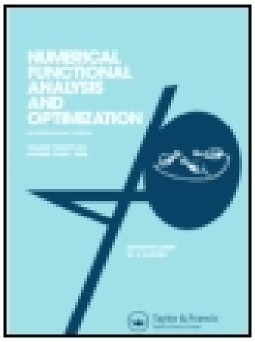

Taylor & Francis
Taylor & Francis Group

Numerical Functional Analysis and Optimization

ISSN: 0163-0563 (Print) 1532-2467 (Online) Journal homepage: http://www.tandfonline.com/loi/lnfa20

New Direct Numerical Methods for Some Multidimensional Problems of the Calculus of Variations

Maksim V. Dolgopolik

To cite this article: Maksim V. Dolgopolik (2017): New Direct Numerical Methods for Some Multidimensional Problems of the Calculus of Variations, Numerical Functional Analysis and Optimization, DOI: 10.1080/01630563.2017.1378231

To link to this article: http://dx.doi.org/10.1080/01630563.2017.1378231

Accepted author version posted online: 12 Sep 2017.
Published online: 12 Sep 2017.

NUMERICAL FUNCTIONAL ANALYSIS AND OPTIMIZATION
https://doi.org/10.1080/01630563.2017.1378231

New Direct Numerical Methods for Some Multidimensional Problems of the Calculus of Variations

Maksim V. Dolgopolik

Department of Applied Mathematics and Control Processes, Saint Petersburg State University, Saint-Petersburg, Russia

ABSTRACT

In this paper, we develop a new approach to the design of direct numerical methods for multidimensional problems of the calculus of variations. The approach is based on a transformation of the problem with the use of a new class of Sobolev-like spaces that is studied in the article. This transformation allows one to analytically compute the direction of steepest descent of the main functional of the calculus of variations with respect to a certain inner product, and, in turn, to construct new direct numerical methods for multidimensional problems of the calculus of variations. In the end of the paper, we point out how the approach developed in the article can be extended to the case of problems with more general boundary conditions, problems for functionals depending on higher order derivatives, and problems with isoperimetric and/or pointwise constraints.

ARTICLE HISTORY
Received 9 December 2014
Revised 20 August 2017
Accepted 6 September 2017

KEYWORDS
Calculus of variations; direct numerical method; direction of steepest descent; method of steepest descent; method of gradient descent

MATHEMATICS SUBJECT CLASSIFICATION
49M05; 49M30

1. Introduction

The main problem of the calculus of variations has the form

$$\min \mathcal{I}(u) = \int_\Omega f(x, u(x), \nabla u(x))\, dx \quad \text{subject to} \quad u|_{\partial\Omega} = \psi, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is an open set, $\partial\Omega$ is the boundary of Ω, and $f \colon \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ and $\psi \colon \partial\Omega \to \mathbb{R}$ are given functions. Various aspects of this problem, such as the existence and regularity of solutions [1–5], qualitative properties of critical points (so-called "the calculus of variations in the large") [6, 7], and necessary and sufficient conditions for a local minimum [8], have been extensively studied by many researches. However, relatively little attention has been paid to the development of direct numerical methods (especially in the multidimensional case) for problems of the calculus of variations.

By direct numerical methods, we mean methods that are not based on direct numerical solution of the Euler–Lagrange equation, and, instead, rely on the variational formulation of the problem. Besides being one of the approaches to numerical solution of some partial differential equations, direct numerical

CONTACT Maksim V. Dolgopolik maxim.dolgopolik@gmail.com Department of Applied Mathematics and Control Processes, Saint Petersburg State University, Universitetskii prospekt 35, Petergof, Saint-Petersburg 198504, Russia.

methods are especially useful and important for those problems of the calculus of variations that arise directly as optimization problems. This type of problems naturally appear, in particular, in image processing [9, 10].

The vast majority of direct numerical methods of the calculus of variations is based on an approximate reduction of problem (1.1) to a finite-dimensional optimization problem. Various types of reduction techniques and corresponding numerical methods in the one-dimensional case (i.e., in the case when $\Omega = (a, b) \subset \mathbb{R}$) were proposed in [11–20]. In the multidimensional case, the range of choice of direct methods is much more narrow, and it includes (although is not exhausted by) the finite elements methods, the Galerkin method, and the Ritz method [21–26]. However, there exist some direct numerical methods of the calculus of variations in the one-dimensional case that do not consist of an approximate reduction to a finite-dimensional problem. Among them are the first- and second-variation methods [27], that are based on straightforward usage of the necessary optimality conditions for problem (1.1), He's variational iteration method [28], the continuous method of steepest descent [29], the discrete steepest descent method [30, 31], Newton's method [32], and the method of hypodifferential descent based on the use of exact penalty functions [33–35]. Let us also mention that some multidimensional problems of the calculus of variations can be solved by standard gradient-based methods for functionals defined on Hilbert or Banach spaces [29, 31, 36–41] with the use of the so-called Sobolev gradients [42].

The main goal of this paper is to develop a new approach to the design of direct numerical methods for multidimensional problems of the calculus of variations. This approach is based on a transformation of problem (1.1) that allows one to analytically compute the direction of steepest descent of the functional $\mathcal{I}$ with respect to a certain norm. Utilizing this direction of steepest descent, one can apply an obvious modification of almost any first-order method of finite-dimensional optimization to problem (1.1). The basic ideas (in a very crude form) that lead to the the development of the approach studied in this article were presented in the two-dimensional case in [43]. It should be noted that the main results of the paper [43] were inspired by the ideas of the late professor Demyanov [44, 45], as well as the method of hypodifferential descent and the method of steepest descent mentioned above.

The paper is organized as follows. In Section 2, we informally discuss the underlying ideas of the approach developed in this article. In Section 3, we introduce and study a new class of Sobolev-like spaces that plays a central role in the formalization of the new direct numerical methods. In particular, we obtain a convenient characterization of a certain function space from this class that is very important for the transformation of problem (1.1). The direction of steepest descent of the functional $\mathcal{I}$ with respect to a certain norm is derived in Section 4. In conclusion, we discuss possible generalizations of the ideas developed in this paper and briefly outline some directions of future research.

2. How to compute the direction of steepest descent?

In this section, we informally discuss a general technique for constructing new minimization methods for the main problem of the calculus of variations. A possible formalization of this technique is presented in the subsequent sections. We suppose that all functions that appear in this section are sufficiently smooth.

Consider the main problem of the calculus of variations

$$\min \mathcal{I}(u) = \int_\Omega f(x, u(x), \nabla u(x))\, dx \quad \text{subject to} \quad u|_{\partial\Omega} = \psi, \tag{2.1}$$

where $\Omega \subset \mathbb{R}^n$ is an open set and $f = f(x, u, z)$. We want to apply infinite-dimensional analogues of standard gradient-based methods of finite-dimensional optimization to the problem above. To do that we need to compute the gradient or, more generally, the direction of steepest descent of the functional $\mathcal{I}$.

It is well known and easy to check that the functional $\mathcal{I}$ is Gâteaux differentiable, and its Gâteaux derivative has the form

$$\mathcal{I}'[u](h) = \int_\Omega \left(\frac{\partial f}{\partial u}(x, u(x), \nabla u(x)) h(x) + \sum_{i=1}^n \frac{\partial f}{\partial z_i}(x, u(x), \nabla u(x)) \frac{\partial h}{\partial x_i}(x) \right) dx. \tag{2.2}$$

Therefore to compute the direction of steepest descent of the functional $\mathcal{I}$, we need to solve the following problem of the calculus of variations

$$\min \mathcal{I}'[u](h) \quad \text{subject to} \quad h|_{\partial\Omega} = 0, \quad \|h\| \le 1,$$

where $\|\cdot\|$ is some norm. However, this problem is, usually, too complicated to be solved analytically, and even in simple cases, it is equivalent to the problem of solving a linear partial differential equation (see, e.g., [42, Chapter 9]).

To overcome this difficulty, let us transform problem (2.1). For the sake of simplicity, suppose that $n = 2$, $\psi \equiv 0$, and let Ω be an open box, i.e., $\Omega = (a_1, b_1) \times (a_2, b_2)$. Let also u be a function such that $u|_{\partial\Omega} = 0$. Observe that

$$u(x_1, x_2) = \int_{a_1}^{x_1} \frac{\partial u}{\partial x_1}(\xi_1, x_2)\, d\xi_1 \quad \forall x \in \Omega. \tag{2.3}$$

Since $u(x_1, a_2) \equiv 0$, then $\partial u / \partial x_1 (x_1, a_2) \equiv 0$. Therefore

$$\frac{\partial u}{\partial x_1}(x_1, x_2) = \int_{a_2}^{x_2} \frac{\partial^2 u}{\partial x_2 \partial x_1}(x_1, \xi_2)\, d\xi_2 \quad \forall x \in \Omega.$$

Hence with the use of (2.3) one gets that

$$u(x_1, x_2) = \int_{a_1}^{x_1} \int_{a_2}^{x_2} \frac{\partial^2 u}{\partial x_2 \partial x_1}(\xi_1, \xi_2)\, d\xi_2 d\xi_1 \quad \forall x \in \Omega.$$

Furthermore, from the latter equality, it follows that

$$\int_{a_1}^{b_1} \frac{\partial^2 u}{\partial x_2 \partial x_1}(\xi_1, \cdot)\, d\xi_1 \equiv 0, \qquad \int_{a_2}^{b_2} \frac{\partial^2 u}{\partial x_2 \partial x_1}(\cdot, \xi_2)\, d\xi_2 \equiv 0$$

due to the fact that $u(x_1, b_2) \equiv 0$ and $u(b_1, x_2) \equiv 0$.

Let now v be a function such that

$$\int_{a_1}^{b_1} v(\xi_1, \cdot)\, d\xi_1 \equiv 0, \qquad \int_{a_2}^{b_2} v(\cdot, \xi_2)\, d\xi_2 \equiv 0.$$

Then it is easy to see that for the function $u = Tv$, where

$$(Tv)(x_1, x_2) = \int_{a_1}^{x_1} \int_{a_2}^{x_2} v(\xi_1, \xi_2)\, d\xi_2 d\xi_1 \quad \forall x \in \Omega,$$

one has $u|_{\partial\Omega} = 0$. Thus, we have that the following result holds true.

Proposition 1. *Let $u \colon [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$ be a sufficiently smooth function. Then $u|_{\partial\Omega} = 0$ if and only if there exists a sufficiently smooth function v such that*
1. $u(x_1, x_2) = (Tv)(x_1, x_2) = \int_{a_1}^{x_1} \int_{a_2}^{x_2} v(\xi_1, \xi_2)\, d\xi_2 d\xi_1$ *for all $x \in \Omega$;*
2. $\int_{a_1}^{b_1} v(\xi_1, \cdot)\, d\xi_1 \equiv 0$ *and* $\int_{a_2}^{b_2} v(\cdot, \xi_2)\, d\xi_2 \equiv 0$.

Moreover, $v = \partial^2 u / \partial x_1 \partial x_2$.

From the proposition above, it follows that problem (2.1) with $\psi \equiv 0$ is equivalent to the following optimization problem: minimize

$$F(v) = \int_\Omega f\Big(x, (Tv)(x), \int_{a_2}^{x_2} v(x_1, \xi_2)\, d\xi_2, \int_{a_1}^{x_1} v(\xi_1, x_2)\, d\xi_1\Big)\, dx_1 dx_2 \qquad (2.4)$$

subject to the constraints

$$\int_{a_1}^{b_1} v(\xi_1, \cdot)\, d\xi_1 \equiv 0, \qquad \int_{a_2}^{b_2} v(\cdot, \xi_2)\, d\xi_2 \equiv 0. \qquad (2.5)$$

As we shall see, one can easily compute the direction of steepest descent for this problem.

Indeed, denote by L_0, the linear space consisting of all functions v satisfying (2.5). Clearly, the functional F is Gâteaux differentiable. Integrating by parts in (2.2) one gets that the Gâteaux derivative of F has the form

$$F'[v](h) = \int_\Omega Q(v)(x) h(x)\, dx,$$

where

$$\begin{aligned} Q(v)(x) = &\int_{x_1}^{b_1} \int_{x_2}^{b_2} \frac{\partial f}{\partial u}(\xi, u(\xi), \nabla u(\xi))\, d\xi_2 d\xi_1 - \\ &- \int_{x_2}^{b_2} \frac{\partial f}{\partial z_1}(x_1, \xi_2, u(x_1, \xi_2), \nabla u(x_1, \xi_2))\, d\xi_2 - \\ &- \int_{x_1}^{b_1} \frac{\partial f}{\partial z_2}(\xi_1, x_2, u(\xi_1, x_2), \nabla u(\xi_1, x_2))\, d\xi_1, \end{aligned} \qquad (2.6)$$

and $u = Tv$. Hence the direction of steepest descent for problem (2.4), (2.5) is a solution of the following optimization problem:

$$\min \int_\Omega Q(v)(x)h(x)\,dx \quad \text{subject to} \quad h \in L_0, \quad \|h\| \leq 1, \tag{2.7}$$

where $\|\cdot\|$ is some norm. We choose the L_2-norm, i.e., $\|h\| = \left(\int_\Omega h^2(x)\,dx\right)^{\frac{1}{2}}$.

Let us solve problem (2.7).

Proposition 2. *Suppose that the function $u = Tv$ does not satisfy the Euler–Lagrange equation for the functional $\mathcal{I}$. Then the direction of steepest descent h^* for problem (2.4), (2.5) has the form $h^*(x) = G(v)(x)/\|G(v)\|_2$, where*

$$G(v)(x) = -Q(v)(x) + \frac{1}{b_1 - a_1}\int_{a_1}^{b_1} Q(v)(\xi_1, x_2)d\xi_1 + \frac{1}{b_2 - a_2}$$
$$\int_{a_2}^{b_2} Q(v)(x_1, \xi_2)d\xi_2 - \frac{1}{(b_1 - a_1)(b_2 - a_2)}$$
$$\int_{a_1}^{b_1}\int_{a_2}^{b_2} Q(v)(\xi_1, \xi_2)d\xi_2 d\xi_1. \tag{2.8}$$

Proof. Applying the Lagrange multipliers rule to problem (2.7), one gets that there exists $\lambda \in \mathbb{R}$ such that

$$\int_\Omega (Q(v)(x) + \lambda h^*(x))h(x)\,dx = 0 \quad \forall h \in L_0.$$

Note that for any infinitely differentiable function φ with compact support one has $\partial^2\varphi/\partial x_1\partial x_2 \in L_0$ (see Proposition 1). Therefore

$$\int_\Omega (Q(v)(x) + \lambda h^*(x))\frac{\partial^2\varphi}{\partial x_1\partial x_2}(x)\,dx = \int_\Omega \frac{\partial^2}{\partial x_1\partial x_2}(Q(v)(x) + \lambda h^*(x))\varphi(x)\,dx$$
$$= 0$$

for any infinitely differentiable function φ with compact support. Hence applying the fundamental lemma of the calculus of variations, one gets that

$$\frac{\partial^2}{\partial x_1\partial x_2}(Q(v)(x) + \lambda h^*(x)) = 0 \quad \forall x \in \Omega.$$

It is easy to verify that $\lambda = 0$ if and only if the function $u = Tv$ satisfies the Euler–Lagrange equation for the functional $\mathcal{I}$ (see Equation (2.6)). Thus, we can suppose that $\lambda \neq 0$. Hence

$$h^*(x) = -\frac{1}{\lambda}Q(v)(x) + r_1(x_1) + r_2(x_2),$$

where r_1 and r_2 are some functions. Taking into account the fact that $h^* \in L_0$ one obtains that

$$\begin{cases} -\dfrac{1}{\lambda}\displaystyle\int_{a_1}^{b_1} Q(v)(\xi_1,\cdot)d\xi_1 + \int_{a_1}^{b_1} r_1(\xi_1)d\xi_1 + (b_1-a_1)r_2(\cdot) = 0, \\ -\dfrac{1}{\lambda}\displaystyle\int_{a_2}^{b_2} Q(v)(\cdot,\xi_2)d\xi_2 + (b_2-a_2)r_1(x_1) + \int_{a_2}^{b_2} r_2(\xi_2)d\xi_2 = 0. \end{cases}$$

Solving this system with respect to r_1 and r_2, one obtains that

$$r_1(x_1) = \frac{1}{\lambda(b_2-a_2)}\int_{a_2}^{b_2} Q(v)(x_1,\xi_2)d\xi_2,$$

$$r_2(x_2) = \frac{1}{\lambda(b_1-a_1)}\int_{a_1}^{b_1} Q(v)(\xi_1,x_2)d\xi_1 - \frac{1}{\lambda(b_1-a_1)(b_2-a_2)}\int_{\Omega} Q(v)(\xi)d\xi.$$

Hence (2.8) holds true. □

Since we know the direction of steepest descent for problem (2.4), (2.5), we can apply an obvious modification of almost any gradient-based algorithm of finite-dimensional optimization to this problem and, in turn, to the initial problem (2.1).

If we look at the way the direction of steepest descent was derived, we can easily see that this direction is the direction of the steepest descent of the functional $\mathcal{I}$ with respect to the norm

$$\|u\| = \left(\int_{\Omega}\left(\frac{\partial^2 u}{\partial x_1 \partial x_2}(x)\right)^2 dx\right)^{\frac{1}{2}}$$

(the fact that this seminorm is, indeed, a norm follows from Proposition 1). However, any space of smooth functions equipped with this norm is incomplete. Therefore, it is natural to consider the original problem in the setting of Sobolev-like spaces and to transfer the main ideas discussed above to this more general setting.

Remark 1. As it is well known, the direction of steepest descent as well as the performance of the method of steepest descent depend on the choice of an underlying Hilbert (Banach) space and an inner product (norm) in this space (see, e.g., [42, 46, 47]). From this point of view, the main goal of this article is to introduce a Hilbert space such that the direction of the steepest descent of the functional $\mathcal{I}(u)$ in this space can be easily computed analytically.

3. Special function spaces

In this section, we introduce a class of function spaces that plays a central role in the formalization of the minimization methods for multidimensional problems of the calculus of variations discussed above. This class of functions is closely related to the Sobolev spaces and possesses many properties of these spaces. We suppose that the reader is familiar with basic results on the Sobolev spaces that can be found in [48–50].

3.1. Main definitions and basic properties

Introduce the notation first. A point in $\mathbb{R}^n$ is denoted by $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, and its norm is denoted by $|x| = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$. As usual, any n-tuple $\alpha = (\alpha_1, \dots, \alpha_n) \in Z_+^n$ of nonnegative integers α_i is called a multi-index; its absolute value $|\alpha| = \alpha_1 + \dots + \alpha_n$. For any multi-index α denote by $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$, a differential operator of order $|\alpha|$, where $D_i = \partial/\partial x_i$ for $i \in \{1, \dots, n\}$. If $\alpha = (0, \dots, 0)$, then $D^\alpha u = u$ for any function u. Define

$$I_k = \left\{\alpha \in Z_+^n \mid |\alpha| = k, \alpha_i = 0 \text{ or } \alpha_i = 1 \ \forall i \in \{1, \dots, n\}\right\}.$$

for any $k \in \{0, \dots, n\}$. It is clear that $I_0 = \{(0, \dots, 0)\}$ and $I_n = \{(1, \dots, 1)\}$. If $\alpha \in I_k$ for some $0 \le k \le n$, then a unique multi-index $\beta \in I_{n-k}$ such that $\alpha + \beta \in I_n$ is denoted by $\overline{\alpha}$. Here the sum of multi-indices is component-wise. The kernel of a linear operator $L \colon X \to Y$ is referred to as $\ker L$ (here X, Y are linear spaces).

Remark 2. We consider only real-valued functions and normed spaces over the field of real numbers. If f is a bounded linear functional defined on a normed space X, then we denote its norm by $\|f\|$ or by $\|f\|_X$ when we want to specify the domain of f.

Hereafter, let $\Omega \subset \mathbb{R}^n$ be a bounded open box, i.e., $\Omega = \prod_{k=1}^n (a_i, b_i)$. Denote by $C^k(\overline{\Omega})$ the set of all those $u \in C^k(\Omega)$ for which all functions $D^\alpha u$ with $0 \le |\alpha| \le k$ are bounded and uniformly continuous on Ω (then there exist unique continuous extensions of functions $D^\alpha u$ with $0 \le |\alpha| \le k$ to the closure $\overline{\Omega}$ of the set Ω). The set of all infinitely continuously differentiable functions $u \colon \Omega \to \mathbb{R}$ with compact support is denoted by $C_0^\infty(\Omega)$.

Let us introduce a new function space. For any $m \in \{1, \dots, n\}$ and $1 \le p \le \infty$ denote by $M^{m,p}(\Omega)$, the set of all $u \in L_p(\Omega)$ such that for any $k \in \{1, \dots, m\}$ and $\alpha \in I_k$, there exists the weak derivative $D^\alpha u$ belonging to $L_p(\Omega)$. Thus, $M^{m,p}(\Omega)$ consists of all functions $u \in L_p(\Omega)$ for which there exist all weak mixed derivatives of the order $k = 1 \colon m$ that belong to $L_p(\Omega)$. The set $M^{m,p}(\Omega)$ is a linear space that can be equipped with the norm

$$\|u; M^{m,p}\| = \begin{cases} \left(\sum_{k=0}^{m} \sum_{\alpha \in I_k} (\|D^\alpha u\|_p)^p\right)^{\frac{1}{p}} & \text{for } 1 \le p < \infty, \\ \max\{\|D^\alpha u\|_\infty \mid \alpha \in I_k, 1 \le k \le m\} & \text{for } p = \infty, \end{cases}$$

where $\|\cdot\|_p$ is the standard norm on $L_p(\Omega)$. The closure of $C_0^\infty(\Omega)$ in the normed space $M^{m,p}(\Omega)$ is denoted by $M_0^{m,p}(\Omega)$.

Let, as usual, $W^{m,p}(\Omega)$ with $m \in \mathbb{N}$ and $1 \le p \le \infty$ be the Sobolev space, and $W_0^{m,p}(\Omega)$ be the closure of $C_0^\infty(\Omega)$ in $W^{m,p}(\Omega)$. It is clear that $W^{m,p}(\Omega) \subset M^{m,p}(\Omega)$ and, analogously, $W_0^{m,p}(\Omega) \subset M_0^{m,p}(\Omega)$. Moreover, these embeddings are continuous. On the other hand, $M^{m,p}(\Omega)$ is dense in $W^{1,p}(\Omega)$ for any $m \in \{1, \dots, n\}$, and $M^{1,p}(\Omega) = W^{1,p}(\Omega)$.

Let us describe some properties of the spaces $M^{m,p}(\Omega)$. Arguing in a similar way to the case of the Sobolev spaces (see, e.g., Theorems 3.2 and 3.5 in [48]) one can easily derive the following results.

Theorem 1. *For any $1 \le m \le n$, the space $M^{m,p}(\Omega)$ is complete in the case $1 \le p \le \infty$, is separable in the case $1 \le p < \infty$, and is reflexive and uniformly convex in the case $1 < p < \infty$. Moreover, $M^{m,2}(\Omega)$ is a separable Hilbert space with the inner product*

$$\langle u, v \rangle_m = \sum_{k=0}^{m} \sum_{\alpha \in I_k} \langle D^\alpha u, D^\alpha v \rangle,$$

where $\langle \varphi, \psi \rangle = \int_\Omega \varphi(x)\psi(x)\,dx$ is the inner product in $L_2(\Omega)$.

Remark 3.

(i) An analogous result holds true for $M_0^{m,p}(\Omega)$.

(ii) Note that for any $\varphi \in C^\infty(\Omega) \cap M^{m,\infty}(\Omega)$ and $u \in M^{m,p}(\Omega)$, one has $\varphi u \in M^{m,p}(\Omega)$. Furthermore, for any $\alpha \in I_k$, $1 \le k \le m$, one has

$$D^\alpha(\varphi u) = \sum_{\beta+\gamma=\alpha} D^\beta \varphi D^\gamma u.$$

It is well known that the space $W_0^{m,p}(\Omega)$ can be equipped with the norm

$$\|u\|_{0,m,p} = \Big(\sum_{|\alpha|=m} (\|D^\alpha u\|_p)^p \Big)^{\frac{1}{p}}, \quad u \in W_0^{m,p}(\Omega) \tag{3.1}$$

which is equivalent to the standard norm $\|\cdot\|_{m,p}$ (see, e.g., [48, Sections 6.25 and 6.26]). Analogously, the space $M_0^{m,p}(\Omega)$ can be equipped with a different norm, which is equivalent to the norm $\|\cdot; M^{m,p}\|$, and is more suitable for our purposes. Set

$$\|u; M_0^{m,p}\| = \Big(\sum_{\alpha \in I_m} (\|D^\alpha u\|_p)^p \Big)^{\frac{1}{p}}, \quad u \in M_0^{m,p}(\Omega).$$

It is clear that $\|\cdot; M_0^{m,p}\|$ is a seminorm on $M_0^{m,p}(\Omega)$. Arguing in a similar way to the case of $W_0^{m,p}(\Omega)$ (cf. [48, Section 6.26]) one can easily verify that the following theorem holds true.

Theorem 2. *The seminorm $\|\cdot; M_0^{m,p}\|$ is a norm on $M_0^{m,p}(\Omega)$ which is equivalent to the norm $\|\cdot; M^{m,p}\|$.*

3.2. Characterization of $M_0^{n,2}(\Omega)$

Our aim now is to give a simple characterization of the space $M_0^{n,2}(\Omega)$ that is crucial for the computation of the direction of steepest descent. To do that, we need to introduce several integral operators that will be useful in the sequel.

Let $v \in L_2(\Omega)$. For $1 \le i \le n$ define the operator

$$(T_i v)(x) = \int_{a_i}^{x_i} v(x_1, \dots x_{i-1}, \xi_i, x_{i+1}, \dots, x_n)\, d\xi_i \quad \text{for a.e. } x \in \Omega.$$

By virtue of the Fubini theorem, one gets that T_i is correctly defined and is a continuous linear operator mapping $L_2(\Omega)$ to $L_2(\Omega)$. Let $\alpha \in I_k$, $1 \le k \le n$, and suppose that $\alpha_{i_j} = 1$, $1 \le j \le k$, where $1 \le i_1 < \dots < i_k \le n$, and $\alpha_l = 0$ iff $l \ne i_j$ for any $1 \le j \le k$, i.e., α_{i_j} are the only nonzero components of the multi-index α. Then define the operator $T_\alpha = T_{\alpha_{i_1}} \circ \dots \circ T_{\alpha_{i_k}}$ mapping continuously $L_2(\Omega)$ to $L_2(\Omega)$. With the use of the Fubini theorem, one gets that for any permutation ℓ of the set $\{1, \dots, k\}$, the following holds

$$T_\alpha = T_{\alpha_{i_{\ell(1)}}} \circ \dots \circ T_{\alpha_{i_{\ell(k)}}}.$$

For the sake of convenience, denote $T = T_{(1,\dots,1)}$. Applying the Lebesgue differentiation theorem (cf., for instance, [51, Corollary 2.9.9]) and integrating by parts, one can verify that for any $v \in L_2(\Omega)$, $1 \le k \le n$ and $\alpha \in I_k$ there exists the weak derivative $D^\alpha(Tv) \in L_2(\Omega)$ and

$$D^\alpha(Tv) = T_{\overline{\alpha}}(v). \tag{3.2}$$

Recall that $\overline{\alpha} \in I_{n-k}$ is a unique multi-index such that $\alpha + \overline{\alpha} \in I_n$. Thus, as it is easy to see, T is a continuous linear operator mapping $L_2(\Omega)$ to $M^{n,2}(\Omega)$.

For any $\alpha \in I_k$, $1 \le k \le n$ and for all $v \in L_2(\Omega)$ define

$$S_\alpha v = \int_{a_{\alpha_{i_1}}}^{b_{\alpha_{i_1}}} \dots \int_{a_{\alpha_{i_k}}}^{b_{\alpha_{i_k}}} v\, d\xi_{\alpha_{i_k}} \dots d\xi_{\alpha_{i_1}},$$

where $1 \le i_1 < \dots < i_k \le n$ and $\alpha_r = 0$ iff $r \ne i_j$ for all $1 \le j \le k$. For the sake of convenience denote $S_i = S_\alpha$ for any $\alpha \in I_1$, where i is the only non-zero component of α.

It is easy to verify that for any $\alpha \in I_k$, $1 \le k \le n$ the linear operator S_α continuously maps $L_2(\Omega)$ to $L_2(\Omega)$. Therefore, in particular, the linear subspace

$$L_0 = \bigcap_{1 \le i \le n} \ker S_i = \left\{ v \in L_2(\Omega) \;\middle|\; \int_{a_k}^{b_k} v\, d\xi_k = 0 \; k \in \{1, \dots, n\} \right\}.$$

of the Hilbert space $L_2(\Omega)$ is closed. Consequently, there exists the orthogonal projector Pr_{L_0} of $L_2(\Omega)$ onto L_0. We will need an explicit formula for the projector Pr_{L_0} (cf. Proposition 2).

Proposition 3. *For any $v \in L_2(\Omega)$, one has*

$$Pr_{L_0} v = v + \sum_{k=1}^{n} \sum_{\alpha \in I_k} (-1)^{|\alpha|} c_\alpha S_\alpha v, \tag{3.3}$$

where $c_\alpha = \prod_{i=1}^n (b_i - a_i)^{-\alpha_i}$.

Proof. Fix an arbitrary $v \in L_2(\Omega)$, and denote the function on the right-hand side of (3.3) by w. A direct computation shows that $\int_{a_k}^{b_k} w \, d\xi_k = 0$ for all $k \in \{1, \dots, n\}$. Consequently, $w \in L_0$. Let us show that $v - w \in L_0^\perp$, where $L_0^\perp$ is the orthogonal complement of L_0, then $Pr_{L_0} v = w$, since the decomposition $v = v_1 + v_2$ for $v_1 \in L_0$ and $v_2 \in L_0^\perp$ is unique. For an arbitrary $h \in L_0$, one has

$$\langle v - w, h \rangle = \sum_{k=1}^{n} \sum_{\alpha \in I_k} (-1)^{|\alpha|} c_\alpha \int_\Omega (S_\alpha v)(x) h(x) \, dx.$$

For any $\alpha \in I_k$, $1 \le k \le n$ there exists $1 \le i \le n$ such that $\alpha_i = 1$. Hence $S_\alpha v$ does not depend on x_i. Denote $\Omega_i = \prod_{k \ne i} (a_i, b_i)$. One has

$$\int_\Omega (S_\alpha v)(x) h(x) \, dx = \int_{\Omega_i} (S_\alpha v)(x) (S_i h)(x) \, dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$
$$= 0,$$

by the Fubini theorem and the fact that $S_i h = 0$ since $h \in L_0$. Therefore $\langle v - w, h \rangle = 0$ for any $h \in L_0$, which means that $v - w \in L_0^\perp$. $\square$

Remark 4.

(i) Note that for any function $\varphi \in C^n(\Omega)$ with compact support, one has $D^{(1,\dots,1)} \varphi \in L_0$.

(ii) It is obvious that if $v \in C^k(\Omega) \cap L_2(\Omega)$, then $Pr_{L_0} v \in C^k(\Omega)$, i.e., the projection operator Pr_{L_0} preserves smoothness.

The following theorem gives a convenient characterization of $M_0^{n,2}(\Omega)$ (cf. Proposition 1).

Theorem 3. *A function $u \colon \Omega \to \mathbb{R}$ belongs to $M_0^{n,2}(\Omega)$ if and only if there exists a function $v \in L_2(\Omega)$ such that*

1. *$v \in L_0$, i.e., for any $1 \le i \le n$ one has $S_i v = \int_{a_i}^{b_i} v \, d\xi_i = 0$,*
2. *$u(x) = (Tv)(x) = \int_{a_n}^{x_n} \dots \int_{a_1}^{x_1} v(\xi) d\xi_1 \dots d\xi_n$ for a.e. $x \in \Omega$.*

Proof. Let us show that for any $u \in M_0^{n,2}(\Omega)$ one has that $v = D^{(1,\dots,1)} u \in L_0$ and $u = Tv$. Indeed, it is clear that for any $\varphi \in C_0^\infty(\Omega)$, one has that $D^{(1,\dots,1)} \varphi \in L_0$

and $\varphi = TD^{(1,\dots,1)}\varphi$. Let $u \in M_0^{n,2}(\Omega)$ be arbitrary, and $\{\varphi_k\} \subset C_0^\infty(\Omega)$ be a sequence such that $\varphi_k \to u$ in $M^{n,2}(\Omega)$. Then, denoting $v = D^{(1,\dots,1)}u$ and $v_k = D^{(1,\dots,1)}\varphi_k$, one gets that for some $C \geq 0$, depending only on n and Ω, the following inequalities holds true

$$\|u - Tv; M^{n,2}(\Omega)\| \leq \|u - \varphi_k; M^{n,2}(\Omega)\| + \|Tv - Tv_k; M^{n,2}(\Omega)\|$$

$$\leq \|u - \varphi_k; M^{n,2}(\Omega)\| + C\|v - v_k\|_2 \leq (C+1)$$

$$\|u - \varphi_k; M^{n,2}(\Omega)\| \to 0$$

as $k \to \infty$. Consequently, $u = T(D^{(1,\dots,1)}u)$. Moreover, since for any $k \in \mathbb{N}$, one has $S_i v_k = 0$, then, as it easy to verify, $S_i v = 0$ for all $1 \leq i \leq n$. Consequently, $v = D^{(1,\dots,1)}u \in L_0$.

Let us prove the converse statement. Fix an arbitrary $v \in L_0$, and set $u = Tv$. We need to prove that there exists a sequence $\{\varphi_m\} \subset C_0^\infty(\Omega)$ such that $\varphi_m \to u$ in $M^{n,2}(\Omega)$. We will prove that there exist functions $u_m \in M^{n,2}(\Omega)$ such that $u_m \to u$ in $M^{n,2}(\Omega)$ and $u_m = 0$ outside some compact set $K_m \subset \Omega$. Then one can mollify u_m to generate a sequence $\{\varphi_m\} \subset C_0^\infty(\Omega)$ such that $\varphi_m \to u$ in $M^{n,2}(\Omega)$ (see, for instance, [49, Theorem 5.3.1 and Appendix C.4]).

Choose an arbitrary function $\zeta \in C_0^\infty(\mathbb{R})$ such that $\zeta(x) = 1$ when $x \in [0,1]$, $\zeta(x) = 0$, when $x \geq 2$ and $0 \leq \zeta(x) \leq 1$ for all $x \in \mathbb{R}$. For any $m \in \mathbb{N}$ define the function

$$u_m(x) = (Tv)(x) \prod_{k=1}^{n} \left[1 - \zeta(m(x_k - a_k))\right]\left[1 - \zeta(m(b_k - x_k))\right] \quad \forall x \in \Omega.$$

It is clear that $u_m \in M^{n,2}(\Omega)$ and $u_m = 0$ outside $\prod_{k=1}^{n}[a_k + 1/m, b_k - 1/m]$ for $m \in \mathbb{N}$ large enough. Let us show that $u_m \to u$ in $M^{n,2}(\Omega)$, then we get the desired result.

Indeed, it is easy to verify that $u_m \to u$ in $L_2(\Omega)$. Fix an arbitrary $\alpha \in I_r$, $1 \leq r \leq n$. Observe that

$$\prod_{k=1}^{n} \big(1 - \zeta(m(x_k - a_k))\big)\big(1 - \zeta(m(b_k - x_k))\big)$$

$$= 1 + \sum_{k=1}^{n} \zeta(m(x_k - a_k))\omega_k^{(1)}(x) + \sum_{k=1}^{n} \zeta(m(b_k - x_k))\omega_k^{(2)}(x),$$

where the functions $\omega_k^{(i)}$ are infinitely continuously differentiable and bounded on Ω. Therefore by Remark 3, one has

$$D^\alpha u_m(x) = (D^\alpha u)(x) + \sum_{k=1}^{n} (D^\alpha u)(x)\zeta(m(x_k - a_k))\omega_k^{(1)}(x)$$

$$+ \sum_{k=1}^{n} (D^\alpha u)(x)\zeta(m(b_k - x_k))\omega_k^{(2)}(x) + \sum_{\beta+\gamma=\alpha, \beta\neq\alpha} (D^\beta u)(x)\zeta_\gamma(x),$$

where $\zeta_\gamma \in C^\infty(\Omega)$ and

$$\zeta_\gamma(x) = D^\gamma\Big(\prod_{k=1}^{n}\big[1-\zeta(m(x_k-a_k)\big]\big[1-\zeta(m(b_k-x_k)\big]\Big). \tag{3.4}$$

Hence there exists $C \geq 0$ depending only on n, Ω and ζ such that

$$\begin{aligned}\int_\Omega |D^\alpha u_m - D^\alpha u|^2 dx \leq C\sum_{k=1}^{n}\int_{a_k}^{a_k+2/m}\int_{\Omega_k}|D^\alpha u|^2\, d\xi^k dx_k \\ + C\sum_{k=1}^{n}\int_{b_k-2/m}^{b_k}\int_{\Omega_k}|D^\alpha u|^2\, d\xi^k dx_k \\ + C\sum_{\beta+\gamma=\alpha,\beta\neq\alpha}\int_\Omega |D^\beta u|^2(x)|\zeta_\gamma|^2(x)\, dx \\ = A_1(m)+A_2(m)+A_3(m)\end{aligned}$$

by virtue of the fact that $\zeta(m(x_k-a_k)) = 0$ for any $x_k > a_k + 2/m$ and $\zeta(m(b_k-x_k)) = 0$ for any $x < b_k - 2/m$. Here $\Omega_k = \prod_{i\neq k}(a_i,b_i)$ and $\xi^k = (\xi_1,\ldots,\xi_{k-1},x_k,\xi_{k+1},\ldots,\xi_n)$. It is clear that $A_1(m)\to 0$ and $A_2(m)\to 0$ as $m\to\infty$. Let us show that $A_3(m)\to 0$ as $m\to\infty$, then $\|D^\alpha u_m - D^\alpha u\|_2 \to 0$ as $m\to\infty$ for any $\alpha \in I_r$, $1\leq r\leq n$ and, consequently, $u_m \to u$ in $M^{n,2}(\Omega)$.

Fix an arbitrary $0\leq k\leq r-1$, $\beta\in I_k$ and $\gamma \in I_{r-k}$ such that $\beta+\gamma=\alpha$. Without loss of generality, we can suppose that $\alpha_1 = \ldots = \alpha_r = 1$, $\gamma_1 = \ldots = \gamma_{r-k} = 1$ and $\beta_{r-k+1} = \ldots = \beta_r = 1$. Our aim is to show that

$$\int_\Omega |D^\beta u|^2(x)|\zeta_\gamma(x)|^2\, dx \to 0 \text{ as } m\to\infty,$$

then $A_3(m)\to 0$ as $m\to\infty$. Taking into account (3.4), one gets that

$$\zeta_\gamma(x) = m^{|\gamma|}\sum_{\eta+\theta=\gamma}\omega_{\eta,\theta}(x)\prod_{l=1}^{n}\zeta'(m(x_l-a_l))^{\eta_l}\prod_{s=1}^{n}\zeta'(m(b_s-x_s))^{\theta_s},$$

where $\omega_{\eta,\theta}\in C^\infty(\Omega)$ and $|\omega_{\eta,\theta}|\leq 1$. Therefore, it is sufficient to show that for any multi-indices η and θ such that $\eta+\theta=\gamma$ one has

$$m^{2|\gamma|}\int_\Omega |D^\beta u|^2(x)\prod_{l=1}^{n}|\zeta'(m(x_l-a_l))|^{2\eta_l}\prod_{s=1}^{n}|\zeta'(m(b_s-x_s))|^{2\theta_s}\, dx \to 0$$

as $m\to\infty$. Fix an arbitrary $0\leq j\leq r-k$, $\eta\in I_j$ and $\theta\in I_{r-k-j}$ such that $\eta+\theta=\gamma$. Without loss of generality, we can suppose that $\eta_1 = \ldots = \eta_j = 1$ and $\theta_{j+1} = \ldots = \theta_{r-k} = 1$. As it was mentioned above, $D^\beta u = D^\beta Tv = T_{\overline{\beta}}v$ (see Equality (3.2)). Then, taking into account the fact that since $v\in L_0$, then

$\int_{a_i}^{x_i} v\, d\xi_i = -\int_{x_i}^{b_i} v\, d\xi_i$ for all $1 \le i \le n$ one has that for a.e. $x \in \Omega$

$$D^\beta u = (-1)^{|\theta|} \int_{a_1}^{x_1} \dots \int_{a_j}^{x_j} \int_{x_{j+1}}^{b_j} \dots \int_{x_{r-k}}^{b_{r-k}} \int_{a_{r+1}}^{x_{r+1}} \dots \int_{a_n}^{x_n} v\, d\xi_n \dots d\xi_{r+1} d\xi_{r-k} \dots d\xi_1.$$

Consequently, with the use of the Hölder inequality one gets that there exists $C \ge 0$, depending only on n and Ω, such that for a.e. $x \in \Omega$

$$|D^\beta u|^2(x) \le C \prod_{l=1}^{j} (x_l - a_l) \prod_{s=j+1}^{r-k} |b_s - x_s| (T_{\overline{\beta}} |v|^2)(x).$$

Therefore, applying the fact that $\zeta'(m(x_k - a_k)) = 0$ for any $x_k > a_k + 2/m$ and $\zeta'(m(b_k - x_k)) = 0$ for any $x_k < b_k - 2/m$ one gets that for some $C_1, C_2 > 0$ that do not depend on m the following holds

$$\begin{aligned} m^{2|\gamma|} &\int_\Omega |D^\beta u|^2(x) \prod_{l=1}^{j} |\zeta'(m(x_l - a_l))| \prod_{s=j+1}^{r-k} |\zeta'(m(b_s - x_s))|\, dx \\ &\le C_1 m^{2|\gamma|} \int_{a_1}^{a_1+2/m} \dots \int_{a_j}^{a_j+2/m} \int_{b_{j+1}-2/m}^{b_{j+1}} \dots \int_{b_{r-k}-2/m}^{b_{r-k}} \prod_{l=1}^{j} (x_l - a_l) \\ &\quad \times \prod_{s=j+1}^{r-k} |b_s - x_s| \Big((S_{\overline{\gamma}} T_{\overline{\beta}} |v|^2)(x) \Big) dx_{r-k} \dots dx_1 \\ &\le C_2 \int_{a_1}^{a_1+2/m} \dots \int_{a_j}^{a_j+2/m} \int_{b_{j+1}-2/m}^{b_{j+1}} \dots \int_{b_{r-k}-2/m}^{b_{r-k}} \int_{a_{r-k+1}}^{b_{r-k+1}} \dots \int_{a_n}^{b_n} |v|^2\, dx \\ &\to 0 \end{aligned}$$

as $m \to \infty$. Thus, the proof is complete. □

Corollary 1. *The operator T is an isometric isomorphism between*

$$L_0 = \left\{ v \in L_2(\Omega) \;\middle|\; \int_{a_i}^{b_i} v d\xi_i = 0, 1 \le i \le n \right\}$$

and $M_0^{n,2}(\Omega)$. Furthermore, the inverse operator of the restriction of T to L_0 is the restriction of the differential operator $D^{(1,\dots,1)}$ to $M_0^{n,2}(\Omega)$.

Remark 5.

(i) The proof of the converse statement of the theorem above is based on the proof of the theorem on the characterization of $W_0^{1,p}(\Omega)$ in terms of the trace operator (see [49, Theorem 5.5.2]).

(ii) Denote by Tr the trace operator defined on $W^{1,2}(\Omega)$. Clearly, if $u \in M_0^{n,2}(\Omega) \subset W_0^{1,2}(\Omega)$, then $Tr(u) = 0$ (see [49, Theorem 5.5.2] and [50, Theorem 15.29]). However, there is an open question whether the converse statement is true, i.e., whether $u \in M^{n,2}(\Omega)$ and $Tr(u) = 0$ implies that $u \in M_0^{n,2}(\Omega)$. This question becomes more reasonable after one notes that if $u \in C^n(\overline{\Omega})$ and $Tr(u) = 0$, i.e., $u|_{\partial\Omega} = 0$, then $u \in M_0^{n,2}(\Omega)$. Indeed, since $Tr(u) = 0$, then

$$u(x) = \int_{a_1}^{x_1} \frac{\partial u}{\partial x_1}(\xi_1, x_2, \ldots, x_n)\, d\xi_1 \quad \forall x \in \Omega. \tag{3.5}$$

Observe that $u(x_1, a_2, x_3, \ldots, x_n) \equiv 0$. Therefore $\partial u/\partial x_1(x_1, a_2, x_3, \ldots, x_n) = 0$ for all $x \in \Omega$. Hence

$$\frac{\partial u}{\partial x_1}(x) = \int_{a_2}^{x_2} \frac{\partial^2 u}{\partial x_2 \partial x_2}(x_1, \xi_2, x_3, \ldots, x_n)\, d\xi_2 \quad \forall x \in \Omega.$$

Taking into account (3.5), one gets that

$$u(x) = \int_{a_1}^{x_1} \int_{a_2}^{x_2} \frac{\partial^2 u}{\partial x_2 \partial x_2}(\xi_1, \xi_2, x_3, \ldots, x_n)\, d\xi_2 d\xi_1 \quad \forall x \in \Omega.$$

Applying mathematical induction one can easily obtain that $u = TD^{(1,\ldots,1)}u$ and $D^{(1,\ldots,1)}u \in L_0$, which by the theorem above yields that $u \in M_0^{n,2}(\Omega)$.

In the following sections, we will use the function space $M^{m,p}(\Omega, \mathbb{R}^d)$ (or $M_0^{m,p}(\Omega, \mathbb{R}^d)$) that consists of all functions $u = (u_1, \ldots, u_d) \colon \Omega \to \mathbb{R}^d$ such that $u_i \in M^{m,p}(\Omega)$ (or $u_i \in M_0^{m,p}(\Omega)$) for all $i \in \{1, \ldots, d\}$. It is clear that the spaces $M^{m,p}(\Omega, \mathbb{R}^d)$ and $M_0^{m,p}(\Omega, \mathbb{R}^d)$ possess the same properties as their "one dimensional" counterparts. In particular, one can easily obtain a characterization of the space $M_0^{n,2}(\Omega, \mathbb{R}^d)$ similar to the characterization of $M_0^{n,2}(\Omega)$.

4. Direction of steepest descent

In this section, we compute the direction of steepest descent of the main functional of the calculus of variations. This direction can be utilized to design new direct numerical methods for solving multidimensional problems of the calculus of variations.

Consider the following problem of the calculus of variations

$$\min \mathcal{I}(u) = \int_\Omega f(x, u(x), \nabla u(x))\, dx \quad \text{subject to} \quad u \in \overline{u} + M_0^{n,2}(\Omega, \mathbb{R}^d), \tag{4.1}$$

where $\overline{u} \in W^{1,2}(\Omega, \mathbb{R}^d)$ is a given function that defines boundary conditions, $u = (u_1, \ldots, u_d)$ and

$$\nabla u = \left\{ \frac{\partial u_j}{\partial x_i} \right\}_{1 \le i \le n}^{1 \le j \le d} \in \mathbb{R}^{d \times n}.$$

We suppose that the function $f\colon \Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times n} \to \mathbb{R}, f = f(x,u,z)$ satisfies the Carathéodory condition, is differentiable with respect to u and z, and the derivatives $\partial f/\partial u_i$, $\partial f/\partial z_{ij}$ satisfy the Carathéodory condition as well. Let also the following growth conditions be valid:

1. there exist $C > 0$ and $g \in L_1(\Omega)$ such that for a.e. $x \in \Omega$ and for all $u \in \mathbb{R}^n$, $z \in \mathbb{R}^{d\times n}$ one has $|f(x,u,z)| \le C(|u|^2 + |z|^2) + g(x)$.
2. there exist $D_1, D_2 > 0$, and $g_1, g_2 \in L_2(\Omega)$ such that for a.e. $x \in \Omega$ and for all $u \in \mathbb{R}^n$, $z \in \mathbb{R}^{d\times n}$ one has

$$\left|\frac{\partial f}{\partial u}(x,u,z)\right| \le D_1(|u| + |z|) + g_1(x),$$

$$\left|\frac{\partial f}{\partial z}(x,u,z)\right| \le D_2(|u| + |z|) + g_2(x).$$

Remark 6. In the general case, problem (4.1) is not equivalent to the standard problem of the calculus of variations

$$\min\ \mathcal{I}(u) = \int_\Omega f(x, u(x), \nabla u(x))\,dx \quad \text{subject to} \quad u \in \overline{u} + W_0^{1,2}(\Omega, \mathbb{R}^d). \tag{4.2}$$

Problems (4.1) and (4.2) are equivalent if and only if there exists an optimal solution u^* of problem (4.2) such that $u^* \in \overline{u} + M_0^{n,2}(\Omega, \mathbb{R}^d)$. In other words, problems (4.1) and (4.2) are equivalent if and only if there exists a "sufficiently smooth" optimal solution u^* of problem (4.2).

It is easy to see that the functional $\mathcal{I}$ is Gâteaux differentiable at every point $u \in W^{1,2}(\Omega, \mathbb{R}^d)$ and its Gâteaux derivative has the form

$$\mathcal{I}'[u](h) = \int_\Omega \left(\left\langle \frac{\partial f}{\partial u}(x, u(x), \nabla u(x)), h(x)\right\rangle + \left\langle \frac{\partial f}{\partial z}(x, u(x), \nabla u(x)), \nabla h(x)\right\rangle\right) dx \tag{4.3}$$

for all $h \in W^{1,2}(\Omega, \mathbb{R}^d)$. Here $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^s$, and

$$\frac{\partial f}{\partial u} = \left(\frac{\partial f}{\partial u_1}, \dots, \frac{\partial f}{\partial u_d}\right), \qquad \frac{\partial f}{\partial z} = \left\{\frac{\partial f}{\partial z_{ji}}\right\}_{1\le i\le n}^{1\le j\le d}.$$

As it was mentioned above, the problem of finding the direction of steepest descent of the functional $\mathcal{I}$ is very complicated. We utilize the characterization of $M_0^{n,2}(\Omega, \mathbb{R}^d)$ (Theorem 3) to solve this problem and to design new direct numerical methods for minimizing this functional.

Recall that the operator T,

$$(Tv)(x) = \int_{a_n}^{x_n} \cdots \int_{a_1}^{x_1} v(\xi_1, \dots, \xi_n) d\xi_1 \dots d\xi_n \quad \text{for a.e. } x \in \Omega,$$

is an isometric isomorphism between $L_0 \subset L_2(\Omega, \mathbb{R}^d)$ and $M_0^{n,2}(\Omega, \mathbb{R}^d)$. Therefore problem (4.1) is equivalent to the problem

$$\min \quad F(v) = \mathcal{I}(\overline{u} + Tv) = \int_\Omega f(x, \overline{u}(x) + (Tv)(x), \nabla \overline{u}(x) + \nabla (Tv)(x))\, dx$$

$$\text{subject to} \quad v \in L_0,$$

We suppose that the functional F is defined on L_0, and will minimize F over the space L_0.

The functional F is Gâteaux differentiable, and integrating by parts in (4.3), one gets that the Gâteaux derivative of the functional F has the form

$$F'[v](h) = \langle Q(u), h \rangle = \int_\Omega \langle Q(u)(x), h(x) \rangle\, dx \quad \forall h \in L_0,$$

where $u = \overline{u} + Tv$,

$$Q(u)(x) = (-1)^n \int_{x_n}^{b_n} \dots \int_{x_1}^{b_1} \frac{\partial f}{\partial u}(\xi, u(\xi), \nabla u(\xi))\, d\xi_1 \dots d\xi_n$$

$$+ (-1)^{n-1} \sum_{i=1}^{n} \int_{x_n}^{b_n} \dots \int_{x_{i+1}}^{b_{i+1}} \int_{x_{i-1}}^{b_{i-1}} \dots \int_{x_1}^{b_1}$$

$$\frac{\partial f}{\partial z_i}(\xi^i, u(\xi^i), \nabla u(\xi^i))\, d\xi_1 \dots d\xi_{i-1} d\xi_{i+1} \dots d\xi_n, \qquad (4.4)$$

$\partial f/\partial z_i = (\partial f/\partial z_{1i}, \dots, \partial f/\partial z_{di})$, and $\xi^i = (\xi_1, \dots, \xi_{i-1}, x_i, \xi_{i+1}, \dots, \xi_n)$. Since the derivatives of f satisfy the growth condition, then $Q(u) \in L_2(\Omega)$ for any $u \in M_0^{n,2}(\Omega, \mathbb{R}^d)$. Hence one has that

$$F'[v](h) = \langle G(u), h \rangle = \int_\Omega \langle G(u)(x), h(x) \rangle\, dx \quad \forall h \in L_0,$$

where $u = \overline{u} + Tv$, and $G(u) = Pr_{L_0} Q(u)$ is the projection of $Q(u)$ onto L_0 in $L_2(\Omega, \mathbb{R}^d)$ (see Proposition 3 above). Note that $G(u) \in L_0$ is the Gâteaux gradient of the functional F at the point v. Consequently, $-G(u)/\|G(u)\|_2$ is the direction of steepest descent of F at the point v, which implies that $-TG(u)/\|G(u)\|_2$ is the direction of steepest descent of the functional $\mathcal{I}$ at the point $u = \overline{u} + Tv$.

Thus, the following result holds true.

Proposition 4. *Let $v \in L_0$ and $u = \overline{u} + Tv$ (or, equivalently, $u = \overline{u} + w$ for some $w \in M_0^{n,2}(\Omega, \mathbb{R}^d)$ and $v = D^{(1,\dots,1)} w$) be such that $F'[v] \neq 0$. Let also $\mathcal{J}(w) = \mathcal{I}(\overline{u} + w)$ for any $w \in M_0^{n,2}(\Omega, \mathbb{R}^d)$. Then $-TG(u)/\|G(u)\|_2$ is the direction of steepest descent of $\mathcal{J}$ at the point $w = Tv$, and $-G(u)/\|G(u)\|_2$ is the direction of steepest descent of F at the point v.*

Remark 7. Note that for any $v \in L_0$, one has $u = Tv \in M_0^{n,2}(\Omega, \mathbb{R}^d) \subset W_0^{1,2}(\Omega, \mathbb{R}^d)$. Therefore, as it was mentioned above, the trace $Tr(u)$ of the function $u = Tv$ is correctly defined and equals zero. Thus, the direction of steepest descent $-TG(u)/\|G(u)\|_2$ does not change the values of a function on the boundary of Ω, i.e., $Tr(u + \alpha TG(u)) = Tr(u)$ for all $\alpha \in \mathbb{R}$.

Let us mention several simple properties of the mappings $G(\cdot)$ and $Q(\cdot)$.

Proposition 5. *Let $v \in L_0$ and $u = \overline{u} + Tv$ (or, equivalently, $u = \overline{u} + w$ for some $w \in M_0^{n,2}(\Omega, \mathbb{R}^d)$ and $v = D^{(1,\ldots,1)}w$). Let also $\mathcal{J}(w) = \mathcal{I}(\overline{u} + w)$ for any $w \in M_0^{n,2}(\Omega, \mathbb{R}^d)$. Then the following statements hold true:*

1. $\|\mathcal{J}'[w]\|^2 = \|F'[v]\|^2 = \mathcal{I}'[u](TG(u)) = F'[v](G(u)) = \big(\|G(u)\|_2\big)^2$;
2. $\mathcal{I}'[u] = 0$ *iff* $F'[v] = 0$ *iff* $D^{(1,\ldots,1)}Q(u) = 0$ *in the weak sense;*
3. *suppose that $u \in C^2(\Omega, \mathbb{R}^d)$ and $f \in C^2(\Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times n})$. Then $F'[v] = 0$ if and only if for all $x \in \Omega$ one has*

$$\frac{\partial f}{\partial u}(x, u(x), \nabla u(x)) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\frac{\partial f}{\partial z_i}(x, u(x), \nabla u(x)) = 0,$$

i.e., $F'[v] = 0$ iff $u = \overline{u} + Tv$ satisfies the Euler–Lagrange equation.

Proof. The validity of the first statement follows directly from definitions. Let us prove the second statement. It is clear that $\mathcal{I}'[u](Th) = F'[v](h)$ for any $h \in L_0$. Therefore $\mathcal{I}'[u] = 0$ if and only if

$$F'[v](h) = \langle Q(u), h\rangle = \int_\Omega \langle Q(u)(x), h(x)\rangle\, dx = 0 \quad \forall h \in L_0. \tag{4.5}$$

Since for any $\varphi \in C_0^\infty(\Omega, \mathbb{R}^d)$, one has $D^{(1,\ldots,1)}\varphi \in L_0$, then (4.5) is equivalent to the fact that $D^{(1,\ldots,1)}Q(u) = 0$ in the weak sense.

Suppose in addition that $u \in C^2(\Omega, \mathbb{R}^d)$ and $f \in C^2(\Omega \times \mathbb{R}^d \times \mathbb{R}^{d\times n})$. Then integrating in (4.5) by parts, one gets that

$$\int_\Omega \langle g(x), \varphi(x)\rangle\, dx = 0 \quad \forall \varphi \in C_0^\infty(\Omega, \mathbb{R}^d),$$

where

$$g(x) = (D^{(1,\ldots,1)}Q(v))(x) = \frac{\partial f}{\partial u}(x, u(x), \nabla u(x)) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\frac{\partial f}{\partial z_i}(x, u(x), \nabla u(x)).$$

Applying the fundamental lemma of the calculus of variations (see, e.g., [5, Theorem 3.40]), one obtains the desired result. □

Observe that the mapping $u \to G(u) \in L_0$ is well defined for any $u \in W^{1,2}(\Omega, \mathbb{R}^d)$.

Proposition 6. *Let $u^* \in \overline{u} + W_0^{1,2}(\Omega, \mathbb{R}^d)$. Then $\mathcal{I}'[u^*](\cdot) = 0$ on $W_0^{1,2}(\Omega, \mathbb{R}^d)$ if and only if $G(u^*) = 0$.*

Proof. Necessity. Suppose that $\mathcal{I}'[u^*](\cdot) = 0$ on $W_0^{1,2}(\Omega, \mathbb{R}^d)$. Since $G(u^*) \in L_0$, then $TG(u^*) \in M_0^{n,2}(\Omega, \mathbb{R}^d) \subset W_0^{1,2}(\Omega, \mathbb{R}^d)$. Hence $\mathcal{I}'[u^*](TG(u^*)) = 0$. Consequently, integrating by parts, one gets

$$0 = \mathcal{I}'[u^*](-TG(u^*)) = -\int_\Omega \langle Q(u^*)(x), G(u^*)(x) \rangle \, dx$$
$$= -\int_\Omega \langle G(u^*)(x), G(u^*(x)) \rangle \, dx = -(\|G(u^*)\|_2)^2,$$

since $G(u^*) \in L_0$ and $Pr_{L_0} Q(u^*) = G(u^*)$. Thus, $G(u^*) = 0$.

Sufficiency. Let $G(u^*) = 0$. Then it is easy to verify that $\mathcal{I}'[u^*](h) = 0$ for all $h \in M_0^{n,2}(\Omega, \mathbb{R}^d)$. The space $M_0^{n,2}(\Omega, \mathbb{R}^d)$ is dense in $W_0^{1,2}(\Omega, \mathbb{R}^d)$, and the linear functional $\mathcal{I}'[u^*]$ is bounded on $W^{1,2}(\Omega, \mathbb{R}^d)$. Therefore $\mathcal{I}'[u^*](h) = 0$ for any $h \in W_0^{1,2}(\Omega, \mathbb{R}^d)$. □

Since we now know the gradient $G(\cdot)$ of the functional F, we can modify most of the gradient-based methods of finite-dimensional optimization to the case of this functional and use these methods to find critical points (or points of global minimum in the convex case) of the functionals F and $\mathcal{I}$. Convergence analysis of these methods can be performed in the standard way (see, e.g., [29, 31, 36, 38–41] for convergence analysis of minimization methods in Banach and Hilbert spaces.).

5. Conclusion

In this article, we developed a new approach to the design of minimization algorithms for the main problem of the calculus of variations. Let us discuss some possible generalizations of this approach and some directions of future research.

5.1. More general boundary conditions

The results developed in this article can be modified to the case when the boundary condition has the form $u|_\Gamma = \psi$, where $\Gamma \subset \partial\Omega$. In other words, one can extend the theory presented in this article to the case when the values of a function u are specified only on a part of the boundary of Ω. However, it should be mentioned that any modification of the methods discussed above to the case of more general boundary conditions requires a different formalization, then the one based on the use of space $M_0^{n,2}(\Omega)$. Therefore, the following discussion has an informal character.

Let, for example, $n = 2$ and the boundary condition has the from

$$u|_\Gamma = \psi, \quad \Gamma = \partial\Omega \setminus [a_1, b_1] \times \{b_2\},$$

i.e., the values of u are not specified on the upper side on the rectangle $\Omega = (a_1, b_1) \times (a_2, b_2)$. Then one can show that the gradient of the functional F has the form

$$Q(u)(x_1, x_2) - \frac{1}{b_1 - a_1} \int_{a_1}^{b_1} Q(u)(\xi_1, x_2)\, d\xi_1 \quad \forall (x_1, x_2) \in \Omega.$$

Here we use the same notation as in Section 4.

In the case when the boundary condition has the form

$$u|_\Gamma = \psi, \quad \Gamma = \big(\{a_1\} \times [a_2, b_2]\big) \cup \big([a_1, b_1] \times \{a_2\}\big),$$

the gradient of the functional F coincides with the function $Q(u)$.

If the boundary condition has the from

$$u|_\Gamma = \psi, \quad \Gamma = \big(\{a_1\} \times [a_2, b_2]\big) \cup \big([a_1, b_1] \times \{b_2\}\big),$$

then one should use a different representation of a function u:

$$u(x_1, x_2) = \int_{a_1}^{x_1} \int_{x_2}^{b_2} \frac{\partial^2 u}{\partial x_1 \partial x_2}(\xi_1, \xi_2) d\xi_2 d\xi_1.$$

This representation can be used to compute the direction of steepest descent.

5.2. Isoperimetric problems

One can easily modify the proposed approach to the case of problems of the calculus of variations with linear isoperimetric constraints. Namely, let $n = 2$, and suppose that there is the additional constraint

$$\mathcal{J}(u) = \int_\Omega \big(g_0(x)u(x) + g_1(x)D_1u(x) + g_2(x)D_2u(x)\big)\, dx = c. \qquad (5.1)$$

Denote by $Q_{\mathcal{I}}(u)$ the function $Q(u)$ for the functional $\mathcal{I}$ (see Equalities (2.6) and (4.4)).

Applying the same argument as in Section 2, one can easily demonstrate that the direction of steepest descent for the problem with additional isoperimetric constraint (5.1) has the same form as for the problem without this constrain with the function $Q_{\mathcal{I}}(u)(x)$ replaced by the function $Q_{\mathcal{I}}(u) + \lambda Q_{\mathcal{J}}(u)$. Here λ is a constant that is chosen so that the direction of steepest descent satisfies constraint (5.1) with $c = 0$.

5.3. Functionals depending on higher order derivatives

Let us also note that the approach developed in this article can be generalized to the case when the functional $\mathcal{I}(u)$ depends on derivatives of the function

u of order greater than 1. Indeed, consider, for instance, the following two-dimensional problem of the calculus of variations:

$$\min \quad \mathcal{I}(u) = \int_\Omega f\big(x, u(x), \nabla u(x), \nabla^2 u(x)\big)\, dx \tag{5.2}$$

$$\text{subject to} \quad u|_{\partial\Omega} = \psi_1, \quad \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega} = \psi_2, \tag{5.3}$$

where $\Omega = (a_1, b_1) \times (a_2, b_2)$, and ν is the outward unit normal at the boundary of Ω. Let us demonstrate how one can easily transform this problem utilizing the same technique as above to easily compute the direction of steepest descent of the functional $\mathcal{I}(u)$ with respect to a certain norm. Here we provide only an informal description of such transformation. A formalization of this transformation can be done in the same way as in the case of the functional that depends only on the first-order derivatives.

Let sufficiently smooth functions u and $\overline{u}$ satisfy boundary conditions (5.3). Then the function $w = u - \overline{u}$ satisfies the same boundary conditions with $\psi_1(\cdot) = \psi_2(\cdot) = 0$. Hence

$$w(x_1, x_2) = \int_{a_1}^{x_1} \frac{\partial w}{\partial x_1}(\xi_1, x_2)\, d\xi_1$$

due to the fact that $w(a_1, \cdot) = 0$. Then applying the fact that $w'_{x_1}(a_1, \cdot) = 0$, one obtains that

$$w(x_1, x_2) = \int_{a_1}^{x_1} \int_{a_1}^{\xi_1} \frac{\partial^2 w}{\partial x_1^2}(\theta_1, x_2)\, d\theta_1\, d\xi_1.$$

Note that since $w(\cdot, a_2) = 0$, then $w''_{x_1 x_1}(\cdot, a_2) = 0$, which implies that

$$w(x_1, x_2) = \int_{a_1}^{x_1} \int_{a_1}^{\xi_1} \int_{a_2}^{x_2} \frac{\partial^3 w}{\partial x_2 \partial x_1^2}(\theta_1, \xi_2)\, d\xi_2\, d\theta_1\, d\xi_1.$$

Finally, from the fact that $w'_{x_2}(\cdot, a_2) = 0$, it follows that $w'''_{x_2 x_1 x_1}(\cdot, a_2) = 0$, which yields

$$w(x) = \big(Tv\big)(x), \quad v = \frac{\partial^4 w}{\partial x_2^2 \partial x_1^2},$$

where

$$\big(Tv\big)(x) = \int_{a_1}^{x_1} \int_{a_1}^{\xi_1} \int_{a_2}^{x_2} \int_{a_2}^{\xi_2} v(\theta_1, \theta_2)\, d\theta_2\, d\xi_2\, d\theta_1\, d\xi_1. \tag{5.4}$$

Furthermore, it is easy to check that

$$w|_{\partial\Omega} = 0, \quad \frac{\partial w}{\partial \nu}\Big|_{\partial\Omega} = 0$$

if and only if

$$\int_{a_1}^{b_1} v(\xi_1, \cdot)\, d\xi_1 = 0, \qquad \int_{a_2}^{b_2} v(\cdot, \xi_2)\, d\xi_2 = 0. \tag{5.5}$$

Thus, the following result holds true (cf. Proposition 1).

Proposition 7. *Let $u\colon [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$ be a sufficiently smooth function. Then u satisfies boundary conditions (5.3) if and only if there exists a sufficiently smooth function v such that*
1. *$u = \overline{u} + Tv$, where the operator T is defined in (5.4);*
2. *$\int_{a_1}^{b_1} v(\xi_1, \cdot)\, d\xi_1 \equiv 0$ and $\int_{a_2}^{b_2} v(\cdot, \xi_2)\, d\xi_2 \equiv 0$.*

Moreover, $v = \partial^4 u / \partial x_1^2 \partial x_2^2$.

Applying the proposition above, one obtains that problem (5.2), (5.3) is equivalent to the problem of minimizing the functional $F(v) = \mathcal{I}(\overline{u} + Tv)$ subject to linear equality constraints (5.5). The direction of steepest descent for this problem with respect to the L_2-norm can be easily computed in the same way as in the proof of Proposition 2.

5.4. Directions of future research

Let us briefly outline some other directions of future research:

- One can modify Newton's method [32] to the multidimensional case, and use the methods developed in this article to perform each iteration of Newton's method. The use of Newton's method might be reasonable in the case of highly nonlinear problems of the calculus of variations.
- In the case of problems of the calculus of variations with nonlinear isoperimetric constraints, one can modify sequential quadratic programming methods [52] to solve these problems. On each iteration of this method, one needs to solve a problem with a quadratic functional and linear constraints that can be solved with the use of the methods discussed above.
- One can use the same augmented Lagrangian method as in [53] (or some other augmented Lagrangian methods) to apply the approach developed in this article to variational problems with pointwise constraints.

Acknowledgments

The author is sincerely grateful to his colleague G. Sh. Tamasyan and the late professor V. F. Demyanov. It should be noted that this article is a straightforward generalization of their works on the calculus of variations. Also, the author wishes to express his thanks to the Editor-in-Chief, professor M. Z. Nashed for the effort he had to put to overcome many difficulties that arose during the reviewing process as well as for the thoughtful and stimulating comments that helped to improve the quality of the article.

Funding

The author was supported by the Russian Foundation for Basic Research under grant no. 16-31-00056.

References

1. C. B. Morrey (1966). *Multiple Integrals in the Calculus of Variations*. Springer-Verlag, Berlin.
2. M. Giaquinta (1983). *Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems*. Princeton University Press, Princeton.
3. E. Giusti (2003). *Direct Methods in the Calculus of Variations*. World Scientific, Singapore.
4. G. Mingione (2006). Regularity of minima: An invitation to the dark side of the calculus of variations. *Appl. Math.* 51:355–426.
5. B. Dacorogna (2008). *Direct Methods in the Calculus of Variations*. Springer Science+ Business Media, New York.
6. M. Morse (1934). *Calculus of Variations in the Large*. American Mathematical Society Colloquium Publications, New York.
7. I. Ekeland and N. Ghoussoub (2002). Selected new aspects of the calculus of variations in the large. *Bull. Am. Math. Soc.* 39:207–265.
8. M. Giaquinta and S. Hildebrandt (1996). *Calculus of Variations I and II*. Springer-Verlag, Berlin, Heidelberg.
9. J.-M. Morel and S. Solimini (1995). *Variational Methods in Image Segmentation: With Seven Image Processing Experiments*. Birkhäuser, Boston.
10. G. Aubert and P. Kornprobst (2006). *Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations*. Springer Science+Business Media, New York.
11. J. Gregory (1988). Numerical methods for extremal problems in the calculus of variations and optimal control theory. *Bull. Am. Math. Soc.* 18:31–34.
12. J. Gregory and R. S. Wang (1990). Discrete variable methods for the m-dependent variable, nonlinear extremal problems in the calculus of variations. *SIAM J. Numer. Anal.* 27:470–487.
13. J. Gregory and R. S. Wang (1993). Discrete variable methods for the m-dependent variable nonlinear, extremal problems in the calculus of variations II. *SIAM J. Numer. Anal.* 30: 871–881.
14. L. Ma and N. J. Walkington (1995). On algorithms for nonconvex optimization in the calculus of variations. *SIAM J. Numer. Anal.* 32:900–923.
15. W. Zhou and Z. S. Chalabi (1996). A new algorithm to solve calculus of variations problems using Wolfe's convergence theory, part 1: Theory and algorithm. *Optimization* 37:253–265.
16. W. Zhou and Z. S. Chalabi (1996). A new algorithm to solve calculus of variations problems using Wolfe's convergence theory, part 2: Implementation. *Optimization* 37:267–277.
17. G. N. Elnagar and M. A. Kazemi (1996). Numerical solvability of nonlinear problems of the calculus of variations. *Numer. Funct. Anal. Optim.* 17:563–575.
18. Y. Levin, M. Nediak, and A. Ben-Israel (2002). A direct Newton method for calculus of variations. *J. Comput. Appl. Math.* 139:197–213.
19. O. P. Agrawal and J. Gregory (2003). $O(h^2)$ global pointwise algorithms for delay quadratic problems in the calculus of variations. *SIAM J. Numer. Anal.* 41:1773–1784.
20. R. Lopez, E. Balsa-Canto, and E. Oñate (2008). Neural networks for variational problems in engineering. *Int. J. Numer. Methods Eng.* 75:1341–1360.
21. S. G. Mikhlin (1964). *Variational Methods in Mathematical Physics*. Pergamon Press, Oxford.
22. J. N. Reddy (2002). *Energy Principles and Variational Methods in Applied Mechanics*. John Wiley & Sons, New York.
23. L. Komzsik (2008). *Applied Calculus of Variations for Engineers*. CRC Press, Taylor & Francis, Boca Raton.

24. K. W. Cassel (2013). *Variational Methods with Applications in Science and Engineering*. Cambridge University Press, Cambridge.
25. C. Carstencen and C. Manigrasso (2008). Advancements in the computational calculus of variations. In *Proceedings of the IUTAM Symposium on Variational Concepts with Applications to the Mechanics of Materials*, Bochum, Germany, pp. 29–50.
26. C. Ortner (2011). Nonconforming finite-element discretization of convex variational problems. *IMA J. Numer. Anal.* 31:847–864.
27. A. Miele and R. E. Pritchard (1972). Numerical solutions in the simples problem of the calculus of variations. *SIAM Rev.* 14:385–398.
28. S. A. Yousefi and M. Dehghan (2010). The use of He's variational iteration method for solving variational problems. *Int. J. Comput. Math.* 87:1299–1314.
29. P. C. Rosenbloom (1956). The method of steepest descent. In *Proceedings of Symposia in Applied Mathematics*, The American Mathematical Society, Providence, vol. 6, pp. 127–176.
30. M. L. Stein (1953). On methods for obtaining solutions of fixed end-point problems in the calculus of variations. *J. Res. Natl. Bur. Stand.* 50:277–297.
31. B. T. Polyak (1963). Gradient methods for minimizing functionals. *Zh. Vychisl. Mat. Mat. Fiz.* 3:643–653 (in Russian).
32. M. Gockenbach and C. Liu (2015). Local convergence of Newton's method in the classical calculus of variations. *Optimization* 64:957–980.
33. V. F. Demyanov and G. Sh. Tamaysan (2011). Exact penalty functions in isoperimetric problems. *Optimization* 60:153–177.
34. G. Sh. Tamasyan (2013). Numerical method in problems of calculus of variations for functionals depending on higher order derivatives. *J. Math. Sci.* 188:299–321.
35. V. F. Demyanov and G. Sh. Tamasyan (2014). Direct methods in the parametric moving boundary variational problem. *Numer. Funct. Anal. Optim.* 35:934–961.
36. J. W. Daniel (1971). *The Approximate Minimization of Functionals*. Prentice-Hall, Englewood Cliffs, NJ.
37. M. M. Vainberg (1964). *Variational Methods of the Study of Nonlinear Operators*. Holden-Day, San Francisco.
38. V. F. Demyanov and A. M. Rubinov (1970). *Approximate Methods in Optimization Problems*. American Elsevier, New York.
39. M. M. Vainberg (1972). *Variational Method and the Method of Monotone Operators in the Theory of Nonlinear Equations*. Nauka, Moscow (in Russian).
40. L. V. Kantorovich and G. P. Akilov (1982). *Functional Analysis*, 2nd ed. Pergamon Press, Oxford.
41. J.-P. Penot (2002). On the convergence of descent algorithms. *Comput. Optim. Appl.* 23: 279–284.
42. J. W. Neuberger (1997). *Sobolev Gradients and Differential Equations*. Springer-Verlag, Berlin, Heidelberg.
43. M. V. Dolgopolik and G. Sh. Tamasyan (2014). Method of steepest descent for two dimensional problems of calculus of variations. In *Constructive Nonsmooth Analysis and Related Topics, Springer Optimization and Its Applications* (V. Demyanov, P. M. Pardalos. and M. Batsyn, eds.), Springer Science+Business Media, New York, vol. 87, pp. 101–113.
44. V. F. Demyanov (1994). Exact penalty functions in problems of nonsmooth optimization. *Vestnik St. Petersburg Univ. Math.* 27:16–22.
45. V. F. Demyanov (2005). *Extremum Conditions and the Calculus of Variations*. Vysh Shkola, Moscow (in Russian).
46. M. Z. Nashed (1965). On general iterative methods for the solutions of a class of nonlinear operator equations. *Math. Comput.* 19:14–24.
47. M. Z. Nashed (1971). Differentiability and related properties of nonlinear operators: Some aspects of the role of differentials in nonlinear functional analysis. In *Nonlinear Functional Analysis and Applications* (L. B. Ball, ed.), Academic Press, New York, pp. 102–309.
48. R. A. Adams (1975). *Sobolev Spaces*. Academic Press, New York.

49. L. C. Evans (2010). *Partial Differential Equations*. American Mathematical Society, Providence, RI.
50. G. Leoni (2009). *A First Course in Sobolev Spaces*. American Mathematical Society, Providence, RI.
51. H. Federer (1969). *Geometric Measure Theory*. Springer-Verlag, New York.
52. P. T. Boggs and J. W. Tolle (1995). Sequential quadratic programming. In *Acta Numerica 1995* (A. Iserles, ed.), Cambridge University Press, Cambridge, pp. 1–51.
53. P. Pedregal (2017). Direct numerical algorithm for constrained variational problems. *Numer. Funct. Anal. Optim.* 38:486–506.