(y)|^q dy \right)^{\frac{1}{q}} \\ &\lesssim \|f\|_{\mathcal{M}_q^p}. \end{aligned}$$

To prove (53.5) we have to keep in mind the following fundamental geometric observation.

If  $R \in \mathcal{Q}$  is a cube that meets both  $Q$  and  $\mathbb{R}^d \setminus 5Q$ , then  $\ell(R) \geq 2\ell(Q)$  and  $2R \supset Q$ .

This geometric observation yields

$$(53.6) \quad Mf_2(y) \lesssim \sup_{Q \subset R \in \mathcal{Q}} \frac{1}{|R|} \int_R |f(y)| dy.$$

If we insert this inequality to (53.5), then we obtain

$$|Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf_2(y)^q dy \right)^{\frac{1}{q}} \lesssim |Q|^{\frac{1}{p}} \sup_{Q \subset R \in \mathcal{Q}} \frac{1}{|R|} \int_R |f(y)| dy.$$

Taking into account  $\mathcal{M}_q^p(\mathbb{R}^d) \hookrightarrow \mathcal{M}_1^p(\mathbb{R}^d)$ , we see that

$$(53.7) \quad |Q|^{\frac{1}{p}-\frac{1}{q}} \left( \int_Q Mf_2(y)^q dy \right)^{\frac{1}{q}} \lesssim \sup_{R \in \mathcal{Q}} |R|^{\frac{1}{p}-1} \int_R |f(y)| dy = \|f\|_{\mathcal{M}_1^p} \leq \|f\|_{\mathcal{M}_q^p}.$$

Consequently (53.5) is proved.  $\square$

53.2. Morrey's lemma.

**Theorem 53.3** (Morrey's lemma). *Suppose that  $u$  is a  $C^1$ -function.*

$$(53.8) \quad |u(y) - u(x)| \lesssim r(B)^{1-\frac{d}{p}} \left( \int_{2B} |Du(z)|^p dz \right)^{\frac{1}{p}}$$

for all  $x \in B$ .

*Proof.* Instead of proving (53.8) directly we have only to prove

$$(53.9) \quad \frac{1}{|B|} \int_B |u(y) - u(x)| dy \lesssim \ell(B)^{1-\frac{d}{p}} \left( \int_{2B} |Du(z)|^p dz \right)^{\frac{1}{p}},$$

where  $B$  is a ball with center  $x$ . To prove this, fix any point  $w \in S^{d-1}$ . Then if  $0 < s < r(B)$ ,

$$|u(x + s\omega) - u(x)| = \left| \int_0^s \frac{d}{dt} u(x + t\omega) dt \right| = \left| \int_0^s Du(x + t\omega) \cdot \omega dt \right| \leq \int_0^s |Du(x + t\omega)| dt.$$

Hence

$$\begin{aligned} \int_{S^{d-1}} |u(x + s\omega) - u(x)| d\omega &\leq \int_0^s \left( \int_{S^{d-1}} |Du(x + t\omega)| d\omega \right) dt \\ &= \int_0^s \left( \int_{S^{d-1}} |Du(x + s\omega)| \frac{t^{d-1}}{t^{d-1}} d\omega \right) dt. \end{aligned}$$

Let  $y = x + t\omega$ , so that  $t = |x - y|$ . Then converting from polar coordinates, we have

$$\int_{S^{d-1}} |u(x + s\omega) - u(x)| d\omega \leq \int_{B(x,s)} \frac{|Du(y)|}{|x - y|^{d-1}} dy \leq \int_{B(x,r)} \frac{|Du(y)|}{|x - y|^{d-1}} dy.$$

Multiply by  $s^{d-1}$  and integrate from 0 to  $r(B)$  with respect to  $s$ :

$$(53.10) \quad \int_B |u(y) - u(x)| dy \lesssim r^d \int_B \frac{|Du(y)|}{|x - y|^{d-1}} dy.$$

Therefore,

$$(53.11) \quad \int_B |u(x) - u(y)| dy \lesssim r^d \left( \int_B |Du(y)|^p dy \right)^{\frac{1}{p}} \left( \int_B |x - y|^{-\frac{p(d-1)}{p-1}} dy \right)^{\frac{p-1}{p}} \lesssim r^{1+d-\frac{d}{p}}.$$

This is the desired result. □

53.3. Fractional integral operators.

Morrey spaces, the BMO space and Hölder spaces stand in a line. More precisely, we formulate this as follows:

**Theorem 53.4** (Boundedness of the fractional integral operators  $I_\alpha$  and the modified fractional integral operators  $\tilde{I}_\alpha$ ). *Suppose that  $1 \leq q \leq p < \infty$  and  $0 < \alpha < d$ . Define*

$$(53.12) \quad I_\alpha f(x) := \int_{\mathbb{R}^d} \frac{f(y)}{|x - y|^{d-\alpha}} dy, \quad \tilde{I}_\alpha f(x) := \int_{\mathbb{R}^d} \left( \frac{1}{|x - y|^{d-\alpha}} - \frac{\chi_{Q_0^c}(y)}{|x_0 - y|^{d-\alpha}} \right) f(y) dy,$$

where  $Q_0$  is a fixed cube centered at  $x_0$ .

(1) (Subcritical case) *Let  $1 < q \leq p < \frac{d}{\alpha}$ . Assume that the parameters  $s, t$  satisfy*

$$(53.13) \quad 1 < t \leq s < \infty, \quad \frac{1}{s} = \frac{1}{p} - \frac{\alpha}{d}, \quad \frac{t}{s} = \frac{q}{p}.$$

Then

$$(53.14) \quad \|I_\alpha f\|_{\mathcal{M}_t^q} \lesssim \|f\|_{\mathcal{M}_s^q}$$

for every (positive)  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ . The same can be said for  $M_\alpha$ , where  $M_\alpha$  is given by (17.19).

(2) (Critical case) Assume that  $1 \leq q \leq p = \frac{d}{\alpha}$ . Then

$$(53.15) \quad \|\tilde{I}_\alpha f\|_* \lesssim \|f\|_{\mathcal{M}_q^p}$$

for every  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ , where  $\|\cdot\|_*$  denotes the  $\text{BMO}(\mathbb{R}^d)$ -norm given by

$$(53.16) \quad \|g\|_* = \sup_{Q \in \mathcal{Q}} m_Q(|g - m_Q(g)|).$$

(3) (Supercritical case) Assume that  $0 < \alpha - \frac{d}{p} < 1$  and that  $1 \leq q \leq p < \infty$ . Then

$$(53.17) \quad \left\| \tilde{I}_\alpha f \right\|_{\text{Lip}_{\alpha - \frac{d}{p}}} \lesssim \|f\|_{\mathcal{M}_q^p}$$

for every  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ , where we denoted

$$(53.18) \quad \|g\|_{\text{Lip}(\theta)} = \sup \left\{ \frac{|g(x) - g(y)|}{|x - y|^\theta} : x, y \in \mathbb{R}^d, x \neq y \right\}, \quad 0 < \theta < 1$$

for a function  $g$ .

Here we content ourselves with proving (2), other results being proved similarly. The assertion for  $M_\alpha$  follows immediately from that for  $I_\alpha$ . So we omit the proof.

*Proof.* We have to prove

$$(53.19) \quad m_Q(|\tilde{I}_\alpha f - m_Q(\tilde{I}_\alpha f)|) \lesssim \|f\|_{\mathcal{M}_q^{d/\alpha}}$$

for all cubes  $Q$ . For the proof we may assume  $q < \frac{d}{\alpha}$  because we always have

$$(53.20) \quad \|f\|_{\mathcal{M}_q^{d/\alpha}} \leq \|f\|_{\mathcal{M}_{d/\alpha}^{d/\alpha}} = \|f\|_{L^{d/\alpha}}.$$

We decompose  $f$  according to  $2Q$  as usual. That is, we split  $f = f_1 + f_2$  with  $f_1 = \chi_{2Q} \cdot f$  and  $f_2 = f - f_1$ . By virtue of the triangle inequality our present task is partitioned into proving

$$(53.21) \quad m_Q(|\tilde{I}_\alpha f_1 - m_Q(\tilde{I}_\alpha f_1)|) + m_Q(|\tilde{I}_\alpha f_2 - m_Q(\tilde{I}_\alpha f_2)|) \lesssim \|f\|_{\mathcal{M}_q^{d/\alpha}}.$$

Then the estimate for  $f_1$  is simple. Indeed, to estimate  $f_1$ , we define an auxiliary index  $w \in (q, \infty)$  by  $\frac{1}{w} = \frac{1}{q} - \frac{\alpha}{d}$ . By the triangle inequality and the Hölder inequality, we have

$$(53.22) \quad m_Q(|I_\alpha f_1 - m_Q(I_\alpha f_1)|) \leq 2m_Q(|I_\alpha f_1|) \leq 2m_Q^{(w)}(|I_\alpha f_1|),$$

where we wrote

$$(53.23) \quad m_Q^{(w)}(F) := \left( \frac{1}{|Q|} \int_Q F(x)^w dx \right)^{\frac{1}{w}}$$

for positive measurable functions  $F$ . By using the  $L^q(\mathbb{R}^d)$ - $L^w(\mathbb{R}^d)$  boundedness of the fractional integral operator we obtain

$$m_Q(|I_\alpha f_1 - m_Q(I_\alpha f_1)|) \lesssim \left( \frac{1}{|Q|} \int_{\mathbb{R}^d} |I_\alpha f_1(x)|^w dx \right)^{1/w} \lesssim m_{2Q}(|f|^q)^{1/w} \lesssim \|f\|_{\mathcal{M}_q^{d/\alpha}}.$$

For the proof of the second inequality, we write the left-side out in full.

$$(53.24) \quad m_Q(|I_\alpha f_2 - m_Q(I_\alpha f_2)|) = \frac{1}{|Q|^2} \int_Q \left| \iint_{Q \times (\mathbb{R}^d \setminus 2Q)} \left( \frac{f(z)}{|x - z|^{d-\alpha}} - \frac{f(z)}{|y - z|^{d-\alpha}} \right) dy dz \right| dx.$$

First, we bound the right-hand side with the triangle inequality and arrange it. Then the right-hand side is majorized by

$$(53.25) \quad \frac{1}{|Q|^2} \iiint_{Q \times Q \times (\mathbb{R}^d \setminus 2Q)} |f(z)| \cdot \left| \frac{1}{|x-z|^{d-\alpha}} - \frac{1}{|y-z|^{d-\alpha}} \right| dx dy dz.$$

By virtue of the mean value theorem, we have

$$(53.26) \quad \left| \frac{1}{|x-z|^{d-\alpha}} - \frac{1}{|y-z|^{d-\alpha}} \right| \lesssim \frac{|x-y|}{|z-z_Q|^{d-\alpha+1}} \leq \frac{\ell(Q)}{|z-z_Q|^{d-\alpha+1}}.$$

Thus, inserting this inequality gives us

$$(53.27) \quad m_Q(|I_\alpha f_2 - m_Q(I_\alpha f_2)|) \lesssim \ell(Q) \int_{\mathbb{R}^d \setminus 2Q} \frac{|f(z)|}{|z-z_Q|^{d-\alpha+1}} dz.$$

By the comparison lemma, the integral of the right-hand side is bounded by

$$(53.28) \quad \int_{2\ell(Q)}^\infty \left( \int_{B(z_Q, \ell)} |f(z)| dz \right) \frac{d\ell}{\ell^{d-\alpha+2}} \lesssim \|f\|_{\mathcal{M}_1^{\frac{d}{\alpha}}} \cdot \int_{\ell(Q)}^\infty \frac{d\ell}{\ell^2} = \ell(Q)^{-1} \|f\|_{\mathcal{M}_1^{\frac{d}{\alpha}}} \leq \ell(Q)^{-1} \|f\|_{\mathcal{M}_q^{\frac{d}{\alpha}}}.$$

Thus, the estimate of the second inequality is complete and the proof is concluded. □

**Remark 53.5.** In the critical case and the subcritical case, we need to use the modified fractional maximal operators. However, if we assume that  $f \in L_c^\infty(\mathbb{R}^d)$ , then we can use the usual fractional maximal operators.

*Exercise 256.* Keep to the same condition as above.

- (1) Prove (1). Hint: Reexamine the argument of the assertion corresponding to  $L^p(\mathbb{R}^d)$ . We can prove

$$(53.29) \quad I_\alpha f(x) \lesssim Mf(x)^{\frac{p}{s}} \|f\|_{\mathcal{M}_1^{\frac{p}{s}}}^{1-\frac{p}{s}} \leq Mf(x)^{\frac{p}{s}} \|f\|_{\mathcal{M}_q^{\frac{p}{s}}}^{1-\frac{p}{s}}.$$

- (2) Prove (3) by mimicking the proof of (2).

**53.4. Singular integral operators.**

Finally we extend the domain of singular integral operators to Morrey spaces. To do this, we use duality. Note that the Morrey space  $\mathcal{M}_q^p(\mathbb{R}^d)$  contains  $L^p(\mathbb{R}^d)$ . Therefore, it is yet for  $T$  to be defined on  $\mathcal{M}_q^p(\mathbb{R}^d)$ . Although the structure of the dual space of  $\mathcal{M}_q^p(\mathbb{R}^d)$  is not known, the predual, whose dual space is  $\mathcal{M}_q^p(\mathbb{R}^d)$ , is known. Note that the operation taking dual reverses the inclusion:  $X \subset Y$  implies  $X^* \supset Y^*$ . Therefore, we are to define singular integral by means of duality. We introduce the predual spaces via blocks, whose definition dates back to [497].

**Definition 53.6.** Let  $1 \leq p \leq q < \infty$ . A measurable function  $A$  is said to be a  $(p, q)$ -block if there is a cube  $Q$  that supports  $A$  and

$$(53.30) \quad \|A\|_q \leq |Q|^{\frac{1}{q} - \frac{1}{p}}.$$

As is easily verified by Hölder's inequality, any  $(p, q)$ -block has  $L^p(\mu)$  norm less than 1.

**Definition 53.7.** Let  $1 \leq p \leq q < \infty$ . Then define a function space  $\mathcal{H}_q^p(\mathbb{R}^d)$  by

$$(53.31) \quad \mathcal{H}_q^p := \left\{ f \in L^p(\mathbb{R}^d) : f = \sum_{j \in \mathbb{N}_0} \lambda_j A_j, \{\lambda_j\} \in \ell^1, A_j \text{ is a } (p, q)\text{-block for all } j \in \mathbb{N}_0 \right\}.$$

Define  $\|f\|_{\mathcal{H}_q^p}$  for  $f \in \mathcal{H}_q^p(\mathbb{R}^d)$  as

$$(53.32) \quad \|f\|_{\mathcal{H}_q^p} = \inf_{\lambda} \|\lambda\|_{\ell^1},$$

where  $\lambda$  runs over the admissible expression

$$(53.33) \quad f = \sum_{j \in \mathbb{N}_0} \lambda_j A_j, \quad \{\lambda_j\} \in \ell^1, \quad A_j \text{ is a } (p, q)\text{-block for all } j \in \mathbb{N}_0.$$

Note that  $\mathcal{H}_q^p(\mathbb{R}^d)$  is embedded continuously into  $L^p(\mathbb{R}^d)$  by its definition. The dual of  $\mathcal{H}_q^p(\mathbb{R}^d)$  is  $\mathcal{M}_{q'}^{p'}(\mathbb{R}^d)$  in the following sense.

**Theorem 53.8.** *Suppose that  $1 < q \leq p < \infty$ .*

- (1) *Let  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ . Then the mapping  $L_f : g \mapsto \int_{\mathbb{R}^d} f(x)g(x) dx$  defines a continuous functional on  $\mathcal{H}_{q'}^{p'}$ .*
- (2) *Conversely every continuous functional  $L$  on  $\mathcal{H}_{q'}^{p'}$  can be realized with  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ .*
- (3) *The correspondence  $f \in \mathcal{M}_q^p(\mathbb{R}^d) \mapsto L_f \in (\mathcal{H}_{q'}^{p'}(\mathbb{R}^d))'$  is an isomorphism. Furthermore*

$$(53.34) \quad \|f\|_{\mathcal{M}_q^p} = \sup_{g \in \mathcal{H}_{q'}^{p'} \setminus \{0\}} \frac{1}{\|g\|_{\mathcal{H}_{q'}^{p'}}} \left| \int f \cdot g \right|, \quad \|g\|_{\mathcal{H}_{q'}^{p'}} = \sup_{f \in \mathcal{M}_q^p \setminus \{0\}} \frac{1}{\|f\|_{\mathcal{M}_q^p}} \left| \int f \cdot g \right|$$

*Proof.* The proof of (1) is simple so we omit it. In fact it is straightforward to show that  $|L_f|_* \leq \|f\|_{\mathcal{M}_q^p}$ .

Take a cube  $Q_0$  and let  $Q_j = 2^j Q_0$ . For the sake of the simplicity we denote  $L^{q'}(Q_j)$  by the set of  $L^{q'}$  functions supported on  $Q_j$ . Since we can regard the element in  $L^{q'}(Q_j)$  as a  $(p', q')$ -block modulo multiplicative constant, the functional  $g \mapsto L(g)$  is well defined and bounded on  $L^{q'}(Q_j)$ . Thus Riesz's representation theorem we can find  $f_j$  such that  $L(g) = \int_{Q_j} f_j \cdot g$  for all  $g \in L^{q'}(Q_j)$ . By the uniqueness of this theorem we can find an  $L_{loc}^q(\mathbb{R}^d)$  function  $f$  such that  $f|_{Q_j} = f_j$  a.e..

We shall prove that  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ . For this purpose we take  $Q$  and estimate

$$(53.35) \quad I := |Q|^{\frac{1}{p} - \frac{1}{q}} \left( \int_Q |f|^q \right)^{\frac{1}{q}}$$

For a fixed cube  $Q$  and a fixed function  $f$  we set  $g(x) := \chi_Q(x) \operatorname{sgn}(f(x)) |f(x)|^{q-1}$ . Then we can write

$$(53.36) \quad I = |Q|^{\frac{1}{p} - \frac{1}{q}} \left( \int_Q f \cdot g \right)^{\frac{1}{q}} = |Q|^{\frac{1}{p} - \frac{1}{q}} (L(g))^{\frac{1}{q}}.$$

Notice that  $\frac{|Q|^{\frac{1}{p} - \frac{1}{q}}}{\|g\|_{q'}} g$  is a  $(p', q')$ -atom. Hence we have  $|L(g)| \leq |L|_* |Q|^{-\frac{1}{p} + \frac{1}{q}} \|g\|_{q'}$ . As a result we have  $I \leq |L|_*$ . This is the desired result. The proof of (3) is included in the proofs of (1) and (2). □

Since  $\mathcal{H}_q^p(\mathbb{R}^d)$  is a subset of  $L^p(\mathbb{R}^d)$ , the domain of a singular integral operator  $T$  contains  $\mathcal{H}_q^p(\mathbb{R}^d)$ .

**Theorem 53.9.** *Let  $1 < p \leq q < \infty$ . Let  $T$  be a singular integral operator. Then  $T$ , which is defined a priori on  $L^p(\mathbb{R}^d)$ , is bounded on  $\mathcal{H}_q^p(\mathbb{R}^d)$ .*



*Proof.* We have only to prove the assertion in the block level. That is, given a block  $a$ , which is supported on a cube  $Q$  and satisfies  $\|a\|_q \leq |Q|^{\frac{1}{q}-\frac{1}{p}}$ , we need only show

$$(53.37) \quad \|Ta\|_{\mathcal{H}_q^p} \lesssim 1.$$

Note that  $T$  is a bounded operator on  $L^q(\mathbb{R}^d)$  and hence  $\chi_{2Q} \cdot Ta$  is a  $(p, q)$ -block modulo multiplicative constants. Furthermore

$$(53.38) \quad |\chi_{\mathbb{R}^d \setminus 2Q}(x) \cdot Ta(x)| \lesssim \frac{|Q|^{1-\frac{1}{p}} \chi_{\mathbb{R}^d \setminus 2Q}(x)}{|x - c(Q)|^d} \lesssim \sum_{j=1}^{\infty} 2^{-j(1-\frac{1}{p})} \chi_{2^j Q}(x) \cdot |2^j Q|^{-\frac{1}{p}}.$$

With this decomposition, we conclude that  $\|Ta\|_{\mathcal{H}_q^p} \lesssim 1$ . □

**Definition 53.10.** Let  $1 < q \leq p < \infty$ . Given a singular operator  $T$ , extend it to  $\mathcal{M}_q^p(\mathbb{R}^d)$  by the formula

$$(53.39) \quad \int_{\mathbb{R}^d} Tf(x)g(x) dx = \int_{\mathbb{R}^d} f(x)T^*g(x) dx$$

for all  $g \in \mathcal{H}_{q'}^{p'}$ , where  $T^*$  is a formal adjoint whose kernel is given by  $(x, y) \rightarrow K(y, x)$ .

**Theorem 53.11.** *Suppose that  $1 < q \leq p < \infty$ . Let  $T$  be a singular integral operator. Then  $T$  is bounded on  $\mathcal{M}_q^p(\mathbb{R}^d)$ .*

*Proof.* By the Hahn-Banach theorem, for all  $f \in \mathcal{M}_q^p(\mathbb{R}^d)$ , we can find  $g \in \mathcal{H}_{q'}^{p'}(\mathbb{R}^d)$  with unit norm such that

$$(53.40) \quad \|Tf\|_{\mathcal{M}_q^p} = \int Tf \cdot g.$$

By using the formal adjoint, we obtain

$$(53.41) \quad \|Tf\|_{\mathcal{M}_q^p} = \int f \cdot T^*g \leq \|f\|_{\mathcal{M}_q^p} \cdot \|T^*g\|_{\mathcal{H}_{q'}^{p'}}.$$

Now that  $T^*$  is bounded on  $\mathcal{H}_{q'}^{p'}$ , we see, taking into account that  $g$  has unit norm, that

$$(53.42) \quad \|Tf\|_{\mathcal{M}_q^p} = \int f \cdot T^*g \lesssim \|f\|_{\mathcal{M}_q^p},$$

which is the desired result. □

## 54. COMMUTATORS

In this section we take up an operator of the form

$$(54.1) \quad [a, T]f(x) = \int_{\mathbb{R}^d} (a(x) - a(y))K(x, y)f(y) dy,$$

where  $T$  is a singular integral operator whose kernel is  $K$  and  $a \in \text{BMO}(\mathbb{R}^d)$ . Such an operator  $[a, T]$  is called a commutator. Assume in addition that  $Kf(x) := \int_{\mathbb{R}^d} K(x, y)f(y) dy$  defines an  $L^p(\mathbb{R}^d)$ -bounded operator. The aim of this section is to apply the sharp maximal inequality (see Theorem 20.5) to prove the boundedness of this operator. Because of the singularity of the kernel of  $[a, T]$ , we have to begin with the task of making clear what (54.1) means.

### 54.1. Commutators generated by BMO and singular integral operators.

Keeping in mind that we are going to use the sharp maximal operator  $M^\sharp$  (see (20.1) for its definition), let us fix a cube  $Q$ . Although  $\text{BMO}(\mathbb{R}^d)$  differs from  $L^\infty(\mathbb{R}^d)$ , it is still quite close to  $L^\infty(\mathbb{R}^d)$ . Once we choose a cube  $Q$ , any element in  $\text{BMO}(\mathbb{R}^d)$  is  $L^u(dx|_Q)$ , for all  $u < \infty$  because of the John Nirenberg inequality. In view of this observation we are tempted to rewrite (54.1) as follows:

$$(54.2) \quad [a, T]f(x) = (a(x) - m_Q(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy - \int_{\mathbb{R}^d} K(x, y)(a(y) - m_Q(a))f(y) dy.$$

The first term of this formula seems nice, because  $K$  is assumed  $L^p(\mathbb{R}^d)$ -bounded. However, we are still at a loss because  $(a - m_Q(a))f$  does not belong to  $L^p(\mathbb{R}^d)$ . Thus, we have to work more. We decompose the above formula once more;

$$\begin{aligned} [a, T]f(x) &= (a(x) - m_Q(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy \\ &\quad - T[(a - m_Q(a))\chi_{2Q} \cdot f](x) - \int_{\mathbb{R}^d \setminus 2Q} K(x, y)(a(y) - m_Q(a))f(y) dy. \end{aligned}$$

Now that  $x \in Q$  and  $(a - m_Q(a))\chi_{2Q} \cdot f \in L^q(\mathbb{R}^d)$  for some  $1 < q < p$  by virtue of the John-Nirenberg inequality, it looks nice. The integral of the third term does not contain the singularity. Speaking more precisely, we have

$$\begin{aligned} \int_{\mathbb{R}^d \setminus 2Q} |K(x, y)(a(y) - m_Q(y))f(y)| dy &\lesssim \sum_{j=1}^{\infty} \frac{1}{(2^j \ell(Q))^d} \int_{2^{j+1}Q \setminus 2^jQ} |a(y) - m_Q(y)| \cdot |f(y)| dy \\ &\lesssim \sum_{j=1}^{\infty} \frac{\|f\|_p}{(2^j \ell(Q))^d} \left( \int_{2^{j+1}Q \setminus 2^jQ} |a(y) - m_Q(y)|^{p'} dy \right)^{\frac{1}{p'}} \\ &\lesssim \sum_{j=1}^{\infty} \frac{j}{(2^j \ell(Q))^{\frac{d}{p}}} \|a\|_* \|f\|_p \\ &\lesssim \frac{\|a\|_* \cdot \|f\|_p}{\ell(Q)^{\frac{d}{p}}}. \end{aligned}$$

Therefore, this is an appropriate candidate of the definition of  $[a, T]f$ .

To justify the observation above, we shall prove the following lemma.

**Lemma 54.1.** *Given a cube  $Q$ , we set*

$$\begin{aligned} S_Q f(x) &:= (a(x) - m_Q(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy \\ &\quad - T[(a - m_Q(a))\chi_{2Q} \cdot f](x) - \int_{\mathbb{R}^d \setminus 2Q} K(x, y)(a(y) - m_Q(a))f(y) dy. \end{aligned}$$

*Then if  $Q$  and  $R$  are cubes with  $Q \subset R$ , then we have*

$$(54.3) \quad S_Q f(x) = S_R f(x)$$

*for almost every  $x \in Q(\subset R)$ .*

*Proof.* It is not so hard to see

$$\begin{aligned} S_Q f(x) &= (a(x) - m_R(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy \\ &\quad - T[(a - m_R(a))\chi_{2Q} \cdot f](x) - \int_{\mathbb{R}^d \setminus 2Q} K(x, y)(a(y) - m_R(a))f(y) dy. \end{aligned}$$

Therefore we have

$$S_Q f(x) - S_R f(x) = T[(a - m_R(a))\chi_{2R \setminus 2Q} \cdot f](x) - \int_{2R \setminus 2Q} K(x, y)(a(y) - m_R(a))f(y) dy.$$

By the kernel condition this is equal to 0. The lemma is therefore proved.  $\square$

**Definition 54.2.** Set  $Q(j) := \{x \in \mathbb{R}^d : \max(|x_1|, |x_2|, \dots, |x_d|) \leq j\}$  for  $j \in \mathbb{N}$ . Given  $a \in \text{BMO}(\mathbb{R}^d)$  and a singular integral operator  $T$  with associated kernel  $K$ , define

$$[a, T]f(x) := \lim_{j \rightarrow \infty} (a(x) - m_{Q(j)}(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy - T[(a - m_{Q(j)}(a))\chi_{2Q(j)} \cdot f](x) - \int_{\mathbb{R}^d \setminus 2Q(j)} K(x, y)(a(y) - m_{Q(j)}(a))f(y) dy.$$

Note that the above definition makes sense because the limit exists for a.e.  $x \in \mathbb{R}^d$ . In view of Lemma 54.1, the following proposition is clear.

**Proposition 54.3.** *If  $R$  is a cube, then we have*

$$[a, T]f(x) = (a(x) - m_R(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy - T[(a - m_R(a))\chi_{2R} \cdot f](x) - \int_{\mathbb{R}^d \setminus 2R} K(x, y)(a(y) - m_R(a))f(y) dy$$

for a.e.  $x \in R$ .

Now we prove the boundedness of commutators, where we present an application of the sharp maximal inequality.

**Theorem 54.4.** *Let  $1 < p < \infty$ . Then  $[a, T]$  is  $L^p(\mathbb{R}^d)$ -bounded and it satisfies the norm estimate*

$$(54.4) \quad \|[a, T]\|_{L^p \rightarrow L^p} \lesssim_T \|a\|_*.$$

*Proof.* We may assume that  $f \in L_c^\infty(\mathbb{R}^d)$ . Now let us estimate of  $M^\sharp([a, T]f)$ . To do this, we use the decomposition of  $[a, T]f$  for a given cube  $Q$  above. First we write

$$\begin{aligned} F_1(x) &:= (a(x) - m_Q(a)) \int_{\mathbb{R}^d} K(x, y)f(y) dy \\ F_2(x) &:= T[(a - m_Q(a))\chi_{2Q} \cdot f](x) \\ F_3(x) &:= \int_{\mathbb{R}^d \setminus 2Q} K(x, y)(a(y) - m_R(a))f(y) dy \end{aligned}$$

for the sake of simplicity. Let  $1 < r < p$  be an auxiliary parameter fixed throughout.

We treat  $F_1$ . Let us recall that we have been writing

$$(54.5) \quad m_Q^{(t)}(G) = \left( \frac{1}{|Q|} \int_Q G(x)^t dx \right)^{\frac{1}{t}}, \quad M^{(t)}F(x) = M[|F|^t](x)^{\frac{1}{t}}$$

for  $t > 0$  and measurable functions  $F, G$  with  $G \geq 0$ . Under this notation we have

$$(54.6) \quad m_Q(|F_1 - m_Q(F_1)|) \leq 2m_Q(|F_1|) \leq m_Q^{(r')}(|a - m_Q(a)|) \cdot m_Q^{(r)}(|Kf|) \lesssim \|a\|_* M^{(r)}[Kf](x).$$

As for  $F_2$ , we use the  $L^{\sqrt{r}}$ -boundedness of  $T$  as well.

$$m_Q(|F_2 - m_Q(F_2)|) \leq 2m_Q(|F_2|) \leq 2m_Q^{(r)}(|F_2|) \leq \frac{2}{|Q|} \|T[(a - m_Q(a))\chi_{2Q} \cdot f]\|_{\sqrt{r}}.$$

We now make use of the John-Nirenberg inequality :

$$m_Q(|F_2 - m_Q(F_2)|) \lesssim m_{2Q}^{(\sqrt{r})}(|(a - m_Q(a))f|) \lesssim \|a\|_* M^{(r)} f(x).$$

Finally for the estimate of  $F_3$ , we take full advantage of the oscillation property of  $M^\sharp$ . Note that if  $x, y \in Q$  then we have

$$(54.7) \quad |F_3(x) - F_3(y)| \lesssim \int_{\mathbb{R}^d \setminus 2Q} |K(x, z) - K(x, y)| \cdot |a(z) - m_R(a)| \cdot |f(z)| dz.$$

Let  $\ell \leq 2\ell(Q)$ . Then, by the Hölder condition of the kernel  $K$ ,

$$(54.8) \quad \frac{1}{|z - c_Q|^{d+1}} = (d+1) \int_0^\infty \frac{\chi_{B(c_Q, \ell)}(z) d\ell}{\ell^{d+2}},$$

and the fact that  $\mathbb{R}^d \setminus 2Q \cap B(x, \ell) = \emptyset$ , we have

$$\begin{aligned} |F_3(x) - F_3(y)| &\lesssim \ell(Q) \int_{\mathbb{R}^d \setminus 2Q} \frac{|a(z) - m_R(a)| \cdot |f(z)|}{|z - c_Q|^{d+1}} dz \\ &\lesssim \ell(Q) \int_{2\ell(Q)}^\infty \left( \frac{1}{\ell^{d+2}} \int_{B(c_Q, \ell)} |a(z) - m_R(a)| \cdot |f(z)| dz \right) d\ell. \end{aligned}$$

If we invoke the John-Nirenberg inequality, then we obtain

$$\begin{aligned} &|F_3(x) - F_3(y)| \\ &\lesssim \ell(Q) \int_{2\ell(Q)}^\infty \left\{ \frac{1}{\ell^{d+2}} \left( \int_{B(c_Q, \ell)} |a(z) - m_R(a)|^{r'} dz \right)^{\frac{1}{r'}} \left( \int_{B(c_Q, \ell)} |f(z)|^r dz \right)^{\frac{1}{r}} \right\} d\ell \\ &\lesssim \ell(Q) \|a\|_* \cdot M^{(r)} f(x) \int_{2\ell(Q)}^\infty \log \left( 2 + \frac{\ell}{\ell(Q)} \right) \frac{d\ell}{\ell^2} \\ &\lesssim \|a\|_* M^{(r)} f(x), \end{aligned}$$

which leads us to a key pointwise estimate

$$(54.9) \quad M^\sharp([a, T]f)(x) \lesssim \|a\|_* \left( M^{(r)}[Tf](x) + M^{(r)}f(x) \right).$$

Now that  $\min(|[a, T]f|, 1) \in L^p(\mathbb{R}^d)$ , we are in the position of applying the sharp maximal inequality to obtain

$$(54.10) \quad \|[a, T]f\|_p \lesssim \|M^\sharp([a, T]f)\|_p.$$

Since  $M$  and  $T$  are bounded on  $L^q(\mathbb{R}^d)$  with  $1 < q < \infty$ , it follows that

$$(54.11) \quad \|M^\sharp([a, T]f)\|_p \lesssim \|a\|_* \left( \|M[|Tf|^r]\|_{p/r'} + \|M[|f|^r]\|_{p/r'} \right) \lesssim \|a\|_* \|f\|_p.$$

Putting together (54.10) and (54.11), we obtain the desired estimate. □

### 54.2. Compactness.

Now we consider commutators which are compact.

**Definition 54.5.** The vanishing mean oscillation space VMO is the set of all  $BMO(\mathbb{R}^d)$  functions  $f$  for which there exists a sequence  $\{f_j\}_{j \in \mathbb{N}}$  of  $C_c^\infty(\mathbb{R}^d)$  that approximates  $f$  in the  $BMO(\mathbb{R}^d)$  norm :

$$(54.12) \quad \lim_{j \rightarrow \infty} \|f - f_j\|_{BMO(\mathbb{R}^d)} = 0.$$

**Theorem 54.6.** Let  $a \in VMO$  and  $T$  a singular integral operator. Then  $[a, T]$  is an  $L^p(\mathbb{R}^d)$ -compact operator.

*Proof.* Reduction Let  $\{a_j\}_{j \in \mathbb{N}}$  be a sequence of compactly supported functions such that  $\lim_{j \rightarrow \infty} \|a - a_j\|_{\text{BMO}(\mathbb{R}^d)} = 0$ . Then we have

$$(54.13) \quad \|[a, T] - [a_j, T]\|_{B(L^p)} \lesssim \|a - a_j\|_{\text{BMO}(\mathbb{R}^d)} \rightarrow 0$$

as  $j \rightarrow \infty$ . Therefore, it follows that  $a$  can be assumed compact.

Truncation procedure 1 Define

$$(54.14) \quad [a, T]_\varepsilon f(x) := \int_{\mathbb{R}^d \setminus B(x, \varepsilon)} (a(x) - a(y))K(x, y)f(y) dy.$$

Then, since we are assuming  $a \in C_c^\infty(\mathbb{R}^d)$ , we have

$$|([a, T] - [a, T]_\varepsilon)f(x)| = \int_{B(x, \varepsilon)} |(a(x) - a(y))K(x, y)f(y)| dy \lesssim \int_{B(x, \varepsilon)} \frac{|f(y)|}{|x - y|^{d-1}} dy.$$

Assume for the time being that  $d \geq 2$ . Now we use the comparison lemma to obtain

$$\begin{aligned} \int_{B(x, \varepsilon)} \frac{|f(y)|}{|x - y|^{d-1}} dy &= (d - 1) \int_0^\infty \frac{1}{\ell^d} \left( \int_{B(x, \min(\varepsilon, \ell))} |f(y)| dy \right) d\ell \\ &\leq (d - 1) \int_0^\infty \frac{|B(x, \min(\varepsilon, \ell))|}{\ell^d} Mf(x) d\ell \\ &\lesssim \varepsilon Mf(x). \end{aligned}$$

Therefore, this pointwise estimate together with the  $L^p(\mathbb{R}^d)$ -boundedness of the maximal operator immediately leads us to

$$(54.15) \quad \|[a, T] - [a, T]_\varepsilon\|_{B(L^p)} \lesssim \varepsilon.$$

Therefore,  $[a, T]_\varepsilon$  approximates  $[a, T]$  in the norm topology. As a consequence the matters are reduced to showing the compactness of  $[a, T]_\varepsilon$ .

Truncation procedure 2 Let  $R > \varepsilon$ . Then define

$$(54.16) \quad [a, T]_\varepsilon^R f(x) := \int_{B(x, R) \setminus B(x, \varepsilon)} (a(x) - a(y))K(x, y)f(y) dy.$$

Set  $S := \text{supp}(a)$  and assume that  $R$  is large enough as to be  $B(x, R) \subset S$ . Then, using a rough estimate  $|a(x) - a(y)| \lesssim (\chi_S(x) + \chi_S(y))$  we have

$$\begin{aligned} |[a, T]_\varepsilon f(x) - [a, T]_\varepsilon^R f(x)| &= \int_{\mathbb{R}^d \setminus B(x, R)} |(a(x) - a(y))K(x, y)f(y)| dy \\ &\lesssim \int_{\mathbb{R}^d \setminus B(x, R)} \frac{\chi_S(x) + \chi_S(y)}{|x - y|^d} |f(y)| dy \\ &\sim \chi_S(x) \int_{\mathbb{R}^d} \frac{|f(y)|}{(R + |x - y|)^d} dy \\ &\sim \chi_S(x) \|f\|_p \left( \int_{\mathbb{R}^d} \frac{dy}{(R + |x - y|)^{dp'}} \right)^{\frac{1}{p'}} \\ &\simeq R^{-\frac{d}{p}} \chi_S(x) \|f\|_p. \end{aligned}$$

As a consequence

$$(54.17) \quad \|[a, T]_\varepsilon - [a, T]_\varepsilon^R\|_{B(L^p)} \lesssim R^{-\frac{d}{p}}.$$

Letting  $R \rightarrow \infty$ , we see that  $[a, T]_\varepsilon^R$  approximates  $[a, T]_\varepsilon$  in the norm topology. Therefore, all we have to prove is the compactness of  $[a, T]_\varepsilon^R$ . However, the kernel of this operator is compactly supported and bounded. Therefore, it is true that  $[a, T]_\varepsilon^R$  is compact.  $\square$

### 54.3. Another type of commutators.

Commutators generated by  $BMO(\mathbb{R}^d)$  functions and fractional integral operators. Now let us replace the singular integral operator  $T$  by the fractional integral operator  $I_\alpha$ . Recall that the definition of  $I_\alpha$  is given by

$$(54.18) \quad I_\alpha f(x) = \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}} dy$$

as long as it makes sense. We will see the analogy with the boundedness of  $I_\alpha$ .

**Definition 54.7.** Let  $0 < \alpha < d$  and  $a \in BMO(\mathbb{R}^d)$ . Then define

$$(54.19) \quad [a, I_\alpha]f(x) := \int_{\mathbb{R}^d} \frac{a(x) - a(y)}{|x-y|^{d-\alpha}} f(y) dy.$$

The following lemma can be shown.

**Lemma 54.8.** Let  $a \in BMO(\mathbb{R}^d)$ ,  $0 < \alpha < d$  and  $0 < r < \infty$ . Then

$$(54.20) \quad M^\sharp[a, I_\alpha]f(x) \lesssim_r M^r[I_\alpha f](x) + M_\alpha f(x),$$

where  $M_\alpha$  is the fractional maximal operator of order  $\alpha$  see (17.19).

**Theorem 54.9.** Let  $a \in BMO(\mathbb{R}^d)$ . Suppose that the parameters  $p, q, \alpha$  satisfy

$$(54.21) \quad 1 < p, q < \infty, 0 < \alpha < d, \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}.$$

Then  $[a, I_\alpha]$  is an  $L^p(\mathbb{R}^d)$ - $L^q(\mathbb{R}^d)$  bounded operator. If we assume in addition that  $a \in VMO$ , then  $[a, I_\alpha]$  is an  $L^p(\mathbb{R}^d)$ - $L^q(\mathbb{R}^d)$  compact operator.

*Proof of Lemma 54.8 and Theorem 54.9.* The proof is analogous to the case of  $[a, T]$ . The proof is omitted.  $\square$

**Theorem 54.10.** Let  $b \in Lip_\gamma$  with  $0 < \gamma < 1$ .

(1) Suppose that  $T$  is a singular integral operator and the parameters  $p, q$  satisfy

$$(54.22) \quad 1 < p, q < \infty, \frac{1}{q} = \frac{1}{p} - \frac{\gamma}{d}.$$

Then  $[b, T]$  is an  $L^p(\mathbb{R}^d)$ - $L^q(\mathbb{R}^d)$  bounded operators.

(2) Suppose that  $0 < \alpha < d - \gamma$  and the parameters  $p, q$  satisfy

$$(54.23) \quad 1 < p, q < \infty, \frac{1}{q} = \frac{1}{p} - \frac{\alpha + \gamma}{d}.$$

Then  $[b, I_\alpha]$  is an  $L^p(\mathbb{R}^d)$ - $L^q(\mathbb{R}^d)$  bounded operators.

*Proof.* We concentrate on (1), the proof of (2) being totally analogous. Note that the kernel is bounded by the constant multiple of  $|x-y|^{d-\gamma}$ . Therefore, we have

$$(54.24) \quad |[b, T]f(x)| \lesssim I_\gamma f(x).$$

Thus, the matters are reduced to the boundedness of  $I_\gamma$ , which is already established.  $\square$

Notes and references for Chapter 23.

For Hardy spaces,  $BMO(\mathbb{R}^d)$  the readers may consult [10, 16, 58] as well.

For more information we refer to [144], [193] and [464].

Section 50. In [182] C. Fefferman and E. M. Stein established the norm equivalences Theorems 50.6, 50.10, 50.11, 50.12 as well as the density results Theorem 50.15. However, the proofs we presented here are different from those in [183]. We depended on the idea by Rychkov [416, 417]. Lemma 50.8, whose clue to the author was the one due to Rychkov [416], was elementary and has several versions. Theorem 50.11, whose proof is rather long, used a famous technique. The one we used in (50.30) is known as the Strömberg-Torchinsky technique. For details we refer to [59]. We refer to [495] for alternative proof of the result in this section.

The atomic decomposition is investigated by Coifman as well (see [141]).

Hardy spaces in connection with complex analysis can be found in [70].

Section 51. Theorem 51.2

Section 52. Theorem 52.4

Theorem 52.5

Theorem 52.7 was obtained by Peetre [390].

Section 53. Theorem 53.3 was originally considered by C. Morrey in connection with partial differential equations (See [360]). Morrey spaces or the Morrey norm do not appeared in this paper. In the early 60's Campanato investigated a function space, which was named the Campanato space later, and investigated the relation between the Morrey norm and the Campanato norm. Theorem 52.2 is a typical example of this attempt. (See [121, 122, 123, 124, 125, 126]). Theorem 52.2 itself can be found in [423]. We refer to [320] for a good generalization of Theorem 52.2 to the spaces of homogeneous type, where we can find a self-contained proof.

We refer to [133, 175, 387] for more details about Theorems 53.2, 53.4, 53.11. Theorem 53.11 is due to Peetre in 1966 and to Chiarenza and Frasca in 1987. Theorem 53.2 is due to Chiarenza and Frasca. In particular, Theorem 53.4 (1), which asserts

$$(54.25) \quad \|I_\alpha f\|_{\mathcal{M}_t^s} \lesssim \|f\|_{\mathcal{M}_q^p}, \quad 1 < q \leq p < \infty, \quad 1 < t \leq s < \infty, \quad \frac{1}{p} - \frac{\alpha}{d} = \frac{1}{s}, \quad \frac{q}{p} = \frac{t}{s},$$

has a little history. Peetre proved a weaker assertion of Theorem 53.4 (1) in [387, Theorem 5.4, p.82]. Adams proved it in the present form in [74]. For more details see [1, p.79 (3.7.2)]. Chiarenza and Frasca considered in [133] in Theorem 53.4 (1). The boundedness of  $M_\alpha$  on Morrey spaces is implicitly due to Chiarenza and Frasca in 1991 [133] in view of  $M_\alpha \lesssim I_\alpha$ . The definition of  $M_\alpha$  seems to date back to Muckenhoupt and Wheeden [363]. Different boundedness of the case (2) was obtained in terms of Orlicz spaces in [425]. As in [387], Morrey spaces are sometimes denoted by  $\mathcal{L}^{p,\lambda}$ . Here the norm is given by

$$(54.26) \quad \|f : \mathcal{L}^{p,\lambda}\| := \sup_{x \in \mathbb{R}^d, r > 0} \left( \frac{1}{r^{d-\lambda}} \int_{B(x,r)} |f|^p \right)^{\frac{1}{p}}, \quad 0 < \lambda < d, \quad 0 < p < \infty.$$

Theorem 53.8 was investigated by Zorko [497]. This notation can be found in [445] and so on. Applications of Morrey spaces to PDE can be found in [289, 331, 333, 420, 421]. The author gave a natural extension of the Morrey norm to general Radon measures [426]. We remark that Morrey spaces do not interpolate well as the work by Luis and Vega shows [99], although there is a partial result on interpolation of the operators in [422]. For example, it is known that  $(\mathcal{M}_q^{p_0}(\mathbb{R}^d), \mathcal{M}_q^{p_1}(\mathbb{R}^d))_{\theta_\infty} \hookrightarrow \mathcal{M}_q^p(\mathbb{R}^d)$ , where  $p, p_0, p_1 \in [1, \infty)$ ,  $q \leq \min(p, p_0, p_1)$ ,  $\theta \in (0, 1)$  and  $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ . Applications of Morrey spaces to Schrödinger equations can be found in [420, 421].

Here is a series of applications to PDEs of Morrey spaces. Let us consider

$$Lu = \sum_{i,j=1}^d a_{ij} u_{x_i} u_{x_j} = f.$$

- (1) The case when  $a_{ij}$  is Hölder continuous. See the book of Gilberg and Trudinger.
- (2) The case when  $a_{ij}$  is uniformly continuous. A. D. N. in 1959.
- (3) Talenti considered the case when  $n = 2$  and each  $a_{ij}$  is measurable in 1966.
- (4) Miranda considered the case when  $d \geq 3$  and each  $a_{ij}$  is measurable in 1963.
- (5) Campanato considered the case when  $a_{ij} \in W^{1,n}(\mathbb{R}^d)$  in 1967.
- (6) Di Fazio and Ragusa dealt with the case when  $a_{ij} \in \text{V.M.O.}$  and  $f \in L_{\text{loc}}^{p,\lambda}(\Omega)$  in 1993. In 1999 they considered the global results. See the paper by Di Fazio and Ragusa in 1991.
- (7) In 1988 Caffelli considered the case when  $f \in L^{p,\lambda}$ .
- (8) Chiarenza, Frasca and Longo considered the case when  $a_{ij} \in \text{V.M.O.}$  and  $u \in W^{2,p}(\mathbb{R}^d)$  with  $1 < p < \infty$ .
- (9) See the paper by Campanato in 1967.
- (10) See the paper by Campanato in 1987.
- (11) Ragusa further considered in 2001 (P.A.M.S.) and in 2002 (Duke).
- (12) Partial regularity is considered by Morrey in 1969.

Theorem 53.9

Section 54. Theorem 54.4 was obtained by R. Coifman, R. Rochberg and G. Weiss in [146], where the converse of this theorem was investigated as well.

Theorem 54.6

As for Theorem 54.9 the converse estimate is known to be true :

$$(54.27) \quad \|[b, I_\alpha]\|_{B(L^p, L^q)} \simeq \|b\|_{\text{BMO}}.$$

The inequality  $c \leq$  is due to Chanillo [127].

Theorem 54.10

The author expresses his gratitude to Y. Han. A discussion with Y. Han helped me a lot [502].



**Part 24. Besov and Triebel-Lizorkin spaces**

Besov spaces and Triebel-Lizorkin spaces are functions spaces that can describe smoothness and integrability very precisely.

Warm-up. Before we go into the details, we recall our main theorem as warm-up.

*Exercise 257.* Construct a smooth function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfying

$$(54.28) \quad \chi_{B(1)} \leq \psi \leq \chi_{B(2)}.$$

Recall that we have obtained Fefferman-Stein's maximal inequality.

(1) (Theorem 12.13) Let  $1 < p \leq \infty$ . Then

$$(54.29) \quad \|Mf\|_p \lesssim_p \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

(2) (Theorem 41.1) Let  $1 < p < \infty$  and  $1 < q \leq \infty$ . Then

$$(54.30) \quad \|Mf_j\|_{L^p(\ell^q)} \lesssim_{pq} \|f_j\|_{L^p(\ell^q)}$$

for all sequences  $\{f_j\}_{j \in \mathbb{N}}$  of the  $L^p(\mathbb{R}^d)$ -functions.

**55. BAND-LIMITED DISTRIBUTIONS**

One of the main purposes of this part is to present the definition of Besov spaces and Triebel-Lizorkin spaces and verify their validness. Let  $\varphi_0, \varphi_1 \in \mathcal{S}(\mathbb{R}^d)$  satisfy

$$(55.1) \quad \chi_{B(2)} \leq \varphi_0 \leq \chi_{B(1)}, \quad \chi_{B(4) \setminus B(2)} \leq \varphi_1 \leq \chi_{B(8) \setminus B(1)}.$$

Furthermore  $\varphi_j(x) = \varphi_1(2^{-j+1}x)$  for  $j \geq 2$ . These function spaces are normed by

$$\|f\|_{B_{pq}^s} = \|2^{js} \varphi_j(D)f\|_{\ell^q(L^p)}, \quad 0 < p, q \leq \infty, s \in \mathbb{R}$$

$$\|f\|_{F_{pq}^s} = \|2^{js} \varphi_j(D)f\|_{L^p(\ell^q)}, \quad 0 < p < \infty, 0 < q \leq \infty, s \in \mathbb{R}$$

respectively. Here, for  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$  we have written  $\varphi(D)f = \mathcal{F}^{-1}(\varphi \mathcal{F}f)$ .

It is not obvious that the norm is independent of the choice of  $\varphi_0$  and  $\varphi_1$ . Furthermore we want to ask ourselves whether the numbers 1, 2, 4, 8 in (55.1) count.

To tackle these problems we have to be familiar with band-limited distributions. Thus, for the time being we deal with band-limited distributions whose Fourier transforms are compactly supported. Guessing from the definition of the Besov norms and the Triebel-Lizorkin norms, we notice that it also counts how large are the supports.

**55.1. Maximal operator control.**

Taking into account the above problem, we are oriented to the systematic and quantitative treatment of the band-limited distributions.

**Definition 55.1.** Let  $A \subset \mathbb{R}^d$  be a bounded set. We define  $(\mathcal{S}'(\mathbb{R}^d))^A$  to be

$$(55.2) \quad (\mathcal{S}'(\mathbb{R}^d))^A := \{f \in \mathcal{S}'(\mathbb{R}^d) : \text{supp}(\mathcal{F}f) \subset \overline{A}\}.$$

We set  $(L^p)^A(\mathbb{R}^d) := L^p(\mathbb{R}^d) \cap (\mathcal{S}'(\mathbb{R}^d))^A$ .

**Remark 55.2.** Let  $f \in (\mathcal{S}'(\mathbb{R}^d))^A$ . Take a compactly supported function  $\psi$  that takes 1 on  $A$ . Then we have  $f = \mathcal{F}^{-1}\mathcal{F}f = \mathcal{F}^{-1}(\psi \cdot \mathcal{F}f) = (2\pi)^d \mathcal{F}^{-1}\psi * f$ , which implies  $f \in C^\infty(\mathbb{R}^d)$ . In particular it is meaningful to evaluate  $f$  at  $x \in \mathbb{R}^d$ .

In view of Remark 55.2 the statement of the following theorem makes sense. This theorem is a milestone of the theory of Besov and Triebel-Lizorkin spaces.

**Theorem 55.3.** *Let  $f \in (\mathcal{S}'(\mathbb{R}^d))^{B(x_0,1)}$ . Then we have*

$$(55.3) \quad \sup_{z \in \mathbb{R}^d} \frac{|\nabla f(x-y)|}{1+|y|^{\frac{d}{r}}} \lesssim \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}}$$

$$(55.4) \quad \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}} \lesssim M^{(r)}f(x),$$

where  $c$  depends on  $r$  and  $d$ .

*Proof of (55.3).* By considering  $e^{ix_0 \cdot *} f$ , we may assume  $x_0 = 0$ . To prove this we take  $\psi \in \mathcal{S}(\mathbb{R}^d)$  so that

$$(55.5) \quad \chi_{B(1)} \leq \psi \leq \chi_{B(2)}.$$

By the similar reasoning as Remark 55.2 we have  $f = (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1}\psi * f$ . Write it out in full:

$$(55.6) \quad f(x) = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} \mathcal{F}^{-1}\psi(y) f(x-y) dy.$$

To prove  $\sup_{z \in \mathbb{R}^d} \frac{|\nabla f(x-y)|}{1+|y|^{\frac{d}{r}}} \lesssim \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}}$  we may replace  $\nabla$  by  $\partial_j$  for fixed  $j$ . That is, we have only to prove it componentwise. Differentiation of (55.6) then yields

$$(55.7) \quad \partial_j f(x) = (2\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} [\partial_j \mathcal{F}^{-1}\psi](y) f(x-y) dy.$$

Let us write  $\partial_j \mathcal{F}^{-1}\psi = \rho$  for simplicity. By the triangle inequality of integral we obtain

$$(55.8) \quad \frac{|\partial_j f(x-y)|}{1+|y|^{\frac{d}{r}}} \leq (2\pi)^d \int_{\mathbb{R}^d} \frac{|\rho(z) f(x-y-z)|}{1+|y|^{\frac{d}{r}}} dz.$$

By the Peetre lemma we have

$$(55.9) \quad (1+|y+z|^{\frac{d}{r}}) \lesssim (1+|z|^{\frac{d}{r}})(1+|y|^{\frac{d}{r}}).$$

Keeping  $\rho \in \mathcal{S}(\mathbb{R}^d)$  in mind, we are led to

$$\begin{aligned} \frac{|\partial_j f(x-y)|}{1+|y|^{\frac{d}{r}}} &\lesssim \int_{\mathbb{R}^d} \frac{(1+|z|^{\frac{d}{r}})|\rho(z) f(x-y-z)|}{1+|y+z|^{\frac{d}{r}}} dz \\ &\lesssim \int_{\mathbb{R}^d} \frac{\{(1+|z|^{\frac{d}{r}})|\rho(z)|\} |f(x-y-z)|}{1+|y+z|^{\frac{d}{r}}} dz \\ &\lesssim \sup_{y \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}}. \end{aligned}$$

This is the desired inequality. □

*Proof of (55.4) Reduction step.* First, we may assume that  $f \in (\mathcal{S}'(\mathbb{R}^d))^{B(1/2)}$  by the dilation argument. Let  $\psi$  be a smooth function such that

$$(55.10) \quad \int \mathcal{F}\psi(\xi) d\xi = (2\pi)^{\frac{d}{2}}, \quad \text{supp}(\mathcal{F}\psi) \subset \chi_{B(1)}.$$

Set  $g_t(x) := \psi(tx)f(x)$ ,  $x \in \mathbb{R}^d$ ,  $0 < t < \frac{1}{2}$ . Then we have

- (1)  $M^{(r)}g_t(x) \leq M^{(r)}f(x)$  for all  $t > 0$  and  $x \in \mathbb{R}^d$ .
- (2)  $\lim_{t \rightarrow +0} g_t(x) = f(x)$  for all  $x \in \mathbb{R}^d$
- (3)  $\text{supp}(\mathcal{F}g_t) \subset t \cdot \text{supp}(\psi) + \text{supp}(f) \subset B\left(\frac{1}{2}\right) + B\left(\frac{1}{4}\right) = B\left(\frac{3}{4}\right)$ .
- (4)  $g_t \in \mathcal{S}(\mathbb{R}^d)$  for each  $0 < t < \frac{1}{4}$ .

Thus we may assume  $f \in \mathcal{S}(\mathbb{R}^d)$ . □

*Proof of (55.4).* To prove this inequality, we first take  $v \in \mathbb{R}^d$  and  $0 < r < 1$ . The precise value of  $r$  will be fixed sufficiently small.

Let  $y_v \in \overline{B}(x, r)$  that attains the minimum of  $|f(\cdot)|$  in  $\overline{B}(v, r)$ . Then by the mean value theorem we have

$$(55.11) \quad |f(v)| \leq |f(y_v)| + |f(v) - f(y_v)| \leq \inf_{z \in B(v, r)} |f(z)| + r \sup_{w \in B(v, r)} |\nabla f(w)|.$$

By replacing  $v$  with  $x - y$  we obtain

$$(55.12) \quad |f(x - y)| \leq \left( \inf_{z \in B(x-y, r)} |f(z)| \right) + \left( r \sup_{w \in B(x-y, r)} |\nabla f(w)| \right), \quad x, y \in \mathbb{R}^d$$

Since  $|B(1)| \geq 1$ , we obtain

$$(55.13) \quad \inf_{z \in B(x-y, r)} |f(z)| \leq \left( \int_{B(x-y, 1)} |f(z)|^r dz \right)^{\frac{1}{r}}.$$

Observe that this is where the integral and hence the maximal operator appears. The inclusion  $B(x - y, 1) \subset B(x, |y| + 1)$  together with (55.13) gives us

$$(55.14) \quad |f(x - y)| \leq \left( \int_{B(x, |y|+1)} |f(z)|^r dz \right)^{\frac{1}{r}} + r \left( \sup_{w \in B(x-y, r)} |\nabla f(w)| \right).$$

Taking supremum over  $y \in \mathbb{R}^d$  we obtain

$$(55.15) \quad \sup_{y \in \mathbb{R}^d} \frac{|f(x - y)|}{1 + |y|^{\frac{d}{r}}} \leq \frac{1}{1 + |y|^{\frac{d}{r}}} \left( \int_{B(x, |y|+1)} |f(z)|^r dz \right)^{\frac{1}{r}} + r \left( \sup_{y, w \in \mathbb{R}^d, |x-y-w| < r} \frac{|\nabla f(w)|}{1 + |y|^{\frac{d}{r}}} \right).$$

Note that, changing variables  $w \mapsto z := x - w$ , we obtain

$$(55.16) \quad \sup_{y, w \in \mathbb{R}^d, |x-y-w| < r} \frac{|\nabla f(w)|}{1 + |y|^{\frac{d}{r}}} = \sup_{y, z \in \mathbb{R}^d, |z-y| < r} \frac{|\nabla f(x - z)|}{1 + |y|^{\frac{d}{r}}}$$

and if  $z \in B(y, r)$  with  $r \leq 1$ , we obtain  $1 + |y|^{\frac{d}{r}} \sim 1 + |z|^{\frac{d}{r}}$ . Meanwhile it is easy to see

$$(55.17) \quad \frac{1}{1 + |y|^{\frac{d}{r}}} \left( \int_{B(x, |y|+1)} |f(z)|^r dz \right)^{\frac{1}{r}} \lesssim M^{(r)}f(x).$$

Consequently we obtain

$$(55.18) \quad \sup_{z \in \mathbb{R}^d} \frac{|f(x - y)|}{1 + |y|^{\frac{d}{r}}} \lesssim \left( M^{(r)}f(x) + r \sup_{z \in B(y, r)} \frac{|\nabla f(x - z)|}{1 + |z|^{\frac{d}{r}}} \right).$$

Since we have shown (55.3), that is,

$$(55.19) \quad \sup_{z \in \mathbb{R}^d} \frac{|\nabla f(x-z)|}{1+|z|^{\frac{d}{r}}} \lesssim \sup_{z \in \mathbb{R}^d} \frac{|f(x-z)|}{1+|z|^{\frac{d}{r}}},$$

it follows that there exists a constant  $c_0 > 0$  such that

$$(55.20) \quad \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}} \lesssim M^{(r)} f(x) + c_0 r \left( \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}} \right).$$

If we take  $r = \min(1, (2c_0)^{-1})$ , we can bring the most right side to the left side. Since  $f \in \mathcal{S}(\mathbb{R}^d)$ , every term in (55.20) is finite. Thus we are allowed to subtract  $c_0 r \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}}$  in (55.20).

Consequently we finally obtain

$$(55.21) \quad \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|y|^{\frac{d}{r}}} \lesssim M^{(r)} f(x).$$

This is the desired result.  $\square$

**Corollary 55.4.** *Let  $0 < p \leq q \leq \infty$ . Then  $(L^p)^{B(1)}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ .*

*Proof.* By the inequality  $\|f\|_q \leq \|f\|_p^{\frac{p}{q}} \cdot \|f\|_\infty^{1-\frac{p}{q}}$  we may assume  $q = \infty$ . Let  $x_0$  be an arbitrary point. Then by virtue of Theorem 55.3

$$\sup_{y \in B(x_0, 1)} |f(y)| \lesssim \sup_{y \in \mathbb{R}^d} \frac{|f(y)|}{1+|x_0-y|^{\frac{d}{r}}} \lesssim \inf_{w \in B(x_0, 1)} \left( \sup_{y \in \mathbb{R}^d} \frac{|f(y)|}{1+|w-y|^{\frac{d}{r}}} \right) \lesssim \inf_{x \in B(x_0, 1)} M^{(r)} f(x).$$

From this we deduce

$$(55.22) \quad \sup_{y \in B(x_0, 1)} |f(y)| \lesssim \left( \int_{B(x_0, 1)} |f(y)|^p dy \right)^{\frac{1}{p}} \lesssim \|f\|_p$$

for all  $x_0 \in \mathbb{R}^d$ . Since

$$(55.23) \quad \|f\|_\infty = \sup_{x_0 \in \mathbb{R}^d} \left( \sup_{y \in B(x_0, 1)} |f(y)| \right),$$

we have  $\|f\|_\infty \lesssim \|f\|_p$ .  $\square$

Scaling. Let  $f \in (\mathcal{S}'(\mathbb{R}^d))^{B(R)}$ . Then  $f(R^{-1}\cdot) \in (\mathcal{S}'(\mathbb{R}^d))^{B(1)}$  and

$$(55.24) \quad \|f(R^{-1}\cdot)\|_p = R^{\frac{d}{p}} \|f\|_p.$$

Via this transform, we obtain

**Theorem 55.5.** *Let  $f \in (\mathcal{S}'(\mathbb{R}^d))^{B(R)}$ . Then for all  $r > 0$  there exists  $c_r > 0$  independent of  $R$  and  $f$  such that*

$$(55.25) \quad \sup_{z \in \mathbb{R}^d} \frac{|f(x-y)|}{1+|Ry|^{\frac{d}{r}}} \leq c_r M^{(r)} f(x).$$

Furthermore let  $0 < p \leq q \leq \infty$ . If  $f \in (L^p)^{B(R)}(\mathbb{R}^d)$ , then

$$(55.26) \quad \|f\|_q \lesssim R^{\frac{d}{p} - \frac{d}{q}} \|f\|_p.$$

55.2. Multiplier theorems.

In this section we prove an important multiplier theorem deduced from Theorem 55.5. Recall that we have proved the Plancherel theorem :

$$(55.27) \quad \int_{\mathbb{R}^d} |f(y)|^2 dy = \int_{\mathbb{R}^d} |\mathcal{F}f(y)|^2 dy = \int_{\mathbb{R}^d} |\mathcal{F}^{-1}f(y)|^2 dy, \quad f \in L^2(\mathbb{R}^d).$$

For  $f \in \mathcal{S}'(\mathbb{R}^d)$  and  $s \in \mathbb{R}$  we define

$$(55.28) \quad (1 - \Delta)^{\frac{s}{2}} f := \mathcal{F}^{-1}(\langle \cdot \rangle^s \cdot \mathcal{F}f).$$

**Lemma 55.6.**  $(1 - \Delta)^{\frac{s}{2}} : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$  is a continuous isomorphism with inverse  $(1 - \Delta)^{-\frac{s}{2}} : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$ .

**Definition 55.7.** Let  $\sigma > 0$ . Then define

$$(55.29) \quad H_2^\sigma(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d) : (1 - \Delta)^{\frac{\sigma}{2}} f \in L^2(\mathbb{R}^d)\}.$$

We equip  $H_2^\sigma$  with the norm

$$(55.30) \quad \|f\|_{H_2^\sigma} := \|(1 - \Delta)^{\frac{\sigma}{2}} f\|_2 = \left( \int_{\mathbb{R}^d} (1 + |\xi|^2)^\sigma |\mathcal{F}f(\xi)|^2 d\xi \right)^{\frac{1}{2}}.$$

To be familiarize with the definitions let us see some properties of this function space.

*Exercise 258.* Let  $\delta, \sigma, \sigma_1, \sigma_2 \in \mathbb{R}$ .

- (1)  $(1 - \Delta)^\delta : H_2^\sigma \rightarrow H_2^{\sigma - \delta}$  is isometry.
- (2) If  $\sigma > 0$ , then  $H_2^\sigma \hookrightarrow H_2^0 = L^2(\mathbb{R}^d)$ .
- (3) In general, if  $\sigma_1 > \sigma_2$ , then  $H_2^{\sigma_1} \hookrightarrow H_2^{\sigma_2}$  in the sense of continuous embedding.

Next we introduce important constants restricting the regularity of functions.

**Definition 55.8.** Define  $\sigma_{pq}$  and  $\sigma_p$  to be

$$(55.31) \quad \sigma_{pq} := d \left( \frac{1}{\min(1, p, q)} - 1 \right), \quad \sigma_p := \sigma_{pp}, \quad 0 < p, q \leq \infty.$$

These constants, as it will turn out, have close connection with the smoothness parameter  $s$ .

Our standard multiplier theorem in this paper is the following.

**Theorem 55.9.** Let  $H \in \mathcal{S}(\mathbb{R}^d)$  and  $\{H_k\}_{k \in \mathbb{N}_0} \subset \mathcal{S}(\mathbb{R}^d)$ .

- (1) Let  $r > 0$ ,  $0 < p \leq \infty$  and  $\sigma > \sigma_p + \frac{d}{2}$ . Then, for all  $f \in (L^p)^{B(r)}(\mathbb{R}^d)$ ,
- $$(55.32) \quad \|H(D)f\|_p \lesssim \|H(r \cdot)\|_{H_2^\sigma} \cdot \|f\|_p.$$
- (2) Let  $\{r_k\}_{k=1}^\infty \subset (0, \infty)$ ,  $0 < p < \infty$ ,  $0 < q \leq \infty$  and  $\sigma > \sigma_{pq} + \frac{3d}{2}$ . Then for all sequence  $\{f_k\}_{k \in \mathbb{N}_0} \subset L^p(\mathbb{R}^d)$  with  $f_k \in (L^p)^{B(r_k)}(\mathbb{R}^d)$  for each  $k$ ,

$$(55.33) \quad \|H_k(D)f_k\|_{L^p(\ell^q)} \lesssim \left( \sup_{j \in \mathbb{N}_0} \|H_j(r_j \cdot)\|_{H_2^\sigma} \right) \|f_k\|_{L^p(\ell^q)}.$$

The meaning of the multiplier operators  $H(D)f$  in Theorem 55.9 is made explicit in the following lemma. Once we accept the following lemma (Lemma 55.10), with the aid of the Fefferman-Stein's vector-valued inequality Theorem 55.9 is completely proved.

**Lemma 55.10.** *Let  $\sigma > n \left( \frac{1}{\eta} + \frac{1}{2} \right)$ ,  $f \in (L^p)^{B(R)}(\mathbb{R}^d)$  and  $H \in \mathcal{S}(\mathbb{R}^d)$ . Then*

$$(55.34) \quad |H(D)f(x)| \lesssim \|H(R\cdot)\|_{H_2^\sigma} \cdot M^{(\eta)}f(x),$$

where the implicit constant is independent of  $r$ ,  $f$  and  $H$ .

Let us take Lemma 55.10 for granted for the time being. With Lemma 55.10 we prove Theorem 55.9 (2), for example. By assumption we can take  $t > 0$  slightly smaller than  $\min(1, p, q)$  so that

$$(55.35) \quad \sigma > n \left( \frac{1}{t} - \frac{1}{2} \right).$$

Then we have by virtue of Lemma 55.10

$$(55.36) \quad |H_k(D)f_k(x)| \lesssim \left( \sup_{j \in \mathbb{N}_0} \|H_j(r_j \cdot)\|_{H_2^\sigma} \right) \cdot M^{(t)}f_k(x).$$

Consequently we obtain

$$\begin{aligned} \|H_k(D)f_k\|_{L^p(\ell^q)} &\lesssim \left( \sup_{j \in \mathbb{N}_0} \|H_j(r_j \cdot)\|_{H_2^\sigma} \right) \|M^{(t)}f_k\|_{L^p(\ell^q)} \\ &= \left( \sup_{j \in \mathbb{N}_0} \|H_j(r_j \cdot)\|_{H_2^\sigma} \right) \left\| M[|f_k|^t] : L^{\frac{p}{t}}(\ell^{\frac{q}{t}}) \right\|^{\frac{1}{t}}. \end{aligned}$$

Since  $p > t$  and  $q > t$ , we are in the position of using the Fefferman-Stein's vector-valued inequality. Thanks to this inequality, we can remove  $M$  in the above formula. Consequently we obtain

$$(55.37) \quad \|H_k(D)f_k\|_{L^p(\ell^q)} \lesssim \left( \sup_{j \in \mathbb{N}_0} \|H_j(r_j \cdot)\|_{H_2^\sigma} \right) \|f_k\|_{L^p(\ell^q)}.$$

Thus, Theorem 55.9 (2) is therefore proved. The proof of Theorem 55.9 (1) is similar. Before we come to the proof of the Theorem, a clarifying remark may be in order.

**Remark 55.11.** The definition of  $H(D)f$  is the same as the usual one, if we assume  $H \in \mathcal{S}(\mathbb{R}^d)$ : we can interpret  $H(D)f$  as  $\mathcal{F}^{-1}(H \cdot \mathcal{F}f)$ .

We now turn to the proof of Lemma 55.10.

*Proof of Lemma 55.10.* First by Hölder's inequality the integral in question is estimated by

$$\int_{\mathbb{R}^d} |\mathcal{F}^{-1}H(y)f(x-y)| dy \leq \left( \int_{\mathbb{R}^d} \langle ry \rangle^{2\sigma} |\mathcal{F}^{-1}H(y)|^2 dy \right)^{\frac{1}{2}} \cdot \left( \int_{\mathbb{R}^d} \langle ry \rangle^{-2\sigma} |f(x-y)|^2 dy \right)^{\frac{1}{2}}.$$

By simple calculation we have

$$\left( \int_{\mathbb{R}^d} \langle ry \rangle^{2\sigma} |\mathcal{F}^{-1}H(y)|^2 dy \right)^{\frac{1}{2}} = r^{-\frac{d}{2}} \left( \int_{\mathbb{R}^d} \langle y \rangle^{2\sigma} |\mathcal{F}^{-1}H(r^{-1}y)|^2 dy \right)^{\frac{1}{2}} = r^{\frac{d}{2}} \|H(r\cdot)\|_{H_2^\sigma}.$$

Thus we are left with the task of proving

$$(55.38) \quad r^{\frac{d}{2}} \left( \int_{\mathbb{R}^d} \langle ry \rangle^{-2\sigma} |f(x-y)|^2 dy \right)^{\frac{1}{2}} \lesssim M^{(r)}f(x).$$

This can be achieved as follows: By Theorem 55.9 we have

$$(55.39) \quad \frac{|f(x-y)|}{\langle ry \rangle^{\frac{d}{R}}} \lesssim M^{(r)}f(x).$$

If we insert this estimate, we obtain

$$(55.40) \quad r^{\frac{d}{2}} \left( \int_{\mathbb{R}^d} \langle ry \rangle^{-2\sigma} |f(x-y)|^2 dy \right)^{\frac{1}{2}} \lesssim r^{\frac{d}{2}} \left( \int_{\mathbb{R}^d} \langle ry \rangle^{-2\sigma + \frac{2d}{R}} dy \right)^{\frac{1}{2}} M^{(r)} f(x).$$

Since we are assuming  $2\sigma - \frac{2d}{R} > d$ , we have  $r^{\frac{d}{2}} \left( \int_{\mathbb{R}^d} \langle ry \rangle^{-2\sigma + \frac{2d}{R}} dy \right)^{\frac{1}{2}} = c < \infty$ . As a result we obtain the desired result.  $\square$

Before we define the function spaces, let us make a brief comment on Theorem 55.9. It is not very useful to learn by heart the condition on  $\sigma$ . It suffices to keep in mind that  $\sigma$  is taken large enough. In our many usage of Theorem 55.9 we always take  $\sigma$  large enough.

**55.3. Application to singular integral operators.**

Denote by  $\mathcal{S}(\mathbb{R}^d)_0$  the set of the closure of the set of all elements  $\mathcal{S}(\mathbb{R}^d)$  whose Fourier transforms do not contain 0 as their support. Note that by virtue of the homogeneous version of the Littlewood-Paley theory, for example,  $\mathcal{S}(\mathbb{R}^d)_0$  is a dense subspace of  $L^p(\mathbb{R}^d)$  for all  $1 < p < \infty$ .

Now we present a powerful application of the Littlewood-Paley theorem.

**Theorem 55.12.** *Let  $1 < p < \infty$  and  $m \in C^\infty(\mathbb{R}^d \setminus \{0\})$ . Suppose that it satisfies the differential inequality*

$$(55.41) \quad |\partial^\alpha m(\xi)| \leq c_\alpha |\xi|^{-|\alpha|}$$

for all  $\alpha \in \mathbb{N}_0$ . Then the linear operator, initially defined on  $\mathcal{S}(\mathbb{R}^d)_0$ ,

$$(55.42) \quad T_m(f) = \lim_{j \rightarrow \infty} \mathcal{F}(m \cdot \varphi_j(D))f$$

extends to a bounded linear operator on  $L^p(\mathbb{R}^d)$ .

Before we come to the proof, two helpful remarks may be in order.

- Remark 55.13.** (1) We remark that the Riesz transform is an operator to which this theorem can be readily applied.  
 (2) If we re-examine the proofs of the theorem in this section, the differential inequality (55.41) for  $|\alpha| \leq \frac{d+1}{2}$  suffices.

This theorem, despite its appearance, is proved very simply once we use our culmination of this book.

*Proof.* Let  $f \in \mathcal{S}(\mathbb{R}^d)_0$ . Note that

$$(55.43) \quad \|T_m f\|_p \simeq \|T_m f\|_{\dot{F}_{p^2}^0}$$

by virtue of the Littlewood-Paley theorem. Therefore, by the multiplier theorem

$$(55.44) \quad \|T_m f\|_{\dot{F}_{p^2}^0} \lesssim \sup_{k \in \mathbb{Z}} \|m(2^k \cdot) \varphi_k\|_{H^M} \cdot \|f\|_{\dot{F}_{p^2}^0} \lesssim \|f\|_{\dot{F}_{p^2}^0}.$$

It remains to use the Littlewood-Paley theorem once more.  $\square$

## 56. BESOV SPACES AND TRIEBEL-LIZORKIN SPACES

In this section we define the function spaces and investigate elementary properties. The spaces we are going to obtain are called nonhomogeneous. Their homogeneous version is considered in the next section.

## 56.1. Definition.

**Definition 56.1.** We pick  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $\psi \in \mathcal{S}(\mathbb{R}^d)$  so that they satisfy

$$(56.1) \quad c_0 \chi_{B(2)} \leq \varphi \leq c_1 \chi_{B(4)}, \quad c_0 \chi_{B(4) \setminus B(2)} \leq \psi \leq c_1 \chi_{B(8) \setminus B(4)}.$$

for some  $c_0, c_1 > 0$ . Set  $\varphi_0 = \varphi$  and  $\varphi_j := \psi(2^{-j+1}\cdot)$  for  $j \in \mathbb{N}$ . Denote  $\Phi := \{\varphi_j\}_{j \in \mathbb{N}_0}$ .

**Example 56.2.** It will turn out that we can freely choose the family  $\Phi$  as in Definition 56.1. For example we can choose  $\Phi$  so that one of the following conditions holds.

- (1) (Radial function) We can take  $(\varphi_j)_{j \in \mathbb{N}_0}$  so that they are radial.
- (2) (Characterization of the unit type 1)  $(\varphi_j)_{j \in \mathbb{N}_0}$  satisfies

$$(56.2) \quad \sum_{j \in \mathbb{N}_0} \varphi_j \equiv 1.$$

- (3) (Characterization of the unit type 2)  $(\varphi_j)_{j \in \mathbb{N}_0}$  satisfies

$$(56.3) \quad \sum_{j \in \mathbb{N}_0} \varphi_j^2 \equiv 1.$$

- (4) (Simple ingredient) If we take  $\psi$  so that  $\varphi_1 = \varphi(2^{-1}\cdot) - \varphi(2\cdot)$ , then  $(\varphi_j)_{j \in \mathbb{N}_0}$  satisfies

$$(56.4) \quad \chi_{B(2)} \leq \varphi \leq \chi_{B(4)}, \quad \chi_{B(4) \setminus B(2)} \leq \psi \leq \chi_{B(8) \setminus B(4)}.$$

*Exercise 259.* Let  $m \in \mathbb{N}$ . Construct a set of functions  $\{\varphi_j\}_{j=0}^\infty \subset \mathcal{S}(\mathbb{R}^d)$  so that it satisfies the following conditions

- (1)  $\sum_{j=1}^\infty \varphi_j^m \equiv 1$ .
- (2)  $0 \leq \varphi_j \leq 1$ .
- (3)  $\text{supp}(\varphi_j) \subset B(2^{j+3}) \setminus B(2^j)$  for  $j \geq 1$  and  $\text{supp}(\varphi_0) \subset B(4)$ .
- (4)  $\varphi_j(x) = \varphi_1(2^{-j+1}x)$  for all  $j \geq 1$ .

**Definition 56.3.** Let  $\Phi$  be the system of functions as in Definition 56.1.

- (1) Let the parameter  $p, q, s$  satisfy

$$(56.5) \quad 0 < p, q \leq \infty, \quad s \in \mathbb{R}.$$

Then define

$$(56.6) \quad \|f\|_{B_{pq}^s}^\Phi := \|2^{js} \varphi_j(D)f\|_{\ell^q(L^p)}.$$

Define  $B_{pq}^s(\mathbb{R}^d)$  to be the set of the Schwartz distributions  $f \in \mathcal{S}'(\mathbb{R}^d)$  for which  $\|f\|_{B_{pq}^s}^\Phi < \infty$ .

- (2) Let the parameter  $p, q, s$  satisfy

$$(56.7) \quad 0 < p < \infty, \quad 0 < q \leq \infty, \quad s \in \mathbb{R}.$$

Then define

$$(56.8) \quad \|f\|_{F_{pq}^s}^\Phi := \|2^{js} \varphi_j(D)f\|_{L^p(\ell^q)}.$$



Define  $F_{pq}^s(\mathbb{R}^d)$  to be the set of the Schwartz distributions  $f \in \mathcal{S}'(\mathbb{R}^d)$  for which  $\|f\|_{F_{pq}^s} < \infty$ .

- (3) Write  $A_{pq}^s(\mathbb{R}^d)$  to denote  $B_{pq}^s(\mathbb{R}^d)$  or  $F_{pq}^s(\mathbb{R}^d)$ . Unless stated otherwise, the letter  $A$  means the same letter throughout the statement.

Before proceeding, some important remarks on notations may be in order.

*Notation .* (1) In what follows we assume that the parameters

$$(56.9) \quad p, p_0, p_1, p_2, q, q_0, q_1, q_2, s, s_0, s_1, s_2$$

satisfy

$$(56.10) \quad 0 < p, p_0, p_1, p_2, q, q_0, q_1, q_2 \leq \infty, \quad s, s_0, s_1, s_2 \in \mathbb{R}$$

to denote  $A_{pq}^s$  and so on, unless otherwise stated. If  $A = F$ , then we further assume

$$(56.11) \quad 0 < p, p_0, p_1, p_2 < \infty.$$

- (2) We shall show that the definitions are independent of  $\Phi$ . For later consideration once we show the independence of  $\Phi$ , we always omit  $\Phi$  in the definition of the norms.

Now we tackle our pending question.

**Theorem 56.4.** *Suppose that  $\Phi^{(1)} = (\varphi_j^{(1)})_{j \in \mathbb{N}_0}$  and  $\Phi^{(2)} = (\varphi_j^{(2)})_{j \in \mathbb{N}_0}$  are the families satisfying the condition in Definition 56.1. Then*

$$(56.12) \quad \|f\|_{A_{pq}^s}^{\Phi^{(2)}} \simeq_{\Phi^{(1)}, \Phi^{(2)}, p, q, s} \|f\|_{A_{pq}^s}^{\Phi^{(1)}}.$$

*Proof.* We have only to show that

$$(56.13) \quad \|f\|_{A_{pq}^s}^{\Phi^{(2)}} \lesssim_{\Phi^{(1)}, \Phi^{(2)}, p, q, s} \|f\|_{A_{pq}^s}^{\Phi^{(1)}}$$

by symmetry.

Let  $\varphi_{-1}^{(1)} = \varphi_{-1}^{(2)} \equiv 0$ . We define  $\{\eta_j\}_{j \in \mathbb{N}_0}$

$$(56.14) \quad \eta_j := \frac{\varphi_j^{(2)}}{\varphi_{j+1}^{(1)} + \varphi_j^{(1)} + \varphi_{j-1}^{(1)}}, \quad j \in \mathbb{N}_0.$$

Note that for every  $j$  the interior of the support of the denominator contains that of the numerator. As a result they are well-defined compactly supported functions. Then

$$(56.15) \quad \eta_j(x) = \eta_2(2^{-j+2}x) \quad j \geq 2, \quad \text{and} \quad \eta_j(x)(\varphi_{j-1}^{(1)}(x) + \varphi_j^{(1)}(x) + \varphi_{j+1}^{(1)}(x)) = \varphi_j^{(2)}(x).$$

Now it is time to invoke Theorem 55.9. Then we obtain

$$\begin{aligned} \|2^{js} \varphi_j^{(2)} f\|_{\ell^q(L^p)} &\lesssim \left( \sup_{k \in \mathbb{N}} \|\eta_k(2^k \cdot)\|_{H_2^\sigma} \right) \cdot \|2^{js} \varphi_j^{(1)} f\|_{\ell^q(L^p)} \\ \|2^{js} \varphi_j^{(2)} f\|_{L^p(\ell^q)} &\lesssim \left( \sup_{k \in \mathbb{N}} \|\eta_k(2^k \cdot)\|_{H_2^\sigma} \right) \cdot \|2^{js} \varphi_j^{(1)} f\|_{L^p(\ell^q)}. \end{aligned}$$

In view of  $\eta_j(x) = \eta_2(2^{-j+2}x)$   $j \geq 2$ , we see that

$$(56.16) \quad \sup_{k \in \mathbb{N}} \|\eta_k(2^k \cdot)\|_{H_2^\sigma} = \sup_{k=0,1,2} \|\eta_k(2^k \cdot)\|_{H_2^\sigma} < \infty.$$

Consequently  $\|f\|_{A_{pq}^s}^{\Phi^{(2)}} \lesssim \|f\|_{A_{pq}^s}^{\Phi^{(1)}}$  is justified. □

In view of this, let us write  $\|f\|_{A_{pq}^s} := \|f\|_{A_{pq}^s}^\Phi$  for some fixed  $\Phi$ . Our new notation reflects the fact that the specifics of  $\Phi$  are no longer central to our arguments.

### 56.2. Elementary properties.

To discuss the property of the norm  $\|\cdot\|_{A_{pq}^s}$ , we use the following family of function.

**Lemma 56.5.** *There exists a family of compactly supported functions  $\{\varphi_j\}_{j \in \mathbb{N}_0}$  with the following properties.*

- (1)  $\text{supp}(\varphi_0) \subset B(4)$  and  $\text{supp}(\varphi_1) \subset B(8) \setminus B(1)$ .
- (2)  $\varphi_j = \varphi(2^{-j+1}\cdot)$ .
- (3)  $\sum_{j=1}^{\infty} \varphi_j \equiv 1$ .
- (4)  $\varphi_j \geq 0$  for all  $j \in \mathbb{N}_0$ .

Quasi-norm properties. Here we prove the quasi-norm property by clarifying its meaning here.

**Theorem 56.6.** *Let  $f, g \in A_{pq}^s(\mathbb{R}^d)$  and  $k \in \mathbb{C}$ . Then*

- (1)  $\|f\|_{A_{pq}^s} \geq 0$ . The equality holds if and only if  $f = 0$ .
- (2)  $\|kf\|_{A_{pq}^s} = |k| \cdot \|f\|_{A_{pq}^s}$ .
- (3) Let  $f, g \in A_{pq}^s$ .

$$(56.17) \quad \|f + g\|_{A_{pq}^s} \leq \max\left(1, 2^{\frac{1}{p}-1}\right) \cdot \max\left(1, 2^{\frac{1}{q}-1}\right) \cdot (\|f\|_{A_{pq}^s} + \|g\|_{A_{pq}^s}).$$

In particular if  $p, q \geq 1$ , then

$$(56.18) \quad \|f + g\|_{A_{pq}^s} \leq \|f\|_{A_{pq}^s} + \|g\|_{A_{pq}^s}.$$

*Proof.* Let  $\{\varphi_j\}_{j \in \mathbb{N}_0}$  be a family constructed in Lemma 5.17. Since  $\sum_{j=0}^J \varphi_j(D)f \rightarrow f$  in  $\mathcal{S}'(\mathbb{R}^d)$ ,

(1) is easy. (2) follows immediately from the definition. (3) is an immediate result from the quasi-triangle inequalities.

$$\begin{aligned} \|f + g\|_p &\leq \max\left(1, 2^{\frac{1}{p}-1}\right) (\|f\|_p + \|g\|_p) \\ \|a + b\|_{\ell^q} &\leq \max\left(1, 2^{\frac{1}{q}-1}\right) (\|a\|_{\ell^q} + \|b\|_{\ell^q}). \end{aligned}$$

If  $p, q \geq 1$ , it is easy to see that the constant  $c$  appearing in the above formula can be taken 1. As a result the theorem is proved.  $\square$

An elementary inclusion. Let us see how the triply parameterized function spaces  $A_{pq}^s(\mathbb{R}^d)$  are related to one another.

*Exercise 260.* Give alternative proof of Lemma 27.10 by using the duality.

**Theorem 56.7.** *Let  $0 < p \leq \infty$ ,  $q, q_1, q_2 \leq \infty$  and  $s \in \mathbb{R}$ .*

- (1)  $A_{pq_1}^s(\mathbb{R}^d) \subset A_{pq_2}^s(\mathbb{R}^d)$ , if  $q_1 \leq q_2$ .
- (2)  $B_{pp}^s(\mathbb{R}^d) = F_{pp}^s(\mathbb{R}^d)$ .
- (3)  $B_{p \min(p,q)}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d) \subset B_{p \max(p,q)}^s(\mathbb{R}^d)$ .

The first two formulae are clear, thus we prove the last one.

*Proof.* The case when  $p \geq q$ . In this case the matters are reduced to

$$(56.19) \quad B_{pq}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d) \subset B_{pp}^s(\mathbb{R}^d).$$

However, with the aid of the first two assertions, we have  $F_{pq}^s(\mathbb{R}^d) \subset F_{pp}^s(\mathbb{R}^d) = B_{pp}^s(\mathbb{R}^d)$ . As a consequence we have only to show that  $B_{pq}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d)$ . By the definition of the norm we have

$$(56.20) \quad \|f_k\|_{\ell^q(L^p)} = \| |f_k|^q \|_{\ell^1(L^{\frac{p}{q}})}^{\frac{1}{q}}, \quad \|f_k\|_{L^p(\ell^q)} = \| |f_k|^q \|_{L^{\frac{p}{q}}(\ell^1)}^{\frac{1}{q}}$$

for all  $\{f_k\}_{k \in \mathbb{N}_0} \subset L^{1,loc}$ . If we invoke Lemma 27.10, we obtain

$$(56.21) \quad \|f_k\|_{\ell^q(L^p)} \geq \|f_k\|_{L^p(\ell^q)}.$$

Seemingly there is no constant in inequality (56.21). However, we should keep in mind that the norms always depend implicitly on the family of functions we choose. Thus, it follows that

$$(56.22) \quad \|f\|_{F_{pq}^s} \leq \|f\|_{B_{pq}^s}.$$

for all  $f \in B_{pq}^s(\mathbb{R}^d)$ , implying  $F_{pq}^s(\mathbb{R}^d) \subset B_{pq}^s(\mathbb{R}^d)$ .

The case when  $p \leq q$ . The proof parallels the one in the previous case. Now we have to show

$$(56.23) \quad B_{pp}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d) \subset B_{pq}^s(\mathbb{R}^d).$$

However, with the aid of the first two assertion again, we have  $F_{pq}^s(\mathbb{R}^d) \supset F_{pp}^s(\mathbb{R}^d) = B_{pp}^s(\mathbb{R}^d)$ . As a consequence we have only to show that  $B_{pq}^s(\mathbb{R}^d) \supset F_{pq}^s(\mathbb{R}^d)$ . Now we note that

$$(56.24) \quad \|f_k\|_{\ell^q(L^p)} = \| |f_k|^p \|_{\ell^{q/p}(L^1)}^{\frac{1}{p}}, \quad \|f_k\|_{L^p(\ell^q)} = \left\| |f_k|^p : L^1 \left( l^{\frac{q}{p}} \right) \right\|^{\frac{1}{p}}$$

for all  $\{f_k\}_{k \in \mathbb{N}_0} \subset L^{1,loc}$ . Consequently

$$(56.25) \quad \|f_k\|_{\ell^q(L^p)} \leq \|f_k\|_{L^p(\ell^q)}.$$

Thus, it follows that

$$(56.26) \quad \|f\|_{F_{pq}^s} \geq \|f\|_{B_{pq}^s}.$$

for all  $f \in F_{pq}^s$ . Note that the same as (56.21) applies for this inequality. And the proof is completely finished.  $\square$

Schwartz-functions. In this paragraph we discuss the relations between our function spaces and Schwartz distributions.

**Theorem 56.8.**  $\mathcal{S}(\mathbb{R}^d) \subset A_{pq}^s(\mathbb{R}^d)$  in the sense of continuous embedding. Speaking precisely, we have

$$(56.27) \quad \|f\|_{A_{pq}^s} \lesssim p_N(f), \quad f \in \mathcal{S}(\mathbb{R}^d)$$

for some large  $N$  depending on  $p, q$  and  $s$ .

*Proof.* Let  $\{\varphi_j\}_{j \in \mathbb{N}_0}$  be a family of function from Lemma 5.17. First of all, since we have established

$$(56.28) \quad B_{p \min(p,q)}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d),$$

we may assume that  $A = B$ . Next, by Theorem 55.5 we have

$$(56.29) \quad \|\varphi_0(D)f\|_p \lesssim \|\varphi_0(D)f\|_\infty \lesssim p_N(f).$$

Thus, our present task is to prove the existence of  $N$  such that

$$(56.30) \quad \|\{2^{js} \varphi_j(D)f\}_{j \in \mathbb{N}}\|_{\ell^q(L^p)} \lesssim p_N(f)$$

for all  $f \in \mathcal{S}(\mathbb{R}^d)$ . We apply Theorem 55.9 with  $H_j = (\varphi_{j-1} + \varphi_j + \varphi_{j+1}) \cdot |\cdot|^{-2N}$ , where  $N \geq 1$  is large enough.

$$\begin{aligned} \|\{2^{js}\varphi_j(D)f\}_{j \in \mathbb{N}}\|_{\ell^q(L^p)} &= \|\{2^{j(s-2N)}|2^{-j}D|^{-2N}\varphi_j(D)(-\Delta)^N f\}_{j \in \mathbb{N}}\|_{\ell^q(L^p)} \\ &\lesssim \|\{2^{js-2N}\varphi_j(D)(-\Delta)^N f\}_{j \in \mathbb{N}}\|_{\ell^q(L^p)}. \end{aligned}$$

As a consequence by replacing  $s$  with  $s - 2N$  we may assume that  $s < 0$ .

Once we assume  $s < 0$ , we have only to show that  $\|\varphi_j(D)f\|_p \lesssim p_{N'}(f)$ , where  $c$  and  $N'$  depend only on  $f$  and  $p$ .

By virtue of the continuity of the Fourier transform we obtain

$$(56.31) \quad \|\varphi_j(D)f\|_p \lesssim p_{N(1)}(\varphi_j \cdot \mathcal{F}f)$$

with  $N(1) > N'$ . By direct computation, there exists a constant  $c > 0$  independent of  $j$  such that

$$(56.32) \quad p_{N(1)}(\varphi_j(\xi)\mathcal{F}f) \lesssim p_{N(1)}(\mathcal{F}f) \lesssim p_{N(2)}(f).$$

If we put our observations together, we obtain the desired result.  $\square$

To proceed further, we prove a lemma showing a property of the Besov functions.

**Lemma 56.9.** *Let  $s > 0$ . Then  $B_{\infty\infty}^s(\mathbb{R}^d) \subset \text{BC}(\mathbb{R}^d)$  in the sense of continuous embedding.*

*Proof.* Again we take  $\{\varphi_j\}_{j \in \mathbb{N}_0}$  as the one described in Lemma 5.17 to fix the norm. Then from the definition we have  $2^{js}\|\varphi_j(D)f\|_{\infty} \leq \|f : B_{\infty\infty}^s\|$  for all  $j \in \mathbb{N}_0$ . Inserting this estimate, we have

$$\|f\|_{\infty} \leq \sum_{j=0}^{\infty} \|\varphi_j(D)f\|_{\infty} \leq \sum_{j=0}^{\infty} 2^{-js} \cdot \|f : B_{\infty\infty}^s\| \lesssim \|f : B_{\infty\infty}^s\|.$$

From this inequality we conclude that  $f \in L^{\infty}(\mathbb{R}^d)$ . Furthermore by replacing  $f$  with  $f - \sum_{j=0}^k \varphi_j(D)f$  and using the multiplier theorem, we conclude that  $\sum_{j=0}^k \varphi_j(D)f$  tends to  $f$  in the

$L^{\infty}(\mathbb{R}^d)$  topology. Since  $\sum_{j=0}^k \varphi_j(D)f$  is continuous, we conclude  $f$  is also continuous. Once we establish  $f$  is continuous, we conclude again from the above inequality  $B_{\infty\infty}^s(\mathbb{R}^d)$  is embedded continuously into  $\text{BC}(\mathbb{R}^d)$ .  $\square$

**Theorem 56.10.** *In the sense of continuous embedding, we have  $A_{pq}^s(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$ . More precisely, there exists a constant  $N$  depending only on  $p, q$  and  $s$  such that*

$$(56.33) \quad |\langle f, \varphi \rangle| \lesssim \|f\|_{A_{pq}^s} \cdot p_N(\varphi)$$

for  $f \in A_{pq}^s(\mathbb{R}^d)$  and  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ .

*Proof.* Let  $\kappa$  be a bump function with  $\chi_{B(8)} \leq \kappa \leq \chi_{B(16)}$ . Then, taking into account  $(L^p)^{B(16)}(\mathbb{R}^d) \subset L^{\infty}(\mathbb{R}^d) \cap A_{pq}^s(\mathbb{R}^d)$ , we obtain  $\kappa(D)f \in L^{\infty}(\mathbb{R}^d)$ . Thus we may assume that  $\text{supp}(f) \cap B(8) = \emptyset$ . By the same argument as before instead of considering  $f$  directly we can consider  $g$ , where  $g$  is a solution of  $(-\Delta)^N g = f$  and  $g \in A_{pq}^{s+N}(\mathbb{R}^d)$ , where  $N$  is large enough. Furthermore by  $F_{pq}^s(\mathbb{R}^d) \subset B_{p \max(p,q)}^s(\mathbb{R}^d) \subset B_{p\infty}^s(\mathbb{R}^d)$  we may assume that  $A = B$ . Since

$$(56.34) \quad \|\varphi_j(D)f\|_p \lesssim 2^{\frac{jd}{p}} \|\varphi_j(D)f\|_{\infty},$$

we may assume even that  $p = \infty$ . Since we have shown  $B_{\infty\infty}^s$  is continuously embedding into  $C$ , the proof of  $A_{pq}^s(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$  is now complete.  $\square$

Completeness. To prove the completeness, we need a lemma.

**Lemma 56.11** (Fatou type estimate). *Let  $\{f_j\}_{j \in \mathbb{N}_0} \subset A_{pq}^s(\mathbb{R}^d)$  be a bounded sequence. Suppose that  $\lim_{j \rightarrow \infty} f_j$  converges to  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then  $f \in A_{pq}^s(\mathbb{R}^d)$  and the estimate \**

$$(56.35) \quad \|f\|_{A_{pq}^s} \leq \liminf_{j \rightarrow \infty} \|f_j\|_{A_{pq}^s}$$

holds.

*Proof.* The proof can be obtained by Fatou's lemma and the definition of the norms. □

Lemma 56.11, obtained by virtue of the Fatou lemma, contains the assertion similar to the Fatou lemma.

**Theorem 56.12.**  $A_{pq}^s$  is complete.

*Proof.* Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a Cauchy sequence in  $A_{pq}^s(\mathbb{R}^d)$ . Since  $A_{pq}^s(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$  by Theorem 56.10, we see that at least  $\lim_{j \rightarrow \infty} f_j = f$  exists in  $\mathcal{S}'(\mathbb{R}^d)$ . By Lemma 56.11 we see that  $\lim_{j \rightarrow \infty} f_j = f$  takes place in  $A_{pq}^s(\mathbb{R}^d)$ . □

Density. Let us show that  $\mathcal{S}(\mathbb{R}^d)$  is dense provided  $p$  and  $q$  are finite.

**Lemma 56.13.** *Let  $f \in (L^p)^{B(r)}(\mathbb{R}^d)$  and  $\varepsilon > 0$ . Then there exists  $g \in \mathcal{S}(\mathbb{R}^d) \cap (L^p)^{B(1.1r)}(\mathbb{R}^d)$  such that*

$$(56.36) \quad \|f - g\|_{L^p} \leq \varepsilon.$$

*Proof.* Let  $\rho \in \mathcal{S}(\mathbb{R}^d)$  be taken so that

$$(56.37) \quad \chi_{B(1)} \leq \rho \leq \chi_{B(2)}.$$

Set  $g_k := \rho_k \cdot f$ , where  $\rho_k(x) = \rho(2^{-k}x)$ . Then

$$(56.38) \quad \mathcal{F}g_k = 2^{kn} \mathcal{F}\rho(2^k \cdot) * \mathcal{F}f.$$

implies  $\text{supp}(g_k) \subset B(1.1r)$  for large  $k \in \mathbb{N}$ . Since  $g_k \in \mathcal{S}(\mathbb{R}^d)$  and  $g_k \rightarrow f$  in  $L^p(\mathbb{R}^d)$ , if  $k$  is large enough, then  $g = g_k$  is the desired function. □

**Theorem 56.14.** *If  $p, q < \infty$ , then  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $A_{pq}^s(\mathbb{R}^d)$ .*

*Proof.* Let  $\rho_k, k \in \mathbb{N}$  be the one in the previous lemma. Let  $f_k = \rho_k(D)f$ . Then

$$(56.39) \quad \varphi_j(D)(f - f_k) = \varphi_j(D)f, \quad j \geq k + 3, \quad \varphi_j(D)(f - f_k) = 0, \quad j \leq k - 3.$$

and

$$(56.40) \quad \varphi_j(D)(f - f_k) = \rho_{k+5}(D)(1 - \rho_k(D))\varphi_j(D)f, \quad k - 3 \leq j \leq k + 3.$$

By Theorem 55.9 we have

$$(56.41) \quad \|\varphi_j(D)(f - f_k)\|_p \lesssim \|\varphi_j(D)f\|_p.$$

As a consequence we obtain

$$\begin{aligned} \|f - f_k\|_{F_{pq}^s} &\lesssim \|\{2^{js} \varphi_j(D)f\}_{j \geq k-3}\|_{L^p(\ell^q)} \\ \|f - f_k\|_{B_{pq}^s} &\lesssim \|\{2^{js} \varphi_j(D)f\}_{j \geq k-3}\|_{\ell^q(L^p)}. \end{aligned}$$

Since we are assuming that  $p, q$  are finite, we are in the position of using the Lebesgue's convergence theorem. By virtue of this theorem, we obtain

$$(56.42) \quad \lim_{k \rightarrow \infty} f_k = f$$

in  $A_{pq}^s(\mathbb{R}^d)$ . Consequently for the purpose of approximating  $f$  we may replace  $f$  by  $\rho_K(D)f$  for some large  $K$ . Given  $\varepsilon > 0$  we take  $\eta \in \mathcal{S}(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$  taken so that

$$(56.43) \quad \|f - \eta\|_p \leq \varepsilon, \quad \text{supp}(\eta) \subset B(2^{K+4}).$$

Then another application of Theorem 55.9 gives us that

$$(56.44) \quad \|f - \eta\|_{A_{pq}^s} \leq c_K \sum_{j=0}^{K+5} \|\varphi_j(D)(f - \eta)\|_p \leq c_K \|f - \eta\|_p.$$

Since  $K$  is taken beforehand, we see that  $f$  can be approximated by  $\mathcal{S}(\mathbb{R}^d)$ .  $\square$

Finally we prove that every element  $\mathcal{S}'(\mathbb{R}^d)$  belongs to  $B_{\infty\infty}^s(\mathbb{R}^d)$  for some  $s \in \mathbb{R}$ .

**Theorem 56.15.** *Let  $s \in \mathbb{R}$ . Then we have*

$$(56.45) \quad \bigcup_{s \in \mathbb{R}} B_{\infty\infty}^s(\mathbb{R}^d) = \mathcal{S}'(\mathbb{R}^d)$$

as a set.

*Proof.* In view of Theorem 33.12 and the lift operator theorem (Theorem 56.18), it suffices to check that  $f \in B_{\infty\infty}^{-M}(\mathbb{R}^d)$  for some  $M \in \mathbb{N}$ , whenever  $f$  is a function which grows polonomially. If we re-examine the proof of Theorem 33.12, it suffices to show

$$(56.46) \quad f \in A_{pq}^s(\mathbb{R}^d) \mapsto x_l \cdot f \in A_{pq}^{s-1}(\mathbb{R}^d).$$

This is an immediate consequence of the fact that

$$(56.47) \quad \varphi_j(D)[x_l \cdot f] = c_d \partial_l [\varphi_j](D)f$$

for some constant  $c_d$  depending only on dimension. Thus, the proof is now complete.  $\square$

### 56.3. Elementary inclusions.

Here we collect elementary inclusions for later considerations.

The first two theorems are already dealt. However, we dare repeat them.

**Theorem 56.16.** *Suppose that  $0 < p, q_1, q_2, q \leq \infty, \varepsilon > 0$  and  $s \in \mathbb{R}$ . The following inclusions holds in the sense of continuous embedding.*

- (1)  $A_{p,q_1}^{s+\varepsilon}(\mathbb{R}^d) \subset A_{p,q_2}^s(\mathbb{R}^d)$ .
- (2)  $B_{p,\min(p,q)}^s(\mathbb{R}^d) \subset F_{pq}^s(\mathbb{R}^d) \subset B_{p,\max(p,q)}^s(\mathbb{R}^d)$ .
- (3) Assume in addition that  $q_1 \leq q_2$ . Then we have  $A_{p,q_1}^s(\mathbb{R}^d) \subset A_{p,q_2}^s(\mathbb{R}^d)$ .

*Proof.* What remains to prove

$$(56.48) \quad B_{p,\infty}^{s+\varepsilon}(\mathbb{R}^d) \subset B_{pq}^s(\mathbb{R}^d).$$

by virtue of (2) and (3), which are already established. Let  $f \in B_{p,\infty}^{s+\varepsilon}(\mathbb{R}^d)$ . Then

$$\begin{aligned} \left( \sum_{j=0}^{\infty} 2^{jq_s} \|\varphi_j(D)f\|_p^q \right)^{\frac{1}{q}} &= \left( \sum_{j=0}^{\infty} 2^{-jq\varepsilon} 2^{jq(s+\varepsilon)} \|\varphi_j(D)f\|_p^q \right)^{\frac{1}{q}} \\ &\leq \sup_{j \in \mathbb{N}_0} \left( 2^{j(s+\varepsilon)} \|\varphi_j(D)f\|_p \right) \cdot \left( \sum_{j=0}^{\infty} 2^{-jq\varepsilon} \right)^{\frac{1}{q}} \\ &\lesssim \sup_{j \in \mathbb{N}_0} \left( 2^{j(s+\varepsilon)} \|\varphi_j(D)f\|_p \right). \end{aligned}$$

Thus, this is the desired result. □

Differential index. The following theorem shows that

$$(56.49) \quad s - \frac{d}{p}$$

corresponds to the Sobolev index in  $W^{m,p}(\mathbb{R}^d)$ .

**Theorem 56.17.** *Let  $s_0 > s_1$ . Assume that  $s_0 - \frac{d}{p_0} = s_1 - \frac{d}{p_1}$ . Then*

$$(56.50) \quad B_{p_0q}^{s_0}(\mathbb{R}^d) \subset B_{p_1q}^{s_1}(\mathbb{R}^d), \quad F_{p_0\infty}^{s_0}(\mathbb{R}^d) \subset F_{p_1q}^{s_1}(\mathbb{R}^d).$$

*Proof.*  $B_{p_0q}^{s_0}(\mathbb{R}^d) \subset B_{p_1q}^{s_1}(\mathbb{R}^d)$ . Note that

$$(56.51) \quad \varphi_j(D)f \in L^{p_2 B(2^{j+3})}.$$

Then we are in the position of the theory in the last part to obtain

$$(56.52) \quad \|\varphi_j(D)f\|_{p_1} \lesssim 2^{j\left(\frac{1}{p_1} - \frac{1}{p_2}\right)} \|\varphi_j(D)f\|_{L^{p_2}}.$$

If we insert this estimate in the definition of the norm, our result follows immediately.\* □

*Proof of  $F_{p_0\infty}^{s_0}(\mathbb{R}^d) \subset F_{p_1q}^{s_1}(\mathbb{R}^d)$ .* The matter is not so simple as the B-scale case. We have to prove

$$(56.54) \quad F_{p_0,\infty}^{s_0}(\mathbb{R}^d) \subset F_{p_1,q}^{s_1}(\mathbb{R}^d).$$

Let  $f \in F_{p_0,\infty}^{s_0}(\mathbb{R}^d)$ . Then we have to estimate

$$\left( \int_{\mathbb{R}^d} \|2^{js_1} \varphi_j(D)f(x)\|_{\ell^q}^{p_1} dx \right)^{\frac{1}{p_1}} = \left( \int_0^\infty p_1 \lambda^{p_1-1} |\{ \|2^{js_1} \varphi_j(D)f\|_{\ell^q} > \lambda \}| d\lambda \right)^{\frac{1}{p_1}}.$$

As for  $2^{js_1} \varphi_j(D)f(x)$ , we need two types of estimates. The first one is

$$|2^{js_1} \varphi_j(D)f(x)| \lesssim \|2^{js_1} \varphi_j(D)f\|_\infty \lesssim 2^{js_1 + \frac{jd}{p_0}} \|\varphi_j(D)f\|_{p_0} \lesssim 2^{\frac{jd}{p_1}} \|f\|_{F_{p_0\infty}^{s_0}},$$

where for the proof of the second inequality we have used the inequality in the footnote again. We also have

$$(56.55) \quad |2^{js_1} \varphi_j(D)f(x)| \leq 2^{j(s_1-s_0)} \|\{2^{js_0} \varphi_j(D)f(x)\}_{j=0}^\infty\|_{\ell^\infty}.$$

\* Recall that we have proved

$$(56.53) \quad \|f\|_q \lesssim R^{\frac{1}{q} - \frac{1}{p}} \|f\|_p$$

for  $f \in L^\infty(\mathbb{R}^d)$  with  $\text{supp}(\mathcal{F}f) \subset B(R)$ .

Set

$$(56.56) \quad A(x) := \|\{2^{js_0} \varphi_j(D)f(x)\}_{j=0}^\infty\|_{\ell^\infty}, \quad B := \|f\|_{F_{p_0}^{s_0} \infty}.$$

Then our observation can be summarized as

$$(56.57) \quad |2^{js_1} \varphi_j(D)f(x)| \lesssim \min(2^{j(s_1-s_0)} A(x), 2^{\frac{jd}{p_1}} B)$$

for all  $j \in \mathbb{N}_0$ . It is convenient to put  $\varphi_{-j} \equiv 0$  for  $j \in \mathbb{N}$ . Then (56.57) is still valid for  $j \in \mathbb{Z}$ . If we insert (56.57) to  $\|2^{js_1} \varphi_j(D)f(x)\|_{\ell^q}$ , then we obtain a pointwise estimate of  $\varphi_j(D)f(x)$  :

$$\begin{aligned} \|2^{js_1} \varphi_j(D)f(x)\|_{\ell^q} &\lesssim \left( \sum_{j \in \mathbb{Z}} \min \left( 2^{jq(s_1-s_0)} A(x)^q, 2^{\frac{jqd}{p_1}} B^q \right) \right)^{\frac{1}{q}} \\ &\lesssim \left( \int_0^\infty \min(R^{q(s_1-s_0)} A(x)^q, R^{qd/p_1} B^q) \frac{dR}{R} \right)^{\frac{1}{q}} \\ &\lesssim A(x)^{\frac{p_0}{p_1}} B^{1-\frac{p_0}{p_1}}. \end{aligned}$$

We shall insert this pointwise estimate to the distribution formula.

By changing the variables we obtain

$$\begin{aligned} \|2^{js_1} \varphi_j(D)f\|_{L^{p_1}(\ell^q)} &\leq \left( \int_0^\infty p_1 \lambda^{p_1-1} |\{ \|2^{js_1} \varphi_j(D)f\|_{\ell^q} > \lambda \}| d\lambda \right)^{\frac{1}{p_1}} \\ &\lesssim \left( \int_0^\infty p_1 \lambda^{p_1-1} |\{ A^{\frac{p_0}{p_1}} B^{1-\frac{p_0}{p_1}} > \lambda \}| d\lambda \right)^{\frac{1}{p_1}} \\ &\simeq B^{1-\frac{p_0}{p_1}} \left( \int_0^\infty p_1 \lambda^{p_1-1} |\{ A^{\frac{p_0}{p_1}} > \lambda \}| d\lambda \right)^{\frac{1}{p_1}} \\ &\simeq B^{1-\frac{p_0}{p_1}} \|A\|_{p_0}^{\frac{p_0}{p_1}} \\ &= \|f\|_{F_{p_0}^{s_0} \infty}. \end{aligned}$$

This is the desired result. □

#### 56.4. Lift operators for nonhomogeneous spaces.

In this subsection we prove

**Theorem 56.18.** *Let  $\delta \in \mathbb{R}$ . Then*

$$(56.58) \quad (1 - \Delta)^\delta : A_{pq}^s(\mathbb{R}^d) \rightarrow A_{pq}^{s-2\delta}(\mathbb{R}^d)$$

*is an isomorphism.*

*Proof.* By induction, for each  $\alpha \in \mathbb{N}_0^d$  we can find  $P_\alpha(x) \in \mathbb{C}[x]$  with  $\deg P_\alpha(x) = |\alpha|$  such that

$$(56.59) \quad \partial_x^\alpha \langle x \rangle^\sigma = P_\alpha(x) \langle x \rangle^{\sigma-2|\alpha|}$$

An immediate consequence of this fact is

$$(56.60) \quad \|\rho(\cdot) \langle 2^k \cdot \rangle^\delta\|_{H_2^\sigma} \lesssim 2^{k\delta},$$

where  $\rho \in \mathcal{S}(\mathbb{R}^d)$  satisfies

$$(56.61) \quad \chi_{B(4) \setminus B(2)} \leq \rho \leq \chi_{B(8) \setminus B(1)}.$$



Set  $\rho_j = \rho(2^{-j}\cdot)$  for  $j \in \mathbb{N}$  and we also take  $\rho_0$  so that  $\chi_{B(4)} \leq \rho_0 \leq \chi_{B(8)}$ . We also assume that  $\Phi := \{\varphi_j\}$  in Definition 56.1 satisfy

$$(56.62) \quad \sum_{j=1}^{\infty} \varphi_j(x) \equiv 1.$$

By Theorem 55.9 again we obtain

$$(56.63) \quad \|(1 - \Delta)^{\frac{\delta}{2}} f\|_{A_{pq}^s} \lesssim \|f\|_{A_{pq}^{s+2\delta}}.$$

In fact if  $A = F$ , then

$$\begin{aligned} \|(1 - \Delta)^{\frac{\delta}{2}} f\|_{A_{pq}^s} &= \|2^{js} \langle D \rangle^{2\delta} \varphi_j(D) f\|_{L^p(\ell^q)} \\ &\lesssim \left( \sup_{k \in \mathbb{N}_0} \|2^{-j\delta} \rho_j(2^j \cdot) \langle 2^j \cdot \rangle^{2\delta}\|_{H_2^\sigma} \right) \cdot \|2^{j(s+2\delta)} \varphi_j(D) f\|_{L^p(\ell^q)} \\ &\lesssim \|2^{j(s+2\delta)} \varphi_j(D) f\|_{L^p(\ell^q)}. \end{aligned}$$

The case when  $A = B$  is proved similarly. □

The theorem below describes well the differential feature of  $A_{pq}^s(\mathbb{R}^d)$ .

**Theorem 56.19.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then*

$$(56.64) \quad \|f\|_{A_{pq}^{s+1}} \sim \|f\|_{A_{pq}^s} + \sum_{j=1}^d \|\partial_j f\|_{A_{pq}^s}.$$

*Proof.* The proof of

$$(56.65) \quad \|f\|_{A_{pq}^s} + \sum_{j=1}^d \|\partial_j f\|_{A_{pq}^s} \lesssim \|f\|_{A_{pq}^{s+1}}$$

is already obtained by Theorem 55.9 again. Thus, we are left with the task of establishing the reverse inequality. By Theorem 56.18 we have only to show that

$$(56.66) \quad \|(1 - \Delta) f\|_{A_{pq}^{s-1}} \lesssim \left( \|f\|_{A_{pq}^s} + \sum_{j=1}^d \|\partial_j f\|_{A_{pq}^s} \right).$$

Then we can prove

$$(56.67) \quad \|\partial_j^2 f\|_{A_{pq}^{s-1}} \lesssim \|\partial_j f\|_{A_{pq}^s}$$

by virtue of Theorem 55.9 once more. Thus, by the quasi-triangle inequality (56.66) is justified. □

*Exercise 261.* Let  $k \in \mathbb{N}$ . Suppose that  $f \in A_{pq}^{s-k}(\mathbb{R}^d)$  and that  $(-\Delta)f \in A_{pq}^s(\mathbb{R}^d)$ . Then prove that  $f \in A_{pq}^{s+2}(\mathbb{R}^d)$ .

*Exercise 262.* Let  $m \in \mathbb{N}$ . Establish the following.

- (1)  $1 + (-\Delta)^m : A_{pq}^{s+2m}(\mathbb{R}^d) \rightarrow A_{pq}^s(\mathbb{R}^d)$  is an isomorphism.
- (2)  $\|f\|_{A_{pq}^{s+m}} \sim \|f\|_{A_{pq}^s} + \sum_{j=1}^d \|\partial_j^m f\|_{A_{pq}^s}.$

57. THE SPACE  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ 

## 57.1. Definition.

**Definition 57.1** ( $\mathcal{P}$  and  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ ). Let  $\mathcal{P} = \mathbb{C}[X]$ , the algebra of complex polynomials. Define an equivalence relation  $\sim$  as follows:

$$(57.1) \quad f \sim g \iff f - g \in \mathcal{P}.$$

Finally equip  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P} := \mathcal{S}'(\mathbb{R}^d)/\sim$  with the quotient topology.

**Lemma 57.2.**  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  is a topological vector space.

*Proof.* We shall prove the continuity of addition. Note that

$$(57.2) \quad \mathcal{S}'(\mathbb{R}^d) \times \mathcal{S}'(\mathbb{R}^d)/(\mathcal{P} \times \mathcal{P}) \simeq \mathcal{S}'(\mathbb{R}^d)/\mathcal{P} \times \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$$

by (9.39). Therefore we have only to show

$$(57.3) \quad [(f, g)] \in \mathcal{S}'(\mathbb{R}^d) \times \mathcal{S}'(\mathbb{R}^d)/(\mathcal{P} \times \mathcal{P}) \rightarrow [f + g] \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$$

is continuous. However,  $(f, g) \in \mathcal{S}'(\mathbb{R}^d) \times \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  is continuous. Thus, the mapping in question is continuous.  $\square$

The following lemma, which the author borrowed from [103, Lemma 6.1], is a key to our observations.

**Lemma 57.3.** Let  $L \in \mathbb{N}$ . Set

$$(57.4) \quad \mathcal{S}(\mathbb{R}^d)_L(\mathbb{R}^d) := \{\varphi \in \mathcal{S}(\mathbb{R}^d) : \partial^\alpha \varphi(0) = 0 \text{ for } |\alpha| \leq L - 1\}.$$

Then there exists a collection  $\{T_\gamma\}_{\gamma \in \mathbb{Z}^d}$  of continuous mappings from  $\mathcal{S}(\mathbb{R}^d)_L(\mathbb{R}^d)$  to  $\mathcal{S}(\mathbb{R}^d)$  such that

$$(57.5) \quad \varphi(x) = \sum_{\gamma \in \mathbb{N}_0^d, |\gamma|=L} x^\gamma T_\gamma \varphi(x)$$

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

*Proof.* If  $d = 1$ , then the result is immediate by using the fundamental theorem of calculus. Suppose that the theorem is true for  $d - 1$  with  $d \geq 2$ .

Pick a function  $a \in C^\infty(\mathbb{R})$  satisfying  $\chi_{[-1,1]} \leq a \leq \chi_{[-2,2]}$ . Set  $\gamma_0 := (0, 0, \dots, 0, d + 1)$ . For  $\varphi \in \mathcal{S}(\mathbb{R}^d)_L$ , we define

$$(57.6) \quad R(\varphi)(x_1, x_2, \dots, x_d) = \frac{a(x_d)}{x_d^L} \left( \varphi(x_1, x_2, \dots, x_d) - \sum_{j=0}^L \frac{\partial_d^j \varphi(x_1, x_2, \dots, x_{d-1}, 0)}{j!} x_1^j \right)$$

for  $x_d \in \mathbb{R} \setminus \{0\}$  and  $x_1, x_2, \dots, x_{d-1} \in \mathbb{R}$ . We extend  $R(\varphi)$  to  $\mathbb{R}^d$  by continuity. By the fundamental theorem of calculus  $\varphi \mapsto R(\varphi)$  is a continuous operation from  $\mathcal{S}(\mathbb{R}^d)_L$  to  $\mathcal{S}(\mathbb{R}^d)$ .

$$(57.7) \quad T_{\gamma_0} \varphi(x_1, x_2, \dots, x_d) = R(\varphi)(x_1, x_2, \dots, x_d) + \frac{1 - a(x_d)}{x_d^L} \varphi(x_1, x_2, \dots, x_d).$$

Then  $\varphi \mapsto T_{\gamma_0}(\varphi)$  is a continuous operation. Observe that

$$(57.8) \quad \varphi(x_1, x_2, \dots, x_d) = x_d^L T_{\gamma_0} \varphi(x_1, x_2, \dots, x_d) + g(x_d) \sum_{j=0}^{d-1} \frac{\partial_d^j \varphi(x_1, x_2, \dots, x_{d-1}, 0)}{j!} x_d^j.$$

It remains to apply the induction assumption. Thus, the lemma is proved completely.  $\square$

By taking the Fourier transform we see the following.

**Corollary 57.4.** *Let  $L \in \mathbb{N}$ . Set*

$$(57.9) \quad \mathcal{S}(\tilde{\mathbb{R}}^d)_L(\mathbb{R}^d) := \{\varphi \in \mathcal{S}(\mathbb{R}^d) : \mathcal{F}\varphi \perp \mathcal{P}_{L-1}\}.$$

*Then there exists a collection  $\{T_\gamma\}_{\gamma \in \mathbb{Z}^d}$  of continuous mappings from  $\mathcal{S}(\tilde{\mathbb{R}}^d)_L(\mathbb{R}^d)$  to  $\mathcal{S}(\mathbb{R}^d)$  such that*

$$(57.10) \quad \varphi(x) = \sum_{\gamma \in \mathbb{N}_0^d, |\gamma|=L} \partial^\gamma T_\gamma \varphi(x)$$

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

We transform this corollary to the one we want to use.

**Corollary 57.5.** *Let  $L \in \mathbb{N}$  and define  $\mathcal{S}(\tilde{\mathbb{R}}^d)_L$  as above. Then there exists a collection of continuous mappings  $\{T_\gamma\}_{\gamma \in \mathbb{N}_0^d, |\gamma|=L}$ , polynomials  $\{P_\gamma\}_{\gamma \in \mathbb{N}_0^d, |\gamma| \leq L-1} \in \mathcal{P}(\mathbb{R}^d)$  and constants  $\{c_\gamma\}_{\gamma \in \mathbb{N}_0^d, |\gamma| \leq L-1}$  such that*

$$(57.11) \quad \varphi(x) = \sum_{\gamma \in \mathbb{N}_0^d, |\gamma| \leq L-1} c_\gamma \cdot \left( \int_{\mathbb{R}^d} \varphi(y) P_\gamma(y) dy \right) \cdot x^\gamma e^{-|x|^2} + \sum_{\gamma \in \mathbb{N}_0^d, |\gamma|=L} \partial^\gamma T_\gamma \varphi(x)$$

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

With this preparation in mind, we shall prove a key theorem in this section.

**Theorem 57.6.** *Let  $L \in \mathbb{N}$ . Suppose that  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence such that*

$$(57.12) \quad \{\partial^\alpha f_j\}_{j \in \mathbb{N}}$$

*is convergent for all  $\alpha$ , provided  $|\alpha| = L$ . Then there exists a sequence of polynomials  $\{P_j\}_{j \in \mathbb{N}}$  such that*

$$(57.13) \quad \lim_{j \rightarrow \infty} (f_j + P_j)$$

*is convergent.*

*Proof.* We have only to let  $P_j(x) := \sum_{\gamma \in \mathbb{N}_0^d, |\gamma| \leq L-1} c_\gamma \langle f, \cdot E_\gamma \rangle \cdot P_\gamma(x)$ , where  $E_\gamma(x) = x^\gamma e^{-|x|^2}$ . □

### 57.2. $\mathcal{S}(\mathbb{R}^d)_0$ and $\mathcal{S}(\mathbb{R}^d)'_0$ .

**Definition 57.7.** Define  $\mathcal{S}(\mathbb{R}^d)_0$  by  $\mathcal{S}(\mathbb{R}^d)_0 := \bigcap_{L=0}^\infty \mathcal{S}(\mathbb{R}^d) \cap \mathcal{P}_L(\mathbb{R}^d)$ . Endow  $\mathcal{S}(\mathbb{R}^d)_0$  with the induced topology of  $\mathcal{S}(\mathbb{R}^d)$ .  $\mathcal{S}(\mathbb{R}^d)'_0$  is a topological dual of  $\mathcal{S}(\mathbb{R}^d)_0$ , that is,

$$(57.14) \quad \mathcal{S}(\mathbb{R}^d)'_0 := \{F \in \text{Hom}_{\mathbb{C}}(\mathcal{S}(\mathbb{R}^d)_0, \mathbb{C}) : F : \mathcal{S}(\mathbb{R}^d)_0 \rightarrow \mathbb{C} \text{ is continuous}\}.$$

We now hesitate to equip  $\mathcal{S}(\mathbb{R}^d)'_0$  with a topology.

Pick a smooth function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  supported on  $B(8) \setminus B(1)$  so that  $\sum_{j=-\infty}^\infty \varphi(2^{-j}x) \equiv 1$  for  $x \in \mathbb{R}^d \setminus \{0\}$ . Set  $\varphi_j(x) = \varphi(2^{-j}x)$  for  $j \in \mathbb{Z}$ .

**Proposition 57.8.** *Keep to the notation above. Then we have the following.*

$$(1) \quad \lim_{J \rightarrow \infty} \sum_{j=-J}^J \varphi_j(D) \tau = \tau \text{ for all } \tau \in \mathcal{S}(\mathbb{R}^d)_0.$$

$$(2) \lim_{J \rightarrow \infty} \sum_{j=-J}^J \varphi_j(D)F = F \text{ for all } F \in \mathcal{S}(\mathbb{R}^d)'.$$

*Proof.* As we did so in the case of  $\mathcal{S}(\mathbb{R}^d)$ , we have only to verify the former formula. Then the latter follows automatically.

Let  $\tau \in \mathcal{S}(\mathbb{R}^d)_0$ . Set  $X_J := B(2^{3-J}) \cup (\mathbb{R}^d \setminus B(2^J))$ . Then

$$\begin{aligned} & p_{\alpha, \beta} \left( \mathcal{F} \left\{ \tau - \sum_{j=-J}^J \varphi_j(D)\tau \right\} \right) \\ &= \sup_{x \in \mathbb{R}^d} \left| x^\alpha \partial^\beta \left( \mathcal{F}\tau(x) - \sum_{j=-J}^J \varphi_j(x)\mathcal{F}\tau(x) \right) \right| \\ &\leq \sup_{x \in \mathbb{R}^d} \left| \left( 1 - \sum_{j=-J}^J \varphi_j(x) \right) x^\alpha \partial^\beta \mathcal{F}\tau(x) \right| + c_\beta \sup_{x \in \mathbb{R}^d} \sum_{\substack{\gamma, \delta \in \mathbb{N}_0^d \\ |\gamma|, |\delta| \leq |\beta|}} \left| x^\alpha \partial^\gamma \mathcal{F}\tau(x) \partial_x^\delta \left[ \sum_{j=-J}^J \varphi(2^{-j}x) \right] \right| \\ &\leq \sup_{x \in X_J} |x^\alpha \partial^\beta \mathcal{F}\tau(x)| + c_\beta \sum_{\substack{\gamma, \delta \in \mathbb{N}_0^d \\ |\gamma|, |\delta| \leq |\beta|}} \sup_{x \in X_J} |x^\alpha \partial^\gamma \mathcal{F}\tau(x)| |x|^{-\delta}. \end{aligned}$$

Since we are assuming  $\tau \in \mathcal{S}(\mathbb{R}^d)_0$ , the most right-hand side of the above chain of inequality tends to 0. Therefore, it follows that  $\lim_{J \rightarrow \infty} \sum_{j=-J}^J \varphi_j(D)\tau = \tau$ . □

**Proposition 57.9.** *Suppose  $\{F_j\}_{j \in \mathbb{N}}$  is a convergent sequence. Then there exists  $N \in \mathbb{N}$  so that*

$$(57.15) \quad |\langle F_j, \varphi \rangle| \leq N p_N(\varphi)$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)_0$ .

*Proof.* We have only to go through the same argument as  $\mathcal{S}(\mathbb{R}^d)$ . □

**Theorem 57.10.** *We have*

$$(57.16) \quad \mathcal{S}'(\mathbb{R}^d)/\mathcal{P} \simeq \mathcal{S}(\mathbb{R}^d)'_0.$$

as linear spaces.

*Proof.* We shall construct linear mappings  $\Phi : \mathcal{S}'(\mathbb{R}^d)/\mathcal{P} \rightarrow \mathcal{S}(\mathbb{R}^d)'_0$  and  $\Psi : \mathcal{S}(\mathbb{R}^d)'_0 \rightarrow \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  which are inverse to each other.

The construction of the mapping  $\Phi : \mathcal{S}'(\mathbb{R}^d)/\mathcal{P} \rightarrow \mathcal{S}(\mathbb{R}^d)'_0$

Let  $P \in \mathcal{P}$ . Then we have, for every  $\tau \in \mathcal{S}(\mathbb{R}^d)_0$ ,  $\int_{\mathbb{R}^d} P(x)\tau(x) dx = 0$ . Therefore the restriction mapping  $f \in \mathcal{S}'(\mathbb{R}^d) \mapsto f|_{\mathcal{S}(\mathbb{R}^d)_0} \in \mathcal{S}(\mathbb{R}^d)'_0$  factors  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ .

The construction of the mapping  $\Psi : \mathcal{S}(\mathbb{R}^d)'_0 \rightarrow \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$

Let  $f \in \mathcal{S}(\mathbb{R}^d)'_0$ . Then there exists  $N \in \mathbb{N}$  so that  $|\langle f, \tau \rangle| \leq N p_N(\tau)$ . Denote by  $\mathcal{S}(\mathbb{R}^d)'_0^{(N)}$  the completion of  $\mathcal{S}(\mathbb{R}^d)_0$  with respect to the norm  $p_N$ . Then  $f$  admits a continuous extension

$f'$  to  $\mathcal{S}(\mathbb{R}^d)_0^{(N)}$  satisfying  $|\langle f', \tau \rangle| \leq N p_N(\tau)$  for all  $\tau \in \mathcal{S}(\mathbb{R}^d)_0^{(N)}$ . By the Hahn-Banach theorem this mapping extends to a mapping  $F$  on  $\mathcal{S}(\mathbb{R}^d)^{(N)}$ , where  $\mathcal{S}(\mathbb{R}^d)^{(N)}$  is a completion of  $\mathcal{S}(\mathbb{R}^d)$  with respect to the norm  $p_N$ . Denote by  $F_0$  the restriction of  $F$  to  $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)^{(N)}$ . From the definition  $F_0$  belongs to  $\mathcal{S}'(\mathbb{R}^d)$ . We shall claim

$$(57.17) \quad f \in \mathcal{S}(\mathbb{R}^d)'_0 \mapsto [F_0] \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$$

is a well-defined mapping. Once this is archived, we see that the mapping is a linear isomorphism.

To see that (57.17) makes sense we have to verify that the definition is independent of the choice of  $F$ . Recall that  $F$  is obtained by the Hahn-Banach theorem.

Let  $G$  be an another extension of  $f : \mathcal{S}(\mathbb{R}^d)_0^{(N)} \rightarrow \mathbb{K}$ . Then  $F_0 - G_0$  annihilates  $\mathcal{S}(\mathbb{R}^d)_0$ . Therefore, if  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  does not contain 0 as its support, then  $\langle \mathcal{F}(F_0 - G_0), \varphi \rangle = \langle F - G, \mathcal{F}\varphi \rangle = 0$ . This means  $\mathcal{F}(F_0 - G_0)$  can be expressed as  $\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq L}} a_\alpha \partial^\alpha \delta_0$ . Therefore  $F_0$  and  $G_0$  coincide

modulo polynomials.

To see that (57.17) is continuous, let us choose a net. □

**Definition 57.11.** Equip  $\mathcal{S}(\mathbb{R}^d)'_0$  with a topology so that (57.16) is an isomorphism.

### 58. SPACES OF HOMOGENEOUS TYPE

Having cleared up the structure of  $\mathcal{S}(\mathbb{R}^d)_0$ , we define  $\dot{F}_{pq}^s$  and  $\dot{B}_{pq}^s$ .

**Definition 58.1.** Let  $0 < p, q \leq \infty$  and  $s \in \mathbb{R}$ . Pick a function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  so that

$$(58.1) \quad \chi_{B(4) \setminus B(2)} \leq \varphi \leq \chi_{B(8) \setminus B(1)}.$$

Let  $\varphi_j(x) = \varphi(2^{-j}x)$  for  $j \in \mathbb{Z}$ .

(1) Let  $f \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ . Then define the seminorm

$$(58.2) \quad \|f : \dot{B}_{pq}^s\| := \|\{\varphi_j(D)f\}_{j \in \mathbb{Z}}\|_{\ell^q(L^p)}.$$

$\dot{B}_{pq}^s$  is the set of elements  $f \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  for which the norm  $\|f : \dot{B}_{pq}^s\| < \infty$ .

(2) Suppose that  $p < \infty$ . Let  $f \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ . Then define the seminorm

$$(58.3) \quad \|f : \dot{F}_{pq}^s\| := \|\{\varphi_j(D)f\}_{j \in \mathbb{Z}}\|_{\ell^q(L^p)}.$$

$\dot{F}_{pq}^s$  is the set of elements  $f \in \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  for which the norm  $\|f : \dot{F}_{pq}^s\| < \infty$ .

(3)  $\dot{A}_{pq}^s$  denotes either  $\dot{B}_{pq}^s$  or  $\dot{F}_{pq}^s$ . In the case when  $A = F$  we tacitly exclude  $p = \infty$ .

Going through the same arguments as the nonhomogeneous case, we can prove the following.

**Theorem 58.2.** Let  $0 < p, q \leq \infty$  and  $s \in \mathbb{R}$ .

(1) The different choice of  $\varphi$  for the definition of  $\dot{A}_{pq}^s$  gives an equivalent norm.

(2)  $\dot{A}_{pq}^s$  is a quasi-normed space.

(3)  $\dot{B}_{p \min(p,q)}^s \subset \dot{F}_{pq}^s \subset \dot{B}_{p \max(p,q)}^s$ .

*Proof.* Reexamine the proofs of Theorems 56.4, 56.6 and 56.7. □

Like the nonhomogenous case, we can prove the following theorems on the lifting properties.

**Theorem 58.3.** Let  $\alpha \in \mathbb{R}$ . Choose a smooth function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  so that

$$(58.4) \quad \text{supp}(\varphi) \subset B(8) \setminus B(1), \quad \sum_{j=-\infty}^{\infty} \varphi(2^{-j}\cdot) \equiv 1 \text{ on } \mathbb{R}^d \setminus \{0\}.$$

Then define

$$(58.5) \quad (-\Delta)^\alpha f := \lim_{J \rightarrow \infty} \sum_{j=-J}^{\infty} \mathcal{F}^{-1}(|\cdot|^\alpha \varphi(2^{-j}\cdot) \mathcal{F}f)$$

for  $f \in \mathcal{S}(\mathbb{R}^d)_0$ .

- (1) If  $f \in \mathcal{S}(\mathbb{R}^d)_0$ , then the convergence takes place in  $\mathcal{S}(\mathbb{R}^d)_0$ . In particular it can be extended by duality to  $\mathcal{S}(\mathbb{R}^d)'_0$ .
- (2) The series defining  $(-\Delta)^\alpha f$  converges in the topology of  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$ . And it is independent of the choice of  $\varphi$  above.

*Proof.* (1) can be checked immediately by taking the Fourier transform of both sides. For the proof of the first assertion (2) we have only to prove

$$(58.6) \quad \partial^\alpha \left( \sum_{j=-J}^{\infty} \mathcal{F}^{-1}(|\cdot|^\alpha \varphi(2^{-j}\cdot) \mathcal{F}f) \right)$$

converges.

Finally the second assertion is immediate because of the duality formula.

$$(58.7) \quad \langle (-\Delta)^\alpha f, \varphi \rangle = \langle f, (-\Delta)^\alpha \varphi \rangle$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)_0$ . □

*Exercise 263.* Suppose that  $F : \mathbb{R} \rightarrow \mathbb{C}$  is a continuous function such that  $F$  is differentiable at every point in  $\mathbb{R} \setminus \{0\}$ . Assume that  $a = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} F'(t)$  exists. Then  $F$  is differentiable even at the

origin and  $F'(0) = a$ .

**Theorem 58.4.**  $\mathcal{S}(\mathbb{R}^d)_0 \subset \dot{A}_{pq}^s \subset \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  in the sense of continuous embedding. That is, there exists  $N = N_{p,q,s} \in \mathbb{N}$  and  $c > 0$  such that

$$(58.8) \quad \|f\|_{\dot{A}_{pq}^s} \lesssim p_N(f) \text{ for all } f \in \mathcal{S}(\mathbb{R}^d)_0, \quad |\langle f, g \rangle| \lesssim \|f\|_{\dot{A}_{pq}^s} \cdot p_N(g) \text{ for all } g \in \mathcal{S}(\mathbb{R}^d)_0.$$

*Proof.* First we shall prove  $\mathcal{S}(\mathbb{R}^d)_0 \subset \dot{A}_{pq}^s$ . Let  $f \in \mathcal{S}(\mathbb{R}^d)_0$  and  $N \in \mathbb{N}$ . By integration by parts and scaling, we can prove

$$(58.9) \quad \|\varphi_j(D)f\|_p \lesssim 2^{-jN} p_N(f),$$

where  $c > 0$ . Once (58.9) is proved, we have only to use  $\dot{B}_{p \min(p,q)}^s \subset \dot{F}_{pq}^s$ .

For the proof of  $\dot{A}_{pq}^s \subset \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  we can reduce the matter to the case when  $A = B$  because  $\dot{F}_{pq}^s \subset \dot{B}_{p \max(p,q)}^s$ . Note that

$$(58.10) \quad \dot{B}_{pq}^s \subset \dot{B}_{\infty\infty}^{s-\frac{d}{p}},$$

whose proof is entirely the same as the nonhomogeneous case. Therefore, the matters are reduced to showing that  $\Phi : \dot{B}_{00}^{-d-1} \rightarrow \mathcal{S}'(\mathbb{R}^d)$  in the sense of continuous embedding, where

$$(58.11) \quad \langle \Phi([f]), \tau \rangle = \sum_{j=-\infty}^{\infty} \langle \varphi_j(D)f, \tau \rangle.$$

Since

$$(58.12) \quad |\langle \varphi_j(D)f, \tau \rangle| \leq \|\varphi_j(D)f\|_\infty \cdot \|\tau\|_1 \leq 2^{(d+1)j} \|f\| : \dot{B}_{\infty\infty}^{-d-1} \| \cdot \|\tau\|_1$$

for  $j \leq 0$  and

$$\begin{aligned} |\langle \varphi_j(D)f, \tau \rangle| &= 2^{-j(d+2)} |\langle (|2^{-j} \cdot |^{2d+2} \varphi_j)(D)f, \Delta^{d+1} \tau \rangle| \\ &\leq 2^{-j(d+2)} \|(|2^{-j} \cdot |^{2d+2} \varphi_j)(D)f\|_\infty \cdot \|\Delta^{d+1} \tau\|_1 \\ &\lesssim 2^{-j(d+2)} \|\varphi_j(D)f\|_\infty \cdot \|\Delta^{d+1} \tau\|_1 \\ &\lesssim 2^{-j(d+1)} \|f\|_{\dot{A}_{pq}^s} \cdot \|\Delta^{d+1} \tau\|_1 \end{aligned}$$

for  $j \geq 0$ , we conclude

$$(58.13) \quad |\langle \Phi([f]), \tau \rangle| \lesssim \|f\|_{\dot{A}_{pq}^s} \cdot p_{d+1}(\varphi).$$

Thus, the theorem is proved. □

**Theorem 58.5.** *Let  $p, q < \infty$ . Then  $\mathcal{S}(\mathbb{R}^d)_0$  is dense in  $\dot{A}_{pq}^s(\mathbb{R}^d)$ .*

*Proof.* Note that  $\sum_{k=-l}^l \varphi_k(D)f$  converges to  $f$  in  $\dot{A}_{pq}^s$  as  $l \rightarrow \infty$ . Thus we may assume that  $f$  is

the element of the form  $\sum_{k=-l}^l \varphi_k(D)f$ . Once we reduce the matter to approximating the element of the above type, the rest is the same as the nonhomogeneous case. □

**Theorem 58.6.** *The space  $\dot{A}_{pq}^s(\mathbb{R}^d)$  is complete. That is, if  $\{f_k\}_{k \in \mathbb{N}}$  is a sequence in  $\dot{A}_{pq}^s(\mathbb{R}^d)$  such that*

$$(58.14) \quad \lim_{K \rightarrow \infty} \left( \sup_{j, k \geq K} \|f_j - f_k\|_{\dot{A}_{pq}^s} \right) = 0,$$

*then there exists  $f \in \dot{A}_{pq}^s(\mathbb{R}^d)$  such that  $\lim_{k \rightarrow \infty} \|f - f_k\|_{\dot{A}_{pq}^s} = 0$ .*

*Proof.* Since any Cauchy sequence has a limit in  $\mathcal{S}'/\mathcal{P}$ , we have only to apply the Fatou lemma as we did in the nonhomogeneous case. □

The following lemma describes the relation between the dotted spaces and the non-dotted spaces.

**Theorem 58.7.** *The projection  $f \in \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  induces a continuous mapping*

$$(58.15) \quad f \in A_{pq}^s(\mathbb{R}^d) \mapsto [f] \in \dot{A}_{pq}^s(\mathbb{R}^d).$$

*Proof.* Let  $\Phi \in \mathcal{S}(\mathbb{R}^d)$  taken so that  $\chi_{B(4)} \leq \Phi \leq \chi_{B(8)}$ . We deal only with the  $F$ -scale, the case of  $B$ -scale being the same. Then

$$(58.16) \quad \|\{2^{js} \varphi_j(D)f\}_{j=-\infty}^0\|_{L^p(\ell^q)} \lesssim \left( \sup_{j \in \mathbb{N}_0} \|\varphi_{-j}\|_{H_2^\sigma} \right) \|\Phi(D)f\|_{L^p(\ell^q)}.$$

This is the desired result. □

## 59. CONCRETE SPACES

## 59.1. Potential spaces and Sobolev spaces.

**Theorem 59.1.** *Let  $1 < p < \infty$  and  $s \in \mathbb{R}$ . Then,*

$$(59.1) \quad \|f\|_{F_{p2}^s} \sim \|f\|_{H_p^s}$$

for all  $f \in \mathcal{S}'(\mathbb{R}^d)$ . In this sense we can say that  $F_{p2}^s(\mathbb{R}^d) = H_p^s$ .

*Proof.* From the definition of  $H_p^s$ , we readily obtain the lifting property of  $H_p^s$ . Since  $F_{pq}^s$  has the same property, we may assume that  $s = 0$ , that is, we can reduce the matter to showing  $L^p(\mathbb{R}^d) \simeq F_{p2}^0(\mathbb{R}^d)$ , which we have struggled to prove.  $\square$

The Littlewood-Paley theorem gives us the following as well.

**Theorem 59.2.** *Let  $1 < p < \infty$ . Then  $\dot{F}_{p2}^0(\mathbb{R}^d) \simeq L^p(\mathbb{R}^d)$  with norm equivalence. Speaking precisely, we have the following.*

(1) *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  such that  $[f] \in \dot{F}_{p2}^0(\mathbb{R}^d)$ . Then  $g := \lim_{J \rightarrow \infty} \sum_{j=-J}^{\infty} \varphi_j(D)f$  exists in  $L^p(\mathbb{R}^d)$  and coincides  $f$  modulo polynomial. Furthermore there exists  $c > 0$  such that*

$$(59.2) \quad \|g\|_p \lesssim \|[f]\|_{\dot{F}_{p2}^0}$$

for all  $f \in \mathcal{S}'(\mathbb{R}^d)$  with  $[f] \in \dot{F}_{p2}^0(\mathbb{R}^d)$ .

(2) *Let  $f \in L^p(\mathbb{R}^d)$ . Then  $[f] \in \dot{F}_{p2}^0(\mathbb{R}^d)$ . Furthermore there exists  $c > 0$  such that*

$$(59.3) \quad \|[f]\|_{\dot{F}_{p2}^0} \lesssim \|f\|_p$$

for all  $f \in L^p(\mathbb{R}^d)$ .

**Definition 59.3.** Let  $1 \leq p \leq \infty$  and  $m \in \mathbb{N}$ . Let  $f \in L^p(\mathbb{R}^d)$ . One defines

$$(59.4) \quad \|f\|_{W^{m,p}} := \|f\|_p + \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_p.$$

Denote by  $W^{m,p}(\mathbb{R}^d)$  the space of all  $L^p(\mathbb{R}^d)$ -functions for which  $\|f\|_{W^{m,p}} < \infty$ .

As well as the fact  $W^{m,p}(\mathbb{R}^d) \simeq F_{p2}^m(\mathbb{R}^d)$  with  $1 < p < \infty$ , it is important to know that lower or mixed derivatives can be estimated by the pure derivatives of the highest order and the lowest derivative. This fact can be stated as follows:

**Theorem 59.4.** *Let  $1 < p < \infty$  and  $m \in \mathbb{N}_0$ . Then*

$$(59.5) \quad \|f\|_{F_{p2}^s} \simeq \|f\|_{W^{m,p}}$$

for all  $f \in \mathcal{S}'(\mathbb{R}^d)$ . In this sense we can say that  $F_{p2}^m(\mathbb{R}^d) = W^{m,p}(\mathbb{R}^d)$ . Furthermore, for every  $f \in W^{m,p}(\mathbb{R}^d)$ ,

$$(59.6) \quad \|f\|_p + \sum_{j=1}^m \|\partial_j^m f\|_p \sim \|f\|_{W^{m,p}}.$$

*Proof.* For the proof we may assume that  $f \in \mathcal{S}(\mathbb{R}^d)$  because  $\mathcal{S}(\mathbb{R}^d)$  is dense both in  $W^{m,p}(\mathbb{R}^d)$  and  $F_{p2}^m(\mathbb{R}^d)$ . Note that

$$(59.7) \quad \left\| \left( 1 + \sum_{j=1}^d \partial_j^{4m} \right) f \right\|_{F_{p2}^{-3m}} \simeq \|f\|_{F_{p2}^m}$$



by the lift operator property. Although, we have just proved that

$$(59.8) \quad \left\| \left( 1 + \sum_{j=1}^d \partial_j^2 \right)^m f \right\|_{F_{p^2}^{-m}} \simeq \|f\|_{F_{p^2}^m}$$

the same proof works for (59.7). Thus, by using  $F_{pq}^s \subset F_{pq}^{s-1}$  and  $\partial_j : F_{pq}^s \rightarrow F_{pq}^{s-1}$  we obtain

$$\|f\|_{W^{m,p}} = \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_p \leq \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{F_{p^2}^0} \lesssim \sum_{|\alpha| \leq m} \|f\|_{F_{p^2}^{|\alpha|}} \lesssim \|f\|_{F_{p^2}^m}.$$

By lifting the smoothness to  $s + 4m$ , we have

$$\|f\|_{F_{p^2}^m} \lesssim \left\| \left( 1 + \sum_{j=1}^d \partial_j^{4m} \right) f \right\|_{F_{p^2}^{-3m}} \lesssim \left( \|f\|_{F_{p^2}^{-3m}} + \sum_{j=1}^d \|\partial_j^{4m} f\|_{F_{p^2}^{-3m}} \right).$$

If we invoke the key relation  $F_{p^2}^0(\mathbb{R}^d) = L^p(\mathbb{R}^d)$  once more, we are led to

$$\|f\|_{F_{p^2}^m} \lesssim \left( \|f\|_{F_{p^2}^0} + \sum_{j=1}^d \|\partial_j^m f\|_{F_{p^2}^0} \right) \lesssim \left( \|f\|_p + \sum_{j=1}^d \|\partial_j^m f\|_p \right).$$

The last quantity being equal to  $\|f\|_{W^{m,p}}$ , a chain of these inequalities gives the desired result.  $\square$

We also have the following variant that seems useful for the partial differential equation when we consider so called “absorbing argument”. The proof is left to the readers as an exercise: we have only to mimic the proof of Theorem 59.4.

*Exercise 264.* Let  $0 < p, q \leq \infty$  and  $s \in \mathbb{R}$ . For all  $\varepsilon > 0$  we can take  $C_\varepsilon > 0$  such that

$$(59.9) \quad \|f\|_{A_{pq}^s} \leq \varepsilon \|f\|_{A_{pq}^{s+1}} + C_\varepsilon \|f\|_{A_{pq}^{s-1}}.$$

Here we present an application. Recall that  $\Delta : L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)$  is defined as follows:

**Definition 59.5.** Define an unbounded operator  $\Delta : L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)$  as follows:

$$(59.10) \quad D(\Delta) = W_p^2(\mathbb{R}^d), \quad \Delta f := \sum_{j=1}^d \partial_j^2 f, \quad f \in W_p^2(\mathbb{R}^d).$$

**Proposition 59.6.**  $\Delta$  is a closed operator in  $L^p(\mathbb{R}^d)$ , if  $1 < p < \infty$ .

*Proof.* Let  $f_k, k = 1, 2, \dots$  be a sequence in  $W_p^2(\mathbb{R}^d)$  such that

$$(59.11) \quad f_k \rightarrow f \text{ in } L^p(\mathbb{R}^d), \quad \Delta f_k \rightarrow g \text{ in } L^p(\mathbb{R}^d).$$

Then

$$(59.12) \quad \|f_k\|_{F_{p^2}^2} \simeq \|f_k\|_{F_{p^2}^0} + \|\Delta f_k\|_{F_{p^2}^0} \simeq \|f_k\|_p + \|\Delta f_k\|_p$$

is bounded. Since  $f_k \rightarrow f$  in  $L^p(\mathbb{R}^d)$  and hence  $f_k \rightarrow f$  in  $\mathcal{S}'(\mathbb{R}^d)$ , the Fatou’s lemma gives us  $f \in F_{p^2}^0(\mathbb{R}^d)$  with

$$(59.13) \quad \|f\|_{F_{p^2}^2} \leq \liminf_{k \rightarrow \infty} \|f_k\|_{F_{p^2}^2}.$$

This means that  $f \in W_p^2(\mathbb{R}^d)$ . Since  $f \in W_p^2(\mathbb{R}^d)$ , we have for every  $\varphi \in C_c^\infty(\mathbb{R}^d)$

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x)g(x) dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(x)\Delta f_k(x) dx \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} \Delta\varphi(x)f_k(x) \\ &= \int_{\mathbb{R}^d} \Delta\varphi(x) \cdot f(x) dx \\ &= \int_{\mathbb{R}^d} \varphi(x)\Delta f(x) dx, \end{aligned}$$

which yields  $\Delta f = g$ . □

In the same way as above we can prove the following theorem. We excluded the case when the function space is not a Banach space for the sake of simplicity.

**Theorem 59.7.** *Let  $1 \leq p, q \leq \infty$ . Then  $\Delta$  is closed on  $A_{pq}^s(\mathbb{R}^d)$  with its domain  $A_{pq}^{s+2}(\mathbb{R}^d)$ .*

## 59.2. Lipschitz spaces.

Finally we present another concrete space.

$$(59.14) \quad \|f\|_{\text{Lip}_\gamma \cap \text{BC}} = \|f\|_{\text{BC}} + \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\gamma}.$$

We are to prove the following theorem.

**Theorem 59.8.** *Let  $\gamma \in (0, 1)$ . Then*

$$(59.15) \quad \text{Lip}_\gamma(\mathbb{R}^d) \cap \text{BC}(\mathbb{R}^d) = B_{\infty\infty}^\gamma(\mathbb{R}^d)$$

with norm equivalence.

*Proof.* We shall show that our definition of  $\text{Lip}^\gamma(\mathbb{R}^d) \cap \text{BC}(\mathbb{R}^d)$  agrees with the usual definition (the usual Lipschitz norm) when  $0 < \gamma < 1$ .

Suppose first that  $f$  is bounded and satisfies

$$(59.16) \quad \sup_{x \in \mathbb{R}^d} |f(x - y) - f(y)| \leq A|y|^\gamma$$

with  $A := \|f\|_{\text{Lip}^\gamma}$ . Then for  $j \in \mathbb{N}$

$$(2\pi)^{\frac{d}{2}} \varphi_j(D)f(x) = f * \mathcal{F}^{-1}\varphi_j(x) = \int_{\mathbb{R}^d} (f(x - y) - f(x))\mathcal{F}^{-1}\varphi_j(y) dy,$$

since  $\int_{\mathbb{R}^d} \mathcal{F}^{-1}\varphi_j(y) dy = 0$ . Thus,

$$(59.17) \quad \|\varphi_j(D)f\|_\infty \leq A \int_{\mathbb{R}^d} |y|^\gamma \cdot |\mathcal{F}^{-1}\varphi_j(y)| dy = A2^{-j\gamma}.$$

Thus,  $f \in B_{\infty\infty}^\gamma(\mathbb{R}^d)$ .

Conversely, if  $f \in B_{\infty\infty}^\gamma(\mathbb{R}^d)$ , we can write

$$(59.18) \quad f = \sum_{j=0}^{\infty} \varphi_j(D)f.$$

We see that  $\sum_{j=0}^{\infty} \varphi_j(D)f$  converges uniformly, since  $\|\varphi_j(D)f\|_{\infty} \leq 2^{-j\gamma} \|f\|_{B_{\infty}^{\gamma}}$ .

Since  $f \in L^{\infty}(\mathbb{R}^d)$ , it is trivial that

$$(59.19) \quad \sup_{x \in \mathbb{R}^d} |f(x-y) - f(x)| \lesssim |y|^{\gamma} \|f\|_{B_{\infty}^{\gamma}}$$

for  $|y| \geq 1$ . Suppose instead that  $|y| \leq 1$ . Then we have

$$(59.20) \quad f(x-y) - f(x) = \sum_{2^j \leq |y|^{-1}} [\varphi_j(D)f(x-y) - \varphi_j(D)f(x)] + \sum_{2^j > |y|^{-1}} [\varphi_j(D)f(x-y) - \varphi_j(D)f(x)].$$

Using the mean-value theorem, we see that the terms in the first sum are dominated by  $|y|2^{-j\gamma+j} \|f\|_{B_{\infty}^{\gamma}}$ , resulting the estimate

$$(59.21) \quad |y| \left( \sum_{2^j \leq |y|^{-1}} 2^{j(1-\gamma)} \|f\|_{B_{\infty}^{\gamma}} \right) \lesssim |y|^{\gamma} \|f\|_{B_{\infty}^{\gamma}},$$

since  $\gamma < 1$ . The terms in the second sum are majorized by  $2 \cdot 2^{-\gamma j} \|f\|_{B_{\infty}^{\gamma}}$  and this gives  $|y|^{\gamma} \|f\|_{B_{\infty}^{\gamma}}$  as an estimate, since  $\gamma > 0$ . Thus,

$$(59.22) \quad \sup_{x \in \mathbb{R}^d} |f(x-y) - f(x)| \lesssim |y|^{\gamma} \|f\|_{B_{\infty}^{\gamma}},$$

and the equivalence of the two definitions are established. □

**Example 59.9.** Let  $\gamma \in (0, 1)$  and  $m$ . Then we have a bilateral estimate

$$(59.23) \quad \|f\|_{C^{\gamma}} + \sum_{j=1}^d \|\partial_j^m f\|_{C^{\gamma}} \simeq \|f\|_{C^{m+\gamma}}.$$

Indeed, this estimate is just the lifting property provided we use the fact  $\mathcal{C}_{\gamma}(\mathbb{R}^d) = \text{Lip}_{\gamma}(\mathbb{R}^d) \cap \text{BC}(\mathbb{R}^d) = B_{\infty}^{\gamma}(\mathbb{R}^d)$ .

## 60. OTHER RELATED FUNCTION SPACES

### 60.1. Modulation spaces.

As an application of Theorem 55.9 and for the purpose of familiarizing ourselves with Theorem 55.9 we deal with modulation spaces.

**Definition 60.1.** Pick a compactly supported function  $\psi \in \mathcal{S}$  so that it satisfies

$$(60.1) \quad \text{supp}(\psi) \subset Q(2) = \{x \in \mathbb{R}^d : \max(|x_1|, |x_2|, \dots, |x_d|) \leq 2\}, \quad \sum_{m \in \mathbf{Z}^d} T_m \psi(x) \equiv 1.$$

Here  $T_m \psi := \psi(* - m)$ . Define

$$(60.2) \quad \|f : M^{pq}\| := \|T_m \psi(D)f\|_{\ell^q(L^p)} := \left( \sum_{m \in \mathbf{Z}^d} \|\mathcal{F}^{-1} T_m \psi * f\|_p^q \right)^{\frac{1}{q}}$$

for  $f \in \mathcal{S}'$ .

*Exercise 265.* A different choice of  $\psi$  will give us an equivalent norm. Mimic the proof of the corresponding assertion for Besov and Triebel-Lizorkin spaces.

*Exercise 266.* Use the Plancherel theorem to prove  $M^{22}(\mathbb{R}^d) = L^2(\mathbb{R}^d)$ .

We do not go into the details of this function space. However, as an application of the results in Besov spaces, we shall prove the following.

**Theorem 60.2.** *Let  $0 < p, q < \infty$ . Then  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $M^{p,q}$ .*

*Proof.* Let  $f \in M^{p,q}$ . Set

$$(60.3) \quad f^M = \sum_{\substack{m \in \mathbb{Z}^d \\ |m| \leq M}} T_m \psi(D) f.$$

Then the multiplier theorem readily gives us

$$(60.4) \quad \|f - f^M\|_{M^{p,q}} \lesssim \left( \sum_{\substack{m \in \mathbb{Z}^d \\ |m| > M-5}} \|T_m \psi(D) f\|_{p^q} \right)^{\frac{1}{q}}.$$

Therefore, assuming  $p$  and  $q$  are finite, we have  $\lim_{M \rightarrow \infty} f^M = f$  in  $M^{p,q}(\mathbb{R}^d)$ .

In view of this paragraph, we may assume  $f$  is supported on a big cube  $Q(M)$ , where  $Q(M) = \{x \in \mathbb{R}^d : \max(|x_1|, |x_2|, \dots, |x_d|) \leq M\}$ . The multiplier theorem gives us a constant  $c = c_M$ , depending only on  $M$  as well as  $p$  and  $q$ , such that

$$(60.5) \quad \|g\|_{M^{p,q}} \lesssim \|g\|_{B_{pq}^0}$$

for all  $g$  with  $\text{supp}(\mathcal{F}g) \subset Q(M+1)$ .

Since  $f \in \mathcal{S}'(\mathbb{R}^d)$  with  $f \in \text{supp}(Q(M))$ , we can approximate it with  $g \in \mathcal{S}(\mathbb{R}^d)$  in the topology of  $B_{pq}^0(\mathbb{R}^d)$ . Thus, in view of (60.5), it follows that  $f$  can be approximated by  $\mathcal{S}(\mathbb{R}^d)$ .  $\square$

*Exercise 267.* Let  $0 < p, q \leq \infty$ . Show that  $\mathcal{S}(\mathbb{R}^d) \subset M^{p,q}$ .

## 60.2. Herz spaces.

Since there is a kin connection between Besov / Triebel Lizorkin spaces and Herz spaces, it is a good chance to introduce Herz spaces. The Herz norm is similar to the Besov norm. Their difference lies in the point that the Besov norm is defined by means of the Fourier multiplier while the Herz norm involves the pointwise multiplier. Here we content ourselves in presenting the definition of the norms.

**Definition 60.3.** Fix  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  so that  $\chi_{B(1)} \leq \varphi \leq \chi_{B(2)}$ . Set

$$(60.6) \quad \varphi_0 = \varphi, \varphi_j = \varphi(2^{-j}\cdot) - \varphi(2^{-j+1}\cdot).$$

Let  $0 < p, q \leq \infty$  and  $\alpha \in \mathbb{R}$ . Then the Herz space  $K_p^{\alpha,q}$  is a function space normed by

$$(60.7) \quad \|f : K_p^{\alpha,q}\| := \|2^{j\alpha} \varphi_j \cdot f\|_{\ell^q(L^p)}.$$

We remark that the following function spaces are of much importance.

$$(60.8) \quad B^p := K_p^{-\frac{d}{p}, \infty}(\mathbb{R}^d), B_{pq} := K_p^{-\frac{d}{q} + \frac{d}{p}, \infty}(\mathbb{R}^d).$$

60.3. **Amalgam spaces.** Finally we take up the amalgam spaces.

**Definition 60.4** (Amalgam space  $(L^p, \ell^q)$ ). Let  $0 < p, q \leq \infty$  and  $s \in \mathbb{R}$ . Set  $Q_z := z + [0, 1]^d$  for  $z \in \mathbb{Z}^d$ , the translation of the unit cube. For a Lebesgue locally integrable function  $f$  define

$$(60.9) \quad \|f\|_{(L^p, \ell^q)} := \|\{\langle z \rangle^s \cdot \|f\|_{L^p(Q_z)}\}_{z \in \mathbb{Z}^d}\|_{\ell^q}.$$

The function space  $(L^p, \ell^q)$  is the set of all  $L^p(\mathbb{R}^d)$ -locally integrable functions  $f$  for which the quasi-norm  $\|f\|_{(L^p, \ell^q)}$  is finite.

*Exercise 268.* Show that  $(L^p, \ell^q)$  is a quasi-Banach space.

*Exercise 269.* Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . Then define

$$(60.10) \quad M_{\text{loc}}f(x) = \sup_{r \leq 1} \frac{1}{r^d} \int_{B(x, r)} |f(z)| dz.$$

Let  $1 < p < \infty$  and  $0 < q \leq \infty$ . Then show that

$$(60.11) \quad \|M_{\text{loc}}f\|_{L^p(\ell^q)} \lesssim \|f\|_{L^p(\ell^q)}$$

for all  $f \in L^p(\ell^q)$ .

Notes and references for Chapter 24.

Besov spaces were defined in [96].

Section 55. Theorems 55.3, 55.5, 55.9 are investigated initially by [399, 371, 47].

Theorem 55.12 was first investigated by Marcinkiewicz in the periodic set up [327]. Later Hörmander, Kree and Mihlin proved the result in the present form (See [238, 290, 348]).

Definition 55.1 and Theorem 55.3 appeared in [62] in a somehow awkward form and they were refined in [63].

Section 56. The theory of function spaces are taken up in [63, 64, 65, 66]. In this part, along with Chapter 19, we have covered the former part of [63, 64]. In [64, 65, 66] decomposition methods are taken up in great detail.

Theorem 56.4

Theorem 56.6

Theorem 56.7

Theorem 56.8

Theorem 56.10

Theorem 56.14

Theorem 56.15

Theorem 56.16

Theorem 56.17

Theorem 56.18

Theorem 56.19

Section 57. We refer to [60] for more details of Theorem 9.22.

Lemma 57.3 and Theorem 57.6 are folklore facts. However, the convincing proofs had been missing, although Lemma 57.3 and Theorem 57.6 are formulated in [49]. As late as 2006 M. Bownik and K. Ho proved them in [103].

Theorem 57.10 is also a well-known fact and few literatures such as [63] made explicit how the topology of  $\mathcal{S}'(\mathbb{R}^d)/\mathcal{P}$  should be induced.

Section 58. Theorem 58.2

Theorem 58.3

Theorem 58.4

Theorem 58.5

Theorem 58.6

Theorem 58.7

Section 59. In [61] the Laplacian is described in great detail.

Theorem 59.1

Theorem 59.2

Theorem 59.4

Theorem 59.7

Theorem 59.8 is essentially due to Taibleson [467, Theorem 4].

Section 60. The modulation space  $M_{p,q}^s$  with  $1 \leq p, q \leq \infty$  was initially defined by Feichtinger [184], which can be found in [32] nowadays. It was subsequently investigated in [186, 189, 190, 481]. Later the modulation space  $M_{p,q}^s$  was defined for  $0 < p, q \leq \infty$  in [276, 54]. In [278] the molecular decomposition was obtained.

Theorem 60.2 was taken up in [276] from a different point of view.

Weak-type Herz spaces are defined and the several boundedness of linear operators is investigated in [437]. Herz-type Hardy spaces are investigated in [206].

**Part 25. Applications to partial differential equations**

This nature of this section is an application. We apply our theory of function spaces and the theory of singular integral operators. First we treat the heat equations and then we treat the pseudo-differential operators.

61. THE HEAT SEMIGROUP

The heat equation and the heat semigroup. Having set down the multiplier property, we shall turn to the heat semigroup. Define

$$(61.1) \quad E(x, t) := \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{4t}\right).$$

**Definition 61.1.** Let  $t \geq 0$ . Define a continuous mapping  $e^{t\Delta} : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$  by

$$(61.2) \quad e^{t\Delta} f := E(\cdot, t) * f, \quad t > 0.$$

For the sake of simplicity set  $e^{0\Delta} = \text{id}_{\mathcal{S}'(\mathbb{R}^d)}$ . The family of mapping  $\{e^{t\Delta}\}_{t>0}$  is called the heat semigroup.

**Lemma 61.2.** Let  $f \in \mathcal{S}(\mathbb{R}^d)$ . We write  $u(t, x) := e^{t\Delta} f(x)$  for  $t \geq 0$  and  $x \in \mathbb{R}^d$ .

- (1)  $u$  is  $C^\infty((0, \infty) \times \mathbb{R}^d)$ .
- (2)  $u$  solves the following heat equation:

$$(61.3) \quad \partial_t u(t, x) - \Delta u(t, x) = 0, \quad t > 0, \quad x \in \mathbb{R}^d, \quad \lim_{t \downarrow 0} u(t, \cdot) = \varphi \text{ in } \mathcal{S}'(\mathbb{R}^d).$$

- (3) Let  $t, s \geq 0$ . Then  $e^{(t+s)\Delta} f = e^{t\Delta}[e^{s\Delta} f]$ .

**Lemma 61.3.** All the statements in Lemma 61.2 remain valid, if we replace  $\mathcal{S}'(\mathbb{R}^d)$  with  $\mathcal{S}(\mathbb{R}^d)$ .

*Exercise 270.* Prove Lemmas 61.2 and 61.3.

The following theorem shows us quantitative information of the smoothing effect, which is frequently used in the Navier-Stokes equations.

**Theorem 61.4.** Suppose that the parameters satisfy  $0 < p, q \leq \infty$ ,  $s \in \mathbb{R}$  and  $\alpha \geq 0$ . Then we have

$$(61.4) \quad \|e^{t\Delta} f\|_{A_{pq}^{s+\alpha}} \lesssim t^{-\frac{\alpha}{2}} \|f\|_{A_{pq}^s}$$

for  $0 \leq t \leq 1$ . Here  $c$  depends only on  $p, q, s, \alpha$ .

*Proof.* We concentrate on the case  $A = F$ , the case  $A = B$  being the same. Suppose that  $\{\varphi_j\}_{j \in \mathbb{N}}$  is a family of functions satisfying

$$(61.5) \quad \chi_{B(2)} \leq \varphi_0 \leq \chi_{B(4)}, \quad \chi_{B(4) \setminus B(2)} \leq \varphi_1 \leq \chi_{B(8) \setminus B(4)}, \quad \varphi_j = \varphi_1(2^{-j+1} \cdot) \text{ for } j \in \mathbb{N}.$$

Choose an auxiliary family  $\{\psi_j\}_{j \in \mathbb{N}}$  satisfying

$$(61.6) \quad \chi_{B(4)} \leq \psi_0 \leq \chi_{B(8)}, \quad \chi_{B(8) \setminus B(4)} \leq \psi_1 \leq \chi_{B(16) \setminus B(8)}, \quad \psi_j = \psi_1(2^{-j+1} \cdot) \text{ for } j \in \mathbb{N}.$$

Set  $\rho_0(x) = \varphi_0(x)$  and  $\rho_j(x) = \frac{\varphi_j(x)}{|2^j \cdot|^\alpha}$ . Then by Theorem 55.9 we have

$$\begin{aligned} \|e^{t\Delta} f\|_{F_{pq}^{s+\alpha}} &= \|2^{j(s+\alpha)} \varphi_j(D) e^{t\Delta} f\|_{L^p(\ell^q)} \\ &= \|2^{j(s+\alpha)} \psi_j(D) \varphi_j(D) e^{t\Delta} f\|_{L^p(\ell^q)} \\ &\lesssim t^{-\frac{\alpha}{2}} M(t) \|2^{js} \rho_j(D) f\|_{L^p(\ell^q)}, \end{aligned}$$

where

$$(61.7) \quad M(t) := \frac{1}{t^{\frac{d}{2}}} \left( \|\psi_0(\cdot)e^{-4^k t |\cdot|^2}\|_{H_2^M} + \sup_{k \in \mathbb{N}} \| |2^k t^{\frac{1}{2}} \cdot |^\alpha \psi_1(\cdot)e^{-4^k t |\cdot|^2}\|_{H_2^M} \right).$$

By change of variables, we see that  $\sup_{0 \leq t \leq 1} M(t) < \infty$ . Therefore the proof is complete.  $\square$

Semigroup property. In this paragraph we shall prove that the heat semigroup is such an example. Until the end of this section we assume  $1 \leq p, q \leq \infty$  and  $s \in \mathbb{R}$ . Namely, for the sake of simplicity we limit ourselves to the case when  $A_{pq}^s(\mathbb{R}^d)$  is a Banach space.

**Theorem 61.5.** *Then  $\{e^{t\Delta}\}_{t \geq 0}$  is a continuous semigroup on  $A_{pq}^s(\mathbb{R}^d)$ .*

*Proof.* By Theorem 61.4,  $\{e^{t\Delta}\}_{t \geq 0}$  is a bounded family of operators. Going through a similar argument, we have

$$(61.8) \quad \|e^{t\Delta} - e^{t'\Delta} : A_{pq}^s\| \lesssim |t - t'|,$$

whenever  $0 \leq t, t' \leq 1$ . As we have seen, for all  $f \in \mathcal{S}'(\mathbb{R}^d)$  we have

$$(61.9) \quad e^{t\Delta}(e^{s\Delta}f) = e^{(t+s)\Delta}f.$$

As a result it follows that  $\{e^{t\Delta}\}_{t \geq 0}$  is a continuous semigroup on  $A_{pq}^s(\mathbb{R}^d)$ .  $\square$

**Theorem 61.6.** *The generator of a semigroup  $e^{t\Delta} : A_{pq}^s(\mathbb{R}^d) \rightarrow A_{pq}^s(\mathbb{R}^d)$  is given by*

$$(61.10) \quad D(A) = A_{pq}^{s+2}(\mathbb{R}^d), \quad Af = \Delta f.$$

*Proof.* Let  $f \in D(A)$ . Since  $\lim_{t \downarrow 0} \frac{e^{t\Delta}f - f}{t} = \Delta f$  in  $\mathcal{S}'(\mathbb{R}^d)$ , if the limit  $\lim_{t \downarrow 0} \frac{e^{t\Delta}f - f}{t}$  must coincide  $\Delta f$ . Therefore,  $f \in D(A)$  is equivalent to  $f \in A_{pq}^s$  and  $\Delta f \in A_{pq}^s$ . However, this is equivalent to  $f \in A_{pq}^{s+2}$ .  $\square$

**Theorem 61.7.** *The spectrum of  $\Delta$  on  $A_{pq}^s$  is  $(-\infty, 0]$ . Let  $0 < \theta < \pi$ . Then*

$$(61.11) \quad \|(\Delta - z)^{-1} : B(A_{pq}^s)\| \leq \frac{M_\theta}{|z|}$$

for all  $z \in \mathbb{C}$  with  $z \neq 0$ ,  $\arg(z) > \pi - \theta$ .

*Proof.* Suppose that  $z \in \mathbb{C} \setminus (-\infty, 0]$  Define  $M_z(\xi) = \frac{1}{|\xi|^2 + z}$ . Then  $M_z$  is smooth with all bounded partial derivatives. Therefore  $M_z(D) = (-\Delta + z)^{-1}$ . Thus,  $\mathbb{C} \setminus (-\infty, 0] \subset \rho(\Delta)$ .

We identify the spectrum Since  $A_{pq}^{s+2}$  is a proper subset of  $A_{pq}^s$ , it cannot be bounded. Therefore, the spectrum is an unbounded set of  $(-\infty, 0]$ . Let  $I_t : f \in \mathcal{S}'(\mathbb{R}^d) \mapsto f(t \cdot) \in \mathcal{S}'(\mathbb{R}^d)$ . Then  $I_t^{-1}(\Delta - z)I_t = (t^2\Delta - z)$ . Therefore, we conclude  $(-\infty, 0)$  is contained in the spectrum. Since the spectrum is closed in  $\mathbb{C}$ , we conclude  $\sigma(\Delta) = (-\infty, 0]$ .

The estimate of resolvent This can be achieved by using Theorem 55.9 as usual.  $\square$

*Exercise 271.* Supply the details of the estimate of the resolvent above.

*Exercise 272.* Let  $0 < p, q \leq \infty$  and  $s \in \mathbb{R}$ . Then,

$$(61.12) \quad \|e^{t\Delta}f - f\|_{A_{pq}^s} \lesssim_\alpha \|f\|_{A_{pq}^{s+\alpha}}$$

for all  $f \in A_{pq}^{s+\alpha}(\mathbb{R}^d)$  and  $0 < t \leq 1$ .



Summarizing the above observations, we obtain;

**Theorem 61.8.** *Let  $1 \leq p, q \leq \infty$  and  $s \in \mathbb{R}$ . The Laplacian  $\Delta$  is a sectorial operator on  $A_{pq}^s(\mathbb{R}^d)$ , whose domain is  $A_{pq}^{s+2}(\mathbb{R}^d)$ . Furthermore it generates a continuous semigroup  $e^{t\Delta}$ .*

It is still important to state what we have learnt for  $L^p(\mathbb{R}^d) \simeq F_{p2}^0(\mathbb{R}^d)$  with  $1 < p < \infty$ . We dare repeat the statement.

**Theorem 61.9.** *Let  $1 < p < \infty$ .*

- (1) *The Laplacian  $\Delta$  is a sectorial operator whose domain is  $W^{2,p}(\mathbb{R}^d)$ .*
- (2) *The heat group is a continuous semigroup whose generator is the Laplacian  $\Delta$ .*

## 62. PSEUDO-DIFFERENTIAL OPERATORS

Before giving precise definitions, we shall make some heuristic observations.

### 62.1. Some heuristics.

Suppose that  $f \in \mathcal{S}(\mathbb{R}^d)$ , then we have

$$\frac{\partial}{\partial x_j} f(x) = i \cdot \int_{\mathbb{R}^d} \xi_j \mathcal{F}f(\xi) e^{ix \cdot \xi} d\xi$$

and

$$x_j f(x) = \int_{\mathbb{R}^d} x_j \mathcal{F}f(\xi) e^{ix \cdot \xi} d\xi.$$

Pseudo differential operator is a mixture of these operator of the form  $\int_{\mathbb{R}^d} a(x, \xi) \mathcal{F}f(\xi) e^{ix \cdot \xi} d\xi$ , where  $a(x, \xi)$  is an appropriately nice function.

### 62.2. Pseudo-differential operators on $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ .

Pseudo-differential operator on  $\mathcal{S}(\mathbb{R}^d)$ . To begin with, let us see how the pseudo-differential operators are defined on the nicest function space  $\mathcal{S}(\mathbb{R}^d)$ .

**Definition 62.1.** Let  $a : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{C}$ . Then define

$$(62.1) \quad T_a f(x) := \int_{\mathbb{R}^d} a(x, \xi) \mathcal{F}f(\xi) e^{ix \cdot \xi} d\xi.$$

**Definition 62.2.** Let  $0 \leq \rho, \delta \leq 1$  and  $m \in \mathbb{R}$ . Define the set of the functions  $S_{\rho, \delta}^m$  by

$$S^m := \left\{ a \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{C}) : \| |A| \|_{S_{\rho, \delta}^m, \alpha, \beta} < \infty \text{ for all } \alpha, \beta \in \mathbb{N}_0^d \right\},$$

where

$$(62.2) \quad \| |A| \|_{S_{\rho, \delta}^m, \alpha, \beta} := \sup_{x, \xi \in \mathbb{R}^d} \frac{|\partial_x^\beta \partial_\xi^\alpha a(x, \xi)|}{(1 + |\xi|)^{m - \delta|\alpha| + \rho|\beta|}}$$

**Example 62.3.** Let  $j = 1, 2, \dots, d$ . Then  $a(x, \xi) = \xi_j$  is an example of  $S^1$ -symbol.

*Exercise 273.* Show that  $a(x, \xi) := \langle \xi \rangle^m \in S^m$ .

**Proposition 62.4.** *Let  $a \in S_{\rho, \delta}^m$  and  $f \in \mathcal{S}(\mathbb{R}^d)$ . Then we have  $T_a f \in \mathcal{S}(\mathbb{R}^d)$ .*

*Proof.* Firstly we note that the integral converge absolutely since  $f \in \mathcal{S}(\mathbb{R}^d)$ .

Next we shall prove that  $T_a f \in \mathcal{S}(\mathbb{R}^d)$ . Let  $L_\xi$  defined by

$$(62.3) \quad L_\xi := \frac{1}{1 + |\xi|^2} (I - \Delta_\xi).$$

Then it follows that  $L_\xi e^{ix \cdot \xi} = e^{ix \cdot \xi}$ . Denote  $(L_\xi)^N$  is  $N$ -times composition of  $L_\xi$ . We intend to use integration by parts.

Using this observation we estimate  $\partial^\alpha T_a f$ . By integration by parts we have

$$T_a f(x) = \int_{\mathbb{R}^d} a(x, \xi) \mathcal{F}f(\xi) (L_\xi)^N e^{ix \cdot \xi} d\xi = \int_{\mathbb{R}^d} (L_\xi)^N (a(x, \xi) \mathcal{F}f(\xi)) e^{ix \cdot \xi} d\xi.$$

Notice that by Leibnitz's formula we have

$$(62.4) \quad |\partial_x^\alpha L_\xi^N (a(x, \xi) \mathcal{F}f(\xi))| \leq C_{M,N,\alpha} (1 + |\xi|)^{-M}.$$

Thus we have  $|x^\alpha D^\beta T_a f(x)| \leq C < \infty$ , which implies  $T_a f \in \mathcal{S}(\mathbb{R}^d)$ . □

**Corollary 62.5.** *Suppose that  $\{a_\varepsilon\}_{\varepsilon \in [0,1]} \subset S^m$  uniformly, that is,*

$$(62.5) \quad |\partial_x^\beta \partial_\xi^\alpha a_\varepsilon(x, \xi)| \lesssim_{\alpha,\beta} (1 + |\xi|)^{m-|\alpha|}$$

for all  $\varepsilon \in [0, 1]$ . Assume further that for any fixed  $x, \xi \in \mathbb{R}^d$

$$(62.6) \quad \partial_x^\beta \partial_\xi^\alpha a_\varepsilon(x, \xi) \rightarrow \partial_x^\beta \partial_\xi^\alpha a_0(x, \xi)$$

as  $\varepsilon \rightarrow 0$ . Then we have  $T_{a_\varepsilon} f \rightarrow T_{a_0} f$  in  $\mathcal{S}(\mathbb{R}^d)$ .

*Proof.* For the proof we use the calculation of the previous proposition. By assumption we can use the dominated convergence theorem. □

**Example 62.6.** Let  $\gamma \in C^\infty(\mathbb{R}^d)$  be such that

$$(62.7) \quad \chi_{Q(1)} \leq \gamma \leq \chi_{Q(2)}.$$

We set

$$(62.8) \quad a_\varepsilon(x, \xi) := a(x, \xi) \gamma(\varepsilon x) \gamma(\varepsilon \xi) \quad (0 \leq \varepsilon \leq 1).$$

Then  $\{a_\varepsilon\}_{\varepsilon \in [0,1]}$  satisfies the condition of the previous corollary.

**Lemma 62.7.** *We have*

$$(62.9) \quad T_a f(x) = \lim_{\varepsilon \rightarrow 0} \iint a_\varepsilon(x, \xi) f(y) e^{i(x-y) \cdot \xi} d\xi dy,$$

where the convergence takes place in  $\mathcal{S}(\mathbb{R}^d)$ .

*Proof.* Use the previous lemma. Then we have

$$\begin{aligned} T_a f(x) &= \lim_{\varepsilon \rightarrow 0} T_{a_\varepsilon} f(x) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} a_\varepsilon(x, \xi) \mathcal{F}f(\xi) e^{ix \cdot \xi} d\xi \\ &= \lim_{\varepsilon \rightarrow 0} \iint a_\varepsilon(x, \xi) f(y) e^{i(x-y) \cdot \xi} d\xi dy. \end{aligned}$$

This is the desired result. □

**Lemma 62.8.** *For all  $f, g \in \mathcal{S}(\mathbb{R}^d)$ , we have*

$$(62.10) \quad \int_{\mathbb{R}^d} T_a f(x) g(x) dx = \int_{\mathbb{R}^d} f(y) T_a^* g(y) dy.$$

(62.10) shows some duality formula holds.

*Proof.* In fact we have

$$\begin{aligned} \int_{\mathbb{R}^d} T_a f(x) \cdot g(x) dx &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} \left( \iint a_\varepsilon(x, \xi) g(x) f(y) e^{i(x-y)\xi} d\xi dy \right) dx \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} \left( \iint a_\varepsilon(y, \xi) g(y) f(x) e^{i(y-x)\xi} d\xi dx \right) dy \\ &= \int_{\mathbb{R}^d} f(x) \cdot T_a^* g(x) dx. \end{aligned}$$

Since we are truncating the integral, the Fubini theorem is applicable. □

The theory of  $\mathcal{S}(\mathbb{R}^d)$  was all stated. In the next section we develop the  $L^p(\mathbb{R}^d)$ -theory. The generalized CZ-integral is a key tool.

### 63. $L^p(\mathbb{R}^d)$ -BOUNDEDNESS OF PSEUDO-DIFFERENTIAL OPERATORS

#### 63.1. $L^2(\mathbb{R}^d)$ -boundedness.

The following result can be obtained by the Fourier transform. This result is applied to the  $\varphi$ -transform which dates back to [195, 196]. Nowadays the  $\varphi$ -transform is widely used in the theory of function spaces.

Sampling theorem.

**Theorem 63.1.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$  with  $\text{supp}(\mathcal{F}) \subset Q(1)$ . Then we have*

$$(63.1) \quad f = \sum_{l \in \mathbb{Z}^d} f(l) \mathcal{F} \kappa(* - l).$$

Here  $\kappa \in \mathcal{S}(\mathbb{R}^d)$  is an even cut-off function such that  $\chi_{Q(1)} \leq \kappa \leq \chi_{Q(2)}$ .

*Proof.* Pick a test function  $\tau \in \mathcal{S}(\mathbb{R}^d)$ . Furthermore take another cut-off function that equals 1 on  $Q(2)$  and supported on  $Q(3)$ . Then we have

$$(63.2) \quad \langle f, \tau \rangle = \langle \mathcal{F} f, \mathcal{F}^{-1} \tau \rangle = \langle \kappa \cdot \mathcal{F} f, \eta \cdot \kappa \cdot \mathcal{F}^{-1} \tau \rangle.$$

Taking into account the support condition, we have

$$\eta \cdot \kappa \cdot \mathcal{F}^{-1} \tau = \eta \cdot \sum_{j \in \mathbb{Z}^d} \kappa(* - 2\pi j) \cdot \mathcal{F}^{-1} \tau(* - 2\pi j).$$

Therefore, if we expand  $\sum_{j \in \mathbb{Z}^d} \kappa(* - 2\pi j) \cdot \mathcal{F}^{-1} \tau(* - 2\pi j)$ , to the Fourier series, then we have

$$(63.3) \quad \sum_{j \in \mathbb{Z}^d} \kappa(x - 2\pi j) \cdot \mathcal{F}^{-1} \tau(x - 2\pi j) = \sum_{l \in \mathbb{Z}^d} a_l \cdot e^{il \cdot x},$$

where  $a_l = \frac{1}{(2\pi)^d} \int_{[0, 2\pi]^d} \left( \sum_{j \in \mathbb{Z}^d} \kappa(y - 2\pi j) \cdot \mathcal{F}^{-1} \tau(y - 2\pi j) \right) e^{-il \cdot y} dy$ . By changing the variable we obtain

$$(63.4) \quad a_l = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \kappa(y) \mathcal{F}^{-1} \tau(y) e^{-il \cdot y} dy = \frac{1}{(2\pi)^{\frac{d}{2}}} \mathcal{F}(\kappa \cdot \mathcal{F}^{-1} \tau)(l) = \frac{1}{(2\pi)^{\frac{d}{2}}} \mathcal{F} \kappa * \tau(l).$$

Since the convergence in (63.3) takes place in  $\mathcal{S}(\mathbb{R}^d)$ , we have

$$(63.5) \quad \langle f, \tau \rangle = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{l \in \mathbb{Z}^d} \langle \mathcal{F}f, \eta e^{il \cdot * } \rangle \mathcal{F}\kappa * \tau(l) = \sum_{l \in \mathbb{Z}^d} f(l) \mathcal{F}\kappa * \tau(l) = \sum_{l \in \mathbb{Z}^d} f(l) \int_{\mathbb{R}^d} \mathcal{F}\kappa(l - y) \tau(y) dy.$$

Therefore, it follows that

$$(63.6) \quad f = \sum_{l \in \mathbb{Z}^d} f(l) \mathcal{F}\kappa(* - l),$$

if we take into account that we assume  $\kappa$  is even. □

As an application of this theorem, we obtain the  $\varphi$ -transform obtained in [195, 196].

**Corollary 63.2.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then one has*

$$(63.7) \quad f = \sum_{m \in \mathbb{Z}^d} \left( \sum_{l \in \mathbb{Z}^d} e^{i(*-l) \cdot m} \varphi_m(D) f(l) \mathcal{F}\kappa(* - l) \right).$$

*Proof.* First, we have

$$(63.8) \quad f = \sum_{m \in \mathbb{Z}^d} \varphi_m(D) f.$$

Since  $\mathcal{F}(e^{-im \cdot * } \varphi_m(D) f) = \varphi_m(D) f(* + m)$ , it follows that  $\text{supp}(\mathcal{F}(e^{-im \cdot * } \varphi_m(D) f)) \subset Q(2)$ . Therefore, we are in the position of applying Theorem 63.1 to obtain

$$(63.9) \quad e^{-im \cdot * } \varphi_m(D) f = \sum_{l \in \mathbb{Z}^d} e^{-im \cdot l} f(l) \mathcal{F}\kappa(* - l).$$

If we arrange the above formula, then we obtain

$$(63.10) \quad \varphi_m(D) f = \sum_{l \in \mathbb{Z}^d} e^{im \cdot (*-l)} \varphi_m(D) f(l) \mathcal{F}\kappa(* - l).$$

Inserting this formula to (63.8), then we obtain (63.7). □

**Lemma 63.3.** *We have*

$$(63.11) \quad \left( \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D) f(l)|^2 \right)^{\frac{1}{2}} \lesssim \|f\|_2.$$

*Proof.* It is convenient to write the left-hand side as

$$(63.12) \quad \left( \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D) f(l)|^2 \right)^{\frac{1}{2}} = \left( \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z}^d} \left( \sum_{l \in \mathbb{Z}^d} |\varphi_m(D) f(l)| \chi_{Q_l} \right)^2 dx \right)^{\frac{1}{2}}.$$

Now we have

$$(63.13) \quad \sum_{l \in \mathbb{Z}^d} |\varphi_m(D) f(l)| \chi_{Q_l} \lesssim M[\varphi_m(D) f](x).$$

If we insert (63.13) to (63.12) and recall that  $M$  is  $L^2(\mathbb{R}^d)$ -bounded, then we obtain

$$\begin{aligned} \left( \sum_{m,l \in \mathbb{Z}^d} |\varphi_m(D)f(l)|^2 \right)^{\frac{1}{2}} &\lesssim \left( \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z}^d} M[\varphi_m(D)f](x)^2 dx \right)^{\frac{1}{2}} \\ &\lesssim \left( \int_{\mathbb{R}^d} \sum_{m \in \mathbb{Z}^d} |\varphi_m(D)f(x)|^2 dx \right)^{\frac{1}{2}} \\ &= \|f\|_2. \end{aligned}$$

This is the desired result. □

**Lemma 63.4.** *Let  $\kappa$  be an even smooth function satisfying  $\chi_{Q(1)} \leq \kappa \leq \chi_{Q(2)}$ . For  $m, l \in \mathbb{Z}^d$ , we set  $\psi_{ml}(x) := e^{ix \cdot m} \mathcal{F}\kappa(x - l)$ . Let  $a \in S_{00}^0$ . Then we have, for all  $N$ ,*

$$(63.14) \quad |\partial^\alpha (e^{-ix \cdot m} a(x, D)\psi_{ml}(x))| \lesssim_{\alpha, N} \langle x - l \rangle^{-N}$$

$$(63.15) \quad |\varphi_k(D)[a(x, D)\psi_{ml}](x)| \lesssim_N \langle m - k \rangle^{-N} \langle x - l \rangle^{-N}.$$

*Proof.* We calculate  $a(x, D)\psi_{ml}(x)$ . Writing it out in full, we obtain

$$\begin{aligned} a(x, D)\psi_{ml}(x) &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} a(x, \xi) e^{ix \cdot \xi} \mathcal{F}\psi_{ml}(\xi) d\xi \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} a(x, \xi) e^{ix \cdot \xi + il \cdot (m - \xi)} \kappa(\xi - m) d\xi \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} a(x, \xi + m) e^{ix \cdot (\xi + m) - il \cdot \xi} \kappa(\xi) d\xi \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} e^{ix \cdot m} \int_{\mathbb{R}^d} a(x, \xi + m) e^{i(x-l) \cdot \xi} \kappa(\xi) d\xi. \end{aligned}$$

Therefore, it follows that

$$(63.16) \quad e^{-ix \cdot m} a(x, D)\psi_{ml}(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} a(x, \xi + m) e^{i(x-l) \cdot \xi} \kappa(\xi) d\xi.$$

Now we use  $(1 - \Delta_\xi)^N e^{i(x-l) \cdot \xi} = \langle x - l \rangle^{2N} e^{i(x-l) \cdot \xi}$ . Inserting this and carrying out integration by parts repeatedly, then we have

$$(63.17) \quad e^{-ix \cdot m} a(x, D)\psi_{ml}(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \langle x - l \rangle^{2N}} \int_{\mathbb{R}^d} (1 - \Delta_\xi)^N (a(x, \xi + m) \kappa(\xi)) e^{i(x-l) \cdot \xi} d\xi.$$

Therefore, we obtain

$$(63.18) \quad |e^{-ix \cdot m} a(x, D)\psi_{ml}(x)| \lesssim \langle x - l \rangle^{-2N},$$

since  $a \in S_{00}^0$  and  $\kappa \in C_c^\infty(\mathbb{R}^d)$ .

A similar argument works for any partial derivative of  $e^{-ix \cdot m} a(x, D)\psi_{ml}(x)$ , (63.14) is therefore proved.

Now we turn to (63.15). To do this, we set  $\eta_{ml}(x) := e^{-ix \cdot m} a(x, D)\psi_{ml}(x)$ . Then, it written out in full, we have

$$(63.19) \quad \varphi_k(D)[a(x, D)\psi_{ml}](x) = e^{ix \cdot k} \int_{\mathbb{R}^d} \mathcal{F}^{-1}\varphi(x - y)\eta_{ml}(y) e^{iy(m-k)} dy.$$

Now we invoke  $(1 - \Delta_y)^N e^{iy(m-k)} = \langle m - k \rangle^{2N} e^{iy(m-k)}$ . Therefore

$$(63.20) \quad \varphi_k(D)[a(x, D)\psi_{ml}](x) = e^{ix \cdot k} \langle m - k \rangle^{-2N} \int_{\mathbb{R}^d} (1 - \Delta_y)^N (\mathcal{F}^{-1}\varphi(x - y)\eta_{ml}(y)) e^{iy(m-k)} dy.$$

Finally we invoke the differential inequality (63.14) to obtain

$$(63.21) \quad |(1 - \Delta_y)^N (\mathcal{F}^{-1} \varphi(x - y) \eta_{ml}(y))| \lesssim \langle x - y \rangle^{-2N} \langle y - l \rangle^{-2N} \lesssim \langle x - l \rangle^{-N} \langle y - l \rangle^{-N}.$$

Inserting this inequality, we obtain (63.15). □

**Theorem 63.5** (Calderón-Vaillancourt). *Let  $a \in S_{00}^0$ . Then  $a(x, D)$  is bounded on  $L^2(\mathbb{R}^d)$ .*

*Proof.* By (63.7) and (63.15), we have

$$(63.22) \quad |\varphi_k(D)[a(x, D)f](x)| \lesssim \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D)f(l)| \cdot \langle m - k \rangle^{-2N} \langle x - l \rangle^{-2N},$$

where  $N$  can be taken as large as we wish. Therefore, by the Hölder inequality, we obtain

$$\begin{aligned} & |\varphi_k(D)[a(x, D)f](x)|^2 \\ & \lesssim \left( \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D)f(l)| \cdot \langle m - k \rangle^{-2N} \langle x - l \rangle^{-2N} \right)^2 \\ & \lesssim \left( \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D)f(l)|^2 \cdot \langle m - k \rangle^{-N} \langle x - l \rangle^{-N} \right) \cdot \left( \sum_{m, l \in \mathbb{Z}^d} \langle m - k \rangle^{-N} \langle x - l \rangle^{-N} \right) \\ & \lesssim \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D)f(l)|^2 \cdot \langle m - k \rangle^{-N} \langle x - l \rangle^{-N}. \end{aligned}$$

Therefore, summing this over  $k \in \mathbb{Z}^d$ , integrating this over  $\mathbb{R}^d$  and then invoking Lemma 63.3, we obtain

$$(63.23) \quad \|a(X, D)f : M^{22}\|^2 \lesssim \sum_{m, l \in \mathbb{Z}^d} |\varphi_m(D)f(l)|^2 \lesssim \|f\|_2^2.$$

This is the desired result. □

**63.2.  $L^p(\mathbb{R}^d)$ -boundedness.**

Dyadic decomposition. In this section we prepare for the proof of  $L^2(\mathbb{R}^d)$  boundedness of pseudo-differential operator with  $S^0(\mathbb{R}^d)$ -symbol. The definitions below we can make a brief look of Littlewood-Paley theory, which we develop in the last part. We introduce functions as follows, whose definitions are valid only in this part.

*Notation .* (1)  $\eta \in C^\infty(\mathbb{R}^d)$  satisfies

$$(63.24) \quad \chi_{Q(1)} \leq \eta \leq \chi_{Q(2)}.$$

(2) We set  $\delta(\xi) = \eta(\xi) - \eta(2\xi)$ .

(3) Let  $a \in S^0$ . Then define

$$(63.25) \quad a_0(x, \xi) := a(x, \xi)\eta(\xi), \quad a_j(x, \xi) := a(x, \xi)\delta(2^{-j}\xi) \quad (j = 1, 2, \dots).$$

(4) We set  $k_j(x, z) := \int a_j(x, \xi)e^{2\pi i \xi z} d\xi$ .

**Proposition 63.6.** *Let  $a \in S^m$ . Then we have*

$$(63.26) \quad a_j(x, D)f(x) = \int k_j(x, z)f(x - z)dz.$$

*Proof.* By change of variables we have

$$\text{R.H.S.} = \iint a_j(x, \xi)f(x - z)e^{2\pi i \xi z} dz d\xi = \iint a_j(x, \xi)f(y)e^{2\pi i \xi(x - y)} dy d\xi = \text{L.H.S.}$$

Thus the proof is complete. □

**Lemma 63.7.** *Suppose that  $a \in S^m$ . Then the following differential inequality*

$$(63.27) \quad |\partial_x^\beta \partial_z^\alpha k_j(x, z)| \lesssim_{M, \alpha, \beta} |z|^{-M} 2^{j(n+m-M+|\alpha|)}$$

*holds for all  $\alpha, \beta \in \mathbb{N}_0^d$  and  $M \in \mathbb{N}_0$ , where the bound is independent of  $j \in \mathbb{N}_0$ .*

By definition we have

$$(63.28) \quad (-iz)^\gamma \partial_x^\beta \partial_z^\alpha k_j(x, z) = \int_{\mathbb{R}^d} \partial_\xi^\gamma [(i\xi)^\alpha \partial_x^\beta a_j(x, \xi)] e^{ix\xi} d\xi.$$

Now we have on  $\text{supp}(f) \subset B(2^{j+2}) \setminus B(2^{j-2})$  that

$$(63.29) \quad \left| \partial_\xi^\gamma [(2\pi i \xi)^\alpha \partial_x^\beta a_j(x, \xi)] \right| = \left| \sum_{\delta \leq \gamma} C_{\gamma, \delta} \partial_\xi^\delta \xi^\alpha \partial_\xi^{\gamma-\delta} a_j(x, \xi) \right| \leq C 2^{j(|\alpha| - |\gamma| + m)}.$$

Thus our desired result follows.

**Proposition 63.8.** *For all  $x$  the series  $k(x, *) := \sum_{j=0}^\infty k_j(x, *)$  converges in  $\mathcal{S}'(\mathbb{R}^d)$ .*

*Proof.* Note that  $k_j(x, \xi) = \mathcal{F}_\xi^{-1}(a_j(x, *))$ . We have  $\sum_{j=0}^\infty a_j(x, *) = a(x, *)$  in  $\mathcal{S}'(\mathbb{R}^d)$ . Thus, this proposition follows by the continuity of the Fourier transform. □

**Corollary 63.9.** *One has*

$$(63.30) \quad T_a f(x) = \langle k(x, *), f(x - *) \rangle$$

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

*Proof.* We have only to prove

$$(63.31) \quad a_j(x, D)f(x) = \int k_j(x, z), f(x - z) dz.$$

For this purpose we shall calculate the right-hand side.

$$\text{R.H.S.} = \int_{\mathbb{R}^d} k_j(x, z), f(x - z) dz = \iint_{\mathbb{R}^d \times \mathbb{R}^d} a(x, y) e^{-izy} f(x - z) dz dy$$

Using the Fubini theorem and changing variables, we obtain

$$\text{R.H.S.} = \iint_{\mathbb{R}^d \times \mathbb{R}^d} a(x, y) e^{-i(w-x)y} f(w) dy dw = a_j(x, D)f(x) = \text{L.H.S.},$$

which is the desired result. □

Having proved the differential inequality (63.27) and the representation formula (63.30), we are now in the position of using the CZ-theory.

**Theorem 63.10.** *Suppose that  $a \in S^0$  and  $1 < p < \infty$ . Then the operator  $a(x, D)$  is an  $L^p(\mathbb{R}^d)$ -bounded operator if it restricts to  $L^p(\mathbb{R}^d)$ .*

*Proof.* The condition on the kernel is cleared due to the differential inequality (63.27). Therefore we have only to apply the CZ-theory. □

Notes and references for Chapter 25.

Section 61. Theorem 61.4

Theorem 61.5

Theorem 61.6

Theorem 61.7

Theorem 61.8

Theorem 61.9

Section 62. There is a huge amount of works on pseudo-differential operators. Let us describe some of them.

We can list the following papers [279, 239] as the pioneering works of the theory of pseudo-differential operators.

In the paper [240] the class  $S_{\rho,\delta}^m$  was introduced.

Fourier integral operators are taken up initially [241]. We can list [93, 241, 339].

Weyl pseudo-differential operators

The operator

$$(63.32) \quad a^w(x, D)u(x) = (2\pi)^{-n} \iint a\left(\frac{x+y}{2}, \xi\right) \exp\{i\langle x-y, \xi \rangle\} u(y) dy d\xi$$

is called a Weyl pseudodifferential operator. We refer to [242] for more information.

We refer to [181] as well.

Applications can be found in [94, 170, 171, 253, 376, 440].

Theorem 63.1 is a breakthrough in the theory of atomic decomposition, which appeared initially in [195]. M. Frazier and B. Jawerth utilized Theorem 63.1 for Triebel-Lizorkin spaces in the celebrated paper [196].

Another quantitative approach of the  $L^2(\mathbb{R}^d)$ -boundedness can be found in the work by Coifman and Meyer [143].

Calderón and Vaillancourt proved Theorem 63.5 (see [115]). Tachizawa extended Theorem 63.5 to the modulation spaces  $M^{p,q}$  with  $1 \leq p, q \leq \infty$ , who showed that operators generated by  $S_{00}^0$  is bounded in the modulation space  $M_{pq}^0$  for  $1 \leq p, q \leq \infty$  (see [466]). A. Miyachi investigated pseudo-differential operators with symbol in  $S_{00}^0$ .

In [441] Sjöstrand established that pseudo-differential operators with symbols in  $M_{\infty,1}^0(\mathbb{R}^d \times \mathbb{R}^d)$  are  $L^2(\mathbb{R}^d)$ -bounded. It is not so hard to prove that

$$(63.33) \quad S_{00}^0 \hookrightarrow M_{\infty,1}^0(\mathbb{R}^d \times \mathbb{R}^d)$$

and the inclusion is, of course, proper. Hence, Sjöstrand's result extends that of Calderón and Vaillancourt.

For the result of commutators generated by Lipschitz functions and pseudo-differential operators, we refer to the works by Coifman, Marshall and many other researchers.

In particular, Kobayashi, Tomita and Sugimoto gave an elegant proof.



Section 63. We refer to [143] for more details in this section.

Theorem 63.10

**Part 26. Supplemental facts  
on measure theory**

**Part 27. Supplemental facts on measure theory**

OVERVIEW

In this section we supplement measure theory. First, we consider the Haar measures and then the product of probability measures.

64. SUPPLEMENTAL OF THE CONSTRUCTION OF MEASURES

64.1. Topological groups and Haar measures.

Locally compact groups occur in many branches of mathematics. Their study falls into two cases: connected groups, which occur as automorphisms of smooth structures such as spheres, and the totally disconnected groups, which occur as automorphisms of discrete structures such as trees. Here we are concerned with the former.

By using the notion of contents, which play a fundamental role in constructing measures on topological spaces, we shall construct a symmetric measure on Lie groups. Let us begin with the definition.

**Definition 64.1** (topological group, Lie group). Let  $G$  be a group, that is,  $G$  is equipped with an operation such that:

- (a) there exists  $e \in G$  such that  $e \cdot g = g \cdot e = e$  for all  $g \in G$ ,
- (b) for any  $g \in G$ , there exists  $h \in G$  such that  $g \cdot h = h \cdot g = e$ ,
- (c) for all  $g, h, k \in G$ ,  $g \cdot (h \cdot k) = (g \cdot h) \cdot k$ .

Below one writes  $g \cdot (h \cdot k) = ghk$  for  $g, h, k \in G$ .

- (1) A group  $G$  is said to be a topological group, if  $G$  comes with a topology such that the mapping  $(a, b) \in G \times G \rightarrow ab^{-1} \in G$  is continuous.
- (2) A locally compact group is a topological group which is locally compact.
- (3) A topological group  $G$  is said to be a Lie group, if  $G$  carries the structure of a  $C^\infty$ -manifold such that the mapping  $(a, b) \in G \times G \rightarrow ab^{-1} \in G$  is smooth.

In this book it is tacitly assumed that all manifolds are separable Hausdorff spaces.

*Exercise 274.* Show that  $\mathbb{T}$  is a Lie group.

**Example 64.2.** Lie groups are typical examples of topological groups. In particular we remark that  $O(d), SO(d)$  are compact topological groups that we encountered in this book.

Actually, in this book, we have encountered many topological groups.

**Example 64.3.** Let  $V$  be a topological vector space. If we disregard the scalar multiplication operation, then  $V$  is a topological group.

There is a curious example of topological groups.

**Example 64.4.** Let  $X$  be a locally compact Hausdorff space. Denote by  $X^X$  the set of all continuous functions from  $X$  to itself. The compact open topology of  $X^X$  is a topology generated by the following family of sets:

$$(64.1) \quad \mathcal{O}(K;U) = \{f \in X^X : f(K) \subset U\}$$

Define by  $\text{Homeo}(X)$  as the set of all homeomorphisms whose topology is induced by  $C(X)$ . Then  $\text{Homeo}(X)$  is a topological group.

We denote by  $\mathcal{B}$  the  $\sigma$ -algebra generated by open sets in  $X$ .

**Theorem 64.5.** *Suppose that  $G$  is a locally compact topological group. Then there exists a Radon measure  $\mu$  so that, for every  $g \in G$  and for every  $E \in \mathcal{B}$ ,*

$$(64.2) \quad \mu(gE) = \mu(E)$$

or equivalently,

$$(64.3) \quad \int_G f(gx) d\mu(x) = \int_G f(x) d\mu(x)$$

for all  $g \in G$  and  $f \in L^1(G)$ . The measure  $\mu$  is unique in the following sense: Fix an open set such that  $U_0^{-1} = U_0$  and that  $U_0$  is relatively-compact. Then  $\mu$  is determined by  $\mu(U_0)$ .

*Proof.* Recall that  $\mathcal{O}$  is defined to be the set of all open sets in  $G$ . Denote by  $\mathcal{K}$  the set of all compact sets in  $K$  and write  $\mathcal{N} := \{U \in \mathcal{O} : e_G \in U\}$  for the neighbourhood system of  $e_G \in G$ . Take  $A \in \mathcal{K}$  with non-empty interior.

Given  $U \in \mathcal{N}$  and  $K \in \mathcal{K}$ , we define

$$(64.4) \quad [K : U] := \min \left\{ k \in \mathbb{N} : \text{there exist } g_1, g_2, \dots, g_k \in G \text{ such that } K \subset \bigcup_{j=1}^k g_j U \right\}.$$

We also denote

$$(64.5) \quad [K : A] := \min \left\{ k \in \mathbb{N} : \text{there exist } g_1, g_2, \dots, g_k \in G \text{ such that } K \subset \bigcup_{j=1}^k g_j A \right\}.$$

In analogy, we define  $[A : U]$  as well. Note that  $[K : A]$  and  $[K : U]$  defined above are finite, since  $K$  is compact. We also note that

$$(64.6) \quad [K : U] \leq [K : A] \cdot [A : U]$$

by (64.4) and (64.5). Thus, we can define

$$(64.7) \quad \lambda : \mathcal{N} \rightarrow \prod_{K \in \mathcal{K}} [0, [K : A]] \text{ by } \lambda(U)_K := \frac{[K : U]}{[A : U]}.$$

Since  $\prod_{K \in \mathcal{K}} [0, [K : A]]$  is compact by the Tikonov theorem, the limit

$$(64.8) \quad \lambda_K := \lim_{U \in \mathcal{N}_0} \lambda(U)$$

exists for some subnet  $\mathcal{N}_0$ .

**Claim 64.6.** *Let  $K, L \in \mathcal{K}$  and  $U \in \mathcal{O}$ . Then we have  $\lambda(U)_{K \cup L} \leq \lambda(U)_K + \lambda(U)_L$ . Furthermore, if we assume in addition that  $K \cap L = \emptyset$ , then we have  $\lambda(U)_K + \lambda(U)_L = \lambda(U)_{K \cup L}$ .*

Once this claim is proved, then it follows that  $\lambda$  is a content.

*Proof.* The first inequality is immediate from the definition of  $\lambda$ . Let us prove the second assertion. By virtue of the Hausdorff property of  $G$ , there exist open sets  $U$  and  $V$  that separate  $K$  and  $L$ ; the open sets  $U$  and  $V$  can be taken  $K \subset U$ ,  $L \subset V$  and  $U \cap V \neq \emptyset$ . If  $W \in \mathcal{N}_0$  is small enough, then  $k \cdot W \subset U$  and  $l \cdot W \subset V$  for all  $k \in K$  and  $l \in L$ , where we defined

$$(64.9) \quad k \cdot W = \{k \cdot w : w \in W\}$$

and  $l \cdot W$  analogously. Thus, if  $W' \gg W$  in the order of  $\mathcal{N}$ , that is, if  $W' \subset W$ , then the above observation yields  $\lambda(W')_K + \lambda(W')_L = \lambda(W)_{K \cup L}$ . A passage to the limit then proves the desired additivity.  $\square$

Denote  $\mu_e$  by the induced outer measure of the content  $\lambda$ . Let  $\mu$  be the restriction of  $\mu_e$  to measurable sets. Then,  $\mu$  is a left-invariant measure which we are looking for.

We need to prove the uniqueness of such a measure  $\mu$ . To this end suppose that we are given a measure  $\mu^*$  which is left-invariant. Let  $F$  be a positive Borel measurable function. We calculate

$$\begin{aligned} \int_G F(g) \mu(g) \times \mu^*(U_0) &= \iint_{G \times G} \chi_{U_0}(g_1) F(g_2) d\mu^*(g_1) d\mu(g_2) \\ &= \iint_{G \times G} \chi_{U_0}(g_1) F(g_1^{-1} g_2) d\mu^*(g_1) d\mu(g_2) \\ &= \iint_{G \times G} \chi_{U_0}(g_2 g_1) F(g_1^{-1}) d\mu^*(g_1) d\mu(g_2). \end{aligned}$$

Note that, if we set

$$H(g_1) = F(g_1^{-1}) \int_G \chi_{U_0}(g_2 g_1) d\mu(g_2),$$

then the mapping  $F \mapsto H$  is invertible and hence we can tell the value of  $\int_G H(g_1) d\mu^*(g_1)$  from  $\mu$  and  $\mu^*(U_0)$ , which shows that  $\mu = \mu^*$ .  $\square$

**Definition 64.7.** A locally compact topological group is said to be a unilocular, if the left Haar measure additionally satisfies

$$(64.10) \quad \mu(Eg) = \mu(E)$$

for all  $g \in G$  and  $E \in \mathcal{B}$ , or equivalently

$$(64.11) \quad \int_G f(gx) d\mu(x) = \int_G f(xg) d\mu(x) = \int_G f(x) d\mu(x)$$

for all  $g \in G$  and  $f \in L^1(G)$ . In this sense  $\mu$  is bi-invariant.

In what follows we always assume that  $G$  is a locally compact topological group.

**Theorem 64.8.** Any compact topological group is unilocular.

*Proof.* Since  $\mu(G) < \infty$  by compactness of  $G$ , we see that

$$(64.12) \quad \nu(E) = \frac{1}{\mu(G)} \int_{E \times G} \mu(xy) d\mu(x) d\mu(y), \quad E \in \mathcal{B}$$

defines a measure. The measure defined by the right-hand side is invariant. Thus, so is  $\mu$ .  $\square$

It is not sufficient to prove the existence of the Haar measure. Below let us see an important example.

**Example 64.9.** The Lebesgue measure  $dx$  is a Haar measure of  $(\mathbb{R}^d, +)$ .

**Example 64.10.** Let  $(0, \infty)$  be a Lie group whose group operation is the multiplication. Then  $\frac{dt}{t}$  is its Haar measure. Indeed, let  $f$  be a positive measurable function. Then the invariance of  $\frac{dt}{t}$  is equivalent to saying

$$(64.13) \quad \int_0^\infty f(at) \frac{dt}{t} = \int_0^\infty f(t) \frac{dt}{t}.$$

It is not only this formula that makes us shed light on this Lie group. We also have

$$(64.14) \quad \int_0^\infty f(t^a) \frac{dt}{t} = \frac{1}{a} \int_0^\infty f(t) \frac{dt}{t}.$$

Regardless of the property of Haar measures, we can prove (64.13) and (64.14) by means of change of variables.

Haar measures are used for averaging procedures. We often utilize them in order to make the matters invariant of the action of Lie group. Below we shall give a typical and important application of averaging procedures.

**Theorem 64.11.** *Let  $G$  be a compact Lie group. Suppose that  $G$  acts on a linear space  $V$  of finite dimension, that is, there exists a mapping  $\Phi : G \rightarrow \text{End}_{\mathbb{R}}(V) = \text{Hom}_{\mathbb{R}}(V, V)$  with the following properties: Below we denote  $\Phi(g)v = g \cdot v$  for  $g \in G$  and  $v \in V$ .*

- (1) *For all fixed  $v \in V$ , the mapping  $g \in G \mapsto g \cdot v \in V$  is continuous.*
- (2) *The mapping  $\Phi$  is associative;  $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$  for all  $g_1, g_2 \in G$  and  $v \in V$ .*
- (3) *Denote by  $e_G$  the unit element of  $G$ . Then  $\Phi(e_G) = \text{id}_V$ .*

*Then there exists an inner product  $\langle *, * \rangle$  on  $V$  such that  $\langle g \cdot u, g \cdot v \rangle = \langle u, v \rangle$  for all  $u, v \in V$  and  $g \in G$ .*

*Proof.* Choose an arbitrary inner product  $\langle *, * \rangle_0$  on  $V$ . Then we define the inner product by

$$(64.15) \quad \langle u, v \rangle := \int_G \langle g \cdot u, g \cdot v \rangle_0 d\mu_G(g) \quad (u, v \in V),$$

where  $\mu_G$  is the normalized Haar measure of  $G$ . It is straightforward that we check  $\langle *, * \rangle$  is an inner product. Let us show the  $G$ -invariance;  $\langle g \cdot u, g \cdot v \rangle = \langle u, v \rangle$  for all  $u, v \in V$ . We utilize the right invariance of  $\mu_G$ : For all  $g \in G$  we have

$$\langle g \cdot u, g \cdot v \rangle = \int_G \langle h \cdot g \cdot u, h \cdot g \cdot v \rangle_0 d\mu_G(h) = \int_G \langle h \cdot u, h \cdot v \rangle_0 d\mu_G(h) = \langle u, v \rangle.$$

Thus, we obtain the desired inner product. □

*Exercise 275.* The aim of this exercise is to see how powerful the above inner product is. Keep to the same setting as above. Suppose  $\langle *, * \rangle$  is an inner product obtained in the theorem. A subspace  $W$  of  $V$  is said to be  $G$ -invariant, if  $g \cdot w \in W$  for all  $w \in W$  and  $g \in G$ . Prove the following.

- (1) Denote by  $W^\perp$  the  $\langle *, * \rangle$ -orthogonal complement of  $W$ . If a subspace  $W$  is  $G$ -invariant, so is  $W^\perp$ .
- (2)  $V$  can be partitioned into an orthogonal sum of  $G$ -invariant spaces.

64.2. Hausdorff measures.

In this book we had been ignored the set of measure zero. In particular when we have developed the maximal operator theory and the singular integral theory, it had been usual that we disregard such sets. However, the attempt has been made to grasp the structure of the null sets. We begin to be able to grasp the feature of them. It seems that we are not totally skillful to describe such a complexity. Here we shall present a method to access such a complexity.

Hausdorff measure. The Hausdorff measure is a continuous extension of the Lebesgue measure. The Hausdorff measure can be used to define and measure the fractional dimension.

**Definition 64.12** ( $\delta$ -covering). Let  $\delta > 0$ . A  $\delta$ -covering of a set  $A$  is a covering of  $A$  consisting of open balls whose radii are less than  $\delta$ .

**Definition 64.13** (Hausdorff measure). Let  $A$  be a set in  $\mathbb{R}^d$ .

$$(1) \text{ Let } s \geq 0. \text{ Then set } \mathcal{H}_\delta^s(A) := \inf \left\{ \sum_{j=1}^\infty \omega_s r_j^s : \{B(x_j, r_j)\}_{j=1}^\infty \text{ is a } \delta\text{-covering of } A \right\}.$$

$$\text{Here } \omega_s \text{ is a constant given by } \omega_s = \pi^{\frac{s}{2}} \Gamma\left(\frac{s+2}{2}\right)^{-1}.$$

$$(2) \text{ Let } s \geq 0. \text{ Then define } \mathcal{H}^s(A) = \lim_{\delta \downarrow 0} \mathcal{H}_\delta^s(A).$$

$$(3) \text{ The Hausdorff dimension of a set } A \text{ is given by } \dim_{\mathcal{H}}(A) = \inf\{s \geq 0 : \mathcal{H}^s(A) = 0\}.$$

Note that  $\omega_s$  is the surface area of  $S^{s-1}$ . As for this definition we have to verify this definition makes sense.

**Lemma 64.14.** *Let  $A \subset \mathbb{R}^d$  be a set.*

- (1) *For all  $s > 0$   $\delta \mapsto \mathcal{H}_\delta^s(A)$  is decreasing. In particular the limit in the definition of  $\mathcal{H}^s(A)$  exists.*
- (2) *If  $s > d$ , then  $\mathcal{H}^s(\mathbb{R}^d) = 0$ . In particular  $\mathcal{H}^s(A) = 0$  whenever  $s > d$ .*

Before we come to the proof, we make a remark above the definition. (1) ensures that the limit exists in  $[0, \infty]$ . Meanwhile (2) says that the set appearing in the infimum defining  $\dim_{\mathcal{H}}(A)$  is not empty.

*Proof.* The first assertion is clear because

$$(64.16) \quad \delta \in (0, \infty) \mapsto \left\{ \sum_{j=1}^\infty \omega_s r_j^s : \{B(x_j, r_j)\}_{j=1}^\infty \text{ is a } \delta\text{-covering of } A \right\} \subset 2^{\mathbb{R}}$$

is increasing. To prove the second assertion, we set  $Q_l = l + [0, 1]^d$ . Rearrange  $\{Q_l\}_{l \in \mathbb{Z}^d}$  to  $R_1, R_2, \dots, R_j, \dots$ . Let  $K \in \mathbb{N}$  be chosen arbitrarily. For each  $j$  we divide equally  $R_j$  into  $(K 2^j)^d$  cubes and obtain  $(K 2^j + 1)^d$  vertices of them. Here we do not count them according to their multiplicity. Cover  $R_j$  with balls centered at such vertices of radius  $\sqrt{d}(K 2^j)^{-1}$ . Then

$$(64.17) \quad \mathcal{H}_{\sqrt{d}K^{-1}}^s(\mathbb{R}^d) \leq \omega_s \sum_{j=1}^\infty (\sqrt{d}(K 2^j)^{-1})^s (K 2^j + 1)^d \lesssim K^{d-s}.$$

Therefore, the integer  $K$  being arbitrary, we conclude  $\mathcal{H}^s(\mathbb{R}^d) = 0$ . □

**Proposition 64.15.**  *$\mathcal{H}^s$  is a content.*

*Proof.* It is easy to see that  $\mathcal{H}^s(\emptyset) = 0$ . Suppose that  $K$  and  $L$  are disjoint compact sets. Then we can tear them apart with  $\delta_0$ -neighborhoods. That is, there exists  $\delta_0 > 0$  such that  $K_{\delta_0} \cap L_{\delta_0} = \emptyset$ . Let  $\delta < \delta_0$ . Suppose that  $\{B_j\}_{j=1}^\infty$  are  $\delta$ -coverings of  $K \cup L$  respectively. For the purpose of estimating  $\mathcal{H}_\delta^s(K \cup L)$  we can assume each ball  $B_j$  meets  $K$  or  $L$  by discarding unnecessary balls,

Set  $J(K) := \{j \in \mathbb{N} : B_j \cap K \neq \emptyset\}$  and  $J(L) := \{j \in \mathbb{N} : B_j \cap L \neq \emptyset\}$ . Then  $\mathbb{N}$  was partitioned into  $J(K)$  and  $J(L)$ . Therefore

$$\begin{aligned} \mathcal{H}_\delta^s(K \cup L) &= \inf \left\{ \sum_{j=1}^\infty \omega_s r(B_j)^s : K \cup L \subset \bigcup_{j=1}^\infty B_j \right\} \\ &= \inf \left\{ \sum_{j=1}^\infty \omega_s r(B_j)^s : K \subset \bigcup_{j=1}^\infty B_j \right\} + \inf \left\{ \sum_{j=1}^\infty \omega_s r(B_j)^s : L \subset \bigcup_{j=1}^\infty B_j \right\} \\ &= \mathcal{H}_\delta^s(K) + \mathcal{H}_\delta^s(L). \end{aligned}$$

Here we have used the fact that  $\delta < \delta_0$  to deduce the second inequality. Letting  $\delta \downarrow 0$ , we obtain  $\mathcal{H}^s(K) + \mathcal{H}^s(L) = \mathcal{H}^s(K \cup L)$

Finally we shall prove the subadditivity. Let  $K$  and  $L$  be compact sets not necessarily disjoint. Then choose  $\delta$ -coverings of  $\{B_j\}_{j \in J_1}$  and  $\{B_j\}_{j \in J_2}$ , where  $J_1$  and  $J_2$  are at most countable. Then  $\{B_j\}_{j \in J_1 \cup J_2}$  is a  $\delta$ -covering of  $K \cup L$ . Therefore, from this observation we conclude  $\mathcal{H}_\delta^s(K \cup L) \leq \mathcal{H}_\delta^s(K) + \mathcal{H}_\delta^s(L)$ .  $\square$

Finally to finish this paragraph, we give a formula of the surface area of a graph.

Set

$$(64.18) \quad G(f) = \{(x, f(x)) : x \in \Omega\} \subset \mathbb{R}^{d+1}$$

for a function defined on a set  $\Omega \subset \mathbb{R}^d$ .

**Lemma 64.16.** *Suppose that  $K$  is a compact set and  $U$  is an open set containing  $K$ . If  $\{f_j\}_{j \in \mathbb{N}}$  is a sequence of functions in  $C^1(U)$  defined on  $U$  such that  $\{f_j\}_{j \in \mathbb{N}}$  and  $\{f'_j\}_{j \in \mathbb{N}}$  converge uniformly over  $K$ . Then*

$$(64.19) \quad \lim_{j \rightarrow \infty} \mathcal{H}^d(G(f_j)) = \mathcal{H}^d(G(f)).$$

*Proof.* The proof is straightforward, if we notice

$$\begin{aligned} &\sqrt{(x-a)^2 + (f(x)-b)^2} \\ &\leq \sqrt{(x-a)^2 + (g(x)-g(a)+f(a)-b)^2} + |f(x)-f(a)-g(x)+g(a)| \\ &\leq \sqrt{(x-a)^2 + (g(x)-g(a)+f(a)-b)^2} (1 + \|f' - g'\|_{L^\infty(K)}) \end{aligned}$$

for all  $f, g \in C^1(U)$ . We can reduce the matter to the case when  $f$  is Affine.  $\square$

Now let us reconsider the notion of surface area.

**Theorem 64.17.** *Let  $f$  be an  $\mathbb{R}$ -valued  $C^1$ -mapping from an open set  $\Omega$  in  $\mathbb{R}^d$ . Then*

$$(64.20) \quad \mathcal{H}^d(G(f)) := \int_\Omega \sqrt{1 + |Df|^2}.$$

*Proof.* By the additivity we have only to prove the assertion with  $\Omega$  replaced by a relatively compact open set  $\Omega^*$  contained in  $\Omega$ . Therefore, we may assume  $f$  is defined on a domain which engulfs the closure of  $\Omega$ . In particular, we have

$$(64.21) \quad M := \sup_{x \in \Omega} |Df(x)| < \infty.$$

Let us define

$$(64.22) \quad \mu(E) := \mathcal{H}^d((E \times \mathbb{R}) \cap G(f))$$

for a Lebesgue measurable set  $E \subset \Omega$ . Note that

$$(64.23) \quad G(f) \cap (B(x, r) \times \mathbb{R}) \subset B((x, f(x)), r\sqrt{1+M^2}).$$

for all  $x \in \Omega$ . Therefore,  $\mu$  is absolutely continuous with respect to the Lebesgue measure  $dx$  on  $\Omega$ . Let us denote by  $\alpha$  the density of  $\mu$  with respect to  $dx$ . Then the Lebesgue convergence theorem gives us

$$(64.24) \quad \alpha(x) = \lim_{r \rightarrow 0} \frac{\mu(B(x, r))}{|B(x, r)|}$$

for almost every  $x \in \Omega$ .

A series of changing variables gives us

$$\begin{aligned} & \mu(B(x, r)) \\ &= \int_{\mathbb{R}^{d+1}} \chi\{(y', y_{d+1}) \in \mathbb{R}^{d+1} : y_{d+1} = f(y'), |y' - x|^2 < r^2\} d\mathcal{H}^d(y', y_{d+1}) \\ &= \int_{\mathbb{R}^{d+1}} \chi\{(z', z_{d+1}) \in \mathbb{R}^{d+1} : z_{d+1} = f(x + z') - f(x), |z'| < r\} d\mathcal{H}^d(z', z_{d+1}) \\ &= r^d \int_{\mathbb{R}^{d+1}} \chi\left\{(w', w_{d+1}) \in \mathbb{R}^{d+1} : w_{d+1} = \frac{f(x + r w') - f(x)}{r}, |w| < 1\right\} d\mathcal{H}^d(w', w_{d+1}). \end{aligned}$$

A passage to the limit using Lemma 64.16 reduces the matter to the case  $f$  is linear. If  $f$  is linear, then the assertion is obtained from a simple geometric argument.  $\square$

### 64.3. Hausdorff metric and fractals.

In general it is very difficult to calculate exactly the value of  $\mathcal{H}^s(A)$  of a set  $A$  with  $s \notin \mathbb{N}$ . The Hausdorff measures are not used in such a way. That is, it is not the value of  $\mathcal{H}^s(A)$  itself that counts. In this section we shall obtain a tool to obtain sets of fractional dimension. To this end, we induce a metric to the set of all compact sets. In this paragraph, we assume that  $X$  is a complete metric space. The Hausdorff distance (Hausdorff-Pompeiu distance) serves as a measure of vicinity.

**Definition 64.18** (Hausdorff distance, Hausdorff-Pompeiu distance). Denote by  $\mathcal{K}(X)$  the set of all compact sets in  $X$ .

- (1) Let  $K \in \mathcal{K}(X)$  and  $\delta > 0$ . The  $\delta$ -body of  $K$ , which is denoted by  $K_\delta$ , is the set of all points in  $X$  whose distance from  $K$  is less than  $\delta$ .
- (2) Define a metric function  $d_{\mathcal{K}}$  on  $\mathcal{K}(X)$  by

$$(64.25) \quad d_{\mathcal{K}}(K, L) := \inf\{\delta > 0 : L \subset K_\delta, K \subset L_\delta\} = \sup_{k \in K} \left( \inf_{l \in L} |k - l| \right).$$

$d_{\mathcal{K}}(K, L)$  is called the Hausdorff distance between  $K$  and  $L$ .

*Exercise 276.* In this exercise we let  $X = \mathbb{R}^2$ .

- (1) Let  $K = \{x \in \mathbb{R}^2 : |x_1| + |x_2| \leq 2\}$ . Display  $K_1$  in the  $(x_1, x_2)$ -plane.



(2) Let  $L = Q(2)$ . Then calculate  $d_{\mathcal{K}}(K, L)$ .

**Theorem 64.19.** *The metric space  $(X, d_{\mathcal{K}})$  is complete.*

*Proof.* The axiom of the complete metric space is clear other than completeness. We concentrate on completeness of  $(X, d_{\mathcal{K}})$ .

Let  $\{K_j\}_{j \in \mathbb{N}}$  be a Cauchy sequence: We need to construct a compact set  $K$  to which  $\{K_j\}_{j \in \mathbb{N}}$  converges. We define

$$(64.26) \quad K := \left\{ x \in X : \lim_{j \rightarrow \infty} x_j = x, \text{ where } \{x_j\} \text{ satisfies (64.27)} \right\}.$$

Here the condition (64.27) is

$$(64.27) \quad x_1 \in K_{n_1}, x_2 \in K_{n_2}, \dots \text{ for some increasing sequence } n_1 < n_2 < \dots$$

We claim that  $K$ , defined above, is actually compact and that  $\{K_j\}_{j \in \mathbb{N}}$  tends to  $K$ .

To prove that  $K$  is compact, it suffices to prove that  $K$  is totally bounded and that  $K$  is closed. Because we are assuming  $X$  is complete.

To prove that  $K$  is totally bounded, we take  $\varepsilon > 0$ . Then there exists  $J_0 > 0$  such that  $d_{\mathcal{K}}(K_j, K_l) < \frac{\varepsilon}{4}$  for all  $j, l \geq J_0$ . Then we have

$$(64.28) \quad K_l \subset K_{J_0 + \frac{\varepsilon}{4}}.$$

In view of compactness of  $K_{J_0}$ , we can cover  $K_{J_0}$  with a finite number of  $\frac{\varepsilon}{4}$ -open balls  $B_1, B_2, \dots, B_N$ . Then the collection  $\{4B_1, 4B_2, \dots, 4B_N\}$  is an open cover of  $\bigcup_{j=J_0}^{\infty} K_j$ . Therefore we can cover  $K$  with  $N$   $\varepsilon$ -open balls.

To prove that  $K$  is closed, we take a sequence  $\{x_j\}_{j=1}^{\infty}$  in  $K$  convergent to  $x \in X$ . Let  $\{x_{j,k}\}_{k=1}^{\infty}$  be a sequence corresponding to  $x_j$ .

First we take  $j_1$  so that  $d(x, x_{j_1}) < 1$  and we choose  $k_1$  and  $l_1$  so large that  $d(x_{j_1}, x_{j_1, k_1}) < 1$  with  $x_{j_1, k_1} \in K_{l_1}$ . Next we take  $j_2 > j_1$  so that  $d(x, x_{j_2}) < \frac{1}{2}$  and we choose  $k_2 > k_1$  and  $l_2 > l_1$  so large that  $d(x_{j_2}, x_{j_2, k_2}) < \frac{1}{2}$  with  $x_{j_2, k_2} \in K_{l_2}$ . Repeat this procedure and then we will obtain three increasing sequences  $\{j_m\}_{m \in \mathbb{N}}, \{k_m\}_{m \in \mathbb{N}}, \{l_m\}_{m \in \mathbb{N}}$  so that they satisfy  $x_{j_m, k_m} \in K_{l_m}$  and  $d(x_{j_m, k_m}, x) \leq \frac{2}{m}$ . Therefore  $x \in K$ .

Consequently  $K$  is compact. It remains to show that  $K$  is a limit of the sequence.

Let  $\varepsilon > 0$  be fixed and then there exists  $J_0 \in \mathbb{N}$  such that  $d_{\mathcal{K}}(K_j, K_l) < \frac{\varepsilon}{2}$  for all  $j, l \geq J_0$ . Let  $j \geq J_0$ .

Let  $x \in K$ . Then there exists a sequence  $\{x_j\}_{j \in \mathbb{N}}$  convergent to  $x$  with (64.27). If we take  $l$  large enough, then  $d(x, x_l) < \frac{\varepsilon}{2}$ . Since  $d_{\mathcal{K}}(K_j, K_l) < \frac{\varepsilon}{2}$ , we can choose  $x_j \in K_j$  so that  $d(x_j, x_k) < \frac{\varepsilon}{2}$ . Therefore, we conclude  $d(x, x_j) < \varepsilon$  and hence  $x \in K_{j\varepsilon}$ . Conversely let  $x \in K_j$ . Then there exists a sequence  $\{x_j\}_{j \in \mathbb{N}}$  convergent to  $x$  with (64.27) as before. Since  $x = \lim_{l \rightarrow \infty} x_l \in \bigcup_{l \geq j} K_l \subset K_{\varepsilon}$ , we conclude that  $K_j \subset K_{\varepsilon}$ . Therefore, it follows that  $d_{\mathcal{K}}(K, K_j) \leq \varepsilon$  for all  $j \geq J_0$ .

Hence  $\{K_j\}_{j \in \mathbb{N}}$  converges to  $K$ . □

Below we shall present an application of the completeness.

**Example 64.20.** The author borrowed this example from [511]. Given  $\alpha \in \mathbb{R}$  and two sets  $K, L$ , we write

$$(64.29) \quad K + L = \{k + l : k \in K, l \in L\}, \quad \alpha K = \{\alpha k : k \in K\}.$$

Suppose that  $K_0$  is a fixed compact set. Define

$$(64.30) \quad \Phi(K) := \frac{1}{2}(K + K_0).$$

Then it is easy to show that  $\Phi$  is a contraction on  $\mathcal{K}(X)$ . Therefore,  $\Phi$  has a unique fixed point. As a result we can say that for all compact sets  $K_0$  there exists a compact set  $K$  such that

$$(64.31) \quad K = \frac{1}{2}(K + K_0).$$

*Exercise 277.* Keep to the same notation as above. For the sake of simplicity let  $X = \mathbb{R}^2$ . Define  $\Phi(K_0)$  as the unique fixed point described above. Show the following.

- (1)  $\Phi$  is unique.
- (2)  $\Phi : \mathcal{K}(\mathbb{R}^2) \rightarrow \mathcal{K}(\mathbb{R}^2)$  is continuous.

*Exercise 278.* Establish that the set of all connected compact sets in  $\mathcal{K}(X)$  is a closed set.

An attractor of the iterating function system. To apply Theorem 64.19 we introduce a set of mappings.

**Definition 64.21** (Iterating function system). A system of Lipschitz functions  $\{S_j\}_{j=1}^m$  is said to be an iterating function system, if each  $S_j$  satisfy

$$(64.32) \quad |S_j(x) - S_j(y)| \leq r|x - y|,$$

where  $r \in (0, 1)$  is independent of  $j = 1, 2, \dots, m$ . In what follows one abbreviates iterating function system to IFS.

**Theorem 64.22.** Suppose that  $\{S_j\}_{j=1}^m$  is an IFS. Then there is a unique compact set  $E \in \mathcal{K}$  such that

$$(64.33) \quad E = \bigcup_{j=1}^m S_j(E).$$

*Proof.* For the proof we reformulate the problem: Define a mapping  $\Phi : \mathcal{K} \rightarrow \mathcal{K}$  by the formula

$$(64.34) \quad \Phi(F) := \bigcup_{j=1}^m S_j(F).$$

Our task is to show the unique existence of a fixed point in  $\mathcal{K}$ . Let  $x \in K, y \in L$ . Then from (64.32) we deduce that  $d_{\mathcal{K}}(\Phi(K), \Phi(L)) \leq r d_{\mathcal{K}}(K, L)$ . Hence,  $\Phi$  is a contraction and the desired result follows from the fixed point theorem of contractions. □

**Definition 64.23** (Attractor). The unique set  $\Gamma$ , whose existence was proved in Theorem 64.22, is called an attractor of IFS  $\{S_j\}_{j=1}^m$ .

*Notation.* Suppose that  $\{S_j\}_{j=1}^m$  is an IFS. Let  $I = (i_1, i_2, \dots, i_k)$  with  $1 \leq i_j \leq m$ . We denote

$$(64.35) \quad S_I(A) := S_{i_k} \circ S_{i_{k-1}} \circ \dots \circ S_{i_1}(A)$$

for all  $A \in \mathcal{K}$ .

As a special case of IFS we will consider similitude functions.

**Definition 64.24.** A similitude is a continuous function  $S$  satisfying

$$(64.36) \quad |S(x) - S(y)| = r|x - y|$$

for some  $0 < r < 1$ .

The following theorem is well-known.

**Proposition 64.25.** Suppose that  $\{S_j\}$  is a set of similitude functions. That is

$$(64.37) \quad |S_j(x) - S_j(y)| = r_j|x - y|.$$

Then we can write  $S_j(x) = r_jA_j(x) + b_j$ , where  $r_j \in (0, \infty)$ ,  $b_j \in \mathbb{R}^d$  and  $A_j \in O(n)$ .

Here we shall not prove Proposition 64.25, which is a review of linear algebra.

*Exercise 279.* The aim of this exercise is to prove Proposition 64.25.

- (1) First reduce the matter to the case  $S_j$  preserves the origin.
- (2) Show that  $S_j$  preserves the middle point. That is, if  $c$  is the middle point of  $a$  and  $b$ , then  $S_j(c)$  is the middle point of  $S_j(a)$  and  $S_j(b)$ .
- (3) Prove that  $S_j$  is a linear operator.
- (4) Finally prove Proposition 64.25.

Weak-convergence of the measures. What nontrivial measure on such fractal sets above do we have ? To answer such a question and to construct a measure, we use the approximation procedure. Recall that  $M(\mathbb{R}^d)$  denotes a totality of the finite Borel measures. Equipped with the weak-\* topology.

It follows from the definition that a sequence  $\{\mu_k\}_{k \in \mathbb{N}} \subset M(\mathbb{R}^d)$  converges to  $\mu \in M(\mathbb{R}^d)$  if

$$(64.38) \quad \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} f(x) d\mu_k(x) = \int_{\mathbb{R}^d} f(x) d\mu(x)$$

for all continuous and compactly supported functions  $f \in C_c$ .

The following is a simple criterion of compactness of sequences of measures.

**Proposition 64.26.** Suppose that  $\{\mu_k\}_{k \in \mathbb{N}} \subset M(\mathbb{R}^d)$ . If we have  $\sup_{k \in \mathbb{N}} \mu_k(A) < \infty$  for all compact sets  $A$ , then we can take a subsequence convergent in the sense of (64.38).

*Proof.* Since  $C_c(\mathbb{R}^d)$  is separable, we can take a dense countable set  $\{f_m\}_{m \in \mathbb{N}} \subset C_c(\mathbb{R}^d)$ . A diagonal argument allows us to arrange even that  $\lim_{l \rightarrow \infty} \int_{\mathbb{R}^d} f_m(x) d\mu_{k_l}(x)$  converge for each  $m \in \mathbb{N}$ , if we pass to a subsequence  $\{\mu_{k_l}\}_{l \in \mathbb{N}}$  of  $\{\mu_j\}_{j \in \mathbb{N}}$ . Thus, if we use the standard density argument, we see that  $\{\mu_{k_l}\}_{l \in \mathbb{N}}$  is the desired subsequence.  $\square$

IFS made of similitudes. Let us apply Theorem 64.19 to obtain compact sets of interest such as the Cantor set, Sierpinski gasket and so on. Below is the general procedure to obtain such an interesting set.

**Definition 64.27** (The open set condition). An IFS  $\{S_j\}_{j=1}^m$  is said to satisfy the open set condition if there exists an open set  $U$  such that

$$(64.39) \quad \sum_{j=1}^m S_j(U) \subset U.$$

**Remark 64.28.** In the sequel we consider similitudes only so  $S_j$ 's admit the expression in Proposition 64.25. Throughout the rest of this report, we use the notation in Proposition 64.25. We call  $\{S_j\}_{j=1}^m$  a similitude IFS.

**Definition 64.29.** Suppose that a similitude IFS  $\{S_j\}_{j=1}^m$  satisfies the open set condition. Define  $\mu_k \in M(\mathbb{R}^d)$  for  $k \in \mathbb{N}$  in the following way: Let  $U$  be an open set associated to the open set condition. Take  $V$  as a nonempty compact set contained in  $U$ . (It may be arbitrary as long as it is contained in  $U$ .) Firstly we temporarily define index sets.

$$(64.40) \quad J_k := \{I = (i_1, \dots, i_k) : 1 \leq i_j \leq m, \forall j \in \{1, 2, \dots, k\}\}$$

and  $J = \bigcup_{k=1}^{\infty} J_k$ . Let  $\mu_k$  be a probability Borel measure whose restriction to  $V_I$  is

$$(64.41) \quad c_d r_{i_1}^{s-d} r_{i_2}^{s-d} \dots r_{i_k}^{s-d} dx,$$

where  $I \in J_k$  and  $c_d$  is a normalization constant.

**Lemma 64.30.** *Suppose that  $\{V_i\}_{i \in I}$  is a family of disjoint open sets such that any  $V_i$  with  $i \in I$  contains a ball with its radius  $ar$  and is contained in a ball with its radius  $br$ . Then any ball  $B(x, r)$  can intersect at most  $(1 + b)^d/a^d$  open sets of  $V_i$  ( $i \in I$ ).*

*Proof.* Define  $I_0$  as

$$(64.42) \quad I_0 := \{i \in I : B(x, r) \cap V_i\}.$$

And denote by  $v_d$  the Lebesgue measure of  $|B(o, 1)|$  in  $\mathbb{R}^d$ . For each  $i \in I_0$  there exists a ball  $B(x_i, ar) \subset V_i$ . Using this ball, we have

$$(64.43) \quad \#I_0 v_d (ar)^d \leq \left| \bigcup_{i \in I_0} V_i \right| \leq |B(x, br + r)|.$$

This implies  $\#I_0 \leq (1 + b)^d/a^d$ . □

**Theorem 64.31.** *Under the same notation as Definition 64.29  $\{\mu_m\}_{m \in \mathbb{N}}$  converges to a measure  $\mu \in M(\mathbb{R}^d)$ . Furthermore  $\mu(F \cap B(x, r)) \sim r^D$  uniformly on  $x \in E$  and  $r \in (0, 1)$ .*

*Proof.* For the proof we fix a continuous function with compact support and fix  $\varepsilon > 0$ . Then by uniform continuity we have  $\delta > 0$  such that  $|f(x) - f(y)| \leq \varepsilon$  for all  $x, y$  with  $|x - y| \leq \delta$ .

If  $m$  is sufficiently large, say  $m \geq M$ , then we have  $\text{diam}(V_I) \leq \delta$  for all  $I \in J_m$ . Using this observation, we have

$$(64.44) \quad \left| \int_{\mathbb{R}^d} f(x) d\mu_m(x) - \int_{\mathbb{R}^d} f d\mu_k(x) \right| < \varepsilon$$

for all  $k, m \leq m$ . Thus the limit  $\lim_{m \rightarrow \infty} \mu_m$  exists and the first assertion follows.

For the proof of the second assertion we introduce some notations. We set

$$(64.45) \quad r_{\max} := \max(r_1, r_2, \dots, r_m), \quad r_{\min} := \min(r_1, r_2, \dots, r_m).$$

We select an integer  $p$  so that  $r_{\max}^p < r \leq r_{\max}^{p-1}$ . Define a set  $J_r$  by

$$(64.46) \quad J_r := \{J := (j_1, j_2, \dots, j_q) : q \leq p, r_{\max}^p < \text{diam}(F_J(O)) \leq r_{\max}^{p-1}\}.$$

Notice that  $\{F_J(O)\}_{J \in J_r}$  satisfies the hypothesis of the lemma. The number  $B(x, r) \cap F_J(O) \neq \emptyset$  with  $J \in J_r$  is majorized by a constant  $c$  depending not on  $r, x$  but on  $O$ . Thus, it follows that

$$\mu(B(x, r)) = \mu \left( B(x, r) \cap \bigcup_{J \in J_r} F_J(O) \right). \text{ This implies that}$$

$$\mu(B(x, r)) = \liminf_{m \rightarrow \infty} \mu_m(B(x, r)) \leq \liminf_{m \rightarrow \infty} \sum_{\substack{J \in J_r \\ B(x, r) \cap F_J(O) \neq \emptyset}} \mu_m(F_J(O)) \lesssim r^D.$$

It remains to show the estimate below:  $\mu(B(x, r)) \gtrsim r^D$  for all  $r > 0$  and  $x$ . But it is easy. We can find  $I$  such that  $F_I(F) \subset B(x, r)$ . We take such  $I$  minimally in the sense that  $\#I$  is minimal. Then  $\text{diam}(F_I(F)) \geq \frac{r_{\min} r}{2}$ . Assume otherwise. Decomposing  $I = (I', i_d)$ , we have  $x \in F_{I'}(E)$  and  $\text{diam}(F_{I'}(F)) \leq \frac{r}{2}$ . This is a contradiction.

Thus, we conclude  $\mu(B(x, r)) \geq \mu(F_I(F)) \gtrsim r^D$ . □

The next theorem is useful when we want to know the Hausdorff dimension of the sets.

**Theorem 64.32.** *Let  $\Gamma$  be a compact set. Assume that  $\mu$  is a measure such that  $\mu(B(x, r)) \sim r^D$  holds uniformly over  $0 < r < 1$  and  $x \in \Gamma$ . Then the Hausdorff dimension of  $\Gamma$  is  $D$  and the  $D$ -dimensional Hausdorff measure of  $\Gamma$  is equivalent to  $\mu$ .*

*Proof.* Let  $A \subset \Gamma$ . Firstly let us cover  $A$  with balls. Suppose that we are given an  $r$ -cover of  $A$ . By the  $5r$ -covering lemma we have  $x_1, x_2, \dots$  such that

$$A \subset \bigcup_j B(x_j, 5r_j) \text{ and that } \{B(x_j, r_j)\}_j \text{ is disjoint.}$$

Then we have

$$(64.47) \quad \mu(A) \leq \sum_j \mu(A \cap B(x_j, 5r_j)) \leq 5^D \sum_j r_j^D.$$

This implies that  $\mu(A) \leq 5^D \mathcal{H}_r^D(A)$  hence a passage to the limit yields  $\mu(A) \leq 5^D \mathcal{H}^D(A)$ . Conversely we also have

$$(64.48) \quad \mathcal{H}_r^D(A) \leq \sum_j \mu(A \cap B(x_j, r_j)) \leq \mu(A).$$

Thus we obtain the desired assertion. □

A measure  $\mu$  on  $\mathbb{R}^d$  is said to be a Frostman measure if  $\mu(B(x, r)) \leq r^D$  for all  $r > 0$  and  $x$ .

Examples of attractors. In the examples below it is convenient to identify  $\mathbb{R}^2$  with  $\mathbb{C}$ . We shall present examples of attractors having special names.

**Example 64.33** (Cantor set). Set

$$(64.49) \quad F_1(z) := \frac{1}{3}z, \quad F_2(z) := \frac{1}{3}z + \frac{2}{3}$$

The attractor is called the Cantor set.

*Exercise 280.* Let  $I = [0, 1] \subset \mathbb{C}$ . Then display  $\Phi(I)$  and  $\Phi(\Phi(I))$ , where  $\Phi : \mathcal{K}(\mathbb{C}) \rightarrow \mathcal{K}(\mathbb{C})$  is given by  $\Phi(K) = F_1(K) \cup F_2(K)$ .

**Example 64.34** (Koch curve). Set

$$(64.50) \quad F_1(z) := \frac{1}{3}z, \quad F_2(z) := \frac{1}{3}e^{\frac{\pi}{3}i}z + \frac{1}{3}, \quad F_3(z) := \frac{1}{3}e^{-\frac{\pi}{3}i}z + \frac{1}{2} \left(1 + \frac{1}{\sqrt{3}}i\right), \quad F_4(z) := \frac{1}{3}z + \frac{2}{3}.$$

The attractor is called the Koch curve.

*Exercise 281.* Let  $I = [0, 1] \subset \mathbb{C}$ . Then display  $\Phi(I)$  and  $\Phi(\Phi(I))$ , where  $\Phi : \mathcal{K}(\mathbb{C}) \rightarrow \mathcal{K}(\mathbb{C})$  is given by  $\Phi(K) = F_1(K) \cup F_2(K)$ .

**Example 64.35** (Sierpinski gasket). Let  $p_1, p_2, p_3$  be distinct points. Set

$$(64.51) \quad F_1(z) := \frac{z + p_1}{2}, \quad F_2(z) := \frac{z + p_2}{2}, \quad F_3(z) := \frac{z + p_3}{2}.$$

The attractor is called the Sierpinski gasket.

Let  $V_0 = \{p_1, p_2, p_3\}$  and set  $V_{j+1} = F_1(V_j) \cup F_2(V_j) \cup F_3(V_j)$  for  $j = 0, 1, 2, \dots$ . Sometimes it is of use to consider each  $V_j$ .

*Exercise 282.* Let  $\Delta$  be a triangle whose vertices are  $p_1, p_2, p_3$ . Then display  $\Phi(\Delta)$  and  $\Phi(\Phi(\Delta))$ , where  $\Phi : \mathcal{K}(\mathbb{C}) \rightarrow \mathcal{K}(\mathbb{C})$  is given by  $\Phi(K) = F_1(K) \cup F_2(K) \cup F_3(K)$ .

**Example 64.36** (Cantor dust). Set

$$(64.52) \quad F_1(z) = \frac{1}{4}z, \quad F_2(z) = \frac{1}{4}z + \frac{3}{4}, \quad F_3(z) = \frac{1}{4}z + \frac{3}{4}i, \quad F_4(z) = \frac{1}{4}z + \frac{3}{4}(1 + i).$$

The attractor is said to be Cantor dust.

*Exercise 283.* Let  $I^2 = \{z \in \mathbb{C} : 0 \leq \operatorname{Re}(z) \leq 1, 0 \leq \operatorname{Im}(z) \leq 1\}$ . Then display  $\Phi(I^2)$  and  $\Phi(\Phi(I^2))$ , where  $\Phi : \mathcal{K}(\mathbb{C}) \rightarrow \mathcal{K}(\mathbb{C})$  is given by  $\Phi(K) = F_1(K) \cup F_2(K) \cup F_3(K) \cup F_4(K)$ .

**Example 64.37** (Hata's tree). Set

$$(64.53) \quad F_1(z) := \frac{1}{4}e^{\frac{\pi}{3}i}z, \quad F_2(z) := \frac{1}{4}e^{-\frac{\pi}{3}i}z + \frac{1}{4}.$$

The set  $K = \bigcup_{j \in \mathbb{N}} K_j$  is said to be Hata's tree. Here  $K_0 = [0, 1]$  and  $K_{j+1} = F_1(K_j) \cup F_2(K_j)$ .

*Exercise 284.* Calculate the Hausdorff dimension of the Cantor set, the Koch curve, the Sierpinski gasket and the Cantor dust after verifying the system of contractions satisfy the separation condition.

## 65. COUNTABLE PRODUCT OF PROBABILITY SPACES

Suppose that we are given a collection of probability space  $(X^{(j)}, \mathcal{B}^{(j)}, \mu^{(j)})$ ,  $j = 1, 2, \dots$ . The aim of this section is to define a measure on  $X := X^{(1)} \times X^{(2)} \times \dots$  which is compatible with  $(X^{(j)}, \mathcal{B}^{(j)}, \mu^{(j)})$ ,  $j = 1, 2, \dots$

**Definition 65.1** (Cylinder set). A cylinder set is a set of the form  $A^{(1)} \times A^{(2)} \times \dots$ , where  $A^{(j)} \in \mathcal{B}^{(j)}$  for all  $j = 1, 2, \dots$  and  $A^{(j)} = X^{(j)}$  for all  $j$  larger than some  $j_0$ . Denote by  $\mathcal{C}$  the set of all cylinder sets.

We shall prove the following theorem in this section:

**Theorem 65.2.** *There exists a measure  $\mu : \sigma(\mathcal{C}) \rightarrow [0, 1]$  such that*

$$(65.1) \quad \mu(A^{(1)} \times A^{(2)} \times \dots \times A^{(j_0)} \times X^{(j_0+1)} \times X^{(j_0+k)} \times \dots) = \mu^{(1)}(A^{(1)}) \times \dots \times \mu^{(j_0)}(A^{(j_0)}).$$

for all  $A^{(j)}$  with  $A^{(j)} \in \mathcal{B}^{(j)}$ ,  $j = 1, 2, \dots, j_0$ .

Suppose  $A \subset \mathcal{C}$  is expressed as

$$A := A^{(1)} \times A^{(2)} \times \dots,$$

where  $A^{(j)} \in \mathcal{B}^{(j)}$  for all  $j = 1, 2, \dots$  and  $A^{(j)} = X^{(j)}$  for all  $j$  larger than some  $j_0$ . Then we define

$$(65.2) \quad \mu_0(A) := \mu^{(1)}(A^{(1)}) \times \dots \times \mu^{(j_0)}(A^{(j_0)}).$$

Define a function  $\Gamma : 2^X \rightarrow [0, \infty]$  by

$$(65.3) \quad \Gamma(A) := \inf \left\{ \sum_{j=1}^{\infty} \mu_0(A_j) : A \subset \bigcup_{j=1}^{\infty} A_j, A_1, A_2, \dots \in \mathcal{C} \right\} \text{ for } A \in 2^X.$$

The following lemma, which we omit the proof, is easy by the definition of  $\Gamma$ .

**Lemma 65.3.**  $\Gamma$  is an outer measure on  $X^{(1)} \times X^{(2)} \times \dots$ .

To prove Theorem 65.2 we have only to show the following proposition:

**Proposition 65.4.** Let  $A \in \mathcal{C}$ . Then  $A$  is  $\Gamma$ -measurable and  $\mu_0(A) = \Gamma(A)$ .

This subsection is devoted to proving Proposition 65.4.

To prove Proposition 65.4 the following lemma is a key.

**Lemma 65.5.** Suppose  $A_1, A_2, \dots$  is decreasing in  $\mathcal{C}$ . Then  $\Gamma \left( \bigcap_{j=1}^{\infty} A_j \right) = \lim_{j \rightarrow \infty} \Gamma(A_j)$ .

*Proof of Lemma 65.5.* It is easy to see  $\Gamma \left( \bigcap_{j=1}^{\infty} A_j \right) \leq \lim_{j \rightarrow \infty} \Gamma(A_j)$  because of the monotonicity of  $\Gamma$ . Let us prove the reverse inequality. Given  $B \in \mathcal{C}$ , we have

$$(65.4) \quad \lim_{k \rightarrow \infty} \mu_0(B \cup A_k) = \mu_0(B).$$

Indeed from the definition of the cylinder set  $B \cup A_k$  can be written as  $B \cup A_k = E_k \times X^{(j_0+1)} \times X^{(j_0+2)} \times \dots$  and similarly  $B = E \times X^{(j_0+1)} \times X^{(j_0+2)} \times \dots$ , where  $E, E_k \in \mathcal{B}^{(1)} \times \dots \times \mathcal{B}^{(j_0)}$ . Note that

$$(65.5) \quad \bigcap_{k=1}^{\infty} E_k \times X^{(j_0+1)} \times X^{(j_0+2)} \times \dots = \bigcap_{k=1}^{\infty} B \cup A_k = B = E \times X^{(j_0+1)} \times X^{(j_0+2)} \times \dots,$$

implying  $\bigcap_{k=1}^{\infty} E_k = E$ . Thus by applying the monotone convergence theorem

$$(65.6) \quad \lim_{k \rightarrow \infty} \mu_0(B \cup A_k) = \lim_{k \rightarrow \infty} \mu^{(1)} \times \dots \times \mu^{(j_0)}(E_k) = \mu^{(1)} \times \dots \times \mu^{(j_0)}(E) = \mu_0(B).$$

As a result (65.4) is established. By virtue of (65.4) we obtain

$$\begin{aligned}
\Gamma\left(\bigcap_{j=1}^{\infty} A_j\right) &= \sup \left\{ \sum_{j=1}^{\infty} \mu_0(B_j) : \bigcup_{j=1}^{\infty} B_j \supset \bigcap_{j=1}^{\infty} A_j, B_j \in \mathcal{C}, j = 1, 2, \dots \right\} \\
&= \sup \left\{ \sum_{j=1}^{\infty} \lim_{k \rightarrow \infty} \mu_0(B_j \cup A_k) : \bigcup_{j=1}^{\infty} B_j \supset \bigcap_{j=1}^{\infty} A_j, B \in \mathcal{C} \right\} \\
&= \sup \left\{ \lim_{J \rightarrow \infty} \lim_{k \rightarrow \infty} \sum_{j=1}^J \mu_0(B_j \cup A_k) : \bigcup_{j=1}^{\infty} B_j \supset \bigcap_{j=1}^{\infty} A_j, B \in \mathcal{C} \right\} \\
&\geq \lim_{k \rightarrow \infty} \mu_0(A_k) \geq \lim_{k \rightarrow \infty} \Gamma(A_k).
\end{aligned}$$

Consequently Lemma 65.5 is proved.  $\square$

$$\text{Denote } \mathcal{C}' := \left\{ \bigcup_{j=1}^{\infty} A_j : A_j \in \mathcal{C} \right\} = \left\{ \prod_{j=1}^{\infty} A_j : A_j \in \mathcal{C} \right\}.$$

*Proof of Proposition 65.4*  $\Gamma(A) = \mu_0(A)$  for  $A \in \mathcal{C}$ . The proof of  $\Gamma(A) = \mu_0(A)$  is easy. Indeed, in the course of the proof of Lemma 65.5 we proved for a decreasing sequence of cylinder sets  $A_1 \supset A_2 \supset \dots$

$$(65.7) \quad \lim_{k \rightarrow \infty} \Gamma(A_k) \geq \Gamma\left(\bigcap_{j=1}^{\infty} A_j\right) \geq \lim_{k \rightarrow \infty} \mu_0(A_k) \geq \lim_{k \rightarrow \infty} \Gamma(A_k).$$

Thus, letting  $A_j = A, j = 1, 2, \dots$  we obtain  $\Gamma(A) = \mu_0(A)$ .  $\square$

As a corollary of Lemma 65.5 we obtain

$$(65.8) \quad \Gamma\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \Gamma(A_j),$$

provided  $A_1, A_2, \dots, A_j, \dots \in \mathcal{C}$  are disjoint. In fact,

$$\begin{aligned}
\Gamma(X^{(1)} \times X^{(2)} \times \dots) - \Gamma\left(\bigcup_{j=1}^{\infty} A_j\right) &= \lim_{J \rightarrow \infty} \Gamma\left(X^{(1)} \times X^{(2)} \times \dots \setminus \bigcup_{j=1}^J A_j\right) \\
&= \lim_{J \rightarrow \infty} \mu_0\left(X^{(1)} \times X^{(2)} \times \dots \setminus \bigcup_{j=1}^J A_j\right) \\
&= \lim_{J \rightarrow \infty} \mu_0\left(X^{(1)} \times X^{(2)} \times \dots\right) - \sum_{j=1}^J \mu_0(A_j) \\
&= \lim_{J \rightarrow \infty} \Gamma\left(X^{(1)} \times X^{(2)} \times \dots\right) - \sum_{j=1}^J \Gamma(A_j).
\end{aligned}$$

Noting that  $\Gamma(B) \leq 1$  for all  $B \subset X^{(1)} \times X^{(2)} \times \dots$ , we obtain (65.8). Thus we obtain

$$(65.9) \quad \Gamma(A) := \inf \left\{ \sum_{j=1}^{\infty} \Gamma(A_j) : A \subset \bigcup_{j=1}^{\infty} A_j, A_1, A_2, \dots \in \mathcal{C} \right\}.$$



**Lemma 65.6.** *A subset  $E \subset X^{(1)} \times X^{(2)} \times \dots$  is measurable, if and only if  $\Gamma(C) \geq \Gamma(A \cap C) + \Gamma(A^c \cap C)$  for all  $C \in \mathcal{C}$ .*

*Proof of Lemma 65.6.* Let  $E \subset X^{(1)} \times X^{(2)} \times \dots$ . By virtue of (65.9) we have

$$\begin{aligned} \Gamma(E) &= \inf \left\{ \sum_{j=1}^{\infty} \Gamma(A_j) : E \subset \bigcup_{j=1}^{\infty} A_j, A_1, A_2, \dots \in \mathcal{C} \right\} \\ &\geq \inf \left\{ \sum_{j=1}^{\infty} \Gamma(A \cap E) + \Gamma(A^c \cap E) : E \subset \bigcup_{j=1}^{\infty} A_j, A_1, A_2, \dots \in \mathcal{C} \right\} \\ &\geq \Gamma(A \cap E) + \Gamma(A^c \cap E). \end{aligned}$$

Since  $\Gamma$  is an outer measure, we have the last inequality. Consequently  $A$  is measurable.  $\square$

It remains to prove the measurability of  $A$  to finish the proof of Theorem 65.2. This can be achieved as follows: Let  $C \in \mathcal{C}$  be fixed. By Lemma 65.6 we have only to show

$$(65.10) \quad \Gamma(C) \geq \Gamma(A \cap C) + \Gamma(A^c \cap C).$$

Now that  $A \in \mathcal{C}$  the above formula can be rephrased as

$$(65.11) \quad \mu_0(C) \geq \mu_0(A \cap C) + \mu_0(A^c \cap C).$$

Since  $\mu_0$  is additive, this is trivial.

The proof of Theorem 65.2 is therefore complete.

Notes and references for Chapter 27.

Section 64. We refer to [13] for details of Hausdorff measures. We refer to [14] for more details of fractals.

Theorem 64.5

Theorem 64.8

Theorem 64.11

Theorem 64.17

Theorem 64.19

Hutchinson proved Theorem 64.22 in [246]. The terminology “IFS” is due originally to Micheal F. Barnsley. Lemma 64.30 and Theorem 64.31, as well as Definition 64.29, are due to Moran [359].

Theorem 64.32

We refer to [231] for details of self-similar sets.

Section 65. As for the proof of Theorem 65.2, the author referred to [15].

## TEXTBOOKS

## REFERENCES

- [1] D. R. Adams and L. I. Hedberg, *Function spaces and potential theory*, Springer Verlag, Berlin, Heiderberg, 1996.
- [2] L. V. Ahlfors, *Complex analysis*, McGraw-Hill Science/Engineering/Math.
- [3] K. Alois, J. Oldvrich and F. Svatopluk, *Function spaces, Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis*, Noordhoff International Publishing, Leyden, Academia, Prague, 1977.
- [4] E. Beckenstein and L. Narici, *Topological vector spaces*.
- [5] C. Bennett and R. Sharpley, *Interpolation of operators*, Academic Press, Boston, 1988.
- [6] J. Bergh and L. Jörgen, *Interpolation spaces. An introduction*, Grundlehren der Mathematischen Wissenschaften, **no. 223**. Springer-Verlag, Berlin-New York, 1976.
- [7] *Fourier integral operators*, Progress in Mathematics, **130**. Birkhäuser Boston, MA, 1996.
- [8] J. Doob, *Stochastic processes*, John Wiley and Sons, Inc., New York; Chapman and Hall, Limited, London, 1953.
- [9] N. Dunford, J. Schwartz, *Linear operators. Part I. General theory. With the assistance of William G. Bade and Robert G. Bartle. Reprint of the 1958 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley Sons, Inc., New York, 1988*.
- [10] J. Duoandikoetxea, *Fourier Analysis*. Translated and revised from the 1995 Spanish original by David Cruz-Urbe. Graduate Studies in Mathematics, 29. American Mathematical Society, Providence, RI, 2001.
- [11] D. Edmunds, V. Kokilashvili and A. Meskhi, *Bounded and compact integral operators*. Kluwer Academic Publishers, Dordrecht, Boston London, 2002.
- [12] C. Evans, *Partial Differential Equations*, American Mathematical Society, Grad. Studies Math. **19**, Providence, RI (1998).
- [13] C. Evans and F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, 1999.
- [14] K. Falconer, *Fractal geometry. Mathematical foundations and applications*, Second edition, John Wiley and Sons, Inc., Hoboken, NJ, 2003.
- [15] T. Funaki, *Stochastic differential equation*, (in Japanese) Iwanami.
- [16] J. García-Cuerva and J. L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland Math. Stud., **116** 1985.
- [17] D. Gilberg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer-Verlag, Berlin, 1983.
- [18] L. Grafakos, *Classical and Modern Fourier Analysis*, Pearson Education, NJ, 2004.
- [19] L. Grafakos, *Classical Fourier Analysis, Second Edition*, Graduate Texts in Math., no 249, Springer, New York, 2008.
- [20] L. Grafakos, *Modern Fourier Analysis, Second Edition*, Graduate Texts in Math., no 250, Springer, New York, 2008.
- [21] K. Gröchenig, *Foundations of time-frequency analysis. Applied and Numerical Harmonic Analysis*, Birkhäuser Boston, Inc., Boston, MA, 2001.
- [22] H. Hedenmalm, B. Boris Korenblum and K. Zhu, *Theory of Bergman spaces*, Springer New York.
- [23] P. Halmos, *Measure theory*, Springer Verlag, New York, (1974).
- [24] E. Hernandez and G. Weiss, *A First Course on Wavelets*, CRC Press, Boca Raton, FL. (1996).
- [25] E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis*, vol 2, Springer-Verlag, New York-Heidelberg-Berlin, (1970).
- [26] E. Hewitt and K. Stromberg, *Real and abstract analysis*, Springer New York, 1965.
- [27] J. Journé, *Calderón-Zygmund Operators, Pseudo differential operators and the Cauchy integral of Calderón*, Lecture Notes in Mathematics, 994. Springer-Verlag, Berlin, 1983.
- [28] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1980.
- [29] J. Kigami, *Analysis on Fractals*, Cambridge university press, 2001.
- [30] V. Kokilashvili, *Weighted estimates for classical integral operators. Nonlinear analysis, function spaces and applications*, **4** (Roudnice nad Labem, 1990), 86–103, Teubner-Texte Math., 119, Teubner, Leipzig, 1990.
- [31] V. Kokilashvili, M. Krbec, *Weighted inequalities in Lorentz and Orlicz spaces*, World Scientific Publishing Co., Inc., River Edge, NJ, 1991.
- [32] *Wavelets and Their Applications*, M. Krishna, R. Radha and S. Thangavelu, editors, Allied Publishers, 99–140, 2003.
- [33] S. G. Krein, Ju. I. Petunin, E. M. Semënov, *Interpolation of linear operators*, Translated from the Russian by J. SzHucs. Translations of Mathematical Monographs, **54**, American Mathematical Society, Providence, R.I., 1982.
- [34] Y. Katznelson, *An introduction to HARMONIC ANALYSIS*, Cambridge University Press, 2004.

- [35] M. A. Krasnosel'skiĭ and J. B. Rutickiĭ, Convex functions and Orlicz spaces, Translated from the first Russian edition by Leo F. Boron. P. Noordhoff Ltd., Groningen 1961.
- [36] S. Krantz, Function theory of several complex variables, Reprint of the 1992 edition, AMS Chelsea Publishing, Providence, RI, 2001.
- [37] R. Larsen, An Introduction to the Theory of Multipliers, Springer-Verlag, New York-Heidelberg-Berlin, 1971.
- [38] A. Lunardi, Analytic semigroups and optimal regularity in parabolic problems, Progress in Nonlinear Differential Equations and their Applications, **16**. Birkhäuser Verlag, Basel, 1995.
- [39] W. Luxenberg, Banach function spaces, Assen, 1955.
- [40] P. Mattila, Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability. Cambridge Studies in Advanced Mathematics, **44**, 1995.
- [41] V. G. Maz'ja, Sobolev spaces, Springer-Verlag, Berlin Heidelberg, New York Tokyo.
- [42] R. Megginson, A introduction to Banach space theory. Graduate Texts in Mathematics, 183. Springer-Verlag, New York, 1998.
- [43] Y. Meyer, Wavelets and Operators, Cambridge University Press, Cambridge 1992.
- [44] Y. Mizuta, Potential theory in Euclidean spaces, Gakkotosho, Tokyo, 1996 (in Japanese).
- [45] H. Nakano, Modulated Semi-Ordered Linear Spaces, Maruzen Co., Ltd., Tokyo, 1950. i+288 pp.
- [46] H. Nakano, Topology of linear topological spaces, Maruzen Co., Ltd., Tokyo, 1951, viii+281 pp.
- [47] S. M. Nikol'skij, Approximation of Functions of Several Variables and Imbedding Theorems. Second ed. (Russian). Moskva: Nauka 1977. (English translation of the first edition: Berlin, Heidelberg, New York:Springer-Verlag, 1975).
- [48] M. A. Pinsky, Introduction to Fourier analysis and wavelets, Brooks/Cole, 2002.
- [49] J. Peetre, New thoughts on Besov spaces. Duke University Mathematics Series, No. 1. Mathematics Department, Duke University, Durham, N. C., 1976.
- [50] T. Ransford, Potential theory in the complex plane, London Mathematical Society Student Texts, **28** (1995).
- [51] M. M. Rao and Z. D. Ren, Theory of Orlicz Spaces, Dekker 1991.
- [52] W. Rudin, Functional analysis, *2nd ed.*, McGraw Hill, 1991.
- [53] S. Saks, Theory of the Integral, Warszawa, 1937.
- [54] H. J. Schmeisser and H. Triebel, Topics in Fourier analysis and function space, Chichester, Wiley, 1987.
- [55] L. Schwartz, Théorie des Distributions, Hermann Press, (2 vols), 1951.
- [56] T. Sobukawa, On interpolation spaces, in preparation.
- [57] E. M. Stein, Singular integral and differential property of functions, Princeton: Princeton University Press, 1970.
- [58] E. M. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton Univ. Press, 1993.
- [59] J. O. Strömberg and A. Torchinsky, Weighted Hardy spaces, Lecture Notes in Math., **1381**, Springer-Verlag, Berlin/NewYork, 1989.
- [60] I. Tamura, Topology, Iwanami (in Japanese).
- [61] H. Tanabe, Functional analysis, Jikkyo.
- [62] H. Triebel, Spaces of Besov-Hardy-Sobolev type. Teubner-Texte zur Mathematik. With German, French and Russian summaries. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1978.
- [63] H. Triebel, Theory of Function Spaces, Birkhäuser Basel, 1983.
- [64] H. Triebel, Theory of Function Spaces II, Birkhäuser Basel, 1992.
- [65] H. Triebel, Fractal and Spectra, Birkhäuser Basel, 1997.
- [66] H. Triebel, The Structure of Functions, Birkhäuser Basel, 2000.
- [67] A. Uchiyama, Hardy spaces on the Euclidean Space, Springer-Verlag, Tokyo, 2001.
- [68] P. Wojtaszczyk, A Mathematical Introduction to Wavelets, Cambridge University Press, Cambridge (1997).
- [69] K. Yajima, Lebesgue integral and functional analysis, Asakura shoten, 2002 (in Japanese).
- [70] N. Young, An introduction to Hilbert space, Cambridge Mathematical Textbooks.
- [71] K. Yoshida, Functional analysis.
- [72] A. H. Zemanian, Distribution Theory and Transform Analysis, McGraw-Hill, New York, (1965).
- [73] A. Zygmund, Trigonometric Series, Cambridge, 1959.

## RESEARCH PAPERS

- [74] D. Adams, A note on Riesz potentials, Duke Math. J. **42** (1975), 765–778.
- [75] D. Adams, Choquet integrals in potential theory, Publ. Mat. J. **42** (1998), no 1, 3–66.
- [76] H. Aïmar, A. Bernardis and F. Martín-Reyes, Multiresolution approximations and wavelet bases of weighted  $L^p$  spaces, J. Fourier Anal. Appl. **9** (2003), 497–510.
- [77] L. Alaoglu, Weak topologies of normed linear spaces, Ann. of Math. (2) **41** (1940), 252–267.

- [78] J. M. Aldaz, Remarks on the Hardy-Littlewood maximal function, *Proceedings of the Royal Society of Edinburgh*, **128 A** (1998), 1–9.
- [79] J. Alvarez, M. Guzmán-Partida and J. Lakey, Spaces of bounded  $\lambda$ -central mean oscillation, Morrey spaces, and  $\lambda$ -central Carleson measures, *Collect. Math.*, **51** (2000), 1–47.
- [80] H. Amann, On the strong solvability of the Navier-Stokes equations, *J. Math. fluid Mech.*, **2** (2000), 16–98.
- [81] Ambrose, Measures on locally compact topological groups, *Ann Arbor*, (1946).
- [82] P. D’Ancona and V. Pierfelice, On the wave equation with a large rough potential, *J. Funct. Anal.* **227** (2005), no. 1, 30–77.
- [83] T. Aoki, Locally bounded linear topological space, *Proc. Imp. Acad. Tokyo* **18** (1942) 588–594.
- [84] N. Aronszajn, Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, **68** (1950), 337–404.
- [85] E. W. Aslaksen and J. R. Klauder, Continuous representation theory using the affine group. *J. Math. Phys.* **10** (1969), 2267–2275.
- [86] A. Baernstein and E.T. Sawyer, Embedding and multiplier theorems for  $H^p(\mathbb{R}^n)$ , *Mem. Amer. Math. Soc.* **318** (1985).
- [87] L. R. Baire, Sur les fonctions de variables réelles, *Ann. Mat. Pura. Appl. (3)* **3** (1899), 1–122.
- [88] S. Banach, Sur les fonctionelles linéaires, *Studia Math.*, **1** (1929), 211–216, 223–239.
- [89] Théorie des opérations Linéaires, *Monografie Matematyczne*, **1**, Warsaw, 1932.
- [90] N. K. Bary and S. B. Stechkin, Best approximations and differential properties of two conjugate functions (in Russian), *Proceedings of Moscow Math. Soc.*, **5** (1956), 483–522.
- [91] G. Battle, A block spin construction of ondelettes. I. Lemarié functions, *Comm. Math. Phys.* **110** (1987), no. 4, 601–615.
- [92] W. Baoxiang, Z. Lifeng, and G. Boling, Isometric decomposition operators, function spaces  $E_{p,q}^\lambda$  and applications to nonlinear evolution equations, *J. Funct. Anal.*, **233**(1), 1–39, (2006).
- [93] R. Beals, A general calculus of pseudodifferential operators, *Duke Math. J.* **42** (1975), 1–42.
- [94] R. Beals and C. Fefferman, On local solvability of linear partial differential equations, *Ann. of Math. (2)* **97** (1973), 482–498.
- [95] A. Benedek, A. Calderón and R. Panzone, Convolution operators on Banach space valued functions, *Proc. Nat. Acad. Sci. U. S. A.* **48** 1962 356–365.
- [96] O. V. Besov, Investigation of a class of function spaces in connection with imbedding and extension theorems, (Russian) *Trudy. Mat. Inst. Steklov.*, **60**, 1961 42–81.
- [97] A. Beurling, Construction and analysis of some convolution algebra, *Ann. Inst. Fourier*, **14** (1964), 1–32.
- [98] R. Beylkin, R. Coifman and V. Rokhlin, Fast wavelet transforms and numerical algorithms I, *Communications on Pura and Applied Mathematics* **44** (1991), 141–183.
- [99] O. Blasco, A. Ruiz and L. Vega, Non-interpolation in Morrey-Campanato and block spaces. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **28** (1999), no. 1, 31–40.
- [100] S. Bloom and R. Kerman, Weighted Orlicz space integral inequalities for the Hardy-Littlewood maximal operator, *Studia Math.* **110** (1994), no. 2, 149–167.
- [101] J. Boman, Supremum norm estimates for partial derivatives of functions of several real variables, *Illinois J. Math.* **16** (1972), 203–216.
- [102] M. Bownik, Boundedness of operators on Hardy spaces via atomic decompositions, *Proc. Amer. Math. Soc.*, **133** (2008), no. 12, 3535–3542.
- [103] M. Bownik and K. Ho, Atomic and molecular decompositions of anisotropic Triebel-Lizorkin spaces, *Trans. Amer. Math. Soc.* **358** (2006), no. 4, 1469–1510.
- [104] D. L. Burkholder, Martingale transforms, *Ann. Math. Statist.* **37** (1966), 1494–1504.
- [105] D. L. Burkholder, Distribution function inequalities for martingales, *Ann. Probability* **1** (1973), 19–42.
- [106] D. L. Burkholder and R. F. Gundy, Extrapolation and interpolation of quasi-linear operators on martingales, *Acta Math.* **124**, no. 1, (1970), 249–304.
- [107] Peter J. Burt and Edward H. Adelson, The Laplacian pyramid as a compact image code, *IEEE Trans. Commun.*, **31** (1983), 532–540.
- [108] P. L. Butzer, P. J. S. G. Ferreira, J. R. Higgins, S. Saitoh, G. Schmeisser and R. L. Steins, E. T. Whittaker, K. Ogura, sampling, their interconnections and followers, *J. Fourier Analysis and Applications*, **17** (2011), 320–354.
- [109] L. Caffarelli, Elliptic second order equations, *Rend. Sem. Mat. Fis. Milano*, **58** (1990), 253–284.
- [110] A. P. Calderón, On the theorems of M. Riesz and Zygmund, *Proc. Amer. Math. Soc.* **1** (1950), 533–535.
- [111] A. P. Calderón, Intermediate spaces and interpolation, *Studia Math.* **1** (1963), 31–34.
- [112] A. P. Calderón, Intermediate spaces and interpolation, the complex method, *Studia Math.* **14** (1964), 113–190.
- [113] A. P. Calderón, Spaces between  $L^1$  and  $L^\infty$  and the theorem of Marcinkiewicz, *Studia Math.* **26** (1966), 273–299.

- [114] A. P. Calderón, Commutators of singular integral operators, Proc. Nat. Acad. Sci. U. S. A. **53**, (1965) 1092–1099.
- [115] A. P. Calderón and R. Vaillancourt, On the boundedness of pseudo-differential operators, J. Math. Soc. Japan **23** (1971) 374–378.
- [116] A. P. Calderón and A. Zygmund, A note on the interpolation of sublinear operations, Amer. J. Math., **78** (1956), 282–288.
- [117] A. P. Calderón and A. Zygmund, On singular integrals, Amer. J. Math., **78** (1956), 289–309.
- [118] A. P. Calderón and A. Zygmund, Algebras of Certain Singular Operators, Amer. J. Math., **78** (1956), 310–320.
- [119] A. P. Calderón and A. Zygmund, Singular integral operators and differential equations, Amer. J. Math., **79** (1957), 901–921.
- [120] A. P. Calderón and A. Zygmund, Local properties of solutions of elliptic partial differential equations. Studia Math. **20** (1961), 171–225.
- [121] S. Campanato, Caratterizzazione delle tracce di funzioni appartenenti ad una classe di Morrey insieme con le loro derivate prime, (Italian) Ann. Scuola Norm. Sup. Pisa (3) **15** (1961) 263–281.
- [122] S. Campanato, Proprietà di hölderianità di alcune classi di funzioni, (Italian) Ann. Scuola Norm. Sup. Pisa (3) **17** 1963 175–188.
- [123] S. Campanato, Proprietà di inclusione per spazi di Morrey, (Italian) Ricerche Mat. **12** (1963) 67–86.
- [124] S. Campanato, Proprietà di una famiglia di spazi funzionali, (Italian) Ann. Scuola Norm. Sup. Pisa (3) **18** (1964) 137–160.
- [125] S. Campanato, Proprietà di una famiglia di spazi funzionali, Ann. Scuola Norm. Sup. Pisa Cl.Sci. **17** (1964), 137–160.
- [126] S. Campanato, Equazioni ellittiche del II deg ordine spazi  $\mathcal{L}^{(2,\lambda)}$ , (Italian) Ann. Mat. Pura Appl. (4) **69** (1965) 321–381.
- [127] S. Chanillo, A note on commutators, Indiana Univ. Math. J. **31** (1982), 7–16.
- [128] C. Carathéodry, Über das lineare Mass von Punktmenge-eine Verallgemeinerung das Längenbegriffs, Nach. Ges. Wiss. Göttingen, 406–426.
- [129] L. Carleson, An interpolation problem for bounded analytic functions, Amer. J. Math. **80** (1958), 921–930.
- [130] L. Carleson, Interpolation by bounded analytic functions and the corona problem, Ann. of Math. **76** (1962), 547–559.
- [131] L. Carleson, On convergence and growth of partial sumas of Fourier series. Acta Math. **116** (1966), 135–157.
- [132] H. Carleson, A new proof of the Hardy-Littlewood maximal theorem, London Math. Soc. **16** (1984), 595–596.
- [133] F. Chiarenza and M. Frasca, Morrey spaces and Hardy-Littlewood maximal function, Rend. Mat., **7** (1987), 273–279.
- [134] F. Chiarenza, M. Frasca and P. Longo,  $W^{2,p}$ -solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients, Transaction of the American Mathematical Society, **336**, no. 2, April 1993.
- [135] Y. Chen and K. Lau, Some new classes of Hardy spaces, J. Funct. Anal., **84** (1989), 255–278.
- [136] W. Chen and E. Sawyer, A note on commutators of fractional integrals with RBMO( $\mu$ ) functions, Illinois J. Math. **46** (2002), no. 4, 1287–1298.
- [137] M. Christ, Lectures on Singular Integral Operators, CBMS Regional Conf. Series in Math. **77**, Amer. Math. Soc, Providence, RI, 1990.
- [138] M. Christ and R. Fefferman, A note on weighted norm inequalities for the Hardy-Littlewood maximal operator, Proc. Amer. Math. Soc. **87** (1983), 447–448.
- [139] A. Cianchi, Strong and weak type inequalities for some classical operators in Orlicz spaces, J. London Math. Soc. (2) **60** (1999), no. 1, 187–202.
- [140] A. Cohen, Ondelettes, analysis multirésolutions et filtres miroirs en quadrature [Wavelets, multiresolution analysis, and quadrature mirror filters], Ann. Inst. H. Poincaré, Anal. Non Linéaire, **7** (1990), 439–459. Translated by Robert D. Ryan.
- [141] R. Coifman, A real variable characterization of  $H^p$ , Studia Math., **51** (1974), 269–274.
- [142] R. R. Coifman and C. Fefferman, Weighted norm inequalities for maximal and singular integrals, Studia Math. **51** (1974), 241–250.
- [143] R. R. Coifman and Y. Meyer, Au delà des opérateurs pseudo-différentiels, (French) (Beyond pseudodifferential operators) With an English summary, Astérisque, 57. Société Mathématique de France, Paris, 1978.
- [144] R. R. Coifman and R. Rochberg, Representation theorems for holomorphic and harmonic functions in  $L^p$ . Representation theorems for Hardy spaces, 11–66, Astérisque, **77**, Soc. Math. France, Paris, 1980.

- [145] R. R. Coifman and R. Rochberg, Another characterization of BMO, Proc. Amer. Math. Soc. **79** (1980), 249–254.
- [146] R. Coifman, R. Rochberg, and G. Weiss, Factorization theorems for Hardy spaces in several variables, Ann. of Math. (2) **103** (1976), no. 3, 611–635.
- [147] R. R. Coifman and G. Weiss, Analyse harmonique non-commutative sur certains espaces homogènes, Lecture Notes in Mathematics, vol. **242**, Springer, Berlin, 1971.
- [148] R. R. Coifman and G. Weiss, Extensions of Hardy spaces and their use in analysis, Bull. Amer. Math. Soc., **83**, (1977), 569–645.
- [149] E. Cordero, S. Pilipović, L. Rodino and N. Teofanov, Quasianalytic Gelfand-Shilov spaces and localization operators, to appear in Rocky Mountain J. Math..
- [150] A. Córdoba and C. Fefferman, A weighted norm inequality for singular integrals, Studia Math. **57** (1976), 97–101.
- [151] M. Cotlar, A combinatorial inequality and its applications to  $L^2$  spaces, Rev. Mat. Cuyana, **1** (1955), 41–55.
- [152] Extrapolation with weights, rearrangement-invariant function spaces, modular inequalities and applications to singular integrals, Advances in Mathematics **203**, (2006), 256–318.
- [153] R. E. Crochierem, S. A. Webber and J. L. Flanagan, Digital coding of speech in sub-bands, Bell System Technical J., **55** (1976), 1069–1085.
- [154] CFN-2003 D. Cruz-Uribe, A. Fiorenza and C. J. Neubebauer, THE MAXIMAL FUNCTION ON VARIABLE  $L^p$  SPACES, Annal. Acad. Sci. Fenn. Math, **Vol 28**, 2003, 223-238.
- [155] CFN-2004 D. Cruz-Uribe, A. Fiorenza and C. J. Neubebauer, CORRECTIONS TO “THE MAXIMAL FUNCTION ON VARIABLE  $L^p$  SPACES”, Annal. Acad. Sci. Fenn. Math, **Vol 29**, 2004, 247-249.
- [156] D. Cruz-Uribe, J. M. Martell and C. Pérez, Extrapolation from  $A_\infty$ -weights and applications, J. Funct. Anal., **213** (2004), 412-439.
- [157] A. David, X. Duong and A. McIntosh, Operator theory and harmonic analysis, Instructional Workshop on Analysis and Geometry, Part II (Canberra 1995), 77-136, Proc. Centre Math. Appl. Austral. Nat. Univ., **34**, Austral. Nat. Univ., Canberra, 1996.
- [158] G. David and J. L. Journé, A boundedness criterion for generalized Calderón-Zygmund operators, Ann. of Math. **120** (1984), 371–397.
- [159] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. **41** (1988), 909-996.
- [160] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inform. Theory, **39** (1990), 961–1005.
- [161] I. Daubechies, A. Grossmann and Y. Meyer, Painless nonorthogonal expansions, J. Math. Physics, **27** (1986), 1271–1283.
- [162] D. Deng, M. Xu and L. Yan, Wavelet characterization of weighted Triebel-Lizorkin spaces, Approx. Theory Appl. **18** (2002), 76–92.
- [163] L. Diening, Maximal functions in generalized  $L^{p(\cdot)}$  spaces, Math. Inequal. Appl. **7** (2004), 245–253.
- [164] L. Diening, Riesz potentials and Sobolev embeddings on generalized Lebesgue and Sobolev spaces  $L^{p(\cdot)}$  and  $W^{k,p(\cdot)}$ , Math. Nachr. **268** (2004), 31–43.
- [165] M. Dimassi and J. Sjöstrand, Spectral asymptotics in the semi-classical limit, volume 268 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, (1999).
- [166] L. Doob, Kolmogorov’s early work on convergence theory and foundations, Ann. Probab. **17** (1989), no. 3, 815–821.
- [167] R. J. Duffin and A. C. Schaeffer, A class of nonharmonic Fourier series, Trans. Amer. Math. Soc., **72**, (1952). 341–366.
- [168] J. Duoandikoetxea, C. Michael, J. L. Rubio de Francia, Maximal operators related to the Radon transform and the Calderón-Zygmund method of rotations, Duke Math. J. **53** (1986), no 1, 189–209.
- [169] J. Duoandikoetxea and J. L. Rubio de Francia, Maximal and singular integral operators via Fourier transform estimates, Invent. Math. **84**, (1986), 541–561.
- [170] J. J. Duistermaat and L. Hörmander, Fourier integral operators. II, Acta Math. **128** (1972), 183–269.
- [171] Ju. V. Egorov and P. R. Popivanov, A certain class of pseudodifferential equations of principal type having no solutions, (Russian) Dokl. Akad. Nauk SSSR **209** (1973), 775–777..
- [172] D. Esteban and C. Galand, Application of quadrature mirror filters to split-band voice coding schemes, ICASSP ’77, IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing, **2**, April 1977, 191–195.
- [173] Eridani and H. Gunawan, On generalized fractional integrals, J. Indonesian Math. Soc. (MIHMI) **8**, no 3, (2002), 25–28.
- [174] On generalized fractional integral operators, Sci. Math. Japan **60**, no 3, (2004), 539–550.
- [175] G.Di Fazio and A. M. Ragusa, Commutators and Morrey spaces, Boll. Un. Mat. Ital. **A (7) 5** (1991), no. 3, 323-332.

- [176] G. Di Fazio, D. K. Palagachev and A. Ragusa, Global Morrey regularity of strong solutions to the Dirichlet problem for elliptic equations with discontinuous coefficients, *J. Funct. Anal.*, **166** (1999), no 2, 179–196.
- [177] C. Fefferman, Characterizations of bounded mean oscillation, *Bull. Amer. Math. Soc.* **77** (1971) 587–588.
- [178] C. Fefferman, The multiplier problem for the ball, *Ann. of Math. (2)* **94** (1971), 330–336.
- [179] C. Fefferman, Pointwise convergence of Fourier series, *Ann. Math.* **98** (1973), 551–571.
- [180] C. Fefferman, The uncertainty principle, *Bull. Amer. Math. Soc.*, **9** (1983), 129–206.
- [181] C. Fefferman and D. H. Phong, On positivity of pseudo-differential operators, *Proc. Nat. Acad. Sci. U. S. A.* **75** (1978), no. 10, 4673–4674.
- [182] C. Fefferman and E. M. Stein, Some maximal inequalities, *Amer. J. Math.* **93** (1971), 107–115.
- [183] C. Fefferman and E. M. Stein,  $H^p$  spaces of several variables, *Acta Math.* **129** (1972), 137–193.
- [184] H. G. Feichtinger, Modulation spaces on locally compact abelian groups, Technical report, University of Vienna.
- [185] H. G. Feichtinger, An elementary approach to Wiener’s third Tauberian theorem for the Euclidean  $n$ -spaces, *Proceedings of Conference at Cortona 1984, Symposia Mathematica*, **29**, Academic Press, New York, 1987, 267–301.
- [186] H. G. Feichtinger, Atomic characterization of modulation spaces through Gabor-type representation, In *Proc. Conf. Constructive Function Theory*, Edmonton, July (1989), 113–126.
- [187] H. G. Feichtinger and F. Weisz, Herz spaces and summability of Fourier transforms, *Math. Nachr.*, **281**, (2008), 309–324.
- [188] H. G. Feichtinger and K. H. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions I, *J. Funct. Anal.*, **86** (1989), 307–340.
- [189] H. Feichtinger and K. Gröchenig, Gabor wavelets and the Heisenberg group: Gabor expansions and short time Fourier transform from the group theoretical point of view. In Charles K. Chui, editor, *Wavelets: A tutorial in theory and applications*, 359–398, Academic Press, Boston, MA, (1992).
- [190] H. Feichtinger and K. Gröchenig, Gabor frames and time-frequency analysis of distributions. *J. Funct. Anal.*, **146** (2), 464–495 (1997).
- [191] L. Fejér, Untersuchungen über Fouriersche Reihen, *Math. Ann.* **58** (1903), no. 1-2, 51–69.
- [192] H. Finbarr, Harmonic analysis on amalgams, *J. London Math. Soc. (2)*, **10** (1975), 295–305.
- [193] G. B. Folland and E. M. Stein, Hardy spaces on homogeneous groups. *Mathematical Notes*, **28**, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.
- [194] P. Franklin, A set of continuous orthogonal functions, *Math. Ann.* **100** (1928), no. 1, 522–529.
- [195] M. Frazier and B. Jawerth, A discrete transform and decompositions of distribution spaces, *Indiana University Math. J.* **34** (1985), no. 4, 777–799.
- [196] M. Frazier and B. Jawerth, A discrete transform and decompositions of distribution spaces. *J. Funct. Anal.* **93** (1990), no. 1, 34–170.
- [197] G. Fubini, Sugli integrali multipli, *Atti della R. Accad. dei Lincei, Rendiconti* **5** (1907), 608–614.
- [198] N. Fujii, A proof of the Fefferman-Stein-Strömberg inequality for the sharp maximal functions, *Proc. Amer. Math. Soc.* **106** (1989), no.2, 371–377.
- [199] T. Futamura, Y. Mizuta and T. Shimomura, Sobolev embeddings for variable exponent Riesz potentials on metric spaces, *Math. Nachr.* **279** (2006), no. 13-14, 1463–1473.
- [200] E. Gagliardo, Interpolation d’espaces de Banach et applications, I–III, *Compt. Rend. Acad. Sci. Paris*, **248** (1959), 1912–1914, 3388–3390, 3517–3518.
- [201] E. Gagliardo, Interpolazione di spazi di Banach e applicazioni, *Edizioni Sci.*, Geneva 1959.
- [202] E. Gagliardo, Interpolazione di spazi di Banach e applicazioni, *Ricerche Mat.* **9** (1960), 58–61.
- [203] Y. V. Galperin and S. Samarah, Time-frequency analysis on modulation spaces  $M_m^{pq}$ ,  $0 < p, q \leq \infty$ , *Appl. Comput. Harmon. Anal.*, **16** (2004), 1–18.
- [204] J. García-Cuerva, Hardy spaces and Beurling algebras, *J. London Math. Soc. (2)*, **39** (1989), 499–513.
- [205] J. García-Cuerva and E. Gatto, Boundedness properties of fractional integral operators associated to non-doubling measures, *Studia Math.* **162** (2004), no. 3, 245–261.
- [206] J. García-Cuerva and M. J. L. Herrero, A theory of Hardy spaces associated to the Herz spaces. *Proc. London Math. Soc. (3)* **69** (1994), no. 3, 605–628.
- [207] J. García-Cuerva and J. M. Martell, Weighted inequalities and vector-valued Calderón-Zygmund operators on nonhomogeneous spaces, *Publ. Mat.* **44** (2000), no. 2, 613–640.
- [208] J. B. Garnett and P. W. Jones, BMO from dyadic BMO, *Pacific J. Math.* **99** (1982), no. 2, 351–371.
- [209] J. García-Cuerva and Martell, Wavelet characterization of weighted spaces, *J. Geom. Anal.* **11** (2001), 241–264.
- [210] B. Gelbaum, Expansions in Banach spaces. *Duke Math. J.* **17**, (1950), 187–196.
- [211] L. Grafakos and S. M. Smith, Best constants for uncentered maximal operators, *Bull. London Math. Soc.* **29** (1997), 60–64.
- [212] L. Grafakos, T. Tao and E. Terwileger,  $L^p$  bounds for a maximal dyadic sum operator, *Math. Z.* **246** (2004), 321–337.

- [213] L. Grafakos and R. H. Torres, Pseudodifferential Operators with Homogeneous Symbols, *Michigan Math. J.* (1999) **46** (2), 261-269.
- [214] L. Grafakos, X. Li and D. Yang, Bilinear operators on Herz-type Hardy spaces, *Trans. Amer. Math. Soc.*, **350** (1998), 1249-1275.
- [215] K. Gröchenig, Composition and spectral invariance of pseudodifferential operators on modulation spaces, *J. Anal. Math.*, **98** 65-82, (2006).
- [216] K. Gröchenig, Time-frequency analysis of Sjöstrands class. *Revista Mat. Iberoam.*, **22** (2), 703-724, (2006), arXiv:math.FA/0409280v1.
- [217] K. Gröchenig and Z. Rzeszotnik, Almost diagonalization of pseudodifferential operators, *Ann. Inst. Fourier*, (2008). to appear.
- [218] K. Gröchenig, Analyse multi-échelle et bases d'ondelettes [Multiscale analysis and wavelet bases], *C. R. Acad. Sci. Paris. Série I*, **305**, (1987), 13-17. Translated by Robert D. Ryan.
- [219] K. Gröchenig and C. Heil, Modulation spaces and pseudo-differential operators, *Integral Equations Operator Theory*, **34** (1999), 439-457.
- [220] A. Grossmann and J. P. Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, **15** (1984), no. 4, 723-736.
- [221] A. Grossmann, J. P. Morlet and T. Paul, Transforms associated to square integrable group representations. I. General results, *J. Math. Phys.* **26** (1985), no. 10, 2473-2479.
- [222] H. Gunawan, A note on the generalized fractional integral operators, *J. Indonesian Math. Soc. (MIHMI)* **9**, no. 1, (2003), 39-43.
- [223] H. Gunawan, On weighted estimates for Stein's maximal function, *Bull. Austral. Math. Soc.*, **54** (1996), 35-39.
- [224] J. Gustavsson and J. Peetre, Interpolation of Orlicz spaces, *Studia Math.* **60** no 1, 33-59.
- [225] A. Haar, Zur theorie der orthogonalen Funktionensysteme, *Math. Annalen* **69** (1910), 331-371.
- [226] H. Hahn, Über Folgen linearer Operationen, *Monatsh. Math. Phys.*, **32** (1922), 3-88.
- [227] G. Hardy, and J. Littlewood, A maximal theorem with function-theoretic applications, *Acta Math.* **54** (1930), no. 1, 81-116.
- [228] G. Hardy and J. Littlewood, Some properties of fractional integrals I, *Math. Z.* **27** (1927), 565-606.
- [229] G. Hardy and J. Littlewood, Some properties of fractional integrals II, *Math. Z.* **34** (1932), 403-439.
- [230] M. Hasumi, Note on the  $n$ -dimension Tempered Ultradistributions, *Tohoku Math. J.* **13** (1961), 94 -104.
- [231] M. Hata, On the structure of Self-Similar Sets, *Japan J. Appl. Math.* **2** (1985), 381-414.
- [232] F. Hausdorff, Dimension und äussers Mass, *Math. Annalen*, **79** (1919), 157-179.
- [233] L. I. Hedberg, On certain convolution inequalities, *Proc. Amer. Math. Soc.*, **36**, No 2 (1972), 505-510.
- [234] E. Hernández, G. Weiss and D. Yang, The  $\varphi$ -transform and wavelet characterizations of Herz-type spaces, *Collect. Math.* **47** (1996), 285-320.
- [235] C. Herz, Lipschitz spaces and Bernstein's theorem on absolutely convergent Fourier transforms, *J. Math. Mech.*, **18** (1968), 283-324.
- [236] I. Hirschman, A convexity theorem for certain groups of transformations, *Journal d'Analyse* **2** (1952), 209-218.
- [237] O. Hölder, Über einen Mittelwerthssatz, *Göttinger Nachr.* (1889), 38-47.
- [238] L. Hörmander, Estimates for translation invariant operators in  $L^p$  spaces, *Acta Math.* **104** (1960), 93-140.
- [239] L. Hörmander, Pseudo-differential operators, *Comm. Pure Appl. Math.* **18** (1965), 501-517.
- [240] L. Hörmander, Pseudo-differential operators and hypoelliptic equations, *Singular integrals (Proc. Sympos. Pure Math., Vol. X, Chicago, Ill., 1966)*, 138-183. Amer. Math. Soc., Providence, R.I., 1967.
- [241] L. Hörmander, Fourier integral operators. I, *Acta Math.* **127** (1971), 79-183.
- [242] L. Hörmander, The Weyl calculus of pseudodifferential operators, *Comm. Pure Appl. Math.* **32** (1979), no. 3, 360-444.
- [243] T. Holmstedt and J. Peetre On certain functionals arising in the theory of interpolation spaces. *J. Funct. Anal.*, **4** (1969), 88-94.
- [244] G. Hu, Y. Meng and D. Yang, Multilinear commutators for fractional integrals in non-homogeneous spaces, *Publ. Mat.* **48** (2004), no. 2, 335-367.
- [245] R. A. Hunt, On the convergence of Fourier series, in: "1968 Orthogonal Expansions and their Continuous Analogues" (Proc. Conf.,Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 235-255.
- [246] J. E. Hutchinson, Fractals and self similarity, *Indiana Math. Journal* **30** (1981), 713-747.
- [247] G. Hu, Y. Meng and D. Yang, Multilinear commutators of singular integrals with non-doubling measures. *Integral Equations Operator Theory* **51** (2005), no. 2, 235-255.
- [248] R. Hunt, D. Kurtz and C. Neugebauer, A note on the equivalence of  $A_p$  and Sawyers condition for equal weights, *Proc. Conf. on Harmonic Analysis in honour of A. Zygmund*, Wadsworth Math. Ser., vol. 1, 156-158.
- [249] R. Hunt, B. Muckenhoupt, R. Wheeden, Weighted norm inequalities for the conjugate function and Hilbert transform, *Trans. Amer. Math. Soc.* **176** (1973), 227-251.
- [250] R. Hunt, On  $L(p, q)$  spaces, *Enseign. Math.* **12** (1966), 249-276.



- [251] R. Hunt, On the convergence of Fourier series, Orthogonal Expansions and Their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., (1968), 235–255.
- [252] Wik Ingemar, On Muckenhoupt's classes of weight functions. *Studia Math.* **94** (1989), no. 3, 245–255.
- [253] V. Ja. Ivriřand V. M. Petkov, Necessary conditions for the correctness of the Cauchy problem for non-strictly hyperbolic equations. (Russian), Collection of articles dedicated to the memory of Ivan Georgievic Petrovskii(1901–1973), III. *Uspehi Mat. Nauk* **29** (1974), no. 5(179), 3–70.
- [254] M. Izuki, The characterizations of weighted Sobolev spaces by wavelets and scaling functions, in preparation.
- [255] R. C. James, Bases in Banach spaces, *Amer. Math. Monthly* **89** (1982), no. 9, 625–640.
- [256] S. Janson, Mean oscillation and commutators of singular integral operators, *Ark. Math.* **16** (1978), 263–270.
- [257] B. Jawerth, Weighted inequalities for maximal operators: linearization, localization and factorization. *Amer. J. Math.* **108** (1986), no. 2, 361–414.
- [258] J. L. W. V. Jensen, Sur les fonctions convexes et lers inéqualités entre les valeurs moyennes, *Acta Math.* **30** (1906), 175–193.
- [259] F. John, 'Quasi-isometric mappings,' *Seminari 1962-1963 di Analisi, Algebra, Geometria e Topologia, Ist Nazarene Alta Matematica 2* (Edizioni Cremonese, Rome, 1965), 462–473.
- [260] J. John and F. Fournier, Vancouver, On the Hausdorff-Young Theorem for amalgams, *Monatsh. Math.* **95** (1983), 117–135.
- [261] F. John and L Nirenberg, On function of bounded mean oscillation, *Comm. Pure Appl. Math.* **14** (1961), 415–426.
- [262] P. Jordan and J. von Neumann, On inner products in linear metric spaces, *Ann. of Math.* **36**(1935), 719–723.
- [263] O. Jorsboe and L. Mejlbro, The Carleson-Hunt Theorem on Fourier Series, *Lecture Notes in Math.* **911**.
- [264] S. Kakutani, Two fixed-point theorems concering bicomact convex sets, *Proc. Imp. Acad. Tokyo* **14** (1938), 242–245.
- [265] S. Karlin, Bases in Banach spaces, *Duke Math. J.* **15** (1948), 971–985.
- [266] G. Karadžov, About the interpolation method of means for quasi-normed spaces, *Blgar. Acad. Naut. Math. Inst.* **15**, 191–207.
- [267] R. Kerman and A. Torchinsky, Integral inequalities with weights for the Hardy maximal function, *Studia Math.* **71** (1982), no. 3, 277–284.
- [268] H. Kita, On maximal functions in Orlicz spaces, *Proc. Amer. Math. Soc.* **124** (1996), 3019–3025.
- [269] H. Kita, On Hardy-Littlewood maximal functions in Orlicz spaces, *Math. Nachr.* **183** (1997), 135–155.
- [270] On regularly convex stes in the space conjugate to a Banach space, *Ann. of Math.*, **41**, 556–583.
- [271] A. N. Kolmogorov, Sur les fonctions harmoniques conjuguées et les séries de Fourier, *Fund. Math.* **7** (1925), 23–28.
- [272] A. N. Kolmogorov, Zur Grössenordnung des Restgliededes Fourierscher Reihen differenzierbarer Funktionen,(German) *Ann. of Math. (2)* **36** (1935), no. 2, 521–526.
- [273] M. A. Kransnoselskii and Ya. B. Rutitskii, Convex functions and Orlicz spaces, *Gos. Izdat. Fiz.-mat. Lit.*, Moscow 1958 (in Russian).
- [274] W. T. Kraynek, Interpolation of sublinear operators in generalized Orlicz and Hardy-Orlicz spaces, *Studia Math.* **43** (1972), 93–123.
- [275] S. G. Kreuin, On an interpolation theorem in operator theory, *Dokl. Akad. Nauk SSSR* **130** 491–494 (Russian); translated as *Soviet Math. Dokl.* **1** (1960), 61–64.
- [276] M. Kobayashi, Modulation spaces  $M^{p,q}$  for  $0 < p, q \leq \infty$ , *J. Funct. Spaces Appl.* **4** (2006), no. 3, 329–341.
- [277] M. Kobayashi, Dual of Modulation spaces, *J. Funct. Spaces Appl.* **5** (2007), no. 1, 1–8.
- [278] M. Kobayashi and Y. Sawano, Molecular decomposition of the modulation spaces  $M^{p,q}$  and its application to the pseudo-differential operators, in preparation.
- [279] J. J. Kohn and L. Nirenberg, An algebra of pseudo-differential operators, *Comm. Pure Appl. Math.* **18** (1965), 269–305.
- [280] A. Knapp and E. M. Stein, Intertwining operators for semi-simple groups, *Ann. of Math.* **93** (1971), 489–578.
- [281] V. Kokilashvili and A. Meskhi, Fractional integrals on measure spaces, *Fract. Calc. Appl. Anal.* **4** (2001), no. 1, 1–24.
- [282] V. Kokilashvili and M. Krbec, Weighted inequalities in Lorentz and Orlicz spaces, World Scientific Publishing Co., Inc., River Edge, NJ, 1991.
- [283] On a trace inequality for one-sided potentials and applications to the solvability of nonlinear integral equations, *Georgian Math. J.*, **8** (2001), no. 3, 521–536.
- [284] Y. Komori, A note on modified maximal functions associated with general measures, *Int. J. Pure Appl. Math.* **12** (2004), no. 2, 169–175.

- [285] Y. Komori, Weighted estimates for operators generated by maximal functions on nonhomogeneous space, *Georgian Math. J.* **12** (2005), no.1, 121-130.
- [286] Y. Komori, Calderón-Zygmund operators on the predual of a Morrey space, *Acta Mathematica Sinica, English Series*, **19** (2), 297-302, (2003).
- [287] Weak type estimates for Calderón-Zygmund operators on Herz spaces at critical indexes, *Math. Nachr.*, **259** (2003), 42–50.
- [288] Y. Komori and T. Mizuhara, Factorization of functions in  $H^1(\mathbb{R}^n)$  and generalized Morrey spaces, *Math. Nachr.* **279** (2006), no. 5-6, 619–624.
- [289] H. Kozono and M. Yamazaki, Semilinear heat equations and the Navier-Stokes equations with distributions in new function spaces as initial data, *Comm. Partial Differential Equations* **19** (1994) no. 5–6, 959–1014.
- [290] P. Kree, Sur les multiplicateurs dans  $\mathcal{F}L^p$ , (French) *Ann. Inst. Fourier (Grenoble)* **16** (1966) 31–89.
- [291] P. C. Kunstman and L. Weis, Maximal  $L^p$  regularity for parabolic equations, Fourier multiplier theorems, and  $H^\infty$ -calculus, *Functional Analytic Methods for Evolution Equations*, 65–311, *Lecture Notes in Math.* 1855, Springer Berlin, 2004.
- [292] H. Kunita and S. Watanabe, On square integrable functions, *Nagoya Math. J.*, **30** (1967), 209–245.
- [293] K. Kurata, S. Nishigaki and S. Sugano, Boundedness of integral operators on generalized Morrey spaces and its application to Schrödinger operators, *Proc. Amer. Math. Soc.* **128** no 4 (2000), 1125-1134.
- [294] S. Kuratsubo, On pointwise convergence of Fourier series of indicator function of  $n$  dimensional ball, *Sci. Rep. Hiroasaki Univ.*, **43** (1996), 199–208.
- [295] S. Kuratsubo, On pointwise convergence of Fourier series of radial functions in several variables, *Proc. Amer. Math. Soc.*, **127** (1999), 2987–2994.
- [296] M. Lacey, Carleson’s theorem: proof, complements, variations, *Publ. Mat.* **48** (2004), 251–307.
- [297] M. T. Lacey, Commutators with Reisz potentials in one and several parameters, *Hokkaido Math. J.* **36** (2007), no. 1, 175–191.
- [298] M. Lacey and C. Thiele, A proof of boundedness of the Carleson operator, *Math. Res. Lett.* **7** (2000), 361-370.
- [299] W. M. Lawton, Tight frames of compactly supported affine wavelets, *J. Math. Phys.*, **31** (1990), 1898–1901.
- [300] H. Lebesgue, *Leçons sur l’intégration et la recherche des fonctions primitives*, Gauthier-Villars, Paris (1904).
- [301] H. Lebesgue, Sur l’intégration des fonctions discontinues, (French) *Ann. Sci. École Norm. Sup.* (3) **27** (1910), 361–450.
- [302] H. Lebesgue, *La mesure des grandeurs*, (French) Nouveau tirage, Librairie Scientifique et Technique Albert Blanchard, Paris, 1975.
- [303] P. G. Lemarié-Rieusset, Ondelettes et poids de Muckenhoupt, *Studia Math.* **108** (1994), 127–147.
- [304] P. G. Lemarié-Rieusset and Malgouyres, Support des fonctions de base dans une analyse multi-résolution, *C. R. Acad. Sci. Paris* **313** (1991), 377-380.
- [305] P. G. Lemarié-Rieusset and Meyer, Ondelettes et bases Hilbertiennes, *Revista Matematica Iberoamericana* **2** (1986), 1–18.
- [306] A. K. Lerner, An elementary approach to several results on the Hardy-Littlewood maximal operator, *Proc. Amer. Math. Soc.* **136** (2008) no. 8, 2829–2833.
- [307] A. K. Lerner, Weighted rearrangement inequalities for local sharp maximal functions, *Transactions of the American Mathematical Society*, **357**, no. 6, 2445-2465.
- [308] A. K. Lerner, S. Ombrosi and C. Pérez,  $A_1$  bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden, *Math. Res. Lett.* **16** (2009), no. 1, 149-156.
- [309] A. K. Lerner, A pointwise estimate for the local sharp maximal function with applications to singular integrals, *Bull. Lond. Math. Soc.* **42** (2010), no. 5, 843–856.
- [310] X. Li and D. Yang, Boundedness of some sublinear operators on Herz spaces, *Illinois J. Math.*, **40** (1996), 484–501.
- [311] J. L. Lions, Théorèmes de trace et d’interpolation. II. (French), *Ann. Scuola Norm. Sup. Pisa* (3) **14** (1960), 317–331.
- [312] J. L. Lions, Une construction d’espaces d’interpolation, *Comptes Rendus Acad. Sc. Paris* **83** (1960), 1853–1855.
- [313] J. L. Lions, Théorèmes de trace et d’interpolation. V. (French), *An. Acad. Brasil. Ci.* **35** (1963), 1–10.
- [314] J. L. Lions, Théorèmes de trace et d’interpolation. IV. (French), *Math. Ann.* **151** (1963), 42–56.
- [315] J. L. Lions, Théorèmes de trace et d’interpolation. III. (French), *J. Math. Pures Appl.* (9) **42** (1963), 195–203.
- [316] J. L. Lions and J. Peetre, Sur une classe d’espaces d’interpolation, (French) *Inst. Hautes Études Sci. Publ. Math.*
- [317] R. Long, The spaces generated by blocks, *Scientia Sinica Ser. A* **27** (1984), 16–26.
- [318] L. H. Loomis, A note on the Hilbert transform, *Bull. Amer. Math. Soc.* **52** (1946), 1082–1086.

- [319] G. G. Lorentz, Some new functional spaces, *Ann. Math.*, **51** (1950), 35–55.
- [320] R. A. Macias and C. Segovia, Lipschitz functions on spaces of homogeneous type, *Adv. Math.* **33** (1979), no. 3, 257–270.
- [321] S. Mallat, Multiresolution approximation and wavelet orthonormal bases of  $L_2(\mathbb{R})$ , *Trans. Amer. Math. Soc.* **315** (1989), 69–88.
- [322] J. Mateu, P. Mattila, A. Nicolau and J. Orobitg, BMO for nondoubling measures, *Duke Math. J.* **102** (2000), no. 3, 533–565.
- [323] J. Marcinkiewicz, Sur les séries de Fourier, *F. M.* **27** (1936), 38–69.
- [324] J. Marcinkiewicz, Quelques théorèmes sur les séries de fonctions, *Bull. du Séminaire Math. de l'Université de Wilno*, **1** (1937).
- [325] J. Marcinkiewicz, Sur quelques intégrals du type de Dini, *Ann. de la Soc. Polonaise de Math* **17** (1938), 42–50.
- [326] J. Marcinkiewicz, Sur l'interpolation d'opérateurs, *C. R. Acad. Sci. Paris* **208** (1939), 1272–1273.
- [327] J. Marcinkiewicz, Sur les multiplicateurs des séries de Fourier, *Studia Math.* **8** (1939), 78–91.
- [328] J. Marcinkiewicz, *Collected papers*, Varsorvie, 1964.
- [329] W. Matuszewska, On generalized Orlicz spaces. (Russian summary), *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **8** (1960), 349–353.
- [330] W. Matuszewska and W. Orlicz, On certain properties of  $\varphi$ -functions, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **8** (1960) 439–443.
- [331] A. Mazzucato, Decomposition of Besov-Morrey spaces. Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), 279–294, *Contemp. Math.*, **320**, Amer. Math. Soc., Providence, RI, 2003.
- [332] A. McIntosh, Operators which have an  $H_\infty$  functional calculus, Miniconference on operator theory and partial differential equations (North Ryde, 1986), 210–231, *Proc. Centre Math. Anal. Austral. Nat. Univ.*, **14**, Austral. Nat. Univ., Canberra, 1986.
- [333] A. Mazzucato, Besov-Morrey spaces: function space theory and applications to non-linear PDE, *Trans. Amer. Math. Soc.* **355** (2003), no. 4, 1297–1364.
- [334] H. Minkowski, *Geometrie der Zahlen*. I. Leipzig, B. G. Teubner, 1896.
- [335] S. Meda, P. Sjogren and M. Vallarino, On the  $H^1$ - $L^1$  boundedness of operators, *Proc. Amer. Math. Soc.*, **136** (2008), no. 8, 2921–2931.
- [336] A. D. Melas, On the centered Hardy-Littlewood maximal operator, *Trans. Amer. Math. Soc.* **354** (2002), no. 8, 3263–3273
- [337] A. D. Melas, The best constant for the centered Hardy-Littlewood maximal inequality, *Ann. of Math. (2)* **157** (2003), no. 2, 647–688.
- [338] A. D. Melas, On a covering problem related to the centered Hardy-Littlewood maximal inequality, *Ark. Mat.* **41** (2003), no. 2, 341–361.
- [339] A. Melin and J. Sjöstrand, Fourier integral operators with complex-valued phase functions, *Fourier integral operators and partial differential equations (Colloq. Internat., Univ. Nice, Nice, 1974)*, 120–223. *Lecture Notes in Math.*, **459**, Springer, Berlin, 1975.
- [340] M. T. Menarguez and F. Soria, Weak type  $(1, 1)$ -inequalities of maximal convolution operators, *Rendiconti del circolo matematico di palermo, Serie II, Tomo XLI* (1992), 342–352.
- [341] M. T. Menarguez and J. L. Torrea, Some inequalities arising from vector-valued dirac deltas, *Math. Scand.* **81** (1997), 301–319.
- [342] Y. Meyer, Principe d'incertitude, bases hilbertiennes et algèbres d'opérateur, *Seminaire Bourbaki* **38** no. 662. (1985–1986).
- [343] Y. Meyer, Wavelets and operators, *Analysis at Urbana, London Math. Soc. Lecture Notes* **137** (1989), 256–365.
- [344] Y. Meyer, *Wavelets Algorithms and Applications*, SIAM, 1993.
- [345] Y. Meyer, Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. (French) [The uncertainty principle, Hilbert bases and operator algebras] *Séminaire Bourbaki, Vol. 1985/86. Astérisque No. 145-146* (1987), **4**, 209–223.
- [346] Y. Meyer, Wavelets with compact support, *Zygmund Lectures*, U. Chicago (1987).
- [347] Y. Meyer, Ondelettes, fonction splines et analysis graduées [Wavelets, spline functions, and multiresolution analysis], *Rend. Sem. Mat. Univ. Politec. Torino*, **45** (1987), 1–42. Translated by John Horváth.
- [348] S. Mihlin, *Singular integral equations*, Amer. Math. Soc. Translation 1950, (1950). no. 24, 116 pp.
- [349] P. W. Millar, Martingale integrals, *Trans. Amer. Math. Soc.*, **133** (1968), 145–166.
- [350] F. Mintzer, Filters for distortion-free two-band multirate filter banks, *IEEE Trans. Acoust., Speech, and Signal Proc.*, **33** (1985), 626–630.
- [351] A. Miyachi, Estimates for pseudo-differential operators of class  $S_{0,0}$ , *Math. Nachr.* **133** (1987), 135–154.
- [352] A. Miyachi and K. Yabuta, On the good  $\lambda$ -inequality, *Bull. Fac. Sci. Ibaraki Univ.* **16** (1984), 1–11.
- [353] T. Mizuhara, Commutators of Singular Integral Operators on Morrey spaces with General Growth Functions, *Suurikaiseki Koukyuroku*, **1102** (1999), 49–63.

- [354] T. Mizuhara, Commutator of singular integral operators on Morrey spaces with some growth functions, Proceedings of the Second ISAAC Congress, **1**, 65–73.
- [355] Y. Mizuta and T. Shimomura, Sobolev's inequality for Riesz potentials of functions in Morrey spaces of integral form, to appear in Math. Nachr.
- [356] Y. Mizuta and T. Shimomura, Exponential integrability for Riesz potentials of functions in Orlicz classes, Hiroshima Math. J. **28** (1998), 355–371.
- [357] Y. Mizuta, T. Shimomura and T. Sobukawa, Sobolev's inequality for Riesz potentials of functions in non-doubling Morrey spaces, Osaka J. Math. **46** no. 1, 255–271.
- [358] E. Moore and H. Smith, A general theory of limits, Amer. J. Math. **44** (1922), 102–121.
- [359] P. A. P. Moran, Additive functions of intervals and Hausdorff measure, Proc. Cambridge Philos. Soc. **42** (1946), 15–23.
- [360] C. Morrey, On solutions of quasi-linear elliptic partial differential equations, Trans. Amer. Math. Soc. **43** (1938), 126–166.
- [361] B. Muckenhoupt, The equivalence of two conditions for weight functions, Studia Math., **55** (1976), no. 3, 279–294.
- [362] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, Trans. Amer. Math. Soc. **165** (1972), 207–226.
- [363] B. Muckenhoupt and R. Wheeden, Weighted norm inequalities for fractional integrals, Trans. Amer. Math. Soc. **192** (1974), 261–274.
- [364] M. Nagumo, Einige analytic Untersuchungen in linearen, metrischen Ringen, Japanese J. Math., **13**, 61–80.
- [365] E. Nakai, Generalized fractional integrals on Orlicz-Morrey Spaces, Proceedings on Banach and Function spaces, Kitakyushu, Japan, Oct. 2-4, 2003 (2004) 323–333.
- [366] E. Nakai, Generalized fractional integrals on Orlicz-Morrey spaces, Proceedings of the International Symposium on Banach and Function spaces, Yokohama Publishers.
- [367] E. Nakai, Hardy-Littlewood maximal operator, singular integral operators and the Riesz potentials on generalized Morrey spaces, Math. Nachr. **166** (1994), 95–103.
- [368] E. Nakai, The Campanato, Morrey and Hölder spaces on spaces of homogeneous type, Studia Math., **176** (2006), 1–19.
- [369] A. Nekvinda, Equivalence of  $l^{p_n}$  norms and shift operators, Math. Inequal. Appl. **5** (2002), 711–723.
- [370] O. Nikodym, Sur une généralisation des intégrales de M. J. Radon, Fund. Math. **15** (1930), 131–179.
- [371] S. M. Nikol'skij, Inequalities for entire function of finite order and their application in the theory of differentiable functions of several variables. (Russian) Trudy Mat. Inst. Steklov **38** (1951), 244–278.
- [372] P. Nilsson, Interpolation of Banach lattices, Studia Math. **82** (1985), no. 2, 135–154.
- [373] F. Nazarov, S. Treil and A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces, Internat. Math. Res. Notices (1997), no. 15, 703–726.
- [374] F. Nazarov, S. Treil and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces, Internat. Math. Res. Notices (1998), 463–487.
- [375] F. Nazarov, S. Treil and A. Volberg, The  $Tb$ -theorem on non-homogeneous spaces, Acta Math. **190** (2003), no. 2, 151–239.
- [376] L. Nirenberg and F. Tréves, On local solvability of linear partial differential equations, II, Sufficient conditions. Comm. Pure Appl. Math. **23** (1970), 459–509.
- [377] K. Ogura, On a certain transcendental integral function in the theory of interpolation, Tohoku Math. J. **17** (1920), 64–72.
- [378] R. O'Neil, Fractional integration in Orlicz spaces. I, Trans. Amer. Math. Soc. **115** (1965), 300–328.
- [379] W. Orlicz, Über Räume ( $L^M$ ), Bull. Acad. Pol. Sc. Letter., Ser. A (1936), 93–107.
- [380] B. Oskar, Remarks on operator BMO spaces, Rev. Un. Mat. Argentina **45** (2004), no. 1, 63–78 (2005).
- [381] B. Oscar, Dyadic BMO, paraproducts and Haar multipliers, Interpolation theory and applications, 11–18, Contemp. Math., **445**, Amer. Math. Soc., Providence, RI, 2007.
- [382] B. Oscar and P. Sandra, Dyadic BMO on the bidisk, Rev. Mat. Iberoamericana **21** (2005), no. 2, 483–510.
- [383] D. Z. Otyrba, On rearrangements of the Haar system in the space BMO, Anal. Math. **21** (1995), no. 2, 125–128.
- [384] R. E. A. C. Paley, A remarkable series of orthogonal functions, Proc. London Math. Soc., **34** (1932), 241–279.
- [385] J. W. Peter, Factorization of  $A_p$ -weights, Ann. of Math. (2) **111** (1980), no 3, 511–530.
- [386] J. Peetre, A theory of interpolation of normed spaces. Lecture notes, Brasilia, (1963).
- [387] J. Peetre, On the theory of  $\mathcal{L}^{p,\lambda}$  spaces, J. Funct. Anal. **4** (1969) 71–87.
- [388] J. Peetre, Sur la transformation de Fourier des fonctions à valeurs vectorielles, (French) Rend. Sem. Mat. Univ. Padova **42** (1969), 15–26.
- [389] J. Peetre, A new approach in interpolation spaces, Studia Math. **18** (1959), 49–65.
- [390] J. Peetre, Interpolation of Lipschitz operators and metric spaces, (Cluj) **12(35)** (1970), 325–334.

- [391] J. Peetre, Sur l'utilisation des suites inconditionnellement sommables dans la théorie des espaces d'interpolation, (French) Rend. Sem. Mat. Univ. **46** (1971), 173–190.
- [392] J. Peetre,  $H_p$ -spaces, Lecture notes, Lund 1974, 2nd ed. (1975).
- [393] C. Pérez, Two weighted norm inequalities for Riesz potentials and uniform  $L^p$ -weighted Sobolev inequalities, Indiana Univ. Math. J. **39** (1990), no 1, 31–44.
- [394] S. Petermichl, S. Treil and A. Volberg, WHY THE RIESZ TRANSFORMS ARE AVERAGES OF THE DYADIC SHIFTS ?, Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations, El Escorial 2000, Publ. Mat. (2002), 209–228.
- [395] L. Pick, Two-weight weak type maximal inequalities in Orlicz classes, Studia Math. **100** (1991), no. 3, 207–218.
- [396] M. A. Pinsky, N. K. Stanton and P. T. Trapa, Fourier series of radial functions in several variables, J. Funct. Anal., **116** (1993), 111–132.
- [397] M. A. Pinsky, Fourier inversion for piecewise smooth functions in several variables, Proc. Amer. Math. Soc., **118** (1993), 903–910.
- [398] Pointwise Fourier inversion and related Eigenfunction expansions, Comm. Pure and Appl. Math., **47** (1994), 653–681.
- [399] M. Plancherel and M. G. Pólya, Fonctions entières et intégrales de fourier multiples, (French) Comment. Math. Helv. **10** (1937), no. 1, 110–163.
- [400] M. Pramanik and E. Terwileger, A weak  $L^2$  estimates for a maximal dyadic sum operator on  $\mathbb{R}^n$ , Illinois J. Math. **47** (2003), 775–813.
- [401] M. M. Rao, Interpolation, ergodicity and martingales, J. Math. Mech. **16** (1966), 543–567.
- [402] L. J. Rogers, An extension of a certain theorem in inequalities, Messenger of Mathematics, **17**, no. 10, 145–150.
- [403] J. L. Rubio de Francia, Factorization theory and  $A_p$  weights, Amer. J. Math. **106** (1984), no. 3, 533–547.
- [404] J. L. Rubio de Francia, A new technique in the theory of  $A_p$  weights, Topics in modern harmonic analysis, Vol. I, II (Turin/Milan, 1982), 571–579.
- [405] V. A. Rvachev, Compactly supported solution of functional-differential equations and their applications, Russian Math. Surveys **45** (1990), 87–120.
- [406] F. Riesz, Sur les systéms orthogonaux de fonctions, C. Acad. Sci. Paris **144** (1907), 1022–1024.
- [407] F. Riesz, Untersuchungen über Systeme integrierbarer Funktionen, Math. Ann. **69** (1910), 449–497.
- [408] M. Riesz, Les fonctions conjuguées et les séries de Fourier, C. R. Acad. Sci. Paris, **178** (1924), 1464–1467.
- [409] M. Riesz, Sur les fonctions conjuguées, Math. Z. **27** (1927), 218–244.
- [410] M. Riesz, Sur les maxima des formes bilinéaires et sur les fonctionelles linéaires Acta. Math. **49** (1927), 465–497.
- [411] S. Rolewicz, On a certain class of linear metric spaces, Bull. Acad. Polon. Sci. **5** (1957), 471–473.
- [412] R. Rochberg and K. Tachizawa, Pseudodifferential operators, Gabor frames, and local trigonometric bases. In Gabor analysis and algorithms, 171–192, Birkhäuser Boston, Boston, MA, 1998.
- [413] J. L. Rubio de Francia, R. Francisco J. , Torrea, José L., Calderón-Zygmund theory for operator-valued kernels, Adv. in Math. **62** (1986), no 1, 7–48.
- [414] A. Ruiz and L. Vega, On local regularity of Schrödinger equation, Internat. Math. Res. Notices, (1993) no 1, 13–27.
- [415] A. Ruiz and L. Vega, Unique continuation for Schrödinger operators with potential in Morrey spaces, Publ. Mat. J. **35** (1991), no 1, 291–298.
- [416] V. S. Rychkov, On a theorem of Bui, Paluszyński, and Taibleson, Proceedings of the Steklov Institute of Mathematics **227** (1999), 280–292.
- [417] V. S. Rychkov, Littlewood-Paley theory and function spaces with  $A_p^{\text{loc}}$  weights, Math. Nachr. **224** (2001), 145–180.
- [418] D. Sarason, Funcations of vanishing mean oscillation, Transations of the American mathematical society, **207**, 1975.
- [419] , I. I. Sharapudinov, The basis property of the Haar system in the space  $L^{p(t)}([0, 1])$  and the principle of localization in the mean, (Russian) Mat. Sb. (N.S.) 130(172) (1986), no. 2, 275–283 and 286.
- [420] Z. Shen, Boundary value problems in Morrey spaces for elliptic systems on Lipschitz domains, Amer. J. Math. **125** (2003), no 5 1079–1115.
- [421] Z. Shen, Boundary value problems operators with potentials in the Morrey class, J. Funct. Anal. **193** (2002), no 2 314–315.
- [422] G. Stampacchia,  $\mathcal{L}^{(p,\lambda)}$ -spaces and interpolation, communications on pure and applied mathematics, **17** (1964), 293–306.
- [423] G. Stampacchia, The spaces  $L(p, \lambda)$ ,  $N(p, \lambda)$  and interpolation, Ann. Scuola Norm. Sup. Pisa (3) **19** 1965 443–462.
- [424] Y. Sawano, Sharp estimates of the modified Hardy-Littlewood maximal operator on the nonhomogeneous space via covering lemmas, Hokkaido Math. J. **34** (2005) 435–458.

- [425] Y. Sawano, T. Sobukawa and H. Tanaka, Limiting case of the boundedness of fractional integral operators on non-homogeneous space, *Journal of Inequalities and Applications*, vol. 2006, Article ID 92470, 16 pages, 2006.
- [426] Y. Sawano and H. Tanaka, Morrey spaces for non-doubling measures, *Acta Math. Sinica*, **21** no.6, 1535–1544.
- [427] Y. Sawano and H. Tanaka, Sharp maximal inequalities and commutators on Morrey spaces with non-doubling measures, *Taiwanese Math. J.* **11** (2007), no. 4, 1091–1112.
- [428] Y. Sawano and H. Tanaka, The John-Nirenberg type inequality for non-doubling measures, *Studia Math.* **181** no 2, (2007), 153–170.
- [429] Y. Sawano and H. Tanaka, Decompositions for Besov-Morrey spaces and Triebel-Lizorkin-Morrey spaces, *Math. Z.* **257** (2007), no. 4, 871–905.
- [430] Y. Sawano and T. Yoneda, Quarkonial decomposition suitable for equations of delay type, to appear in *Math. Nach.*
- [431] Y. Sawano and T. Yoneda, On the Young theorem for amalgams and Besov spaces, *International Journal of Pure and Applied Mathematics*, **36** (2007), no 2, 199–208.
- [432] J. Schauder, Über die Umkehrung linearer, stetiger funktionaloperationen, *Studia Math.*, **2**, 1–6.
- [433] J. Schwartz, A remark on inequalities of Calderon-Zygmund type for vector-valued functions, *Comm. Pure Appl. Math.*, **14** (1961), 785–799.
- [434] A. Seeger, Estimates near  $L^1$  for Fourier multipliers and maximal functions, *Arch. Math. (Basel)* **53** (1989), no. 2, 188–193.
- [435] C. Segovia, On the area function of Lusin, *Studia Math.*, **33** (1969), 311–343.
- [436] J. M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, *IEEE Trans. Signal Processing*, **41** (1993), 3445–3462.
- [437] L. Shanzhen, K. Yabuta and Y. Dachun, Boundedness of some sublinear operators in weighted Herz-type spaces, *Kodai Math. J.*, **23** (2000), no. 3, 391–410.
- [438] T. Shimomura,  $L^q$ -mean limits of Taylor's expansion of Riesz potentials of functions in Orlicz spaces, *Hiroshima Math. J.*, **27** (1997), 159–175.
- [439] T. Shimomura and Y. Mizuta, Taylor expansion of Riesz potentials, *Hiroshima Math. J.*, **25** (1995), 595–621.
- [440] J. Sjöstrand, Parametrices for pseudodifferential operators with multiple characteristics, *Ark. Mat.*, **12** (1974), 85–130.
- [441] J. Sjöstrand, An algebra of pseudodifferential operators, *Math. Res. Lett.* **1** (1994), no. 2, 185–192.
- [442] J. Sjöstrand, Wiener type algebras of pseudodifferential operators, In *Séminaire sur les Équations aux Dérivées Partielles*, 1994–1995, pages Exp. No. IV, 21. École Polytech., Palaiseau, 1995.
- [443] J. Sjöstrand, Pseudodifferential operators and weighted normed symbol spaces, Preprint, 2007. arXiv:0704.1230v1.
- [444] S. Spanne, Sur l'interpolation entre les espaces  $\mathcal{L}_k^{p\Phi}$ , *Ann. Scuola Norm. Sup. Pisa*, **20** (1966), 625–648.
- [445] S. Shirai, Necessary and sufficient conditions for boundedness of commutators of fractional integral operators on classical Morrey spaces, *Hokkaido Math. J.* **35** (2006), no. 3, 683–696.
- [446] M. J. T. Smith and T. P. Barnwell III, A procedure for designing exact reconstruction filter banks for tree-structured subband coders, ICASSP '84, *IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing*, **9**, March 1984, 421–424.
- [447] S. Sobolev, On a theorem in functional analysis (Russian), *Math. Sob.* **46** (1938), 471–490.
- [448] O. J. Strömberg, A modified Franklin system and higher-order spline systems on  $\mathbf{R}^n$  as unconditional bases for Hardy spaces, *Conference on harmonic analysis in honor of Antoni Zygmund*, Vol. I, II (Chicago, Ill., 1981), 475–494, *Wadsworth Math. Ser.*, Wadsworth, Belmont, CA, 1983.
- [449] P. Stein, On a theorem of M. Riesz, *J. London Math. Soc.* **8** (1933), 242–247.
- [450] E. M. Stein, Interpolation of linear operators, *Trans. Amer. Math. Soc.* **83** (1956), 482–492.
- [451] E. M. Stein, Note on singular integrals, *Proc. Amer. Math. Soc.* **8** (1957), 250–254.
- [452] E. M. Stein, On the functions of Littlewood-Paley, Lusin, and Marcinkiewicz, *Trans. Amer. Math. Soc.* **88** (1958), 430–466.
- [453] E. M. Stein, A maximal function with applications to Fourier series, *Ann. of Math. (2)* **68** (1958), 584–603.
- [454] E. M. Stein, Singular integrals, harmonic functions, and differentiability properties of functions of several variables. *Singular integrals (Proc. Sympos. Pure Math., Chicago, Ill., 1966)*, 316–335. *Amer. Math. Soc.*, Providence, R.I., (1967).
- [455] E. M. Stein, Note on the Class  $L \log L$ , *Studia Math.* **31** (1969), 305–310.
- [456] E. M. Stein, Topics in harmonic analysis related to the Littlewood-Paley theory, *Ann. Math. Studies* **63**. Princeton, 1970.
- [457] E. M. Stein, The development of square functions in the work of A. Zygmund, *Bull. Amer. Math. Soc.* **7** (1982), no. 2, 359–376.

- [458] E. M. Stein and G. Weiss, On the theory of harmonic functions of several variables. I. The theory of  $H^p$ -spaces, *Acta Math.* **103** (1960), 25–62.
- [459] E. M. Stein and G. Weiss, Interpolation of operators with change of measures, *Trans. Amer. Math. Soc.* **87** (1958), 159–172.
- [460] E. M. Stein and J. Strömberg, Behavior of maximal functions in  $R^n$  for large  $n$ , *Ark. Mat.* **21** (1983), no. 2, 259–269.
- [461] R. S. Strichartz, A functional calculus for elliptic pseudo-differential operators, *Amer. J. Math.* **94** (1972), 711–722.
- [462] G. Stampacchia,  $\mathcal{L}^{(p,\lambda)}$ -spaces and interpolation, *Comm. Pure Appl. Math.* **17** (1964) 293–306.
- [463] J. O. Strömberg, A modified Franklin system and higher-order spline systems on  $\mathbb{R}^n$  as unconditional bases for Hardy spaces, *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II* (Chicago, Ill., 1981), 475–494, *Wadsworth Math. Ser.*, Wadsworth, Belmont, CA, 1983.
- [464] J. O. Strömberg and A. Torchinsky, Weights, sharp maximal functions and Hardy spaces, *Bull. Amer. Math. Soc. (N.S.)* **3** (1980), no. 3, 1053–1056.
- [465] S. Sugano and H. Tanaka, Boundedness of fractional integral operators on generalized Morrey spaces, *Sci. Math. Jpn.* **58** (2003), no. 3, 531–540.
- [466] K. Tachizawa, The boundedness of pseudodifferential operators on modulation spaces, *Math. Nachr.*, **168**, 263–277, 1994.
- [467] M. H. Taibleson, On the theory of Lipschitz spaces of distributions on Euclidean  $n$ -space. I. Principal properties, *J. Math. Mech.*, **13** (1964), 407–479.
- [468] L. Tang and J. Xu, Some properties of Morrey type Besov-Triebel spaces, *Math. Nachr.*, **278**, no 7-8, 904–917.
- [469] H. Tanaka, Kakeya conjecture, (Japanese) *Sugaku* **57** (2005), no. 3, 225–241.
- [470] M. Taylor, Analysis on Morrey spaces and applications to Navier-Stokes and other evolution equations, *Comm. Partial Differential Equations*, **17** (1992), no 9–10, 1407–1456.
- [471] Y. Terasawa, Outer measure and weak- $(1, 1)$  estimates of the modified Hardy-Littlewood maximal operators, *J. Inequal. Appl.* (2006), Art. ID 15063, 13 pp.
- [472] G. Thorin, An extension of a convexity theorem due to M. Riesz, *Kungl. Fysiogr. Sällsk i Lund Förh.* **8** (1938), 166–170.
- [473] X. Tolsa, Littlewood-Paley theory and the  $T(1)$  theorem with non-doubling measures, *Adv. Math.* **164** (2001), 57–116.
- [474] X. Tolsa,  $BMO$ ,  $H^1$  and Calderón-Zygmund operators for non doubling measures, *Math. Ann.* **319** (2001), 89–149.
- [475] X. Tolsa, The space  $H^p$  for nondoubling measures in terms of a grand maximal operator, *Trans. Amer. Math. Soc.* **355** (2003), no. 1, 315–348 (electronic).
- [476] X. Tolsa, Littlewood-Paley theory and the  $T(1)$  theorem with non-doubling measures, *Adv. Math.* **164** (2001), 57–116.
- [477] X. Tolsa, Painlevé’s problem and the semiadditivity of analytic capacity. *Acta Math.* **190** (2003), no. 1, 105–149.
- [478] X. Tolsa, Bilipschitz maps, analytic capacity, and the Cauchy integral, *Ann. of Math. (2)* **162** (2005), no. 3, 1243–1304.
- [479] T. Tsuchikura, Sample properties of martingales and their arithmetic means, *Tohoku Math. J.*, **20** (1968), 400–415.
- [480] H. Triebel, Spaces of distributions of Besov type on Euclidean  $n$ -space. Duality, interpolation, *Ark. Mat.* **11**, 13–64. (1973).
- [481] H. Triebel, Modulation spaces on the Euclidean  $n$ -space, *Z. Anal. Anwendungen*, **2** (5), (1983) 443–457.
- [482] S. Trudinger, On imbeddings into Orlicz spaces and some applications, *J. Math. Mech.*, **17** (1967) 473–483.
- [483] A. Uchiyama, A constructive proof of the Fefferman-Stein decomposition of  $BMO(\mathbb{R}^n)$ , *Acta Math.* **148**(1982), 215–241.
- [484] A. Uchiyama, On the compactness of operators of Hankel type, *Tôhoku Math. J.*, **30** (1978), 163–171.
- [485] P. P. Vaidyanathan, Theory and design of  $M$ -channel maximally decimated quadrature mirror filters with arbitrary  $M$ , having the perfect-reconstruction property, *IEEE Trans. Acoust., Speech, and Signal Proc.*, **35** (1987), 476–492.
- [486] Martin Vetterli, Filter banks allowing perfect reconstruction, *Signal Processing*, **10** (1986), 219–244.
- [487] B. Virot, Extensions vectorielles d’opérateurs linéaires bornés sur  $L^p$ . (French) [Vector extensions of bounded linear operators on  $L^p$ ] *C. R. Acad. Sci. Paris Sér. I Math.* **293** (1981), no. 8, 413–415.
- [488] W. H. Young, On classes of summable functions and their Fourier series, *Proc. Roy. Soc. (A)*, **87** (1912), 225–229.
- [489] T. Yoneda, On the functional-differential equation of advanced type  $f'(x) = af(2x)$  with  $f(0) = 0$ , *J. Math. Anal. Appl.* **317** (2006), no. 1, 320–330.

- [490] A. Yoshikawa, An operator theoretical remark on the Hardy-Littlewood-Sobolev inequality, *J. Fac. Sci. Univ. Tokyo Sect. I A Math.* **17** (1970) 559–566.
- [491] G. V. Welland, Weighted norm inequalities for fractional integrals, *Proc. Amer. Math. Soc.* **51**, (1975), 143–148.
- [492] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.* **36** (1934), no. 1, 63–89.
- [493] H. Widom, A complete symbolic calculus for pseudodifferential operators, *Bull. Sci. Math. (2)* **104** (1980), no. 1, 19–63.
- [494] N. Wiener, The ergodic theorem, *Duke Mathematical Journal* **5** (1939), 1–18.
- [495] M. Wilson, A simple proof of the atomic decomposition for  $H^p(\mathbb{R}^n)$ ,  $0 < p \leq 1$ , *Studia Math.* **74** (1982), no. 1, 25–33.
- [496] T. Wolff, A note on interpolation spaces, *Harmonic analysis (Minneapolis, Minn., 1981)*, 199–204, *Lecture Notes in Math.*, **908**, Springer, Berlin-New York, 1982.
- [497] C. Zorko, Morrey space, *Proc. Amer. Math. Soc.* **98** (1986), no. 4, 586–592.
- [498] A. Zygmund, On a theorem of Marcinkiewicz concerning interpolation theorem, *J. Math. Pures. Appl. (9)* **35** (1956), 223–248.
- [499] A. Zygmund, On certain lemmas of Marcinkiewicz and Carleson, *J. Approx Theory*, **2** (1969), 249–257.

## PERSONAL COMMUNICATIONS

- [500] M. Frazier, personal communication.
- [501] Y. Giga, personal communication.
- [502] Y. Han, personal communication.
- [503] M. Izuki, personal communication.
- [504] A. Miyachi, personal communication.
- [505] E. Nakai, personal communication.
- [506] K. Oguiso, personal communication.
- [507] T. Sobukawa, personal communication.
- [508] H. Tanaka, personal communication.
- [509] N. Tomita, personal communication.
- [510] K. Yoneda, personal communication.
- [511] T. Yoneda, personal communication.