

Chapter 14

Differentiability of Inverse Operators

Simon Y. Serovajsky

Abstract The Inverse Function Theorem is a mighty tool of the local nonlinear analysis. It guarantees the existence of the inverse function and its differentiability. However the first property is sometimes not used. It is true, for example, for the optimal control theory and the inverse problems of mathematical physics. The inverse operator can be interpreted as a control-state mapping here. Its existence is a corollary of the state equation properties, and the differentiability of the inverse operator is used for the differentiation of the minimizing functional or the discrepancy. We establish a differentiability criterion of the inverse operator. Moreover, we prove a property which can be interpreted as a weak form of the operator differentiability. The Dirichlet problem for a nonlinear elliptic equation is considered as an example.

Mathematics Subject Classification 58C20 · 46T20 · 35J60 · 49K20

14.1 Introduction

Consider an operator $A: V \rightarrow Y$, where V and Y are Banach spaces. Suppose that it is continuously differentiable at a neighborhood of a point $y_0 \in Y$. Denote by $A'(y_0)$ the derivative of the operator A at the point y_0 . It is well known that the following result holds (see, for example, [1]).

The Inverse Function Theorem *Assume that there exists the continuous inverse operator $A'(y_0)^{-1}$. Then there exists an open neighborhood O of the point y_0 such that the set $O' = A(O)$ is an open neighborhood of the point $v_0 = Ay_0$; moreover, there exists the continuously differentiable inverse map $A^{-1}: O' \rightarrow O$, and its derivative is defined by the formula*

$$(A^{-1})'(v) = \{A'[A^{-1}(v)]\}^{-1} \quad \forall v \in O'.$$

This result has very important applications. It has relationships to the Implicit Function Theorem [2], Newton–Kantorovich Method [1, 2], Lusternik Smooth

S.Y. Serovajsky (✉)

al-Farabi Kazakh National University, 71 al Farabi av., Almaty 050078, Kazakhstan
e-mail: serovajskys@mail.ru

Manifold Approximation Theorem [3, 4], Brouwer Fixed Point Theorem [1, 5], Morse Smooth Function Singularity Lemma [1], Graves Cover Theorem [4], etc. Extensions of the Inverse Function Theorem to high orders differentiability [6], non-smooth operators [7–9], multiple-valued maps [1, 7, 10], etc., are also known.

In reality the Inverse Function Theorem involve two different results. These are the invertibility of the given operator and the differentiability of the corresponding inverse operator. Sometimes only the second property is important. It is true, for example, for the extremum theory and the inverse problems theory. In particular, consider the system described by the equation

$$Ay = v. \tag{1}$$

The term v can be interpreted here as a control or an identifiable parameter, and y is a state function. Suppose that (1) has a unique solution $y = y(v)$ from the space Y for all values $v \in V$. Then the operator A is invertible. This result can be proved by some tools which are applicable to the given equation. Therefore, it is not necessary to use of the Inverse Function Theorem here.

Let U be a convex closed subset of the space V . The state functional is defined by the formula

$$I(v) = J(v) + K[y(v)],$$

where J is a functional on the set V , and K is a functional on the set Y . We have the following optimization control problem.

Problem 1 Minimize the functional I on the set U .

A necessary condition for the minimum of a smooth functional F on a convex set W at a point v_0 is the variational inequality (see [11])

$$\langle F'(v_0), v - v_0 \rangle \geq 0 \quad \forall v \in W, \tag{2}$$

where $\langle \lambda, \varphi \rangle$ is the value of the linear continuous functional λ at the point φ .

The functional I is the sum of J and the map $v \rightarrow K[y(v)]$. The last mapping is the superposition of the functional K and the map $v \rightarrow y(v)$, which is, in fact, the inverse operator A^{-1} . Then the proof of the differentiability of the given functional requires the differentiation of the inverse operator. This result can be obtained using the Inverse Function Theorem.

Lemma 1 *Suppose that the operator A has a continuous inverse operator, which is continuously differentiable at an open neighborhood of the point $y_0 = y(v_0)$, and there exists the continuous inverse operator $A'(y_0)^{-1}$. Then the map $y(\cdot) : V \rightarrow Y$ is Gateaux differentiable at the point v_0 , and its derivative satisfies the formula*

$$\langle \mu, y'(v_0)h \rangle = \langle p_\mu(v_0), h \rangle \quad \forall \mu \in Y^*, h \in V, \tag{3}$$

where Y^* is the adjoint space of Y , and $p_\mu(v_0)$ is the solution of the equation

$$[A'(y_0)]^* p_\mu(v_0) = \mu. \tag{4}$$

Proof By the Inverse Function Theorem the map $y(\cdot) : V \rightarrow Y$ is differentiable at the point v_0 , and, moreover,

$$y'(v_0) = [A'(y_0)]^{-1}.$$

Then we get

$$\langle \mu, y'(v_0)h \rangle = \langle \mu, [A'(y_0)]^{-1}h \rangle = \langle \{[A'(y_0)]^{-1}\}^* \mu, h \rangle \quad \forall \mu \in Y^*, h \in V.$$

It is known that each linear operator and its adjoint operator are invertible at the same time (see p. 460 in [2]). Therefore (4) has a unique solution

$$p_\mu(v_0) = \{[A'(y_0)]^{-1}\}^* \mu$$

from the space V^* . So the previous formula can be transformed to (4), and the equality (3) is true. \square

Now we can prove the differentiability of the functional I and obtain necessary conditions of optimality. Let v_0 be the solution of the minimization problem for the functional I on the set U . Define $y_0 = y(v_0)$.

Lemma 2 *Under the conditions of Lemma 1 suppose that the functional J is Gateaux differentiable at the point v_0 , and the functional K is Frechet differentiable at the point y_0 . Then the control v_0 satisfies the variational inequality*

$$\langle J'(v_0) - p_0, v - v_0 \rangle \geq 0 \quad \forall v \in U, \quad (5)$$

where p_0 is a solution of the adjoint equation

$$[A'(y_0)]^* p_0 = -K'(y_0). \quad (6)$$

Proof Using the Composite Function Theorem (see p. 637 in [2]), we obtain that the Gateaux derivative of the map $v \rightarrow K[y(v)]$ exists such that

$$(Ky)'(v_0) = K'(y_0)y'(v_0).$$

By equality (3) we get

$$\langle (Ky)'(v_0), h \rangle = \langle K'(y_0), y'(v_0)h \rangle = -\langle p_0, h \rangle \quad \forall h \in V,$$

where p_0 is the solution of (4) for $\mu = -K'(y_0)$. Thus, we obtain the adjoint equation (6). So the derivative of the map $v \rightarrow K[y(v)]$ at the point v_0 equals to $-p_0$. Then the functional I has the derivative

$$I'(v_0) = J'(v_0) - p_0$$

at this point. Using (2), we obtain the variational inequality (5). \square

Thus the Inverse Function Theorem is a good tool for proving the differentiability of the control-state mapping. This result is the basis for obtaining necessary optimality conditions. Note that we use now the serious assumption of the invertibility of the operator's derivative. It is equivalent to the existence of the unique solution $y \in Y$ for the linearized equation

$$A'(y_0)y = v \tag{7}$$

for all $v \in V$.

Now we have the following questions:

- How large is the class of operators that satisfy the mentioned assumption?
- What is the criterion of the differentiability of the inverse operator at a concrete point?
- Could we prove the differentiability of the inverse operator without using the Inverse Function Theorem?
- Could we prove a weaker form of the differentiability of the inverse operator for obtaining optimality conditions in the case of non-invertibility of the operator's derivative?

We will try to answer these questions.

14.2 Criterion for the Differentiability of the Inverse Operator

Consider an operator $A : Y \rightarrow V$. Let it be continuous and differentiable at a neighborhood of a point $y_0 \in Y$.

Theorem 1 *Suppose the existence of an open neighborhood O of the point y_0 such that the set $O' = A(O)$ is an open neighborhood of the point $v_0 = Ay_0$. Suppose that there exists the inverse operator $A^{-1} : O' \rightarrow O$, and that (7) has not more than one solution. Then this inverse operator is Gateaux differentiable at v_0 if and only if the derivative $A'(y_0)$ is a surjection.*

Proof Let the derivative $A'(y_0)$ be a surjection. Then it is invertible by the assumptions of the theorem. By Banach Inverse Operator Theorem there exists the continuous inverse operator $A'(y_0)^{-1}$. Therefore, the differentiability of the operator A^{-1} at the point v_0 follows from the Inverse Function Theorem directly.

Suppose now that the operator A^{-1} has the Gateaux derivative D at y_0 , and that the derivative $A'(y_0)$ is not a surjection. We get the equality

$$Ay(v_0 + \sigma v) - Ay(v_0) = \sigma v$$

for all $v \in V$ and small enough number σ . Dividing it by σ and passing to the limit as $\sigma \rightarrow 0$, using the Composite Function Theorem and differentiability of A^{-1} , we get

$$A'(y_0)Dv = v.$$

Then there exists a point $y = Dv$ from Y such that $A'(y_0)y = v$. So the derivative $A'(y_0)$ is a surjection. However this conclusion contradicts our assumption. Hence, the operator $A'(y_0)$ is a surjection whenever the inverse operator is differentiable. \square

Thus Gateaux differentiability of the inverse operator is equivalent to the following property: *the operator $A'(y_0)$ is a surjection*. It is called *Lusternik Condition* [4].

Consider as an example the homogeneous Dirichlet Problem for the equation

$$-\Delta y + |y|^\rho y = v \tag{8}$$

in the n -dimensional bounded set Ω , where $\rho > 0$. Denote the space

$$Y = H_0^1(\Omega) \cap L_q(\Omega),$$

where $q = \rho + 2$. Using Monotone Operators Theory [12], we obtain that this boundary problem has the unique solution $y \in Y$ for all v from the set V , which is the adjoint space

$$Y^* = H^{-1}(\Omega) + L_{q'}(\Omega),$$

where $1/q + 1/q' = 1$. Denote the operator $A : Y \rightarrow V$ such that Ay equals to the left side of the equality (8). The existence of the operator A^{-1} follows from the one-valued solvability of the boundary problem. Its differentiability can be obtained by using the properties of the linearized equation. It is the homogeneous Dirichlet Problem for the equation

$$-\Delta y + (\rho + 1)|y_0|^\rho y = v. \tag{9}$$

Corollary 1 *The solution of the Dirichlet problem for (8) is Gateaux differentiable with respect to the absolute term at the point $v = v_0$ iff (9) has a solution $y \in Y$ for all $v \in V$.*

Indeed, the continuous differentiability of the given operator A is obvious. The existence of the inverse operator follows from the one-valued solvability of the given boundary problem. It is obvious that the Dirichlet problem for the linear equation (9) cannot have two solutions. Then the criterion for the invertibility of the inverse operator is the Lusternik condition, by Theorem 1.

Now we obtain a criterion for the differentiability of the solution of (8) with respect to the absolute term on the space V .

Corollary 2 *The solution of the Dirichlet problem for (8) is Gateaux differentiable with respect to the absolute term at an arbitrary point if and only if the embedding $H_0^1(\Omega) \subset L_q(\Omega)$ is true.*

Proof Multiply equality (9) by the function y and integrate the result in $x \in \Omega$ using the Green formula and the boundary condition. We get

$$\int_{\Omega} |\nabla y|^2 dx + (\rho + 1) \int_{\Omega} |y_0|^\rho y^2 dx = \int_{\Omega} v y dx.$$

We have $Y = H_0^1(\Omega)$ by the given assumption, hence $V = H^{-1}(\Omega)$. So the a-priori estimate of the solution of (9) in the sense of Y for all $v \in V$ follows from the obtained equality. Now we get the one-valued solvability of the linearized equation by means of the standard theory of elliptic equations (see, for example, [11]). Thus the differentiability of the solution of (8) with respect to the absolute term at an arbitrary point follows from Corollary 1.

We prove now that the solution of (8) is not differentiable with respect to its absolute term, if the mentioned embedding does not hold. Let y_0 be a continuous function from the space Y . Then the left side of the equality (9) is a point of the space $H^{-1}(\Omega)$ for all $y \in Y$. Therefore, the image of the derivative $A'(y_0)$ is narrower than the set V , if the mentioned embedding does not hold. So (9) does not have any solutions from the space Y for all function v from the difference $V \setminus H^{-1}(\Omega)$. Therefore, the solution of the homogeneous Dirichlet problem for (8) is not Gateaux differentiable at the point

$$v_0 = -\Delta y_0 + |y_0|^\rho y_0$$

by Corollary 1. This completes the proof of Corollary 2. □

By Sobolev Theorem the embedding $H_0^1(\Omega) \subset L_q(\Omega)$ is true if $n = 2$ or $\rho \leq 4/(n - 2)$ for $n > 2$. Then the solution of (8) is differentiable with respect to the absolute term for small enough values of the set dimension n and nonlinearity parameter ρ . These characteristics determine a degree of the difficulty for the given equation. It is clear that the differentiability of the inverse operator (but not the absence of this property) follows from the Inverse Function Theorem. We will show soon that there exists another technique for proving this property. It is applicable even in the case of nondifferentiability in the sense of Gateaux. However it is important, that it satisfies some property which can be interpreted as a weak form of the differentiability.

The obtained result can be used for the analysis of optimization control problems for the system described by (8). Consider as an example the functional

$$I(v) = \frac{\alpha}{2} \|v\|_*^2 + \frac{1}{2} \|y(v) - y_d\|^2,$$

where $\alpha > 0$, $y_d \in H^{-1}(\Omega)$, and $y(v)$ is the solution of the Dirichlet problem (8) for the control v , besides $\|\cdot\|$ and $\|\cdot\|_*$ are the norms of the spaces $H_0^1(\Omega)$ and $H^{-1}(\Omega)$. Consider the following optimization problem.

Problem 2 Minimize the functional I on the convex closed subset U of the space V .

The solvability of this problem can be proved by a standard method (see, for example Chap. 1, Theorem 1.1 in [11]) using the weak continuity of the state function with respect to the absolute term. Note that the indeterminacy of the functional I on the complete set U is not an obstacle for the analysis of the optimization problem [13].

Corollary 3 *If $H_0^1(\Omega) \subset L_q(\Omega)$, then the solution v_0 of Problem 2 satisfies the inequality*

$$\int_{\Omega} (\alpha \Lambda v_0 - p_0)(v - v_0) dx \geq 0 \quad \forall v \in U, \quad (10)$$

where Λ is the canonical isomorphism of the spaces $H^{-1}(\Omega)$ and $H_0^1(\Omega)$, and p_0 is the solution of the homogeneous Dirichlet problem for the equation

$$-\Delta p_0 + (\rho + 1)|y_0|^\rho p_0 = \Delta y_0 - \Delta y_d. \quad (11)$$

Proof The derivative of the functional J (first term of the minimizing functional) is defined by the equality

$$\langle J'(v_0), h \rangle = \alpha(v_0, h)_* \quad \forall h \in H^{-1},$$

where $(\cdot, \cdot)_*$ is the scalar product of the space $H^{-1}(\Omega)$. By Riesz theorem there exists the canonical isomorphism $\Lambda : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$. Then we get

$$J'(v_0) = \alpha \Lambda v_0.$$

The derivative of the functional K (second term of the minimizing functional) is defined by the equality

$$\langle K'(y_0), h \rangle = (y_0 - y_d, h) \quad \forall h \in H_0^1(\Omega),$$

where (\cdot, \cdot) is the scalar product of the space $H_0^1(\Omega)$. Using Green formula, we obtain

$$K'(y_0) = \Delta y_d - \Delta y_0.$$

The operator $A'(y_0)$ is self-adjoint. Then the adjoint equation (6) transforms to (11), and the variational inequality (5) transforms to (10). This completes the proof of the corollary. \square

14.3 Differentiation of the Inverse Operator

We will try to prove the differentiability of the inverse operator directly without using of the Inverse function Theorem. Consider again an operator $A : V \rightarrow Y$ and a point $v_0 \in V$. Suppose the following assumption.

Property 1 The operator A is invertible in a neighborhood O of the point v_0 .

Choose a small enough positive number σ such that the point $v_\sigma = v_0 + \sigma h$ is in O for all $h \in V$. Denote by $y(v)$ the value $A^{-1}v$. Using the equalities $Ay(v_\sigma) = v_\sigma$, $Ay(v_0) = v_0$, we get

$$Ay(v_\sigma) - Ay(v_0) = \sigma h.$$

Assume the following property.

Property 2 The operator A is Gateaux differentiable.

By the Mean Value Theorem we obtain

$$Ay - Ay_0 = \left\{ \int_0^1 A' [y_0 + \theta(y - y_0)] d\theta \right\} (y - y_0),$$

where $y_0 = y(v_0)$. Then we have

$$G(v_\sigma)[y(v_\sigma) - y(v_0)] = \sigma h,$$

where the linear continuous operator $G(v) : Y \rightarrow V$ is defined by the formula

$$G(v) = \int_0^1 A' \{y_0 + \theta[y(v) - y_0]\} d\theta$$

for all $v \in V$. We get

$$\langle G(v_\sigma)^* \lambda, [y(v_\sigma) - y(v_0)]/\sigma \rangle = \langle \lambda, h \rangle \quad \forall \lambda \in V^*. \quad (12)$$

Consider the linear operator equation

$$G(v)^* p_\mu(v) = \mu. \quad (13)$$

It transforms to

$$A'(y_0)^* p_\mu(v_0) = \mu \quad (14)$$

for $v = v_0$. We will use the following assumption.

Property 3 Equation (13) has a unique solution $p_\mu(v) \in V^*$ for all $\mu \in Y^*$, $v \in O$.

Defining $\lambda = p_\mu(v_\sigma)$ for small enough σ in (12) we get

$$\langle \mu, [y(v_0 + \sigma h) - y(v_0)]/\sigma \rangle = \langle p_\mu(v_\sigma), h \rangle \quad \forall \mu \in Y^*, h \in V. \quad (15)$$

Define

$$M = \{ \mu \in Y^* \mid \|\mu\| = 1 \}.$$

Property 4 The convergence $p_\mu(v_\sigma) \rightarrow p_\mu(v_0)$ *-weakly in V^* uniformly with respect to $\mu \in M$ as $\sigma \rightarrow 0$ is true for all $v \in V$.

Theorem 2 *Let us suppose the Properties 1–4. Then the operator A^{-1} has the Gateaux derivative D at the point v_0 such that*

$$\langle \mu, Dh \rangle = \langle p_\mu(v_0), h \rangle \quad \forall \mu \in Y^*, h \in V. \tag{16}$$

Proof Let the operator D be defined by (16). It is a map from V to Y . Besides it is linear continuous. Using (15) and (16) we get

$$\begin{aligned} \left\| [y(v_0 + \sigma h) - y(v_0)]/\sigma - Dh \right\|_V &= \sup_{\mu \in M} \left| \langle \mu, [y(v_0 + \sigma h) - y(v_0)]/\sigma - Dh \rangle \right| \\ &= \sup_{\mu \in M} \left| \langle p_\mu(v_\sigma) - p_\mu(v_0), h \rangle \right| \end{aligned}$$

by the definition of the norm. Then we obtain $p_\mu(v_\sigma) \rightarrow p_\mu(v_0)$ *-weakly in V^* uniformly with respect to $\mu \in M$ for all $h \in V$ because of Property 4. Passing to the limit in the last equality as $\sigma \rightarrow 0$, we get the convergence

$$[y(v_0 + \sigma h) - y(v_0)]/\sigma \rightarrow Dh \quad \text{in } V \text{ for all } h \in V.$$

So the operator D is the Gateaux derivative of the operator A^{-1} at the point v_0 . \square

Let us explain applications of this result.

Lemma 3 *The operator A for (8) satisfies the Properties 1–4 if $H_0^1(\Omega) \subset L_q(\Omega)$.*

Proof Property 1 is the one-valued solvability of (8). The differentiability of the operator A (Property 2) is obvious, moreover, its derivative is defined by the equality

$$A'(y)h = -\Delta h + (\rho + 1)|y|^{\rho}h \quad \forall h \in Y.$$

Thus it is necessary to use Properties 3 and 4 and properties of the adjoint equation (13).

We have

$$\begin{aligned} G(v)y &= \left\{ \int_0^1 A' \{ y_0 + \theta [y(v) - y_0] \} d\theta \right\} y \\ &= -\Delta y + \left\{ \int_0^1 |y_0 + \theta [y(v) - y_0]|^\rho d\theta \right\} y \\ &= -\Delta y + |y_0 + \varepsilon [y(v) - y_0]|^\rho y \quad \forall y \in Y, \end{aligned}$$

where $\varepsilon \in [0, 1]$. Define

$$g(v) = |y_0 + \varepsilon [y(v) - y_0]|^{\rho/2},$$

so that we get

$$G(v)y = -\Delta y + g(v)^2 y.$$

Then we obtain the equality

$$\langle G(v)^* p, y \rangle = \langle p, G(v)y \rangle = \int_{\Omega} [-\Delta y + g(v)^2 y] p dx = \int_{\Omega} [-\Delta p + g(v)^2 p] y dx$$

for all $y \in Y, p \in V^*, v \in V$. So we get

$$G(v)^* p = -\Delta p + g(v)^2 p,$$

and (13) is transformed to

$$-\Delta p_{\mu}(v_{\sigma}) + g(v_{\sigma})^2 p_{\mu}(v_{\sigma}) = \mu. \tag{17}$$

Multiplying (17) by $p_{\mu}(v_{\sigma})$ and integrating in $x \in \Omega$ we have

$$\int_{\Omega} |\nabla p_{\mu}(v_{\sigma})|^2 dx + \int_{\Omega} |g(v_{\sigma}) p_{\mu}(v_{\sigma})|^2 dx = \int_{\Omega} \mu p_{\mu}(v_{\sigma}) dx.$$

Then we obtain the inequality

$$\|p_{\mu}(v_{\sigma})\|^2 + \|g(v_{\sigma}) p_{\mu}(v_{\sigma})\|_2^2 \leq \|\mu\|_* \|p_{\mu}(v_{\sigma})\|,$$

where $\|\cdot\|_p$ is the norm in $L_p(\Omega)$. So we get

$$\|p_{\mu}(v_{\sigma})\| \leq \|\mu\|_*, \quad \|g(v_{\sigma}) p_{\mu}(v_{\sigma})\|_2 \leq \|\mu\|_*. \tag{18}$$

Then (17) has the unique solution $p_{\mu}(v_{\sigma}) \in V^*$ for all $\mu \in Y^*, h \in V$, and σ , and hence Property 3 holds.

The space V is reflexive, so it is sufficient to prove that $p_{\mu}(v_{\sigma}) \rightarrow p_{\mu}(v_0)$ weakly in V^* uniformly with respect to μ as $\sigma \rightarrow 0$ for all $h \in V$. The set $\{p_{\mu}(v_{\sigma})\}$ is bounded in the space $H_0^1(\Omega)$, and the set $\{g(v_{\sigma}) p_{\mu}(v_{\sigma})\}$ is bounded in the space $L_2(\Omega)$ uniformly with respect to $\mu \in M$ for all $h \in V$ because of the inequalities (18). Using the Banach–Alaogly Theorem we get $p_{\mu}(v_{\sigma}) \rightarrow p$ weakly in $H_0^1(\Omega)$ uniformly with respect to $\mu \in M$ for all $h \in V$. Applying the Rellich–Kondrashov Theorem we get $p_{\mu}(v_{\sigma}) \rightarrow p$ strongly in $L_2(\Omega)$ and a.e. on Ω . Using the continuity of the solution of (8) with respect to the absolute term, we obtain $y(v_{\sigma}) \rightarrow y(v_0)$ in $H_0^1(\Omega)$ and a.e. on Ω . Then

$$|g(v_{\sigma})|^2 p_{\mu}(v_{\sigma}) \rightarrow (\rho + 1)|y_0|^{\rho} p \quad \text{a.e. on } \Omega.$$

The sets $\{p_{\mu}(v_{\sigma})\}, \{y(v_{\sigma})\},$ and $\{g(v_{\sigma})^{2/\rho}\}$ are uniformly bounded in $L_q(\Omega)$. We have

$$\begin{aligned} \|g(v_{\sigma})^2 p_{\mu}(v_{\sigma})\|_{q'} &\leq \|g(v_{\sigma}) p_{\mu}(v_{\sigma})\|_2 \|g(v_{\sigma})\|_{2q/\rho} \\ &= \|g(v_{\sigma}) p_{\mu}(v_{\sigma})\|_2 \|y_0 + \varepsilon[y(v_{\sigma}) - y_0]\|_q^{\rho/2}. \end{aligned}$$

So the set $\{g(v_\sigma)^2 p_\mu(v_\sigma)\}$ is uniformly bounded in $L_{q'}(\Omega)$. Using Lemma 1.3 (see Chap. 1 in [12]), we get

$$g(v_\sigma)^2 p_\mu(v_\sigma) \rightarrow (\rho + 1)|y_0|^\rho p \quad \text{weakly in } L_{q'}(\Omega)$$

uniformly with respect to $\mu \in M$ for all $h \in V$.

Let us multiply (16) by a function $\lambda \in H_0^1(\Omega)$. After integration we get

$$\int_\Omega [-\Delta p_\mu(v_\sigma) + g(v_\sigma)^2 p_\mu(v_\sigma)] \lambda dx = \int_\Omega \lambda \mu dx.$$

Passing to the limit as $\sigma \rightarrow 0$, we obtain, that the function $p = p_\mu(v_0)$ satisfies the equation

$$-\Delta p_\mu(v_0) + (\rho + 1)|y_0|^\rho p_\mu(v_0) = \mu. \quad (19)$$

Thus $p_\mu(v_\sigma) \rightarrow p_\mu(v_0)$ weakly in $H_0^1(\Omega)$ uniformly with respect to $\mu \in M$ for all $h \in V$, notably the Property 4 is true. \square

By Lemma 3 the differentiability of the solution of (8) with respect to the absolute term follows from Theorem 2 if the embedding $H_0^1(\Omega) \subset L_q(\Omega)$ holds.

Lemma 4 *Properties 1–4 follow from the assumptions of the Inverse Function Theorem.*

Proof The existence of the inverse operator is a corollary of the Inverse Function Theorem. The differentiability of the operator A is the assumption of this theorem. So our general difficulty is the analysis of (13), namely the justification of Assumptions 3 and 4. Equation (13) can be transformed to

$$G(v_0)^* p_\mu(v_\sigma) = A'(y_0)^* p_\mu(v_\sigma) = [G(v_0)^* - G(v_\sigma)^*] p_\mu(v_\sigma) + \mu.$$

The derivative $A'(y_0)$ is invertible by the Inverse Function Theorem. So its adjoint operator is invertible too. Then (13) can be transformed to the equality

$$p_\mu(v_\sigma) = L_\mu(\sigma h) p_\mu(v_\sigma), \quad (20)$$

where the map $L_\mu(\sigma h) : V^* \rightarrow V^*$ is defined by the formula

$$L_\mu(\sigma h) p = [A'(y_0)^*]^{-1} \{ [G(v_0)^* - G(v_\sigma)^*] p + \mu \}.$$

Using properties of the operator norm we get the inequality

$$\begin{aligned} \|L_\mu(\sigma h) p_1 - L_\mu(\sigma h) p_2\|_{V^*} &= \|[A'(y_0)^*]^{-1} [G(v_0)^* - G(v_\sigma)^*] (p_1 - p_2)\|_{V^*} \\ &\leq \|[A'(y_0)^*]^{-1}\| \|G(v_0)^* - G(v_\sigma)^*\| \|p_1 - p_2\|_{V^*} \end{aligned}$$

for all $p_1, p_2 \in V^*$. Then we obtain

$$\begin{aligned} & \|L_\mu(\sigma h)p_1 - L_\mu(\sigma h)p_2\|_{V^*} \\ & \leq \|A'(y_0)^{-1}\| \|G(v_0) - G(v_\sigma)\| \|p_1 - p_2\|_{V^*} \quad \forall p_1, p_2 \in V^* \end{aligned}$$

because of the equality of the norms for adjoint operators. The operator A^{-1} is continuous at the point v_0 by the Inverse Function Theorem. Therefore we get the convergence $y(v_0 + \sigma h) \rightarrow y_0$ in Y as $\sigma \rightarrow 0$ for all $h \in V$. Using the continuous differentiability of the operator A at the point y_0 , we get $G(v_\sigma) \rightarrow G(v_0)$ in the sense of the corresponding operator norm. The value σ can be chosen small enough such that

$$\|G(v_\sigma) - G(v_0)\| \leq \chi \|A'(y_0)^{-1}\|^{-1},$$

where $0 < \chi < 1$. So we obtain the estimate

$$\|L_\mu(\sigma h)p_1 - L_\mu(\sigma h)p_2\|_{V^*} \leq \chi \|p_1 - p_2\|_{V^*} \quad \forall p_1, p_2 \in V^*.$$

Thus the operator $L_\mu(\sigma h)$ is contracting. Then (20) has a unique solution $p_\mu(v_\sigma) \in V^*$ because of the Contracting Mapping Theorem.

We get $G(v_\sigma) \rightarrow G(v_0)$ as $\sigma \rightarrow 0$. So $G(v_\sigma)\lambda \rightarrow G(v_0)\lambda$ in V for all $\lambda \in Y$. Using the obtained inequalities, we get

$$\begin{aligned} \|p_\mu(v_\sigma)\|_{V^*} &= \|L_\mu(\sigma h)p_\mu(v_\sigma)\|_{V^*} \\ &\leq \|[A'(y_0)^*]^{-1}\| \|[G(v_0)^* - G(v_\sigma)^*]p_\mu(v_\sigma) + \mu\|_{Y^*} \\ &\leq \|A'(y_0)^{-1}\| \|[G(v_0) - G(v_\sigma)]\| \|p_\mu(v_\sigma)\|_{V^*} + \|\mu\|_{Y^*} \\ &\leq \chi \|p_\mu(v_\sigma)\|_{V^*} + \|A'(y_0)^{-1}\| \|\mu\|_{Y^*}. \end{aligned}$$

So we have

$$(1 - \chi) \|p_\mu(v_\sigma)\|_{V^*} \leq \|A'(y_0)^{-1}\| \|\mu\|_{Y^*}.$$

Then $p_\mu(v_\sigma) \rightarrow p$ *-weakly in V^* for all $h \in V$ as $\sigma \rightarrow 0$.

Using inequality (13) we get

$$\langle p_\mu(v_\sigma), G(v_\sigma)\lambda \rangle = \langle \mu, \lambda \rangle \quad \forall \lambda \in Y.$$

As a consequence $\{p_\mu(v_\sigma)\}$ converges *-weakly, and $\{G(v_\sigma)\}$ converges strongly. After passing to the limit we have $A'(y_0)^* p = \mu$, and $p = p_\mu(v_0)$. \square

Thus the assumptions of Theorem 2 follow from the assumptions of the Inverse Operator Theorem. However assertions of Theorem 2 may be true if assumptions of the Inverse Operator Theorem are not satisfied.

14.4 Extended Differentiation of the Inverse Operator

The solution of (8) is differentiable with respect to the absolute term for small enough values of the set dimension n and nonlinearity parameter ρ . But it is not differentiable for large enough values of these parameters. Suppose $n \geq 3$. By Sobolev Theorem the embedding $H_0^1(\Omega) \subset L_q(\Omega)$ is true if $\rho \leq 4/(n-2)$. It guarantees the differentiability of the considered inverse operator. However this embedding fails if the parameter ρ increases. Then the solution of the equation becomes non-differentiable with respect to the absolute term. It seems to be a strange situation. Properties of the inverse operator change with a jump at the neighborhood of some value ρ . The differentiability of the operator disappears after the passage of this value. This situation seems not likely. We could suppose the existence of a weaker operator differentiability than the Gateaux derivative. We would like also to determine the extension of the operator derivative because the solvability of our optimization problem was proved for all values of the set dimension and the nonlinearity parameter.

There exist extensions of classical operator differentiation, for example, subdifferential calculus [14], Clarke derivatives [15], quasidifferential calculus [7]. They are used also for the resolution of nonsmooth optimization problems. These results are effective enough for the analysis of operators with nonsmooth terms, for example, the absolute value or the maximum of functions. However similar terms are absent in our case. So we will try to define another form of operator derivatives extension.

It is known that “*the general idea of the differential calculus is a local approximation of a function by a linear function*” (see p. 170 in [16]). The differentiation is a tool of the local approximation of the analyzed object. The desired form of an operator derivative can be obtained by weakening of topological approximation properties of the differentiation. Then we get the extended operator derivative (see [17–19]).

Definition An operator $L : V \rightarrow Y$ is called $(V_0, Y_0; V_1, Y_1)$ -*extended differentiable in the sense of Gateaux* at the point $v_0 \in V$ if there exist linear topological spaces V_0, Y_0, V_1, Y_1 with continuous embeddings

$$V_1 \subset V_0 \subset V, \quad Y \subset Y_0 \subset Y_1,$$

and a linear continuous operator $D : V_0 \rightarrow Y_0$ such that

$$[L(v_0 + \sigma h) - L(v_0)]/\sigma \rightarrow Dh \quad \text{in } Y_1 \text{ for all } h \in V_1$$

as $\sigma \rightarrow 0$.

It is obvious that the $(V, Y; V, Y)$ -derivative is the standard Gateaux derivative. The following result is known (see Theorem 4 in [18]; Theorem 5.4 in [19]).

Lemma 5 *The operator A^{-1} for (8) is $(V_0, Y_0; V_1, Y_1)$ -extended differentiable in the sense of Gateaux at an arbitrary point $v_0 \in V$, where*

$$Y_1 = H_0^1(\Omega), \quad Y_0 = Y_1 \cap \{y \mid |y_0|^{\rho/2} y \in L_2(\Omega)\},$$

$$V_1 = H^{-1}(\Omega), \quad V_0 = V_1 + \{v \mid v = |y_0|^{\rho/2} \varphi, \varphi \in L_2(\Omega)\}, \quad y_0 = y(v_0),$$

moreover, its derivative D satisfies the equality

$$\int_{\Omega} \mu D h dx = \int_{\Omega} p_{\mu}(v_0) h dx \quad \forall \mu \in Y_0^*, h \in V_0, \tag{21}$$

and $p_{\mu}(v_0)$ is the solution of the homogeneous Dirichlet problem for (19).

Thus the inverse operator for the given example is extended differentiable for all values of the set dimensions and nonlinearity parameters. Its extended derivative is transformed to the Gateaux one for small enough values of these characteristics. However the Gateaux derivative does not exist for its large enough values, notably in the case of the high enough degree of the difficulty for the problem. Besides the difference between standard derivative and extended one is determined by this degree of the difficulty. Thus the inverse operator is extended differentiable without any constraints. However the extended derivative differs from the classical one after the augmentation of the parameters that determine the degree of the difficulty for the problem. Then we obtain the gradual change of the inverse operator properties after the gradual change of its parameters, although the standard derivatives theory permits the change with a jump.

We will prove that the obtained result is sufficient for the analysis of the given optimization problem without any constraints.

Corollary 4 *The solution of the minimization problem of the functional I on the set U for (8) satisfies the variational inequality*

$$\int_{\Omega} (\alpha \Lambda v_0 - p_0)(v - v_0) dx \geq 0 \quad \forall v \in U_1, \tag{22}$$

where $U_1 = U \cap (v_0 + V_1)$, and p_0 is a solution of (11).

Indeed, if v_0 is a solution of the optimization problem, then

$$I[v_0 + \sigma(v - v_0)] - I(v_0) \geq 0 \quad \forall v \in U.$$

Let us choose $v \in v_0 + V_1$. Passing to the limit and using Lemma 5 after division by σ we get

$$\int_{\Omega} \alpha \Lambda v_0 (v - v_0) dx + \int_{\Omega} \nabla(y_0 - y_d) \nabla D(v - v_0) dx \geq 0 \quad \forall v \in U_1.$$

Then the inequality (22) is true.

If $H_0^1(\Omega) \subset L_q(\Omega)$, then $U_1 = U$, and the variational inequalities (10) and (22) are equal. Thus necessary conditions of optimality can be obtained without any assumptions by means of the extended derivatives theory. Optimization problems for elliptic equations with power nonlinearity without Gateaux differentiability of the control-state mapping were considered in [18, 19]. But the control space was narrower, and the state functional was more regular there. This technique was used for the analysis of optimization problems for others equations in [20].

Note that Lemma 5 uses the technique of the proof of Theorem 2. We can suppose that it is possible to obtain the extended differentiability of the inverse operator in the general case. Consider Banach spaces Y, V , a map $A : Y \rightarrow V$, and points $y_0 \in Y, v_0 = Ay_0$. Let V_1 be a Banach subspace of V with a neighborhood O_1 of zero. Then $O = v_0 + O_1$ is a neighborhood of v_0 . We suppose the following assertion.

Property 5 The operator A is invertible on the set O .

Define $y(v) = A^{-1}v$. We get the equality

$$Ay(v_\sigma) - Ay(v_0) = \sigma h$$

for all $v \in V_1$ and small enough σ , where $v_\sigma = v_0 + \sigma h$. Let $G(v)$ be the operator from the proof of Theorem 2. We have

$$G(v_\sigma)[y(v_\sigma) - y(v_0)] = \sigma h,$$

so

$$\langle \lambda, G(v_\sigma)[y(v_\sigma) - y(v_0)] \rangle = \sigma \langle \lambda, h \rangle \quad \forall \lambda \in V^*.$$

Consider Banach spaces $V(v)$ and $Y(v)$ such that the embeddings of the spaces Y, Y_1 and $Y(v)$ to $V(v), Y(v)$ and V , respectively, are continuous for all $v \in O$. Let the following assumption be true.

Property 6 The operator A is Gateaux differentiable, moreover, there exists the continuous extension $\overline{G}(v)$ of the operator $G(v)$ to $Y(v)$ such that its image is a subset of $V(v)$ for all $v \in O$.

Using the properties $y(v) \in y_0 + Y(v)$ and $V(v)^* \subset V^*$ we get

$$\langle \overline{G}(v_\sigma)^* \lambda, [y(v_\sigma) - y(v_0)] \rangle = \sigma \langle \lambda, h \rangle \quad \forall \lambda \in V^*. \tag{23}$$

It is an analogue of (12). Consider the linear operator equation

$$\overline{G}(v_\sigma)^* p_\mu(v_\sigma) = \mu, \tag{24}$$

which is an analogue of (13). It can be transformed to

$$\overline{A}'(y_0) p_\mu(v_0) = \mu$$

for $v = v_0$, where $\overline{A'}(y_0) = \overline{G}(v_0)$ is the extension of the operator $A'(y_0) = G(v_0)$ to the set $Y(v_0)$.

Consider Banach space V_1 such that the embedding $Y(v) \subset V_1$ is continuous and dense for all $v \in O$. We suppose the following condition.

Property 7 Equation (24) has the unique solution $p_\mu(v) \in V(v)^*$ for all $v \in O$, $\mu \in Y(v)^*$.

Defining in (23) $\lambda = p_\mu(v_\sigma)$ for a small enough σ we get

$$\langle \mu, [y(v_0 + \sigma h) - y(v_0)]/\sigma \rangle = \langle p_\mu(v_\sigma), h \rangle \quad \forall \mu \in Y(v_\sigma)^*, h \in V_1. \quad (25)$$

We will use the additional assumption.

Property 8 The convergence $p_\mu(v_\sigma) \rightarrow p_\mu(v_0)$ holds *-weakly in V_1^* uniformly with respect to $\mu \in M$ as $\sigma \rightarrow 0$ for all $h \in V_1$.

The extended differentiability of the inverse operator is guaranteed by the following result.

Theorem 3 *Let us suppose the Properties 5–8. Then the operator A^{-1} has the $(V(v_0), Y(v_0); V_1, Y_1)$ -extended Gateaux derivative D at the point v_0 such that*

$$\langle \mu, Dh \rangle = \langle p_\mu(v_0), h \rangle \quad \forall \mu \in Y(v_0)^*, h \in V(v_0). \quad (26)$$

Proof By (25), (26) we get

$$\begin{aligned} & \langle \mu, [y(v_0 + \sigma h) - y(v_0)]/\sigma - Dh \rangle \\ &= \langle p_\mu(v_\sigma) - p_\mu(v_0), h \rangle \quad \forall \mu \in M, h \in V(v_1). \end{aligned} \quad (27)$$

Then

$$\| [y(v_0 + \sigma h) - y(v_0)]/\sigma - Dh \|_{V_1} = \sup_{\mu \in M} | \langle p_\mu(v_\sigma) - p_\mu(v_0), h \rangle |.$$

We have $p_\mu(v_\sigma) \rightarrow p_\mu(v_0)$ *-weakly in V_1 uniformly with respect to $\mu \in M$ for all $h \in V_1$ as $\sigma \rightarrow 0$ by Property 8. Passing to the limit in the last equality we obtain

$$[y(v_0 + \sigma h) - y(v_0)]/\sigma \rightarrow Dh \quad \text{in } Y_1$$

for all $h \in V_1$. Thus D is an extended derivative of the inverse operator. □

A result of the extended differentiability of the inverse operator for nonnormalized spaces was obtained in [18].

Let us prove that the assumptions of the Theorem 3 are true for the considered example.

Lemma 6 *The operator A , which is defined by (8), satisfies the Properties 5–8.*

Proof The Property 5 is the solvability of (8) at a neighborhood of the given point. It is obviously that this assumption is true. The differentiability of the operator A is clear. The operator $G(v)$ for our case is determined by the equality

$$G(v)y = -\Delta y + g(v)^2 y \quad \forall y \in Y,$$

where

$$g(v)^2 = (\rho + 1)|y_0 + \varepsilon[y(v) - y(v_0)]|^\rho, \quad \varepsilon \in [0, 1].$$

Let the spaces $Y_1, V_1, Y(v), V(v)$ be those defined in the proof of Lemma 5. Define the map $\overline{G}(v)$ by the equality

$$\overline{G}(v)y = -\Delta y + g(v)^2 y \quad \forall y \in Y(v).$$

Then Property 6 is true. Reliability of Properties 7 and 8 was obtained in the proof of Lemma 5. \square

Thus extended differentiability of the inverse operator for (8) follows from Theorem 3.

Lemma 7 *the Properties 5–8 follow from the assumptions of the Inverse Function Theorem.*

Indeed, Property 5 is a direct corollary of this theorem. Let us define the spaces $V = V_1, Y = Y_1$. Then we get $Y(v) = Y, V(v) = V$. So the operator $\overline{G}(v)$ is equal to $G(v)$, and Property 6 is trivial. Therefore Properties 7 and 8 are transformed to Properties 3 and 4. Its validity was proved before.

Thus Theorem 3 is a generalization of the Theorem 2. The obtained results can be used for other applications if it is necessary to differentiate an inverse operator. For example the extended differentiable submanifolds of Banach spaces are defined in [21, 22]. Optimization control problems on differentiable submanifolds are considered there. Analogical results could be obtained for the implicit operator, including the case of nonnormalized spaces (see [23]). Banach spaces with extended differentiable operators form a category, and necessary conditions of optimality have a category interpretation (see [24]).

References

1. Aubin, J.P., Ekeland, I.: Applied Nonlinear Analysis. Wiley, New York (1984)
2. Kantorovich, L.V., Akilov, G.P.: Functional Analysis. Nauka, Moscow (1977)
3. Lusternik, L.A.: About extremum functionals conditions. Math. USSR Sb. **3**, 390–401 (1934)
4. Dmitruk, A.V., Milutin, A.A., Osmolovsky, N.P.: Lusternik's theorem and extremum theory. Usp. Mat. Nauk **35**(6), 11–46 (1980)

5. Milnor, J.: Analytic proof of the hairy ball theorem and the Brouwer fixed point theorem. *Am. Math. Mon.* **85**, 521–524 (1978)
6. Frankowska, H.: High order inverse function theorems. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **6**, 283–303 (1989)
7. Dem'yaniv, F.V., Rubinov, A.M.: *Basis of the Nonsmooth Analysis and Quasidifferential Calculus*. Nauka, Moscow (1990)
8. Radulescu, R., Radulescu, M.: Local inversion theorems without assuming continuous differentiability. *J. Math. Anal. Appl.* **138**, 581–590 (1989)
9. Cristea, M.: Local inversion theorems without assuming continuous differentiability. *J. Math. Anal. Appl.* **143**, 259–263 (1989)
10. Aubin, J.P., Frankowska, H.: On inverse function theorems for set-valued maps. *J. Math. Pures Appl.* **66**, 71–89 (1987)
11. Lions, J.L.: *Contrôle Optimal de Systèmes Gouvernés par des Equations aux Derivées Partielles*. Dunod, Paris (1968)
12. Lions, J.L.: *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*. Dunod, Paris (1969)
13. Lions, J.L.: *Contrôle de Systèmes Distribués Singuliers*. Gauthier-Villars, Paris (1983)
14. Ekeland, I., Temam, R.: In: *Convex Analysis and Variational Problems*. North-Holland, Amsterdam (1976)
15. Clarke, F.H.: *Optimization and Nonsmooth Analysis*. Wiley, New York (1983)
16. Dieudonne, J.: *Foundation of Modern Analysis*. Nauka, Moscow (1964)
17. Serovajsky, S.: Differentiation of the inverse function for nonnormalised spaces. *Funct. Anal. Appl.* **27**(4), 84–87 (1993)
18. Serovajsky, S.: Calculation of functional gradients and extended differentiation of operators. *J. Inverse Ill-Posed Probl.* **13**(4), 383–396 (2005)
19. Serovajsky, S.: *Optimization and Differentiation*, vol. 1. Print-S, Almaty (2006)
20. Serovajsky, S.: Necessary conditions of optimality for the case of nondifferentiability of the control-state mapping. *Differ. Equ.* **31**(6), 1055–1059 (1995)
21. Serovajsky, S.: Extended differentiability of the implicit function in the spaces without norms. *Russ. Math. (Izv. VUZ)* **12**, 55–63 (1991)
22. Serovajsky, S.: Extremal problems on differentiable manifolds of Banach spaces. *Russ. Math. (Izv. VUZ)* **5**, 83–86 (1996)
23. Serovajsky, S.: Extended differentiable submanifolds. *Russ. Math. (Izv. VUZ)* **1**, 56–65 (1997)
24. Serovajsky, S.: Differentiation of operators and conditions of the extremum with category interpretation. *Russ. Math. (Izv. VUZ)* **2**, 66–76 (2010)