*AMO - Advanced Modeling and Optimization, Volume 14, Number 3, 2012*

# OPTIMIZATION FOR SYSTEMS GOVERNED BY PARTIAL DIFFERENTIAL EQUATIONS

VICTOR K. TOLSTYKH*

**Abstract.** This paper develops an algorithm for obtaining the gradient as a Fréchet derivative in optimization problems with Partial Differential Equations (PDE). Proposed algorithm uses the calculus of variations and a direct minimization of the objective functional. The algorithm is based on the rational use of the formalism for the Lagrange multipliers. It does not demand the multipliers for boundary conditions and optimality constraints. There are two examples for parabolic and hyperbolic PDEs. This paper, besides providing some new results, has a didactic orientation. It can be useful for mathematicians and engineers, who intend to solve PDE optimization problems.

**Key words.** Optimization, gradient, PDE

**AMS subject classifications.** Primary 65K10, 49K20; Secondary 49J20, 93B30

**1. Introduction.** There are known three categories of optimization problems for PDE [1]: optimal control, optimal design, and parameter estimation. One of the motives for arising the different categories is a lack of a universal optimization method. In different categories the different optimization methods are sovereign. If unknown variable $u$ depends on time $t$, then one speaks about optimal control and uses the maximum principle [3], [6]. If $u$ depends from space variable $x$, then one often speaks about optimal design and uses necessary conditions and methods based on the calculations of variations and evolutions of this technic [8], [9]. If $u$ is looked for on the basis of experimental data, then one speaks about parameter estimation (identification of a model) and uses the methods type of least square. For non-steady PDE, the unknown variable $u$ can depends on both $t$ and $x$. That erases differences between the categories. Actually, all mathematical statements for optimization problems are the same in principle.

The purpose of this paper is to present all categories of optimization problems in a uniform style. This style can be formed on the basis of non-linear optimization methods with account of technical constraints inqluding PDF. We think that a uniform style (*direct optimization*) can be useful for mathematicians and engineers, who intend to solve diverse PDE optimization problems.

Let us formulate a typical optimization problem (it can be optimal design, optimal control or parameter estimation) for systems described by PDE:

$$\text{minimize} \quad J(u) = \int_\Omega I\left(v\right) d\Omega, \tag{1.1}$$

$$\begin{gathered}\text{subject to } \mathbb{D}\left(\tau, v\right) v + F\left(\tau, v, u\right) = 0, \qquad \tau \in \text{int}\Omega, \\ G\left(\tau, v, \overline{u}\right) = 0, \qquad \tau \in \Gamma \subset \partial\Omega.\end{gathered} \tag{1.2}$$

Here $\tau \equiv \{t, \mathbf{x}\}$ is a time-space variable, differential operator $\mathbb{D}$ of order $k$ and free term $F$ represent a system of $m$ PDEs, which functions within the closed time-space domain $\Omega$, $G$ is a system of initial and boundary conditions on $\Gamma \subset \partial\Omega$, $\partial\Omega$ is a closer of $\Omega$. The state of system (1.2) is $m$-vector-function $v(\tau)$ in linear real space $V^m(\Omega)$. The state can be controlled by $n$-vector-function $u(\tau)$ in linear real space $U^n(\text{int}\Omega)$ i.e., on interior of $\Omega$. For this reason we will call $u$ as the control for all optimization problem. Control $\overline{u}(\tau), \tau \in \overline{S} \subset \Gamma$ is a boundary

*Computer Science Department, Donetsk National University, Universitetskaya, 24, Donetsk 83055, Ukraine (mail@tolstykh.com, http://www.tolstykh.com
***AMO - Advanced Modeling and Optimization. ISSN: 1841-4311)**

piece of $u$. Thus we have the control on $S = \text{int}\Omega \cup \overline{S}$. The quality of controlling is defined by objective functional $J(u)$ on $\Omega$. The integrand $I$ depends explicitly from state $v$.

Optimization problem (1.1), (1.2) can have constraints:

| | | |
|---|---|---|
| $u(\tau) \in [u_{\min}(\tau), u_{\max}(\tau)], \quad \tau \in S$ | - controlling in a convex admissible set (infinite-dimensional parallelepiped); | (1.3a) |
| $\int_S u(\tau) dS = C = const$ | - isoperimetric condition; | (1.3b) |
| $u(\tau) = \{u_i \text{ on } S_i \subset S\}_{i=1\ldots N}$ | - piecewise control; | (1.3c) |
| $v(\tau) \in [v_{\min}(\tau), v_{\max}(\tau)], \quad \tau \in \text{int}\Omega$ | - constrained system state; | (1.3d) |

These constraints are typical and widely distributed for PDE optimization but of course are not comprehensive. Let us take note of the fact that constraint (1.3c) is a condition of existence for a finite-dimensional control $u \in R^N$.

Sometimes we can meet the restrictions in view of equations as connections between the components of vector $v$ and its derivatives. In practice those equations can be found as a special boundary condition $G = 0$ and not optimization constraints actually. We have to differentiate the restrictions, especial conditions and so on to solve the system $\mathbb{D}v + F = 0$ and the constraints to minimize the objective functional $J(u)$.

Optimization problem (1.1)–(1.3) is formulated as the following: we have an object (1.2), the state of its $v$ can be controlled by parameter $u$; the quality of control of object is described by the criterion $J$; for optimizing (minimizing) the criterion $J$ there appear the constraints (1.3), which prevent us to obtain the best state of object. The goal of a mathematician or an engineer is a looking for the parameter $u$, which minimizes $J$.

We shall consider a direct optimization approach to solve the problem (1.1)–(1.3). This approach implies step-by-step minimization of $J(u)$ by means of derivative $J'_u$. In so doing we shall not make a preliminary discretization and other transforms of PDE and $J$, we shall not substitute function $u(\tau)$ by a vector $u = \{u_i\}$ of coefficients in series, shall not use the necessary conditions for optimality (maximum principle, dynamic programming, Karush-Kuhn-Tucker system...). In short, we do not replace optimization problem (1.1)–(1.3) similar to it, yet another. Discretization of problem (1.1)–(1.3) at finite-dimensional optimization problem is not advisable [10]: "...certain infinite-dimensional effects are better retained then in the discretized version". This is absolutely correctly. You can in it be convinced if look examples of infinite- and finite-dimensional optimization for thermal processes "Minimizing the Functionals" in [13].

Difficulties in direct minimizing $J(u)$ consists of:

1. functional $J$ depends from $u$ *implicitly*;
2. argument $u$ of $J$ is *infinite*-dimensional i.e., functions.

To overcome the first difficulty, and to compute derivative $J'_u$, there are used various methods of adjoint variables (Lagrange's multipliers) . This procedure is often called "adjoint approach" [7]. To overcome the second difficulty we use heuristic infinite-dimensional optimization methods [4], [5], [11], [12], [13], [14] based on derivative $J'_u$.

In this paper we consider some view point (method) to compute the gradient as a Fréchet derivative $J'_u \equiv \nabla J \equiv \nabla J(\tau; u)$ for problem (1.1)–(1.3). This derivative is often used by wide range of optimization and analysis methods. We will use it for direct optimization of PDE.

**2. Searching for the gradient in PDE optimization. General algorithm.** Proposed algorithm based on the calculus of variations and functional analysis and it consists of the following eight steps.

**2.1. Allowance for constraints.** First, the constraints for control (1.3a)–(1.3c) we put aside. We will implement its by optimization methods later. Second, we are going to implement a constraint on the state (1.3d) of the traditional way in the form of penalty functional. So the objective functional becomes

$$J(u) = \int_{\Omega} I(v)d\Omega + \kappa \int_{\text{int}\Omega} P(v)d\Omega \to \min, \tag{2.4}$$

where $\kappa$ is a weighting coefficient, $P$ is a penalty function from $v$ with parameters $v_{\min}, v_{\max}$. It can be inner (barrier), outer or mixed penalty. Now we have unconstrained optimization problem (1.2), (2.4) instead of (1.1)–(1.3).

**2.2. Linearizing the optimization problem.** Let we have an initial point $u$ to start optimization and let $J(u)$ be a Fréchet differentiable at $u$. According to PDE (1.2) variation $\delta u \in U^n$ leads to variation $\delta v \in V^m$, from that we have following variation of PDE at point $u$:

$$\delta\left(\mathbb{D}v + F\right) = \mathbb{A}\delta v + \mathbb{B}\delta u = 0 \in V^m(\text{int}\Omega), \tag{2.5}$$

where $\mathbb{A}$, $\mathbb{B}$ are linear operators over $V^m$ and $U^n$ respectively. The boundary conditions we will consider later.

For instance, if differential operator $\mathbb{D} = A(x,v)\dfrac{\partial^k \cdot}{\partial x^k}$, where $A$ is $m \times m$-matrix, then linear operator $\mathbb{A}$ is

$$\mathbb{A}\cdot\ = A\frac{\partial^k \cdot}{\partial x^k} + \sum_{i=1}^{m} \frac{\partial A}{\partial v_i} \cdot\big|_i \frac{\partial^k v}{\partial x^k} + F'_v \quad : \quad V^m(\text{int}\Omega) \to V^m(\text{int}\Omega).$$

Linear operator $\mathbb{B}$ is

$$\mathbb{B}\cdot\ = F'_u \cdot \quad : \quad U^n(\text{int}\Omega) \to V^m(\text{int}\Omega).$$

Operators $F'_v$ and $F'_u$ are the matrixes of derivatives $m \times m$ and $m \times n$ respectively.

The variation of objective (2.4) at $u$ we can represent as inner product:

$$\delta J(u) = \langle I'_v, \delta v\rangle_{V^*(\Omega)} + \langle \kappa P'_v, \delta v\rangle_{V^*(\text{int}\Omega)} \in R, \tag{2.6}$$

where linear functional $I'_v \in V^*(\Omega)$ and $P'_v \in V^*(\text{int}\Omega)$, upper asterisk denotes a conjugacy, $R$ is a real number space. We notice, if $m > 1$, then in inner product $\langle \cdot, \cdot \rangle$ the first vector is a row-vector and the second vector is a column-vector. Variation $\delta J$ in (2.6) we can treat as a sum of values of linear functionals $I'_v$, $P'_v$ at element $\delta v$.

The expressions (2.5) and (2.6) belong to different spaces $V^m(\text{int}\Omega)$ and $R$. We have to unit these expressions, but first we need to make the necessary converting.

**2.3. Converting the linearized PDE.** Let us write PDE (2.5) and variation (2.6) in like linear spaces, namely in $R$. To convert (2.5) into $R$ let us introduce a linear functional $\widetilde{f} \in V^*(\Omega)$ and write its value at element $\delta\left(\mathbb{D}v + F\right)$:

$$\left\langle \widetilde{f}, \delta\left(\mathbb{D}v + F\right)\right\rangle_{V^*(\text{int}\Omega)} = \left\langle \widetilde{f}, \mathbb{A}\delta v\right\rangle_{V^*(\text{int}\Omega)} + \left\langle \widetilde{f}, \mathbb{B}\delta u\right\rangle_{V^*(\text{int}\Omega)} = 0 \in R. \tag{2.7}$$

Equation (2.7) breaks the tie between all components of vector $\delta v$ and $\delta u$ because of vague of $m$-vector-function $\widetilde{f}(\tau)$. If system (1.2) had $n$ independent parameters, then equation (2.7) has $m+n$ independent parameters $\{\delta v_i, \delta u_j\}_{i=1\ldots m, j=1\ldots n}$. For this reason we can convert equation (2.7) so that to distinguish the operators from new independent parameters. This converting looks like a Lagrange identity. So we have:

$$\begin{aligned}
&\left\langle \widetilde{f}, \mathbb{A}\delta v \right\rangle_{V^*(\mathrm{int}\Omega)} + \left\langle \widetilde{f}, \mathbb{B}\delta u \right\rangle_{V^*(\mathrm{int}\Omega)} \\
&= \left\langle \mathbb{A}^*\widetilde{f}, \delta v \right\rangle_{V^*(\mathrm{int}\Omega)} + \left\langle \mathbb{B}^*\widetilde{f}, \delta u \right\rangle_{U^*(\mathrm{int}\Omega)} \\
&\quad + \left\langle a_0^*\widetilde{f}, \delta v \right\rangle_{V^*(\partial\Omega)} + \sum_j \left\langle a_{1_j}^*\widetilde{f}, \frac{\partial \delta v}{\partial \tau_j} \right\rangle_{V^*(\partial\Omega)} + \ldots = 0 \in R,
\end{aligned} \tag{2.8}$$

where $\mathbb{A}^*$, $\mathbb{B}^*$ are adjoint operators on $\mathrm{int}\Omega$ and set $\{a_{k-1}^*\}_{k=1,2\ldots}$ are adjoint operators on boundary $\partial\Omega$. The lower line is a by-product on $\partial\Omega$ after application the Green formula. Dots mean the boundary terms, which contain the high derivatives for the case $k > 2$. To understand the expression (2.8) see example in the appendix.

**2.4. Joining the problem.** Expressions (2.6) and (2.8) were written in the same space $R$. Because the conjugate spaces are linear, we can unite linear functionals in the spaces $V^*$, $U^*$ on the elements $\delta v$, $\delta u$. So we have:

$$\begin{aligned}
\delta J(u) = &\left\langle \mathbb{A}^*\widetilde{f} + I_v' + \kappa P_v', \delta v \right\rangle_{V^*(\mathrm{int}\Omega)} + \left\langle \mathbb{B}^*\widetilde{f}, \delta u \right\rangle_{U^*(\mathrm{int}\Omega)} \\
&+ \left\langle a_0^*\widetilde{f} + I_v', \delta v \right\rangle_{V^*(\partial\Omega)} + \sum_j \left\langle a_{1_j}^*\widetilde{f}, \frac{\partial \delta v}{\partial \tau_j} \right\rangle_{V^*(\partial\Omega)}.
\end{aligned} \tag{2.9}$$

Here we have limited the boundary terms by the problem case $k \le 2$.

**2.5. Taking into account the initial-boundary conditions.** Two ways can be derived for take into account the initial-boundary conditions.

First way. We have to do all steps described above: linearize the initial-boundary conditions, introduce the additional linear functionals on $\Gamma$ and receive the adjoint operators at independent variations $\delta v$ and $\delta\overline{u}$. After this we can unify obtained initial-boundary elements with problem (2.9), what is more precisely – with bottom line in problem (2.9). This line contains boundary elements on closer $\partial\Omega$ obtained from interior of $\Omega$.

Second way is more comfortable. We recommend using it. On the one hand the expression (2.9) contains independent variations $\delta v_i$ and may be its derivatives (e.g., $\dfrac{\partial \delta v_i}{\partial \tau_j}$ for case $k = 2$) on $\partial\Omega$. On the other hand, from the initial-boundary conditions of primal problem (1.2), we can get the equations for variations:

$$\delta G = G_v'\delta v + G_u'\delta\overline{u} + \sum_j G_{v_j'}' \frac{\partial \delta v}{\partial \tau_j} = 0 \in V^m(\Gamma). \tag{2.10}$$

Here the variations are not independent. We need to find variations $\delta v$ and $\dfrac{\partial \delta v}{\partial \tau_j}$ on the corresponding pieces of $\Gamma$ from equations (2.10) and substitute its into bottom line (2.9).

The solving of equations (2.10) and substituting in (2.9) in general terms is a cumbersome and useless procedure. So let us consider some typical examples for the case of $\Omega = [t_{initial}, t_{final}] \times [x_{left}, x_{right}]$.

Let we know an initial condition

$$v = v_{initial} \in V^m(\Gamma_a = t_{initial} \times x).$$

In this case from (2.10) on $\Gamma_a$ we obtain $G'_v = \mathbf{1}$ (unit matrix), $G'_u = G'_{v'_x} = 0$, i.e., $\delta v = 0$ on $\Gamma_a \subset \partial\Omega$. This denotes, that in first item of bottom line (2.9), the all adjoint elements on $\Gamma_a$ disappear. On opposite side of time, i.e., on $\Gamma^*_a = t_{final} \times x$ the analogous adjoint elements still stay. We can say that the initial condition is "turned" in adjoint space.

The analogous "turn" effect holds for boundary conditions too. However, here can be additional peculiarity – the presence of boundary control $\overline{u}$. For example, let for $m = n = 1$ exists a boundary condition

$$\frac{\partial v}{\partial x} = \overline{u} \in L_2\left(\Gamma_b = t \times x_{left} = \overline{S}\right).$$

In this case the operators $G'_v = 0$, $G'_u = -1$, $G'_{v'_x} = 1$. This denotes, that in second term of bottom line (2.9), the derivative $\dfrac{\partial \delta v}{\partial x}$ on $\Gamma_b$ disappears and variation $\delta\overline{u}$ on $\overline{S} \subset \partial\Omega$ appears. On opposite side of $x$, i.e., on $\Gamma^*_b = t \times x_{right}$ the all adjoint elements still stay.

The third example. Let we have a one boundary equation ($m_1 = 1$) for a vector-state $v \in V^m(\Gamma_c = t \times x_{left})$, $m = 2$:

$$av_1 + bv_2 = 0 \quad \text{on } \Gamma_c, \tag{2.11}$$

where $m_1$ is a number of equations on $\Gamma_c$. In this case $G'_u = G'_{v'_x} = 0$ and matrix $G'_v = (a, b)$, i.e., $G'_v$ is not a square matrix of $m \times m$. Thus we can not find a vector $\delta v$ from condition (2.11). Therefore, we express one component of variation by other – $\delta v_1 = -\frac{b}{a}\delta v_2$ – and substitute in (2.9), where we obtain a vector $\delta v = \left(-\frac{b}{a}\delta v_2, \delta v_2\right)$ on $\Gamma_c$. In so doing, we remove adjoint elements in (2.9) at $\delta v_1$ on $\Gamma_c$ and keep analogous elements at $\delta v_1$ on opposite set $t \times x_{right}$. For third example $\Gamma^*_c$ is the uniting $\{t \times x_{left}\} \cup \{t \times x_{right}\}$. In general case, for each $m$ with $m_1 \le m$, we will have $m - m_1$ adjoint elements (components) in $\delta J$ at variations $\delta v_{i=m_1+1,\ldots,m}$ on $\Gamma_c$. If $m_1 = m$, then we will have nothing in $\delta J$ on $\Gamma_c$, for this case $\Gamma^*_c = t \times x_{right}$.

We finally obtain the following general view of the first variation $\delta J$:

$$\begin{aligned}\delta J = \left\langle \mathbb{A}^*\widetilde{f} + I'_v + \kappa P'_v, \delta v\right\rangle_{V^*(\mathrm{int}\Omega)} + \left\langle \mathbb{B}^*\widetilde{f}, \delta u\right\rangle_{U^*(\mathrm{int}\Omega)} \\ + \left\langle a^*_0\widetilde{f} + I'_v, \delta v\right\rangle_{V^*(\Gamma^*)} + \sum_j \left\langle a^*_{1_j}\widetilde{f}, \frac{\partial \delta v}{\partial \tau_j}\right\rangle_{V^*(\Gamma^*_j)} \\ + \left\langle g_0\left(a^*_0\widetilde{f} + I'_v\right) + \sum_j g_{1_j} a^*_{1_j}\widetilde{f}, \delta\overline{u}\right\rangle_{U^*(\overline{S})}.\end{aligned} \tag{2.12}$$

Here $\Gamma^* \subset \partial\Omega$ is a set of the "turned" initial-boundary conditions at $\delta v$, $\Gamma^*_j \subset \partial\Omega$ are the analogues sets at $\dfrac{\partial \delta v}{\partial \tau_j}$. Operator $g_0 \neq 0$ if a boundary condition for state $v$ contains the control $\overline{u}$, and $g_{1_j} \neq 0$ if a boundary condition for derivatives $\dfrac{\partial v}{\partial \tau_j}$ contains $\overline{u}$.

**2.6. Extracting the gradient.** First variation $\delta J$ (2.12) has $m + n$ independent parameters $\delta v_{i=1\ldots m}$ and $\delta u_{j=1\ldots n}$. So we can represent $\delta J$ as the following

$$\delta J = \langle \nabla_v J, \delta v\rangle_{V^*(\mathrm{int}\Omega + \Gamma^* + \Gamma^*_j)} + \langle \nabla_u J, \delta u\rangle_{U^*(S)},$$

where $\nabla_v J$ and $\nabla_u J$ are the components of a gradient

$$\nabla J = (\nabla_v J, \nabla_u J) \in V^* \times U^*.$$

This form of the gradient is not comfortable in practice because of components $\nabla_v J$. On this step it is expedient to restore the connection between variation $\delta v$ and $\delta u$. Let us take $\widetilde{f} = f$ such that $\nabla_v J = 0$. We receive

$$\begin{array}{l} \mathbb{A}^* f + I_v' + \kappa P_v' = 0 \quad \text{on int}\Omega; \\ a_0^* f + I_v' = 0 \quad \text{on } \Gamma^*, \qquad a_{1_j}^* f = 0 \quad \text{on } \Gamma_j^*. \end{array} \tag{2.13}$$

Equations (2.13) is named the adjoint problem.

Finally we receive the gradient:

$$\nabla J(\tau; u) = \left\{ \begin{array}{l} \mathbb{B}^* f \quad \text{on int}\Omega, \\ g_0 \left( a_0^* f + I_v' \right) + \sum\limits_j g_{1_j} a_{1_j}^* f \quad \text{on } \overline{S} \end{array} \right\} \in U^*(S). \tag{2.14}$$

The gradient depends from linear functional $f$. To compute the gradient at point $u$ we have to solve primal problem (1.2) and adjoint problem (2.13).

**2.7. Analyzing the adjoint problem.** Adjoint problem (2.13) is a linear PDE relative to $f$. Consider the type of adjoint problem (parabolic, hyperbolic...) relative to a primal problem. For that we are going to compare the operator $\mathbb{D}$ and $\mathbb{A}^*$. First consider a stationary primal PDE with operator

$$\mathbb{D} = A_{11} \frac{\partial^2}{\partial x_1^2} + 2A_{12} \frac{\partial^2}{\partial x_1 x_2} + A_{22} \frac{\partial^2}{\partial x_2^2},$$

where $A_{ij}(x_1, x_2, v)$ are real functions on int$\Omega$. Define $d = A_{11}A_{22} - A_{12}^2$. It is known if $d < 0$, then PDE is hyperbolic, if $d = 0$, then PDE is parabolic, and if $d > 0$, then PDE is elliptic. From the example of Appendix we can see that adjoint operator $\mathbb{A}^*$ is

$$\mathbb{A}^* = \mathbb{D} + ...$$

Dots denote the terms with derivatives fewer then 2. Since the type of PDE is defined by coefficients to largest derivatives therefore primal and adjoint PDEs have the same type.

Consider non-stationary PDEs.

Parabolic PDE:

$$\mathbb{D} = \frac{\partial}{\partial t} - A \frac{\partial^2}{\partial x^2},$$

where function $A(t, x, v) > 0$ on int$\Omega$. According to Appendix the adjoint operator $\mathbb{A}^*$ is

$$\mathbb{A}^* = -\frac{\partial}{\partial t} - A \frac{\partial^2}{\partial x^2} + ...$$

Dots denote the terms with space derivatives fewer then 2. We have a negative sign before the time derivative. For example, if for primal PDE an initial condition happens to be the case for $\Gamma_0 = t_{initial} \times [x_a, x_b]$, then for adjoin PDE an initial condition appears on $\Gamma_0^* = t_{final} \times [x_a, x_b]$. So we have to solve an adjoint PDE in inverse time direction, that is, for $dt < 0$. That fits to negative time derivative in $\mathbb{A}^*$ very well. Therefore adjoint PDE has the same parabolic type with consideration of inverse time.

Hyperbolic PDE, case $k = 2$, $m = 1$:

$$\mathbb{D} = \frac{\partial^2}{\partial t^2} - A\frac{\partial^2}{\partial x^2},$$

where function $A(t, x, v) > 0$ on int$\Omega$. Formally this PDE coincides with the first example with $d < 0$. Adjoint operator

$$\mathbb{A}^* = \frac{\partial^2}{\partial t^2} - A\frac{\partial^2}{\partial x^2} + ...$$

Dots denote the terms with derivatives fewer then 2. Both PDEs have the same characteristics $\frac{dx}{dt} = \pm\sqrt{A}$, but remember the adjoint problem always has inverse time. If for primal problem a characteristic transfers a disturbance from $x_{left}$ to $x_{right}$ at $t \to t_{final}$, then for adjoint problem the same characteristic transfers a disturbance from $x_{right}$ to $x_{left}$ at $t \to t_{initial}$. That fits to adjoint boundary conditions. Really, for instance, if a boundary condition happens to be the case for $\Gamma_0 = x_{left} \times (t_{initial}, t_{final})$, then an adjoint boundary condition happens to be the case for $\Gamma_0^* = x_{right} \times (t_{initial}, t_{final})$.

Hyperbolic PDEs, case $k = 1$, $m > 1$:

$$\mathbb{D} = \frac{\partial}{\partial t} + A\frac{\partial}{\partial x},$$

where $m \times m$-matrix-function $A(t, x, v)$ on int$\Omega$ has $m$ real eigenvalues $\lambda_{1...m}$ . Adjoint operator

$$\mathbb{A}^* = -\left(\frac{\partial}{\partial t} + A^T\frac{\partial}{\partial x} + ...\right).$$

Dots denote the terms without derivatives. The eigenvalues of matrixes $A$ and $A^T$ are the same ones. Therefore both PDEs have the same characteristics $\frac{dx}{dt} = \lambda_{1...m}$ that is similarly for the previous case $k = 2$.

So, primal and adjoint problems have the same type and characteristics. The adjoint problem is always linear. If PDE is non-stationary, then adjoint problem has inverse time.

If we look at adjoint problem (2.13) we can see that the sources of disturbance are the linear functionals $I'_v$, $P'_v$. If $I'_v = 0$ and $P'_v = 0$, then the adjoint problem has trivial solution $f = 0$ on $\Omega$ and according to (2.14) the gradient $\nabla J = 0$ on $S$. The trivial solution is a necessary condition for unconstrained $\min J(u)$.

**2.8. Taking into account the constraints.** Constraints (1.3a)–(1.3c) for control we have excluded on step 1 in section 2.1. Now we are going to implement constraints by direct optimization methods.

Unconstrained direct minimization for $J(u)$ has iterative methods, which look as [14]

$$u^{K+1}(\tau) = u^K(\tau) + b^K d(\tau; u^K), \quad b^K > 0,\ K = 0, 1, ... \tag{2.15}$$

where minimization direction $d(\tau; u^K)$ is computed through the gradient $\nabla J(\tau; u^K)$ by various methods [4], [5], [11], [12], [13]. In the case of steepest descent method $d = -\nabla J$.

The constraint (1.3a) can be easily implemented on the base of (2.15) by projection of point $u^{K+1}$onto admissible parallelepiped $[u_{\min}(\tau), u_{\max}(\tau)]$. So we need to make additional correction for unconstrained method (2.15):

$$\begin{cases} \text{if } u^{K+1}(\tau) < u_{\min}(\tau), \text{ then } u^{K+1}(\tau) \leftarrow u_{\min}(\tau), \\ \text{if } u^{K+1}(\tau) > u_{\max}(\tau), \text{ then } u^{K+1}(\tau) \leftarrow u_{\max}(\tau). \end{cases} \tag{2.16}$$

Isoperimetric constraint (1.3b) is implemented by method [12] additionally to (2.15):

$$u^{K+1}(\tau) \leftarrow u^{K+1}(\tau) - \frac{C^{K+1} - C}{\mathbf{S}}, \tag{2.17}$$

where $C^{K+1} = \int_S u^{K+1}(\tau)dS, \quad \mathbf{S} = \int\limits_S dS.$

If the control is a piecewise (1.3c), then we receive $N$-dimensional minimization based on the gradient

$$\nabla J = \left\{ \int_{S_i} \nabla J(\tau; u)dS \right\}_{i=1 \ldots N} \in R^N$$

for the following method

$$u_i^{K+1} = u_i^K + b^K d_i(u^K), \quad i = 1, ..., N. \tag{2.18}$$

General algorithm for direct minimization of the objective (1.1) subject to PDE (1.2) with constraints (1.3) is finished. Let us consider two examples to explain the general algorithm.

**3. Example 1. Parabolic PDE.** We consider a one-dimensional process of heat transfer into the chemical reactor through a steel wall. Corresponding PDE has the view:

$$\mathbb{D}v + F = C\rho\frac{\partial T}{\partial t} - \lambda\frac{\partial^2 T}{\partial x^2} = 0, \quad \Omega = [t_0, t_1] \times [x_a, x_b], \tag{3.1}$$

where $T(t, x)$ is a temperature in wall, $C, \rho, \lambda$ is a heat capacity, a density and a heat conducting coefficient accordingly. The initial-boundary conditions on $\Gamma = \Gamma_0 + \Gamma_a + \Gamma_b$ are the following:

$$G = \begin{cases} T(t_0, x) - T_0 \quad = 0 & \text{on } \Gamma_0 = t_0 \times [x_a, x_b], \\ \lambda\dfrac{\partial T}{\partial x} - q \quad = 0 & \text{on } \Gamma_a = (t_0, t_1) \times x_a, \\ \lambda\dfrac{\partial T}{\partial x} - \overline{u} \quad = 0 & \text{on } \Gamma_b = (t_0, t_1) \times x_b = S. \end{cases}$$

The chemical reaction generates heat flux $q(t)$, but the reaction should be with a nominal temperature $T_*(t)$. So, we have to find the heat flux $\overline{u}(t)$ at the boundary $x_b$, which gives a minimum to the following objective functional at $x_a$ :

$$J(u) = \int_{\Gamma_a} I\,(T)dt, \quad I(T) = (T - T_*)^2. \tag{3.2}$$

The category of this optimization problem is a synthesis of optimal control.

General definitions in previous sections here has a view: state $v = T(t, x) \in L_2(\Omega)$, operator $\mathbb{D} = C\rho\frac{\partial}{\partial t} - \lambda\frac{\partial^2}{\partial x^2}$, free term $F = 0$, control $\overline{u}(t) \in L_2(S)$, dimensions $m = n = 1$, the constraints for a control are absent.

For the first variation of $\delta J$ in (2.9) we have the operators:

$$\mathbb{A}^* = -C\rho\frac{\partial}{\partial t} - \lambda\frac{\partial^2}{\partial x^2} \quad \text{and} \quad \mathbb{B}^* = 0 \quad \text{on int}\Omega,$$

$$a_0^* = C\rho\Big|_{\Gamma_0}^{\Gamma_1} + \lambda\frac{\partial}{\partial x}\Big|_{\Gamma_a}^{\Gamma_b} \qquad \text{— at } \delta T,$$

$$a_1^* = -\,\lambda\Big|_{\Gamma_a}^{\Gamma_b} \qquad \text{— at } \frac{\partial \delta T}{\partial x}.$$

Linear functional $I'_v = 2\,(T - T_*)|_{\Gamma_a}$. Initial-boundary conditions $G = 0$ give us the variations:

$$\delta\, T|_{\Gamma_0} = 0, \quad \left.\frac{\partial \delta T}{\partial x}\right|_{\Gamma_a} = 0, \quad \left.\frac{\partial \delta T}{\partial x}\right|_{\Gamma_b = S} = \lambda^{-1}\delta\overline{u}.$$

From last equation on $S$ we have in (2.12) $g_0 = 0$, $g_1 = \lambda^{-1}$.

The first variation (2.12) here takes the form:

$$\begin{aligned}\delta J = \left\langle \mathbb{A}^*\widetilde{f}, \delta T\right\rangle_{L_2(\mathrm{int}\Omega)} + \left\langle C\rho\widetilde{f}, \delta T\right\rangle_{L_2(\Gamma_1)} \\ + \left\langle -\lambda\frac{\partial\widetilde{f}}{\partial x} + I'_v, \delta T\right\rangle_{L_2(\Gamma_a)} + \left\langle \lambda\frac{\partial\widetilde{f}}{\partial x}, \delta T\right\rangle_{L_2(\Gamma_b)} \\ + \left\langle -\widetilde{f}, \delta\overline{u}\right\rangle_{L_2(S)}.\end{aligned}$$

From this follows the adjoint problem (2.13):

$$\begin{aligned}\mathbb{A}^* f \quad &= -C\rho\frac{\partial f}{\partial t} - \lambda\frac{\partial^2 f}{\partial x^2} = 0 \quad \text{on int}\Omega, \\ f \quad &= 0 \quad \text{on } \Gamma_0^* = \Gamma_1 = t_1 \times [x_a, x_b], \\ \lambda\frac{\partial f}{\partial x} \quad &= 0 \quad \text{on } \Gamma_a^* = \Gamma_b, \\ \lambda\frac{\partial f}{\partial x} \quad &= 2(T - T_*) \quad \text{on } \Gamma_b^* = \Gamma_a.\end{aligned} \tag{3.3}$$

Gradient (2.14) has the form:

$$\nabla J(t; u) = -f \in L_2(S). \tag{3.4}$$

Notice, that $L_2^* = L_2$ for real functions.

Adjoint problem (3.3) is a parabolic PDE with consideration of inverse time. Solution $f$ on right bound $x_b$ gives us the gradient. The source of disturbance $I'_v = 2(T - T_*)$ is lumped at the left bound $x_a$. If $I'_v = 0$ on $x_a$, then we have the trivial adjoint state $f = 0$ on $x_b$, wich tells us that the corresponding value of $u$ satisfies to the necessary condition for optimality, here $\|\nabla J\|_{L_2(S)} = 0$.

**4. Example 2. Hyperbolic PDEs.** Consider the second optimization problem about identifying the friction coefficient $u(x)$ for a water flow in open channel [2]:

$$\mathbb{D}v + F = \left\{\begin{array}{l}\dfrac{\partial Q}{\partial t} + 2w\dfrac{\partial Q}{\partial x} + b(c^2 - w^2)\dfrac{\partial H}{\partial x} + F_{fr} - gsi = 0, \\ \dfrac{\partial H}{\partial t} + \dfrac{1}{b}\dfrac{\partial Q}{\partial x} - \dfrac{q}{b} = 0, \quad \Omega = [t_0, t_1] \times [x_a, x_b]\,,\end{array}\right. \tag{4.1}$$

where $Q(t, x)$ and $H(t, x)$ are water discharge and depth, $w = Q/s$ is a velocity of a flow, $s(x, H)$ and $b(x, H) = \partial s/\partial H$ are across sectional area and upper width of flow, $c = \sqrt{gs/b}$ is a velocity of small disturbances, $R(x, H)$ is a hydraulic radius, $i(x)$ is an inclination of a channel bottom, $q(x)$ is a lateral flow, $F_{fr} = F_{fr}(Q, H, u)$ is a friction term between a water and a bed of channel.

Initial-boundary conditions on $\Gamma$ are the following:

$$G = \begin{cases} Q(t_0,x) - Q_0(x) & = 0, \quad H(t_0,x) - H_0(x) = 0 & \text{on } \Gamma_0 = t_0 \times [x_a, x_b], \\ H(t,x_a) - H_a(t) & = 0 & \text{on } \Gamma_a = (t_0,t_1) \times x_a, \\ Q - \chi(H - H_\chi) & = 0 & \text{on } \Gamma_b = (t_0,t_1) \times x_b. \end{cases} \tag{4.2}$$

where $\chi$, $H_\chi$ are parameters of a hydro-technical equipment at the right end of a channel. We notice, the equation on $\Gamma_b$ is not a constraint for control, it is only a boundary condition for PDEs (4.1).

The objective functional here is

$$J(u) = \frac{1}{L}\int_{\text{int}\Omega} I(H)\,dtdx + \int_{\Gamma_b} I\,(H)dt, \quad I(H) = (H - H_*)^2, \tag{4.3}$$

where $H_*$ is an experimental observed depth, $L$ is a length of a channel.

General definitions here has a view: vector-state $v = (Q,H) \in V^m(\Omega) = L_2^m(\Omega)$, $m = 2$, operator $\mathbb{D} = \dfrac{\partial}{\partial t} + A\dfrac{\partial}{\partial x}$, where matrix $A(x,v) = \begin{pmatrix} 2w & b(c^2 - w^2) \\ 1/b & 0 \end{pmatrix}$ with eigenvalues $\lambda_{1,2} = c \pm w$, free term is a column-vector $F = (F_{fr} - gsi, -q/b)$, control $u(x) \in U^n(S) = L_2(S)$, $S = (x_a, x_b)$, $n = 1$. The constraints for a control are absent.

To receive the gradient $\nabla J$ we are going to start from step 2 (section 2.2). We need to linearize PDEs (4.1) and objective (4.3):

$$\mathbb{A}\delta v = \frac{\partial \delta v}{\partial t} + A\frac{\partial \delta v}{\partial x} + \left(\frac{\partial A}{\partial Q}\delta Q + \frac{\partial A}{\partial H}\delta H\right)\frac{\partial v}{\partial x} + F'_v\delta v \in L_2^2(\text{int}\Omega), \tag{4.4}$$

$$\mathbb{B}\delta u = F'_u\delta u \in L_2^2(\text{int}\Omega), \qquad \mathbb{B} \text{ is } 2\times 1 \text{ matrix (a colomn-vector)}, \tag{4.5}$$

$$\delta J(u) = \frac{1}{L}\left\langle I'_v, \delta v\right\rangle_{L_2(\text{int}\Omega)} + \left\langle I'_v, \delta v\right\rangle_{L_2(\Gamma_b)}, \quad I'_v = (0, 2[H - H_*]).$$

At step 3 (section 2.3) with vector-function $\widetilde{f} = (f_1, f_2) \in V^*(\Omega) = L_2^2(\Omega)$ we have in (2.8) the following adjoint operators:

$$\mathbb{A}^* = -\frac{\partial}{\partial t} - A^T\frac{\partial}{\partial x} - \frac{\partial A^T}{\partial x} + \mathbb{C}^* + F'^T_v,$$

$$\mathbb{B}^* = F'^T_u = \left(\frac{\partial F_{fr}}{\partial u}, 0\right)^T \quad \text{is a row-vector.}$$

Here matrix $\mathbb{C}^* = \begin{pmatrix} \left(\dfrac{\partial A}{\partial Q}\dfrac{\partial v}{\partial x}\right)^T \\ \left(\dfrac{\partial A}{\partial H}\dfrac{\partial v}{\partial x}\right)^T \end{pmatrix}$. Let us denote

$$
\begin{aligned}
\mathbb{C}_1^* &= \mathbb{C}^* - \frac{\partial A^T}{\partial x} + F_v'^T \\
&= \begin{pmatrix} \left(\dfrac{\partial A_{12}}{\partial Q} - \dfrac{\partial A_{11}}{\partial H}\right)\dfrac{\partial H}{\partial x} + \dfrac{\partial F_1}{\partial Q} & \left(\dfrac{\partial A_{22}}{\partial Q} - \dfrac{\partial A_{21}}{\partial H}\right)\dfrac{\partial H}{\partial x} + \dfrac{\partial F_2}{\partial Q} \\ \left(\dfrac{\partial A_{11}}{\partial H} - \dfrac{\partial A_{12}}{\partial Q}\right)\dfrac{\partial Q}{\partial x} + \dfrac{\partial F_1}{\partial H} & \left(\dfrac{\partial A_{21}}{\partial H} - \dfrac{\partial A_{22}}{\partial Q}\right)\dfrac{\partial Q}{\partial x} + \dfrac{\partial F_2}{\partial H} \end{pmatrix} \\
&\qquad - \left.\frac{\partial A^T}{\partial x}\right|_{v=const} \\
&= \begin{pmatrix} 2\dfrac{w}{s}\left.\dfrac{\partial s}{\partial x}\right|_H + \dfrac{\partial F_{fr}}{\partial Q} & \dfrac{1}{b^2}\dfrac{\partial b}{\partial x} \\ w^2\left.\dfrac{\partial b}{\partial x}\right|_H - \dfrac{b}{s}\left(c^2+2w^2\right)\left.\dfrac{\partial s}{\partial x}\right|_H + \dfrac{\partial F_{fr}}{\partial H} - gbi & \dfrac{1}{b^2}\left(q - \dfrac{\partial Q}{\partial x}\right)\dfrac{\partial b}{\partial H} \end{pmatrix}.
\end{aligned}
$$

So, we have

$$\mathbb{A}^* = -\frac{\partial}{\partial t} - A^T\frac{\partial}{\partial x} + \mathbb{C}_1^*.$$

The boundary adjoint operators in (2.8) are the following:

$$
\begin{aligned}
a_0^* &= \left.\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right|_{\Gamma_0}^{\Gamma_1} + A^T\Big|_{\Gamma_a}^{\Gamma_b} \qquad \text{— at } \delta v, \\
a_1^* &= 0
\end{aligned}
$$

At step 5 (section 2.5) we need to take into account the variations of initial-boundary conditions (4.2). We have:

$$\delta Q = \delta H = 0 \quad \text{on } \Gamma_0, \qquad \delta H = 0 \quad \text{on } \Gamma_a$$

and from the boundary equation on $\Gamma_b$ of type (2.11) we have:

$$\delta Q = \chi \delta H \quad \text{on } \Gamma_b.$$

The first variation $\delta J$ in (2.12) becomes:

$$
\begin{aligned}
\delta J = &\left\langle \mathbb{A}^*\widetilde{f} + \frac{1}{L}I_v', \delta v\right\rangle_{L_2^2(\mathrm{int}\Omega)} + \left\langle \mathbb{B}^*\widetilde{f}, \delta u\right\rangle_{L_2(\mathrm{int}\Omega)} \\
&+ \left\langle \widetilde{f}, \delta v\right\rangle_{L_2^2(\Gamma_1)} + \left\langle A^T\widetilde{f} + I_v', (\chi\delta H, \delta H)\right\rangle_{L_2^2(\Gamma_b)} + \left\langle -A^T\widetilde{f}, (\delta Q, 0)\right\rangle_{L_2^2(\Gamma_a)}.
\end{aligned}
\tag{4.6}
$$

So the adjoint problem (2.13) takes the view:

$$\mathbb{A}^* f + \frac{1}{L} I_v' = \begin{cases} -\dfrac{\partial f_1}{\partial t} - 2w\dfrac{\partial f_1}{\partial x} - \dfrac{1}{b}\dfrac{\partial f_2}{\partial x} \\ \qquad + \left[ 2\dfrac{w}{s}\left.\dfrac{\partial s}{\partial x}\right|_H + \dfrac{\partial F_{fr}}{\partial Q} \right] f_1 + \dfrac{1}{b^2}\dfrac{\partial b}{\partial x} f_2 = 0, \\ -\dfrac{\partial f_2}{\partial t} - b(c^2 - w^2)\dfrac{\partial f_1}{\partial x} \\ \qquad + \left[ w^2 \left.\dfrac{\partial b}{\partial x}\right|_H - \dfrac{b}{s}\left(c^2 + 2w^2\right)\left.\dfrac{\partial s}{\partial x}\right|_H + \dfrac{\partial F_{fr}}{\partial H} - gbi \right] f_1 \\ \qquad\quad + \dfrac{1}{b^2}\left( q - \dfrac{\partial Q}{\partial x} \right)\dfrac{\partial b}{\partial H} f_2 + \frac{2}{L}\left(H - H_*\right) = 0 \end{cases} \tag{4.7}$$

with initial-boundary conditions:

$$\begin{aligned} f_1 = f_2 &= 0 \quad \text{on } \Gamma_0^* = \Gamma_1, \\ \left( 2wf_1 + \frac{1}{b} f_2 \right)\chi + b(c^2 - w^2) + 2\left(H - H_*\right) &= 0 \quad \text{on } \Gamma_a^* = \Gamma_b \quad \text{— at } \delta H, \\ -2wf_1 - \frac{1}{b} f_2 &= 0 \quad \text{on } \Gamma_b^* = \Gamma_a \quad \text{— at } \delta Q. \end{aligned}$$

Variation (4.6) becomes:

$$\delta J(u) = \left\langle \int_{t_0}^{t_1} \mathbb{B}^* f dt, \delta u \right\rangle_{L_2(S)},$$

from that follows the gradient of the objective functional (4.3):

$$\nabla J(x;u) = \int_{t_0}^{t_1} \frac{\partial F_{fr}}{\partial u} f_1 dt \in L_2(S). \tag{4.8}$$

Adjoint PDE (4.7) is linear hyperbolic. Both PDE (4.1) and (4.7) have the same two characteristics $\lambda_{1,2} = c \pm w$. The adjoint problem has two sources of disturbance: first $\frac{2}{L}\left(H - H_*\right)$ is distributed in interval $(x_a, x_b)$, second $2(H - H_*)$ is lumped at the right bound $x_b$. From the right bound the disturbance is spread to the left bound along the family of characteristics $\frac{dx}{dt} = \lambda_1 = c + w$ at $t \to t_0$. Just time $t_1 = t_0 + \int_{x_a}^{x_b} \frac{dx}{\lambda_1}$ is a minimum necessary time to compute the gradient $\nabla J(x;u)$ on interval $(x_a, x_b)$. Given $H = H_*$, we have the trivial adjoint state $f = 0$ over $\Omega$ and a friction coefficient value satisfies to the necessary condition for optimality, here $\|\nabla J\|_{L_2(S)} = 0$.

**5. Conclusions.** We have described the direct optimization approach (method) for solving the different optimization categories with PDE. The essence of approach is transparent: find the gradient of objective functional and go to minimum.

The approach allows us:

- find an infinite-dimensional gradient for optimizing any determinate smooth PDE problem by using universal (formalized) algorithm;
- separate and implement optimization constraints easily, without Lagrange multipliers;
- use infinite-dimensional or finite-dimensional gradient (if a problem admits the simplification) for constructing a direct minimization sequence.

**Appendix A. Adjoint operator.**

Consider an example showing the adjoint operator $\mathbb{A}^*$ and its boundary terms $a^*$. Let be $\mathbb{D} = A(x,v)\dfrac{\partial^k}{\partial x^k}$ on int$\Omega$, $k \le 2$, $A$ is $m \times m$ real matrix, $m$-vector-function $v(x) \in L_2^m(\Omega)$, $\Omega = [x_a, x_b]$. Remind the product $\langle a, b\rangle_{L_2^m(\text{int}\Omega)} = \int_{x_a}^{x_b} \sum_{i=1}^m a_i b_i dx$.

At first, to receive adjoint operator $\mathbb{A}^*$, we need to remove the derivatives of $\delta v$ on $\Omega$:

$$\left\langle \widetilde{f}, \mathbb{A}\delta v \right\rangle_{V^*(\text{int}\Omega)} = \left\langle \widetilde{f}, A\frac{\partial^k \delta v}{\partial x^k} + \sum_{i=1}^m \frac{\partial A}{\partial v_i}\delta v_i \frac{\partial^k v}{\partial x^k} + F'_v \right\rangle_{V^*(\text{int}\Omega)}$$
$$= \int_{\text{int}\Omega} \frac{\partial}{\partial x}\left[\widetilde{f}^T A \frac{\partial^{k-1}\delta v}{\partial x^{k-1}}\right] dx - \left\langle \frac{\partial \widetilde{f}}{\partial x}, A\frac{\partial^{k-1}\delta v}{\partial x^{k-1}} \right\rangle_{V^*(\text{int}\Omega)}$$
$$- \left\langle \widetilde{f}, \frac{\partial A}{\partial x}\frac{\partial^{k-1}\delta v}{\partial x^{k-1}} \right\rangle_{V^*(\text{int}\Omega)} + \left\langle \widetilde{f}, \sum_{i=1}^m \frac{\partial A}{\partial v_i}\delta v_i \frac{\partial^k v}{\partial x^k} + F'_v \right\rangle_{V^*(\text{int}\Omega)}.$$

Here and next upper index $T$ denotes either transposition for a matrix or emphasize a *row*-vector if it is necessary for more accurate definitions. Remember, the first vector in inner product is a row-vector, therefore we do not mark it by index $T$. Received expression removes the derivatives of $\delta v$ for case $k = 1$. If $k = 2$ we need to reduce derivatives additionally in the second and third terms:

$$- \left\langle \frac{\partial \widetilde{f}}{\partial x}, A\frac{\partial^{k-1}\delta v}{\partial x^{k-1}} \right\rangle_{V^*(\text{int}\Omega)} = -\int_{\text{int}\Omega} \frac{\partial}{\partial x}\left[\frac{\partial \widetilde{f}}{\partial x}^T A \frac{\partial^{k-2}\delta v}{\partial x^{k-2}}\right] dx$$
$$+ \left\langle \frac{\partial^2 \widetilde{f}}{\partial x^2}, A\frac{\partial^{k-2}\delta v}{\partial x^{k-2}} \right\rangle_{V^*(\text{int}\Omega)} + \left\langle \frac{\partial \widetilde{f}}{\partial x}, \frac{\partial A}{\partial x}\frac{\partial^{k-2}\delta v}{\partial x^{k-2}} \right\rangle_{V^*(\text{int}\Omega)}$$

$$- \left\langle \widetilde{f}, \frac{\partial A}{\partial x}\frac{\partial^{k-1}\delta v}{\partial x^{k-1}} \right\rangle_{V^*(\text{int}\Omega)} = -\int_{\text{int}\Omega} \frac{\partial}{\partial x}\left[\widetilde{f}^T \frac{\partial A}{\partial x} \frac{\partial^{k-2}\delta v}{\partial x^{k-2}}\right] dx$$
$$+ \left\langle \frac{\partial \widetilde{f}}{\partial x}, \frac{\partial A}{\partial x}\frac{\partial^{k-2}\delta v}{\partial x^{k-2}} \right\rangle_{V^*(\text{int}\Omega)} + \left\langle \widetilde{f}, \frac{\partial^2 A}{\partial x^2}\frac{\partial^{k-2}\delta v}{\partial x^{k-2}} \right\rangle_{V^*(\text{int}\Omega)}.$$

Wee can see that the first terms in all expressions are easily integrated from int$\Omega$ onto $\partial\Omega$. So we have for $k = 1$:

$$\left\langle \widetilde{f}, \mathbb{A}\delta v \right\rangle_{V^*(\text{int}\Omega)}$$
$$= \left\langle A^T\widetilde{f}, \delta v \right\rangle_{V^*(\partial\Omega)} + \left\langle -A^T\frac{\partial \widetilde{f}}{\partial x} - \frac{\partial A^T}{\partial x}\widetilde{f} + \mathbb{C}^*\widetilde{f}, \delta v \right\rangle_{V^*(\text{int}\Omega)},$$

where operator-matrix $\mathbb{C}^*$ is:

$$\mathbb{C}^* = \begin{pmatrix} \left(\dfrac{\partial A}{\partial v_1}\dfrac{\partial^k v}{\partial x^k}\right)^T \\ \cdots \\ \left(\dfrac{\partial A}{\partial v_m}\dfrac{\partial^k v}{\partial x^k}\right)^T \end{pmatrix}.$$

This matrix consists of $m$ row-vectors $\left(\frac{\partial A}{\partial v_i}\frac{\partial^k v}{\partial x^k}\right)^T$, $i = 1, ..., m$.

For $k = 2$ we have:

$$\left\langle \widetilde{f}, \mathbb{A}\delta v \right\rangle_{V^*(\mathrm{int}\Omega)} = \left\langle A^T\widetilde{f}, \frac{\partial \delta v}{\partial x} \right\rangle_{V^*(\partial\Omega)} + \left\langle -A^T\frac{\partial \widetilde{f}}{\partial x} - \frac{\partial A^T}{\partial x}\widetilde{f}, \delta v \right\rangle_{V^*(\partial\Omega)} + \left\langle A^T\frac{\partial^2 \widetilde{f}}{\partial x^2} + \frac{\partial A^T}{\partial x}\frac{\partial \widetilde{f}}{\partial x} + \frac{\partial^2 A^T}{\partial x^2}\widetilde{f} + \mathbb{C}^*\widetilde{f}, \delta v \right\rangle_{V^*(\mathrm{int}\Omega)}.$$

From that we can see adjoint operator $\mathbb{A}^*$ on int$\Omega$:

$$\mathbb{A}^* = (-1)^k \sum_{l=0}^{k} \frac{\partial^l A^T}{\partial x^l}\frac{\partial^{k-l}}{\partial x^{k-l}} + \mathbb{C}^*, \quad k = 1, 2 \quad \text{— at } \delta v,$$

and boundary operators:

$$\begin{aligned} a_0^* &= (-1)^{k-1} \sum_{l=0}^{k-1} \frac{\partial^l A^T}{\partial x^l}\frac{\partial^{k-1-l}}{\partial x^{k-1-l}}\Bigg|_{x_a}^{x_b}, && k = 1, 2 && \text{— at } \delta v, \\ a_1^* &= A^T\Big|_{x_a}^{x_b}, && k = 2 && \text{— at } \frac{\partial \delta v}{\partial x}. \end{aligned}$$

Notice, that $A = A(x, v(x))$, so $\frac{\partial A}{\partial x} = \frac{\partial A}{\partial x}\Big|_{v=const} + \sum_{i=1}^{m} \frac{\partial A}{\partial v_i}\frac{\partial v_i}{\partial x}$.

## REFERENCES


[1] Alexandrov, N.M., (2000) Optimization of Engineering Systems Governed by Differential Equations. *SIAG/OPT Views-and-News*, **Vol.11, nr.2**, pp.1–4.

[2] Atanov ,G.A., Voronin, S.T., and Tolstykh, V.K., (1986) On Parameter Identification Problem of Open Channel (Russian). *Water Resources*, **no.4**, pp.69-78.

[3] Butkovskiy, A.G., (1969) *Theory of Optimal Control of Distributed Parameter Systems*. American Elsevier, New York.

[4] Chavent, G., Dupuy, M., and Lemonnier, P., (1975) History Matching by Use of Optimal Theory. *Society of Petroleum Engineers Journal*, **Vol.15, nr.1**, pp.74–86.

[5] Chen, W.H., Gavalas, G.R., Seinfeld, J.H. and Wasserman, M.L., (1974) A New Algorithm for Automatic History Matching. *Society of Petroleum Engineers Journal*, **Vol.14, nr.6**, pp.593–608.

[6] Egorov, A.I., (1978) *Optimal Control of Heat and Diffusion Equations* (Russian). Nauka, Moscow.

[7] Lewis, R.M., (2000) The Adjoint Approach in a Nutshell. *SIAG/OPT Views-and-News*, **Vol.11, nr.2**, pp.9–12.

[8] Lions, J.L., (1968) *Contrôle Optimal de Systemes Gouvernes par des Equations aux Derivees Partielles*. Dunod Gauthier-Villars, Paris.

[9] Miele, A., (1965) *Theory of Optimum Aerodynamic Shapes*. New-York-London: Akademic Press.

[10] Sachs, Ekkehard W., (2003) PDE Constrained Optimization, *SIAG/OPT Views-and-News*, **Vol.14, nr.1**, pp.7–10.

[11] Tolstykh V.K., (1986) Applying the Gradient Method for Optimization Problems of Distributed Parameter Systems (Russian). *J. of Calcul. Mathem. and Mathem. Physics*, **Vol.26, nr.1**, pp.137-140.

[12] Tolstykh, V.K., (1997) *Direct Extreme Approach for Optimization of Distributed Parameter Systems* (Russian). Yugo–Vostok, Donetsk.

[13] Tolstykh, V.K., (2003) Efficient Method of Optimization of Physical Processes, *Journal of Engineering Physics and Thermophysics*,**Vol.76, nr.2**, pp.424-427.

[14] Tolstykh, V.K., (2012) Optimality Conditions and Algorithms for Direct Optimizing the Partial Differential Equations, *Engineering*, **Vol.4**, **nr.7**, pp.390-394.