8 Weak convergence

In many cases the concept of convergence with respect to the norm turns out to be too restrictive. That is why in this chapter we will introduce a weaker notion of convergence which will enable us to solve minimum problems under far weaker assumptions.

In 4.3 we proved the projection theorem in Hilbert spaces and noted subsequently that the same result cannot be expected to hold in general Banach spaces. The difficulty lies in finding a convergent subsequence within a given minimal sequence, something that is in general not possible with respect to the norm convergence, as balls in infinite-dimensional spaces are not precompact (see 4.10). However, we will see (in 8.10) that closed balls are sequentially compact with respect to weak convergence, at least for the class of reflexive spaces (see 8.8). Here we lose the continuity of the norm, but we nonetheless retain its lower semicontinuity (see 8.3(4)). This property will play a crucial role in the proofs of the existence results 8.15 and 8.17. Hence the class of reflexive spaces, which lies between the class of Hilbert spaces and the class of general Banach spaces, plays a significant role in applications.

In this chapter all the spaces are assumed to be complete, except in 8.12-8.14. In the following, we will always use the notation $\langle x\,,\,x'\rangle_X := x'(x)$ for $x \in X$ and $x' \in X'$ from 7.4. We will also write $\langle x\,,\,x'\rangle := \langle x\,,\,x'\rangle_X$. This simple notation is used in the case when only one Banach space X is involved.

8.1 Definition (weak convergence). Let X be a Banach space.

(1) A sequence $(x_k)_{k\in\mathbb{N}}$ in X ***converges weakly*** to $x \in X$ (we write $x_k \to x$ weakly in X as $k \to \infty$, or $x_k \rightharpoonup x$ as $k \to \infty$) if for all $x' \in X'$

$$\langle x_k\,,\,x'\rangle_X \to \langle x\,,\,x'\rangle_X \qquad \text{as } k \to \infty.$$

(2) A sequence $(x'_k)_{k\in\mathbb{N}}$ in X' ***converges weakly**** to $x' \in X'$ (we write $x'_k \to x'$ weakly* in X' as $k \to \infty$, or $x'_k \overset{*}{\rightharpoonup} x'$ as $k \to \infty$) if for all $x \in X$

$$\langle x\,,\,x'_k\rangle_X \to \langle x\,,\,x'\rangle_X \qquad \text{as } k \to \infty.$$

(3) Analogously to (1) and (2) we define weak and weak* Cauchy sequences.

(4) A set $M \subset X$ (X') is called ***weakly sequentially compact*** (***weakly* sequentially compact***) if every sequence in M contains a weakly (weakly*) convergent subsequence whose weak (weak*) limit lies in M.

H.W. Alt, *Linear Functional Analysis*, Universitext,
DOI 10.1007/978-1-4471-7280-2_8

Warning: It is possible to define a corresponding weak (weak*) topology (see 8.7). However, if X is not separable, this topology does not have a countable basis of neighbourhoods. It follows that “covering compact” and “sequentially compact” are not equivalent properties (see the example 8.7(4)).
Note: As a complement to weak convergence, convergence with respect to a norm, i.e. norm convergence, will also be referred to as ***strong convergence***. This reduces confusion.

The weak convergence may be interpreted as weak* convergence in the bidual space:

8.2 Embedding into the bidual space.

(1) Defining

$$\langle x' \, , \, J_X x \rangle_{X'} := \langle x \, , \, x' \rangle_X \qquad \text{for } x \in X,\ x' \in X'$$

yields an isometric map $J_X \in \mathscr{L}(X; X'')$. Here

$$X'' := (X')' = \mathscr{L}(X'; \mathbb{K})$$

is the ***bidual space*** of X.
(2) Let $x_k, x \in X$ for $k \in \mathbb{N}$. Then:

$$\begin{matrix} x_k \to x \text{ weakly} \\ \text{in } X \text{ as } k \to \infty \end{matrix} \quad \Longleftrightarrow \quad \begin{matrix} J_X x_k \to J_X x \text{ weakly}^* \\ \text{in } X'' \text{ as } k \to \infty. \end{matrix}$$

(3) Let $x'_k, x' \in X'$ for $k \in \mathbb{N}$. Then:

$$\begin{matrix} x'_k \to x' \text{ weakly} \\ \text{in } X' \text{ as } k \to \infty \end{matrix} \quad \Longrightarrow \quad \begin{matrix} x'_k \to x' \text{ weakly}^* \\ \text{in } X' \text{ as } k \to \infty. \end{matrix}$$

Proof (1). See 6.17(3). □

Proof (2). For $x' \in X'$ we have that $\langle x_k \, , \, x' \rangle_X = \langle x' \, , \, J_X x_k \rangle_{X'}$ and $\langle x \, , \, x' \rangle_X = \langle x' \, , \, J_X x \rangle_{X'}$. □

Proof (3). Because $\langle x \, , \, x'_k \rangle_X = \langle x'_k \, , \, J_X x \rangle_{X'}$ for all $x \in X$. □

8.3 Remarks.

(1) It follows from 6.17(2) that the weak limit of a sequence is unique. For the weak* limit this holds trivially.
(2) Strong convergence (i.e. norm convergence) of a sequence implies weak convergence and weak* convergence.
(3) If $x'_k \to x'$ weakly* in X' as $k \to \infty$, then

$$\|x'\|_{X'} \le \liminf_{k \to \infty} \|x'_k\|_{X'} \, .$$

(4) If $x_k \to x$ weakly in X as $k \to \infty$, then

$$\|x\|_X \le \liminf_{k\to\infty} \|x_k\|_X \,.$$

(5) Weakly convergent sequences and weakly* convergent sequences are bounded.

(6) Let $x_k \to x$ (strongly) in X and $x'_k \to x'$ weakly* in X' as $k \to \infty$. Then

$$\langle x_k \,,\, x'_k\rangle_X \to \langle x \,,\, x'\rangle_X \qquad \text{as } k \to \infty. \tag{8-1}$$

The same holds if $x_k \to x$ weakly in X and $x'_k \to x'$ (strongly) in X'.

Remark: Assertion (4) means that the norm is ***lower semicontinuous*** with respect to the weak convergence of sequences (see also E8.5). Assertion (6) is often used when considering convergence in function spaces.

Proof (3). For all $x \in X$ we have that as $k \to \infty$

$$|\langle x \,,\, x'\rangle_X| \longleftarrow \big|\langle x \,,\, x'_k\rangle_X\big| \le \|x'_k\|_{X'} \cdot \|x\|_X \,,$$

which implies that

$$|\langle x \,,\, x'\rangle_X| \le \liminf_{k\to\infty} \|x'_k\|_{X'} \cdot \|x\|_X \,.$$

Therefore, by the definition of the X'-norm,

$$\|x'\|_{X'} = \sup_{\|x\|_X \le 1} |\langle x \,,\, x'\rangle_X| \le \liminf_{k\to\infty} \|x'_k\|_{X'} \,.$$

□

Proof (4). Analogously to the proof of (3) it holds for all $x' \in X'$ that

$$|\langle x \,,\, x'\rangle_X| \le \|x'\|_{X'} \cdot \liminf_{k\to\infty} \|x_k\|_X \,.$$

If $x \neq 0$, we can choose x' with $\|x'\|_{X'} = 1$ and $\langle x \,,\, x'\rangle_X = \|x\|_X$ (see 6.17(1)) to obtain the desired result. For $x = 0$ the result holds trivially. □

Proof (5). If $x'_k \to x'$ weakly* in X', then

$$\sup_{k\in\mathbb{N}} |\langle x \,,\, x'_k\rangle_X| < \infty \qquad \text{for all } x \in X,$$

and so it follows from the Banach-Steinhaus theorem (see 7.3) that

$$\sup_{k\in\mathbb{N}} \|x'_k\|_{X'} < \infty \,.$$

If $x_k \to x$ weakly in X, then $J_X x_k \to J_X x$ weakly* in X'' (with J_X as in 8.2), and so it follows from the above that $J_X x_k$ is bounded in X'', and hence also x_k in X. □

Proof (6). The first claim follows on noting that

$$\begin{aligned}|\langle x\,,\,x'\rangle_X-\langle x_k\,,\,x_k'\rangle_X|&\le|\langle x\,,\,x'-x_k'\rangle_X|+|\langle x_k-x\,,\,x_k'\rangle_X|\\&\le\underbrace{|\langle x\,,\,x'-x_k'\rangle_X|}_{\to0\text{ as }k\to\infty}+\underbrace{\|x-x_k\|_X}_{\to0\text{ as }k\to\infty}\cdot\underbrace{\|x_k'\|_{X'}}_{\text{bounded in }k}\,,\end{aligned}$$

since, by (5), the sequence $(x_k')_{k\in\mathbb{N}}$ is bounded in X'. The second claim follows analogously. □

We now give some characterizations of weak convergence in function spaces.

8.4 Examples.

(1) Let $1\le p<\infty$ with $\frac{1}{p}+\frac{1}{p'}=1$ (where in the case $p=1$ we assume that the measure space is σ-finite). Then for $f_k,f\in L^p(\mu)$

$$\begin{aligned}&f_k\to f\quad\text{weakly in }L^p(\mu)\text{ as }k\to\infty\\\iff\quad&\int_S f_k\overline{g}\,\mathrm{d}\mu\longrightarrow\int_S f\overline{g}\,\mathrm{d}\mu\quad\text{as }k\to\infty\text{ for all }g\in L^{p'}(\mu).\end{aligned}$$

(2) Let $S\subset\mathbb{R}^n$ be compact. Then for $f_k,f\in C^0(S)$ (see also E8.4)

$$\begin{aligned}&f_k\to f\quad\text{weakly in }C^0(S)\text{ as }k\to\infty\\\iff\quad&\int_S f_k\,\mathrm{d}\lambda\longrightarrow\int_S f\,\mathrm{d}\lambda\quad\text{as }k\to\infty\text{ for all }\lambda\in rca(S).\end{aligned}$$

(3) Let $\Omega\subset\mathbb{R}^n$ be open, let $m\in\mathbb{N}$ and let $1\le p\le\infty$. Then for $u_k,u\in W^{m,p}(\Omega)$

$$\begin{aligned}&u_k\to u\quad\text{weakly in }W^{m,p}(\Omega)\text{ as }k\to\infty\\\iff\quad&\partial^s u_k\to\partial^s u\quad\text{weakly in }L^p(\Omega)\text{ as }k\to\infty\text{ for all }|s|\le m.\end{aligned}$$

The same result holds for the subspace $W_0^{m,p}(\Omega)$.

Proof (1) *and* (2). Follow directly from Theorem 6.12 and Theorem 6.23, respectively. □

Proof (3). Let $X=W^{m,p}(\Omega)$ or $X=W_0^{m,p}(\Omega)$. Then

$$(Jv)(x):=\left(\partial^s v(x)\right)_{|s|\le m}\in\mathbb{K}^M\quad\text{for }v\in X\text{ and almost all }x\in\Omega$$

defines a linear map $J:X\to L^p(\Omega;\mathbb{K}^M)$, where $M:=\binom{n+m}{n}$ is the number of multi-indices s with $|s|\le m$. In addition, $\|Jv\|_{L^p(\Omega;\mathbb{K}^M)}$ can be bounded from above and from below by $\|v\|_{W^{m,p}(\Omega)}$, and so the completeness of X yields that the subspace $Y:=J(X)\subset L^p(\Omega;\mathbb{K}^M)$ is closed. Therefore, J is a bijective continuous linear map between X and $Y=J(X)$ with a continuous inverse $J^{-1}\in\mathscr{L}(Y;X)$.

If $u_k\to u$ weakly in X as $k\to\infty$ and $R\in L^p(\Omega;\mathbb{K}^M)'$, then $T:=RJ\in X'$ and

$$R(Ju_k) = T(u_k) \longrightarrow T(u) = R(Ju) \quad \text{as } k \to \infty,$$

that is, $Ju_k \to Ju$ weakly in $L^p(\Omega; \mathbb{K}^M)$. On the other hand, if this is true and $T \in X'$, then $\bar{R} := TJ^{-1} \in Y'$. Applying the Hahn-Banach theorem 6.15 we obtain an extension $R \in L^p(\Omega; \mathbb{K}^M)'$ of $\bar{R}$ and therefore

$$T(u_k) = \bar{R}(Ju_k) = R(Ju_k) \longrightarrow R(Ju) = \bar{R}(Ju) = T(u) \quad \text{as } k \to \infty,$$

that is, $u_k \to u$ weakly in X. Finally, with $v_k^s := \partial^s u_k$ and $v^s := \partial^s u$, it is clear that

$$\begin{gathered}(v_k^s)_{|s|\le m} \longrightarrow (v^s)_{|s|\le m} \text{ weakly in } L^p(\Omega; \mathbb{K}^M) \text{ as } k \to \infty \\ \Longleftrightarrow \\ \text{for all } |s| \le m : \left(v_k^s \to v^s \text{ weakly in } L^p(\Omega; \mathbb{K}) \text{ as } k \to \infty \right),\end{gathered}$$

a property that is true in general. □

Weak convergence can be interpreted as a generalization of convergence of all coordinates or coordinatewise convergence, as we know it for finite-dimensional spaces. As an analogy of this we replace in the infinite-dimensional case the "coordinates of a point" $x \in X$ by the values $\langle x\,,\, x' \rangle$ for $x' \in X'$. This is the idea behind the proof of the following theorem, which is the main functional analysis result of this chapter.

8.5 Theorem. Let X be separable. Then the closed unit ball $\overline{\mathrm{B}_1(0)}$ in X' is weakly* sequentially compact.

Remark: This then also holds for every other closed ball $\overline{\mathrm{B}_R(x)}$ in X'.

Proof. Let $\{x_n \,;\, n \in \mathbb{N}\}$ be dense in X. If $(x_k')_{k\in\mathbb{N}}$ is a sequence in X' with $\|x_k'\| \le 1$, then $(\langle x_n\,,\, x_k'\rangle)_{k\in\mathbb{N}}$ are bounded sequences in $\mathbb{K}$. Applying a diagonalization procedure we produce a subsequence $k \to \infty$ such that for all n

$$\lim_{k\to\infty} \langle x_n\,,\, x_k' \rangle \quad \text{exists in } \mathbb{K}.$$

Hence we have that for all $y \in Y := \operatorname{span}\{x_n \,;\, n \in \mathbb{N}\}$ the limit

$$x'(y) := \lim_{k\to\infty} \langle y\,,\, x_k' \rangle \quad \text{exists in } \mathbb{K},$$

and $x' : Y \to \mathbb{K}$ is linear. It follows from

$$|x'(y)| = \lim_{k\to\infty} |\langle y\,,\, x_k'\rangle| \le \|y\|$$

that x' is uniformly continuous on Y and so it can be uniquely extended to a continuous linear map x' on $\overline{Y} = X$ (see E5.3). Therefore, $x' \in X'$ with $\|x'\| \le 1$, and for all $x \in X$ and $y \in Y$

$$\begin{aligned}|\langle x\,,\, x' - x_k'\rangle| &\le |\langle x - y\,,\, x' - x_k'\rangle| + |\langle y\,,\, x' - x_k'\rangle| \\ &\le 2\,\|x - y\| + |\langle y\,,\, x' - x_k'\rangle|\,.\end{aligned}$$

The second term, for every y, converges to zero as $k \to \infty$, while the first term can be made arbitrarily small because $\overline{Y} = X$. □

8.6 Examples.

(1) If $X = L^1(\mu)$ is separable, then we obtain from 6.12 (see proof below) the following result: If $(f_k)_{k\in\mathbb{N}}$ is bounded in $L^\infty(\mu)$, then there exists a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ and an $f \in L^\infty(\mu)$ such that

$$\int_S f_{k_i}\overline{g}\,\mathrm{d}\mu \longrightarrow \int_S f\overline{g}\,\mathrm{d}\mu \qquad \text{as } i\to\infty \text{ for all } g \in L^1(\mu).$$

Note: $L^1(\mu)$ is separable, for example, if $S \subset \mathbb{R}^n$ is Lebesgue measurable and μ is the Lebesgue measure, or if $S \subset \mathbb{R}^n$ is compact and $\mu \in rca(S)$.

(2) If $X = C^0(S)$ with $S \subset \mathbb{R}^n$ being compact, then 4.18(3) and 6.23 yield the following result: If $(\mu_k)_{k\in\mathbb{N}}$ is bounded in $rca(S)$, then there exist a subsequence $(\mu_{k_i})_{i\in\mathbb{N}}$ and a measure $\mu \in rca(S)$ such that

$$\int_S g\,\mathrm{d}\mu_{k_i} \longrightarrow \int_S g\,\mathrm{d}\mu \qquad \text{as } i\to\infty \text{ for all } g \in C^0(S).$$

Proof (1) *Note.* If μ is the Lebesgue measure on $S \subset \mathbb{R}^n$, then $L^1(\mu)$ is separable (see 4.18(4)). This also holds for $\mu \in rca(S)$, when $S \subset \mathbb{R}^n$ is compact, because every function in $L^1(\mu)$ can be approximated in the L^1-norm by step functions, and, as μ is regular, every μ-measurable set can be approximated in measure by relatively open sets (with respect to S). But every open set is a countable union of semi-open cuboids, with each cuboid having its center on the lattice $2^{-i}\cdot\mathbb{Z}^n$ and side length 2^{1-i} for an $i \in \mathbb{N}$. □

Proof (1). Let $L^1(\mu)$ be separable. On recalling that functions in $L^1(\mu)$ can be approximated by step functions, it follows from 4.17(2) that there exists a subset $\{g_i\,;\ i\in\mathbb{N}\}$ of step functions which is dense in $L^1(\mu)$, e.g.

$$g_i := \sum_{j=1}^{m_i} \alpha_{ij}\mathcal{X}_{E_{ij}} \qquad \text{with } \mu(E_{ij}) < \infty\,.$$

Let

$$\widetilde{S} := \bigcup_{i,j} E_{ij} \qquad \text{and} \qquad \widetilde{\mu}(E) := \mu(E\cap\widetilde{S}) \text{ for } E \in \mathcal{B}\,.$$

Then $\widetilde{\mu}$ is σ-finite, and so 6.12 can be applied to $L^1(\widetilde{\mu})$. This yields the desired result, because

$$f \in L^1(\mu) \quad\Longrightarrow\quad f = 0\ \mu\text{-almost everywhere in } S\setminus\widetilde{S}\,.$$

To see the above, observe that there exists a sequence $(i_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that $\|f - g_{i_k}\|_{L^1(\mu)} \to 0$ as $k\to\infty$, and so

$$\int_{S\setminus\widetilde{S}} |f|\,\mathrm{d}\mu = \int_{S\setminus\widetilde{S}} |f - g_{i_k}|\,\mathrm{d}\mu \le \|f - g_{i_k}\|_{L^1(\mu)} \longrightarrow 0 \qquad \text{as } k\to\infty\,,$$

which implies that $f = 0$ almost everywhere in $S\setminus\widetilde{S}$. □

8.7 Weak topology. The following results serve to illustrate the concept of weak sequential compactness. They will not be used in the remainder of this book.

(1) ***Weak topology.*** Let X be a Banach space. For triples (n, z', ε) with $n \in \mathbb{N}$, $z' = (z'_k)_{k=1,\dots,n}$, $z'_1, \dots, z'_n \in X'$ and $\varepsilon > 0$ define

$$U_{n,z',\varepsilon} := \left\{ x \in X \; ; \; |\langle x \, , \, z'_k \rangle| < \varepsilon \text{ for } k = 1, \dots, n \right\} ,$$

and

$$\mathcal{T}_w := \left\{ A \subset X \; ; \quad x \in A \Longrightarrow x + U_{n,z',\varepsilon} \subset A \text{ for some } U_{n,z',\varepsilon} \right\} .$$

Then X equipped with $\mathcal{T}_w$ (called the ***weak topology***) is a locally convex topological vector space (as in 5.21), and $\mathcal{T}_w$ is the weakest topology for which all $x' \in X'$ are continuous maps $x' : X \to \mathbb{K}$ with respect to $\mathcal{T}_w$.

(2) ***Weak* topology.*** Let X be a Banach space. For triples (n, z, ε) with $n \in \mathbb{N}$, $z = (z_k)_{k=1,\dots,n}$, $z_1, \dots, z_n \in X$ and $\varepsilon > 0$ define

$$U_{n,z,\varepsilon} := \left\{ x' \in X' \; ; \; |\langle z_k \, , \, x' \rangle| < \varepsilon \text{ for } k = 1, \dots, n \right\} ,$$

and

$$\mathcal{T}'_w := \left\{ A \subset X' \; ; \; x' \in A \Longrightarrow x' + U_{n,z,\varepsilon} \subset A \text{ for some } U_{n,z,\varepsilon} \right\} .$$

Then X' equipped with $\mathcal{T}'_w$ (called the ***weak* topology***) is a locally convex topological vector space (as in 5.21).
Moreover, it holds that: If $\mathcal{T}''_w$ is the weak* topology on $(X')'$ and if J_X is as in 8.2(1), then $\mathcal{T}_w = \{J_X^{-1}(A) \, ; \; A \in \mathcal{T}''_w\}$.

(3) ***Alaoglu's theorem.*** Let X be a Banach space. Then $\overline{\mathrm{B}_1(0)} \subset X'$ (the closed unit ball with respect to the norm on X') is covering compact with respect to the weak* topology on X'.
On the proof: We omit the proof. The result can be shown with the help of Tychonoff's theorem (according to A. N. Tikhonov), see e.g. [Conway].

(4) ***Counterexample to compactness theorems.*** Theorem 8.5 does not hold without the separability of X, that is: In general "weak* sequential compactness" and "cover compactness with respect to the weak* topology" need to be distinguished.
Example: Let $X = L^\infty(]0,1[)$ and for $\varepsilon > 0$ define

$$T_\varepsilon f := \frac{1}{\varepsilon} \int_0^\varepsilon f(x) \, \mathrm{d}x \qquad \text{for } f \in L^\infty(]0,1[) \, .$$

Then $T_\varepsilon \in L^\infty(]0,1[)'$ with $\|T_\varepsilon\| = 1$, and the following holds: There exists no null sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ such that $(T_{\varepsilon_k})_{k\in\mathbb{N}}$ is weakly* convergent in $L^\infty(]0,1[)'$.

Proof (4) *Example.* Assume that $(T_{\varepsilon_k})_{k\in\mathbb{N}}$ is weakly* convergent. By choosing a subsequence (which is then also weakly* convergent and which we again denote by $(T_{\varepsilon_k})_{k\in\mathbb{N}}$), we can assume that

$$1 > \tfrac{\varepsilon_{k+1}}{\varepsilon_k} \to 0 \qquad \text{as } k \to \infty.$$

Now consider the function $f \in L^\infty(]0,1[)$ defined by

$$f(x) := (-1)^j \qquad \text{for } \varepsilon_{j+1} < x < \varepsilon_j \text{ and } j \in \mathbb{N}.$$

Then

$$T_{\varepsilon_k} f = \frac{1}{\varepsilon_k}\Big((\varepsilon_k - \varepsilon_{k+1})(-1)^k + \int_0^{\varepsilon_{k+1}} f(x)\,\mathrm{d}x\Big)\,,$$

and so

$$\big|T_{\varepsilon_k} f - (-1)^k\big| \le \frac{1}{\varepsilon_k}\Big(\varepsilon_{k+1} + \int_0^{\varepsilon_{k+1}} |f(x)|\,\mathrm{d}x\Big) \le \frac{2\varepsilon_{k+1}}{\varepsilon_k} \longrightarrow 0$$

as $k \to \infty$. This shows that the sequence $(T_{\varepsilon_k} f)_{k\in\mathbb{N}}$ has the two cluster points ± 1. Hence $(T_{\varepsilon_k})_{k\in\mathbb{N}}$ cannot be weakly* convergent. □

Reflexive spaces

In the following we consider the class of reflexive spaces. A reflexive space X is characterized by the fact that the bidual space X'' is isometrically isomorphic to the space X itself, however not (!) with respect to an arbitrary isometry, but precisely with respect to the isometry J_X defined in 8.2(1). The class of reflexive spaces contains all Hilbert spaces (see 8.11(1)).

8.8 Reflexivity. Let X be a Banach space and let J_X be the isometry from 8.2(1). Then we call

$$X \textbf{\textit{ reflexive}} \quad :\Longleftrightarrow \quad J_X \text{ is surjective}\ .$$

We have the following results:

(1) If X is reflexive, then weak* and weak sequence convergence in X' coincide.

(2) If X is reflexive, then every closed subspace of X is reflexive.

(3) If $T : X \to Y$ is an isomorphism, then

$$X \text{ reflexive} \quad \Longleftrightarrow \quad Y \text{ reflexive}\ .$$

(4) It holds that

$$X \text{ reflexive} \quad \Longleftrightarrow \quad X' \text{ reflexive}\ .$$

Proof (2). Let $Y \subset X$ be a closed subspace. Given a $y'' \in Y''$, let

$$\langle x' \,,\, x''\rangle_{X'} := \langle x'|_Y \,,\, y''\rangle_{Y'} \qquad \text{for } x' \in X'.$$

Then $x'' \in X''$. Let $x := J_X^{-1} x''$. Now for all $x' \in X'$ with $x' = 0$ on Y we have that

$$\langle x\,,\, x'\rangle_X = \langle x'\,,\, x''\rangle_{X'} = \big\langle x'\big|_Y\,,\, y''\big\rangle_{Y'} = 0\,,$$

which, on recalling 6.16, implies that $x \in Y$. Now let $y' \in Y'$, and let $x' \in X'$ denote an extension of y' as in the Hahn-Banach theorem (see 6.15). Then we conclude that

$$\langle x\,,\, y'\rangle_Y = \langle x\,,\, x'\rangle_X = \big\langle x'\big|_Y\,,\, y''\big\rangle_{Y'} = \langle y'\,,\, y''\rangle_{Y'}\ ,$$

i.e. $y'' = J_Y x$. This shows that J_Y is surjective. □

Proof (3). The claim is symmetric in X and Y, and so it is sufficient to consider the case where X is reflexive. We need to show the reflexivity of Y. Let $y'' \in Y''$. Then

$$\langle x'\,,\, x''\rangle_{X'} := \big\langle x'\circ T^{-1}\,,\, y''\big\rangle_{Y'} \qquad \text{for } x' \in X'$$

defines an $x'' \in X''$, and for $y' \in Y'$ (setting $x' := y'\circ T$)

$$\langle y'\,,\, y''\rangle_{Y'} = \langle y'\circ T\,,\, x''\rangle_{X'} = \big\langle J_X^{-1}x''\,,\, y'\circ T\big\rangle_X = \big\langle TJ_X^{-1}x''\,,\, y'\big\rangle_Y\ ,$$

and so $y'' = J_Y T J_X^{-1} x''$. □

Proof (4)⇒. If $x''' \in X'''$ then $x'''\circ J_X \in X'$, and it holds for all $x'' \in X''$ that

$$\langle x''\,,\, x'''\rangle_{X''} = \big\langle J_X^{-1}x''\,,\, x'''\circ J_X\big\rangle_X = \langle x'''\circ J_X\,,\, x''\rangle_{X'}\ ,$$

i.e. $x''' = J_{X'}(x'''\circ J_X)$. □

Proof (4)⇐. Employing the established implication "⇒" for the Banach space X' yields that X'' is reflexive. As J_X is isometric, $J_X(X)$ is a closed subspace of X'', which according to (2) is also reflexive. Hence (3) implies that X is reflexive. □

The proof of theorem 8.10 below employs the following:

8.9 Lemma. For every Banach space X,

$$X' \text{ separable} \quad\Longrightarrow\quad X \text{ separable}\ .$$

Observe: The converse is false, as shown by the very important example $X = L^1(\mu)$ (see 6.12 and 4.18(4)).

Proof. Let $\{x'_n\,;\ n \in \mathbb{N}\}$ be dense in X'. Choose $x_n \in X$ with

$$|\langle x_n\,,\, x'_n\rangle_X| \geq \tfrac{1}{2}\|x'_n\| \quad \text{and} \quad \|x_n\| = 1$$

and define $Y := \operatorname{clos}(\operatorname{span}\{x_n\,;\ n \in \mathbb{N}\})$. Now if $x' \in X'$ with $x' = 0$ on Y, then for all n

$$\begin{aligned}&\|x' - x'_n\| \geq |\langle x_n\,,\, x' - x'_n\rangle_X| = |\langle x_n\,,\, x'_n\rangle_X|\\ &\geq \tfrac{1}{2}\|x'_n\| \geq \tfrac{1}{2}(\|x'\| - \|x'_n - x'\|)\end{aligned}$$

and so

$$\|x'\| \leq 3 \inf_n \|x' - x'_n\| = 0\,,$$

since $\{x'_n\,;\ n \in \mathbb{N}\}$ is a dense subset. Hence it follows from 6.16 that $Y = X$. ☐

We now prove the main theorem for reflexive spaces.

8.10 Theorem. Let X be a reflexive Banach space. Then the closed unit ball $\overline{\mathrm{B}_1(0)} \subset X$ is weakly sequentially compact.
Remark: This then also holds for every other closed ball $\overline{\mathrm{B}_R(x)}$.

Proof. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\overline{\mathrm{B}_1(0)} \subset X$ and set

$$Y := \overline{\operatorname{span}\{x_k\,;\ k \in \mathbb{N}\}}\,.$$

Then Y is reflexive (see 8.8(2)) and, by definition, separable. It follows that $Y'' = J_Y Y$ is separable, and hence so is Y' (see 8.9). That means that we can apply 8.5 to the space Y' and to the sequence $(J_Y x_k)_{k\in\mathbb{N}}$ in Y''. In particular, there exists a $y'' \in Y''$ such that for a subsequence $k \to \infty$

$$\langle y'\,,\, J_Y x_k\rangle_{Y'} \to \langle y'\,,\, y''\rangle_{Y'} \qquad \text{for all } y' \in Y'.$$

Setting $x := J_Y^{-1} y'' \in Y$, it follows that

$$\langle x_k\,,\, y'\rangle_Y = \langle y'\,,\, J_Y x_k\rangle_{Y'} \longrightarrow \langle y'\,,\, y''\rangle_{Y'} = \langle x\,,\, y'\rangle_Y \qquad \text{as } k \to \infty$$

for all $y' \in Y'$. Since for $x' \in X'$ the map $x'\big|_Y$ lies in Y', it follows that also $\langle x_k\,,\, x'\rangle_X \to \langle x\,,\, x'\rangle_X$ as $k \to \infty$, and so $x_k \to x$ weakly in X as $k \to \infty$. ☐

8.11 Examples of reflexive spaces. Here are several consequences of theorem 8.10.

(1) Every Hilbert space X is reflexive. Together with the Riesz representation theorem 6.1 we obtain: If $(x_k)_{k\in\mathbb{N}}$ is a bounded sequence in X, then there exists a subsequence $(x_{k_i})_{i\in\mathbb{N}}$ and an $x \in X$ such that

$$(y\,,\, x_{k_i})_X \to (y\,,\, x)_X \qquad \text{as } i \to \infty \text{ for all } y \in X.$$

(2) $L^p(\mu)$ for $1 < p < \infty$ is reflexive. It follows from 6.12 that: If $(f_k)_{k\in\mathbb{N}}$ is a bounded sequence in $L^p(\mu)$, then there exists a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ and an $f \in L^p(\mu)$ such that

$$\int_S g f_{k_i}\,\mathrm{d}\mu \longrightarrow \int_S g f\,\mathrm{d}\mu \qquad \text{as } i \to \infty \text{ for all } g \in L^{p'}(\mu).$$

(3) $W^{m,p}(\Omega)$ for $1 < p < \infty$ is reflexive. It holds that: If $(f_k)_{k\in\mathbb{N}}$ is a bounded sequence in $W^{m,p}(\Omega)$, then there exist a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ and an $f \in W^{m,p}(\Omega)$ such that for all $|s| \le m$

$$\int_\Omega g\partial^s f_{k_i}\,\mathrm{dL}^n \longrightarrow \int_\Omega g\partial^s f\,\mathrm{dL}^n \quad \text{as } i\to\infty \text{ for all } g \in L^{p'}(\Omega).$$

(4) $L^1(\mu)$ and $L^\infty(\mu)$ (with the measure μ being σ-finite) are not reflexive if the underlying σ-algebra $\mathcal{B}$ contains infinitely many disjoint sets with positive measure, i.e. if and only if $L^1(\mu)$ and $L^\infty(\mu)$, respectively, are infinite-dimensional.

(5) $C^0(S)$ and $rca(S)$ are not reflexive if $S \subset \mathbb{R}^n$ is compact and contains more than finitely many points, i.e. if and only if $C^0(S)$ and $rca(S)$, respectively, are infinite-dimensional.

Proof (1). Let $R_X : X \to X'$ be the (conjugate linear) isomorphism from the Riesz representation theorem. Then for $x'' \in X''$ letting

$$\langle y\,,\,x'\rangle_X := \overline{\langle R_X y\,,\,x''\rangle_{X'}} \quad \text{for } y \in X$$

defines an $x' \in X'$. Set $x := R_X^{-1}x'$. Then for all $y \in X$

$$\langle R_X y\,,\,x''\rangle_{X'} = \overline{\langle y\,,\,R_X x\rangle_X} = \overline{(y\,,\,x)_X} = \langle x\,,\,R_X y\rangle_X\ ,$$

i.e. $x'' = J_X x$, which shows that J_X is surjective.
Remark: Hence in the real case, i.e. $\mathbb{K} = \mathbb{R}$, it holds that $J_X^{-1} = R_X^{-1}R_X'$, with $R_X' : X'' \to X'$ denoting the adjoint map (see 5.5(8)) of R_X. ☐

Proof (2). The isometries

$$J_p : L^p(\mu) \to L^{p'}(\mu)' \quad \text{and} \quad J_{p'} : L^{p'}(\mu) \to L^p(\mu)'$$

from 6.12 satisfy

$$\overline{\langle f\,,\,J_{p'}g\rangle_{L^p(\mu)}} = \langle g\,,\,J_p f\rangle_{L^{p'}(\mu)} \quad \text{for all } f \in L^p(\mu),\ g \in L^{p'}(\mu).$$

For $f'' \in L^p(\mu)''$ letting

$$\langle g\,,\,g'\rangle_{L^{p'}(\mu)} := \overline{\langle J_{p'}g\,,\,f''\rangle_{L^p(\mu)'}} \quad \text{for } g \in L^{p'}(\mu)$$

defines a $g' \in L^{p'}(\mu)'$. Set $f := J_p^{-1}g'$. Then for all $g \in L^{p'}(\mu)$

$$\langle g\,,\,g'\rangle_{L^{p'}(\mu)} = \langle g\,,\,J_p f\rangle_{L^{p'}(\mu)} = \overline{\langle f\,,\,J_{p'}g\rangle_{L^p(\mu)}} = \overline{\left\langle J_{p'}g\,,\,J_{L^p(\mu)}f\right\rangle_{L^p(\mu)'}}\,,$$

where $J_{L^p(\mu)} : L^p(\mu) \to L^p(\mu)''$ denotes the embedding from 8.2. Consequently,

$$\langle J_{p'}g\,,\,f''\rangle_{L^p(\mu)'} = \left\langle J_{p'}g\,,\,J_{L^p(\mu)}f\right\rangle_{L^p(\mu)'} \quad \text{for all } g \in L^{p'}(\mu).$$

As $J_{p'}$ is surjective, it follows that $f'' = J_{L^p(\mu)}f$, which proves the reflexivity of $L^p(\mu)$.

Remark: Hence in the real case, i.e. $\mathbb{K} = \mathbb{R}$, it holds that $J_{L^p(\mu)}^{-1} = J_p^{-1} J'_{p'}$, with $J'_{p'} : L^p(\mu)'' \to L^{p'}(\mu)'$ denoting the adjoint map (see 5.5(8)) of $J_{p'}$. □

Proof (3). Let $J : W^{m,p}(\Omega) \to L^p(\Omega; \mathbb{K}^M)$ be defined as in the proof of 8.4(3). Then combining (2) and 8.8(2) yields that the closed subspace $J\big(W^{m,p}(\Omega)\big)$ is reflexive (the proof of (2) is the same for functions with values in $\mathbb{K}^M$). The claim now follows from 8.8(3). □

Proof (4). On noting 8.8(4), 6.12 for $p = 1$ and 8.8(3), it is sufficient to show this for $L^1(\mu)$. Let $F \in L^\infty(\mu)'$. If $J_\infty : L^\infty(\mu) \to L^1(\mu)'$ denotes the isomorphism from 6.12, then setting

$$\langle f', G\rangle_{L^1(\mu)'} := \overline{\big\langle J_\infty^{-1} f', F\big\rangle_{L^\infty(\mu)}} \quad \text{for } f' \in L^1(\mu)'$$

defines a $G \in L^1(\mu)''$. If $G = J_{L^1(\mu)}f$ for an $f \in L^1(\mu)$, with $J_{L^1(\mu)}$ denoting the embedding from 8.2, then it holds for all $g \in L^\infty(\mu)$ that

$$\begin{aligned} \overline{\langle g, F\rangle_{L^\infty(\mu)}} &= \langle J_\infty g, G\rangle_{L^1(\mu)'} = \big\langle J_\infty g, J_{L^1(\mu)} f\big\rangle_{L^1(\mu)'} \\ &= \langle f, J_\infty g\rangle_{L^1(\mu)} = \textstyle\int_S f\overline{g}\,\mathrm{d}\mu\,, \end{aligned}$$

that is,

$$\langle g, F\rangle_{L^\infty(\mu)} = \textstyle\int_S g\overline{f}\,\mathrm{d}\mu \quad \text{for all } g \in L^\infty(\mu). \tag{8-2}$$

Under the assumption that $L^1(\mu)$ is infinite-dimensional, we now construct an F which does not satisfy this property. To this end, let $E_k \in \mathcal{B}$ be such that

$$E_k \subset E_{k+1},\ \mu(E_k) < \mu(E_{k+1}) \quad \text{and} \quad E := \textstyle\bigcup_{k\in\mathbb{N}} E_k\,.$$

Consider the subspace

$$Y := \mathrm{clos}\big(\big\{\, g \in L^\infty(\mu)\,;\ g = 0 \text{ on } S \setminus E_k \text{ for some } k \,\big\}\big) \subset L^\infty(\mu)\,.$$

Then $\mathcal{X}_E \notin Y$, and so 6.16 implies that there exists an $F \in L^\infty(\mu)'$ with $F = 0$ on Y and $F(\mathcal{X}_E) = 1$. Hence,

$$F(\mathcal{X}_{E_k}) = 0 \quad \text{and} \quad F(\mathcal{X}_E) = 1\,,$$

but for every $f \in L^1(\mu)$ we have that

$$\textstyle\int_S \mathcal{X}_{E_k}\overline{f}\,\mathrm{d}\mu \longrightarrow \int_S \mathcal{X}_E \overline{f}\,\mathrm{d}\mu\,.$$

Therefore, F cannot have the representation (8-2). □

Proof (5). Let $C^0(S)$ be reflexive. Then analogously to the proof of (4), and on using 6.23, there exists for every functional $F \in rca(S)'$ an $f \in C^0(S)$ with

$$\langle \nu \,,\, F \rangle_{rca(S)} = \int_S f \,\mathrm{d}\nu \quad \text{for all } \nu \in rca(S)\,. \tag{8-3}$$

If S is not finite, then there exist points $x_k \in S$ for $k \in \mathbb{N}$ with $x_k \to x \in S$ as $k \to \infty$ and with $x_k \neq x$ for all k. Consider the Dirac measures δ_{x_k} and δ_x and set $Y := \{\nu \in rca(S)\,;\ \nu(\{x\}) = 0\}$. It holds that $Y \subset rca(S)$ is a closed subspace with $\delta_{x_k} \in Y$ and $\delta_x \notin Y$. It follows from 6.16 that there exists an $F \in rca(S)'$ with $F(\delta_{x_k}) = 0$ for all k and $F(\delta_x) = 1$. But for every $f \in C^0(S)$ we have that

$$\int_S f \,\mathrm{d}\delta_{x_k} = f(x_k) \longrightarrow f(x) = \int_S f \,\mathrm{d}\delta_x\,.$$

Hence F cannot have the representation (8-3). $\square$

Minkowski's functional

In 4.3 we solved the minimal distance problem for closed convex sets in Hilbert spaces, and we saw in E4.3 that in general this is not possible in Banach spaces. We will now show that in reflexive spaces the distance to such sets is attained (see 8.15). This is based on the fact that convex side constraints for elements of an arbitrary Banach space remain valid for limits of weakly convergent sequences, see theorem 8.13. For closed balls this theorem can be obtained directly from 8.3(4), and for general closed convex sets it follows from the following

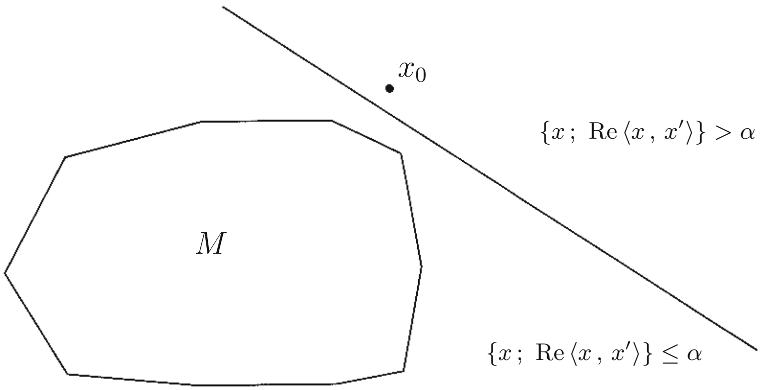

Fig. 8.1. *Separation theorem*

8.12 Separation theorem. Let X be a normed space, let $M \subset X$ be nonempty, closed and convex, and let $x_0 \in X \setminus M$. Then there exist an $x' \in X'$ and an $\alpha \in \mathrm{I\!R}$ with

$$\mathrm{Re}\,\langle x\,,\,x'\rangle \le \alpha \text{ for } x \in M \quad \text{ and } \quad \mathrm{Re}\,\langle x_0\,,\,x'\rangle > \alpha\,.$$

Remark: It follows that $x' \neq 0$, and hence $\{x \in X\,;\ \mathrm{Re}\,\langle x\,,\,x'\rangle = \alpha\}$ is a hyperplane.

Proof. First we consider the case $\mathrm{I\!K} = \mathrm{I\!R}$. We may assume with no loss of generality that

$$0 \in \overset{\circ}{M}\,.$$

Justification: Choose an $\widetilde{x} \in M$ and consider $\widetilde{x}_0 := x_0 - \widetilde{x}$ and $\widetilde{M} := \overline{\mathrm{B}_r(M - \widetilde{x})}$ with $0 < r < \mathrm{dist}(x_0, M)$. Then if the theorem is established for $\widetilde{M}$ and $\widetilde{x}_0$ with x' and $\widetilde{\alpha}$, it follows that the theorem holds for M and x_0 with x' and $\alpha := \widetilde{\alpha} + \langle \widetilde{x}\,,\,x'\rangle$. Consider the ***Minkowski functional***

$$p(x) := \inf\Big\{\, r > 0\ ;\ \frac{x}{r} \in M\,\Big\} \qquad \text{for } x \in X.$$

Since $0 \in \overset{\circ}{M}$, it follows that $0 \le p(x) < \infty$ for all $x \in X$. Moreover,

$$p \le 1 \text{ on } M\,, \quad p(x_0) > 1\,, \quad p(0) = 0\,.$$

In addition, we have for $x, y \in X$ that

$$\begin{aligned} p(ax) &= ap(x) \qquad \text{for } a \ge 0\,,\\ p(x+y) &\le p(x) + p(y)\,, \end{aligned}$$

i.e. p is sublinear. To see this, note that for $\alpha > 0$

$$\frac{x}{r} \in M \quad \Longleftrightarrow \quad \frac{\alpha x}{\alpha r} \in M\,,$$

and that the convexity of M implies that

$$\frac{x}{r} \in M,\ \frac{y}{s} \in M \quad \Longrightarrow \quad \frac{x+y}{r+s} = \frac{r}{r+s}\frac{x}{r} + \frac{s}{r+s}\frac{y}{s} \in M\,.$$

Now let $f : \mathrm{span}\{x_0\} \to \mathrm{I\!R}$ be defined by

$$f(ax_0) := ap(x_0) \qquad \text{for } a \in \mathrm{I\!R}.$$

Then

$$\begin{aligned} f(ax_0) &= p(ax_0) &&\text{for } a \ge 0\,,\\ f(ax_0) &\le 0 \le p(ax_0) &&\text{for } a \le 0\,. \end{aligned}$$

It follows from the Hahn-Banach theorem (see 6.14), applied to the subspace $\mathrm{span}\{x_0\}$, that there exists a linear extension F of f with $F \le p$ on X. Hence

$$F \le p \le 1 \text{ on } M, \quad F(x_0) = f(x_0) = p(x_0) > 1\,.$$

On recalling that $\overline{\mathrm{B}_\varrho(0)} \subset M$ for some $\varrho > 0$, we note that

$$x \in X \implies \frac{x}{\frac{1}{\varrho}\|x\|} \in M \implies p(x) \le \frac{1}{\varrho}\|x\| \implies F(x) \le \frac{1}{\varrho}\|x\|\,.$$

Similarly, $-F(x) = F(-x) \le \frac{1}{\varrho}\|x\|$, which implies that $F \in X'$. Hence we have shown the desired result for $x' := F$ and $\alpha = 1$.

In the case $\mathbb{K} = \mathbb{C}$ consider X as an $\mathbb{R}$-vector space $X_{\mathbb{R}}$ and obtain an $F_{\mathbb{R}} \in X'_{\mathbb{R}}$ with the desired properties. Then, as in the proof of 6.15, proceed to the function $F(x) := F_{\mathbb{R}}(x) - \mathrm{i}F_{\mathbb{R}}(\mathrm{i}x)$. □

8.13 Theorem. Let X be a normed space and let $M \subset X$ be closed and convex. Then M is ***weakly sequentially closed***, i.e. if $x_k, x \in X$ for $k \in \mathbb{N}$, then

$$\left.\begin{array}{l} x_k \to x \text{ weakly in } X \text{ as } k \to \infty, \\ x_k \in M \text{ for } k \in \mathbb{N} \end{array}\right. \implies \quad x \in M\,.$$

Proof. If $x \notin M$, then by the separation theorem 8.12 there exist an $x' \in X'$ and an $\alpha \in \mathbb{R}$ such that

$$\mathrm{Re}\,\langle y\,,\,x'\rangle \le \alpha \text{ for } y \in M \text{ and } \mathrm{Re}\,\langle x\,,\,x'\rangle > \alpha.$$

Now we have that $\mathrm{Re}\,\langle x_k\,,\,x'\rangle \le \alpha$, and the weak convergence to x yields that also $\mathrm{Re}\,\langle x\,,\,x'\rangle \le \alpha$, a contradiction. □

The following two results are consequences of this theorem.

8.14 Mazur's lemma. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in a normed space X that converges weakly to x. Then $x \in \mathrm{clos}\,(\mathrm{conv}\,\{x_k\,;\; k \in \mathbb{N}\})$.

Proof. $M := \mathrm{conv}\,\{x_k\,;\; k \in \mathbb{N}\}$ is a convex set, and hence so is $\overline{M}$. Now apply theorem 8.13. □

8.15 Theorem. Let X be a reflexive Banach space and let $M \subset X$ be nonempty, closed and convex. Then for $x_0 \in X$ there exists an $x \in M$ with

$$\|x - x_0\| = \mathrm{dist}(x_0, M)\,.$$

Proof. Let $(x_k)_{k\in\mathbb{N}}$ be a ***minimal sequence***, i.e.

$$x_k \in M \quad \text{and} \quad \|x_k - x_0\| \to \mathrm{dist}(x_0, M) \quad \text{as } k \to \infty.$$

Then $(x_k)_{k\in\mathbb{N}}$ is a bounded sequence, and so it follows from 8.10 that there exists a subsequence $k \to \infty$ such that $x_k \to x$ weakly in X as $k \to \infty$. Hence 8.13 yields $x \in M$. On noting that also $x_k - x_0 \to x - x_0$ weakly in X, it follows from the lower semicontinuity of the norm (see 8.3(4)) that $\|x - x_0\| = \mathrm{dist}(x_0, M)$. □

Variational methods

Closed convex sets play an important role in existence proofs for elliptic partial differential equations. We now provide applications of theorem 8.13 on closed convex sets to variational problems with side constraints (see 8.17–8.18), where a generalization of the Poincaré inequality 6.7 is needed (see 8.16). The results on partial differential equations will rely on Sobolev spaces, and the theorems required for these spaces will be derived in Appendix A8. Moreover, we always consider open sets $\Omega \subset \mathbb{R}^n$ which are connected.

Remark: An open set $\Omega \subset \mathbb{R}^n$ is connected if and only if it is ***path connected***, i.e. if for every two points $x_0, x_1 \in \Omega$ there exists a (continuous) path in Ω from x_0 to x_1, i.e. a continuous map $\gamma : [0,1] \to \Omega$ with $\gamma(0) = x_0$ and $\gamma(1) = x_1$. In the following we will always only make use of this property (see e.g. 10.4). In a general topological space X a subset $A \subset X$ is said to be ***connected*** if A is not the union of two disjoint, nonempty and relatively in A open sets.

8.16 Generalized Poincaré inequality. Let $\Omega \subset \mathbb{R}^n$ be open, bounded and connected with Lipschitz boundary $\partial\Omega$ (see definition A8.2). Moreover, let $1 < p < \infty$ and let $M \subset W^{1,p}(\Omega)$ be nonempty, closed and convex. Then the following are equivalent for every $u_0 \in M$:

(1) There exists a constant $C_0 < \infty$ such that for all $\xi \in \mathbb{R}$,

$$u_0 + \xi \in M \quad \Longrightarrow \quad |\xi| \le C_0 \,.$$

(2) There exists a constant $C < \infty$ with

$$\|u\|_{L^p(\Omega)} \le C \cdot \big(\|\nabla u\|_{L^p(\Omega)} + 1\big) \quad \text{for all } u \in M \,.$$

Note: If M, in addition, is a ***cone with apex* 0**, i.e. if

$$u \in M,\ r \ge 0 \quad \Longrightarrow \quad ru \in M \,,$$

then the inequality in (2) can be replaced with

$$\|u\|_{L^p(\Omega)} \le C \cdot \|\nabla u\|_{L^p(\Omega)} \quad \text{for all } u \in M \,.$$

Proof Note. Replace u in (2) with ru and let $r \nearrow \infty$. □

Proof (2)⇒(1). Let $\xi \in \mathbb{R}$ with $u := u_0 + \xi \in M$. Then $\nabla u = \nabla u_0$, and hence the inequality in (2) for u implies that

$$C \cdot (\|\nabla u_0\|_{L^p} + 1) \ge \|u_0 + \xi\|_{L^p} \ge |\xi| \cdot \|1\|_{L^p} - \|u_0\|_{L^p} \,.$$

This yields the desired result with a C_0 that depends on C and u_0. □

Proof (1)⇒(2). Without loss of generality we may assume that $u_0 = 0$. To see this, note that if the desired inequality holds for $\widetilde{u} \in \widetilde{M} := M - u_0$ with a constant $\widetilde{C}$, then it follows for $u := \widetilde{u} + u_0$ that

$$\|u\|_{L^p} \le \|\widetilde{u}\|_{L^p} + \|u_0\|_{L^p} \le \widetilde{C} \cdot \big(\|\nabla u\|_{L^p} + \|\nabla u_0\|_{L^p} + 1\big) + \|u_0\|_{L^p}.$$

Now let $u_0 = 0$ and assume that the conclusion is false. Then there exist $u_k \in M$, $k \in \mathbb{N}$, with

$$\|\nabla u_k\|_{L^p} + 1 \le \frac{1}{k}\|u_k\|_{L^p}. \tag{8-4}$$

In particular, $\|u_k\|_{L^p} \to \infty$, and so for every given $R > 0$ (for k sufficiently large)

$$\delta_k := \frac{R}{\|u_k\|_{L^p}} \longrightarrow 0 \quad \text{as } k \to \infty.$$

Hence we have that $0 < \delta_k \le 1$ for k sufficiently large, and combining the fact that $0 \in M$ and the convexity of M then yields that $v_k := \delta_k u_k \in M$. Further,

$$\|v_k\|_{L^p} = \delta_k \|u_k\|_{L^p} = R,$$

and the inequality (8-4) yields that

$$\|\nabla v_k\|_{L^p} + \delta_k \le \frac{1}{k}\|v_k\|_{L^p} = \frac{R}{k} \longrightarrow 0 \quad \text{as } k \to \infty.$$

Thus, the v_k are bounded in $W^{1,p}(\Omega)$. Then 8.11(3) implies that there exist a subsequence, again denoted by $(v_k)_{k\in\mathbb{N}}$, and a $v \in W^{1,p}(\Omega)$, such that $v_k \to v$ weakly in $W^{1,p}(\Omega)$ as $k \to \infty$, and so $v \in M$ on recalling 8.13. In particular, $\nabla v_k \to \nabla v$ weakly in $L^p(\Omega)$ (see 8.4(3)). However, the above inequality yields that $\nabla v_k \to 0$ strongly in $L^p(\Omega)$, and hence $\nabla v = 0$. As Ω is connected, it follows that v is (almost everywhere) a constant function (see E8.9). This means that $v = \xi$ almost everywhere in Ω for some $\xi \in \mathbb{R}$, and the assumptions yield that $|\xi| \le C_0$. On the other hand, by Rellich's embedding theorem (see A8.4), the weak convergence in $W^{1,p}(\Omega)$ implies that $v_k \to v$ strongly in $L^p(\Omega)$, and so

$$R = \|v_k\|_{L^p} \longrightarrow \|v\|_{L^p} = |\xi| \cdot \|1\|_{L^p} \le C_0 \|1\|_{L^p}.$$

This yields a contradiction, on initially choosing R sufficiently large. □

In the above result we have considered domains $\Omega \subset \mathbb{R}^n$ with a local Lipschitz boundary $\partial\Omega$. It turns out that the class of such "Lipschitz domains" is mathematically very robust (see, for example, the trace theorem A8.6 or the embedding theorem 10.9, which for Lipschitz domains holds in Sobolev spaces of arbitrary order). And it is the class of domains that is appropriate for applications, as the boundary can have edges and corners (e.g. cubes are allowed, and more general domains with piecewise smooth boundaries, where the pieces meet at nondegenerate angles). We now consider Sobolev functions on Lipschitz domains and solve the

8.17 Elliptic minimum problem. Let $\Omega \subset \mathbb{R}^n$ be open, bounded and connected with Lipschitz boundary (see A8.2). Let $\mathbb{K} = \mathbb{R}$. Then

$$E(u) := \int_\Omega \Big(\frac{1}{2} \sum_{i,j=1}^n \partial_i u \cdot a_{ij} \partial_j u + fu \Big) \, \mathrm{dL}^n \quad \text{for } u \in W^{1,2}(\Omega)$$

defines a map $E : W^{1,2}(\Omega) \to \mathbb{R}$, where we assume that $f \in L^2(\Omega)$ and $a_{ij} \in L^\infty(\Omega)$. In addition, we assume that $(a_{ij})_{i,j=1,\dots,n}$ is ***elliptic*** (as in (6-8)), i.e. that there exists a positive constant c_0 such that for all $x \in \Omega$

$$\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq c_0 |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^n. \tag{8-5}$$

Without loss of generality we may assume ***symmetry***, i.e. that

$$a_{ij} = a_{ji} \quad \text{for } i,j = 1, \dots, n. \tag{8-6}$$

(Otherwise replace a_{ij} with $\widetilde{a}_{ij} := \frac{1}{2}(a_{ij} + a_{ji})$.) Then for every nonempty, closed and convex subset $M \subset W^{1,2}(\Omega)$ with the property in 8.16 (the property (8-10), below, is stronger) it holds that:

(1) E has an ***absolute minimum*** u on M, i.e. there exists a $u \in M$ such that

$$E(u) \leq E(v) \quad \text{for all } v \in M. \tag{8-7}$$

(2) The absolute minima u of E on M are precisely the solutions of the following ***variational inequality*** of E on M:

$$\int_\Omega \Big(\sum_{i,j=1}^n \partial_i (u - v) \cdot a_{ij} \partial_j u + (u - v) f \Big) \, \mathrm{dL}^n \leq 0 \quad \text{for all } v \in M. \tag{8-8}$$

(3) If M is a closed affine subspace, that is, if $M = u_0 + M_0$ for some $u_0 \in M$ and a closed subspace $M_0 \subset W^{1,2}(\Omega)$, then the variational inequality (8-8) for $u \in M$ is equivalent to

$$\int_\Omega \Big(\sum_{i,j=1}^n \partial_i v \cdot a_{ij} \partial_j u + v f \Big) \, \mathrm{dL}^n = 0 \quad \text{for all } v \in M_0. \tag{8-9}$$

(4) If M satisfies

$$v \in M, \ \xi \in \mathbb{R}, \ v + \xi \in M \quad \Longrightarrow \quad \xi = 0\,, \tag{8-10}$$

then there exists a unique absolute minimum and a unique solution of the variational inequality of E on M.

Proof (1). We begin by showing that there exist positive constants c and C such that

$$E(u) \geq c \int_\Omega |\nabla u|^2 \,\mathrm{dL}^n - C \quad \text{for all } u \in M. \tag{8-11}$$

On noting the elementary ***Young's inequality***

$$a \cdot b \leq \delta a^2 + \tfrac{1}{4\delta} b^2 \quad \text{for } a, b \geq 0 \text{ and } \delta > 0, \tag{8-12}$$

it follows from the ellipticity in (8-5) that

$$\begin{aligned} E(u) &\geq c_0 \int_\Omega |\nabla u|^2 \,\mathrm{dL}^n - \|f\|_{L^2} \|u\|_{L^2} \\ &\geq c_0 \|\nabla u\|_{L^2}^2 - \delta \|u\|_{L^2}^2 - \frac{1}{4\delta} \|f\|_{L^2}^2 \,. \end{aligned}$$

Letting C_1 denote the constant from the Poincaré inequality 8.16(2),

$$\|u\|_{L^2}^2 \leq 2C_1^2 \|\nabla u\|_{L^2}^2 + 2 \,,$$

and so

$$E(u) \geq (c_0 - 2C_1^2 \delta) \|\nabla u\|_{L^2}^2 - C(\delta, f) \,,$$

where $C(\delta, f)$ is a quantity depending on δ and f. On choosing δ sufficiently small, we obtain (8-11) with $c = \frac{c_0}{2}$.

It follows from (8-11) that $E(u) \geq -C$ for all $u \in M$, i.e. E is bounded from below on M. Now choose a ***minimal sequence*** $(u_k)_{k \in \mathbb{N}}$ in M, i.e.

$$E(u_k) \longrightarrow d := \inf_{v \in M} E(v) > -\infty \quad \text{as } k \to \infty.$$

By (8-11), the sequence $(\nabla u_k)_{k \in \mathbb{N}}$ is bounded in $L^2(\Omega)$. Together with the Poincaré inequality 8.16(2) we obtain that $(u_k)_{k \in \mathbb{N}}$ is a bounded sequence in $W^{1,2}(\Omega)$. It follows from 8.11(3) that there exists a $u \in W^{1,2}(\Omega)$ such that $u_k \to u$ weakly in $W^{1,2}(\Omega)$ for a subsequence $k \to \infty$. Since M is closed and convex, it follows from theorem 8.13 that $u \in M$. Moreover, it follows from 8.4(3) that the weak convergence implies that

$$\int_\Omega f(u_k - u) \,\mathrm{dL}^n \longrightarrow 0 \quad \text{and} \quad \sum_{i,j=1}^n \int_\Omega a_{ij} \partial_i u \partial_j (u_k - u) \,\mathrm{dL}^n \longrightarrow 0 \,.$$

Hence we have that

$$\begin{aligned} E(u_k) &= E(u + u_k - u) \\ &= E(u) + \underbrace{\sum_{i,j=1}^n \int_\Omega a_{ij} \partial_i u \partial_j (u_k - u) \,\mathrm{dL}^n + \int_\Omega f(u_k - u) \,\mathrm{dL}^n}_{\to 0 \text{ as } k \to \infty} \\ &\qquad + \int_\Omega \frac{1}{2} \underbrace{\sum_{i,j=1}^n a_{ij} \partial_i (u_k - u) \partial_j (u_k - u)}_{\geq 0} \,\mathrm{dL}^n \,, \end{aligned}$$

which yields that $E(u) \le \liminf_{k\to\infty} E(u_k) = d$. On the other hand, $u \in M$ implies that $E(u) \ge \inf_{v\in M} E(v) = d$, and so $E(u) = d$. □

Proof (2). If u is an absolute minimum and if $v \in M$, then, since M is convex, $(1-\varepsilon)u + \varepsilon v \in M$ for $0 < \varepsilon \le 1$, and so

$$\begin{aligned} E(u) &\le E\big((1-\varepsilon)u+\varepsilon v\big) = E\big(u+\varepsilon(v-u)\big) \\ &= E(u) + \varepsilon \int_\Omega \Big(\sum_{i,j=1}^n \partial_i(v-u) a_{ij} \partial_j u + (v-u) f \Big) \, \mathrm{dL}^n \\ &\qquad + \frac{\varepsilon^2}{2} \int_\Omega \underbrace{\sum_{i,j=1}^n \partial_i(v-u) a_{ij} \partial_j (v-u)}_{\ge 0} \, \mathrm{dL}^n \,. \end{aligned} \tag{8-13}$$

Subtracting $E(u)$, dividing by ε and letting $\varepsilon \searrow 0$ then yields the desired variational inequality.

Conversely, if $u \in M$ then the identity in (8-13) (with $\varepsilon = 1$) yields for all $v \in M$ that

$$E(v) \ge E(u) + \int_\Omega \Big(\sum_{i,j=1}^n \partial_i(v-u) a_{ij} \partial_j u + (v-u) f \Big) \, \mathrm{dL}^n \,.$$

Now if u is a solution of the variational inequality, then the above integral is nonnegative. Hence u is an absolute minimum of E on M. □

Proof (3). In (8-8) choose $v = u \pm \widetilde{v}$ with $\widetilde{v} \in M_0$ (cf. the proof of 4.4(1)). □

Proof (4). If u_1 and u_2 are two solutions of the variational inequality, then choose $v = u_2$ in the variational inequality for u_1 and $v = u_1$ in the variational inequality for u_2 to obtain

$$\int_\Omega \Big(\sum_{i,j=1}^n \partial_i(u_1-u_2) \cdot a_{ij} \partial_j u_1 + (u_1-u_2) f \Big) \, \mathrm{dL}^n \le 0 \,,$$

$$\int_\Omega \Big(\sum_{i,j=1}^n \partial_i(u_2-u_1) \cdot a_{ij} \partial_j u_2 + (u_2-u_1) f \Big) \, \mathrm{dL}^n \le 0 \,.$$

Adding these two inequalities yields that

$$0 \ge \int_\Omega \sum_{i,j=1}^n \partial_i(u_1-u_2) \cdot a_{ij} \partial_j (u_1-u_2) \, \mathrm{dL}^n \ge c_0 \int_\Omega |\nabla(u_1-u_2)|^2 \, \mathrm{dL}^n \,,$$

and so $\nabla(u_1-u_2) = 0$ in $L^2(\Omega)$. As in the proof of 8.16 it now follows for some $\xi \in \mathbb{R}$ that $u_1 - u_2 = \xi \in \mathbb{R}$ almost everywhere in Ω, with the assumptions implying that $\xi = 0$. □

We remark that the techniques used in the proof for the minimum problem in 8.17 carry over to nonquadratic functionals. We now give some important examples of the set M for this minimum problem. Here all of the occurring boundary values are defined with the help of the trace theorem A8.6.

8.18 Examples of minimum problems.

(1) Let

$$M := \left\{ v \in W^{1,2}(\Omega) \;;\; v = 0 \ \mathrm{H}^{n-1}\text{-almost everywhere on } \partial\Omega \right\} .$$

Then it holds: There exists a unique absolute minimum u in 8.17. It satisfies (8-9) with $M_0 = M$. Hence u is the weak solution of the homogeneous ***Dirichlet problem*** in 6.5(1) (for $h_i = 0$, $b = 0$).
Note: It holds that $M = W_0^{1,2}(\Omega)$. Hence this is a special case of theorem 6.8, which was shown there for general open and bounded sets $\Omega \subset \mathbb{R}^n$.

(2) Let

$$M := \left\{ v \in W^{1,2}(\Omega) \;;\; \int_\Omega v \,\mathrm{dL}^n = 0 \right\} .$$

In addition, we assume that $\int_\Omega f \,\mathrm{dL}^n = 0$. Then it holds: There exists a unique absolute minimum u in 8.17. It satisfies the equality (8-9) for all $v \in W^{1,2}(\Omega)$. Hence u is a weak solution of the homogeneous ***Neumann problem*** in 6.5(2) (for $h_i = 0$, $b = 0$). The solution to this problem is unique up to an additive constant.
Observe: This result differs from theorem 6.6, as there the Neumann problem was solved for $b > 0$.

(3) Let $u_0, \psi \in W^{1,2}(\Omega)$ be given and let $u_0(x) \ge \psi(x)$ for almost all $x \in \Omega$. Define

$$M := \Big\{ v \in W^{1,2}(\Omega) \;;\; v = u_0 \ \mathrm{H}^{n-1}\text{-almost everywhere on } \partial\Omega, \quad v \ge \psi \ \mathrm{L}^n\text{-almost everywhere in } \Omega \Big\} .$$

The corresponding minimum problem is called an ***obstacle problem***. Then it holds: There exists a unique solution u to the obstacle problem. It satisfies the variational inequality (8-8).
Special case: For the case $n = 1$, see also E8.8.

(4) Let Lebesgue measurable sets $E_1, E_2 \subset \Omega$ with $\mathrm{L}^n(E_1) > 0$ and $\mathrm{L}^n(E_2) > 0$, and $\psi_1, \psi_2 \in W^{1,2}(\Omega)$ with $\psi_1 \le \psi_2$ almost everywhere in Ω be given. Define

$$M := \Big\{ v \in W^{1,2}(\Omega) \;;\; v \ge \psi_1 \ \mathrm{L}^n\text{-almost everywhere in } E_1, \quad v \le \psi_2 \ \mathrm{L}^n\text{-almost everywhere in } E_2 \Big\} .$$

The corresponding minimum problem is called a ***double obstacle problem***. Then it holds: There exists a solution u to this obstacle problem and it satisfies the variational inequality (8-8).
Remark: The solution need not be unique.

(5) Let $u_0 \in W^{1,2}(\Omega)$ and let $\Gamma \subset \partial\Omega$ be a closed subset with measure $\mathrm{H}^{n-1}(\Gamma) > 0$. Define

$$M := \left\{ v \in W^{1,2}(\Omega) \ ; \ v = u_0 \ \mathrm{H}^{n-1}\text{-almost everywhere on } \Gamma \right\} .$$

Then it holds: There exists a unique absolute minimum $u \in M$ in 8.17. It satisfies (8-9) with

$$M_0 = \{v \in W^{1,2}(\Omega) \, ; \ v = 0 \ \mathrm{H}^{n-1}\text{-almost everywhere on } \Gamma\} .$$

Definition: Then $u \in W^{1,2}(\Omega)$ is called a ***weak solution*** of the ***mixed boundary value problem***

$$\begin{aligned} -\textstyle\sum_{i,j=1}^n \partial_i(a_{ij}\partial_j u) + f = 0 \quad & \text{in } \Omega, \\ u = u_0 \quad & \text{on } \Gamma, \\ \textstyle\sum_{i,j=1}^n \nu_i a_{ij}\partial_j u = 0 \quad & \text{on } \partial\Omega \setminus \Gamma, \end{aligned}$$

where ν is the outer normal to Ω defined in A8.5(3). The weak solution in $W^{1,2}(\Omega)$ to this boundary value problem is unique.

Proof (1). The continuity of the trace operator yields that $M \subset W^{1,2}(\Omega)$ is a closed subspace (with S as in A8.6 it holds that $M = \mathscr{N}(S)$). Clearly M is nonempty and satisfies (8-10) (from $v \in M$ and $v + \xi \in M$ it follows for the traces that $v = 0$ and $v + \xi = 0$ almost everywhere on $\partial\Omega$, and so $\xi = 0$). Now 8.17 yields the existence of a unique solution u, which satisfies (8-9) with $M_0 = M$. □

Proof (2). M is a subspace and contains 0 as the only constant function. In addition, M is closed (the embedding from $W^{1,2}(\Omega)$ into $L^1(\Omega)$ is continuous and the side constraint is continuous on $L^1(\Omega)$). Hence M satisfies the property (8-10), and so 8.17 yields the existence of a unique solution u, which satisfies (8-9) with $M_0 = M$.

For arbitrary $v \in W^{1,2}(\Omega)$ it holds that $\widetilde{v} := v - m(v) \in M$, where

$$m(g) := \fint_\Omega g \, \mathrm{dL}^n := \frac{1}{\mathrm{L}^n(\Omega)} \int_\Omega g \, \mathrm{dL}^n \quad \text{for } g \in L^1(\Omega) \tag{8-14}$$

denotes the ***mean*** of g on Ω.

On recalling that $m(f) = 0$, we obtain that (8-9) holds for constant functions, and hence it also holds for $v = \widetilde{v} + m(v)$, as claimed.

Now if $\widetilde{u} \in M$ is another function that satisfies (8-9) for all $v \in W^{1,2}(\Omega)$, then

$$\int_\Omega \sum_{i,j=1}^n \partial_i v \cdot a_{ij}\partial_j(u - \widetilde{u}) \, \mathrm{dL}^n = 0 \quad \text{for all } v \in W^{1,2}(\Omega).$$

Set $v = u - \widetilde{u}$. Then

$$0 = \int_\Omega \sum_{i,j=1}^n \partial_i(u - \widetilde{u}) \cdot a_{ij}\partial_j(u - \widetilde{u}) \, \mathrm{dL}^n \geq c_0 \int_\Omega |\nabla(u - \widetilde{u})|^2 \, \mathrm{dL}^n .$$

Hence we have that $\nabla(u-\widetilde{u}) = 0$ almost everywhere in Ω. As Ω is connected, it follows that there exists a $\xi \in \mathbb{R}$ such that $\widetilde{u} = u + \xi$ almost everywhere in Ω. $\square$

Proof (3). M is convex and $u_0 \in M$. We show that M is closed. Let $(u_k)_{k\in\mathbb{N}}$ be a sequence in M that converges in $W^{1,2}(\Omega)$ to a $u \in W^{1,2}(\Omega)$. Then it follows from the trace theorem A8.6 that $u_k \to u$ in $L^2(\partial\Omega)$. On noting that $u_k = u_0$ in $L^2(\partial\Omega)$, we also have that $u = u_0$ in $L^2(\partial\Omega)$. In addition, $u_k \to u$ in $L^2(\Omega)$. Hence there exists a subsequence $k \to \infty$ such that $u_k \to u$ almost everywhere in Ω. Now $u_k \geq \psi$ almost everywhere implies that $u \geq \psi$.

Moreover, (8-10) holds. Indeed, it follows from $v \in M$ and $\widetilde{v} := v+\xi \in M$ that $\xi = \widetilde{v} - v = 0$ almost everywhere on $\partial\Omega$, and so $\xi = 0$. By 8.17, there exists a unique solution to the variational inequality. $\square$

Proof (4). We have that M is convex and that $\psi_1, \psi_2 \in M$. The closedness of M follows as in the proof of (3). In addition, 8.16(1) is satisfied, e.g. with $u_0 = \psi_1$. To see this, note that if $v := \psi_1 + \xi \in M$ with $\xi \in \mathbb{R}$, then $\xi \geq 0$, since $\mathrm{L}^n(E_1) > 0$. Similarly, we have that $\xi \leq \psi_2 - \psi_1$ on E_2, and so it follows from $\mathrm{L}^n(E_2) > 0$ (on applying either the Hölder inequality (see 3.18) or Jensen's inequality (see E4.10)) that

$$\xi \leq \fint_{E_2} |\psi_2 - \psi_1| \, \mathrm{dL}^n \leq \left(\fint_{E_2} |\psi_2 - \psi_1|^2 \, \mathrm{dL}^n \right)^{\frac{1}{2}}$$
$$= \left(\mathrm{L}^n(E_2)\right)^{-\frac{1}{2}} \|\psi_2 - \psi_1\|_{L^2(E_2)} < \infty\,.$$

By 8.17, there exists a solution to the minimum problem.
On the uniqueness: In general, there exist several solutions. For example, if $\psi_1 = -1$, $\psi_2 = +1$, $f = 0$, then every constant function $u = \xi$ with $\xi \in [-1, 1]$ is a solution. This would no longer be the case if, in addition, Dirichlet data were prescribed on $\partial\Omega$ (e.g. as in (3)). $\square$

Proof (5). M is convex and $u_0 \in M$. The closedness of M follows as in the proof of (3), on restricting the pointwise argument to the subset $\Gamma \subset \partial\Omega$. The same holds for the proof of (8-10), where now we use that $\mathrm{H}^{n-1}(\Gamma) > 0$. Then 8.17 yields the existence of a unique solution. On noting that $M_0 := M - u_0$ is a subspace, we conclude that (8-9) holds. $\square$

E8 Exercises

Throughout these exercises we let $\mathbb{K} = \mathbb{R}$.

E8.1 Weak limit in $L^p(\mu)$. Let μ be a σ-finite measure and let $f_j, f \in L^p(\mu)$ with $1 \leq p \leq \infty$. Then it holds: If $f_j \to f$ weakly in $L^p(\mu)$ and $f_j \to \widetilde{f}$ μ-almost everywhere as $j \to \infty$, then $\widetilde{f} = f$ μ-almost everywhere.

Solution. Let S_m be as in 3.9(4). It follows from Egorov's theorem A3.18 that for $\varepsilon > 0$ there exists a measurable set $E_\varepsilon \subset S_m$ such that $\mu(S_m \setminus E_\varepsilon) \le \varepsilon$ and $f_j \to \widetilde{f}$ uniformly on E_ε as $j \to \infty$. Given $\zeta \in L^\infty(\mu)$, the map

$$g \longmapsto \int_{E_\varepsilon} \zeta g \,\mathrm{d}\mu$$

defines a continuous linear functional on $L^p(\mu)$ (for $p < \infty$ this follows from $\mu(E_\varepsilon) < \infty$ and the Hölder inequality), i.e. an element of $L^p(\mu)'$. Hence we have that

$$\int_{E_\varepsilon} \zeta(f_j - f)\,\mathrm{d}\mu \longrightarrow 0 \quad \text{as } j \to \infty.$$

Since $f_j \to \widetilde{f}$ uniformly on E_ε,

$$\int_{E_\varepsilon} \zeta(\widetilde{f} - f)\,\mathrm{d}\mu = 0 \quad \text{for all } \zeta \in L^\infty(\mu).$$

Now set $\zeta(x) = \psi\big(\widetilde{f}(x) - f(x)\big)$, where

$$\psi(z) := \begin{cases} \dfrac{z}{|z|} & \text{for } z \ne 0, \\ 0 & \text{for } z = 0. \end{cases}$$

Then $\zeta(\widetilde{f} - f) = |\widetilde{f} - f|$, and hence we obtain that $\widetilde{f} = f$ almost everywhere on E_ε. Letting $\varepsilon \searrow 0$ and $m \nearrow \infty$ yields the desired result. ☐

E8.2 Weak limit of a product. Let μ be a σ-finite measure and let $1 < p < \infty$. Moreover, let $f_j \to f$ in $L^p(\mu)$ as $j \to \infty$, let $(g_j)_{j\in\mathbb{N}}$ be bounded in $L^{p'}(\mu)$ and let $g_j \to g$ almost everywhere. Then

$$g_j f_j \longrightarrow g f \quad \text{weakly in } L^1(\mu) \text{ as } j \to \infty.$$

In particular,

$$\int_S g_j f_j \,\mathrm{d}\mu \longrightarrow \int_S g f \,\mathrm{d}\mu \quad \text{as } j \to \infty.$$

Solution. Otherwise it follows from theorem 6.12 that there exists a $\zeta \in L^\infty(\mu)$ such that for a subsequence $j \to \infty$ and a $\delta > 0$ we have that

$$\left| \int_S g_j f_j \zeta \,\mathrm{d}\mu - \int_S g f \zeta \,\mathrm{d}\mu \right| \ge \delta \quad \text{for all } j. \tag{E8-1}$$

On recalling from 8.11(2) that $L^{p'}(\mu)$ is reflexive for $1 < p' < \infty$, it follows from theorem 8.10 that there exists a $\widetilde{g} \in L^{p'}(\mu)$ such that for a further subsequence $g_j \to \widetilde{g}$ weakly in $L^{p'}(\mu)$ as $j \to \infty$. Now E8.1 yields that $\widetilde{g} = g$, and hence $g_j \to g$ weakly in $L^{p'}(\mu)$. Moreover, $f_j\zeta \to f\zeta$ converges (strongly) in $L^p(\mu)$ as $j \to \infty$. In this situation we can apply 8.3(6):

If $J : L^p(\mu) \to \big(L^{p'}(\mu)\big)'$ denotes the isomorphism from 6.12, then $J(f_j\zeta) \to J(f\zeta)$ converges (strongly) in $\big(L^{p'}(\mu)\big)'$ and hence the second result in 8.3(6) yields that $\langle g_j\,,\, J(f_j\zeta)\rangle_{L^{p'}} \to \langle g\,,\, J(f\zeta)\rangle_{L^{p'}}$, in contradiction to (E8-1). $\square$

E8.3 Weak limit of a product. Let $\mu(S) < \infty$ and let $1 < p \le \infty$. Assume that $f_j \to f$ converges weakly in $L^p(\mu)$ as $j \to \infty$. In addition, let $g_j : S \to \mathbb{R}$ be measurable and uniformly bounded, and let $g_j \to g$ almost everywhere as $j \to \infty$. Then

$$g_j f_j \longrightarrow gf \quad \text{weakly in } L^1(\mu) \text{ as } j \to \infty.$$

Solution. Since $|g_j - g|^{p'}$ are uniformly bounded and $\mu(S) < \infty$, it follows for a constant C that

$$|g_j - g|^{p'} \le C \in L^1(\mu).$$

Since these functions converge almost everywhere to 0, it follows from Lebesgue's convergence theorem 3.25 that $|g_j - g|^{p'} \to 0$ in $L^1(\mu)$, and hence $\zeta g_j \to \zeta g$ (strongly) in $L^{p'}(\mu)$ as $j \to \infty$ for all $\zeta \in L^\infty(\mu)$. Moreover, the assumptions state that $f_j \to f$ weakly in $L^p(\mu)$. In this situation we can apply the first result in 8.3(6) (analogously to the solution of E8.2). $\square$

E8.4 Weak convergence in $\boldsymbol{C^0}$. Let $S \subset \mathbb{R}^n$ be compact and let $f_j, f \in C^0(S)$. Then

$$\begin{array}{l} f_j \longrightarrow f \text{ weakly in } C^0(S) \\ \text{as } j \to \infty \end{array} \quad\Longleftrightarrow\quad \begin{array}{l} \sup\limits_{x\in S}\sup\limits_{j\in\mathbb{N}} |f_j(x)| < \infty \text{ and} \\ f_j(x) \longrightarrow f(x) \text{ as } j \to \infty \\ \text{for all } x \in S. \end{array}$$

Remark: It holds that $\sup_{x\in S}\sup_{j\in\mathbb{N}} |f_j(x)| = \sup_{j\in\mathbb{N}}\sup_{x\in S} |f_j(x)|$.

Solution $\Rightarrow$. By 8.3(5), the sequence $(f_j)_{j\in\mathbb{N}}$ is bounded in $C^0(S)$. Moreover, it follows from 6.23 that the weak convergence is equivalent to

$$\int_S f_j \,\mathrm{d}\nu \longrightarrow \int_S f \,\mathrm{d}\nu \quad \text{as } j \to \infty \tag{E8-2}$$

for all $\nu \in rca(S)$. Now choose $\nu = \delta_x$ for $x \in S$, where δ_x denotes the Dirac measure at the point x. $\square$

Solution $\Leftarrow$. We have to show (E8-2). Let $\mu \in rca(S)$ be nonnegative. It follows from Egorov's theorem A3.18 that for $\varepsilon > 0$ there exists a measurable set $E_\varepsilon \subset S$ with $\mu(S \setminus E_\varepsilon) \le \varepsilon$ such that $f_j \to f$ uniformly on E_ε as $j \to \infty$. On recalling that the functions f_j are uniformly bounded, say $|f_j| \le C$, we have that

$$\left|\int_S (f_j - f)\,\mathrm{d}\mu\right| \le \underbrace{\mu(E_\varepsilon)\sup_{x\in E_\varepsilon}|f_j(x)-f(x)|}_{\substack{\to 0 \text{ as } j\to\infty \\ \text{for every } \varepsilon}} + C\cdot \underbrace{\mu(S\setminus E_\varepsilon)}_{\to 0 \text{ as } \varepsilon\to 0}\,.$$

This yields (E8-2) for μ.

Note: The desired result also holds for arbitrary measures in $rca(S;\mathbb{R})$, as they can be decomposed into their real and imaginary parts, and these further into their positive and negative parts (the nonnegative and nonpositive parts, see the Hahn decomposition A6.2). $\square$

E8.5 Strong convergence in Hilbert spaces. Let X be a Hilbert space. Then it holds for every sequence $(x_k)_{k\in\mathbb{N}}$ in X that:

$$\begin{array}{lcl} x_k \longrightarrow x \text{ (strongly) in } X & & x_k \longrightarrow x \text{ weakly in } X \text{ and} \\ \text{as } k\to\infty & \Longleftrightarrow & \|x_k\|_X \longrightarrow \|x\|_X \text{ as } k\to\infty. \end{array}$$

Solution $\Leftarrow$. We have that

$$\|x_k\|_X^2 = \|x\|_X^2 + 2\mathrm{Re}\,(x_k - x\,,\,x)_X + \|x_k - x\|_X^2\,.$$

It follows from the Riesz representation theorem that $(x_k - x\,,\,x)_X \to 0$ as $k\to\infty$, and so the convergence $\|x_k\|_X \to \|x\|_X$ yields the desired result. $\square$

E8.6 Strong convergence in L^p spaces. Prove that the equivalence in E8.5 also holds for the Banach space $X = L^p(\mu)$ with $1 < p < \infty$.

Solution $\Leftarrow$. Let $f_k, f \in L^p(\mu)$ be such that $f_k \to f$ weakly in $L^p(\mu)$ as $k\to\infty$, which on recalling theorem 6.12 means that

$$\int_S f_k g\,\mathrm{d}\mu \longrightarrow \int_S fg\,\mathrm{d}\mu \quad \text{for all } g\in L^{p'}(\mu)\,,$$

and such that $\|f_k\|_{L^p} \to \|f\|_{L^p}$ as $k\to\infty$. We employ the elementary inequality

$$|b|^p \ge |a|^p + p\cdot(b-a)\bullet\big(|a|^{p-2}a\big) + c\cdot\big(|b|+|a|\big)^{p-2}|b-a|^2 \tag{E8-3}$$

for $a,b\in\mathbb{R}^m, a\ne 0$, with a constant $c>0$ depending on m and p (proof see below).

Set $a = f(x)$, if $f(x)\ne 0$, and $b = f_k(x)$. With $g(x) := |f(x)|^{p-2}f(x)$ (we consider the real case), if $f(x)\ne 0$, and $g(x) := 0$ otherwise, it follows that

$$\int_S |f_k|^p\,\mathrm{d}\mu \ge \int_S |f|^p\,\mathrm{d}\mu + p\cdot\mathrm{Re}\Big(\int_S (f_k - f)g\,\mathrm{d}\mu\Big) + c\cdot\delta_k \tag{E8-4}$$

with

$$\delta_k := \int_{S_k} \big(|f_k| + |f|\big)^{p-2} |f_k - f|^2 \, \mathrm{d}\mu \,,$$

where $S_k := \{x \in S \, ; \; |f_k(x)| + |f(x)| > 0\}$. On noting that $g \in L^{p'}(\mu)$, it follows from the assumptions that the second term on the right-hand side of (E8-4) converges to 0, and that the left-hand side converges to the first term on the right-hand side. We conclude that $\delta_k \to 0$ as $k \to \infty$. For $p \geq 2$ this yields the desired result, since

$$\delta_k \geq \int_S |f_k - f|^p \, \mathrm{d}\mu \,.$$

For $1 < p < 2$ and $\varepsilon > 0$ let

$$E_{\varepsilon,k} := \big\{ \, x \in S_k \; ; \; |f_k(x) - f(x)| \geq \varepsilon(|f_k(x)| + |f(x)|) \, \big\} \,.$$

Then

$$|f_k - f|^p \leq \begin{cases} \varepsilon^{p-2}(|f_k| + |f|)^{p-2} |f_k - f|^2 & \text{on } E_{\varepsilon,k} \,, \\ \varepsilon^p(|f_k| + |f|)^p \leq 2^{p-1}\varepsilon^p(|f_k|^p + |f|^p) & \text{on } S_k \setminus E_{\varepsilon,k} \,, \end{cases}$$

whence

$$\begin{aligned} &\int_S |f_k - f|^p \, \mathrm{d}\mu = \int_{S_k} |f_k - f|^p \, \mathrm{d}\mu \\ &\leq 2^{p-1}\varepsilon^p \int_{S_k \setminus E_{\varepsilon,k}} (|f_k|^p + |f|^p) \, \mathrm{d}\mu + \varepsilon^{p-2} \int_{E_{\varepsilon,k}} (|f_k| + |f|)^{p-2} |f_k - f|^2 \, \mathrm{d}\mu \\ &\leq 2^{p-1}\varepsilon^p \underbrace{(\|f_k\|_{L^p}^p + \|f\|_{L^p}^p)}_{\text{bounded in } k} + \varepsilon^{p-2}\delta_k \end{aligned}$$

for all ε and k, which yields the desired result.

For the proof of (E8-3) let $a_s := (1-s)a + sb$. As (E8-3) depends continuously on b, we may assume that $a_s \neq 0$ for $0 \leq s \leq 1$. Then

$$|a_1|^p - |a_0|^p = p \int_0^1 |a_s|^{p-2} a_s \bullet (a_1 - a_0) \, \mathrm{d}s \,,$$

and hence

$$\begin{aligned} &|a_1|^p - |a_0|^p - p|a_0|^{p-2} a_0 \bullet (a_1 - a_0) \\ &= p \; (a_1 - a_0) \bullet \int_0^1 \int_0^s \frac{\mathrm{d}}{\mathrm{d}t} \big(|a_t|^{p-2} a_t\big) \, \mathrm{d}t \, \mathrm{d}s \\ &= p \int_0^1 \int_0^s |a_t|^{p-2} \Big(|a_1 - a_0|^2 + (p-2)\Big((a_1 - a_0) \bullet \frac{a_t}{|a_t|} \Big)^2 \Big) \, \mathrm{d}t \, \mathrm{d}s \\ &\geq p \; \big(1 + \min(p-2, 0)\big) \cdot \psi(a_0, a_1) \cdot |a_1 - a_0|^2 \,, \end{aligned}$$

with

$$\psi(a_0,a_1) := \int_0^1 \int_0^s |a_t|^{p-2}\,\mathrm{d}t\,\mathrm{d}s\,.$$

Observe that $\psi(a_0,a_1) = \big(|a_0|+|a_1|\big)^{p-2}\,\psi(b_0,b_1)$ with $b_l := \big(|a_0|+|a_1|\big)^{-1}a_l$ for $l=0,1$. Hence we need to show that

$$\inf\{\psi(b_0,b_1)\,;\ |b_0|+|b_1|=1\} > 0\,.$$

For $1<p\le 2$ we have that $\psi(b_0,b_1)\ge \frac{1}{2}$, because $|(1-t)b_0+tb_1|\le 1$, and for $p>2$ the value $\psi(b_0,b_1)$ can converge to 0 only if $b_0\to 0$ and $b_1\to 0$. ☐

E8.7 Weak convergence of oscillating functions. Let $I\subset\mathbb{R}$ be an open, bounded interval and let $1<p<\infty$.

(1) If $g\in L^\infty(\mathbb{R})$ is a ***periodic function*** with ***period*** $\kappa>0$, i.e. $g(x+\kappa)=g(x)$ for almost all x, and if

$$\frac{1}{\kappa}\int_0^\kappa g(x)\,\mathrm{d}x = \lambda\,,$$

then the functions $f_n(x) := g(nx)$ converge weakly in $L^p(I)$ to λ as $n\to\infty$.

(2) Let $\alpha,\beta\in\mathbb{R}$, $0<\theta<1$, and

$$f_n(x) := \begin{cases} \alpha & \text{for } k<nx<k+\theta,\ k\in\mathbb{Z},\\ \beta & \text{for } k+\theta<nx<k+1,\ k\in\mathbb{Z}.\end{cases}$$

Then the functions f_n converge weakly in $L^p(I)$ to the constant function $\theta\alpha+(1-\theta)\beta$ as $n\to\infty$.

(3) Find functions $f_n,f,g_n,g\in L^\infty(I)$ such that $f_n\to f$, $g_n\to g$ weakly in $L^p(I)$ as $n\to\infty$, but such that f_ng_n does not converge weakly to fg.

Solution (1). Without loss of generality let $\lambda=0$ (otherwise replace g with $g-\lambda$). Then the assumptions on g yield that

$$h(x) := \int_0^x g(y)\,\mathrm{d}y$$

defines a continuous function that is bounded on all of $\mathbb{R}$. If $[a,b]\subset I$, then

$$\int_a^b f_n(x)\,\mathrm{d}x = \frac{1}{n}\,(h(nb)-h(na)) \longrightarrow 0 \qquad \text{as } n\to\infty.$$

Consequently,

$$\int_I f_n(x)\zeta(x)\,\mathrm{d}x \longrightarrow 0 \qquad \text{as } n\to\infty$$

for all step functions ζ. As these step functions are dense in $L^{p'}(I)$, and as the functions f_n are bounded in $L^p(I)$, we obtain the same result also for all $\zeta\in L^{p'}(I)$ (see E5.4). ☐

Solution (2). This follows from (1), on noting that

$$\int_0^1 f_1(x)\,\mathrm{d}x = \theta\alpha + (1-\theta)\beta\,.$$

$\square$

Solution (3). Let f_n be as in (2) and define g_n correspondingly for the values $\widetilde{\alpha}, \widetilde{\beta} \in \mathbb{R}$ and the same value θ. Then (2) yields the following weak convergence results in $L^p(I)$:

$$\begin{aligned} f_n &\longrightarrow \theta\alpha + (1-\theta)\beta\,,\\ g_n &\longrightarrow \theta\widetilde{\alpha} + (1-\theta)\widetilde{\beta}\,,\\ f_n g_n &\longrightarrow \theta\alpha\widetilde{\alpha} + (1-\theta)\beta\widetilde{\beta}\,. \end{aligned}$$

Now the equation

$$\theta\alpha\widetilde{\alpha} + (1-\theta)\beta\widetilde{\beta} = (\theta\alpha + (1-\theta)\beta)\left(\theta\widetilde{\alpha} + (1-\theta)\widetilde{\beta}\right)$$

is equivalent to $(\alpha-\beta)(\widetilde{\alpha}-\widetilde{\beta}) = 0$, and so for $\alpha \neq \beta$ and $\widetilde{\alpha} \neq \widetilde{\beta}$ we obtain the desired example. $\square$

E8.8 Variational inequality. Find the solution $u \in W^{1,2}(\Omega)$ of the obstacle problem in 8.18(3) for $n = 1$, $\Omega =]-1,1[\, \subset \mathbb{R}$, $u_0 \geq 0$, $\psi = 0$, $f = 1$ and $a = 1$.

Solution. (On recalling E3.6, we use the fact that for $n = 1$ functions in $W^{1,2}(\Omega)$ can be identified with functions in $C^0(\overline{\Omega})$.) Let

$$\begin{aligned} M = \big\{\, v \in W^{1,2}(\Omega)\ ;\ & v \geq 0 \text{ almost everywhere in } \Omega,\\ & v(\pm 1) = u_\pm := u_0(\pm 1)\ \big\}\,. \end{aligned}$$

Then $u \in M \cap C^0([-1,1])$ and

$$\int_{-1}^1 \big((u-v)'u' + (u-v)\big)\,\mathrm{dL}^1 \leq 0 \qquad \text{for all } v \in M\,.$$

First we consider an interval $]a,b[$ in which $u > 0$. If $\zeta \in C_0^\infty(]a,b[)$, then $u \geq c$ in $\operatorname{supp}\zeta$ for a $c > 0$, and hence $u + \varepsilon\zeta \in M$ for small $|\varepsilon|$. It follows that

$$0 = \int_a^b (\zeta' u' + \zeta)\,\mathrm{dL}^1 = \int_a^b \zeta' v'\,\mathrm{dL}^1\,,$$

where $v(x) := u(x) - \frac{1}{2}x^2$. This implies (see E8.9) that v is linear in $]a,b[$, and hence there exist $d_0, d_1 \in \mathbb{R}$ such that

$$u(x) = \frac{x^2}{2} + d_1 x + d_0 \qquad \text{for } a < x < b.$$

On choosing $]a,b[\ \subset \{u > 0\}$ maximally, i.e. $u(a) = 0$, if $a > -1$, and $u(b) = 0$, if $b < 1$, the obtained characterization of u implies that we have to distinguish the following cases:

$$\begin{aligned} &a = -1\,,\ b = 1\,, &&\text{and so } u > 0 \text{ in }]-1,1[\,, \\ &a > -1\,,\ b = 1\,, &&\text{and so } u > 0 \text{ in }]a,1] \text{ with } u(a) = 0, \\ &a = -1\,,\ b < 1\,, &&\text{and so } u > 0 \text{ in } [-1,b[\text{ with } u(b) = 0. \end{aligned}$$

Hence overall we obtain the following two cases for u:

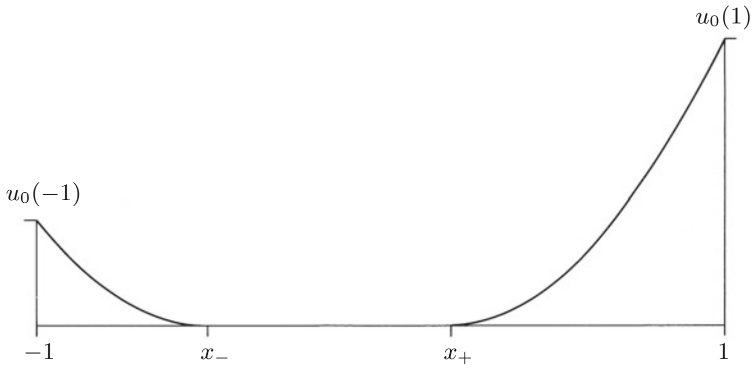

Fig. 8.2. *Solution of the obstacle problem*

(1) $u > 0$ in $]-1,1[$,
(2) There exist $-1 \le x_- \le x_+ \le 1$ such that $u(x) = 0$ for $x_- \le x \le x_+$ and $u(x) > 0$ otherwise.

In the case (1) the values d_0 and d_1 are determined by the boundary conditions, and we obtain

$$u(x) = \tfrac{1}{2}\big(x^2 - 1 + (u_+ - u_-)x + u_+ + u_-\big)$$

and the necessary condition

$$|u_+ - u_-| \ge 2 \quad \text{or} \quad u_+ + u_- > 1 + \tfrac{1}{4}(u_+ - u_-)^2\,. \tag{E8-5}$$

Correspondingly, in the case (2) we obtain for certain $s_\pm \ge 0$ that

$$\begin{aligned} &u(x) = \tfrac{1}{2}(x - x_+)^2 + s_+(x - x_+) \\ &\qquad \text{for } x \ge x_+ \text{ with } (1 - x_+)s_+ = u_+ - \tfrac{1}{2}(1 - x_+)^2 \ge 0\,, \\ &u(x) = \tfrac{1}{2}(x_- - x)^2 + s_-(x_- - x) \\ &\qquad \text{for } x \le x_- \text{ with } (1 + x_-)s_- = u_- - \tfrac{1}{2}(1 + x_-)^2 \ge 0\,. \end{aligned}$$

The uniqueness of the solution means that $x_\pm$ are uniquely determined by $u_\pm$. Hence we further investigate the variational inequality. For $\zeta \in C_0^\infty(]-1,1[)$ with $\zeta \geq 0$ it holds that $u+\zeta \in M$, and so the variational inequality yields that

$$\begin{aligned}0 &\leq \int_{-1}^{1} (\zeta' u' + \zeta)\,\mathrm{dL}^1 \\ &= \int_{x_+}^{1} \zeta'(x)(x - x_+ + s_+)\,\mathrm{d}x + \int_{-1}^{x_-} \zeta'(x)(x - x_- - s_-)\,\mathrm{d}x + \int_{-1}^{1} \zeta\,\mathrm{dL}^1 \\ &= -\zeta(x_+)s_+ - \zeta(x_-)s_- + \int_{x_-}^{x_+} \zeta\,\mathrm{dL}^1\,.\end{aligned}$$

If $x_+ < 1$ set $\zeta(x) := \max(0, 1 - \frac{1}{\delta}|x - x_+|)$ and obtain as $\delta \to 0$ that $s_+ \leq 0$. Together with the above inequality for s_+ we obtain that $s_+ = 0$, and similarly for $s_- = 0$. Therefore,

$$u(x) = \begin{cases} \frac{1}{2}(x_- - x)^2 & \text{for } x \leq x_-, \\ 0 & \text{for } x_- \leq x \leq x_+, \\ \frac{1}{2}(x - x_+)^2 & \text{for } x \geq x_+, \end{cases}$$

where

$$u_+ - \tfrac{1}{2}(1 - x_+)^2 = 0 \quad \text{and} \quad u_- - \tfrac{1}{2}(1 + x_-)^2 = 0\,.$$

Apart from $(u_-, u_+) = (0,2)$ or $(2,0)$, this case is complementary to the case (E8-5). □

E8.9 A fundamental lemma. Let $\Omega \subset \mathbb{R}^n$ be open and connected, and suppose that $u \in L^1_{\mathrm{loc}}(\Omega)$ satisfies

$$\int_\Omega u \cdot \partial_i \zeta\,\mathrm{dL}^n = 0 \qquad \text{for } \zeta \in C_0^\infty(\Omega) \text{ and } i = 1, \ldots, n.$$

Then u is (almost everywhere) a constant function.

Solution. Let B be a ball with $\overline{B} \subset \Omega$ and let $(\varphi_\varepsilon)_{\varepsilon>0}$ be a standard Dirac sequence. On setting $\widetilde{\varphi}_\varepsilon(y) := \varphi_\varepsilon(-y)$ we have that $\zeta * \widetilde{\varphi}_\varepsilon \in C_0^\infty(\Omega)$ for $\zeta \in C_0^\infty(B)$ and $\varepsilon < \operatorname{dist}(B, \partial\Omega)$, and so

$$-\int_\Omega \partial_i(u * \varphi_\varepsilon)\,\zeta\,\mathrm{dL}^n = \int_\Omega (u * \varphi_\varepsilon)\,\partial_i\zeta\,\mathrm{dL}^n = \int_\Omega u\,\partial_i(\zeta * \widetilde{\varphi}_\varepsilon)\,\mathrm{dL}^n = 0\,.$$

Hence $\nabla(u * \varphi_\varepsilon) = 0$ in B, which yields that $u * \varphi_\varepsilon$ is constant in B. On recalling that $u * \varphi_\varepsilon \to u$ in $L^1(B)$ as $\varepsilon \to 0$, it follows that u is also a constant almost everywhere in B. As Ω is path connected (see remark above 8.16), this constant does not depend on B. □

A8 Properties of Sobolev functions

Here we will derive properties of functions in $W^{m,p}(\Omega)$, where we treat bounded sets Ω with Lipschitz boundary $\partial\Omega$ (see definition A8.2). This class of domains, on one hand, allows a functional analytically uniform presentation of the theory of Sobolev spaces, and on the other hand, this class is of major importance in applications, because it contains domains with edges and corners, as they occur in flow domains and also in workpieces.

In applications to boundary value problems on such domains, e.g. on cuboids, often different boundary conditions are prescribed on different sides of the domain (see the mixed boundary value problem in 8.18(5)). For the weak formulation of these boundary value problems we need to prove that functions in $W^{1,p}(\Omega)$ have weak boundary values on $\partial\Omega$ (see A8.6). Then we show (see A8.10) that $W_0^{1,p}(\Omega)$ consists precisely of those functions in $W^{1,p}(\Omega)$ that have weak boundary values 0. This belatedly justifies the weak formulation of the homogeneous Dirichlet problem in 6.5.

We begin with Rellich's embedding theorem A8.1 for $W_0^{m,p}(\Omega)$ and A8.4 for $W^{m,p}(\Omega)$.

A8.1 Rellich's embedding theorem. Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $1 \le p < \infty$ and let $m \ge 1$. If $u_k \in W_0^{m,p}(\Omega)$ for $k \in \mathbb{N}$ and if $u \in W_0^{m-1,p}(\Omega)$, then

$$\begin{array}{lcl} (u_k)_{k\in\mathbb{N}} \text{ bounded in } W_0^{m,p}(\Omega), & & u_k \to u \text{ (strongly) in } W_0^{m-1,p}(\Omega) \\ u_k \to u \text{ weakly in } W_0^{m-1,p}(\Omega) & \Longrightarrow & \text{as } k \to \infty\,. \\ \text{as } k \to \infty & & \end{array}$$

Remark: On recalling 8.3(5), it follows if $u_k, u \in W_0^{m,p}(\Omega)$ for $k \in \mathbb{N}$ that

$$\begin{array}{lcl} u_k \to u \text{ weakly in } W_0^{m,p}(\Omega) & \Longrightarrow & u_k \to u \text{ (strongly) in } W_0^{m-1,p}(\Omega) \\ \text{as } k \to \infty & & \text{as } k \to \infty\,. \end{array}$$

Proof. Let $m = 1$. Hence u_k are bounded in $W_0^{1,p}(\Omega)$ and converge weakly in $L^p(\Omega)$ towards u. (For $m > 1$ apply the proof below for $|s| \le m-1$ to $\partial^s u_k$ in place of u_k. It holds that $\partial^s u_k$ are bounded in $W_0^{1,p}(\Omega)$ and, by 8.4(3), they converge weakly in $L^p(\Omega)$ to $\partial^s u$.)

Extend u_k, u to $\mathbb{R}^n \setminus \Omega$ by 0. Then, by assumption, $u_k \in W^{1,p}(\mathbb{R}^n)$ (see 3.29), with support in $\overline{\Omega}$, and moreover u_k are bounded in $W^{1,p}(\mathbb{R}^n)$ converging by 8.4(1) weakly in $L^p(\mathbb{R}^n)$ towards u.

Now if $(\varphi_\varepsilon)_{\varepsilon>0}$ is a standard Dirac sequence, then $\varphi_\varepsilon * u_k \in C_0^\infty(\mathbb{R}^n)$ and for every $\varepsilon > 0$

$$\varphi_\varepsilon * u_k \to \varphi_\varepsilon * u \quad \text{as } k \to \infty \text{ in } L^p(\mathbb{R}^n)\,. \tag{A8-1}$$

To see this, consider for $x \in \mathbb{R}^n$ the functionals $\Psi_\varepsilon(x) \in L^p(\mathbb{R}^n)'$ defined by

$$\langle v\,,\,\Psi_\varepsilon(x)\rangle_{L^p} := \int_{\mathbb{R}^n} v(y)\varphi_\varepsilon(x-y)\,\mathrm{d}y \quad \text{for } v \in L^p(\mathbb{R}^n).$$

If $x_k \to x$ converges as $k \to \infty$, then $\varphi_\varepsilon(x_k - \bullet) \to \varphi_\varepsilon(x - \bullet)$ converges uniformly on $\mathbb{R}^n$ for $\varepsilon > 0$, hence $\Psi_\varepsilon(x_k) \to \Psi_\varepsilon(x)$ converges in $L^p(\mathbb{R}^n)'$. Since, by assumption, $u_k \to u$ weakly in $L^p(\mathbb{R}^n)$, we obtain, see the second result in 8.3(6),

$$(\varphi_\varepsilon * u_k)(x_k) = \langle u_k\,,\,\Psi_\varepsilon(x_k)\rangle_{L^p} \longrightarrow \langle u\,,\,\Psi_\varepsilon(x)\rangle_{L^p} = (\varphi_\varepsilon * u)(x)\,.$$

This shows that $\varphi_\varepsilon * u_k \to \varphi_\varepsilon * u$ locally uniformly on $\mathbb{R}^n$. As $\varphi_\varepsilon * u_k$ and $\varphi_\varepsilon * u$ vanish outside the bounded set $\overline{\mathrm{B}_\varepsilon(\Omega)}$, we obtain the result (A8-1). Moreover,

$$\|v - \varphi_\varepsilon * v\|_{L^p} \le \varepsilon \|\nabla v\|_{L^p} \tag{A8-2}$$

for all $v \in W^{1,p}(\mathbb{R}^n)$ with compact support. For the proof of (A8-2) observe that the left- and right-hand sides depend continuously on v with respect to the $W^{1,p}$-norm. Hence on approximating v (e.g. by convolution as in 4.23), it is sufficient to show (A8-2) for $v \in C_0^\infty(\mathbb{R}^n)$. Then

$$\begin{aligned}(v - \varphi_\varepsilon * v)(x) &= \int_{\mathbb{R}^n} \varphi_\varepsilon(y)\big(v(x) - v(x-y)\big)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \varphi_\varepsilon(y)\Big(\int_0^1 \nabla v(x - sy)\bullet y\,\mathrm{d}s\Big)\,\mathrm{d}y\,,\end{aligned}$$

and so it follows from 4.13(1) that

$$\begin{aligned}\|v - \varphi_\varepsilon * v\|_{L^p} &\le \sup_{h \in \operatorname{supp}\varphi_\varepsilon} \Big\| \int_0^1 \nabla v(\bullet - sh)\bullet h\,\mathrm{d}s \Big\|_{L^p} \\ &\le \varepsilon \sup_{|h|\le\varepsilon} \int_0^1 \|\nabla v(\bullet - sh)\|_{L^p}\,\mathrm{d}s = \varepsilon\|\nabla v\|_{L^p}\,.\end{aligned}$$

Combining (A8-1) and (A8-2) yields that

$$\|u - u_k\|_{L^p} \le \|u - \varphi_\varepsilon * u\|_{L^p} + \underbrace{\|\varphi_\varepsilon * u - \varphi_\varepsilon * u_k\|_{L^p}}_{\substack{\to 0 \text{ as } k \to \infty \\ \text{for every } \varepsilon}} + \varepsilon\|\nabla u_k\|_{L^p}\,.$$

Noting that ∇u_k are bounded in $L^p(\mathbb{R}^n)$ and recalling from 4.15(2) that $\varphi_\varepsilon * u \to u$ in $L^p(\mathbb{R}^n)$ as $\varepsilon \to 0$, we obtain the desired result. ☐

A8.2 Lipschitz boundary. Let $\Omega \subset \mathbb{R}^n$ be open and bounded. We say that Ω has a ***Lipschitz boundary*** if $\partial\Omega$ can be covered by finitely many open sets $U^1, \dots, U^l$ such that $\partial\Omega \cap U^j$ for $j = 1, \dots, l$ is the graph of a Lipschitz continuous function with $\Omega \cap U^j$ in each case lying on one side of this graph. This means the following: There exists an $l \in \mathbb{N}$ and for

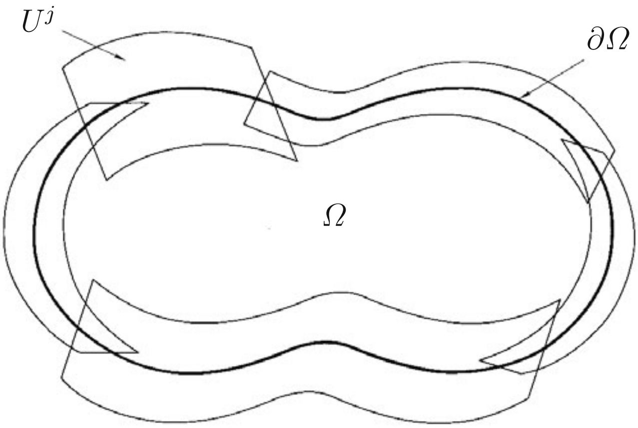

Fig. 8.3. *Cover of the boundary*

$j = 1, \ldots, l$ there exist a Euclidean coordinate system $e_1^j, \ldots, e_n^j$ in $\mathbb{R}^n$, a reference point $y^j \in \mathbb{R}^{n-1}$, numbers $r^j > 0$ and $h^j > 0$ and a Lipschitz continuous function $g^j : \mathbb{R}^{n-1} \to \mathbb{R}$, such that with the notation

$$x_{,n}^j := (x_1^j, \ldots, x_{n-1}^j), \qquad \text{where } x = \sum_{i=1}^n x_i^j e_i^j \,,$$

it holds that

$$U^j = \left\{ x \in \mathbb{R}^n \; ; \; \left| x_{,n}^j - y^j \right| < r^j \text{ and } \left| x_n^j - g^j(x_{,n}^j) \right| < h^j \right\} ,$$

and for $x \in U^j$

$$\begin{aligned} x_n^j = g^j(x_{,n}^j) \quad &\Longrightarrow \quad x \in \partial\Omega \,, \\ 0 < x_n^j - g^j(x_{,n}^j) < h^j \quad &\Longrightarrow \quad x \in \Omega \,, \\ 0 > x_n^j - g^j(x_{,n}^j) > -h^j \quad &\Longrightarrow \quad x \notin \Omega \end{aligned} \tag{A8-3}$$

(hence $U^j \cap \Omega = Q^j$, see Fig. 8.4), and

$$\partial\Omega \subset \bigcup_{j=1}^{l} U^j \,.$$

Furthermore, we may then add another open set U^0 with $\overline{U^0} \subset \Omega$ such that $U^0, \ldots, U^l$ cover all of $\overline{\Omega}$.

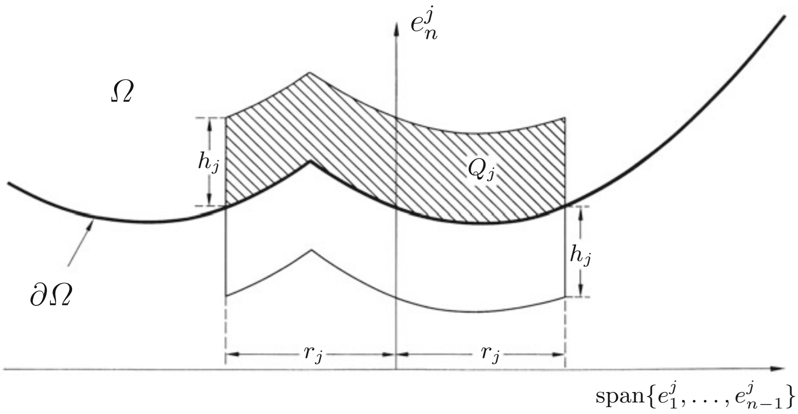

Fig. 8.4. *Local boundary neighbourhood*

A8.3 Localization. Let Ω be as in A8.2. We prove results for Sobolev functions by localizing these functions with respect to the open cover U^j, $j = 0, \ldots, l$ in A8.2. We choose a partition of unity $\eta^0, \ldots, \eta^l$ on $\overline{\Omega}$ with respect to this cover (see 4.20), i.e. $\eta^j \in C^\infty(\mathbb{R}^n)$ with compact support $\mathrm{supp}\,(\eta^j) \subset U^j$ (this means $\eta^j \in C_0^\infty(U^j)$) and

$$0 \le \eta^j \le 1 \text{ in } \mathbb{R}^n \text{ and } \sum_{j=0}^{l} \eta^j = 1 \text{ on } \overline{\Omega}\,.$$

Now if $u \in W^{m,p}(\Omega)$, then

$$u = \sum_{j=0}^{l} \eta^j u \text{ in } \Omega\,.$$

In particular, $\eta^0 u \in W^{m,p}(\Omega)$ with compact support in Ω and for $j = 1, \ldots, l$ we have that $\eta^j u \in W^{m,p}(\Omega^j)$, where

$$\Omega^j := \{x \in \mathbb{R}^n\,;\; 0 < x_n^j - g^j(x_{,n}^j)\}\,,$$

with $(\eta^j u)(x) = 0$ if $\left|x_{,n}^j - y^j\right| \ge r^j$ or $x_n^j - g^j(x_{,n}^j) \ge h^j$.

A8.4 Rellich's embedding theorem. Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary, let $1 \le p < \infty$ and let $m \ge 1$. If $u_k \in W^{m,p}(\Omega)$ for $k \in \mathbb{N}$ and if $u \in W^{m-1,p}(\Omega)$, then

$$\begin{array}{l} (u_k)_{k\in\mathbb{N}} \text{ bounded in } W^{m,p}(\Omega) \\ u_k \to u \text{ weakly in } W^{m-1,p}(\Omega) \\ \text{as } k \to \infty \end{array} \quad \Longrightarrow \quad \begin{array}{l} u_k \to u \text{ (strongly) in } W^{m-1,p}(\Omega) \\ \text{as } k \to \infty\,. \end{array}$$

Remark: On recalling 8.3(5), it follows if $u_k, u \in W^{m,p}(\Omega)$ for $k \in \mathbb{N}$ that

$$\begin{array}{lcl} u_k \to u \text{ weakly in } W^{m,p}(\Omega) & \Longrightarrow & u_k \to u \text{ (strongly) in } W^{m-1,p}(\Omega) \\ \text{as } k \to \infty & & \text{as } k \to \infty\,. \end{array}$$

Proof. Similarly to A8.1, it is sufficient to prove the theorem for $m = 1$. With the notations as in A8.3, we have that the assumptions are also satisfied by $u_k^j := \eta^j u_k$ and $u^j := \eta^j u$, and we need to show that $u_k^j \to u^j$ in $L^p(\Omega)$ as $k \to \infty$. For $j = 0$ this follows from A8.1.

For $j \geq 1$ this follows on replicating the proof of A8.1. The proofs of (A8-1) and (A8-2) carry over (for the proof of (A8-2) use 4.24 for the approximation) if we replace the integration domain $\mathbb{R}^n$ with Ω^j. Here we have to make sure that in the convolution

$$(\varphi_\varepsilon * v)(x) = \int_{\mathbb{R}^n} \varphi_\varepsilon(x-y)v(y)\,\mathrm{d}y \qquad \text{for } v \in W^{1,p}(\Omega^j)$$

the function $y \mapsto \varphi_\varepsilon(x-y)$ has compact support in Ω^j for $x \in \Omega^j$. By the definition of Ω^j this means that

$$x_n^j > g^j(x_{,n}^j)\,,\ \varphi_\varepsilon(x-y) \neq 0 \quad \Longrightarrow \quad y_n^j > g^j(y_{,n}^j)\,.$$

If λ denotes the Lipschitz constant of g^j, then the above holds if

$$\varphi_\varepsilon(z) \neq 0 \quad \Longrightarrow \quad z_n^j < -\lambda\big|z_{,n}^j\big|\,,$$

i.e. we need to choose the function φ, on which the Dirac sequence $(\varphi_\varepsilon)_{\varepsilon>0}$ is based, so that

$$\varphi \in C_0^\infty\big(\{z \in \mathrm{B}_1(0)\ ;\ z_n^j < -\lambda\big|z_{,n}^j\big|\}\big)\,,$$

which is satisfied, for example, for $\varphi \in C_0^\infty\big(\mathrm{B}_\delta\big(-\frac{1}{2}e_n^j\big)\big)$ with $0 < \delta < \frac{1}{2}(1+\lambda)^{-1}$. This choice has the property that for $x \in \Omega^j$ and $\varphi_\varepsilon(x-y) \neq 0$ the segment connecting x and y lies in Ω^j.

Remark: Another possibility (for $m = 1$) is to extend the functions u_k^j, u^j to functions in $W^{1,p}(\mathbb{R}^n)$ with compact support (see the proof of A8.12), and then apply A8.1. □

The corresponding result for $p = \infty$ plays a special role, because for domains Ω with Lipschitz boundary it holds that $W^{m,\infty}(\Omega) = C^{m-1,1}(\overline{\Omega})$ (see theorem 10.5(2)). The assertion of Rellich's embedding theorem for $p = \infty$ then follows from the Arzelà-Ascoli theorem. The argument for $m = 1$ is as follows: Every sequence bounded in $C^{0,1}(\overline{\Omega})$ contains a subsequence that converges in $C^0(\overline{\Omega})$. But as every cluster point has to coincide with the weak limit, the whole sequence converges strongly in $C^0(\overline{\Omega})$.

Now we want to show that Sobolev functions in $W^{1,p}(\Omega)$ in a weak sense have boundary values in $L^p(\partial\Omega)$. To this end, we first define spaces of functions that are integrable on $\partial\Omega$.

A8.5 Boundary integral. Let Ω be open and bounded with Lipschitz boundary and let Y be a Banach space.

(1) We call $f : \partial\Omega \to Y$ measurable and integrable, respectively, if with the notations as in A8.3 it holds for $j = 1, \dots, l$ that the functions

$$y \longmapsto (\eta^j f)\Big(\sum_{i=1}^{n-1} y_i e_i^j + g^j(y) e_n^j\Big) \quad \text{for } y \in \mathbb{R}^{n-1} \text{ with } \big|y - y^j\big| < r^j$$

are measurable and integrable, respectively, with respect to the $(n-1)$-dimensional Lebesgue measure. The boundary integral of f on $\partial\Omega$ is then defined by

$$\int_{\partial\Omega} f \, \mathrm{dH}^{n-1} := \sum_{j=1}^{l} \int_{\partial\Omega} \eta^j f \, \mathrm{dH}^{n-1} ,$$

where we define, if $\operatorname{supp} h \subset U^j$,

$$\int_{\partial\Omega} h \, \mathrm{dH}^{n-1} := \int_{\mathbb{R}^{n-1}} h\Big(\sum_{i=1}^{n-1} y_i e_i^j + g^j(y) e_n^j\Big) \sqrt{1 + |\nabla g^j(y)|^2} \, \mathrm{dL}^{n-1}(y) .$$

Here $\nabla g^j \in L^\infty_{\mathrm{loc}}(\mathbb{R}^{n-1}; \mathbb{R}^{n-1})$, since theorem 10.5(2) implies that the Lipschitz continuous function $g^j : \mathbb{R}^{n-1} \to \mathbb{R}$ lies in $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^{n-1})$. Hence the last integral represents a generalization of the surface integral on smooth hypersurfaces as introduced in 3.10(4). Claim: This definition of the integral is independent of the local partition and independent of the representation of the boundary.

(2) For $1 \le p \le \infty$, let

$$L^p(\partial\Omega; Y) := \big\{ f : \partial\Omega \to Y \ ; \ f \text{ is measurable and } \|f\|_{L^p(\partial\Omega)} < \infty \big\} ,$$

where for $1 \le p < \infty$

$$\|f\|_{L^p(\partial\Omega)} := \Big(\int_{\partial\Omega} |f|^p \, \mathrm{dH}^{n-1}\Big)^{\frac{1}{p}} , \quad \text{and} \quad \|f\|_{L^\infty(\partial\Omega)} := \operatorname*{ess\,sup}_{\partial\Omega} |f|$$

with the ess sup-norm defined analogously to 3.15. Then $L^p(\partial\Omega; Y)$ with this norm is a Banach space for $1 \le p \le \infty$, and for $p < \infty$ the set $\big\{ f\big|_{\partial\Omega} \ ; \ f \in C^\infty(\mathbb{R}^n; Y) \big\}$ is dense in $L^p(\partial\Omega; Y)$.

(3) We define the ***outer normal*** to Ω at the point $x \in \partial\Omega$ as

$$\nu_\Omega(x) := \big(1 + \big|\nabla g^j(y)\big|^2\big)^{-\frac{1}{2}} \Big(\sum_{i=1}^{n-1} \partial_i g^j(y) e_i^j - e_n^j\Big)$$

$$\text{for } x = \sum_{i=1}^{n-1} y_i e_i^j + g^j(y) e_n^j \in U^j \text{ with } \big|y - y^j\big| < r^j .$$

It holds that ν_Ω is measurable on $\partial\Omega$ with $|\nu_\Omega| = 1$, and hence $\nu_\Omega \in L^\infty(\partial\Omega; \mathbb{R}^n)$. The definition of ν_Ω is independent of the local representation of the boundary. With the above representation of x, the normal $\nu_\Omega(x)$ is perpendicular to the tangent vectors

$$\tau_k(x) := \partial_{y_k}\Big(\sum_{i=1}^{n-1} y_i e_i^j + g^j(y)e_n^j\Big) = e_k^j + \partial_k g^j(y)e_n^j \quad \text{for } 1 \le k \le n-1.$$

In addition, $\nu_\Omega(x)$ points outward, i.e. $x + \varepsilon\nu_\Omega(x) \notin \Omega$ for $\varepsilon > 0$ sufficiently small, if g is differentiable in y.

Proof (1). In a small open set $U \subset \mathbb{R}^n$ we consider two different representations of $\partial\Omega$ as defined in A8.2, i.e. we consider two coordinate systems $e_1, \dots, e_n$ and $\widetilde{e}_1, \dots, \widetilde{e}_n$, two Lipschitz continuous functions $g : \mathbb{R}^{n-1} \to \mathbb{R}$ and $\widetilde{g} : \mathbb{R}^{n-1} \to \mathbb{R}$ and two bounded open sets $V, \widetilde{V} \subset \mathbb{R}^{n-1}$, such that with $\Gamma := \partial\Omega \cap U$

$$\Big\{\sum_{i=1}^{n-1} y_i e_i + g(y)e_n \;;\; y \in V\Big\} = \Big\{\sum_{i=1}^{n-1} \widetilde{y}_i\widetilde{e}_i + \widetilde{g}(\widetilde{y})\widetilde{e}_n \;;\; \widetilde{y} \in \widetilde{V}\Big\} = \Gamma\,.$$

On setting

$$\psi(y) := \sum_{i=1}^{n-1} y_i e_i + g(y)e_n \quad \text{for } y \in \mathbb{R}^{n-1},$$

and similarly for $\widetilde{\psi}$, we need to show that for every function $f : \Gamma \to \mathbb{R}$ with $\operatorname{supp} f \subset U$ it holds that:

$$f \circ \psi \text{ integrable} \quad \Longleftrightarrow \quad f \circ \widetilde{\psi} \text{ integrable}$$

and

$$\int_V f\big(\psi(y)\big)\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y = \int_{\widetilde{V}} f\big(\widetilde{\psi}(\widetilde{y})\big)\sqrt{1 + |\nabla\widetilde{g}(\widetilde{y})|^2}\,\mathrm{d}\widetilde{y}\,. \tag{A8-4}$$

Consider the transformation $\tau := \widetilde{\psi}^{-1} \circ \psi$, hence $y \mapsto \widetilde{y} = \tau(y)$. Since

$$\big|y^1 - y^2\big| \le \big|\psi(y^1) - \psi(y^2)\big| \le \sqrt{1 + \mathrm{Lip}(g)^2}\big|y^1 - y^2\big|,$$

$\psi : V \to \Gamma$ is a Lipschitz continuous map with a Lipschitz continuous inverse $\psi^{-1} : \Gamma \to V$, and the same holds for $\widetilde{\psi}$. This implies that $\tau : V \to \widetilde{V}$ is bijective and that τ and τ^{-1} are Lipschitz continuous. Hence $f \circ \psi$ is measurable if and only if $f \circ \widetilde{\psi}$ is measurable (use 4.27).

In order to prove the integral identity, we first consider the case where $f \in C_0^0(U)$ and $g \in C^1(V)$. Then also $\widetilde{g}$ is continuously differentiable. To see this, note that the differentiability of g, and therefore ψ, is equivalent to

$$\psi(y) - \psi(y_0) - P_{\psi(y_0)}(\psi(y) - \psi(y_0)) = o(|\psi(y) - \psi(y_0)|)$$

as $y \to y_0$. Hereby P_{x_0} is the orthogonal projection on the tangent space of Γ in $x_0 := \psi(y_0)$. Now, we have $\psi(y) = \widetilde{\psi}(\widetilde{y})$ if $\widetilde{y} = \tau(y)$, and since τ is continuous, it follows that $y \to y_0$ implies $\widetilde{y} \to \widetilde{y}_0$. Hence

$$\widetilde{\psi}(\widetilde{y}) - \widetilde{\psi}(\widetilde{y}_0) - P_{x_0}(\widetilde{\psi}(\widetilde{y}) - \widetilde{\psi}(\widetilde{y}_0)) = o(|\widetilde{\psi}(\widetilde{y}) - \widetilde{\psi}(\widetilde{y}_0)|)$$

as $\widetilde{y} \to \widetilde{y}_0$. But this is equivalent to the differentiability of $\widetilde{\psi}$ and thus also $\widetilde{g}$. Also the differentiability of τ and τ^{-1} is shown. It follows from the (classical) change-of-variables theorem for C^1-transformations that for every function $\widetilde{f} \in C_0^0(\widetilde{V})$

$$\int_{\widetilde{V}} \widetilde{f} \, \mathrm{dL}^{n-1} = \int_V \widetilde{f} \circ \tau \, |\det D\tau| \, \mathrm{dL}^{n-1} .$$

Let $\widetilde{f}(\widetilde{y}) := f\big(\widetilde{\psi}(\widetilde{y})\big)\sqrt{1 + |\nabla \widetilde{g}(\widetilde{y})|^2}$. Then we need to show that

$$\sqrt{1 + |\nabla \widetilde{g} \circ \tau|^2} \, |\det D\tau| = \sqrt{1 + |\nabla g|^2} .$$

But since $(D\widetilde{\psi}) \circ \tau \; D\tau = D\psi$, this reduces to a purely algebraic result for determinants. Hence in this case the integral identity (A8-4) is proved.

If g is only Lipschitz continuous and $f : \Gamma \to \mathbb{R}$ with $f \in C_0^0(U)$, we shall approximate g by continuously differentiable functions. Let $\operatorname{supp} f \circ \psi \subset V_0$ with an open connected subset V_0 satisfying $\overline{V_0} \subset V$. With $\tau = (\tau_1, \ldots, \tau_{n-1})$ we have for $y \in V$ that

$$\sum_{j=1}^{n-1} \tau_j(y)\widetilde{e}_j + \widetilde{g}\big(\tau(y)\big)\widetilde{e}_n \;=\; \sum_{i=1}^{n-1} y_i e_i + g(y) e_n . \tag{A8-5}$$

In the case that $\widetilde{e}_n \neq e_n$, an $(n-2)$-dimensional subspace of $\mathbb{R}^n$ is given by $\operatorname{span}\{e_1, \ldots, e_{n-1}\} \cap \operatorname{span}\{\widetilde{e}_1, \ldots, \widetilde{e}_{n-1}\}$. As L^{n-1} is invariant under orthogonal transformations, we may assume that $\widetilde{e}_i = e_i$ for $1 < i < n$, hence $\operatorname{span}\{e_1, e_n\} = \operatorname{span}\{\widetilde{e}_1, \widetilde{e}_n\}$. (If $\widetilde{e}_n = e_n$ there is nothing to show due to the invariance.) Then

$$\begin{aligned} \tau_j(y) &= y_j \quad \text{ for } 1 < j \le n-1 , \\ \tau_1(y) &= y_1 \widetilde{e}_1 \bullet e_1 + g(y)\, \widetilde{e}_1 \bullet e_n , \\ \widetilde{g}\big(\tau(y)\big) &= y_1 \widetilde{e}_n \bullet e_1 + g(y)\, \widetilde{e}_n \bullet e_n . \end{aligned} \tag{A8-6}$$

Now let $g_\varepsilon := \varphi_\varepsilon * g$ for a standard Dirac sequence $(\varphi_\varepsilon)_{\varepsilon > 0}$ and define continuously differentiable functions $\tau_\varepsilon = (\tau_{\varepsilon 1}, \ldots, \tau_{\varepsilon n-1})$ and ψ_ε by

$$\begin{aligned} \tau_{\varepsilon j}(y) &:= y_j \quad \text{ for } 1 < j \le n-1 , \\ \tau_{\varepsilon 1}(y) &:= y_1 \widetilde{e}_1 \bullet e_1 + g_\varepsilon(y) \widetilde{e}_1 \bullet e_n , \\ \psi_\varepsilon(y) &:= \sum_{i=1}^{n-1} y_i e_i + g_\varepsilon(y) e_n . \end{aligned} \tag{A8-7}$$

We want to show that τ_ε is a diffeomorphism. We have shown that τ^{-1} is Lipschitz continuous, wich implies that there exists a constant $c > 0$ such that for $y \in V_0$ and $h > 0$ sufficiently small ($\{\mathbf{e}_1, \dots, \mathbf{e}_{n-1}\}$ is the canonical basis of $\mathbb{R}^{n-1}$)

$$\begin{aligned} c &\le \tfrac{1}{h}|\tau(y + h\mathbf{e}_1) - \tau(y)| = \tfrac{1}{h}|\tau_1(y + h\mathbf{e}_1) - \tau_1(y)| \\ &= \left|\widetilde{e}_1 \bullet e_1 + \tfrac{1}{h}\big(g(y + h\mathbf{e}_1) - g(y)\big)\, \widetilde{e}_1 \bullet e_n\right| . \end{aligned}$$

The term inside the modulus has to have a fixed sign $\sigma \in \{\pm 1\}$ which by the continuity of g is independent of y. It follows that

$$c \le \sigma\widetilde{e}_1 \bullet e_1 + \frac{1}{h}\big(g(y + h\mathbf{e}_1) - g(y)\big) \cdot \sigma\widetilde{e}_1 \bullet e_n .$$

Since $g_\varepsilon = \varphi_\varepsilon * g$ is a convolution of g, it follows that this convex inequality also holds for g_ε, that is, for ε small,

$$c \le \sigma\widetilde{e}_1 \bullet e_1 + \frac{1}{h}\big(g_\varepsilon(y + h\mathbf{e}_1) - g_\varepsilon(y)\big) \cdot \sigma\widetilde{e}_1 \bullet e_n ,$$

hence also

$$c \le \sigma\widetilde{e}_1 \bullet e_1 + \partial_1 g_\varepsilon(y) \cdot \sigma\widetilde{e}_1 \bullet e_n .$$

Then it follows for $y \in V_0$

$$\sigma \det \mathrm{D}\tau_\varepsilon(y) = \sigma\partial_1\tau_{\varepsilon 1}(y) = \sigma\widetilde{e}_1 \bullet e_1 + \partial_1 g_\varepsilon(y) \cdot \sigma\widetilde{e}_1 \bullet e_n \ge c . \tag{A8-8}$$

This implies that τ_ε is a diffeomorphism because τ_ε is defined as in (A8-7). Hence τ_ε^{-1} exists and therefore, with $\widetilde{y} = \tau_\varepsilon(y)$,

$$\begin{aligned} \widetilde{g}_\varepsilon\big(\tau_\varepsilon(y)\big) &:= y_1\widetilde{e}_n \bullet e_1 + g_\varepsilon(y)\widetilde{e}_n \bullet e_n = \widetilde{e}_n \bullet \psi_\varepsilon(y) , \\ \widetilde{\psi}_\varepsilon(\widetilde{y}) &:= \sum_{j=1}^{n-1} \widetilde{y}_j\widetilde{e}_j + g_\varepsilon(\widetilde{y})\widetilde{e}_n = \psi_\varepsilon(y) , \end{aligned} \tag{A8-9}$$

defines continuously differentiable functions $\widetilde{g}_\varepsilon$ and $\widetilde{\psi}_\varepsilon$.

Now we can show that the integral identity holds. If we define the function $f_\varepsilon := f \circ \widetilde{\psi} \circ \widetilde{\psi}_\varepsilon^{-1}$ on the C^1-surface $\Gamma_\varepsilon := \psi_\varepsilon(V_0)$ we see that for $\varepsilon \to 0$

$$\begin{aligned} &\int_{\widetilde{V}} f\big(\widetilde{\psi}(\widetilde{y})\big)\sqrt{1 + |\nabla\widetilde{g}(\widetilde{y})|^2}\, \mathrm{d}\widetilde{y} \longleftarrow \int_{\widetilde{V}} \underbrace{f\big(\widetilde{\psi}(\widetilde{y})\big)}_{=f_\varepsilon(\widetilde{\psi}_\varepsilon(\widetilde{y}))} \sqrt{1 + |\nabla\widetilde{g}_\varepsilon(\widetilde{y})|^2}\, \mathrm{d}\widetilde{y} \\ &= \int_V \underbrace{f_\varepsilon\big(\psi_\varepsilon(y)\big)}_{=f(\widetilde{\psi}\circ\tau_\varepsilon(y))} \sqrt{1 + |\nabla g_\varepsilon(y)|^2}\, \mathrm{d}y \longrightarrow \int_V \underbrace{f\big(\widetilde{\psi}\circ\tau(y)\big)}_{=f(\psi(y))} \sqrt{1 + |\nabla g(y)|^2}\, \mathrm{d}y . \end{aligned}$$

Indeed, the equality is an equation on Γ_ε and follows from the above step for the C^1-case. The first convergence follows from the fact that $\nabla\widetilde{g}_\varepsilon \to \nabla\widetilde{g}$ with respect to the L^p-norm for every $p < \infty$. In fact, the definition (A8-9) of $\widetilde{g}_\varepsilon$ implies

$$(\mathrm{D}\tau_\varepsilon)^T(\nabla\widetilde{g}_\varepsilon)\circ\tau_\varepsilon = (\mathrm{D}\psi_\varepsilon)^T\widetilde{e}_n$$

and by computing the derivative of $\mathrm{D}\tau_\varepsilon$ using (A8-7) and (A8-8) we obtain that $\nabla\widetilde{g}_\varepsilon$ is bounded in L^∞. Since g_ε is bounded in C^0 it follows from the Arzela-Ascoli theorem that for a subsequence $\varepsilon \to 0$ the uniform limit $g_\varepsilon \to \widehat{g}$ exists. Hence by the definition of g_ε

$$\widehat{g}(\tau(y)) \leftarrow \widetilde{g}_\varepsilon(\tau_\varepsilon(y)) = \widetilde{e}_n \bullet \psi_\varepsilon(y) \to \widetilde{e}_n \bullet \psi(y) = \widetilde{g}(\tau(y))$$

for $\varepsilon \to 0$, that is, $\widehat{g} = \widetilde{g}$. This proves the convergence of the gradients of $\widetilde{g}_\varepsilon$ for a subsequence. The second convergence follow from the uniform convergence $\tau_\varepsilon \to \tau$ and from the convergence $\nabla g_\varepsilon \to \nabla g$ with respect to the L^p-norm for every $p < \infty$. Hence the integral identity (A8-4) holds.

Finally, let f be arbitrary. Since $\widehat{f} := f\circ\psi$ has compact support in V, we can approximate $\widehat{f}$ in $L^1(V)$ by functions $\widehat{f}_i \in C_0^\infty(V)$ as $i \to \infty$. Then we can apply the results above to the functions $f_i := \widehat{f}_i\circ\psi^{-1}$, i.e.,

$$\int_V f_i\circ\psi(y)\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y = \int_{\widetilde{V}} f_i\circ\widetilde{\psi}(\widetilde{y})\sqrt{1 + |\nabla\widetilde{g}(\widetilde{y})|^2}\,\mathrm{d}\widetilde{y}\,.$$

Moreover, we have that

$$\begin{aligned}
&\int_{\widetilde{V}} \Big|f_i\circ\widetilde{\psi}(\widetilde{y}) - f_j\circ\widetilde{\psi}(\widetilde{y})\Big|\sqrt{1 + |\nabla\widetilde{g}(\widetilde{y})|^2}\,\mathrm{d}\widetilde{y} \\
&= \int_V |f_i\circ\psi(y) - f_j\circ\psi(y)|\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y \\
&\le C\Big\|\widehat{f}_i - \widehat{f}_j\Big\|_{L^1(V)} \longrightarrow 0 \qquad \text{as } i, j \to \infty.
\end{aligned}$$

Hence the functions $f_i\circ\widetilde{\psi}$ converge as $i \to \infty$ to a limit in $L^1(\widetilde{V})$. But as $\widehat{f}_i(y) \to \widehat{f}(y)$ for almost all $y \in V$ for a subsequence $i \to \infty$, it follows that also $f_i\circ\widetilde{\psi}(\widetilde{y}) \to f\circ\widetilde{\psi}(\widetilde{y})$ for almost all $\widetilde{y} \in \widetilde{V}$, because τ maps null sets into null sets (see 4.27). This implies that the above limit in $L^1(\widetilde{V})$ must be the function $f\circ\psi$. Hence we obtain the desired integral formula in the general case as well. □

Proof (2). On choosing $f = \mathcal{X}_E$ in (1) for Borel sets $E \subset \partial\Omega$, we obtain that

$$E \longmapsto \mu(E) := \int_{\partial\Omega} \mathcal{X}_E\,\mathrm{d}\mathrm{H}^{n-1}$$

is the $(n-1)$-dimensional Hausdorff measure on $\partial\Omega$ (also denoted by $\mathrm{H}^{n-1}\llcorner\partial\Omega$). Then $L^p(\partial\Omega)$ coincides with the space $L^p(\mu)$ for $\mu = \mathrm{H}^{n-1}\llcorner\partial\Omega$ from Chapter 3. □

Proof (3). In the proof (1) consider the approximation g_ε of g. On noting that $(D\widetilde{\psi}_\varepsilon)\circ\tau_\varepsilon\; D\tau_\varepsilon = D\psi_\varepsilon$, it follows that $(D\widetilde{\psi})\circ\tau\; D\tau = D\psi$ almost everywhere, and so

$$(D\psi)^T = (D\tau)^T (D\widetilde{\psi})^T \circ\tau \quad \text{almost everywhere.} \tag{A8-10}$$

We have that $\nu = \nu_\Omega$, with respect to g, is uniquely defined by

$$(D\psi)^T\nu = 0\,,\ |\nu| = 1\,,\ \nu\bullet e_n < 0\,.$$

Similarly, $\widetilde{\nu}$ is uniquely defined with respect to $\widetilde{g}$. It follows from (A8-10) and (A8-3) that $\widetilde{\nu}\circ\tau = \nu$ almost everywhere. □

A8.6 Trace theorem. Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary and let $1 \le p \le \infty$. Then there exists a unique continuous linear map

$$S : W^{1,p}(\Omega) \longrightarrow L^p(\partial\Omega) \quad (\textit{\textbf{trace operator}})$$

such that

$$Su = u\big|_{\partial\Omega} \quad \text{for } u \in W^{1,p}(\Omega)\cap C^0(\overline{\Omega}).$$

We call Su the ***trace*** or the ***weak boundary values*** of u on $\partial\Omega$.
Notation: In general we write $u(x)$ in place of $(Su)(x)$ for $x \in \partial\Omega$.

Proof. In the case $p = \infty$ it follows from theorem 10.5 that $W^{1,\infty}(\Omega)$ is embedded in $C^{0,1}(\overline{\Omega})$, and so the claim holds trivially. Now let $p < \infty$ and $u \in W^{1,p}(\Omega)$. With the notations as in A8.3, we have that $v := \eta^j u \in W^{1,p}(\Omega^j)$ and for some $\delta > 0$ it holds that

$$v(x) = 0 \quad \text{for } \left|x^j_{,n} - y^j\right| \ge r^j - \delta \text{ and for } x^j_n - g^j(x^j_{,n}) \ge h^j - \delta\,.$$

For $0 < s < h_j$ we define the functions $v_s : \mathbb{R}^{n-1} \to \mathbb{R}$ via

$$v_s(y) := v(y, g^j(y) + s)\,, \qquad \text{where } (y,h) := \sum_{i=1}^{n-1} y_i e^j_i + h e^j_n\,.$$

Being a Lipschitz transformation, $(y,h) \mapsto (y, g^j(y) + h)$ maps measurable functions into measurable functions (recall 4.27), and so it follows from Fubini's theorem that the v_s are measurable functions for almost all s. In addition, $v_s = 0$ for $s \ge h^j - \delta$. Now the essential observation is that for almost all $s_1, s_2 > 0$ and then for almost all $y \in \mathbb{R}^{n-1}$ we have

$$\begin{aligned} v_{s_2}(y) - v_{s_1}(y) &= v(y, g^j(y) + s_2) - v(y, g^j(y) + s_1) \\ &= \int_{g^j(y)+s_1}^{g^j(y)+s_2} \partial_{e^j_n} v(y,h)\,\mathrm{d}h\,. \end{aligned} \tag{A8-11}$$

In order to prove this, we approximate v by functions $w_k \in W^{1,p}(\Omega^j) \cap C^\infty(\Omega^j)$ using theorem 4.24. The identity (A8-11) holds for w_k, and setting $D := \mathrm{B}_{r^j}\left(y^j\right)$ we have that

$$\begin{aligned}&\int_0^{h^j}\int_D \big|v\big(y,g^j(y)+s\big)-w_k\big(y,g^j(y)+s\big)\big|\,\mathrm{d}y\,\mathrm{d}s\\&=\int_{\Omega^j}|v-w_k|\,\mathrm{dL}^n\longrightarrow 0\end{aligned}$$

as $k\to\infty$, and

$$\begin{aligned}&\int_0^{h^j}\int_D\int_{g^j(y)}^{g^j(y)+s}\Big|\partial_{e_n^j}v(y,h)-\partial_{e_n^j}w_k(y,h)\Big|\,\mathrm{d}h\,\mathrm{d}y\,\mathrm{d}s\\&\le h^j\int_{\Omega^j}\Big|\partial_{e_n^j}(v-w_k)\Big|\,\mathrm{dL}^n\longrightarrow 0\end{aligned}$$

as $k\to\infty$. Hence the integrands converge for a subsequence $k\to\infty$ for almost all (y,s). This proves (A8-11). Then the Hölder inequality implies for $s_1<s_2$ that

$$\begin{aligned}&\int_D|v_{s_2}-v_{s_1}|^p\,\mathrm{dL}^{n-1}\le\int_D|s_2-s_1|^{p-1}\int_{g^j(y)+s_1}^{g^j(y)+s_2}\Big|\partial_{e_n^j}v(y,h)\Big|^p\,\mathrm{d}h\,\mathrm{d}y\\&\le|s_2-s_1|^{p-1}\int_{D^j(s_1,s_2)}|\nabla v|^p\,\mathrm{dL}^n\end{aligned}$$

with $D^j(s_1,s_2):=\{x\in\Omega^j\,;\ s_1<x_n^j-g^j(x_{,n}^j)<s_2\}$, and hence

$$\|v_{s_2}-v_{s_1}\|_{L^p(D)}\le|s_2-s_1|^{1-\frac1p}\|\nabla v\|_{L^p(D^j(s_1,s_2))}\,. \tag{A8-12}$$

Since the norm on the right-hand side converges to 0 as $s_1,s_2\to0$, the functions v_s form a Cauchy sequence in $L^p(\mathbb{R}^{n-1})$ as $s\to0$, and hence

$$v_s\to v_0\quad\text{ in } L^p(\mathbb{R}^{n-1}) \text{ as } s\to 0$$

for some $v_0\in L^p(\mathbb{R}^{n-1})$. Now let

$$S^jv(y,g^j(y)):=v_0(y)\,. \tag{A8-13}$$

That is, the weak boundary values are defined as the limit of the function values on hypersurfaces which are a translation of $\partial\Omega$. It follows from A8.5 that $S^jv\in L^p(\partial\Omega)$ with the bound $\big\|S^jv\big\|_{L^p(\partial\Omega)}\le C_j\|v_0\|_{L^p(D)}$. Then on choosing a fixed s^j with $h^j-\delta<s^j<h^j$, so that we then have $v_{s^j}=0$, we obtain from (A8-12), by setting $[s_1,s_2]=[s,s^j]$, that

$$\begin{aligned}\big\|S^jv\big\|_{L^p(\partial\Omega)}&\le C_j\|v_0\|_{L^p(D)}=C_j\|v_{s^j}-v_0\|_{L^p(D)}\\&=C_j\lim_{s\searrow0}\|v_{s^j}-v_s\|_{L^p(D)}\le C_j\cdot(s^j)^{1-\frac1p}\|\nabla v\|_{L^p(\Omega^j)}\,.\end{aligned}$$

In addition, $\|\nabla v\|_{L^p(\Omega^j)}\le C(\eta^j)\cdot\|u\|_{W^{1,p}(\Omega^j)}$. For $u\in W^{1,p}(\Omega)$ we now define

$$Su := \sum_{j=1}^{l} S^j(\eta^j u)\,. \tag{A8-14}$$

In particular, we have that $Su = u\big|_{\partial\Omega}$ if u is continuous on $\overline{\Omega}$. This proves the existence of S. The uniqueness of S follows by establishing that $W^{1,p}(\Omega) \cap C^0(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$, which will be done in A8.7. □

A8.7 Lemma. Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary and let $1 \le p < \infty$ and $m \ge 0$. Then

$$\left\{\, u\big|_{\Omega}\ ;\ u \in C_0^\infty(\mathbb{R}^n)\,\right\} \quad \text{is dense in } W^{m,p}(\Omega).$$

Proof. Following A8.3, we partition u as

$$u = \sum_{j=0}^{l} \eta^j u\,.$$

For the part $\eta^0 u$ choose a standard Dirac sequence $(\varphi_\varepsilon)_{\varepsilon>0}$. Since $\eta^0 \in C_0^\infty(\Omega)$, it follows that $\varphi_\varepsilon * (\eta^0 u) \in C_0^\infty(\Omega)$ for ε sufficiently small, and hence $\varphi_\varepsilon * (\eta^0 u) \to \eta^0 u$ in $W^{m,p}(\Omega)$ as $\varepsilon \to 0$. For $j \ge 1$ let Ω^j and $e_1^j, \ldots, e_n^j$ be as in A8.3. For $\delta > 0$ define

$$\begin{aligned} v_\delta(x) &:= (\eta^j u)(x + \delta e_n^j) \quad \text{for } x \in \Omega_\delta^j,\\ \Omega_\delta^j &:= \left\{\, x \in \mathbb{R}^n\ ;\ \left|x_{,n}^j - y^j\right| < r^j \text{ and } -\delta < x_n^j - g^j(x_{,n}^j) < h^j \,\right\}. \end{aligned}$$

Then $v_{\delta,\varepsilon} := \varphi_\varepsilon * \left(\mathcal{X}_{\Omega_\delta^j} v_\delta\right) \in C_0^\infty(\mathbb{R}^n)$ and, on recalling 4.23, it holds on Ω that $v_{\delta,\varepsilon} = \varphi_\varepsilon * v_\delta \in W^{m,p}(\Omega)$ for ε sufficiently small (so that $\left(1+\mathrm{Lip}(g^j)\right)\cdot\varepsilon < \delta$) with

$$\varphi_\varepsilon * v_\delta \to \eta^j u \quad \text{in } W^{m,p}(\Omega),$$

when first $\varepsilon \searrow 0$ and then $\delta \searrow 0$. This shows that $\eta^j u$ can be approximated in the $W^{m,p}(\Omega)$-norm by functions in $C_0^\infty(\mathbb{R}^n)$, and hence overall also u. □

We now prove some frequently used results on weak boundary values, beginning with integration by parts for Sobolev functions.

A8.8 Weak Gauß's theorem (Weak divergence theorem). Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary.

(1) If $u \in W^{1,1}(\Omega)$, then for $i = 1, \ldots, n$

$$\int_\Omega \partial_i u \,\mathrm{dL}^n = \int_{\partial\Omega} u\nu_i \,\mathrm{dH}^{n-1}\,,$$

where ν is the outer normal to $\partial\Omega$ as defined in A8.5.

(2) Let $1 \le p \le \infty$. If $u \in W^{1,p}(\Omega)$ and $v \in W^{1,p'}(\Omega)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, then for $i = 1, \dots, n$

$$\int_\Omega (u\partial_i v + v\partial_i u)\,\mathrm{dL}^n = \int_{\partial\Omega} uv\nu_i\,\mathrm{dH}^{n-1}\,.$$

Proof (2). It follows from 4.25 that $uv \in W^{1,1}(\Omega)$ with $\partial_i(uv) = u\partial_i v + v\partial_i u$. On recalling A8.7, for $1 < p < \infty$ we approximate u and v by functions in $C^\infty(\mathbb{R}^n)$ and obtain (with S denoting the operator from A8.6) that

$$S(uv) = S(u)\cdot S(v) \quad \text{in } L^1(\partial\Omega)\,. \tag{A8-15}$$

For $p = 1$ we have that $p' = \infty$, and so after modification on a null set, v is in $C^{0,1}(\overline{\Omega})$ (see theorem 10.5). Hence the boundary values of v are well defined and are attained continuously. Now (A8-15) follows from the proof of A8.6. Thus, (2) is reduced to (1). □

Proof (1). On recalling A8.7 and A8.6, we may assume that $u \in C_0^\infty(\mathbb{R}^n)$. Following A8.3, we partition u into $\eta^j u$, $j = 0, \dots, l$. For $\eta^0 u \in C_0^\infty(\Omega)$ the boundary integral vanishes and the formula follows from integration by parts in the i-th coordinate direction. For $j \ge 1$ the function $\eta^j u$ is defined on the local set Ω^j. Hence on applying an orthogonal transformation to the canonical Euclidean coordinate system, we need to prove the desired result for functions $u \in C_0^\infty(\mathbb{R}^n)$ and the domain

$$\Omega = \left\{ (y,h) \in \mathbb{R}^n \;;\; h > g(y) \right\}$$

with a Lipschitz continuous function $g : \mathbb{R}^{n-1} \to \mathbb{R}$. By A8.5(3), the normal ν is then defined by

$$\nu(y,g(y)) := \frac{(\nabla g(y), -1)}{\sqrt{1 + |\nabla g(y)|^2}} \qquad \text{for } y \in \mathbb{R}^{n-1}.$$

Hence we need to show that

$$\begin{aligned} \int_\Omega \nabla u(x)\,\mathrm{d}x &= \int_{\mathbb{R}^{n-1}} (u\nu)\big(y,g(y)\big)\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y \\ &= \int_{\mathbb{R}^{n-1}} u(y,g(y))(\nabla g(y), -1)\,\mathrm{d}y\,. \end{aligned} \tag{A8-16}$$

When g is continuously differentiable, this is the classical Gauß's theorem, which can be shown for instance as follows: Let $v(y,s) := u\big(y, g(y) + s\big)$. Then

$$\begin{aligned} \partial_n v(y,s) &= \partial_n u\big(y, g(y) + s\big)\,, \\ \partial_i v(y,s) &= \partial_i u\big(y, g(y) + s\big) + \partial_i g(y)\partial_n u\big(y, g(y) + s\big) \qquad \text{for } i < n, \end{aligned}$$

and hence

$$
\begin{aligned}
\int_\Omega \nabla u(x)\,\mathrm{d}x &= \int_{\mathbb{R}^{n-1}} \int_0^\infty \nabla u\big(y, g(y)+s\big)\,\mathrm{d}s\,\mathrm{d}y \\
&= \int_{\mathbb{R}^{n-1}} \int_0^\infty \big(\nabla v - \partial_n v \cdot (\nabla g, 0)\big)(y,s)\,\mathrm{d}s\,\mathrm{d}y \\
&= \sum_{i=1}^{n-1} \Big(\int_0^\infty \Big(\int_{\mathbb{R}^{n-1}} \partial_i v(y,s)\,\mathrm{d}y\Big)\,\mathrm{d}s\Big)\mathbf{e}_i \\
&\quad - \int_{\mathbb{R}^{n-1}} \Big(\int_0^\infty \partial_n v(y,s)\,\mathrm{d}s\Big)\big(\nabla g(y), -1\big)\,\mathrm{d}y\,.
\end{aligned}
$$

Integration by parts with respect to y_i yields for $i < n$, since the support of $v(\bullet, s)$ is compact, that

$$
\int_{\mathbb{R}^{n-1}} \partial_i v(y,s)\,\mathrm{d}s = 0\,,
$$

and integration by parts with respect to s gives

$$
\int_0^\infty \partial_n v(y,s)\,\mathrm{d}s = -v(y,0) = -u\big(y, g(y)\big)\,.
$$

Now we use convolution to approximate the Lipschitz continuous function g by continuously differentiable functions g_k. Letting $\Omega_k := \{(y,h) \in \mathbb{R}^n\,;\ h > g_k(y)\}$ we have that $\mathcal{X}_{\Omega_k} \to \mathcal{X}_\Omega$ as $k \to \infty$ in $L^1(\mathbb{R}^n) \cap \mathrm{B}_R(0)$ for every R and $u(\bullet, g_k) \to u(\bullet, g)$ uniformly, because $g_k \to g$ locally uniformly, and also (recall 4.15)

$$
\nabla g_k \to \nabla g \quad \text{in } L^p(\mathrm{B}_R(0)) \quad \text{for every } p < \infty \text{ and every } R.
$$

Hence in (A8-16) we can pass to the limit for g_k. This yields the desired result. ☐

The following result is a generalization of E3.7 to the n-dimensional case.

A8.9 Lemma. Let $g : \mathbb{R}^{n-1} \to \mathbb{R}$ be Lipschitz continuous, let

$$
\Omega_\pm := \{(y,h) \in \mathbb{R}^n\,;\ \pm(h - g(y)) > 0\}\,,
$$

and let $u : \mathbb{R}^n \to \mathbb{R}$ with $u\big|_{\Omega_+} \in W^{1,1}(\Omega_+)$ and $u\big|_{\Omega_-} \in W^{1,1}(\Omega_-)$. Then, on denoting by $S_\pm$ the trace operators with respect to the domains $\Omega_\pm$ from A8.6,

$$
u \in W^{1,1}(\mathbb{R}^n) \quad \Longleftrightarrow \quad S_+\big(u\big|_{\Omega_+}\big) = S_-\big(u\big|_{\Omega_-}\big)\,.
$$

Corollary: Concerning the removability of singularities in Sobolev spaces we have the following result: If $N \subset \mathbb{R}^{n-1}$ is a closed Lebesgue null set and $A := \{(y, g(y))\,;\ y \in N\}$ with g as above, then for every open set $\Omega \subset \mathbb{R}^n$

$$
u \in W^{1,1}(\Omega \setminus A) \quad \Longleftrightarrow \quad u \in W^{1,1}(\Omega)\,.
$$

Proof $\Rightarrow$. Setting $u_s(y) := u(y, g(y)+s)$ for $s \in \mathrm{I\!R}$, it holds that (see (A8-12) with $p = 1$)

$$\int_{\mathrm{I\!R}^{n-1}} |u_\varepsilon - u_{-\varepsilon}| \,\mathrm{dL}^{n-1} \le \int_{\mathrm{I\!R}^{n-1}} \int_{g(y)-\varepsilon}^{g(y)+\varepsilon} |\nabla u(y,h)| \,\mathrm{d}h \,\mathrm{d}y \longrightarrow 0$$

as $\varepsilon \searrow 0$, and so $S_+\bigl(u\big|_{\Omega_+}\bigr) = S_-\bigl(u\big|_{\Omega_-}\bigr)$ by the definition of the trace operator in (A8-13). $\square$

Proof $\Leftarrow$. Define $u_+ = u\big|_{\Omega_+}$ and $u_- = u\big|_{\Omega_-}$. Let $\nu_\pm$ denote the outer normal to $\Omega_\pm$. Then it follows from A8.8(2) for $\zeta \in C_0^\infty(\mathrm{I\!R}^n)$ that

$$\begin{aligned}
\int_{\mathrm{I\!R}^n} (u\nabla\zeta + \zeta\nabla u) \,\mathrm{dL}^n &= \int_{\Omega_+} (u\nabla\zeta + \zeta\nabla u) \,\mathrm{dL}^n + \int_{\Omega_-} (u\nabla\zeta + \zeta\nabla u) \,\mathrm{dL}^n \\
&= \int_{\partial\Omega_+} \zeta S_+(u_+)\nu_+ \,\mathrm{dH}^{n-1} + \int_{\partial\Omega_-} \zeta S_-(u_-)\nu_- \,\mathrm{dH}^{n-1} \\
&= \int_{\mathrm{graph}(g)} \zeta \cdot \underbrace{(S_+(u_+)\nu_+ + S_-(u_-)\nu_-)}_{=0} \,\mathrm{dH}^{n-1} = 0\,,
\end{aligned}$$

because $\nu_- = -\nu_+$ and $S_+(u_+) = S_-(u_-)$. $\square$

We now show that functions in $W_0^{1,p}(\Omega)$ have weak boundary values 0.

A8.10 Lemma. Let $\Omega \subset \mathrm{I\!R}^n$ be open and bounded with Lipschitz boundary and let $1 \le p < \infty$. Let S be the trace operator from A8.6. Then

$$W_0^{1,p}(\Omega) = \{u \in W^{1,p}(\Omega)\,;\ Su = 0\}\,.$$

Proof $\subset$. Every function $u \in W_0^{1,p}(\Omega)$ can be approximated by $C_0^\infty(\Omega)$-functions u_i as $i \to \infty$. The properties of the trace operator then imply that $0 = Su_i \to Su$ in $L^p(\partial\Omega)$. $\square$

Proof $\supset$. Let $u \in W^{1,p}(\Omega)$ with $Su = 0$. Choosing η^j as in A8.3, it follows (see (A8-15)) that $S(\eta^j u) = \eta^j S(u) = 0$ on $\partial\Omega$ for $j = 1, \dots, l$. Now define for $j = 1, \dots, l$

$$v_j(x) := \begin{cases} (\eta^j u)(x) & \text{for } x \in \Omega^j, \\ 0 & \text{otherwise.} \end{cases}$$

Then A8.9 implies that $v_j \in W^{1,p}(\mathrm{I\!R}^n)$, and hence for $\delta > 0$ also $v_{j\delta} \in W^{1,p}(\mathrm{I\!R}^n)$, where

$$v_{j\delta}(x) := v_j(x - \delta e_n^j)\,,$$

and $v_{j\delta} \to v_j$ in $W^{1,p}(\mathrm{I\!R}^n)$ as $\delta \to 0$. Consequently,

$$u_\delta := \eta^0 u + \sum_{j=1}^{l} v_{j\delta} \longrightarrow u \quad \text{in } W^{1,p}(\Omega) \text{ as } \delta \to 0.$$

Since u_δ has compact support in Ω, it can be approximated in $W^{1,p}(\Omega)$ with the help of convolution by functions in $C_0^\infty(\Omega)$. $\square$

A8.11 Remark. Results for Sobolev functions on domains with Lipschitz boundary can also be proved by locally straightening the boundary. In the local situation at the boundary, i.e. $\Omega = \Omega_+$ with the notations as in A8.9, this means that we consider

$$\begin{aligned} \widetilde{\Omega} &:= \{(y,h) \in \mathbb{R}^n ;\ h > 0\} , \\ \widetilde{u}(y,h) &:= u\big(y, g(y) + h\big) \quad \text{ for } (y,h) \in \widetilde{\Omega} . \end{aligned}$$

It holds that: If $1 \le p \le \infty$ and $u \in W^{1,p}(\Omega)$, then $\widetilde{u} \in W^{1,p}(\widetilde{\Omega})$ with the ***chain rule***

$$\begin{aligned} \partial_n \widetilde{u}(y,h) &= \partial_n u\big(u, g(y) + h\big) , \\ \partial_i \widetilde{u}(y,h) &= \partial_i u\big(y, g(y) + h\big) + \partial_i g(y) \partial_n u\big(y, g(y) + h\big) \end{aligned} \tag{A8-17}$$

for $i < n$.

Proof. Let $\tau(y,h) := \big(y, g(y)+h\big)$. For $v \in L^p(\Omega)$ with $p < \infty$ it follows from Fubini's theorem that $v\circ\tau \in L^p(\widetilde{\Omega})$, with

$$\begin{aligned} &\int_\Omega |v|^p \, \mathrm{dL}^n = \int_{\mathbb{R}^{n-1}} \int_{g(y)}^\infty |v(y,h)|^p \, \mathrm{d}h \, \mathrm{d}y \\ &= \int_{\mathbb{R}^{n-1}} \int_0^\infty \big|v\big(y, g(y) + h\big)\big|^p \, \mathrm{d}h \, \mathrm{d}y = \int_{\widetilde{\Omega}} |v\circ\tau|^p \, \mathrm{dL}^n . \end{aligned} \tag{A8-18}$$

Hence we have that $\|v\|_{L^p(\Omega)} = \|v\circ\tau\|_{L^p(\widetilde{\Omega})}$ for $1 \le p \le \infty$. This shows that the right-hand sides in (A8-17) lie in $L^p(\widetilde{\Omega})$, and so (A8-17) (by the definition of the weak derivatives) only needs to be shown locally in $\widetilde{\Omega}$ for the case $p = 1$.

We approximate g by $g_\varepsilon := \varphi_\varepsilon * g$ with a standard Dirac sequence $(\varphi_\varepsilon)_{\varepsilon>0}$. On setting $\tau_\varepsilon(y,h) := \big(y, g_\varepsilon(y) + h\big)$, let

$$\widetilde{u}_\varepsilon := u\circ\tau_\varepsilon \quad \text{ on } \widetilde{\Omega}_\varepsilon := \tau_\varepsilon^{-1}(\Omega) .$$

By 4.26, we have $\widetilde{u}_\varepsilon \in W^{1,1}(\widetilde{\Omega}_\varepsilon)$ and the chain rule (A8-17) holds for $\widetilde{u}_\varepsilon$. We note that $g_\varepsilon \to g$ locally uniformly as $\varepsilon \to 0$ and $\nabla g_\varepsilon \to \nabla g$ in $L^q_{\mathrm{loc}}(\mathbb{R}^{n-1})$ for every $q < \infty$, and so $\nabla g_\varepsilon \to \nabla g$ almost everywhere for a subsequence $\varepsilon \to 0$. Moreover, the ∇g_ε are bounded in $L^\infty_{\mathrm{loc}}(\mathbb{R}^{n-1})$. If we can show that for $v \in L^1_{\mathrm{loc}}(\Omega)$ and for every $D \subset\subset \Omega$

$$v\circ\tau_\varepsilon \to v\circ\tau \quad \text{ as } \varepsilon \to 0 \text{ in } L^1\big(\tau^{-1}(D)\big) , \tag{A8-19}$$

then this implies the convergence of $u\circ\tau_\varepsilon$ and $(\partial_i u)\circ\tau_\varepsilon$, and we can pass to the limit in the chain rule (A8-17).

Now it follows from (A8-18) that (A8-19) is equivalent to

$$v\circ\tau_\varepsilon\circ\tau^{-1} \to v \quad \text{ as } \varepsilon \to 0 \text{ in } L^1(D) . \tag{A8-20}$$

Here, we approximate v in the L^1-norm by continuous functions v_k. These functions satisfy (A8-20) and therefore we have (cf. the proof of 4.15(1))

$$\begin{aligned}\big\|v\circ\tau_\varepsilon\circ\tau^{-1}-v\big\|_{L^1(D)} &\le \big\|v\circ\tau_\varepsilon\circ\tau^{-1}-v_k\circ\tau_\varepsilon\circ\tau^{-1}\big\|_{L^1(D)}\\ &\qquad+\big\|v_k\circ\tau_\varepsilon\circ\tau^{-1}-v_k\big\|_{L^1(D)}+\|v_k-v\|_{L^1(D)}\\ &\le (C(\tau_\varepsilon\circ\tau^{-1})+1)\|v_k-v\|_{L^1(D)}+\big\|v_k\circ\tau_\varepsilon\circ\tau^{-1}-v_k\big\|_{L^1(D)}\,,\end{aligned}$$

where $C(\tau_\varepsilon\circ\tau^{-1})$ converges to 1 as $\varepsilon\to 0$. Thus (A8-20) also holds for v. ☐

A further consequence of A8.9 is:

A8.12 Extension theorem. Let $\Omega\subset\mathbb{R}^n$ be open and bounded with Lipschitz boundary and let $1\le p\le\infty$. Then, for $\delta>0$, there exists an extension operator

$$E: W^{1,p}(\Omega)\longrightarrow W_0^{1,p}\big(\mathrm{B}_\delta(\Omega)\big)\,,$$

i.e. E is linear, continuous, and such that $(Eu)\big|_\Omega=u$ for all $u\in W^{1,p}(\Omega)$.

Proof. We treat E similarly to the operator S in (A8-14). Hence it is sufficient to consider the local situation near the boundary (cf. the proof of A8.8(1)). Let $\Omega=\Omega_+$ with $\Omega_\pm$ as in A8.9. Choose a cut-off function $\eta\in C^\infty(\mathbb{R}^n)$ with $\eta=1$ in $\mathrm{B}_{\frac{\delta}{2}}(\Omega)$ and $\eta=0$ in $\mathbb{R}^n\setminus\mathrm{B}_\delta(\Omega)$. Then, define $Eu:=\eta\widetilde{u}$ with

$$\widetilde{u}(y,h):=\begin{cases}u(y,h) & \text{for } h>g(y),\\ u\big(y,2g(y)-h\big) & \text{for } h<g(y).\end{cases}$$

(For $p=\infty$ it follows from theorem 10.5 that this defines a $C^{0,1}$-extension of u.) For $p<\infty$ it follows similarly to the proof of A8.11 that $\widetilde{u}\in W^{1,p}(\Omega_-)$, with

$$\begin{aligned}\|\widetilde{u}\|_{L^p(\Omega_-)} &= \|u\|_{L^p(\Omega_+)}\,,\\ \|\nabla\widetilde{u}\|_{L^p(\Omega_-)} &\le \big(2+\mathrm{Lip}(g)\big)\ \|\nabla u\|_{L^p(\Omega_+)}\,.\end{aligned}$$

Consequently, $Eu\in W^{1,p}(\Omega_-)$ with

$$\|Eu\|_{W^{1,p}(\Omega_-)}\le C\|u\|_{W^{1,p}(\Omega_+)}\,.$$

Then by the definition of the trace operator in (A8-13) it holds that for a sequence $\varepsilon\searrow 0$ and for almost all y

$$\begin{aligned}S_-(Eu)\big(y,g(y)\big) &\longleftarrow Eu\big(y,g(y)-\varepsilon\big)\\ &= u\big(y,g(y)+\varepsilon\big)\longrightarrow S_+(u)\big(y,g(y)\big)\end{aligned}$$

Now A8.9 yields that $Eu\in W^{1,p}(\mathbb{R}^n)$. ☐

The following theorem implies that sets of the form

$$M := \{u \in W^{1,2}(\Omega)\,;\ \varphi(u) = g \text{ on } \partial\Omega\} \tag{A8-21}$$

are weakly sequentially closed in $W^{1,2}(\Omega)$, if $\varphi : \mathbb{R} \to \mathbb{R}$ is continuous and $g : \partial\Omega \to \mathbb{R}$ is measurable.

A8.13 Embedding theorem onto the boundary. If $\Omega \subset \mathbb{R}^n$ is open and bounded with Lipschitz boundary, then for $1 \le p < \infty$ and $u_k, u \in W^{1,p}(\Omega)$ it holds that:

$$\begin{matrix} u_k \to u \text{ weakly in } W^{1,p}(\Omega) \\ \text{as } k \to \infty \end{matrix} \quad \Longrightarrow \quad \begin{matrix} u_k \to u \text{ (strongly) in } L^p(\partial\Omega) \\ \text{as } k \to \infty\,. \end{matrix}$$

Proof. Without loss of generality let $u = 0$. If $\eta \in C^\infty(\mathbb{R}^n)$, then also $\eta u_k \to 0$ weakly in $W^{1,p}(\Omega)$, and so it follows from A8.3 and A8.1 that we only need to consider the local situation on the boundary. Hence let $\Omega = \Omega_+$ as in A8.9 and let the supports of u_k, u be contained in a bounded set of $\mathbb{R}^n$. On recalling (A8-12), the functions $u_{ks}(y) := u_k(y, g(y) + s)$ satisfy for almost all ε, s with $0 < \varepsilon < s$ the bound

$$\int_{\mathbb{R}^{n-1}} |u_{ks} - u_{k\varepsilon}|^p \, \mathrm{dL}^{n-1} \le |s - \varepsilon|^{p-1} \int_{E_{\varepsilon,s}} |\nabla u_k|^p \, \mathrm{dL}^n\,,$$

where $E_{\varepsilon,s} := \{(y,h) \in \mathbb{R}^n\,;\ \varepsilon < h - g(y) < s\}$. Let $\delta > 0$. Then for almost all s with $0 < \varepsilon \le s \le \delta$, on setting $C = 2^{p-1}$ (see (3-13)), we have that

$$\int_{\mathbb{R}^{n-1}} |u_{k\varepsilon}|^p \, \mathrm{dL}^{n-1} \le C \int_{\mathbb{R}^{n-1}} |u_{ks}|^p \, \mathrm{dL}^{n-1} + C\delta^{p-1} \int_{E_{0,\delta}} |\nabla u_k|^p \, \mathrm{dL}^n\,.$$

On letting $\varepsilon \to 0$, we have that $u_{k\varepsilon} \to u_{k0}$ in $L^p(\mathbb{R}^{n-1})$, where u_{k0} are the weak boundary values of u_k. Then integrating this inequality over $s \in [\frac{\delta}{2}, \delta]$ and dividing by $\frac{\delta}{2}$ yields that

$$\int_{\mathbb{R}^{n-1}} |u_{k0}|^p \, \mathrm{dL}^{n-1} \le \frac{2C}{\delta} \int_{E_{\frac{\delta}{2},\delta}} |u_k|^p \, \mathrm{dL}^n + C\delta^{p-1} \int_{E_{0,\delta}} |\nabla u_k|^p \, \mathrm{dL}^n\,.$$

It follows from Rellich's embedding theorem A8.4 that the first term on the right-hand side converges to 0 for every δ. If $p > 1$ then the second term converges to 0 as $\delta \to 0$, since the functions ∇u_k are bounded in $L^p(\Omega_+)$. In the case $p = 1$ it follows from the following theorem that the integral in the second term converges to 0 uniformly in k as $\delta \to 0$, because the ∇u_k converge weakly to 0 in $L^1(\Omega_+; \mathbb{R}^n)$ and because they have supports in a bounded set. This yields the desired result. ☐

A8.14 Weak sequential compactness in $L^1(\mu)$. Let $(S, \mathcal{B}, \mu)$ be a measure space and let $M \subset L^1(\mu; \mathbb{R}^m)$. Then every sequence in M contains a subsequence that converges weakly in $L^1(\mu; \mathbb{R}^m)$ if and only if

(1) M is bounded in $L^1(\mu; \mathbb{R}^m)$.

(2) It holds that

$$\sup_{f \in M} \int_E |f| \, \mathrm{d}\mu \longrightarrow 0 \quad \text{ as } \mu(E) \to 0 \, .$$

(3) There exist sets $S_k \in \mathcal{B}$, for $k \in \mathbb{N}$, with $\mu(S_k) < \infty$, such that

$$\sup_{f \in M} \int_{S \setminus S_k} |f| \, \mathrm{d}\mu \longrightarrow 0 \quad \text{ as } k \to \infty \, .$$

Remark: If $\mu(S) < \infty$, condition (3) is not necessary, choose $S_k = S$.

Proof $\Rightarrow$. (1) follows via an indirect argument from 8.3(5).

Assume that (2) is false. Hence there exist a $c > 0$ and measurable sets E_n as well as $f_n \in M$ for $n \in \mathbb{N}$ such that

$$\mu(E_n) \to 0 \text{ as } n \to \infty \quad \text{and} \quad \int_{E_n} |f_n| \, \mathrm{d}\mu \geq c \text{ for all } n \, .$$

From this it follows that there exist $\widetilde{E}_n \in \mathcal{B}$ with $\mu(\widetilde{E}_n) \to 0$ as $n \to \infty$ and

$$\left| \int_{\widetilde{E}_n} f_n \, \mathrm{d}\mu \right| \geq \frac{c}{2m} \, . \tag{A8-22}$$

To see this, let $A_j^{\pm} := \{ x \in S \, ; \; \pm f_n(x) \bullet \mathbf{e}_j > 0 \}$ for $j = 1, \ldots, m$. Then

$$\begin{aligned} \int_{E_n} |f_n| \, \mathrm{d}\mu &\leq \sum_{j=1}^m \left(\int_{E_n \cap A_j^+} |f_n \bullet \mathbf{e}_j| \, \mathrm{d}\mu + \int_{E_n \cap A_j^-} |f_n \bullet \mathbf{e}_j| \, \mathrm{d}\mu \right) \\ &= \sum_{j=1}^m \left(\left| \int_{E_n \cap A_j^+} f_n \bullet \mathbf{e}_j \, \mathrm{d}\mu \right| + \left| \int_{E_n \cap A_j^-} f_n \bullet \mathbf{e}_j \, \mathrm{d}\mu \right| \right) \, , \end{aligned}$$

which means that for some j (which depends on n) we have that

$$\left| \int_{E_n \cap A_j^+} f_n \bullet \mathbf{e}_j \, \mathrm{d}\mu \right| \geq \frac{c}{2m} \quad \text{or} \quad \left| \int_{E_n \cap A_j^-} f_n \bullet \mathbf{e}_j \, \mathrm{d}\mu \right| \geq \frac{c}{2m} \, .$$

Let $\widetilde{E}_n := E_n \cap A_j^+$ in the first case, and $\widetilde{E}_n := E_n \cap A_j^-$ in the second case. Then

$$\frac{c}{2m} \leq \left| \int_{\widetilde{E}_n} f_n \bullet \mathbf{e}_j \, \mathrm{d}\mu \right| = \left| \mathbf{e}_j \bullet \int_{\widetilde{E}_n} f_n \, \mathrm{d}\mu \right| \leq \left| \int_{\widetilde{E}_n} f_n \, \mathrm{d}\mu \right| ,$$

and $\mu(\widetilde{E}_n) \le \mu(E_n) \to 0$ as $n \to \infty$. This proves (A8-22). It follows from the assumption on M that there exists a subsequence $n \to \infty$ (there is no extra notation for the subsequence) such that for all μ-measurable E the limit

$$\lim_{n\to\infty} \lambda_n(E) \quad \text{exists, with} \quad \lambda_n(E) := \int_E f_n \,\mathrm{d}\mu \,. \tag{A8-23}$$

Since for every n we have $\lambda_n(E) \to 0$ as $\mu_n(E) \to 0$, the following theorem A8.15 yields a contradiction to $\left|\lambda_n(\widetilde{E}_n)\right| \ge \frac{c}{2m}$.

Now assume that (3) is false, i.e. there exists a $c > 0$ such that for all $E \in \mathcal{B}$ with $\mu(E) < \infty$

$$\int_{S\setminus E} |f| \,\mathrm{d}\mu \ge c \quad \text{for an } f \in M \,. \tag{A8-24}$$

Moreover, for all $f \in L^1(\mu; \mathbb{R}^m)$ and $\varepsilon > 0$,

$$\int_{S\setminus E} |f| \,\mathrm{d}\mu \le \varepsilon \quad \text{for an } E \in \mathcal{B} \text{ with } \mu(E) < \infty \,, \tag{A8-25}$$

because there exists a step function g with $\|f - g\|_{L^1} \le \varepsilon$, and then $E := \{x \in S\,;\ g(x) \ne 0\}$ has finite measure.

On combining (A8-25) and (A8-24) we inductively choose $f_n \in M$ and $E_n \in \mathcal{B}$ with $\mu(E_n) < \infty$ and $E_n \subset E_{n+1}$ such that

$$\int_{S\setminus E_{n+1}} |f_n| \,\mathrm{d}\mu \le \frac{1}{n} \quad \text{and} \quad \int_{S\setminus E_{n+1}} |f_{n+1}| \,\mathrm{d}\mu \ge c \,.$$

Then it holds for $n \ge \frac{2}{c}$ that

$$\int_{E_{n+1}\setminus E_n} |f_n| \,\mathrm{d}\mu = \int_{S\setminus E_n} |f_n| \,\mathrm{d}\mu - \int_{S\setminus E_{n+1}} |f_n| \,\mathrm{d}\mu \ge \frac{c}{2} \,.$$

Next, as in the proof of (A8-22), there exist measurable sets $\widetilde{E}_n \subset E_{n+1} \setminus E_n$ such that

$$\left| \int_{\widetilde{E}_n} f_n \,\mathrm{d}\mu \right| \ge \frac{c}{4m} \,,$$

and for a subsequence $n \to \infty$ the corresponding λ_n satisfy the above property (A8-23). We now consider the measure space $(\widetilde{S}, \widetilde{\mathcal{B}}, \widetilde{\mu})$ with

$$\widetilde{S} := \bigcup_{n\in\mathbb{N}} E_n \,, \quad \widetilde{\mathcal{B}} := \{E \cap \widetilde{S}\,;\ E \in \mathcal{B}\} \,,$$

$$\widetilde{\mu}(E) := \sum_{j\in\mathbb{N}} 2^{-j} \frac{\mu(E \cap E_j \setminus E_{j-1})}{1 + \mu(E_j \setminus E_{j-1})} \,,$$

where $E_0 := \emptyset$. Since $\widetilde{\mu}(E) \to 0$ implies $\mu(E \cap E_j \setminus E_{j-1}) \to 0$ for all j, and since, for fixed n, for $E \subset \widetilde{S} \setminus E_j$ it holds that

$$|\lambda_n(E)| \le \int_{\widetilde{S}\setminus E_j} |f_n| \,\mathrm{d}\mu \longrightarrow 0 \quad \text{as } j \to \infty\,,$$

we obtain that $|\lambda_n(E)| \to 0$ for $E \in \widetilde{\mathcal{B}}$ with $\widetilde{\mu}(E) \to 0$ for fixed n. Combining the following theorem A8.15 applied to the measure $\widetilde{\mu}$ and the facts that

$$\left|\lambda_n(\widetilde{E}_n)\right| \ge \frac{c}{4m} \quad \text{and} \quad \widetilde{\mu}(\widetilde{E}_n) \le 2^{-n} \to 0 \text{ as } n \to \infty,$$

we arrive at a contradiction. □

Proof $\Leftarrow$ *for regular measures.* Let $S \subset \mathbb{R}^n$ be compact and let μ be a non-negative measure in $rca(S)$. We may assume that $m = 1$. For every sequence $(f_n)_{n\in\mathbb{N}}$ in M it follows from (1) and A3.17(2) that

$$\lambda_n(E) := \int_E f_n \,\mathrm{d}\mu$$

defines a bounded sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $rca(S)$. By 8.6(2), there exists a $\lambda \in rca(S)$ such that for a subsequence $n \to \infty$,

$$\int_S g \,\mathrm{d}\lambda_n \to \int_S g \,\mathrm{d}\lambda \quad \text{for all } g \in C^0(S)\,. \tag{A8-26}$$

If E is a μ-null set, then, on recalling that μ is regular, for $\varepsilon > 0$ there exists a relatively in S open set U with $E \subset U$ and $\mu(U) \le \varepsilon$. Moreover, as λ is regular, there exist finitely many disjoint closed sets $K_j \subset U$ such that

$$|\lambda|(U) \le \varepsilon + \sum_j |\lambda(K_j)|\,.$$

For $\delta > 0$ choose $g_j \in C^0(S)$ with $\mathcal{X}_{K_j} \le g_j \le \mathcal{X}_{\mathrm{B}_\delta(K_j)}$. Then it follows that

$$|\lambda|(U) \le \varepsilon + \sum_j |\lambda|\big(\mathrm{B}_\delta(K_j) \setminus K_j\big) + \sum_j \left|\int_S g_j \,\mathrm{d}\lambda\right|$$

and

$$\begin{aligned}\sum_j \left|\int_S g_j \,\mathrm{d}\lambda\right| &\longleftarrow \sum_j \left|\int_S g_j \,\mathrm{d}\lambda_n\right| \quad (\text{as } n \to \infty)\\ &= \sum_j \left|\int_S g_j f_n \,\mathrm{d}\mu\right| \le \int_S \Big(\sum_j g_j\Big)|f_n| \,\mathrm{d}\mu \le \int_U |f_n| \,\mathrm{d}\mu\,,\end{aligned}$$

where we observe that the last inequality holds for δ sufficiently small, because then the sets $\mathrm{B}_\delta(K_j)$ are disjoint subsets of U. Letting $\delta \searrow 0$ and noting assumption (2) we get

$$|\lambda|(U) \leq \varepsilon + \sup_{f \in M} \int_U |f| \, \mathrm{d}\mu \longrightarrow 0 \quad \text{as } \varepsilon \to 0.$$

This shows that E is also a $|\lambda|$-null set. Hence we can apply the Radon-Nikodým theorem 6.11 and obtain that there exists an $f \in L^1(\mu)$ with

$$\lambda(E) = \int_E f \, \mathrm{d}\mu$$

for all μ-measurable sets E. It follows from (A8-26) that

$$\int_S g f_n \, \mathrm{d}\mu \longrightarrow \int_S g f \, \mathrm{d}\mu \quad \text{as } n \to \infty \tag{A8-27}$$

for all $g \in C^0(S)$. On recalling 6.12, we need to show that this also holds for all $g \in L^\infty(\mu)$. First let $g = \mathcal{X}_E$ with a μ-measurable set E. For $\varepsilon > 0$ choose K closed and U relatively open in S such that $K \subset E \subset U$ with $\mu(U \setminus K) \leq \varepsilon$ and $\widetilde{g} \in C^0(S)$ with $\mathcal{X}_K \leq \widetilde{g} \leq \mathcal{X}_U$. Then

$$\left| \int_E f_n \, \mathrm{d}\mu - \int_E f \, \mathrm{d}\mu \right| \leq \left| \int_S \widetilde{g}(f_n - f) \, \mathrm{d}\mu \right| + \sup_{n'} \int_{U \setminus K} \big(|f_{n'}| + |f| \big) \, \mathrm{d}\mu \,,$$

where, thanks to (2), the second term converges to 0 as $\varepsilon \to 0$. Since the first term converges to 0 as $n \to \infty$ by (A8-27), we obtain

$$\int_E f_n \, \mathrm{d}\mu \longrightarrow \int_E f \, \mathrm{d}\mu \quad \text{as } n \to \infty.$$

Recalling that the characteristic functions span a dense subspace of $L^\infty(\mu)$ then yields that (A8-27) also holds for all $g \in L^\infty(\mu)$. □

Proof $\Leftarrow$ *for bounded measures.* The idea is to use a separable analogue of $C^0(S)$ in the above proof. As before, let $m = 1$. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in M, and let

$$g_n = \sum_{j=1}^{k_n} \alpha_{nj} \mathcal{X}_{E_{nj}} \quad \text{with } \mu(E_{nj}) < \infty$$

be step functions with $\| f_n - g_n \|_{L^1} \leq \frac{1}{n}$. On noting that for every $n_0 \in \mathbb{N}$ it holds that

$$\int_E |g_n| \, \mathrm{d}\mu \leq \max_{i \leq n_0} \int_E |g_i| \, \mathrm{d}\mu + \frac{1}{n_0} + \sup_{f \in M} \int_E |f| \, \mathrm{d}\mu \,,$$

we have that $\{g_n \,;\, n \in \mathbb{N}\}$ also satisfies the assumption (2), and it is sufficient to show that $(g_n)_{n \in \mathbb{N}}$ contains a weakly convergent subsequence.

Now the algebra $\mathcal{B}_0$ induced by the set $\{E_{nj} \,;\, j \leq k_n \,,\, n \in \mathbb{N}\}$ is countable. Hence it follows from (1) that with the help of a diagonalization procedure we obtain a subsequence such that (without special notation for the subsequence)

$$\lambda(E) := \lim_{n\to\infty} \int_E g_n \,\mathrm{d}\mu$$

exists for all $E \in \mathcal{B}_0$. It holds that λ is additive on $\mathcal{B}_0$. Let $\mathcal{B}_1$ be the smallest σ-algebra that contains $\mathcal{B}_0$ and all μ-null sets, and define $\mu_1 := \mu \llcorner \mathcal{B}_1$. Then $(S, \mathcal{B}_1, \mu_1)$ is a finite measure space. Since $\mu_1(S) < \infty$, we can show that λ admits a σ-additive extension to $\mathcal{B}_1$. To see this, let $(E_k)_{k\in\mathbb{N}}$ be a shrinking sequence of sets in $\mathcal{B}_1$ for which the above limit exists, and let

$$E := \bigcap_{k\in\mathbb{N}} E_k\,.$$

Then

$$\left|\int_E (g_n - g_l)\,\mathrm{d}\mu\right| \le \underbrace{\left|\int_{E_k} (g_n - g_l)\,\mathrm{d}\mu\right|}_{\substack{\to 0 \text{ as } n,l\to\infty \\ \text{for any } k}} + \underbrace{2\sup_j \int_{E_k\setminus E} |g_j|\,\mathrm{d}\mu}_{\substack{\to 0 \text{ as } k\to\infty, \\ \text{recall (2)}}}\,,$$

which shows that the above limit defines λ on all of $\mathcal{B}_1$. On noting that, in addition,

$$|\lambda(E_k \setminus E)| = \lim_{n\to\infty} \left|\int_{E_k\setminus E} g_n \,\mathrm{d}\mu\right| \le \sup_n \int_{E_k\setminus E} |g_n|\,\mathrm{d}\mu \longrightarrow 0 \qquad \text{as } k\to\infty,$$

we see that λ is even σ-additive on $\mathcal{B}_1$ and that $\lambda(E) = 0$ if $\mu(E) = 0$. Hence it follows from the Radon-Nikodým theorem that there exists an $f \in L^1(\mu_1)$ with

$$\lambda(E) = \int_E f\,\mathrm{d}\mu_1 \qquad \text{for all } E \in \mathcal{B}_1.$$

As the characteristic functions span a dense subspace of $L^\infty(\mu_1)$, this means, on recalling 6.12, that $g_n \to f$ weakly in $L^1(\mu_1)$. Now $L^1(\mu_1) \subset L^1(\mu)$ implies that $g_n \to f$ weakly also in $L^1(\mu)$. ☐

Proof $\Leftarrow$ the general case. As before, let $m = 1$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in M and let S_k for $k \in \mathbb{N}$ be the sets from (3), which we can choose such that $S_k \subset S_{k+1}$. We apply the result just shown to the sets

$$M_k := \{\mathcal{X}_{S_k} f\,;\ f \in M\}$$

(with the measure μ being restricted to S_k). Hence a diagonalization procedure yields a subsequence $n \to \infty$ and $h_k \in L^1(\mu)$ with $h_k = 0$ on $S \setminus S_k$, such that

$$\int_{S_k} f_n\, g\,\mathrm{d}\mu \longrightarrow \int_{S_k} h_k\, g\,\mathrm{d}\mu \qquad \text{as } n\to\infty \text{ for all } g \in L^\infty(\mu)\,.$$

Then $h_{k+1} = h_k$ almost everywhere on S_k, and on setting $\widetilde{S} := \bigcup_{k\in\mathbb{N}} S_k$ we have that

$$h(x) := \begin{cases} h_k(x) & \text{for } x \in S_k,\ k \in \mathbb{N}, \\ 0 & \text{for } x \in S \setminus \widetilde{S}, \end{cases}$$

defines a μ-measurable function. Now it holds for $k < l$ and for all $g \in L^\infty(\mu)$ that as $n \to \infty$

$$\begin{aligned} \left| \int_S (h_l - h_k) g \, \mathrm{d}\mu \right| &= \left| \int_{S_l} h_l \, \mathcal{X}_{S \setminus S_k} \, g \, \mathrm{d}\mu \right| \\ &\longleftarrow \left| \int_{S_l} f_n \, \mathcal{X}_{S \setminus S_k} \, g \, \mathrm{d}\mu \right| \le \delta_k \| g \|_{L^\infty} , \end{aligned}$$

where

$$\delta_k := \sup_{f \in M} \int_{S \setminus S_k} |f| \, \mathrm{d}\mu .$$

It follows that $\| h_l - h_k \|_{L^1} \le \delta_k \to 0$ as $k \to \infty$, on recalling (3). Hence $h \in L^1(\mu)$ and for $g \in L^\infty(\mu)$

$$\int_{\widetilde{S}} (h - f_n) \, g \, \mathrm{d}\mu \le \| g \|_{L^\infty} \cdot \underbrace{\Big(\int_{\widetilde{S} \setminus S_k} |h| \, \mathrm{d}\mu + \delta_k \Big)}_{\to 0 \text{ as } k \to \infty} + \underbrace{\left| \int_{S_k} (h_k - f_n) \, g \, \mathrm{d}\mu \right|}_{\substack{\to 0 \text{ as } n \to \infty \\ \text{for any } k}} .$$

This shows that $f_n \to h$ weakly in $L^1(\mu)$ as $n \to \infty$. To see this, note that if $\widetilde{\mu}$ is the measure μ restricted to $\widetilde{S}$, then $\widetilde{\mu}$ is σ-finite and

$$Jf(x) := \begin{cases} f(x) & \text{for } x \in \widetilde{S}, \\ 0 & \text{for } x \in S \setminus \widetilde{S}, \end{cases}$$

defines an embedding $J : L^1(\widetilde{\mu}) \to L^1(\mu)$. Hence for $F \in L^1(\mu)'$ we have that $\widetilde{F} := F \circ J \in L^1(\widetilde{\mu})'$, which by 6.12 can be represented by means of $g \in L^\infty(\widetilde{\mu})$. Consequently,

$$F(h - f_n) = \widetilde{F}(h - f_n) = \int_{\widetilde{S}} (h - f_n) g \, \mathrm{d}\mu \longrightarrow 0 \qquad \text{as } n \to \infty .$$

$\square$

A8.15 Theorem (Vitali-Hahn-Saks). Let $(S, \mathcal{B}, \mu)$ be a measure space and let $\lambda_n : \mathcal{B} \to \mathbb{K}$ be σ-additive for $n \in \mathbb{N}$. Suppose that

$$\forall\ n \in \mathbb{N} : \Big(\ |\lambda_n(E)| \to 0 \text{ as } \mu(E) \to 0 \ \Big),$$

and that the limit

$$\lim_{n \to \infty} \lambda_n(E) \in \mathbb{K} \text{ exists for all } E \in \mathcal{B}.$$

Then

$$\sup_{n \in \mathbb{N}} |\lambda_n(E)| \to 0 \text{ as } \mu(E) \to 0.$$

Proof. The set

$$\mathcal{M} := \{E \in \mathcal{B}\,;\ \mu(E) < \infty\}\,,$$

equipped with the distance

$$d(E_1, E_2) := \int_S |\mathcal{X}_{E_1} - \mathcal{X}_{E_2}|\,\mathrm{d}\mu\,,$$

is a complete metric space if the equivalence relation

$$E_1 = E_2 \text{ in } \mathcal{M} \quad :\Longleftrightarrow \quad \mathcal{X}_{E_1} = \mathcal{X}_{E_2}\ \mu\text{-almost everywhere}$$

is used in $\mathcal{M}$. The completeness follows from the fact that the limit of characteristic functions in $L^1(\mu)$ is again a characteristic function (this follows from A3.11). The assumptions yield that the λ_n are continuous on $\mathcal{M}$. Indeed, $d(E_k, E) \to 0$ as $k \to \infty$ implies that $\mu(E_k \setminus E) \to 0$ and $\mu(E \setminus E_k) \to 0$, and so

$$\begin{aligned}|\lambda_n(E_k) - \lambda_n(E)| &= |\lambda_n(E_k \setminus E) - \lambda_n(E \setminus E_k)| \\ &\le |\lambda_n(E_k \setminus E)| + |\lambda_n(E \setminus E_k)| \to 0\,.\end{aligned}$$

Hence for $\varepsilon > 0$ and $k \in \mathbb{N}$ the sets

$$\mathcal{A}_k^\varepsilon := \big\{\ E \in \mathcal{M}\ ;\ |\lambda_k(E) - \lambda_j(E)| \le \varepsilon \text{ for all } j \ge k\ \big\}$$

are closed subsets in $\mathcal{M}$ and the assumptions of the theorem imply that

$$\bigcup_{k\in\mathbb{N}} \mathcal{A}_k^\varepsilon = \mathcal{M}$$

for all $\varepsilon > 0$. It follows from the Baire category theorem 7.1 that at least one $\mathcal{A}_k^\varepsilon$ has a nonempty interior, i.e. there exist $k_\varepsilon \in \mathbb{N}$, $A_\varepsilon \in \mathcal{M}$, $\delta_\varepsilon > 0$ with

$$d(E, A_\varepsilon) \le \delta_\varepsilon \quad \Longrightarrow \quad |\lambda_{k_\varepsilon}(E) - \lambda_j(E)| \le \varepsilon \text{ for all } j \ge k_\varepsilon.$$

Now for $E \in \mathcal{M}$ arbitrary and $E_1 := A_\varepsilon \cup E$, $E_2 := A_\varepsilon \setminus E$

$$E = E_1 \setminus E_2\,, \quad d(E_1, A_\varepsilon) \le \mu(E)\,, \quad d(E_2, A_\varepsilon) \le \mu(E)\,.$$

If $\mu(E) \le \delta_\varepsilon$ it then follows for $j \ge k_\varepsilon$ that

$$\begin{aligned}|\lambda_j(E)| &\le |\lambda_{k_\varepsilon}(E)| + |(\lambda_{k_\varepsilon}(E_1) - \lambda_j(E_1)) - (\lambda_{k_\varepsilon}(E_2) - \lambda_j(E_2))| \\ &\le |\lambda_{k_\varepsilon}(E)| + 2\,\varepsilon\,,\end{aligned}$$

and so

$$\sup_{j\in\mathbb{N}} |\lambda_j(E)| \le 2\,\varepsilon + \underbrace{\max_{j\le k_\varepsilon} |\lambda_j(E)|}_{\substack{\to 0 \text{ as } \mu(E)\to 0\\ \text{for any } \varepsilon > 0.}}$$

This proves the desired result. $\square$