6 Linear functionals

In this chapter we deal with the representations of dual spaces, i.e. we will state canonical isomorphisms between the most important dual spaces and already known spaces. We will use this to solve boundary value problems for partial differential equations.

The most important case is that of a Hilbert space, for which the dual space is isomorphic to the space itself (theorem 6.1). As a consequence, we obtain the Lax-Milgram theorem (see 6.2), with the help of which elliptic boundary value problems can be solved (see 6.4 – 6.9).

In the second part, we state representations of the dual spaces of $L^p(\mu)$ for $p < \infty$ (see 6.12) and of $C^0(S)$ (see 6.23). The proof of 6.23 will employ the Hahn-Banach theorem (see 6.14 – 6.15). This theorem states that continuous linear maps $f : Y \to \mathbb{K}$ can be extended from a subspace $Y \subset X$ to the full space X such that the norm of the map is maintained, which is one of the general principles of functional analysis. A constructive proof of the Hahn-Banach theorem for separable spaces X will be given in 9.2.

Lax-Milgram's theorem

We start with an existence theory, which is based on the following result.

6.1 Riesz representation theorem. If X is a Hilbert space, then

$$J(x)(y) := (y\,,\,x)_X \qquad \text{for } x, y \in X$$

defines an isometric conjugate linear isomorphism $J : X \to X'$.
Notation: In the remainder of this book we will also denote this isomorphism by $R_X : X \to X'$.
Definition: Here a map J is called ***conjugate linear*** if for all $x, y \in X$ and $\alpha \in \mathbb{K}$ it holds that $J(\alpha x + y) = \overline{\alpha} J(x) + J(y)$. In the case $\mathbb{K} = \mathbb{R}$ this reduces to J being linear.

Proof. By the Cauchy-Schwarz inequality,

$$|J(x)(y)| \leq \|x\|_X \cdot \|y\|_Y\,,$$

i.e. $J(x) \in X'$ with $\|J(x)\|_{X'} \leq \|x\|_X$. On noting that $|J(x)(x)| = \|x\|_X^2$, we see that $\|J(x)\|_{X'} \geq \|x\|_X$. Hence J is isometric, and in particular injective.

H.W. Alt, *Linear Functional Analysis*, Universitext,
DOI 10.1007/978-1-4471-7280-2_6

Now the crucial step is to show that J is surjective. Let $0 \neq x_0' \in X'$ and let P be the orthogonal projection from 4.3 onto the closed null space $\mathscr{N}(x_0')$. Choose $e \in X$ with $x_0'(e) = 1$ and define

$$x_0 := e - Pe, \qquad \text{hence also } x_0'(x_0) = 1,$$

and in particular $x_0 \neq 0$. Now it follows from 4.3 (see 4.4(2)) that

$$(y\,,\, x_0)_X = 0 \quad \text{for all } y \in \mathscr{N}(x_0'). \tag{6-3}$$

For all $x \in X$,

$$x = \underbrace{x - x_0'(x)x_0}_{\in \mathscr{N}(x_0')} + x_0'(x)x_0\,,$$

which together with (6-3) yields that

$$(x\,,\, x_0)_X = (x_0'(x)x_0\,,\, x_0)_X = x_0'(x)\|x_0\|^2\,,$$

and hence

$$x_0'(x) = \left(x\,,\, \frac{x_0}{\|x_0\|^2}\right)_X = J\Big(\frac{x_0}{\|x_0\|^2}\Big)(x)\,.$$

□

An application of the Riesz representation theorem is the

6.2 Lax-Milgram theorem. Let X be a Hilbert space over $\mathbb{K}$ and let $a : X \times X \to \mathbb{K}$ be sesquilinear. Assume that there exist constants c_0 and C_0 with $0 < c_0 \leq C_0 < \infty$ such that for all $x, y \in X$

(1) $|a(x,y)| \leq C_0\|x\|_X\|y\|_X$ (***Continuity***),
(2) $\mathrm{Re}\,a(x,x) \geq c_0\|x\|_X^2$ (***Coercivity***).

Then there exists a unique map $A : X \to X$ with

$$a(y,x) = (y\,,\, Ax)_X \quad \text{for all } x, y \in X.$$

In addition, $A \in \mathscr{L}(X)$ is an invertible operator with

$$\|A\| \leq C_0 \quad \text{and} \quad \left\|A^{-1}\right\| \leq \frac{1}{c_0}\,.$$

Proof. For every $x \in X$ it follows from (1) that the function $a(\bullet, x)$ lies in X' and satisfies

$$\|a(\bullet, x)\|_{X'} \leq C_0\|x\|_X\,.$$

Hence, by the Riesz representation theorem 6.1, there exists a unique element $A(x) \in X$ such that

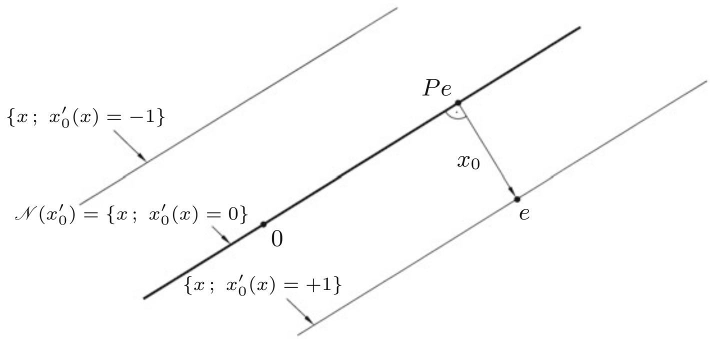

Fig. 6.1. *Proof of the Riesz representation theorem*

$$a(y,x) = (y\,,\,A(x))_X \qquad \text{for all } y \in X$$

and moreover

$$\|A(x)\|_X = \|a(\bullet,x)\|_{X'} \le C_0\|x\|_X\,.$$

Since a and the scalar product are conjugate linear in the second argument, it follows that A is linear. Hence $A \in \mathscr{L}(X)$ with $\|A\| \le C_0$. Moreover,

$$c_0\|x\|_X^2 \le \operatorname{Re}\, a(x,x) = \operatorname{Re}\,(x\,,\,A(x))_X \le \|x\|_X \cdot \|Ax\|_X\,,$$

and so

$$c_0\|x\|_X \le \|A(x)\|_X \qquad \text{for all } x \in X, \tag{6-4}$$

which implies that $\mathscr{N}(A) = \{0\}$. In addition, it follows that the image space $\mathscr{R}(A)$ is closed, on noting that for $x_k, x \in X$

$$\begin{aligned}
&A(x_k) \to y \qquad \text{as } k \to \infty\\
\Longrightarrow\quad &\|x_k - x_l\|_X \le \tfrac{1}{c_0}\|A(x_k - x_l)\|_X \qquad \text{(recall (6-4))}\\
&\qquad\qquad = \tfrac{1}{c_0}\|A(x_k) - A(x_l)\|_X \to 0 \qquad \text{as } k,l \to \infty\\
\Longrightarrow\quad &x_k \to x \qquad \text{as } k \to \infty \text{ for an } x \in X\\
\Longrightarrow\quad &A(x_k) \to A(x) \qquad \text{(as } A \text{ is continuous)}\\
\Longrightarrow\quad &y = Ax \in \mathscr{R}(A)\,.
\end{aligned}$$

It remains to show that $\mathscr{R}(A) = X$. If $\mathscr{R}(A) \ne X$, then, on recalling that $\mathscr{R}(A)$ is a closed subspace, the projection theorem 4.3 yields that there exists an $x_0 \in X \setminus \mathscr{R}(A)$ such that (recall 4.4(2))

$$(y\,,\,x_0)_X = 0 \qquad \text{for all } y \in \mathscr{R}(A)$$

(choose an $\widetilde{x}_0 \in X \setminus \mathscr{R}(A)$ and set $x_0 := \widetilde{x}_0 - P\widetilde{x}_0$, where P is the orthogonal projection onto $\mathscr{R}(A)$). This yields, on setting $y = A(x_0)$, that

$$0 = \mathrm{Re}\,(A(x_0)\,,\,x_0)_X = \mathrm{Re}\,(x_0\,,\,A(x_0))_X = \mathrm{Re}\;a(x_0,x_0) \geq c_0\|x_0\|_X^2 > 0\,,$$

a contradiction. Hence we have shown that A is bijective. It follows from (6-4) that $\left\|A^{-1}\right\| \leq \frac{1}{c_0}$. □

6.3 Consequences.

(1) Let A be the operator from 6.2 and let R_X be the isometry from theorem 6.1. For a given $x' \in X'$, the unique solution of

$$a(y,x) = x'(y) \quad \text{for all } y \in X \tag{6-5}$$

is then $x := A^{-1}R_X^{-1}x'$.

(2) The solution in (1) has the ***stability property***

$$\|x\|_X \leq \frac{1}{c_0}\|x'\|_{X'}\,. \tag{6-6}$$

Interpretation: If we consider two "right-hand sides" x'_1 and x'_2 and the corresponding solutions x_1 and x_2 in (1), then it follows from (6-6), due to the linearity of the problem ($x_1 - x_2$ is the solution to $x'_1 - x'_2$), that

$$\|x_1 - x_2\|_X \leq \tfrac{1}{c_0}\|x'_1 - x'_2\|_{X'}\,.$$

Hence the error in the solutions can be estimated by the error in the data. This justifies the term stability.

(3) Formulated for the operator A, the Lax-Milgram theorem reads as follows: Let X be a Hilbert space and let $A \in \mathscr{L}(X)$ be ***coercive***, i.e. there exists a constant $c_0 > 0$ such that

$$\mathrm{Re}\,(x\,,\,Ax)_X \geq c_0\|x\|_X^2 \quad \text{for all } x \in X.$$

Then A is invertible, with $\left\|A^{-1}\right\| \leq \frac{1}{c_0}$.

(4) If a in 6.2 is a scalar product, then the solution x in statement (1) is, in addition, the uniquely determined absolute minimum of the functional

$$E(y) := \tfrac{1}{2}a(y,y) - \mathrm{Re}x'(y)\,.$$

Proof (1) *and* (2). By the definition of A and R_X, for all $x, y \in X$

$$a(y,x) = (y\,,\,Ax)_X = (R_X Ax)(y)\,,$$

and $R_X A : X \to X'$ is bijective. If $x = (R_X A)^{-1}x'$, then it follows from (6-4) that

$$c_0\|x\|_X \leq \|Ax\|_X = \left\|R_X^{-1}x'\right\|_X = \|x'\|_{X'}\,.$$

□

Proof (3). The product $a(y,x) := (y\,,\,Ax)_X$ satisfies the properties in 6.2 with $C_0 = \|A\|$. Moreover, A is the operator corresponding to a from 6.2. □

Proof (4). Let $y \in X$. Then

$$\begin{aligned} E(y) - E(x) &= \tfrac{1}{2}\big(a(y,y) - a(x,x)\big) - \mathrm{Re}\, x'(y-x) \\ &= \tfrac{1}{2}\big(a(y,y) - a(x,x)\big) - \mathrm{Re}\, a(y-x,x) \\ &= \tfrac{1}{2}\big(a(y,y) - a(y,x) - a(x,y) + a(x,x)\big) \\ &= \tfrac{1}{2}a(y-x,y-x) \geq \tfrac{c_0}{2}\|y-x\|_X^2 \,. \end{aligned}$$

□

The Lax-Milgram theorem has applications for integral operators (see E6.3) and for differential operators, which will be discussed in the following. First we consider the classical case in spaces of continuous functions.

6.4 Elliptic boundary value problems. Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $\mathbb{K} = \mathbb{R}$. We want to find functions $u \in C^2(\Omega)$ satisfying the differential equation

$$-\sum_{i=1}^{n} \partial_i \Big(\sum_{j=1}^{n} a_{ij}\partial_j u + h_i \Big) + bu + f = 0 \quad \text{in } \Omega\,. \tag{6-7}$$

Here $a_{ij}, h_i \in C^1(\Omega)$ for $i,j = 1,\dots,n$ and $f, b \in C^0(\Omega)$ are given real-valued functions, and we assume that there exists a $c_0 > 0$ such that for all $x \in \Omega$,

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq c_0|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^n. \tag{6-8}$$

We then say that the matrix $\big(a_{ij}(x)\big)_{i,j}$ is ***uniformly elliptic*** in x. (For every $c > 0$, the set of points $\xi \in \mathbb{R}^n$, for which $\sum_{i,j} a_{ij}(x)\xi_i\xi_j = c$, is an ellipsoid.) Let us emphasize here that the matrix $\big(a_{ij}(x)\big)_{i,j}$ need not be symmetric.

It turns out that, under certain assumptions, there exists a unique function u solving (6-7), which in addition satisfies suitable boundary conditions on $\partial\Omega$. The two most frequently occurring boundary conditions in mathematical physics are:

(1) ***Dirichlet boundary condition.*** Let $g \in C^0(\partial\Omega)$ be given. Find a function $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ which solves the following ***Dirichlet boundary value problem***:

$$u \text{ satisfies (6-7) in } \Omega\,, \quad u = g \text{ on } \partial\Omega.$$

(2) ***Neumann boundary condition.*** We assume that Ω has a C^1-boundary, i.e. that the boundary $\partial\Omega$ can be locally represented as the graph of a C^1-function in an appropriately chosen coordinate system (as in A8.2). In addition, we assume that $a_{ij}, h_i \in C^0(\overline{\Omega})$. Let $g \in C^0(\partial\Omega)$ be given. Find a function $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ which solves the following ***Neumann boundary value problem***:

$$u \text{ satisfies (6-7) in } \Omega\,, \quad -\sum_{i=1}^{n} \nu_i \Big(\sum_{j=1}^{n} a_{ij}\partial_j u + h_i\Big) = g \text{ on } \partial\Omega.$$

Here $\nu = (\nu_i)_{i=1,\dots,n}$ is the ***outer normal*** to $\partial\Omega$.

Remark: For the boundary value problem (1) to be at all solvable, there must exist some function $u_0 \in C^2(\Omega) \cap C^0(\overline{\Omega})$ with $u_0 = g$ on $\partial\Omega$. Then the boundary value problem can be transformed to one for $\widetilde{u} := u - u_0$, by replacing g with $\widetilde{g} := 0$, h_i with $\widetilde{h_i} := h_i + \sum_j a_{ij}\partial_j u_0$, and f with $\widetilde{f} := f + bu_0$. Analogously, for (2) there must exist a function $u_0 \in C^2(\Omega) \cap C^1(\overline{\Omega})$ with $-\sum_i \nu_i \big(\sum_j a_{ij}\partial_j u_0 + h_i\big) = g$ on $\partial\Omega$. Then the boundary value problem can be transformed to one for $\widetilde{u} := u - u_0$, by replacing g with $\widetilde{g} := 0$, h_i with $\widetilde{h_i} := 0$, and f with $\widetilde{f} := f - \sum_i \partial_i \big(\sum_j a_{ij}\partial_j u_0 + h_i\big) + bu_0$. We then call the boundary conditions ***homogeneous***.

We now give an equivalent definition of the boundary value problem with the help of test functions (this gives a connection to distributions, which were treated at the end of section 5).

In the Dirichlet case, if we multiply the differential equation (6-7) by functions $\zeta \in C_0^\infty(\Omega)$, then we obtain after integration by parts that

$$\int_\Omega \Big(\sum_i \partial_i\zeta \Big(\sum_j a_{ij}\partial_j u + h_i\Big) + \zeta(bu + f)\Big)\, \mathrm{dL}^n = 0\,. \tag{6-9}$$

Conversely, if this integral identity is satisfied for all $\zeta \in C_0^\infty(\Omega)$, then we obtain, on reversing the integration by parts, that

$$\int_\Omega \zeta w \, \mathrm{dL}^n = 0 \quad \text{with} \quad w := -\sum_i \partial_i \Big(\sum_j a_{ij}\partial_j u + h_i\Big) + bu + f\,.$$

If we assume that $w(x_0) \neq 0$ for some $x_0 \in \Omega$, then we can choose an $\varepsilon_0 > 0$ with $w > 0$ or $w < 0$ in $\mathrm{B}_{\varepsilon_0}(x_0) \subset \Omega$, and then a nontrivial $\zeta \in C_0^\infty\big(\mathrm{B}_{\varepsilon_0}(x_0)\big)$ with $\zeta \geq 0$, in order to obtain a contradiction. Hence it follows that $w = 0$ in Ω (this also follows directly from 4.22), i.e. the differential equation (6-7) holds in Ω.

Similarly, in the Neumann case, if we multiply the differential equation (6-7) by functions $\zeta \in C^\infty(\overline{\Omega})$, on assuming that $a_{ij}, h_i \in C^1(\overline{\Omega})$, we obtain after integration by parts that

$$\int_\Omega \Big(\sum_i \partial_i\zeta \Big(\sum_j a_{ij}\partial_j u + h_i\Big) + \zeta(bu + f)\Big)\, \mathrm{dL}^n + \int_{\partial\Omega} \zeta g \, \mathrm{dH}^{n-1} = 0\,. \tag{6-10}$$

Conversely, if this holds for all $\zeta \in C^\infty(\overline{\Omega})$, then as before we obtain the differential equation in Ω (here it is sufficient to consider $\zeta \in C_0^\infty(\Omega)$), and then it holds for $\zeta \in C^\infty(\overline{\Omega})$ that

$$\int_{\partial\Omega} \zeta w \, \mathrm{dH}^{n-1} = 0 \quad \text{with} \quad w := \sum_i \nu_i \Big(\sum_j a_{ij} \partial_j u + h_i \Big) + g.$$

Similarly to the argumentation above, it now follows that the Neumann boundary condition is satisfied.

The basic idea for the solution of these boundary value problems with the help of Hilbert space methods is to interpret the integral terms in (6-9) and (6-10) as an L^2-bilinear form, and enlarge the spaces for test functions and solutions accordingly. As the test function appears with ζ and $\partial_i\zeta$, the appropriate test space for (6-9) is the closure of $C_0^\infty(\Omega)$ in the space $W^{1,2}(\Omega)$, i.e. the space $W_0^{1,2}(\Omega)$ (see 3.29). Since functions in $W_0^{1,2}(\Omega)$, when Ω has a C^1-boundary, have in a weak sense boundary values 0 (see A8.10), $W_0^{1,2}(\Omega)$ is also the appropriate enlarged solution space. For (6-10) the appropriate test space is the closure of $C^\infty(\overline{\Omega})$ in the space $W^{1,2}(\Omega)$, i.e. for sets Ω with a C^1-boundary the space $W^{1,2}(\Omega)$ itself (see A8.7), which is also the appropriately enlarged solution space.

For the resulting weak formulations of the problem it is no longer necessary to assume that the data a_{ij}, h_i, b, f of the problem are continuous functions in Ω. However, it is necessary to make assumptions on their integrability, for instance as formulated in the following:

6.5 Weak boundary value problems. With $\mathbb{K} = \mathbb{R}$ it is assumed in the following that $\Omega \subset \mathbb{R}^n$ is open and bounded, that $a_{ij} \in L^\infty(\Omega)$ satisfy the ellipticity condition (6-8) for almost all $x \in \Omega$, and that $b \in L^\infty(\Omega)$ and $h_i, f \in L^2(\Omega)$. The weak formulation of the boundary value problem in 6.4 is defined as follows (where we consider only the case $g = 0$):

(1) We call $u : \Omega \to \mathbb{R}$ a ***weak solution*** of the ***Dirichlet problem*** if

$$\begin{aligned} & u \in W_0^{1,2}(\Omega) \text{ and} \\ & \int_\Omega \Big(\sum_i \partial_i \zeta \Big(\sum_j a_{ij} \partial_j u + h_i \Big) + \zeta (bu + f) \Big) \, \mathrm{dL}^n = 0 \\ & \text{for all } \zeta \in W_0^{1,2}(\Omega). \end{aligned}$$

Here, as remarked above, if Ω has a C^1-boundary, then the condition $u \in W_0^{1,2}(\Omega)$ in a weak sense contains the homogeneous boundary conditions, and it is irrelevant whether ζ varies in the space $W_0^{1,2}(\Omega)$, or only in the space $C_0^\infty(\Omega)$.

(2) We call $u : \Omega \to \mathbb{R}$ a ***weak solution*** of the ***Neumann problem*** if

$$\begin{aligned} & u \in W^{1,2}(\Omega) \text{ and} \\ & \int_\Omega \Big(\sum_i \partial_i \zeta \Big(\sum_j a_{ij} \partial_j u + h_i \Big) + \zeta (bu + f) \Big) \, \mathrm{dL}^n = 0 \\ & \text{for all } \zeta \in W^{1,2}(\Omega). \end{aligned}$$

Here, as explained above, if Ω has a C^1-boundary, then the integral term in a weak sense contains the homogeneous boundary conditions (for $g = 0$ in 6.4(2) the boundary integral in (6-10) vanishes), and it is irrelevant whether ζ varies in the space $W^{1,2}(\Omega)$, or only in the space $C^\infty(\overline{\Omega})$.

We will now prove the existence of solutions to these weak boundary value problems.

6.6 Existence theorem for the Neumann problem. Let the assumptions in 6.5 hold and let $b_0 > 0$ with $b(x) \geq b_0$ for almost all $x \in \Omega$. Then there exists a unique solution $u \in W^{1,2}(\Omega)$ for the Neumann problem in 6.5(2). Moreover,

$$\|u\|_{W^{1,2}} \leq C\big(\|h\|_{L^2} + \|f\|_{L^2}\big)\,,$$

with a constant C that is independent of h and f.

Proof. For $u, v \in W^{1,2}(\Omega)$ we define

$$a(u,v) := \sum_{i,j} \int_\Omega \partial_i u \cdot a_{ij} \partial_j v \, \mathrm{dL}^n + \int_\Omega u \cdot bv \, \mathrm{dL}^n\,. \tag{6-11}$$

(We mention that in general a does not need to be a scalar product, for $(a_{ij})_{ij}$ can be asymmetric.) Then a is bilinear, with

$$\begin{aligned} |a(u,v)| &\leq \sum_{i,j} \|a_{ij}\|_{L^\infty} \|\partial_i u\|_{L^2} \|\partial_j v\|_{L^2} + \|b\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} \\ &\leq C\|u\|_{W^{1,2}} \|v\|_{W^{1,2}} \qquad \text{with } C := \sum_{i,j} \|a_{ij}\|_{L^\infty} + \|b\|_{L^\infty}\,. \end{aligned}$$

In addition, it follows from the assumptions on a_{ij} and b that

$$a(u,u) \geq c_0 \int_\Omega |\nabla u|^2 \, \mathrm{dL}^n + b_0 \int_\Omega |u|^2 \, \mathrm{dL}^n \geq c \cdot \|u\|^2_{W^{1,2}}$$

with $c := \min(c_0, b_0)$. Hence a satisfies the assumptions of the Lax-Milgram theorem 6.2 on the Hilbert space $W^{1,2}(\Omega)$. We want to find a $u \in W^{1,2}(\Omega)$ such that

$$a(v,u) = F(v) \qquad \text{for all } v \in W^{1,2}(\Omega),$$

where

$$F(v) := -\int_\Omega \Big(\sum_i \partial_i v \cdot h_i + vf\Big) \, \mathrm{dL}^n\,. \tag{6-12}$$

It follows from 6.3(1) that there exists a unique such u if F belongs to the dual space of $W^{1,2}(\Omega)$. But this is the case, since F is linear, with

$$|F(v)| \leq \|h\|_{L^2} \|\nabla v\|_{L^2} + \|f\|_{L^2} \|v\|_{L^2} \leq (\|h\|_{L^2} + \|f\|_{L^2})\, \|v\|_{W^{1,2}}\,.$$

In addition, the solution u can be estimated by the data, since, by 6.3(2),

$$\|u\|_{W^{1,2}} \le \frac{1}{c}\|F\| \le \frac{1}{c}\left(\|h\|_{L^2} + \|f\|_{L^2}\right).$$

□

The Dirichlet problem can also be solved in the case $b = 0$. Here we need the following

6.7 Poincaré inequality. If $\Omega \subset \mathbb{R}^n$ is open and bounded, then there exists a constant C_0 (which depends on Ω), such that

$$\int_\Omega |u|^2 \,\mathrm{dL}^n \le C_0 \int_\Omega |\nabla u|^2 \,\mathrm{dL}^n \qquad \text{for all } u \in W_0^{1,2}(\Omega).$$

Note: See also 8.16 and E10.10.

Proof. On noting that both sides of the inequality depend continuously on u in the $W^{1,2}$-norm, and on recalling the definition of $W_0^{1,2}(\Omega)$, it is sufficient to prove the estimate for functions $u \in C_0^\infty(\Omega)$. In the case $n = 1$, let $\Omega \subset [a,b] \subset \mathbb{R}$. Then the Hölder inequality yields for $a \le x \le b$, on setting $u = 0$ in $\mathbb{R} \setminus \Omega$, that

$$\begin{aligned}|u(x)|^2 = |u(x) - u(a)|^2 &= \left|\int_a^x \partial_x u(y)\,\mathrm{d}y\right|^2 \\ &\le (x-a)\int_a^x |\partial_x u(y)|^2 \,\mathrm{d}y \le (b-a)\int_a^b |\partial_x u(y)|^2\,\mathrm{d}y\,.\end{aligned}$$

Integration over x gives

$$\int_a^b |u|^2 \,\mathrm{dL}^1 \le (b-a)^2 \int_a^b |\partial_x u|^2 \,\mathrm{dL}^1\,. \tag{6-13}$$

In the case $n > 1$, let $\Omega \subset [a,b] \times \mathbb{R}^{n-1}$. Then we obtain (6-13) by integrating over x_1. Integration over the remaining $n-1$ coordinates then yields the desired result. (Hence the Poincaré inequality also holds for infinite slab domains.) □

6.8 Existence theorem for the Dirichlet problem. Let the assumptions in 6.5 hold and let $b \ge 0$. Then there exists a unique weak solution $u \in W_0^{1,2}(\Omega)$ for the Dirichlet problem in 6.5(1). Moreover,

$$\|u\|_{W^{1,2}} \le C\big(\|h\|_{L^2} + \|f\|_{L^2}\big)$$

with a constant C that is independent of h and f.

Proof. Consider the bilinear form a in (6-11), now on the Hilbert space $W_0^{1,2}(\Omega)$. As in the proof of 6.6,

$$|a(u,v)| \le C\|u\|_{W^{1,2}}\|v\|_{W^{1,2}}$$

and the assumptions on the coefficients yield that

$$a(u,u) \ge c_0 \int_\Omega |\nabla u|^2 \,\mathrm{dL}^n = c_0 \|\nabla u\|_{L^2}^2 \quad \text{for } u \in W_0^{1,2}(\Omega)\,.$$

Then it follows, with the constant C_0 from 6.7, that

$$\|u\|_{W^{1,2}}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 \le (C_0+1)\|\nabla u\|_{L^2}^2 \le \frac{C_0+1}{c_0} a(u,u)\,,$$

and so $a(u,u) \ge c\,\|u\|_{W^{1,2}}^2$ with $c = c_0 \cdot (C_0+1)^{-1}$. Hence a satisfies the assumptions of the Lax-Milgram theorem 6.2 on the Hilbert space $W_0^{1,2}(\Omega)$. The functional F in (6-12), restricted to the space $W_0^{1,2}(\Omega)$, then lies in its dual space. Hence, by 6.3(1), there exists a unique $u \in W_0^{1,2}(\Omega)$ with

$$a(v,u) = F(v) \quad \text{for all } v \in W_0^{1,2}(\Omega)\,.$$

The estimate follows again from 6.3(2) (see the proof 6.6). □

6.9 Remark (Regularity of the solution). Based on the existence proofs in 6.6 and 6.8 for weak solutions of the boundary value problem, it is possible to show a posteriori that a weak solution is indeed a classical solution of the boundary value problem in the sense of 6.4, provided the data a_{ij}, h_i, b, f and $\partial\Omega$ satisfy certain regularity conditions (by the regularity theory for partial differential equations, see e.g. [GilbargTrudinger]). If we assume, for instance, that $a_{ij} \in C^{m,1}(\Omega)$, $h_i \in W^{m+1,2}(\Omega)$ and $f \in W^{m,2}(\Omega)$ with $m \ge 0$, then it follows that $u \in W_{\mathrm{loc}}^{m+2,2}(\Omega)$ (see Friedrichs' theorem A12.2). If in addition $\partial\Omega$ is locally given by graphs of $C^{m+1,1}$-functions, then one can correspondingly show that $u \in W^{m+2,2}(\Omega)$ (see A12.3). These two theorems constitute the L^2-regularity theory. This compares with the L^p-theory, which is based on the Calderón-Zygmund inequality in 10.20, and the Schauder theory, which on the basis of the Hölder-Korn-Lichtenstein inequality in 10.19 gives regularity results in Hölder spaces.

Radon-Nikodým's theorem

After we have shown in 6.1 that the dual space of a Hilbert space is canonically isomorphic to the Hilbert space itself, we now want to consider specific Banach spaces, $L^p(\mu)$ and $C^0(S)$, and characterize their dual spaces. (a list of dual spaces can be found in [DunfordSchwartz : IV 15, S. 374-379]). First we state a characterization of $L^p(\mu)'$, for which we will need the Radon-Nikodým theorem 6.11.

6.10 Definition (Variational measure). Let $\mathcal{B}$ be a ring over a set S (see A3.1) and let $\lambda : \mathcal{B} \to \mathbb{K}^m$ be additive. For $E \in \mathcal{B}$ define

$$|\lambda|(E) := \sup \Big\{ \sum_{i=1}^{k} |\lambda(E_i)| \ ; \ k \in \mathbb{N},\ E_i \in \mathcal{B} \text{ pairwise disjoint},\ E_i \subset E \Big\} .$$

It holds that $|\lambda| : \mathcal{B} \to [0, \infty]$ is additive. We also call $|\lambda|$ the ***variational measure*** for λ. In addition, in the case where $\mathcal{B}$ contains the set S, we call

$$\|\lambda\|_{\mathrm{var}} := |\lambda|(S)$$

the ***total variation*** of λ. The measure λ is called a ***bounded measure*** if $\|\lambda\|_{\mathrm{var}} < \infty$.

Proof. We prove the additivity of $|\lambda|$. If $B_1, B_2 \in \mathcal{B}$ are disjoint, then it is easy to see that

$$|\lambda|(B_1) + |\lambda|(B_2) \le |\lambda|(B_1 \cup B_2) .$$

Moreover, for $\varepsilon > 0$ choose disjoint $E_i \in \mathcal{B}$, $i = 1, \ldots, k$, with $E_i \subset B_1 \cup B_2$, such that

$$\begin{aligned} |\lambda|(B_1 \cup B_2) - \varepsilon \le \sum_{i=1}^{k} |\lambda(E_i)| &= \sum_{i=1}^{k} |\lambda(E_i \cap B_1) + \lambda(E_i \cap B_2)| \\ &\le |\lambda|(B_1) + |\lambda|(B_2) . \end{aligned}$$

□

6.11 Radon-Nikodým theorem. Let $(S, \mathcal{B}, \mu)$ be a σ-finite measure space and let

$$\nu : \mathcal{B} \to \mathbb{K} \quad \text{be } \sigma\text{-additive with } \|\nu\|_{\mathrm{var}} < \infty.$$

In addition, let ν be ***absolutely continuous*** with respect to μ, i.e. for all $E \in \mathcal{B}$

$$\mu(E) = 0 \quad \Longrightarrow \quad \nu(E) = 0 .$$

Then there exists a unique function $f \in L^1(\mu)$ such that

$$\nu(E) = \int_E f \, \mathrm{d}\mu \quad \text{for all } E \in \mathcal{B}.$$

Remark: The function f is called the ***Radon-Nikodým derivative*** of ν with respect to μ, and is also denoted by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

Proof. Let $f_1, f_2 \in L^1(\mu)$ be two such representing functions and let $f := f_1 - f_2$. Let $E := \{x \in S ;\ f(x) \bullet e \ge \delta\}$, where $e \in \mathbb{K} \setminus \{0\}$ and $\delta > 0$. Then (recall 5.11)

$$0 = \Big(\int_E f_1 \,\mathrm{d}\mu - \int_E f_2 \,\mathrm{d}\mu\Big) \bullet e = \int_E f \bullet e \,\mathrm{d}\mu \ge \delta\mu(E)\,,$$

and so $\mu(E) = 0$ for all e, δ, which implies that $f_1 = f_2$ μ-almost everywhere. This proves the uniqueness.

In order to prove the existence, we may assume that ν is real-valued (otherwise consider the real and imaginary part separately). It follows from the Hahn decomposition (see A6.2) that we may further assume that ν is nonnegative. Then $(S, \mathcal{B}, \mu + \nu)$ is also a measure space, since for $N \in \mathcal{B}$ and $E \subset S$

$$\begin{aligned} &(\mu+\nu)(N) = 0,\ E \subset N \\ \Longrightarrow\quad &\mu(N) = 0,\ E \subset N \quad\Longrightarrow\quad E \in \mathcal{B},\ \mu(E) = 0\,. \end{aligned}$$

Now ν induces a measure space $(S, \widehat{\mathcal{B}}, \nu)$ with $\mathcal{B} \subset \widehat{\mathcal{B}}$, where the sets from $\widehat{\mathcal{B}}$ are unions of sets from $\mathcal{B}$ with ν-null sets. Since $\nu \le \mu + \nu$, it holds that $L^1(\mu+\nu)$ is contained in $L^1(\nu)$. On recalling that $\nu(S) < \infty$, it follows from the Hölder inequality that $L^2(\nu) \subset L^1(\nu)$. Hence if $g \in L^2(\mu+\nu)$, then

$$\Big|\int_S g \,\mathrm{d}\nu\Big| \le \sqrt{\nu(S)}\|g\|_{L^2(\nu)} \le \sqrt{\nu(S)}\|g\|_{L^2(\mu+\nu)}\,.$$

As $L^2(\mu+\nu)$ is a Hilbert space, the Riesz representation theorem 6.1 then implies that there exists an $h \in L^2(\mu+\nu)$ such that, for all $g \in L^2(\mu+\nu)$,

$$\int_S g \,\mathrm{d}\nu = (g\,,\,h)_{L^2(\mu+\nu)} = \int_S gh \,\mathrm{d}(\mu+\nu)\,,$$

i.e.

$$\int_S g(1-h)\,\mathrm{d}\nu = \int_S gh \,\mathrm{d}\mu \qquad \text{for all } g \in L^2(\mu+\nu). \tag{6-14}$$

We now show that

$$0 \le h < 1 \qquad (\mu+\nu)\text{-almost everywhere.}$$

On setting $g = \mathcal{X}_{\{h<0\}\cap S_m}$, where $\{h < 0\} := \{x \in S\,;\ h(x) < 0\}$ and S_m is as in 3.9(4), it follows from (6-14) that

$$0 \le \int_{\{h<0\}\cap S_m} (1-h)\,\mathrm{d}\nu = \int_{\{h<0\}\cap S_m} h \,\mathrm{d}\mu \le -\varepsilon\mu\left(\{h < -\varepsilon\} \cap S_m\right)\,.$$

This implies that $\mu\left(\{h < -\varepsilon\} \cap S_m\right) = 0$ for all $\varepsilon > 0$ and all m, and hence also $\mu\left(\{h < 0\}\right) = 0$. Since ν is absolutely continuous with respect to μ, it follows that also $\nu\left(\{h < 0\}\right) = 0$. Similarly, it follows from (6-14) that, when $g = \mathcal{X}_{\{h\ge 1\}\cap S_m}$,

$$0 \ge \int_{\{h\ge 1\}\cap S_m} (1-h)\,\mathrm{d}\nu = \int_{\{h\ge 1\}\cap S_m} h \,\mathrm{d}\mu \ge \mu\left(\{h \ge 1\} \cap S_m\right)\,,$$

and so $\mu(\{h \geq 1\}) = 0$, which by assumption yields that $\nu(\{h \geq 1\}) = 0$. This shows that $0 \leq h < 1$ almost everywhere with respect to $\mu + \nu$. In particular, it follows that for $E \in \mathcal{B}$ with $\mu(E) < \infty$ we can in (6-14) choose

$$g = \frac{1-h^k}{1-h}\mathcal{X}_E = \Big(\sum_{i=0}^{k-1} h^i\Big)\mathcal{X}_E \in L^\infty(\mu+\nu)\,,$$

which yields that

$$\int_E (1-h^k)\,\mathrm{d}\nu = \int_E \frac{h}{1-h}(1-h^k)\,\mathrm{d}\mu\,.$$

On noting that $\mu+\nu$-almost everywhere $0 \leq (1-h^k)\mathcal{X}_E \nearrow \mathcal{X}_E \in L^1(\mu+\nu)$ as $k \nearrow \infty$, we conclude from the monotone convergence theorem that $\frac{h}{1-h}\mathcal{X}_E \in L^1(\mu)$ and

$$\nu(E) = \int_E \frac{h}{1-h}\,\mathrm{d}\mu\,,$$

i.e. $\frac{h}{1-h}$ is the desired function. The fact that $\frac{h}{1-h} \in L^1(\mu)$ follows again from the monotone convergence theorem, upon setting $E = \bigcup_{j \leq m} S_j$, taking the limit $m \to \infty$, and recalling that $\nu(S) < \infty$. (A purely measure theoretical proof of the Radon-Nikodým theorem can be found in e.g. [Halmos].) □

6.12 Theorem (Dual space of L^p for $p < \infty$). Let $(S, \mathcal{B}, \mu)$ be a measure space and let $1 \leq p < \infty$ (the ***dual exponent*** p' is given by $\frac{1}{p} + \frac{1}{p'} = 1$, if $p = 1$ then $p' = \infty$). In the case $p = 1$, we assume in addition that μ is σ-finite. For $f \in L^{p'}(\mu)$ let

$$J(f)(g) := \int_S g\overline{f}\,\mathrm{d}\mu \qquad \text{for all } g \in L^p(\mu)\,.$$

Then $J : L^{p'}(\mu) \to L^p(\mu)'$ is a conjugate linear isometric isomorphism.
Special case: In the Hilbert space case $p = 2 = p'$, the isometry J coincides with the isometry in 6.1.

Proof. It follows from the Hölder inequality that J is well defined and that $\|J(f)\|_{(L^p)'} \leq \|f\|_{L^{p'}}$. Clearly, J is conjugate linear. Moreover, J is injective, since $J(f) = 0$ implies in the case $p > 1$ with $g := |f|^{p'-2}f \in L^p(\mu)$ that

$$0 = J(f)(g) = \int_S |f|^{p'}\,\mathrm{d}\mu\,,$$

and so that $f = 0$ in $L^{p'}(\mu)$. In the case $p = 1$ set $g = \mathcal{X}_{S_m} f \in L^1(\mu)$ with S_m as in 3.9(4) and obtain that $f = 0$ almost everywhere in S_m. Letting $m \to \infty$ we conclude that $f = 0$ in $L^\infty(\mu)$.

Now let $F \in L^p(\mu)'$. We need to show that there exists an $f \in L^{p'}(\mu)$ with

$$F = J(f) \quad \text{and} \quad \|f\|_{L^{p'}} \le \|F\|_{(L^p)'} .$$

First we consider the special case $\mu(S) < \infty$. Then

$$\nu(E) := F(\mathcal{X}_E) \quad \text{for } E \in \mathcal{B}$$

satisfies the assumptions of the Radon-Nikodým theorem. To see this, note that for disjoint sets $E_1, \dots, E_m$ in $\mathcal{B}$ with $\nu(E_i) \ne 0$ it holds that

$$\begin{aligned}\sum_{i=1}^m |\nu(E_i)| &= \sum_{i=1}^m \sigma_i \nu(E_i) \quad \text{with } \sigma_i := \frac{\overline{\nu(E_i)}}{|\nu(E_i)|} \\ &= F\Big(\sum_{i=1}^m \sigma_i \mathcal{X}_{E_i}\Big) \le \|F\|_{(L^p)'} \cdot \Big\|\sum_{i=1}^m \sigma_i \mathcal{X}_{E_i}\Big\|_{L^p} \qquad (6\text{-}15) \\ &= \|F\|_{(L^p)'} \cdot \Big(\sum_{i=1}^m \mu(E_i)\Big)^{\frac{1}{p}} \le \|F\|_{(L^p)'} \cdot \mu(S)^{\frac{1}{p}} ,\end{aligned}$$

i.e. $\|\nu\|_{\text{var}} < \infty$. In addition, for $E = \bigcup_{i \in \mathbb{N}} E_i$ with $E_i \in \mathcal{B}$, $E_i \subset E_{i+1}$

$$|\nu(E) - \nu(E_i)| = \big|F(\mathcal{X}_{E \setminus E_i})\big| \le \|F\|_{(L^p)'} \mu(E \setminus E_i)^{\frac{1}{p}} \to 0 \quad \text{as } i \to \infty,$$

i.e. ν is σ-additive. By the way, ν is absolutely continuous w.r.t. μ, since for μ-null sets E we have $\mathcal{X}_E = 0$ in $L^p(\mu)$, and therefore $\nu(E) = F(\mathcal{X}_E) = 0$.

Hence, by the Radon-Nikodým theorem 6.11, there exists a function $f \in L^1(\mu)$ with

$$F(\mathcal{X}_E) = \int_S \mathcal{X}_E \overline{f} \, d\mu \quad \text{for all } E \in \mathcal{B}.$$

It follows that

$$F(g) = \int_S g \overline{f} \, d\mu \qquad (6\text{-}16)$$

for all functions $g \in L^\infty(\mu)$, because such functions can be uniformly approximated by finite linear combinations of characteristic functions $\mathcal{X}_E$ with measurable $E \subset S$ (see the note in 3.26(1)). Now for $m \in \mathbb{N}$ and $1 \le q < \infty$ we choose in particular

$$g = \mathcal{X}_{A_m} |f|^{q-2} f , \quad \text{where } A_m := \{x \in S ;\ 0 < |f(x)| \le m\},$$

and obtain from (6-16) that

$$\int_{A_m} |f|^q \, d\mu = F(g) \le \|F\|_{(L^p)'} \|g\|_{L^p} = \|F\|_{(L^p)'} \Big(\int_{A_m} |f|^{p(q-1)} \, d\mu\Big)^{\frac{1}{p}} .$$

In the case $p > 1$, setting $q = p'$ (so that $p(q-1) = p'$), yields after cancellation that

$$\Big(\int_{A_m} |f|^{p'} \, d\mu\Big)^{\frac{1}{p'}} \le \|F\|_{(L^p)'} .$$

On letting $m \to \infty$, it follows from the monotone convergence theorem that $f \in L^{p'}(\mu)$ and $\|f\|_{L^{p'}} \le \|F\|_{(L^p)'}$. In the case $p = 1$, choose $q \in \mathbb{N}$ and obtain inductively that

$$\int_{A_m} |f|^q \,\mathrm{d}\mu \le \|F\|_{(L^p)'} \int_{A_m} |f|^{q-1} \,\mathrm{d}\mu \le \|F\|^q_{(L^p)'} \cdot \mu(A_m)\,,$$

i.e.

$$\Big(\int_{A_m} |f|^q \,\mathrm{d}\mu\Big)^{\frac{1}{q}} \le \|F\|_{(L^p)'} \cdot \mu(A_m)^{\frac{1}{q}}\,.$$

Then, on letting $q \to \infty$, it follows from E3.4 (for the function $\mathcal{X}_{A_m} f$) that $|f| \le \|F\|_{(L^p)'}$ almost everywhere in A_m, which implies that $\|f\|_{L^\infty} \le \|F\|_{(L^p)'}$.

On noting that the functions g, for which (6-16) originally held, are dense in $L^p(\mu)$, it now follows from the Hölder inequality that (6-16) holds for all $g \in L^p(\mu)$, and so $F = J(f)$, which is what we wanted to show.

We now consider the case of a general measure space, and define $\widetilde{\mathcal{B}} := \{A \in \mathcal{B}\,;\ \mu(A) < \infty\}$. For $A \in \widetilde{\mathcal{B}}$ let

$$\mu_A(E) := \mu(A \cap E), \quad F_A(g) := F(\mathcal{X}_A g)\,.$$

Then $\mu_A(S) < \infty$ with $\mu_A(S \setminus A) = 0$, and $F_A \in L^p(\mu_A)'$, with $\|F_A\|_{(L^p)'} \le \|F\|_{(L^p)'}$. Hence it follows from what we have shown so far that there exists a unique $f_A \in L^{p'}(\mu_A)$ with

$$F_A(g) = \int_S g\overline{f_A} \,\mathrm{d}\mu_A \quad \text{for all } g \in L^p(\mu_A) \tag{6-17}$$

and $\|f_A\|_{L^{p'}} = \|F_A\|_{(L^p)'}$. On defining $f_A(x) := 0$ for $x \in S \setminus A$, we have that $f_A \in L^{p'}(\mu)$. As in the proof of the injectivity of J, it follows that $f_{A_1} = f_{A_2}$ μ-almost everywhere in $A_1 \cap A_2$ for $A_1, A_2 \in \widetilde{\mathcal{B}}$. Hence, $|f_{A_1}| \le |f_{A_2}|$ μ-almost everywhere if $A_1 \subset A_2$, and then

$$\|f_{A_1}\|_{L^{p'}} \le \|f_{A_2}\|_{L^{p'}} = \|F_{A_2}\|_{(L^p)'} \le \|F\|_{(L^p)'} < \infty\,.$$

It follows that there exist $B_m \in \widetilde{\mathcal{B}}$ with $B_m \subset B_{m+1}$ for $m \in \mathbb{N}$, such that

$$\|f_{B_m}\|_{L^{p'}} \longrightarrow s := \sup_{A \in \widetilde{\mathcal{B}}} \|f_A\|_{L^{p'}} \quad \text{as } m \to \infty.$$

If $p = 1$, then the B_m can be chosen such that $S_m \subset B_m$, where the S_m are as in 3.9(4). Then

$$B := \bigcup_{m \in \mathbb{N}} B_m\,, \quad f(x) := \begin{cases} f_{B_m}(x) & \text{for } x \in B_m,\ m \in \mathbb{N}, \\ 0 & \text{for } x \in S \setminus B, \end{cases}$$

(for $p > 1$ by the monotone convergence theorem) defines an $f \in L^{p'}(\mu)$ with

$$\|f\|_{L^{p'}} = s = \sup_{A \in \widetilde{\mathcal{B}}} \|f_A\|_{L^{p'}} = \sup_{A \in \widetilde{\mathcal{B}}} \|F_A\|_{(L^p)'} \leq \|F\|_{(L^p)'} \,.$$

Now

$$f_A = f \text{ almost everywhere in } A, \text{ if } A \in \mathcal{B} \text{ with } A \subset B,$$

since in $A \cap B_m$ it holds almost everywhere that $f = f_{B_m} = f_{A \cap B_m} = f_A$. We claim that

$$f_A = 0 \text{ almost everywhere in } S, \text{ if } A \in \mathcal{B} \text{ with } A \cap B = \emptyset.$$

In the case $p = 1$, this trivially follows from $B = S$. In the case $p > 1$, on noting that $A \cap B_m = \emptyset$, it follows that

$$|f_{A \cup B_m}|^{p'} = |f_A|^{p'} + |f_{B_m}|^{p'}, \quad \text{and so} \quad s^{p'} \geq \|f_A\|_{L^{p'}}^{p'} + \|f_{B_m}\|_{L^{p'}}^{p'} \,.$$

Letting $m \to \infty$ yields that $s^{p'} \geq \|f_A\|_{L^{p'}}^{p'} + s^{p'}$, and hence our claim.

Now let $g \in L^p(\mu)$ with $g = 0$ almost everywhere in $S \setminus A$ for an $A \in \mathcal{B}$. Then, by (6-17),

$$F(g) = F_A(g) = \int_S g\overline{f_A} \, \mathrm{d}\mu_A = \int_A g\overline{f_A} \, \mathrm{d}\mu \,.$$

Since, as shown above, $f_A = f_{A \setminus B} = 0$ in $A \setminus B$ and $f_A = f_{A \cap B} = f$ in $A \cap B$, this is in turn equal to

$$\int_{A \cap B} g\overline{f} \, \mathrm{d}\mu = \int_A g\overline{f} \, \mathrm{d}\mu = \int_S g\overline{f} \, \mathrm{d}\mu = J(f)(g) \,.$$

On noting that such functions g are dense in $L^p(\mu)$ (approximating g, for example, by $\mathcal{X}_{A_n} g$, $n \in \mathbb{N}$, with $A_n := \{x \in S\,;\ |g(x)| \geq \frac{1}{n}\}$), it follows that $F(g) = J(f)(g)$ for all $g \in L^p(\mu)$. □

With the help of the result in theorem 6.12, we can establish a distributional characterization of L^p-functions:

6.13 Corollary. Let $\Omega \subset \mathbb{R}^n$ be open and let $1 \leq p \leq \infty$. Then it holds for functions $f : \Omega \to \mathbb{K}$ that

$$f \in L^p(\Omega) \quad \Longleftrightarrow \quad \begin{cases} f \in L^1_{\mathrm{loc}}(\Omega) \text{ and there exists a } C \text{ with} \\ \left| \int_\Omega \zeta f \, \mathrm{dL}^n \right| \leq C \|\zeta\|_{L^{p'}(\Omega)} \text{ for all } \zeta \in C_0^\infty(\Omega). \end{cases}$$

The constant C on the right-hand side satisfies $\|f\|_{L^p(\Omega)} \leq C$.

Notation: Here $L^1_{\mathrm{loc}}(\Omega)$ is the space of locally integrable functions in Ω, defined in 5.13(2). Moreover, $1 \leq p' \leq \infty$ is the ***dual exponent***, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$.

Note: For a generalization of the result to Sobolev functions, see E6.7.

Proof $\Rightarrow$. The Hölder inequality yields that

$$\left|\int_\Omega \zeta f \,\mathrm{dL}^n\right| \le \|\zeta\|_{L^{p'}(\Omega)} \cdot \|f\|_{L^p(\Omega)}\,.$$

□

Proof $\Leftarrow$. The estimate yields that on $C_0^\infty(\Omega)$ equipped with the $L^{p'}$-norm,

$$F(\zeta) := \int_\Omega \zeta f \,\mathrm{dL}^n$$

is linear and continuous. In the case $p > 1$, we have that $C_0^\infty(\Omega)$ is dense in $L^{p'}(\Omega)$ (this follows from 4.15(3) as $p' < \infty$), and so F can be uniquely extended to $L^{p'}(\Omega)$, as a functional $F \in L^{p'}(\Omega)'$ (see E5.3). Hence it follows from 6.12 that there exists an $\widetilde{f} \in L^p(\Omega)$ with

$$F(g) = \int_\Omega g\widetilde{f} \,\mathrm{dL}^n \quad \text{for all } g \in L^{p'}(\Omega).$$

Since

$$\int_\Omega \zeta f \,\mathrm{dL}^n = \int_\Omega \zeta\widetilde{f} \,\mathrm{dL}^n \quad \text{for all } \zeta \in C_0^\infty(\Omega),$$

$f = \widetilde{f}$ almost everywhere in Ω (see 4.22). In the case $p = 1$, set

$$g(x) := \begin{cases} \dfrac{\overline{f(x)}}{|f(x)|}, & \text{if } f(x) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let $D \subset\subset \Omega$ and let $(\varphi_\varepsilon)_{\varepsilon>0}$ be a standard Dirac sequence. Then $\zeta_\varepsilon := \varphi_\varepsilon * \big(\mathcal{X}_D g\big) \in C_0^\infty(\Omega)$ for sufficiently small $\varepsilon > 0$, and

$$\left|\int_\Omega \zeta_\varepsilon f \,\mathrm{dL}^n\right| \le C\|\zeta_\varepsilon\|_{L^\infty} \le C\,.$$

Letting $\varepsilon \to 0$, we obtain from Lebesgue's convergence theorem (as $\zeta_\varepsilon \to \mathcal{X}_D g$ almost everywhere for a subsequence $\varepsilon \to 0$) that

$$\int_D |f| \,\mathrm{dL}^n = \left|\int_D g f \,\mathrm{dL}^n\right| \le C\,,$$

where the constant C is independent of D. Hence $f \in L^1(\Omega)$. □

Hahn-Banach's theorem

For the characterization of $C^0(S)'$ we will use the fact that functionals on $C^0(S)$ can be extended norm-preservingly to $B(S)$ (see the proof 6.23). The existence of such extensions in more general situations is guaranteed by the following two theorems.

6.14 Hahn-Banach theorem. Let X be an $\mathbb{R}$-vector space and let the following hold:

(1) $p : X \to \mathbb{R}$ is ***sublinear***, i.e. for all $x, y \in X$ and $\alpha \in \mathbb{R}$,

$$p(x+y) \leq p(x) + p(y) \quad \text{and} \quad p(\alpha x) = \alpha p(x) \text{ for } \alpha \geq 0.$$

(2) $f : Y \to \mathbb{R}$ is linear with a subspace $Y \subset X$.

(3) $f(x) \leq p(x)$ for $x \in Y$.

Then there exists a linear map $F : X \to \mathbb{R}$ such that

$$F(x) = f(x) \text{ for } x \in Y \quad \text{and} \quad F(x) \leq p(x) \text{ for } x \in X.$$

Proof. We consider the class of all extensions of f, that is,

$$\begin{aligned} \mathcal{M} := \Big\{ (Z, g) \; ; \; & Z \text{ subspace}, Y \subset Z \subset X, \\ & g : Z \to \mathbb{R} \text{ linear}, g = f \text{ on } Y, g \leq p \text{ on } Z \Big\} . \end{aligned}$$

Consider an arbitrary $(Z, g) \in \mathcal{M}$ with $Z \neq X$ and a $z_0 \in X \setminus Z$. We want to extend g at least to

$$Z_0 := \operatorname{span}(Z \cup \{z_0\}) = Z \oplus \operatorname{span}\{z_0\} .$$

We attempt the ansatz

$$g_0(z + \alpha z_0) := g(z) + c\alpha \quad \text{for } z \in Z \text{ and } \alpha \in \mathbb{R}.$$

Here c still needs to be suitably chosen, so that $(Z_0, g_0) \in \mathcal{M}$. Clearly, g_0 is linear on Z_0. Moreover, $g_0 = g = f$ on Y. It remains to show that

$$g(z) + c\alpha \leq p(z + \alpha z_0) \quad \text{for } z \in Z \text{ and } \alpha \in \mathbb{R}.$$

Since $g \leq p$ on Z, this is satisfied for $\alpha = 0$. For $\alpha > 0$ the inequality is equivalent to

$$c \leq \tfrac{1}{\alpha}\big(p(z + \alpha z_0) - g(z)\big) = p\big(\tfrac{z}{\alpha} + z_0\big) - g\big(\tfrac{z}{\alpha}\big)$$

and for $\alpha < 0$ to

$$c \geq \tfrac{1}{\alpha}\big(p(z + \alpha z_0) - g(z)\big) = g\big(-\tfrac{z}{\alpha}\big) - p\big(-\tfrac{z}{\alpha} - z_0\big) .$$

Hence we need to find a number c such that

$$\sup_{z \in Z} \left(g(z) - p(z - z_0)\right) \leq c \leq \inf_{z \in Z} \left(p(z + z_0) - g(z)\right) .$$

This is possible, because for $z, z' \in Z$ we have

$$\begin{aligned} g(z') + g(z) &= g(z' + z) \leq p(z' + z) \\ &= p(z' - z_0 + z + z_0) \leq p(z' - z_0) + p(z + z_0) , \end{aligned}$$

and hence

$$g(z') - p(z' - z_0) \le p(z + z_0) - g(z)\,.$$

We now hope that this extension procedure yields an $(X, F) \in \mathcal{M}$. To this end, we make use of

Zorn's lemma: Let $(\mathcal{M}, \le)$ be a nonempty ***partially ordered set*** (i.e. if $m_1 \le m_2$ and $m_2 \le m_3$, then $m_1 \le m_3$, and $m \le m$ for all $m \in \mathcal{M}$) such that every totally ordered subset $\mathcal{N}$ (i.e. for all $n_1, n_2 \in \mathcal{N}$ it holds that $n_1 \le n_2$ or $n_2 \le n_1$) has an upper bound (i.e. there exists an $m \in \mathcal{M}$ with $n \le m$ for all $n \in \mathcal{N}$). Then $\mathcal{M}$ contains a ***maximal element*** (i.e. there exists an $m_0 \in \mathcal{M}$ such that for all $m \in \mathcal{M}$ it holds that $m_0 \le m \Longrightarrow m \le m_0$).

In our case, an order is defined by

$$(Z_1, g_1) \le (Z_2, g_2) \quad :\Longleftrightarrow \quad Z_1 \subset Z_2 \text{ and } g_2 = g_1 \text{ on } Z_1.$$

We need to verify the assumptions of Zorn's lemma. Let $\mathcal{N} \subset \mathcal{M}$ be totally ordered and define

$$\begin{aligned} Z_* \quad &:= \bigcup_{(Z,g)\in\mathcal{N}} Z\,, \\ g_*(x) &:= g(x)\,, \quad \text{if } x \in Z \text{ and } (Z, g) \in \mathcal{N}. \end{aligned}$$

We need to show that $(Z_*, g_*) \in \mathcal{M}$. Now $Y \subset Z_* \subset X$, and g_* is a well defined function, because

$$\begin{aligned} & x \in Z_1 \cap Z_2\ ,\ (Z_1, g_1) \in \mathcal{N}\ ,\ (Z_2, g_2) \in \mathcal{N} \\ \Longrightarrow\ & (Z_1, g_1) \le (Z_2, g_2) \text{ or } (Z_2, g_2) \le (Z_1, g_1) \quad (\text{total order of } \mathcal{N}) \\ \Longrightarrow\ & Z_1 \subset Z_2 \text{ and } g_2 = g_1 \text{ on } Z_1 \quad (\text{in the first case}) \\ \Longrightarrow\ & g_2(x) = g_1(x) \quad (\text{as } x \in Z_1)\,. \end{aligned}$$

The properties $g_* = f$ on Y and $g_* \le p$ on Z_* carry over. The linearity of Z_* and g_* can be seen as follows:

$$\begin{aligned} & x, y \in Z_*\ ,\ \alpha \in \mathbb{R} \\ \Longrightarrow\ & \text{There exist } (Z_x, g_x) \in \mathcal{N}, (Z_y, g_y) \in \mathcal{N} \text{ with } x \in Z_x \text{ and } y \in Z_y \\ \Longrightarrow\ & (Z_x, g_x) \le (Z_y, g_y) \text{ or } (Z_y, g_y) \le (Z_x, g_x) \\ \Longrightarrow\ & x, y \in Z_\xi \text{ with } \xi = y \text{ in the first and } \xi = x \text{ in the second case,} \\ & \text{hence also } x + \alpha y \in Z_\xi \subset Z_* \text{ and} \\ & g_*(x + \alpha y) = g_\xi(x + \alpha y) = g_\xi(x) + \alpha g_\xi(y) = g_*(x) + \alpha g_*(y)\,. \end{aligned}$$

Hence it follows from Zorn's lemma that $\mathcal{M}$ has a maximal element (Z, g). If we assume that $Z \ne X$, then the extension procedure from the beginning of the proof yields a $(Z_0, g_0) \in \mathcal{M}$ with

$$(Z, g) \le (Z_0, g_0) \quad \text{and} \quad Z_0 \ne Z\,,$$

which contradicts the maximality of (Z, g). □

The Hahn-Banach theorem has the following version for linear functionals.

6.15 Hahn-Banach theorem (for linear functionals). Let X be a normed $\mathbb{K}$-vector space and Y be a subspace (with the norm of X !). Then for $y' \in Y'$ there exists an $x' \in X'$ with

$$x' = y' \text{ on } Y \quad \text{and} \quad \|x'\|_{X'} = \|y'\|_{Y'} .$$

Proof for $\mathbb{K} = \mathbb{R}$. Choose

$$p(x) := \|y'\|_{Y'} \|x\|_X \quad \text{for } x \in X$$

in 6.14, so that for $y \in Y$

$$y'(y) \le \|y'\|_{Y'} \|y\|_Y = \|y'\|_{Y'} \|y\|_X = p(y) .$$

Then, by 6.14, there exists a linear map $x' : X \to \mathbb{R}$ with

$$x' = y' \text{ on } Y \quad \text{and} \quad x' \le p \text{ on } X.$$

The second property implies that

$$\pm x'(x) = x'(\pm x) \le p(\pm x) = \|y'\|_{Y'} \|x\|_X ,$$

i.e. $x' \in X'$ with $\|x'\|_{X'} \le \|y'\|_{Y'}$, and the first property implies that

$$\|y'\|_{Y'} = \sup_{\substack{y \in Y \\ \|y\|_X \le 1}} |y'(y)| = \sup_{\substack{y \in Y \\ \|y\|_X \le 1}} |x'(y)| \le \|x'\|_{X'} .$$

□

Proof for $\mathbb{K} = \mathbb{C}$. Consider X and Y as normed $\mathbb{R}$-vector spaces $X_{\mathbb{R}}$ and $Y_{\mathbb{R}}$ (i.e. scalar multiplication is defined only for real numbers, but the norms remain the same). Let $X'_{\mathbb{R}}$ and $Y'_{\mathbb{R}}$ be the corresponding dual spaces. For $y' \in Y'$ it then holds that

$$y'_{\text{re}} := \text{Re}y' \in Y'_{\mathbb{R}} \quad \text{with} \quad \|y'_{\text{re}}\|_{Y'_{\mathbb{R}}} \le \|y'\|_{Y'}$$

and

$$y'(x) = \text{Re}y'(x) + \text{iIm}y'(x) = y'_{\text{re}}(x) - \text{i}y'_{\text{re}}(\text{i}x) .$$

It follows from the real case treated above that there exists an extension x'_{re} of y'_{re} to $X_{\mathbb{R}}$ with $\|x'_{\text{re}}\|_{X'_{\mathbb{R}}} = \|y'_{\text{re}}\|_{Y'_{\mathbb{R}}}$. Define

$$x'(x) := x'_{\text{re}}(x) - \text{i}x'_{\text{re}}(\text{i}x) .$$

Then $x' = y'$ on Y, and $x' : X \to \mathbb{C}$ is $\mathbb{C}$-linear, because x' is $\mathbb{R}$-linear and for $x \in X$ we have that

$$\begin{aligned} x'(\mathrm{i}x) &= x'_{\mathrm{re}}(\mathrm{i}x) - \mathrm{i}x'_{\mathrm{re}}(-x) = x'_{\mathrm{re}}(\mathrm{i}x) + \mathrm{i}x'_{\mathrm{re}}(x) \\ &= \mathrm{i}\,(-\mathrm{i}x'_{\mathrm{re}}(\mathrm{i}x) + x'_{\mathrm{re}}(x)) = \mathrm{i}x'(x)\,. \end{aligned}$$

Now let $x \in X$. Then $x'(x) \in \mathbb{C}$ can be written as $x'(x) = r\mathrm{e}^{\mathrm{i}\theta}$ with $\theta \in \mathbb{R}$ and $r \geq 0$. Therefore,

$$\begin{aligned} |x'(x)| &= r = \mathrm{Re}\big(\mathrm{e}^{-\mathrm{i}\theta}x'(x)\big) = \mathrm{Re}x'(\mathrm{e}^{-\mathrm{i}\theta}x) \\ &= x'_{\mathrm{re}}(\mathrm{e}^{-\mathrm{i}\theta}x) \leq \|x'_{\mathrm{re}}\|_{X'_{\mathbb{R}}}\|x\|_{X'}\,, \end{aligned}$$

and we recall that $\|x'_{\mathrm{re}}\|_{X'_{\mathbb{R}}} = \|y'_{\mathrm{re}}\|_{Y'_{\mathbb{R}}} \leq \|y'\|_{Y'}$. This shows that $x' \in X'$ with $\|x'\|_{X'} \leq \|y'\|_{Y'}$. As x' is an extension of y', it must also hold that $\|x'\|_{X'} \geq \|y'\|_{Y'}$. □

As an application, we show that points in a normed space can be separated from subspaces with the help of linear functionals (see the generalization of suspaces to closed convex sets in 8.12). This separation property is often used in order to show that a given subspace is dense in the ambient space X.

6.16 Theorem. Let Y be a closed subspace of the normed space X and let $x_0 \notin Y$. Then there exists an $x' \in X'$ with

$$x' = 0 \text{ on } Y\,, \quad \|x'\|_{X'} = 1\,, \quad x'(x_0) = \mathrm{dist}(x_0, Y)\,.$$

Remark: Then there also exists an $x' \in X'$ with

$$x' = 0 \text{ on } Y\,, \quad \|x'\|_{X'} = \frac{1}{\mathrm{dist}(x_0, Y)}\,, \quad x'(x_0) = 1\,.$$

Proof. On

$$Y_0 := \mathrm{span}\,(Y \cup \{x_0\}) = Y \oplus \mathrm{span}\{x_0\}$$

define

$$y'_0(y + \alpha x_0) := \alpha \cdot \mathrm{dist}(x_0, Y) \quad \text{for } y \in Y \text{ and } \alpha \in \mathbb{K}.$$

Then $y'_0 : Y_0 \to \mathbb{K}$ is linear and $y'_0 = 0$ on Y. We want to show that $y'_0 \in Y'_0$ with $\|y'_0\|_{Y'_0} = 1$, as 6.15 then yields the desired result.

Let $y \in Y$ and $\alpha \neq 0$. Then

$$\mathrm{dist}(x_0, Y) \leq \left\| x_0 - \frac{-y}{\alpha} \right\|_X,$$

and so

$$|y'_0(y + \alpha x_0)| \leq |\alpha| \left\| x_0 - \frac{-y}{\alpha} \right\|_X = \|\alpha x_0 + y\|_X\,,$$

and hence $y_0 \in Y'_0$ with $\|y'_0\|_{Y'_0} \leq 1$. The closedness of Y yields that $\mathrm{dist}(x_0, Y) > 0$, and so for $\varepsilon > 0$ we can choose a $y_\varepsilon \in Y$ such that

$$\|x_0 - y_\varepsilon\|_X \le (1+\varepsilon)\mathrm{dist}(x_0, Y)\,.$$

Then

$$y_0'(x_0 - y_\varepsilon) = \mathrm{dist}(x_0, Y) \ge \tfrac{1}{1+\varepsilon}\|x_0 - y_\varepsilon\|_X\,,$$

which, since $x_0 - y_\varepsilon \neq 0$, implies that $\|y_0'\|_{Y_0'} \ge \frac{1}{1+\varepsilon} \to 1$ as $\varepsilon \searrow 0$. ☐

6.17 Corollaries. Let X be a normed space and let $x_0 \in X$. Then:

(1) If $x_0 \neq 0$, then there exists an $x_0' \in X'$ with

$$\|x_0'\|_{X'} = 1 \quad \text{and} \quad x_0'(x_0) = \|x_0\|_X\,.$$

(2) If $x'(x_0) = 0$ for all $x' \in X'$, then $x_0 = 0$.

(3) Setting $Tx' := x'(x_0)$ for $x' \in X'$ defines an element T of $\mathscr{L}(X'; \mathbb{K}) = (X')'$, the bidual space (see 8.2), with $\|T\| = \|x_0\|_X$.

Proof. (1) is the result in 6.16 with $Y = \{0\}$, and (2) follows from (1). In (3) we have that $|Tx'| \le \|x'\|_{X'}\|x_0\|_X$, and if $x_0 \neq 0$ it holds that $|Tx_0'| = \|x_0\|_X$ with x_0' as in (1). Hence $\|T\| = \|x_0\|_X$. ☐

6.18 Remark. The result 6.16 may also be interpreted as a generalization of the projection theorem for Hilbert spaces in the linear case. To see this, assume that X is a Hilbert space and define

$$x'(x) := \left(x\,,\ \frac{x_0 - Px_0}{\|x_0 - Px_0\|}\right)_X\,,$$

where P is the orthogonal projection onto Y from 4.3. It follows from 4.4(2) that $x' = 0$ on Y and hence

$$x'(x_0) = x'(x_0 - Px_0) = \|x_0 - Px_0\|_X\,,$$

and moreover $|x'(x)| \le \|x\|_X$. Hence x' has all the properties in 6.16.

Riesz-Radon's theorem

As we have seen in 6.12 the dual space of the function space $L^p(\mu)$, if $1 \le p < \infty$, is isomorphic to a space that is again a function space. We will now show that the dual space of $C^0(S)$ is isomorphic to a space of measures. To this end, we need the following definitions (the notations are the same as in [DunfordSchwartz : IV 2]).

6.19 Definition (Borel sets). Let X be a topological space. The set of ***Borel sets*** is defined as the smallest σ-algebra that contains the closed subsets of X (or, equivalently, the open subsets of X).

6.20 Spaces of additive measures. Let $S \subset \mathbb{R}^n$ be equipped with the relative topology of $\mathbb{R}^n$ (see 2.11). Let $\mathcal{B}_0$ be the smallest Boolean algebra that contains the closed (or, equivalently, open) subsets of S, and let $\mathcal{B}_1$ be the set of Borel sets of S, i.e. the smallest σ-algebra containing $\mathcal{B}_0$. Then

$$ba(S;\mathbb{K}^m) := \{\lambda : \mathcal{B}_0 \to \mathbb{K}^m\,;\ \lambda \text{ is additive and } \|\lambda\|_{\mathrm{var}} < \infty\}\,,$$
$$ca(S;\mathbb{K}^m) := \{\lambda : \mathcal{B}_1 \to \mathbb{K}^m\,;\ \lambda \text{ is } \sigma\text{-additive and } \|\lambda\|_{\mathrm{var}} < \infty\}$$

are $\mathbb{K}$-vector spaces and, equipped with the total variation as the norm, also Banach spaces. In the definition, ba stands for "bounded additive" and ca stands for "countably additive". As usual, we set $ba(S) := ba(S;\mathbb{K})$ and $ca(S) := ca(S;\mathbb{K})$.

Proof. We prove the completeness. Let $(\lambda_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $ba(S;\mathbb{K}^m)$. Then it holds for $E \in \mathcal{B}_0$ that

$$|\lambda_l(E) - \lambda_k(E)| \le \|\lambda_l - \lambda_k\|_{\mathrm{var}} \to 0 \quad \text{as } k,l \to \infty,$$

and so there exists

$$\lambda(E) := \lim_{l\to\infty} \lambda_l(E) \quad \text{for } E \in \mathcal{B}_0$$

and the additivity carries over to λ. In addition,

$$\|\lambda - \lambda_k\|_{\mathrm{var}} \le \liminf_{l\to\infty} \|\lambda_l - \lambda_k\|_{\mathrm{var}} \longrightarrow 0 \quad \text{as } k \to \infty.$$

Analogously, for Cauchy sequences in $ca(S;\mathbb{K}^m)$ there exists a limit λ on $\mathcal{B}_1$. If $E_i \in \mathcal{B}_1$ with $E_i \supset E_{i+1}$ and $\bigcap_{i\in\mathbb{N}} E_i = \emptyset$, then for $l \ge k$ and as $l \to \infty$

$$|\lambda(E_i)| \longleftarrow |\lambda_l(E_i)| \le \underbrace{|\lambda_k(E_i)|}_{\substack{\to 0 \text{ as } i\to\infty \\ \text{for every } k}} + \underbrace{\|\lambda_l - \lambda_k\|_{\mathrm{var}}}_{\to 0 \text{ as } l \ge k \to \infty}\,,$$

i.e. λ is σ-additive. ☐

6.21 Spaces of regular measures. Let $S \subset \mathbb{R}^n$, $\mathcal{B}_0$, and $\mathcal{B}_1$ be as in 6.20. A measure λ in $ba(S;\mathbb{K}^m)$ or $ca(S;\mathbb{K}^m)$ is called ***regular*** if for all $E \in \mathcal{B}_0$ or $E \in \mathcal{B}_1$, respectively,

$$\inf\Big\{\, |\lambda|(U \setminus K)\ ;\ K \subset E \subset U,\ K \text{ is closed in } S \text{ and } U \text{ is open in } S \,\Big\} = 0\,.$$

Here $|\lambda|$ is the variational measure from 6.10 and in S we consider the relative topology from 2.11, i.e. a set $U \subset S$ is called open in S if it is of the form $U = S \cap V$ for an open set $V \subset \mathbb{R}^n$, and a set $K \subset S$ is called closed in S if $S \setminus K$ is open in S. We define

$$rba(S;\mathbb{K}^m) := \{\lambda \in ba(S;\mathbb{K}^m)\,;\ \lambda \text{ is regular}\}\,,$$
$$rca(S;\mathbb{K}^m) := \{\lambda \in ca(S;\mathbb{K}^m)\,;\ \lambda \text{ is regular}\}\,.$$

These sets are $\mathbb{K}$-vector spaces and, equipped with the total variation as the norm, also Banach spaces. In the definition, rba stands for "regular bounded additive" and rca stands for "regular countably additive". As usual, we set $rba(S) := rba(S;\mathbb{K})$ and $rca(S) := rca(S;\mathbb{K})$.

Proof. For the completeness we need to show that for regular measures λ_k it follows from $\lambda_k \to \lambda$ in $ba(S;\mathbb{K}^m)$ as $k \to \infty$ that λ is also regular. To prove this we note that for $K \subset E \subset U$, as in the definition of regularity,

$$|\lambda|(U \setminus K) \le |\lambda_k|(U \setminus K) + \|\lambda - \lambda_k\|_{\mathrm{var}}\,.$$

The first term on the right-hand side can be made arbitrarily small for every k, by choosing U and K appropriately. □

In the following we need the fact that for regular measures $\mu : \mathcal{B}_1 \to [0,\infty]$, continuous functions are integrable, i.e. that they lie in $L^1(\mu)$. The proof of this result is the construction of the Riemann integral, which for our purposes we give here for vector-valued measures $\lambda : \mathcal{B}_0 \to \mathbb{K}^m$.

6.22 Integral of continuous functions (Riemann integral). Let $\mathcal{B}_0$ be as in 6.20. In addition, assume that $\lambda : \mathcal{B}_0 \to \mathbb{K}^m$ is additive with $\|\lambda\|_{\mathrm{var}} < \infty$. For step functions

$$f = \sum_{i=1}^{k} \mathcal{X}_{E_i}\alpha_i\,, \quad k \in \mathbb{N}\,,\ \alpha_i \in \mathbb{K}\,,\ E_i \in \mathcal{B}_0\,,$$

it holds that

$$\int_S f\,\mathrm{d}\lambda := \sum_{i=1}^{k} \alpha_i \lambda(E_i)$$

is independent of the representation of f. Moreover, we have that (choose E_i in the representation of f disjoint)

$$\left|\int_S f\,\mathrm{d}\lambda\right| \le \|f\|_{\sup} \cdot \|\lambda\|_{\mathrm{var}}\,.$$

Every continuous and bounded function $f : S \to \mathbb{K}$ can be approximated by such step functions in the supremum norm. To see this, cover the bounded set $\overline{f(S)}$ with open sets U_i, $i = 1,\dots,l$, with diameter $\le \frac{1}{k}$. Then one can construct another cover by (cf. the proof of A3.19(2))

$$V_i := U_i \setminus \bigcup_{j<i} U_j \quad \text{for } i = 1,\dots,l,$$

where now the sets V_i are pairwise disjoint. In addition,

$$E_i := f^{-1}(V_i) = f^{-1}(U_i) \setminus \bigcup_{j<i} f^{-1}(U_j) \in \mathcal{B}_0 \,.$$

On choosing $\alpha_i \in V_i$, if V_i is nonempty, it follows that

$$\Bigl\| \sum_{i=1}^{l} \alpha_i \mathcal{X}_{E_i} - f \Bigr\|_{\sup} \le \frac{1}{k} \,,$$

which proves the desired approximation property.

Now, if $(f_k)_{k\in\mathbb{N}}$ is a sequence of step functions that converges uniformly to f, then it follows that

$$\Bigl| \int_S f_k \,\mathrm{d}\lambda - \int_S f_l \,\mathrm{d}\lambda \Bigr| \le \| f_k - f_l \|_{\sup} \cdot \| \lambda \|_{\mathrm{var}} \longrightarrow 0 \quad \text{as } k, l \to \infty.$$

Hence there exists

$$\int_S f \,\mathrm{d}\lambda := \lim_{k\to\infty} \int_S f_k \,\mathrm{d}\lambda \,,$$

and the limit is independent of the choice of approximating sequence $(f_k)_{k\in\mathbb{N}}$.

6.23 Riesz-Radon theorem (Dual space of C^0). Let $S \subset \mathbb{R}^n$ be compact. Then

$$J(\nu)(f) := \int_S f \,\mathrm{d}\nu$$

defines a linear isometric isomorphism

$$J : rca(S) \to C^0(S)' \,.$$

Here $rca(S)$ is the space defined in 6.21 and the integral for continuous functions is defined as in 6.22.

Proof. For $\nu \in rca(S)$ and $f \in C^0(S)$ it follows from the definition of the Riemann integral that

$$|J(\nu)(f)| = \Bigl| \int_S f \,\mathrm{d}\nu \Bigr| \le \| f \|_{\sup} \cdot \| \nu \|_{\mathrm{var}} \,,$$

and hence J is continuous. Moreover, J is isometric. To see this, note that for $\nu \in rca(S)$ and $\varepsilon > 0$ there exists a partitioning of S into Borel sets E_i, $i = 1, \dots, m$, with

$$\| \nu \|_{\mathrm{var}} \le \varepsilon + \sum_{i=1}^{m} | \nu(E_i) | \,.$$

As ν is regular, there exist compact sets $K_i \subset E_i$ with $|\nu|(E_i \setminus K_i) \le \frac{\varepsilon}{m}$. Then $\mathrm{B}_\delta(K_i)$ are disjoint sets for sufficiently small $\delta > 0$, and

$$|\nu|\bigl(S \cap \mathrm{B}_\delta(K_i) \setminus K_i\bigr) \to 0 \quad \text{as } \delta \searrow 0,$$

which follows once again from the regularity of ν. On defining

$$f_i(x) := \max\bigl(1 - \tfrac{1}{\delta}\operatorname{dist}(x, K_i), 0\bigr)$$

and

$$\sigma_i := \begin{cases} \dfrac{\overline{\nu(K_i)}}{|\nu(K_i)|}, & \text{if } \nu(K_i) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

it holds, if δ is sufficiently small, that

$$\Bigl\|\sum_{i=1}^{m} \sigma_i f_i\Bigr\|_{\sup} \leq 1$$

and

$$\begin{aligned} \Bigl|J(\nu)\Bigl(\sum_{i=1}^{m} \sigma_i f_i\Bigr)\Bigr| &= \Bigl|\sum_{i=1}^{m} \sigma_i \int_S f_i \,\mathrm{d}\nu\Bigr| \\ &= \Bigl|\sum_{i=1}^{m}\Bigl(|\nu(K_i)| + \sigma_i \int_{S\cap \mathrm{B}_\delta(K_i)\setminus K_i} f_i \,\mathrm{d}\nu\Bigr)\Bigr| \\ &\geq \sum_{i=1}^{m} |\nu(K_i)| - \sum_{i=1}^{m} |\nu|\,(S \cap \mathrm{B}_\delta(K_i) \setminus K_i) \\ &\geq \|\nu\|_{\mathrm{var}} - 2\varepsilon - \sum_{i=1}^{m} |\nu|\,(S \cap \mathrm{B}_\delta(K_i) \setminus K_i) \\ &\longrightarrow \|\nu\|_{\mathrm{var}} \quad \text{on letting } \delta \searrow 0 \text{ and then } \varepsilon \searrow 0. \end{aligned}$$

Now the crucial step is to show that for $F \in C^0(S)'$ there exists a $\nu \in rca(S)$ with $J(\nu) = F$. It follows from the Hahn-Banach theorem that F can be extended norm-preservingly to $F \in B(S)'$ ($B(S)$ is the space defined in 3.1). Define

$$\lambda(E) := F(\mathcal{X}_E) \quad \text{for } E \subset S.$$

Then λ is additive and $\|\lambda\|_{\mathrm{var}} \leq \|F\|_{B(S)'}$, which follows as in (6-15). Therefore, by the definition of the Riemann integral,

$$F(f) = \int_S f \,\mathrm{d}\lambda$$

for all $f \in C^0(S)$. Hence we want to find a $\nu \in rca(S)$ such that

$$\int_S f \,\mathrm{d}\nu = \int_S f \,\mathrm{d}\lambda \quad \text{for all } f \in C^0(S).$$

The proof that such a ν exists is given in Appendix A6 (see A6.6). □

With the help of the result in theorem 6.23, we can provide a distributional characterization of regular measures.

6.24 Corollary. Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $C \geq 0$ and let

$$T : C_0^0(\Omega) \to \mathbb{K} \quad \text{be linear with} \quad |T(\zeta)| \leq C \cdot \|\zeta\|_{\sup} \text{ for all } \zeta \in C_0^0(\Omega)\,.$$

Then there exists a unique $\lambda \in rca(\Omega)$ with

$$\|\lambda\|_{\mathrm{var}} = \sup\left\{ |T(\zeta)| \;;\; \zeta \in C_0^0(\Omega),\ \|\zeta\|_{\sup} = 1 \right\} \leq C\,,$$
$$T(\zeta) = \int_\Omega \zeta \,\mathrm{d}\lambda \quad \text{for all } \zeta \in C_0^0(\Omega)\,.$$

Remark: It is sufficient to assume that

$$T \in \mathscr{D}'(\Omega) \quad \text{with} \quad |T(\zeta)| \leq C \cdot \|\zeta\|_{\sup} \text{ for all } \zeta \in C_0^\infty(\Omega)\,.$$

That is because T can then be uniquely extended to a linear map on $C_0^0(\Omega)$, which satisfies the above estimate (approximate functions in $C_0^0(\Omega)$ by means of convolutions).

Proof. Consider the open sets

$$\Omega_m := \left\{ x \in \Omega \;;\; \mathrm{dist}(x, \partial\Omega) > \tfrac{1}{m} \right\}\,.$$

For $m \geq m_0$, with m_0 sufficiently large, Ω_m is nonempty and $S_m := \overline{\Omega_m} \subset \Omega_{m+1}$ is compact. For $m > m_0$ choose $\eta_m \in C_0^\infty(\Omega_m)$ with $0 \leq \eta_m \leq 1$ and $\eta_m = 1$ on S_{m-1}. Then

$$T_m(g) := T(\eta_m g) \quad \text{for } g \in C^0(S_m)$$

defines a $T_m \in C^0(S_m)'$ with

$$\|T_m\| \leq C_T := \sup\{|T(\zeta)|\,;\ \zeta \in C_0^0(\Omega),\ \|\zeta\|_{\sup} = 1\} \leq C\,.$$

Hence it follows from 6.23 that there exist uniquely determined $\nu_m \in rca(S_m)$ with $\|\nu_m\|_{\mathrm{var}} \leq C_T$ and

$$T_m(g) = \int_{S_m} g \,\mathrm{d}\nu_m \quad \text{for } g \in C^0(S_m).$$

For $\zeta \in C_0^0(\Omega_m)$ and $l > m$ it holds that $\eta_l \zeta = \zeta$ (here we set $\zeta = 0$ outside of Ω_m), and so

$$\int_{S_m} \zeta \,\mathrm{d}\nu_l = \int_{S_l} \zeta \,\mathrm{d}\nu_l = T(\eta_l \zeta) = T(\zeta)$$

independently of l. We claim that

$$\nu_l(E) \text{ is independent of } l > m \text{ for Borel sets } E \subset S_{m-1}. \tag{6-18}$$

Indeed, let $K \subset S_{m-1}$ be compact. Then $\zeta_\delta(x) := \max\bigl(1-\frac{1}{\delta}\mathrm{dist}(x,K),0\bigr)$ for small $\delta > 0$ defines a $\zeta_\delta \in C_0^0(\Omega_m)$. Since ν_l is a regular measure, $|\nu_l|(\mathrm{B}_\delta(K)\setminus K) \searrow 0$ as $\delta \searrow 0$, and hence

$$\int_{S_m} \zeta_\delta \,\mathrm{d}\nu_l \longrightarrow \nu_l(K) \quad \text{as } \delta \searrow 0,$$

i.e. (6-18) holds for compact sets in S_{m-1}. The regularity of ν_l then implies that (6-18) holds for all Borel sets. For Borel sets E with $\overline{E} \subset \Omega$ we have that $E \subset S_m$ for some $m \in \mathbb{N}$, and it follows from (6-18) that

$$\lambda(E) := \nu_l(E) \quad \text{for } l,m \in \mathbb{N} \text{ with } E \subset S_m,\ l \geq m+2$$

is well defined. For $\zeta \in C_0^0(\Omega)$ it holds that $\mathrm{supp}(\zeta) \subset \Omega_m$ for some $m \in \mathbb{N}$ and

$$T(\zeta) = \int_{S_m} \zeta \,\mathrm{d}\lambda$$

independently of m.

We need to show that λ can be extended to a $\lambda \in rca(\Omega)$. If E_i, $i = 1,\dots,k$, are pairwise disjoint with $\overline{E_i} \subset \Omega$, then, as above, there exists an m with $E_i \subset S_m$ for $i = 1,\dots,k$ and

$$\sum_{i=1}^{k} |\lambda(E_i)| = \sum_{i=1}^{k} |\nu_{m+2}(E_i)| \leq \|\nu_{m+2}\|_{\mathrm{var}} \leq C_T\,.$$

In addition, for every Borel set $E \subset \Omega$ the limit

$$\lambda(E) := \lim_{m\to\infty} \lambda(E \cap S_m) \tag{6-19}$$

exists. To see this, let $E_m := E \cap S_m \setminus S_{m-1}$ for $m > m_0$ and $E_{m_0} := E \cap S_{m_0}$. Then

$$E \cap S_m = \bigcup_{i=m_0}^{m} E_i\,, \quad \lambda(E \cap S_m) = \sum_{i=m_0}^{m} \lambda(E_i)$$

and, as shown above,

$$\sum_{i=m_0}^{m} |\lambda(E_i)| \leq C_T\,.$$

Hence (6-19) defines an extension of λ to the Borel sets of Ω. Then it easily follows that $\lambda \in rca(\Omega)$ with $\|\lambda\|_{\mathrm{var}} \leq C_T$. From the representation of T it then easily follows that $C_T \leq \|\lambda\|_{\mathrm{var}}$. ☐

As an application of theorem 6.23 (and in particular of 6.24), we consider the space $BV(\Omega)$. This space plays an important role in the functional analysis treatment of certain geometric differential equations, because it replaces the space $W^{1,p}(\Omega)$ for $p = 1$, which is not reflexive (see 8.11(4)). The functions in $BV(\Omega)$ have the advantage that their weak derivatives (see 6.25, below) can be interpreted as elements of a dual space. For existence proofs in reflexive spaces one employs theorem 8.10, however in the space $BV(\Omega)$ one can apply theorem 8.5.

6.25 Functions of bounded variation. Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Consider pairs (f, λ) with $f \in L^1(\Omega)$ and $\lambda \in rca(\Omega; \mathbb{K}^n)$ such that the following rule of integration by parts holds:

$$\int_\Omega \partial_i \zeta \cdot f \, \mathrm{dL}^n + \int_\Omega \zeta \, \mathrm{d}\lambda_i = 0 \quad \text{for all } \zeta \in C_0^\infty(\Omega) \tag{6-20}$$

for $i = 1, \ldots, n$. This is equivalent to

$$\partial_i[f] = [\lambda_i] \quad \text{in } \mathscr{D}'(\Omega)$$

for $i = 1, \ldots, n$.
Notation: The λ_i-integral is defined in 6.22, while the distributions $[f]$ and $[\lambda_i]$ are defined in 5.15.

In the spirit of the analogous definition in Sobolev spaces, we call $\partial_i f := \lambda_i$ the ***weak derivative*** of f. We have that:

(1) The set

$$BV(\Omega) := \Big\{ f \in L^1(\Omega) \; ; \; \text{there exists a } \lambda \in rca(\Omega; \mathbb{K}^n), \text{ such that (6-20) holds} \Big\}$$

of functions of ***bounded variation*** is a $\mathbb{K}$-vector space, and it becomes a Banach space with the norm

$$\|f\|_{BV(\Omega)} := \|f\|_{L^1(\Omega)} + \|\lambda\|_{\mathrm{var}} .$$

(2) $W^{1,1}(\Omega) \subset BV(\Omega)$ with a continuous inclusion.
(3) $W^{1,1}(\Omega)$ is a proper subset of $BV(\Omega)$.

Proof (2). For $f \in W^{1,1}(\Omega)$ the corresponding measure $\lambda \in rca(\Omega; \mathbb{K}^n)$ is given by

$$\lambda(E) := \int_E \nabla f \, \mathrm{dL}^n .$$

Moreover, $\|\lambda\|_{\mathrm{var}} \leq \|\nabla f\|_{L^1(\Omega)}$. □

Proof (3). The fact that the space $BV(\Omega)$ is larger than $W^{1,1}(\Omega)$ follows from the existence of measures that have no representation as a function. For instance, for $\Omega = \,]-1, 1[\, \subset \mathbb{R}$ the ***Heaviside function***

$$f(x) := \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0, \end{cases}$$

lies in $BV(]-1,1[)$ with

$$\int_{-1}^{1} \zeta' f \,\mathrm{d}\mathrm{L}^1 = -\zeta(0) = -\int_{-1}^{1} \zeta \,\mathrm{d}\delta_0 \,,$$

i.e. the weak derivative is the Dirac measure δ_0 at the point 0, and so

$$[f]' = [\delta_0] \quad \text{in } \mathscr{D}'(\Omega).$$

This example can be generalized to an arbitrary Ω. □

The following theorem yields an equivalent definition of the space $BV(\Omega)$, which is formulated with the help of the distribution $[f] \in \mathscr{D}'(\Omega)$ for $f \in L^1(\Omega)$ (see 5.15). An additional possible definition in the case $n = 1$ is presented in E6.9.

6.26 Theorem. Let $\Omega \subset \mathbb{R}^n$ be open and bounded, and for $f \in L^1(\Omega)$ let

$$\|f\|_{\mathrm{grad}} := \sup\left\{ \left| \int_\Omega f \operatorname{div} g \,\mathrm{d}\mathrm{L}^n \right| \,;\; g \in C_0^\infty(\Omega;\mathbb{K}^n) \text{ with } |g(x)| \le 1 \text{ for } x \in \Omega \right\} \in [0,\infty] \,.$$

Here the ***divergence*** of a vector field is defined by

$$\operatorname{div} v := \sum_{i=1}^{n} \partial_i v_i \qquad \text{for } v \in C^1(\Omega;\mathbb{K}^n).$$

Then

$$BV(\Omega) = \left\{ f \in L^1(\Omega) \,;\; \|f\|_{\mathrm{grad}} < \infty \right\}$$

and for $f \in BV(\Omega)$ with $\nabla f := (\partial_i f)_{i=1,\ldots,n} \in rca(\Omega;\mathbb{K}^n)$,

$$\|f\|_{\mathrm{grad}} = \|\nabla f\|_{\mathrm{var}}.$$

Proof. For $g \in C_0^0(\Omega;\mathbb{K}^n)$ let

$$\int_\Omega g \bullet \mathrm{d}\lambda := \sum_{i=1}^{n} \int_\Omega g_i \,\mathrm{d}\overline{\lambda_i} \,, \qquad \text{so that} \qquad \left| \int_\Omega g \bullet \mathrm{d}\lambda \right| \le \|g\|_{\sup} \cdot \|\lambda\|_{\mathrm{var}} \,,$$

which follows by approximating g with step functions as in 6.22.

For $f \in BV(\Omega)$ with $\lambda_i := \partial_i f$ as in 6.25 and g as in the above definition of $\|f\|_{\mathrm{grad}}$ it then holds that

$$\left| \int_\Omega f \operatorname{div} g \,\mathrm{d}\mathrm{L}^n \right| = \left| \sum_{i=1}^{n} \int_\Omega g_i \,\mathrm{d}\lambda_i \right| = \left| \int_\Omega \overline{g} \bullet \mathrm{d}\lambda \right| \le \|\lambda\|_{\mathrm{var}} \,,$$

and so $\|f\|_{\mathrm{grad}} \le \|\lambda\|_{\mathrm{var}}$.

Now let $f \in L^1(\Omega)$ with $\|f\|_{\text{grad}} < \infty$ and put

$$T_i(\zeta) := -\int_\Omega f\partial_i\zeta \,\mathrm{dL}^n = -\int_\Omega f \operatorname{div}(\zeta \mathbf{e}_i)\,\mathrm{dL}^n \quad \text{for } \zeta \in C_0^\infty(\Omega).$$

By the definition of $\|f\|_{\text{grad}}$ it holds that $|T_i(\zeta)| \le \|\zeta\|_{\sup} \cdot \|f\|_{\text{grad}}$. This estimate shows that T_i can be uniquely extended onto $C_0^0(\Omega)$. Hence, by 6.24, there exists a $\lambda_i \in rca(\Omega)$ with

$$T_i(\zeta) = \int_\Omega \zeta \,\mathrm{d}\lambda_i \quad \text{for } \zeta \in C_0^0(\Omega).$$

This shows that $f \in BV(\Omega)$ with $\partial_i f = \lambda_i$. On setting $\lambda := (\lambda_i)_{i=1,\dots,n}$ it then holds for $g \in C_0^\infty(\Omega;\mathbb{K}^n)$ that

$$\overline{\int_\Omega g \bullet \mathrm{d}\lambda} = \sum_{i=1}^n \int_\Omega \overline{g_i}\,\mathrm{d}\lambda_i = \sum_{i=1}^n T_i(\overline{g_i}) = -\int_\Omega f \operatorname{div}(\overline{g})\,\mathrm{dL}^n\,,$$

and so

$$\left|\int_\Omega g \bullet \mathrm{d}\lambda\right| \le \|g\|_{\sup} \cdot \|f\|_{\text{grad}}\,.$$

Similarly to the proof of the isometry property in 6.23, this implies the inequality $\|\lambda\|_{\text{var}} \le \|f\|_{\text{grad}}$. ☐

E6 Exercises

E6.1 Dual norm on $\mathbb{R}^n$. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$, i.e. we consider the normed space $(\mathbb{R}^n, \|\cdot\|)$.

(1) Show that

$$J(x)(y) := \sum_{i=1}^n y_i x_i \quad \text{for } x, y \in \mathbb{R}^n$$

defines a linear map $J : (\mathbb{R}^n, \|\cdot\|) \to (\mathbb{R}^n, \|\cdot\|)'$.

(2) Show that

$$\|x\|' := \|J(x)\| \quad \text{for } x \in \mathbb{R}^n$$

defines a norm on $\mathbb{R}^n$ (we call it the ***dual norm*** to $\|\cdot\|$).

(3) $J : (\mathbb{R}^n, \|\cdot\|') \to (\mathbb{R}^n, \|\cdot\|)'$ is an isometric isomorphism.

(4) For $1 \le p \le \infty$, find the dual norm to the p-norm in 2.5.

E6.2 Dual space of the cross product. Let X_1 and X_2 be normed spaces and

$$J : X_1' \times X_2' \to (X_1 \times X_2)' ,$$
$$J\big((x_1', x_2')\big)\big((x_1, x_2)\big) := x_1'x_1 + x_2'x_2 .$$

Show that J is an isometric isomorphism if the norms in $X_1 \times X_2$ and $X_1' \times X_2'$ are defined as in E4.12(1) with respect to $|\cdot|$ and $|\cdot|'$, respectively.
Remark: Here $|\cdot|'$ is the dual norm to $|\cdot|$ from E6.1(2).
Show that this dual norm is also a monotone norm on $\mathbb{R}^2$.

E6.3 Integral equation. Let $K \in L^2(\Omega \times \Omega)$ and let $f \in L^2(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is Lebesgue measurable. For $\lambda \in \mathbb{R}$ consider the ***integral equation***

$$\int_\Omega K(x, y)u(y)\,\mathrm{d}y = \lambda u(x) + f(x) \qquad \text{for almost all } x \in \Omega.$$

Show that for $\lambda > \|K\|_{L^2(\Omega\times\Omega)}$ there exists a unique solution $u \in L^2(\Omega)$.

Solution. It follows from 5.12 that

$$(Tu)(x) := \int_\Omega K(x, y)u(y)\,\mathrm{d}y$$

defines an operator $T \in \mathscr{L}\big(L^2(\Omega)\big)$ with $\|T\|_{\mathscr{L}(L^2(\Omega))} \le \|K\|_{L^2(\Omega\times\Omega)}$. Then also $A := \lambda\mathrm{Id} - T \in \mathscr{L}\big(L^2(\Omega)\big)$ and for $u \in L^2(\Omega)$

$$\begin{aligned}\operatorname{Re}(u\,,\,Au)_{L^2} &= \lambda\|u\|_{L^2}^2 - \operatorname{Re}(u\,,\,Tu)_{L^2} \\ &\ge \lambda\|u\|_{L^2}^2 - \|T\|_{\mathscr{L}(L^2(\Omega))}\cdot\|u\|_{L^2}^2 \\ &\ge \underbrace{\Big(\lambda - \|K\|_{L^2(\Omega\times\Omega)}\Big)}_{=:c_0>0}\|u\|_{L^2}^2 .\end{aligned}$$

It follows from the Lax-Milgram theorem (see the equivalent result 6.3(3)) that A is invertible, and so $u := A^{-1}(-f)$ is the solution of the integral equation. □

E6.4 Examples of elements from $C^0([0,1])'$. Show that the following maps T are linear and continuous on $C^0([0,1])$ and calculate their norm.

(1) $T : C^0([0,1]) \to C^0([0,1])$, for a given $g \in C^0([0,1])$ defined by

$$(Tf)(x) := g(x)\cdot f(x) .$$

(2) $T : C^0([0,1]) \to \mathbb{K}$, with $\alpha_i \in \mathbb{R}$ and pairwise distinct $x_i \in [0,1]$, $i = 1, \dots, m$, defined by

$$Tf := \textstyle\sum_{i=1}^m \alpha_i f(x_i) .$$

(3) $T : C^0([0,1]) \to \mathbb{K}$, with points x_i and coefficients α_i as in (2), defined by

$$Tf := \int_0^1 f(x)\,\mathrm{d}x - \sum_{i=1}^m \alpha_i f(x_i)\,.$$

Solution (1). On noting that $|(Tf)(x)| \le \|g\|_{\sup}\|f\|_{\sup}$, we have that T is continuous, with $\|T\| \le \|g\|_{\sup}$. As $\|Tg\|_{\sup} = \left\|g^2\right\|_{\sup} = \|g\|_{\sup}^2$, it holds that $\|T\| \ge \|g\|_{\sup}$. □

Solution (2). Since

$$|Tf| \le \sum_{i=1}^m |\alpha_i| \cdot \|f\|_{\sup}\,,$$

T is continuous, with $\|T\| \le \sum_{i=1}^m |\alpha_i|$. As the x_i are pairwise distinct, there exists a continuous function f with $|f| \le 1$ and $f(x_i) = \mathrm{sign}(\alpha_i)$ for $i = 1, \dots, m$. Then

$$|Tf| = \sum_{i=1}^m |\alpha_i|\,, \quad \text{and so} \quad \|T\| \ge \sum_{i=1}^m |\alpha_i|\,.$$

□

Solution (3). Since

$$|Tf| \le \Big(1 + \sum_{i=1}^m |\alpha_i|\Big)\|f\|_{\sup}\,,$$

T is continuous, with $\|T\| \le 1 + \sum_{i=1}^m |\alpha_i|$. Now for small $\delta > 0$, chosen so that $\delta < \frac{1}{2}|x_i - x_j|$ for all $i \ne j$, consider the continuous function

$$f(x) := \begin{cases} (1 - \frac{|x-x_i|}{\delta})\mathrm{sign}(-\alpha_i) + \frac{|x-x_i|}{\delta} & \text{if } x \in I_{i\delta} \text{ for an } i, \\ 1 & \text{otherwise,} \end{cases}$$

where $I_{i\delta} := [x_i - \delta, x_i + \delta]$ are disjoint intervals. Then $\|f\|_{\sup} = 1$ and

$$\begin{aligned} |Tf| &= \left| \int_0^1 (f(x) - 1)\,\mathrm{d}x + 1 + \sum_{i=1}^m |\alpha_i| \right| \\ &= \left| \sum_{i=1}^m \Big(\int_{[0,1]\cap I_{i\delta}} (f(x) - 1)\,\mathrm{d}x \Big) + 1 + \sum_{i=1}^m |\alpha_i| \right| \\ &\ge -4m\delta + 1 + \sum_{i=1}^m |\alpha_i|\,, \end{aligned}$$

which shows that $\|T\| \ge 1 + \sum_{i=1}^m |\alpha_i|$.
Result: This means that no such ***quadrature formula*** can approximate the integral over $[0,1]$ for all (!) continuous functions. □

E6.5 Dual space of $C^m(I)$. Let $I \subset \mathbb{R}$ be a closed interval and let $x_0 \in I$. Then, for $m \ge 1$,

$$J(\xi,\nu)(f) := \sum_{i=1}^m \xi_i f^{(i-1)}(x_0) + \int_I f^{(m)}\,\mathrm{d}\nu$$

defines an isomorphism $J : \mathbb{K}^m \times rca(I) \to C^m(I)'$.

Solution. It holds that

$$|J(\xi,\nu)(f)| \le \Big(\max_{i=1,\dots,m} |\xi_i| + \|\nu\|_{\mathrm{var}}\Big)\|f\|_{C^m(I)}\,,$$

and hence J is continuous with $\|J\| \le 1$ if on $\mathbb{K}^m \times rca(I)$ we introduce the norm

$$\|(\xi,\nu)\| := \max_{i=1,\dots,m} |\xi_i| + \|\nu\|_{\mathrm{var}}$$

and if the C^m-norm is defined as in 3.6. Now for every function $f \in C^m(I)$ we have

$$f(x) = \sum_{i=0}^{m-1} \frac{1}{i\,!} f^{(i)}(x_0)(x-x_0)^i + \frac{1}{(m-1)!}\int_{x_0}^x f^{(m)}(y)(x-y)^{m-1}\,\mathrm{d}y\,.$$

This can be shown by induction on m. First, note that for $m = 1$ this is the fundamental theorem of calculus. The following identity then proves the formula inductively:

$$\begin{aligned}\int_{x_0}^x f^{(m)}(y)(x-y)^{m-1}\,\mathrm{d}y &= -\frac{1}{m}\int_{x_0}^x f^{(m)}(y)\frac{\mathrm{d}}{\mathrm{d}y}(x-y)^m\,\mathrm{d}y\\ &= \frac{1}{m} f^{(m)}(x_0)(x-x_0)^m + \frac{1}{m}\int_{x_0}^x f^{(m+1)}(y)(x-y)^m\,\mathrm{d}y\,.\end{aligned}$$

Hence, for every $F \in C^m(I)'$ we have

$$Ff = \sum_{i=0}^{m-1} f^{(i)}(x_0)Fp_i + FTf^{(m)}\,,$$

where

$$p_i(x) := \frac{(x-x_0)^i}{i\,!} \quad \text{and} \quad Tg(x) := \int_{x_0}^x g(y)\frac{(x-y)^{m-1}}{(m-1)!}\,\mathrm{d}y\,.$$

For $i = 0,\dots,m-1$ it follows inductively that

$$(Tg)^{(i)}(x) = \int_{x_0}^x g(y)\frac{(x-y)^{m-1-i}}{(m-1-i)!}\,\mathrm{d}y\,,$$

since the integrand vanishes at the upper limit x. In particular,

$$(Tg)^{(m-1)}(x) = \int_{x_0}^x g(y)\,\mathrm{d}y\,, \quad \text{and so} \quad (Tg)^{(m)}(x) = g(x)\,.$$

Hence we have the estimate $\|Tg\|_{C^m(I)} \le C \cdot \|g\|_{C^0(I)}$ and it follows that $T \in \mathscr{L}(C^0(I); C^m(I))$, which implies that $FT \in C^0(I)'$. By theorem 6.23, there exists a $\nu \in rca(I)$ with $\|\nu\|_{\mathrm{var}} = \|FT\|$ and

$$FTg = \int_I g \, \mathrm{d}\nu \qquad \text{for } g \in C^0(I).$$

Setting $\xi_i := Fp_{i-1}$ for $i = 1, \ldots, m$, we have that

$$F = J(\xi, \nu)$$

and

$$\|(\xi, \nu)\| \leq \Big(\max_{i=0,\ldots,m-1} \|p_i\|_{C^m(I)} + \|T\|\Big)\|F\| .$$

This shows that J is surjective. If in addition we can show that J is injective, then this estimate yields that the inverse J^{-1} is also continuous. If $J(\xi, \nu) = 0$, then it holds for $i = 1, \ldots, m$ that

$$0 = J(\xi, \nu)p_{i-1} = \xi_i$$

and for all $g \in C^0(I)$ that

$$0 = J(\xi, \nu)Tg = \int_I g \, \mathrm{d}\nu ,$$

which yields $\nu = 0$, thanks to theorem 6.23. Hence J is injective. □

Remark: If

$$J_1(\xi)(z) := z \bullet \xi$$

is the isometry $J_1 : \mathbb{K}^m \to (\mathbb{K}^m)'$ from 6.1 and

$$J_2(\nu)(g) := \int_I g \, \mathrm{d}\nu$$

is the isometry $J_2 : rca(I) \to C^0(I)'$ from 6.23, then it follows from E6.2 that

$$J_0(\xi, \nu)(z, g) := J_1(\xi)(z) + J_2(\nu)(g)$$

defines an isomorphism $J_0 : \mathbb{K}^m \times rca(I) \to (\mathbb{K}^m \times C^0(I))'$. Moreover,

$$S(f) := \Big(\big(f^{(i)}(x_0)\big)_{i=0,\ldots,m-1} \, , \, f^{(m)}\Big)$$

defines a continuous linear map from $C^m(I)$ to $\mathbb{K}^m \times C^0(I)$. With these definitions

$$J = S'J_0 ,$$

where S' is the adjoint map of S (see 5.5(8)). Hence J being an isomorphism is equivalent to the isomorphy of S' and, by theorem 12.5, equivalent to the isomorphy of S.

E6.6 Dual space of c_0 and c. Let

$$c_0 := \left\{ x \in \ell^\infty(\mathbb{R})\ ;\ \lim_{i\to\infty} x_i = 0 \right\} ,$$
$$c\ := \left\{ x \in \ell^\infty(\mathbb{R})\ ;\ \text{it exists } \lim_{i\to\infty} x_i \right\} .$$

The sets c_0 and c, equipped with the $\ell^\infty(\mathbb{R})$-norm, are Banach spaces. Characterize the dual spaces c_0' and c'.

Solution. For every $y \in \ell^1(\mathbb{R})$, setting

$$J(y)(x) := \sum_{i=1}^\infty y_i x_i \qquad \text{for } x \in c_0$$

defines a $J(y) \in c_0'$ with $\|J(y)\| \le \|y\|_{\ell^1}$, because

$$|J(y)(x)| \le \sup_i |x_i| \cdot \sum_{i=1}^\infty |y_i| = \|x\|_{\ell^\infty} \|y\|_{\ell^1} .$$

If we define for $n \in \mathbb{N}$

$$x_i := \begin{cases} \operatorname{sign}(y_i) & \text{for } i \le n, \\ 0 & \text{for } i > n, \end{cases}$$

then $\left\|(x_i)_{i\in\mathbb{N}}\right\|_{\ell^\infty} = 1$ and

$$J(y)(x) = \sum_{i\le n} |y_i| \to \|y\|_{\ell^1} \qquad \text{as } n \to \infty.$$

Hence $J : \ell^1(\mathbb{R}) \to c_0'$ is isometric. Now let $F \in c_0'$. Since for all $x \in c_0$ we have that

$$x = \sum_{i=1}^\infty x_i e_i \qquad \text{in the } \ell^\infty\text{-norm,}$$

it follows that

$$F(x) = \sum_{i=1}^\infty x_i F e_i ,$$

and so $F = J(y)$, where $y_i := Fe_i$, provided that $y \in \ell^1(\mathbb{R})$. But this is indeed the case, since

$$\sum_{i\le n} |y_i| = F\Big(\sum_{i\le n} \operatorname{sign}(y_i)\, e_i\Big) \le \|F\| \cdot \Big\|\sum_{i\le n} \operatorname{sign}(y_i)\, e_i\Big\|_{\ell^\infty} = \|F\| .$$

This shows that J is an isomorphism. Then the dual space c' can be characterized as follows:

$$Sx := (\lim_{i\to\infty} x_i, x_1 - \lim_{i\to\infty} x_i, x_2 - \lim_{i\to\infty} x_i, \ldots)$$

defines an $S \in \mathscr{L}(c; c_0)$, and S is in fact an isomorphism, with

$$S^{-1}x = (x_2 + x_1, x_3 + x_1, x_4 + x_1, \ldots) .$$

Therefore

$$\widetilde{J}(y) := J(y)S$$

defines an isomorphism $\widetilde{J} : \ell^1(\mathrm{I\!R}) \to c'$. □

E6.7 Characterization of Sobolev functions. Let $\Omega \subset \mathrm{I\!R}^n$ be open. For $m \in \mathrm{I\!N} \cup \{0\}$ and $1 < p \le \infty$ (if $m = 0$ then also for $p = 1$) it holds for functions $f : \Omega \to \mathrm{I\!R}$ that

$$f \in W^{m,p}(\Omega) \quad \Longleftrightarrow \quad \begin{cases} f \in L^1_{\mathrm{loc}}(\Omega) \text{ and there exists a constant } C \text{ with} \\ \left| \int_\Omega f \partial^s \zeta \,\mathrm{dL}^n \right| \le C \|\zeta\|_{L^{p'}(\Omega)} \\ \text{for all } |s| \le m \text{ and all } \zeta \in C_0^\infty(\Omega)\,. \end{cases}$$

Here p' is the dual exponent to p.
Note: For this characterization in the case $m = 0$, see 6.13. In case $m > 0$ we have to assume $p > 1$, see the space $BV(\Omega)$ and 6.26.

Solution $\Rightarrow$.

$$\left| \int_\Omega f \partial^s \zeta \,\mathrm{dL}^n \right| = \left| \int_\Omega \partial^s f \cdot \zeta \,\mathrm{dL}^n \right| \le \|\partial^s f\|_{L^p(\Omega)} \|\zeta\|_{L^{p'}(\Omega)}\,.$$

□

Solution $\Leftarrow$. It follows from 6.13 that $f \in L^p(\Omega)$. For $0 < |s| \le m$ let

$$F_s(\zeta) := \int_\Omega f \partial^s \zeta \,\mathrm{dL}^n \quad \text{for } \zeta \in C_0^\infty(\Omega).$$

The estimate $|F_s(\zeta)| \le C\|\zeta\|_{L^{p'}(\Omega)}$ says, since $p' < \infty$, that F_s can be extended to a functional on $L^{p'}(\Omega)$. Then it follows from 6.12, again since $p' < \infty$, that there exists a function $f_s \in L^p(\Omega)$ with

$$F_s(g) = \int_\Omega g \cdot f_s \,\mathrm{dL}^n \quad \text{for } g \in L^{p'}(\Omega).$$

Therefore,

$$\int_\Omega f \partial^s \zeta \,\mathrm{dL}^n = \int_\Omega f_s \zeta \,\mathrm{dL}^n \quad \text{for } \zeta \in C_0^\infty(\Omega),$$

which yields that $f \in W^{m,p}(\Omega)$ (with $\partial^s f = (-1)^{|s|} f_s$). □

E6.8 Positive functionals on C_0^0. Let $\Omega \subset \mathrm{I\!R}^n$ be open and let $F : C_0^0(\Omega; \mathrm{I\!R}) \to \mathrm{I\!R}$ be a linear map with

$$f \ge 0 \text{ in } \Omega \quad \Longrightarrow \quad F(f) \ge 0\,.$$

Then there exists a nonnegative locally bounded regular σ-additive measure μ on the Borel sets of Ω (μ is then also called a ***Radon measure***) such that

$$F(f) = \int_\Omega f \,\mathrm{d}\mu \quad \text{for all } f \in C_0^0(\Omega; \mathrm{I\!R}).$$

Solution. Here $\mathbb{K} = \mathbb{R}$. Let $D \subset \Omega$ be open and bounded with

$$d := \tfrac{1}{2}\mathrm{dist}(D, \partial\Omega) > 0\,.$$

In addition, let $S := \overline{\mathrm{B}_d(D)}$. Choose a cut-off function $\eta \in C_0^0(\Omega)$ (see 4.19) with

$$0 \le \eta \le 1,\ \eta = 1 \text{ on } D,\ \eta = 0 \text{ outside of } \mathrm{B}_d(D)\,,$$

e.g.

$$\eta(x) = \max(0, 1 - \tfrac{1}{d}\mathrm{dist}(x, D))\,.$$

For nonnegative functions $f \in C^0(S)$ we then have that $\eta f \in C_0^0(\Omega)$, with

$$0 \le \eta f \le \eta \sup_S f\,,$$

and so

$$0 \le F(\eta f) \le F(\eta) \cdot \sup_S f\,.$$

Then it follows for all $f \in C^0(S)$, on setting $f^+ := \max(f, 0)$ and $f^- := \max(-f, 0)$, that

$$\begin{aligned} |F(\eta f)| &= \big|F(\eta f^+) - F(\eta f^-)\big| \\ &\le (\sup_S f^+ + \sup_S f^-)F(\eta) \le \|f\|_{C^0(S)} \cdot F(\eta)\,. \end{aligned}$$

Hence $f \mapsto F(\eta f)$ is a continuous functional on $C^0(S)$, and 6.23 yields the existence of a $\mu \in rca(S)$ with

$$F(\eta f) = \int_S f\,\mathrm{d}\mu \quad \text{for all } f \in C^0(S).$$

For $f \in C_0^0(D)$ it holds that $\eta f = f$, and hence

$$F(f) = \int_S f\,\mathrm{d}\mu \quad \text{for all } f \in C_0^0(D).$$

We need to show that $\mu \ge 0$. As μ is regular, it is sufficient to show that $\mu(K) \ge 0$ for compact sets $K \subset D$. Now, define

$$\eta_\varepsilon(x) := \max\big(0, 1 - \tfrac{1}{\varepsilon}\mathrm{dist}(x, K)\big)\,,$$

so we have $\eta_\varepsilon \in C_0^0(D)$ for sufficiently small ε. Since $1 \ge \eta_\varepsilon \searrow \mathcal{X}_K$ pointwise as $\varepsilon \searrow 0$, we obtain that

$$0 \le F(\eta_\varepsilon) = \int_S \eta_\varepsilon\,\mathrm{d}\mu \longrightarrow \mu(K)\,.$$

A similar argument shows that $\widetilde{\mu} = \mu$ in D, if $\widetilde{\mu}$ is the measure in $rca(\widetilde{S})$ for a $\widetilde{D}$ as above with $D \subset \widetilde{D}$. Exhausting Ω with countably many (not necessarily connected) domains D then yields the desired result. □

As an alternative to the space $BV(\Omega)$ in 6.25 we define the following:

E6.9 Functions of bounded variation. In the one-dimensional case we define for $S := [a,b] \subset \mathbb{R}$

$$\widetilde{BV}(S) := \left\{ f : [a,b] \to \mathbb{K}\ ;\ \|f\|_{\widetilde{BV}} := |f(a)| + \mathrm{var}(f,S) < \infty \right\},$$

where the ***variation*** of f on $[a,b]$ is defined by

$$\mathrm{var}(f,[a,b]) := \sup\Big\{ \sum_{i=1}^{m} |f(a_i) - f(a_{i-1})|\ ; \quad m \in \mathbb{N},\ a = a_0 < a_1 < \ldots < a_m = b \Big\}.$$

Show that for $f \in \widetilde{BV}(S)$ it holds that:

(1) For $a \le x_1 < x_2 < x_3 \le b$,

$$\mathrm{var}(f,[x_1,x_3]) = \mathrm{var}(f,[x_1,x_2]) + \mathrm{var}(f,[x_2,x_3]).$$

(2) The following limits exist

$$f_+(x) := \lim_{\varepsilon\searrow 0} f(x+\varepsilon) \quad \text{for } a \le x < b,$$
$$f_-(x) := \lim_{\varepsilon\searrow 0} f(x-\varepsilon) \quad \text{for } a < x \le b.$$

(3) Every function in $\widetilde{BV}(S)$ has at most countably many discontinuity points.

Solution (1). The "$\le$" part in the identity follows from adding x_2 to the interval partitionings of $[x_1,x_3]$. □

Solution (2). Noting that

$$|f(x)| \le |f(a)| + |f(x) - f(a)| \le \|f\|_{\widetilde{BV}}$$

yields that f is bounded. Hence for $x < b$ there exists a sequence $(\kappa_i)_{i\in\mathbb{N}}$ with $\kappa_i \searrow x$ for $i \to \infty$, such that

$$\xi := \lim_{i\to\infty} f(\kappa_i)$$

exists. Now it follows from (1) that for all m

$$\textstyle\sum_{i=1}^{m} \mathrm{var}(f,[\kappa_{i+1},\kappa_i]) = \mathrm{var}(f,[\kappa_{m+1},\kappa_1]) \le \|f\|_{\widetilde{BV}} < \infty,$$

and hence

$$\textstyle\sum_{i=1}^{\infty} \mathrm{var}(f,[\kappa_{i+1},\kappa_i]) \le \|f\|_{\widetilde{BV}} < \infty,$$

which implies that $\operatorname{var}(f, [\kappa_{i+1}, \kappa_i]) \to 0$ as $i \to \infty$. Hence also

$$\begin{aligned} &\sup_{\kappa_{i+1} \le y \le \kappa_i} |f(y) - \xi| \\ &\le |f(\kappa_i) - \xi| + \sup_{\kappa_{i+1} \le y \le \kappa_i} |f(y) - f(\kappa_i)| \\ &\le |f(\kappa_i) - \xi| + \operatorname{var}(f, [\kappa_{i+1}, \kappa_i]) \to 0 \quad \text{as } i \to \infty, \end{aligned}$$

which shows that $\xi = f_+(x)$. □

Solution (3). If $a < x_1 < \ldots < x_m < b$ are discontinuity points of f, for which $|f_+(x_i) - f_-(x_i)| \ge \delta$, then it holds for small $\varepsilon \to 0$ that

$$\begin{aligned} \operatorname{var}(f, S) &\ge \textstyle\sum_{i=1}^m |f(x_i + \varepsilon) - f(x_i - \varepsilon)| \\ &\to \textstyle\sum_{i=1}^m |f_+(x_i) - f_-(x_i)| \ge m\delta\,, \end{aligned}$$

and so $m \le \delta^{-1} \|f\|_{\widetilde{BV}}$. On choosing a null sequence for δ, it follows that the discontinuity points of f are countable. □

Riemann-Stieltjes integral: Let $S = [a, b] \subset \mathbb{R}$ and $f \in \widetilde{BV}(S)$. Consider for $g \in C^0(S)$ and for partitionings $a = s_0 < s_1 < \ldots < s_n = b$ the sum

$$\sum_{i=1}^n g(s_i)\big(f(s_i) - f(s_{i-1})\big)\,.$$

If $(t_j)_{j=1,\ldots,m}$ is a finer partitioning of S, say $t_{k_i} = s_i$ with $k_{i-1} < k_i$, then, on setting $\delta_s := \max_i |s_i - s_{i-1}|$,

$$\begin{aligned} &\left| \sum_{i=1}^n g(s_i)\big(f(s_i) - f(s_{i-1})\big) - \sum_{j=1}^m g(t_j)\big(f(t_j) - f(t_{j-1})\big) \right| \\ &= \left| \sum_{i=1}^n \sum_{j=k_{i-1}+1}^{k_i} \big(g(s_i) - g(t_j)\big)\big(f(t_j) - f(t_{j-1})\big) \right| \\ &\le \sup_{|x_1 - x_2| \le \delta_s} |g(x_1) - g(x_2)| \cdot \|f\|_{\widetilde{BV}} \longrightarrow 0 \quad \text{as } \delta_s \to 0. \end{aligned}$$

Hence the ***Riemann-Stieltjes integral***

$$\int_S g \, \mathrm{d}f := \lim_{\delta_s \to 0} \sum_{i=1}^n g(s_i)\big(f(s_i) - f(s_{i-1})\big)$$

exists for $f \in \widetilde{BV}(S)$ and $g \in C^0(S)$.

E6.10 Representation of the Riemann-Stieltjes integral. Suppose that $f \in \widetilde{BV}(S)$. Then the following holds for the above defined integral.

(1) There exists a $\lambda \in rca(S)$ with

$$\int_S g\,\mathrm{d}f = \int_S g\,\mathrm{d}\lambda \quad \text{for all } g \in C^0(S)\,.$$

(2) The measure λ in (1) satisfies for $a \le x < b$

$$\lambda([a,x]) = \lim_{\varepsilon\searrow 0}\big(f(x+\varepsilon) - f(a)\big)\,.$$

Solution (1). The map

$$T_f(g) := \int_S g\,\mathrm{d}f$$

satisfies

$$\left|\int_S g\,\mathrm{d}f\right| \le \|g\|_{C^0}\cdot\|f\|_{\widetilde{BV}}\,.$$

It follows that $T_f \in C^0(S)'$ and hence theorem 6.23 yields the existence of a $\lambda \in rca(S)$ such that

$$\int_S g\,\mathrm{d}\lambda = T_f(g) = \int_S g\,\mathrm{d}f \quad \text{for all } g \in C^0(S)$$

and $\|\lambda\|_{\mathrm{var}} = \|T_f\|_{C^0(S)'}$. □

Solution (2). For $a < x_0 < b$ and sufficiently small $\varepsilon > 0$, consider the continuous function

$$g_\varepsilon(x) := \begin{cases} 1 & \text{for } x \le x_0+\varepsilon, \\ 1 - \frac{x-x_0-\varepsilon}{\varepsilon} & \text{for } x_0+\varepsilon \le x \le x_0+2\varepsilon, \\ 0 & \text{for } x_0+2\varepsilon \le x. \end{cases}$$

Then by the σ-additivity of $|\lambda|$

$$\int_{[a,x_0+\varepsilon]} g_\varepsilon\,\mathrm{d}\lambda = \lambda\big([a,x_0+\varepsilon]\big) \longrightarrow \lambda\big([a,x_0]\big) \quad \text{as } \varepsilon\to 0$$

and the definition of the Riemann integral gives

$$\left|\int_S g_\varepsilon\,\mathrm{d}\lambda - \lambda\big([a,x_0+\varepsilon]\big)\right| \le |\lambda|\big([x_0+\varepsilon,x_0+2\varepsilon]\big) \longrightarrow 0$$

for a sequence $\varepsilon\to 0$, since $\|\lambda\|_{\mathrm{var}} < \infty$. Moreover, by the definition of the Riemann-Stieltjes integral,

$$\int_{[a,x_0+\varepsilon]} g_\varepsilon\,\mathrm{d}f = f(x_0+\varepsilon) - f(a)$$

which converges to $\lim_{\varepsilon\searrow 0}\big(f(x_0+\varepsilon)-f(a)\big)$, and

$$\left|\int_S g_\varepsilon\,\mathrm{d}f - \big(f(x_0+\varepsilon)-f(a)\big)\right| \le \mathrm{var}(f,[x_0+\varepsilon,x_0+2\varepsilon]) \longrightarrow 0$$

as $\varepsilon\to 0$. □

Consider the functions

$$f_\varepsilon(x) := \begin{cases} 1 \text{ for } |x| \leq \varepsilon, \\ 0 \text{ otherwise}, \end{cases} \qquad f(x) := \begin{cases} 1 \text{ for } x = 0, \\ 0 \text{ otherwise}. \end{cases}$$

Then $f_\varepsilon \to f$ pointwise as $\varepsilon \to 0$ and $f \neq 0$ in $\widetilde{BV}([-1,1])$. Also,

$$\operatorname{var}(f, [-1,1]) = 2, \text{ but } \int_{[-1,1]} g \,\mathrm{d}f = 0 \text{ for all } g \in C^0([-1,1]).$$

In fact, with respect to the L^1-measure we have $f_\varepsilon \to 0$ almost everywhere as $\varepsilon \to 0$. As a consequence one considers function spaces

$$\begin{aligned} BV_{rc}([a,b]) := \{ f \in \widetilde{BV}([a,b])\,;\ & f(x) = f_+(x) \text{ for } a \leq x < b, \\ & f(b) = f_-(b) \}\,, \\ BV_{lc}([a,b]) := \{ f \in \widetilde{BV}([a,b])\,;\ & f(a) = f_+(a)\,, \\ & f(x) = f_-(x) \text{ for } a < x \leq b \}\,, \end{aligned}$$

which consist of right-continuous and left-continuous functions, respectively. Both spaces are bijective (isomorphic) to $BV(]a,b[)$ in 6.25.

E6.11 Normalized *BV* functions. With $S := [a,b] \subset \mathbb{R}$ and the notations as in E6.9, let

$$\begin{aligned} NBV(S) := \{ f \in \widetilde{BV}(S)\,;\ & f(x) = f_+(x) \text{ for } a \leq x < b, \\ & f(a) = 0 \text{ and } f(b) = f_-(b) \} \end{aligned}$$

be the space of ***normalized functions of bounded variation***, equipped with the norm of $\widetilde{BV}(S)$. Show that

$$(J\lambda)(x) := \lambda\big([a,x]\big) \quad \text{for } a \leq x \leq b$$

defines an isometric isomorphism

$$J : \{\lambda \in rca([a,b])\,;\ \lambda(\{a\}) = 0,\ \lambda(\{b\}) = 0\} \to NBV([a,b])\,.$$

Solution. The σ-additivity of λ yields that $f := J\lambda$ is right-continuous. Since $\lambda(\{a\}) = 0$ it follows that $f(a) = 0$, and since $\lambda(\{b\}) = 0$ the σ-additivity gives that $f(x) \to f(b)$ as $x \nearrow b$.

Moreover, for every partitioning $a = a_0 < a_1 < \ldots < a_m = b$,

$$\sum_{i=1}^{m} |f(a_i) - f(a_{i-1})| = \sum_{i=1}^{m} \big|\lambda\big(]a_{i-1}, a_i]\big)\big| \leq \|\lambda\|_{\mathrm{var}}\,,$$

i.e. $\|f\|_{\widetilde{BV}} \leq \|\lambda\|_{\mathrm{var}}$.

In addition, J is injective. In order to prove surjectivity, we use the previous exercise, which for a given $f \in NBV([a,b])$ yields a $\lambda \in rca([a,b])$, for which $\|\lambda\|_{\text{var}} \le \text{var}(f, [a,b]) = \|f\|_{\widetilde{BV}}$. It follows from E6.10(2) that $J\lambda = f$, since for $a \le x < b$

$$(J\lambda)(x) = \lambda([a,x]) = \lim_{\varepsilon \searrow 0} \big(f(x+\varepsilon) - f(a)\big) = f(x)$$

and also $(J\lambda)(b) = \lambda([a,b]) = \lim_{\varepsilon \searrow 0} \lambda([a, b-\varepsilon]) = f(b)$. □

A6 Results from measure theory

The purpose of this appendix is to complete the proof of the representation theorem 6.23 (see A6.6). The necessary construction of regular measures can be found in A6.3.

Subsequently, we also present versions of Luzin's theorem (see A6.7) and Fubini's theorem (see A6.10).

In the following two results, S is an arbitrary set.

A6.1 Jordan decomposition. Let $\mathcal{B}$ be a ring of subsets of the set S and let $\lambda : \mathcal{B} \to \mathbb{R}$ be additive and bounded. Then

$$\lambda^+ := \tfrac{1}{2}(|\lambda| + \lambda)\,, \quad \lambda^- := \tfrac{1}{2}(|\lambda| - \lambda)$$

are additive, bounded and nonnegative on $\mathcal{B}$. It holds that

$$\lambda = \lambda^+ - \lambda^-\,, \quad |\lambda| = \lambda^+ + \lambda^-\,,$$

and, in addition,

$$\lambda^+(E) = \sup_{A \in \mathcal{B}\,:\,A \subset E} \lambda(A) \quad \text{and} \quad \lambda^-(E) = -\inf_{A \in \mathcal{B}\,:\,A \subset E} \lambda(A)\,.$$

Proof. On recalling 6.10, we only need to show that the last identity holds for λ^+.

If $A \subset E$, then $|\lambda|(A) \ge |\lambda(A)|$, and so

$$\lambda^+(E) \ge \lambda^+(A) \ge \tfrac{1}{2}\big(|\lambda(A)| + \lambda(A)\big) \ge \lambda(A)\,.$$

Now for a given $\varepsilon > 0$ choose disjoint sets $E_1, \ldots, E_m$ with $E_i \subset E$ and

$$|\lambda|(E) \le \varepsilon + \sum_{i=1}^{m} |\lambda(E_i)|\,.$$

On setting $E_{m+1} := E \setminus \bigcup_{i=1}^m E_i$, we have

$$\lambda(E) = \sum_{i=1}^{m+1} \lambda(E_i)\,,$$

and so

$$\lambda^+(E) = \frac{1}{2}\big(|\lambda|(E) + \lambda(E)\big) \le \frac{\varepsilon}{2} + \frac{1}{2}\sum_{i=1}^{m+1}\big(|\lambda(E_i)| + \lambda(E_i)\big)$$
$$= \frac{\varepsilon}{2} + \sum_{i\,:\,\lambda(E_i)>0} \lambda(E_i) = \frac{\varepsilon}{2} + \lambda\Big(\bigcup_{i\,:\,\lambda(E_i)>0} E_i\Big) \le \frac{\varepsilon}{2} + \sup_{A\in\mathcal{B}\,:\,A\subset E} \lambda(A)\,.$$

□

A6.2 Hahn decomposition. Let $\mathcal{B}$ be a σ-ring on the set S and let $\nu : \mathcal{B} \to \mathbb{R}$ be σ-additive and bounded. Then there exists an $E^+ \in \mathcal{B}$ such that

$$\nu(E \cap E^+) \ge 0 \quad \text{and} \quad \nu(E \setminus E^+) \le 0 \quad \text{for all } E \in \mathcal{B}\,.$$

Proof. We assume that there exists an $E \in \mathcal{B}$ with $\nu(E) > 0$ (otherwise choose $E^+ := \emptyset$). We now want to find an $E^+ \in \mathcal{B}$ such that

$$\nu(E^+) = s_0 := \sup\nolimits_{E\in\mathcal{B}} \nu(E)\,. \tag{A6-1}$$

Such an E^+ satisfies the desired result. To see this, assume that $\nu(E\setminus E^+) > 0$ for some $E \in \mathcal{B}$. Then

$$\nu(E^+ \cup E) = \nu(E^+) + \nu(E \setminus E^+) > \nu(E^+) = s_0\,,$$

which contradicts the definition of s_0. Similarly, if $\nu(E \cap E^+) < 0$ for some $E \in \mathcal{B}$, then

$$\nu(E^+ \setminus E) = \nu(E^+) - \nu(E \cap E^+) > \nu(E^+) = s_0\,,$$

which again contradicts the definition of s_0.

For the construction of E^+, define for $k \in \mathbb{N}$

$$\mathcal{M}_k := \big\{\, E \in \mathcal{B}\,;\ \nu(E) \ge \big(1 - \tfrac{1}{k}\big)s_0 \,\big\}$$

with the partial order

$$E_1 \le E_2 \quad :\Longleftrightarrow \quad \big(\, E_1 \supset E_2 \text{ and } \nu(E_1) < \nu(E_2) \,\big) \text{ or } E_1 = E_2\,.$$

Let $\mathcal{N} \subset \mathcal{M}_k$ be totally ordered and let

$$s := \sup_{E\in\mathcal{N}} \nu(E)\,.$$

Then there exist $E_i \in \mathcal{N}$, $i \in \mathbb{N}$, with

$$\nu(E_i) \le \nu(E_{i+1}) \to s \quad \text{as } i \to \infty\,. \tag{A6-2}$$

As $\mathcal{N}$ is totally ordered, it follows that $E_i \le E_{i+1}$ or $E_{i+1} \le E_i$. If $E_i \le E_{i+1}$ then (A6-2) implies $E_i \supset E_{i+1}$, and if $E_{i+1} \le E_i$ it implies $E_i = E_{i+1}$. Therefore the sets E_i are decreasing and

$$E_0 := \bigcap_{i\in\mathbb{N}} E_i \in \mathcal{M}_k \,, \quad \nu(E_0) = \lim_{i\to\infty} \nu(E_i) = s\,.$$

The found set $E_0 \in \mathcal{M}_k$ is an upper bound of $\mathcal{N}$. This follows from the fact that if $E \in \mathcal{N}$ with $E_0 \leq E$, then $E \subset E_0$ and $\nu(E) > \nu(E_0)$, or $E = E_0$, where the former case contradicts the definition of s, since $\nu(E) > \nu(E_0) = s$, therefore $E = E_0$.

Hence, by Zorn's lemma (see the proof of 6.14), there exists a maximal element $M_k^+ \in \mathcal{M}_k$. It satisfies

$$\nu(M_k^+) \geq \big(1 - \tfrac{1}{k}\big)s_0\,,$$

and in addition it holds for all $A \in \mathcal{B}$ that

$$A \subset M_k^+ \quad \Longrightarrow \quad \nu(A) \geq 0\,. \tag{A6-3}$$

To see this, assume that $\nu(A) < 0$. Then $\nu(M_k^+ \setminus A) > \nu(M_k^+)$, and so $M_k^+ \setminus A \in \mathcal{M}_k$ with $M_k^+ \setminus A \geq M_k^+$. Then the maximality of M_k^+ yields that $M_k^+ \setminus A \leq M_k^+$, a contradiction.

Then the property (A6-3) also holds with M_k^+ replaced by the sets

$$E_k^+ := \bigcup_{j\leq k} M_j^+\,,$$

because if $A \in \mathcal{B}$, $A \subset E_k^+$, then $A_j := A \cap M_j^+ \setminus \bigcup_{i<j} M_i^+ \subset M_j^+$ form a partition of A, and hence

$$\nu(A) = \sum_{j=1}^{k} \nu(A_j) \geq 0\,.$$

In particular,

$$\nu(E_k^+) \geq \nu(M_k^+) \geq \big(1 - \tfrac{1}{k}\big)s_0\,.$$

Hence

$$E^+ := \bigcup_{k\in\mathbb{N}} E_k^+ \in \mathcal{B} \quad \text{with} \quad \nu(E^+) = \lim_{k\to\infty} \nu(E_k^+) = s_0\,.$$

Therefore E^+ satisfies (A6-1). □

In the following, let $S \subset \mathbb{R}^n$ be a closed set and let $\mathcal{B}_0, \mathcal{B}_1$ for S be defined as in 6.20. Furthermore, let $ba(S)$ etc. be the spaces defined in 6.20 and 6.21.

A6.3 Lemma. Let $\lambda \in ba(S)$ be nonnegative and let

$$\mu(E) := \sup_{\substack{A\,:\,A\subset E\\ A \text{ closed}}} \; \inf_{\substack{U\,:\,A\subset U\\ U \text{ open}}} \lambda(U) \quad \text{for } E \in \mathcal{B}_0.$$

Then $\mu \in rba(S)$ and

$$\int_S f \,\mathrm{d}\mu = \int_S f \,\mathrm{d}\lambda \quad \text{for all } f \in C^0(S).$$

Proof. (All occurring sets are in $\mathcal{B}_0$.) μ is nonnegative and monotone, i.e. $E_1 \subset E_2$ implies that $\mu(E_1) \le \mu(E_2)$. For closed sets A

$$\mu(A) = \inf_{\substack{U\,:\,A \subset U \\ U \text{ open}}} \lambda(U)\,, \qquad \text{and so} \qquad \mu(E) = \sup_{\substack{A\,:\,A \subset E \\ A \text{ closed}}} \mu(A) \tag{A6-4}$$

for all E. Define

$$\mathcal{M} := \{B \in \mathcal{B}_0\,;\ \mu(E) = \mu(E \cap B) + \mu(E \setminus B) \text{ for all } E \in \mathcal{B}_0\}\,.$$

We want to show that $\mathcal{M} = \mathcal{B}_0$. Obviously, $\emptyset, S \in \mathcal{M}$ and from $B \in \mathcal{M}$ it follows that $S \setminus B \in \mathcal{M}$. If $A, B \in \mathcal{M}$, then it follows that for all $E \in \mathcal{B}_0$

$$\begin{aligned}
&\mu\big(E \cap (A \cap B)\big) + \mu\big(E \setminus (A \cap B)\big) \\
&= \mu\big(\underbrace{E \cap (A \cap B)}_{=(E\cap B)\cap A}\big) + \mu\big(\underbrace{(E \setminus (A \cap B)) \cap B}_{=(E\cap B)\setminus A}\big) + \mu\big(\underbrace{(E \setminus (A \cap B)) \setminus B}_{=E\setminus B}\big) \\
&= \mu(E \cap B) + \mu(E \setminus B) = \mu(E)\,,
\end{aligned}$$

and so $A \cap B \in \mathcal{M}$. Hence $\mathcal{M}$ is a Boolean algebra. It remains to show that $\mathcal{M}$ contains the closed sets. If A_1, A_2 are closed and disjoint, then there exist open disjoint sets U_i with $A_i \subset U_i$. Then it holds for every open set $U \supset A_1 \cup A_2$ that

$$\lambda(U) \ge \lambda\big(U \cap (U_1 \cup U_2)\big) = \lambda(U \cap U_1) + \lambda(U \cap U_2) \ge \mu(A_1) + \mu(A_2)\,,$$

and combining with (A6-4) yields that

$$\mu(A_1 \cup A_2) \ge \mu(A_1) + \mu(A_2)\,.$$

Now let B be closed and let E be arbitrary. Then if $A_1 \subset E \cap B$, $A_2 \subset E \setminus B$ are closed sets,

$$\mu(A_1) + \mu(A_2) \le \mu(A_1 \cup A_2) \le \mu(E)\,,$$

and so (A6-4) implies that

$$\mu(E \cap B) + \mu(E \setminus B) \le \mu(E)\,.$$

On the other hand, if $A \subset E$ is closed and U_1, U_2 are open with $A \cap B \subset U_1$ and $A \setminus U_1 \subset U_2$, then $A \subset U_1 \cup U_2$, and hence

$$\lambda(U_1) + \lambda(U_2) \ge \lambda(U_1 \cup U_2) \ge \mu(A)\,.$$

Taking the infimum over all U_2, and noting that $A \setminus U_1$ is closed, we obtain

$$\lambda(U_1) + \mu(A \setminus U_1) \ge \mu(A)\,.$$

Since $A \setminus U_1$ is a closed subset of $E \setminus B$, it follows that

$$\lambda(U_1) + \mu(E \setminus B) \ge \mu(A)\,.$$

Now noting that $A \cap B$ is closed, and taking the infimum over all U_1, we obtain

$$\mu(A \cap B) + \mu(E \setminus B) \geq \mu(A)\,,$$

and so, since $A \cap B$ is a closed subset of $E \cap B$,

$$\mu(E \cap B) + \mu(E \setminus B) \geq \mu(A)\,.$$

On taking the supremum over all A, it finally follows that

$$\mu(E \cap B) + \mu(E \setminus B) \geq \mu(E)\,.$$

This shows that $B \in \mathcal{M}$, and hence $\mathcal{M} = \mathcal{B}_0$.

It follows that μ is additive on $\mathcal{M}$, for if $E_1, E_2 \in \mathcal{M}$ are disjoint, then it holds for all E that

$$\mu(E) = \mu(E \cap E_1) + \mu(E \setminus E_1)\,,$$

and for $E = E_1 \cup E_2$ we obtain that

$$\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)\,.$$

Moreover, μ is regular, because for E and $\varepsilon > 0$ there exist closed sets $A_1 \subset E$ and $A_2 \subset S \setminus E$ with

$$\mu(E) \leq \mu(A_1) + \varepsilon \quad \text{and} \quad \mu(S \setminus E) \leq \mu(A_2) + \varepsilon\,.$$

Then $A_1 \subset E \subset S \setminus A_2$ and, on recalling that $|\mu| = \mu$, it follows that

$$|\mu|\big((S \setminus A_2) \setminus A_1\big) \leq 2\varepsilon\,.$$

It remains to show that the integral identity holds. Without loss of generality let $0 \leq f \leq 1$. For $n \in \mathbb{N}$ define

$$E_i := \left\{\tfrac{i}{n} \leq f < \tfrac{i+1}{n}\right\} \in \mathcal{B}_0 \qquad \text{for } i = 0, \ldots, n.$$

For a given $\varepsilon > 0$ choose $A_i \subset E_i$ closed with $\mu(E_i \setminus A_i) \leq \varepsilon$. Since the A_i are disjoint and f is continuous, there exist disjoint open sets U_i with

$$A_i \subset U_i \quad \text{and} \quad \inf_{U_i} f \geq \frac{i}{n} - \varepsilon\,.$$

As $\mu(A_i) \leq \lambda(U_i)$, it follows that

$$\begin{aligned}
\int_S f \,\mathrm{d}\mu &\leq \sum_i \frac{i+1}{n}\mu(E_i) \leq \frac{1}{n}\mu(S) + \sum_i \frac{i}{n}\mu(E_i) \\
&\leq \frac{\mu(S)}{n} + n\varepsilon + \sum_i \frac{i}{n}\lambda(U_i) \\
&\leq \underbrace{\frac{\mu(S)}{n}}_{\to 0 \text{ as } n \to \infty} + \underbrace{n\varepsilon + \varepsilon\lambda(S)}_{\substack{\to 0 \text{ as } \varepsilon \to 0 \\ \text{for any } n}} + \int_S f \,\mathrm{d}\lambda\,.
\end{aligned}$$

Replacing f by $1-f$ yields, on noting that $\mu(S)=\lambda(S)$, that

$$\mu(S)-\int_S f\,\mathrm{d}\mu=\int_S(1-f)\,\mathrm{d}\mu\le\int_S(1-f)\,\mathrm{d}\lambda=\lambda(S)-\int_S f\,\mathrm{d}\lambda\,,$$

and hence the desired result. □

A6.4 Corollary. For $\lambda\in ba(S)$ there exists a $\nu\in rba(S)$ such that

$$\int_S f\,\mathrm{d}\lambda=\int_S f\,\mathrm{d}\nu\qquad\text{for all } f\in C^0(S).$$

Proof. Since we can split λ into a real and an imaginary part, we may assume without loss of generality that λ is real-valued. Let $\lambda=\lambda^+-\lambda^-$ be the Jordan decomposition of λ and let $\mu^\pm$ be the measures from A6.3 corresponding to $\lambda^\pm$. Set $\nu:=\mu^+-\mu^-$. It obviously holds that $|\nu|\le\mu^++\mu^-$, and so the regularity of $\mu^\pm$ implies that ν is regular. □

A6.5 Lemma (Alexandrov). If $S\subset\mathbb{R}^n$ is compact, then

$$\nu\in rba(S)\quad\Longrightarrow\quad\nu\text{ is }\sigma\text{-additive (on }\mathcal{B}_0\text{ !).}$$

Proof (Compare A3.3*).* Let $E_i\in\mathcal{B}_0$, $i\in\mathbb{N}$, be disjoint and let $E:=\bigcup_i E_i\in\mathcal{B}_0$. As ν is regular, we can choose for $\varepsilon>0$ a closed set A with $A\subset E$ and $|\nu|(E\setminus A)\le\varepsilon$ and open sets U_i with $E_i\subset U_i$ and $|\nu|(U_i\setminus E_i)\le\varepsilon 2^{-i}$. On noting that $(U_i)_{i\in\mathbb{N}}$ is a cover of A with A being compact, we see that

$$A\subset\bigcup_{i=1}^m U_i\quad\text{for an } m,$$

and hence, since $|\nu|$ is nonnegative and additive (see 6.10), that

$$|\nu|(E)\le\varepsilon+|\nu|(A)\le\varepsilon+\sum_{i=1}^m|\nu|(U_i)\le\varepsilon+\varepsilon\sum_{i=1}^\infty 2^{-i}+\sum_{i=1}^\infty|\nu|(E_i)\,.$$

In addition, for all m

$$|\nu|(E)\ge|\nu|\Big(\bigcup_{i=1}^m E_i\Big)=\sum_{i=1}^m|\nu|(E_i)\,,$$

which proves that

$$|\nu|(E)=\sum_{i=1}^\infty|\nu|(E_i)\,.$$

Similarly, for all m

$$|\nu|\Big(\bigcup_{i>m}E_i\Big)=\sum_{i>m}|\nu|(E_i)\longrightarrow 0\qquad\text{as } m\to\infty.$$

We conclude that

$$\Big|\nu(E) - \sum_{i=1}^{m} \nu(E_i)\Big| = \Big|\nu\big(E \setminus \bigcup_{i\le m} E_i\big)\Big|$$
$$= \Big|\nu\big(\bigcup_{i>m} E_i\big)\Big| \le |\nu|\big(\bigcup_{i>m} E_i\big) \longrightarrow 0 \quad \text{as } m \to \infty.$$

□

A6.6 Lemma. Let $S \subset \mathbb{R}^n$ be compact. For $\lambda \in ba(S)$ there exists a $\nu \in rca(S)$ with

$$\int_S f \,\mathrm{d}\nu = \int_S f \,\mathrm{d}\lambda \quad \text{for all } f \in C^0(S).$$

Proof. We may assume without loss of generality that λ is real-valued and nonnegative (see the proof of A6.4). Let $\mu \in rba(S)$ be the measure corresponding to λ as in A6.3. It follows from lemma A6.5 that μ is σ-additive on $\mathcal{B}_0$. Then by A3.15 there exists an extension of $(\mathcal{B}_0, \mu)$ to $(\mathcal{B}, \bar{\mu})$ with a σ-algebra $\mathcal{B}$ and a σ-additive measure $\bar{\mu}$ on $\mathcal{B}$. As $\mathcal{B}_1$ is the smallest σ-algebra that contains $\mathcal{B}_0$, it follows that $\mathcal{B}_1 \subset \mathcal{B}$. Hence $\bar{\mu}$ is σ-additive on $\mathcal{B}_1$.

We now show that $\bar{\mu}$ is also regular. To this end, let

$$\mathcal{M} := \{E \in \mathcal{B}_1\,;\ \text{For } \varepsilon > 0 \text{ there exist sets } A \text{ and } U \text{ with}$$
$$A \subset E \subset U,\ A \text{ closed}, U \text{ open}, \bar{\mu}(U \setminus A) \le \varepsilon\}\,.$$

Clearly $\mathcal{M}$ is an algebra, and since $\bar{\mu}$ is an extension of μ, it holds that $\mathcal{B}_0 \subset \mathcal{M}$. Then it follows that $\mathcal{M} = \mathcal{B}_1$, if we can show that

$$E_i \in \mathcal{M} \text{ for } i \in \mathbb{N} \text{ with } E_i \subset E_{i+1} \quad \Longrightarrow \quad E := \bigcup_{i \in \mathbb{N}} E_i \in \mathcal{M}\,.$$

To this end, choose a closed set A_i with $A_i \subset E_i$ and an open set U_i with $E_i \subset U_i$ such that $\bar{\mu}(U_i \setminus A_i) \le \varepsilon 2^{-i}$. Then

$$\bigcup_{i \le m} A_i \subset E \subset \bigcup_{i \in \mathbb{N}} U_i =: U$$

and

$$\bar{\mu}\big(U \setminus \bigcup_{i\le m} A_i\big) \le \bar{\mu}\big(U \setminus \bigcup_{i\le m} U_i\big) + \bar{\mu}\big(\bigcup_{i\le m} U_i \setminus \bigcup_{i\le m} A_i\big)\,.$$

The first term is smaller than ε, if we choose m sufficiently large, and the second term is

$$\le \bar{\mu}\big(\bigcup_{i\le m} (U_i \setminus A_i)\big) \le \sum_{i\le m} \bar{\mu}(U_i \setminus A_i) \le \varepsilon\,.$$

The integral identity follows as in the proof of A6.3. □

We present the following result on measurable functions. Here S can be replaced with any compact topological space.

A6.7 Luzin's theorem. Let $S \subset \mathbb{R}^n$ be compact, $\mu \in rca(S)$ be nonnegative, and Y be a Banach space. Then every μ-measurable function $f : S \to Y$ is ***μ-almost continuous***, i.e. for every μ-measurable set E and every $\varepsilon > 0$ there exists a compact set $K \subset E$ with $\mu(E \setminus K) \le \varepsilon$ such that $f\big|_K$ is a continuous function on K.

Proof. First we recall that for every μ-measurable set E there exist an $\widetilde{E} \in \mathcal{B}_1$ and a μ-null set N with $E \setminus N = \widetilde{E} \setminus N$ (see A3.14(2)). Moreover, for every μ-null set N and every $\varepsilon > 0$ there exists an $N_\varepsilon \in \mathcal{B}_1$ with $N \subset N_\varepsilon$ and $\mu(N_\varepsilon) \le \varepsilon$ (see A3.4). As μ is regular, there exist a compact set $\widetilde{K} \subset \widetilde{E}$ and an open set $\widetilde{U} \supset \widetilde{E}$ with $\mu(\widetilde{U} \setminus \widetilde{K}) \le \varepsilon$, as well as an open set $V \supset N_\varepsilon$ with $\mu(V) \le 2\varepsilon$. Then $K := \widetilde{K} \setminus V \subset E$ is compact and $U := \widetilde{U} \cup V \supset E$ is open with $\mu(U \setminus K) \le 3\varepsilon$.

There exists a μ-null set N such that $f(S \setminus N)$ is separable (see 3.11(2)). Choose a countable dense subset $\{y_j \,;\, j \in \mathbb{N}\}$ of $f(S \setminus N)$. For every i it holds that the sets $\mathrm{B}_{\frac{1}{i}}(y_j)$, $j \in \mathbb{N}$, form a cover of $f(S \setminus N)$, and hence also

$$B_{ij} := \mathrm{B}_{\frac{1}{i}}(y_j) \setminus \textstyle\bigcup_{k<j} \mathrm{B}_{\frac{1}{i}}(y_k) \ .$$

This implies that

$$E_{ij} := E \cap f^{-1}(B_{ij}) \setminus N \quad \text{ for } j \in \mathbb{N}$$

form a disjoint partitioning of $E \setminus N$ into μ-measurable sets. It follows from the remark at the beginning of the proof that there exist compact sets $K_{ij} \subset E_{ij}$ with $\mu(E_{ij} \setminus K_{ij}) \le \varepsilon 2^{-i-j-1}$. Consequently, $\mu\big(E \setminus \bigcup_j K_{ij}\big) \le \varepsilon 2^{-i-1}$, and hence there exists a j_i with

$$\mu(E \setminus K_i) \le \varepsilon 2^{-i}\,, \qquad \text{where} \qquad K_i := \textstyle\bigcup_{j \le j_i} K_{ij} \ .$$

K_i is a compact subset of $E \setminus N$, and by construction it is the disjoint union of the compact sets K_{ij} for $j \le j_i$. Hence

$$g_i(x) := y_j \quad \text{for } x \in K_{ij} \ (\text{if } K_{ij} \ne \emptyset)$$

defines a $g_i \in C^0(K_i;Y)$ with

$$\sup_{x \in K_i} \|g_i(x) - f(x)\|_Y \le \frac{1}{i} \ .$$

Set $K := \bigcap_i K_i$. Then the functions $g_i\big|_K \in C^0(K;Y)$, and on K they converge uniformly to f as $i \to \infty$, which yields that $f\big|_K \in C^0(K;Y)$. In addition, K is a compact subset of E and

$$\mu(E \setminus K) \le \sum_{i \in \mathbb{N}} \mu(E \setminus K_i) \le \varepsilon \ .$$

□

We add now a functional analysis formulation of Fubini's theorem, where we restrict ourselves to the case of bounded regular measures.

A6.8 Product measure. Let $S^l \subset \mathbb{R}^{n_l}$ be compact, $l = 1, 2$, and let $(S^l, \mathcal{B}^l, \mu^l)$ be measure spaces. Let $\mathcal{B}^l$ contain the Borel sets of S^l and let $\mu^l \in rca(S^l)$. Define

$$\begin{aligned}&\mathcal{B}^1 \times \mathcal{B}^2 := \{E^1 \times E^2\,;\; E^1 \in \mathcal{B}^1 \text{ and } E^2 \in \mathcal{B}^2\}\,,\\&(\mu^1 \times \mu^2)(E^1 \times E^2) := \mu^1(E^1) \cdot \mu^2(E^2) \quad \text{for } E^1 \times E^2 \in \mathcal{B}^1 \times \mathcal{B}^2\,.\end{aligned}$$

Denote by $\mathcal{B}^0$ the Boolean algebra induced by $\mathcal{B}^1 \times \mathcal{B}^2$. Then $\mathcal{B}^0$ consists of finite disjoint unions of sets in $\mathcal{B}^1 \times \mathcal{B}^2$, and $\mu^1 \times \mu^2$ can be canonically extended to an additive measure on $\mathcal{B}^0$.

Proposition: $\mu^1 \times \mu^2$ is σ-subadditive on $\mathcal{B}^0$, so that all the properties in A3.1 are satisfied.

Proof of proposition. Let $E, E_i \in \mathcal{B}^0$, $i \in \mathbb{N}$, with $E \subset \bigcup_{i\in\mathbb{N}} E_i$. We have to show that for $\mu := \mu^1 \times \mu^2$ it holds that

$$\mu(E) \le \sum_{i\in\mathbb{N}} \mu(E_i)\,.$$

By the definitions of $\mathcal{B}^0$ and μ, we may assume that

$$E_i = E_i^1 \times E_i^2 \in \mathcal{B}^1 \times \mathcal{B}^2\,.$$

As the μ^l are regular, it follows that for $\varepsilon > 0$ there exist open sets $U_i^l \in \mathcal{B}^l$ with (see the beginning of the proof in A6.7)

$$E_i^l \subset U_i^l \quad \text{and} \quad \mu^l(U_i^l \setminus E_i^l) \le \varepsilon 2^{-i}\,.$$

Then

$$\begin{aligned}\mu(U_i^1 \times U_i^2) &\le \mu(E_i^1 \times E_i^2) + \mu\big((U_i^1 \setminus E_i^1) \times U_i^2\big) + \mu\big(E_i^1 \times (U_i^2 \setminus E_i^2)\big)\\&\le \mu(E_i^1 \times E_i^2) + \mu^1(U_i^1 \setminus E_i^1)\mu^2(S^2) + \mu^1(S^1)\mu^2(U_i^2 \setminus E_i^2)\\&\le \mu(E_i^1 \times E_i^2) + C2^{-i}\varepsilon \quad \text{with } C := \mu^1(S^1) + \mu^2(S^2)\,.\end{aligned}$$

Similarly, there exists a compact set $K \in \mathcal{B}^0$ with

$$K \subset E \quad \text{and} \quad \mu(E) \le \mu(K) + \varepsilon\,.$$

(E is the disjoint union of elements in $\mathcal{B}^1 \times \mathcal{B}^2$, and each of these subsets can be approximated in measure by compact subsets to an arbitrary accuracy. K is then the disjoint union of Cartesian products of compact sets.) Since the sets $U_1^l \times U_2^l$ form a cover of the set K, there exists an m with

$$K \subset \bigcup_{i=1}^{m} U_i^1 \times U_i^2\,,$$

and hence

$$\mu(E) \leq \mu(K) + \varepsilon \leq \sum_{i=1}^{m} \mu(U_i^1 \times U_i^2) + \varepsilon \leq \sum_{i=1}^{m} \mu(E_i^1 \times E_i^2) + (C+1)\varepsilon \,.$$

□

Therefore the Lebesgue integral for $(S^1 \times S^2, \mathcal{B}^0, \mu^1 \times \mu^2)$ can be constructed as in Appendix A3. In particular, there exists a measure extension to a measure space $(S^1 \times S^2, \mathcal{B}, \mu^1 \times \mu^2)$. We now characterize the Lebesgue space $L^p(\mu^1 \times \mu^2; Y)$ with the help of iterated integration. But first we consider the following special case:

A6.9 Lemma. If N is a $\mu^1 \times \mu^2$-null set, then for μ^1-almost all $x_1 \in S^1$

$$\{x_2 \in S^2 \,;\; (x_1, x_2) \in N\}$$

is a μ^2-null set.

Proof. It follows from the definition of null sets in A3.4 that for $\varepsilon > 0$ there exist sets $E_i^l \in \mathcal{B}^l$, $i \in \mathbb{N}$, $l = 1, 2$ with

$$N \subset \bigcup_{i \in \mathbb{N}} E_i^1 \times E_i^2 \quad \text{and} \quad \sum_{i \in \mathbb{N}} \mu(E_i^1 \times E_i^2) \leq \varepsilon \,,$$

where $\mu := \mu^1 \times \mu^2$. Consider the functions

$$g_{\varepsilon n}(x_1)(x_2) := \sum_{i \leq n} \mathcal{X}_{E_i^1}(x_1) \mathcal{X}_{E_i^2}(x_2) \,.$$

For all x_1 we have that $g_{\varepsilon n}(x_1) \in L^1(\mu^2)$ satisfying the following equation

$$G_{\varepsilon n}(x_1) := \int_{S^2} g_{\varepsilon n}(x_1) \, \mathrm{d}\mu^2 = \sum_{i \leq n} \mathcal{X}_{E_i^1}(x_1) \mu^2(E_i^2) \,.$$

The function $G_{\varepsilon n} \in L^1(\mu^1)$ with

$$\int_{S^1} G_{\varepsilon n} \, \mathrm{d}\mu^1 = \sum_{i \leq n} \mu^1(E_i^1) \mu^2(E_i^2) \leq \varepsilon \,.$$

On noting that

$$G_{\varepsilon n}(x_1) \nearrow G_\varepsilon(x_1) := \sum_{i \in \mathbb{N}} \mathcal{X}_{E_i^1}(x_1) \mu^2(E_i^2)$$

as $n \nearrow \infty$, it follows from the monotone convergence theorem (see A3.12(3)) that $G_\varepsilon \in L^1(\mu^1)$ with

$$\int_{S^1} G_\varepsilon \, \mathrm{d}\mu^1 = \lim_{n\to\infty} \int_{S^1} G_{\varepsilon n} \, \mathrm{d}\mu^1 \le \varepsilon \, .$$

But this means that $G_\varepsilon \to 0$ in $L^1(\mu^1)$ as $\varepsilon \to 0$. Hence there exists a subsequence $\varepsilon \to 0$ such that $G_\varepsilon(x_1) \to 0$ for μ^1-almost all $x_1 \in S^1$. In the following we consider such x_1. On noting that for small ε and as $n \nearrow \infty$ we have that

$$\int_{S^2} g_{\varepsilon n}(x_1) \, \mathrm{d}\mu^2 = G_{\varepsilon n}(x_1) \nearrow G_\varepsilon(x_1) < \infty$$

and

$$g_{\varepsilon n}(x_1)(x_2) \nearrow g_\varepsilon(x_1)(x_2) := \sum_{i\in\mathbb{N}} \mathcal{X}_{E_i^1}(x_1)\mathcal{X}_{E_i^2}(x_2) \, ,$$

it follows once again from the monotone convergence theorem that the function $g_\varepsilon(x_1) \in L^1(\mu^2)$ satisfies

$$\int_{S^2} g_\varepsilon(x_1) \, \mathrm{d}\mu^2 = G_\varepsilon(x_1) \, .$$

Therefore, $g_\varepsilon(x_1) \to 0$ in $L^1(\mu^2)$, and so there exists a subsequence $\varepsilon \to 0$ (depending on x_1!) with $g_\varepsilon(x_1)(x_2) \to 0$ for μ^2-almost all $x_2 \in S^2$. But noting that $g_\varepsilon(x_1)(x_2) \ge \mathcal{X}_N(x_1, x_2)$ implies that $\mathcal{X}_N(x_1, x_2) = 0$ for μ^2-almost all $x_2 \in S^2$. ☐

A6.10 Fubini's theorem. Let Y be a Banach space and let $1 \le p < \infty$. Consider the product measure in A6.8. Then

$$(Jf)(x_1)(x_2) := f(x_1, x_2)$$

defines a linear isometric isomorphism

$$J : L^p(\mu^1 \times \mu^2; Y) \longrightarrow L^p(\mu^1; L^p(\mu^2; Y)) \, .$$

In particular, for $f \in L^p(\mu^1 \times \mu^2; Y)$ there exists

$$F(x_1) := \int_{S^2} f(x_1, x_2) \, \mathrm{d}\mu^2(x_2) \quad \text{for } \mu^1\text{-almost all } x_1 \in S^1$$

and $F \in L^p(\mu^1; Y)$ with

$$\int_{S^1} F(x_1) \, \mathrm{d}\mu^1(x_1) = \int_{S^1\times S^2} f(x_1, x_2) \, \mathrm{d}(\mu^1 \times \mu^2)(x_1, x_2) \, .$$

A symmetry argument then yields that

$$\int_{S_1} \Big(\int_{S_2} f(x_1, x_2) \, \mathrm{d}\mu^2(x_2) \Big) \, \mathrm{d}\mu^1(x_1) = \int_{S_2} \Big(\int_{S_1} f(x_1, x_2) \, \mathrm{d}\mu^1(x_1) \Big) \, \mathrm{d}\mu^2(x_2) \, .$$

Proof. Let $f \in L^p(\mu^1 \times \mu^2; Y)$ (we suppress in the following proof the argument Y). Since $p < \infty$, it follows from the construction of the Lebesgue integral (see the proof of 3.26(1)) that f can be approximated by step functions

$$f_k = \sum_{i=1}^{n} \mathcal{X}_{E_i} \alpha_i \quad \text{with } E_i \in \mathcal{B}^0 \text{ and } \alpha_i \in Y,$$

where n, E_i and α_i depend on k. The definition of $\mathcal{B}^0$ then yields that f_k can also be represented as

$$f_k = \sum_{i,j=1}^{n} \mathcal{X}_{E_i^1 \times E_j^2} \alpha_{ij} \quad \text{with } E_i^1 \in \mathcal{B}^1,\ E_j^2 \in \mathcal{B}^2,\ \alpha_{ij} \in Y$$

with a new n, where both the E_i^1 and the E_j^2 are disjoint. Then for all x_1

$$(Jf_k)(x_1) = \sum_{i,j=1}^{n} \mathcal{X}_{E_i^1}(x_1) \mathcal{X}_{E_j^2} \alpha_{ij} \in L^p(\mu^2)\,,$$

and $Jf_k \in L^p(\mu^1; L^p(\mu^2))$, with

$$\begin{aligned} \int_{S^1} \| Jf_k \|_{L^p(\mu^2)}^p \, \mathrm{d}\mu^1 &= \sum_{i=1}^{n} \mu^1(E_i^1) \left\| \sum_{j=1}^{n} \mathcal{X}_{E_j^2} \alpha_{ij} \right\|_{L^p(\mu^2)}^p \\ &= \sum_{i,j=1}^{n} \mu^1(E_i^1) \mu^2(E_j^2) \| \alpha_{ij} \|_Y^p = \int_{S^1 \times S^2} \| f_k \|_Y^p \, \mathrm{d}\mu\,, \end{aligned}$$

where $\mu := \mu^1 \times \mu^2$. Similarly, we observe that

$$\int_{S^2} (Jf_k)(x_1) \, \mathrm{d}\mu^2$$

as a function of x_1 lies in $L^1(\mu^1)$ and satisfies

$$\int_{S^1} \Big(\int_{S^2} (Jf_k)(x_1) \, \mathrm{d}\mu^2 \Big) \, \mathrm{d}\mu^1(x_1) = \int_{S^1 \times S^2} f_k \, \mathrm{d}\mu\,.$$

These properties, which we have derived for f_k, are of course also valid for the step functions $f_k - f_l$. Therefore,

$$\| Jf_k - Jf_l \|_{L^p(\mu^1; L^p(\mu^2))} = \| f_k - f_l \|_{L^p(\mu)} \to 0 \quad \text{as } k, l \to \infty.$$

By completeness of $L^p(\mu^1; L^p(\mu^2))$, there exists an F such that

$$Jf_k \to F \quad \text{in } L^p(\mu^1; L^p(\mu^2)) \text{ as } k \to \infty.$$

Hence there exists a subsequence such that for μ^1-almost all x_1

$$Jf_k(x_1) \to F(x_1) \quad \text{in } L^p(\mu^2).$$

On the other hand, since $f_k \to f$ in $L^p(\mu)$, $\mu = \mu^1 \times \mu^2$, there exists a subsequence such that

$$f_k(x_1, x_2) \to f(x_1, x_2) \quad \text{for } \mu\text{-almost all } (x_1, x_2).$$

It follows from A6.9 that then for μ^1-almost all x_1

$$f_k(x_1, x_2) \to f(x_1, x_2) \quad \text{for } \mu^2\text{-almost all } x_2.$$

On recalling that $f_k(x_1, x_2) = Jf_k(x_1)(x_2)$, we then obtain that

$$F(x_1) = f(x_1, \bullet) \quad \text{in } L^p(\mu^2)$$

for μ^1-almost all x_1, i.e. $F = Jf$. In addition, it follows from the above that

$$\|Jf\|_{L^p(\mu^1;L^p(\mu^2))} = \|f\|_{L^p(\mu)}\,.$$

This shows that J is well defined and isometric. Consequently the image of J is closed. Hence, in order to show the surjectivity, it is sufficient to show that the image is dense. Every element in $L^p(\mu^1; L^p(\mu^2))$ can be approximated by linear combinations of functions $\mathcal{X}_{E^1} g$ with $E^1 \in \mathcal{B}^1$ and $g \in L^p(\mu^2)$, and similarly g can be approximated by linear combinations of $\mathcal{X}_{E^2}\alpha$ with $E^2 \in \mathcal{B}^2$ and $\alpha \in Y$. But functions $F(x_1)(x_2) = \mathcal{X}_{E^1}(x_1)\mathcal{X}_{E^2}(x_2)\alpha$ in $L^p(\mu^1; L^p(\mu^2))$ clearly lie in the image of J.

In order to prove the integral formula, we exploit the fact that the integral with respect to μ^2 is a linear continuous map from $L^1(\mu^2)$ to Y. If $f \in L^1(\mu)$, then $Jf \in L^1(\mu^1; L^1(\mu^2))$, and hence (see theorem 5.11)

$$x_1 \longmapsto \int_{S^2} Jf(x_1)\,\mathrm{d}\mu^2$$

is a function in $L^1(\mu^1)$. On noting that in addition $Jf_k \to Jf$ in $L^1(\mu^1; L^1(\mu^2))$ as $k \to \infty$, if the f_k are chosen as above, we obtain with the help of 5.11 that as $k \to \infty$

$$\begin{aligned}
&\int_{S^1}\Big(\int_{S^2} Jf(x_1)\,\mathrm{d}\mu^2\Big)\,\mathrm{d}\mu^1(x_1) = \int_{S^2}\Big(\int_{S^1} Jf\,\mathrm{d}\mu^1\Big)\,\mathrm{d}\mu^2 \\
&\longleftarrow \int_{S^2}\Big(\int_{S^1} Jf_k\,\mathrm{d}\mu^1\Big)\,\mathrm{d}\mu^2 = \int_{S^1}\Big(\int_{S^2} Jf_k(x_1)\,\mathrm{d}\mu^2\Big)\,\mathrm{d}\mu^1(x_1) \\
&= \int_{S^1\times S^2} f_k\,\mathrm{d}\mu \longrightarrow \int_{S^1\times S^2} f\,\mathrm{d}\mu\,.
\end{aligned}$$

□

A6.11 Remark on the case $p = \infty$. With the above assumptions, let $f \in L^\infty(\mu^1 \times \mu^2; Y)$. Then $f \in L^q(\mu^1 \times \mu^2; Y)$ for every $1 \le q < \infty$, so that the result shown in A6.10 yields that

$$Jf \in \bigcap_{1 \le q < \infty} L^q(\mu^1; L^q(\mu^2; Y)) \,.$$

Moreover, it follows easily from E3.4 and A6.9 that

$$\| f \|_{L^\infty(\mu^1 \times \mu^2)} = \| g \|_{L^\infty(\mu^1)} \,,$$

where $g(x_1) := \| f(x_1, \bullet) \|_{L^\infty(\mu^2)} = \| (Jf)(x_1) \|_{L^\infty(\mu^2)}$.

However, in general Jf is not (!) an element of $L^\infty(\mu^1; L^\infty(\mu^2; Y))$, as can be seen from the example $\mu^1 = \mu^2 = \mathrm{L}^1 \llcorner [0,1]$, $Y = \mathbb{R}$, $f = \mathcal{X}_E$, $E := \{(x_1, x_2)\,;\; x_1 < x_2\}$. In this case the function

$$x_1 \mapsto \mathcal{X}_{[x_1,1]} \in L^\infty(\mu^2; Y)$$

is not μ^1-measurable.