

## IX. $L^p$ Spaces, 448-489

---

DOI: [10.3792/euclid/9781429799997-9](https://doi.org/10.3792/euclid/9781429799997-9)

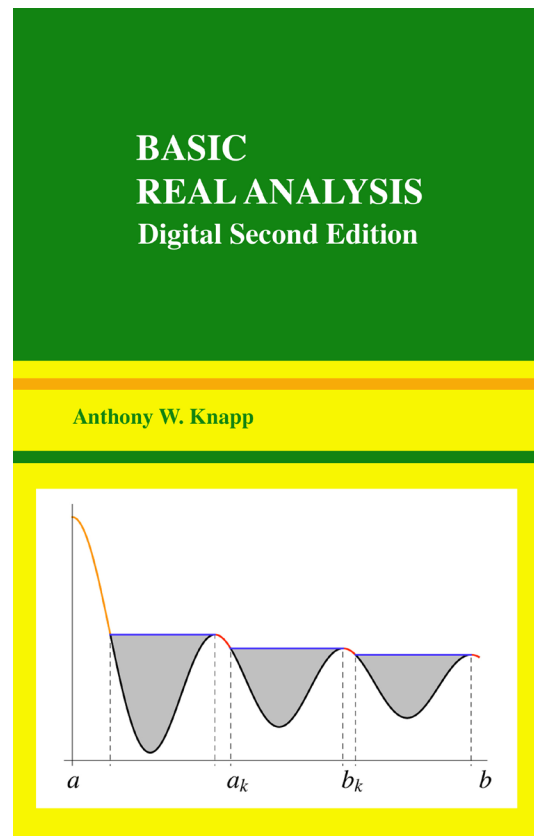
from

### *Basic Real Analysis* *Digital Second Edition*

Anthony W. Knapp

Full Book DOI: [10.3792/euclid/9781429799997](https://doi.org/10.3792/euclid/9781429799997)

ISBN: 978-1-4297-9999-7



Anthony W. Knapp  
81 Upper Sheep Pasture Road  
East Setauket, N.Y. 11733-1729, U.S.A.  
Email to: [aknapp@math.stonybrook.edu](mailto:aknapp@math.stonybrook.edu)  
Homepage: [www.math.stonybrook.edu/~aknapp](http://www.math.stonybrook.edu/~aknapp)

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"

Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28-01, 26-01, 42-01, 54-01, 34-01, 30-01, 32-01.

First Edition, ISBN-13 978-0-8176-3250-2

©2005 Anthony W. Knapp

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

©2016 Anthony W. Knapp

Published by the author

All rights reserved. This file is a digital second edition of the above named book. The text, images, and other data contained in this file, which is in portable document format (PDF), are proprietary to the author, and the author retains all rights, including copyright, in them. The use in this file of trade names, trademarks, service marks, and similar items, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

All rights to print media for the first edition of this book have been licensed to Birkhäuser Boston, c/o Springer Science+Business Media Inc., 233 Spring Street, New York, NY 10013, USA, and this organization and its successor licensees may have certain rights concerning print media for the digital second edition. The author has retained all rights worldwide concerning digital media for both the first edition and the digital second edition.

The file is made available for limited noncommercial use for purposes of education, scholarship, and research, and for these purposes only, or for fair use as understood in the United States copyright law. Users may freely download this file for their own use and may store it, post it online, and transmit it digitally for purposes of education, scholarship, and research. They may not convert it from PDF to any other format (e.g., EPUB), they may not edit it, and they may not do reverse engineering with it. In transmitting the file to others or posting it online, users must charge no fee, nor may they include the file in any collection of files for which a fee is charged. Any exception to these rules requires written permission from the author.

Except as provided by fair use provisions of the United States copyright law, no extracts or quotations from this file may be used that do not consist of whole pages unless permission has been granted by the author (and by Birkhäuser Boston if appropriate).

The permission granted for use of the whole file and the prohibition against charging fees extend to any partial file that contains only whole pages from this file, except that the copyright notice on this page must be included in any partial file that does not consist exclusively of the front cover page. Such a partial file shall not be included in any derivative work unless permission has been granted by the author (and by Birkhäuser Boston if appropriate).

Inquiries concerning print copies of either edition should be directed to Springer Science+Business Media Inc.

# CHAPTER IX

## $L^p$ Spaces

**Abstract.** This chapter extends the theory of the spaces  $L^1$ ,  $L^2$ , and  $L^\infty$  to include a whole family of spaces  $L^p$ ,  $1 \leq p \leq \infty$ , in order to be able to capture finer quantitative facts about the size of measurable functions and the effect of linear operators on such functions.

Sections 1–2 give the basics about  $L^p$ . For general measure spaces these consist of Hölder’s inequality, Minkowski’s inequality, a completeness theorem, and related results. For Euclidean space they include also facts about convolution.

Sections 3–4 develop some tools that at first may seem quite unrelated to  $L^p$  spaces but play a significant role in Section 5. These are the Radon–Nikodym Theorem and two decomposition theorems for additive set functions. The Radon–Nikodym Theorem gives a sufficient condition for writing a measure as a function times another measure.

Section 5 identifies the space of continuous linear functionals on  $L^p$  for  $1 \leq p < \infty$  when the underlying measure is  $\sigma$ -finite. For one thing this identification makes Alaoglu’s Theorem in Chapter V concrete enough so as to be quite useful.

Section 6 establishes the Riesz–Thorin Convexity Theorem, which asserts that linear operators that are bounded between two pairs of  $L^p$  spaces are bounded between suitable intermediate pairs of  $L^p$  spaces as well. Immediate corollaries include the Hausdorff–Young Theorem concerning the Euclidean Fourier transform and Young’s inequality concerning convolution of functions in two  $L^p$  spaces.

Section 7 discusses the Marcinkiewicz Interpolation Theorem, which allows one to reinterpret bounded sublinear operators between two pairs of  $L^p$  spaces as bounded between suitable intermediate pairs of  $L^p$  spaces as well. The theorem has immediate corollaries for the Hardy–Littlewood maximal function and an approximation to the Hilbert transform, and Section 7 goes on to use each of these corollaries to derive interesting consequences.

### 1. Inequalities and Completeness

In the context of any measure space, we introduced in Section V.9 the spaces  $L^1$ ,  $L^2$ , and  $L^\infty$ . Since then, we have used these three spaces to capture quantitative facts about the size of measurable functions. The construction in each case involved introducing a certain pseudonorm in a vector space of functions, thereby making the vector space into a pseudo normed linear space and in particular a pseudometric space. The corresponding metric space obtained from the construction of Proposition 2.12 was  $L^1$ ,  $L^2$ , or  $L^\infty$  in the respective cases. For each of

the three, the vector-space structure for the pseudometric space yielded a vector-space structure for the metric space, and  $L^1$ ,  $L^2$ , and  $L^\infty$  were normed linear spaces. As was true in Chapters V and VI, it continues in the present chapter to be largely a matter of indifference whether the functions in question are real valued or complex valued, hence whether the scalars for these vector spaces are real or complex.

Now we shall enlarge the family consisting of  $L^1$ ,  $L^2$ ,  $L^\infty$  to a family  $L^p$  for  $1 \leq p \leq \infty$  in order to be able to capture finer quantitative facts about the size of measurable functions. Enlarging the family in this way makes it possible to get better insight into the behavior of specific operators and to make more helpful estimates with partial differential equations.

Let  $(X, \mathcal{A}, \mu)$  be a measure space. We have already dealt with  $p = \infty$ . For  $1 \leq p < \infty$ , we consider the set  $V = V_p$  of measurable functions  $f$  on  $X$  such that  $\int_X |f|^p d\mu$  is finite. This integral is well defined; in fact,  $f$  measurable implies  $|f|$  measurable, and also, for  $c > 0$ ,  $(|f|^p)^{-1}(c, +\infty) = |f|^{-1}(c^{1/p}, +\infty)$ . The set  $V$  is in fact a vector space of functions. It is certainly closed under scalar multiplication; let us see that it is closed under addition. If  $f$  and  $g$  are in  $V$ , then we have

$$\begin{aligned} (|f(x)| + |g(x)|)^p &\leq (\max\{|f(x)|, |g(x)|\} + \max\{|f(x)|, |g(x)|\})^p \\ &= 2^p \max\{|f(x)|^p, |g(x)|^p\} \leq 2^p |f(x)|^p + 2^p |g(x)|^p \end{aligned}$$

for every  $x$  in  $X$ . Integrating over  $X$ , we see that  $f + g$  is in  $V$  if  $f$  and  $g$  are in  $V$ .

Following the construction of the prototypes  $L^1$  and  $L^2$  in Section V.9, we introduce the expression  $\|f\|_p = (\int_X |f|^p d\mu)^{1/p}$  for  $f$  in  $V_p$ . We would like  $\|\cdot\|_p$  to be a pseudonorm in the sense of satisfying

- (i)  $\|x\|_p \geq 0$  for all  $x \in V$ ,
- (ii)  $\|cx\|_p = |c|\|x\|_p$  for all scalars  $c$  and all  $x \in V$ ,
- (iii)  $\|x + y\|_p \leq \|x\|_p + \|y\|_p$  for all  $x$  and  $y$  in  $V$ .

Properties (i) and (ii) are certainly satisfied, but a little argument is needed to verify (iii). We return to this matter in a moment. Once the function  $\|\cdot\|_p$  on the vector space  $V_p$  is known to be a pseudonorm,  $V_p$  meets the conditions of being a pseudo normed linear space in the sense of Section V.9.

We can pass to the set of equivalence classes just as in that section, and this set is defined to be  $L^p$  or  $L^p(X)$  or  $L^p(X, \mu)$ . The equivalence class of 0 is again the set of all functions vanishing almost everywhere. The function  $\|\cdot\|_p$  is well defined on  $L^p$ , and  $L^p$  is a normed linear space. In particular, it has the structure of a metric space. This handles  $1 \leq p < \infty$ , and the space  $L^\infty$  was constructed in Section V.9.

As is true with  $L^1$ ,  $L^2$ , and  $L^\infty$ , one sometimes relaxes the terminology and works with the members of  $L^p(X)$  as if they were functions, saying, “Let the function  $f$  be in  $L^p(X)$ ” or “Let  $f$  be an  $L^p$  function.” There is little possibility of ambiguity in using such expressions.

Let us return to property (iii) above. This will be proved as Minkowski’s inequality below. But first we prove a numerical lemma and then “Hölder’s inequality,” which is a version for  $L^p$  of the Schwarz inequality for  $L^2$ . Hölder’s inequality makes use of the **dual index**  $p'$  to  $p$ , defined by the equality  $\frac{1}{p} + \frac{1}{p'} = 1$ . The dual index to 1 is  $\infty$ , and vice versa. The index 2 is its own dual.

**Lemma 9.1.** If  $s, t, \alpha$ , and  $\beta$  are real numbers  $\geq 0$  with  $\alpha + \beta = 1$ , then

$$s^\alpha t^\beta \leq \alpha s + \beta t.$$

PROOF. If any of  $s, t, \alpha, \beta$  is 0, the result is certainly true. If all are nonzero, consider the function

$$f(x) = \alpha x^{\alpha-1} + (1 - \alpha)x^\alpha,$$

defined for  $x > 0$ . The derivative  $f'(x) = (1 - \alpha)\alpha x^{\alpha-2}(x - 1)$  is  $< 0$  for  $0 < x < 1$ , is  $= 0$  for  $x = 1$ , and is  $> 0$  for  $x > 1$ . Therefore  $f(x)$  assumes its absolute minimum value for  $x = 1$ . Since  $f(1) = 1$ , we have

$$1 \leq \alpha x^{\alpha-1} + (1 - \alpha)x^\alpha = \alpha x^{-\beta} + \beta x^\alpha$$

for all  $x > 0$ . The lemma follows by putting  $x = t/s$  in this inequality and by multiplying both sides by  $s^\alpha t^\beta$ .  $\square$

REMARK. Alternatively, this lemma can be proved by Lagrange multipliers in the same way that Problem 20 at the end of Chapter III suggested using for the arithmetic-geometric mean inequality.

**Theorem 9.2** (Hölder’s inequality). Let  $(X, \mathcal{A}, \mu)$  be any measure space, let  $1 \leq p \leq \infty$ , and let  $p'$  be the dual index to  $p$ . If  $f$  is in  $L^p$  and  $g$  is in  $L^{p'}$ , then  $fg$  is in  $L^1$ , and

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}.$$

REMARK. The inequality holds trivially if  $\|f\|_p = +\infty$  or  $\|g\|_{p'} = +\infty$ .

PROOF. We already know the result if  $p = 1$  and  $p' = \infty$  or the other way around. Thus suppose that  $p > 1$  and  $p' > 1$ . We may assume that neither  $f$

nor  $g$  is 0 almost everywhere. Then we can apply Lemma 9.1 with  $\alpha = p^{-1}$ ,  $\beta = p'^{-1}$ ,

$$s = \frac{|f(x)|^p}{\int_X |f|^p d\mu}, \quad \text{and} \quad t = \frac{|g(x)|^{p'}}{\int_X |g|^{p'} d\mu},$$

getting

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_{p'}} \leq \frac{|f(x)|^p}{p \int_X |f|^p d\mu} + \frac{|g(x)|^{p'}}{p' \int_X |g|^{p'} d\mu}.$$

Integrating, we obtain

$$\frac{\int_X |fg| d\mu}{\|f\|_p \|g\|_{p'}} \leq \frac{1}{p} + \frac{1}{p'} = 1,$$

and the conclusions of the theorem follow.  $\square$

**Theorem 9.3** (Minkowski's inequality). Let  $(X, \mathcal{A}, \mu)$  be any measure space, and let  $1 \leq p \leq \infty$ . If  $f$  and  $g$  are in  $L^p$ , then  $f + g$  is in  $L^p$  and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

REMARK. The theorem assumes the usual convention that  $f + g$  is made to be 0 at any point  $x$  where  $f(x) + g(x)$  is not defined. The set where this change occurs is of measure 0 since  $f$  and  $g$  have to be finite almost everywhere to be in  $L^p$ .

PROOF. We have already seen that  $f + g$  is in  $L^p$ , and we know the inequality for  $p = 1$  and  $p = \infty$  from Section V.9. For  $1 < p < \infty$ , let  $p'$  be the dual index. We apply Hölder's inequality (Theorem 9.2) to  $f$  and  $|f + g|^{p-1}$  and to  $g$  and  $|f + g|^{p-1}$  to obtain

$$\begin{aligned} \int_X |f + g|^p d\mu &\leq \int_X |f + g| |f + g|^{p-1} d\mu \\ &\leq \int_X |f| |f + g|^{p-1} d\mu + \int_X |g| |f + g|^{p-1} d\mu \\ &\leq \|f\|_p \left( \int_X |f + g|^{(p-1)p'} d\mu \right)^{1/p'} + \|g\|_{p'} \left( \int_X |f + g|^{(p-1)p'} d\mu \right)^{1/p'} \\ &= \left( \int_X |f + g|^p d\mu \right)^{1/p'} (\|f\|_p + \|g\|_{p'}), \end{aligned}$$

the last step holding because  $(p-1)p' = p$ . If  $\|f + g\|_p = 0$ , the inequality of the theorem is certainly true. Otherwise the inequality of the theorem follows after dividing the inequality of the display by  $\left( \int_X |f + g|^p d\mu \right)^{1/p'}$ , which we know to be finite, and using the fact that  $1 - \frac{1}{p'} = \frac{1}{p}$ .  $\square$

Thus  $L^p$  is a normed linear space for  $1 \leq p \leq \infty$ . Let us derive some of its properties.

**Proposition 9.4.** Let  $(X, \mathcal{A}, \mu)$  be a measure space, and let  $1 \leq p < \infty$ . Then every indicator function of a set of finite measure is in  $L^p(X)$ , and the smallest closed subspace of  $L^p(X)$  containing all such indicator functions is  $L^p(X)$  itself. Consequently

- (a) the set of simple functions built from sets of finite measure lies in every  $L^p(X)$  for  $1 \leq p \leq \infty$  and is dense in  $L^p(X)$  if  $1 \leq p < \infty$ ,
- (b)  $1 \leq p_1 \leq p \leq p_2 \leq \infty$  and  $p < \infty$  together imply that  $L^{p_1}(X) \cap L^{p_2}(X)$  is dense in  $L^p(X)$ .

In addition,

- (c)  $1 \leq p_1 \leq p \leq p_2 \leq \infty$  implies that  $L^p(X) \subseteq L^{p_1}(X) + L^{p_2}(X)$ .

PROOF. The conclusion in the second sentence of the proposition is proved by the same argument as for Proposition 5.56. Part (a) then follows from Proposition 5.55d. Part (b) follows by combining these two results once it is known that  $L^{p_1}(X) \cap L^{p_2}(X) \subseteq L^p(X)$ . For this inclusion let  $f$  be in  $L^{p_1}(X) \cap L^{p_2}(X)$ . We may assume that  $p < \infty$ . If  $p_2 < \infty$ , then

$$\begin{aligned} \int_X |f|^p d\mu &= \int_{\{|f|>1\}} |f|^p d\mu + \int_{\{|f|\leq 1\}} |f|^p d\mu \\ &\leq \int_{\{|f|>1\}} |f|^{p_2} d\mu + \int_{\{|f|\leq 1\}} |f|^{p_1} d\mu < +\infty, \end{aligned}$$

and hence  $f$  is in  $L^p(X)$ . If  $p_2 = \infty$ , then  $\{|f| > 1\}$  has finite measure since  $f$  is in  $L^{p_1}$  and  $p_1 < \infty$ . Thus

$$\begin{aligned} \int_X |f|^p d\mu &= \int_{\{|f|>1\}} |f|^p d\mu + \int_{\{|f|\leq 1\}} |f|^p d\mu \\ &\leq \|f\|_\infty^p \mu(\{|f| > 1\}) + \int_{\{|f|\leq 1\}} |f|^{p_1} d\mu < +\infty, \end{aligned}$$

and again  $f$  is in  $L^p(X)$ . This completes the proof of (b).

For (c), let  $f$  be in  $L^p$ , and write  $f = f_1 + f_2$ , where

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Then

$$\int_X |f_1|^{p_1} d\mu = \int_{\{|f|>1\}} |f|^{p_1} d\mu \leq \int_{\{|f|>1\}} |f|^p d\mu < \infty$$

shows that  $f_1$  is in  $L^{p_1}(X)$ . It is apparent that  $f_2$  is in  $L^\infty(X)$ , and thus  $f_2$  is certainly in  $L^{p_2}(X)$  if  $p_2 = \infty$ . If  $p_2 < \infty$ , then

$$\int_X |f_2|^{p_2} d\mu = \int_{\{|f|\leq 1\}} |f|^{p_2} d\mu \leq \int_{\{|f|\leq 1\}} |f|^p d\mu < \infty$$

shows that  $f_2$  is in  $L^{p_2}(X)$ . This proves (c).  $\square$

Hölder's inequality allows us to prove the following supplement to the conclusions of Proposition 9.4.

**Proposition 9.5.** Let  $(X, \mathcal{A}, \mu)$  be any measure space. Let  $1 \leq p_1 < p < p_2$ , and define  $t$  with  $0 \leq t \leq 1$  by  $\frac{1}{p} = \frac{1-t}{p_1} + \frac{t}{p_2}$ . Then

$$\|f\|_p \leq \|f\|_{p_1}^{1-t} \|f\|_{p_2}^t.$$

PROOF. First suppose that  $p_2 < \infty$ . Since  $\frac{1}{p} > \frac{1-t}{p_1}$ , we can find  $b$  with  $1 < b < +\infty$  such that  $\frac{1}{bp} = \frac{1-t}{p_1}$ . If  $b'$  denotes the dual index, then  $\frac{1}{b'p} = \frac{1}{p} - \frac{1}{bp} = \frac{1}{p} - \frac{1-t}{p_1} = \frac{t}{p_2}$ . Define  $a$  by the equation  $ab = p_1$ . Then  $(p-a)b' = (p - \frac{p_1}{b})\frac{p_2}{p} = p_2(\frac{1}{t} - \frac{p_1}{bt}) = p_2(\frac{1}{t} - \frac{1-t}{t}) = p_2$ .

We write  $|f|^p = |f|^a |f|^{p-a}$ . Application of Hölder's inequality with index  $b$  and dual index  $b'$  gives  $\int |f|^p d\mu \leq (\int |f|^{ab} d\mu)^{1/b} (\int |f|^{(p-a)b'} d\mu)^{1/b'}$ , and hence

$$\|f\|_p \leq (\int |f|^{ab} d\mu)^{1/(bp)} (\int |f|^{(p-a)b'} d\mu)^{1/(b'p)}.$$

We have seen that  $ab = p_1$ ,  $1/(bp) = (1-t)/p_1$ ,  $(p-a)b' = p_2$ , and  $1/(b'p) = t/p_2$ . Thus the inequality reads  $\|f\|_p \leq \|f\|_{p_1}^{1-t} \|f\|_{p_2}^t$ , and the proof is complete when  $p_2 < \infty$ .

When  $p_2 = \infty$ , we write  $|f|^p = |f|^{p_1} |f|^{p-p_1}$ . Replacing  $|f|^{p-p_1}$  by its essential supremum gives  $\int |f|^p d\mu \leq \|f\|_{\infty}^{p-p_1} \int |f|^{p_1} d\mu$  and hence  $\|f\|_p$  is

$$\leq (\int |f|^{p_1} d\mu)^{1/p} \|f\|_{\infty}^{(p-p_1)/p} = (\int |f|^{p_1} d\mu)^{(1-t)/p_1} \|f\|_{\infty}^{1-p_1/p} = \|f\|_{p_1}^{1-t} \|f\|_{\infty}^t.$$

This completes the proof when  $p_2 = \infty$ . □

We have already made serious use of the completeness of  $L^p$  for  $p$  equal to 1, 2, and  $\infty$  as proved in Theorem 5.58. As might be expected, this result extends to be valid for the other values of  $p$ .

**Theorem 9.6.** Let  $(X, \mathcal{A}, \mu)$  be any measure space, and let  $1 \leq p \leq \infty$ . Any Cauchy sequence  $\{f_k\}$  in  $L^p$  has a subsequence  $\{f_{k_n}\}$  such that  $\|f_{k_n} - f_{k_m}\|_p \leq C_{\min\{m,n\}}$  with  $\sum_n C_n < +\infty$ . A subsequence  $\{f_{k_n}\}$  with this property is necessarily Cauchy pointwise almost everywhere. If  $f$  denotes the almost-everywhere limit of  $\{f_{k_n}\}$ , then the original sequence  $\{f_k\}$  converges to  $f$  in  $L^p$ . Consequently the space  $L^p$ , when regarded as a metric space, is complete in the sense that every Cauchy sequence converges.

REMARK. As in the case with  $p$  equal to 1, 2, and  $\infty$ , the detail is important. The detailed statement of the theorem allows us to conclude, among other things, that if a sequence of functions is convergent in  $L^{p_1}$  and in  $L^{p_2}$ , then the limit functions in the two spaces are equal almost everywhere.



PROOF. We may assume that  $p < \infty$ , the case  $p = \infty$  having been handled in Theorem 5.58. The argument for  $1 \leq p < \infty$  is word-for-word the same as in the proof for  $p = 1$  and  $p = 2$  of Theorem 5.58.  $\square$

In Section V.9 the inequality  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$  for  $p$  equal to 1, 2, or  $\infty$  says in words that “the norm of a sum is  $\leq$  the sum of the norms.” In that section we obtained a generalization for those values of  $p$ , saying that “the norm of an integral is  $\leq$  the integral of the norms.” The generalization continues to be valid for the other  $p$ 's under study; the proof amounts to a direct derivation from Hölder's inequality.

**Theorem 9.7** (Minkowski's inequality for integrals). Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, and let  $1 \leq p \leq \infty$ . If  $f$  is measurable on  $X \times Y$ , then

$$\left\| \int_X f(x, y) d\mu(x) \right\|_{p, d\nu(y)} \leq \int_X \|f(x, y)\|_{p, d\nu(y)} d\mu(x)$$

in the following sense: The integrand on the right side is measurable. If the integral on the right is finite, then for almost every  $y$  [ $d\nu$ ] the integral on the left is defined; when it is redefined to be 0 for the exceptional  $y$ 's, then the formula holds.

PROOF. Theorem 5.60 handles  $p = 1$  and  $p = \infty$ , and we may assume that  $1 < p < \infty$ . The measurability question is handled for  $1 < p < \infty$  in the same way as in Theorem 5.60 for  $p = 2$ . In proving the inequality, we may assume without loss of generality that  $f \geq 0$ . The generalization of the computation in the proof of Theorem 9.3 makes use of Fubini's Theorem and proceeds as follows:

$$\begin{aligned} & \int_Y \left| \int_X f(x, y) d\mu(x) \right|^p d\nu(y) \\ &= \int_Y \left| \int_X f(x, y) d\mu(x) \right| \left| \int_X f(x', y) d\mu(x') \right|^{p-1} d\nu(y) \\ &= \int_X \left\{ \int_Y f(x, y) \left| \int_X f(x', y) d\mu(x') \right|^{p-1} d\nu(y) \right\} d\mu(x) \\ &\leq \int_X \left( \int_Y |f(x, y)|^p d\nu(y) \right)^{1/p} \\ &\quad \times \left( \int_Y \left| \int_X f(x', y) d\mu(x') \right|^{(p-1)p'} d\nu(y) \right)^{1/p'} d\mu(x) \\ &= \left( \int_X \|f(x, y)\|_{p, d\nu(y)}^p d\mu(x) \right) \left( \int_Y \left| \int_X f(x', y) d\mu(x') \right|^p d\nu(y) \right)^{1/p'}. \end{aligned}$$

The next-to-last step uses Hölder's inequality (Theorem 9.2), and the last step uses the fact that  $(p-1)p' = p$ .

In order to complete the proof, we need to be able to divide by the factor  $\left( \int_Y \left| \int_X f(x', y) d\mu(x') \right|^p d\nu(y) \right)^{1/p'}$ . There is no problem with the theorem if

this factor is 0, since then the left side of the inequality of the theorem is 0. A problem occurs if this factor is infinite. Instead of trying to prove directly that this factor is finite (and hence the division is allowable), let us retreat to the special case that  $f$  is bounded and is equal to 0 off an abstract rectangle of finite  $\mu \times \nu$  measure. Then the factor in question is certainly finite, the division is allowable, and we obtain the inequality of the theorem. To handle general measurable  $f \geq 0$ , we do not attempt to justify this division. Instead, we observe that the validity of the inequality in the theorem when  $f$  is bounded and is equal to 0 off a set of finite  $\mu \times \nu$  measure implies the validity of the inequality in general, by a routine application of monotone convergence. This completes the proof.  $\square$

The last basic fact about  $L^p$  spaces is the identification of continuous linear functionals on  $L^p$ , at least when  $p$  is finite. Deriving the necessary tools for this analysis will require a digression, and we shall return to this topic in Section 5. Meanwhile, we can easily obtain one part of the identification of continuous linear functionals, as in Proposition 9.8 below. It amounts to a combination of Hölder's inequality and a converse, and it gives a way of computing  $L^p$  norms by starting with computations that are linear.

**Proposition 9.8.** Let  $(X, \mathcal{A}, \mu)$  be any measure space, let  $1 \leq p \leq \infty$ , and let  $p'$  be the dual index. If  $p < \infty$ , then

$$\|f\|_p = \sup_{\substack{g \in L^{p'} \\ \|g\|_{p'} \leq 1}} \left| \int_X fg \, d\mu \right|,$$

and this equality remains valid for  $p = \infty$  if  $\mu$  is  $\sigma$ -finite.

REMARK. The equality can fail when  $p = \infty$  and  $\mu$  is not  $\sigma$ -finite. Problem 4 at the end of the chapter gives an example.

PROOF. With  $1 \leq p \leq \infty$ , if  $g$  is in  $L^{p'}$  with  $\|g\|_{p'} \leq 1$ , then Hölder's inequality gives  $|\int fg \, d\mu| \leq \int |fg| \, d\mu \leq \|f\|_p \|g\|_{p'} \leq \|f\|_p$ . Taking the supremum over  $g$  with  $\|g\|_{p'} \leq 1$  shows that  $\sup_g |\int fg \, d\mu| \leq \|f\|_p$ .

For the reverse inequality we may assume that  $\|f\|_p \neq 0$ . First suppose that  $1 < p < \infty$ . Define  $g(x)$  by

$$g(x) = \begin{cases} \|f\|_p^{-(p-1)} \overline{f(x)} |f(x)|^{p-2} & \text{if } f(x) \neq 0, \\ 0 & \text{if } f(x) = 0. \end{cases}$$

Then  $\int |g(x)|^{p'} \, d\mu = \|f\|_p^{-(p-1)p'} \int |f(x)|^{(p-1)p'} \, d\mu = \|f\|_p^{-p} \int |f(x)|^p \, d\mu = 1$ . For this  $g$ , we have  $|\int fg \, d\mu| = \|f\|_p^{-(p-1)} \int |f|^p \, d\mu = \|f\|_p$ . Thus the supremum over the relevant  $g$ 's of  $|\int fg \, d\mu|$  is  $\geq \|f\|_p$ .

Next suppose that  $p = 1$ . If we define  $g(x)$  to be  $\overline{f(x)}/|f(x)|$  when  $f(x) \neq 0$  and to be 0 when  $f(x) = 0$ , then  $\|g\|_\infty = 1$  and  $|\int fg d\mu| = \int |f|^2/|f| d\mu = \|f\|_1$ , and the supremum over  $g$  of  $|\int fg d\mu|$  is  $\geq \|f\|_1$ .

Finally suppose that  $p = \infty$ . Let  $\epsilon > 0$  be given with  $\epsilon \leq \|f\|_\infty$ , and let  $E$  be the set where  $|f(x)| \geq \|f\|_\infty - \epsilon$ . Since  $\mu$  is  $\sigma$ -finite, there must exist a subset of  $E$  with nonzero finite measure. If  $F$  is such a subset and if  $g(x)$  is defined to be  $\mu(F)^{-1}\overline{f(x)}/|f(x)|$  when  $x$  is in  $F$  and to be 0 when  $x$  is in  $F^c$ , then  $\|g\|_1 = 1$  and  $|\int_X fg d\mu| = \mu(F)^{-1} \int_F |f| d\mu \geq \|f\|_\infty - \epsilon$ . Thus the supremum over  $g$  of  $|\int_X fg d\mu|$  is  $\geq \|f\|_\infty - \epsilon$ . Since  $\epsilon$  is arbitrary, the supremum over  $g$  of  $|\int_X fg d\mu|$  is  $\geq \|f\|_\infty$ .  $\square$

## 2. Convolution Involving $L^p$

In this section we collect results about  $L^p$  spaces that extend facts proved about  $L^1$ ,  $L^2$ , and  $L^\infty$  in the first three sections of Chapter VI.

**Proposition 9.9.** If  $\mu$  is a Borel measure on a nonempty open set  $V$  in  $\mathbb{R}^N$  and if  $1 \leq p < \infty$ , then

- (a)  $C_{\text{com}}(V)$  is dense in  $L^p(V, \mu)$ ,
- (b) the smallest closed subspace of  $L^p(V, \mu)$  containing all indicator functions of compact subsets of  $V$  is  $L^p(V, \mu)$  itself,
- (c)  $L^p(V, \mu)$  is separable.

PROOF. Parts (a) and (b) are proved from Lemma 6.22c, the regularity of  $\mu$  (Theorem 6.25), Proposition 9.4, and Proposition 5.56 by the same kind of argument as for Corollary 6.4. Part (c) is obtained as a consequence in the same way that Corollary 6.27d follows from the other parts of that corollary.  $\square$

The remaining results in this section concern Lebesgue measure in  $\mathbb{R}^N$ , and the  $L^p$  spaces are understood to be  $L^p(\mathbb{R}^N, \{\text{Borel sets}\}, dx)$ .

**Proposition 9.10.** Let  $1 < p < \infty$ , and let  $p'$  be the dual index. Convolution is defined in the following additional cases beyond those listed in Proposition 6.14, and the indicated inequalities hold:

- (e) for  $f$  in  $L^1(\mathbb{R}^N, dx)$  and  $g$  in  $L^p(\mathbb{R}^N, dx)$ , and then  $\|f * g\|_p \leq \|f\|_1 \|g\|_p$ ;  
for  $f$  in  $L^p(\mathbb{R}^N, dx)$  and  $g$  in  $L^1(\mathbb{R}^N, dx)$ , and then  $\|f * g\|_p \leq \|f\|_p \|g\|_1$ ;
- (f) for  $f$  in  $L^p(\mathbb{R}^N, dx)$  and  $g$  in  $L^{p'}(\mathbb{R}^N, dx)$ , and then  $\|f * g\|_\infty \leq \|f\|_p \|g\|_{p'}$ ;  
for  $f$  in  $L^{p'}(\mathbb{R}^N, dx)$  and  $g$  in  $L^p(\mathbb{R}^N, dx)$ , and then  $\|f * g\|_\infty \leq \|f\|_{p'} \|g\|_p$ .

PROOF. The two conclusions in (e) follow from Minkowski's inequality for integrals (Theorem 9.7) in the same way that the special case of  $p = 2$  was proved in Proposition 6.14 from Theorem 5.60. The two conclusions in (f) follow from Hölder's inequality (Theorem 9.2) in the same way that the special case  $p = p' = 2$  was proved in Proposition 6.14 from the Schwarz inequality.  $\square$

**Proposition 9.11.** If  $1 \leq p < \infty$ , then translation of a function is continuous in the translation parameter in  $L^p(\mathbb{R}^N, dx)$ . In other words, if  $f$  is in  $L^p(\mathbb{R}^N, dx)$ , then  $\lim_{h \rightarrow 0} \|\tau_{t+h}f - \tau_t f\|_p = 0$  for all  $t$ .

PROOF. This follows from the denseness of  $C_{\text{com}}(\mathbb{R}^N)$  in  $L^p(\mathbb{R}^N, dx)$  (Proposition 9.9a) and is proved in the same way that Proposition 6.16 is derived from Corollary 6.4a.  $\square$

**Proposition 9.12.** Let  $1 \leq p \leq \infty$ , and let  $p'$  be the dual index. Then the convolution of an  $L^p$  function with an  $L^{p'}$  function results in an everywhere-defined bounded uniformly continuous function, not just an  $L^\infty$  function. Moreover,

$$\|f * g\|_{\text{sup}} \leq \|f\|_p \|g\|_{p'}.$$

PROOF. This extends Proposition 6.18 and is derived for  $1 < p < \infty$  from Propositions 9.10 and 9.11 in the same way that Proposition 6.18 is derived for  $p = 2$  from Propositions 6.14 and 6.16.  $\square$

**Theorem 9.13.** Let  $\varphi$  be in  $L^1(\mathbb{R}^N, dx)$ , define

$$\varphi_\varepsilon(x) = \varepsilon^{-N} \varphi(\varepsilon^{-1}x) \quad \text{for } \varepsilon > 0,$$

and put  $c = \int_{\mathbb{R}^N} \varphi(x) dx$ . If  $f$  is in  $L^p(\mathbb{R}^N, dx)$  with  $1 \leq p < \infty$ , then

$$\lim_{\varepsilon \downarrow 0} \|\varphi_\varepsilon * f - cf\|_p = 0.$$

PROOF. This is derived from Minkowski's inequality for integrals (Theorem 9.7) and the continuity of translation in  $L^p$  (Proposition 9.11) in the same way that Theorem 6.20a is derived for  $p = 2$  from Theorem 5.60 and Proposition 6.16.  $\square$

### 3. Jordan and Hahn Decompositions

Now we digress before returning in Section 5 to the subject of continuous linear functionals on  $L^p$  spaces. The subject of the present section is decompositions of additive and completely additive real-valued set functions into positive and negative parts. This material will be applied in Section 4 to obtain the Radon–Nikodym Theorem, an abstract generalization of some consequences of Lebesgue’s theory of differentiation of integrals. In turn, we shall use the Radon–Nikodym Theorem in Section 5 to address the subject of continuous linear functionals on  $L^p$  spaces.

A real-valued additive set function  $\nu$  on an algebra of sets is said to be **bounded** if  $|\nu(E)| \leq C$  for all  $E$  in the algebra. A real-valued completely additive set function on a  $\sigma$ -algebra of sets is said to be a **signed measure**.

**Theorem 9.14** (Jordan decomposition). Let  $\nu$  be a bounded additive set function on an algebra  $\mathcal{A}$  of sets, and define set functions  $\nu^+$  and  $\nu^-$  on  $\mathcal{A}$  by

$$\nu^+(E) = \sup_{\substack{F \subseteq E, \\ F \in \mathcal{A}}} \nu(F) \quad \text{and} \quad \nu^-(E) = - \inf_{\substack{F \subseteq E, \\ F \in \mathcal{A}}} \nu(F).$$

Then  $\nu^+$  and  $\nu^-$  are nonnegative bounded additive set functions on  $\mathcal{A}$  such that  $\nu = \nu^+ - \nu^-$ . They are completely additive if  $\nu$  is completely additive. In any event, the decomposition  $\nu = \nu^+ - \nu^-$  is minimal in the sense that an equality  $\nu = \mu^+ - \mu^-$  in which  $\mu^+$  and  $\mu^-$  are nonnegative bounded additive set functions must have  $\nu^+ \leq \mu^+$  and  $\nu^- \leq \mu^-$ .

PROOF. First let us see that  $\nu^+$  is additive always. In fact, let  $E_1$  and  $E_2$  be disjoint members of  $\mathcal{A}$ . If  $F \subseteq E_1 \cup E_2$ , then the additivity of  $\nu$  implies that  $\nu(F) = \nu(F \cap E_1) + \nu(F \cap E_2) \leq \nu^+(E_1) + \nu^+(E_2)$ . Hence

$$\nu^+(E_1 \cup E_2) \leq \nu^+(E_1) + \nu^+(E_2).$$

On the other hand, if  $F_1 \subseteq E_1$  and  $F_2 \subseteq E_2$ , then  $\nu(F_1) + \nu(F_2) = \nu(F_1 \cup F_2) \leq \nu^+(E_1 \cup E_2)$ . Taking the supremum over  $F_1$  and then over  $F_2$  gives

$$\nu^+(E_1) + \nu^+(E_2) \leq \nu^+(E_1 \cup E_2).$$

Thus  $\nu^+$  is additive.

Second let us see that  $\nu^+$  is completely additive if  $\nu$  is completely additive. Let  $E_n$  be a disjoint sequence of sets in  $\mathcal{A}$  whose union  $E$  is in  $\mathcal{A}$ . If  $F \subseteq E$ , then the complete additivity of  $\nu$  implies that  $\nu(F) = \sum_n \nu(F \cap E_n) \leq \sum_n \nu^+(E_n)$ . Hence  $\nu^+(E) \leq \sum_{n=1}^{\infty} \nu^+(E_n)$ . On the other hand, the fact that  $\nu^+$  is nonnegative additive implies for every  $N$  that  $\sum_{n=1}^N \nu^+(E_n) = \nu^+(E_1 \cup \dots \cup E_N) \leq \nu^+(E)$ . Thus  $\sum_{n=1}^{\infty} \nu^+(E_n) \leq \nu^+(E)$ . Therefore  $\nu^+$  is completely additive.

Third let us see that  $\nu = \nu^+ - \nu^-$ . This equality will imply also that  $\nu^-$  is additive and that  $\nu^-$  is completely additive if  $\nu$  is completely additive. Form  $\nu(E) + \nu^-(E) = \nu(E) + \sup_{F \subseteq E} \{-\nu(F)\}$ ; we are to show that this equals  $\nu^+(E)$ . For any  $F \subseteq E$ , we have  $\nu(E) + (-\nu(F)) = \nu(E - F) \leq \nu^+(E)$ . Taking the supremum over  $F$  gives  $\nu(E) + \nu^-(E) \leq \nu^+(E)$ . In the reverse direction,  $F \subseteq E$  implies that  $\nu(F) = \nu(E) - \nu(E - F) \leq \nu(E) + \sup_{G \subseteq E} \{-\nu(G)\} = \nu(E) + \nu^-(E)$ . Taking the supremum over  $F$  gives  $\nu^+(E) \leq \nu(E) + \nu^-(E)$ . This proves the decomposition  $\nu = \nu^+ - \nu^-$ .

Finally we prove the minimality of the decomposition. Let  $\nu = \mu^+ - \mu^-$  with  $\mu^+$  and  $\mu^-$  nonnegative additive. If  $F \subseteq E$ , then we can write  $\nu(F) = \mu^+(F) - \mu^-(F) \leq \mu^+(F) \leq \mu^+(E)$ . Taking the supremum over  $F$  gives  $\nu^+(E) \leq \mu^+(E)$ . Similarly  $\nu^- \leq \mu^-$ .  $\square$

**Theorem 9.15** (Hahn decomposition). If  $\nu$  is a bounded signed measure on a  $\sigma$ -algebra  $\mathcal{A}$  of subsets of  $X$ , then there exist disjoint measurable sets  $P$  and  $N$  in  $\mathcal{A}$  with  $X = P \cup N$  such that  $\nu(E) \geq 0$  for all sets  $E \subseteq P$  and  $\nu(E) \leq 0$  for all sets  $E \subseteq N$ .

PROOF. Write  $\nu = \nu^+ - \nu^-$  as in Theorem 9.14. If  $\epsilon > 0$  is given, choose  $A$  in  $\mathcal{A}$  with  $\nu(A) \geq \nu^+(X) - \epsilon$ . Then

$$\nu^-(A) = \nu^+(A) - \nu(A) \leq \nu^+(A) - \nu^+(X) + \epsilon \leq \epsilon$$

$$\text{and } \nu^+(A^c) = \nu^+(X) - \nu^+(A) \leq \nu(A) + \epsilon - \nu^+(A) \leq \epsilon.$$

By taking  $P_0 = A$  and  $N_0 = A^c$ , we see that for any  $\epsilon > 0$  we can write  $X = P_0 \cup N_0$  disjointly with  $\nu^+(N_0) \leq \epsilon$  and  $\nu^-(P_0) \leq \epsilon$ .

For  $n \geq 1$ , write  $X = P_n \cup N_n$  disjointly with  $\nu^+(N_n) \leq 2^{-n}$  and  $\nu^-(P_n) \leq 2^{-n}$ . Define

$$P = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} P_m \quad \text{and} \quad N = P^c = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} N_m.$$

These sets are in  $\mathcal{A}$  since  $\mathcal{A}$  is a  $\sigma$ -algebra. Theorem 9.14 shows that  $\nu^-$  is completely additive, and hence  $\nu^-(P) \leq \sum_{n=1}^{\infty} \nu^-(\bigcap_{m=n}^{\infty} P_m)$ . The right side is 0 since  $\nu^-(\bigcap_{m=n}^{\infty} P_m) \leq \nu^-(P_{n+k}) \leq 2^{-(n+k)}$  for all  $k \geq 0$ , and therefore  $\nu^-(P) = 0$ . In addition, every  $n$  has  $\nu^+(N) \leq \nu^+(\bigcup_{m=n}^{\infty} \nu^+(N_m)) \leq \sum_{m=n}^{\infty} \nu^+(N_m) \leq \sum_{m=n}^{\infty} 2^{-m} = 2^{-n+1}$ , and therefore  $\nu^+(N) = 0$ .  $\square$

#### 4. Radon-Nikodym Theorem

The Lebesgue decomposition of Chapter VII says that any Stieltjes measure  $\mu$  on the line decomposes as  $\mu(E) = \int_E f dx + \mu_s$  with  $\mu_s = \mu_{cs} + \mu_d$  concentrated

on a Borel set of Lebesgue measure 0. The function  $f$  is obtained in that chapter as the derivative almost everywhere of the distribution function of  $\mu$ , hence as the limit of  $\mu(I)/m(I)$  as intervals  $I$  shrink to a point; here  $m$  is Lebesgue measure. In this formulation of the result, the geometry of the line plays an essential role, and attempts to generalize to abstract settings the construction of  $f$  from limits of  $\mu(I)/m(I)$  have not been fruitful.

Nevertheless, the Lebesgue decomposition itself turns out to be a general measure-theory theorem, valid for any two measures in place of  $\mu$  and  $dx$ , as long as suitable finiteness conditions are satisfied. For a reinterpretation of the results of Chapter VII, the heart of the matter is that one can tell in advance which  $\mu$ 's have  $\mu(E) = \int_E f dx$  with the singular term  $\mu_s$  absent. The answer is given by the equivalent conditions of Proposition 7.11, which are taken in that chapter as a definition of “absolute continuity” of  $\mu$  with respect to  $dx$ . The remarkable fact is that those conditions continue to be equivalent when any two finite measures replace  $\mu$  and  $dx$ . This is the content of the Radon–Nikodym Theorem, which we shall prove in this section, and then a version of the Lebesgue decomposition will follow as a consequence.

Let  $X$  be a nonempty set, and let  $\mathcal{A}$  be a  $\sigma$ -algebra of subsets of  $X$ . If  $\mu$  and  $\nu$  are measures defined on  $\mathcal{A}$ , we say that  $\nu$  is **absolutely continuous** with respect to  $\mu$ , written  $\nu \ll \mu$ , if  $\nu(E) = 0$  whenever  $\mu(E) = 0$ .

**Theorem 9.16** (Radon–Nikodym Theorem). Let  $(X, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space, and let  $\nu$  be a  $\sigma$ -finite measure on  $\mathcal{A}$  with  $\nu \ll \mu$ . Then there exists a measurable  $f \geq 0$  such that  $\nu(E) = \int_E f d\mu$  for all  $E$  in  $\mathcal{A}$ , and  $f$  is unique up to a set of  $\mu$  measure 0.

The Radon–Nikodym Theorem has two chief initial applications. One is to the identification of continuous linear functionals on  $L^p$  for  $1 \leq p < \infty$ , and the other is to the construction of “conditional expectation” in probability theory. The application to  $L^p$  will be given in Section 5, and the application to conditional expectation appears in Problems 23–26 at the end of the chapter.

In both applications one needs a version of the theorem in which the completely additive set function  $\nu$  is complex-valued but not necessarily  $\geq 0$ . We take up this extension of the theorem later in this section.

Most of the effort in the proof goes into showing existence when  $\mu$  and  $\nu$  are both finite measures, as we shall see. In this setting we can quickly use the Hahn decomposition (Theorem 9.15) to get an idea how to construct  $f$ : Imagine that  $\nu(E) = \int_E f d\mu$  for all  $E$ . Fix  $c$  and  $d$ , and let  $S$  be the set of  $x$ 's where  $c \leq f(x) < d$ . On any subset  $E$  of  $S$ , we then have  $c\mu(E) \leq \nu(E) \leq d\mu(E)$ . In other words, the bounded signed measure  $\nu - c\mu$  is  $\geq 0$  on every subset of  $S$ , and the bounded signed measure  $\nu - d\mu$  is  $\leq 0$  on every subset of  $S$ . Let

$X = P_c \cup N_c$  and  $X = P_d \cup N_d$  be Hahn decompositions of  $\nu - c\mu$  and  $\nu - d\mu$  with respect to  $\mu$ . Then it is reasonable to expect  $S$  to be  $P_c \cap N_d$ . In particular,  $c$  is a good lower bound for the values of  $f$  on  $S$ . It is easy to imagine that we can use this process repeatedly to obtain a monotone sequence of functions  $f_n \geq 0$  tending to the desired function  $f$ .

Actually, this argument can be pushed through, but handling the details is a good deal more complicated than one might at first suppose. The reason is that a Hahn decomposition is not necessarily unique. Sets of measure 0 account for the nonuniqueness, and the particular measures yielding these sets of measure 0 are constantly changing. The complication is that one has to adjust all the Hahn decompositions to satisfy various compatibility conditions. We shall not pursue this idea because a simpler proof is available.

PROOF OF UNIQUENESS IN THEOREM 9.16. Suppose that  $f$  and  $g$  are non-negative measurable functions with  $\int_E f d\mu = \int_E g d\mu$  for every measurable  $E$ . If  $F$  is a set where the equal integrals  $\int_F f d\mu$  and  $\int_F g d\mu$  are finite, then  $\int_{E \cap F} (f - g) d\mu = 0$  for every measurable subset  $E \cap F$  of  $F$ . If  $E$  is taken as the set where  $f > g$ , then Corollary 5.23 shows that  $f = g$  a.e. on  $E \cap F$ . Similarly  $f = g$  a.e. on the set  $E^c \cap F$ , where  $f \leq g$ . Thus  $f = g$  a.e. on  $F$ . By  $\sigma$ -finiteness of  $\mu$  and  $\nu$ , we can write  $X = \bigcup_{n=1}^{\infty} X_n$  disjointly with  $\mu(X_n)$  and  $\nu(X_n)$  finite for all  $n$ . Taking  $F$  equal to each  $X_n$  in turn, we see that  $f = g$  a.e. on each  $X_n$ , and we conclude that  $f = g$  a.e. on  $X$ .  $\square$

PROOF OF EXISTENCE IN THEOREM 9.16 WHEN  $\mu$  AND  $\nu$  ARE FINITE. Let  $\mathcal{F}(\nu)$  be the set of all  $f \geq 0$  in  $L^1(X, \mu)$  such that  $\int_E f d\mu \leq \nu(E)$  for all sets  $E$  in  $\mathcal{A}$ . The zero function is in  $\mathcal{F}(\nu)$ , and thus it makes sense to define

$$C = \sup_{f \in \mathcal{F}(\nu)} \int_X f d\mu.$$

Let  $\{f_n\}$  be a sequence in  $\mathcal{F}(\nu)$  with  $\lim_n \int_X f_n d\mu = C$ .

Let us show that there is no loss of generality in assuming that the  $f_n$  satisfy  $f_1 \leq f_2 \leq \dots$ . To show this, it is enough to show that  $g$  and  $h$  in  $\mathcal{F}(\nu)$  implies that  $\max\{g, h\}$  is in  $\mathcal{F}(\nu)$ . We have

$$\begin{aligned} \int_E \max\{g, h\} d\mu &= \int_{E \cap \{g \geq h\}} g d\mu + \int_{E \cap \{g < h\}} h d\mu \\ &\leq \nu(E \cap \{g \geq h\}) + \nu(E \cap \{g < h\}) = \nu(E), \end{aligned}$$

and hence  $\max\{g, h\}$  is indeed in  $\mathcal{F}(\nu)$ .

With the  $f_n$ 's now increasing with  $n$ , put  $f(x) = \lim_n f_n(x)$ . Monotone convergence shows that  $f$  is in  $\mathcal{F}(\nu)$  and  $\int_X f d\mu = C$ . Define

$$\nu_0(E) = \nu(E) - \int_E f d\mu.$$



Then  $\nu_0$  is a measure,  $\nu_0 \ll \mu$ , and the class  $\mathcal{F}(\nu_0)$  for  $\nu_0$  consists of 0 alone. We shall complete this part of the proof by showing that  $\nu_0 = 0$ .

If  $\nu_0 \neq 0$ , choose  $n$  large enough so that  $\nu_0(X) - \frac{1}{n}\mu(X) > 0$ , and put  $\nu'_0 = \nu_0 - \frac{1}{n}\mu$ . Let  $X = P \cup N$  be a Hahn decomposition for  $\nu'_0$  as in Theorem 9.15, and define  $g = \frac{1}{n}I_P$ . Then the calculation

$$\int_E \frac{1}{n}I_P d\mu = \frac{1}{n}\mu(P \cap E) = \nu_0(P \cap E) - \nu'_0(P \cap E) \leq \nu_0(P \cap E) \leq \nu_0(E)$$

shows that  $g$  is in  $\mathcal{F}(\nu_0)$ . Hence  $g = 0$  a.e.  $[d\mu]$ , and  $\mu(P) = 0$ . Since  $\nu_0 \ll \mu$ , we obtain  $\nu_0(P) = 0$  and therefore also  $\nu'_0(P) = 0$ . Then  $\nu'_0 \leq 0$ , and we must have  $\nu_0(X) - \frac{1}{n}\mu(X) \leq 0$ . This contradicts the choice of  $n$ , and the proof of existence is complete when  $\mu$  and  $\nu$  are finite.  $\square$

PROOF OF EXISTENCE IN THEOREM 9.16 WHEN  $\mu$  AND  $\nu$  ARE  $\sigma$ -FINITE. Write  $X$  as the countable disjoint union of sets  $X_n$  such that  $\mu(X_n)$  and  $\nu(X_n)$  are both finite. If we put  $\mu_n(E) = \mu(E \cap X_n)$  and  $\nu_n(E) = \nu(E \cap X_n)$ , then  $\mu_n$  and  $\nu_n$  are finite measures such that  $\nu_n \ll \mu_n$ , and the above special case produces functions  $f_n \geq 0$  such that  $\nu_n(E) = \int_E f_n d\mu_n$  for all  $E$ . Since  $\nu_n(X_n^c) = 0$ , we may assume that  $f_n(x) = 0$  for  $x \notin X_n$ . Let  $f \geq 0$  be the measurable function that equals  $f_n$  on  $X_n$  for each  $n$ . Then our formula reads  $\nu(E \cap X_n) = \int_{E \cap X_n} f d\mu$  for all  $n$  and for all  $E$ . Summing on  $n$ , we obtain  $\nu(E) = \int_E f d\mu$  for all  $E$  in  $\mathcal{A}$ .  $\square$

**Corollary 9.17.** Let  $(X, \mathcal{A}, \mu)$  be a finite measure space, and let  $\nu$  be a (real-valued) bounded signed measure on  $\mathcal{A}$  with  $\nu \ll \mu$  in the sense that  $\mu(E) = 0$  implies  $\nu(E) = 0$ . Then there exists a function  $f$  in  $L^1(X, \mu)$  such that  $\nu(E) = \int_E f d\mu$  for all  $E$  in  $\mathcal{A}$ , and  $f$  is unique up to a set of  $\mu$  measure 0.

PROOF. Let  $\nu = \nu^+ - \nu^-$  be the Jordan decomposition of  $\nu$  as in Theorem 9.14, and let  $X = P \cup N$  be a Hahn decomposition of  $\nu$  as in Theorem 9.15. Suppose  $\mu(E) = 0$ . Since  $\mu$  is nonnegative, we obtain  $\mu(E \cap P) = 0$  and  $\mu(E \cap N) = 0$ , and the assumption  $\nu \ll \mu$  forces

$$0 = \nu(E \cap P) = \nu^+(E \cap P) = \nu^+(E)$$

and 
$$0 = \nu(E \cap N) = -\nu^-(E \cap N) = -\nu^-(E).$$

Therefore  $\nu^+ \ll \mu$  and  $\nu^- \ll \mu$ , and the corollary follows by applying Theorem 9.16 to  $\nu^+$  and  $\nu^-$  separately.  $\square$

**Corollary 9.18.** Let  $(X, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space, and let  $\nu$  be a  $\sigma$ -finite measure on  $\mathcal{A}$ . Then there exist a measurable  $f \geq 0$  and a set  $S$  in  $\mathcal{A}$  with  $\mu(S) = 0$  such that  $\nu = f d\mu + \nu_s$ , where  $\nu_s(E) = \nu(E \cap S)$ . The measure  $\nu_s$  is unique, and the function  $f$  is unique up to a set of  $\mu$  measure 0.

REMARK. The measure  $\nu_s$ , being carried on a set of  $\mu$  measure 0, is said to be **singular** with respect to  $\mu$ . The measure  $f d\mu$  is, of course, absolutely continuous with respect to  $\mu$ . The decomposition of  $\nu$  into the sum of an absolutely continuous part and a singular part is called the **Lebesgue decomposition** of  $\nu$  with respect to  $\mu$ . The corollary asserts that this decomposition of measures exists and is unique.

PROOF. As in the proof of Theorem 9.16, we can reduce matters to the case that  $\nu$  and  $\mu$  are both finite, and it is therefore enough to handle this special case. Among all sets  $E$  in  $\mathcal{A}$  with  $\mu(E) = 0$ , let  $C$  be the supremum of  $\nu(E)$ . The number  $C$  is finite, being  $\leq \nu(X)$ . Choose a sequence of sets  $E_n$  in  $\mathcal{A}$  with  $\mu(E_n) = 0$  and  $\nu(E_n)$  increasing to  $C$ . Without loss of generality, we may assume that  $E_1 \subseteq E_2 \subseteq \dots$ . Put  $S = \bigcup_n E_n$ . Proposition 5.2 shows that  $\mu(S) = 0$  and  $\nu(S) = C$ . Define  $\nu_a(E) = \nu(E \cap S^c)$  and  $\nu_s(E) = \nu(E \cap S)$ . Then  $\nu_a$  and  $\nu_s$  are measures, and  $\nu = \nu_a + \nu_s$ .

Certainly  $\nu_s$  is singular with respect to  $\mu$ , being carried on the set  $S$  of  $\mu$  measure 0. Let us see that  $\nu_a$  is absolutely continuous. Thus suppose that  $\mu(E) = 0$ . Then  $\mu(S \cup E) \leq \mu(S) + \mu(E) = 0$ , and the construction of  $C$  shows that  $\nu(S \cup E) \leq C = \nu(S)$ . Therefore  $\nu(S \cup E) - \nu(S) \leq 0$  and  $\nu(S \cup E) - \nu(S) = 0$ . Hence  $0 = \nu(S \cup E) - \nu(S) = \nu(E - S) = \nu(E \cap S^c) = \nu_a(E)$ , and  $\nu_a$  is indeed absolutely continuous. Applying the Radon–Nikodym Theorem (Theorem 9.16), we obtain  $\nu = \nu_a + \nu_s = f d\mu + \nu_s$ . This proves existence.

For uniqueness, suppose that we have  $\nu = f d\mu + \nu_s = f^\# d\mu + \nu_s^\#$  with  $\nu_s$  and  $\nu_s^\#$  carried on respective sets  $S$  and  $S^\#$  of  $\mu$  measure 0. The functions  $f$  and  $f^\#$  are integrable with respect to  $\mu$ , and we have  $\int_E (f - f^\#) d\mu = \nu_s^\#(E) - \nu_s(E)$ . Taking  $E$  to be any subset  $T$  in  $\mathcal{A}$  of  $S \cup S^\#$ , we see that  $0 = \nu_s^\#(T) - \nu_s(T)$ . Therefore  $\nu_s^\#(T) = \nu_s(T)$  whenever  $T \subseteq S \cup S^\#$ . On  $(S \cup S^\#)^c$ , we have  $\nu_s^\#((S \cup S^\#)^c) = \nu_s((S \cup S^\#)^c) = 0$ . Therefore  $\nu_s^\# = \nu_s$ . The uniqueness of the function part follows from the uniqueness in the Radon–Nikodym Theorem, which is part of the statement of that theorem (Theorem 9.16).  $\square$

## 5. Continuous Linear Functionals on $L^p$

We return to the question of identifying the continuous linear functionals on  $L^p$  spaces. Let  $(X, \mathcal{A}, \mu)$  be a fixed  $\sigma$ -finite measure space. The space  $L^p(X, \mu)$  is a normed linear space and, as such, is both a vector space and a metric space. The scalars may be real or complex.

Recall from Section V.9 that a linear functional on  $L^p(X, \mu)$  is a linear function from  $L^p(X, \mu)$  into the scalars. Proposition 5.57 shows that a linear functional  $x^*$  is continuous if and only if it is bounded in the sense that  $|x^*(f)| \leq C \|f\|_p$  for some constant  $C$  and all  $f$  in  $L^p$ . The inequality  $|x^*(f)| \leq C \|f\|_p$  holds for

all  $f$  in  $L^p$  if and only if it holds for all  $f$  with  $\|f\|_p \leq 1$ , if and only if it holds for all  $f$  with  $\|f\|_p = 1$ . If there is such a constant  $C$ , then the finite number

$$\|x^*\| = \sup_{\|f\|_p \leq 1} |x^*(f)| = \sup_{\|f\|_p = 1} |x^*(f)|$$

is the least such constant  $C$  and is called the **norm** of  $x^*$ . Since  $\|x^*\|$  is one such constant  $C$ , we have

$$|x^*(f)| \leq \|x^*\| \|f\|_p.$$

Let  $p'$  be the dual index to  $p$ , defined by  $\frac{1}{p} + \frac{1}{p'} = 1$ . Each member  $g$  of  $L^{p'}(X, \mu)$  provides an example of a continuous linear functional on  $L^p$  by the formula  $x^*(f) = \int_X fg \, d\mu$ . The linear functional  $x^*$  is bounded, hence continuous, as a consequence of Hölder's inequality:  $|\int_X fg \, d\mu| \leq \|g\|_{p'} \|f\|_p$ . This inequality shows that  $\|x^*\| \leq \|g\|_{p'}$ . Proposition 9.8 shows that equality  $\|x^*\| = \|g\|_{p'}$  holds if  $\mu$  is  $\sigma$ -finite and  $1 \leq p \leq \infty$ .

Theorem 9.19 gives a converse when  $1 \leq p < \infty$ , saying that there are no other examples of continuous linear functionals if  $\mu$  is  $\sigma$ -finite. By contrast, there *can* be other examples in the case of  $L^\infty(X, \mu)$ . For example, for the situation in which  $X$  is the set of positive integers and  $\mathcal{A}$  consists of all subsets of  $X$  and  $\mu$  is the counting measure, Problems 39–43 at the end of Chapter V show how to construct a bounded additive set function on  $\mathcal{A}$  that is not completely additive, and they show how this set function leads to a notion of integration (hence a linear functional) on this  $L^\infty$  space; this linear functional is not given by an  $L^1$  function.

**Theorem 9.19** (Riesz Representation Theorem for  $L^p$ ). Let  $(X, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space, let  $1 \leq p < \infty$ , and let  $p'$  be the dual index to  $p$ . If  $x^*$  is a continuous linear functional on  $L^p(X, \mu)$ , then there exists a unique member  $g$  of  $L^{p'}(X, \mu)$  such that

$$x^*(f) = \int_X fg \, d\mu$$

for all  $f$  in  $L^p$ . For this function  $g$ ,  $\|x^*\| = \|g\|_{p'}$ .

REMARKS. For  $1 \leq p < \infty$ , Proposition 9.9 shows that  $L^p(V, \mu)$  is separable if  $\mu$  is a Borel measure on an open subset of  $\mathbb{R}^N$ . For this or any other setting in which any of these  $L^p$  spaces is separable, Alaoglu's Theorem (Theorem 5.58) says that any bounded sequence in  $L^p(V, \mu)^*$  has a weak-star convergent subsequence. Because of Theorem 9.19 we know what the members of the dual space are. Thus any bounded sequence in  $L^{p'}$  has a subsequence that is convergent weak-star against  $L^p$ . In effect we obtain a nonconstructive way of producing members of  $L^{p'}$ . Problem 8 at the end of the chapter will illustrate the usefulness of this technique.

PROOF OF UNIQUENESS. Write  $X = \bigcup_{n=1}^{\infty} X_n$  disjointly with  $\mu(X_n)$  finite for all  $n$ . If  $\int_X fg \, d\mu = 0$  for all  $f$  in  $L^p$ , then  $\int_X I_{A \cap X_n} g \, d\mu = 0$  for every measurable subset  $A$  of  $X$ . Taking  $A$  successively to be each of the sets where  $\operatorname{Re} g$  or  $\operatorname{Im} g$  is  $\geq 0$  or is  $\leq 0$  and applying Corollary 5.23, we see that  $g$  is 0 almost everywhere on  $X_n$  for each  $n$ . Hence  $g$  is 0 almost everywhere.  $\square$

PROOF OF EXISTENCE IF  $\mu(X)$  IS FINITE. Temporarily let us suppose that the underlying scalars are real. Define a set function  $\nu$  on  $\mathcal{A}$  by  $\nu(E) = x^*(I_E)$ ;  $\nu$  is well defined because every  $I_E$  is in  $L^p$ , and  $\nu$  is additive because  $x^*$  is linear. If  $E_n$  is an increasing sequence of measurable sets with union  $E$ , then  $\lim_n I_{E_n} = I_E$  pointwise, and hence  $\lim_n |I_E - I_{E_n}|^p = 0$  pointwise. By dominated convergence,  $\lim_n \|I_E - I_{E_n}\|_p = 0$ . Thus

$$|\nu(E) - \nu(E_n)| = |x^*(I_E - I_{E_n})| \leq \|x^*\| \|I_E - I_{E_n}\|_p,$$

and the right side has limit 0. By Proposition 5.2,  $\nu$  is completely additive. The set function  $\nu$  is bounded because  $|\nu(E)| = |x^*(I_E)| \leq \|x^*\| \|I_E\|_p = \|x^*\| (\mu(E))^{1/p} \leq \|x^*\| (\mu(X))^{1/p}$ , and it satisfies  $\nu \ll \mu$  because if  $\mu(E) = 0$ , then  $I_E$  is the 0 function of  $L^p$  and thus  $\nu(E) = x^*(I_E) = x^*(0) = 0$ . By the Radon–Nikodym Theorem in the form of Corollary 9.17, there exists an integrable real-valued function  $g$  such that  $\nu(E) = \int_E g \, d\mu$  for all  $E$ , i.e.,

$$x^*(I_E) = \int_X I_E g \, d\mu \quad \text{for every measurable set } E.$$

By linearity, this equality extends to show that  $x^*(s) = \int_X s g \, d\mu$  for every simple function  $s$ . Let  $f \geq 0$  be in  $L^p$ , and choose an increasing sequence  $\{s_n\}$  of simple functions  $\geq 0$  with pointwise limit  $f$ . We shall show that  $fg$  is integrable and  $x^*(f) = \int fg \, d\mu$ . In fact, let  $A$  be the set where  $g(x) \geq 0$ . Then  $\lim_n |f I_A - s_n I_A|^p = 0$  pointwise, and hence  $\lim_n \|f I_A - s_n I_A\|_p = 0$  by dominated convergence. Since

$$|x^*(f I_A) - x^*(s_n I_A)| \leq \|x^*\| \|f I_A - s_n I_A\|_p$$

and since the right side tends to 0, the set  $\{x^*(s_n I_A)\}$  of numbers is bounded. Thus the set  $\{\int_X s_n I_A g \, d\mu\}$  of equal numbers is bounded. Since  $g \geq 0$  on  $A$ , the functions  $s_n I_A g$  increase to  $f I_A g$ , and thus  $\int_X f I_A g \, d\mu$  is finite by monotone convergence. In other words,  $fg^+$  is integrable. Similarly  $fg^-$  is integrable, and thus  $fg$  is integrable. Since  $\lim_n x^*(s_n I_A) = x^*(f I_A)$  and  $\lim_n \int_X s_n I_A g \, d\mu = \int_X f I_A g \, d\mu$  and since a similar result holds for  $g^-$ , we conclude that

$$x^*(f) = \int_X fg \, d\mu \quad \text{for all } f \geq 0 \text{ in } L^p.$$

This conclusion, now proved for  $f \geq 0$ , immediately extends by linearity to all  $f$  in  $L^p$  and completes the verification that  $x^*(f) = \int_X fg d\mu$  in the case that the scalars are real.

If the scalars are complex, we apply the above argument to the restrictions of  $\operatorname{Re} x^*$  and  $\operatorname{Im} x^*$  to the real-valued functions in  $L^p$ , obtaining real-valued functions  $g_1$  and  $g_2$  in  $L^{p'}$  with  $\operatorname{Re} x^*(f) = \int_X fg_1 d\mu$  and  $\operatorname{Im} x^*(f) = \int_X fg_2 d\mu$  for all real-valued  $f$ . Then  $x^*(f) = \int_X f(g_1 + ig_2) d\mu$  for all real-valued  $f$ , and it follows that this same equality is valid for all complex-valued  $f$ . Since  $g_1$  and  $g_2$  are in  $L^{p'}$ , so is  $g_1 + ig_2$ . This completes the verification that  $x^*(f) = \int_X fg d\mu$  for a suitable  $g$  in the case that the scalars are complex.

Finally Proposition 9.8 shows that  $\|x^*\| = \|g\|_{p'}$  and completes the proof of the theorem under the assumption that  $\mu(X)$  is finite.  $\square$

PROOF OF EXISTENCE IF  $\mu(X)$  IS  $\sigma$ -FINITE. Again we temporarily suppose that the underlying scalars are real. Since  $\mu$  is  $\sigma$ -finite, we can write  $X$  as the increasing union of sets  $E_n$  of finite measure. Let  $L_n^p$  be the set of members of  $L^p$  that vanish off  $E_n$ , and let  $x_n^*$  be the restriction of  $x^*$  to  $L_n^p$ . Find, by the special case just completed, a function  $g_n$  for each  $n$  such that  $x_n^*(f_n) = \int_{E_n} f_n g_n d\mu$  for all  $f_n$  in  $L_n^p$ . The already proved uniqueness result implies that the restriction of  $g_{n+1}$  to  $E_n$  equals  $g_n$  almost everywhere  $[d\mu]$ . Let  $g$  be the measurable function equal to  $g_1$  on  $E_1$  and equal to  $g_n$  on  $E_n - E_{n-1}$  if  $n \geq 2$ . Let  $A$  be the set where  $g(x) \geq 0$ , and let  $f \geq 0$  be in  $L^p$ . Then  $fI_{E_n \cap A}$  increases to  $fI_A$ , and dominated convergence implies that  $\lim_n \|fI_{E_n \cap A} - fI_A\|_p = 0$ . Since  $fI_{E_n \cap A} g$  increases pointwise to  $fI_A g$ , monotone convergence gives

$$\begin{aligned} \int_X fI_A g d\mu &= \lim_n \int_X fI_{E_n \cap A} g d\mu = \lim_n \int_X fI_{E_n \cap A} g_n d\mu \\ &= \lim_n x_n^*(fI_{E_n \cap A}) = \lim_n x^*(fI_{E_n \cap A}) = x^*(fI_A), \end{aligned}$$

the last equality holding since  $\|fI_{E_n \cap A} - fI_A\|_p$  tends to 0. Hence  $fg^+$  is integrable. By proceeding similarly with the set where  $g(x) < 0$  and by writing a general  $f$  as  $f = f^+ - f^-$ , we conclude that  $fg$  is integrable for every  $f$  in  $L^p$  and  $x^*(f) = \int_X fg d\mu$ , provided the scalars are real.

Again there is no difficulty in extending the argument to the case that the scalars are complex, and Proposition 9.8 shows that  $\|x^*\| = \|g\|_{p'}$ .  $\square$

## 6. Riesz–Thorin Convexity Theorem

This section and the next concern linear functions and some almost-linear functions between  $L^p$  spaces. We saw evidence in Proposition 9.5 that the  $L^p$  spaces behave collectively like a well-behaved family of spaces. That result specifically gave an upper bound for  $\|f\|_p$  in terms of  $\|f\|_{p_1}$  and  $\|f\|_{p_2}$  when  $p_1 \leq p \leq p_2$ .

It turns out that linear functions between pairs of  $L^p$  spaces satisfy inequalities of a similar sort.

At the level of this book, there are two classes of results in this direction. Results of the first kind, which are the subject of this section, touch on methods of complex analysis, address bounded linear operators only, and give estimates for a one-parameter family of operators that are sharp at the ends. The main result of this kind is the “Riesz<sup>1</sup>–Thorin Convexity Theorem,” given below as Theorem 9.19A. Results of the second kind use methods of real analysis and will be considered in Section 7.

Before formulating the Riesz–Thorin Convexity Theorem precisely, we give some definitions and make some observations. Let  $(X, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space, and let  $p$  be an index such that  $1 \leq p \leq \infty$ . Recall from Section 9.1 that members of  $L^p(X, \mu)$  are really equivalence classes of functions, two functions being equivalent if they differ on a set of measure 0. In spite of the formal definition in terms of equivalence classes, we use language that treats members of  $L^p(X, \mu)$  as genuine functions, and we expect no confusion to result. We shall make use of the fact that the vector subspace of simple functions in  $L^p(X, \mu)$  is dense in  $L^p(X, \mu)$ . The reason is that the closure of this vector subspace has to be a vector subspace and that all nonnegative functions in  $L^p(X, \mu)$  are approximable by nonnegative simple functions in  $L^p(X, \mu)$ . For  $1 \leq p < \infty$ , this approximation property is proved in Proposition 9.4a, which is a consequence of Proposition 5.11 and the Monotone Convergence Theorem. For  $p = \infty$ , monotone convergence is not invoked, but the convergence of the approximations in Proposition 5.11 is manifestly uniform for bounded measurable functions.

If  $1 \leq p < \infty$ , then the simple functions in  $L^p(X, \mu)$  are exactly all simple functions that vanish off some set of finite measure, the set depending on the function. The space of them is independent of  $p$ . For  $p = \infty$  and the case that  $\mu(X)$  is infinite, the constant function 1 is a simple function that is in  $L^\infty(X, \mu)$  but is not in the space of all simple functions that vanish off a set of finite measure.

Suppose that  $(Y, \mathcal{B}, \nu)$  is a second  $\sigma$ -finite measure space and that  $q$  is an index with  $1 \leq q \leq \infty$ . A linear operator  $T : L^p(X, \mu) \rightarrow L^q(Y, \nu)$  will be said to be of **type**  $(p, q)$  or **strong type**  $(p, q)$  if it is bounded, i.e., if  $\|Tf\|_q \leq M\|f\|_p$  for all  $f$  in  $L^p(X, \mu)$  and for some finite constant  $M$  independent of  $f$ . The least  $M$  for which this inequality holds is called the **norm** or **operator norm** of  $T$ . We allow the same terminology if  $T$  is defined only on a dense subspace of  $L^p(X, \mu)$  and if the same estimate holds for all  $f$  in that subspace. This usage on the dense subspace is consistent with the usage on the whole space because of the uniform

---

<sup>1</sup>The person in question here is Marcel Riesz, whose name is associated also with convergence of the partial sums of the Fourier series of an  $L^p$  function in  $L^p$  for  $1 < p < \infty$ . The other mentions of the name “Riesz” in this book, namely in connection with the Rising Sun Lemma of Section VII.1 and various results known as the Riesz Representation Theorem, refer to Frigyes Riesz.

continuity of  $T$  given in Proposition 5.57, the completeness of  $L^q(Y, \nu)$  given in Theorem 9.6, and the extendability of  $T$  boundedly to all of  $L^p(X, \mu)$  given in Proposition 2.47.

We are interested in the situation where we have two such estimates for the same linear function  $T$ . Thus let  $1 \leq p_0 \leq p_1 \leq \infty$ , and let  $q_0$  and  $q_1$  be two indices between 1 and  $\infty$ . We want to be able to say that  $T$  is of type  $(p_0, q_0)$  and also type  $(p_1, q_1)$ , and we have to make the two versions of  $T$  agree on the common domain. We can formulate matters this way: We suppose that  $T$  is a linear function from  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$  to the vector space of equivalence classes of measurable functions on  $Y$ , two functions being equivalent on  $Y$  if they differ on a set of  $\nu$  measure 0. If  $p_1 < \infty$ , then the vector subspace  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$  is dense in  $L^{p_0}(X, \mu)$  and dense in  $L^{p_1}(X, \mu)$  because it contains the space of simple functions that vanish off a set of finite measure. It is then meaningful to suppose that  $T$  is of type  $(p_0, q_0)$  and also type  $(p_1, q_1)$ , i.e., that  $T$  satisfies the estimates  $\|Tf\|_{q_0} \leq \|f\|_{p_0}$  and  $\|Tf\|_{q_1} \leq \|f\|_{p_1}$  for  $f$  in  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ .

We can ask about any other pairs  $(p, q)$  such that  $T$  is automatically of type  $(p, q)$ . Proposition 9.4c shows that members of  $L^p(X, \mu)$  are in the sum of  $L^{p_0}(X, \mu)$  and  $L^{p_1}(X, \mu)$  if  $p_0 \leq p \leq p_1$ . The thrust of the Riesz–Thorin Convexity Theorem is that if a linear operator  $T$  is of type  $(p_0, q_0)$  and type  $(p_1, q_1)$ , then  $T$  is also of type  $(p, q)$  for all pairs  $(p, q)$  such that  $(\frac{1}{p}, \frac{1}{q})$  lies on the line segment in the  $(\frac{1}{p}, \frac{1}{q})$  plane from  $(\frac{1}{p_0}, \frac{1}{q_0})$  to  $(\frac{1}{p_1}, \frac{1}{q_1})$ . The conclusion gives also some specific information about the norm of  $T$  on  $L^p(X, \mu)$ .

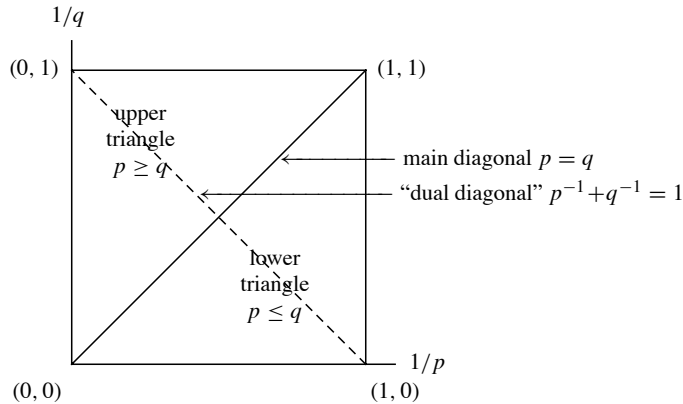


FIGURE 9.1. Geometric description of pairs  $(1/p, 1/q)$  occurring in the Riesz–Thorin Convexity Theorem.

**Theorem 9.19A** (Riesz–Thorin Convexity Theorem). Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, let  $1 \leq p_0 \leq p_1 \leq \infty$  be given, and

let  $q_0$  and  $q_1$  be indices between 1 and  $\infty$ . Suppose that  $T$  is a complex-linear function with domain  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$  taking values in the vector space of equivalence classes of measurable complex-valued functions on  $Y$ , two functions being equivalent if they differ on a set of  $\nu$  measure 0. Suppose further that  $T$  is of type  $(p_0, q_0)$  with bound  $M_{p_0, q_0}$  and type  $(p_1, q_1)$  with bound  $M_{p_1, q_1}$ . Then for  $0 \leq t \leq 1$ ,  $T$  is of type  $(p, q)$  if  $p$  and  $q$  satisfy

$$\frac{1}{p} = \frac{1-t}{p_0} + \frac{t}{p_1} \quad \text{and} \quad \frac{1}{q} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

Moreover, the bound  $M_{p, q}$  as an operator of type  $(p, q)$  has  $M_{p, q} \leq M_{p_0, q_0}^{1-t} M_{p_1, q_1}^t$ .

REMARKS.

(1) Domain indices in the proof are consistently labeled as  $p$ , possibly with a subscript, and range indices are consistently labeled as  $q$ , possibly with a subscript. See Figure 9.1, in which various pairs  $(p, q)$  are plotted in a plane with coordinates as  $(p^{-1}, q^{-1})$ . The theorem says that if  $T$  is bounded for two pairs  $(p, q)$ , then it is of type  $(p, q)$  at all points on the straight line segment between them in the  $(p^{-1}, q^{-1})$  plane. In other words, the set of points  $(p^{-1}, q^{-1})$  such that  $T$  is of type  $(p, q)$  is convex.

(2) The main diagonal in the figure has  $p = q$ , and any corresponding operator is to carry a space  $L^p(X, \mu)$  to itself. The other diagonal has  $p^{-1} + q^{-1} = 1$ , and any corresponding operator carries a space  $L^p(X, \mu)$  to its dual (except that  $L^\infty(X, \mu)$  is to be carried to its predual  $L^1(X, \mu)$ ). That is why the figure refers to this diagonal as the “dual diagonal.”

(3) In practice when  $\mu(X) = \infty$ , as is the case when  $X = \mathbb{R}^N$  and  $\mu$  is Lebesgue measure, we usually have  $p \leq q$ , and the points of interest are in the lower triangle. Of particular importance are the main diagonal, the parallels below it, and the part of the dual diagonal that lies in the lower triangle. On  $\mathbb{R}^N$ , examples of operators yielding boundedness in these important cases are convolution with a fixed  $L^1$  function (in the case of the main diagonal), convolution with a fixed function in a class  $L^r$  (in the case of a line parallel to the main diagonal), and the Fourier transform (in the case of the lower half of the dual diagonal). For precise statements of boundedness in these cases, see the convolution result in Proposition 9.10, Young’s Inequality in Theorem 9.19D, and the Hausdorff–Young Theorem in Theorem 9.19C.

In order to give the proof of the theorem, we require the lemma below, which is known as the **Three Lines Theorem**. The lemma refers to “analytic functions” and is really a result in elementary complex analysis, but in keeping with the design of this book, we shall prove it using real analysis, avoiding using any results in



Appendix B that depend on the Cauchy Integral Theorem and its consequences. The corresponding argument via complex analysis will appear in a footnote.

A complex-valued function  $f$  on a connected open subset  $U$  of  $\mathbb{C}$  is said to be **analytic** if  $f$  has a complex derivative at each point of  $U$ . The definition of “complex derivative” to use is the usual one for functions of real variables, except that the domain and range are now allowed to be complex. Alternatively one can write  $f$  in terms of its real and imaginary parts as  $f(z) = u(x, y) + iv(x, y)$  and require the function  $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$  to be differentiable in the sense of Chapter III in such a way that the Jacobian matrix of the derivative represents a complex number, i.e., has the form  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ . This condition on the Jacobian matrix is the same as the condition that  $u$  and  $v$  satisfy the **Cauchy–Riemann equations**:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Such a function  $f$  is actually  $C^\infty$  on  $U$ . This fact is not so easy to prove in general, but it will be evident for the particular functions that enter the proof of the Three Lines Theorem. Thus we omit the argument.

By the usual proofs given in calculus, analytic functions are closed under addition, subtraction, multiplication, and division, provided division by 0 is not involved. Then we see from the Cauchy–Riemann equations that

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x} \left( \frac{\partial v}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial v}{\partial x} \right) = \frac{\partial}{\partial y} \left( -\frac{\partial u}{\partial y} \right) = -\frac{\partial^2 u}{\partial y^2},$$

i.e., that  $u$  is **harmonic** in the sense of satisfying **Laplace’s equation**

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Similarly  $v$  is harmonic, and hence the complex-valued function  $f = u + iv$  is harmonic.

Although our lemma will be stated in terms of complex analysis, it will make use of a **maximum principle** in real analysis, namely that for a complex-valued harmonic function  $f(z)$  on a connected open subset  $U$  of  $\mathbb{C}$ ,  $|f(z)|$  cannot attain a maximum on  $U$  unless  $f$  is constant. We give a conversational proof. Harmonic functions were treated in Problem 14 at the end of Chapter III and Problems 10–12 at the end of Chapter IV, and we make use of the results of those problems. The idea is to show that the set where  $|f(z)|$  attains its maximum is open and closed in  $U$ . Since  $U$  is connected, the set either must then be empty or be all of  $U$ . The set where  $|f(z)|$  attains its maximum is certainly closed, since  $f$  is continuous, and

we have to show that the set is a neighborhood of each of its points. If  $|f|$  attains a maximum, we may assume without loss of generality that it does so at  $z = 0$ . The mentioned problems in Chapters III and IV show that  $f$  has an expansion  $f(re^{i\theta}) = \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta}$  valid for  $r \geq 0$  sufficiently close to 0. On every sufficiently small circle  $r = r_0$  about 0, term-by-term integration shows that the function  $f$  has  $\frac{1}{2\pi} \int_{-\pi}^{\pi} f(r_0 e^{i\theta}) d\theta = c_0$ , i.e., its average value equals its value at 0, where  $|f|$  attains its maximum. Hence  $f$  is constant on all such circles, and 0 is an interior point of the set where  $|f|$  attains the value  $|c_0|$ . This completes the argument and establishes the maximum principle.<sup>2</sup>

Armed with the maximum principle for analytic functions on connected open subsets of  $\mathbb{C}$ , we can prove the Three Lines Theorem.

**Lemma 9.19B** (Three Lines Theorem). Let  $\Phi$  be analytic on an open subset of  $\mathbb{C}$  containing the closed vertical strip  $0 \leq \operatorname{Re}(z) \leq 1$ , and suppose that  $|\Phi(z)|$  is bounded in the strip. Define  $M_t = \sup_{-\infty < y < \infty} |\Phi(t + iy)|$  for  $0 \leq t \leq 1$ . Then

$$M_t \leq M_0^{1-t} M_1^t$$

for all  $y$ .

REMARK. In other words,  $\log M_t \leq (1-t) \log M_0 + t \log M_1$ . Scaling the domain, we see that this kind of inequality must persist for any three  $t$  values between 0 and 1. Briefly the conclusion is that  $\log M_t$  is convex as a function of  $t$ .

PROOF. First we handle the special case that  $M_0 = M_1 = 1$ . We are to prove that  $M_t \leq 1$ . It is enough to prove that  $|\Phi(t)| \leq 1$  since the case of  $\Phi(t + iy_0)$  can be handled by considering a vertical translate of  $\Phi$ . Since  $\Phi(z)$  is by assumption bounded in the strip, we can write  $|\Phi(z)| \leq A$  for some constant  $A$ . For each  $n \geq 1$ , define  $\Phi_n(z) = e^{z^2/n} \Phi(z)$ . This is analytic on the same open subset as for  $\Phi$ , and it has

$$|\Phi_n(z)| \leq |\Phi(z)| e^{z^2/n} \leq A e^{(x^2 - y^2)/n} \leq A e^{(1 - y^2)/n}$$

since  $z = x + iy$  has  $0 \leq x \leq 1$ . Fix  $n$ . By taking  $y$  large enough, we can arrange that  $A e^{(1 - y^2)/n} \leq 1$ . Thus on a suitably large vertical rectangle with  $0 \leq x \leq 1$  and  $y$  symmetric about 0,  $|\Phi_n(z)|$  is bounded by 1 on the top and bottom, and it has

$$|\Phi_n(z)| \leq |\Phi(z)| e^{(x^2 - y^2)/n} \leq e^{x^2/n} \leq e^{1/n}$$

<sup>2</sup>If elementary complex analysis is allowed in the proof of the Three Lines Theorem, then the previous page can be replaced by the simple remark that the Maximum Modulus Theorem (Corollary B.24) applies to analytic functions  $f$  on connected open subsets of  $\mathbb{C}$ ; for such a function,  $|f(z)|$  cannot attain its maximum value unless  $f$  is a constant function.

on the left and right sides. By the maximum principle,  $|\Phi_n(z)| \leq e^{1/n}$  in the interior of the rectangle. In particular,  $|\Phi_n(t)| \leq e^{1/n}$  for real  $t$  satisfying  $0 \leq t \leq 1$ . That is,  $e^{t^2/n}|\Phi(t)| \leq e^{1/n}$ . Letting  $n$  tend to infinity, we conclude that  $|\Phi(t)| \leq 1$  for  $0 \leq t \leq 1$ . This completes the discussion of the special case.

The case that  $M_0 = 0$  or  $M_1 = 0$  requires separate comment. In this case,  $\Phi(z)$  vanishes on an entire vertical line. An analytic function on a connected open set cannot vanish on a vertical line without vanishing identically,<sup>3</sup> and hence the lemma is valid if  $M_0 = 0$  or  $M_1 = 0$ .

For the general case with  $M_0$  and  $M_1$  nonzero, we modify our given analytic function  $\Phi$  by defining  $\Phi_0(z) = \Phi(z)M_0^{z-1}M_1^{-z}$ . The function  $\Phi_0$  is bounded on the vertical strip and has

$$|\Phi_0(iy)| = |\Phi(iy)||M_0^{iy-1}||M_1^{-iy}| \leq M_0M_0^{-1}M_1^0 = 1$$

and

$$|\Phi_0(1+iy)| = |\Phi(1+iy)||M_0^{iy}||M_1^{-1-iy}| \leq M_1M_0^0M_1^{-1} = 1.$$

The special case applies and shows that  $|\Phi_0(t)| \leq 1$ . Therefore  $|\Phi(t)| \leq M_0^{1-t}M_1^t$ .  $\square$

PROOF OF THEOREM 9.19A. Let  $M_0 = M_{p_0, q_0}$  and  $M_1 = M_{p_1, q_1}$ . We may assume that  $0 < t < 1$ . Let  $f \neq 0$  be simple on  $X$  with  $\|f\|_p \leq 1$ , and write  $f = \sum_m a_m I_{E_m}$  uniquely with the sets  $E_m$  disjoint and the complex numbers  $a_m$  distinct and nonzero. By Proposition 9.8 and the denseness of simple functions,

$$\|Tf\|_q = \sup_{\substack{g \text{ simple,} \\ \|g\|_{q'} \leq 1}} \left| \int_Y (Tf)g \, d\nu \right|, \quad (*)$$

where  $q'$  is the dual index to  $q$  defined by  $q^{-1} + q'^{-1} = 1$ . Fix  $g$  simple with  $\|g\|_{q'} \leq 1$ , and write  $g = \sum_n b_n I_{F_n}$  uniquely with the sets  $F_n$  disjoint and the complex numbers  $b_n$  distinct and nonzero. We shall prove that

$$\left| \int_Y (Tf)g \, d\nu \right| \leq M_0^{1-t}M_1^t. \quad (**)$$

For complex  $z$ , define

$$\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1} \quad \text{and} \quad \beta(z) = \frac{1-z}{q_0} + \frac{z}{q_1}.$$

<sup>3</sup>It is possible to come to this conclusion as a consequence of properties of harmonic functions in the mentioned problems in Chapters III and IV, but we shall not bother to do so.

Put

$$\alpha = \alpha(t) = p^{-1} \quad \text{and} \quad \beta = \beta(t) = q^{-1}.$$

Observe that  $1 - \beta = 1 - q^{-1} = q'^{-1}$ . The cases that  $\alpha = 0$ , i.e., that  $p = \infty$ , and that  $1 - \beta = 0$ , i.e., that  $q' = \infty$ , require special treatment and are deferred to the end of the proof. In the remaining cases, We define

$$f_z = |f|^{\frac{\alpha(z)}{\alpha}} \frac{f}{|f|} \quad \text{and} \quad g_z = |g|^{\frac{1-\beta(z)}{1-\beta}} \frac{g}{|g|}.$$

Here  $f_z(x)$  is understood to be 0 whenever  $f(x) = 0$ , and  $g_z(y)$  is understood to be 0 whenever  $g(y) = 0$ . Observe that  $f_t = f$  and  $g_t = g$ .

For each complex  $z$ , the functions  $f_z$  and  $g_z$  defined above are simple functions on  $X$  and  $Y$  given by

$$f_z = \sum_m |a_m|^{\frac{\alpha(z)}{\alpha}} \frac{a_m}{|a_m|} I_{E_m} \quad \text{and} \quad g_z = \sum_n |b_n|^{\frac{1-\beta(z)}{1-\beta}} \frac{b_n}{|b_n|} I_{F_n}.$$

The operations on the coefficients make sense because  $a_m \neq 0$  and  $b_n \neq 0$  for all  $m$  and  $n$ . Let

$$\Phi(z) = \int_Y (Tf_z) g_z \, d\nu = \sum_{m,n} |a_m|^{\frac{\alpha(z)}{\alpha}} |b_n|^{\frac{1-\beta(z)}{1-\beta}} \frac{a_m}{|a_m|} \frac{b_n}{|b_n|} \int_Y (T I_{E_m}) I_{F_n} \, d\nu.$$

The expansion on the right shows that  $\Phi(z)$  is analytic for all complex  $z$ . There are only finitely many terms, and each term is unaffected in absolute value by changing the imaginary part of  $z$ . Therefore  $\Phi(z)$  is bounded for  $0 \leq \operatorname{Re} z \leq 1$ . Moreover,

$$\Phi(t) = \int_Y (Tf) g \, d\nu$$

because  $f_t = f$  and  $g_t = g$ .

Let us see that

$$|\Phi(iy)| \leq M_0 = M_{p_0, q_0} \quad \text{and} \quad |\Phi(1 + iy)| \leq M_1 = M_{p_1, q_1}. \quad (\dagger)$$

For the first inequality we have

$$|\Phi(iy)| = \left| \int_Y (Tf_{iy}) g_{iy} \, d\nu \right| \leq \|Tf_{iy}\|_{q_0} \|g_{iy}\|_{q'_0} \leq M_0 \|f_{iy}\|_{p_0} \|g_{iy}\|_{q'_0}. \quad (\dagger\dagger)$$

Now  $\frac{\alpha(iy)p_0}{\alpha} = p \left( (1 - iy) + \frac{iy p_0}{p_1} \right)$ , and hence

$$|f_{iy}|^{p_0} = \left| |f|^{\frac{\alpha(iy)p_0}{\alpha}} \right| = |f|^p.$$

So  $\|f_{iy}\|_{p_0} = (\int_X |f|^p d\mu)^{1/p_0} = \|f\|_p^{p/p_0}$ , and this is  $\leq 1$  since  $\|f\|_p \leq 1$ .

Also  $\frac{1-\beta(iy_0)}{1-\beta} = q'(1 - \beta(iy_0)) = q'(1 - \frac{1-iy_0}{q_0} - \frac{iy_0}{q_1}) = q'(\frac{1}{q_0} + \frac{iy_0}{q_0} - \frac{iy_0}{q_1})$ ,  
and hence

$$|g_{iy}|^{q'_0} = |g|^{\frac{(1-\beta(iy))q'_0}{1-\beta}} = |g|^{q'}.$$

So  $\|g_{iy}\|_{q'_0} = (\int_Y |g|^{q'} d\nu)^{1/q'_0} = \|g\|_{q'}^{q'/q'_0}$ , and this is  $\leq 1$  since  $\|g\|_{q'} \leq 1$ .

The first inequality in  $(\dagger)$  follows from  $(\dagger\dagger)$  and the estimates  $\|f_{iy}\|_{p_0} \leq 1$  and  $\|g_{iy}\|_{q'_0} \leq 1$  that we have just established. The second inequality in  $(\dagger)$  is proved similarly.

Applying Lemma 9.19B and using  $(\dagger)$  and the boundedness of  $\Phi(z)$  for  $0 \leq \operatorname{Re} z \leq 1$ , we conclude that  $|\Phi(t)| \leq M_0^{1-t} M_1^t$ . This is inequality  $(**)$ , which we are trying to prove. Taking the supremum over all  $g$  simple with  $\|g\|_{q'} \leq 1$  and applying  $(*)$ , we find that  $\|Tf\|_q \leq M_0^{1-t} M_1^t$ . Here  $f$  is nonzero simple with  $\|f\|_p \leq 1$ .

We conclude that  $T$  satisfies the inequality

$$\|Tf\|_q \leq M_0^{1-t} M_1^t \|f\|_p \quad (\ddagger)$$

for all simple  $f$  in  $L^p(X, \mu)$ . This is the boundedness estimate required by the theorem, except that we have proved it only for simple functions  $f$  in  $L^p(X, \mu)$ , whereas the theorem is asserting this estimate for all  $f$  in  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ .

The fact that  $T$  was given as well defined on  $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$  implies that  $T$  is well defined on the set of sums of members of  $L^{p_0}(X, \mu) + L^{p_1}(X, \mu)$ , the definition being that  $Tf = Tf_{p_0} + Tf_{p_1}$  whenever  $f \in L^p(X, \mu)$  is decomposed as a sum  $f = f_{p_0} + f_{p_1}$  in  $L^{p_0}(X, \mu) + L^{p_1}(X, \mu)$ . The space  $L^p(X, \mu)$  is contained in this space of sums,<sup>4</sup> and thus we were already given a definition of  $T$  on  $L^p(X, \mu)$ .

Inequality  $(\ddagger)$  and Proposition 2.47 give us a second definition of  $T$  on  $L^p(X, \mu)$ , namely as the continuous extension of  $T$  to  $L^p(X, \mu)$  from the subspace of simple functions in  $L^p(X, \mu)$ , and we have to check that the two definitions of  $T$  on  $L^p(X, \mu)$  coincide. For this purpose let  $f$  be given in  $L^p(X, \mu)$ . The function  $f$  is a linear combination of four functions  $\geq 0$ . For each one, Proposition 5.11 gives us a sequence of simple functions increasing monotonically to the function  $\geq 0$ , and monotone convergence shows that the convergence takes place simultaneously in  $L^{p_0}$ ,  $L^p$ , and  $L^{p_1}$ . The linear combination of the sequences of simple functions is therefore a sequence of simple functions  $\{s_n\}$  converging to  $f$  simultaneously in  $L^{p_0}$ ,  $L^p$ , and  $L^{p_1}$ . On  $L^{p_0}(X, \mu)$  and  $L^{p_1}(X, \mu)$ ,  $T$  was given as continuous, and thus  $Tf = \lim T s_n$  in  $L^{q_0}(Y, \nu)$  and  $L^{q_1}(Y, \nu)$ . Using

<sup>4</sup>Under the identification of its members with genuine functions rather than functions modulo sets of measure 0.

Theorem 9.6, we can pass to a subsequence, which we still call  $\{s_n\}$ , that is convergent almost everywhere to  $Tf$  in  $L^{q_0}(Y, \nu)$  and  $L^{q_1}(Y, \nu)$ . This version of  $Tf$  is the result of the operation of  $T$  on  $L^{p_0}(X, \mu) + L^{p_1}(X, \mu)$ .

The second version of  $Tf$  is the one obtained by using  $(\ddagger)$  and the resulting continuity on  $L^p(X, \mu)$ . For it we have  $Tf = \lim T s_n$  in  $L^p(X, \mu)$ . Again we can pass to a subsequence by Theorem 9.6 and obtain  $Tf$  as a limit almost everywhere. The two almost-everywhere limits must be equal, and thus the two definitions of  $Tf$  coincide.

To complete the proof, we must handle the deferred cases that  $\alpha = 0$  or that  $1 - \beta = 0$ , or both. If  $\alpha = 0$ , then  $p = \infty$ . The fact that  $0 < t < 1$  forces  $p_0 = p_1 = \infty$ . For this situation we define  $f_z = f$  for all  $z$  instead of using the earlier definition  $f_z = \sum_m |a_m|^{\frac{\alpha(z)}{\alpha}} \frac{a_m}{|a_m|} I_{E_m}$ . If  $1 - \beta = 0$ , then  $q' = \infty$  and  $q = 1$ . The fact that  $0 < t < 1$  forces  $q_0 = q_1 = 1$ . For this situation we define  $g_z = g$  for all  $z$  instead of using the earlier definition  $g_z = \sum_n |b_n|^{\frac{1-\beta(z)}{1-\beta}} \frac{b_n}{|b_n|} I_{F_n}$ . Again we let  $\Phi(z) = \int_Y (Tf_z) g_z d\nu$ , and the earlier argument that  $(\ddagger)$  holds for  $f$  simple in  $L^p(X, \mu)$  goes through. Arguing as earlier, we find that  $(\ddagger)$  holds for general  $f$  in  $L^p(X, \mu)$ , and the proof is complete.  $\square$

Now that we have the the Riesz–Thorin Convexity Theorem in hand, we shall obtain two consequences—the Hausdorff–Young Theorem and Young’s inequality. For the Hausdorff–Young Theorem our theory is to be applied with  $T$  equal to the Euclidean Fourier transform.

**Corollary 9.19C** (Hausdorff–Young Theorem). If  $1 \leq p \leq 2$  and if  $p'$  is the dual index, then the Euclidean Fourier transform  $\mathcal{F}$ , whose definition on  $L^1(\mathbb{R}^N, dx) \cap L^2(\mathbb{R}^N, dx)$  makes it well defined on  $L^1(\mathbb{R}^N, dx) + L^2(\mathbb{R}^N, dx)$  and hence on the subspace  $L^p(\mathbb{R}^N, dx)$ , satisfies

$$\|\mathcal{F}(f)\|_{p'} \leq \|f\|_p$$

for all  $f$  in  $L^p(\mathbb{R}^N, dx)$ .

PROOF. The linear operator  $T$  in the Hausdorff–Young Theorem is the Fourier transform  $\mathcal{F}$ , and the instances of the theorem that we know from earlier are when  $(p, p')$  equals  $(1, \infty)$  or  $(2, 2)$ . The numerology that allows the Riesz–Thorin Convexity Theorem (Theorem 9.19A) to apply is that

$$\frac{1}{p} = \frac{1-t}{1} + \frac{t}{2} \quad \text{and} \quad \frac{1}{p'} = \frac{1-t}{\infty} + \frac{t}{2}$$

for the same  $t$ . Corollary 9.19C follows immediately.  $\square$

For Young's inequality our theory is to be applied with  $T$  equal to the convolution operator  $g \mapsto f * g$  with  $f$  fixed in  $L^p(\mathbb{R}^N, dx)$ .

**Corollary 9.19D** (Young's inequality). Let  $p, q,$  and  $r$  be three indices  $\geq 1$  and  $\leq \infty$  such that  $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$ . Then convolution  $f * g$  is well defined for  $f$  in  $L^p(\mathbb{R}^N, dx)$  and  $g$  in  $L^q(\mathbb{R}^N, dx)$ , and it satisfies

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

PROOF. The linear operator  $T$  in Young's inequality is the convolution operator  $g \mapsto f * g$  with  $f$  fixed in  $L^p(\mathbb{R}^N, dx)$ . The instances of the inequality that we know from earlier are when  $(q, r)$  equals  $(1, p)$  or  $(p', \infty)$ . The numerology that allows the Riesz–Thorin Convexity Theorem (Theorem 9.19A) to apply is that

$$\frac{1}{q} = \frac{1-t}{1} + \frac{t}{p'} \quad \text{and} \quad \frac{1}{r} = \frac{1-t}{p} + \frac{t}{\infty}$$

for the same  $t$ . Corollary 9.19D follows immediately.  $\square$

REMARK. It is instructive to consider how the pairs  $(q, r)$  for fixed  $p$  appear in the  $(1/q, 1/r)$  plane. Subtraction of the two displayed equations in the above proof gives  $\frac{1}{q} - \frac{1}{r} = 1 - \frac{1}{p}$ , and thus the set is the intersection of the lower triangle with a line parallel to the main diagonal.

## 7. Marcinkiewicz Interpolation Theorem

A second class of results concerning linear operators between  $L^p$  spaces is built around the the Marcinkiewicz Interpolation Theorem. This result actually applies to a somewhat wider class of operators than linear operators, and the extra generality is important. To fix the notation, let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, and let  $T$  be a function from a vector subspace of measurable functions on  $X$ , modulo sets<sup>5</sup> of  $\mu$  measure 0, into measurable functions on  $Y$ , modulo sets of  $\nu$  measure 0. We say that  $T$  is a **sublinear** operator if  $|T(f + g)| \leq |T(f)| + |T(g)|$  for all  $f$  and  $g$  in the domain of  $T$ .

The two examples of  $T$  to keep in mind are the sublinear operator  $f \mapsto f^*$  in  $\mathbb{R}^N$  of passing to the Hardy–Littlewood maximal function, as in Section VI.6, and the linear operator  $f \mapsto H_1 f$  in  $\mathbb{R}^1$  of forming a certain approximation  $H_1$  to the Hilbert transform, as in Section VIII.7. More specifically the Hardy–Littlewood maximal function of a locally integrable function  $f$  on  $\mathbb{R}^N$  is defined as

$$f^*(x) = \sup_{0 < r < \infty} m(B_r)^{-1} \int_{B_r} |f(x - y)| dy, \quad \text{where } B_r = B(r; 0) \text{ in } \mathbb{R}^N,$$

<sup>5</sup>This condition means that the domain of  $T$  is to be regarded as a vector subspace of measurable functions, except that two functions are identified if they differ only on a set of measure 0.

and the sublinear operator  $T$  is  $Tf = f^*$ . The approximation  $H_1$  to the Hilbert transform is defined for  $f$  in  $L^1 + L^2$  by

$$H_1 f(x) = h_1 * f(x) = \frac{1}{\pi} \int_{|t| \geq 1} \frac{f(x-t)}{t} dt$$

as the convolution with a fixed  $L^2$  function.

Let  $1 \leq p, q \leq \infty$ . We generalize the notion of boundedness of a linear operator between  $L^p(X, \mu)$  and  $L^q(Y, \nu)$  so that we can work with sublinear operators as well as linear ones. A sublinear operator  $T$  is said to be of **type**  $(p, q)$  or **strong type**  $(p, q)$  if  $\|Tf\|_q \leq M\|f\|_p$  with  $M$  finite and independent of  $f$ . The least  $M$  for which this inequality holds is called the **norm** or **operator norm** of  $T$ . If  $q < \infty$ , then Chebyshev's inequality from Section VI.10 gives

$$\nu(\{y \in Y \mid |Tf(y)| > \xi\}) \leq \frac{\int_Y |Tf|^q d\nu}{\xi^q},$$

and for any  $M$  such that  $\|Tf\|_q \leq M\|f\|_p$  for all  $f$ , it follows that

$$\nu(\{y \in Y \mid |Tf(y)| > \xi\}) \leq \left( \frac{M\|f\|_p}{\xi} \right)^q.$$

If  $q < \infty$ , a sublinear operator  $T$  is said to be of **weak type**  $(p, q)$  if it satisfies

$$\nu(\{y \in Y \mid |Tf(y)| > \xi\}) \leq \left( \frac{M\|f\|_p}{\xi} \right)^q$$

for some  $M$ . In this case the least such  $M$  is called the **weak-type norm** of  $T$ . We already encountered the definition of weak type  $(1, 1)$  in Section VI.6. If  $q = \infty$ , the convention is that weak type  $(p, \infty)$  is the same as strong type  $(p, \infty)$ .

Consider our two examples. The operation  $T(f) = f^*$  of passing to the Hardy–Littlewood maximal function in  $\mathbb{R}^N$  is of weak type  $(1, 1)$  by the Hardy–Littlewood Maximal Theorem (Theorem 6.38), and the evident inequality

$$\left\| \sup_{0 < r < \infty} m(B_r)^{-1} \int_{B_r} |f(x-y)| dy \right\|_{\infty} \leq \|f\|_{\infty}$$

shows that  $f \mapsto f^*$  is of type  $(\infty, \infty)$  as well. The linear operator  $T(f) = H_1 f$  of passing to the approximation  $H_1$  to the Hilbert transform in  $\mathbb{R}^1$  is of weak type  $(1, 1)$  and type  $(2, 2)$  by Theorem 8.25.



We include below a statement of the Marcinkiewicz Interpolation Theorem in general and the proof in a special case of exceptional interest. The Marcinkiewicz theorem imposes some restrictions on the pairs  $(p_1, q_1)$  and  $(p_2, q_2)$  that are not needed in the Riesz–Thorin Convexity Theorem, but situations that do not satisfy these restrictions are of comparatively little interest in applications. In any event, in the situations where the Marcinkiewicz theorem applies, it is only the specific information about the operator bound in the Riesz–Thorin theorem that does not come out of the real-analysis proof of the Marcinkiewicz theorem.

**Theorem 9.20** (Marcinkiewicz Interpolation Theorem). Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, and let  $(p_1, q_1)$  and  $(p_2, q_2)$  be two pairs of indices between 1 and  $\infty$ . Suppose that  $1 \leq p_1 \leq q_1 \leq \infty$ ,  $1 \leq p_2 \leq q_2 \leq \infty$ , and  $p_1 \neq p_2$ . Let  $T$  be a sublinear operator from  $L^{p_1}(X, \mu) + L^{p_2}(X, \mu)$  to the space of measurable functions on  $Y$  modulo sets of  $\nu$  measure 0, and suppose that  $T$  is of weak types  $(p_1, q_1)$  and  $(p_2, q_2)$  with respective weak-type norms  $M_1$  and  $M_2$ . Fix  $t$  with  $0 < t < 1$ , and define  $(p, q)$  by

$$\frac{1}{p} = \frac{1-t}{p_1} + \frac{t}{p_2} \quad \text{and} \quad \frac{1}{q} = \frac{1-t}{q_1} + \frac{t}{q_2}.$$

Then  $T$  is of strong type  $(p, q)$  with

$$\|Tf\|_q \leq C \|f\|_p \quad \text{for all } f \in L^p(X, \mu),$$

with the constant  $C$  depending only on  $t, M_1, M_2, p_1, q_1, p_2, q_2$  and with  $C$  bounded as a function of  $t$  as long as  $t$  is bounded away from 0 and 1.

Before discussing the proof, let us apply the theorem to the two examples mentioned at the beginning of the section, the Hardy–Littlewood maximal function and the approximation  $H_1$  to the Hilbert transform. Then let us draw some consequences of these applications. As was said before the statement of Theorem 9.20, the sublinear operator  $f \mapsto f^*$  is of weak type  $(1, 1)$  and strong type  $(2, 2)$ . The theorem immediately gives the following corollary.

**Corollary 9.21.** If  $1 < p \leq \infty$ , then there exists a constant  $A_p$  such that the Hardy–Littlewood maximal function satisfies

$$\|f^*\|_p \leq A_p \|f\|_p$$

for all  $f$  in  $L^p(\mathbb{R}^N)$ .

The case of this result in one dimension implies something in  $N$  dimensions that we have not obtained earlier. If  $f$  is locally integrable on  $\mathbb{R}^N$ , one says that **strong differentiation** holds for  $f$  at  $x$  if

$$\lim_{\substack{\text{diam}(R) \rightarrow 0, \\ R = \text{geometric rectangle} \\ \text{centered at } x}} \frac{1}{m(R)} \int_R f(y) dy = f(x).$$

A consequence of Corollary 9.21 is that strong differentiation holds almost everywhere for each  $f$  in  $L^p(\mathbb{R}^N, dx)$  for  $p > 1$ . The proof is outlined in Problems 13–15 at the end of the chapter. By contrast, it is known that there are functions in  $L^1(\mathbb{R}^N, dx)$  for which strong differentiation fails everywhere.

In the second example the operator  $H_1$  that approximates the Hilbert transform is of weak type  $(1, 1)$  and strong type  $(2, 2)$ , and Theorem 9.20 allows us to conclude that it is of strong type  $(p, p)$  for  $1 < p \leq 2$ . But we can do better. The operator  $H_1$  is convolution by the function  $h_1$  with  $h_1(x) = 1/(\pi x)$  for  $|x| \geq 1$  and  $h_1(x) = 0$  for  $|x| < 1$ . The function  $h_1$  is in  $L^p$  for all  $p > 1$ , and Proposition 9.10f shows that  $h_1 * f$  is well defined as a bounded continuous function whenever  $f$  is in some  $L^q$  with  $1 \leq q < \infty$ . Thus  $H_1$  is defined on all  $L^p$  classes for  $1 < p < \infty$ , and a general result that we prove below as Lemma 9.22 shows that an inequality  $\|H_1 f\|_p \leq A_p \|f\|_p$  for all  $f$  in  $L^p$  implies  $\|H_1 g\|_{p'} \leq A_p \|g\|_{p'}$  for all  $g$  in  $L^{p'}$ , provided  $p'$  is the dual index to  $p$  and  $1 < p < \infty$ . Thus the boundedness result for  $H_1$  on  $L^p$  extends to  $1 < p < \infty$ .

Next, we define the dilate  $h_\varepsilon$  in the usual way by  $h_\varepsilon(x) = \varepsilon^{-1} h_1(x)$ , and we put  $H_\varepsilon f = h_\varepsilon * f$ . In Theorem 9.23 below we shall see for every  $\varepsilon > 0$  that  $\|H_\varepsilon f\|_p \leq A_p \|f\|_p$  with the same constant  $A_p$ . In addition, we shall see that we can let  $\varepsilon$  decrease to 0 and obtain the Hilbert transform  $H$  as a well-defined linear operator on all  $L^p$  classes for  $1 < p < \infty$ ; the estimate is  $\|Hf\|_p \leq A_p \|f\|_p$ , again for the same  $A_p$ . Problems 20–22 at the end of the chapter indicate how to use this boundedness to prove that the Fourier series of any  $L^p$  function on  $[-\pi, \pi]$  converges to the function in  $L^p$  if  $1 < p < \infty$ ; this is the convergence result of M. Riesz that was mentioned in a footnote near the beginning of Section 6.

**Lemma 9.22.** Fix  $p$  with  $1 < p < \infty$ , let  $p'$  be the dual index, and suppose that  $h$  is in  $L^p(\mathbb{R}^N) \cap L^{p'}(\mathbb{R}^N)$ . If  $\|h * f\|_p \leq A_p \|f\|_p$  for all  $f$  in  $L^p(\mathbb{R}^N)$ , then  $\|h * g\|_{p'} \leq A_p \|g\|_{p'}$  for all  $g$  in  $L^{p'}$ .

REMARKS. Since  $h$  is in  $L^{p'}$ ,  $h * f$  is in  $L^\infty$  when  $f$  is in  $L^p$ . Thus  $h * f$  is well defined, and it is meaningful to say that  $h * f$  is actually in  $L^p$ . When  $h * f$  is in  $L^p$ , the integral  $\int (h * f)g dx$  is well defined for  $g$  in  $L^{p'}$ . A little care is required in working with this integral in the proof because  $\int (|h| * f)g dx$  need not be well defined and Fubini's Theorem may not directly applicable.

PROOF. For any function  $F$  on  $\mathbb{R}^N$ , define  $F^\#(x) = F(-x)$  and observe that  $\|F^\#\|_r = \|F\|_r$  for  $1 \leq r \leq \infty$ . If  $g$  is an integrable simple function, then  $(h^\# * g)(x) = \int h(y-x)g(y) dy = \int h(-y-x)g^\#(y) dy = (h * g^\#)(-x)$ . Thus this  $g$  and an integrable simple function  $f$  together satisfy

$$\begin{aligned} \int (h * f^\#)(x)g(x) dx &= \iint h(x-y)f(-y)g(x) dy dx \\ &= \iint h(x+y)f(y)g(x) dy dx \end{aligned}$$

and

$$\begin{aligned} \int (h * g^\#)(y)f(y) dy &= \int (h^\# * g)(-y)f(y) dy \\ &= \iint h^\#(-y-x)g(x)f(y) dx dy \\ &= \iint h(x+y)g(x)f(y) dx dy. \end{aligned}$$

Because  $f$  and  $g$  are in every  $L^r$  class, the right sides of these two displays are finite when absolute value signs are inserted in the integrands. Thus Fubini's Theorem applies and shows that the two right sides are equal. Combining this fact with Hölder's inequality and the hypothesis about  $h$ , we obtain

$$\begin{aligned} \left| \int (h * g^\#)(y)f(y) dy \right| &= \left| \int (h * f^\#)(x)g(x) dx \right| \\ &\leq \|h * f^\#\|_p \|g\|_{p'} \leq A_p \|f^\#\|_p \|g\|_{p'} = A_p \|f\|_p \|g\|_{p'} \end{aligned}$$

whenever  $f$  and  $g$  are integrable simple functions. If a general  $f_0$  in  $L^p$  is given, we can find a sequence  $f_n$  of integrable simple functions such that  $\|f_n - f_0\|_p \rightarrow 0$ , and we apply this inequality to each  $f_n$ . Then the left side of the inequality tends to  $\left| \int (h * g^\#)(y)f_0(y) dy \right|$ , and the right side tends to  $A_p \|f_0\|_p \|g\|_{p'}$ . Taking the supremum over all  $f_0$  with  $\|f_0\|_p \leq 1$  and applying Proposition 9.8, we find that  $\|h * g^\#\|_{p'} \leq A_p \|g\|_{p'} = A_p \|g^\#\|_{p'}$ . In other words,

$$\|h * g_n\|_{p'} \leq A_p \|g_n\|_{p'}$$

for every integrable simple function  $g_n$ . For a general  $g$  in  $L^{p'}$ , choose a sequence of integrable simple functions  $g_n$  with  $\|g_n - g\|_{p'} \rightarrow 0$ . Since  $h$  is in  $L^p$ , it follows from Proposition 9.10f that  $h * g_n$  converges to  $h * g$  uniformly. On the other hand, the inequality  $\|h * (g_m - g_n)\|_{p'} \leq A_p \|g_m - g_n\|_{p'}$  shows that  $\{h * g_n\}$  is Cauchy in  $L^{p'}$ . By Theorem 5.58,  $\{h * g_n\}$  converges to some function in  $L^{p'}$  and has an almost-everywhere convergent subsequence to this function. Since  $h * g_n$  converges uniformly to  $h * g$ , we conclude that  $h * g_n$  converges to  $h * g$  in  $L^{p'}$ . Therefore  $\|h * g\|_{p'} \leq A_p \|g\|_{p'}$ , and the proof is complete.  $\square$

Again let  $h_1$  be the function on  $\mathbb{R}^1$  equal to  $1/(\pi x)$  for  $|x| \geq 1$  and equal to 0 for  $|x| < 1$ . This is in  $L^r(\mathbb{R}^1)$  for every  $r > 1$ . Our operator giving an

approximation to the Hilbert transform is  $H_1 f = h_1 * f$ . Using our results from Chapter VIII along with the Marcinkiewicz Interpolation Theorem, we saw earlier in this section that  $H_1$  satisfies  $\|H_1 f\|_p \leq A_p \|f\|_p$  for  $1 < p \leq 2$  and all  $f$  in  $L^p(\mathbb{R}^1)$ . Lemma 9.22 shows that this inequality remains valid for  $1 < p < \infty$ . From this result we can extend the Hilbert transform to  $L^p(\mathbb{R}^1)$  for all  $p$  with  $1 < p < \infty$ , as follows.

**Theorem 9.23.** Let  $1 < p < \infty$ , let

$$h_\varepsilon(x) = \varepsilon^{-1} h_1(\varepsilon^{-1}x) = \begin{cases} 1/(\pi x) & \text{for } |x| \geq \varepsilon, \\ 0 & \text{for } |x| < \varepsilon, \end{cases}$$

and define  $H_\varepsilon f = h_\varepsilon * f$  for  $f$  in  $L^p$  and  $\varepsilon > 0$ . Then

- (a) there exists a constant  $A_p$  independent of  $\varepsilon$  such that  $\|H_\varepsilon f\|_p \leq A_p \|f\|_p$  for all  $f$  in  $L^p$ ,  
 (b) the limit

$$Hf(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{|t| \geq \varepsilon} \frac{f(x-t) dt}{t}$$

exists in  $L^p$  for every  $f$  in  $L^p$ ,

- (c) the operator  $H$  satisfies  $\|Hf\|_p \leq A_p \|f\|_p$  for every  $f$  in  $L^p$ .

PROOF. Convolution with  $h_\varepsilon$  is well defined on  $L^p$  because  $h_\varepsilon$  is in  $L^{p'}$ ,  $p'$  being the dual index for  $p$ . The three computations

$$\begin{aligned} H_\varepsilon f(x) &= (f * h_\varepsilon)(x) = \int f(x-y) \varepsilon^{-1} h_1(\varepsilon^{-1}y) dy = \int f(x-\varepsilon y) h_1(y) dy \\ &= \int \varepsilon^{-1} f_{\varepsilon^{-1}}(\varepsilon^{-1}x - y) h_1(y) dy = \varepsilon^{-1} (H_1 f_{\varepsilon^{-1}})(\varepsilon^{-1}x), \end{aligned}$$

$$\int |(H_\varepsilon f)(x)|^p dx = \varepsilon^{-p} \int |(H_1 f_{\varepsilon^{-1}})(\varepsilon^{-1}x)|^p dx = \varepsilon^{1-p} \int |(H_1 f_{\varepsilon^{-1}})(x)|^p dx,$$

$$\text{and} \quad \int |g_{\varepsilon^{-1}}(x)|^p dx = \varepsilon^p \int |g(\varepsilon x)|^p dx = \varepsilon^{-1+p} \int |g(x)|^p dx$$

allow us to write

$$\|H_\varepsilon f\|_p^p = \varepsilon^{1-p} \|H_1 f_{\varepsilon^{-1}}\|_p^p \leq A_p^p \varepsilon^{1-p} \|f_{\varepsilon^{-1}}\|_p^p = A_p^p \|f\|_p^p.$$

This proves (a), the constant  $A_p$  being any constant that works for  $H_1$ .

In Lemma 9.24 below we show by a direct computation that (b) holds for the dense subset of  $C^1$  functions  $f$  of compact support. Let us deduce (b) for general  $f$  in  $L^p$  from this fact and (a). In fact, if we are given  $f$ , we choose a sequence  $f_n$  in this dense set with  $f_n \rightarrow f$  in  $L^p$ . Then

$$\begin{aligned} \|H_\varepsilon f - H_{\varepsilon'} f\|_p &\leq \|H_\varepsilon(f - f_n)\|_p + \|H_\varepsilon f_n - H_{\varepsilon'} f_n\|_p + \|H_{\varepsilon'}(f_n - f)\|_p \\ &\leq A_p \|f_n - f\|_p + \|H_\varepsilon f_n - H_{\varepsilon'} f_n\|_p + A_p \|f_n - f\|_p. \end{aligned}$$

Choose  $n$  to make the first and third terms small on the right, and then choose  $\varepsilon$  and  $\varepsilon'$  sufficiently close to 0 so that the second term on the right is small. The result is that  $H_{\varepsilon_n} f$  is Cauchy in  $L^p$  along any sequence  $\varepsilon_n$  tending to 0. This proves (b), apart from the direct computation for the dense subset.

In (b), we proved that  $H_\varepsilon f \rightarrow Hf$  in  $L^p$ . Then (a) gives  $\|Hf\|_p = \lim_{\varepsilon \downarrow 0} \|H_\varepsilon f\|_p \leq \limsup_{\varepsilon \downarrow 0} A_p \|f\|_p = A_p \|f\|_p$ . This proves (c) and completes the proof of Theorem 9.23 except for the following lemma.  $\square$

**Lemma 9.24.** If  $f$  is a  $C^1$  function of compact support on  $\mathbb{R}^1$ , then

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{|t| \geq \varepsilon} \frac{f(x-t) dt}{t}$$

exists uniformly and in  $L^p$  for every  $p > 1$ .

PROOF. Let  $\|\cdot\|$  denote the supremum norm or the  $L^p$  norm. By the Cauchy criterion it is enough to show that

$$\left\| \int_{\varepsilon_1 \leq |t| \leq \varepsilon_2} \frac{f(x-t) dt}{t} \right\|$$

tends to 0 for the above interpretations of  $\|\cdot\|$  as  $\varepsilon_1$  and  $\varepsilon_2$  tend to 0. Since  $|f'(u)| \leq M$ , use of the Mean Value Theorem on  $\operatorname{Re} f$  and  $\operatorname{Im} f$  shows that  $|f(x-t) - f(x)| \leq 2M|t|$ . Suppose that  $0 < \varepsilon_1 \leq \varepsilon_2 \leq 1$ . If  $E$  is a compact set containing the sum of any member of the support of  $f$  and any  $x$  with  $|x| \leq 1$ , then it follows that

$$\begin{aligned} \left\| \int_{\varepsilon_1 \leq |t| \leq \varepsilon_2} \frac{f(x-t) dt}{t} \right\| &= \left\| \int_{\varepsilon_1 \leq |t| \leq \varepsilon_2} \frac{[f(x-t) - f(x)] dt}{t} \right\| \\ &\leq \int_{\varepsilon_1 \leq |t| \leq \varepsilon_2} \frac{\|f(x-t) - f(x)\|_x dt}{|t|} \\ &\leq \int_{\varepsilon_1 \leq |t| \leq \varepsilon_2} \frac{2M|t| \|I_E\| dt}{|t|} \\ &= 4M \|I_E\| (\varepsilon_2 - \varepsilon_1). \end{aligned}$$

The right side tends to 0 as  $\varepsilon_1$  and  $\varepsilon_2$  tend to 0, and the proof of the lemma is complete.  $\square$

Having now completely proved Theorem 9.23, let us return to a discussion of the proof of the Marcinkiewicz theorem, Theorem 9.20. The proof is considerably simplified by assuming that  $q_1 = p_1$  and  $q_2 = p_2$ , which happens to be the special case of most interest to us, and we shall give a proof only under this additional hypothesis.<sup>6</sup> The idea in the special case will be to estimate integrals of powers

<sup>6</sup>A proof in the general case may be found in Appendix B of Stein's *Singular Integrals and Differentiability Properties of Functions*.

of functions by using Proposition 6.56b to reduce the estimates to facts about distribution functions.

The proof in general has the same flavor as the argument we give, but it involves also a subtler decomposition of  $f$  into two parts, a nonobvious application of Hölder's inequality, and a clever use of Proposition 9.8.

PROOF OF THEOREM 9.20 WHEN  $p_1 = q_1 < p_2 = q_2$ . We divide matters into two cases, the first when  $p_2 < \infty$  and the second when  $p_2 = \infty$ .

We begin with the case with  $p_2 < \infty$ . Let

$$\lambda(\xi) = \lambda_{Tf}(\xi) = \nu(\{y \mid |Tf(y)| > \xi\})$$

be the distribution function of  $Tf$  as in Section VI.10. Proposition 6.56b shows that

$$\|Tf\|_p^p = p \int_0^\infty \xi^{p-1} \lambda(\xi) d\xi = 2^p p \int_0^\infty \xi^{p-1} \lambda(2\xi) d\xi. \quad (*)$$

With  $\xi > 0$  fixed, we shall estimate  $\lambda(2\xi)$ . We decompose  $f$  as  $f = f_1 + f_2$  with

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > \xi \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq \xi \\ 0 & \text{otherwise} \end{cases}.$$

Just as in the proof of Proposition 9.4c,  $f_1$  is in  $L^{p_1}(X, \mu)$  and  $f_2$  is in  $L^{p_2}(X, \mu)$ . Because  $f = f_1 + f_2$ , sublinearity of  $T$  gives  $|Tf| \leq |Tf_1| + |Tf_2|$ . If  $\lambda_1$  and  $\lambda_2$  are the distribution functions of  $Tf_1$  and  $Tf_2$  and if  $\alpha > 0$  is given, then

$$\lambda(2\alpha) \leq \lambda_1(\alpha) + \lambda_2(\alpha)$$

because  $|Tf|$  can be  $> 2\alpha$  only if at least one of  $|Tf_1|$  and  $|Tf_2|$  is  $> \alpha$ . For every  $\alpha > 0$ , the assumption that  $T$  is of weak types  $(p_1, p_1)$  and  $(p_2, p_2)$  gives us

$$\lambda_1(\alpha) \leq \left( \frac{M_1 \|f_1\|_{p_1}}{\alpha} \right)^{p_1} \quad \text{and} \quad \lambda_2(\alpha) \leq \left( \frac{M_2 \|f_2\|_{p_2}}{\alpha} \right)^{p_2}.$$

For  $\alpha = \xi$ , we therefore obtain

$$\begin{aligned} \lambda(2\xi) &\leq \lambda_1(\xi) + \lambda_2(\xi) \leq M_1^{p_1} \xi^{-p_1} \int_X |f_1|^{p_1} d\mu + M_2^{p_2} \xi^{-p_2} \int_X |f_2|^{p_2} d\mu \\ &= M_1^{p_1} \xi^{-p_1} \int_{\{|f|>\xi\}} |f|^{p_1} d\mu + M_2^{p_2} \xi^{-p_2} \int_{\{|f|\leq\xi\}} |f|^{p_2} d\mu. \end{aligned} \quad (**)$$

With the estimate for  $\lambda(2\xi)$  in hand, we can now let  $\xi$  vary and estimate  $\|Tf\|_p^p$ . From (\*) and (\*\*) we obtain  $\|Tf\|_p^p \leq I_1 + I_2$ , where

$$I_1 = 2^p p M_1^{p_1} \int_0^\infty \xi^{p-p_1-1} \int_{\{|f(x)|>\xi\}} |f(x)|^{p_1} d\mu(x) d\xi$$

and

$$I_2 = 2^p p M_2^{p_2} \int_0^\infty \xi^{p-p_2-1} \int_{\{|f(x)|\leq\xi\}} |f(x)|^{p_2} d\mu(x) d\xi.$$

Fubini's Theorem gives

$$I_1 = 2^p p M_1^{p_1} \int_X |f|^{p_1} \left[ \int_0^{|f|} \xi^{p-p_1-1} d\xi \right] d\mu = \frac{2^p p M_1^{p_1}}{p-p_1} \int_X |f|^p d\mu.$$

Similarly

$$I_2 = \frac{2^p p M_2^{p_2}}{p_2-p} \int_X |f|^p d\mu,$$

and thus  $\|Tf\|_p^p \leq C^p \|f\|_p^p$  as required.

The remaining case to handle has  $p_2 = \infty$ . The general line of the argument is the same as above, but there are small differences. With  $\xi$  fixed, the definitions of  $f_1$  and  $f_2$  are adjusted to be

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > \xi/\|T\|_\infty, \\ 0 & \text{otherwise,} \end{cases}$$

and  $f_2 = f - f_1$ . Then  $\|f_2\|_\infty \leq \xi/\|T\|_\infty$ ,  $\|Tf_2\|_\infty \leq \xi$ , and  $\lambda_2(\xi) = 0$ . Hence

$$\lambda(2\xi) \leq \lambda_1(\xi) + \lambda_2(\xi) = \lambda_1(\xi) \leq M_1^{p_1} \xi^{-p_1} \int_{\{|f|>\xi/\|T\|_\infty\}} |f|^{p_1} d\mu,$$

and then the proof can proceed along the lines above.  $\square$

## 8. Problems

1. For a measure space of finite measure, prove that  $L^p \subseteq L^q$  whenever  $p \geq q \geq 1$ . More particularly prove, for the case that the total measure is 1, that  $\|f\|_q \leq \|f\|_p$  whenever  $p \geq q \geq 1$ .
2. Let  $p, q, r$  be real numbers in  $[1, +\infty]$  with  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ . Using the equality  $\frac{r'}{p} + \frac{r'}{q} = 1$  and Hölder's inequality, prove that  $\int_X |fgh| d\mu \leq \|f\|_p \|g\|_q \|h\|_r$ .
3. For a measure space of finite measure, let  $\{f_n\}$  be a sequence of measurable functions converging pointwise to  $f$ . Suppose that  $1 \leq q < p < \infty$ , and suppose that the sequence of numbers  $\{\|f\|_p\}$  is bounded. Using Egoroff's Theorem (Problem 17, Chapter V) or uniform integrability (Problem 21, Chapter V), prove that  $f_n \rightarrow f$  in  $L^q$ .
4. This problem produces an example of a measure space in which two distinct members of  $L^\infty$  act as the same linear functional on  $L^1$ . The measure space  $(X, \mathcal{A}, \mu)$  has  $X$  consisting of a single point  $p$ ,  $\mathcal{A} = \{\emptyset, X\}$ , and  $\mu(X) = +\infty$ .
  - (a) Show that  $\dim L^1(X) = 0$  and  $\dim L^\infty(X) = 1$ .
  - (b) Proposition 9.8 assumed  $\sigma$ -finiteness to ensure its conclusion when  $p = \infty$ . Show that the conclusion of Proposition 9.8 fails for  $p = \infty$  in this example.

5. If  $f$  is real-valued and integrable on the measure space  $(X, \mathcal{A}, \mu)$ , what are all the Hahn decompositions for the signed measure  $\nu(E) = \int_E f d\mu$ ?
6. Provide examples of each of the following. Each example can be produced on one of the following three algebras of subsets of a set  $X$ : the finite subsets of a  $X$  and their complements, all subsets of a countable set  $X$ , the Borel sets of  $X = [0, 1]$ .
  - (a) An additive set function  $\nu$  on an algebra of sets with  $|\nu(X)| < \infty$  but with  $\sup_E |\nu(E)| = \infty$ .
  - (b) A counterexample to the Hahn decomposition if the assumption “ $\sigma$ -algebra” is relaxed to “algebra” but the other assumptions are left in place.
  - (c) A finite measure  $\nu$  and a non  $\sigma$ -finite measure  $\mu$ , both defined on a  $\sigma$ -algebra, such that  $\nu \ll \mu$  but  $\nu$  is not given by an integral with respect to  $\mu$ .

Problems 7–8 concern harmonic functions and the Poisson integral formula for the unit disk in  $\mathbb{R}^2$ . These matters were the subject of Problems 27–29 at the end of Chapter I, Problems 14–15 at the end of Chapter III, Problems 10–13 at the end of Chapter IV, and Problems 18–20 at the end of Chapter VI. Problem 7 updates the results from Chapter VI so that they apply for  $1 \leq p < \infty$ , and Problem 8 uses weak-star convergence to establish a converse result.

7. If  $1 \leq p < \infty$  and if  $f$  is in  $L^p([-\pi, \pi], \frac{1}{2\pi} d\theta)$ , prove that the Poisson integral  $u(r, \theta)$  of  $f$  has the properties that  $\|u(r, \cdot)\|_p \leq \|f\|_p$  for  $0 \leq r < 1$  and that  $u(r, \cdot)$  tends to  $f$  in  $L^p$  in the sense that  $\lim_{r \uparrow 1} \|u(r, \cdot) - f\|_p = 0$ .
8. Suppose that  $1 < p' \leq \infty$  and that  $u(r, \theta)$  is a harmonic function on the open unit disk such that  $\sup_{0 \leq r < 1} \|u(r, \cdot)\|_{p'}$  is finite. By using Problem 13 at the end of Chapter IV and taking a weak-star limit of a suitable sequence of functions  $u(r_n, \theta)$  with  $\{r_n\}$  increasing to 1, prove that  $u(r, \theta)$  is the Poisson integral of a function in  $L^{p'}([-\pi, \pi], \frac{1}{2\pi} d\theta)$ .

Problems 9–12 concern decomposing any bounded nonnegative additive set function on an algebra into a completely additive part and a “purely finitely additive” part. They make use of Zorn’s Lemma (Section A9 of Appendix A). A bounded nonnegative additive set function  $\mu$  will be called **purely finitely additive** if there is no nonzero completely additive set function  $\nu$  such that  $0 \leq \nu(E) \leq \mu(E)$  for all  $E$ .

9. Suppose that  $\mu$  is an additive set function on the  $\sigma$ -algebra of all subsets of the integers such that  $\mu$  has image  $[0, 1]$  and  $\mu(\{n\}) = 0$  for every integer  $n$ . Prove that  $\mu$  is purely finitely additive. (Such a  $\mu$  was constructed by means of a nontrivial ultrafilter in Problems 39–41 at the end of Chapter V.)
10. Use Zorn’s Lemma to show that any bounded nonnegative additive set function is the sum of a nonnegative completely additive set function and a purely finitely additive set function.



11. Prove that if  $\nu$  is a bounded nonnegative completely additive set function and if  $\mu$  is bounded nonnegative and purely finitely additive with  $0 \leq \mu(E) \leq \nu(E)$  for all  $E$ , then  $\mu = 0$ .
12. Deduce from the previous problem and the Jordan Decomposition Theorem that the decomposition of Problem 10 is unique.

Problems 13–15 prove the theorem, for the case of  $\mathbb{R}^2$ , of Jessen–Marcinkiewicz–Zygmund concerning strong differentiation of integrals of  $L^p$  functions almost everywhere when  $p > 1$ . Strong differentiation holds at  $(x, y)$  for the locally integrable function  $f$  on  $\mathbb{R}^2$  if

$$\lim_{\substack{\text{diam}(R) \rightarrow 0, \\ R = \text{geometric rectangle} \\ \text{centered at } (x, y)}} \frac{1}{m(R)} \int_R f(u, v) \, dv \, du = f(x, y).$$

Let  $f^{**}$  be the associated maximal function, given by

$$f^{**}(x, y) = \sup_{\substack{\text{diam}(R) \rightarrow 0, \\ R = \text{geometric rectangle} \\ \text{centered at } (x, y)}} \frac{1}{m(R)} \int_R |f(u, v)| \, dv \, du.$$

13. Let  $f_1(x, y)$  be the value of the one-dimensional Hardy–Littlewood maximal function of  $y \mapsto f(x, y)$ , and let  $f_2(x, y)$  be the value of the one-dimensional Hardy–Littlewood maximal function of  $x \mapsto f_1(x, y)$ . Prove that  $f^{**}(x, y) \leq f_2(x, y)$ .
14. Using Corollary 9.21 and the previous problem, prove that  $\|f^{**}\|_p \leq A_p^2 \|f\|_p$  if  $1 < p \leq \infty$ .
15. Conclude that strong differentiation holds almost everywhere for each  $f$  in  $L^p(\mathbb{R}^2)$  if  $1 < p \leq \infty$ .

Problems 16–19 concern the Hilbert transform  $H$  defined in Section VIII.7 and Theorem 9.23. The operator  $H$  is defined on  $L^p(\mathbb{R}^1)$  for  $1 < p < \infty$ . Recall the functions  $h_\varepsilon$ ,  $Q_\varepsilon$ , and  $\psi_\varepsilon$  on  $\mathbb{R}^1$  satisfying  $Q_\varepsilon = h_\varepsilon + \psi_\varepsilon$ . Let  $f$  be in  $L^p$ , and let  $f^*$  be the Hardy–Littlewood maximal function of  $f$ .

16. Prove that there exists a continuous integrable function  $\Phi \geq 0$  on  $\mathbb{R}^1$  of the form  $\Phi(x) = \Phi_0(|x|)$ , where  $\Phi_0$  is a decreasing  $C^1$  function on  $[0, \infty)$ , such that the function  $\psi_\varepsilon$  for  $\varepsilon = 1$  satisfies  $|\psi_1| \leq \Phi$ .
17. Deduce from the previous problem and Corollary 6.42 that  $\sup_{\varepsilon > 0} |(\psi_\varepsilon * f)(x)| \leq C f^*(x)$ . How does it follow that  $\lim_{\varepsilon \downarrow 0} (\psi_\varepsilon * f)(x) = 0$  almost everywhere for all  $f$  in  $L^p$ ,  $1 \leq p \leq \infty$ ?
18. Prove that  $Q_\varepsilon * f = P_\varepsilon * (Hf)$  for  $f \in L^p$  with  $1 < p < \infty$ , where  $P_\varepsilon(x) = P(x, \varepsilon)$  is the Poisson kernel.

19. Deduce from the previous two problems that the limit in the equality

$$Hf(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{|t| \geq \varepsilon} \frac{f(x-t) dt}{t}$$

of Theorem 9.23 may be interpreted as an almost-everywhere limit if  $f$  is in  $L^p(\mathbb{R}^1)$  and  $1 < p < \infty$ .

Problems 20–22 prove the theorem of M. Riesz that the partial sums of the Fourier series of a function in  $L^p([-\pi, \pi])$  converge to the function in  $L^p$  if  $1 < p < \infty$ . Recall from Sections I.10 and VI.7 that if  $f$  is integrable on  $[-\pi, \pi]$ , then the  $n^{\text{th}}$  partial sum of the Fourier series of  $f$  is given by  $(S_n f)(x) = (D_n * f)(x)$ , where  $D_n$  is the Dirichlet kernel  $D_n(t) = \frac{\sin(n + \frac{1}{2})t}{\sin \frac{1}{2}t}$  and the convolution is taken relative to  $\frac{1}{2\pi} dt$ .

20. Suppose it can be proved that  $\|S_n f\|_p \leq A_p \|f\|_p$  for  $1 < p < \infty$  with  $A_p$  independent of  $n$  and  $f$ . Prove that  $S_n f \rightarrow f$  in  $L^p$  for all  $f$  in  $L^p$ , provided  $1 < p < \infty$ .

21. Define  $E_n(t) = \frac{2 \sin(n + \frac{1}{2})t}{t}$  for  $\frac{1}{2n+1} \leq |t| \leq \pi$  and  $E_n(t) = 0$  for  $|t| < \frac{1}{2n+1}$ . Then extend  $E_n(t)$  periodically. Show that  $D_n - E_n = \varphi_n$  is integrable on  $[-\pi, \pi]$  with  $\|\varphi_n\|_1 \leq C$  independently of  $n$ , and say why it is therefore enough to prove that the operators  $T_n$  with  $T_n f = E_n * f$  satisfy  $\|T_n f\|_p \leq B_p \|f\|_p$  for  $1 < p < \infty$  with  $B_p$  independent of  $n$  and  $f$ .

22. In  $E_n(t)$ , write  $\sin(n + \frac{1}{2})t$  as a linear combination of two exponentials  $e^{ikt}$ , rewrite each exponential as  $e^{-ik(x-t)} e^{ikx}$ , and decompose the operator  $T_n$  as the corresponding sum of two operators. By relating these two operators separately to the operators  $H_\varepsilon$  in Theorem 9.23, prove that the  $T_n$ 's satisfy the desired estimate  $\|T_n f\|_p \leq B_p \|f\|_p$ .

Problems 23–26 develop a kind of function-valued integration known as **conditional expectation** in probability theory. They make use of the Radon–Nikodym Theorem (Theorem 9.16). Let  $(X, \mathcal{A}, \mu)$  be a measure space with  $\mu(X) = 1$ .

23. If  $f$  is integrable and if  $\mathcal{B}$  is a  $\sigma$ -algebra contained in  $\mathcal{A}$ , prove that there exists a function  $E[f|\mathcal{B}]$  that

- (i) is measurable with respect to  $\mathcal{B}$  and
- (ii) has  $\int_B f d\mu = \int_B E[f|\mathcal{B}] d\mu$  for all  $B$  in  $\mathcal{B}$ .

Show further that  $E[f|\mathcal{B}]$  is unique in this sense: any two functions satisfying (i) and (ii) differ only on a set in  $\mathcal{B}$  of  $\mu$  measure 0.

24. Suppose that  $X$  is a countable disjoint union of sets  $X_n$  in  $\mathcal{A}$  and that  $\mathcal{B}$  consists of all possible unions of the  $X_n$ 's. Give an explicit formula for  $E[f|\mathcal{B}]$ .

25. Show that if  $\mathcal{B} = \mathcal{A}$ , then  $E[f|\mathcal{B}] = f$  almost everywhere.

26. Let  $\mathcal{B}$  and  $\mathcal{C}$  be  $\sigma$ -algebras with  $\mathcal{C} \subseteq \mathcal{B} \subseteq \mathcal{A}$ . Prove the following:
- $E[E[f|\mathcal{B}]|\mathcal{C}] = E[f|\mathcal{C}]$  almost everywhere.
  - If  $f$  and  $g$  are integrable and everywhere finite, then
 
$$E[f+g|\mathcal{B}] = E[f|\mathcal{B}] + E[g|\mathcal{B}]$$
 almost everywhere.
  - If  $g$  is measurable with respect to  $\mathcal{B}$  and if  $f$  and  $fg$  are integrable, then
 
$$E[fg|\mathcal{B}] = gE[f|\mathcal{B}]$$
 almost everywhere.
  - If  $f$  and  $g$  are in  $L^2(X, \mathcal{A}, \mu)$ , then  $\int_X f E[g|\mathcal{B}] d\mu = \int_X E[f|\mathcal{B}] g d\mu$ .

Problems 27–33 concern bounded linear operators  $A : L^p(\mathbb{R}^N, dx) \rightarrow L^q(\mathbb{R}^N, dx)$  that commute with translations. Recall that the translation operators are defined on functions on  $\mathbb{R}^N$  by  $(\tau_x f)(y) = f(y - x)$  and that the condition of commuting with translations means that  $A\tau_x f = \tau_x Af$  for all  $x \in \mathbb{R}^N$  and all functions in question.

27. Using Theorem 8.14, prove that every bounded linear operator from  $L^2(\mathbb{R}^1, dx)$  to itself commuting with translations and dilations is a linear combination of the identity and the Hilbert transform. (The **dilation**  $\delta_r$  for  $r > 0$  is defined on a function  $f$  on  $\mathbb{R}^1$  by  $(\delta_r f)(x) = f(r^{-1}x)$ .)
28. Suppose that  $1 \leq p < \infty$  and  $1 \leq q < \infty$ . Going over the proof of Lemma 8.13 and the results in Chapter VI on which it depends, prove that if  $A : L^p(\mathbb{R}^N, dx) \rightarrow L^q(\mathbb{R}^N, dx)$  is a bounded linear operator that commutes with translations, then  $A$  commutes with convolution by  $L^1$  functions in the following sense: if  $f$  is in  $L^1$  and  $g$  is in  $L^p$ , then the members  $f * Ag$  and  $A(f * g)$  of  $L^q$  are equal.
29. Suppose that  $1 < p < \infty$ , that  $A$  is bounded linear from  $L^p$  to itself, and that  $A$  commutes with translations. Let  $p'$  be the dual index.
- Prove the following for every pair of simple functions  $f$  and  $g$  that vanish off a set of finite measure:  $\int_{\mathbb{R}^N} (Af)(x)g(-x) dx = \int_{\mathbb{R}^N} (Ag)(x)f(-x) dx$ .
  - Taking into account that the space of simple functions vanishing off a set of finite measure is dense in  $L^{p'}$ , prove that  $A$  extends to a bounded linear operator from  $L^{p'}$  to itself commuting with translations and that the norm of  $A : L^{p'} \rightarrow L^{p'}$  equals the norm of  $A : L^p \rightarrow L^p$ .
  - Explain how (b) generalizes in the case of a linear operator bounded from  $L^p$  to  $L^q$  if also  $1 < q < \infty$ .
30. Let  $1 < p \leq 2$ , let  $p'$  be the dual index, and let  $\mathcal{F} : L^p \rightarrow L^{p'}$  be the Fourier transform as defined in the Hausdorff–Young Theorem (Corollary 9.19C). Prove the following generalization of Theorem 8.14: If  $A$  is bounded linear from  $L^p$  to itself and if  $A$  commutes with translations, then there exists an  $L^\infty$  function  $m$  such that  $\mathcal{F}(Af) = m\mathcal{F}(f)$  for all  $f$  in  $L^p$ .

31. Take for granted the **Helly–Bray Theorem**, i.e., the statement that if  $\{\mu_n\}$  is a sequence of finite measures on  $\mathbb{R}^N$  with  $\{\mu_n(\mathbb{R}^N)\}$  bounded, then there is a subsequence convergent to some finite measure  $\mu$  weak-star against  $C_{\text{com}}(\mathbb{R}^N)$ . This result was assumed and used previously for Problems 6–12 in Chapter VIII; it will be proved in something like the stated form in Chapter XI. Carry out the following steps to prove that each bounded linear  $A : L^1 \rightarrow L^1$  commuting with translations is given by convolution with a finite *signed* measure on  $\mathbb{R}^N$ , convolution with a finite measure being defined in Problem 5 at the end of Chapter VIII and the notion being extended from finite measures to finite signed measures by the Jordan decomposition (Theorem 9.14) and linearity:
- Let  $\varphi$  be a function in  $C_{\text{com}}(\mathbb{R}^N)$  with integral 1, and form the approximate identity  $\{\varphi_\varepsilon\}$  as in Theorem 6.20. Show that there is a finite signed measure  $\rho$  on  $\mathbb{R}^N$  such that some sequence  $\{\varphi_{\varepsilon_k}\}$  with  $\varepsilon_k$  decreasing to 0 has  $\lim_k \int_{\mathbb{R}^N} h(x)(A\varphi_{\varepsilon_k})(x) dx = \int_{\mathbb{R}^N} h(x) d\rho(x)$  for all  $h \in C_{\text{com}}(\mathbb{R}^N)$ .
  - With  $g \in C_{\text{com}}(\mathbb{R}^N)$  and with  $g * \rho$  defined as the function  $(g * \rho)(x) = \int_{\mathbb{R}^N} g(x - y) d\rho(y)$ , prove that  $g * \rho$  is a continuous function.
  - For any function  $f$  on  $\mathbb{R}^N$ , define  $f^\#(x) = f(-x)$ . Prove for every  $h$  in  $C_{\text{com}}(\mathbb{R}^N)$  that  $\lim_k (Ah^\# * \varphi_{\varepsilon_k})(y) = (h^\# * \rho)(y)$  for every  $y \in \mathbb{R}^N$ .
  - Conclude that  $Af = f * \rho$  for all  $f$  in  $L^1$ , hence that  $A$  is given by convolution with a finite signed measure.
32. Suppose that  $1 \leq p < q < \infty$ . Give an example of a nonzero bounded linear operator  $A : L^p(\mathbb{R}^N, dx) \rightarrow L^q(\mathbb{R}^N, dx)$  commuting with translations.
33. Suppose that  $1 \leq q < p < \infty$ . Prove that there is no nonzero bounded linear operator  $A : L^p(\mathbb{R}^N, dx) \rightarrow L^q(\mathbb{R}^N, dx)$  commuting with translations. Carry out the following steps to do so:
- Make use of the fact that  $C_{\text{com}}(\mathbb{R}^N)$  is dense in  $L^p$  to prove that  $\lim_{h \rightarrow \infty} \|\tau_h f + f\|_p = 2^{1/p} \|f\|_p$ .
  - Arguing by contradiction, suppose that such an  $A$  has norm  $M > 0$ . Let  $f$  be arbitrary in  $L^p$ . Obtain an estimate  $\|\tau_h(Af) + Af\|_q \leq M \|\tau_h f + f\|_p$  for all  $f \in L^p$ , let  $h$  tend to infinity, and derive a contradiction.