Chapter 10
The Dirichlet Principle. Variational Methods for the Solution of PDEs (Existence Techniques III)

10.1 Dirichlet's Principle

We consider the Dirichlet problem for harmonic functions once more.

We want to find a solution $u : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ a domain, of

$$\begin{aligned} \Delta u &= 0 \quad \text{in } \Omega, \\ u &= f \quad \text{on } \partial\Omega, \end{aligned} \tag{10.1.1}$$

with given f.

Dirichlet's principle is based on the following observation: Let $u \in C^2(\Omega)$ be a function with $u = f$ on $\partial\Omega$ and

$$\int_\Omega |\nabla u(x)|^2 \, \mathrm{d}x = \min \left\{ \int_\Omega |\nabla v(x)|^2 \, \mathrm{d}x : v : \Omega \to \mathbb{R} \text{ with } v = f \text{ on } \partial\Omega \right\}. \tag{10.1.2}$$

We now claim that u then solves (10.1.1). To show this, let

$$\eta \in C_0^\infty(\Omega).^1$$

According to (10.1.2), the function

$$\alpha(t) := \int_\Omega |\nabla(u + t\eta)(x)|^2 \, \mathrm{d}x$$

possesses a minimum at $t = 0$, because $u + t\eta = f$ on $\partial\Omega$, since η vanishes on $\partial\Omega$. Expanding this expression, we obtain

[1] $C_0^\infty(A) := \{\varphi \in C^\infty(A) : \text{the closure of } \{x : \varphi(x) \neq 0\} \text{ is compact and contained in } A\}$.

J. Jost, *Partial Differential Equations*, Graduate Texts in Mathematics 214,
DOI 10.1007/978-1-4614-4809-9_10,

$$\alpha(t) = \int_\Omega |\nabla u(x)|^2 \,\mathrm{d}x + 2t \int_\Omega \nabla u(x) \cdot \nabla \eta(x) \mathrm{d}x + t^2 \int_\Omega |\nabla \eta(x)|^2 \,\mathrm{d}x. \quad (10.1.3)$$

In particular, α is differentiable with respect to t, and the minimality at $t = 0$ implies

$$\dot{\alpha}(0) = 0. \quad (10.1.4)$$

By (10.1.3) this implies

$$\int_\Omega \nabla u(x) \cdot \nabla \eta(x) \mathrm{d}x = 0, \quad (10.1.5)$$

and this holds for all $\eta \in C_0^\infty(\Omega)$.

Integrating (10.1.5) by parts, we obtain

$$\int_\Omega \Delta u(x) \eta(x) \mathrm{d}x = 0 \quad \text{for all } \eta \in C_0^\infty(\Omega). \quad (10.1.6)$$

We now recall the following well-known and elementary fact:

Lemma 10.1.1. *Suppose $g \in C^0(\Omega)$ satisfies*

$$\int_\Omega g(x) \eta(x) \mathrm{d}x = 0 \quad \textit{for all } \eta \in C_0^\infty(\Omega).$$

Then $g \equiv 0$ in Ω.

Applying Lemma 10.1.1 to (10.1.6) (which is possible, since $\Delta u \in C^0(\Omega)$ by our assumption $u \in C^2(\Omega)$), we indeed obtain

$$\Delta u(x) = 0 \quad \text{in } \Omega,$$

as claimed.

This observation suggests that we try to minimize the so-called Dirichlet integral

$$D(u) := \int_\Omega |\nabla u(x)|^2 \,\mathrm{d}x \quad (10.1.7)$$

in the class of all functions $u : \Omega \to \mathbb{R}$ with $u = f$ on $\partial\Omega$. This is Dirichlet's principle.

It is by no means evident, however, that the Dirichlet integral assumes its infimum within the considered class of functions. This constitutes the essential difficulty of Dirichlet's principle. In any case, so far, we have not specified which class of functions $u : \Omega \to \mathbb{R}$ (with the given boundary values) we allow for competition; the possibilities include functions of class C^∞, which would be natural, since we have shown already in Chap. 2 that any solution of (10.1.1) automatically is of regularity class C^∞; functions of class C^2, which would be natural, since then the differential equation $\Delta u(x) = 0$ would have a meaning; and functions of class C^1 because then at least (assuming Ω bounded and f sufficiently regular, e.g., $f \in C^1$) the Dirichlet integral $D(u)$ would be finite. Posing the question somewhat differently, should we try to minimize $D(U)$ in a space of functions that is as large as possible, in order to increase the chance that a minimizing sequence possesses a limit in that space that then would be a natural candidate for a minimizer, or should we rather select a smaller space in order to facilitate the verification that a tentative solution is a minimizer?

In order to analyze this question, we consider a minimizing sequence $(u_n)_{n\in\mathbb{N}}$ for D, i.e.,

$$\lim_{n\to\infty} D(u_n) = \inf\{D(v) : v : \Omega \to \mathbb{R}, v = f \text{ on } \partial\Omega\} =: \kappa, \tag{10.1.8}$$

where, of course, we assume $u_n = f$ on $\partial\Omega$ for all u_n. To find properties of such a minimizing sequence, we shall employ the following simple lemma:

Lemma 10.1.2. *Dirichlet's integral is convex, i.e.,*

$$D(tu + (1-t)v) \le tD(u) + (1-t)D(v) \tag{10.1.9}$$

for all u, v, *and* $t \in [0,1]$.

Proof.

$$\begin{aligned} D(tu + (1-t)v) &= \int_\Omega |t\nabla u + (1-t)\nabla v|^2 \\ &\le \int_\Omega \left\{t\,|\nabla u|^2 + (1-t)\,|\nabla v|^2\right\} \\ &\qquad \text{because of the convexity of } w \mapsto |w|^2 \\ &= tD(u) + (1-t)Dv. \end{aligned}$$

□

Now let $(u_n)_{n\in\mathbb{N}}$ be a minimizing sequence. Then

$$\begin{aligned} D(u_n - u_m) &= \int_\Omega |\nabla(u_n - u_m)|^2 \\ &= 2\int_\Omega |\nabla u_n|^2 + 2\int_\Omega |\nabla u_m|^2 - 4\int_\Omega \left|\nabla\left(\frac{u_n + u_m}{2}\right)\right|^2 \\ &= 2D(u_n) + 2D(u_m) - 4D\left(\frac{u_n + u_m}{2}\right). \end{aligned} \tag{10.1.10}$$

We now have

$$
\begin{aligned}
\kappa &\le D\left(\frac{u_n+u_m}{2}\right) \text{ by definition of } \kappa \text{ (10.1.8)} \\
&\le \frac{1}{2}D(u_n)+\frac{1}{2}D(u_m) \text{ by Lemma 10.1.2} \\
&\to \kappa \quad \text{for } n,m\to\infty, \qquad (10.1.11)
\end{aligned}
$$

since (u_n) is a minimizing sequence. This implies that the right-hand side of (10.1.10) converges to 0 for $n, m \to \infty$, and so then does the left-hand side. This means that $(\nabla u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to the topology of the space $L^2(\Omega)$. (Since ∇u_n has d components, i.e., is vector-valued, this says that $\frac{\partial u_n}{\partial x^i}$ is a Cauchy sequence in $L^2(\Omega)$ for $i = 1, \ldots, d$.) Since $L^2(\Omega)$ is a Hilbert space, hence complete; ∇u_n thus converges to some $w \in L^2(\Omega)$. The question now is whether w can be represented as the gradient ∇u of some function $u : \Omega \to \mathbb{R}$. At the moment, however, we know only that $w \in L^2(\Omega)$, and so it is not clear what regularity properties u should possess. In any case, this consideration suggests that we seek a minimum of D in the space of those functions whose gradient is in $L^2(\Omega)$. In a subsequent step we would then have to analyze the regularity properties of such a minimizer u. For that step, the starting point would be relation (10.1.5), i.e.,

$$
\int_\Omega \nabla u(x)\cdot\nabla\eta(x)\mathrm{d}x = 0 \quad \text{for all } \eta\in C_0^\infty(\Omega), \tag{10.1.12}
$$

which continues to hold in the context presently considered. By Corollary 2.2.1 this already implies $u \in C^\infty(\Omega)$. In the next chapter, however, we shall investigate this problem in greater generality.

Dividing the problem into two steps as just sketched, namely, first proving the existence of a minimizer and afterwards establishing its regularity, proves to be a fruitful approach indeed, as we shall find in the sequel. For that purpose, we first need to investigate the space of functions just considered in more detail. This is the task of the next section.

10.2 The Sobolev Space $W^{1,2}$

Definition 10.2.1. Let $\Omega \subset \mathbb{R}^d$ be open and $u \in L^1_{\mathrm{loc}}(\Omega)$. A function $v \in L^1_{\mathrm{loc}}(\Omega)$ is called weak derivative of u in the direction x^i $(x = (x^1, \ldots, x^d) \in \mathbb{R}^d)$ if

$$
\int_\Omega \phi v = -\int_\Omega u\frac{\partial\phi}{\partial x^i}\mathrm{d}x \tag{10.2.1}
$$

for all $\phi \in C^1_0(\Omega)$.[2] We write $v = D_i u$.

[2] $C^k_0(\Omega) := \{f \in C^k(\Omega) : \text{the closure of } \{x : f(x) \neq 0\} \text{ is a compact subset of } \Omega\}$ $(k = 1, 2, \ldots)$.

A function u is called weakly differentiable if it possesses a weak derivative in the direction x^i for all $i \in \{1, \ldots, d\}$.

It is obvious that each $u \in C^1(\Omega)$ is weakly differentiable, and the weak derivatives are simply given by the ordinary derivatives. Equation (10.2.1) is then the formula for integrating by parts. Thus, the idea behind the definition of weak derivatives is to use the integration by parts formula as an abstract axiom.

Lemma 10.2.1. *Let $u \in L^1_{\text{loc}}(\Omega)$, and suppose $v = D_i u$ exists. If* $\operatorname{dist}(x, \partial\Omega) > h$, *we have*

$$D_i(u_h(x)) = (D_i u)_h(x).$$

Proof. By differentiating under the integral, we obtain

$$\begin{aligned}
D_i(u_h(x)) &= \frac{1}{h^d}\int \frac{\partial}{\partial x^i}\varrho\left(\frac{x-y}{h}\right)u(y)\mathrm{d}y \\
&= \frac{-1}{h^d}\int \frac{\partial}{\partial y^i}\varrho\left(\frac{x-y}{h}\right)u(y)\mathrm{d}y \\
&= \frac{1}{h^d}\int \varrho\left(\frac{x-y}{h}\right)D_i u(y)\mathrm{d}y \text{ by (10.2.1)} \\
&= (D_i u)_h(x).
\end{aligned}$$

□

Lemmas A.3 and 10.2.1 and formula (10.2.1) imply the following theorem:

Theorem 10.2.1. *Let $u, v \in L^2(\Omega)$. Then*

$$v = D_i u$$

precisely if there exists a sequence $(u_n) \subset C^\infty(\Omega)$ with

$$u_n \to u, \quad \frac{\partial}{\partial x^i}u_n \to v \quad \textit{in } L^2(\Omega') \quad \textit{for any } \Omega' \subset\subset \Omega.$$

Definition 10.2.2. The Sobolev space $W^{1,2}(\Omega)$ is defined as the space of those $u \in L^2(\Omega)$ that possess a weak derivative of class $L^2(\Omega)$ for each direction x^i ($i = 1, \ldots, d$).

In $W^{1,2}(\Omega)$ we define a scalar product

$$(u, v)_{W^{1,2}(\Omega)} := \int_\Omega uv + \sum_{i=1}^d \int_\Omega D_i u \cdot D_i v$$

and a norm

$$\|u\|_{W^{1,2}(\Omega)} := (u, u)^{\frac{1}{2}}_{W^{1,2}(\Omega)}.$$

We also define $H^{1,2}(\Omega)$ as the closure of $C^\infty(\Omega) \cap W^{1,2}(\Omega)$ with respect to the $W^{1,2}$-norm, and $H_0^{1,2}(\Omega)$ as the closure of $C_0^\infty(\Omega)$ with respect to this norm.

Corollary 10.2.1. *$W^{1,2}(\Omega)$ is complete with respect to $\|\cdot\|_{W^{1,2}}$, and is hence a Hilbert space. $W^{1,2}(\Omega) = H^{1,2}(\Omega)$.*

Proof. Let $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $W^{1,2}(\Omega)$. Then $(u_n)_{n\in\mathbb{N}}$, $(D_i u_n)_{n\in\mathbb{N}}$ $(i = 1,\dots,d)$ are Cauchy sequences in $L^2(\Omega)$. Since $L^2(\Omega)$ is complete, there exist $u, v^i \in L^2(\Omega)$ with

$$u_n \to u, \quad D_i u_n \to v^i \quad \text{in } L^2(\Omega) \quad (i = 1,\dots,d).$$

For $\phi \in C_0^1(\Omega)$, we have

$$\int D_i u_n \cdot \phi = -\int u_n D_i \phi,$$

and the left-hand side converges to $\int v^i \cdot \phi$, the right-hand side to $-\int u \cdot D_i \phi$. Therefore, $D_i u = v^i$, and thus $u \in W^{1,2}(\Omega)$. This shows completeness.

In order to prove the equality $H^{1,2}(\Omega) = W^{1,2}(\Omega)$, we need to verify that the space $C^\infty(\Omega) \cap W^{1,2}(\Omega)$ is dense in $W^{1,2}(\Omega)$. For $n \in \mathbb{N}$, we put

$$\Omega_n := \left\{ x \in \Omega : \|x\| < n, \operatorname{dist}(x, \partial\Omega) > \frac{1}{n} \right\},$$

with $\Omega_0 := \Omega_{-1} := \emptyset$. Thus,

$$\Omega_n \subset\subset \Omega_{n+1} \quad \text{and} \quad \bigcup_{n\in\mathbb{N}} \Omega_n = \Omega.$$

We let $\{\varphi_j\}_{j\in\mathbb{N}}$ be a partition of unity subordinate to the cover

$$\left\{ \Omega_{n+1} \setminus \bar{\Omega}_{n-1} \right\}$$

of Ω. Let $u \in W^{1,2}(\Omega)$. By Theorem 10.2.1, for every $\varepsilon > 0$, we may find a positive number h_n for any $n \in \mathbb{N}$ such that

$$h_n \le \operatorname{dist}(\Omega_n, \partial\Omega_{n+1}),$$

$$\|(\varphi_n u)_{h_n} - \varphi_n u\|_{W^{1,2}(\Omega)} < \frac{\varepsilon}{2^n}.$$

Since the φ_n constitute a partition of unity, on any $\Omega' \subset\subset \Omega$, at most finitely many of the smooth functions $(\varphi_n u)_{h_n}$ are non-zero. Consequently,

$$\tilde{u} := \sum_n (\varphi_n u)_{h_n} \in C^\infty(\Omega).$$

We have

$$\|u - \tilde{u}\|_{W^{1,2}(\Omega)} \leq \sum_n \|(\varphi_n u)_{h_n} - \varphi_n u\| < \varepsilon,$$

and we see that every $u \in W^{1,2}(\Omega)$ can be approximated by C^∞-functions. □

Corollary 10.2.1 answers one of the questions raised in Sect. 10.1, namely, whether the function w considered there can be represented as the gradient of an L^2-function.

Examples:

(i) We consider $\Omega = (-1, 1) \subset \mathbb{R}$ and $u(x) := |x|$.
In that case, $u \in W^{1,2}((-1, 1))$, and

$$Du(x) = \begin{cases} 1 & \text{for } 0 < x < 1, \\ -1 & \text{for } -1 < x < 0, \end{cases}$$

because for every $\phi \in C_0^1((-1, 1))$,

$$\int_{-1}^{0} -\phi(x)\mathrm{d}x + \int_0^1 \phi(x)\mathrm{d}x = -\int_{-1}^{1} \phi'(x) \cdot |x| \, \mathrm{d}x.$$

(ii) Again, $\Omega = (-1, 1) \subset \mathbb{R}$, and

$$u(x) := \begin{cases} 1 & \text{for } 0 \leq x < 1, \\ 0 & \text{for } -1 < x < 0, \end{cases}$$

is not weakly differentiable, for if it were, necessarily $Du(x) = 0$ for $x \neq 0$; hence as an L^1_{loc} function $Du \equiv 0$, but we do not have, for every $\phi \in C_0^1((-1, 1))$,

$$0 = \int_{-1}^{1} \phi(x) \cdot 0 \, \mathrm{d}x = -\int_{-1}^{1} \phi'(x) u(x)\mathrm{d}x = -\int_0^1 \phi'(x)\mathrm{d}x = \phi(0).$$

Remark. Any $u \in L^1_{\text{loc}}(\Omega)$ defines a distribution (cf. Sect. 2.1) l_u by

$$l_u[\varphi] := \int_\Omega u(x)\varphi(x)\mathrm{d}x \quad \text{for } \varphi \in C_0^\infty(\Omega).$$

Every distribution l possesses distributional derivatives $D_i l$, $i = 1, \ldots, d$, defined by

$$D_i l[\varphi] := -l\left[\frac{\partial \varphi}{\partial x^i}\right].$$

If $v = D_i u \in L^1_{\mathrm{loc}}(\Omega)$ is the weak derivative of u, then

$$D_i l_u = l_v,$$

because

$$l_v[\varphi] = \int_\Omega D_i u(x)\varphi(x)\mathrm{d}x = -\int_\Omega u(x)\frac{\partial\varphi}{\partial x^i}(x)\mathrm{d}x = D_i l_u[\varphi]$$

for all $\varphi \in C_0^\infty(\Omega)$.

Whereas the distributional derivative $D_i l_u$ always exists, the weak derivative need not exist. Thus, in general, the distributional derivative is not of the form l_v for some $v \in L^1_{\mathrm{loc}}(\Omega)$, i.e., not represented by a locally integrable function. This is what happens in Example (ii). Here, $Dl_u = \delta_0$, the delta distribution at 0, because

$$Dl_u[\varphi] = -l_u[\varphi'] = -\int_{-1}^1 u(x)\varphi'(x)\mathrm{d}x = -\int_0^1 \varphi'(x)\mathrm{d}x = \varphi(0).$$

The delta distribution cannot be represented by some locally integrable function v, because, as one easily verifies, there is no function $v \in L^1_{\mathrm{loc}}((-1,1))$ with

$$\int_{-1}^1 v(x)\varphi(x)\mathrm{d}x = \varphi(0) \quad \text{for all } \varphi \in C_0^\infty(\Omega).$$

This explains why u from Example (ii) is not weakly differentiable.

(iii) This time, $\Omega = B(0,1) \subset \mathbb{R}^d$, and $u(x) := \frac{x}{|x|}$, i.e., we consider a mapping from $B(0,1)$ to $\mathbb{R}^d$. u is smooth except at $x = 0$ where it is discontinuous. For $d = 1$, of course, $u(x) = 1$ for $x > 0$ and $u(x) = -1$ for $x < 0$. Hence, in that case, as in Example (ii), u is not weakly differentiable. We now consider the case $d > 1$. We let e_i be the ith unit vector, i.e., $x = \sum_i x^i e_i$. For $x \neq 0$, we have

$$\frac{\partial}{\partial x^i}\frac{x}{|x|} = \frac{e_i}{|x|} - \frac{x^i x}{|x|^3}. \tag{10.2.2}$$

We claim that, for $d > 1$, this extends as the weak derivative $D_i u$ across the singularity at $x = 0$. To check this, we need to verify that

$$\int_{B(0,1)} \left(\frac{e_i}{|x|} - \frac{x^i x}{|x|^3}\right)\phi = -\int_{B(0,1)} \frac{x}{|x|}\frac{\partial\phi}{\partial x^i} \tag{10.2.3}$$

for all test functions $\phi \in H_0^{1,2}(B(0,1),\mathbb{R}^d)$. (Note that (10.2.3) has to be understood in the vector sense; for instance, $\frac{x}{|x|}\frac{\partial\phi}{\partial x^i}$ stands for $\sum_{\alpha=1}^d \frac{x^\alpha}{|x|}\frac{\partial\phi^\alpha}{\partial x^i}$.) First of all, we observe that (10.2.3) holds for all such ϕ that vanish in the vicinity of 0, because u is smooth away from 0. In order to handle the discontinuity at 0, we introduce the Lipschitz cut-off functions

$$\eta_m := \begin{cases} 1 & \text{if } |x| \le 2^{-m} \\ \frac{1}{2^{m-1}}\left(\frac{1}{|x|} - 2^{m-1}\right) & \text{if } 2^{-m} \le |x| \le 2^{-(m-1)} \\ 0 & \text{if } 2^{-(m-1)} \le |x| \end{cases}$$

and write $\phi = (1-\eta_m)\phi + \eta_m\phi$. The first term then vanishes near 0, and therefore, as just explained, this term is fine in (10.2.3). In order to verify (10.2.3) for a general ϕ, we therefore only have to verify that the contributions coming from $\eta_m\phi$ go to 0 for $m \to \infty$. When we insert $\eta_m\varphi$ in (10.2.3), the only contribution that does not obviously go to 0 for $m \to \infty$ (and $d > 1$) is

$$\int \frac{x}{|x|}\frac{\partial\eta_m}{\partial x^i}\phi. \tag{10.2.4}$$

However, since

$$\frac{\partial\eta_m}{\partial x^i} = \begin{cases} 2^{1-m}\frac{x^i}{|x|^3} & \text{for } 2^{-m} \le |x| \le 2^{-(m-1)} \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\left|\frac{\partial\eta_m}{\partial x^i}\right| \le \frac{2}{|x|} \text{ since } 2^{-m} \le |x|.$$

Therefore, (10.2.4) does go to 0 for $m \to \infty$. We conclude that u possesses weak derivatives for $d > 1$, indeed. We note, however, from (10.2.2) that

$$\left|D\frac{x}{|x|}\right|^2 = \frac{d-1}{|x|^2}, \tag{10.2.5}$$

and so,

$$\int_{B(0,1)} |Du|^2 < \infty \text{ for } d \ge 3,$$

i.e., $u \in W^{1,2}(B(0,1))$ for $d \ge 3$, but not for $d = 2$.

(iv) We leave it to the reader to check that

$$u(x) := \log\log\frac{1}{|x|} \tag{10.2.6}$$

is in $W^{1,2}$ on the ball $B(0,\frac{1}{2}) \subset \mathbb{R}^2$. Again, this u is discontinuous (and unbounded) at $x = 0$. Thus, also for $d = 2$, functions in the Sobolev space $W^{1,2}$ need not be continuous.

We now prove a replacement lemma exhibiting a characteristic property of Sobolev functions:

Lemma 10.2.2. *Let $\Omega_0 \subset\subset \Omega$, $g \in W^{1,2}(\Omega)$, $u \in W^{1,2}(\Omega_0)$, $u - g \in H_0^{1,2}(\Omega_0)$. Then*

$$v(x) := \begin{cases} u(x) & \textit{for } x \in \Omega_0, \\ g(x) & \textit{for } x \in \Omega \setminus \Omega_0, \end{cases}$$

is contained in $W^{1,2}(\Omega)$, and

$$D_i v(x) = \begin{cases} D_i u(x) & \textit{for } x \in \Omega_0, \\ D_i g(x) & \textit{for } x \in \Omega \setminus \Omega_0. \end{cases}$$

Proof. By Corollary 10.2.1, there exist $g_n \in C^\infty(\Omega)$, $u_n \in C^\infty(\Omega_0)$ with

$$\begin{aligned} g_n &\to g \quad \text{in } W^{1,2}(\Omega), \\ u_n &\to u \quad \text{in } W^{1,2}(\Omega_0), \\ u_n - g_n &= 0 \quad \text{on } \partial\Omega_0. \end{aligned} \tag{10.2.7}$$

We put

$$w_n^i(x) := \begin{cases} D_i u_n(x) & \text{for } x \in \Omega_0, \\ D_i g_n(x) & \text{for } x \in \Omega \setminus \Omega_0, \end{cases}$$

$$v_n(x) := \begin{cases} u_n(x) & \text{for } x \in \Omega_0, \\ g_n(x) & \text{for } x \in \Omega \setminus \Omega_0, \end{cases}$$

$$w^i(x) := \begin{cases} D_i u(x) & \text{for } x \in \Omega_0, \\ D_i g(x) & \text{for } x \in \Omega \setminus \Omega_0. \end{cases}$$

We then have for $\varphi \in C_0^1(\Omega)$,

$$\begin{aligned} \int_\Omega \varphi w_n^i = \int_{\Omega_0} \varphi w_n^i + \int_{\Omega\setminus\Omega_0} \varphi w_n^i &= \int_{\Omega_0} \varphi D_i u_n + \int_{\Omega\setminus\Omega_0} \varphi D_i g_n \\ &= -\int_{\Omega_0} u_n D_i \varphi - \int_{\Omega\setminus\Omega_0} g_n D_i \varphi \end{aligned}$$

since the two boundary terms resulting from integrating the two integrals by parts have opposite signs and thus cancel because of $g_n = u_n$ on $\partial\Omega_0$

$$= -\int_\Omega v_n D_i \varphi$$

by (10.2.7). Now for $n \to \infty$,

$$\int_\Omega \varphi w_n^i \to \int_{\Omega_0} \varphi D_i u + \int_{\Omega \setminus \Omega_0} \varphi D_i g,$$

$$\int_\Omega v_n D_i \varphi \to \int_\Omega v D_i \varphi,$$

and the claim follows. □

The next lemma is a chain rule for Sobolev functions:

Lemma 10.2.3. *For $u \in W^{1,2}(\Omega)$, $f \in C^1(\mathbb{R})$, suppose*

$$\sup_{y \in \mathbb{R}} \left| f'(y) \right| < \infty.$$

Then $f \circ u \in W^{1,2}(\Omega)$, and the weak derivative satisfies $D(f \circ u) = f'(u) Du$.

Proof. Let $u_n \in C^\infty(\Omega)$, $u_n \to u$ in $W^{1,2}(\Omega)$ for $n \to \infty$. Then

$$\int_\Omega |f(u_n) - f(u)|^2 \, \mathrm{d}x \le \sup \left| f' \right|^2 \int_\Omega |u_n - u|^2 \, \mathrm{d}x \to 0$$

and

$$\int_\Omega \left| f'(u_n) Du_n - f'(u) Du \right|^2 \mathrm{d}x \le 2 \sup \left| f' \right|^2 \int_\Omega |Du_n - Du|^2 \, \mathrm{d}x + 2 \int_\Omega \left| f'(u_n) - f'(u) \right|^2 |Du|^2 \, \mathrm{d}x.$$

By a well-known result about L^2-functions, after selection of a subsequence, u_n converges to u pointwise almost everywhere in Ω.[3] Since f' is continuous, $f'(u_n)$ then also converges pointwise almost everywhere to $f'(u)$, and since f' is also bounded, the last integral converges to 0 for $n \to \infty$ by Lebesgue's theorem on dominated convergence.

Thus

$$f(u_n) \to f(u) \quad \text{in } L^2(\Omega)$$

and

$$D(f(u_n)) = f'(u_n) Du_n \to f'(u) Du \quad \text{in } L^2(\Omega),$$

and hence $f \circ u \in W^{1,2}(\Omega)$ and $D(f \circ u) = f'(u) Du$. □

---

[3] See p. 240 of [19].

Corollary 10.2.2. *If $u \in W^{1,2}(\Omega)$, then also $|u| \in W^{1,2}(\Omega)$, and $D|u| = \operatorname{sign} u \cdot Du$.*

Proof. We consider $f_\varepsilon(u) := (u^2 + \varepsilon^2)^{\frac{1}{2}} - \varepsilon$, apply Lemma 10.2.3, and let $\varepsilon \to 0$, using once more Lebesgue's theorem on dominated convergence to justify the limit as before. □

We next prove the Poincaré inequality (see also Corollary 11.5.1 below).

Theorem 10.2.2. *For $u \in H_0^{1,2}(\Omega)$, we have*

$$\|u\|_{L^2(\Omega)} \leq \left(\frac{|\Omega|}{\omega_d}\right)^{\frac{1}{d}} \|Du\|_{L^2(\Omega)} \tag{10.2.8}$$

where $|\Omega|$ denotes the (Lebesgue) measure of Ω and ω_d is the measure of the unit ball in $\mathbb{R}^d$. In particular, for any $u \in H_0^{1,2}(\Omega)$, its $W^{1,2}$-norm is controlled by the L^2-norm of Du:

$$\|u\|_{W^{1,2}(\Omega)} \leq \left(1 + \left(\frac{|\Omega|}{\omega_d}\right)^{\frac{1}{d}}\right) \|Du\|_{L^2(\Omega)}$$

Proof. Suppose first $u \in C_0^1(\Omega)$; we put $u(x) = 0$ for $x \in \mathbb{R}^d \setminus \Omega$. For $\omega \in \mathbb{R}^d$ with $|\omega| = 1$, by the fundamental theorem of calculus, we obtain by integrating along the ray $\{r\omega : 0 \leq r < \infty\}$ that

$$u(x) = -\int_0^\infty \frac{\partial}{\partial r} u(x + r\omega) \mathrm{d}r.$$

Integrating with respect to ω then yields, as in the proof of Theorem 2.2.1,

$$\begin{aligned} u(x) &= -\frac{1}{d\omega_d} \int_0^\infty \int_{|\omega|=1} \frac{\partial}{\partial r} u(x + r\omega)\, \mathrm{d}\omega \mathrm{d}r \\ &= -\frac{1}{d\omega_d} \int_0^\infty \int_{\partial B(x,r)} \frac{1}{r^{d-1}} \frac{\partial u}{\partial \nu}(z) \mathrm{d}\sigma(z) \mathrm{d}r \\ &= -\frac{1}{d\omega_d} \int_\Omega \frac{1}{|x-y|^{d-1}} \sum_{i=1}^d \frac{\partial}{\partial y^i} u(y) \frac{x^i - y^i}{|x-y|} \mathrm{d}y, \end{aligned} \tag{10.2.9}$$

and thus with the Schwarz inequality,

$$|u(x)| \leq \frac{1}{d\omega_d} \int_\Omega \frac{1}{|x-y|^{d-1}} \cdot |Du(y)|\, \mathrm{d}y. \tag{10.2.10}$$

We now need a lemma:

Lemma 10.2.4. *For* $f \in L^1(\Omega)$, $0 < \mu \leq 1$, *let*

$$(V_\mu f)(x) := \int_\Omega |x-y|^{d(\mu-1)} f(y) \mathrm{d}y.$$

Then

$$\left\| V_\mu f \right\|_{L^2(\Omega)} \leq \frac{1}{\mu} \omega_d^{1-\mu} |\Omega|^\mu \left\| f \right\|_{L^2(\Omega)}.$$

Proof. $B(x,R) := \{y \in \mathbb{R}^d : |x-y| \leq R\}$. Let R be chosen such that $|\Omega| = |B(x,R)| = \omega_d R^d$. Since in that case

$$|\Omega \setminus (\Omega \cap B(x,R))| = |B(x,R) \setminus (\Omega \cap B(x,R))|$$

and

$$\begin{aligned} |x-y|^{d(\mu-1)} &\leq R^{d(\mu-1)} \quad \text{for } |x-y| \geq R, \\ |x-y|^{d(\mu-1)} &\geq R^{d(\mu-1)} \quad \text{for } |x-y| \leq R; \end{aligned}$$

it follows that

$$\int_\Omega |x-y|^{d(\mu-1)} \, \mathrm{d}y \leq \int_{B(x,R)} |x-y|^{d(\mu-1)} \, \mathrm{d}y = \frac{1}{\mu} \omega_d R^{d\mu} = \frac{1}{\mu} \omega_d^{1-\mu} |\Omega|^\mu . \tag{10.2.11}$$

We now write

$$|x-y|^{d(\mu-1)} |f(y)| = \left(|x-y|^{\frac{d}{2}(\mu-1)} \right) \left(|x-y|^{\frac{d}{2}(\mu-1)} |f(y)| \right)$$

and obtain, applying the Cauchy Schwarz inequality,

$$\begin{aligned} \left| (V_\mu f)(x) \right| &\leq \int_\Omega |x-y|^{d(\mu-1)} |f(y)| \, \mathrm{d}y \\ &\leq \left(\int_\Omega |x-y|^{d(\mu-1)} \, \mathrm{d}y \right)^{\frac{1}{2}} \left(\int_\Omega |x-y|^{d(\mu-1)} |f(y)|^2 \, \mathrm{d}y \right)^{\frac{1}{2}}, \end{aligned}$$

and hence

$$\int_\Omega |V_\mu f(x)|^2 \, \mathrm{d}x \leq \frac{1}{\mu} \omega_d^{1-\mu} |\Omega|^\mu \int_\Omega \int_\Omega |x-y|^{d(\mu-1)} |f(y)|^2 \, \mathrm{d}y \, \mathrm{d}x$$

by estimating the first integral of the preceding inequality with (10.2.11)

$$\leq \left(\frac{1}{\mu}\omega_d^{1-\mu}|\Omega|^{\mu}\right)^2 \int_{\Omega} |f(y)|^2 \mathrm{d}y$$

by interchanging the integrations with respect to x and y and applying (10.2.11) once more, whence the claim. □

We may now complete the proof of Theorem 10.2.2: Applying Lemma 10.2.4 with $\mu = \frac{1}{d}$ and $f = |Du|$ to the right-hand side of (10.2.10), we obtain (10.2.8) for $u \in C_0^1(\Omega)$. Since by definition of $H_0^{1,2}(\Omega)$, it contains $C_0^1(\Omega)$ as a dense subspace, we may approximate u in the $H^{1,2}$-norm by some sequence $(u_n)_{n\in\mathbb{N}} \subset C_0^1(\Omega)$. Thus, u_n converges to u in L^2 and Du_n to u. Thus, the inequality (10.2.8) that has been proved for u_n extends to u. □

Remark. The assumption that u is contained in $H_0^{1,2}(\Omega)$, and not only in $H^{1,2}(\Omega)$, is necessary for Theorem 10.2.2, since otherwise the nonzero constants would constitute counterexamples. However, the assumption $u \in H_0^{1,2}(\Omega)$ may be replaced by other assumptions that exclude nonzero constants, for example,by $\int_{\Omega} u(x)\mathrm{d}x = 0$.

For our treatment of eigenvalues of the Laplace operator in Sect. 11.5, the fundamental tool will be the compactness theorem of Rellich:

Theorem 10.2.3. *Let $\Omega \in \mathbb{R}^d$ be open and bounded. Then $H_0^{1,2}(\Omega)$ is compactly embedded in $L^2(\Omega)$; i.e., any sequence $(u_n)_{n\in\mathbb{N}} \subset H_0^{1,2}(\Omega)$ with*

$$\|u_n\|_{W^{1,2}(\Omega)} \leq c_0 \tag{10.2.12}$$

contains a subsequence that converges in $L^2(\Omega)$.

Proof. The strategy is to find functions $w_{n,\varepsilon} \in C^1(\Omega)$, for every $\varepsilon > 0$, with

$$\|u_n - w_{n,\varepsilon}\|_{W^{1,2}(\Omega)} < \frac{\varepsilon}{2} \tag{10.2.13}$$

and

$$\|w_{n,\varepsilon}\|_{W^{1,2}(\Omega)} \leq c_1 \tag{10.2.14}$$

(the constant c_1 will depend on ε, but not on n). By the Arzela–Ascoli theorem, $(w_{n,\varepsilon})_{n\in\mathbb{N}}$ then contains a subsequence that converges uniformly, hence also in L^2. Since this holds for every $\varepsilon > 0$, one may appeal to a general theorem about compact subsets of metric spaces to conclude that the closure of $(u_n)_{n\in\mathbb{N}}$ is compact in $L^2(\Omega)$ and thus contains a convergent subsequence. That theorem[4] states that a subset of

[4]See, for example, [18], Theorem 7.38.

a metric space is compact precisely if it is complete and totally bounded, i.e., if for any $\varepsilon > 0$, it is contained in the union of a finite number of balls of radius ε. Applying this result to the (closure of the) sequence $(w_{n,\varepsilon})_{n\in\mathbb{N}}$, we infer that there exist finitely many z_ν, $\nu = 1, \dots, N$, in $L^2(\Omega)$ such that for every $n \in \mathbb{N}$,

$$\|w_{n,\varepsilon} - z_\nu\|_{L^2(\Omega)} < \frac{\varepsilon}{2} \quad \text{for some } \nu \in \{1, \dots, N\}. \tag{10.2.15}$$

Hence, from (10.2.13) and (10.2.15), for every $n \in \mathbb{N}$,

$$\|u_n - z_\nu\|_{L^2(\Omega)} < \varepsilon \quad \text{for some } \nu.$$

Since this holds for every $\varepsilon > 0$, the sequence $(u_n)_{n\in\mathbb{N}}$ is totally bounded, and so its closure is compact in $L^2(\Omega)$, and we get the desired convergent subsequence in $L^2(\Omega)$.

It remains to construct the $w_{n,\varepsilon}$. First of all, by definition of $H_0^{1,2}(\Omega)$, there exists $w_n \in C_0^1(\Omega)$ with

$$\|u_n - w_n\|_{W^{1,2}(\Omega)} < \frac{\varepsilon}{4}. \tag{10.2.16}$$

By (10.2.12), then also

$$\|w_n\|_{W^{1,2}(\Omega)} \le c_0' \quad \text{for some constant } c_0'. \tag{10.2.17}$$

We then define $w_{n,\varepsilon}$ as the mollification of w_n with a parameter $h = h(\varepsilon)$ to be determined subsequently:

$$w_{n,\varepsilon}(x) = \frac{1}{h^d} \int_\Omega \varrho\left(\frac{x-y}{h}\right) w_n(y)\mathrm{d}y.$$

The crucial step now is to control the L^2-norm of the difference $w_n - w_{n,\varepsilon}$ with the help of the $W^{1,2}$-bound on the original u_n. This goes as follows:

$$\begin{aligned}
\int_\Omega |w_n(x) - w_{n,\varepsilon}(x)|^2 \mathrm{d}x &= \int_\Omega \left(\int_{|y|\le 1} \varrho(y)(w_n(x) - w_n(x - hy))\mathrm{d}y \right)^2 \mathrm{d}x \\
&\le \int_\Omega \left(\int_{|y|\le 1} \varrho(y) \int_0^{h|y|} \left| \frac{\partial}{\partial r} w_n(x - r\omega) \right| \mathrm{d}r\, \mathrm{d}y \right)^2 \mathrm{d}x \quad \text{with } \omega = \frac{y}{|y|} \\
&= \int_\Omega \left(\int_{|y|\le 1} \varrho(y)^{\frac12} \varrho(y)^{\frac12} \int_0^{h|y|} \left| \frac{\partial}{\partial r} w_n(x - r\omega) \right| \mathrm{d}r\, \mathrm{d}y \right)^2 \mathrm{d}x \\
&\le \left(\int_{|y|\le 1} \varrho(y)\mathrm{d}y \right) \left(\int_{|y|\le 1} \varrho(y) h^2 |y|^2 \int |Dw_n(x)|^2 \,\mathrm{d}x\, \mathrm{d}y \right)
\end{aligned}$$

by Hölder's inequality [(A.4) of the appendix] and Fubini's theorem. Since $\int_{|y|\le 1} \varrho(y)\,\mathrm{d}y = 1$, we obtain the estimate

$$\|w_n - w_{n,\varepsilon}\|_{L^2(\Omega)} \le h\,\|Dw_n\|_{L^2(\Omega)}\,.$$

Because of (10.2.17), we may then choose h such that

$$\|w_n - w_{n,\varepsilon}\|_{L^2(\Omega)} < \frac{\varepsilon}{4}. \tag{10.2.18}$$

Then (10.2.16) and (10.2.18) yield the desired estimate (10.2.13). □

10.3 Weak Solutions of the Poisson Equation

As before, let Ω be an open and bounded subset of $\mathbb{R}^d$, $g \in H^{1,2}(\Omega)$. With the concepts introduced in the previous section, we now consider the following version of the Dirichlet principle. We seek a solution of

$$\begin{aligned} \Delta u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{for } \partial\Omega \quad \Big(\text{meaning } u - g \in H_0^{1,2}(\Omega)\Big), \end{aligned}$$

by minimizing the Dirichlet integral

$$\int_\Omega |Dv|^2 \quad (\text{here, } Dv = (D_1 v, \ldots, D_d v))$$

among all $v \in H^{1,2}(\Omega)$ with $v - g \in H_0^{1,2}(\Omega)$. We want to convince ourselves that this approach indeed works. Let

$$\kappa := \inf\left\{\int_\Omega |Dv|^2 : v \in H^{1,2}(\Omega), v - g \in H_0^{1,2}(\Omega)\right\},$$

and let $(u_n)_{n\in\mathbb{N}}$ be a minimizing sequence, meaning that $u_n - g \in H_0^{1,2}(\Omega)$, and

$$\int_\Omega |Du_n|^2 \to \kappa.$$

We have already argued in Sect. 10.1 that for a minimizing sequence $(u_n)_{n\in\mathbb{N}}$, the sequence of (weak) derivatives (Du_n) is a Cauchy sequence in $L^2(\Omega)$. Theorem 10.2.2 implies

$$\|u_n - u_m\|_{L^2(\Omega)} \le \text{const}\,\|Du_n - Du_m\|_{L^2(\Omega)}\,.$$

Thus, (u_n) also is a Cauchy sequence in $L^2(\Omega)$. We conclude that $(u_n)_{n\in\mathbb{N}}$ converges in $W^{1,2}(\Omega)$ to some u. This u satisfies

$$\int_\Omega |Du|^2 = \kappa$$

as well as

$$u - g \in H_0^{1,2}(\Omega),$$

because $H_0^{1,2}(\Omega)$ is a closed subspace of $W^{1,2}(\Omega)$. Furthermore, for every $v \in H_0^{1,2}(\Omega)$, $t \in \mathbb{R}$, putting $Du \cdot Dv := \sum_{i=1}^d D_i u \cdot D_i v$, we have

$$\kappa \le \int_\Omega |D(u+tv)|^2 = \int_\Omega |Du|^2 + 2t\int_\Omega Du \cdot Dv + t^2 \int_\Omega |Dv|^2 ,$$

and differentiating with respect to t at $t = 0$ yields

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |D(u+tv)|^2 |_{t=0} = 2\int_\Omega Du \cdot Dv \quad \text{for all } v \in H_0^{1,2}(\Omega).$$

Definition 10.3.1. A function $u \in H^{1,2}(\Omega)$ is called weakly harmonic, or a weak solution of the Laplace equation, if

$$\int_\Omega Du \cdot Dv = 0 \quad \text{for all } v \in H_0^{1,2}(\Omega). \tag{10.3.1}$$

Any harmonic function obviously satisfies (10.3.1). In order to obtain a harmonic function from the Dirichlet principle one has to show that, conversely, any solution of (10.3.1) is twice continuously differentiable, hence harmonic. In the present case, this follows directly from Corollary 2.2.1:

Corollary 10.3.1. *Any weakly harmonic function is smooth and harmonic. In particular, applying the Dirichlet principle yields harmonic functions. More precisely, for any open and bounded Ω in $\mathbb{R}^d$, $g \in H^{1,2}(\Omega)$, there exists a function $u \in H^{1,2}(\Omega) \cap C^\infty(\Omega)$ with*

$$\Delta u = 0 \quad \textit{in } \Omega$$

and

$$u - g \in H_0^{1,2}(\Omega).$$

The proof of Corollary 10.3.1 depends on the rotational invariance of the Laplace operator and therefore cannot be generalized. For that reason, in the sequel, we want to develop a more general approach to regularity theory. Before turning to that theory, however, we wish to slightly extend the situation just considered.

Definition 10.3.2. Let $f \in L^2(\Omega)$. A function $u \in H^{1,2}(\Omega)$ is called a weak solution of the Poisson equation $\Delta u = f$ if for all $v \in H_0^{1,2}(\Omega)$,

$$\int_\Omega Du \cdot Dv + \int_\Omega f v = 0. \tag{10.3.2}$$

Remark. For given boundary values g (meaning $u - g \in H_0^{1,2}(\Omega)$), a solution can be obtained by minimizing

$$\frac{1}{2}\int_\Omega |Dw|^2 + \int_\Omega f w$$

inside the class of all $w \in H^{1,2}(\Omega)$ with $w - g \in H_0^{1,2}(\Omega)$. Note that this expression is bounded from below by the Poincaré inequality (Theorem 10.2.2), because we are assuming fixed boundary values g.

Lemma 10.3.1 (Stability lemma). *Let $u_{i=1,2}$ be a weak solution of $\Delta u_i = f_i$ with $u_1 - u_2 \in H_0^{1,2}(\Omega)$. Then*

$$\|u_1 - u_2\|_{W^{1,2}(\Omega)} \le \text{const} \, \|f_1 - f_2\|_{L^2(\Omega)} \,.$$

In particular, a weak solution of $\Delta u = f$, $u - g \in H_0^{1,2}(\Omega)$ is uniquely determined.

Proof. We have

$$\int_\Omega D(u_1 - u_2)Dv = -\int_\Omega (f_1 - f_2)v \quad \text{for all } v \in H_0^{1,2}(\Omega),$$

and thus in particular,

$$\begin{aligned}\int_\Omega D(u_1 - u_2)D(u_1 - u_2) &= -\int_\Omega (f_1 - f_2)(u_1 - u_2) \\ &\le \|f_1 - f_2\|_{L^2(\Omega)} \|u_1 - u_2\|_{L^2(\Omega)} \\ &\le \text{const} \, \|f_1 - f_2\|_{L^2(\Omega)} \|Du_1 - Du_2\|_{L^2(\Omega)}\end{aligned}$$

by Theorem 10.2.2, and hence

$$\|Du_1 - Du_2\|_{L^2(\Omega)} \le \text{const} \, \|f_1 - f_2\|_{L^2(\Omega)} \,.$$

The claim follows by applying Theorem 10.2.2 once more. □

We have thus obtained the existence and uniqueness of weak solutions of the Poisson equation in a very simple manner. The task of regularity theory then consists in showing that (for sufficiently well-behaved f) a weak solution is of class C^2 and thus also a classical solution of $\Delta u = f$.

We shall present three different methods, namely, the so-called L^2-theory, the theory of strong solutions, and the C^α-theory. The L^2-theory will be developed in Chap. 11, the theory of strong solutions in Chap. 12, and the C^α-theory in Chap. 13.

10.4 Quadratic Variational Problems

We may ask whether the Dirichlet principle can be generalized to obtain solutions of other PDEs. In general, of course, a minimizer u of some variational problem has to satisfy the corresponding Euler–Lagrange equations, first in the weak sense, and if u is regular, also in the classical sense. In the general case, however, regularity theory encounters obstacles, and weak solutions of Euler–Lagrange equations need not always be regular. We therefore restrict ourselves to quadratic variational problems and consider

$$\begin{aligned} I(u) := \int_\Omega \Bigg\{ &\sum_{i,j=1}^d a^{ij}(x) D_i u(x) D_j u(x) \\ &+ 2\sum_{j=1}^d b^j(x) D_j u(x) u(x) + c(x) u(x)^2 \Bigg\} \mathrm{d}x. \end{aligned} \tag{10.4.1}$$

We require the symmetry condition $a^{ij} = a^{ji}$ for all i, j. In addition, the coefficients $a^{ij}(x)$, $b^j(x)$, $c(x)$ should all be bounded. Then $I(u)$ is defined for $u \in H^{1,2}(\Omega)$. As before, we compute, for $\varphi \in H_0^{1,2}(\Omega)$,

$$\begin{aligned} I(u + t\varphi) = I(u) + 2t \int_\Omega \Bigg\{ &\sum_{i,j} a^{ij} D_i u D_j \varphi + \sum_j b^j u D_j \varphi \\ &+ \Big(\sum_j b^j D_j u + cu \Big) \varphi \Bigg\} \mathrm{d}x + t^2 I(\varphi). \end{aligned} \tag{10.4.2}$$

A minimizer u thus satisfies, as before,

$$\frac{\mathrm{d}}{\mathrm{d}t} I(u + t\varphi)|_{t=0} = 0 \quad \text{for all } \varphi \in H_0^{1,2}(\Omega); \tag{10.4.3}$$

hence

$$\int_\Omega \Bigg\{ \sum_j \Big(\sum_i a^{ij} D_i u + b^j u \Big) D_j \varphi + \Big(\sum_j b^j D_j u + cu \Big) \varphi \Bigg\} \mathrm{d}x = 0 \tag{10.4.4}$$

for all $\varphi \in H_0^{1,2}(\Omega)$.

If $u \in C^2(\Omega)$ and $a^{ij}, b^j \in C^1(\Omega)$, then (10.4.4) implies the differential equation

$$\sum_{j=1}^{d} \frac{\partial}{\partial x^j}\left(\sum_{i=1}^{d} a^{ij}(x)\frac{\partial u}{\partial x^i} + b^j(x)u\right) - \sum_{j=1}^{d} b^j(x)\frac{\partial u}{\partial x^j} - c(x)u = 0. \qquad (10.4.5)$$

As the Euler–Lagrange equation of a quadratic variational integral, we thus obtain a linear PDE of second order. This equation is elliptic when we assume that the matrix $(a^{ij}(x))_{i,j=1,\ldots,d}$ is positive definite at every $x \in \Omega$.

In the next chapter we should see that weak solutions of (10.4.5) [i.e., solutions of (10.4.4)] are regular, provided that appropriate assumptions for the coefficients a^{ij}, b^j, and c hold. The direct method of the calculus of variations, as this generalization of the Dirichlet principle is called, consists in finding a weak solution of (10.4.5) by minimizing $I(u)$, and then demonstrating its regularity. We finally wish to study the transformation behavior of the Dirichlet integral and the Laplace operator with respect to changes of the independent variables. We shall also need that transformation rule for our investigation of boundary regularity in the next chapter.

Thus let

$$\xi \to x(\xi)$$

be a diffeomorphism from Ω' to Ω. We put

$$g_{ij} := \sum_{\alpha=1}^{d} \frac{\partial x^\alpha}{\partial \xi^i}\frac{\partial x^\alpha}{\partial \xi^j}, \qquad (10.4.6)$$

$$g^{ij} := \sum_{\alpha=1}^{d} \frac{\partial \xi^i}{\partial x^\alpha}\frac{\partial \xi^j}{\partial x^\alpha}, \qquad (10.4.7)$$

i.e.,

$$\sum_{k=1}^{d} g_{ki}g^{kj} = \delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j, \end{cases}$$

and

$$g := \det\left(g_{ij}\right)_{i,j=1,\ldots,d}. \qquad (10.4.8)$$

We then have, for $u(\xi(x))$,

$$\sum_{\alpha=1}^{d}\left(\frac{\partial u}{\partial x^\alpha}\right)^2 = \sum_{\alpha=1}^{d}\sum_{i,j=1}^{d} \frac{\partial u}{\partial \xi^i}\frac{\partial \xi^i}{\partial x^\alpha}\frac{\partial u}{\partial \xi^j}\frac{\partial \xi^j}{\partial x^\alpha} = \sum_{i,j=1}^{d} g^{ij}\frac{\partial u}{\partial \xi^i}\frac{\partial u}{\partial \xi^j}. \qquad (10.4.9)$$

The Dirichlet integral thus transforms via

$$\int_{\Omega} \sum_{\alpha=1}^{d} \left(\frac{\partial u}{\partial x^{\alpha}} \right)^2 \mathrm{d}x = \int_{\Omega'} \sum_{i,j=1}^{d} g^{ij} \frac{\partial u}{\partial \xi^i} \frac{\partial u}{\partial \xi^j} \sqrt{g} \mathrm{d}\xi. \tag{10.4.10}$$

By (10.4.5), the Euler–Lagrange equation for the integral on the right-hand side is

$$\frac{1}{\sqrt{g}} \sum_{j=1}^{d} \left(\frac{\partial}{\partial \xi^j} \left(\sqrt{g} \sum_{i=1}^{d} g^{ij} \frac{\partial u}{\partial \xi^i} \right) \right) = 0, \tag{10.4.11}$$

where we have added the normalization factor $1/\sqrt{g}$. This means that under our substitution $x = x(\xi)$ of the independent variables, the Laplace equation, i.e., the Euler–Lagrange equation for the Dirichlet integral, is transformed into (10.4.11).

Likewise, (10.4.5) is transformed into

$$\begin{aligned} \frac{1}{\sqrt{g}} \sum_{j=1}^{d} \frac{\partial}{\partial \xi^j} \Bigg(\sqrt{g} \Bigg(\sum_{i,\alpha,\beta=1}^{d} a^{\alpha\beta}(x) \frac{\partial \xi^i}{\partial x^{\alpha}} \frac{\partial \xi^j}{\partial x^{\beta}} \frac{\partial}{\partial \xi^i} u + \sum_{\alpha} b^{\alpha}(x) \frac{\partial \xi^j}{\partial x^{\alpha}} u \Bigg) \Bigg) \\ - \sum_{j,\alpha} b^{\alpha}(x) \frac{\partial \xi^j}{\partial x^{\alpha}} \frac{\partial u}{\partial \xi^j} - c(x) u = 0, \end{aligned} \tag{10.4.12}$$

where $x = x(\xi)$ has to be inserted, of course.

10.5 Abstract Hilbert Space Formulation of the Variational Problem. The Finite Element Method

The present section presents an abstract version of the approach described in Sect. 10.3 together with a method for constructing an approximate solution.

We again set out from some model problem, the Poisson equation with homogeneous boundary data

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{10.5.1}$$

In Definition 10.3.2 we introduced a weak version of that problem, namely the problem of finding a solution u in the Hilbert space $H_0^{1,2}(\Omega)$ of

$$\int_{\Omega} Du\, D\varphi + \int_{\Omega} f\varphi = 0 \quad \text{for all } \varphi \in H_0^{1,2}(\Omega). \tag{10.5.2}$$

This problem can be generalized as an abstract Hilbert space problem that we now wish to describe:

Definition 10.5.1. Let $(H, (\cdot,\cdot))$ be a Hilbert space with associated norm $\|\cdot\|$, $A : H \times H \to \mathbb{R}$ a continuous symmetric bilinear form. Here, continuity means that there exists a constant C such that for all $u, v \in H$,

$$A(u,v) \le C\, \|u\|\, \|v\| .$$

Symmetry means that for all $u, v \in H$,

$$A(u,v) = A(v,u).$$

The form A is called elliptic, or coercive, if there exists a positive λ such that for all $v \in H$,

$$A(v,v) \ge \lambda\, \|v\|^2 . \tag{10.5.3}$$

In our example, $H = H_0^{1,2}(\Omega)$, and

$$A(u,v) = \frac{1}{2}\int_\Omega Du \cdot Dv. \tag{10.5.4}$$

Symmetry is obvious here, continuity follows from Hölder's inequality, and ellipticity results from

$$\frac{1}{2}\int Du \cdot Du = \frac{1}{2}\, \|Du\|^2_{L^2(\Omega)}$$

and the Poincaré inequality (Theorem 10.2.2), which implies for $u \in H_0^{1,2}(\Omega)$,

$$\|u\|_{H_0^{1,2}(\Omega)} \le \text{ const } \|Du\|_{L^2(\Omega)} .$$

Moreover, for $f \in L^2(\Omega)$,

$$L : H_0^{1,2}(\Omega) \to \mathbb{R}, \qquad v \mapsto \int_\Omega f v,$$

yields a continuous linear map on $H_0^{1,2}(\Omega)$ (even on $L^2(\Omega)$).

Namely,

$$\|L\| := \sup_{v \ne 0} \frac{|Lv|}{\|v\|_{W^{1,2}(\Omega)}} \le \|f\|_{L^2(\Omega)} ,$$

for by Hölder's inequality,

$$\int_\Omega f v \le \|f\|_{L^2(\Omega)}\, \|v\|_{L^2(\Omega)} \le \|f\|_{L^2(\Omega)}\, \|v\|_{W^{1,2}(\Omega)} .$$

Of course, the purpose of Definition 10.5.1 is to isolate certain abstract assumptions that allow us to treat not only the Dirichlet integral, but also more general variational problems as considered in Sect. 10.4. However, we do need to impose certain restrictions, in particular for satisfying the ellipticity condition. We consider

$$A(u,v) := \frac{1}{2}\int_{\Omega}\left\{\sum_{i,j=1}^{d} a^{ij}(x)D_i u(x)D_j v(x) + c(x)u(x)v(x)\right\}\mathrm{d}x,$$

with $u, v \in H = H_0^{1,2}(\Omega)$, where we assume:

(A) Symmetry:

$$a^{ij}(x) = a^{ji}(x) \quad \text{for all } i, j, \text{ and } x \in \Omega.$$

(B) Ellipticity: There exists $\lambda > 0$ with

$$\sum_{i,j=1}^{d} a^{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2 \quad \text{for all } x \in \Omega, \xi \in \mathbb{R}^d.$$

(C) Boundedness: There exists $\Lambda < \infty$ with

$$|c(x)|, |a^{ij}| \leq \Lambda \quad \text{for all } i, j, \text{ and } x \in \Omega.$$

(D) Nonnegativity:

$$c(x) \geq 0 \quad \text{for all } x \in \Omega.$$

The ellipticity condition (B) and the nonnegativity (D) imply that

$$A(v,v) \geq \frac{1}{2}\lambda\int_{\Omega} Dv \cdot Dv \quad \text{for all } v \in H_0^{1,2}(\Omega),$$

and using the Poincaré inequality, we obtain

$$A(v,v) \geq \frac{\lambda}{2}\|v\|_{H^{1,2}(\Omega)} \quad \text{for all } v \in H_0^{1,2}(\Omega);$$

i.e., A is elliptic in the sense of Definition 10.5.1. The continuity of A of course follows from the boundedness condition (C), and the symmetry is condition (A).

Theorem 10.5.1. *Let $(H, (\cdot,\cdot))$ be a Hilbert space with norm $\|\cdot\|$, $V \subset H$ convex and closed, $A : H \times H \to \mathbb{R}$ a continuous symmetric elliptic bilinear form, $L : H \to \mathbb{R}$ a continuous linear map. Then*

$$J(v) := A(v,v) + L(v)$$

has precisely one minimizer u in V.

Remark. The solution u depends not only on A and L but also on V, for it solves the problem

$$J(u) = \inf_{v \in V} J(v).$$

Proof. By ellipticity of A, J is bounded from below, namely,

$$J(v) \geq \lambda \|v\|^2 - \|L\| \|v\| \geq -\frac{\|L\|^2}{4\lambda}.$$

We put

$$\kappa := \inf_{v \in V} J(v).$$

Now let $(u_n)_{n \in \mathbb{N}} \subset V$ be a minimizing sequence, i.e.,

$$\lim_{n \to \infty} J(u_n) = \kappa. \tag{10.5.5}$$

We claim that $(u_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, from which we then deduce, since V is closed, the existence of a limit

$$u = \lim_{n \to \infty} u_n \in V.$$

The Cauchy property is verified as follows: By definition of κ,

$$\kappa \leq J\left(\frac{u_n + u_m}{2}\right) = \frac{1}{2} J(u_n) + \frac{1}{2} J(u_m) - \frac{1}{4} A(u_n - u_m, u_n - u_m).$$

(Here, we have used that if u_n and u_m are in V, so is $\frac{u_n + u_m}{2}$, because V is convex.)

Since $J(u_n)$ and $J(u_m)$ by (10.4.5) for $n, m \to \infty$ both converge to κ, we deduce that

$$A(u_n - u_m, u_n - u_m)$$

converges to 0 for $n, m \to \infty$. Ellipticity then implies that $\|u_n - u_m\|$ converges to 0 as well, and hence the Cauchy property.

Since J is continuous, the limit u satisfies

$$J(u) = \lim_{n \to \infty} J(u_n) = \inf_{v \in V} J(v)$$

by the choice of the sequence $(u_n)_{n \in \mathbb{N}}$.

The preceding proof yields uniqueness of u, too. It is instructive, however, to see this once more as a consequence of the convexity of J: Thus, let u_1, u_2 be two minimizers, i.e.,

$$J(u_1) = J(u_2) = \kappa = \inf_{v \in V} J(v).$$

Since together with u_1 and u_2, $\frac{u_1+u_2}{2}$ is also contained in the convex set V, we have

$$\begin{aligned}\kappa &\le J\left(\frac{u_1+u_2}{2}\right) = \frac{1}{2}J(u_1) + \frac{1}{2}J(u_2) - \frac{1}{4}A(u_1-u_2, u_1-u_2)\\ &= \kappa - \frac{1}{4}A(u_1-u_2, u_1-u_2),\end{aligned}$$

and thus $A(u_1-u_2, u_1-u_2) = 0$, which by ellipticity of A implies $u_1 = u_2$. □

Remark. Theorem 10.5.1 remains true without the symmetry assumption for A. This is the content of the Lax–Milgram theorem, proved in Appendix A.

This remark allows us also to treat variational integrands that in addition to the symmetric terms

$$\sum_{i,j=1}^{d} a^{ij}(x) D_i D_j v(x) \quad (a^{ij} = a^{ji})$$

and $c(x)u(x)v(x)$ also contain terms of the form $2\sum_{j=1}^{d} b^j(x) D_j u(x) v(x)$ as in (10.4.1). Of course, we need to impose conditions on the function $b^j(x)$ so as to guarantee boundedness and nonnegativity (the latter requires bounds on $|b^j(x)|$ depending on λ and a lower bound for $|c(x)|$). We leave the details to the reader.

Corollary 10.5.1. *The other assumptions of the previous theorem remaining in force, now let V be a closed linear (hence convex) subspace of H. Then there exists precisely one $u \in V$ that solves*

$$2A(u,\varphi) + L(\varphi) = 0 \quad \textit{for all } \varphi \in V. \tag{10.5.6}$$

Proof. The point u is a critical point (e.g., a minimum) of the functional

$$J(v) = A(v,v) + L(v)$$

in V precisely if

$$2A(v,\varphi) + L(\varphi) = 0 \quad \text{for all } \varphi \in V.$$

Namely, that u is a critical point means here that

$$\frac{\mathrm{d}}{\mathrm{d}t} J(u + t\varphi)_{|t=0} = 0 \quad \text{for all } \varphi \in V.$$

This, however, is equivalent to

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(A(u+t\varphi, u+t\varphi) + L(u+t\varphi))_{|t=0} = 2A(u,\varphi) + L(\varphi).$$

Conversely, if that holds, then

$$J(u+t\varphi) = J(u) + t(2A(u,\varphi) + L(\varphi)) + t^2 A(\varphi,\varphi) \ge J(u)$$

for all $\varphi \in V$, and u thus is a minimizer. The existence and uniqueness of a minimizer established in the theorem thus yields the corollary. □

For our example $A(u,v) = \frac{1}{2}\int Du \cdot Dv, L(v) = \int fv$ with $f \in L^2(\Omega)$, Corollary 10.5.1 thus yields the existence of some $u \in H_0^{1,2}(\Omega)$ satisfying

$$\int_\Omega Du \cdot D\varphi + \int_\Omega f\varphi = 0, \tag{10.5.7}$$

i.e, a weak solution of the Poisson equation in the sense of Definition 10.3.2.

As explained above, the assumptions apply to more general variational problems, and we deduce the following result from Corollary 10.5.1:

Corollary 10.5.2. *Let $\Omega \subset \mathbb{R}^d$ be open and bounded, and let the functions $a^{ij}(x)$ $(i,j = 1,\dots,d)$ and $c(x)$ satisfy the above assumptions (A)–(D). Let $f \in L^2(\Omega)$. Then there exists a unique $u \in H_0^{1,2}(\Omega)$ satisfying*

$$\int_\Omega \left\{ \sum_{i,j=1}^d a^{ij}(x) D_i u(x) D_j \varphi(x) + c(x)u(x)\varphi(x) \right\} \mathrm{d}x$$

$$= \int_\Omega f(x)\varphi(x)\mathrm{d}x \quad \textit{for all } \varphi \in H_0^{1,2}(\Omega).$$

Thus, we obtain a weak solution of

$$-\sum_{i,j=1}^d \frac{\partial}{\partial x^i}\left(a^{ij}(x)\frac{\partial}{\partial x^j}u(x)\right) + c(x)u(x) = f(x)$$

with $u = 0$ on $\partial\Omega$. Of course, so far, this equation does not yet make sense, since we do not know yet whether our weak solution u is regular, i.e., of class $C^2(\Omega)$. This issue, however, will be addressed in the next chapter.

We now want to compare the solution of our variational problem $J(v) \to \min$ in H with the one obtained in the subspace V of H.

Lemma 10.5.1. *Let $A : H \times H \to \mathbb{R}$ be a continuous, symmetric, elliptic, bilinear form in the sense of Definition 10.5.1, and let $L : H \to \mathbb{R}$ be linear and continuous. We consider once more the problem*

$$J(v) := A(v,v) + L(v) \to \min. \tag{10.5.8}$$

Let u be the solution in H u_V the solution in the closed linear subspace V. Then

$$\|u - u_V\| \le \frac{C}{\lambda} \inf_{v \in V} \|u - v\| \tag{10.5.9}$$

with the constants C and λ from Definition 10.5.1.

Proof. By Corollary 10.5.1,

$$2A(u,\varphi) + L(\varphi) = 0 \quad \text{for all } \varphi \in H,$$
$$2A(u_V,\varphi) + L(\varphi) = 0 \quad \text{for all } \varphi \in V,$$

hence also

$$2A(u - u_V,\varphi) = 0 \quad \text{for all } \varphi \in V. \tag{10.5.10}$$

For $v \in V$, we thus obtain

$$\begin{aligned}
\|u - u_V\|^2 &\le \frac{1}{\lambda} A(u - u_V, u - u_V) \text{ by ellipticity of } A \\
&= \frac{1}{\lambda} A(u - u_V, u - v) + \frac{1}{\lambda} A(u - u_V, v - u_V) \\
&= \frac{1}{\lambda} A(u - u_V, u - v) \text{ from (10.5.10) with } \varphi = v - u_V \in V \\
&\le \frac{C}{\lambda} \|u - u_V\| \, \|u - v\| ,
\end{aligned}$$

and since the inequality holds for arbitrary $v \in V$, (10.5.9) follows. □

This lemma is the basis for an important numerical method for the approximative solution of variational problems. Since numerically only finite-dimensional problems can be solved, it is necessary to approximate infinite-dimensional problems by finite-dimensional ones. Thus, $J(v) \to \min$ cannot be solved in an infinite-dimensional Hilbert space like $H = H_0^{1,2}(\Omega)$, but one needs to replace H by some finite-dimensional subspace V of H that on the one hand can easily be handled numerically and on the other hand possesses good approximation properties. These requirements are satisfied well by the finite element spaces. Here, the region Ω is subdivided into polyhedra that are as uniform as possible, for example, triangles or squares in the two-dimensional case (if the boundary of Ω is curved, of course, it can only be approximated by such a polyhedral subdivision). The finite elements then are simply piecewise polynomials of a given degree. This means that the restriction of such a finite element ψ onto each polyhedron occurring in the subdivision is a polynomial. In addition, one usually requires that across the boundaries between the polyhedra, ψ be continuous or even satisfy certain specified differentiability properties. The simplest such finite elements are piecewise linear functions on triangles, where the continuity requirement is satisfied by choosing the coefficients on neighboring triangles approximately. The theory of numerical mathematics then derives several approximation theorems of the type sketched above. This is not particularly difficult and rather elementary, but somewhat lengthy and therefore not pursued here. We rather refer to the corresponding textbooks like Strang–Fix [30] or Braess [3].

The quality of the approximation of course depends not only on the degree of the polynomials but also on the scale of the subdivision employed. Typically, it makes sense to work with a fixed polynomial degree, for example, admitting only piecewise linear or quadratic elements, and make the subdivision finer and finer.

As presented here, the method of finite elements depends on the fact that according to some abstract theorem, one is assured of the existence (and uniqueness) of a solution of the variational problem under investigation and that one can approximate that solution by elements of cleverly chosen subspaces. Even though that will not be necessary for the theoretical analysis of the method, for reasons of mathematical consistency it might be preferable to avoid the abstract existence result and to convert the finite-dimensional approximations into a constructive existence proof instead. This is what we now wish to do.

Theorem 10.5.2. *Let $A : H \times H \to \mathbb{R}$ be a continuous, symmetric, elliptic, bilinear form on the Hilbert space $(H, (\cdot,\cdot))$ with norm $\|\cdot\|$, and let $L : H \to \mathbb{R}$ be linear and continuous. We consider the variational problem*

$$J(v) = A(v, v) + L(v) \to \min.$$

Let $(V_n)_{n\in\mathbb{N}} \subset H$ be an increasing (i.e., $V_n \subset V_{n+1}$ for all n) sequence of closed linear subspaces exhausting H in the sense that for all $v \in H$ and $\delta > 0$, there exist $n \in \mathbb{N}$ and $v_n \in V_n$ with

$$\|v - v_n\| < \delta.$$

Let u_n be the solution of the problem

$$J(v) \to \min \text{ in } V_n$$

obtained in Theorem 10.5.1. Then $(u_n)_{n\in\mathbb{N}}$ converges for $n \to \infty$ towards a solution of

$$J(v) \to \min \text{ in } H.$$

Proof. Let

$$\kappa := \inf_{v\in H} J(v).$$

We want to show that

$$\lim_{n\to\infty} J(u_n) = \kappa.$$

In that case, $(u_n)_{n\in\mathbb{N}}$ will be a minimizing sequence for J in H, and thus it will converge to a minimizer of J in H by the proof of Theorem 10.5.1. We shall proceed by contradiction and thus assume that for some $\varepsilon > 0$ and all $n \in \mathbb{N}$,

$$J(u_n) \geq \kappa + \varepsilon \tag{10.5.11}$$

(since $V_n \subset V_{n+1}$, we have $J(u_{n+1}) \leq J(u_n)$ for all n, by the way).

By definition of κ, there exists some $u_0 \in H$ with

$$J(u_0) < \kappa + \varepsilon/2. \tag{10.5.12}$$

For every $\delta > 0$, by assumption, there exist some $n \in \mathbb{N}$ and some $v_n \in V_n$ with

$$\|u_0 - v_n\| < \delta.$$

With $w_n := v_n - u_0$, we then have

$$\begin{aligned}
|J(v_n) - J(u_0)| &\le |A(v_n, v_n) - A(u_0, u_0)| + |L(v_n) - L(u_0)| \\
&\le A(w_n, w_n) + 2|A(w_n, u_0)| + \|L\|\,\|w_n\| \\
&\le C\,\|w_n\|^2 + 2C\,\|w_n\|\,\|u_0\| + \|L\|\,\|w_n\| \\
&< \varepsilon/2
\end{aligned}$$

for some appropriate choice of δ.

Thus

$$J(v_n) < J(u_0) + \varepsilon/2 < \kappa + \varepsilon \quad \text{by (10.5.12)} < J(u_n) \quad \text{by (10.5.11)},$$

contradicting the minimizing property of u_n.

This contradiction shows that $(u_n)_{n\in\mathbb{N}}$ indeed is a minimizing sequence, implying the convergence to a minimizer as already explained. □

We thus have a constructive method for the (approximative) solution of our variational problem when we choose all the V_n as suitable finite-dimensional subspaces of H. For each V_n, by Corollary 10.5.1, one needs to solve only a finite linear system, with $\dim V_n$ equations; namely, let $e_1, \ldots, e_N$ be a basis of V_n. Then (10.5.6) is equivalent to the N linear equations for $u_n \in V_n$,

$$2A(u_n, e_j) + L(e_j) = 0 \quad \text{for } j = 1, \ldots, N. \tag{10.5.13}$$

Of course, the more general quadratic variational problems studied in Sect. 10.4 can also be covered by this method; we leave this as an exercise.

10.6 Convex Variational Problems

In the preceding sections, we have studied quadratic variational problems, and we provided an abstract Hilbert space interpretation of Dirichlet's principle. In this section, we shall find out that what is essential is not the quadratic structure of the integrand, but rather the fact that the integrand satisfies suitable bounds. In addition, we need the key assumption of convexity of the integrand, and hence, as we shall see, also of the variational integral.

For simplicity, we consider only variational integrals of the form

$$I(u) = \int_\Omega f(x, Du(x))\mathrm{d}x, \tag{10.6.1}$$

where $Du = (D_1 u, \ldots, D_d u)$ denotes the weak derivatives of $u \in H^{1,2}(\Omega)$, instead of admitting more general integrands of the type

$$f(x, u(x), Du(x)). \tag{10.6.2}$$

The additional dependence on the function u itself, instead of just on its derivatives, does not change the results significantly, but it makes the proofs technically more complicated. In Sect. 14.4 below, when we address the regularity of minimizers, we shall even drop the dependence on x and consider only integrands of the form

$$f(Du(x)),$$

in order to make the proofs as transparent as possible while still preserving the essential features.

The main result of this section then is the following theorem:

Theorem 10.6.1. *Let $\Omega \subset \mathbb{R}^d$ be open, and consider a function*

$$f : \Omega \times \mathbb{R}^d \to \mathbb{R}$$

satisfying:

(i) $f(\cdot, v)$ *is measurable for all* $v \in \mathbb{R}^d$.
(ii) $f(x, \cdot)$ *is convex for all* $x \in \Omega$.
(iii) $f(x, v) \geq -\gamma(x) + \kappa|v|^2$ *for almost all* $x \in \Omega$, *all* $v \in \mathbb{R}^d$, *with* $\gamma \in L^1(\Omega)$, $\kappa > 0$.

We let $g \in H^{1,2}(\Omega)$, and we consider the variational problem

$$I(u) := \int_\Omega f(x, Du(x))\mathrm{d}x \to \min$$

among all $u \in H^{1,2}(\Omega)$ with $u - g \in H_0^{1,2}(\Omega)$ (thus, g are boundary values prescribed in the Sobolev sense).

Then I assumes its infimum; i.e., there exists such a u_0 with

$$I(u_0) = \inf_{u-g\in H_0^{1,2}(\Omega)} I(u).$$

To simplify our further considerations, we first observe that it suffices to consider the case $g = 0$. Namely, otherwise, we consider, for $w = u - g$,

$$\tilde{f}(x, w(x)) := f(x, w(x) + g(x)).$$

The function $\tilde{f}$ satisfies the same structural assumptions that f does; this is clear for (i) and (ii), and for (iii), we observe that

$$\tilde{f}(x, w(x)) \geq -\gamma(x) + \kappa |w(x) + g(x)|^2 \geq -\gamma(x) + \kappa \left(\frac{1}{2} |w(x)|^2 - |g(x)|^2 \right),$$

and so $\tilde{f}$ satisfies the analogue of (iii) with

$$\tilde{\gamma}(x) := \gamma(x) + \kappa |g(x)|^2 \in L^1$$

and $\tilde{\kappa} := \frac{1}{2}\kappa$. Thus, for the rest of this section, we assume

$$g = 0. \tag{10.6.3}$$

In order to prepare the proof of the Theorem 10.6.1, we shall first derive some properties of the variational integral I. We point out that in the next two lemmas the function v takes its values in $\mathbb{R}^d$, i.e., is vector- instead of scalar-valued, but that will not influence our reasoning at all.

Lemma 10.6.1. *Suppose that f is as in Theorem 10.6.1, but with (ii) weakened to (ii') $f(x, \cdot)$ is continuous for all $x \in \Omega$,*
and supposing in (iii) only $\kappa \in \mathbb{R}$, but not necessarily $\kappa > 0$.

Then

$$J(v) := \int_\Omega f(x, v(x)) \mathrm{d}x$$

is a lower semicontinuous functional on $L^2(\Omega; \mathbb{R}^d)$.

Proof. We first observe that if v is in L^2, it is measurable, and since $f(x, v)$ is continuous with respect to v, $f(x, v(x))$ then is measurable by a basic result in Lebesgue integration theory.[5] Now let $(v_n)_{n\in\mathbb{N}}$ converge to v in $L^2(\Omega; \mathbb{R}^d)$. By another basic result in Lebesgue integration theory,[6] after selection of a subsequence, (v_n) also converges to v pointwise almost everywhere. (It is legitimate to select a subsequence here, because the subsequent arguments can be applied to any subsequence of (v_n).) By continuity of f,

$$f(x, v(x)) - \kappa |v(x)|^2 = \lim_{n\to\infty} (f(x, v_n(x)) - \kappa |v_n(x)|^2).$$

Since $f(x, v_n(x)) - \kappa |v(x)|^2 \geq -\gamma(x)$ and γ is integrable, we may apply Fatou's lemma[7] to obtain

[5] See p. 214 of [19].

[6] See Lemma A.1 or p. 240 of [19].

[7] See p. 202 of [19].

$$\int_\Omega \left(f(x, v(x)) - \kappa |v(x)|^2\right) \mathrm{d}x \le \liminf_{n\to\infty} \int_\Omega \left(f(x, v_n(x)) - \kappa |v_n(x)|^2\right) \mathrm{d}x,$$

and since (v_n) converges to v in L^2, then also

$$\int_\Omega f(x, v(x))\mathrm{d}x \le \liminf_{n\to\infty} \int_\Omega f(x, v_n(x))\mathrm{d}x.$$

□

Lemma 10.6.2. *Let f be as in Theorem 10.6.1, without necessarily requiring κ in (iii) to be positive. Then*

$$J(v) = \int_\Omega f(x, v(x))\mathrm{d}x$$

is convex on $L^2(\Omega; \mathbb{R}^d)$.

Proof. Let $v_0, v_1 \in L^2(\Omega, \mathbb{R}^d)$, $0 \le t \le 1$. We have

$$\begin{aligned} J(tv_0 + (1-t)v_1) &= \int f(x, tv_0(x) + (1-t)v_1(x)) \\ &\le \int (tf(x, v_0(x)) + (1-t)f(x, v_1(x))) \quad \text{by (ii)} \\ &= tJ(v_0) + (1-t)J(v_1). \end{aligned}$$

Thus, J is convex. □

Lemmas 10.6.1 and 10.6.2 imply the following result:

Lemma 10.6.3. *Let f be as in Theorem 10.6.1, still not necessarily requiring $\kappa > 0$. With our previous simplification $g = 0$ (10.6.3), the functional*

$$I(u) = \int_\Omega f(x, Du(x))\mathrm{d}x$$

is a convex and lower semicontinuous functional on $H_0^{1,2}(\Omega)$.

With Lemma 10.6.3, Theorem 10.6.1 is a consequence of the following abstract result:

Theorem 10.6.2. *Let H be a Hilbert space, with norm $\|\cdot\|$,*

$$I : H \to \mathbb{R} \cup \{\infty\}$$

be bounded from below, not identically equal to $+\infty$, convex and lower semicontinuous. Then, for every $\lambda > 0$, and $u \in H$,

$$I_\lambda(u) := \inf_{y \in H} \left(I(y) + \lambda \|u - y\|^2 \right) \tag{10.6.4}$$

is realized by a unique $u_\lambda \in H$, i.e.,

$$I_\lambda(u) = I(u_\lambda) + \lambda \|u - u_\lambda\|^2, \tag{10.6.5}$$

and if $(u_\lambda)_{\lambda>0}$ remains bounded as $\lambda \searrow 0$, then

$$u_0 := \lim_{\lambda \to 0} u_\lambda$$

exists and minimizes I, i.e.,

$$I(u_0) = \inf_{u \in H} I(u).$$

Proof. We first verify the auxiliary statement about the uniqueness and existence of u_λ. We let $(y_n)_{n\in\mathbb{N}}$ be a minimizing sequence for (10.6.4), i.e.,

$$I(y_n) + \lambda \|u - y_n\|^2 \to \inf_{y \in H} \left(I(y) + \lambda \|u - y\|^2 \right).$$

For $m, n \in \mathbb{N}$, we put

$$y_{m,n} := \frac{1}{2}(y_m + y_n).$$

We then have

$$\begin{aligned} I(y_{m,n}) + \lambda \|u - y_{m,n}\|^2 \le \frac{1}{2} &\left(I(y_m) + \lambda \|u - y_m\|^2 \right) \\ &+ \frac{1}{2} \left(I(y_n) + \lambda \|u - y_n\|^2 \right) - \frac{\lambda}{4} \|y_m - y_n\|^2 \end{aligned} \tag{10.6.6}$$

by the convexity of I and the general Hilbert space identity

$$\left\| x - \frac{1}{2}(y_1 + y_2) \right\|^2 = \frac{1}{2} \left(\|x - y_1\|^2 + \|x - y_2\|^2 \right) - \frac{1}{4} \|y_1 - y_2\|^2 \tag{10.6.7}$$

for any $x, y_1, y_2 \in H$, which is easily derived from expressing the norm squares as scalar products and expanding these scalar products.

Now, by definition of $I_\lambda(u)$, the left-hand side of (10.6.6) has to be $\ge I_\lambda(u)$, whereas for $k = m$ and n, $I(y_k) + \lambda \|u - y_k\|^2$ converges to $I_\lambda(u)$, by choice of the sequence (y_k), for $k \to \infty$. This implies that

$$\|y_m - y_n\|^2 \to 0$$

for $m, n \to \infty$. Thus, $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and it converges to a unique limit u_λ. Since $\|\cdot\|^2$ is continuous, and I is lower semicontinuous, u_λ realizes the infimum in (10.6.4); i.e., (10.6.5) holds.

If (u_λ) then remains bounded for $\lambda \to 0$, this minimizing property implies that

$$\lim_{\lambda\to 0} I(u_\lambda) = \inf_{y\in H} I(y). \tag{10.6.8}$$

Thus, for any sequence $\lambda_n \to 0$, (u_{λ_n}) is a minimizing sequence for I.

We now let $0 < \lambda_1 < \lambda_2$. From the definition of u_{λ_1},

$$I(u_{\lambda_2}) + \lambda_1 \|u - u_{\lambda_2}\|^2 \geq I(u_{\lambda_1}) + \lambda_1 \|u - u_{\lambda_1}\|^2 ,$$

and so

$$\begin{aligned} I(u_{\lambda_2}) + \lambda_2 \|u - u_{\lambda_2}\|^2 \geq I(u_{\lambda_1}) &+ \lambda_2 \|u - u_{\lambda_1}\|^2 \\ &+ (\lambda_1 - \lambda_2)\left(\|u - u_{\lambda_1}\|^2 - \|u - u_{\lambda_2}\|^2\right). \end{aligned}$$

Since u_{λ_2} minimizes $I(y) + \lambda_2 \|u - y\|^2$, we conclude from this and $\lambda_1 < \lambda_2$ that

$$\|u - u_{\lambda_1}\|^2 \geq \|u - u_{\lambda_2}\|^2 .$$

This means that

$$\|u - u_\lambda\|^2$$

is a decreasing function of λ, or in other words, it increases as $\lambda \searrow 0$. Since this expression is also bounded by assumption, it has to converge as $\lambda \searrow 0$. In particular, for any $\varepsilon > 0$, we may find $\lambda_0 > 0$ such that for $0 < \lambda_1, \lambda_2 < \lambda_0$,

$$\left| \|u - u_{\lambda_1}\|^2 - \|u - u_{\lambda_2}\|^2 \right| < \frac{\varepsilon}{2}. \tag{10.6.9}$$

We put

$$u_{1,2} := \frac{1}{2}\left(u_{\lambda_1} + u_{\lambda_2}\right).$$

If we assume, without loss of generality, $I(u_{\lambda_1}) \geq I(u_{\lambda_2})$, the convexity of I implies

$$I(u_{1,2}) \leq I(u_{\lambda_1}). \tag{10.6.10}$$

We then have

$$\begin{aligned} I(u_{1,2}) &+ \lambda_1 \|u - u_{1,2}\|^2 \\ &\le I(u_{\lambda_1}) + \lambda_1 \left(\frac{1}{2}\|u - u_{\lambda_1}\| + \frac{1}{2}\|u - u_{\lambda_2}\|^2 - \frac{1}{4}\|u_{\lambda_1} - u_{\lambda_2}\|^2 \right) \\ &\hspace{18em} \text{by (10.6.10) and (10.6.7)} \\ &< I(u_{\lambda_1}) + \lambda_1 \left(\|u - u_{\lambda_1}\|^2 + \frac{\varepsilon}{4} - \frac{1}{4}\|u_{\lambda_1} - u_{\lambda_2}\|^2 \right) \quad \text{by (10.6.9).} \end{aligned}$$

Since u_{λ_1} minimizes $I(y) + \lambda_1 \|u - y\|^2$, we conclude that

$$\|u_{\lambda_1} - u_{\lambda_2}\|^2 < \varepsilon.$$

So, we have shown the Cauchy property of u_λ for $\lambda \searrow 0$, and therefore, we obtain the existence of

$$u_0 = \lim_{\lambda \to 0} u_\lambda.$$

By (10.6.8) and the lower semicontinuity of I, we see that

$$I(u_0) = \inf_{y \in H} I(y).$$

Thus, we have shown the existence of a minimizer of I. This concludes the proof of Theorem 10.6.2, as well as that of Theorem 10.6.1. □

While we shall see in Chap. 11 that the minimizers of the quadratic variational problems studied in the preceding sections of this chapter are smooth, we have to wait until Chap. 14 until we can derive a regularity theorem for minimizers of a class of variational integrals that satisfy similar structural conditions as in Theorem 10.6.1. Let us anticipate here Theorem 14.4.1 below:
Let $f : \mathbb{R}^d \to \mathbb{R}$ be of class C^∞ and satisfy:

(i) There exists a constant $K < \infty$ with

$$\left| \frac{\partial f}{\partial v_i}(v) \right| \le K|v| \quad \textit{for } i = 1, \dots, d \quad (v = (v^1, \dots, v^d) \in \mathbb{R}^d).$$

(ii) There exist constants $\lambda > 0$, $\Lambda < \infty$ with

$$\lambda|\xi|^2 \le \sum_{i,j=1}^{d} \frac{\partial^2 f(v)}{\partial v_i v_j} \xi_i \xi_j \le \Lambda|\xi|^2 \quad \textit{for all } \xi \in \mathbb{R}^d.$$

Let $\Omega \subset \mathbb{R}^d$ be open and bounded. Let $u_0 \in W^{1,2}(\Omega)$ minimize

$$I(u) := \int_\Omega f(Du(x))\mathrm{d}x$$

among all $u \in W^{1,2}(\Omega)$ *with* $u - u_0 \in H_0^{1,2}(\Omega)$. *Then*

$$u_0 \in C^\infty(\Omega).$$

In order to compare the assumptions of this result with those of Theorem 10.6.1, we first observe that (i) implies that there exist constants c and k with

$$|f(v)| \leq c + k\,|v|^2 .$$

Thus, in place of the lower bound in (iii) of Theorem 10.6.1, here we have an upper bound with the same asymptotic growth as $|v| \to \infty$. Thus, altogether, we are considering integrands with quadratic growth. In fact, it is also possible to consider variational integrands that asymptotically grow like $|v|^p$, with $1 < p < \infty$. The existence of a minimizer follows with similar techniques as described here, by working in the Banach space $H_0^{1,p}(\Omega)$ and exploiting a crucial geometric property of those particular Banach spaces, namely, that the unit ball is uniformly convex. The first steps of the regularity proof also do not change significantly, but higher regularity poses a problem for $p \neq 2$.

The lower bound in assumption (ii) above should be compared with the convexity assumption in Theorem 10.6.1. For $f \in C^2(\mathbb{R}^d)$, convexity means

$$\frac{\partial^2 f(v)}{\partial v^i \partial v^j}\xi_i \xi_j \geq 0 \quad \text{for all } \xi = (\xi_1, \dots, \xi_d).$$

Thus, in contrast to the assumption in the regularity theorem, we are not summing here with respect i, and j, and so this is a stronger assumption. On the other hand, we are not requiring a positive lower bound as in the regularity theorem, but only nonnegativity.

The existence of minimizers of variational problems is discussed in more detail in Jost and Li-Jost [21]. The minimizing scheme presented here is put in a broader context in Jost [16].

Summary

The Dirichlet principle consists in finding solutions of the Dirichlet problem

$$\begin{aligned} \Delta u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{on } \partial\Omega, \end{aligned}$$

by minimizing the Dirichlet integral

$$\int_\Omega |Du(x)|^2 \mathrm{d}x$$

among all functions u with boundary values g in the function space $W^{1,2}(\Omega)$ (Sobolev space) (which turns out to be the appropriate space for this task). More generally, one may also treat the Poisson equation

$$\Delta u = f \quad \text{in } \Omega$$

this way, namely, minimizing

$$\int_\Omega |Du(x)|^2 \, \mathrm{d}x + 2\int_\Omega f(x)u(x)\, \mathrm{d}x.$$

A minimizer then satisfies the equation

$$\int_\Omega Du(x)\, D\varphi(x)\, \mathrm{d}x = 0$$

(respectively $\int_\Omega Du(x)D\varphi(x)\,\mathrm{d}x + \int f(x)\varphi(x)\,\mathrm{d}x = 0$ for the Poisson equation) for all $\varphi \in C_0^\infty(\Omega)$. If one manages to show that a minimizer u is regular (e.g., of class $C^2(\Omega)$), then this equation results from integrating the original differential equation (Laplace or Poisson equation, respectively) by parts. However, since the Sobolev space $W^{1,2}(\Omega)$ is considerably larger than the space $C^2(\Omega)$, we first need to show in the next chapter that a solution of this equation (called a "weak" differential equation) is indeed regular.

The Dirichlet principle also works for a more general class of elliptic equations, and it admits an abstract Hilbert space formulation.

Exercises

10.1. Show that the norm

$$|||u||| := \|u\|_{L^2(\Omega)} + \|Du\|_{L^2(\Omega)}$$

is equivalent to the norm $\|u\|_{W^{1,2}(\Omega)}$ (i.e., there are constants $0 < \alpha \le \beta < \infty$ satisfying

$$\alpha|||u||| \le \|u\|_{W^{1,2}(\Omega)} \le \beta|||u||| \quad \text{for all } u \in W^{1,2}(\Omega)).$$

Why does one prefer the norm $\|u\|_{W^{1,2}(\Omega)}$?

10.2. What would be a natural definition of k-times weak differentiability? (The answer will be given in the next chapter, but you might wish to try yourself at this point to define Sobolev spaces $W^{k,2}(\Omega)$ of k-times weakly differentiably functions that are contained in $L^2(\Omega)$ together with all their weak derivatives and to prove results analogous to Theorem 10.2.1 and Corollary 10.2.1 for them.)

10.3. Consider a variational problem of the type

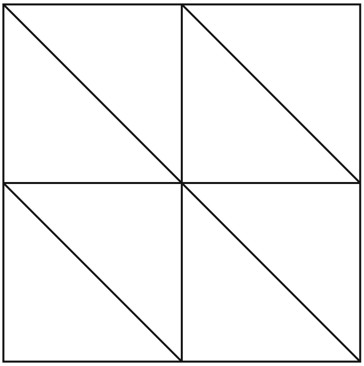

Fig. 10.1

$$I(u) = \int_{\Omega} F(Du(x))\mathrm{d}x$$

with a smooth function $F : \mathbb{R}^d \to \Omega$ satisfying an inequality of the form

$$|F(p)| \leq c_1|p|^2 + c_2 \quad \text{for all } p \in \mathbb{R}^d.$$

Derive the corresponding Euler–Lagrange equations for a minimizer [in the weak sense; cf. (10.4.4)]. Try more generally to find conditions for integrands of the type $F(x, u(x), Du(x))$ that allow one to derive weak Euler–Lagrange equations for minimizers.

10.4. Following R. Courant, as a model problem for finite elements, we consider the Poisson equation

$$\Delta u = f \quad \text{in } \Omega,$$

$$u = 0 \quad \text{on } \partial\Omega$$

in the unit square $\overline{\Omega} = [0, 1]\times[0, 1] \subset \mathbb{R}^2$. For $h = \frac{1}{2^n}$ $(n \in \mathbb{N})$, we subdivide $\overline{\Omega}$ into $\frac{1}{h^2}(= 2^{2n})$ subsquares of side length h, and each such square in turn is subdivided into two right-angled symmetric triangles by the diagonal from the upper left to the lower right vertex (see Fig. 10.1). We thus obtain triangles Δ_i^h, $i = 1, \ldots, 2^{2n+1}$. What is the number of interior vertices p_j of this triangulation?

We consider the space of continuous triangular finite elements

$$S^h := \{\varphi \in C^0(\overline{\Omega}) : \varphi_{|\Delta_i^h} \quad \text{linear for all } i, \varphi = 0 \text{ on } \partial\overline{\Omega}\}.$$

The triangular elements φ_j with

$$\varphi_j(p_i) = \delta_{ij}$$

constitute a basis of S^h (proof?).

Compute

$$a_{ij} := \int_{\overline{\Omega}} D\varphi_i \cdot D\varphi_j \quad \text{for all pairs } i, j$$

land establish the system of linear equations for the approximating solution of the Poisson equation in S^h, i.e., for the minimizer φ^h of

$$\int_{\overline{\Omega}} |D\varphi|^2 + 2\int_{\overline{\Omega}} f\varphi$$

for $\varphi \in S^h$, with respect to the above basis φ_j of S^h (for that purpose, you have just computed the coefficients a_{ij} !).