

ON NONCLASSICAL OPTIMALITY CONDITIONS OR ON REGULARIZATION OF THE LAGRANGE MULTIPLIER RULE FOR STABLE APPROXIMATION OF SOLUTIONS TO CONVEX CONSTRAINED OPTIMIZATION PROBLEMS

MIKHAIL I. SUMIN^{1,*}

¹*Department of Mathematics, Derzhavin Tambov State University, 33 Internatsionalnaya St., Tambov 392000, Russian Federation*

ABSTRACT. This paper deals with a convex optimization problem with a strongly convex objective functional, an operator equality-constraint (i.e., defined by an operator with an infinite-dimensional image) in a Hilbert space, and a finite number of functional inequality-constraints. In relation to this problem, using the combination of methods of constrained optimization, convex analysis, and the theory of ill-posed problems, we discuss necessary and sufficient nonclassical optimality conditions (NCOCs), which are inextricably linked to the classical Lagrange multiplier rule (LMR). The necessary (sufficient) NCOCs discussed are necessary (sufficient) conditions for ordinary optimality, but they are written in a nonclassical form. Unlike the LMR, the necessary NCOCs are satisfied under only one condition of the solvability of the constrained optimization problem. Simultaneously, NCOCs represent regularizing algorithms for the problem under consideration, providing stable to input data errors approximation of its solution by extremals of a regular Lagrange function while the dual variable is selected in accordance with Tikhonov's regularization procedure. Thus, the NCOCs, particularly in the form of the so-called regularized LMR, overcome the fundamental difficulties associated with the ill-posedness of classical analogs, which shows itself in the possible unsatisfiability (impracticability) and instability of the latter. The paper concludes by showing how a classical LMR, in the case of a convex (possibly not strongly convex) objective functional, can be obtained through a limit transition using NCOC relations in the form of a regularized LMR.

Keywords. Convex constrained optimization, Lagrange multiplier rule, Nonclassical optimality conditions, Regularizing algorithm.

© Optimization Eruditorum

1. INTRODUCTION

The main object of our attention in this paper will be the family of convex constrained optimization problems (see, for example, [1, Item 3.3.1]) with a continuous strongly convex objective functional, with an operator-equality (i.e., defined by an operator with an infinite-dimensional image) constraint that acts into a Hilbert space, and with a finite number of functional inequality-constraints

$$(P_{p,r}^\delta) \quad f^\delta(z) \rightarrow \inf, \quad A^\delta z = h^\delta + p, \quad g_i^\delta(z) \leq r_i, \quad i = 1, \dots, m, \quad z \in \mathcal{D} \subseteq Z.$$

Here: $p \in H$, $r = (r_1, \dots, r_m)^* \in \mathbb{R}^m$ are parameters, Z, H are Hilbert spaces, $f^\delta : \mathcal{D} \rightarrow \mathbb{R}^1$ is a continuous strongly convex functional with a strong convexity constant κ independent of $\delta \in [0, \delta_0]$, $A^\delta : \mathcal{D} \rightarrow H$ is an operator (mapping) such that $A^\delta(\alpha z_1 + (1 - \alpha)z_2) = \alpha A^\delta z_1 + (1 - \alpha)A^\delta z_2 \quad \forall z_1, z_2 \in \mathcal{D}, \alpha \in [0, 1]$ (for example, a bounded linear operator from Z to H can be used as the operator A^δ), $h^\delta \in H$ is a given element, $g_i^\delta : \mathcal{D} \rightarrow \mathbb{R}^1, i = 1, \dots, m$, are continuous convex functionals, $g^\delta(z) \equiv (g_1^\delta(z), \dots, g_m^\delta(z))^*$, $\mathcal{D} \subseteq Z$ is a convex closed set. Problems of the family $(P_{p,r}^\delta)$ depend on the numerical parameter $\delta \in [0, \delta_0]$ which characterizes the error in specifying the input data, where $\delta_0 > 0$ is a certain number, as well as on the parameters additively included in the constraints.

*Corresponding author.

E-mail address: m.sumin@mail.ru (M.-I. Sumin)

2020 Mathematics Subject Classification: 49K27, 49N15, 47A52, 90C46, 90C25.

Accepted: August 19, 2026.

Problem $(P_{p,r}^\delta)$ with $\delta > 0$ will be called perturbed; with $\delta = 0$, accordingly, we will speak of an exact problem. We denote the unique solution to Problem $(P_{p,r}^0)$, if it exists, by $z_{p,r}^0$. The numerical parameter $\delta \in [0, \delta_0]$ characterizing the error in the input data and the parameters $p \in H$, $r \in \mathbb{R}^m$ carry fundamentally different loads in Problem $(P_{p,r}^\delta)$. The presence of the parameters $p \in H$, $r \in \mathbb{R}^m$ in the problem means that below the so-called perturbation method (see, for example, [1, Item 3.3.2]) will be used for its study.

Strictly speaking, from a formal point of view, when we deal with an extremal problem of the form $(P_{p,r}^0)$, we need to determine how we will interpret it. This is due to the fact that extremal problems with constraints [1, 34] have a significant characteristic: each specific problem, depending on its substantive meaning and the researcher's perspective, can be viewed from two fundamentally different perspectives simultaneously. This duality in the interpretation of optimization problems with constraints is inherent in their nature and should be considered a natural property.

In the first, classical approach to extremal problems with constraints, which can be considered traditional for the *theory of extremal problems* [1], the input data of Problem $(P_{p,r}^0)$ should be considered “rigidly” defined and fixed. This approach underlies the aforementioned theory, the core of which is the Lagrange multiplier rule (LMR), or, in other words, the Lagrange principle [1].

At the same time, the “mathematical nature” of constrained extremal problems is such that their entire set can be regarded as a typical class of ill-posed problems. Various examples of such problems, in which at least one of Hadamard's three well-posedness conditions is violated, are well known [31, 34]. Moreover, it seems appropriate to discuss not only the ill-posedness of the constrained optimization problems themselves, but also the ill-posedness of the corresponding classical optimality conditions (COCs). The ill-posedness of the COCs, and in particular the LMR, primarily includes their possible unsatisfiability (impracticability) and instability [15, 16, 17, 18, 19]. Recall that we speak of the unsatisfiability (impracticability) of the multiplier rule when it is known that this classical rule cannot be written down in a given constrained extremal problem. In turn, we understand the instability of the classical rule in the sense that feasible elements satisfying all the relations that make up the multiplier rule behave unstable with respect to perturbations of the input data of the extremal problem. At the same time, it can be argued that the requirement for precise knowledge of the input data when posing many specific constrained optimization problems arising in modern natural science applications often contradicts the physical essence of these problems. The latter is explained by the fact that the input data of such extremal problems are often inextricably linked with the results of a particular physical experiment, the conduct of which is associated with inevitable errors [5, 8, 31]. In such situations, we are forced to adhere to the second approach to problems involving a constrained extremal problem of the form $(P_{p,r}^0)$ within the framework of *constrained optimization theory* [34], which allows for errors in the input data of the extremal problems. It is also appropriate to recall here that the constrained optimization theory itself owes its appearance almost two and a half centuries ago to the needs of solving problems of a wide variety of nature related to human activity [1, 4].

The ill-posedness of extremal problems with constraints, in general, speaks in favor of the naturalness of combining traditional methods of their study [1] with methods of the theory of ill-posed problems [31, 34]. Below in the paper, we show that the LMR in the nondifferential form — the basic result in the constrained extremal problems — serves as a natural bridge between the two approaches mentioned above. On the one hand, in the theory of extremal problems, in accordance with the classical approach, when analyzing and solving a specific extremal problem, it is customary to act in accordance with the well-known classical paradigm [1, §1.6], which is based on the following sequence of actions: 1) formalizing the problem; 2) writing down the Lagrange principle for it; 3) studying the relations that make up the Lagrange principle; 4) writing out the answer in the original problem. The application of this scheme made it possible to solve many extremal problems, related in particular to classical analysis and approximation theory; for detail see [28, 29, 30].

On the other hand, there are important and representative classes of extremal problems whose analysis and solving within the framework of this paradigm encounters fundamental difficulties. Section 2 is devoted to a detailed discussion of these difficulties, as applied to the “simplest” constrained extremal problem. In particular, one such difficulty is the impossibility of writing optimality conditions in LMR form for the element $z_{p,r}^0$, which is the solution to an extremal problem of the form $(P_{p,r}^0)$. However, at the same time, in the same problem, it is possible to write (ordinary) optimality conditions for such an element that are not COCs in the form of the LMR, but to write them, only the existence of an optimal element $z_{p,r}^0$ in the problem is required. More specifically, in this case, the *consequence of the optimality of $z_{p,r}^0$* is a constructive indication of a one-parameter family of admissible elements $z^{\delta,\alpha(\delta)} \in \mathcal{D}$, depending on the parameter $\delta \in (0, \delta_0]$, such that the following circumstances hold:

1) the elements $z^{\delta,\alpha(\delta)}$ as $\delta \rightarrow 0$ converge on function and on constraints, and, under condition the subdifferentiability of the objective functional f^0 , also on argument, to the solution $z_{p,r}^0$ of the problem (see Subsection 4.1.1 and Theorem 4.3). In the language of ill-posed problem theory, this means nothing more than the presence and constructive presentation of a regularizing algorithm (operator) in the sense of Definition 3.3 in Problem $(P_{p,r}^0)$;

2) the elements $z^{\delta,\alpha(\delta)}$ represent extremals of the regular Lagrange function in the perturbed Problem $(P_{p,r}^\delta)$, while the dual variable of the Lagrange function is chosen in accordance with the stabilization procedure (regularization, in one sense or another), using the regularization parameter $\alpha = \alpha(\delta)$, in the problem that is dual to Problem $(P_{p,r}^0)$ (see Section 3), which, as is well known, can always be treated as a concave optimization problem of general form (without equality and inequality constraints);

3) in the “ideal” situation, when $\delta = 0$, the role of the elements $z^{\delta,\alpha(\delta)}$ is taken over by the elements z^α , which are extremals of the regular Lagrange function in the exact Problem $(P_{p,r}^0)$, where $\alpha > 0$ is the regularization parameter of Tikhonov stabilization in the problem dual to Problem $(P_{p,r}^0)$;

4) the elements $z^{\delta,\alpha(\delta)}$ for $\delta > 0$, taken together, satisfy relations (see Theorem 4.3), which are naturally called *nonclassical optimality conditions* (NCOCs) for the element $z_{p,r}^0$, and the passage to the limit as $\delta \rightarrow 0$ in these relations leads to the relations of the classical LMR for the optimal element $z_{p,r}^0$ (see Section 5) provided that the latter satisfies them;

5) if the half-segment $(0, \delta_0]$ is replaced by its “discrete analog” in the form of a sequence of positive numbers $\delta^k \leq \delta_0$, $k = 1, 2, \dots$ converging to zero, then one should talk about NCOCs in sequential form for the element $z_{p,r}^0$, which are written in terms of generalized minimizing sequences (GMS) (see Definition 3.1) in Problem $(P_{p,r}^0)$ (see Theorem 4.6), consisting extremals $z^{\delta^k, \alpha(\delta^k)}$, $k = 1, 2, \dots$, of the Lagrange function, and also in the limit lead to the classical LMR for the element $z_{p,r}^0$ if this COC is satisfied for this optimal element.

To summarize the above, it can be stated that the necessary (sufficient) *nonclassical optimality conditions* (NCOCs), which are necessary (sufficient) optimality conditions in the usual sense: 1) are written in a nonclassical form, while the necessary NCOCs are satisfied for any problem of the form $(P_{p,r}^0)$, if only it has a solution; 2) define regularizing algorithms (operators) in the sense of Definitions 3.3, 3.5 in this problem, which allow one to approach its solution arbitrarily accurately in a stable manner with respect to errors in the input data, see Theorems 4.3, 4.6; 3) lead to COCs in the form of LMR in the limit, provided that the latter are satisfiable in a problem of the form $(P_{p,r}^0)$, see Section 5.

The class of problems of the form $(P_{p,r}^0)$ for which problems with the satisfiability and stability of the LMR arise is quite broad. First of all, we encounter such problems among extremal problems with operator equality-constraints, which arise, in particular, in applications related to optimal control problems and the inverse problems that reduce to them [5, 8, 31]. In the next Section, we will consider in detail one such problem, which, on the one hand, is an ordinary, very simple-looking extremal problem of the type $(P_{p,r}^0)$, and on the other hand, is a classical ill-posed problem of finding a normal solution to

an operator equation of the first kind on a pair of Hilbert spaces, i.e., a problem whose study [31] laid the foundations of the modern theory of ill-posed problems.

An approach to constrained extremal problems of the form $(P_{p,r}^0)$ associated with the regularization of the LMR for such problems was proposed and developed in our papers [15, 16, 17, 18, 19]. In these papers, regularized LMRs were formulated as theorems on necessary and sufficient conditions for the existence of GMSs in the problems under consideration. In this paper, we approach this issue from a slightly different angle and formulate our results first of all in terms of NCOCs, thereby emphasizing that the issue of LMR regularization is inextricably linked with the theory of (ordinary) optimality conditions, and that the (ordinary) optimality conditions themselves can also be expressed in the form of regularizing algorithms (operators).

Thus, below in the paper, based on the conjugation of methods of constrained optimization, convex analysis and the theory of ill-posed problems, we sequentially: 1) discuss the fact that in the theory of extremal problems there exist simple in formulation, and at the same time important in applied respect, problems, which, on the one hand, are typical representatives of the class of extremal problems with constraints, “it would seem no different from other similar extremal problems”, and on the other hand, cause fundamental difficulties when using standard methods of the theory for their analysis (Section 2); 2) present concepts, definitions, statements necessary for the presentation of the main results (Section 3); 3) prove necessary NCOC in Problem $(P_{p,r}^0)$ in the form of a dual regularization algorithm and also prove a sequential nonclassical optimality criterion in the form of a regularized LMR (Section 4); 4) show that NCOCs lead to classical LMR in Problem $(P_{p,r}^0)$ in the limit (Section 5).

Finally, in conclusion of Introduction, we note that various aspects of the numerical realization, including numerical experiments, of regularized COCs, or, in other words, NCOCs, in constrained optimization and optimal control problems, were considered earlier in [10, 11, 12, 13].

The following standard notations are used in the paper: $\text{int } X$ is the interior of a set X , ∂X is the boundary of a set X , $Pr_X(x)$ is the projection of a point x onto a set X , $N_X(x)$ is the normal cone (in the sense of convex analysis) to a convex set X at a point $x \in X$, $\text{dom } f \equiv \{x \in X : f(x) < +\infty\}$, $\text{epi } f \equiv \{(x, \alpha) \in X \times \mathbb{R}^1 : f(x) \leq \alpha\}$, $\partial f(x)$ and $\partial^\infty f(x)$ are the domain, epigraph, subdifferential, and singular subdifferential of a convex function $f : X \rightarrow \mathbb{R}^1 \cup \{+\infty\}$.

2. ON THE “SIMPLEST” CONSTRAINED EXTREMAL PROBLEM, WHICH UNDERLIES THE ENTIRE THEORY OF ILL-POSED PROBLEMS

As noted in Introduction, the standard paradigm [1, §1.6] underlying the theory of extremal problems, according to which “familiar” optimality conditions are written in terms of optimal and corresponding dual (adjoint) variables, fails in a vast number of cases. An important, typical, and “simplest” example of an extremal problem that falls outside the scope of this paradigm is the classical ill-posed problem of finding the norm-minimal solution to an operator equation of the first kind on a pair of Hilbert spaces.

2.1. The “simplest” constrained extremal problem. Thus, the object of our attention in this Section will be the “simplest” form of the constrained extremal problem with an operator-equality (i.e., defined by an operator with an infinite-dimensional range) constraint in a Hilbert space

$$(P) \quad \|z\|^2 \rightarrow \inf, \quad Az = h, \quad z \in \mathcal{D} \subseteq Z,$$

where $A : Z \rightarrow H$ is a bounded linear operator, $h \in H$ is a given element, $\mathcal{D} \subseteq Z$ is a convex closed set, Z, H are Hilbert spaces. Problem (P) can be reformulated equivalently in another form as the inverse problem of finding a norm-minimal solution to an abstract operator equation of the first kind on a pair of Hilbert spaces

$$(IP) \quad Az = h, \quad z \in \mathcal{D} \subseteq Z.$$

Historically, it was in this equivalent form that the far from simple Problem (P) became known at the turn of the 1950s and 1960s as the “main” ill-posed problem [31], the intensive study of which ultimately served as the basis for the development of the modern theory of ill-posed problems [5, 8]. Within the theory of ill-posed problems, the search for a norm-minimal solution to equation (IP) was replaced by an equivalent search for a minimum of the functional $\|Az - h\|^2$, $z \in \mathcal{D}$. In this approach, the “main” ill-posed problem was in no way considered as Problem (P) itself, i.e., a constrained minimum problem with an operator equality-constraint, using the corresponding apparatus of the Lagrange principle theory. The author believes that perhaps the most important reason for this was the insufficient development in those years of the apparatus of the Lagrange principle theory for extremal problems with complicated operator constraints. Let us consider the main difficulties associated with the LMR that arise in the formal approach to the ill-posed Problem (P) as a constrained minimum problem with an operator-equality constraint.

2.2. Parametric and individual LMRs in nondifferential form for problem (P) and difficulties associated with their applicability to the problem solving. Unfortunately, the mathematical nature of constrained extremal problems is such that the application of the LMR for the practical solving many pressing problems in natural science encounters, as noted above, fundamental difficulties caused by the ill-posedness of both constrained optimization problems in general [34] and the classical rule itself [15, 16, 17, 18, 19]. With regard to problem (P) , we will highlight three main points in this context related to the properties of the LMR ill-posedness, such as unsatisfiability and instability. [15, 16, 17, 18, 19].

2.2.1. When it is possible to write the relations of the LMR, its satisfiability. Recall that there are two approaches to obtaining the LMR in nondifferential form for problems of the form $(P_{p,r}^0)$. First, one can use the perturbation method [1, Item 3.3.2] and, relying on the subdifferentiability properties of convex and lower semicontinuous value function, arrive at the so-called parametric (i.e., dependent on the parameters p, r) LMR. For detail, see [15, 24]. Secondly, following, for example, [1], one can strive to obtain the LMR “personally” in each individual problem $(P_{p,r}^0)$, for example, in Problem $(P_{0,0}^0)$.

2.2.1.1. When it is possible to write the relations of the parametric LMR, its satisfiability. The approach to deriving the LMR for Problem (P) using the perturbation method (see, e.g., [1, Item 3.3.2]), which exploits its inclusion in a family of similar problems depending on the parameter $p \in H$,

$$(P_p) \quad \|z\|^2 \rightarrow \inf, \quad Az = h + p, \quad z \in \mathcal{D} \subseteq Z,$$

establishes a rigid connection between the LMR relations and the subdifferential properties of the value function of the problem. Namely, as shown in [15, Theorem 2.1], [24, Theorem 1.1], this approach allows us to obtain formally a nondegenerate (regular or irregular) LMR in Problem $(P) = (P_0)$ if and only if at least one of the two relations holds: $\partial\beta(0) \neq \emptyset$ or $\partial^\infty\beta(0) \neq \{0\}$. Here $\partial\beta(0)$ and $\partial^\infty\beta(0)$ are, respectively, the subdifferential and singular subdifferential (in the sense of convex analysis, see, for example, [14]) at zero of the convex lower semicontinuous value function $\beta(p) \equiv \min_{z \in \mathcal{D}, Az=h+p} \|z\|^2$, $p \in H$.

Proposition 2.1. [Parametric nondifferential LMR] *Let $p \in H$ be a point such that $\beta(p) < +\infty$. Then:*

1. *If $z_p^0 \in \mathcal{D}_p^0 \equiv \{z \in \mathcal{D} : Az - h - p = 0\}$ is an optimal element in Problem (P_p) , i.e. $\|z_p^0\|^2 = \beta(p)$, and $\zeta \in \partial\beta(p)$, where $\partial\beta(p)$ is a subdifferential in the sense of convex analysis, then for the Lagrange multiplier $\lambda \in H$, $\lambda = -\zeta$, for $\mu_0 = 1$ the relation holds*

$$L_p(z_p^0, \mu_0, \lambda) \leq L_p(z, \mu_0, \lambda) \quad \forall z \in \mathcal{D}, \quad L_p(z, \mu_0, \lambda) \equiv \mu_0 \|z\|^2 + \langle \lambda, Az - h - p \rangle, \quad (2.1)$$

and $-\zeta = \lambda$ is the Kuhn–Tucker vector of Problem (P_p) .

Conversely, if $\tilde{z} \in \mathcal{D}_p^0$ is an element such that for some $\mu_0 > 0$, $\lambda \in H$ the relation (2.1) holds with z_p^0 replaced by $\tilde{z}$, then this element is optimal in Problem (P_p) , i.e. $\tilde{z} = z_p^0$, the element λ/μ_0 is a Kuhn–Tucker vector for it and simultaneously $-\lambda/\mu_0 \in \partial\beta(p)$.

2. If $z_p^0 \in \mathcal{D}_p^0$ is an optimal element in Problem (P_p) , $p \in \partial \text{dom}\beta$, and $\zeta \in \partial^\infty\beta(p)$, $\zeta \neq 0$, where $\partial^\infty\beta(p)$ is a singular subdifferential defined by the formula $\partial^\infty\beta(p) \equiv \{\lambda \in H : (\lambda, 0) \in N_{\text{epi}\beta}(p, \beta(p))\}$, then for the Lagrange multiplier $\lambda \in H$, $\lambda = -\zeta$, the relation (2.1) holds for $\mu_0 = 0$.

Conversely, if $\tilde{z} \in \mathcal{D}_p^0$ is an element such that for $\mu_0 = 0$ and some $\lambda \in H$, $\lambda \neq 0$, the relation (2.1) holds with z_p^0 replaced by $\tilde{z}$, then $p \in \partial \text{dom}\beta$ and simultaneously $-\lambda \in \partial^\infty\beta(p)$.

Remark 2.2. Unfortunately, verifying the satisfiability of the required subdifferential properties of the value function is itself a complex mathematical problem.

2.2.1.2. When it is possible to write relations of an individual LMR, its satisfiability. In parallel with Proposition 2.1, known approaches to deriving the LMR in individual problems [1, 4] require the closedness of the image of the operator A . As applied to Problem (P) in the case $\mathcal{D} = Z$, the classical LMR for smooth problems with equalities from [1, Item 3.2.2, Theorem], see also [4, Section 2, Corollary 1], when Z, H are Hilbert spaces, is formulated as follows.

Proposition 2.3. *If the point z^0 is a solution to Problem (P) and $R(A) = \overline{R(A)}$, then there exists a nondegenerate set $(\mu_0, \lambda) \in \mathbb{R}_+^1 \times H$ such that $2\mu_0 z^0 + A^*\lambda = 0$. If $R(A) = H$, then in the last equality we can take $\mu_0 > 0$.*

Remark 2.4. Since Problem (P) is convex, in this case the equality $2\mu_0 z^0 + A^*\lambda = 0$ is equivalent to the inequality (see also [1, Item 3.3.1, Corollary 1])

$$L_0(z^0, \mu_0, \lambda) \leq L_0(z, \mu_0, \lambda) \quad \forall z \in Z, \quad L_0(z, \mu_0, \lambda) \equiv \mu_0 \|z\|^2 + \langle \lambda, Az - h \rangle.$$

Remark 2.5. It is important to note that the main requirement $R(A) = \overline{R(A)}$ of Proposition 2.3 is very strict, and, as noted in [1, Item 3.2.4], failure to satisfy this closure condition may lead to the LMR not being satisfied at all. The closure condition cannot be satisfied, for example, for completely continuous operators A [33, p. 225, Theorem 1], “most frequently” encountered when considering various substantial optimization problems related to differential equations (optimal control, inverse problems) [8, 31]. Thus, for Problem (P_p) of finding a normal solution to an operator equation with a completely continuous operator, for example, in the case $\mathcal{D} = Z$, the individual LMR [1, Item 3.2.2, Theorem]) is completely inapplicable for any $p \in H$, while the parametric LMR of Proposition 2.1 is applicable for p from a dense set of points in $\text{dom}\beta$.

2.2.1.3. Comparison of parametric and individual LMRs. In this context, it is also of interest to compare two LMRs, namely, the parametric LMR [15, Theorem 2.1], [24, Theorem 1.1] as applied to Problem (P_p) and the individual LMR formulated for the same problem, i.e., Proposition 2.3. To simplify the situation, we consider Problem (P) under the condition $\mathcal{D} = Z$.

First, note that the condition $R(A) = \overline{R(A)}$, $R(A) \neq H$ of Proposition 2.3 implies the presence of a nonzero element in the singular subdifferential $\partial^\infty\beta(0)$. Indeed, in this case $\text{dom}\beta = R(A)$ and for any $x \in R(A)$ the normal cone $N_{R(A)}(x)$ coincides [2, Proposition 4.1.9] with the orthogonal subspace $R(A)^\perp$ containing a nonzero element. This circumstance, in accordance with the first assertion of the second part of Proposition 2.1, ensures the satisfiability of the irregular ($\mu_0 = 0$) LMR (2.1) for $\mathcal{D} = Z$.

Secondly, the condition $R(A) = \overline{R(A)} = H$, as can be seen, ensures that the subdifferential $\partial\beta(0)$ is nonempty. This is due to the fact that the functional β in this case, in accordance with Lemma 3.2, is convex and finite at each point $x \in H$, i.e. $\text{dom}\beta = H$, and therefore $\partial\beta(x) \neq \emptyset$ due to Corollary 2.3 and Theorem 4.2 in [3], which, in turn, due to the validity of the first assertion of the first part of Proposition 2.1 in this situation, ensures the satisfiability of the regular ($\mu_0 > 0$) LMR (2.1) when

$\mathcal{D} = Z$. Thus, the conditions $R(A) = \overline{R(A)}$, $R(A) \neq H$ and $R(A) = \overline{R(A)} = H$ of Proposition 2.3 (the classical LMR for smooth problems with equalities [1, Item 3.2.2, Theorem]) are also sufficient for the applicability of the corresponding necessary conditions for the extremum of Proposition 2.1.

Remark 2.6. It can be argued that there exists a broad class of problems, for example of the form (P_0) , for each of which a nondegenerate classical parametric LMR of Proposition 2.1 can be written (i.e., in (2.1) the set $(\mu_0, \lambda) \neq 0$), but, at the same time, the LMR of Proposition 2.3 cannot be applied (for smooth problems with equalities [1, Item 3.2.2, Theorem]). In the case of $\mathcal{D} = Z$, such problems include, for example, problems in which $R(A) \neq \overline{R(A)}$, but, simultaneously, either $\partial^\infty \beta(0) \neq \{0\}$, or $\partial \beta(0) \neq \emptyset$, or both hold. First of all, such problems include problems with integral operators A , for which, in a large number of cases that are crucial for various natural science applications, the equality $R(A) = \overline{R(A)}$ does not hold.

2.2.2. Unsatisfiability of the LMR when it is unknown whether the LMR relations can be written down or not; possible unsatisfiability of the LMR. In turn, we speak of the unsatisfiability of the LMR when it is known that this classical rule cannot be written down for a given constrained extremal problem. In fact, a situation “close” to this also arises when we do not know whether a multiplier rule can be written down for a given constrained extremal problem. For example, this happens when we obtain a multiplier rule under certain assumptions and do not know how to do this when these assumptions are relaxed. In particular, as applied to Problem (P) , the classical LMR for smooth problems with equalities from [1, Item 3.2.2, Theorem], see also [4, Section 2, Corollary 1], in the case $\mathcal{D} = Z$, when Z, H are Hilbert spaces, is formulated as a consequence of the satisfiability of a certain set of corresponding conditions (see Subsection 2.2.1.2). If these conditions are not satisfied, then the corresponding statements [1, Item 3.2.2, Theorem], [4, Section 2, Corollary 1] do not convey any information about the satisfiability of the Lagrange principle.

A classic example of the unsatisfiability of the Lagrange principle can be found in [1, Item 3.2.4]. We give another similar example related to the “simplest” minimization Problem (P) with an operator equality-constraint. It is intended to demonstrate that in this class of important problems (including those from an applied perspective), there are substantive problems for which the Lagrange principle is simply “impossible to write down”, making its formal application in such situations completely “meaningless”. At the same time, we note that the problem of finding a normal (minimal in norm) solution to an operator equation of the first kind considered in this example is an ill-posed problem, a detailed analysis of the instability with respect to perturbations of the input data for which can be found, essentially, in any textbook on the theory of ill-posed problems (see, for example, [31, 34]).

Example 2.7. Let us consider the problem of minimizing a strongly convex functional with an operator equality-constraint

$$\|z\|^2 \rightarrow \min, \quad Az = p, \quad z \in Z \quad (2.2)$$

with an injective and self-adjoint operator $A : Z \rightarrow Z$, Z is a Hilbert space, such that $R(A) \neq Z$ (for example, A can be a Fredholm integral operator with a closed symmetric kernel). We show that in the problem (2.2), for certainly chosen p , the Lagrange principle fails, which, as is easily seen, also leads to the unsolvability of the corresponding dual problem. Let $z^0 \in Z$, but $z^0 \notin R(A)$. Then we consider the problem (2.2) with $p = Az^0$. In this problem, the classical Lagrange principle in differential form [1, § 3.2.2], and, consequently, in nondifferential form, does not hold. Indeed, if this were not the case, then there would exist a nondegenerate pair of factors $(\lambda_0, \lambda) \in \mathbb{R}_+^1 \times Z$ such that

$$2\lambda_0 z^0 + A\lambda = 0.$$

In this case, for $\lambda_0 = 0$, we obtain $\lambda = 0$ due to the injectivity of A , and for $\lambda_0 = 1$, we obtain the contradictory equality $z^0 = -1/2A\lambda$, which proves the impossibility of satisfying the classical Lagrange principle in problem (2.2) with the chosen p .

As a specific constrained optimization problem of the form (2.2), we can take, for example, the problem of finding a normal solution to a Fredholm integral equation of the first kind (such problems are among the classical ones in the theory of ill-posed problems [31])

$$\|z\|_{2,(0,1)}^2 \rightarrow \min, \quad \int_0^1 K(x,s)z(s)ds = h(x), \quad 0 \leq x \leq 1, \quad z \in L_2(0,1) \quad (2.3)$$

with symmetric kernel $K(x,s) = \{(1-x)s, 0 \leq s \leq x; x(1-s), x \leq s \leq 1\}$ and with $Z = L_2(0,1)$. In this case, in accordance with Picard's theorem (see, for example, [9, p. 148]), the equation (2.3) can only be uniquely solvable (see the analysis of Example 1 on the page 149 in [9]) due to the closedness of the kernel. Therefore, in accordance with the above, it can be stated that for any $h(\cdot) = \int_0^1 K(\cdot, s)z(s)ds$ with $z \in L_2(0,1)$ that are not continuous functions (the corresponding equivalence class does not contain a continuous function), the Lagrange principle is not satisfied in the constrained optimization problem (2.3).

Remark 2.8. The problem considered in Example 2.7 for $p = Az^0$ has the peculiarity that for the corresponding value function β at the chosen point p , the relations $\partial\beta(p) = \emptyset$ and $\partial^\infty\beta(p) = \{0\}$ simultaneously hold.

For various examples of the unsatisfiability of the Lagrange principle, see also in [15, 16, 17].

2.2.3. Instability of the LMR. Finally, if Problem (P) is such that the LMR can be “written” in it, then the “practical” use of the classical result (for example, when finding approximations to the solution of the problem) inevitably runs into the problem of its instability with respect to perturbations of the input data [15, 16, 17]. Let us consider an example of the simplest finite-dimensional quadratic programming problem, which, on the one hand, can be easily solved using the Kuhn-Tucker theorem (however, due to the simplicity of the problem, it can be easily solved without this classical optimization theorem), but, on the other hand, effectively demonstrates the instability of the usual optimality conditions formulated in this classical theorem if we formally apply it in the case where the input data of the problem are subject to perturbation.

Example 2.9. Let us consider the problem of minimizing a strongly convex quadratic function of two variables on a set defined by an affine equality-constraint. This problem is equivalent to the problem of finding a normal (minimal in norm) solution to a linear algebraic system of two equations in two unknowns,

$$|x|^2 \rightarrow \min, \quad Ax = y, \quad x \in \mathbb{R}^2, \quad A \equiv \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad y \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (2.4)$$

whose exact normal solution is $x^* = (0.5; 0.5)$. The problem dual to (2.4) has the form

$$V(\lambda) \equiv L(x(\lambda), \lambda) = -\frac{1}{4}\langle AA^*\lambda, \lambda \rangle - \langle y, \lambda \rangle \rightarrow \max, \quad \lambda \in \mathbb{R}^2,$$

where $L(x, \lambda) \equiv |x|^2 + \langle \lambda, Ax - y \rangle$, $x(\lambda) \equiv \operatorname{argmin}\{L(x, \lambda), x \in \mathbb{R}^2\} = -\frac{1}{2}A^*\lambda$. Its solution is the vector $(-1, \alpha) \forall \alpha \in \mathbb{R}^1$. It is easy to verify that any such vector is also a Kuhn-Tucker vector of the problem (2.4). Consider the perturbed problem (2.4) for $\delta > 0$

$$|x|^2 \rightarrow \min, \quad A^\delta x = y^\delta, \quad x \in \mathbb{R}^2, \quad A^\delta \equiv \begin{pmatrix} 1 & 1 \\ 0 & \delta^2 \end{pmatrix}, \quad y^\delta \equiv \begin{pmatrix} 1 \\ \delta \end{pmatrix}. \quad (2.5)$$

The corresponding dual problem

$$V^\delta(\lambda) \equiv L^\delta(x^\delta(\lambda), \lambda) \rightarrow \max, \quad \lambda \in \mathbb{R}^2$$

has a solution $\lambda^\delta = (\frac{2-2\delta}{\delta}, \frac{2\delta-4}{\delta^3})$ and the vector $x^\delta(\lambda^\delta) \equiv \operatorname{argmin}\{L^\delta(x, \lambda^\delta), x \in \mathbb{R}^2\} = (1 - \frac{1}{\delta}, \frac{1}{\delta})$ is, in accordance with the classical Kuhn-Tucker theorem, a solution to the perturbed problem (2.5)

and simultaneously an “approximate” solution to the original problem (2.4), which as $\delta \rightarrow 0$ does not converge to its unique exact solution. We emphasize that both the exact problem and the perturbed one are regular here, and therefore the Kuhn–Tucker theorem is used to solve them. However, as is easily seen, the value function of this problem, as a function of the parameter $\delta \geq 0$, suffers a discontinuity of the second kind at the point $\delta = 0$.

3. NECESSARY CONCEPTS, DEFINITIONS, CONSTRUCTIONS, AND AUXILIARY RESULTS

3.1. Necessary concepts, definitions, constructions. We now return to the study of the family of Problems $(P_{p,r}^\delta)$, recalling that the unique solution of Problem $(P_{p,r}^0)$, if it exists, is denoted by $z_{p,r}^0$. Below, the following condition will play a central role:

Condition A. The input data f^δ , A^δ , g^δ of Problem $(P_{p,r}^\delta)$ are continuous on $\mathcal{D}$, and their values are bounded uniformly for $z \in \mathcal{D} \cap B_M$ on $\mathcal{D} \cap B_M$ for each fixed $M > 0$, where $B_M \equiv \{z \in Z : \|z\| \leq M\}$.

We further assume that Condition A is satisfied. In deriving duality-based NCOCs for Problem $(P_{p,r}^0)$ in the form of an algorithm for solving it that is stable to errors in the input data, the concept of a generalized minimizing sequence (GMS) is crucial. To define this central concept, let $\beta(p, r)$ denote the generalized value (generalized lower bound) of Problem $(P_{p,r}^0)$:

$$\beta(p, r) \equiv \beta_{+0}(p, r) \equiv \lim_{\epsilon \rightarrow +0} \beta_\epsilon(p, r), \quad \beta_\epsilon(p, r) \equiv \inf_{z \in \mathcal{D}_{p,r}^{0,\epsilon}} f^0(z), \quad \beta_\epsilon(p, r) \equiv +\infty, \text{ if } \mathcal{D}_{p,r}^{0,\epsilon} = \emptyset,$$

where

$$\mathcal{D}_{p,r}^{\delta,\epsilon} \equiv \{z \in \mathcal{D} : \|A^\delta z - h^\delta - p\| \leq \epsilon, \quad g_i^\delta(z) \leq r_i + \epsilon, \quad i = 1, \dots, m\}, \quad \delta \geq 0, \quad \epsilon \geq 0, \quad \mathcal{D}_{p,r}^{0,\epsilon} \equiv \mathcal{D}_{p,r}^\epsilon.$$

Definition 3.1. A sequence $z^i \in \mathcal{D}$, $i = 1, 2, \dots$ such that

$$f^0(z^i) \leq \beta(p, r) + \delta^i, \quad z^i \in \mathcal{D}_{p,r}^{0,\epsilon^i}$$

for some sequences of nonnegative numbers δ^i , ϵ^i , $i = 1, 2, \dots$, converging to zero, is called a generalized minimizing sequence (GMS) in Problem $(P_{p,r}^0)$.

Let us also recall that the concept of GMS coincides with the well-known concept of a generalized optimal plan in the sense of [7] in mathematical programming, as well as with the concept of a minimizing approximate solution in the sense of J. Warga [35, Chapter III] used in optimal control.

Obviously, in the general situation, $\beta(p, r) \leq \beta_0(p, r) \equiv \inf_{z \in \mathcal{D}_{p,r}^0} f^0(z)$, where the quantity

$$\beta_0(p, r) \equiv \{f^0(z_{p,r}^0), \text{ if } z_{p,r}^0 \text{ exists; } +\infty \text{ otherwise}\}$$

is the classical value (the classical lower bound) of the problem. But in the case of Problem $(P_{p,r}^0)$ posed above, the equality $\beta(p, r) = \beta_0(p, r)$ holds. Indeed, since Problem $(P_{p,r}^0)$ is a convex programming one, and the functional f^0 is also strongly convex and continuous on $\mathcal{D}$, it can be stated that the following lemma holds [24, Lemma 1.2], [25, Lemma 2.4.1].

Lemma 3.2. *If $\beta(p, r) < +\infty$, then the generalized lower bound $\beta(p, r)$ in Problem $(P_{p,r}^0)$ is attained. In this case, the equality $\beta_0(p, r) = \beta(p, r)$ holds. The functional $\beta : H \times \mathbb{R}^m \rightarrow \mathbb{R}^1 \cup \{+\infty\}$ is convex and lower semicontinuous.*

Lemma 3.2 and Definition 3.1 ensure the validity of the limit relation for a GMS $z^i \in \mathcal{D}$, $i = 1, 2, \dots$,

$$f^0(z^i) \rightarrow f^0(z_{p,r}^0) = \beta_0(p, r) = \beta(p, r), \quad i \rightarrow \infty, \quad (3.1)$$

if a solution to Problem $(P_{p,r}^0)$ exists.

Thus, we have sets of unperturbed and perturbed input data respectively: $\{f^0, A^0, h^0, g^0\}$ and $\{f^\delta, A^\delta, h^\delta, g^\delta\}$, $\delta \in (0, \delta_0]$, $\delta_0 > 0$ is some number. We assume that

$$|f^\delta(z) - f^0(z)| \leq C\delta(1 + \|z\|^2), \quad \|A^\delta z - A^0 z\| \leq C\delta(1 + \|z\|) \quad \forall z \in \mathcal{D}, \quad (3.2)$$

$$\|h^\delta - h^0\| \leq C\delta, \quad |g^\delta(z) - g^0(z)| \leq C\delta(1 + \|z\|^2) \quad \forall z \in \mathcal{D},$$

where $C > 0$ is independent of δ , $g^\delta(\cdot) \equiv (g_1^\delta(\cdot), \dots, g_m^\delta(\cdot))^*$.

We introduce an important definition of a regularizing operator (algorithm) in convex Problem $(P_{p,r}^0)$ [20, 21], see also [18].

Definition 3.3. An operator (algorithm) $R(\cdot, \cdot, \cdot, \cdot, \delta)$, depending on $\delta \in (0, \delta_0)$, that associates with each set of the input data $(f^\delta, A^\delta, h^\delta, g^\delta)$ satisfying the estimates (3.2) an element $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) \equiv z^\delta \in \mathcal{D}$ such that $f^0(z^\delta) \rightarrow \beta(p, r)$, $\|A^0 z^\delta - h^0 - p\| \rightarrow 0$, $\min_{x \in \mathbb{R}_+^m} |g^0(z^\delta) - x - r| \rightarrow 0$, $\delta \rightarrow 0$, is called a regularizing operator (algorithm) in Problem $(P_{p,r}^0)$.

Remark 3.4. As it is well known, the concept of a regularizing operator (algorithm) is fundamental in the theory of ill-posed problems [31] (see also [34]). A similar concept, but applied to constrained extremal problems, for which the same phrase “regularizing operator (algorithm)” is used, was introduced in [20, 21]. In [22] it is shown that in the case of the “simplest” problem (P) , i.e., in the case of an operator equation of the first kind (IP) , the concepts of a regularizing algorithm in the sense of [31] and in the sense of [20, 21] are equivalent.

Since our main goal is ultimately to construct a GMS in Problem $(P_{p,r}^0)$, and the family $\{z^\delta \in \mathcal{D} : \delta \in (0, \delta_0)\}$ from Definition 3.3 is not a sequence, then, in addition to the above definition of a regularizing operator (algorithm) in Problem $(P_{p,r}^0)$, introduce its “trace”, namely, the definition of the GMS-generating operator (algorithm) in Problem $(P_{p,r}^0)$ [20, 21], see also [18].

Definition 3.5. Let $\delta^k \in (0, \delta_0)$, $k = 1, 2, \dots$, be a sequence of positive numbers converging to zero. An operator (algorithm) $R(\cdot, \cdot, \cdot, \cdot, \delta^k)$, depending on δ^k , $k = 1, 2, \dots$, that associates with each set of the input data $(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k})$ satisfying the estimates (3.2) for $\delta = \delta^k$ an element $z^{\delta^k} \in \mathcal{D}$ is called GMS-generating in Problem $(P_{p,r}^0)$ if the sequence z^{δ^k} , $k = 1, 2, \dots$, is a GMS in this problem.

We introduce the Lagrange functional

$$L_{p,r}^\delta(z, \lambda, \mu) \equiv f^\delta(z) + \langle \lambda, A^\delta z - h^\delta - p \rangle + \langle \mu, g^\delta(z) - r \rangle, \quad z \in \mathcal{D}, \quad \lambda \in H, \quad \mu \in \mathbb{R}^m,$$

concave dual functional

$$V_{p,r}^\delta(\lambda, \mu) \equiv \inf_{z \in \mathcal{D}} L_{p,r}^\delta(z, \lambda, \mu), \quad \lambda \in H, \quad \mu \in \mathbb{R}^m,$$

as well as the dual problem

$$V_{p,r}^\delta(\lambda, \mu) \rightarrow \sup, \quad \lambda \in H, \quad \mu \in \mathbb{R}_+^m.$$

Due to the strong convexity and continuity on z of the Lagrange functional for any pair $(\lambda, \mu) \in H \times \mathbb{R}_+^m$, where $\mathbb{R}_+^m \equiv \{x = (x_1, \dots, x_m)^* \in \mathbb{R}^m : x_i \geq 0, i = 1, \dots, m\}$, the value $V_{p,r}^\delta(\lambda, \mu)$ is achieved at a single element $z^\delta[\lambda, \mu] \equiv \operatorname{argmin}\{L_{p,r}^\delta(z, \lambda, \mu), z \in \mathcal{D}\}$, and this element does not depend on h^δ .

3.2. Auxiliary results. In this section, we formulate and prove auxiliary assertions that are of primary importance in substantiating the main results of the paper. First, we present the following two lemmas concerning the properties of strongly convex functions. The proof of the first can be found in [34, Theorems 4.3.1, 8.2.10].

Lemma 3.6. *Let $\mathcal{H}$ be a Hilbert space, the set $X \subseteq \mathcal{H}$ be convex and closed, the function $\varphi : X \rightarrow \mathbb{R}^1$ be strongly convex with constant κ and lower semicontinuous, and $z^* = \operatorname{argmin}\{\varphi(z), z \in X\}$. Then we have the estimate*

$$\frac{\kappa}{2} \|z - z^*\|^2 \leq \varphi(z) - \varphi(z^*) \quad \forall z \in X.$$

In proving the following lemma we will rely on [15, Lemma 3.2], as well as [34, Theorem 4.3.2].

Lemma 3.7. *Let $f : X \rightarrow \mathbb{R}^1$ be a convex, lower semicontinuous, and subdifferentiable function at the points of a convex closed set $X \subseteq \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space. Then, if it is strongly convex on X with constant κ , then*

$$f(z) \geq f(y) + \langle \partial f(y), z - y \rangle + \frac{\kappa}{2} \|z - y\|^2 \quad \forall z, y \in X. \quad (3.3)$$

Proof. We use the fact that the general definition of strong convexity [34, Chapter 8, p. 650] is analogous to the definition of strong convexity of a function of a finite number of variables. The Hilbert space analogue of the proposition [34, Lemma 4.3.1] allows us to say that f is strongly convex on X with constant $\kappa > 0$ if and only if there exists a convex function $g(\cdot) : X \rightarrow \mathbb{R}^1$ such that $f(z) = g(z) + (\kappa/2)\|z\|^2$, $z \in X$. Moreover (see [3, Corollary 4.3]) the functions f and g are subdifferentiable at the points of X simultaneously and

$$\partial f(y) = \partial g(y) + \kappa y, \quad y \in X. \quad (3.4)$$

Thus, the convex function $g(z) \equiv f(z) - (\kappa/2)\|z\|^2$, $z \in X$, is subdifferentiable on X , i.e.

$$g(z) \geq g(y) + \langle \partial g(y), z - y \rangle, \quad z, y \in X. \quad (3.5)$$

It remains to note that from the relations (3.4), (3.5) and the obvious identity $\|z\|^2 = \|y\|^2 + \langle 2y, z - y \rangle + \|z - y\|^2$, $z, y \in X$, (3.3) follows. $\square$

We now show that, given the strong convexity of the functional f^0 , we can assert that the limit relation (3.1) has an important consequence.

Lemma 3.8. *Let $\beta(p, r) < +\infty$. Then, for any GMS z^i , $i = 1, 2, \dots$, in Problem $(P_{p,r}^0)$, which is solvable in this case, the limit relation holds*

$$z^i \rightarrow z_{p,r}^0 \text{ weakly in } Z, \quad i \rightarrow \infty.$$

If, in addition to the assumptions made above, the strongly convex functional f^0 is subdifferentiable (in the sense of convex analysis) at the points of $\mathcal{D}$, then the weak convergence indicated above is strong, i.e.

$$\|z^i - z_{p,r}^0\| \rightarrow 0, \quad i \rightarrow \infty.$$

Proof. First, it can be noted that since $\beta(p, r) < +\infty$, $f^0(z^i) \rightarrow \beta(p, r)$, $i \rightarrow \infty$, and the functional f^0 is continuous and strongly convex on $\mathcal{D}$, the sequence z^i , $k = 1, 2, \dots$ discussed in the lemma is bounded. Indeed, let $z^* \in \mathcal{D}$ be the minimum point of f^0 on $\mathcal{D}$ (a strongly convex continuous functional on a convex closed set of a Hilbert space attains a minimum, and at a unique point). Then, using the estimate from Lemma 3.6 for strongly convex functions with constant strong convexity κ , we can write that $(\kappa/2)\|z^i - z^*\|^2 \leq f^0(z^i) - f^0(z^*)$, which implies the boundedness of the sequence z^i , $k = 1, 2, \dots$

Secondly, using standard arguments based on the weak compactness of a bounded convex closed set in a Hilbert space (see [34, Theorem 8.2.3]), the fact that weak convergence in such a space is equivalent to convergence in the so-called weak norm $|\cdot|_w$ (see [35, Theorem I.3.11]) and weak lower semicontinuity of continuous and convex functionals $f^0(z)$, $\|A^0 z - h^0 - p\|$, $g_k^0(z)$, $k = 1, \dots, m$, $z \in \mathcal{D}$ (see [34, Theorem 8.2.7]) we obtain that the elements z^i as $i \rightarrow \infty$ converge weakly to the solution $z_{p,r}^0$ of Problem $(P_{p,r}^0)$.

And finally, thirdly, since simultaneously $f^0(z^i) \rightarrow f^0(z_{p,r}^0)$ as $i \rightarrow \infty$, then, based on the necessary condition for strong convexity of Lemma 3.7 (see also [15, Lemma 3.2]), we can write in this case the inequality

$$f^0(z^i) \geq f^0(z_{p,r}^0) + \langle \partial f^0(z_{p,r}^0), z^i - z_{p,r}^0 \rangle + \frac{\kappa}{2} \|z^i - z_{p,r}^0\|^2 \quad \forall i = 1, 2, \dots,$$

from which we obtain the strong convergence of z^i to $z_{p,r}^0$ as $i \rightarrow \infty$. $\square$

The following important, in the context of this paper, assertion about the density of subdifferentiability in the sense of convex analysis also holds (see, for example, [3, Theorem 4.3]).

Lemma 3.9. *The subdifferential of a proper convex lower semicontinuous function $\varphi : \mathcal{H} \rightarrow \mathbb{R}^1 \cup \{+\infty\}$, where $\mathcal{H}$ is a Hilbert space, is nonempty at points of a set dense in $\text{dom } \varphi$.*

Below we use the notations $B_\epsilon(z) \equiv \{y \in Z : \|y - z\| \leq \epsilon\}$, $\partial B_\epsilon(z) \equiv \{y \in Z : \|y - z\| = \epsilon\}$.

Lemma 3.10. *If Condition A is satisfied, then the element $z^\delta[\lambda, \mu]$ depends continuously on $(\lambda, \mu) \in H \times \mathbb{R}_+^m$.*

Proof. Let $\epsilon > 0$ be an arbitrary fixed number. Due to the strong convexity of the Lagrange functional $L_{p,r}^\delta(z, \lambda, \mu)$, $z \in \mathcal{D}$, with a constant κ independent of δ , its continuity in z (due to the continuity on $\mathcal{D}$ of the input data, see Condition A), and Lemma 3.6, we can write

$$\frac{\kappa}{2} \|z - z^\delta[\lambda, \mu]\|^2 \leq L_{p,r}^\delta(z, \lambda, \mu) - L_{p,r}^\delta(z^\delta[\lambda, \mu], \lambda, \mu) \quad \forall z \in \mathcal{D} \cap \partial B_\epsilon(z^\delta[\lambda, \mu])$$

or

$$L_{p,r}^\delta(z^\delta[\lambda, \mu], \lambda, \mu) \leq L_{p,r}^\delta(z, \lambda, \mu) - \frac{\kappa}{2} \epsilon^2 \quad \forall z \in \mathcal{D} \cap \partial B_\epsilon(z^\delta[\lambda, \mu]).$$

Hence, for all $(\lambda', \mu') \in H \times \mathbb{R}_+^m$ such that $\|\lambda' - \lambda\| \leq \phi(\epsilon)$, $|\mu' - \mu| \leq \phi(\epsilon)$, $\phi(\epsilon) \rightarrow 0$, $\epsilon \rightarrow 0$,

$$L_{p,r}^\delta(z^\delta[\lambda, \mu], \lambda', \mu') \leq L_{p,r}^\delta(z, \lambda', \mu') - \frac{\kappa}{4} \epsilon^2 < L_{p,r}^\delta(z, \lambda', \mu') \quad \forall z \in \mathcal{D} \cap \partial B_\epsilon(z^\delta[\lambda, \mu]).$$

From the last inequality, due to the convexity of $\mathcal{D}$, strong convexity and continuity of $L_{p,r}^\delta(\cdot, \lambda', \mu')$ (see Condition A), we deduce that for any point $z \in \mathcal{D} \setminus B_\epsilon(z^\delta[\lambda, \mu])$, the strongly convex function of one variable $L_{p,r}^\delta((1-\alpha)z^\delta[\lambda, \mu] + \alpha z, \lambda', \mu')$, $\alpha \in [0, 1]$, at the point $\alpha = 0$ takes a value strictly smaller than the value at the point $\bar{\alpha} \in (0, 1)$, for which the equality holds $\|(1-\bar{\alpha})z^\delta[\lambda, \mu] + \bar{\alpha} z - z^\delta[\lambda, \mu]\| = \bar{\alpha} \|z - z^\delta[\lambda, \mu]\| = \epsilon$. The latter means that the strongly convex functional $L_{p,r}^\delta(\cdot, \lambda', \mu')$ on the interval $(1-\alpha)z^\delta[\lambda, \mu] + \alpha z \subset \mathcal{D}$, $\alpha \in [0, 1]$, for $\alpha = 0$ takes a value strictly smaller than its value on the same interval for $\alpha = \bar{\alpha}$. Thus, taking into account the well-known properties of strongly convex functions, we can state that $L_{p,r}^\delta(z^\delta[\lambda, \mu], \lambda', \mu') < L_{p,r}^\delta(z, \lambda', \mu')$ for any point $z \in \mathcal{D} \setminus B_\epsilon(z^\delta[\lambda, \mu])$, i.e. $z^\delta[\lambda', \mu'] \in \mathcal{D} \cap B_\epsilon(z^\delta[\lambda, \mu])$. Taking advantage of the arbitrariness of the choice of ϵ , we conclude that $z^\delta[\lambda', \mu'] \rightarrow z^\delta[\lambda, \mu]$ as $(\lambda', \mu') \rightarrow (\lambda, \mu)$. Thus, the lemma is proved. $\square$

Next, we obtain an important estimate for the deviation of the values of the perturbed objective function of the dual problem from the corresponding values of the unperturbed one.

Lemma 3.11. *In both the case of a bounded and an unbounded set $\mathcal{D}$, the following estimate holds*

$$|V_{p,r}^\delta(\lambda, \mu) - V_{p,r}^0(\lambda, \mu)| \leq C_V \delta (1 + \|\lambda\| + |\mu|), \quad (3.6)$$

in which the constant $C_V > 0$ in the case of a bounded $\mathcal{D}$ depends on $\sup_{z \in \mathcal{D}} \|z\|$, but does not depend on

$\delta, p, r, (\lambda, \mu) \in H \times \mathbb{R}^m$, and in the case of an unbounded $\mathcal{D}$, it depends on a constant $M > 0$ such that the norms of the elements $z^\delta[\lambda, \mu], z^0[\lambda, \mu]$ are bounded uniformly in $\delta \in [0, \delta_0)$ by a number M : $\|z^\delta[\lambda, \mu]\| \leq M, \|z^0[\lambda, \mu]\| \leq M$.

Proof. To prove this estimate, assume without loss of generality that $V_{p,r}^\delta(\lambda, \mu) \geq V_{p,r}^0(\lambda, \mu)$. Then we can write the following chain of equalities and inequalities

$$\begin{aligned}
|V_{p,r}^\delta(\lambda, \mu) - V_{p,r}^0(\lambda, \mu)| &= V_{p,r}^\delta(\lambda, \mu) - V_{p,r}^0(\lambda, \mu) \\
&= V_{p,r}^\delta(\lambda, \mu) - \inf_{z \in \mathcal{D} \cap B_M} (L_{p,r}^0(z, \lambda, \mu) - L_{p,r}^\delta(z, \lambda, \mu) + L_{p,r}^\delta(z, \lambda, \mu)) \\
&\leq V_{p,r}^\delta(\lambda, \mu) - V_{p,r}^\delta(\lambda, \mu) - \inf_{z \in \mathcal{D} \cap B_M} (L_{p,r}^0(z, \lambda, \mu) - L_{p,r}^\delta(z, \lambda, \mu)) \\
&= - \inf_{z \in \mathcal{D} \cap B_M} (L_{p,r}^0(z, \lambda, \mu) - L_{p,r}^\delta(z, \lambda, \mu)) \\
&\leq \sup_{z \in \mathcal{D} \cap B_M} |L_{p,r}^\delta(z, \lambda, \mu) - L_{p,r}^0(z, \lambda, \mu)|,
\end{aligned}$$

an obvious consequence of which, taking into account the estimates (3.2), is the estimate (3.6). $\square$

Next, we present a formula for the gradient of the functional $V_{p,r}^\delta$ on $H \times \mathbb{R}_+^m$ as $\delta \geq 0$. Its proof can be found in [18, Lemma 8].

Lemma 3.12. *The Frechét derivative of the functional $V_{p,r}^\delta$, considered only on the set $H \times \mathbb{R}_+^m$, at an interior point (λ, μ) of this set, is equal to*

$$\partial V_{p,r}^\delta(\lambda, \mu) = (A^\delta z^\delta[\lambda, \mu] - h^\delta - p, g^\delta(z^\delta[\lambda, \mu]) - r).$$

Simultaneously, for all points $(\lambda, \mu) \in H \times \mathbb{R}_+^m$, the inclusion holds

$$(A^\delta z^\delta[\lambda, \mu] - h^\delta - p, g^\delta(z^\delta[\lambda, \mu]) - r) \in \partial V_{p,r}^\delta(\lambda, \mu),$$

where $\partial V_{p,r}^\delta(\lambda, \mu)$ is the superdifferential in the sense of convex analysis of the functional $V_{p,r}^\delta$ at the point $(\lambda, \mu) \in H \times \mathbb{R}_+^m$. The element $(A^\delta z^\delta[\lambda, \mu] - h^\delta - p, g^\delta(z^\delta[\lambda, \mu]) - r)$ depends on (λ, μ) on the set $H \times \mathbb{R}_+^m$ continuously.

The following lemma is a special case of the convergence theorem [34, Theorem 9.4.2] of Tikhonov's stabilization method for a general convex optimization problem [32].

Lemma 3.13. *Let $\mathcal{H}$ be a Hilbert space, the set $X \subseteq \mathcal{H}$ be convex and closed, the function $\varphi : X \rightarrow \mathbb{R}^1$ be convex and continuous on X , and the set X^* of its minimum points on X be nonempty. Let there also be given a family of convex continuous functions $\varphi^\gamma : X \rightarrow \mathbb{R}^1$ depending on a numerical parameter $\gamma \in (0, \gamma_0]$, $\gamma_0 > 0$, such that $|\varphi^\gamma(z) - \varphi(z)| \leq \gamma(1 + \|z - \tilde{z}\|^2) \forall z \in X$, where $\tilde{z} \in \mathcal{H}$ is an arbitrary fixed element. In this case, if*

$$z_{\gamma,\alpha} = \operatorname{argmin}\{\varphi^\gamma(z) + \alpha\|z - \tilde{z}\|^2, z \in X\}, \alpha > 0,$$

then, under the condition of consistent tending to zero $\gamma/\alpha \rightarrow 0, \gamma \rightarrow 0, \alpha \rightarrow 0$, the limit relation $z_{\gamma,\alpha} \rightarrow z^*$ holds, where $z^* \in X^*$ is a point that provides a minimum in the problem $\|z - \tilde{z}\|^2 \rightarrow \min, z \in X^*$.

Finally, the statement of the following lemma is a special case of the more general statement contained in Corollary 4.3.6(c) in [2].

Lemma 3.14. *Let $\mathcal{H}$ be a Hilbert space, $f : \mathcal{H} \rightarrow \mathbb{R}^1 \cup \{+\infty\}$ be a proper lower semicontinuous convex function, $\Omega \subset \mathcal{H}$ is a closed convex set. In this case, if $0 \in \operatorname{int}(\Omega - \operatorname{dom} f)$, then $x \in \Omega$ is a minimum point of f on Ω if and only if $0 \in \partial f(x) + N_\Omega(x)$.*

4. NONCLASSICAL OPTIMALITY CONDITIONS IN A CONVEX CONSTRAINED EXTREMAL PROBLEM WITH A STRONGLY CONVEX OBJECTIVE FUNCTIONAL

4.1. Necessary optimality condition in the form of the dual regularization algorithm. The main goal of this Subsection is to obtain a necessary optimality condition for Problem $(P_{p,r}^0)$ in the form of a regularized dual algorithm, previously proposed in [26, 27] (see also [15, 16, 17]). The presence of the parameters p, r in the problem means that it is possible to study the properties of this necessary optimality condition depending on the subdifferential properties of the value function β . In other words, the dual regularization algorithm is considered below from the perspective of the perturbation method [1, Item 3.3.2].

From a formal point of view, the dual regularization method is a sequence of actions according to which, in a constrained minimization problem, the concave dual problem is solved using the standard Tikhonov stabilization procedure (see [34, Chapter 9]) and, as a consequence, approximations to the exact solution $z_{p,r}^0$ of the original Problem $(P_{p,r}^0)$ are constructed. For a bounded set of admissible elements $\mathcal{D}$, such an approach allows one to rely on the corresponding convergence theorems of the Tikhonov stabilization method when justifying the convergence of the dual method; see Lemma 3.13, as well as [34, Theorem 9.4.2]. In this case, the required a priori estimates for the deviation of the objective function values of the dual problem depending on the perturbation of the input data are used, following from the inequality (3.6) of Lemma 3.11. In the case of an unbounded $\mathcal{D}$, there are no such a priori estimates for the dual problem, and for this reason, it is impossible to use standard convergence theorems of the stabilization method to justify the convergence of the dual regularization procedure. Therefore, below, to solve the convergence problem of this dual procedure in the case of an unbounded $\mathcal{D}$, relying on the special properties of the primal and dual problems, we apply the corresponding result from [18, Item 2.2.2].

4.1.1. Basic auxiliary estimates, the relationship between the regularization parameter of the dual problem and the dual variables. Let Problem $(P_{p,r}^0)$ be solvable. Let us denote by $(\lambda_{p,r}^{\delta,\alpha}, \mu_{p,r}^{\delta,\alpha})$ the only point in $H \times \mathbb{R}_+^m$ that gives a maximum to the strongly concave functional on this set for $\alpha > 0$

$$R_{p,r}^{\delta,\alpha}(\lambda, \mu) \equiv V_{p,r}^\delta(\lambda, \mu) - \alpha \|\lambda - \tilde{\lambda}\|^2 - \alpha |\mu - \tilde{\mu}|^2, \quad (\lambda, \mu) \in H \times \mathbb{R}_+^m,$$

where $(\tilde{\lambda}, \tilde{\mu}) \in H \times \mathbb{R}^m$ is an arbitrary fixed element.

Let the matching condition be satisfied

$$\frac{\delta}{\alpha(\delta)} \rightarrow 0, \quad \alpha(\delta) \rightarrow 0, \quad \delta \rightarrow 0. \quad (4.1)$$

We show that the regularized elements $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$ strongly converge to the solution $z_{p,r}^0$ of Problem $(P_{p,r}^0)$ as $\delta \rightarrow 0$. It is these elements that were designated in Introduction as $z^{\delta,\alpha(\delta)}$.

First of all, note that, taking into account Lemma 3.12 and Lemma 3.14, we can write the inequality

$$\langle (A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p, g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r) - 2\alpha(\delta)(\lambda_{p,r}^{\delta,\alpha(\delta)} - \tilde{\lambda}, \mu_{p,r}^{\delta,\alpha(\delta)} - \tilde{\mu}), \quad (4.2)$$

$$(\lambda, \mu) - (\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rangle \leq 0 \quad \forall (\lambda, \mu) \in H \times \mathbb{R}_+^m.$$

Indeed, if $(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})$ is an interior point of the set $H \times \mathbb{R}_+^m$, then this inequality is obvious due to the fact that the first factor in the scalar product in (4.2) is equal to zero. If $(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})$ is a boundary point, then for $\epsilon > 0$ we consider the auxiliary problem

$$R_{p,r}^{\delta,\alpha(\delta)}(\lambda, \mu) \rightarrow \sup, \quad (\lambda, \mu) \in H \times \mathbb{R}_{+\epsilon}^m,$$

$$\mathbb{R}_{+\epsilon}^m \equiv \{x = (x_1, \dots, x_m) \in \mathbb{R}_+^m : x_i \geq \epsilon, i = 1, \dots, m\}.$$

An admissible element it will be the element $(\lambda^\epsilon, \mu^\epsilon) = Pr_{H \times \mathbb{R}_+^m}(\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)})$, which as $\epsilon \rightarrow 0$ tends to $(\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)})$. Then, due to the continuity of $V_{p,r}^\delta$ on $H \times \mathbb{R}_+^m$ (this follows from the equality $V_{p,r}^\delta(\lambda, \mu) = L_{p,r}^\delta(z^\delta[\lambda, \mu], \lambda, \mu)$ by Lemma 3.10) we have the inequality

$$\sup_{(\lambda, \mu) \in H \times \mathbb{R}_+^m} R_{p,r}^{\delta, \alpha(\delta)}(\lambda, \mu) - \phi(\epsilon) \leq R_{p,r}^{\delta, \alpha(\delta)}(\lambda^\epsilon, \mu^\epsilon), \quad \phi(\epsilon) \rightarrow 0, \quad \epsilon \rightarrow 0.$$

Due to this inequality, the continuity of the functional $R_{p,r}^{\delta, \alpha(\delta)}$, its boundedness from above on $H \times \mathbb{R}_+^m$ and the Ekeland variational principle [6] for each sufficiently small ϵ there exists a pair $(\bar{\lambda}^\epsilon, \bar{\mu}^\epsilon) \in H \times \mathbb{R}_+^m$ satisfying the inequalities

$$\|\lambda^\epsilon - \bar{\lambda}^\epsilon\| + |\mu^\epsilon - \bar{\mu}^\epsilon| \leq \sqrt{\phi(\epsilon)}, \quad R_{p,r}^{\delta, \alpha(\delta)}(\bar{\lambda}^\epsilon, \bar{\mu}^\epsilon) \geq R_{p,r}^{\delta, \alpha(\delta)}(\lambda^\epsilon, \mu^\epsilon)$$

and such that it is a solution to the problem

$$-R_{p,r}^{\delta, \alpha(\delta)}(\lambda, \mu) + \sqrt{\phi(\epsilon)}(\|\lambda - \bar{\lambda}^\epsilon\| + |\mu - \bar{\mu}^\epsilon|) \rightarrow \min, \quad (\lambda, \mu) \in H \times \mathbb{R}_+^m.$$

Then, based on the minimum criterion in Lemma 3.14, we can assert that the corresponding optimality conditions are satisfied.

$$\langle -(A^\delta z^\delta[\bar{\lambda}^\epsilon, \bar{\mu}^\epsilon] - h^\delta - p, g^\delta(z^\delta[\bar{\lambda}^\epsilon, \bar{\mu}^\epsilon]) - r) + 2\alpha(\delta)(\bar{\lambda}^\epsilon - \tilde{\lambda}, \bar{\mu}^\epsilon - \tilde{\mu}) + \sqrt{\phi(\epsilon)}(\zeta^\epsilon, \theta^\epsilon),$$

$$(\lambda, \mu) - (\bar{\lambda}^\epsilon, \bar{\mu}^\epsilon) \rangle \geq 0 \quad \forall (\lambda, \mu) \in H \times \mathbb{R}_+^m, \quad \|\zeta^\epsilon\| \leq 1, \quad |\theta^\epsilon| \leq 1.$$

Passing to the limit in these relations as $\epsilon \rightarrow 0$ based on the continuity of $z^\delta[\lambda, \mu]$ (see Lemma 3.10) on the set $H \times \mathbb{R}_+^m$, we conclude that inequality (4.2) is also true for the boundary points $(\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)})$.

The inequality (4.2) directly implies the relations

$$A^\delta z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}] - h^\delta - p = 2\alpha(\delta)(\lambda_{p,r}^{\delta, \alpha(\delta)} - \tilde{\lambda}), \quad (4.3)$$

$$\langle g^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r - 2\alpha(\delta)(\mu_{p,r}^{\delta, \alpha(\delta)} - \tilde{\mu}), \mu - \mu_{p,r}^{\delta, \alpha(\delta)} \rangle \leq 0 \quad \forall \mu \in \mathbb{R}_+^m. \quad (4.4)$$

From (4.4) it follows that if $\mu_{p,r,j}^{\delta, \alpha(\delta)} > 0$ for some $j \in \{1, \dots, m\}$, then

$$g_j^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r_j - 2\alpha(\delta)(\mu_{p,r,j}^{\delta, \alpha(\delta)} - \tilde{\mu}_j) = 0. \quad (4.5)$$

If $\mu_{p,r,j}^{\delta, \alpha(\delta)} = 0$ for some $j \in \{1, \dots, m\}$, then from (4.4) it follows that

$$g_j^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r_j + 2\alpha(\delta)\tilde{\mu}_j \leq 0. \quad (4.6)$$

From (4.3) and (4.5) we get equality

$$\begin{aligned} & \langle (A^\delta z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}] - h^\delta - p, g^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r), (\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}) \rangle \\ &= 2\alpha(\delta)(\|\lambda_{p,r}^{\delta, \alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta, \alpha(\delta)}|^2) - 2\alpha(\delta)\langle \lambda_{p,r}^{\delta, \alpha(\delta)}, \tilde{\lambda} \rangle - 2\alpha(\delta)\langle \mu_{p,r}^{\delta, \alpha(\delta)}, \tilde{\mu} \rangle \end{aligned} \quad (4.7)$$

and, as a consequence, the inequality

$$\begin{aligned} & \langle (A^\delta z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}] - h^\delta - p, g^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r), (\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}) \rangle \\ &+ 2\alpha(\delta)\langle \lambda_{p,r}^{\delta, \alpha(\delta)}, \tilde{\lambda} \rangle + 2\alpha(\delta)\langle \mu_{p,r}^{\delta, \alpha(\delta)}, \tilde{\mu} \rangle \geq 0. \end{aligned} \quad (4.8)$$

Next, since

$$\begin{aligned} & L_{p,r}^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}], \lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}) \\ & \equiv f^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) + \langle \lambda_{p,r}^{\delta, \alpha(\delta)}, A^\delta z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}] - h^\delta - p \rangle + \langle \mu_{p,r}^{\delta, \alpha(\delta)}, g^\delta(z^\delta[\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}]) - r \rangle \\ & \leq L_{p,r}^\delta(z_{p,r}^0, \lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)}) \equiv f^\delta(z_{p,r}^0) + \langle \lambda_{p,r}^{\delta, \alpha(\delta)}, A^\delta z_{p,r}^0 - h^\delta - p \rangle + \langle \mu_{p,r}^{\delta, \alpha(\delta)}, g^\delta(z_{p,r}^0) - r \rangle \\ & \leq f^0(z_{p,r}^0) + (f^\delta(z_{p,r}^0) - f^0(z_{p,r}^0)) + \|\lambda_{p,r}^{\delta, \alpha(\delta)}\| \|A^\delta z_{p,r}^0 - A^0 z_{p,r}^0 + h^0 - h^\delta\| + |\mu_{p,r}^{\delta, \alpha(\delta)}| |g^\delta(z_{p,r}^0) - g^0(z_{p,r}^0)|, \end{aligned}$$

then by virtue of the estimates (3.2) we obtain

$$\begin{aligned} & f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \\ & + \langle \lambda_{p,r}^{\delta,\alpha(\delta)}, A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p \rangle + \langle \mu_{p,r}^{\delta,\alpha(\delta)}, g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r \rangle \\ & \leq f^0(z_{p,r}^0) + C\delta(1 + \|z_{p,r}^0\|^2) + C\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|(\delta\|z_{p,r}^0\| + \delta + \delta) + C|\mu_{p,r}^{\delta,\alpha(\delta)}|(\delta(1 + \|z_{p,r}^0\|^2)), \end{aligned} \quad (4.9)$$

and therefore, taking into account the inequality (4.8),

$$\begin{aligned} & f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \\ & \leq f^0(z_{p,r}^0) + C\delta(1 + \|z_{p,r}^0\|^2) + C\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|(\delta\|z_{p,r}^0\| + \delta + \delta) + C|\mu_{p,r}^{\delta,\alpha(\delta)}|(\delta(1 + \|z_{p,r}^0\|^2)) \\ & \quad + 2\alpha(\delta)\langle \lambda_{p,r}^{\delta,\alpha(\delta)}, \tilde{\lambda} \rangle + 2\alpha(\delta)\langle \mu_{p,r}^{\delta,\alpha(\delta)}, \tilde{\mu} \rangle. \end{aligned} \quad (4.10)$$

Rewriting also inequality (4.9) as

$$\begin{aligned} & \langle \lambda_{p,r}^{\delta,\alpha(\delta)}, A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p \rangle + \langle \mu_{p,r}^{\delta,\alpha(\delta)}, g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r \rangle \\ & \leq f^0(z_{p,r}^0) + C\delta(1 + \|z_{p,r}^0\|^2) + C\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|(\delta\|z_{p,r}^0\| + \delta + \delta) + C|\mu_{p,r}^{\delta,\alpha(\delta)}|(\delta(1 + \|z_{p,r}^0\|^2)) \\ & \quad - f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]), \end{aligned}$$

we get from it, in turn, taking into account (4.7),

$$\begin{aligned} & 2\alpha(\delta)(\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta,\alpha(\delta)}|^2) \\ & \leq f^0(z_{p,r}^0) + C\delta(1 + \|z_{p,r}^0\|^2) + C\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|(\delta\|z_{p,r}^0\| + \delta + \delta) + C|\mu_{p,r}^{\delta,\alpha(\delta)}|(\delta(1 + \|z_{p,r}^0\|^2)) \\ & \quad + 2\alpha(\delta)\langle \lambda_{p,r}^{\delta,\alpha(\delta)}, \tilde{\lambda} \rangle + 2\alpha(\delta)\langle \mu_{p,r}^{\delta,\alpha(\delta)}, \tilde{\mu} \rangle - f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \\ & \leq f^0(z_{p,r}^0) + C\delta(1 + \|z_{p,r}^0\|^2) + \sqrt{2}C_1(\delta)\sqrt{\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta,\alpha(\delta)}|^2} - f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]), \end{aligned} \quad (4.11)$$

where

$$C_1(\delta) \equiv \max\{C(\delta\|z_{p,r}^0\| + \delta + \delta), C\delta(1 + \|z_{p,r}^0\|^2), 2\alpha(\delta)\|\tilde{\lambda}\|, 2\alpha(\delta)\|\tilde{\mu}\|\},$$

i.e.

$$\begin{aligned} & 2\alpha(\delta)(\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta,\alpha(\delta)}|^2) \\ & - \sqrt{2}C_1(\delta)\sqrt{\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta,\alpha(\delta)}|^2} - f^0(z_{p,r}^0) - C\delta(1 + \|z_{p,r}^0\|^2) + f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \leq 0. \end{aligned}$$

From the last estimate, we obtain the estimate

$$\begin{aligned} & \sqrt{\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 + |\mu_{p,r}^{\delta,\alpha(\delta)}|^2} \\ & \leq \frac{\sqrt{2}C_1(\delta) + \sqrt{2(C_1(\delta))^2 - 8\alpha(\delta)(f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - f^0(z_{p,r}^0) - C\delta(1 + \|z_{p,r}^0\|^2))}}{4\alpha(\delta)}, \end{aligned}$$

from which, in turn, it follows that

$$\begin{aligned} & \alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\| + \alpha(\delta)|\mu_{p,r}^{\delta,\alpha(\delta)}| \leq \sqrt{2}C_1(\delta) \\ & + \sqrt{2(C_1(\delta))^2 - 8\alpha(\delta)(f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - f^0(z_{p,r}^0) - C\delta(1 + \|z_{p,r}^0\|^2))} \\ & \leq \sqrt{2}C_1(\delta) + \sqrt{2(C_1(\delta))^2 - 8\alpha(\delta)(f^\delta(z_\delta^*) - f^0(z_{p,r}^0) - C\delta(1 + \|z_{p,r}^0\|^2))}, \end{aligned} \quad (4.12)$$

where $z_\delta^* = z^\delta[0, 0] \equiv \operatorname{argmin}\{f^\delta(z), z \in \mathcal{D}\}$.

Remark 4.1. It is important to note that all relations (4.3) – (4.12) are valid for any $\delta \in [0, \delta_0]$.

From (4.12), provided that the quantity $f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}])$ is bounded below uniformly in $\delta \in [0, \delta_0]$, we obtain the limit relations

$$\alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\| \rightarrow 0, \quad \alpha(\delta)|\mu_{p,r}^{\delta,\alpha(\delta)}| \rightarrow 0, \quad \delta \rightarrow 0. \quad (4.13)$$

Let us show that the quantity $f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}])$ is bounded from below uniformly in $\delta \in [0, \delta_0]$ without loss of generality. More precisely, we show that without loss of generality (for all sufficiently small δ), the following estimate holds (recall that $z_\delta^* = z^\delta[0, 0] \equiv \operatorname{argmin}\{f^\delta(z), z \in \mathcal{D}\}$)

$$f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \geq f^\delta(z_\delta^*) \geq K \quad \forall \delta \in [0, \delta_0], \quad (4.14)$$

where K does not depend on δ , but depends, strictly speaking, on the strong convexity constant κ . The last estimate leads, by virtue of (4.12), to the estimate

$$\begin{aligned} & \alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\| + \alpha(\delta)|\mu_{p,r}^{\delta,\alpha(\delta)}| \\ & \leq \sqrt{2C_1(\delta)} + \sqrt{2(C_1(\delta))^2 - 8\alpha(\delta)(K - f^0(z_{p,r}^0) - C\delta(1 + \|z_{p,r}^0\|^2))}. \end{aligned} \quad (4.15)$$

Obviously, to prove the inequality (4.14) holds for all $\delta \in [0, \delta_0]$, without loss of generality, it suffices to show that the elements z_δ^* are uniformly bounded for all sufficiently small δ . To this end, we show, using the estimates (3.2), that $z_\delta^* \rightarrow z_0^*$ as $\delta \rightarrow 0$.

Indeed, let $\varepsilon > 0$ be an arbitrary fixed number. Taking into account the strong convexity of the functional f^δ with a constant κ independent of δ and its continuity (Condition A), we can write (see Lemma 3.6)

$$\frac{\kappa}{2}\|z - z_0^*\|^2 \leq f^0(z) - f^0(z_0^*) \quad \forall z \in \mathcal{D} \cap \partial B_\varepsilon(z_0^*)$$

or

$$f^0(z_0^*) \leq f^0(z) - \frac{\kappa}{2}\varepsilon^2 \quad \forall z \in \mathcal{D} \cap \partial B_\varepsilon(z_0^*),$$

where, as in the proof of Lemma 3.10, $B_\varepsilon(z) \equiv \{z' \in Z : \|z' - z\| \leq \varepsilon\}$, $\partial B_\varepsilon(z) \equiv \{z' \in Z : \|z' - z\| = \varepsilon\}$. Hence, by virtue of the first estimate (3.2), for all $\delta \leq \phi(\varepsilon)$, $\phi(\varepsilon) \rightarrow 0$, $\varepsilon \rightarrow 0$,

$$f^\delta(z_0^*) \leq f^\delta(z) - \frac{1}{4}\kappa\varepsilon^2 < f^\delta(z) \quad \forall z \in \mathcal{D} \cap \partial B_\varepsilon(z_0^*).$$

From the last inequality, due to the convexity of $\mathcal{D}$, strong convexity, and continuity of f^δ (see Condition A), arguing as in the proof of Lemma 3.10, we obtain

$$f^\delta(z_0^*) < f^\delta(z) \quad \forall z \in \mathcal{D} \setminus B_\varepsilon(z_0^*)$$

that is, $z_\delta^* \in \mathcal{D} \cap B_\varepsilon(z_0^*)$. Taking advantage of the arbitrariness of the choice of ε , we conclude that $z_\delta^* \rightarrow z_0^*$ as $\delta \rightarrow 0$.

So, summing up the above, we can say that the estimate (4.14), without loss of generality for all $\delta \in [0, \delta_0]$, and along with it the estimate (4.15) and as consequence the limit relations (4.13) are fully justified.

4.1.2. Estimates of the deviation of approximate solutions on function and constraints. First of all, note that by (4.3) – (4.6) and the limit relations (4.13), we can write

$$A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p = 2\alpha(\delta)(\lambda_{p,r}^{\delta,\alpha(\delta)} - \tilde{\lambda}) \rightarrow 0, \quad \delta \rightarrow 0, \quad (4.16)$$

$$g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r \leq 2\alpha(\delta)(\mu_{p,r}^{\delta,\alpha(\delta)} - \tilde{\mu}) \rightarrow 0, \quad \delta \rightarrow 0,$$

where the equality and the inequality in (4.16) hold without loss of generality for any $\delta \in (0, \delta_0]$.

Further, from the inequality (4.10) we obtain for all $\delta \in (0, \delta_0]$

$$f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \leq f^0(z_{p,r}^0) + \psi(\delta), \quad (4.17)$$

where

$$\begin{aligned} \psi(\delta) & \equiv C\delta(1 + \|z_{p,r}^0\|^2) + C(\delta\|z_{p,r}^0\| + \delta + \delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\| + C\delta(1 + \|z_{p,r}^0\|^2)|\mu_{p,r}^{\delta,\alpha(\delta)}| \\ & \quad + 2\alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|\|\tilde{\lambda}\| + 2\alpha(\delta)|\mu_{p,r}^{\delta,\alpha(\delta)}|\|\tilde{\mu}\| \geq 0, \end{aligned} \quad (4.18)$$

i.e., $\psi(\delta) \rightarrow 0$, $\delta \rightarrow 0$ by virtue of the limit relations (4.13) and the matching condition (4.1). At the same time, in view of the well-known estimate for strongly convex functions of Lemma 3.6 we can write for the functional f^δ with a strong convexity constant κ independent of δ

$$\frac{\kappa}{2} \|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - z_\delta^*\|^2 \leq f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - f^\delta(z_\delta^*), \quad (4.19)$$

where $z_\delta^* \equiv \operatorname{argmin}\{f^\delta(z), z \in \mathcal{D}\}$, $\delta \in [0, \delta_0]$.

Since it was shown above in the proof of uniform on δ estimate (4.14) that $z_\delta^* \rightarrow z_0^*$ as $\delta \rightarrow 0$, then, by virtue of the estimates (4.17) and (4.19), the family of the elements $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, $\delta \in (0, \delta_0]$, can also be considered uniformly bounded on δ , i.e., without loss of generality, the estimate holds

$$\|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]\| \leq L \quad \forall \delta \in (0, \delta_0], \quad (4.20)$$

in which the constant $L > 0$ does not depend on δ , but depends, generally speaking, on the strong convexity constant κ .

Due to this boundedness and the first of the estimates (3.2), we derive from the estimate (4.17)

$$\begin{aligned} f^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) &\leq f^0(z_{p,r}^0) + \psi(\delta) + f^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - f^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \\ &\leq f^0(z_{p,r}^0) + \psi(\delta) + C\delta(1 + \|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]\|^2) \leq f^0(z_{p,r}^0) + \psi_1(\delta), \end{aligned} \quad (4.21)$$

where $\psi_1(\delta) \equiv \psi(\delta) + C\delta(1 + L^2) \geq 0$, $\psi_1(\delta) \rightarrow 0$, $\delta \rightarrow 0$, and the quantity $\psi(\delta)$ is given by the expression (4.18).

In the obtained estimate (4.21), the quantity $\psi_1(\delta)$ is bounded from above using the quantity L , which implicitly depends on the strong convexity constant κ and, by virtue of the estimate (4.20), bounds from above the norm of elements from the family $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, $\delta \in (0, \delta_0]$.

Further, taking into account the estimates (3.2), the equalities (4.3), (4.5) and the proven boundedness of the elements $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, by virtue of (4.16) we can write for any $\delta \in (0, \delta_0]$

$$\begin{aligned} &\|A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^0 - p\| \\ &\leq \|A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p\| + \|A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]\| + \|h^0 - h^\delta\| \\ &\leq \|A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p\| + C\delta(1 + \|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]\|) + C\delta \\ &\leq 2\alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)} - \tilde{\lambda}\| + C(\delta(1 + L) + \delta) \rightarrow 0, \quad \delta \rightarrow 0, \\ &\quad g_i^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r_i \\ &= g_i^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r_i + g_i^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - g_i^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \\ &\leq g_i^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r_i + C\delta(1 + \|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]\|^2) \\ &\leq 2\alpha(\delta)\|\mu_{p,r}^{\delta,\alpha(\delta)} - \tilde{\mu}\| + C\delta(1 + L^2) \equiv \phi_1(\delta), \quad \phi_1(\delta) \rightarrow 0, \quad \delta \rightarrow 0, \quad i = 1, \dots, m. \end{aligned} \quad (4.22)$$

Thus, we have obtained estimates for the deviation of approximate solutions on function and on constraints in which the value L , by virtue of the estimate (4.20), limits from above the norms of elements from the family $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, $\delta \in (0, \delta_0]$, and implicitly depends on the constant strong convexity κ .

4.1.3. Convergence of approximate solutions on function, constraints, and argument. We now show the convergence as $\delta \rightarrow 0$ of the elements $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$ to $z_{p,r}^0$ on function and on constraints, and, as a consequence, in the case of subdifferentiability of f^0 , also on argument. By virtue of the estimate (4.21), the relations (4.22), the matching condition (4.1), and the limit relations (4.13), for any sequence δ_s , $s = 1, 2, \dots$, $\delta_s \rightarrow 0$, $s \rightarrow \infty$, we obtain the relations

$$\begin{aligned} f^0(z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}]) &\leq f^0(z_{p,r}^0) + \psi_1(\delta_s), \quad \|A^0 z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}] - h^0 - p\| \rightarrow 0, \\ g_i^0(z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}]) - r_i &\leq \phi_1(\delta_s), \quad \phi_1(\delta_s) \rightarrow 0, \quad i = 1, \dots, m, \quad \delta_s \rightarrow 0, \quad s \rightarrow \infty. \end{aligned}$$

From these relations it follows that any sequence $z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}]$, $s = 1, 2, \dots$, is a GMS in Problem $(P_{p,r}^0)$ and, therefore, by virtue of the limit relation (3.1) and the first statement of Lemma 3.8 we have the limit relations

$$f^0(z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}]) \rightarrow f^0(z_{p,r}^0), \quad z^{\delta_s}[\lambda_{p,r}^{\delta_s,\alpha(\delta_s)}, \mu_{p,r}^{\delta_s,\alpha(\delta_s)}] \rightarrow z_{p,r}^0 \quad \text{weakly in } Z, \quad \delta_s \rightarrow 0,$$

whence, due to the arbitrariness of the choice of the sequence δ_s , $s = 1, 2, \dots$, $\delta_s \rightarrow 0$, $s \rightarrow \infty$ we obtain the limit relation

$$f^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \rightarrow f^0(z_{p,r}^0), \quad \delta \rightarrow 0, \quad (4.23)$$

and also, if we take into account the fact that weak convergence in a Hilbert space is equivalent to convergence in the so-called weak norm $|\cdot|_w$ (see [35, Theorem I.3.11]), weak convergence also

$$z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] \rightarrow z_{p,r}^0 \text{ weakly in } Z, \quad \delta \rightarrow 0.$$

Now assume that the strongly convex functional f^0 is subdifferentiable (in the sense of convex analysis) at the points of $\mathcal{D}$. Then, by the second assertion of Lemma 3.8, we finally obtain

$$\|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - z_{p,r}^0\| \rightarrow 0, \quad \delta \rightarrow 0.$$

Furthermore, the boundedness of the family $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, $\delta \in (0, \delta_0]$, the limit relation (4.23), estimates (4.11), and limit relations (4.13) lead to the limit relations

$$\alpha(\delta)\|\lambda_{p,r}^{\delta,\alpha(\delta)}\|^2 \rightarrow 0, \quad \alpha(\delta)\|\mu_{p,r}^{\delta,\alpha(\delta)}\|^2 \rightarrow 0, \quad \delta \rightarrow 0. \quad (4.24)$$

Finally, a consequence of the limit relations (4.24) and the equality (4.7) is the limit relation

$$\langle (A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p, g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}])) - r, (\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rangle \rightarrow 0, \quad \delta \rightarrow 0.$$

4.1.4. Convergence of regularized solutions of the dual problem. Next, we consider the convergence of regularized solutions $(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})$ on function and on argument of the problem that is dual to Problem $(P_{p,r}^0)$. Convergence on argument in the situation of solvability of the dual problem under the boundedness of $\mathcal{D}$ is proved by a simple reference to the corresponding convergence theorem of Tikhonov's stabilization method, see Lemma 3.13. This is made possible by the estimate (3.6), which, due to the boundedness of $\mathcal{D}$, implies an estimate for the deviation of the objective function of the dual problem under perturbations of the input data, which is necessary for applying the stabilization method. In the same case of solvability of the dual problem, but with unbounded $\mathcal{D}$, we do not have the required estimate of the deviation of the objective function of the dual problem under perturbations of the input data. Therefore, convergence on argument of the dual procedure in this situation is justified in [18], taking into account the special properties of the primal and dual problems proved above.

4.1.4.1. Convergence of regularized solutions of the dual problem on function. In parallel with the construction of the GMS in the original problem $(P_{p,r}^0)$, the dual algorithm under discussion simultaneously produces a maximizing sequence in the corresponding dual problem. Namely, it can be asserted that the limit relation and equality hold:

$$\lim_{\delta \rightarrow +0} V_{p,r}^0(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) = \sup_{(\lambda,\mu) \in H \times \mathbb{R}_+^m} V_{p,r}^0(\lambda, \mu) = f^0(z_{p,r}^0), \quad (4.25)$$

the proof of which can be found in [15, Subsection 3.3]; see the proof of equality (3.33) in [15].

Remark 4.2. The equality

$$\sup_{(\lambda,\mu) \in H \times \mathbb{R}_+^m} V_{p,r}^0(\lambda, \mu) = f^0(z_{p,r}^0)$$

can also be considered a consequence of the classical nonsymmetric minimax theorem in [2, Chapter 6, §2, Theorem 7].

We also note that, by virtue of the proved limit relation (4.25), if the functional $V_{p,r}^0$ does not attain its maximum on $H \times \mathbb{R}_+^m$, then, obviously, the limit relation $\|(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})\| \rightarrow \infty$, $\delta \rightarrow 0$ is valid.

4.1.4.2. Convergence of regularized solutions of the dual problem on argument. If the functional $V_{p,r}^0$ has a maximum point on $H \times \mathbb{R}_+^m$, then in the case of bounded $\mathcal{D}$, by virtue of the estimate (3.6) and the matching condition (4.1), the regularization procedure of the dual problem, which is the usual Tikhonov regularization procedure for a general maximization problem, in accordance with Lemma 3.13, yields strong convergence (see also [34, Theorem 9.4.2])

$$(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rightarrow (\lambda_{p,r}^0, \mu_{p,r}^0), \quad \delta \rightarrow 0, \quad (4.26)$$

where $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$ is the element that provides a minimum to the functional $\|(\lambda, \mu) - (\tilde{\lambda}, \tilde{\mu})\|^2$, $(\lambda, \mu) \in H \times \mathbb{R}_+^m$ among all elements delivering the maximum functional $V_{p,r}^0$. Thus, due to the arbitrariness in the choice of the point $(\tilde{\lambda}, \tilde{\mu}) \in H \times \mathbb{R}_+^m$, in the case of solvability of the problem that is dual to $(P_{p,r}^0)$, we can

assume without loss of generality that the limit relation $(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rightarrow (\lambda_{p,r}^0, \mu_{p,r}^0)$, $\delta \rightarrow 0$ holds, where $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$ can be taken to be any pre-chosen and fixed solution of the dual problem, for example, normal (i.e., minimal in norm) in the case $(\tilde{\lambda}, \tilde{\mu}) = 0$. Moreover, since the set of all solutions of the dual problem, or, in other words, the set of all Kuhn-Tucker vectors for Problem $(P_{p,r}^0)$, coincides with subdifferential $\partial\beta(p, r)$ taken with the opposite sign, i.e., with $-\partial\beta(p, r)$ (see, for example, [24, Theorem 1.1]), any pre-selected and fixed element from $-\partial\beta(p, r)$ can be taken as $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$.

If there is no boundedness condition for $\mathcal{D}$, then we cannot directly apply Lemma 3.13 on the convergence of the stabilization method to prove the limit relation (4.26) due to the absence in this case of the estimate needed to apply the indicated lemma, which follows in the case of bounded $\mathcal{D}$ from the estimate (3.6). Nevertheless, as shown in [18, Subsection 2.2.2], everything said regarding the convergence on argument of regularized solutions of the problem dual to $(P_{p,r}^0)$ remains valid even in the absence of the boundedness condition for $\mathcal{D}$. This is a consequence of the convergence properties of the duality-based algorithm proved above and, in particular, the boundedness of the family $z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]$, $\delta \in (0, \delta_0)$, established above.

4.1.5. *A nonclassical necessary optimality condition in the form of a dual regularization algorithm.* As a result of the above reasoning, we formulate a nonclassical necessary optimality condition for Problem $(P_{p,r}^0)$ in the form of a duality-based regularization algorithm. This result can be used as a convergence theorem for the dual regularization algorithm.

Theorem 4.3. *[A nonclassical necessary optimality condition in the form of a dual regularization algorithm] Suppose that Problem $(P_{p,r}^0)$ is solvable, Condition A is satisfied, estimates (3.2) for the deviation of perturbed input data from exact ones hold, and the consistency condition (4.1) holds. Let also $\delta \in [0, \delta_0]$, $k = 1, 2, \dots$, $\delta_0 > 0$ is a sufficiently small fixed number. Then:*

1) *The following relations hold:*

$$\begin{aligned} f^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) &\rightarrow f^0(z_{p,r}^0), \quad A^0 z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^0 - p \rightarrow 0, \quad \delta \rightarrow 0, \\ g_i^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) - r_i &\leq \xi(\delta), \quad i = 1, \dots, m, \quad \xi(\delta) \rightarrow 0, \quad \delta \rightarrow 0, \end{aligned}$$

indicating that, regardless of whether the dual problem to $(P_{p,r}^0)$ is solvable or not, the dual regularization algorithm defined by the equality

$$R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) = z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}],$$

is regularizing, in the sense of Definition 3.3, in Problem $(P_{p,r}^0)$. In addition to those mentioned above, the limiting relations also hold:

$$\begin{aligned} \alpha(\delta) \|(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})\|^2 &\rightarrow 0, \\ \langle (A^\delta z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^\delta - p, g^\delta(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}])) - r, (\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rangle &\rightarrow 0, \quad \delta \rightarrow 0. \end{aligned}$$

If a strongly convex functional f^0 is also subdifferentiable (in the sense of convex analysis) at the points of $\mathcal{D}$, then the limit relation holds

$$\|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - z_{p,r}^0\| \rightarrow 0, \quad \delta \rightarrow 0.$$

2) *The limit relation is satisfied*

$$\lim_{\delta \rightarrow +0} V_{p,r}^0(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) = \sup_{(\lambda, \mu) \in H \times \mathbb{R}_+^m} V_{p,r}^0(\lambda, \mu) = f^0(z_{p,r}^0),$$

and besides, in the case of solvability of the dual problem, the limit relation is true also

$$(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}) \rightarrow (\lambda_{p,r}^0, \mu_{p,r}^0), \quad \delta \rightarrow 0, \quad (4.27)$$

where $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$ is an element that minimizes the functional $\|(\lambda, \mu) - (\tilde{\lambda}, \tilde{\mu})\|^2$, $(\lambda, \mu) \in H \times \mathbb{R}_+^m$ among all elements that maximize the functional $V_{p,r}^0$ on $H \times \mathbb{R}_+^m$. Due to the arbitrariness in the choice of the point $(\tilde{\lambda}, \tilde{\mu}) \in H \times \mathbb{R}_+^m$, in the case of solvability of the problem dual to $(P_{p,r}^0)$, we can assume without loss of generality that in the limit relation (4.27) $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$ can be taken as any pre-chosen and fixed solution of the dual problem (for example, a normal solution, i.e., minimal in norm, in the case $(\tilde{\lambda}, \tilde{\mu}) = 0$). Moreover, since the set of all solutions of the problem that is dual to $(P_{p,r}^0)$, or, in other words, the set of all Kuhn-Tucker vectors of Problem $(P_{p,r}^0)$, coincides with the subdifferential $\partial\beta(p, r)$ taken with the opposite sign, i.e., with $-\partial\beta(p, r)$ (see, for example, [24, Theorem 1.1]), any pre-selected and fixed element from $-\partial\beta(p, r)$ can be taken as $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$.

Thus, depending on whether the problem dual to $(P_{p,r}^0)$ has a solution or not, the dual algorithm constructed above behaves in two ways. If a solution to the dual problem exists, the family of dual variables $(\lambda_{p,r}^{\delta, \alpha(\delta)}, \mu_{p,r}^{\delta, \alpha(\delta)})$, $\delta \rightarrow 0$ generated by the algorithm is bounded in $H \times \mathbb{R}_+^m$. If such a solution does not exist, then the norms of the elements of this family increase indefinitely as $\delta \rightarrow 0$.

Remark 4.4. From Theorem 4.3 one can easily obtain its “sequential version” or, in other words, a sequential version of the nonclassical necessary optimality conditions in Problem $(P_{p,r}^0)$. To do this, one must take a sequence of nonnegative numbers δ^k , $k = 1, 2, \dots$ converging to zero, and replace δ with δ^k in the formulation of the theorem, and the limit relation $\delta \rightarrow 0$ with $\delta^k \rightarrow 0$, $k \rightarrow \infty$. Moreover, the sequential version of the dual regularization algorithm defined by the equality

$$R(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}, \delta^k) = z^{\delta^k} [\lambda_{p,r}^{\delta^k, \alpha(\delta^k)}, \mu_{p,r}^{\delta^k, \alpha(\delta^k)}],$$

is an GMS-generating (regularizing) algorithm in the sense of Definition 3.5 in Problem $(P_{p,r}^0)$.

Remark 4.5. It is important to note that both Theorem 4.3 and Remark 4.4 “serve” under the sole condition of the existence of a solution to Problem $(P_{p,r}^0)$ all three possible situations associated with the subdifferential properties of the value function $\beta(p, r)$, $p \in H$, $r \in \mathbb{R}_+^m$, namely: 1. $\partial\beta(p, r) \neq \emptyset$; 2. $\partial^\infty\beta(p, r) \neq \{0\}$; 3. $\partial\beta(p, r) = \emptyset$, $\partial^\infty\beta(p, r) = \{0\}$. In other words, the NCOCs discussed in Theorem 4.3 and in Remark 4.4 are valid, unlike the classical LMR (see [15, Theorem 2.1], [24, Theorem 1.1]), for any “subdifferential properties” of the value function β at the point (p, r) , if Problem $(P_{p,r}^0)$ is solvable.

4.2. A nonclassical optimality criterion in the form of a regularized LMR. From the nonclassical necessary optimality conditions of Theorem 4.3, taking into account Remark 4.4, we move on to a nonclassical optimality criterion in Problem $(P_{p,r}^0)$ in the form of a regularized LMR. Here, we must take into account, keeping in mind Lemma 3.8, that the existence of a bounded GMS in Problem $(P_{p,r}^0)$ is equivalent to its solvability, i.e., to the existence of a solution $z_{p,r}^0$. It can be argued that the following theorem is valid; for the proof, see [18, Section 3, Theorem 1, Remark 6]. Before formulating this theorem, let us recall that $\mathcal{D}_{p,r}^{\delta, \epsilon} \equiv \{z \in \mathcal{D} : \|A^\delta z - h^\delta - p\| \leq \epsilon, g_i^\delta(z) \leq r_i + \epsilon, i = 1, \dots, m\}$.

Theorem 4.6. [Nonclassical optimality criterion in the form of regularized LMR] *Let Problem $(P_{p,r}^0)$ satisfy Condition A and be given an arbitrary sequence $\delta^k \in \mathbb{R}_+^1$, $k = 1, 2, \dots$ of nonnegative numbers converging to zero. For the existence of a bounded GMS in Problem $(P_{p,r}^0)$, it is necessary and sufficient that there exists a sequence $(\lambda^k, \mu^k) \in H \times \mathbb{R}_+^m$, $k = 1, 2, \dots$, such that the relations hold*

$$\delta^k \|(\lambda^k, \mu^k)\|^2 \rightarrow 0, \quad z^{\delta^k} [\lambda^k, \mu^k] \in \mathcal{D}_{p,r}^{\delta^k, \epsilon^k}, \quad \epsilon^k \rightarrow 0,$$

$$\langle (\lambda^k, \mu^k), A^{\delta^k} z^{\delta^k} [\lambda^k, \mu^k] - h^{\delta^k} - p, g^{\delta^k}(z^{\delta^k} [\lambda^k, \mu^k]) - r \rangle \rightarrow 0, \quad k \rightarrow \infty,$$

and the sequence $z^{\delta^k} [\lambda^k, \mu^k]$, $k = 1, 2, \dots$ was bounded. This sequence $z^{\delta^k} [\lambda^k, \mu^k]$, $k = 1, 2, \dots$ is the (desired) GMS in Problem $(P_{p,r}^0)$. In the case of subdifferentiability of f^0 on $\mathcal{D}$, regardless of whether the problem dual to $(P_{p,r}^0)$ is solvable or not, $z^{\delta^k} [\lambda^k, \mu^k] \rightarrow z_{p,r}^0$, $k \rightarrow \infty$. In other words, the operator $R(\cdot, \cdot, \cdot, \cdot, \delta^k)$, which is defined by the equality $R(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}, \delta^k) = z^{\delta^k} [\lambda^k, \mu^k]$, is a GMS-generating (regularizing) in Problem $(P_{p,r}^0)$ in the sense of Definition 3.5 and generates a GMS that converges to $z_{p,r}^0$ as $k \rightarrow \infty$ in the case of subdifferentiability of f^0 on $\mathcal{D}$.

In addition, the limit relation holds

$$V_{p,r}^0(\lambda^k, \mu^k) \rightarrow \sup_{(\lambda, \mu) \in H \times \mathbb{R}_+^m} V_{p,r}^0(\lambda, \mu) = f^0(z_{p,r}^0), \quad k \rightarrow \infty.$$

As a sequence $(\lambda^k, \mu^k) \in H \times \mathbb{R}_+^m$, $k = 1, 2, \dots$, we can take the sequence $(\lambda_{p,r}^{\delta^k, \alpha(\delta^k)}, \mu_{p,r}^{\delta^k, \alpha(\delta^k)})$, $k = 1, 2, \dots$, $\delta^k / \alpha(\delta^k) \rightarrow 0$, $k \rightarrow \infty$, in which $(\lambda_{p,r}^{\delta^k, \alpha(\delta^k)}, \mu_{p,r}^{\delta^k, \alpha(\delta^k)}) \equiv \operatorname{argmax}\{V_{p,r}^0(\lambda, \mu) - \alpha(\delta^k) \|(\lambda, \mu) - (\tilde{\lambda}, \tilde{\mu})\|^2, (\lambda, \mu) \in H \times \mathbb{R}_+^m\}$, $(\tilde{\lambda}, \tilde{\mu}) \in H \times \mathbb{R}_+^m$ is an arbitrary fixed vector.

In the case of existence of a bounded GMS and solvability of the problem that is dual to $(P_{p,r}^0)$, the limit relation $(\lambda^k, \mu^k) \rightarrow (\lambda_{p,r}^0, \mu_{p,r}^0)$, $k \rightarrow \infty$ holds, where due to the arbitrariness in the choice of $(\tilde{\lambda}, \tilde{\mu})$, any pre-chosen and fixed solution of the dual problem (for example, a normal one, i.e., minimal in norm in the case $(\tilde{\lambda}, \tilde{\mu}) = 0$) can be

taken as $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$. Moreover, since the set of all solutions of the problem that is dual to $(P_{p,r}^0)$, or, in other words, the set of all Kuhn-Tucker vectors of Problem $(P_{p,r}^0)$, coincides with subdifferential $\partial\beta(p, r)$ taken with the opposite sign, i.e., with $-\partial\beta(p, r)$ (see, for example, [24, Theorem 1.1]), any pre-selected and fixed element from $-\partial\beta(p, r)$ can be taken as $(\lambda_{p,r}^0, \mu_{p,r}^0) \in H \times \mathbb{R}_+^m$.

5. THE CLASSICAL LAGRANGE MULTIPLIER RULE AS A LIMIT VERSION OF ITS NONCLASSICAL ANALOG

One of the important features of the NCOC (regularized LMR) is that it yields its classical analog “in the limit”. We demonstrate this using the perturbation method employed in this paper. We thereby establish a strong connection between the NCOC and its classical analog and substantiate a new sequential method for obtaining the LMR in convex constrained extremal problems. This new method for proving the LMR as a necessary optimality condition (the proof begins with the assumption of the existence of an optimal element) relies on the regularized LMR of Theorem 4.6 formulated above for Problem $(P_{p,r}^0)$ with a strongly convex objective functional. For simplicity, we assume that the convex, possibly not strongly convex, objective function f^0 is subdifferentiable on $\mathcal{D}$. Note that in the case of a convex, perhaps not strongly, objective functional f^0 , the value function β , as in the case of a strongly convex f^0 , is convex and lower semicontinuous [24, Lemma 1.2].

Thus, we show that the limit passage using the relations of the regularized LMR of Theorem 4.6 leads to the necessary conditions of the classical LMR in nondifferential form in Problem $(P_{p,r}^0)$ (see [24, Theorem 1.1]), and the following theorem holds.

Theorem 5.1. *Let $f^0 : \mathcal{D} \rightarrow \mathbb{R}^1$ be a continuous convex and subdifferentiable function on $\mathcal{D}$, $(p, r) \in H \times \mathbb{R}^m$ be a point such that $\beta(p, r) < +\infty$, $z_{p,r}^0 \in \mathcal{D}_{p,r}^0 \equiv \{z \in \mathcal{D} : A^0 z - h^0 - p = 0, g_i^0(z) \leq r_i, i = 1, \dots, m\}$ be an optimal element in Problem $(P_{p,r}^0)$, and $f^0(z_{p,r}^0) = \beta(p, r) = \beta_0(p, r)$. Then the following two statements are true.*

1. *If $\zeta \in \partial\beta(p, r)$, where $\partial\beta(p, r)$ is a subdifferential in the sense of convex analysis, then for the Lagrange multipliers $\lambda \in H$, $\mu \in \mathbb{R}_+^m$, $(\lambda, \mu) = -\zeta$, for $\mu_0 = 1$ the relations are satisfied*

$$L_{p,r}^0(z_{p,r}^0, \mu_0, \lambda, \mu) \leq L_{p,r}^0(z, \mu_0, \lambda, \mu) \quad \forall z \in \mathcal{D}, \quad \mu_i(g_i(z_{p,r}^0) - r_i) = 0, \quad i = 1, \dots, m, \quad (5.1)$$

$$L_{p,r}^0(z, \mu_0, \lambda, \mu) \equiv \mu_0 f^0(z) + \langle \lambda, A^0 z - h^0 - p \rangle + \langle \mu, g^0(z) - r \rangle.$$

2. *If $\zeta \in \partial^\infty\beta(p, r)$, $\zeta \neq 0$, where $\partial^\infty\beta(p, r)$ is a singular subdifferential (see, for example, [14]), defined by the formula*

$$\partial^\infty\beta(p, r) \equiv \{(\lambda, \mu) \in H \times \mathbb{R}^m : ((\lambda, \mu), 0) \in N_{\text{epi}\beta}((p, r), \beta(p, r))\},$$

then for the Lagrange multipliers $\lambda \in H$, $\mu \in \mathbb{R}_+^m$, $(\lambda, \mu) = -\zeta$, the relations (5.1) are satisfied for $\mu_0 = 0$.

Proof. Consider an arbitrary fixed point $(p, r) \in \text{dom}\beta$; for example, (p, r) can equal to $(0, 0)$. At this point, at least one of the following three cases occurs:

1. $\partial\beta(p, r) \neq \emptyset$;
2. $\partial^\infty\beta(p, r) \neq \{0\}$;
3. $\partial\beta(p, r) = \emptyset$, $\partial^\infty\beta(p, r) = \{0\}$.

We will show that in the first two cases, passing to the limit using the relations of the regularized LMR of Theorem 4.6 leads to the classical LMR in nondifferential form in Problem $(P_{p,r}^0)$. The third case corresponds to the situation where the classical LMR in Problem $(P_{p,r}^0)$ degenerates (see Theorem 1.1 in [24]).

Let us fix an arbitrary element $z_{p,r}^0 \in Z_{p,r}^0$, where $Z_{p,r}^0$ is the set of all solutions to Problem $(P_{p,r}^0)$, which we assume to be nonempty. Consider auxiliary problem that depends on $l \in [0, 1]$ $((p', r') \in H \times \mathbb{R}^m)$

$$(\tilde{P}_{p',r'}^0) \quad f^0(z) + l\|z - z_{p,r}^0\|^2 \rightarrow \inf, \quad A^0 z = h^0 + p', \quad g_i^0(z) \leq r'_i, \quad i = 1, \dots, m, \quad z \in \mathcal{D},$$

which is a convex programming problem with a strongly convex and subdifferentiable objective functional $f^0(z) + l\|z - z_{p,r}^0\|^2$, $z \in \mathcal{D}$, with the value function $\tilde{\beta}$, which has the same properties as β . Here and below, the “wave” denotes the dependence of Problem $(\tilde{P}_{p',r'}^0)$, as well as the “elements” associated with it, on l . The (unique) solutions $\tilde{z}_{p',r'}^0$ of this problem (they depend on l) exist for any $(p', r') \in \text{dom}\tilde{\beta}$, and, obviously, $\tilde{z}_{p,r}^0 = z_{p,r}^0$. Its peculiarity is that for all $(p', r') \in H \times \mathbb{R}^m$ the inequality $\tilde{\beta}(p', r') \geq \beta(p', r')$ holds, and at the point (p, r) we have the equality $\tilde{\beta}(p, r) = \beta(p, r)$. Thus, $\text{epi}\tilde{\beta} \subseteq \text{epi}\beta$ and, therefore, any normal (in the sense of convex analysis) to $\text{epi}\beta$ at the point $(p, r, \beta(p, r))$ is simultaneously a normal (in the same sense) at the same point to $\text{epi}\tilde{\beta}$. Thus, for $(p', r') = (p, r)$ we have the inclusions $\partial\beta(p, r) \subseteq \partial\tilde{\beta}(p, r)$, $\partial^\infty\beta(p, r) \subseteq \partial^\infty\tilde{\beta}(p, r)$.

Let us first consider the first situation, assuming that $\partial\beta(p, r) \neq \emptyset$ and fixing an arbitrary element $\zeta \in \partial\beta(p, r)$. For this reason, in Problem $(\tilde{P}_{p,r}^0)$ with a strongly convex objective functional, the inequality $\partial\tilde{\beta}(p, r) \neq \emptyset$ holds, and its only optimal element is $z_{p,r}^0$. In this situation, the regularized LMR of Theorem 4.6 can be applied to Problem $(\tilde{P}_{p',r'}^0)$; see also Theorem 1 in [18]. Let (recall that “wave” means dependence on l)

$$\tilde{L}_{p',r'}^\delta(z, \lambda, \mu) \equiv f^\delta(z) + l\|z - z_{p,r}^0\|^2 + \langle \lambda, A^\delta z - h^\delta - p' \rangle + \langle \mu, g^\delta(z) - r' \rangle$$

is the Lagrange functional of the problem

$$(\tilde{P}_{p',r'}^\delta) \quad f^\delta(z) + l\|z - z_{p,r}^0\|^2 \rightarrow \inf, \quad A^\delta z = h^\delta + p', \quad g_i^\delta(z) \leq r'_i, \quad i = 1, \dots, m, \quad z \in \mathcal{D},$$

where $\lambda \in H$, $\mu \in \mathbb{R}_+^m$, $p' \in H$, $r' \in \mathbb{R}^m$, $\delta \in [0, \delta_0]$. For any $\lambda \in H$, $\mu \in \mathbb{R}_+^m$ and each $\delta \in (0, \delta_0]$, the function $\tilde{L}_{p',r'}^\delta(z, \lambda, \mu)$, $z \in \mathcal{D}$, is strongly convex with constant l and continuous. Therefore, it attains a minimum on the convex and closed set $\mathcal{D} \subseteq Z$, and at a unique point (see, e.g., [34, Chapter 8, §2, Theorem 10]), which we denote by $\tilde{z}^\delta[\lambda, \mu]$.

The dual to Problem $(\tilde{P}_{p',r'}^\delta)$ is the problem

$$\tilde{V}_{p',r'}^\delta(\lambda, \mu) \equiv \min_{z \in \mathcal{D}} \tilde{L}_{p',r'}^\delta(z, \lambda, \mu) \rightarrow \sup, \quad \lambda \in H, \quad \mu \in \mathbb{R}_+^m.$$

Let $(\tilde{\lambda}_{p',r'}^{\delta,\alpha}, \tilde{\mu}_{p',r'}^{\delta,\alpha})$ be the unique point (it depends on l) that gives a maximum on $H \times \mathbb{R}_+^m$ to the strongly concave functional

$$\tilde{R}_{p',r'}^{\delta,\alpha}(\lambda, \mu) \equiv \tilde{V}_{p',r'}^\delta(\lambda, \mu) - \alpha\|\lambda\|^2 - \alpha\|\mu\|^2, \quad \lambda \in H, \quad \mu \in \mathbb{R}_+^m, \quad \alpha \in \mathbb{R}_+^1, \quad \delta \in (0, \delta_0].$$

Thus, in the case of solvability of the problem that is dual to $(\tilde{P}_{p',r'}^0)$, i.e., in the case $\partial\tilde{\beta}(p', r') \neq \emptyset$, in accordance with the statement of Theorem 4.6, due to the strong convergence of the corresponding GMSs $\tilde{z}^{\delta j}[\tilde{\lambda}_{p',r'}^{\delta j, \alpha(\delta^j)}, \tilde{\mu}_{p',r'}^{\delta j, \alpha(\delta^j)}]$, $j = 1, 2, \dots$, ($\tilde{z}^\delta[\lambda, \mu] \equiv \operatorname{argmin}\{\tilde{L}_{p',r'}^\delta(z, \lambda, \mu), z \in \mathcal{D}\}$) to the optimal element $\tilde{z}_{p',r'}^0$ of Problem $(\tilde{P}_{p',r'}^0)$, and the sequence of dual variables to some (arbitrarily chosen and fixed) solution of the problem that is dual to $(\tilde{P}_{p',r'}^0)$

$$(\tilde{\lambda}_{p',r'}^{\delta j, \alpha(\delta^j)}, \tilde{\mu}_{p',r'}^{\delta j, \alpha(\delta^j)}) \rightarrow (\tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0), \quad j \rightarrow \infty, \quad (5.2)$$

we obtain in the limit as $j \rightarrow \infty$ the inequality

$$\tilde{L}_{p',r'}^0(\tilde{z}_{p',r'}^0, \tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0) \leq \tilde{L}_{p',r'}^0(z, \tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0), \quad z \in \mathcal{D}, \quad (5.3)$$

and the complementary slackness condition $(\tilde{\mu}_{p',r'}^0 = (\tilde{\mu}_{p',r',1}^0, \dots, \tilde{\mu}_{p',r',m}^0))$

$$\langle \tilde{\mu}_{p',r'}^0, g^0(\tilde{z}_{p',r'}^0) - r' \rangle = 0. \quad (5.4)$$

Here it is important to note that at (5.2) the point $(\tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0)$ for $(p', r') = (p, r)$, without loss of generality (taking into account the indicated Theorem 4.6, see also Theorem 1 in [18]) can be assumed to be independent of l and equal to any pre-chosen fixed element $(\lambda_{p,r}^0, \mu_{p,r}^0)$ from $-\partial\beta(p, r) \subseteq -\partial\tilde{\beta}(p, r)$ for any $l \in (0, 1]$. In particular, the element $-(\lambda_{p,r}^0, \mu_{p,r}^0)$ can be taken equal to the above-chosen and fixed element $\zeta \in \partial\beta(p, r)$.

Thus, for $(p', r') = (p, r)$, $\tilde{z}_{p,r}^0 = z_{p,r}^0$ in the first of the three situations under consideration, we obtain the relations of the classical regular LMR in Problem $(P_{p,r}^0)$ (see [24, Theorem 1.1]) after passing to the limit as $l \rightarrow 0$ in (5.3), (5.4):

$$L_{p,r}^0(z_{p,r}^0, \lambda_{p,r}^0, \mu_{p,r}^0) \leq L_{p,r}^0(z, \lambda_{p,r}^0, \mu_{p,r}^0), \quad z \in \mathcal{D}; \quad \langle \mu_{p,r}^0, g^0(z_{p,r}^0) - r \rangle = 0,$$

where $L_{p,r}^0(z, \lambda, \mu) \equiv f^0(z) + \langle \lambda, A^0 z - h^0 - p \rangle + \langle \mu, g^0(z) - r \rangle$. So, the first statement of the theorem is proven.

Next, before moving on to the second situation, we emphasize that in each Problem $(\tilde{P}_{p',r'}^0)$ with $\partial\tilde{\beta}(p', r') \neq \emptyset$ the sequence of dual variables $(\tilde{\lambda}_{p',r'}^{\delta j, \alpha(\delta^j)}, \tilde{\mu}_{p',r'}^{\delta j, \alpha(\delta^j)})$, $j = 1, 2, \dots$, can be assumed without loss of generality to be such that in the limit relation (5.2) the element $(\tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0)$ is equal to any point that gives a maximum in the dual problem $\tilde{V}_{p',r'}^0(\lambda, \mu) \rightarrow \sup$, $\{\lambda, \mu\} \in H \times \mathbb{R}_+^m$. In other words, since $-\partial\tilde{\beta}(p', r')$ coincides with the set of all maximum points of this dual problem, then in the limit relation (5.2) the element $(\tilde{\lambda}_{p',r'}^0, \tilde{\mu}_{p',r'}^0)$ can be

considered equal to any fixed pre-chosen element of the subdifferential $\partial\tilde{\beta}(p', r')$, taken with the opposite sign. We use this property of the limit relation (5.2) in analyzing the second of the three situations mentioned above.

So, we move to the second situation, when $\partial^\infty\beta(p, r) \neq \{0\}$, and as ζ we now fix an arbitrary nonzero element from $\partial^\infty\beta(p, r)$. Then, in accordance with the above, we also have the inequality $\partial^\infty\tilde{\beta}(p, r) \neq \{0\}$. In this case, to pass to the limit with using the relations of Theorem 4.6, we proceed somewhat differently, using two important facts related to the subdifferentiability properties of a convex lower semicontinuous value function $\tilde{\beta}$.

The first is that every such function in a Hilbert space is subdifferentiable on a dense subset of its domain (effective set), see Lemma 3.9. The second is related to the well-known representation for the singular subdifferential of a convex lower semicontinuous function (see, for example, [14, Proposition 4C2]), which in our case is the value function $\tilde{\beta}$ (see Lemma 3.2):

$$\partial^\infty\tilde{\beta}(p, r) = \limsup_{(p', r') \xrightarrow{\tilde{\beta}} (p, r), t \downarrow 0} t\partial\tilde{\beta}(p', r') \equiv \{w - \lim_{j \rightarrow \infty} t_j \zeta_j : t_j \downarrow 0, \zeta_j \in \partial\tilde{\beta}(p^j, r^j), (p^j, r^j) \xrightarrow{\tilde{\beta}} (p, r)\},$$

where the symbol $(p', r') \xrightarrow{\tilde{\beta}} (p, r)$ means that $((p', r'), \tilde{\beta}(p', r')) \rightarrow ((p, r), \tilde{\beta}(p, r))$, the symbol $t \downarrow 0$ means convergence to zero from the right, and the symbol $w - \lim_{j \rightarrow \infty}$ means weak limit transition.

Now we multiply the inequality (5.3) by $s > 0$:

$$\begin{aligned} \tilde{L}_{p', r'}^0(z_{p', r'}^0, s, s\tilde{\lambda}_{p', r'}^0, s\tilde{\mu}_{p', r'}^0) &\leq \tilde{L}_{p', r'}^0(z, s, s\tilde{\lambda}_{p', r'}^0, s\tilde{\mu}_{p', r'}^0), \\ \tilde{L}_{p', r'}^0(z, s, \lambda, \mu) &\equiv s(f^0(z) + l\|z - z_{p, r}^0\|^2) + \langle \lambda, A^0 z - h^0 - p' \rangle + \langle \mu, g^0(z) - r' \rangle. \end{aligned} \quad (5.5)$$

For any weak limit point of the form

$$(\hat{\lambda}_{p, r}, \hat{\mu}_{p, r}) = w - \lim_{j \rightarrow \infty, (p', r') \xrightarrow{\tilde{\beta}} (p, r), s_j \downarrow 0} s_j(\tilde{\lambda}_{p^j, r^j}^0, \tilde{\mu}_{p^j, r^j}^0)$$

with $(\tilde{\lambda}_{p^j, r^j}^0, \tilde{\mu}_{p^j, r^j}^0) \in (-\partial\tilde{\beta}(p^j, r^j)) \neq \emptyset$ we can write after the obvious limit transition in (5.5) with $(p', r') = (p^j, r^j)$, $(\tilde{\lambda}_{p', r'}^0, \tilde{\mu}_{p', r'}^0) = (\tilde{\lambda}_{p^j, r^j}^0, \tilde{\mu}_{p^j, r^j}^0)$, $s = s_j \rightarrow 0$, $j \rightarrow \infty$:

$$L_{p, r}^0(z_{p, r}^0, 0, \hat{\lambda}_{p, r}, \hat{\mu}_{p, r}) \leq L_{p, r}^0(z, 0, \hat{\lambda}_{p, r}, \hat{\mu}_{p, r}) \quad \forall z \in \mathcal{D}. \quad (5.6)$$

This weak limit point $(\hat{\lambda}_{p, r}, \hat{\mu}_{p, r}) \in -\partial^\infty\tilde{\beta}(p, r)$, in view of the above representation for the singular subdifferential, can be chosen equal to any fixed point from $-\partial^\infty\tilde{\beta}(p, r)$. Moreover, given the inclusion $\partial^\infty\beta(p, r) \subseteq \partial^\infty\tilde{\beta}(p, r)$ mentioned above, we can assume, without loss of generality, that it is any fixed element belonging to $-\partial^\infty\beta(p, r)$ and independent of l . In particular, the element $-(\hat{\lambda}_{p, r}, \hat{\mu}_{p, r})$ can be taken equal to the above-chosen and fixed element $\zeta \in \partial^\infty\beta(p, r)$.

Note that in this limit passage we used the limit relation $\tilde{z}_{p^j, r^j}^0 \rightarrow z_{p, r}^0$, $j \rightarrow \infty$, which is a consequence of the weak convergence of $\tilde{z}_{p^j, r^j}^0$ to $z_{p, r}^0$, numerical convergence (for each fixed $l \in (0, 1]$) $\tilde{\beta}(p^j, r^j) = f^0(\tilde{z}_{p^j, r^j}^0) + l\|\tilde{z}_{p^j, r^j}^0 - z_{p, r}^0\|^2$ to $\tilde{\beta}(p, r) = f^0(z_{p, r}^0)$ as $j \rightarrow \infty$, subdifferentiability at the points $\mathcal{D}$ and strong convexity (for $l \in (0, 1]$) of the functional $f^0(z) + l\|z - z_{p, r}^0\|^2$, $z \in \mathcal{D}$. Simultaneously, by virtue of the complementary slackness condition $\tilde{\mu}_{p^j, r^j}^0(g_i^0(\tilde{z}_{p^j, r^j}^0) - r_i^j) = 0$, $i = 1, \dots, m$ (see (5.4)) as a result of the limit passage in it as $j \rightarrow \infty$ and already applied in obtaining (5.6) the limit relation $\tilde{z}_{p^j, r^j}^0 \rightarrow z_{p, r}^0$, $j \rightarrow \infty$ we obtain $\hat{\mu}_{p, r, i}(g_i^0(z_{p, r}^0) - r_i) = 0$, $i = 1, \dots, m$, which together with (5.6) means that the nonregular nondegenerate LMR is satisfiable (see [24, Theorem 1.1]). Thus, the second statement of the theorem is also proven. In this case, we approximated the solution $z_{p, r}^0$ of Problem $(P_{p, r}^0)$ by points $\tilde{z}_{p^j, r^j}^0$, which provide the minimum value of the Lagrange functions $\tilde{L}_{p^j, r^j}^0(z, \lambda_{p^j, r^j}^0, \mu_{p^j, r^j}^0)$, $z \in \mathcal{D}$, and took into account the independence of $z_{p, r}^0$ and $(\hat{\lambda}_{p, r}, \hat{\mu}_{p, r})$ from l . Thus, for $(p', r') = (p, r)$, $\tilde{z}_{p, r}^0 = z_{p, r}^0$, we obtain the relations of the classical nonregular LMR in Problem $(P_{p, r}^0)$.

And finally, let the last of the three situations identified above be realized, i.e., we have the equalities $\partial\beta(p, r) = \emptyset$, $\partial^\infty\beta(p, r) = \{0\}$. In this case, the limit passage described above formally leads to a degenerate LMR in Problem $(P_{p, r}^0)$, i.e., in other words, it can be stated that the LMR is not satisfied in this problem, see [24, Theorem 1.1], for corresponding examples see [15, 16].

□

6. CONCLUSION

The crucial role of the Lagrange multiplier rule (LMR) in nondifferential form in constrained optimization theory is demonstrated, particularly in a convex constrained extremal problem with an operator equality-constraint and a finite number of functional inequality-constraints, as well as with a strongly convex objective functional. It allows us to naturally link the traditional approach to obtaining classical optimality conditions (COCs) within the framework of the classical theory of extremal problems with the approach according to which we regard the “same” extremal problems as ill-posed. Necessary and sufficient so-called nonclassical optimality conditions (NCOCs) in the problem under consideration are proven. Necessary NCOCs are necessary conditions for (ordinary) optimality, but they are written in nonclassical form, and besides they are satisfied under the sole assumption of the solvability of the problem. The main purpose of NCOCs is to overcome the problems of COCs ill-posedness, manifested in their possible unsatisfiability and instability. They are regularizing algorithms that provide approximation, stable to errors in the input data, of the solution to the problem under consideration by extremals of its regular Lagrangian function while the dual variable in it is chosen in accordance with the Tikhonov regularization procedure. The characteristic property of the NCOCs is that the limit transition in the relations that comprise them leads to classical analogues.

Finally, we note that in the constrained extremal Problem $(P_{p,r}^0)$ considered in this paper, the objective functional f^0 is strongly convex. In the case of a convex, but not strongly convex, functional f^0 , when obtaining NCOCs in Problem $(P_{p,r}^0)$, similar to those obtained in this paper, one can use the approach [20, 21], associated with the approximation of the original problem by a family of auxiliary problems, in each of which the functional f^0 is replaced by the Tikhonov regularized functional $f^0(z) + \epsilon\|z\|^2$, $z \in \mathcal{D}$ with the regularization parameter ϵ . Then, for each $\epsilon > 0$, the auxiliary problem has the same properties as Problem $(P_{p,r}^0)$ of this paper, and the consistent convergence of the regularization parameter ϵ and the dual regularization parameter α (see Subsection 4.1.1) to zero in Problem $(P_{p,r}^0)$ with not strongly convex f^0 ensures the corresponding NCOCs (see [20, 21] for detail). Note also that the application of the regularized LMR or, in other words, NCOCs for the stable approximate solving the classical ill-posed problem of finding a normal solution to an operator equation of the first kind, i.e. for Problem (IP) of Section 2, is discussed in detail in [22, 23].

STATEMENTS AND DECLARATIONS

The author declares that he has no conflict of interest, and the manuscript has no associated data.

REFERENCES

- [1] V. M. Alekseev, V. M. Tikhomirov, and S. V. Fomin. *Optimal Control*. Nauka, Moscow, 1979.
- [2] J.-P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. Wiley, New York, 1984.
- [3] J.-P. Aubin. *L'analyse Non Linéaire et ses Motivations Économiques*. Masson, Paris, 1984.
- [4] E. R. Avakov, G. G. Magaril-Il'yayev, and V. M. Tikhomirov. Lagrange's principle in extremum problems with constraints. *Russian Mathematical Surveys*, 68(3):401-433, 2013.
- [5] A. B. Bakushinsky and A. V. Goncharsky. *Ill-Posed Problems: Numerical Methods and Applications*. Moscow University Publishing House, Moscow, 1989.
- [6] I. Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324-353, 1974.
- [7] E. G. Gol'shtein. *Duality Theory in Mathematical Programming and Its Applications*. Nauka, Moscow, 1971.
- [8] A. N. Tikhonov and A. V. Goncharsky, editors. *Ill-Posed Problems in Natural Science: Collection of Articles*. Moscow University Publishing House, Moscow, 1987.
- [9] M. L. Krasnov, A. I. Kiselev, and G. I. Makarenko. *Integral Equations*. Nauka, Moscow, 1976.
- [10] F. A. Kuterin and M. I. Sumin. Stable iterative Lagrange principle in convex programming as a tool for solving unstable problems. *Computational Mathematics and Mathematical Physics*, 57:71-82, 2017.
- [11] F. A. Kuterin and M. I. Sumin. On the regularized Lagrange principle in the iterative form and its application for solving unstable problems. *Mathematical Models and Computer Simulations*, 9:328-338, 2017.
- [12] F. A. Kuterin and M. I. Sumin. The regularized iterative Pontryagin maximum principle in optimal control. I. Optimization of a lumped system. *The Bulletin of Udmurt University. Mathematics. Mechanics. Computer Science*, 26(4):474-489, 2016.
- [13] F. A. Kuterin and M. I. Sumin. The regularized iterative Pontryagin maximum principle in optimal control. II. Optimization of a distributed system. *The Bulletin of Udmurt University. Mathematics. Mechanics. Computer Science*, 27(1):26-41, 2017.

- [14] P. D. Loewen. *Optimal Control via Nonsmooth Analysis*. volume 2 of *CRM Proceedings & Lecture Notes*. American Mathematical Society, Providence, RI, 1993.
- [15] M. I. Sumin. Regularized parametric Kuhn–Tucker theorem in a Hilbert space. *Computational Mathematics and Mathematical Physics*, 51:1489–1509, 2011.
- [16] M. I. Sumin. Regularized Lagrange principle and Pontryagin maximum principle in optimal control and in inverse problems. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 25(1):279–296, 2019.
- [17] M. I. Sumin. Lagrange principle and its regularization as a theoretical basis of stable solving optimal control and inverse problems. *Russian Universities Reports. Mathematics*, 26:151–171, 2021.
- [18] M. I. Sumin. The perturbation method and regularization of the Lagrange multiplier rule in convex constrained optimization problems. *Proceedings of the Steklov Institute of Mathematics*, 325(Suppl. 1):S194–S211, 2024.
- [19] M. I. Sumin. On regularization of the Lagrange multiplier rule in convex constrained extremum problems and on its universality. *Communications in Optimization Theory*, 2026:Article ID 9, 2026.
- [20] M. I. Sumin. On regularization of the classical optimality conditions in convex optimal control problems. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 26(2):252–269, 2020.
- [21] M. I. Sumin. On the regularization of the Lagrange principle and on the construction of the generalized minimizing sequences in convex constrained optimization problems. *The Bulletin of Udmurt University. Mathematics. Mechanics. Computer Science*, 30:410–428, 2020.
- [22] M. I. Sumin. On ill-posed problems, extremals of the Tikhonov functional and the regularized Lagrange principles. *Russian Universities Reports. Mathematics*, 27:58–79, 2022.
- [23] M. I. Sumin. On the role of Lagrange multipliers and duality in ill-posed problems for constrained extremum. To the 60th anniversary of the Tikhonov regularization method. *Russian Universities Reports. Mathematics*, 28:414–435, 2023.
- [24] M. I. Sumin. Nondifferential Kuhn–Tucker theorems in constrained optimization problems and subdifferentials of nonsmooth analysis. *Russian Universities Reports. Mathematics*, 25:307–330, 2020.
- [25] M. I. Sumin. *Ill-Posed Problems and Their Solving Methods. Lecture Materials for Senior Students*. Publishing House of Nizhniy Novgorod State University, Nizhniy Novgorod, 2009.
- [26] M. I. Sumin. A regularized gradient dual method for the inverse problem of a final observation for a parabolic equation. *Computational Mathematics and Mathematical Physics*, 44:1903–1921, 2004.
- [27] M. I. Sumin. Duality-based regularization in a linear convex mathematical programming problem. *Computational Mathematics and Mathematical Physics*, 47:602–625, 2007.
- [28] V. M. Tikhomirov. *Some Questions of Approximation Theory*. Moscow University Publishing House, Moscow, 1976.
- [29] V. M. Tikhomirov. Extremum theory and extremal problems of the classical calculus. *Journal of Mathematical Sciences*, 104:889–931, 2001.
- [30] V. M. Tikhomirov. Specific problems and general extremum theory. *Proceedings of the Steklov Institute of Mathematics*, Suppl. 2:S154–S167, 2004.
- [31] A. N. Tikhonov and V. Ya. Arsenin. *Solutions of Ill-Posed Problems*. Nauka, Moscow, 1974.
- [32] A. N. Tikhonov. On the stability of the functional optimization problem. *USSR Computational Mathematics and Mathematical Physics*, 6(4):28–33, 1966.
- [33] V. A. Trenogin. *Functional Analysis*. Nauka, Moscow, 1980.
- [34] F. P. Vasil’ev. *Optimization Methods*. Volumes 1, 2, Moscow Center for Continuous Mathematical Education, Moscow, 2011.
- [35] J. Warga. *Optimal Control of Differential and Functional Equations*. Academic Press, New York, 1972.