A Newton derivative scheme for shape optimization problems constrained by variational inequalities

Nico Goldammer *, Volker H. Schulz † and Kathrin Welker ‡

Abstract
Shape optimization problems constrained by variational inequalities (VI) are non-smooth and non-convex optimization problems. The non-smoothness arises due to the variational inequality constraint, which makes it challenging to derive optimality conditions. Besides the non-smoothness there are complementary aspects due to the VIs as well as distributed, non-linear, non-convex and infinite-dimensional aspects due to the shapes which complicate to set up an optimality system and, thus, to develop fast and higher order solution algorithms. In this paper, we consider Newton-derivatives in order to formulate optimality conditions. In this context, we set up a Newton-shape derivative scheme. Examples show the application of the proposed scheme.

Key words. Newton derivative, variational inequality, shape optimization, material derivative, shape derivative, optimality conditions

AMS subject classifications. 49Q10, 49J40, 35Q93, 65K15

1 Introduction
Optimal control problems with constraints in the form of variational inequalities (VI) are challenging, since classical constraint qualifications for deriving Lagrange multipliers generically fail. Therefore, not only the development of stable numerical solution schemes but also the development of suitable first order optimality conditions is an issue. By usage of tools of modern analysis, such as monotone operators in Banach spaces, significant results on properties of the solution operator of variational inequalities have been achieved since the 1960s (cf. [6, 7, 31]). Comprehensive studies of variational inequalities and more references can be found in [15, 26, 27, 44]. The generic non-smoothness and non-convexity in the feasible set described by variational inequalities causes difficulties already in finite dimensional versions of the problem. In fact, finite dimensional bilevel optimization (i.e., optimization with optimization problems in the constraints) is its own field of research since the 1970s (cf., e.g., [5]) and has been generalized to mathematical programming with equilibrium constraints (MPECs) for the optimization of stationary systems of constrained problems in [10]. For a survey on bilevel programming and MPECs see, e.g., [35]. In [45], the authors concentrate on the typical complementarity structure

*Helmut-Schmidt-University / University of the Federal Armed Forces Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany, goldammer@hsu-hh.de
†Trier University, Universitätsring 15, 54296 Trier, Germany, volker.schulz@uni-trier.de
‡Helmut-Schmidt-University / University of the Federal Armed Forces Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany, welker@hsu-hh.de
of variational inequalities and derive a hierarchy of stationarity concepts (depending on constraint qualification conditions) for the more general problem class of mathematical programs with complementarity constraints (MPCCs). During the last decade, these concepts have partly been transferred to respective concepts in function space in [19, 20, 21]. The optimal control of variational inequalities that are posed in function space has been studied since the 1970s and necessary stationarity conditions have been derived by use of penalty and smoothing techniques and strengthened by the usage of instruments from convex analysis and differentiability, see, e.g., [2, 38, 42]. The conditions that a solution can be shown to verify have a complex structure and the problem to find candidates for solutions leads to a system of non-linear and non-smooth equations. This demands for the development of numerical algorithms and a proper mathematical analysis on their convergence behavior, see, e.g., the discussion in [28, 43].

In this paper, we consider shape optimization problems constrained by variational inequalities. These problems are non-smooth and non-convex optimization problems. The non-smoothness arises due to the variational inequality constraint, which makes it challenging to derive optimality conditions. Moreover, besides the non-smoothness there are complementary aspects due to the VIs as well as distributed, non-linear, non-convex and infinite dimensional aspects due to the shapes which complicate to set up an optimality system. In particular, one cannot expect for an arbitrary shape functional depending on solutions to VIs the existence of the shape derivative or to obtain the shape derivative as a linear mapping. In addition, the adjoint state can generally not be introduced and, thus, an optimality system cannot be setted up. In this paper, we consider Newton derivatives instead of classical derivatives in order to formulate optimality conditions. The proposed Newton derivative scheme enables the analytical and computational treatment of shape optimization problems constrained by VIs which are non-shape differentiable in the classical sense such that these can handled and solved without any regularization techniques leading often only to approximated shape solutions.

So far, there are only very few approaches in literature to the problem class of VI constrained shape optimization problems. In [29], shape optimization of 2D elasto-plastic bodies is studied, where the shape is simplified to a graph such that one dimension can be written as a function of the other. In [47, Chap. 4], shape derivatives of elliptic variational inequality problems are presented in the form of solutions to again variational inequalities. In [39], shape optimization for 2D graph-like domains are investigated. Also [32, 33] present existence results for shape optimization problems which can be reformulated as optimal control problems, whereas [12, 14] show existence of solutions in a more general set-up. In [40, 41], level-set methods are proposed and applied to graph-like two-dimensional problems. Moreover, [22] presents a regularization approach to the computation of shape and topological derivatives in the context of elliptic variational inequalities and, thus, circumventing the numerical problems in [47, Chap. 4]. In [17], the analysis of state material derivatives is significantly refined over [47, Chap. 4]. All these mentioned problems have in common that one cannot expect for an arbitrary shape functional depending on solutions to VIs to obtain the shape derivative as a linear mapping (cf. [47, Example in Chap. 1]). In general, the shape derivative for the obstacle problems fails to be linear with respect to the normal component of the vector field defined on the boundary of the domain under consideration. In order to circumvent the problems related to the non-linearity of the shape derivative and in particular the non-existence of the shape derivative of a VI constrained shape optimization problem, this paper sets up a Newton (shape) derivative scheme in shape spaces. This would also open the door for formulating higher order optimization methods on shape spaces, e.g., the semi-smooth Newton method. With the help of a Newton derivative scheme, VI constrained shape optimization problems which
are non-shape differentiable in the classical sense can handled and solved without any regularization techniques.

This paper is structured as follows. In section 2, we set up a novel Newton derivative framework (including its necessary rules like the chain rule) for non-smooth shape optimization problems. In particular, we define a Newton material derivative and also a Newton shape derivative. The developed framework allows the formulation of optimality conditions to VI constrained shape optimization problems in section 3. Finally, we apply the Newton shape derivative framework to a shape optimization problem constrained by VIs of first kind in section 4. More precisely, based on the Newton derivative framework of section 2 we set up the optimality system of section 3 for the obstacle-type problem.

2 Newton derivative calculus

A main focus in shape optimization is in the investigation of shape functionals. A shape functional on an arbitrary shape space $\mathcal{U}$ is given by a function $J: \mathcal{U} \rightarrow \mathbb{R}$, $u \mapsto J(u)$. In general, a shape optimization problem can be formulated by

$$
\min_{u \in \mathcal{U}} J(u).
$$

Often, shape optimization problems are constrained by equations, e.g., equations involving an unknown function of two or more variables and at least one partial derivative of this function. The objective may depend not only on the shapes $u$ but also on the state variable $y$, where the state variable is the solution of the underlying constraint. We concentrate on constraints in the form of variational inequalities in this paper. These problems are in particular highly challenging because one cannot expect for an arbitrary shape functional depending on solutions to VIs the existence of the adjoint and of the shape derivative or to obtain the shape derivative as a linear mapping. In this section, we consider Newton derivatives instead of classical derivatives in order to be able to formulate optimality conditions in section 3.

We start by defining generally a Newton derivative and its necessary rules in subsection 2.1, which ends in formulating necessary conditions of optimality for Newton differentiable objective functions (cf. theorem 2.11). In subsection 2.2, we concentrate on shape optimization and extend the Newton derivative concepts to shape calculus.

2.1 Newton derivative calculus

In this subsection, we first define generally a Newton derivative and formulate some rules like the chain and product rule. Based on the Newton derivative, we introduce the Newton gradient. The new objects are demonstrated on some examples.

We follow [9] for the definition of a Newton derivative.

**Definition 2.1 (Newton derivative).** Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Let $L(X; Y)$ denote the set of all linear mappings from $X$ to $Y$. We call $F: X \rightarrow Y$ Newton differentiable in $x \in X$ if there exists a neighborhood $U \subset X$ of $x$ and a
mapping $D_N F: U \to \mathbb{L}(X;Y)$ such that

$$\lim_{\|h\|_X \to 0} \frac{\|F(x + h) - F(x) - D_N F(x + h)h\|_Y}{\|h\|_X} = 0.$$  

We then call $D_N F(x)$ a Newton derivative of $F$ at $x$.

**Remark 2.2.** We notice that the Newton derivative differs from the Fréchet derivative. At first, the evaluation in the Newton derivative is in $x + h$ and not in $x$. In addition, there are no requirements on a connection between $D_N F$ and $F$, while there is a link for the Fréchet derivative of $F$ due to the directional derivative of $F$. Furthermore a Newton derivative is not unique, e.g., changing one point in a Newton derivative $D_N F$ does not change the fact that the resulting function $D_N F$ is still a Newton derivative.

**Remark 2.3.** Obviously, if a function is Fréchet differentiable then it is also Newton differentiable. Thanks to [4, Theorem 14.8] we also know that a piecewise differentiable function $F: \mathbb{R}^N \to \mathbb{R}$ is Newton differentiable for all $x \in \mathbb{R}^N$.

As stated in [9], there is a chain rule for the Newton derivative.

**Lemma 2.4 (Chain rule).** Let $X, Y$ and $Z$ be Banach spaces. Moreover, let $F: X \to Y$ be Newton differentiable at $x \in X$ with Newton derivative $D_N F(x)$ and $G: Y \to Z$ be Newton differentiable at $y := F(x) \in Y$ with Newton derivative $D_N G(y)$. If $D_N F(x), D_N G(y)$ are uniformly bounded in a neighborhood of $x, y$, respectively, then $G \circ F$ is also Newton differentiable at $x$ with Newton derivative

$$D_N (G \circ F)(x) = D_N G(y) \circ D_N F(x).$$

In addition, we provide a product rule in the following lemma.

**Lemma 2.5.** Let $X$ and $Y$ be Banach spaces such that $\|xy\|_Y = \|x\|_X \|y\|_Y$ holds for all $x, y \in Y$. Moreover, let $F, G: X \to Y$ be Newton differentiable at $x \in X$ with Newton derivative $D_N F(x), D_N G(x)$, respectively. If

$$\frac{\|F(x + h) - F(x)\|_Y \|D_N G(x + h)h\|_X}{\|h\|_X} \to 0 \quad \text{for} \quad \|h\|_X \to 0$$

and $F, D_N G$ are uniformly bounded in a neighborhood of $x$, then

$$D_N (EF)(x) = F(x)D_N G(x) + D_N F(x)G(x).$$

**Proof.** If we considering the numerator and adding a zero we obtain

$$\|F(x + h)G(x + h) - F(x)G(x) - F(x)D_N G(x + h)h - D_N F(x + h)hG(x)\|_Y$$

$$= \|F(x + h)G(x + h) - F(x)G(x) - F(x + h)G(x) + F(x + h)G(x)\|_Y$$

$$- F(x)D_N G(x + h)h - D_N F(x + h)hG(x)\|_Y$$

$$= \|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h\|_Y$$

$$+ (F(x + h) - F(x) - D_N F(x + h))G(x)\|_Y.$$  

Adding $F(x + h)D_N G(x + h)h - F(x + h)D_N G(x + h)h$ after using the triangle inequality leads to

$$\|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h\|_Y$$

$$+ (F(x + h) - F(x) - D_N F(x + h))G(x)\|_Y$$

$$\leq \|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h\|_Y$$

$$+ \|((F(x + h) - F(x) - D_N F(x + h))G(x)\|_Y$$

$$= \|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h + F(x + h)D_N G(x + h)h - F(x + h)D_N G(x + h)h\|_Y$$

$$+ \|((F(x + h) - F(x) - D_N F(x + h))G(x)\|_Y.$$
Figure 1: Graph of $F$ (left) and $D_N F$ (right) in $[-2, 2]$. The blue line represents a choice of $\delta \in \mathbb{R}$ at 0.

The triangle inequality yields again

\[ \|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h + F(x + h)D_N G(x + h)h - F(x + h)D_N F(x + h)hG(x)\|_Y + \|(F(x + h) - F(x) - D_N F(x + h)h)G(x)\|_Y \]
\[ \leq \|F(x + h)(G(x + h) - G(x)) - F(x)D_N G(x + h)h\|_Y + \|F(x + h)D_N G(x + h)h\|_Y + \|(F(x + h) - F(x) - D_N F(x + h)h)G(x)\|_Y \]
\[ = \|F(x + h)\|_Y \|G(x + h) - G(x) - D_N G(x + h)h\|_Y \]
\[ + \|F(x + h)\|_Y \|D_N G(x + h)h\|_Y \]
\[ + \|F(x + h) - F(x) - D_N F(x + h)h\|_Y \|G(x)\|_Y \]

Since $F$ as well as $G$ are Newton differentiable at $x$ and $F$ as well as $D_N G$ are uniformly bounded near $x$ and $\|F(x + h)\|_Y \|D_N G(x + h)h\|_Y \to 0$ for $\|h\| \to 0$, we finally obtain

\[ \lim_{\|h\| \to 0} \frac{\|F(x + h)G(x + h) - F(x)G(x) - F(x)D_N G(x + h)h - D_N F(x + h)hG(x)\|_Y}{\|h\|_X} \]
\[ \leq \lim_{\|h\| \to 0} \left( \frac{\|F(x + h)\|_Y \|G(x + h) - G(x) - D_N G(x + h)h\|_Y}{\|h\|_X} \right) \]
\[ + \frac{\|F(x + h)\|_Y \|D_N G(x + h)h\|_Y}{\|h\|_X} \]
\[ + \frac{\|F(x + h) - F(x) - D_N F(x + h)h\|_Y \|G(x)\|_Y}{\|h\|_X} \]
\[ = \lim_{\|h\| \to 0} \frac{\|F(x + h)\|_Y \|G(x + h) - G(x) - D_N G(x + h)h\|_Y}{\|h\|_X} \]
\[ + \lim_{\|h\| \to 0} \frac{\|F(x + h)\|_Y \|D_N G(x + h)h\|_Y}{\|h\|_X} \]
\[ + \lim_{\|h\| \to 0} \frac{\|F(x + h) - F(x) - D_N F(x + h)h\|_Y \|G(x)\|_Y}{\|h\|_X} \]
\[ = 0. \]

For a better understanding we give some examples of the novel objects.

**Example 2.6.** Let $F(x) := \max\{0, x\}$ with $x \in \mathbb{R}$. The Newton derivative $D_N F$ of $F$ is given by

\[ D_N F(x) = \begin{cases} 
1, & x > 0 \\
\delta, & x = 0 \\
0, & x < 0 
\end{cases} \]  \hspace{1cm} (2.2)
In order to verify (2.2), thanks to definition 2.1 we need to show
\[ \lim_{h \downarrow 0} \frac{|F(x + h) - F(x) - D_N F(x + h)h|}{h} = 0, \]
where \( |\cdot| \) denotes the absolute value. We consider three cases:

- **If** \( x > 0 \), **we know** \( x + h > 0 \) for any \( h > 0 \). As a result we have
  \[ \lim_{h \downarrow 0} \frac{|F(x + h) - F(x) - D_N F(x + h)h|}{h} = \lim_{h \downarrow 0} \frac{|x + h - x - h|}{h} = 0. \]

- **If** \( x < 0 \), **then** \( x + h < 0 \) for sufficiently small \( h \). Therefore, we have
  \[ \lim_{h \downarrow 0} \frac{|F(x + h) - F(x) - D_N F(x + h)h|}{h} = \lim_{h \downarrow 0} \frac{0 - 0}{h} = 0. \]

- **If** \( x = 0 \), **then** \( x + h = h \) and so we have
  \[ \lim_{h \downarrow 0} \frac{|F(x + h) - F(x) - D_N F(x + h)h|}{h} = \lim_{h \downarrow 0} \frac{|h - h|}{h} = 0. \]

Therefore we can choose any value \( \delta \in \mathbb{R} \) and define \( D_N F(0) = \delta \).

**Example 2.7.** Let \( F \) be given as in example 2.6, i.e., \( F(x) := \max\{0, x\} \) with \( x \in \mathbb{R} \). In addition, let \( H \) be the periodic continuation of \( h : [-2, 2] \rightarrow \mathbb{R}, x \mapsto |x| - 1 \).

A calculation similar to the one of \( D_N F \) in example 2.6 gives the Newton derivative of \( H \):

\[ D_N H(x) = \begin{cases} 
1, & x \in (2z, 2z + 1) \\
-1, & x \in (2z - 1, 2z) \\
\delta, & \text{else} \end{cases} \]

for any value \( \delta \in \mathbb{R} \) and for \( z \in \mathbb{Z} \); see figure 4 for a visualization. We should note here that \( \delta \) may change from point to point. In this example, we focus on the calculation of the Newton derivative of the composition \( F \circ H \). The composition \( F \circ H \) is given by

\[ F(H(x)) = \begin{cases} 
0, & x \in \bigcup_{z \in \mathbb{Z}} [-1 + 4z, 1 + 4z] \\
H(x), & x \in \bigcup_{z \in \mathbb{Z}} (-3 + 4z, -1 + 4z) \end{cases}. \]

Without using the chain rule (cf. lemma 2.4) the Newton derivative of \( F(H(x)) \) is given by

\[ G_1(x) = \begin{cases} 
0, & x \in \bigcup_{z \in \mathbb{Z}} (-1 + 4z, 1 + 4z) \\
1, & x \in \bigcup_{z \in \mathbb{Z}} (-3 + 4z, -2 + 4z) \\
-1, & x \in \bigcup_{z \in \mathbb{Z}} (-2 + 4z, -1 + 4z) \\
\delta, & \text{else} \end{cases} \]
for some $\delta \in \mathbb{R}$; see figure 3 for a visualization. Using the chain rule given in lemma 2.4 leads to

$$G_2(x) = \begin{cases} 
0, & x \in \bigcup_{z \in \mathbb{Z}} (-1 + 4z, 1 + 4z) \\
1, & x \in \bigcup_{z \in \mathbb{Z}} (-3 + 4z, -2 + 4z) \\
-1, & x \in \bigcup_{z \in \mathbb{Z}} (-2 + 4z, -1 + 4z) \\
\delta, & x = 2z + 1 \text{ for } z \in \mathbb{Z} \\
\sigma, & \text{else}
\end{cases}$$

for some $\delta, \sigma \in \mathbb{R}$. Both $G_1$ and $G_2$ are Newton derivatives of $F \circ H$; see figure 4 for a visualization.

**Remark 2.8.** Example 2.7 shows that changing any separate point of a Newton derivative $D_N F$ of a function $F$ does generally not change the fact that $D_N F$ is a Newton derivative of $F$. More precisely, a Newton derivative is only unique up to a set of measure zero and, thus, every change of a measure zero set does not change the Function $D_N F$ in a Newton derivative sense. Since a Newton derivative is unique up to a set of measure zero, integral equations containing the Newton derivative hold for every point.

Thanks to the Newton derivative definition we are able to formulate the Newton gradient.

**Definition 2.9 (Newton gradient).** Let $(X, \langle \cdot, \cdot \rangle_X)$ be a Hilbert space. Moreover, let $F : X \to \mathbb{R}$ be Newton differentiable in $x \in X$. The Newton gradient of $F$ at $x \in X$ is denoted by $\nabla_N F(x)$ and defined by

$$\langle \nabla_N F(x), v \rangle_X = D_N F(x)v \quad \forall v \in X.$$ 

Thanks to the chain rule in lemma 2.4 and the product rule in lemma 2.5 we obtain a chain rule and a product rule for the Newton gradient.

**Lemma 2.10.** Let $X$ be a Banach space. Moreover, let $F, G : X \to \mathbb{R}$ be Newton differentiable functions at $x \in X$ with Newton derivative $D_N F(x), D_N G(x)$, respectively. Moreover, let $H : \mathbb{R} \to \mathbb{R}$ be Newton differentiable at $y := F(x) \in \mathbb{R}$ with Newton derivative $D_N H(y)$. 

![Figure 3: Graph of $F \circ H$ (left) and $D_N(F \circ H)$ in $[-2, 2]$ without using the chain rule. The violet lines represents a choice of $\delta_i \in \mathbb{R}$ for $i \in \mathbb{N}$.](image)

![Figure 4: Graph of $D_N(F \circ H)$ in $[-2, 2]$ using the chain rule. The violet lines represents a choice of $\delta_i \in \mathbb{R}$ for $i \in \mathbb{N}$ and the blue line represent an additional choice of $\sigma \in \mathbb{R}$ for $i \in \mathbb{N}$.](image)
(i) Let $D_N F(x)$, $D_N H(y)$ be uniformly bounded in a neighborhood of $x$, $y$, respectively. Then $\nabla_N (H \circ F)(x) = D_N H(F(x)) \nabla_N F(x)$ holds.

(ii) If $F$ as well as $D_N G$ are uniformly bounded in a neighborhood of $x$ and if $\|F(x) - F(x+h)\|_X \rightarrow 0$ for $\|h\|_X \rightarrow 0$, then $\nabla_N (F G)(x) = F(x) \nabla_N G(x) + \nabla_N F(x) G(x)$ holds.

Next, we consider an optimization problem in the following form:

$$\min_{x \in X} J(x), \quad (2.3)$$

where $X$ is a Banach space and $J : X \rightarrow \mathbb{R}$ is assumed to be Newton differentiable. Following standard reasoning, now we can formulate necessary conditions of optimality for the optimization problem (2.3).

**Theorem 2.11.** We assume that the function $J : X \rightarrow \mathbb{R}$, where $X$ is a Banach space, is Newton differentiable and that $\hat{x} \in X$ is a local minimum of $J$. Then, there holds

$$\lim_{t \rightarrow 0} D_N J(\hat{x} + tv)v \geq 0 \quad \forall v \in X.$$

**Proof.** Since $\hat{x}$ is minimum, there holds $J(z) \geq J(\hat{x})$, for all $z \in X$. We choose in particular $z := \hat{x} + tv$ for an arbitrary $v \in X$ and $t > 0$. From this, we conclude

$$\frac{1}{t} (J(\hat{x} + tv) - J(\hat{x})) \geq 0,$$

and thus by definition of the Newton derivative

$$\lim_{t \rightarrow 0} D_N J(\hat{x} + tv)v \geq 0.$$

It is no surprise that the necessary condition of optimality does not come in the form of a point-wise condition on the Newton derivative, since the Newton derivative is not unique at a specific point.

### 2.2 Newton shape derivative calculus

In this subsection, we concentrate on shape optimization and extend the Newton derivative concepts from subsection [2.1] to shape calculus. We set up notation and terminology of novel Newton shape optimization concepts. For a detailed introduction into classical shape calculus, the reader is referred to the monographs [11, 47].

We will introduce the concept of Newton shape derivatives. In order to define these derivatives, we define shapes $u$ as (Lebesgue-)measurable subdomains of a hold-all domain $D \subset \mathbb{R}^n$ in the following. In addition, we consider a family $\{F_t\}_{t \in [0,T]}$ of mappings $F_t : \overline{D} \rightarrow \mathbb{R}^n$ such that $F_0 = \text{id}$, where $\overline{D}$ denotes the closure of $D$ and $T > 0$. This family transforms shapes $u$ into new perturbed shapes

$$F_t(u) = \{F_t(x) : x \in u\}.$$

Such a transformation can be described by the velocity method or by the perturbation of identity; cf. [47] pages 45 and 49. In the following, the perturbation of identity is considered. It is defined by $F_t = F_t^V := \text{id} + tV$, where $V : \overline{D} \rightarrow \mathbb{R}^n$ denotes a sufficiently smooth vector field.
Definition 2.12 (Newton shape derivative). Let $k \in \mathbb{N} \cup \{\infty\}$. We call a shape functional $J: u \to \mathbb{R}$ Newton shape differentiable of class $C^k$ at $u$ if there exists a family of mappings $D_N J$ such that
\[
\lim_{t \to 0} \frac{|J(u_t) - J(u) - D_N J(u)[t V]|}{|t|} = 0,
\]
where $D_N J(u): C^k_0(D_\infty, \mathbb{R}^n) \to \mathbb{R}$ is linear and continuous and $u_t := F_t(u)$. In this case, the expression $D_N J(u)[V]$ is called the Newton shape derivative of $J$ at $u$ in direction $V \in C^k_0(D_\infty, \mathbb{R}^n)$.

Our next step is to generalize some well-known concepts from classical shape calculus. We start by generalizing the material and shape derivative of a generic function to the Newton derivative setting.

Definition 2.13 (Material Newton derivative). Let $\{p_t: u \to \mathbb{R}: t \leq T\}$ denote a family of mappings and $p(\equiv p_0): u \to \mathbb{R}$. In addition, let $\Phi_p(t, x) := (p_t \circ F_t)(x)$. The material Newton derivative at $x \in u$ is a family of mappings of Newton derivatives $D_N \Phi_p(0, x)$ such that
\[
\lim_{t \to 0} \frac{|(p_t \circ F_t)(x) - p(x) - D_N \Phi_p(t, x)[t]|}{|t|} = 0
\]
holds. We denote the material Newton derivative of $p$ by $D_{Nm} p$.

Lemma 2.14 (Shape Newton derivative). The shape Newton derivative of a function $p: u \to \mathbb{R}$ in the direction of a sufficiently smooth vector field $V$ is denoted by $D_{Ns} p$ and given by
\[
D_{Ns} p = D_{Nm} p - V^\top \nabla_{NP}.
\]

Proof. Let $t \in [0, T]$ and $x \in u$. We define
\[
\Phi_p(t, x) := (p_t \circ F_t)(x),
R(t, x) := p_t(x),
A(t, x) := (t, x + tV(x))^\top.
\]
In the following, we denote the Newton derivative of the i-th component of a function by $D_{Ni}$. Then, for all $x \in u$ we get
\[
D_{Nm} p(x) = D_N \Phi_p(0, x) = D_N (R \circ A)(0, x) = D_N R(A(0, x)) D_N A(0, x)
= D_N R(A(0, x)) \begin{pmatrix} 1 \\ V(x) \end{pmatrix} = \begin{pmatrix} D_{Ni} R(A(0, x)) \\ D_{Nz} R(A(0, x)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ V(x) \end{pmatrix}
= \begin{pmatrix} D_{Ni} R(A(0, x)) \\ \nabla_{NP}(x) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ V(x) \end{pmatrix} = D_{Ns} p(x) + V^\top(x) \nabla_{NP}(x),
\]
for $D_{Ni} R(A(0, x)) := D_{Ns} p(x)$. 

Next, we formulate material Newton derivative rules, which are generalizations of the results in [4] regarding the material derivative.

Lemma 2.15. Let $p, q: u \to \mathbb{R}$ denote two functions. derivative. Let the Newton derivative of $\Phi_p(t) := p_t \circ F_t$ be continuous in a neighborhood of $0 \in \mathbb{R}$. Moreover, let the Newton derivative of $\Phi_p(t)$ and the Newton derivative of $\Phi_q(t) := q_0 \circ F_t$ be uniformly bounded in a neighborhood of $0 \in \mathbb{R}$.
(i) For the material Newton derivative the product rule holds, i.e.,
\[ D_{Nm}(pq) = D_{Nm}p \, q + p \, D_{Nm}q. \]

(ii) Let \( V_j \) denote the \( j \)th component of \( V \). The Newton material derivative does not change with the gradient but the following equality holds:
\[ D_{Nm}(\partial_i p) = \partial_i D_{Nm}(p) - \sum_{j=1}^{n} \partial_j p \, \partial_i V_j \quad \text{for } i = 1, \ldots, n, \]
which is equivalent to
\[ D_{Nm} \nabla_N p = \nabla_N D_{Nm}p = \nabla_N V^\top \nabla_N p. \]

(iii) The following equality holds:
\[ D_{Nm}(\nabla_N q^\top \nabla_N p) = \nabla_N p^\top \nabla_N D_{Nm}q - \nabla_N q^\top (\nabla_N V + \nabla_N V^\top) \nabla_N q + \nabla_N q^\top \nabla_N D_{Nm}p. \]

Proof. (i) Thanks to Definition 2.13, the material Newton derivative of \( p, q \) is given by \( D_N \Phi_p(0), D_N \Phi_q(0) \), respectively. Using the product rule lemma 2.5, we get for all \( x \in \mathbb{U} \)
\[ D_{Nm}(pq)(x) = D_N(\Phi_p \Phi_q)(s, x)|_{s=0} = (\Phi_p(s, x)D_N \Phi_q(s, x) + D_N \Phi_p(s, x)\Phi_q(s, x))|_{s=0} = p(x) D_{Nm}q(x) + D_{Nm}p(x) q(x). \]

(ii) Definition 2.13 gives
\[ 0 = \lim_{t \downarrow 0} \frac{|\partial_i (p_t \circ F_t) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t}. \]
Therefore, we have
\[ 0 = \partial_i \lim_{t \downarrow 0} \frac{|p_t \circ F_t - p - D_N(p_t \circ F_t)t|}{t} \]
for \( i = 1, \ldots, n \). We can use the Moore-Osgood theorem to obtain
\[ 0 = \lim_{t \downarrow 0} \frac{|\partial_i(p_t \circ F_t) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t}. \]
Let \( \delta_{i,j} \) denote the Kronecker delta. The (multi-valued) chain rule leads to
\[ 0 = \lim_{t \downarrow 0} \frac{|\partial_i(p_t \circ F_t) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t} \]
\[ = \lim_{t \downarrow 0} \frac{|\sum_{j=1}^{n} ((\partial_j p_t \circ F_t)\partial_i F_t) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t} \]
\[ = \lim_{t \downarrow 0} \frac{|\sum_{j=1}^{n} ((\partial_j p_t \circ F_t)(\delta_{i,j} + t\partial_i V_j)) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t} \]
\[ = \lim_{t \downarrow 0} \frac{|\sum_{j=1}^{n} ((\partial_j p_t \circ F_t)\delta_{i,j} + (\partial_j p_t \circ F_t)t\partial_i V_j) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t} \]
\[ = \lim_{t \downarrow 0} \frac{|(\partial_i p_t \circ F_t) + \sum_{j=1}^{n} ((\partial_j p_t \circ F_t)t\partial_i V_j) - \partial_i p - \partial_i D_N(p_t \circ F_t)t|}{t} \]
\[ = \lim_{t \downarrow 0} \frac{|\partial_i p_t \circ F_t - \partial_i p - (\partial_i D_N(p_t \circ F_t)t - \sum_{j=1}^{n} ((\partial_j p_t \circ F_t)t\partial_i V_j))t|}{t} \]
and, thus, to
\[ \partial_i D Nm p = \partial_i D N(p) - \sum_{j=1}^{n} (\partial_j p \partial_i V_j) \]
for \( i = 1, \ldots, n \).

(iii) The combination of the product rule in (i) together with equality in (ii) proofs the claim. \( \square \)

As already mentioned above, shape optimization problems are often constrained by equations. Thus, we provide the following lemma.

**Lemma 2.16.** Let assumptions and notation be as in lemma 2.15. Given coefficient functions \( a_{i,j}, d_i, b \in L^\infty(u) \) of a strongly elliptic bilinear form that fulfill the weak maximum principle, the following equality holds:

\[
D Nm \left( \sum_{i,j} a_{i,j} \partial_i p \partial_j q + \sum_i d_i (\partial_i p q + p \partial_i q) + b p q \right) = \\
\sum_{i,j} a_{i,j} \left( \partial_i D Nm p - \sum_{l=1}^{n} \partial_h p \partial_i V_l \right) \partial_j q + \partial_i p \left( \partial_j D Nm q - \sum_{l=1}^{n} (\partial_h q \partial_i V_l) \right) + \\
\sum_i d_i \left( \partial_i D Nm p - \sum_{l=1}^{n} \partial_h p \partial_i V_l \right) q + \partial_i p \left( D Nm p + D Nm q \right) \partial_i q + \\
b(D Nm p q + p D Nm q) + \\
+ b(D Nm p q + p D Nm q)
\]

**Proof.** The combination of the product rule in lemma 2.15 (i) together with equality in lemma 2.15 (ii) proofs the claim. \( \square \)

Next, we formulate a Newton shape derivative formula. For this purpose, we need to consider perturbed objective functions due to definition 2.12.

**Lemma 2.17.** Let the mappings \( F_t \) of transformations be differentiable in the usual sense. We define the domain integral \( J(u) = \int_u p d x \) for a function \( p: u \to \mathbb{R} \). Then, we have

\[ D N J(u)[V] = \int_u D Nm p + \text{div } V p \ d x. \]

**Proof.** A straight forward calculation like in the proof of [49, Theorem 4.11] using common analysis, the product rule 2.5 as well as definition 2.13 gives the result. \( \square \)

### 3 Newton optimality system

In this section, we focus on necessary optimality conditions for non-smooth shape optimization problems. In theorem 2.11, we observe that the (shape) Newton derivative can be used to characterize necessary optimality conditions. We consider constrained shape optimization problems of the following form:

\[
\min_{u \in \mathcal{U}} \ J(u, y) \quad (3.1)
\]

s.t. \( b(c(u, y), p) = 0 \quad \forall p \in \mathcal{H}(u) \) \quad (3.2)

\[ \text{Such a bilinear form is defined in 4.15.} \]
Here, $\mathcal{H}(u)$ is a Hilbert space defined on the shape $u$ containing the state variable $y \in \mathcal{H}(u)$ and $b(\cdot, \cdot)_u$ is a bilinear and in $\mathcal{H}(u)$ a coercive form. Moreover, $\mathcal{U}$ is the set of admissible shapes, i.e., an appropriate shape space. In this section, we follow the linear deformation space framework as investigated in [46], i.e., we consider a space $Y$ as an appropriate vector space of deformations, such that the set $\mathcal{U}$ of admissible shapes $u$ is constructed as $\mathcal{S}_{\text{adm}} = \{ W(u_0^0) : W \in Y \}$, where $u_0^0$ is a reference starting domain, which is assumed to be a subset of the hold-all domain $D$. We assume that the mapping $c$ is semismooth and Newton differentiable, and that the constraint (3.2) defines a unique solution $y(u,f)$ on any shape $u$ under consideration.

Because $y(u)$ is assumed to satisfy the constraint, we may write for arbitrary $p(u) \in \mathcal{H}(u)$

$$J(u,y(u)) = J(u,y(u)) + b(c(u,y(u)),p(u))_u.$$  

In order to derive necessary conditions of optimality, we differentiate the right-hand side with respect to $u$ and simplify the expressions by introducing the notation $L(u,y,p) := J(u,y) + b(c(u,y),p)_u$, where we keep in mind the implicit dependence of $y,p$ on $u$.

Thus, the chain rule yields

$$D_N L(u,y,p)[V] = \partial^1_N L(u,y,p)[V] + \partial^2_N L(u,y,p)D_{Nm}y + \partial^3_N L(u,y,p)D_{Nm}p$$

for all $V \in Y$, where $\partial^i_N$ denote the Newton partial derivative with respect to the $i$-th argument ($i \in \{1, 2, 3\}$).

Since $y$ satisfies the state equation (3.2) in variational form, which is linear in $p$, we observe

$$\partial^3_N L(u,y,p)D_{Nm}p = 0. \tag{3.3}$$

Furthermore, we obtain

$$\partial^2_N L(u,y,p)D_{Nm}y = \partial^3_N J(u,y)D_{Nm}y + b(\partial^3_N c(u,y)D_{Nm}y,p)_u$$

and, thus, we may obtain $p$ from the Newton adjoint equation in variational form:

$$\partial^2_N J(u,y)p + b(\partial^3_N c(u,y)p)_u = 0 \quad \forall \tilde{y} \in \mathcal{H}(u) \tag{3.4}$$

The solvability of the Newton adjoint equation is in question in this rather general set-up. Therefore, we take it for granted now and show solvability, when confronted with a particular model problem as in the next section. Now, if $y$ satisfies the state equation (3.2) and $p$ satisfies the Newton adjoint equation (3.4), then the Newton shape derivative is given by

$$D_N L(u,y,p)[V] = \partial^1_N L(u,y,p)[V].$$

Nevertheless, it is a manually easier way to compute the Newton shape derivative of the full Lagrangian by employing shape and Newton calculus and later on eliminate expressions relating to the state and Newton adjoint equation, as exemplified in the next section. From Theorem 2.11 we conclude now the necessary condition of optimality for an optimal shape $u$ as

$$\lim_{t \searrow 0} D_N L((id + tv)(u,y,p))[V] \geq 0 \quad \forall V \in Y, \tag{3.5}$$

where $y$ satisfies the state equation (3.2) and $p$ satisfies the Newton adjoint equation (3.4).
In many cases, as is demonstrated in the next section, the Newton shape derivative is continuous in \( t \downarrow 0 \), although the constraints of the shape optimization problem are only semismooth. Then, the necessary condition is just the usual homogeneity of the shape derivative. In this case, the (Newton) shape derivative can be used in order to define a descent direction for algorithmical purposes. Nevertheless, finding a descent direction from the Newton derivative is a challenge in general.

In the next section, we study weak formulations of elliptic problems. These are typically formulated in the Sobolev space \( H^1(u) \) of weakly differentiable \( L^2 \)-functions. For the standard elliptic heat-equation-type problem, the solution is mostly in \( H^2(u) \subset H^1(u) \) and, thus, their material derivative again in \( H^1(u) \) due to Lemma 2.14. However, in the context of variational inequalities, the solution is only piecewise \( H^2(u) \), which means that material derivatives cannot be used as test functions like in (3.3). A similar problem arises in discontinuous Galerkin approximations, from where we borrow the notion of a “broken” Sobolev space here, which is analyzed in detail in [8]. This concept is based on a disjoint partitioning \( u_h \) of open subsets \( K \subset u \) with Lipschitz boundaries such that \( \bigcup K \in u_h \). Then one defines

\[
H^1(u_h) := \{ u \in L^2(u) : u|_K \in H^1(K), K \in u_h \}
\]

In [8], this space is used as test space and shown that a resulting weak formulation of the standard elliptic problem exists which inherits stability and, thus, existence of a unique solution. Thus, we mean this more general weak formulation in the following, whenever a test function is used, which is only piecewise \( H^1 \).

### 4 Application of the Newton derivative scheme

We consider a tracking-type shape optimization problem constrained by a variational inequality of the first kind, a so-called obstacle-type problem. Applications are manifold and arise, whenever a shape is to be constructed in a way not to violate constraints for the state solutions of partial differential equation depending on a geometry to be optimized. Just think of a heat equation depending on a shape, where the temperature is not allowed to surpass a certain threshold. This example is basically the model problem already considered in [34] and that we are formulating in the following. In contrast to [34], which formulates an optimization approach based on the convergence of state, adjoint and shape derivative of the regularized problem to limit objects, we do not consider regularized versions of the VI. We apply our new developed Newton derivative scheme to the model problem and are able to formulate the optimality system. We will see that this system is in line with the limit objects of [34].

**Problem class.** Let \( X \subset \mathbb{R}^n \) be a bounded domain equipped with a sufficiently smooth boundary \( \partial X \). This domain is assumed to be partitioned in a subdomain \( X_{\text{out}} \subset X \) and an interior domain \( X_{\text{int}} \subset X \) with boundary \( \Gamma := \partial X_{\text{int}} \) such that \( X_{\text{out}} \cup \Gamma \sqcup X_{\text{int}} = X \), where \( \sqcup \) denotes the disjoint union. The closure of \( X \) is denoted by \( \bar{X} \). We consider \( X \) depending on \( \Gamma \), i.e., \( X = X(\Gamma) \). In the following, the boundary \( \Gamma \) of the interior domain is called the *interface*. In the setting above, the shape \( u \) is represented by the interior domain \( X_{\text{int}} \). In contrast to the outer boundary \( \partial X \), which is assumed to be fixed, the inner boundary is variable. If \( \Gamma(= \partial u) \) changes, then the subdomains \( u, X_{\text{out}} \subset X \) change in a natural manner.

Let \( \nu > 0 \) be an arbitrary constant. For the objective function \( J(y,u) := \)
\( J(y,u) + J_{\text{reg}}(\Gamma) \) with

\[
J(y,u) := \frac{1}{2} \int_X (y - \bar{y})^2 \, dx, \quad (4.1)
\]

\[
J_{\text{reg}}(\Gamma) := \nu \int_{\Gamma} 1 \, ds (4.2)
\]

we consider

\[
\min_{u \in U} J(y,u) \quad (4.3)
\]

constrained by the obstacle type variational inequality

\[
a(y,v - y) \geq \langle f,v - y \rangle \quad \forall v \in K := \{ \theta \in H^1_0(X) : \theta(x) \leq \varphi(x) \text{ in } X \}, \quad (4.4)
\]

where \( y \in K \) is the solution of the VI, \( f \in L^2(X) \) is explicitly dependent on the shape, \( \langle \cdot, \cdot \rangle \) denotes the duality pairing and \( a(\cdot, \cdot) \) is a general strongly elliptic, i.e. coercive, symmetric bilinear form

\[
a: H^1_0(X) \times H^1_0(X) \rightarrow \mathbb{R}
\]

\[
(y,v) \mapsto \int_X \sum_{i,j} a_{i,j} \partial_i y \partial_j v + \sum_i d_i (\partial_i y v + y \partial_i v) + b v \, dx \quad (4.5)
\]

defined by coefficient functions \( a_{i,j}, d_j, b \in L^\infty(X) \), fulfilling the weak maximum principle.

With the tracking-type objective \( J \) the model is fitted to data measurements \( \bar{y} \in H^1(X) \). The second term \( J_{\text{reg}} \) in the objective function \( J \) is a perimeter regularization. In \( 4.3 \), \( \varphi \) denotes an obstacle which needs to be an element of \( L^1_{\text{loc}}(X) \) such that the set of admissible functions \( K \) is non-empty (cf. \cite{47}). If additionally \( \partial X \) is Lipschitz and \( \varphi \in H^1(X) \) with \( \varphi_{|\partial X} \geq 0 \), then there is a unique solution to \( 4.4 \) satisfying \( y \in H^1_0(X) \), given that the assumptions from above hold (cf. \cite{23, 10, 48}). Further, \( 4.4 \) can be equivalently expressed as

\[
a(y,v) + (\lambda,v)_{L^2(X)} = \langle f,v \rangle_{L^2(X)} \quad \forall v \in H^1_0(X) \quad (4.6)
\]

\[
\lambda \geq 0 \quad \text{in } X
\]

\[
y \leq \varphi \quad \text{in } X
\]

\[
\lambda(y - \varphi) = 0 \quad \text{in } X
\]

with \( \langle \cdot, \cdot \rangle_{L^2(X)} \) denoting the \( L^2 \)-scalar product and \( \lambda \in L^2(X) \). It is well-known, e.g., from \cite{10}, that under these assumptions there exists a unique solution \( y \) to the obstacle type variational inequality \( 4.4 \) and an associated Lagrange multiplier \( \lambda \). We assume this situation, which is also found in \cite{24}, giving us \( \lambda \in L^2(X) \). It can be easily verified that this in turn gives the possibility to summarize the conditions \( 4.7 \) equivalently into a single condition of the form

\[
\lambda = \max \left( 0, \lambda + c(y - \varphi) \right) \quad \text{for any } c > 0. \quad (4.8)
\]

In the following, we denote the active set corresponding to \( 4.6 \) and \( 4.7 \) by

\[
A := \{ x \in X : y - \varphi \geq 0 \}.
\]
Newton adjoint equation. Since the perimeter regularization (4.2) is only used due to technical reasons to overcome ill-posedness of inverse problems (cf., e.g., [1]) and does not influence the adjoint system, we omit it for our investigations in the following. Thus, we consider the (reduced) Lagrangian function to the minimization of (4.1) constrained by
\[
a(y, v) + (\max (0, \lambda + c(y - \varphi)), v)_{L^2(\mathcal{X})} = (f, v)_{L^2(\mathcal{X})} \quad \forall v \in H^1_0(\mathcal{X}),
\]
which is given by
\[
\mathcal{L}(u, y, v) = \frac{1}{2} \int_{\mathcal{X}} (y - \bar{y})^2 \, dx - a(y, v) - \int_{\mathcal{X}} f v \, dx - \int_{\mathcal{X}} \max\{0, \lambda + c(y - \varphi)\} v \, dx,
\]
(4.10)
to formulate the Newton adjoint equation to the model problem (4.3)–(4.4) by computing \(\partial^N_2 \mathcal{L}(u, y, v)\).

In order to compute \(\partial^N_2 \mathcal{L}(u, y, v)\), we consider a variation of \(y\). Let \(t > 0\) and \(\bar{y} \in H^1_0(\mathcal{X})\). Then, we get
\[
\partial^N_2 \mathcal{L}(u, y, v) = \frac{\partial^N}{\partial t} \bigg|_{t=0^+} \mathcal{L}(u, y + t\bar{y}, v)
\]
\[
= \int_{\mathcal{X}} (y - \bar{y}) \bar{y} \, dx - a(\bar{y}, v) + \int_{\mathcal{X}} (\max\{0, \lambda + c(y + t\bar{y} - \varphi)\} v) \, dx.
\]
Using the chain rule in lemma 2.3, we obtain
\[
\frac{\partial^N}{\partial t} \bigg|_{t=0^+} (\max\{0, \lambda + c(t\bar{y} - \varphi)\} v) = D_N(\max\{0, \cdot\})(\lambda + c(t\bar{y} - \varphi)) v
\]
\[
= D_N(\max\{0, \cdot\})(\lambda + c(y - \varphi)) c\varphi v.
\]
Combining the Newton derivative of the maximum function given in (2.2) with the equality
\[
\max\{0, \lambda - c(y - \varphi)\} = \begin{cases} 
\lambda + c(y - \varphi) & \text{in } A,
0 & \text{in } \mathcal{X} \setminus A
\end{cases}
\]
(4.11)
gives
\[
\partial^N_2 \mathcal{L}(u, y, v) = \int_{\mathcal{X}} (y - \bar{y}) \bar{y} \, dx - a(\bar{y}, v) + \int_{\mathcal{X}} \mathbb{1}_A cv \bar{y} \, dx,
\]
where \(\mathbb{1}_A\) denotes the indicator function on the active set \(A\). Thus, the Newton adjoint equation is given in its weak form by
\[
\int_{\mathcal{X}} (y - \bar{y}) \bar{y} \, dx - a(\bar{y}, v) = - \int_{\mathcal{X}} \mathbb{1}_A cv \bar{y} \, dx \quad \forall \bar{y} \in H^1_0(\mathcal{X}).
\]
(4.12)

Newton shape derivative. In order to set up the optimality system to the model problem (4.3)–(4.4), we need the Newton shape derivative of the (full) Lagrangian \(\mathcal{L}_{\text{full}}(y, u, v) = \mathcal{L}(y, u, v) + \mathcal{J}_{\text{reg}}(\Gamma)\), where \(\mathcal{L}\) denotes the (reduced) Lagrangian (4.10). The Newton shape derivative of \(\mathcal{L}_{\text{full}}\) is given by the sum of the Newton shape derivative of the (reduced) Lagrangian (4.10) and the shape derivative of \(\mathcal{J}_{\text{reg}}\). Standard calculation techniques yield the shape derivative of \(\mathcal{J}_{\text{reg}}\).
Lemma 4.1. Let $\varphi \in H^2(\mathcal{X})$, $f \in L^2(\mathcal{X})$, $\bar{y} \in H^1(\mathcal{X})$, $v \in H^1_0(\mathcal{X})$ and $\lambda \in L^2(\mathcal{X})$. Then,

$$D_N\mathcal{L}(u, y, v)[V] = \int_{\mathcal{X}} \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \nabla y^T \nabla v - f v \right] \, dx$$

$$- \int_{\mathcal{X}} (y - \bar{y}) \nabla \bar{y}^T V \, dx + \int_{\mathcal{X}} \nabla f^T V v \, dx$$

$$- \int_{\mathcal{X}} \sum_{i,j} a_{i,j} \left( -\partial_j v \sum_l \partial_l y \partial_i V_l - \partial_j y \sum_l \partial_l v \partial_i V_l \right) \, dx \quad (4.13)$$

$$- \int_{\mathcal{X}} \sum_i d_i \left( -v \sum_l \partial_l y \partial_i V_j - y \sum_l \partial_l v \partial_i V_j \right) \, dx$$

$$+ \int_A (\varphi - \bar{y}) \nabla \varphi^T V \, dx.$$ 

Proof. For an easier understanding and notation purpose we define

$$\chi(y, v) := \sum_{i,j} a_{i,j} \partial_i y \partial_j v + \sum_i d_i (\partial_i y v + y \partial_i v) + b y v$$

such that

$$a(y, v) = \int_{\mathcal{X}} \chi(y, v) \, dx.$$ 

Let

$$G(u)[V] := \int_{\mathcal{X}} D_Nm \left( \frac{1}{2} (y - \bar{y})^2 - \chi(y, v) + f v + \max\{0, \lambda + c(y - \varphi)\} v \right)$$

$$+ \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \chi(y, v) - f v \right] \, dx \quad (4.14)$$

We consider a variation $u_t = F_t(u)$ of $u$ in the following. Since $\mathcal{X}$ depends on $u$, we also use the notation $\mathcal{X}_t := F_t(\mathcal{X})$. We get

$$\lim_{t \searrow 0} \left| \frac{\mathcal{L}(u, y, v) - \mathcal{L}(u_t, y, v) - G(u_t)[V]}{t} \right|$$

$$= \lim_{t \searrow 0} \left| \frac{1}{2} \int_{\mathcal{X}_t} (y - \bar{y})^2 - \frac{1}{2} \int_{\mathcal{X}_t} (y_t - \bar{y}_t)^2 \, dx - \int_{\mathcal{X}_t} \chi(y, v) \, dx - \int_{\mathcal{X}_t} \chi(y_t, v_t) \, dx \right.$$ 

$$- \int_{\mathcal{X}} f v \, dx - \int_{\mathcal{X}_t} f_t v_t \, dx$$

$$+ \int_{\mathcal{X}} \text{max}(0, \lambda + c(y - \varphi)) v \, dx - \int_{\mathcal{X}_t} \text{max}(0, \lambda_t + c_t(y_t - \varphi)) v_t \, dx$$

$$- G(u)[V] \right|$$

Using the definition of $G$ together with the definition of the material Newton derivative (cf. definition 2.13) as well as lemma 2.17 we see

$$\lim_{t \searrow 0} \left| \frac{\mathcal{L}(\mathcal{X}) - \mathcal{L}(\mathcal{X}_t) - G(\mathcal{X}_t)[V]}{t} \right| = 0.$$ 

Therefore, the Newton derivative $D_N\mathcal{L}$ is given by $G$. Combining lemma 2.15 and lemma 2.14 with the state equation (4.9) and adjoint equation (2.2) leads to the expression on the right-hand side of (4.13) for $G(u)[V]$. We refer to the appendix A for the calculation and details.
Remark 4.2. It is worth to mention that the Newton adjoint equation (4.12) and the Newton shape derivative given in lemma 4.1 are the limit object in [34, Theorem 3.3] and [34, Theorem 3.5], respectively. Consequently, if we consider the special case $a(y,v) := \int_X \nabla y^\top \nabla v \, dx$ as in [34, section 4], the Lagrangian is given by $L(y,u,v) = \int_X \frac{1}{2} (y - \bar{y})^2 - \nabla y^\top \nabla v - fv + \max \{0, \lambda + c(y - \varphi)\} v \, dx$. Then, lemma 4.1 yields the Newton shape derivative

$$D_N L(u,y,v)[V] = \int_X - (y - \bar{y})\nabla \bar{y}^\top V - \nabla y^\top (\nabla V^\top + \nabla V) \nabla p$$

$$+ \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \nabla y^\top \nabla v - fv \right] \, dx + \int_A (\varphi - \bar{y}) \nabla \varphi^\top V \, dx,$$

which confirms the limit object given in [34, equality (43)].

Newton optimality system. Here, we summarize the optimality conditions. For a solution shape $u$ to problem (4.3)–(4.4), there holds the Newton adjoint variational equation (4.12). Since the Newton shape derivative $D_N L(u,y,v)[V]$ given in lemma 4.1 is continuous in $V$, we obtain from (3.5) the following necessary condition for the optimal shape $u$:

$$0 = D_N L(u,y,v)[V] \quad \forall V \in H^1(u, \mathbb{R}^n)$$

The Newton adjoint equation, this necessary condition and the state equation (4.4) define together a set of equations, which is used for the computation of the solution in [34], where a perturbation approach is used for construction of $D_N L(u,y,v)[V]$. We observe also that $D_N L(u,y,v)[V]$ is an integral on $\mathcal{X}(u)$, where the integrand is Newton differentiable with respect to $u$ and which lacks standard differentiability only at the the boundary of the active set $A$, which is a set of Lebesgue measure zero. Thus, $D_N L(u,y,v)[V]$ is a shape derivative and can therefore be used, in order to define a descent direction by employing an appropriate scalar product.

In [34], the same expression has been derived in a perturbation approach, which necessitates a safeguard technique. The fact observed here that the resulting step is a descent direction justifies also theoretically, why the safeguard technique has never been activated in the numerical computations in [34].

5 Conclusions

The Newton derivative for semismooth functions is a cornerstone for the semismooth Newton method which is used for the numerical solution of necessary conditions for optimization problems for variational inequalities. In this paper, the concept of Newton derivative is generalized to a Newton shape derivative calculus. This enables the definition of a Newton shape derivative, a Newton material derivative as well as a Newton gradient. In addition to the chain rule given in [9], we provide a product rule. Further, rules for the material derivative calculus given in [4] are generalized to the Newton material derivative setting. This paper provides also an equation to calculate the Newton material derivative of two functions twisted by a general strongly elliptic bilinear form. One of the major advantages of the Newton scheme developed in this paper is that one no longer need to regularize variational inequality contained optimization problems and a limit process as in [34] can be avoided. Beyond that, this paper explains the limiting expression for the shape derivative given in [34] now as an expression derived from a Newton adjoint and, thus, safeguarding techniques as in [34] are no longer necessary. These observations
open the door for further potential usage of Newton derivatives for more complicated variations inequalities than studied in section 4 and of Newton shape derivatives of even higher order.

Acknowledgment

This work has been partly supported by the German Research Foundation (DFG) within the priority program SPP1962/2 under contract numbers WE 6629/1-1 and SCHU 804/19-1.

A Derivation Newton shape derivative expression

We consider the problem class stated in section 4. Let $\varphi \in H^2(\mathcal{X})$, $f \in L^2(\mathcal{X})$, $\bar{y} \in H^1(\mathcal{X})$, $v \in H^1_0(\mathcal{X})$ and $\lambda \in L^2(\mathcal{X})$. In this appendix, we show that $G$ defined in (4.14) is given by the right-hand side of (4.13).

Let $\chi(y, v) := \sum_{i,j} a_{i,j} \partial_i y \partial_j v + \sum_i d_i (\partial_i y v + y \partial_i v) + byv$ such that $a(y, v) = \int_{\mathcal{X}} \chi(y, v) \, dx$.

Then, lemma 2.15 and lemma 2.16 lead us to

$$G(u)[V] = \int_{\mathcal{X}} \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \chi(y, v) - f v \right]$$

$$+ D_{Nm} \left( \frac{1}{2} (y - \bar{y})^2 \right) - D_{Nm}(\chi(y, v)) + D_{Nm}(f v)$$

$$+ D_{Nm}(\max\{0, \lambda + c(y - \varphi)\}) \, dx$$

$$= \int_{\mathcal{X}} \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \chi(y, v) - f v \right]$$

$$+ (y - \bar{y}) D_{Nm} y - (y - \bar{y}) D_{Nm} \bar{y} + v D_{Nm} f + f D_{Nm} v$$

$$- D_{Nm} \left( \sum_{i,j} a_{i,j} \partial_i y \partial_j v + \sum_i d_i (\partial_i y v + y \partial_i v) + byv \right)$$

$$+ D_{Nm}(\max\{0, \lambda + c(y - \varphi)\}) v + \max\{0, \lambda + c(y - \varphi)\} D_{Nm} v \, dx$$

$$= \int_{\mathcal{X}} \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \chi(y, v) - f v \right]$$

$$+ (y - \bar{y}) D_{Nm} y - (y - \bar{y}) D_{Nm} \bar{y} + v D_{Nm} f + f D_{Nm} v$$

$$- \chi(D_{Nm} y, v) + \chi(y, D_{Nm} v)$$

$$- \sum_{i,j} a_{i,j} \left( -\partial_i v \sum_l \partial_l y \partial_l V_l - \partial_i y \sum_l \partial_l v \partial_l V_l \right)$$

$$- \sum_i d_i (\partial_i y D_{Nm} v + D_{Nm} y \partial_i v)$$

$$+ D_{Nm}(\max\{0, \lambda + c(y - \varphi)\}) v + \max\{0, \lambda + c(y - \varphi)\} D_{Nm} v \, dx$$

Combining the Newton derivative of the maximum function given in 2.2 with the
equality (4.11) gives
\[
\int_X D_{Nm} \left( \max \{0, \lambda + c(y - \varphi)\} \right) v \, dx = \int_X \mathbb{1}_A (D_{Nm} \lambda + c(D_{Nm} y - D_{Nm} \varphi)) v \, dx.
\]
In addition, lemma 2.14 gives
\[
D_{Nm} \bar{y} = \nabla \bar{y}^\top V \quad \text{and} \quad D_{Nm} f = \nabla f^\top V
\]
if we assume \( \bar{y} \) and \( f \) independent of the shape. Thus, thanks to the state equation (4.9) and the Newton equation (4.12) we get
\[
G(u)[V] = \int_X \text{div}(V) \left[ \frac{1}{2} (y - \bar{y})^2 + \nabla y^\top \nabla v - f v \right] - (y - \bar{y}) \nabla \bar{y}^\top V + v \nabla f^\top V
\]
\[
- \sum_{i,j} a_{i,j} \left( -\partial_j v \sum_l \partial_i y \partial_l V_l - \partial_i y \sum_l \partial_j v \partial_l V_l \right)
\]
\[
- \sum_l d_l \left( -v \sum_l \partial_i y \partial_l V_j - y \sum_l \partial_i v \partial_l V_j \right) \, dx
\]
\[
+ \int_A (y - \bar{y}) D_{Nm} y \, dx.
\]
In the active set \( A \), we have \( y = \varphi \). Moreover, \( D_m \varphi = \nabla \varphi^\top V \) thanks to lemma 2.14. Thus, the integral over the active set is given by
\[
\int_A (y - \bar{y}) D_{Nm} y \, dx = \int_A (\varphi - \bar{y}) \nabla \varphi^\top V \, dx.
\]

References

[1] H.B. Ameur, M. Burger, and B. Hackl. Level Set Methods for Geometric Inverse Problems in Linear Elasticity. *Inverse Problems*, 20(3):673–696, 2004.

[2] V. Barbu. Optimal control of variational inequalities. *Research Notes in Mathematics*, 100, 1984.

[3] M. Bauer, P. Harms, and P.M. Michor. Sobolev metrics on shape space of surfaces. *Journal of Geometric Mechanics*, 3(4):389–438, 2011.

[4] M. Berggren. A unified discrete-continuous sensitivity analysis method for shape optimization. In W. Fitzgibbon et al., editors, *Applied and Numerical Partial Differential Equations*, volume 15 of *Computational Methods in Applied Sciences*, pages 25–39. Springer, 2010.

[5] J. Bracken and J. McGill. Mathematical programs with optimization problems in the constraints. *Operations Research*, 21(1):37–44, 1973.

[6] H. Brézis. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. *Contributions to Nonlinear Functional Analysis*, pages 101–156, 1971.

[7] H. Brézis and G. Stampacchia. Sur la regularité de la solution d’inéquations elliptiques. *Bulletin de la Société Mathématique de France*, 96:153–180, 1968.
[8] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan. Breaking spaces and forms for the dpg method and applications including maxwell equations. *Computers & Mathematics with Applications*, 72(3):494–522, 2016.

[9] C. Clason and T. Valkonen. Introduction to nonsmooth analysis and optimization. *arXiv:2001.00216*, 2020.

[10] G. Stampacchia D. Kinderlehrer. *An Introduction to Variational Inequalities and Their Applications*, volume 31. SIAM, 1980.

[11] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, volume 22 of *Adv. Des. Control*. SIAM, 2nd edition, 2001.

[12] Z. Denkowski and S. Migorski. Optimal shape design for hemivariational inequalities. *Universitatis Iagellonicae Acta Mathematica*, 36:81–88, 1998.

[13] M. Droske and M. Rumpf. Multi scale joint segmentation and registration of image morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2181–2194, 2007.

[14] L. Gasinski. Mapping method in optimal shape design problems governed by hemivariational inequalities. In J. Cagnol, M. Polis, and J.-P. Zolésio, editors, *Shape optimization and optimal design*, number 216 in Lecture Notes in Pure and Applied Mathematics, pages 277–288. Marcel Dekker, 2001.

[15] R. Glowinski. *Numerical methods for nonlinear variational problems*. Springer, 1984.

[16] P. Harker and J.-S. Pang. Existence of optimal solutions to mathematical programs with equilibrium constraints. *Operations Research Letters*, 7(2):61–64, 1988.

[17] C. Heinemann and K. Sturm. Shape optimization for a class of semilinear variational inequalities with applications to damage models. *SIAM Journal on Mathematical Analysis*, 48(5):3579–3617, 2016.

[18] R. Herzog and E. Loayza-Romero. A manifold of planar triangular meshes with complete Riemannian metric.

[19] R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM Journal on Control and Optimization*, 50(5):3052–3082, 2012.

[20] R. Herzog, C. Meyer, and G. Wachsmuth. B- and strong stationarity for optimal control of static plasticity with hardening. *SIAM Journal on Control and Optimization*, 23(1):312–352, 2013.

[21] M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20:868–902, 2009.

[22] M. Hintermüller and A. Laurain. Optimal shape design subject to elliptic variational inequalities. *SIAM Journal on Control and Optimization*, 49(3):1015–1047, 2011.

[23] K. Ito and K. Kunisch. Optimal Control of Elliptic Variational Inequalities. *Applied Mathematics and Optimization*, 41(3):343–364, 2000.
[24] K. Ito and K. Kunisch. Semi-Smooth Newton Methods for Variational Inequalities of the First Kind. *ESIAM: Mathematical Modelling and Numerical Analysis*, 37:41–62, 2003.

[25] D.G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.

[26] N. Kikuchi and J. Oden. *Contact problems in elasticity: A study of variational inequalities and finite element methods*. SIAM, 1988.

[27] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*. Academic Press, 1980.

[28] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods, and Applications*. Kluwer Academic, 2002.

[29] M. Kocvara and J. Outrata. Shape optimization of elasto-plastic bodies governed by variational inequalities. In J.-P. Zolésio, editor, *Boundary Control and Variation*, number 163 in Lecture Notes in Pure and Applied Mathematics, pages 261–271. Marcel Dekker, 1994.

[30] H. Ling and D.W. Jacobs. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):286–299, 2007.

[31] J. Lions and G. Stampacchia. Variational inequalities. *Communications on Pure and Applied Mathematics*, 20:493–519, 1967.

[32] W. Liu and J. Rubio. Optimal shape design for systems governed by variational inequalities, part 1: Existence theory for the elliptic case. *Journal of Optimization Theory and Applications*, 69(2):351–371, 1991.

[33] W. Liu and J. Rubio. Optimal shape design for systems governed by variational inequalities, part 2: Existence theory for the evolution case. *Journal of Optimization Theory and Applications*, 69(2):373–396, 1991.

[34] D. Luft, V.H. Schulz, and K. Welker. Efficient techniques for shape optimization with variational inequalities using adjoints. *SIAM Journal on Optimization*, 30(3):1922–1953, 2020.

[35] Z. Luo, J. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

[36] P.M. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Mathematica*, 10:217–245, 2005.

[37] P.M. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.

[38] F. Mignot and J. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.

[39] A. Myśliński. Domain optimization for unilateral problems by an embedding domain method. In J. Cagnol, M. Polis, and J.-P. Zolésio, editors, *Shape optimization and optimal design*, number 216 in Lecture Notes in Pure and Applied Mathematics, pages 355–370. Marcel Dekker, 2001.
[40] A. Myśliński. Level set approach for shape optimization of contact problems. In P. Neittaanmäki, T. Rossi, K. Majava, O. Pironneau, and I. Lasiecka, editors, European Congress on Computational Methods in Applied Sciences and Engineering ECCOMAS, 2004.

[41] A. Myśliński. Level set method for shape and topology optimization of contact problems. In IFIP Conference on System Modeling and Optimization, pages 397–410. Springer, 2007.

[42] P. Neittaanmäki, J. Sprekels, and D. Tiba. Optimization of Elliptic Systems: Theory and Applications. Springer Monographs in Mathematics, 2006.

[43] J. Outrata, M. Kočvara, and J. Zowe. Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications, and Numerical Results. Number 152 in Nonconvex optimization and its applications. Kluwer Academic Publisher, 1998.

[44] P. Panagiotopoulos. Inequality problems in mechanics and applications. Birkhäuser, 1985.

[45] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. Mathematics of Operations Research, 25(1):1–22, 2000.

[46] S. Schmidt and V.H. Schulz. A linear view on shape optimization. Technical report, 2022.

[47] J. Sokolowski and J.-P. Zolésio. Introduction to Shape Optimization, volume 16 of Computational Mathematics. Springer, 1992.

[48] G.M. Troianiello. Elliptic Differential Equations and Obstacle Problems. Springer Science & Business Media, 2013.

[49] K. Welker. Efficient PDE Constrained Shape Optimization in Shape Spaces. PhD thesis, Universität Trier, 2016.

[50] B. Wirth, L. Bar, M. Rumpf, and G. Sapiro. A continuum mechanical approach to geodesics in shape space. International Journal of Computer Vision, 93(3):293–318, 2011.

[51] B. Wirth and M. Rumpf. A nonlinear elastic shape averaging approach. SIAM Journal on Imaging Sciences, 2(3):800–833, 2009.

[52] J.-P. Zolésio. Control of moving domains, shape stabilization and variational tube formulations. International Series of Numerical Mathematics, 155:329–382, 2007.