22. Variational Principles and Partial Differential Equations

Dirichlet's principle consists in constructing harmonic functions by minimizing the Dirichlet integral in an appropriate class of functions. This idea is generalized, and minimizers of variational integrals are weak solutions of the associated differential equations of Euler and Lagrange. Several examples are discussed.

We shall first consider a special example, in order to make prominent the basic idea of the following considerations. The generalization of these reflections will then later present no great difficulty.

The equation to be treated in this example is perhaps the most important partial differential equation for mathematics and physics, namely the Laplace equation.

In the following, Ω will be an open, bounded subset of $\mathbb{R}^d$. A function $f : \Omega \to \mathbb{R}$ is said to be harmonic if it satisfies in Ω the Laplace equation

$$\Delta f(x) = \frac{\partial^2 f(x)}{(\partial x^1)^2} + \ldots + \frac{\partial^2 f(x)}{(\partial x^d)^2} = 0.$$

Harmonic functions occur, for example, in complex analysis. If $\Omega \subset \mathbb{C}$ and $z = x + iy \in \Omega$, and if $f(z) = u(z) + iv(z)$ is holomorphic on Ω, then the so-called Cauchy-Riemann differential equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \; \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{1}$$

hold, and as a holomorphic function is in the class C^∞, we can differentiate (1) and obtain

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

and similarly

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

Thus the real and imaginary parts of a holomorphic function are harmonic.

Conversely, two harmonic functions which satisfy (1) are called conjugate and a pair of conjugate harmonic functions gives precisely the holomorphic function $f = u + iv$.

In case (1) holds, one can interpret $(u(x,y), -v(x,y))$ as the velocity field of a two dimensional rotation-free incompressible fluid. For $d = 3$ the harmonic functions describe likewise the velocity field of a rotation-free incompressible fluid, as well as electrostatic and gravitational fields (outside attracting or repelling charges or attracting masses), temperature distribution in thermal equilibrium, equilibrium states of elastic membranes, etc.

The most important problem in harmonic functions is the Dirichlet problem: Here, a function $g : \partial\Omega \to \mathbb{R}$ is given and one seeks $f : \overline{\Omega} \to \mathbb{R}$ with

$$\begin{aligned} \Delta f(x) &= 0 \quad \text{for } x \in \Omega \\ f(x) &= g(x) \quad \text{for } x \in \partial\Omega. \end{aligned} \tag{2}$$

For example, this models the state of equilibrium of a membrane which is fixed at the boundary of Ω.

There exist various methods to solve the Dirichlet problem for harmonic functions. Perhaps the most important and general is the so-called Dirichlet principle, which we want to introduce now.

In order to pose (2) sensibly, one must make certain assumptions on Ω and g. For the moment we only assume that $g \in W^{1,2}(\Omega)$. As already said, Ω is an open and bounded subset of $\mathbb{R}^d$. Further restrictions will follow in due course in our study of the boundary condition $f = g$ on $\partial\Omega$.

The Dirichlet principle consists in finding a solution of

$$\begin{aligned} \Delta f &= 0 \quad \text{in } \Omega \\ f &= g \quad \text{on } \partial\Omega \text{ (in the sense that } f - g \in H_0^{1,2}(\Omega)) \end{aligned}$$

by minimizing the Dirichlet integral

$$\frac{1}{2}\int_\Omega |Dv|^2 \quad (\text{here } Dv = (D_1 v, \dots D_d v))$$

over all $v \in H^{1,2}(\Omega)$ for which $v - g \in H_0^{1,2}(\Omega)$.

We shall now verify that this method really works.

Let

$$m := \inf\Big\{\frac{1}{2}\int_\Omega |Dv|^2 : v \in H^{1,2}(\Omega), v - g \in H_0^{1,2}(\Omega)\Big\}.$$

We must show that m is assumed and that the function for which it is assumed is harmonic. (Notation: By corollary 20.10, $W^{1,2} = H^{1,2}$ and in the sequel we shall mostly write $H^{1,2}$ for this space.)

Let $(f_n)_{n\in\mathbb{N}}$ be a minimizing sequence, so $f_n - g \in H_0^{1,2}(\Omega)$ and

$$\int_\Omega |Df_n|^2 \to 2m.$$

By corollary 20.16 we have

$$\begin{aligned}||f_n||_{L^2(\Omega)} &\leq ||g||_{L^2(\Omega)} + ||f_n - g||_{L^2(\Omega)} \\ &\leq ||g||_{L^2(\Omega)} + \text{ const. } ||Df_n - Dg||_{L^2(\Omega)} \\ &\leq ||g||_{L^2(\Omega)} + c_1||Dg||_{L^2(\Omega)} + c_2||Df_n||_{L^2(\Omega)} \\ &\leq \text{ const. } + c_2||Df_n||_{L^2(\Omega)},\end{aligned}$$

as g has been chosen to be fixed.

Without loss of generality let

$$||Df_n||^2_{L^2(\Omega)} \leq m + 1.$$

It follows that

$$||f_n||_{H^{1,2}(\Omega)} \leq \text{ const. (independent of } n).$$

By theorem 21.8 f_n converges weakly, after a choice of a subsequence, to an $f \in H^{1,2}(\Omega)$ with $f - g \in H_0^{1,2}(\Omega)$ (this follows from corollary 21.12) and corollary 21.9 gives

$$\int_\Omega |Df|^2 \leq \liminf_{n\to\infty} \int_\Omega |Df_n|^2 = 2m.$$

By the theorem of Rellich (theorem 20.20) the remaining term of $||f_n||^2_{H^{1,2}}$, namely $\int |f_n|^2$ is even continuous, so $\int_\Omega |f|^2 = \lim_{n\to\infty} \int_\Omega |f_n|^2$, after choosing a subsequence of (f_n).

Because of $f - g \in H_0^{1,2}(\Omega)$, it follows from the definition of m that

$$\int_\Omega |Df|^2 = 2m.$$

Furthermore, for every $v \in H_0^{1,2}, t \in \mathbb{R}$ we have

$$m \leq \int_\Omega |D(f+tv)|^2 = \int_\Omega |Df|^2 + 2t \int_\Omega Df \cdot Dv + t^2 \int_\Omega |Dv|^2$$

(where $Df \cdot Dv := \sum_{i=1}^{d} D_i f \cdot D_i v$) and differentiation by t at $t = 0$ gives

$$0 = \frac{d}{dt} \int_\Omega |D(f+tv)|^2|_{t=0} = 2 \int_\Omega Df \cdot Dv$$

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

By the way, this calculation also shows that the map

$$\begin{aligned} E : H^{1,2}(\Omega) &\to \mathbb{R} \\ f &\mapsto \int_\Omega |Df|^2 \end{aligned}$$

is differentiable, with

$$DE(f)(v) = 2 \int_\Omega Df \cdot Dv.$$

Definition 22.1 A function $f \in H^{1,2}(\Omega)$ is called weakly harmonic or a weak solution of the Laplace equation if

$$\int_\Omega Df \cdot Dv = 0 \quad \text{for all } v \in H_0^{1,2}(\Omega). \tag{3}$$

Obviously, every harmonic function satisfies (3). In order to obtain a harmonic function by applying the Dirichlet principle, one has now to show conversely that a solution of (3) is twice continuously differentiable and therefore, in particular, harmonic. This will be achieved in §23.

However, we shall presently treat a more general situation:

Definition 22.2 Let $\varphi \in L^2(\Omega)$. A function $f \in H^{1,2}(\Omega)$ is called a weak solution of the Poisson equation ($\Delta f = \varphi$) if for all $v \in H_0^{1,2}(\Omega)$

$$\int_\Omega Df \cdot Dv + \int_\Omega \varphi \cdot v = 0 \tag{4}$$

holds.

Remark. For a preassigned boundary value g (in the sense of $f-g \in H_0^{1,2}(\Omega)$) a solution of (4) can be obtained by minimizing

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

in the class of all $w \in H^{1,2}(\Omega)$ for which $w - g \in H_0^{1,2}(\Omega)$. One notices that this expression is bounded from below by the Poincaré inequality (corollary 20.16), as we have fixed the boundary value g.

Another possibility of finding a solution of (4) for a preassigned $f - g \in H_0^{1,2}$ is the following:

If one sets $w := f - g \in H_0^{1,2}$, then w has to solve

$$\int_\Omega Dw \cdot Dv = - \int_\Omega \varphi \cdot v - \int_\Omega Dg \cdot Dv \tag{5}$$

for all $v \in H_0^{1,2}$.

The Poincaré inequality (corollary 20.16) implies that a scalar product on $H_0^{1,2}(\Omega)$ is already given by

$$((f,v)) := (Df, Dv)_{L^2(\Omega)} = \int_\Omega Df \cdot Dv.$$

With this scalar product, $H_0^{1,2}(\Omega)$ becomes a Hilbert space. Furthermore,

$$\int_\Omega \varphi \cdot v \leq \|\varphi\|_{L^2} \cdot \|v\|_{L^2} \leq \text{ const. } \|\varphi\|_{L^2} \cdot \|Dv\|_{L^2},$$

again by corollary 20.16. It follows that

$$Lv := -\int_\Omega \varphi \cdot v - \int_\Omega Dg \cdot Dv$$

defines a bounded linear functional on $H_0^{1,2}(\Omega)$. By theorem 21.6 there exists a uniquely determined $w \in H_0^{1,2}(\Omega)$ with

$$((w,v)) = Lv \quad \text{for all } v \in H_0^{1,2},$$

and w then solves (5).

This argument also shows that a solution of (4) is unique. This also follows from the following general result.

Lemma 22.3 *Let $f_i, i = 1,2$, be weak solutions of $\Delta f_i = \varphi_i$ with $f_1 - f_2 \in H_0^{1,2}(\Omega)$. Then*

$$\|f_1 - f_2\|_{W^{1,2}(\Omega)} \leq \text{ const. } \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}.$$

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

Proof. We have

$$\int_\Omega D(f_1 - f_2)Dv = -\int_\Omega (\varphi_1 - \varphi_2)v$$

for all $v \in H_0^{1,2}(\Omega)$ and therefore in particular

$$\begin{aligned}\int_\Omega D(f_1 - f_2)D(f_1 - f_2) &= -\int_\Omega (\varphi_1 - \varphi_2)(f_1 - f_2)\\ &\leq \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}\|f_1 - f_2\|_{L^2(\Omega)}\\ &\leq \text{ const. } \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}\|Df_1 - Df_2\|_{L^2(\Omega)}\end{aligned}$$

by corollary 20.16, and consequently

$$\|Df_1 - Df_2\|_{L^2(\Omega)} \leq \text{ const. } \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}.$$

The assertion follows by another application of corollary 20.16. □

We have thus obtained the existence and uniqueness of weak solutions of the Poisson equation in a very simple manner.

The aim of the regularity theory consists in showing that (for sufficiently well behaved φ) a weak solution is already of class C^2, and thus also a classical solution of $\Delta f = \varphi$. In particular we shall show that a solution of $\Delta f = 0$ is even of class $C^\infty(\Omega)$.

Besides, we must investigate in which sense, if for example $\partial\Omega$ is of class C^∞ – in a sense yet to be made precise – and $g \in C^\infty(\overline{\Omega})$, the boundary condition $f - g \in H_0^{1,2}(\Omega)$ is realized. It turns out that in this case, a solution of $\Delta f = 0$ is of class C^∞ and for all $x \in \partial\Omega \quad f(x) = g(x)$ holds.

We shall now endeavour to make a generalization of the above ideas. For this we shall first summarize the central idea of these considerations:

In order to minimize the Dirichlet integral, we had first observed that there exists a bounded minimizing sequence in $H^{1,2}$. From this we could then choose a weakly convergent subsequence. As the Dirichlet integral is lower semicontinuous with respect to weak convergence the limit of this sequence then yields a minimum. Thus, with this initial step, the existence of a minimum is established. The second important observation then was that a minimum must satisfy, at least in a weak form, a partial differential equation.

We shall now consider a variational problem of the form

$$I(f) := \int_\Omega H(x, f(x), D(f(x)))dx \to \min .$$

under yet to be specified conditions on the real valued function H; here Ω is always an open, bounded subset of $\mathbb{R}^d$ and f is allowed to vary in the space $H^{1,2}(\Omega)$.

Similar considerations could be made in the spaces $H^{1,p}(\Omega)$, but we have introduced the concept of weak convergence only in Hilbert and not in general Banach spaces.

Theorem 22.4 *Let $H : \Omega \times \mathbb{R}^d \to \mathbb{R}$ be non-negative, measurable in the first and convex in the second argument, so $H(x, tp + (1-t)q) \leq tH(x,p) + (1-t)$ $H(x,q)$ holds for all $x \in \Omega, \quad p, q \in \mathbb{R}^d$ and $0 \leq t \leq 1$.*

For $f \in H^{1,2}(\Omega)$ we define

$$I(f) := \int_\Omega H(x, Df(x))dx \leq \infty.$$

Then

$$I : H^{1,2}(\Omega) \to \mathbb{R} \cup \{\infty\}$$

is convex and lower semicontinuous (relative to strong convergence, i.e. if $(f_n)_{n\in\mathbb{N}}$ converges in $H^{1,2}$ to f then

$$I(f) \leq \liminf_{n\to\infty} I(f_n)).$$

(As H is continuous in the second argument (see below) and Df is measurable, $H(x, Df(x))$ is again measurable (by corollary 17.12), so $I(f)$ is well-defined).

Proof. The convexity of I follows from that of H, as the integral is a linear function: Let $f, g \in H^{1,2}(\Omega), 0 \leq t \leq 1$. Then

$$\begin{aligned} I(tf + (1-t)g) &= \int_\Omega H(x, tDf(x) + (1-t)Dg(x))dx \\ &\leq \int_\Omega \{tH(x, Df(x)) + (1-t)H(x, Dg(x))\}dx \\ &= tI(f) + (1-t)I(g). \end{aligned}$$

It remains to show the lower semicontinuity. Let $(f_n)_{n\in\mathbb{N}}$ converge to f in $H^{1,2}$. By choosing a subsequence, we may assume that $\liminf_{n\to\infty} I(f_n) = \lim_{n\to\infty} I(f_n)$. By a further choice of a subsequence, Df_n then converges pointwise to Df almost everywhere. By theorem 19.12 this follows from the fact that Df_n converges in L^2 to Df. As H is continuous in the second variable (see lemma 22.5 infra), $H(x, Df_n(x))$ converges pointwise to $H(x, Df(x))$ almost everywhere on Ω. By the assumption $H \geq 0$ we can apply Fatou's lemma and obtain

$$\begin{aligned} I(f) &= \int_\Omega H(x, Df(x))dx = \int_\Omega \lim_{n\to\infty} H(x, Df_n(x))dx \\ &\leq \liminf_{n\to\infty} \int_\Omega H(x, Df_n(x))dx \\ &= \liminf_{n\to\infty} I(f_n). \end{aligned}$$

(As $\lim I(f_n) = \liminf I(f_n)$, by choice of the first subsequence, $\liminf I(f_n)$ does not change anymore in choosing the second subsequence). Thereby, the lower semicontinuity has been shown. $\square$

We append further the following result:

Lemma 22.5 *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be convex. Then φ is continuous.*

Proof. We must control the difference $|\varphi(y+h) - \varphi(y)|$ for $h \to 0$. We set $\ell := \frac{h}{|h|}$ (we may assume $h \neq 0$) and choose $t \in [0,1]$ with

$$h = (1-t)\ell.$$

By convexity, we have

$$\varphi(ty + (1-t)(y+\ell)) \le t\varphi(y) + (1-t)\varphi(y+\ell)$$

so

$$\varphi(y+h) \le t\varphi(y) + (1-t)\varphi(y+\ell),$$

and therefore

$$\varphi(y+h) - \varphi(y) \le \frac{1-t}{t}(-\varphi(y+h) + \varphi(y+\ell)). \tag{6}$$

The convexity of φ also gives

$$\varphi(y) \le t\varphi(y+h) + (1-t)\varphi(y - t\ell),$$

so

$$\varphi(y+h) - \varphi(y) \ge \frac{1-t}{t}(\varphi(y) - \varphi(y - t\ell)). \tag{7}$$

We now let h approach 0, so $t \to 1$, and obtain the continuity of φ at y from (6) and (7). □

We now prove

Lemma 22.6 *Let A be a convex subset of a Hilbert space, $I : A \to \mathbb{R} \cup \{\pm\infty\}$ be convex and lower semicontinuous. Then I is also lower semicontinuous relative to weak convergence.*

Proof. Let $(f_n)_{n\in\mathbb{N}} \subset A$ be weakly convergent to $f \in A$. We then have to show that

$$I(f) \le \liminf_{n\to\infty} I(f_n). \tag{8}$$

By choosing a subsequence, we may assume that $I(f_n)$ is convergent, say

$$\liminf_{n\to\infty} I(f_n) = \lim_{n\to\infty} I(f_n) =: \omega. \tag{9}$$

By choosing a further subsequence and using the Banach-Saks lemma (corollary 21.10) the convex combination

$$g_k := \frac{1}{k}\sum_{\nu=1}^{k} f_{N+\nu}$$

converges strongly to f as $k \to \infty$, and indeed for every $N \in \mathbb{N}$.

The convexity of I gives

$$I(g_k) \le \frac{1}{k}\sum_{\nu=1}^{k} I(f_{N+\nu}). \tag{10}$$

Now we choose, for $\varepsilon > 0$, N so large that for all $\nu \in \mathbb{N}$

$$I(f_{N+\nu}) < \omega + \varepsilon$$

holds (compare (9)). By (10) it then follows that

$$\limsup_{k\to\infty} I(g_k) \leq \omega.$$

The lower semicontinuity of I relative to strong convergence now gives

$$I(f) \leq \liminf_{k\to\infty} I(g_k) \leq \limsup_{k\to\infty} I(g_k) \leq \omega = \liminf_{n\to\infty} I(f_n).$$

Thereby (8) has been verified. □

We obtain now the important

Corollary 22.7 *Let $H : \Omega \times \mathbb{R}^d \to \mathbb{R}$ be non-negative, measurable in the first and convex in the second argument. For $f \in H^{1,2}(\Omega)$, let*

$$I(f) := \int_\Omega H(x, Df(x))dx.$$

Then I is lower semicontinuous relative to weak convergence in $H^{1,2}$.

Let A be a closed convex subset of $H^{1,2}(\Omega)$.

If there exists a bounded minimizing sequence $(f_n)_{n\in\mathbb{N}} \subset A$, that is,

$$I(f_n) \to \inf_{g\in A} I(g) \text{ with } \|f\|_{H^{1,2}} \leq K,$$

then I assumes its minimum on A, i.e. there is an $f \in A$ with

$$I(f) = \inf_{g\in A} I(g).$$

Proof. The lower semicontinuity follows from theorem 22.4 and lemma 22.6. Now let $(f_n)_{n\in\mathbb{N}}$ be a bounded minimizing sequence. By theorem 21.8, after choosing a subsequence, the sequence f_n converges weakly to an f, which by corollary 21.12 is in A. Due to weak lower semicontinuity it follows that

$$I(f) \leq \liminf_{n\to\infty} I(f_n) = \inf_{g\in A} I(g),$$

and, as trivially $\inf_{g\in A} I(g) \leq I(f)$ holds, the assertion follows. □

Remarks.

1) In corollary 22.7, H depends only on x and $Df(x)$, but not on $f(x)$. In fact, in the general case

$$I(f) = \int_\Omega H(x, f(x), Df(x))dx$$

there are lower semicontinuity results under suitable assumptions on H, but these are considerably more difficult to prove. The only exception is the following statement:
Let $H : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be measurable in the first and jointly convex in the second and third argument, i.e. for $x \in \Omega, f, g \in \mathbb{R}, p, q \in \mathbb{R}^d, 0 \le t \le 1$ one has

$$H(x, tf + (1-t)g, tp + (1-t)q) \le tH(x, f, p) + (1-t)H(x, g, q).$$

Then the results of corollary 22.7 also hold for

$$I(f) := \int_\Omega H(x, f(x), Df(x))dx.$$

The proof of this result is the same as that of corollary 22.7.

2) Weak convergence was a suitable concept for the above considerations due to the following reasons. One needs a convergence concept which, on the one hand, should allow lower semicontinuity statements and so should be as strong as possible, and on the other hand, it should admit a selection principle, so that every bounded sequence contains a convergent subsequence and therefore should be as weak as possible. The concept of weak convergence unites these two requirements.

Example. We now want to consider an important example:
For $i, j = 1, \dots, d$, let $a_{ij} : \Omega \to \mathbb{R}$ be measurable functions with

$$\sum_{i,j=1}^d a_{ij}(x)\xi^i\xi^j \ge \lambda|\xi|^2 \tag{11}$$

for all $x \in \Omega$, $\xi = (\xi^1, \dots, \xi^d) \in \mathbb{R}^d$, with a $\lambda > 0$.
The condition (11) is called an ellipticity condition.

We consider

$$I(f) := \int_\Omega \sum_{i,j=1}^d a_{ij}(x)D_i f(x)D_j f(x)dx$$

for $f \in H^{1,2}(\Omega)$.
We shall also assume that

$$\operatorname*{ess\,sup}_{\substack{x\in\Omega \\ i,j=1,\dots d}} |a_{ij}(x)| \le m. \tag{12}$$

Then $I(f) < \infty$ for all $f \in H^{1,2}(\Omega)$.
By (11) and (12),

$$\lambda \int_{\Omega} |Df(x)|^2 dx \le I(f) \le md \int_{\Omega} |Df(x)|^2 dx \tag{13}$$

holds.

We now observe that

$$\langle f, g \rangle := \frac{1}{2} \int_{\Omega} \sum_{i,j=1}^{d} a_{ij}(x)(D_i f(x) D_j g(x) + D_j f(x) D_i g(x)) dx$$

is bilinear, symmetric and positive semi-definite (so $\langle f, f \rangle \ge 0$ for all f) on $H^{1,2}(\Omega)$. Therefore the Schwarz inequality holds:

$$\langle f, g \rangle \le I(f)^{\frac{1}{2}} \cdot I(g)^{\frac{1}{2}}. \tag{14}$$

It now follows easily that I is convex:

$$\begin{aligned} I(tf + (1-t)g) = \int_{\Omega} \sum_{i,j=1}^{d} & a_{ij}(x)(t^2 D_i f(x) D_j f(x) \\ & + t(1-t)(D_i f(x) D_j g(x) + D_j f(x) D_i g(x)) \\ & + (1-t)^2 D_i g(x) D_j g(x)) dx, \end{aligned}$$

thus

$$\begin{aligned} I(tf + (1-t)g) &= t^2 I(f) + 2t(1-t)\langle f, g \rangle + (1-t)^2 I(g) \\ &\le t^2 I(f) + 2t(1-t) I(f)^{\frac{1}{2}} I(g)^{\frac{1}{2}} + (1-t)^2 I(g) \text{ by (14)} \\ &\le t^2 I(f) + t(1-t)(I(f) + I(g)) + (1-t)^2 I(g) \\ &= tI(f) + (1-t)I(g). \end{aligned}$$

Finally, we also observe that if we restrict ourselves to the space $H_0^{1,2}(\Omega)$, then every minimizing sequence for I is bounded. Namely, for $f \in H_0^{1,2}(\Omega)$, the Poincaré inequality (corollary 20.16) holds:

$$\begin{aligned} \|f\|^2_{H^{1,2}(\Omega)} &\le c \int_{\Omega} |Df(x)|^2 dx \text{ where } c \text{ is a constant} \\ &\le \frac{c}{\lambda} I(f) \text{ by (13)}, \end{aligned} \tag{15}$$

and thereby a minimizing sequence is bounded in $H^{1,2}(\Omega)$.

In general, for a fixed $g \in H^{1,2}(\Omega)$, we can also consider the space

$$A_g := \{ f \in H^{1,2}(\Omega) : f - g \in H_0^{1,2}(\Omega) \}$$

The space A_g is closed and convex and for $f \in A_g$ we have

$$\begin{aligned}\|f\|_{H^{1,2}(\Omega)} &\le \|f-g\|_{H^{1,2}(\Omega)} + \|g\|_{H^{1,2}(\Omega)}\\ &\le \left(\frac{c}{\lambda}I(f-g)\right)^{\frac12} + \|g\|_{H^{1,2}(\Omega)} \text{ since } f-g\in H_0^{1,2}(\Omega)\\ &\le \left(\frac{c}{\lambda}(I(f)+I(g))^2\right)^{\frac12} + \|g\|_{H^{1,2}(\Omega)}\end{aligned}$$

(using the triangle inequality implied by the Schwarz inequality for $I(f)^{\frac12} = \langle f, f\rangle^{\frac12}$)

$$= \left(\frac{c}{\lambda}\right)^{\frac12} I(f) + \left(\frac{c}{\lambda}\right)^{\frac12} I(g) + \|g\|_{H^{1,2}(\Omega)}.$$

As g is fixed, the $H^{1,2}$-norm for a minimizing sequence for I in A_g is again bounded.

We deduce from corollary 22.7 that I assumes its minimum on A_g, i.e. for any $g \in H^{1,2}(\Omega)$ there exists an $f \in H^{1,2}(\Omega)$ with $f-g \in H_0^{1,2}(\Omega)$ and

$$I(f) = \inf\{I(h) : h \in H^{1,2}(\Omega), h-g \in H_0^{1,2}(\Omega)\}.$$

This generalizes the corresponding statements for the Dirichlet integral. In the same manner we can treat, for a given $\varphi \in L^2(\Omega)$

$$J(f) = \int_\Omega \Big(\sum_{i,j=1}^d a_{ij}(x) D_i f(x) D_j f(x) + \varphi(x) f(x)\Big) dx$$

and verify the existence of a minimum with given boundary conditions.

However, not every variational problem admits a minimum:

Examples.

1) We consider, for $f : [-1,1] \to \mathbb{R}$

$$I(f) := \int_{-1}^{1} (f'(x))^2 x^4 dx$$

with boundary conditions $f(-1) = -1, f(1) = 1$. Consider

$$f_n(x) = \begin{cases} -1 & \text{for } -1 \le x < -\frac1n \\ nx & \text{for } -\frac1n \le x \le \frac1n \\ 1 & \text{for } \frac1n < x \le 1 \end{cases}$$

Then $\lim_{n\to\infty} I(f_n) = 0$, but for every f we have $I(f) > 0$. Thus the infimum of $I(f)$, with the given boundary conditions, is not assumed.

2) We shall now consider an example related to the question of realization of boundary values:
Let $\Omega := U(0,1)\backslash\{0\} = \{x \in \mathbb{R}^d : 0 < \|x\| < 1\}, d \ge 2$.
We choose $g \in C^1(\overline{\Omega})$ with

$$g(x) = 0 \quad \text{for } \|x\| = 1$$
$$g(0) = 1$$

We want to minimize the Dirichlet integral over $A_g = \{f \in H^{1,2}(\Omega) : f - g \in H_0^{1,2}(\Omega)\}$. Consider, for $0 < \varepsilon < 1$, $(r = \|x\|)$,

$$f_\varepsilon(r) := \begin{cases} 1 & \text{for } 0 \le r \le \varepsilon \\ \frac{\log(r)}{\log(\varepsilon)} & \text{for } \varepsilon < r \le 1 . \end{cases}$$

By the computation rules given in §20, $f_\varepsilon(r)$ is in $H_0^{1,2}(\Omega)$ and

$$\begin{aligned} \int_\Omega |Df_\varepsilon(r)|^2 dx &= \frac{1}{(\log \varepsilon)^2} \int_{\varepsilon \le r \le 1} \frac{1}{r^2} dx \\ &= \frac{d\omega_d}{(\log \varepsilon)^2} \int_\varepsilon^1 \frac{r^{d-1}}{r^2} dr \quad \text{(theorem 13.21)} \\ &= \begin{cases} \frac{2\pi}{\log \frac{1}{\varepsilon}} & \text{for } d = 2 \\ \frac{1}{(\log \varepsilon)^2} \frac{d\omega_d}{d-2}(1 - \varepsilon^{d-2}) & \text{for } d > 2 \end{cases} \end{aligned}$$

It follows that

$$\lim_{\varepsilon \to 0} \int_\Omega |Df_\varepsilon|^2 = 0,$$

and thereby

$$\inf\{\int_\Omega |Df|^2, f \in A_g\} = 0.$$

Now, it follows from the Poincaré inequality (corollary 20.16) as usual that for a minimizing sequence $(f_n)_{n \in \mathbb{N}} \subset A_g$

$$\|f_n\|_{H^{1,2}} \to 0 \quad \text{for } n \to \infty.$$

Thus f_n converges in $H^{1,2}$ to zero. So the limit $f \equiv 0$ does not fulfil the prescribed boundary condition $f(0) = 1$. The reason for this is that an isolated point is really too small to play a role in the minimizing of Dirichlet integrals. We shall later even see that there exists no function h at all such that

$$h : B(0,1) \to \mathbb{R},\ \Delta h(x) = 0 \text{ for } 0 < \|x\| < 1,$$
$$h(x) = 0 \text{ for } \|x\| = 1 \text{ and } h(0) = 1$$

(see example after theorem 24.4).
The phenomenon which has just appeared can be easily formulated abstractly.

Definition 22.8 Let Ω be open in $\mathbb{R}^d$, $K \subset \Omega$ compact. We define the capacity of K with respect to Ω by

$$\mathrm{cap}_\Omega(K) := \inf\{\int_\Omega |Df|^2 : f \in H_0^{1,2}(\Omega), f \geq 1 \text{ on } K\}.$$

So the capacity of an isolated point in $\mathbb{R}^d$ vanishes for $d \geq 2$.

In general we have

Theorem 22.9 *Let $\Omega \subset \mathbb{R}^d$ be open, $K \subset \Omega$ compact with $\mathrm{cap}_\Omega(K) = 0$. Then the Dirichlet principle cannot give a solution of the problem*

$$\begin{aligned} f &: \overline{\Omega\backslash K} \to \mathbb{R} \\ \Delta f(x) &= 0 \textit{ for } x \in \Omega\backslash K \\ f(x) &= 0 \textit{ for } x \in \partial\Omega \\ f(x) &= 1 \textit{ for } x \in \partial K. \end{aligned}$$

□

For an arbitrary $A \subset \Omega$ one can also define

$$\mathrm{cap}_\Omega(A) := \sup_{\substack{K \subset A \\ K \text{ compact}}} \mathrm{cap}\,(K)$$

(as for an A with e.g. $\mathrm{vol}\,(A) = \infty$ there is no $f \in H_0^{1,2}(\Omega)$ with $f \geq 1$ on A, we cannot define the capacity directly as in definition 22.8).

We shall now derive the so-called Euler-Lagrange differential equations as necessary conditions for the existence of a minimum of a variational problem.

Theorem 22.10 *Consider $H : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, with H measurable in the first and differentiable in the other two arguments. We set*

$$I(f) := \int_\Omega H(x, f(x), Df(x))dx$$

for $f \in H^{1,2}(\Omega)$. Assume that

$$|H(x, f, p)| \leq c_1|p|^2 + c_2|f|^2 + c_3 \tag{16}$$

with constants c_1, c_2, c_3 for almost all $x \in \Omega$ and all $f \in \mathbb{R}, p \in \mathbb{R}^d$. ($I(f)$ is therefore finite for all $f \in H^{1,2}(\Omega)$).

(i) *Let $A \subset H^{1,2}(\Omega)$ and let $f \in A$ satisfy*

$$I(f) = \inf\{I(g) : g \in A\}.$$

Let A be such that for every $\varphi \in C_0^\infty(\Omega)$ there is a $t_0 > 0$ with

$$f + t\varphi \in A \quad \textit{for all } t \textit{ with } |t| < t_0. \tag{17}$$

Assume that H satisfies for almost all x and all f, p

$$|H_f(x,f,p)| + \sum_{i=1}^{d} |H_{p_i}(x,f,p)| \leq c_4|p|^2 + c_5|f|^2 + c_6 \tag{18}$$

with constants c_4, c_5, c_6; here, the subscripts denote partial derivatives and $p = (p_1, \ldots p_d)$. Then for all $\varphi \in C_0^\infty(\Omega)$ we have

$$\int_\Omega \{H_f(x,f(x),Df(x))\varphi(x) + \sum_{i=1}^{d} H_{p_i}(x,f(x),Df(x))D_i\varphi(x)\}dx = 0 \tag{19}$$

(ii) *Under the same assumptions as in (i) assume that even for any $\varphi \in H_0^{1,2}(\Omega)$ there is a t_0 such that (17) holds. Furthermore, assume instead of (18) the inequality*

$$|H_f(x,f,p)| + \sum_{i=1}^{d} |H_{p_i}(x,f,p)| \leq c_7|p| + c_8|f| + c_9, \tag{20}$$

with constants c_7, c_8, c_9. Then the condition (19) holds for all $\varphi \in H_0^{1,2}(\Omega)$.

(iii) *Under the same assumptions as in (i), let now H be continuously differentiable in all the variables. Then, if f is also twice continuously differentiable, we have*

$$\begin{aligned}\sum_{i,j=1}^{d} & H_{p_ip_j}(x,f(x),Df(x)) \cdot \frac{\partial^2 f(x)}{\partial x^i \partial x^j} \\ &+ \sum_{i=1}^{d} H_{p_if}(x,f(x),Df(x))\frac{\partial f(x)}{\partial x^i} + \\ &+ \sum_{i=1}^{d} H_{p_ix^i}(x,f(x),Df(x)) - H_f(x,f(x),Df(x)) = 0\end{aligned} \tag{21}$$

or, abbreviated,

$$\sum_{i=1}^{d} \frac{d}{dx^i}(H_{p_i}(x,f(x),Df(x)) - H_f(x,f(x),Df(x)) = 0 \tag{22}$$

(here $\frac{d}{dx^i}$ is to be distinguished from $\frac{\partial}{\partial x^i}$!).

Definition 22.11 The equations (21) are called the Euler-Lagrange equations of the variational problem $I(f) \to \min$.

The equation (21) was first established by Euler for the case $d = 1$ by means of approximation by difference equations and then by Lagrange in the general case by a method essentially similar to the one used here.

Proof of theorem 22.10

(i) We have

$$I(f) \le I(f + t\varphi) \quad \text{for } |t| < t_0. \tag{23}$$

Now

$$I(f + t\varphi) = \int_\Omega H(x, f(x) + t\varphi(x), Df(x) + tD\varphi(x))dx.$$

As for $\varphi \in C_0^\infty(\Omega)$, φ and $D\varphi$ are bounded we can apply theorem 16.10 on account of (16) and (18) and conclude that $I(f + t\varphi)$ is differentiable in t for $|t| < t_0$ with derivative

$$\frac{d}{dt} I(f + t\varphi) = \int_\Omega \{H_f(x, f(x) + t\varphi(x), Df(x) + tD\varphi(x))\varphi(x)$$
$$+ \sum_{i=1}^d H_{p_i}(x, f(x) + t\varphi(x), Df(x) + tD\varphi(x)) \cdot D_i\varphi(x)\}dx.$$

From (23) it follows that

$$0 = \frac{d}{dt} I(f + t\varphi)_{|t=0} =$$
$$\int_\Omega \{H_f(x, f(x), Df(x))\varphi(x) + \sum_{i=1}^d H_{p_i}(x, f(x), Df(x)) \cdot D_i\varphi(x)\}dx.$$

This proves (i).

If (20) holds, we can differentiate under the integral with respect to t in case $\varphi \in H_0^{1,2}(\Omega)$, for then the integrand of the derivative is bounded by

$$(c_7|Df(x) + tD\varphi(x)| + c_8|f(x) + t\varphi(x)| + c_9)(|\varphi(x)| + |D\varphi(x)|)$$

the integral of which, by the Schwarz inequality, is bounded by

$$\text{const.}\ \|f + t\varphi\|_{H^{1,2}} \cdot \|\varphi\|_{H^{1,2}}.$$

Therefore theorem 16.10 can indeed again be applied to justify differentiation under the integral sign. Thus (ii) follows.

For the proof of (21) we notice that, due to the assumptions of continuous differentiability, there exists for every $x \in \Omega$ a neighborhood $U(x)$ in which $H_{p_ip_j} \frac{\partial^2 f}{\partial x^i \partial x^j}$, $H_{p_if} \frac{\partial f}{\partial x^i}$ and $H_{p_ix^i}$ are bounded.

For $\varphi \in C_0^\infty(U(x))$ we can then integrate (19) by parts and obtain

$$0 = \int_{\Omega} \{H_f(x, f(x), Df(x)) - \sum_{i,j=1}^{d} H_{p_i p_j}(x, f(x), Df(x)) \frac{\partial}{\partial x^i}(\frac{\partial f(x)}{\partial x^j})$$

$$- \sum_{i=1}^{d} H_{p_i f}(x, f(x), Df(x)) \frac{\partial f}{\partial x^i}(x) - \sum_{i=1}^{d} H_{p_i x^i}(x, f(x), Df(x))\}\varphi(x)dx.$$

As this holds for all $\varphi \in C_0^\infty(U(x))$ it follows from corollary 19.19 that the expression in the curly brackets vanishes in $U(x)$, and as this holds for every $x \in \Omega$, the validity of (21) in Ω follows. □

Remark. By the Sobolev embedding theorem, one can substitute the term $c_2|f|^2$ in (16) by $c_2|f|^{\frac{2d}{d-2}}$ for $d > 2$ and by $c_2|f|^q$ with arbitrary $q < \infty$ for $d = 2$, and similarly $c_5|f|^2$ in (18) etc., without harming the validity of the conclusions. (Note, however that the version of the Sobolev embedding theorem proved in the present book is formulated only for $H_0^{1,2}$ and not for $H^{1,2}$, and so is not directly applicable here.)

One can also consider more general variational problems for vector-valued functions: Let

$$H : \Omega \times \mathbb{R}^c \times \mathbb{R}^{dc} \to \mathbb{R}$$

be given, and for $f : \Omega \to \mathbb{R}^c$ consider the problem

$$I(f) := \int_{\Omega} H(x, f(x), Df(x))dx \to \min .$$

In this case, the Euler-Lagrange differential equations are

$$\sum_{i=1}^{d} \frac{d}{dx^i}(H_{p_i^\alpha}(x, f(x), Df(x))) - H_{f^\alpha}(x, f(x), Df(x)) = 0 \text{ for } \alpha = 1, \ldots c$$

or, written out,

$$\sum_{i,j=1}^{d} \sum_{\beta=1}^{c} H_{p_i^\alpha p_j^\beta}(x, f(x), Df(x)) \frac{\partial^2}{\partial x^i \partial x^j} f^\beta$$

$$+ \sum_{\beta=1}^{c} \sum_{i=1}^{d} H_{p_i^\alpha f^\beta}(x, f(x), Df(x)) \frac{\partial f^\beta}{\partial x^i}$$

$$+ \sum_{i=1}^{d} H_{p_i^\alpha x^i}(x, f(x), Df(x)) - H_{f^\alpha}(x, f(x), Df(x)) = 0 \quad \text{for } \alpha = 1, \ldots, c.$$

So this time, we obtain a *system* of partial differential equations.
For the rest of this paragraph, H will always be of class C^2.

Examples. We shall now consider a series of examples:

1) For $a, b \in \mathbb{R}$, $f : [a,b] \to \mathbb{R}$, we want to minimize the arc length of the graph of f, thus of the curve $(x, f(x)) \subset \mathbb{R}^2$, hence

$$\int_a^b \sqrt{1+f'(x)^2}\,dx \to \min .$$

The Euler-Lagrange equations are

$$\begin{aligned} 0 = \frac{d}{dx}\frac{f'(x)}{\sqrt{1+f'(x)^2}} \\ &= \frac{f''(x)}{\sqrt{1+f'(x)^2}} - \frac{f'(x)^2 f''(x)}{[1+f'(x)^2]^{\frac{3}{2}}} \\ &= \frac{f''(x)}{(1+f'(x)^2)^{\frac{3}{2}}}, \end{aligned}$$

so

$$f''(x) = 0. \tag{24}$$

Of course, the solutions of (24) are precisely the straight lines, and we shall see below that these indeed give the minimum for given boundary conditions $f(a) = \alpha, f(b) = \beta$.

2) The so-called Fermat principle says that a light ray traverses its actual path between two points in less time than any other path joining those two points. Thus the path of light in an inhomogeneous two dimensional medium with speed $\gamma(x, f)$ is determined by the variational problem

$$I(f) = \int_a^b \frac{\sqrt{1+f'(x)^2}}{\gamma(x, f(x))}dx \to \min .$$

The Euler-Lagrange equations are

$$\begin{aligned} 0 &= \frac{d}{dx}\frac{f'(x)}{\gamma(x,f(x))\sqrt{1+f'(x)^2}} + \frac{\gamma_f}{\gamma^2}\sqrt{1+f'(x)^2} \\ &= \frac{f''(x)}{\gamma\sqrt{1+f'(x)^2}} - \frac{(f'(x))^2 f''(x)}{\gamma(1+f'(x)^2)^{\frac{3}{2}}} - \frac{\gamma_x}{\gamma^2}\frac{f'(x)}{\sqrt{1+f'(x)^2}} \\ &\quad - \frac{\gamma_f}{\gamma^2}\frac{f'(x)^2}{\sqrt{1+f'(x)^2}} + \frac{\gamma_f}{\gamma^2}\sqrt{1+f'(x)^2}, \end{aligned}$$

so

$$0 = f''(x) - \frac{\gamma_x}{\gamma}f'(x)(1+f'(x)^2) + \frac{\gamma_f}{\gamma}(1+f'(x)^2). \tag{25}$$

Obviously, example 2 is a generalization of example 1.

3) The brachistochrone problem is formally a special case of the preceding example. Here, two points $(x_0, 0)$ and (x_1, y_1) are joined by a curve on which a particle moves, without friction, under the influence of a gravitational field directed along the y-axis, and it is required that the particle moves from one point to the other in the shortest possible time.
Denoting acceleration due to gravity by g, the particle attains the speed $(2gy)^{\frac{1}{2}}$ after falling the height y and the time required to fall by the amount $y = f(x)$ is therefore

$$I(f) = \int_{x_0}^{x_1} \sqrt{\frac{1 + f'(x)^2}{2gf(x)}}\, dx.$$

We consider this as the problem $I(f) \to \min.$ subject to the boundary conditions $f(x_0) = 0, f(x_1) = y_1$. Setting $\gamma = \sqrt{2gf(x)}$, equation (25) becomes

$$0 = f''(x) + (1 + f'(x)^2)\frac{1}{2f(x)}. \tag{26}$$

We shall solve (26) explicitly. Consider the integrand

$$H(f(x), f'(x)) = \sqrt{\frac{1 + f'(x)^2}{2gf(x)}}.$$

From the Euler-Lagrange equations

$$\frac{d}{dx}H_p - H_f = 0$$

it follows, as H does not depend explicitly on x, that

$$\begin{aligned}\frac{d}{dx}(f' \cdot H_p - H) &= f'' \cdot H_p + f'\frac{d}{dx}H_p - H_p \cdot f'' - H_f \cdot f' \\ &= f'(\frac{d}{dx}H_p - H_f) = 0,\end{aligned}$$

so $f' \cdot H_p - H \equiv$ const. $\equiv c$.

From this, f' can be expressed as a function of f and c, and in case $f' \neq 0$, the inverse function theorem gives, with $f' = \varphi(f, c)$,

$$x = \int \frac{df}{\varphi(f, c)}.$$

In our case

$$c = f' \cdot H_p - H = -\frac{1}{\sqrt{2gf(1 + f'^2)}},$$

so

$$f' = \pm\sqrt{\frac{1}{2gc^2 f} - 1}.$$

We set $2gc^2 f = \frac{1}{2}(1-\cos t)$, so that $f' = \sqrt{\frac{1+\cos t}{1-\cos t}} = \frac{\sin t}{1-\cos t}$, and then

$$\begin{aligned} x &= \int \frac{df}{f'} = \int \frac{1-\cos t}{\sin t}\frac{df}{dt}dt \\ &= \frac{1}{4gc^2}\int (1-\cos t)dt = \frac{1}{4gc^2}(t-\sin t) + c_1. \end{aligned} \tag{27}$$

Thereby f and x have been determined as functions of t. If one solves (27) for $t = t(x)$ and puts this in the equation for f, then one also obtains $f(x)$.

In the preceding example, we have learnt an important method for solving ordinary differential equations, namely, that of finding an expression which by the differential equation, must be constant as a function of the independent variable. From the constancy of this expression x and $f(x)$ can then be obtained as a function of a parameter. One can proceed similarly in the case where h does not contain the dependent variable f; then the Euler-Lagrange equation is simply

$$\frac{d}{dx}H_p = 0$$

and therefore $H_p =$ const., and from this one can again obtain f' and then x and $f(x)$ by integration.

All the above examples were concerned with the simplest possible situation, namely the case where only one independent and one dependent variable occured. If one considers, for example, in 1) an arbitrary curve $g(x) = (g_1(x), \dots, g_c(x))$ in $\mathbb{R}^c$, then we have to minimize

$$I(g) = \int_a^b \|g'(x)\|dx = \int_a^b \left(\sum_{i=1}^c \left(\frac{d}{dx}g_i(x)\right)^2\right)^{\frac{1}{2}} dx,$$

and we obtain as the Euler-Lagrange equations

$$0 = \frac{d}{dx}\frac{g_i'(x)}{\left(\sum\limits_{j=1}^c g_j'(x)^2\right)^{\frac{1}{2}}} = \frac{g_i'' \sum\limits_{j=1}^c (g_j')^2 - g_i' \cdot \sum\limits_{j=1}^c g_j' g_j''}{\left(\sum\limits_{j=1}^c (g_j')^2\right)^{\frac{3}{2}}} \quad \text{for } i = 1, \dots, c. \tag{28}$$

From this, one can at first not see too much, and this is not surprising as we had already seen earlier that the length $I(g)$ of the curve $g(x)$ is invariant under reparametrizations. Thus, if $x \mapsto g(x)$ is a solution of (28) then so is $\gamma(t) := g(x(t))$ for every bijective map $t \mapsto x(t)$. In other words, there are just too many solutions. On the other hand, we know that for a smooth curve g we can always arrange $\|\frac{d}{dt}g(x(t))\| \equiv 1$ by a reparametrization $x = x(t)$. The equations (28) then become

$$\frac{d}{dt}(\frac{d}{dt}g_i(x(t))) = 0 \quad \text{for } i = 1, \dots, c,$$

and it follows that $g(x(t))$ is a straight line. Then $g(x)$ is also a straight line, only here $g(x)$ does not necessarily describe the arc length.

In physics, stable equilibria are characterized by the principle of minimal potential energy, whereas dynamical processes are described by Hamilton's principle. In both, it is a question of variational principles. Let a physical system with d degrees of freedom be given; let the parameters be $q^1, \dots, q^d$. We want to determine the state of the system by expressing the parameters as functions of the time t. The mechanical properties of the system may be described by:

— the kinetic energy $T = \sum_{i,j=1}^{d} A_{ij}(q^1, \dots q^d, t)\dot{q}^i\dot{q}^j$

(thus T is a function of the velocities $\dot{q}^1, \dots, \dot{q}^d$ – a point " · " always denotes derivative with respect to time –, the coordinates $q^1, \dots, q^d$, and time t; often, T does not depend anymore explicitly on t (see below): Here, T is a quadratic form in the generalized velocities $q^1, \dots, q^d$)

— and the potential energy $U = U(q^1, \dots q^d, t)$.

Both U and T are assumed to be of class C^2.

Hamilton's principle now postulates that motion between two points in time t_0 and t_1 occurs in such a way that the integral

$$I(q) := \int_{t_0}^{t_1} (T - U)dt \tag{29}$$

is stationary in the class of all functions $q(t) = (q^1(t), \dots, q^d(t))$ with fixed initial and final states $q(t_0)$ and $q(t_1)$ respectively .

Thus one does not necessarily look for a minimum under all motions which carry the system from an initial state to a final state, rather only for a stationary value of the integral. For a stationary value, the Euler-Lagrange equations must hold exactly as for a minimum, thus

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}^i} - \frac{\partial}{\partial q^i}(T - U) = 0 \quad \text{for } i = 1, \dots, d. \tag{30}$$

If U and T do not depend explicitly on time t, then equilibrium states are characterized by all the quantities being moreover constant in time, so in particular $\dot{q}^i = 0$ for $i = 1, \dots, d$, and thereby $T = 0$, therefore by (30)

$$\frac{\partial U}{\partial q^i} = 0 \quad \text{for } i = 1, \dots, d. \tag{31}$$

Thus in a state of equilibrium, U must have a critical point and in order for this equilibrium to be stable U must even have a minimum there.

We shall now derive the theorem of conservation of energy in the case where T and U do not depend explicitly on time (though they depend implicitly as they depend on q^i, $\dot{q}^i$ which in turn depend on t).

By observing that

$$\sum_{i,j=1}^{d} A_{ij}\dot{q}^i\dot{q}^j = \frac{1}{2}\sum_{i,j=1}^{d}(A_{ij}+A_{ji})\dot{q}^i\dot{q}^j$$

and, if necessary, replacing A_{ij} by $\frac{1}{2}(A_{ij}+A_{ji})$, we may assume that

$$A_{ij} = A_{ji}.$$

Now

$$T = \sum_{i,j=1}^{d} A_{ij}(q^1,\dots,q^d)\dot{q}^i\dot{q}^j$$

$$U = U(q^1,\dots,q^d).$$

Introducing the Lagrangian

$$L = T - U,$$

the Euler-Lagrange equations become

$$0 = \frac{d}{dt}L_{\dot{q}^i} - L_{q^i} \quad (i = 1,\dots,d).$$

As above, one calculates that

$$\frac{d}{dt}\left(\sum_{i=1}^{d}\dot{q}^i L_{\dot{q}^i} - L\right) = \sum_{i=1}^{d}\left(\ddot{q}^i L_{\dot{q}^i} + \dot{q}^i\frac{d}{dt}L_{\dot{q}^i} - L_{\dot{q}^i}\ddot{q}^i - L_{q^i}\dot{q}^i\right) = 0,$$

so

$$\sum_{i=1}^{d}\dot{q}^i L_{\dot{q}^i} - L = \text{ const. (independent of } t).$$

On the other hand

$$\sum_{i=1}^{d}\dot{q}^i L_{\dot{q}^i} = \sum_{i=1}^{d} 2\dot{q}^i\sum_{k=1}^{d} A_{ik}\dot{q}^k = 2T,$$

and it follows that

$$2T - L = T + U$$

is constant in t. $T + U$ is called the total energy of the system and we have therefore shown the time conservation of energy, in case T and U do not depend explicitly on t.

A special case is the motion of a point of mass m in three dimensional space; let its path be $q(t) = (q^1(t), q^2(t), q^3(t))$. In this case

$$T = \frac{m}{2} \sum_{i=1}^{3} \dot{q}^i(t)^2$$

and U is determined by Newton's law of gravitation, for example,

$$U = -m \frac{g}{\|q\|}$$

in case an attracting mass is situated at the origin of coordinates ($g =$ const.)

We shall now consider motion in the neighborhood of a stable equilibrium. Here we will again assume that T and U do not depend explicitly on time t. Without loss of generality, assume that the equilibrium point is at $t = 0$ and also that $U(0) = 0$ holds. As motion occurs in a neighborhood of a stationary state, we ignore terms of order higher than two in the $\dot{q}^i$ and q^i; thus, we set

$$\begin{aligned} T &= \sum_{i,j=1}^{d} a_{ij} \dot{q}^i \dot{q}^j \\ U &= \sum_{i,j=1}^{d} b_{ij} q^i q^j \end{aligned} \tag{32}$$

with constant coefficients a_{ij}, b_{ij}. We have therefore substituted U by the second order terms of its Taylor series (the first order terms vanish because of (31)). In particular, we can assume $b_{ij} = b_{ji}$. By writing

$$T = \sum_{i,j=1}^{d} \frac{1}{2}(a_{ij} + a_{ji}) \dot{q}^i \dot{q}^j,$$

we can likewise assume that the coefficients of T are symmetric. As U is to have a minimum at 0, we shall also assume that the matrix

$$B = (b_{ij})_{i,j=1,\dots,d}$$

is positive definite.

Finally, we also assume that

$$A = (a_{ij})_{i,j=1,\dots,d}$$

is positive definite.

Equation (30) transforms to

$$\sum_{j=1}^{d} a_{ij} \ddot{q}^j + \sum_{j=1}^{d} b_{ij} q^j = 0 \quad \text{for } i = 1, \dots, d, \tag{33}$$

so in vector notation to

$$\ddot{q} + Cq = 0 \tag{34}$$

with the positive definite symmetric matrix $C = A^{-1}B$. As C is symmetric, it can be transformed to a diagonal matrix by an orthogonal matrix, hence

$$S^{-1}CS =: D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{pmatrix}$$

for an orthogonal matrix S. As C is positive definite, all the eigenvalues $\lambda_1, \ldots \lambda_d$ are positive. We set $y = S^{-1}q$, and (34) then becomes

$$\ddot{y} + Dy = 0,$$

thus

$$\ddot{y}^i + \lambda_i y^i = 0 \quad \text{for } i = 1, 2, \ldots, d. \tag{35}$$

The general solution of (35) is

$$y^i(t) = \alpha_i \cos(\sqrt{\lambda_i} t) + \beta_i \sin(\sqrt{\lambda_i} t)$$

with arbitrary real constants $\alpha_i, \beta_i (i = 1, \ldots, d)$.

We now come to the simplest problems of continuum mechanics. States of equilibrium and motion can be characterized formally as before, however the state of a system can no longer be determined by finitely many coordinates. Instead of $q^1(t), \ldots, q^d(t)$ we now must determine a (real or vector-valued) function $f(x,t)$ or $f(x)$ describing states of motion or rest, respectively.

First we consider the simplest example of a homogeneous vibrating string. The string is under a constant tension μ and executes small vibrations about a stable state of equilibrium. This state corresponds to the segment $0 \leq x \leq \ell$ of the x-axis and the stretching perpendicular to the x-axis is described by the function $f(x,t)$. The string is fixed at the end points and therefore $f(0,t) = 0 = f(\ell, t)$ for all t.

Now the kinetic energy is

$$T = \frac{\rho}{2} \int_0^\ell f_t^2 dx \quad (\rho \text{ means density of the string}), \tag{36}$$

and the potential energy is

$$U = \mu\{ \int_0^\ell \sqrt{1 + f_x^2} dx - \ell \},$$

thus proportional to the increase in length relative to the state of rest. We shall consider a small stretching from the equilibrium position and therefore ignore terms of higher order and set, as before,

$$U = \frac{\mu}{2} \int_0^\ell f_x^2 dx. \tag{37}$$

By Hamilton's principle, the motion is characterized by

$$I(f) = \int_{t_0}^{t_1} (T - U)dt = \frac{1}{2} \int_{t_0}^{t_1} \int_0^\ell (\rho f_t^2 - \mu f_x^2) dx dt \tag{38}$$

being stationary in the class of all functions with $f(0, t) = f(\ell, t) = 0$ for all t.

The Euler-Lagrange equation is now

$$\rho f_{tt} - \mu f_{xx} = 0. \tag{39}$$

This is the so-called *wave equation*. For simplicity we shall take $\rho = \mu = 1$. The weak form of the Euler-Lagrange equation is then

$$\int_{t_0}^{t_1} \int_0^\ell (f_t \varphi_t - f_x \varphi_x) dx dt = 0 \text{ for all } \varphi \in C_0^\infty((0, \ell) \times (t_0, t_1)) \tag{40}$$

(we have not required any boundary conditions for $t = t_0$ and $t = t_1$ and therefore this holds even for functions φ which do not necessarily vanish at $t = t_0$ and $t = t_1$, but this we do not want to investigate here in detail).

Now let $\gamma \in C^1(\mathbb{R})$. Then the function g defined by

$$g(x, t) := \gamma(x - t)$$

is in $C^1([0, \ell] \times [t_0, t_1])$ and satisfies

$$g_x = -g_t.$$

Therefore, for all $\varphi \in C_0^\infty((0, \ell) \times (t_0, t_1))$ we have

$$\int_{t_0}^{t_1} \int_0^\ell (g_t \varphi_t - g_x \varphi_x) dx dt = \int_{t_0}^{t_1} \int_0^\ell (-g_x \varphi_t + g_t \varphi_x) dx dt$$

$$= \int_{t_0}^{t_1} \int_0^\ell g(\varphi_{tx} - \varphi_{xt}) dx dt = 0.$$

Thus g is a solution of (40) although g is not necessarily twice differentiable and therefore not necessarily a classical solution of the Euler-Lagrange equation

$$f_{tt} - f_{xx} = 0. \tag{41}$$

Hence a weak solution of the Euler-Lagrange equation need not necessarily be a classical solution.

In this example, the integrand

$$H(p) = p_2^2 - p_1^2 \quad (p_1 \text{ stands for } \frac{\partial f}{\partial x}, p_2 \text{ for } \frac{\partial f}{\partial t})$$

is analytic indeed, but has an indefinite Hessian $(H_{p_i p_j})$, namely

$$\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Moreover, the fact behind this example is that

$$g(x,t) = \gamma(x-t) + \delta(x+t)$$

is the general solution of the wave equation

$$g_{tt} - g_{xx} = 0.$$

If the string is subjected to an additional external force $k(x,t)$ then the potential energy becomes

$$U = \frac{\mu}{2}\int_0^\ell f_x^2 dx + \int_0^\ell k(x,t) f(x,t) dx,$$

and the equation of motion becomes

$$\rho f_{tt} - \mu f_{xx} + k = 0. \tag{42}$$

Correspondingly, an equilibrium state (assuming that k depends no longer on t) is given by

$$\mu f_{xx}(x) - k(x) = 0. \tag{43}$$

The situation looks similar for a plane membrane – i.e. an elastic surface that at rest covers a portion Ω of the xy-plane and can move vertically. The potential energy is proportional to the difference of the surface area to the surface area at rest. We set the factor of proportionality as well as the subsequent physical constants equal to 1. If $f(x,y,t)$ denotes the vertical stretching of the surface then

$$U = \int_\Omega \sqrt{1 + f_x^2 + f_y^2}\, dx dy - \operatorname{Vol}(\Omega). \tag{44}$$

We shall again restrict ourselves to small pertubations and therefore substitute U as before by

$$U = \frac{1}{2}\int_\Omega (f_x^2 + f_y^2) dx dy. \tag{45}$$

The kinetic energy is

$$T = \frac{1}{2} \int\limits_{\Omega} f_t^2 dxdy. \tag{46}$$

The equation of motion is then

$$f_{tt} - \Delta f = 0 \quad (\Delta f = f_{xx} + f_{yy}) \tag{47}$$

and its state of rest is characterized by

$$\Delta f = 0. \tag{48}$$

We had already derived this earlier. Under the influence of an external force $k(x)$, its state of rest is correspondingly given by

$$\Delta f(x,y) = k(x,y). \tag{49}$$

Thus, if the membrane is fixed at the boundary, we have to solve the Dirichlet problem

$$\begin{aligned} \Delta f(x,y) &= k(x,y) \quad \text{for } (x,y) \in \Omega \\ f(x,y) &= 0 \quad \text{for } (x,y) \in \partial\Omega. \end{aligned}$$

We shall now derive the Euler-Lagrange equations for the area functional

$$I(f) = \int\limits_{\Omega} \sqrt{1 + f_x^2 + f_y^2} dxdy.$$

Setting $H(p_1, p_2) = \sqrt{1 + p_1^2 + p_2^2}$ we have

$$H_{p_i} = \frac{p_i}{\sqrt{1 + p_1^2 + p_2^2}}$$

and

$$H_{p_i p_j} = \frac{\delta_{ij}}{\sqrt{1 + p_1^2 + p_2^2}} - \frac{p_i p_j}{(1 + p_1^2 + p_2^2)^{\frac{3}{2}}} \quad \left(\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \right).$$

Thereby, the Euler-Lagrange equations become

$$\begin{aligned} 0 = \sum_{i,j=1}^{2} H_{p_i p_j} f_{x^i x^j} = \frac{1}{(1 + f_x^2 + f_y^2)^{\frac{3}{2}}} \{ &(1 + f_y^2) f_{xx} \\ &- 2 f_x f_y f_{xy} + (1 + f_x^2) f_{yy} \}, \end{aligned}$$

so

$$(1 + f_y^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + f_x^2) f_{yy} = 0.$$

This is the so-called *minimal surface equation.* It describes surfaces with stationary area that can be represented as graphs over a domain Ω in the (x,y)-plane.

Finally, we consider quadratic integrals of the form

$$Q(f) = \int_\Omega \Big\{ \sum_{i,j=1}^d a_{ij}(x) f_{x^i} f_{x^j} + \sum_{i=1}^d 2b_i(x) f \cdot f_{x^i} + c(x) f(x)^2 \Big\} dx; \tag{50}$$

again, without loss of generality, let $a_{ij} = a_{ji}$. The Euler-Lagrange equations are now

$$-\sum_{i=1}^d \frac{\partial}{\partial x^i} \Big(\sum_{j=1}^d a_{ij}(x) \frac{\partial f}{\partial x^j} + b_i(x) f \Big) + \sum_{i=1}^d b_i(x) \frac{\partial f}{\partial x^i} + c(x) f = 0. \tag{51}$$

The Euler-Lagrange equations for a quadratic variational problem are therefore linear in f and its derivatives.

We shall now study the behaviour of the Euler-Lagrange equations under transformations of the independent variables.

So let $\xi \mapsto x(\xi)$ be a diffeomorphism of Ω' onto Ω; we set $D_x f = (\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^d})$, $D_\xi x = (\frac{\partial x^i}{\partial \xi^j})_{i,j=1,\ldots,d}$ etc.

$$H(x, f, D_x f) = H(x(\xi), f, D_\xi f \cdot (D_\xi x)^{-1}) =: \Phi(\xi, f, D_\xi f).$$

By the change of variables in integrals we have

$$\int_\Omega H(x, f, D_x f) dx = \int_{\Omega'} \Phi(\xi, f, D_\xi f) |\det(D_\xi x)| d\xi. \tag{52}$$

We now write for the sake of abbreviation

$$[H]_f = -\Big(\sum_{i=1}^d \frac{d}{dx^i} H_{p^i} - H_f \Big). \tag{53}$$

We then have for $\varphi \in C_0^\infty(\Omega)$, on account of the derivation of Euler-Lagrange equations,

$$\begin{aligned} \int_\Omega [H]_f \varphi dx &= \frac{d}{dt} \int_\Omega H(x, f + t\varphi, D_x f + t D_x \varphi) dx_{|t=0} \\ &= \frac{d}{dt} \int_{\Omega'} \Phi(\xi, f + t\varphi, D_\xi f + t D_\xi \varphi) |\det(D_\xi x)| d\xi_{|t=0} \\ &= \int_{\Omega'} [\Phi |\det(D_\xi x)|]_f \varphi d\xi \\ &= \int_\Omega [\Phi |\det(D_\xi x)|]_f \varphi |\det(D_x \xi)| dx. \end{aligned}$$

As this holds for all $\varphi \in C_0^\infty(\Omega)$, it follows, as usual, from corollary 19.20 that

$$[H]_f = [\Phi |\det(D_\xi x)|]_f |\det(D_x \xi)|. \tag{54}$$

(Under the assumption $H \in C^2$, we consider

$$I(f) = \int\limits_{\Omega} H(x, f(x), Df(x))dx$$

as a function

$$I : C^2(\Omega) \to \mathbb{R}$$

and $[H]_f$ is then the gradient of I, as the derivative of I is given by

$$\varphi \mapsto DI(\varphi) = \int\limits_{\Omega} [H]_f \varphi dx.$$

Thus equation (54) expresses that the behaviour under transformations of this gradient is quite analogous to that of a gradient in the finite dimensional case.)

We shall use this to study the transformation of the Laplace operator; the advantage of (54) lies precisely in this that one does not have to transform derivatives of second order. Now the Laplace equation, as we have already seen at the beginning, is precisely the Euler-Lagrange equation for the Dirichlet integral.

So let $\xi \mapsto x(\xi)$ be again a diffeomorphism of Ω' onto Ω; we set

$$g_{ij} := \sum_{k=1}^{d} \frac{\partial x^k}{\partial \xi^i} \frac{\partial x^k}{\partial \xi^j}$$

and

$$g^{ij} := \sum_{k=1}^{d} \frac{\partial \xi^i}{\partial x^k} \frac{\partial \xi^j}{\partial x^k}.$$

Thus

$$\sum_{i=1}^{d} g_{ik} g^{i\ell} = \delta_{k\ell} \left(= \begin{cases} 1 & \text{for } k = \ell \\ 0 & \text{for } k \neq \ell \end{cases} \right).$$

Furthermore, let

$$g := \det(g_{ij}).$$

Now

$$\sum_{i=1}^{d} \left(\frac{\partial f}{\partial x^i} \right)^2 = \sum_{i=1}^{d} \sum_{j,k=1}^{d} \frac{\partial f}{\partial \xi^j} \frac{\partial \xi^j}{\partial x^i} \frac{\partial f}{\partial \xi^k} \frac{\partial \xi^k}{\partial x^i} = \sum_{j,k=1}^{d} g^{jk} \frac{\partial f}{\partial \xi^j} \frac{\partial f}{\partial \xi^k}.$$

Formula (54) now gives directly, together with (50) and (51),

$$\Delta f(x) = \frac{1}{\sqrt{g}} \sum_{j=1}^{d} \frac{\partial}{\partial \xi^j} \left(\sqrt{g} \sum_{k=1}^{d} g^{jk} \frac{\partial f}{\partial \xi^k} \right). \tag{55}$$

This is the desired transformation formula for the Laplace operator.

For plane polar coordinates

$$x = r\cos\varphi, y = r\sin\varphi$$

one calculates from this

$$\Delta f(x,y) = \frac{1}{r}\left(\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial\varphi}\left(\frac{1}{r}\frac{\partial f}{\partial\varphi}\right)\right), \tag{56}$$

and for spatial polar coordinates

$$x = r\cos\varphi\sin\theta, y = r\sin\varphi\sin\theta, z = r\cos\theta$$

$$\Delta f(x,y,z) = \frac{1}{r^2\sin\theta}\left(\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}\frac{\partial f}{\partial\varphi}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right)\right) \tag{57}$$

(cf. §18 for the discussion of polar coordinates).

Exercises for §22

1) Let $\Omega \subset \mathbb{R}^d$ be open and bounded. For $f \in H^{2,2}(\Omega)$, put

$$E(f) := \int_\Omega |D^2 f(x)|^2 dx.$$

(Here, $D^2 f$ is the matrix of weak second derivatives $D_i D_j f$, $i,j = 1,\ldots,d$, and

$$|D^2 f(x)|^2 = \sum_{i,j=1}^{d} |D_i D_j f(x)|^2.)$$

Discuss the following variational problem: For given $g \in H^{2,2}(\Omega)$, minimize $E(f)$ in the class

$$A_g := \{f \in H^{2,2}(\Omega) : f - g \in H_0^{1,2}(\Omega), D_i f - D_i g \in H_0^{1,2}(\Omega), i = 1,\ldots,d\}.$$

2) Let $H : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be nonnegative, measurable w.r.t. the first variable, and convex w.r.t. the second and third variables jointly, i.e. for all $f, g \in \mathbb{R}, p, q \in \mathbb{R}^d, 0 \le t \le 1, x \in \Omega$, we have

$$H(x, tf + (1-t)g, tp + (1-t)q) \le tH(x,f,p) + (1-t)H(x,g,q).$$

For $f \in H^{1,2}(\Omega)$, we put

$$I(f) := \int_\Omega H(x, f(x), Df(x))dx.$$

Show that I is lower semicontinuous w.r.t. weak $H^{1,2}$ convergence.

3)

a) Let A be a $(d \times d)$ matrix with $\det(A) \neq 0$. Consider the coordinate transformation

$$\xi \mapsto x = A\xi.$$

How does the Laplacian $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{(\partial x^i)^2}$ transform under this coordinate transformation?

b) Discuss the coordinate transformation $(\xi, \eta) \mapsto (x, y)$ with

$$x = \sin\xi \cosh\eta$$
$$y = \cos\xi \sinh\eta$$

(planar elliptic coordinates) and express the Laplacian in these coordinates.

4) Determine all rotationally symmetric harmonic functions $f\colon \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$.

5) For $m \in \mathbb{N}$, define the Legendre polynomial as

$$P_m(t) := \frac{1}{2^m m!}\left(\frac{d}{dt}\right)^m (t^2 - 1)^m.$$

Show that

$$f(r, \theta) := r^m P_m(\cos\theta)$$

satisfies $\Delta f = 0$ (in spatial polar coordinates).

6) Let $a, b \in \mathbb{R}, g_1, g_2 > 0$. For functions $f : [a, b] \to \mathbb{R}$ with $f(a) = g_1, f(b) = g_2$, we consider

$$K(f) := 2\pi \int_a^b f(x)\sqrt{1 + f'(x)^2}dx \to \min.$$

($I(f)$ yields the area of the surface obtained by revolving the graph of f about the x-axis. Thus, we are seeking a surface of revolution with smallest area with two circles given as boundary.) Solve the corresponding Euler-Lagrange equations!

7) We define a plate to be a thin elastic body with a planar rest position. We wish to study small transversal vibrations of such a body, induced by an exterior force K. Let us first consider the equilibrium position. Let $f(x, y)$ be the vertical displacement. The potential energy of a deformation is

$$U = U_1 + U_2,$$

where

$$U_1 = \int_\Omega \left(\left(\frac{1}{2}\Delta f(x,y)\right)^2 + \mu\left(f_{xx}f_{yy} - f_{xy}^2\right)\right) dxdy$$

(here, $\Omega \subset \mathbb{R}^2$ is the rest position, $\mu =$ const.),

$$U_2 = \int_\Omega K(x,y)f(x,y)dxdy.$$

Derive the Euler-Lagrange equations

$$\Delta(\Delta f) + K = 0.$$

For the motion, $f(x,y,t)$ is the vertical displacement, and the kinetic energy is

$$T = \frac{1}{2}\int_\Omega f_t^2 dxdy.$$

Derive the differential equation that describes the motion of the plate.