Diploma thesis

3D Electrical Impedance Tomography with sparsity constraints
Algorithm and implementation in application to the complete electrode model

Tobias Kluth

1. Gutachter
Dr. Kamil Kazimierski

2. Gutachter
Prof. Dr. Peter Maaß

05. August 2011
Acknowledgements

First I would like to thank Prof. Dr. Peter Maaß for his motivation and for giving me the opportunity to learn as much as possible in my scientific work. I am heartily thankful to Bangti Jin, Ph.D., whose teaching, guidance and many comments enabled me to develop an understanding of the subject. This thesis would not have been possible without him. I am grateful to Dr. Kamil Kazimierski for supervision, helpful advice and support. I would like to thank Prof. Jari Kaipio, Ph.D., for his kind invitation to Finland and Aku Seppänen, Ph.D., and Antti Lipponen for their hospitality and experimental assistance.

I am indebted to my student colleagues for their support and many discussions. I wish to thank my sister, Sarah Kluth, and my fellow student, Nicolas Schilling, for proofreading and comments. In this context I would like to thank Valerie Scholes for proofreading.

At last and most importantly I wish to thank my partner Karina Heidt who supports me in everything, cares for me and has complete understanding.

Bremen, 05. August 2011

Tobias Kluth
# Contents

1. **Introduction** 5

2. **Basic principles and notations** 7

3. **Model** 10
   - 3.1. Physics ................................................................. 10
   - 3.1.1. Interior .............................................................. 13
   - 3.1.2. Boundary ............................................................ 13
   - 3.2. Complete Electrode Model ........................................... 15

4. **The inverse problem** 29
   - 4.1. Tikhonov functional ................................................ 31
   - 4.2. Sparse regularization functional .................................... 32

5. **Algorithm** 39
   - 5.1. Step 1: Gradient $D'(\sigma)$ ....................................... 41
   - 5.2. Step 2: Sobolev smoothed gradient ................................. 44
   - 5.3. Step 3: Choice of the step size ...................................... 46
   - 5.4. Step 4: Applying shrinkage operator $S_\alpha$ ...................... 48
   - 5.5. Step 5: Stopping criterion ........................................... 49

6. **Implementation** 50
   - 6.1. Galerkin method ..................................................... 50
   - 6.2. Finite element method .............................................. 50
   - 6.3. Forward operator ..................................................... 55
   - 6.3.1. Stiffness matrix $A$ ............................................... 57
   - 6.3.2. Matrix $B$ ............................................................ 59
   - 6.3.3. Matrix $C$ ............................................................ 61
   - 6.3.4. Matrix $D$ ............................................................ 61
   - 6.4. Implementation Step 1: Gradient $\tilde{D}'(\sigma)$ ................ 62
   - 6.5. Implementation Step 2: Sobolev smoothed gradient $\tilde{D}'_s(\sigma)$ .... 64
   - 6.5.1. Matrix $B_s$ .......................................................... 65
   - 6.6. Implementation Step 3: Choice of the step size ................... 67
   - 6.7. Implementation Step 4: Shrinkage operator $\tilde{S}_\alpha$ .......... 69
   - 6.8. Implementation Step 5: Stopping criterion ....................... 70
7. Application 71
   7.1. Simulated data ............................................ 71
   7.2. Real data ..................................................... 77

8. Summary and outlook 81

A. Proofs 83
   A.1. Model ............................................................. 83
       A.1.1. Uniform continuity ........................................ 83
       A.1.2. Differentiability in $L^p$ sense ............................ 84
1. Introduction

Electrical impedance tomography (EIT) is a nondestructive imaging method in which the electrical conductivity of an object is determined by using information on the boundary. It is a noninvasive technique because it uses only measurements at the boundary of the object and in practice the current is applied to the electrodes which are attached to the surface and the resulting potential differences are measured. EIT covers a big field of applications, for example geophysical exploration [8], medical diagnostics [5] and the nondestructive analysis of concrete [21]. Many approaches for solving this nonlinear inverse problem and obtaining numerically computed reconstructions of the conductivity have been discussed, e.g. a statistical approach [20]. The electrodes, in particular, can be modeled in several ways [32]. In this thesis we focus on the complete electrode model (CEM) and an iterated soft shrinkage approach which is used to minimize a Tikhonov functional with sparsity constraints. This approach and the application to Tikhonov regularization have also been considered more generally in [19, 18, 7, 4]. The comparison between sparse and smooth reconstructions in the two-dimensional case using the same approach as presented in this thesis can also be found in [12]. In this thesis we consider only the nonlinear model and the sparse penalty term such that we can do it in more detail and finally we apply the algorithm to a three-dimensional setting and can compare it with the results from [12, 13].

This thesis is divided into a theoretical and a numerical part. At first we consider the theory of EIT and the sparse Tikhonov regularization and then with this background knowledge the numerical part, which includes the iterated soft shrinkage algorithm and the Galerkin approach. This means in detail that Section 2 summarizes the necessary basic principles. In Section 3 we then discuss the physical derivation of the model and present the important abilities of the model which includes the CEM and the forward operator which maps the current to the potentials. The main properties like the weak formulation of the problem, the differentiability of the forward operator and the continuity of the derivative are discussed and partially proven. As mentioned above we minimize a Tikhonov functional with a sparsity promoting penalty term. This procedure and some important properties of the discrepancy and penalty term are discussed in Section 4. The following Section 5 contains the general considerations of the algorithm in detail. The different steps
like the gradient calculation, the Sobolev smoothed gradient, the choice of step size and the shrinkage operator are discussed. We then discuss the Galerkin approach in Section 6. With this approach we can use the finite element method to calculate the solution of the forward operator. We also apply the other steps to the finite element space. In Section 7 we first discuss the reconstructions from simulated data and subsequently from real data. And lastly we give the summary and outlook in Section 8.
2. Basic principles and notations

For the subsequent topics of this thesis we present the main basic principles and notations which are used to describe the model, the inverse problem and the algorithm. The proofs can be found in general literature, e.g [1].

**Definition 2.0.1** (Lipschitz domain). Let $\Omega \subset \mathbb{R}^n$ be open and bounded. We call $\Omega$ a *Lipschitz domain* if the following holds:

For every point $x \in \partial\Omega$ there exists a radius $r > 0$ and a mapping $h_x : B_r(x) \to B_1(0)$ such that $h_x$ is a bijection and $h_x$, $h_x^{-1}$ are Lipschitz continuous functions with

\[
    h_x(\partial\Omega \cap B_r(x)) = \{(x_1, \ldots, x_n) \in B_1(0)|x_n = 0\} \quad \text{and} \quad h_x(\Omega \cap B_r(x)) = \{(x_1, \ldots, x_n) \in B_1(0)|x_n > 0\}.
\]

**Example.** Lipschitz domains are, for example, spheres and polygonal bounded domains where the domain is defined on only one side of the boundary. A non Lipschitz domain is, for example, a circle with a slot, like

\[
    \Omega := B_1(0) \setminus \{(x, y)| x > 0, y = 0\}.
\]

**Remark.** In this thesis every $\Omega \subset \mathbb{R}^n$ is a connected Lipschitz domain, especially a cylindrical or a polygonal bounded domain.

**Definition 2.0.2** (Sobolev space $W^{k,p}$). Let $\Omega \subset \mathbb{R}^n$ be open and $1 \leq p \leq \infty$ then the Sobolev space $W^{k,p}$ is defined by

\[
    W^{k,p}(\Omega) := \{u \in L^p(\Omega)|\forall \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n, |\alpha| \leq k : \partial^{\alpha}u \in L^p(\Omega)\}.
\]

With the norm

\[
    \|u\|_{W^{k,p}(\Omega)} := \begin{cases} 
        \left( \sum_{|\alpha| \leq k} \|\partial^{\alpha}u\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}}, & p < \infty \\
        \max_{|\alpha| \leq k} \|\partial^{\alpha}u\|_{L^\infty(\Omega)}, & p = \infty
    \end{cases}
\]

$W^{k,p}(\Omega)$ is a Banach space. We also define

\[
    W^{k,p}_0(\Omega) := \{u \in W^{k,p}(\Omega)|u = 0 \text{ on } \partial\Omega\}.
\]
Especially for $p = 2$ we have $H^k(\Omega) := W^{k,2}(\Omega)$. Equipped with the scalar product

$$\langle u, v \rangle_{H^k(\Omega)} := \sum_{|\alpha| \leq k} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2(\Omega)} = \sum_{|\alpha| \leq k} \int_\Omega \partial^\alpha u \cdot \partial^\alpha v \, dx$$

$H^k(\Omega)$ is a Hilbert space.

**Theorem 2.0.3** (Sobolev embedding theorem). Let $\Omega \subset \mathbb{R}^n$ be open and be bounded by a Lipschitz boundary. Let $m_1, m_2 \in \mathbb{N}$ and $1 \leq p_1, p_2 < \infty$. The following then holds:

1. If

   $$m_1 - \frac{n}{p_1} \geq m_2 - \frac{n}{p_2} \quad \text{and} \quad m_1 \geq m_2,$$

   then the embedding

   $$i : W^{m_1, p_1}(\Omega) \hookrightarrow W^{m_2, p_2}(\Omega)$$

   exists and is continuous.

2. If

   $$m_1 - \frac{n}{p_1} > m_2 - \frac{n}{p_2} \quad \text{and} \quad m_1 > m_2,$$

   then the embedding

   $$i : W^{m_1, p_1}(\Omega) \hookrightarrow W^{m_2, p_2}(\Omega)$$

   exists and is continuous and compact.

3. For arbitrary open, bounded sets $\Omega \subset \mathbb{R}^n$ the two statements above hold for $W^{m_1, p_1}_0(\Omega)$ instead of $W^{m_1, p_1}(\Omega)$ with the corresponding norms.

**Theorem 2.0.4** (Trace theorem). Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain and $1 \leq p \leq \infty$. There then exists exactly one linear and continuous mapping

$$S : H^{1,p}(\Omega) \to L^p(\partial \Omega)$$

such that

$$Su = u_{|\partial \Omega} \quad \text{for} \ u \in H^{1,p}(\Omega) \cap C(\overline{\Omega}).$$
**Theorem 2.0.5** (Kondrashov embedding theorem). Suppose $Ω ⊂ \mathbb{R}^n$ is an open, bounded Lipschitz-domain and let $1 ≤ p < n$. Set

$$p^* := \frac{np}{n - p}.$$

Then the space $W^{1,p}(Ω)$ is continuously embedded in $L^{p^*}(Ω)$. For $1 ≤ q < p^*$ the space $W^{1,p}(Ω)$ is compactly embedded in $L^q(Ω)$.

**Theorem 2.0.6** (Eberlein–Šmulian theorem). For a subset $A$ of the Banach space $X$ the following conditions are equivalent:

- $A$ is weakly compact, i.e. $A$ is compact with respect to the weak topology $\{φ(U) | φ ∈ X^*, U$ open subset in $\mathbb{R}\}$.
- $A$ is weakly sequentially compact, i.e. for every sequence in $A$ exists a weak convergent subsequence with limit in $A$.

**Lemma 2.0.7** (Fundamental lemma of variational calculus). Let $Ω ⊂ \mathbb{R}^n$ be open. Then for $g ∈ L^1(Ω)$ the following is equivalent:

- $\int_Ω gv \, dx = 0$, $∀v ∈ C_0^∞(Ω)$.
- $g = 0$, a.e. in $Ω$.

**Definition 2.0.8** (Generic constant). A generic constant $C$ can take different values every time it is used. We use this notation in some proofs to avoid unnecessary numeration of constants.
3. Model

Several models are described in general theory for electrical impedance tomography. In this section we will describe the so-called complete electrode model (CEM). We first discuss the basic principles and quantities of electrostatics for a better understanding and derive the main equations for the CEM step by step. For further information about other models, like the continuums model, shunt model, etc. we cite the paper by Somersalo, Cheney and Isaacson [32]. They also present a derivation of the model by applying the general Maxwell's equations. The second part of this section deals with the properties of the CEM, like existence and uniqueness of a solution, continuity and differentiability.

3.1. Physics

In this section we introduce the important physical quantities from the electrostatics and formulate the model which is presented later. For further physical informations we refer to [24]. The main assumption in electrostatics is that only static charge exists and therefore no magnetic field results. Additionally we assume that the quantities are not disperse, i.e. they are independent of the frequency. This especially holds by using co-current flow or alternating current with a very small frequency.

Electric charge $Q$

The electric charge $Q$ depends on the fundamental building blocks of matter which consists of electrons, protons and neutrons. The electrons have a negative elementary charge and the protons the same positive charge. The neutrons have no charge. For matter we can assume that the protons are much heavier than the electrodes. The electrons are able to move and the protons and neutrons are fixed. Electrically charged particles are subjected to electromagnetic interactions. This affects the bonding of atoms, molecules and solids.

The absolute charge of a body is the sum of all elementary charges. The body is neutral if the number of electrons equals the number of protons. Charge conservation is the effect that in a bounded system the amount of electric charge is constant. Opposite charges attract each other and like charges repel each other. The unit of electric charge is $[Q] = 1C$ (Coulomb).

Moreover, we introduce the charge density $\rho$ which describes the electric charge per
volume and has the unit \( [\rho] = 1 \text{C/m}^3 \).

This gives us the coherence

\[
Q(t) = \int_V \rho(t, x, y, z) \, dV
\]

where \( V \) is a three dimensional body. In electrostatics, steady state processes are observed in which charge density does not change with respect to the time:

\[
\dot{\rho}(t, x, y, z) = 0.
\]

**Electric field strength \( E \)**

The *electric field strength* \( E \) is a vectorial quantity which assigns every point in a body \( V \) a direction and a magnitude. The electric field strength describes the ability to exert force on charge.

\[
E = \frac{F}{q}
\]

where \( F \) is the force exerted on the electric charge \( q \). Because of the proportionality between force and charge, \( E \) is independent of the charge. The unit is \( [E] = 1 \text{N/C} = 1 \text{V/m} \).

**Electric potential \( \Phi \) / Voltage \( U \)**

The electric field exerts a force on the electric charge. We therefore have to do work to move the charge through the electric field. Assuming that electrostatic fields are irrotational this work is independent of the path. We can therefore define the potential work as the product of the charge and the *potential* \( \Phi \), which depends only on the start and end points.

\[
W_{\text{pot}}(x) = \int_{L_x} F \, dl = q\Phi(x) \quad \Rightarrow \quad \Phi(x) = \int_{L_x} E \, dl
\]

where \( l \) is the parametric parameter and \( L_x \) denotes the line between an arbitrary point \( x \) and a point with the potential energy zero. The difference in the potentials at two points is called the *voltage* \( U \). The unit is \( [\Phi] = [U] = 1 \text{V} \) (Volt).

The relationship between electric field strength and potential is as follows

\[
E = -\nabla \Phi.
\]
Current $I$ / Current density $j$

The current $I$ describes the amount of charge per time which flows through an area $A$ with the current density $j$, which is a vectorial quantity. The current density describes the direction and the current per unit area of cross section. The units are $[I] = 1\, \text{C} = 1\, \text{A}$ (Ampere) and $[j] = 1\, \text{A/m}^2$.

The relationship between current and current density is

$$ I = \int_A j \cdot n \, dS $$

where $n$ is the normal vector to the area $A$.

Conductivity $\sigma$ / Ohm’s law

The conductivity $\sigma$ is a quantity which depends on the material. It describes the conducting ability of the material. The conductivity of a material depends on the number of charge carriers, which can be electrons or ions. The more charge carriers are available, the better is the conductivity. In general the conductivity is a tensor but we assume isotropic materials, where the conductivity is a scalar. The unit is $[\sigma] = \Omega^{-1}\, \text{m}$. The conductivity is related closely with Ohm’s law. In general it describes the relationship between the current and the voltage for a specific conductor.

$$ U = RI $$

where $R$ is called the resistance with the unit $[R] = \Omega$ (Ohm).

For local observations it becomes

$$ j = \sigma E. $$

It describes the connection between the current density and the electric field strength. The better the conductor, the stronger the electric field affects a current density.

Kirchhoff’s current law

The current law by Kirchhoff says that the current flowing out of a closed surface is equal to the loss of charge in the interior of the enclosed volume. From this it follows

$$ \nabla \cdot j = \dot{\rho}. $$
With the assumption of charge conservation, i.e. a stationary process, we have
\[ \nabla \cdot \mathbf{j} = 0. \]

### 3.1.1. Interior

We assume a bounded Lipschitz domain \( \Omega \) as an isotropic, nondisperse conductor. Additionally we assume a stationary process. The main aim is to obtain a condition for the interior of the conductor which depends on the conductivity and the potential. First we take Kirchhoff’s current law with the charge conservation for an arbitrary \( x \in \Omega \)
\[ \nabla \cdot \mathbf{j}(x) = 0. \]
Applying Ohm’s law we get
\[ \nabla \cdot (\sigma(x) \mathbf{E}(x)) = 0. \]
With the relationship between electric field strength and the potential we get the main equation
\[ -\nabla \cdot (\sigma \nabla \Phi) = 0 \quad \text{in } \Omega. \quad (3.1) \]

### 3.1.2. Boundary

On the boundary we have to differentiate between two areas \( A_e \) and \( A_0 \), the first one with electrodes and the second one without. On the area without electrodes we assume that the current \( I \) is zero. We therefore get
\[ \int_{A_0} \mathbf{j} \cdot \mathbf{n} \, dS = 0. \]
Replacing the current density as before we get
\[ -(\sigma \nabla \Phi) \cdot \mathbf{n} = -\sigma \frac{\partial \Phi}{\partial n} = 0 \quad \text{on } A_0. \]

We define a metal electrode \( e \subset A_e \) on the boundary as can be seen in Figure 1. For the electrode we assume a constant conductivity \( \sigma_e \). \( A_f \) denotes the area where the cable connects to the electrode. Additionally, we assume a thin layer on the metal surface where chemical reactions lead to a different resistance and denote the
conductivity in this layer by $\sigma_r$. The area between metal electrode and the thin layer is denoted by $A_r$. We assume that a current $I$ flows through the area $A_I$. With charge conservation and no current density except on $A_I$, $A_r$ and $e$ we therefore have

$$I = \int_{A_I} j \cdot n \, dS = \int_{A_r} j \cdot n \, dS = \int_e j \cdot n \, dS.$$ 

The potential difference $U_I$ between one point $x_1$ on area $A_I$ and $x_2$ from $A_r$ is given by

$$U_I = \Phi(x_2) - \Phi(x_1) = \int_L E \, dl = \frac{1}{\sigma_e} \int_L j \, dl,$$

which depends on the length of the connecting line $L$. We now consider the thin layer. Because of the thinness, we assume the electric field lines to be parallel and orthogonal to the area $A_r$. Since $A_r$ and $e$ have the same size we have constant current density along a line orthogonal to the area. We therefore get

$$U_r = \Phi(x_3) - \Phi(x_2) = \int_{L_r} E \, dl = \frac{1}{\sigma_r} \int_{L_r} j \, dl = j \cdot n \frac{|L_r|}{\sigma_r}$$
if we assume the conductivity is constant in the thin layer and \( n \) is the orthonormal vector to \( e \). Adding the two equations together we get a voltage drop of

\[
U = U_l + U_r = \frac{1}{\sigma_e} \int_L j \ dl + j \cdot n \frac{|L_r|}{\sigma_r}.
\]

With the measured potential \( \Phi_e \) which is assumed to be constant on area \( A_I \) we get the boundary condition

\[
\Phi_e = \Phi(x) + U(x) = \Phi(x) + \frac{1}{\sigma_e} \int_L j \ dl + j(x) \cdot n \frac{|L_r|}{\sigma_r}
\]

where \( L \) depends on \( x \), too. With the assumption that \( \frac{|L|}{\sigma_e} << \frac{|L_r|}{\sigma_r} \) we disregard the middle term. That means the charge need not to do work to get from area \( A_I \) to area \( A_r \). We now replace the current density as done above and get

\[
\Phi_e = \Phi(x) + \sigma \frac{\partial \Phi}{\partial n}(x) \frac{|L_r|}{\sigma_r}
\]

where the minus sign disappears because the point of view changes from the thin layer to the whole conductor \( \Omega \) and thus the definition of the normal to the boundary changes the sign. The characteristic constant is called contact impedance \( z_e \) and is defined for every electrode by

\[
z_e := \frac{|L_r|}{\sigma_r}
\]

with the unit \([z_e] = 1\Omega m^2\).

In summary we have the following boundary conditions for an electrode \( e \)

\[
\int_e \sigma \frac{\partial \Phi}{\partial n} \ dS = I \quad \text{on } e, \quad (3.2)
\]

\[
\sigma \frac{\partial \Phi}{\partial n} = 0 \quad \text{on } A_0, \quad (3.3)
\]

\[
\Phi + z_e \sigma \frac{\partial \Phi}{\partial n} = \Phi_e \quad \text{on } e. \quad (3.4)
\]

### 3.2. Complete Electrode Model

For a better overview Figure 2 gives two examples for experimental settings. The first example is a water tank where 16 electrodes are attached on the boundary and the second one is a block of concrete with 16 electrodes, too. In the following we describe equivalent experimental settings in a mathematical way.
Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain which represents a body with the conductivity $\sigma$ and $L$ electrodes are attached on the surface $\partial \Omega$. The current is injected through the electrodes into the body and we measure the resulting voltages on the same electrodes.

The electrodes are identified with the part connected to the surface and denoted by $e_l$, $l \in \{1, \ldots, L\}$. The current applied to electrode $e_l$ is denoted by $I_l$. If $I = (I_1, \ldots, I_L)^T$ satisfies the charge condition

$$\sum_{i=1}^{L} I_i = 0,$$

we call $I$ a current pattern. The corresponding voltage pattern is denoted by $U = (U_1, \ldots, U_L)^T$. The ground potential is chosen such that

$$\sum_{i=1}^{L} U_i = 0.$$

The relationship between current pattern and voltage pattern is a linear mapping

$$U = RI,$$

where $R \in \mathbb{R}^{L \times L}$ is symmetric, which is shown in [32]. $R$ is called the resistance matrix. Let

$$\Sigma := \{x \in \mathbb{R}^L : \sum_{i=1}^{L} x_i = 0\}.$$

(3.5)
With these notations and the derived boundary conditions for the electrodes from section 3.1.2 we can formulate the model. We now denote the potential by \( u \).

**Definition 3.2.1 (Complete Electrode Model).** Let \( \Omega \subset \mathbb{R}^n \ (n = 2, 3) \) be an open bounded domain. The \( L \in \mathbb{N} \) electrodes \( e_l \subset \partial \Omega \) are nonempty, connected and have a disjoint closure. \( U, I \in \Sigma \), \( z \in \mathbb{R}^L \) with contact impedances \( z_l > Z \) and \( Z > 0 \). \( \sigma \in \mathcal{A} \) and \( u \in H^1(\Omega) \) are real-valued functions with \( \mathcal{A} := \{ \sigma \in L^\infty(\Omega) : \lambda \leq \sigma \leq \lambda^{-1} \text{ a.e. with } \lambda \in (0, 1) \text{ and } \text{supp}(\sigma - \sigma^\dagger) \subset \Omega' \} \), where \( \sigma^\dagger \) denotes the real conductivity and \( \Omega' \) is an open subset with smooth boundary compactly contained in \( \Omega \). With these assumptions the complete electrode model is described in the following way:

\[
-\nabla \cdot (\sigma \nabla u) = 0 \quad \text{in } \Omega, \tag{3.6}
\]

\[
u + z_l \sigma \frac{\partial u}{\partial n} = U_l \quad \text{on } e_l, \quad l = 1, \ldots, L, \tag{3.7}
\]

\[
\sigma \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial \Omega \setminus \bigcup_{l=1}^L e_l, \tag{3.8}
\]

\[
\int_{e_l} \sigma \frac{\partial u}{\partial n} dS = I_l \quad \text{on } e_l, \quad l = 1, \ldots, L. \tag{3.9}
\]

**Remark.** The boundary Conditions (3.7), (3.8) and (3.9) result from the following assumptions about the electrodes:

- We have \( L \) separated electrodes. That implies \( \bigcup_{l=1}^L e_l \) is not connected and the current density is zero on the non-electrode area as can be seen in Equation (3.8).

- Current prefers the path of least resistance. So what we really know about the current is Equation (3.9).

- Between electrode and liquid is an extra resistance resulting from an electro-chemical effect. This fact and the fact that we neglect the conductivity from the metal part of the electrode implies Equation (3.7).

For later considerations in relation to the Galerkin method and the finite element method, we introduce the weak formulation of the complete electrode model. A similar proof with a complex valued conductivity \( \sigma \) can be found in [32] and the real valued case is also discussed in [12].
We are looking for a formulation in the space $H = H^1(\Omega) \oplus \Sigma$ which includes the electrical potential and the measurable voltages on the electrodes. A norm on the space $H$ is obviously given by

$$\| (u, U) \|_H^2 = \| u \|_{H^1(\Omega)}^2 + \| U \|_{\mathbb{R}^L}^2. \quad (3.10)$$

**Proposition 3.2.2** (Weak formulation complete electrode model). Let $\Omega, \sigma, I$ and $z$ satisfy the conditions from Definition 3.2.1. $(u, U) \in H$ is a (weak) solution of

1. $-\nabla \cdot (\sigma \nabla u) = 0$ in $\Omega,$ \quad (3.11)
2. $u + z_l \frac{\partial u}{\partial n} = U_l$ on $e_l,$ \quad $l = 1, \ldots, L,$ \quad (3.12)
3. $\sigma \frac{\partial u}{\partial n} = 0$ on $\partial \Omega \setminus \bigcup_{l=1}^L e_l,$ \quad (3.13)
4. $\int_{e_l} \sigma \frac{\partial u}{\partial n} dS = I_l$ on $e_l,$ \quad $l = 1, \ldots, L$ \quad (3.14)

if and only if

$$B((u, U), (v, V)) = \langle I, V \rangle_{\mathbb{R}_L \times \mathbb{R}_L}, \quad \forall (v, V) \in H, \quad (3.15)$$

with $B : H \times H \to \mathbb{R}$ is the bilinear form with

$$B((u, U), (v, V)) := \int_\Omega \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) dS. \quad (3.16)$$

**Proof.** First we assume $(u, U) \in H$ satisfies Equations (3.11), (3.12), (3.13) and (3.14). Let $(v, V) \in H$ be arbitrary. With the Lemma of variational calculus 2.0.7 and $\langle 0, v \rangle_{L^2 \times L^2} = 0$ we get

$$\langle -\nabla \cdot (\sigma \nabla u), v \rangle_{L^2 \times L^2} + \langle (\int_{e_l} \sigma \frac{\partial u}{\partial n} dS)_l, V \rangle_{\mathbb{R}_L \times \mathbb{R}_L} = \langle I, V \rangle_{\mathbb{R}_L \times \mathbb{R}_L}, \quad \forall (v, V) \in H. \quad (3.17)$$

We now apply Greens formula

$$\int_\Omega (\phi [\nabla \cdot \nabla \psi] + \nabla \phi \cdot \nabla \psi) \, dx = \int_{\partial \Omega} \phi \frac{\partial \psi}{\partial n} \, dS \quad (3.18)$$

on $\langle -\nabla \cdot (\sigma \nabla u), v \rangle_{L^2 \times L^2}.$ We thus get

$$-\int_\Omega [\nabla \cdot (\sigma \nabla u)] v \, dx = \int_\Omega \sigma \nabla u \cdot \nabla v \, dx - \int_{\partial \Omega} \sigma \frac{\partial u}{\partial n} v \, dS. \quad (3.19)$$
From Equations (3.12) and (3.13) we obtain
\[
\int_{\partial \Omega} \sigma \frac{\partial u}{\partial n} v \, dS = \int_{\partial \Omega \setminus \bigcup_{l=1}^L e_l} \sigma \frac{\partial u}{\partial n} v \, dS + \sum_{l=1}^L \int_{e_l} \sigma \frac{\partial u}{\partial n} v \, dS \\
= \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (U_l - u) v \, dS.
\] (3.20)

Equations (3.19) and (3.20) together become
\[
\langle -\nabla \cdot (\sigma \nabla u), v \rangle_{L^2 \times L^2} = \int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l) v \, dS.
\] (3.21)

Adding Equation (3.12) gives
\[
\langle \int_{e_l} \sigma \frac{\partial u}{\partial n} dS_l, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L} = - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (U_l - u) V \, dS.
\] (3.22)

Equations (3.21) and (3.22) added to (3.17) yields
\[
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) \, dS = \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L}
\]
for an arbitrary \((v, V) \in H\). It thus holds for all \((v, V) \in H\).

To prove the converse, we assume \((u, U) \in H\) satisfies Equation (3.15) for all \((v, V) \in H\).

Especially for \(v \in C_0^\infty(\Omega)\) and \(V = 0\) we get from Equation (3.15)
\[
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx = 0.
\] (3.23)

With Green’s formula (3.18) and \(v = 0\) on \(\partial \Omega\) Equation (3.23) is equivalent to
\[
\int_{\Omega} [\nabla \cdot (\sigma \nabla u)] v \, dx = 0,
\]
i.e. \(u\) satisfies \(-\nabla \cdot (\sigma \nabla u) = 0\) in \(\Omega\) in the weak sense. From Greens formula we therefore again get
\[
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx = \int_{\partial \Omega} \sigma \frac{\partial u}{\partial n} v \, dS
\] (3.24)
for an arbitrary \(v \in H^1(\Omega)\). Since \(v \in H^1(\Omega)\), \(V = 0\) and Equation (3.15) it follows
\[
\int_{\partial \Omega} \sigma \frac{\partial u}{\partial n} v \, dS = \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (U_l - u) v \, dS \\
= \int_{\partial \Omega \setminus \bigcup_{l=1}^L e_l} 0 \cdot v \, dS + \sum_{l=1}^L \frac{1}{z_l} (U_l - u) v \, dS.
\] (3.25)
Therefore $u$ holds $\sigma \frac{\partial u}{\partial n} = 0$ on $\partial \Omega \setminus \bigcup_{l=1}^{L} e_l$ and $u + z_l \sigma \frac{\partial u}{\partial n} = U_l$ on $e_l$, $l = 1, \ldots, L$ in the weak sense.

We now can use Equations (3.24) and (3.25) such that with Equation (3.15) and an arbitrary $(v, V) \in H$ we get

$$\sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} V_l(U_l - u) \, dS = \sum_{l=1}^{L} I_l V_l,$$

i.e. the following holds in the weak sense

$$\frac{1}{z_l} \int_{e_l} (U_l - u) \, dS = I_l, \quad l = 1, \ldots, L.$$

With Equation (3.12) which holds in the weak sense we get

$$\int_{e_l} \sigma \frac{\partial u}{\partial n} \, dS = I_l, \quad l = 1, \ldots, L.$$

Everything combined $(u, U)$ satisfies Equations (3.11), (3.12), (3.13) and (3.14) in the weak sense.

\[ \square \]

The operator $F(\sigma) : \Sigma \to H^1(\Omega) \oplus \Sigma$ with $I \mapsto (u, U)$ as the solution of the weak formulated partial differential equation is called the forward operator of the complete electrode model.

**Remark.** $F(\sigma)$ is linear with respect to current pattern $I$.

In the following we give a proof for the existence and the uniqueness of a solution from the forward operator. The proof follows the techniques presented in [32] for complex valued conductivities and a similar one for real valued conductivities can be found in [12]. This also includes the proofs for continuity and coercivity. We therefore use Lax Milgram’s theorem presented in the following lemma. The proof can be found in almost every functional analysis textbook, at this point we refer to [1].

**Lemma 3.2.3 (Lax-Milgram).** Let $X$ be a Hilbert space over $\mathbb{R}$ and $B : X \times X \to \mathbb{R}$ bilinear. There exist constants $c_0$ and $C_0$ with $0 < c_0 \leq C_0 < \infty$, such that for every $x, y \in X$ the following holds

- **Continuity:**
  $$|B(x, y)| \leq C_0 \|x\|_X \|y\|_X.$$
- Coercivity:
  \[ B(x, x) \geq c_0 \| x \|_X^2. \]

There then exists exactly one mapping \( A : X \to X \)
with
\[ B(y, x) = \langle y, Ax \rangle_X, \quad \forall x, y \in X. \]

Additionally \( A \in L(X, X) \) is an invertible operator with
\[ \| A \| \leq C_0 \quad \text{and} \quad \| A^{-1} \| \leq \frac{1}{c_0}. \]

To show the existence of a solution we satisfy continuity and coercivity for the bilinear form (3.16) in the following propositions.

**Proposition 3.2.4 (Continuity).** For the bilinear form \( B : H \times H \to \mathbb{R} \) with

\[ B((u, U), (v, V)) := \int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) \, dS \]

there exists one constant \( 0 < C_0 < \infty \) such that
\[ |B((u, U), (v, V))| \leq C_0 \| (u, U) \|_H \| (v, V) \|_H, \quad \forall (u, U), (v, V) \in H. \]

**Proof.** Remember
\[ \| (u, U) \|_H^2 = \| u \|_{H^1(\Omega)}^2 + \| U \|_{H^1}^2. \]

We now consider for arbitrary \( (u, U), (v, V) \in H \)
\[ |B((u, U), (v, V))| = \left| \int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) \, dS \right| \]
\[ \leq \int_{\Omega} |\sigma \nabla u \cdot \nabla v| \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} |(u - U_l)(v - V_l)| \, dS. \]

With \( \sigma \) bounded from above by \( \lambda^{-1} \) and \( z_l \) bounded from below by \( Z \) we get
\[ |B((u, U), (v, V))| \leq C \left( \int_{\Omega} |\nabla u \cdot \nabla v| \, dx + \sum_{l=1}^{L} \int_{e_l} |(u - U_l)(v - V_l)| \, dS \right). \]
According to Hölder’s inequality it follows
\[
\int_{\Omega} |\nabla u \cdot \nabla v| \, dx \leq \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}
\]
\[
\int_{e_l} |(u - U_l)(v - V_l)| \, dS \leq \|u - U_l\|_{L^2(e_l)} \|v - V_l\|_{L^2(e_l)}, \quad l = 1, \ldots, L.
\]

With this and with the Cauchy Schwarz’s inequality applied on the \(\mathbb{R}^{L+1}\) we get
\[
|B((u, U), (v, V))| \leq C \left( \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \sum_{l=1}^{L} \|u - U_l\|_{L^2(e_l)} \|v - V_l\|_{L^2(e_l)} \right)
\]
\[
\leq C \left( \|\nabla u\|^2_{L^2(\Omega)} + \sum_{l=1}^{L} \|u - U_l\|^2_{L^2(e_l)} \right)^{\frac{1}{2}} \left( \|\nabla v\|^2_{L^2(\Omega)} + \sum_{l=1}^{L} \|v - V_l\|^2_{L^2(e_l)} \right)^{\frac{1}{2}}
\]

The last step to finish the proof is to show
\[
\|\nabla u\|^2_{L^2(\Omega)} + \sum_{l=1}^{L} \|u - U_l\|^2_{L^2(e_l)} \leq C(\|u\|^2_{H^1(\Omega)} + \|U\|^2_{L^2})
\]

We therefore apply the triangle inequality and the Young’s inequality in the special case \(ab \leq \frac{1}{2}(a^2 + b^2)\) such that
\[
\sum_{l=1}^{L} \|u - U_l\|^2_{L^2(e_l)} \leq \sum_{l=1}^{L} \left[ \|u\|^2_{L^2(e_l)} + \|U_l\|^2_{L^2(e_l)} + 2\|u\|_{L^2(e_l)} \|U_l\|_{L^2(e_l)} \right]
\]
\[
\leq 2 \sum_{l=1}^{L} \left[ \|u\|^2_{L^2(e_l)} + \|U_l\|^2_{L^2(e_l)} \right].
\]

With the trace Theorem 2.0.4 we obtain the relationship
\[
\|u\|_{L^2(e_l)} \leq \|u\|_{L^2(\partial \Omega)} \leq C \|u\|_{H^1(\Omega)}
\]

and with \(e_{\max} := \max_l |e_l|\) we get
\[
\sum_{l=1}^{L} \|U_l\|^2_{L^2(e_l)} \leq e_{\max} \sum_{l=1}^{L} U_l^2 = e_{\max} \|U\|^2_{L^2}.
\]

Everything combined yields
\[
\sum_{l=1}^{L} \|u - U_l\|^2_{L^2(e_l)} \leq C \|u\|^2_{H^1(\Omega)} + e_{\max} \|U\|^2_{L^2}.
\]
We thus have
\[
\|\nabla u\|_{L^2(\Omega)}^2 + \sum_{l=1}^L \|u - U_l\|_{L^2(\epsilon_l)}^2 \leq \|\nabla u\|_{L^2(\Omega)}^2 + C\|u\|_{H^1(\Omega)}^2 + \epsilon_{\text{max}}\|U_l\|_{R^L}^2 \\
\leq C\|u\|_{H^1(\Omega)}^2 + \epsilon_{\text{max}}\|U_l\|_{R^L}^2 \\
\leq C(\|u\|_{H^1(\Omega)}^2 + \|U_l\|_{R^L}^2).
\]
This holds for \((v,V) \in H\), too. We therefore get
\[
|B((u,U),(v,V))| \leq C(\|u\|_{H^1(\Omega)}^2 + \|U_l\|_{R^L}^2)^{\frac{1}{2}}(\|v\|_{H^1(\Omega)}^2 + \|V_l\|_{R^L}^2)^{\frac{1}{2}}
= C\|(u,U)\|_H\|(v,V)\|_H.
\]

\[\square\]

**Proposition 3.2.5 (Coercivity).** For the bilinear form \(B : H \times H \to \mathbb{R}\) with
\[
B((u,U),(v,V)) := \int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{\epsilon_l} (u - U_l)(v - V_l) \, dS
\]
there exists one constant \(0 < c_0 < \infty\) such that
\[
B((u,U),(u,U)) \geq c_0\|(u,U)\|^2_H \quad \forall (u,U) \in H.
\]

**Proof.** We assume the opposite
\[
B((u,U),(u,U)) < \|(u,U)\|^2_H.
\]
Then we pick the sequence \(\{(u^n,U^n)\} \subset H\) such that
\[
\|(u^n,U^n)\|_H = 1 \quad \text{and} \quad B((u,U),(u,U)) < \frac{1}{n}.
\]
With the boundedness of the sequence and the reflexivity of \(H^1(\Omega)\) we can pick a weak convergent sub-sequence according to the theorem of Eberlein-Smulian 2.0.6 which is also denoted by \(\{(u^n,U^n)\}\). With the compact embedding \(H^1(\Omega) \hookrightarrow L^2(\Omega)\) we get strong convergence in \(L^2(\Omega)\)
\[
u^n \rightharpoonup u \quad \text{in} \quad H^1(\Omega) \Rightarrow u^n \rightarrow u \quad \text{in} \quad L^2(\Omega).
\]
We now consider

$$B((u^n, U^n), (u^n, U^n)) = \int_\Omega \sigma \nabla u^n \cdot \nabla u^n \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u^n - U^n_l)^2 \, dS$$

$$\geq \int_\Omega \sigma \nabla u^n \cdot \nabla u^n \, dx$$

$$\geq \lambda \int_\Omega \nabla u^n \cdot \nabla u^n \, dx = \lambda \|\nabla u^n\|_{L^2(\Omega)}^2$$

and we therefore have

$$\|\nabla u^n\|_{L^2(\Omega)}^2 \leq \frac{1}{\lambda n} \quad \Rightarrow \|\nabla u\|_{L^2(\Omega)} = 0.$$

Hence

$$u^n \to u \quad \text{in } H^1(\Omega) \quad \text{with } u = c = \text{const. in } \Omega.$$

From the continuity of the trace operator we can write

$$u^n_{|\partial \Omega} \to u_{|\partial \Omega} \quad \text{in } L^2(\partial \Omega).$$

We therefore consider

$$B((u^n, U^n), (u^n, U^n)) = \int_\Omega \sigma \nabla u^n \cdot \nabla u^n \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u^n - U^n_l)^2 \, dS$$

$$\geq \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u^n - U^n_l)^2 \, dS$$

$$\geq \frac{1}{z} \int_{e_l} (u^n - U^n_l)^2 \, dS = \frac{1}{z} \|u^n - U^n_l\|_{L^2(e_l)}^2, \quad l = 1, \ldots, L,$$

with $z := \max_l z_l$. This gives us

$$\|u^n - U^n_l\|_{L^2(e_l)}^2 \leq \frac{z}{n}, \quad l = 1, \ldots, L.$$

With $U^n$ converging to $U$, we get $U_l = u$ on $e_l$. Together with $u = c$ on $\partial \Omega$ we have

$$0 = \sum_{l=1}^L U_l = Lc \quad \Rightarrow U_l = 0, \quad l = 1, \ldots, L, \quad \text{and } u = 0 \text{ on } \partial \Omega.$$

Hence

$$u^n \to 0 \quad \text{in } H^1(\Omega),$$

$$U^n \to 0 \quad \text{in } \mathbb{R}^L$$

which contradicts the assumption $\|(u^n, U^n)\|_H = 1$. \qed
**Theorem 3.2.6** (Existence and uniqueness). For any $I \in \Sigma$ there exists a unique $(u, U) \in H$ satisfying

$$F(\sigma)I = (u, U)$$

(3.26)

with $(u, U) \in H$ as the solution of

$$B(((u, U), (v, V)) = \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \ \forall (v, V) \in H$$

with $B : H \times H \to \mathbb{R}$ is the bilinear form with

$$B(((u, U), (v, V)) := \int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) \, dS.$$

**Proof.** For the existence we apply Lax-Milgram’s Lemma 3.2.3 with the continuity and coercivity result from Propositions 3.2.4 and 3.2.5. This gives us the existence of exactly one linear mapping $A : H \to H$ such that

$$B(((u, U), (v, V)) = \langle A(u, U), (v, V) \rangle_H, \ \forall (u, U), (v, V) \in H.$$  

From Riesz’s representation theorem we obtain:

For every $R \in H^*$ there exists an $(u, U) \in H$ such that

$$R((u, U))((v, V)) = \langle (u, U), (v, V) \rangle_H, \ \forall (v, V) \in H,$$

where $R$ is an isometric and linear isomorphism. Therefore

$$B(((u, U), (v, V)) = \langle A(u, U), (v, V) \rangle_H = R(A(u, U))((v, V)), \ \forall (v, V) \in H.$$  

With $A \circ R$ bijective we get

$$\forall f \in H^* : \ \exists (u, U) \in H \ s.t. \ B(((u, U), (v, V)) = f(v, V), \ \forall (v, V) \in H.$$  

This especially holds for the linear and continuous mapping $f(v, V) := \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L}$. For the uniqueness we first show

$$B(((u, U), (v, V)) = 0, \ \forall (v, V) \in H \ \iff \ (u, U) = 0 \ in \ H. \ \ (3.27)$$

We assume $B(((u, U), (v, V)) = \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L} = 0, \ \forall (v, V) \in H$. Especially for an arbitrary $V$ we get $I = 0$. With Proposition 3.2.2 we get from Equations (3.13) and (3.14)

$$\sigma \frac{\partial u}{\partial n} = 0 \ \text{on} \ \partial \Omega.$$  

25
Together with Equation (3.12) we have

\[ u = U_l \quad \text{on } e_l, \ l = 1, \ldots, L. \]

Thus, particularly with \((v, V) = (u, U)\), we get

\[ B((u, U), (u, U)) = \int_{\Omega} \sigma(\nabla u)^2 \, dx = 0. \]

With \(\sigma\) bounded from below by a positive constant we get

\[ \nabla u = 0 \quad \text{in } \Omega \]

which implies \(u\) is a constant function on \(\Omega\). This results in

\[ u = U_1 = \ldots = U_L \]

and combined with the zero mean of \(U\) we get \((u, U) = 0\).

The converse part follows directly by insertion.

To proof the uniqueness we assume there are two solutions \((u^1, U^1), (u^2, U^2) \in H\) such that

\[ B((u^1, U^1), (v, V)) = \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \quad \forall (v, V) \in H. \]

With the bilinearity of \(B\) we get

\[ B((u^1, U^1) - (u^2, U^2), (v, V)) = 0, \quad \forall (v, V) \in H, \]

which implies with Equation (3.27)

\[ (u^1, U^1) = (u^2, U^2) \quad \text{in } H. \]

\[ \square \]

For the following results about uniform continuity and differentiability with respect to \(\sigma\) of the forward operator \(F(\sigma)\) remember

\[ A := \{ \sigma \in L^\infty(\Omega) : \lambda \leq \sigma \leq \lambda^{-1} \text{ a.e. with } \lambda \in (0, 1) \text{ and supp}(\sigma - \sigma^\dagger) \subset \Omega' \} \] (3.28)

where \(\sigma^\dagger\) denotes the real conductivity and \(\Omega'\) is an open subset with smooth boundary compactly contained in \(\Omega\). The proofs for uniform continuity and differentiability are not direct contents of this thesis but the interested reader can find one variant
in Appendix A.1.1 and A.1.2. The proofs can also be found in [12, 19]. For proving the uniform continuity, differentiability and later presented results the following proposition from [29], which depends on the main theorem by Meyers in [25], is used. We therefore get a relationship between $\lambda$ and differentiability which has to be considered carefully with respect to later convergence results of the algorithm.

**Proposition 3.2.7.** Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^n$, $n \geq 2$. For a fixed $\lambda \in (0, 1)$, there exists a constant $Q$, $2 < Q < \infty$, which depends on $\lambda$ and $n$ only, $Q \to 2$ as $\lambda \to 0$ and $Q \to \infty$ as $\lambda \to 1$ such that for any $2 < q < Q$ and any $\sigma \in \mathcal{A}$ it is satisfied the following:

If $f \in L^q(\Omega, \mathbb{R}^n)$, $h \in L^q(\Omega)$ and $u \in H^1(\Omega)$ is a weak solution to

$$\nabla \cdot (\sigma \nabla u) = \nabla \cdot f + h \text{ in } \Omega,$$

then $u \in W^{1,q}_{\text{loc}}(\Omega)$ and for any compact subset $\Omega' \subset \Omega$ the following estimate holds

$$\|u\|_{W^{1,q}(\Omega')} \leq C(\|u\|_{H^1(\Omega)} + \|f\|_{L^q(\Omega)} + \|h\|_{L^q(\Omega)})$$

where the constant $C$ depends on $\lambda, n, q, \Omega'$ and $\Omega$.

**Theorem 3.2.8 (Uniform continuity in $L^p$-norm).** Let $\sigma \in \mathcal{A}$, where $\delta \sigma$ is compactly supported in $\Omega$ such that $\sigma + \delta \sigma \in \mathcal{A}$, $Q > 2$ according to Proposition 3.2.7 and $p > \frac{2Q}{Q-2}$. Then $F_\sigma$ is uniformly continuous with respect to $L^p$-norm, i.e.

$$\|F(\sigma + \delta \sigma)I - F(\sigma)I\|_H \leq C\|\delta \sigma\|_{L^p(\Omega)}, \quad \forall I \in \Sigma \text{ with } \|I\|_2 < k, 0 < k < \infty. \quad (3.29)$$

**Theorem 3.2.9 (Differentiability in $L^p$-sense).** Let $\sigma \in \mathcal{A}$, where $\delta \sigma$ is compactly supported in $\Omega$ such that $\sigma + \delta \sigma \in \mathcal{A}$, $Q > 2$ according to Proposition 3.2.7 and $p > \frac{2Q}{Q-2}$. With $(u, U) = F(\sigma)I, I \in \Sigma, \|I\|_2 < k, 0 < k < \infty$ we define a bounded linear operator $F'(\sigma)[\delta \sigma]I = (\delta u, \delta U)$ as the solution to

$$\int_\Omega \sigma \nabla u \cdot \nabla v dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (\delta u - \delta U_l)(v - V_l) dS = - \int_\Omega \delta \sigma \nabla u \cdot \nabla v dx, \quad \forall (v, V) \in H. \quad (3.30)$$

Then $F(\sigma)$ is Fréchet-differentiable in $\sigma$ with respect to $L^p$-norm in the following sense

$$\frac{\|F(\sigma + \delta \sigma)I - F(\sigma)I - F'(\sigma)[\delta \sigma]I\|_H}{\|\delta \sigma\|_{L^p(\Omega')}} \to 0 \quad \text{as } \|\delta \sigma\|_{L^p(\Omega')} \to 0, \quad \forall I \in \Sigma.$$
The next theorem gives a continuity result for the derivative $F'$ presented in Theorem 3.2.9 which is necessary for later convergence results and concludes this section. For theoretical justification we refer to [19].

**Theorem 3.2.10.** Let $\sigma \in \mathcal{A}$, where $\delta \sigma$ compactly supported in $\Omega$ such that $\sigma + \delta \sigma \in \mathcal{A}$, $Q > 2$ according to Proposition 3.2.7 and $p > \frac{4Q}{Q-2}$. Then the operator $F'(\sigma)$ as defined in Theorem 3.2.9 is continuous in the sense

$$
\| F'(\sigma + \delta \sigma) - F'(\sigma) \|_{L^p(\Omega), H} \leq C \| \delta \sigma \|_{L^p(\Omega)}.
$$
4. The inverse problem

At this point we have the complete electrode model with its forward operator $F(\sigma) : \Sigma \rightarrow H$, which maps the current pattern $I$ to the voltage pattern $U$ and the potential function $u$. In practice we are only able to measure the voltage pattern. So we have to consider $\gamma F(\sigma) : \Sigma \rightarrow \Sigma$ with $\gamma$ defined as follows

$$\gamma : H \rightarrow \Sigma,$$

$$(u, U) \mapsto U.$$

The mapping $\gamma F(\sigma)$ is also called the Neumann-to-Dirichlet mapping, which is linear. From a physical point of view this current to voltage mapping is described with the resistance matrix $R$ as introduced in Section 3.2. The main aim of the electrical impedance tomography is to reconstruct the conductivity $\sigma$ from a given data set of current patterns and corresponding voltage patterns. We therefore define the nonlinear operator $G$ for an arbitrary fixed current pattern $I$ which depends on the conductivity $\sigma$ as follows

$$G : A \rightarrow \Sigma,$$

$$\sigma \mapsto \gamma F(\sigma)I.$$

We therefore want to solve the inverse problem $G(\sigma) = U$. In the following we present the general considerations about nonlinear inverse problems from [28]. In this context $G(\sigma) = U$ is an ill-posed inverse problem in the following sense.

**Definition 4.0.11** (Ill-posed inverse problem). A nonlinear inverse Problem $A : X \rightarrow Y$, $A(f) = g$ with $X, Y$ Banach spaces is termed locally ill-posed in $f^\dagger \in D(A)$ (domain of $A$) if the following holds:

For all $r > 0$ there exists a sequences $\{f^\sigma_n\} \subset B_r(f^\dagger) \cap D(A)$ which gives

$$\lim_{k \rightarrow \infty} \|A(f^\sigma_k) - A(f^\dagger)\|_Y = 0 \quad \text{but} \quad \lim_{k \rightarrow \infty} \|f^\sigma_k - f^\dagger\|_X \neq 0.$$

With the following proposition which is also presented and proven in [28] we introduce a class of ill-posed nonlinear operators.

**Lemma 4.0.12.** Let $A : X \rightarrow Y$ be a nonlinear, compact and continuous operator with $X, Y$ Banach spaces and $D(A)$ be weakly closed, e.g. closed and convex. If there exists a subsequence $\{f_n\} \subset D(A)$ such that

$$f_n \rightharpoonup f^\dagger \quad \text{with} \quad f_n \nrightarrow f^\dagger,$$

(4.1)
then the inverse problem is locally ill-posed in $f^\dagger$ and, moreover, the inverse $A^{-1}$ is not continuous in $g = A(f^\dagger)$.

**Proof.** We therefore assume the inverse problem is well-posed in $f^\dagger \in \mathcal{D}(A)$, i.e. there exists an $r > 0$ such that for all sequences $\{f_k\}_{k \in \mathbb{N}} \subset B_r(f^\dagger) \cap \mathcal{D}(A)$ we get
\[
\lim_{k \to \infty} \|A(f_k) - A(f^\dagger)\|_Y = 0 \Rightarrow \lim_{k \to \infty} \|f_k - f^\dagger\|_X = 0.
\]
This gives the existence of an $\epsilon > 0$ such that $A(f) = \tilde{g}$ has a unique solution for all $\tilde{g} \in \mathcal{R}(A) \cap B_\epsilon(g)$ where $g = A(f^\dagger)$. The compact and weak closed operator $A$ maps weak convergent sequences onto strong convergent ones. Thus the weak convergent sequence $\{f_k\}$ as defined in Equation (4.1) becomes
\[
g_k := A(f_k) \to A(f^\dagger),
\]
which implies the strong convergence of $\{f_k\}$ to $f^\dagger$ by the well-posedness of the problem. That contradicts the definition of the sequence. For sufficiently large $k$ we have $g_k \in \mathcal{R}(A) \cap B_\epsilon(g)$ and therefore $f_k$ is the unique solution of $A(f) = g_k$, i.e. $f_k = A^{-1}(g_k)$. With the definition of the sequence the operator $A^{-1}$ is not continuous in $g$.

**Remark.** When $X$ is a separable Hilbert space with infinite dimension, in particular, a weak but not strong convergent sequence is given by the orthonormal base $\{b_i\}$ of $X$.

The operator $G$ is a compact operator, i.e. it is continuous and the image of every bounded subset in $\mathcal{A}$ is relative compact which follows from the continuity result and the boundedness by $\|I\|_2$. The boundedness results from the coercivity. In conjunction with the infinite dimension of $\mathcal{A}$ and the weak closure it implies the ill-posed inverse problem if there exists a weak but not strong convergent sequence. Especially noisy data produces large errors and the reconstructions of the conductivity are not useful. To avoid this we have to use special inversion techniques which convert the ill-posed problem in a well-posed one. This results in the definition of a regularization.

**Definition 4.0.13** (Regularization). Let $X, Y$ be Banach spaces and $A : X \to Y$ be a nonlinear operator. With $f^\dagger$ we denote the minimum norm solution with respect to $\|\cdot\|_X$ of $A(f) = g$ and $f^* \in X$, i.e. $f^\dagger = \arg \min_{f \in \{u \in \mathcal{D}(A) | A(u) = g\}} \|f - f^*\|_X$.
A family of continuous operators $T_\alpha : Y \to X$ is called a regularisation of $A(f) = g$ if there exists a parameter choice $\alpha = \alpha(\delta, g^\delta)$ such that for all $g \in \mathcal{R}(A)$ the following holds

$$\sup\{\|f^\dagger - T_\alpha(\delta, g^\delta)\|_X \mid g^\delta \in Y, \|g - g^\delta\|_Y \leq \delta\} \to 0.$$  

For fixed $g \in \mathcal{R}(A)$ a pair $(T_\alpha, \alpha)$ is called a regularization method.

In the following we present the most common regularization method, the Tikhonov regularization.

### 4.1. Tikhonov functional

Andrey N. Tikhonov presented a method to replace an ill-posed problem with a well-posed problem as presented in [33] and the application to EIT can also be found in [19, 12]. We therefore formulate the following Tikhonov functional with one given data set $(I, U^\delta)$ and the noise level $\|U - U^\delta\|_{R_L} < \delta$ with $U$ as the zero noise potential on the electrodes.

$$J_\alpha(\sigma) = \frac{1}{2} \|G(\sigma) - U^\delta\|_{R_L}^2 + \alpha R(\sigma)$$  

(4.2)

where $R(\sigma)$ is a regularization functional depending on the a priori knowledge of the solution. $D(\sigma) := \frac{1}{2}\|G(\sigma) - U^\delta\|_{R_L}^2$ is the discrepancy term and $\alpha$ weights the regularization term with respect to the discrepancy term. Since we minimize the functional we get an approximation $\sigma^*_\alpha$ with

$$\sigma^*_\alpha = \arg \min_{\sigma \in A} J_\alpha(\sigma).$$

For later considerations we present the following continuity result for the discrepancy term. A similar result without the specification of the Hölder exponent is also given in [19].

**Proposition 4.1.1.** The discrepancy term $D(\sigma) := \frac{1}{2}\|G(\sigma) - U^\delta\|_{R_L}^2$ is Hölder-continuous with respect to $L^p(\Omega')$, $1 \leq p \leq \infty$, i.e. for $\sigma \in A$, where $\delta\sigma$ is compactly supported such that $\sigma + \delta\sigma \in A$, holds

$$|D(\sigma + \delta\sigma) - D(\sigma)| \leq C\|\delta\sigma\|_{L^p(\Omega')}^q, \quad \forall \eta \in I(p),$$

with

$$I(p) := \begin{cases} 
(0, 1] & \text{if } p > \frac{2Q}{Q-2} \\
(0, \frac{Q-2}{2Q}) & \text{if } 1 \leq p \leq \frac{2Q}{Q-2} 
\end{cases}$$

where $Q$ is according to Proposition 3.2.7.
Proof. With Cauchy Schwarz’s inequality we estimate

\[
|D(\sigma + \delta\sigma) - D(\sigma)| = \frac{1}{2} \| G(\sigma + \delta\sigma) - U^\delta \|_{L^2}^2 - \| G(\sigma) - U^\delta \|_{L^2}^2 \\
= \frac{1}{2} \langle G(\sigma + \delta\sigma) - U^\delta, G(\sigma + \delta\sigma) - U^\delta \rangle_{L^2 \times L^2} \\
+ \langle G(\sigma + \delta\sigma) - U^\delta, G(\sigma) - U^\delta \rangle_{L^2 \times L^2} \\
+ \langle G(\sigma) - U^\delta, G(\sigma + \delta\sigma) - U^\delta \rangle_{L^2 \times L^2} \\
- \langle G(\sigma) - U^\delta, G(\sigma) - U^\delta \rangle_{L^2 \times L^2} \\
= \| G(\sigma + \delta\sigma) - G(\sigma), G(\sigma + \delta\sigma) + G(\sigma) - 2U^\delta \|_{L^2 \times L^2} \\
\leq \| G(\sigma + \delta\sigma) - G(\sigma) \|_{L^2} \| G(\sigma + \delta\sigma) + G(\sigma) - 2U^\delta \|_{L^2}.
\]

Since the operator \( B \) is coercive, \( \|(u, U)\|_H \leq C \| I \|_2 \) follows and we can estimate

\[
\| G(\sigma + \delta\sigma) + G(\sigma) - 2U^\delta \|_{L^p} \leq C(\| I \|_2 + \| U^\delta \|_2) \leq C.
\]

With the assumption \( p > \frac{2Q}{Q-2} \) we therefore get Lipschitz continuity from Theorem 3.2.8

\[
|D(\sigma + \delta\sigma) - D(\sigma)| \leq C \| \delta\sigma \|_{L^p(\Omega)}.
\]

Since \( A \) is bounded with respect to \( L^\infty(\Omega) \), we get the Hölder continuity for all \( \eta \in (0, 1) \) for \( p > \frac{2Q}{Q-2} \). With the interpolation inequality and \( q \) such that \( \frac{1}{p} = \frac{1-\xi}{\infty} + \frac{\xi}{q} \) we can estimate

\[
|D(\sigma + \delta\sigma) - D(\sigma)| \leq C \| \delta\sigma \|_{L^\infty(\Omega)}^{1-\xi} \| \delta\sigma \|_{L^q(\Omega)}^\xi \leq C \| \delta\sigma \|_{L^q(\Omega)}^\xi.
\]

We therefore get \( \frac{q}{\xi} = p > \frac{2Q}{Q-2} \) which implies the Hölder continuity for \( \xi \leq \frac{Q-2}{2Q} \) and especially for \( 1 \leq q \leq \frac{2Q}{Q-2} \), which concludes the proof. \( \square \)

4.2. Sparse regularization functional

We assume that the real conductivity consists of a background plus nonzero conductivity on a bounded subset of the whole domain \( \Omega \). Moreover, the reality is not smooth with respect to the conductivity hence we want to allow sharp edges in the reconstruction. We describe this attribute in a mathematical way with the term sparsity.
**Definition 4.2.1** (Sparsity). Let $X$ be an infinite dimensional Hilbert space. There thus exists the orthonormal basis $\{b_i\}_{i \in I}$ where $I$ is a countable index set. $x \in X$ is termed sparse with respect to $\{b_i\}_{i \in I}$ if $\langle x, b_i \rangle_X$ is nonzero for a finite number of basis elements, i.e. $x$ is finitely supported in $\{b_i\}_{i \in I}$.

**Remark.** The orthonormal basis $\{b_i\}_{i \in I}$ can be replaced by a frame.

This leads to the formulation of the following regularization term:

$$R(x) = \|x\|^p_\ell_p = \sum_{i \in I} |\langle x, b_i \rangle_X|^p$$

with $1 \leq p < 2$. The closer $p$ is to 1, the more sparsity-promoting is the regularization term because many small but nonzero values are penalized more and few bigger values are preferred. For the application we use $p = 1$ and assume that the reconstructed conductivity is one constant background $\sigma_0$ plus one sparse $\delta \sigma \in H^1_0(\Omega')$. The choice of $H^1_0(\Omega')$ with higher regularity than, for example, the $L^2$ seems to be a step away from sharp edges but is necessary as can be seen in later considerations.

Otherwise the $l^1$ is a subspace of the $l^2$ and consequently $R$ penalizes more than $\|\cdot\|_{H^1(\Omega')}$. We therefore have to accept this compromise at the moment. Finally, the penalty term depends on $\delta \sigma$ with the orthonormal basis $\{b_i\}$ of the $H^1_0(\Omega')$ and becomes

$$R(\delta \sigma) = \|\delta \sigma\|_{l^1} = \sum_{i \in I} |\langle \delta \sigma, b_i \rangle_{H^1_0(\Omega')}|.$$  \hspace{1cm} (4.3)

For later considerations we show the convexity and therefore the weak lower semi-continuity of $R(\delta \sigma)$. More general relations between convexity and weak lower semi-continuity are presented and proven in [31].

**Lemma 4.2.2.** The penalty term $R(\delta \sigma)$ is convex and weakly lower semi-continuous, i.e. for all $x_1, x_2 \in H^1_0(\Omega')$

$$R(t x_1 + (1 - t)x_2) \leq t R(x_1) + (1 - t) R(x_2), \quad \forall t \in (0, 1),$$

holds and for all $\{x_n\} \subset H^1_0(\Omega')$ with $x_n \to x$

$$\liminf_{n \to \infty} R(x_n) \geq R(x)$$

is satisfied.
Proof. With the triangle inequality the convexity follows directly. To proof the weak lower semi-continuity we assume the converse. There exists a sequence \( \{x_n\} \subset H^1_0(\Omega') \) with \( x_n \rightharpoonup x \) for which holds \( \liminf_{n \to \infty} R(x_n) < R(x) \). The set
\[
S := \{(y, k) \in H^1_0(\Omega') \times \mathbb{R} | R(y) \leq k\}
\]
is convex, which follows from the convexity of \( R \). With \( \liminf_{n \to \infty} R(x_n) < R(x) \) there exists an \( l \in \mathbb{R} \) such that \( \liminf_{n \to \infty} R(x_n) < l < R(x) \). Therefore \( (x, l) \) is not in \( S \). According to the separation theorem of Hahn-Banach [1] there exists a linear continuous functional \( \phi \in (H^1_0(\Omega') \times \mathbb{R})^* \) such that
\[
\phi((x, l)) < \inf \{ \phi((y, k)) | (y, k) \in S \}.
\]
The inequality holds especially for \( (x_n, R(x_n)) \in S \) and with an \( \alpha \in \mathbb{R} \) and \( v \in H^1_0(\Omega') \) we get
\[
\alpha l + \langle x, v \rangle_{H^1_0(\Omega')} < \alpha R(x_n) + \langle x_n, v \rangle_{H^1_0(\Omega')}.
\]
Using the lower limit together with the weak convergence of \( \{x_n\} \) we obtain
\[
l < \liminf_{n \to \infty} R(x_n),
\]
which is a contradiction and concludes the proof.

Finally, the minimizing functional is
\[
J_\alpha(\sigma) = \frac{1}{2} \|G(\sigma_0 + \delta \sigma) - U^\delta\|_{L^2}^2 + \alpha \|\delta \sigma\|_{L^1}.
\]
(4.4)

where \( \sigma = \sigma_0 + \delta \sigma \).

In the following theorem we present the regularization abilities for minimizing the functional \( J_\alpha \). The theorem is also presented in [19] and more general results and their proofs can be found in [16].

Theorem 4.2.3.

Existence: There exists at least one minimizer \( \delta \sigma_\alpha^\delta \) to the functional \( J_\alpha(\sigma) \) on the admissible set \( A \).

Stability: Let \( \{U^n\} \subset \Sigma \) be a sequence of noisy data converging to \( U^\delta \) and \( \{\delta \sigma^n\} \) a sequence satisfying
\[
\delta \sigma^n \in \arg \min \left\{ \frac{1}{2} \|G(\sigma_0 + \delta \sigma) - U^n\|_{L^2}^2 + \alpha \|\delta \sigma\|_{L^1} | \sigma = \sigma_0 + \delta \sigma \in A \right\}.
\]

Then the sequence \( \{\delta \sigma^n\} \) has a sub-sequence converging in \( H^1(\Omega') \) to a minimizer of \( J_\alpha \).
Convergence: If $\alpha = \alpha(\delta)$ satisfies
\[
\lim_{\delta \to 0} \alpha(\delta) = 0 \quad \text{and} \quad \lim_{\delta \to 0} \frac{\delta^2}{\alpha(\delta)} = 0,
\]
then the sequence of minimizers $\{\delta \sigma^n\}$ has a sub-sequence converging in $H^1(\Omega')$ to a $\| \cdot \|_{L^1}$-minimizing solution $\delta \sigma^\dagger$ as $\delta \to 0$. Moreover, if $\delta \sigma^\dagger$ is unique, then the whole sequence converges.

Proof. Existence: We consider the nonnegativity of the functional $J_\alpha(\sigma)$. Thus it is bounded from below and there exists a minimizing sequence $\{\delta \sigma^n\}$ with $\{\sigma^n\} \subset A$ where $\sigma^n = \sigma_0 + \delta \sigma^n$ such that
\[
\lim_{n \to \infty} J_\alpha(\sigma^n) = \inf_{\sigma \in A} J_\alpha(\sigma) \quad \text{with} \quad J_\alpha(\sigma^n) \leq C.
\]
This implies
\[
\|\delta \sigma^n\|_{L^1} \leq C.
\]
The triangle inequality applied to $\| \cdot \|_2$ yields $\|\delta \sigma^n\|_2 \leq \|\delta \sigma^n\|_{L^1} \leq C$. With
\[
\|\delta \sigma^n\|_2 = \left( \sum_{i=1}^{\infty} |\langle \delta \sigma^n, b_i \rangle_{H^1(\Omega')}|^2 \right)^{\frac{1}{2}}
\]
\[
= \left( \sum_{i=1}^{\infty} \langle \sum_{j=1}^{\infty} \langle \delta \sigma^n, b_i \rangle_{H^1(\Omega')} b_j, \delta \sigma^n \rangle_{H^1(\Omega')} \right)^{\frac{1}{2}}
\]
\[
= \left( \langle \delta \sigma^n, \delta \sigma^n \rangle_{H^1(\Omega')} \right)^{\frac{1}{2}} = \|\delta \sigma^n\|_{H^1(\Omega')}
\]
we get the uniform boundedness of $\{\delta \sigma^n\}$ in $H^1(\Omega')$. The reflexivity of $H^1(\Omega')$ and the Riesz representation theorem result in the existence of a sub-sequence also denoted by $\{\delta \sigma^n\}$ and a $\delta \sigma^* \in H^1_0(\Omega')$ such that
\[
\delta \sigma^n \rightharpoonup \delta \sigma^* \quad \text{in} \quad H^1(\Omega').
\]
According to Kondrashov embedding Theorem 2.0.5 and the three-dimensional domain $\Omega'$ we get the compact embedding in $L^p(\Omega')$ for $p < 6$. We therefore get
\[
\delta \sigma^n \to \delta \sigma^* \quad \text{in} \quad L^p(\Omega'), \ p < 6.
\]
With the Hölder continuity of the discrepancy term we have
\[
|D(\sigma^n) - D(\sigma^*)| \leq C \|\delta \sigma^n - \delta \sigma^*\|_{L^p(\Omega')}.
\]
Since we have strong convergence in $L^p(\Omega')$, $p < 6$, it follows that

$$D(\sigma^n) \to D(\sigma^*)$$

The weak lower semi-continuity of $\|\delta \sigma\|_\mu$ implies

$$\liminf_{n \to \infty} \|\delta \sigma^n\|_\mu \geq \|\delta \sigma^*\|_\mu$$

and combining the above yields

$$J_\alpha(\sigma^*) \leq \lim\inf_{n \to \infty} J_\alpha(\sigma^n) = \lim_{n \to \infty} J_\alpha(\sigma^n) = \inf_{\sigma \in \mathcal{A}} J_\alpha(\sigma).$$

This proves the existence of a minimizer.

**Stability:** The definition of $\delta \sigma^n$ with $\sigma^n = \sigma_0 + \delta \sigma^n$ gives

$$\|G(\sigma^n) - U^n\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu \leq \|G(\sigma) - U^n\|^2_{R^L} + \alpha \|\delta \sigma\|_\mu,$$  

We can therefore use Young’s inequality and one fixed $\bar{\sigma} \in \mathcal{A}$ for the following estimate. For easier notation we denote $U = U^\delta$.

$$\|G(\sigma^n) - U^n\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu \leq \|G(\sigma^n) - U\|^2_{R^L} + \|U - U^n\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu$$

$$+ 2 \|G(\sigma^n) - U\|_{R^L} \|U - U^n\|_{R^L}$$

$$\leq 2 \|G(\sigma^n) - U\|^2_{R^L} + 2 \|U - U^n\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu.$$  

We thus obtain

$$\|G(\sigma^n) - U\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu \leq 2 \|G(\sigma^n) - U\|^2_{R^L} + \|\sigma - \sigma^n\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu$$

$$\leq 2 \|G(\sigma) - U\|^2_{R^L} + \|\sigma - \sigma^n\|^2_{R^L} + \alpha \|\delta \sigma\|_\mu.$$  

According to the convergence $U^n \to U^\delta$ for every $\epsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ holds

$$\|G(\sigma^n) - U\|^2_{R^L} + \alpha \|\delta \sigma^n\|_\mu \leq 2 \|G(\sigma) - U\|^2_{R^L} + \epsilon + \alpha \|\delta \sigma\|_\mu =: C.$$  

As mentioned in the existence proof there exists a sub-sequence also denoted by $\{\delta \sigma^n\}$ and a $\delta \sigma^* \in H^1(\Omega')$ such that

$$\delta \sigma^n \rightharpoonup \delta \sigma^* \quad \text{in} \quad H^1(\Omega')$$

which converges strongly in $L^p(\Omega')$ for $p < 6$. Using the uniform continuity of the forward operator yields

$$G(\sigma^n) \to G(\sigma^*)$$
and with the strong convergence of \( \{U^n\} \) we obtain

\[
\|G(\sigma^n) - U^n\|_{RL} \to \|G(\sigma^*) - U\|_{RL}.
\]

Since \( \|\cdot\|_1 \) is weak lower semi-continuous, the result is

\[
\|G(\sigma^*) - U\|_{RL}^2 + \alpha \|\delta \sigma^*\|_1 \leq \lim_{n \to \infty} \|G(\sigma^n) - U^n\|_{RL}^2 + \alpha \liminf_{n \to \infty} \|\delta \sigma^n\|_1
\]

\[
\leq \limsup_{n \to \infty} (\|G(\sigma^n) - U^n\|_{RL}^2 + \alpha \|\delta \sigma^n\|_1)
\]

\[
\leq \lim_{n \to \infty} (\|G(\sigma) - U^n\|_{RL}^2 + \alpha \|\delta \sigma\|_1)
\]

\[
= \|G(\sigma) - U\|_{RL}^2 + \alpha \|\delta \sigma\|_1, \quad \sigma \in \mathcal{A}.
\]

Therefore \( \sigma^* = \sigma_0 + \delta \sigma^* \) is a minimizer and initiation as \( \sigma = \sigma^* \) yields

\[
\|G(\sigma^*) - U\|_{RL}^2 + \alpha \|\delta \sigma^*\|_1 = \lim_{n \to \infty} (\|G(\sigma^n) - U^n\|_{RL}^2 + \alpha \|\delta \sigma^n\|_1) \tag{4.5}
\]

The last step is to show that \( \|\delta \sigma^n\|_1 \) converges to \( \|\delta \sigma^*\|_1 \). To prove we assume the converse. With weak lower semi-continuity it follows

\[
k := \limsup_{n \to \infty} \|\delta \sigma^n\|_1 > \|\delta \sigma^*\|_1.
\]

We now pick a sub-sequence \( \{\delta \sigma^n\} \) such that \( \|\delta \sigma^n\|_1 \to k \). With Equation (4.5) we have

\[
\|G(\sigma^*) - U\|_{RL}^2 + \alpha \|\delta \sigma^*\|_1 = \lim_{n \to \infty} (\|G(\sigma^n) - U^n\|_{RL}^2 + \alpha \|\delta \sigma^n\|_1)
\]

\[
= \lim_{n \to \infty} (\|G(\sigma^n) - U^n\|_{RL}^2 + \alpha \lim_{n \to \infty} (\|\delta \sigma^n\|_1))
\]

\[
= \lim_{n \to \infty} \|G(\sigma^n) - U^n\|_{RL}^2 + \alpha k.
\]

This results in

\[
\lim_{n \to \infty} \|G(\sigma^n) - U^n\|_{RL}^2 = \|G(\sigma^*) - U\|_{RL}^2 + \alpha (\|\delta \sigma^*\|_1 - k) < \|G(\sigma^*) - U\|_{RL}^2,
\]

which is a contradiction to \( \|G(\sigma^n) - U^n\|_{RL} \to \|G(\sigma^*) - U\|_{RL} \). We thus have norm convergence in the \( l^1 \)-norm which implies norm convergence in the \( H^1(\Omega') \)-norm as can be seen in the existence proof. Combined with the weak convergence we have

\[
\delta \sigma^n \to \delta \sigma^* \quad \text{in} \quad H^1(\Omega').
\]
**Convergence:** Let \( \delta \sigma^\dagger \) be a \( \| \cdot \|_{l^1} \)-minimizing solution with \( G(\sigma_0 + \delta \sigma^\dagger) = U \) and \( \{ \delta \sigma^\delta \} \) be a sequence of minimizers. This gives

\[
\| G(\sigma_0 + \delta \sigma^\delta) - U^\delta \|_{R^L} + \alpha \| \delta \sigma^\delta \|_{l^1} \leq \| G(\sigma_0 + \delta \sigma^\dagger) - U^\delta \|_{R^L} + \alpha \| \delta \sigma^\dagger \|_{l^1} \\
\leq \delta^2 + \alpha \| \delta \sigma^\dagger \|_{l^1}.
\]

With \( \delta \to 0 \) and \( \lim_{\delta \to 0} \alpha(\delta) = 0 \) we therefore get

\[
\lim_{\delta \to 0} \| G(\sigma_0 + \delta \sigma^\delta - U^\delta) \|_{R^L} = 0 \quad \Rightarrow \quad G(\sigma_0 + \delta \sigma^\delta) \to U \quad \text{as} \quad \delta \to 0.
\]

We thus consider

\[
\frac{\| G(\sigma_0 + \delta \sigma^\delta) - U^\delta \|_{R^L}}{\alpha} + \| \delta \sigma^\delta \|_{l^1} \leq \frac{\delta^2}{\alpha} + \| \delta \sigma^\dagger \|_{l^1}
\]

and get uniform boundedness of \( \{ \delta \sigma^\delta \} \) and

\[
\limsup_{\delta \to 0} \| \delta \sigma^\delta \|_{l^1} \leq \| \delta \sigma^\dagger \|_{l^1}.
\]

As mentioned in the existence proof there exists a sub-sequence which is also denoted by \( \{ \delta \sigma^\delta \} \) and a \( \delta \sigma^* \in H^1_0(\Omega') \) such that

\[
\delta \sigma^n \rightharpoonup \delta \sigma^* \quad \text{in} \quad H^1(\Omega')
\]

which converges strongly in \( L^p(\Omega') \), \( p < 6 \). With uniform continuity of \( G \) follows \( \lim_{\delta \to 0} G(\sigma_0 + \delta \sigma^\delta) = U \). Since \( \| \cdot \|_{l^1} \) is weak lower semi-continuous, this results in

\[
\| \delta \sigma^* \|_{l^1} \leq \liminf_{\delta \to 0} \| \delta \sigma^\delta \|_{l^1} \leq \limsup_{\delta \to 0} \| \delta \sigma^\delta \|_{l^1} \leq \| \delta \sigma^\dagger \|_{l^1} \leq \| \delta \sigma \|_{l^1}
\]

for all \( \delta \sigma \) with \( G(\sigma_0 + \delta \sigma) = U \). Especially with \( \delta \sigma = \delta \sigma^* \) we obtain \( \| \delta \sigma^* \|_{l^1} = \| \delta \sigma^\dagger \|_{l^1} \) and thus \( \delta \sigma^* \) is a \( \| \cdot \|_{l^1} \)-minimizing solution. The convergence in \( l^1 \)-norm and the weak convergence in \( H^1(\Omega') \) imply the strong convergence of \( \{ \delta \sigma^\delta \} \) in \( H^1(\Omega') \), as mentioned in the stability proof.

If the \( \| \cdot \|_{l^1} \)-minimizing solution \( \delta \sigma^\dagger \) is unique, every convergent sub-sequence of \( \{ \delta \sigma^\delta \} \) has the limit point \( \delta \sigma^\dagger \). We therefore get convergence of the whole sequence. \( \square \)
Algorithm 5.1 Generalized conditional gradient method

Choose $f_0 \in H$ with $\Phi(f_0) < \infty$;

while stopping criterion do

Determine the descent direction $v_n$ as a solution of

\[ \min_{v \in H} \langle F'(f_n), v \rangle_H + \Phi(v) \]

Determine the step size $s_n = \arg \min_{s \in [0,1]} F(f_n + s(v_n - f_n)) + \Phi(f_n + s(v_n - f_n))$;

Update $f_{n+1} = f_n + s_n(v_n - f_n)$;

end while

5. Algorithm

The algorithm is based on the iterated soft shrinkage approach presented for linear operators in the pioneering work [9]. Applications to the linearized model are discussed in [19]. A generalization to nonlinear inverse problems is given in [7, 4, 27] and especially the generalized conditional gradient method in particular will be applied to our problem in the following as presented in [4]. In a general setting, let $H$ be a Hilbert space and $T : H \to H$ a nonlinear operator which implies the inverse problem $T(f) = g$. As mentioned in the preceding section we want to minimize the Tikhonov functional by using sparsity constraints. In [7] a more general approach is presented by minimizing the functional

\[ \min_{f \in H} F(f) + \Phi(f) \]

with the generalized conditional gradient method, cf. Algorithm 5.1. $F : H \to \mathbb{R}$ is in general not convex but Gâteaux differentiable and to state convergence results we must assume that $\Phi$ is proper, lower semi-continuous and coercive with respect to the norm in $H$. But especially in our case the functional

\[ \frac{1}{2} \|T(f) - g\|^2_H + \alpha \Psi(f) \]

with $\Psi(f) := \|f\|_{\ell^1}$ as the sparsity promoting penalty term does not fulfill the assumed coercivity criterion and we have to extend the functional to

\[ \Gamma(f) = \frac{1}{2} \|T(f) - g\|^2_H - \xi \Theta(f) + \xi \Theta(f) + \alpha \Psi(f). \quad (5.1) \]

The following convergence result verifies the application of this algorithm to electrical impedance tomography and formulates necessary assumptions on $\Theta$. The proof can be found in [4].
Theorem 5.0.4. Let $\Psi$ be proper and weakly coercive, i.e. $\lim_{\|f\|_H \to \infty} |\Psi(f)| = \infty$. Assume $E_s = \{f \in H : \Psi(f) \leq s\}$ to be compact for all $s \in \mathbb{R}$. Let $\Theta$ be positive, coercive, continuously Fréchet differentiable, bounded on bounded sets and convex such that $\Phi = \Theta + \Psi$ is convex. Let $T : H \to H$ denote a nonlinear, continuous Fréchet differentiable operator which is bounded on bounded sets. Assume that $f_0 \in H$ obeys $\Phi(f_0) < \infty$. Let $\{f_n\}_{n \in \mathbb{N}}$ denote the sequence generated by Algorithm 5.1. The sequence $\{f_n\}$ then has a convergent sub-sequence and every convergent sub-sequence of $\{f_n\}$ converges to a stationary point of the functional $\Gamma$.

In the following we use the generalized conditional gradient method with the functionals

$$\Theta(f) := \frac{1}{2} \|f\|_H^2, \quad \Psi(f) := \|f\|_\nu = \sum_{k=1}^\infty |\langle f, \phi_k \rangle_H|,$$

where $\{\phi_k\}$ denotes an orthonormal base of $H$. Especially for $H := H^1(\Omega)$ the weak coercivity results from the inequality $\|f\|_{L^2} \leq \|f\|_\nu$. The assumptions for the operator $T = G$ are satisfied as can be seen in Section 3.2.

We therefore consider the minimization problem to determine the direction of descent. The Fréchet derivative of $F$, cf. Equation (5.1), is given by

$$T(f)[\delta f] = \langle T'(f)^*(T(f) - g^\delta) - \xi f, \delta f \rangle_H.$$

Thus the minimization problem becomes

$$\min_{v \in H} \langle T'(f_n)^*(T(f_n) - g^\delta) - \xi f_n, v \rangle_H + \frac{\lambda}{2} \|v\|_H^2 + \alpha \sum_{k=1}^\infty |\langle v, \phi_k \rangle_H|.$$

As presented in [4] a unique minimizer of the functional is given by

$$v_n = S_{\alpha/\xi}[f_n - \xi^{-1}T'(f_n)^*(T(f_n) - g^\delta)]$$

where $S_{\alpha/\xi}$ is the threshold or also called soft shrinkage operator with

$$S_\mu[f] = \sum_{i=1}^\infty f_i \phi_i \quad \text{with} \quad f_i := \begin{cases} \frac{|\langle f, \phi_i \rangle_H| - \mu}{\|\langle f, \phi_i \rangle_H\|} \text{sign}(\langle f, \phi_i \rangle_H), & \text{if } |\langle f, \phi_i \rangle_H| > \mu \\ 0, & \text{else}. \end{cases}$$

(5.2)

For detailed information about the shrinkage operator and the relationship $S_\mu[f] = (I + \mu \partial \Psi)^{-1}(f)$ we refer to [9].
Algorithm 5.2 Specified sparse gradient method

Choose $f_0 \in H$ with $\Phi(f_0) < \infty$;

\[ \textbf{while stopping criterion do} \]

\[ \text{Determine the direction of descent } v_n := S_{\alpha \tau}[f_n - \tau T'(f_n)^*(T(f_n) - g^\delta)]; \]

\[ \text{Update } f_{n+1} = v_n; \]

\[ \textbf{end while} \]

The determination of the step size $s_n$ is not necessary if $\tau := 1/\xi$ is sufficiently small enough, as presented in [4]. We thus assume $s_n = 1$ as a constant and concentrate on an intelligent selection of $\tau$ such that the weak monotonicity of the sequence is satisfied later. The specification of the generalized conditional gradient method can be seen in Algorithm 5.2.

The application to the functional presented in the preceding section

\[ J_\alpha(\sigma) = \frac{1}{2}\|G(\sigma_0 + \delta \sigma) - U^\delta\|_{RL}^2 + \alpha \|\delta \sigma\|_1 \]  

(5.3)

yields

\[ \delta \sigma^{n+1} = S_{\alpha \tau}[\delta \sigma^n - \tau G'(\sigma^n)^*(G(\sigma^n) - U^\delta)] \]  

(5.4)

where $\sigma^n = \sigma_0 + \delta \sigma^n$.

Remark. The iteration consists of two steps

- In squared brackets is the iteration step of a gradient descent method with $G'(\sigma^n)^*(G(\sigma^n) - U^\delta)$ as the derivative of the discrepancy term $\frac{1}{2}\|G(\sigma_0 + \delta \sigma) - U^\delta\|_{RL}^2$.

- The shrinkage step by the operator $S_{\alpha \tau}$ which promotes sparsity as it sets all small coefficients to zero.

From here on the algorithm consists of the following steps.

5.1. Step 1: Gradient $D'(\sigma)$

The main idea is to obtain the gradient of the whole discrepancy term and to avoid calculating $G'(\sigma)^*$ in the iteration Rule (5.4). The discrepancy term is a mapping
defined in the following way

\[
D : \mathcal{A} \to \mathbb{R} \\
\sigma \mapsto \frac{1}{2} \| G(\sigma) - U^\delta \|_{L^2}^2.
\]

This therefore gives

\[
D'(\sigma) = G'(\sigma)^*(G(\sigma) - U^\delta).
\]

For the calculation of \(D'(\sigma)\) we use the adjoint method which is presented in [35] and incurs lower calculation costs. This can be achieved by solving an adjoint problem as below. This result is also presented similarly in [12].

**Proposition 5.1.1** (Calculation \(D'(\sigma)\)).

Let \(\delta \sigma\) be compactly supported such that \(\sigma + \delta \sigma \in \mathcal{A}, Q > 2\) according to Proposition 3.2.7, \(p > \frac{2Q}{Q - 2}\) and \(\frac{1}{p} + \frac{1}{q} = 1\), i.e. \(q < \frac{2Q}{Q + 2}\). The gradient of \(D\) is then defined by

\[
D' : \mathcal{A} \to L^q(\Omega) \cong (L^p(\Omega))^*,
\]

\[
\sigma \mapsto \langle D'(\sigma), \delta \sigma \rangle_{L^q \times L^p} = \int_{\Omega} D'(\sigma) \delta \sigma \, dx
\]

with

\[
D'(\sigma) = -\nabla u(\sigma) \cdot \nabla p(\sigma) \in L^q(\Omega)
\]  \hspace{1cm} (5.5)

and \(p \in H^1(\Omega)\) with \(P \in \Sigma\) such that

\[
-\nabla \cdot (\sigma \nabla p) = 0 \quad \text{in } \Omega, \hspace{1cm} (5.6)
\]

\[
p + z_l \sigma \frac{\partial p}{\partial n} = P_l \quad \text{on } e_l, \quad l = 1, \ldots, L, \hspace{1cm} (5.7)
\]

\[
\int_{e_l} \sigma \frac{\partial p}{\partial n} \, dS = (G(\sigma))_l - U^\delta_l \quad \text{for } l = 1, \ldots, L, \hspace{1cm} (5.8)
\]

\[
\sigma \frac{\partial p}{\partial n} = 0 \quad \text{on } \partial \Omega \setminus \bigcup_{l=1}^L e_l \hspace{1cm} (5.9)
\]

with \((u(\sigma), G(\sigma)) = F(\sigma)I\).

**Proof.** First, we calculate the gradient of the discrepancy term \(D\) by applying the chain rule. We therefore get

\[
D'(\sigma)[\delta \sigma] = (G(\sigma) - U^\delta) \cdot G'(\sigma)[\delta \sigma] = \langle G(\sigma) - U^\delta, G'(\sigma)[\delta \sigma] \rangle_{\mathbb{R}^L \times \mathbb{R}^L}.
\]
For easier notations we denote $(u, U) := F(\sigma)I$ and $(\delta u, \delta U) := F'(\sigma)[\delta\sigma]I$ with $F'(\sigma)$ as defined in Theorem 3.2.9. According to this theorem we get the $L^p$-differentiability of $F(\sigma)$. Also remember $G(\sigma) = \gamma F(\sigma)I$. We thus get

$$D'(\sigma)[\delta\sigma] = \langle U - U^\delta, \delta U \rangle_{\mathbb{R}L \times \mathbb{R}L}.$$  

We now consider the weak formulation for $(\delta u, \delta U)$ given by Equation (3.30) in Theorem 3.2.9 for $(v, V) = (p, P)$

$$\int_{\Omega} \sigma \nabla \delta u \cdot \nabla p \, dx + \sum_{i=1}^{L} \frac{1}{\varepsilon_i} \int_{\varepsilon_i} (\delta u - \delta U_i)(p - P_i) \, dS = - \int_{\Omega} \delta \sigma \nabla u \cdot \nabla p \, dx$$

and the weak formulation for $(p, P) = F(\sigma)(U - U^\delta)$ given by Equation (3.15) in Proposition 3.2.2 for $(v, V) = (\delta u, \delta U)$

$$\int_{\Omega} \sigma \nabla p \cdot \nabla \delta u \, dx + \sum_{i=1}^{L} \frac{1}{\varepsilon_i} \int_{\varepsilon_i} (p - P_i)(\delta u - \delta U_i) \, dS = \langle U - U^\delta, \delta U \rangle_{\mathbb{R}L \times \mathbb{R}L}.$$  

Combining them we get

$$D'(\sigma)[\delta\sigma] = \langle U - U^\delta, \delta U \rangle_{\mathbb{R}L \times \mathbb{R}L} = - \int_{\Omega} \delta \sigma \nabla u \cdot \nabla p \, dx = \langle -\nabla u \cdot \nabla p, \delta\sigma \rangle_{L^q \times L^p}.$$  

We now have to show $-\nabla u \cdot \nabla p \in L^q(\Omega), \; q < \frac{2Q}{Q+2}$. According to Proposition 3.2.7 we have $\mathbf{u}, \mathbf{p} \in W^{1,\tilde{q}}(\Omega)$ for all $2 < \tilde{q} < Q$. We now can estimate by Minkowski’s and Hölder’s inequality with $\frac{2}{\tilde{q}} = \frac{1}{q} + \frac{1}{\tilde{q}}$

$$\|\nabla u \nabla p\|_{L^{\tilde{q}/2}(\Omega)} \leq \sum_{i=1}^{3} \|\partial_{\varepsilon_i} u \partial_{\varepsilon_i} p\|_{L^{\tilde{q}/2}(\Omega)}$$

$$\leq \sum_{i=1}^{3} \|\partial_{\varepsilon_i} u\|_{L^{\tilde{q}}(\Omega)} \|\partial_{\varepsilon_i} p\|_{L^{\tilde{q}}(\Omega)}.$$  

It therefore follows

$$D'(\sigma) = -\nabla u \cdot \nabla p \in L^q(\Omega), \; 1 < q < \frac{Q}{2},$$

where $u = u(\sigma)$ and $p = p(\sigma)$ still depend on $\sigma$. With $Q > 2$ we get $\frac{2Q}{Q+2} < \frac{Q}{2}$ and conclude the proof. 

With this proposition we could rewrite the iteration formula as

$$\delta\sigma^{n+1} = S_{\alpha \tau}[\delta\sigma^n - \tau D'(\sigma^n)] \quad (5.10)$$
but with $D'(\sigma^n) \in L^q(\Omega)$, $q < \frac{2Q}{Q+2}$, we cannot satisfy $\delta\sigma^{n+1} \in L^p(\Omega')$, $p > \frac{2Q}{Q-2}$, which is necessary for the differentiability of the forward operator $F$. Additionally, $\delta\sigma^{n+1}$ could not be compactly supported and $\sigma_0 + \delta\sigma^{n+1} \notin A$ results. Taken together this means that the next iteration cannot be admissible or the gradient does not exist. For it we have to guarantee higher regularity and a compact support on $\delta\sigma^{n+1}$. To solve this problem we use the Sobolev smoothed gradient presented in the next section.

5.2. Step 2: Sobolev smoothed gradient

The problem is that we only ensure $\delta\sigma^{n+1} \in L^q(\Omega)$, $q < \frac{2Q}{Q+2}$, by an update with $D'(\sigma^n) \in L^q(\Omega)$. To satisfy the convergence of the algorithm we need at least a gradient in a space with better regularity. We therefore look for a gradient in the space $H^1_0(\Omega)$. For the two-dimensional case the embedding into $L^p(\Omega)$ with $p < \infty$ holds. The three-dimensional case needs slightly more restriction because the embedding of $H^1_0(\Omega)$ into $L^p(\Omega)$ holds only for $p < 6$. With the differentiability Theorem 3.2.9 we have to assume a sufficiently large $Q$ and therefore a $\lambda$ closer to one such that $\frac{2Q}{Q-2} < p < 6$. If this is satisfied, the differentiability is given and the iteration can proceed. Additionally, it emerged in practice that the gradient $D'(\sigma)$ has unnatural oscillating properties, which can be avoided by using the smoother gradient. This process is also called denoising. We therefore look for a Sobolev smoothed gradient $D'_s(\sigma) \in H^1_0(\Omega)$ defined as

$$D'(\sigma)[\delta\sigma] = \langle D'_s(\sigma), \delta\sigma \rangle_{H^1_0(\Omega)}. \tag{5.11}$$

This approach to Sobolev gradients for weak solutions is presented in more detail in [26]. With the Sobolev embedding Theorem 2.0.3 the space $H^1_0(\Omega)$ is compactly embedded in the $L^p(\Omega)$ with $p < 6$ for a three-dimensional domain $\Omega$. With Equation (5.11) and $\langle \cdot, \cdot \rangle_{H^1_0(\Omega)}$ as a scalar product weighted by $\beta$ we get

$$\langle D'(\sigma), \delta\sigma \rangle_{L^q \times L^p} = \langle D'_s(\sigma), \delta\sigma \rangle_{L^2(\Omega)} + \beta \langle \nabla D'_s(\sigma), \nabla \delta\sigma \rangle_{L^2(\Omega)}$$
which is equivalent to
\[
\int_\Omega D'(\sigma) \delta \sigma \, dx = \int_\Omega D'_s(\sigma) \delta \sigma \, dx + \beta \int_\Omega \nabla D'_s(\sigma) \cdot \nabla \delta \sigma \, dx
\]
\[
= \int_\Omega D'_s(\sigma) \delta \sigma \, dx - \beta \int_\Omega \Delta D'_s(\sigma) \delta \sigma \, dx
\]
\[
= \int_\Omega [-\beta \Delta D'_s(\sigma) + D'_s(\sigma)] \delta \sigma \, dx, \quad \delta \sigma \in H^1_0(\Omega).
\]

For more details about using the weighted scalar product and therefore a restriction to weighted Sobolev spaces we refer to [23]. Therefore \(D'_s(\sigma)\) is the solution of the following Dirichlet boundary problem:
\[
-\beta \Delta D'_s(\sigma) + D'_s(\sigma) = D'(\sigma) \quad \text{in } \Omega,
\]
\[
D'_s(\sigma) = 0 \quad \text{on } \partial \Omega
\]
where \(\beta\) is a scalar for controlling the degree of smoothing. For \(\beta = 1\), in particular, we get \(D'_s(\sigma) \in H^1_0(\Omega)\) and for \(\beta = 0\) we get the gradient in \(L^q(\Omega)\). Using the smoothed gradient is an implicit restriction of the admissible solution to a smoother subset. For theoretical justification we refer to [18]. For computation it is necessary to present the weak formulation of the partial differential equation and the boundary condition.

**Proposition 5.2.1 (Sobolev smoothing equation - weak formulation).**
\(D_s(\sigma) \in H^1_0(\Omega)\) is a weak solution of
\[
-\beta \Delta D'_s(\sigma) + D'_s(\sigma) = D'(\sigma) \quad \text{in } \Omega,
\]
\[
D'_s(\sigma) = 0 \quad \text{on } \partial \Omega
\]
if and only if
\[
\int_\Omega D'(\sigma) v \, dx = \int_\Omega D'_s(\sigma) v \, dx + \beta \int_\Omega \nabla D'_s(\sigma) \cdot \nabla v \, dx, \forall v \in H^1_0(\Omega). \quad (5.14)
\]

**Proof.** First we assume \(D'_s \in H^1_0(\Omega)\) is a solution of the partial differential equation given by (5.12) and (5.13). With the fundamental lemma of variational calculus we get
\[
\int_\Omega D'(\sigma) v \, dx = \int_\Omega [-\beta \Delta D'_s(\sigma) + D'_s(\sigma)] v \, dx, \forall v \in H^1_0(\Omega).
\]
Conversion and applying Green’s formula with the zero boundary of \(v\) yields in equation (5.14).
• For the reverse we assume Equation (5.14) holds. Equation (5.12) follows by applying Green’s formula to \( \int_\Omega \nabla D'_s(\sigma) \cdot \nabla v \, dx \) and \( v \in H^1_0(\Omega) \). The homogeneous boundary condition follows from the choice of test functions.

With this proposition we can rewrite the iteration formula by

\[
\delta \sigma^{n+1} = S_{\alpha \tau}[\delta \sigma^n - \tau D'_s(\sigma^n)]
\]

(5.15)

where admissibility and differentiability are satisfied, especially for \( \beta = 1 \). We have to choose \( \beta \) carefully and sufficiently close to one such that we do not lose differentiability for the next iteration.

5.3. Step 3: Choice of the step size

The step size \( \tau \) can be chosen so it speeds up the algorithm but with the assumption \( s_n = 1 \) we additionally have to satisfy the monotonicity of the sequence \( \delta \sigma^n \), which is not given if \( \tau_n \) is not sufficiently small enough. A compromise between speed and monotonicity is the use of weak monotonicity. The motivation for speeding up is the comparison with the classical Landweber iteration whose slow convergence results from using a constant step size, especially by using a very small step size. We thus use an adaptive selection of the step size to tune the convergence speed, wherefore we consider only the steepest descent operation \( \delta \sigma^n - \tau_n D'_s(\sigma^n) \) of the algorithm. The selection is done by the two-point rule of Barzilai and Borwein [2] which calculates the step size as

\[
\tau_n = \arg \min_{\tau} \| \tau (\delta \sigma^n - \delta \sigma^{n-1}) - (D'_s(\sigma^n) - D'_s(\sigma^{n-1})) \|^2_{H^1(\Omega)}.
\]

The choice of this functional is motivated by the secant equation \( D'(\sigma^n) = D'(\sigma^{n-1}) + B(\delta \sigma^n - \delta \sigma^{n-1}) \) with \( \tau \) times the identity operator as the approximation of the Hessian \( B \) and describes the approximation performance of the last iteration step. This equation does not necessarily have a solution, so it is solved in a least-squares sense. In the one-dimensional real case this procedure implies the secant method. By minimizing the functional we get the following formula for the step size

\[
\tau_n = \frac{\langle \delta \sigma^n - \delta \sigma^{n-1}, D'_s(\sigma^n) - D'_s(\sigma^{n-1}) \rangle_{H^1(\Omega)}}{\langle \delta \sigma^n - \delta \sigma^{n-1}, \delta \sigma^n - \delta \sigma^{n-1} \rangle_{H^1(\Omega)}}.
\]

(5.16)
Another approach, also from Barzilai and Borwein, is that $\tau$ times the identity operator imitates the inverse of the Hessian $B$ over the last step. This results in

$$\tau_n = \arg \min_\tau \| (\delta \sigma^n - \delta \sigma^{n-1}) - \tau(D'_s(\sigma^n) - D'_s(\sigma^{n-1})) \|^2_{H^1(\Omega)}$$

and we therefore get

$$\tau_n = \frac{\langle \delta \sigma^n - \delta \sigma^{n-1}, D'_s(\sigma^n) - D'_s(\sigma^{n-1}) \rangle_{H^1(\Omega)}}{\langle D'_s(\sigma^n) - D'_s(\sigma^{n-1}), D'_s(\sigma^n) - D'_s(\sigma^{n-1}) \rangle_{H^1(\Omega)}}. \quad (5.17)$$

In the implementation we use this step size as an initial guess and it is decreased geometrically until the following variant of Armijo condition applied to the functional $J_\alpha(\sigma)$ is satisfied. More detailed information about Armijo conditions are given, for example, in [15].

$$J_\alpha(\sigma_0 + S_\tau(\delta \sigma^n - \tau D'_s(\sigma^n))) \leq \max_{n-M+1 \leq k \leq n} J_\alpha(\sigma^k) - \tau s \| S_\tau(\delta \sigma^n - \tau D'_s(\sigma^n)) - \delta \sigma^n \|^2_{L^2(\Omega)} \quad (5.18)$$

where $s$ is a small number and $M$ is an integer. This equation is motivated by the Taylor approximation of $J_\alpha(\sigma^{n+1})$ which is given by

$$J_\alpha(\sigma^{n+1}) = J_\alpha(\sigma^n + (\delta \sigma^{n+1} - \delta \sigma^n)) = J_\alpha(\sigma^n) + J'_\alpha(\sigma^n + \xi(\delta \sigma^{n+1} - \delta \sigma^n))[\delta \sigma^{n+1} - \delta \sigma^n]$$

with one $\xi \in (0, 1)$. Remember $\delta \sigma^{n+1}$ depends on $\tau$ and therefore $\xi$ too. We now restrict the gradient to the negative direction, i.e. we have to find a $\tau$ such that $J_\alpha(\sigma^{n+1}) \leq J_\alpha(\sigma^n)$. This yields

$$J_\alpha(\sigma^{n+1}) = J_\alpha(\sigma^n) - |J'_\alpha(\sigma^n + \xi(\delta \sigma^{n+1} - \delta \sigma^n))[\delta \sigma^{n+1} - \delta \sigma^n]|.$$

The calculation of the gradient and the $\xi$ is not advantageous and we therefore take a sufficiently small and fixed $s > 0$ such that

$$s\tau \| \langle \delta \sigma^{n+1} - \delta \sigma^n, \delta \sigma^{n+1} - \delta \sigma^n \rangle_{L^2(\Omega)} \| \leq |J'_\alpha(\sigma^n + \xi(\delta \sigma^{n+1} - \delta \sigma^n))[\delta \sigma^{n+1} - \delta \sigma^n]|,$$

i.e. it is only important that the direction is negative. We thus get the estimate

$$J_\alpha(\sigma^{n+1}) \leq J_\alpha(\sigma^n) - \tau s \| \langle \delta \sigma^{n+1} - \delta \sigma^n, \delta \sigma^{n+1} - \delta \sigma^n \rangle_{L^2(\Omega)} \|.$$

In Figure 3 the grey area above the red line denotes all step sizes which are accepted for the next iteration. This criterion causes slow convergence because it implies
strong monotonicity and we therefore can get very small step sizes. We replace \( J_\alpha(\sigma^n) \) by \( \max_{n-M+1 \leq k \leq n} J_\alpha(\sigma^k) \) with a positive integer \( M \). This is a weaker monotonicity which implies bigger step sizes and we thus can get faster convergence. In Figure 3 this replacement is a parallel shift of the straight line. Combining the above we get Equation (5.18) by adding \( \delta \sigma^{n+1} = S_\tau \alpha (\delta \sigma^n - \tau D'_s(\sigma^n)) \). For further information we refer to [17, 14].

5.4. Step 4: Applying shrinkage operator \( S_\alpha \)

The shrinkage operator \( S_\alpha \) is defined componentwise with respect to an orthonormal basis or a frame \( \{\phi_k\} \) of the space \( H^1_0(\Omega') \). With the definition of \( S_\mu \) from Equation (5.2) and \( x := \delta \sigma^n - \tau D'_s(\sigma^n) \), where we use the Sobolev smoothed gradient \( D'_s(\sigma) \) and the step size \( \tau \) from the Barzilai and Borwein rule with satisfied monotonicity criterion, we get the shrinkage step

\[
S_\tau \alpha (x) = \sum_{i=1}^{\infty} x_i \phi_i
\]

with

\[
x_i := \begin{cases} 
(\|\langle x, \phi_i \rangle_{H^1_0(\Omega')}\| - \tau \alpha) \text{sign}(\langle x, \phi_i \rangle_{H^1_0(\Omega')}), & \text{if } \|\langle x, \phi_i \rangle_{H^1_0(\Omega')}\| > \tau \alpha \\
0, & \text{else.}
\end{cases}
\]
5.5. Step 5: Stopping criterion

The algorithm is terminated if the following holds

\[
\max_{n-\zeta+1 \leq k \leq n} \frac{\|\delta \sigma^{k+1} - \delta \sigma^k\|_{H^1(\Omega)}}{\|\delta \sigma^{k+1}\|_{H^1(\Omega)}} < s_{\text{stop}}.
\]  

(5.21)

This means the algorithm is terminated if the last \(\zeta\) steps have sufficiently small changes. Another possible stopping criterion is \(\tau\) falling below a positive constant which would be a similar criterion.

The whole algorithm for sparsity reconstruction which computes an estimate by iterated soft shrinkage with Sobolev smoothed gradient and step size by Barzilai and Borwein can be seen in Algorithm 5.3.

**Algorithm 5.3 Sparse reconstruction algorithm**

- Set \(\delta \sigma^1 = 0\)

- for \(n = 1, \ldots, N\) do
  - Compute \(\sigma^n = \sigma^0 + \delta \sigma^n\);
  - Step 1: Compute the gradient \(D'(\sigma^n)\);
  - Step 2: Compute the smoothed gradient \(D'_s(\sigma^n)\);
  - Step 3: Determine the step size \(\tau_n\);
  - Update inhomogeneity \(\delta \sigma^{n+1} = \delta \sigma^n - \tau_n D'_s(\sigma^n)\);
  - Step 4: Threshold \(\delta \sigma^{n+1}\) by \(S_{\tau_n, \alpha}(\delta \sigma^{n+1})\);
  - Step 5: Check stopping criterion.

- end for

- output Approximated minimizer \(\delta \sigma\)

- return \(\sigma^0 + \delta \sigma\)
6. Implementation

In this section we discuss the main parts of the implementation of Algorithm 5.3. We therefore present the general Galerkin method and introduce the basic principles for the finite element method. With these tools we describe the whole implementation and give the parts of the algorithm in detail such that a direct implementation is possible. The main sources for the Galerkin approach and the finite element method are [3, 6, 11].

6.1. Galerkin method

The Galerkin method is the bridge between calculus and numerical mathematics because it can be used for existence proofs and to calculate an approximated solution of an operator equation. Especially for weakly formulated partial differential equations we have to find a $u \in H$, where $H$ is a Hilbert space, for which holds

$$B(u, v) = f(v), \quad \forall v \in H,$$

where $B$ is bilinear, symmetric and coercive and $f \in H^*$ is linear and bounded. This is the continuous problem. Mostly the solution of the continuous problem cannot be calculated. We therefore consider an alternative problem on a finite dimensional subset $H_n \subset H$ and calculate an approximated solution $u_n \in H_n$ of the discrete problem

$$B(u_n, v_n) = f(v_n), \quad \forall v_n \in H_n.$$  

This procedure is called the Galerkin approach. When

$$\|u_n\|_H^2 \leq CB(u_n, u_n) = Cf(u_n) \leq C\|u_n\|_H \|f\|_{H^*},$$

it follows the stability of the Galerkin method

$$\|u_n\|_H \leq C\|f\|_{H^*}.$$  

6.2. Finite element method

For the computation and especially to solve the partial differential equation with the finite element method we need a partition of the domain $\Omega$. In our case $\Omega$ is a bounded and connected subset of $\mathbb{R}^3$ and we additionally assume that it is
polyhedral. This motivates us to define a tetrahedral triangulation of the domain. For a better overview we use a general defined index set $I_X := \{1, \ldots, X\}$, $X \in \mathbb{N}$, in this section.

**Definition 6.2.1 (Triangulation).** Let $\Omega \subset \mathbb{R}^3$ be a polyhedral domain. The finite set $\mathcal{T}$ of tetrahedrons $T \subset \overline{\Omega}$ is called a triangulation of $\Omega$ if the following holds:

- $\text{vol}(T) > 0$ for all $T \in \mathcal{T}$,
- $\bigcup_{T \in \mathcal{T}} T = \overline{\Omega}$,
- $\text{Int}(T_1) \cap \text{Int}(T_2) = \emptyset$ for all $T_1, T_2 \in \mathcal{T}$ with $T_1 \neq T_2$.

A triangulation $\mathcal{T}$ is termed **consistent** if the intersection of two simplices $T_1, T_2 \in \mathcal{T}$ is empty or equals with one facet of $T_1$ or $T_2$, cf. Figure 4.

As mentioned in the preceding section we consider finite dimensional subspaces and we therefore choose piecewise polynomial functions, i.e. $f_T \in P_m$ for all $T \in \mathcal{T}$ with

$$P_m := \{f(x, y, z) = \sum_{0 \leq i+j+k \leq m} \alpha_{ijk} x^i y^j z^k\}$$

as the space of polynomials with degree $m$ or less. In this thesis we always consider the $H^1(\Omega)$ and the following theorem from [6] justifies the choice of the approach space with piecewise linear functions.
Theorem 6.2.2. Let $\mathcal{T}$ be a triangulation of the domain $\Omega$. The mapping $v : \Omega \to \mathbb{R}$ satisfies $f|_T \in C^k(T)$ for all $T \in \mathcal{T}$ with $k \geq 1$. $f \in H^k(\Omega)$ then holds if and only if $f \in C^{k-1}(\Omega)$.

In the following we use the discretization parameter $h$ with

$$h := \max_{T \in \mathcal{T}} h(T)$$

where $h(T)$ is the length of the longest edge of the tetrahedron $T$ to describe the mesh amplitude of a triangulation $\mathcal{T}$. The triangulation is then termed by $\mathcal{T}_h$. The following definition of a mesh summarizes all information which we need for the implementation.

Definition 6.2.3 (Mesh). Let $\Omega$ be a connected, bounded and polyhedral domain in $\mathbb{R}^3$. The decomposition of $\Omega$ into $N$ tetrahedrons and $\partial \Omega$ into $M$ triangles is given by a triangulation $\mathcal{T}_h$. The mesh $(p,e,t)$ of the decomposition of $\Omega$ is then defined by the following:

- Set of vertices $p := \{v_i | i \in I_K\}$,
- $t : n \mapsto \{i_1, \ldots, i_4\}$, which maps the index of $\Omega_n$ to the four vertex indices belonging to this tetrahedron,
- $e : m \mapsto \{i_1, \ldots, i_3\}$, which maps the index of $\Gamma_m$ to the three vertex indices belonging to this triangle.

Approach functions on tetrahedrons

We consider tetrahedral elements with polynomial functions of the first degree. We therefore use the four vertices $v_1, \ldots, v_4$ of every tetrahedron as the interpolation nodes for the linear functions, as can be seen in Figure 5. The polynomials which have the value 1 on one node and disappear on the others are the nodal basis. With this choice of interpolation nodes we automatically have the continuity over the whole domain. The restriction of the polynomials to one facet is given uniquely because we have three interpolation nodes and three degrees of freedom for the coefficients of the polynomial. With the same input from the neighbouring element we get the continuity over the facet. The same argument holds for every edge. We thus have global continuity and from Theorem 6.2.2 it follows that the polynomial
Interpolation nodes for polynomial functions of the first degree.

Approximation is in $H^1(\Omega)$. We define the nodal basis elements, also termed *hat functions*, more theoretically below because for the later application to numerical integration rules it is more useful.

**Definition 6.2.4** (Hat function 3-dimensional). Let $\Omega \subset \mathbb{R}^3$ be a connected, bounded and polyhedral domain with the triangulation $T_h = \{\Omega_i\}_{i \in I_N}$ and the corresponding mesh $(p,e,t)$. $v_k \in p$ is arbitrary and $I_k := \{i \in I_N | v_k \in \Omega_i\}$ is the index set of all tetrahedrons to which this vertex belongs. The hat function is piecewise defined on every $\Omega_n$, $n \in I_k$. We thus denote the vertices on one element $\Omega_n$ as $v^1_n, \ldots, v^4_n$ with $v^1_n = (v_k)^T$ and the other vertices are also transposed. With

$$f_n(x) := \begin{vmatrix} 1 & x^T \\ 1 & v^2_n \\ 1 & v^3_n \\ 1 & v^4_n \\ 1 & v^1_n \\ 1 & v^2_n \\ 1 & v^3_n \\ 1 & v^4_n \end{vmatrix}, x \in \Omega_n$$  \hspace{1cm} (6.3)

we get the following definition of the hat function $\phi_k$:

$$\phi_k(x) := \begin{cases} f_i(x), & x \in \Omega_i, \ i \in I_k \\ 0, & \text{else} \end{cases} \hspace{1cm} (6.4)$$

with $\text{supp}(\phi_k) = \bigcup_{i \in I_k} \Omega_i$.  

53
Remark. The function $\phi_k$ is a piecewise linear function on every $\Omega_i$, $i \in I_k$. It is $\phi_k(v_k) = 1$ and on every edge $l$ of the tetrahedron where $v_k \notin l$ it takes the value zero.

We can now define the approach space

$$\tilde{H}^1(\Omega) := \left\{ \sum_{k=1}^{K} \omega_k \phi_k(x) | \omega_i \in \mathbb{R}, i \in I_K \right\}$$

(6.5)

with a given triangulation of $\Omega$ and the corresponding hat functions $\phi_k$. With Theorem 6.2.2 we have $\tilde{H}^1(\Omega) \subset H^1(\Omega)$. In the following description of the implementation we consider all functions to be approximated in $\tilde{H}^1(\Omega)$ as described above. We define the piecewise linear interpolation with respect to the triangulation $T_h$ as follows.

**Definition 6.2.5.** Let $\Omega \subset \mathbb{R}^3$ be a polyhedral bounded and connected domain and let there be a mesh $(p, e, t)$ which depends on the triangulation $T_h$. Then the *approximation operator* $\mathcal{P}_h$ is defined as follows:

$$\mathcal{P}_h f := \sum_{k=1}^{K} f_k \phi_k(x)$$

(6.6)

where $f_k = f(v_k)$, $v_k \in p$ and $\phi_k$ as the hat functions from Definition 6.2.4. The vector $(f_i)_{i \in I_K}$ is termed the *numerical representation* of $f$.

**Remark.** To apply this to the CEM bilinear form $B$ we denote $\tilde{H} := \tilde{H}^1(\Omega) \oplus \mathbb{R}^L$. 

Figure 6: Hat function on one facet $\Gamma$. 

Remark. To apply this to the CEM bilinear form $B$ we denote $\tilde{H} := \tilde{H}^1(\Omega) \oplus \mathbb{R}^L$. 

54
6.3. Forward operator

We can now describe the forward problem $F(\sigma)$ from Definition 3.2.1 with a conductivity $\sigma \in \tilde{H}^1(\Omega)$ which is also interpolated linearly in a numerical way. This is necessary because the bounded calculating capacity means we cannot compute a solution in the whole $H^1(\Omega) \oplus \mathbb{R}^L$. We thus consider the discrete forward operator $\tilde{F}(\sigma) : \Sigma \rightarrow \tilde{H}$ which maps the current pattern to a linearly interpolated potential function and the approximated potentials on the electrodes. For easier notations we denote the approximations $(P_h u, U_h)$ also by $(u, U)$ and define the vector $(u_i)_{i \in I_K}$ as the numerical representation of the corresponding element in $H^1(\Omega)$. With the Galerkin approach presented in Section 6.1 and the weak formulation of the differential equation from Proposition 3.2.2 we get

$$\sum_{i=1}^K u_i \int_{\Omega} \sigma \nabla \phi_i \cdot \nabla v \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (\sum_{i=1}^K u_i \phi_i - U_l) (v - V_l) dS = \langle I, V \rangle_{\mathbb{R}^L \times \mathbb{R}^L} \quad (6.7)$$

for all test functions $(v, V) \in \tilde{H}$ so that we differentiate between two cases:

1. Case: $v = \phi_j, \ j \in I_K$ and $V = 0$.

Thus equation (6.7) becomes

$$\sum_{i=1}^K u_i \int_{\Omega} \sigma \nabla \phi_i \cdot \nabla \phi_j \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \phi_i \phi_j \, dS - \sum_{l=1}^L U_l \frac{1}{z_l} \int_{e_l} \phi_j dS = 0, \ \forall j \in I_K. \quad (6.8)$$

With

$$A := \left( \int_{\Omega} \sigma \nabla \phi_i \cdot \nabla \phi_j \, dx \right)_{i,j \in I_K} \in \mathbb{R}^{K \times K},$$

$$B := \left( \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \phi_i \phi_j \, dS \right)_{i,j \in I_K} \in \mathbb{R}^{K \times K},$$

$$C := \left( -\frac{1}{z_l} \int_{e_l} \phi_j \, dS \right)_{(j,l) \in I_K \times I_L} \in \mathbb{R}^{K \times L},$$

Equation (6.8) is equivalent to

$$\begin{bmatrix} A + B & C \end{bmatrix} \begin{bmatrix} u \\ U \end{bmatrix} = 0. \quad (6.9)$$
2. Case: \( v = 0 \) and \( V = (\delta_{k,l})_{k \in I_L} \) for all \( l \in I_L \), i.e. \( V \) are the basis unit vectors of \( \mathbb{R}^L \).

Therefore Equation (6.7) becomes

\[
- \sum_{i=1}^K u_i \frac{1}{z_k} \int_{e_k} \phi_i \, dS + U_k \frac{1}{z_k} \int_{e_k} 1 \, dS = I_k, \quad \forall k \in I_L. \tag{6.10}
\]

With

\[
C^T = \left( -\frac{1}{z_k} \int_{e_k} \phi_i \, dS \right)_{(k,i) \in I_L \times I_K} \in \mathbb{R}^{L \times K},
\]

\[
D := \left( \delta_{i,k} \frac{1}{z_k} \int_{e_k} 1 \, dS \right)_{i,k \in I_L} \in \mathbb{R}^{L \times L}. \tag{6.11}
\]

Equation (6.10) becomes

\[
\begin{bmatrix}
C^T & D
\end{bmatrix}
\begin{bmatrix}
u \\
U
\end{bmatrix} = I. \tag{6.12}
\]

Combining the two cases yields

\[
\begin{bmatrix}
A + B & C \\
C^T & D
\end{bmatrix}
\begin{bmatrix}
u \\
U
\end{bmatrix} = \begin{bmatrix} 0 \\
I
\end{bmatrix}. \tag{6.13}
\]

as a linear equation system. To guarantee the existence of a unique solution we have to add the following zero mean condition

\[
\sum_{l=1}^L U_l = 0
\]

and then summarizing with Equation (6.13) the linear equation system to solve the forward problem becomes

\[
\begin{bmatrix}
A + B & C \\
C^T & D
\end{bmatrix}
\begin{bmatrix}
u \\
U
\end{bmatrix} = \begin{bmatrix} 0 \\
I
\end{bmatrix}. \tag{6.14}
\]

We therefore have

\[
\hat{F}(\sigma)I = (K_F^{-1}R_I)^T. \tag{6.15}
\]

The calculation of the matrices \( A, B, C \) and \( D \) is not trivial so that we present the procedure below. We assume a given mesh \((p, e, t)\) which is based on the triangulation \( T_h \) of the domain \( \Omega \).
6.3.1. Stiffness matrix $A$

For the matrix $A = (a_{ij})_{i,j \in I}$ we calculate the integral on every tetrahedron $\Omega_k$ where $k \in I_{ij} := \{ n \in I_{N} | \Omega_n \subset \text{supp}(\phi_i,\phi_j) \}$ denotes the indices of tetrahedrons which belong to the support of the integrand, and sum them:

$$a_{ij} = \sum_{k \in I_{ij}} \int_{\Omega_k} \sigma \nabla \phi_i \cdot \nabla \phi_j \, dx.$$

We now consider one fixed tetrahedron $\Omega_k$ and calculate the element stiffness matrix $E^k \in \mathbb{R}^{4 \times 4}$ defined by

$$E^k = (a^k_{ij})_{i,j \in t(k)}.$$

(6.16)

For the following we denote the four vertices of $\Omega_k$ by $v^1, \ldots, v^4$, where $v^i = (x_i, y_i, z_i)$, and the corresponding hat functions by $\psi_1, \ldots, \psi_4$. With

$$h := \begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}$$

(6.17)

and

$$g_1(x, y, z) := \begin{vmatrix} 1 & x & y & z \end{vmatrix}, \ldots, g_4(x, y, z) := \begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}$$

(6.18)

it follows that

$$\psi_i = \frac{g_i}{h}, \quad i = 1, \ldots, 4.$$

(6.19)

Developing the determinant in $g_i$ after the $i$th line yields

$$g_i(x, y, z) = d_{i1}x + d_{i2}y + d_{i3}z + d_{i4}, \quad i = 1, \ldots, 4,$$

(6.20)

where $d_{ij}$ are the corresponding sub-determinants. For example $i = 1$:

$$d_{11} = \begin{vmatrix} 1 & y_2 & z_2 \\ 1 & y_3 & z_3 \\ 1 & y_4 & z_4 \end{vmatrix} ; \quad d_{12} = \begin{vmatrix} 1 & x_2 & z_2 \\ 1 & x_3 & z_3 \\ 1 & x_4 & z_4 \end{vmatrix} ; \quad d_{13} = \begin{vmatrix} 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ 1 & x_4 & y_4 \end{vmatrix} ; \quad d_{14} = \begin{vmatrix} 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}.$$
Therefore the derivative of one hat function $\psi_i$ is

$$\nabla \psi_i = \frac{1}{h} \begin{pmatrix} d_{i1} \\ d_{i2} \\ d_{i3} \end{pmatrix}.$$  \hspace{1cm} (6.21)

With $d := (d_{ij})_{j,i}$, $i = 1, \ldots, 4$, $j = 1, \ldots, 3$ Equation (6.16) becomes

$$E^k = \frac{1}{h^2} d^T d \int_{\Omega_k} \sigma \, dx.$$ \hspace{1cm} (6.22)

With the given conductivity in $\tilde{H}^1(\Omega)$ it can be rewritten as

$$\sigma = \sum_{i=1}^{4} \sigma_i \psi_i$$

where $\sigma_i = \sigma(v^i)$.

**Remark.** In general we can use another arbitrary finite basis set for $\sigma$ but for easier computation and a later update with $\delta \sigma \in \tilde{H}^1(\Omega)$ we use also an approximation in $\tilde{H}^1(\Omega)$.

With a quadrature rule of degree one from [22] we can calculate the integral for $\sigma \in \tilde{H}^1(\Omega)$ exactly by using

$$\int_{\Omega_k} \sigma \, dx = \int_{\Omega_k} \sigma \left( \frac{1}{4} \sum_{j=4}^{4} v^j \right) \, dx = \int_{\Omega_k} \frac{1}{4} \sum_{j=4}^{4} \sigma_i \sum_{j=1}^{4} \psi_i(v^j) \, dx = |\Omega_k| \int_{\Omega} \frac{1}{4} \sum_{i=1}^{4} \sigma_i.$$

Therefore Equation (6.22) becomes

$$E^k = \frac{1}{h^2} d^T d \frac{1}{4} \sum_{i=1}^{4} \sigma_i |\Omega_k|.$$ \hspace{1cm} (6.23)

With the given element stiffness matrix of every tetrahedron we can formulate the calculation of $A$ as follows

$$A = \sum_{n=1}^{N} \tilde{E}^n$$

where $\tilde{E}^n \in \mathbb{R}^{K \times K}$ is a sparse matrix with $\tilde{E}^n_{ij} = E^m_{ij}$ for $i, j \in t(n)$ and zero else. The whole algorithm to calculate the stiffness matrix $A$ is described in Algorithm 6.1.
Remark. The volume of an arbitrary tetrahedron $T$ with the vertices $v^1, \ldots, v^4$ is given by

$$|T| = \frac{1}{6} |[(v^2 - v^1) \times (v^3 - v^1)] \cdot (v^4 - v^1)|.$$  

Remark. Because of the commutative multiplication of $\nabla \phi_i$ and $\nabla \phi_j$ the stiffness matrix is symmetric.

**Algorithm 6.1** Stiffness matrix algorithm

Initialize $A$ as empty matrix;

for $n = 1, \ldots, N$ do

Set $\text{ind} = t(n)$;

Compute subdeterminant matrix $d$;

Compute $h$;

Compute element stiffness matrix $E$ by equation (6.23);

$A(\text{ind, ind}) = A(\text{ind, ind}) + E$;

end for

return Stiffness matrix $A$

6.3.2. Matrix $B$

For the matrix $B \in \mathbb{R}^{K \times K}$ defined by

$$B := \left( \sum_{l=1}^{L} \frac{1}{Z_l} \int_{e_l} \phi_i \phi_j \ dS \right)_{i,j \in I_K}$$

we only have to consider the boundary triangle elements belonging to the electrodes. With $Z_l$ as the set of indices from the triangles $\Gamma_m \subset e_l$ and the assumption that the electrodes are disjoint we can rewrite the definition of $B = (b_{ij})_{i,j \in I_K}$ as

$$b_{ij} = \begin{cases} \frac{1}{Z_l} \int_{e_l} \phi_i \phi_j \ dS, & i,j \in \bigcup_{m \in Z_l} e(m) \text{ and } l \in I_L \\ 0, & \text{else.} \end{cases}$$

On one fixed electrode $e_l$ we calculate the integral on every triangle $\Gamma_k$ where $k \in T_{ij}^l := \{m \in Z_l| \Gamma_m \subset \text{supp}(\phi_i \phi_j)\}$, i.e. $\Gamma_k$ belongs to the support of $\phi_i \phi_j$, and sum them up. This yields

$$b_{ij} = \sum_{k \in T_{ij}^l} \frac{1}{Z_l} \int_{\Gamma_k} \phi_i \phi_j \ dS.$$
For the implementation we consider one fixed triangle $\Gamma_k \subset e_l$ and calculate the element matrix $E^k \in \mathbb{R}^{3 \times 3}$ defined by

$$E^k = (b^k_{ij})_{i,j \in e(k)}.$$ 

To compute the integral we use a two-dimensional quadrature rule of degree two \[10\] which is given in general by

$$\int_T f(x) \, dx = \frac{|T|}{3} \sum_{i=1}^{3} f(\tilde{v}^i)$$

where $T$ is a triangle and $\tilde{v}^i$ are the midpoints of the edges connecting the vertices $v^i$. We denote $\tilde{v}^i$ as the midpoint obverse to the vertex $v^i$. With

$$V := (\phi_i(\tilde{v}^j))_{i,j=1,...,3} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

we get

$$E^k = \frac{|\Gamma_k|}{3z_l} V^T V \quad (6.24)$$

where $|\Gamma_k|$ is the area of the triangle $\Gamma_k$. It is easily computed by using

$$|\Gamma_k| = \frac{\| (v^2 - v^1) \times (v^3 - v^1) \|_2}{2} \quad (6.25)$$

Thus the computation of $B$ becomes

$$B = \sum_{l=1}^{L} \sum_{k \in Z_l} \tilde{E}^k$$

where $\tilde{E}^k \in \mathbb{R}^{K \times K}$ is defined as

$$\tilde{E}^k_{ij} := \begin{cases} E^k_{ij}, & i,j \in e(k) \\ 0, & \text{else.} \end{cases}$$

The complete algorithm to calculate matrix $B$ can be found in Algorithm 6.2.

**Remark.** Because of the commutative multiplication of $\phi_i$ and $\phi_j$ the matrix $B$ is symmetric.
Algorithm 6.2 Matrix B algorithm

Initialize $B$ as empty matrix;

for $l = 1, \ldots, L$ do

for $k \in Z_l$ do

Set $ind = e(k)$;

Compute area $|\Gamma_k|$;

Compute element stiffness matrix $E$;

$B(ind, ind) = B(ind, ind) + E$;

end for

end for

return Matrix $B$

6.3.3. Matrix $C$

The matrix $C \in \mathbb{R}^{K \times L}$ is defined by

$$C := \left( -\frac{1}{z_l} \int_{e_l} \phi_j \, dS \right)_{(j,l) \in I_K \times I_L} = [c_1, \ldots, c_L]$$

with

$$c_l := \sum_{k \in Z_l} -\frac{|\Gamma_k|}{3z_l} c_l^k \in \mathbb{R}^K \quad (6.26)$$

where $Z_l$ is the index set of triangles belonging to $e_l$ and

$$\left(\hat{c}\right)_l^k := \begin{cases} 1, & j \in e(k) \\ 0, & \text{else} \end{cases}$$

Equation (6.26) results from the integration rule in Section 6.3.2.

6.3.4. Matrix $D$

The matrix $D \in \mathbb{R}^{L \times L}$ with

$$D := \left( \delta_{i,k} \frac{1}{z_k} \int_{e_k} 1 \, dS \right)_{i,k \in I_L}$$

can be computed by

$$D_{ik} = \begin{cases} \frac{|e_k|}{z_k}, & i = k \\ 0, & \text{else} \end{cases} \quad (6.27)$$
\[|e_l| = \sum_{k \in \mathbb{Z}} |\Gamma_k|, \quad \forall l \in I_L.\]

The area of one triangle \(\Gamma_k\) can be calculated by Equation (6.25).

**6.4. Implementation Step 1: Gradient \(\tilde{D}'(\sigma)\)**

In Section 5.1 we presented the adjoint problem to calculate the \(L^q\) gradient \(D'(\sigma)\) of the discrepancy term which is given by

\[D'(\sigma) = -\nabla p \cdot \nabla u\]

where \(p \in H^1(\Omega)\) is the solution of

\[
\int_{\Omega} \sigma \nabla p \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (p-P_l)(v-V_l) dS = \langle U(\sigma) - U^\delta, V \rangle_{\mathbb{R}L \times \mathbb{R}L}, \quad \forall (v, V) \in H.
\]

In the following we consider the linearly interpolated approximations \(P_h\sigma, (P_hu, U_h)\) and \((P_hp, P_h)\) and denote them by \(\sigma, (u, U)\) and \((p, P)\). In this case we calculate the approximated gradient \(\tilde{D}(\sigma)\) which is also an element of the \(L^q(\Omega)\) as \(\tilde{H}^1(\Omega) \subset H^1(\Omega)\).

As can be seen, the weak formulation is equivalent to the forward problem. We thus get a numerical solution \((p, P) \in \tilde{H}\) by applying the discrete forward operator \(\tilde{F}(\sigma)\) from Section 6.3 to \((U(\sigma) - U^\delta)\) where \(U(\sigma) = \gamma \tilde{F}(\sigma)I\). This leads to

\[(\tilde{p}(\sigma), \tilde{P}(\sigma)) = \tilde{F}(\sigma)(\gamma \tilde{F}(\sigma)I - U^\delta).\]

The next step is to compute \(\nabla \tilde{p}(\sigma)\) and \(\nabla u(\sigma)\) where the potential function \(u\) is automatically given by evaluating \(\tilde{F}(\sigma)I = (u(\sigma), U(\sigma))\). We therefore calculate the gradient on the centre points of every tetrahedron and approximate the values on the vertices. In general with an \(f \in \tilde{H}^1(\Omega)\), an arbitrary tetrahedron \(\Omega_n\) and the corresponding definition of the hat functions given by Equations (6.17), (6.18) and (6.19) we get

\[
\nabla f_n(c) = \sum_{i=1}^{4} f_i \nabla \psi_i(c)
\]
where \(c\) is the centre point of the tetrahedron with 
\[ c = \frac{1}{4} \sum_{i=1}^{4} v^i, \quad i = 1, \ldots, 4, \]
the four vertices of the tetrahedron. With Equations (6.20) and (6.21) this gives
\[
\nabla f_n(c) = \sum_{i=1}^{4} \frac{f_i}{h} \begin{pmatrix} d_{i1} \\ d_{i2} \\ d_{i3} \end{pmatrix}
\]
which is equivalent to the shorter matrix vector multiplication
\[
\nabla f_n(c) = \frac{1}{h} d^T \tilde{f}
\]
with 
\[ d := (d_{ij})_{j,i}, \quad i = 1, \ldots, 4, \quad j = 1, \ldots, 3, \]
and 
\[ \tilde{f} := (f_1, \ldots, f_4)^T. \]

The function values of \(\nabla f\) on one vertex are then calculated by taking the arithmetic mean of all centre points of the tetrahedrons where the vertex is an element from. This is defined by the mapping
\[
c_2 v : \mathbb{R}^{N \times 3} \to \mathbb{R}^{K \times 3}
\]
\[
f^c \mapsto f^v := \left( \frac{1}{|I_i|} \sum_{k \in I_i} f_k^c \right)_{i \in I_K}
\]
where \(f_k^c\) denotes the \(k\)th line of \(f^c\) and 
\(I_i\) is the index set of all tetrahedrons, of which the vertex \(v^i\) is an element.

**Remark.** If the mesh is locally refined, the distances to the centre points can differ greatly. It is therefore better to use a weighted arithmetic mean like
\[
f^v := \left( \frac{1}{\sum_{k \in I_i} ||v^i - c_k||_2} \sum_{k \in I_i} f_k^c ||v^i - c_k||_2 \right)_{i \in I_K}
\]

Applying this to \(u\) and \(p\) yields
\[
\tilde{D}'(\sigma) = - (u_k^v \cdot (p_k^v)^T)_{k \in I_K} \in \mathbb{R}^K
\]
where \(u_k^v\) and \(p_k^v\) denote the \(k\)th line of \(u^v\) and \(p^v\). For subsequent application to the Sobolev smoothed gradient we use the piecewise linear interpolation of \(\tilde{D}'(\sigma)\) in \(\tilde{H}^1(\Omega)\) and denote it as equal. The algorithm can be found in Algorithm 6.3.
Remark (Using multiple data sets). Let $I^k$ denote different current patterns and $U^k$ are the corresponding measured voltage patterns. The minimized Tikhonov functional then becomes

$$J_\alpha(\sigma) = \sum_k \omega_k \| G(\sigma) - U^k \|_{RL} + \alpha \| \delta \sigma \|_{l1}$$

where $\omega_k$ are weights. Minimizing this functional is nearly equivalent to the algorithm presented with one exception. The gradient calculation of the discrepancy term differs. We therefore have to calculate the gradient $D'_k(\sigma)$ for every voltage pattern $U^k$ and sum them up. This yields

$$D'(\sigma) = \sum_k \omega_k D'_k(\sigma). \quad (6.31)$$

The remaining algorithm proceeds as described before.

Algorithm 6.3 Gradient $\tilde{D}'(\sigma)$ algorithm

- Solve $(u(\sigma), U(\sigma)) = \tilde{F}(\sigma) I$;
- Solve $(p(\sigma), P(\sigma)) = \tilde{F}(\sigma)(U(\sigma) - U^\delta)$;
- Compute $u^c, p^c$ with Equation (6.28);
- Compute $u^v, p^v$ with Equation (6.29);
- Compute $\tilde{D}'(\sigma)$ with Equation (6.30);
- return Gradient $\tilde{D}'(\sigma)$

6.5. Implementation Step 2: Sobolev smoothed gradient $\tilde{D}'_s(\sigma)$

In Section 5.2 we presented the Sobolev smoothed gradient $D'_s(\sigma)$ to guarantee differentiability and admissibility for the next iterate. We therefore have to solve the Dirichlet boundary problem given by Equations (5.12) and (5.13). Here we use the Galerkin approach again to find a solution $\tilde{D}'_s(\sigma) = P_h D'_s(\sigma) \in \tilde{H}^1_0(\Omega)$. From the preceding section we have the $L^q$ gradient $\tilde{D}'(\sigma) \in \tilde{H}^1(\Omega)$. We thus consider the weak formulation of the problem in $\tilde{H}^1(\Omega)$

$$\beta \int_{\Omega} \nabla \tilde{D}'_s(\sigma) \cdot \nabla v \, dx + \int_{\Omega} \tilde{D}'_s(\sigma)v \, dx = \int_{\Omega} \tilde{D}'(\sigma)v \, dx, \quad \forall v \in \tilde{H}^1_0(\Omega). \quad (6.32)$$
With
\[ \tilde{D}'(\sigma) := \sum_{k=1}^{K} s_k \phi_k, \quad \tilde{D}'(\sigma) := \sum_{k=1}^{K} t_k \phi_k, \]
where \( s := (s_k)_{k \in I_K} \) and \( t := (t_k)_{k \in I_K} \) denote the corresponding numerical representations, we can rewrite Equation (6.32) as
\[ K \sum_{i=1}^{K} s_i \left[ \beta \int_{\Omega} \nabla \phi_i \cdot \nabla v \, dx + \int_{\Omega} \phi_i v \, dx \right] = \sum_{i=1}^{K} t_i \int_{\Omega} \phi_i v \, dx, \quad \forall v \in \tilde{H}_0^1(\Omega). \]
This especially holds for all \( v = \phi_j, \ j \in I_K \) and with
\[
A_s := \left( \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx \right)_{i,j \in I_K} \in \mathbb{R}^{K \times K}, \\
B_s := \left( \int_{\Omega} \phi_i \cdot \phi_j \, dx \right)_{i,j \in I_K} \in \mathbb{R}^{K \times K},
\]
we get the following linear equation system
\[ [\beta A_s + B_s]s = B_s t. \quad (6.35) \]
The numerical representation of \( \tilde{D}'(\sigma) \) is then given by \( s = [\beta A_s + B_s]^{-1} B_s t. \)

**Remark.** The test space \( \tilde{H}_0^1(\Omega) \) restricts the linear equation system to a smaller one. It has only to be solved for the basis functions which are not defined on a vertex on the boundary.

The calculation of the matrix \( A_s \) is equivalent to the calculation of the matrix \( A \) in Section 6.3.1 with the assumption \( \sigma \equiv 1 \). The calculation of \( B_s \) is described in the following section.

### 6.5.1. Matrix \( B_s \)

For the implementation of the matrix \( B_s = (b_{ij})_{i,j \in I_K} \) defined by
\[ b_{ij} = \int_{\Omega} \phi_i \phi_j \, dx \]
we calculate the integral on the supported tetrahedrons from \( \phi_i \phi_j \) and sum them up. Since every tetrahedron \( \Omega_m \) is only supported by combinations of four functions \( \phi_i, \ i \in t(m) \), we introduce the element matrix \( E^m = (b^m_{ij})_{i,j=1,...,4} \) with
\[ b^m_{ij} := \int_{\Omega_m} \psi_i \psi_j \, dx \]
where $\psi_i$ are the corresponding functions to $\phi_i$, $i \in t(m)$, as defined in Section 6.3.1. In this case we have to integrate a polynomial function of degree two over a tetrahedron. This requires a quadrature rule of degree two for an exact calculation. We therefore use a rule given in [22] as follows.

For a given function $f$ we want to calculate the integral over one tetrahedron $T$ with the vertices $v^1, \ldots, v^4$. The quadrature rule is

$$\int_T f \, dx = \sum_{i=1}^{4} \omega_i f(\sum_{j=1}^{4} \kappa_{ij} v^j)$$

where $\omega \in \mathbb{R}^4$ and $\kappa \in \mathbb{R}^{4 \times 4}$ denote the weight vector or matrix. For a better overview we calculate the four points $p^1, \ldots, p^4$, on which the function has to be evaluated, using

$$\begin{bmatrix} p^1 \\ p^2 \\ p^3 \\ p^4 \end{bmatrix} = \kappa \cdot \begin{bmatrix} v^1 \\ v^2 \\ v^3 \\ v^4 \end{bmatrix}.$$  

With $T = \Omega_m$, $f = \psi_i \psi_j$ and the matrix $c := (\psi_i(p^j))_{i,j=1,\ldots,4}$ the element matrix becomes

$$E^m = c \cdot \text{diag}(\omega) \cdot c^T.$$  

(6.36)

The calculation of the matrix $B_s$ therefore becomes

$$B_s = \sum_{n=1}^{N} \tilde{E}^n$$

with $\tilde{E}^n \in \mathbb{R}^{K \times K}$ defined as follows

$$\tilde{E}^n_{ij} := \begin{cases} E^m_{ij}, & i,j \in t(n) \\ 0, & \text{else}. \end{cases}$$

**Remark.** The weight vector and matrix are given in [22] by

$$\omega = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix}, \quad \kappa = \begin{pmatrix} 0.58541 & 0.13819 & 0.13819 & 0.13819 \\ 0.13819 & 0.58541 & 0.13819 & 0.13819 \\ 0.13819 & 0.13819 & 0.58541 & 0.13819 \\ 0.13819 & 0.13819 & 0.13819 & 0.58541 \end{pmatrix}.$$
6.6. Implementation Step 3: Choice of the step size

As presented in Section 5.3 we first calculate an initial estimate for the step size using

$$\tau_n = \frac{\langle \delta \sigma^n - \delta \sigma^{n-1}, D_s'(\sigma^n) - D_s'(\sigma^{n-1}) \rangle_{H^1(\Omega)}}{\langle D_s'(\sigma^n) - D_s'(\sigma^{n-1}), D_s'(\sigma^n) - D_s'(\sigma^{n-1}) \rangle_{H^1(\Omega)}}. \tag{6.37}$$

With the linearly interpolated functions in $\tilde{H}^1(\Omega)$ and their numerical representations $s_i$ for $\tilde{D}_s'(\sigma^i)$ and $\delta \sigma^i$ for $\delta \sigma^i$ we consider the $H^1$ scalar product as follows. Let $u, v \in \tilde{H}^1(\Omega)$ and $\tilde{u}, \tilde{v}$ be the corresponding numerical representations. Then the $H^1$ scalar product

$$\langle u, v \rangle_{H^1(\Omega)} = \int_{\Omega} \nabla u \cdot \nabla v \, dx + \int_{\Omega} uv \, dx$$

is equal to the left side of Equation (6.30) with $\beta = 1$. We thus get

$$\langle u, v \rangle_{H^1(\Omega)} = \tilde{u}^T \left[ A_s + B_s \right] \tilde{v}, \quad \tilde{u}, \tilde{v} \in \mathbb{R}^K$$

where the matrices $A_s$ and $B_s$ are defined in Equations (6.33) and (6.34). The computation is also described in Section 6.5. Therefore Equation (6.37) becomes

$$\tau_n = \frac{(\delta c^n - \delta c^{n-1})^T \cdot K_s \cdot (s^n - s^{n-1})}{(s^n - s^{n-1})^T \cdot K_s \cdot (s^n - s^{n-1})}. \tag{6.38}$$

This step size is used as an initial estimate $\tau_0^n$ and is decreased until the weak monotonicity criterion from Equation (5.18) holds for the linearly interpolated functions in $\tilde{H}^1(\Omega)$.

$$J_\alpha(\sigma_0 + \tilde{S}_\tau(\delta \sigma^n - \tau \tilde{D}_s'(\sigma^n))) \leq \max_{n-M+1 \leq k \leq n} J_\alpha(\sigma^k) - \tau s \| \tilde{S}_\tau(\delta \sigma^n - \tau \tilde{D}_s'(\sigma^n)) - \delta \sigma^n \|^2_{L^2(\Omega)}. \tag{6.39}$$

We therefore have to compute the following terms:

1. Term: $J_\alpha(\sigma_0 + \delta \sigma)$.

   The functional is defined as

   $$J_\alpha(\sigma_0 + \delta \sigma) = \frac{1}{2} \| G(\sigma_0 + \delta \sigma) - U^s \|^2_{\mathbb{R}^L} + \alpha \| \delta \sigma \|_{H^1}. $$

67
With the discrete forward operator from Section 6.3 we get \( U = \gamma \tilde{F}(\sigma_0 + \delta \sigma)I \).

The discrepancy term therefore becomes
\[
\frac{1}{2} \| G(\sigma_0 + \delta \sigma) - U^\delta \|_{L^2}^2 = \frac{1}{2} (U - U^\delta)^T \cdot (U - U^\delta).
\]

To calculate the discrete regularization term we can use the discrete \( H^1 \) scalar product as described before and we therefore get
\[
\| \delta \sigma \|_{L^1} = \sum_{j=1}^K | \langle \delta \sigma, \phi_j \rangle_{H^1(\Omega)} | = \| K_s \delta c \|_1.
\]

The discrepancy and regularization term together become
\[
J_\alpha(\sigma_0 + \delta \sigma) = \frac{1}{2} (U - U^\delta)^T \cdot (U - U^\delta) + \alpha \| K_s \delta c \|_1.
\]

2. Term: \( \| \tilde{S}_{\tau_0}(\delta \sigma^n + \tau \tilde{D}'(\sigma^n)) - \delta \sigma^n \|_{L^2(\Omega)}^2 \).

The application of the shrinkage operator \( \tilde{S}_{\tau \sigma} \) can be found in Section 6.7.

With \( x := S_{\tau \sigma}(\delta \sigma^n + \tau D'(\sigma^n)) - \delta \sigma^n \) and \( \tilde{x} \) as the corresponding numerical representation we can rewrite the \( L^2 \) scalar product as follows
\[
\| x \|_{L^2(\Omega)} = | \langle x, x \rangle_{L^2(\Omega)} | = | \tilde{x}^T B_s \tilde{x} |.
\]

with \( B_s \) as defined and computed in Section 6.5.

We can therefore rewrite the monotonicity criterion as follows
\[
\frac{1}{2} (U^{n+1} - U^\delta)^T \cdot (U^{n+1} - U^\delta) \leq \alpha \| K_s \delta c^{n+1} \|_1
\]
\[
\leq \max_{n-M+1 \leq k \leq n} \left( \frac{1}{2} (U^k - U^\delta)^T \cdot (U^k - U^\delta) + \alpha \| K_s \delta c^k \|_1 \right) - \tau \| \tilde{x}^T B_s \tilde{x} \| \quad (6.40)
\]
with \( U^i = \gamma \tilde{F}(\sigma^0 + \delta \sigma^i)I \). Note that \( U^{n+1}, \delta c^{n+1} \) and \( \tilde{x} \) depend on \( \tau \). We now take the \( \tau_n^0 \) as an initial estimate and check the weak monotonicity criterion. If it holds, we accept \( \delta \sigma^{n+1} \). Otherwise we decrease the step size by
\[
\tau_{n+1} = \frac{\tau_n}{\eta} \quad (6.41)
\]
where \( \eta \) defines the grade of decreasing. If the step size is smaller than a lower boundary \( \tau_{\text{stop}} \) we accept the iterate, too. In our implementation we use \( \eta = 2 \), \( \tau_{\text{stop}} = 1 \times 10^{-3} \), \( M = 5 \) and \( s = 1 \times 10^{-10} \).
Remark. • For the first iteration step of the whole algorithm we cannot use the rule developed by Barzilai and Borwein and the monotonicity criterion. Hence we use a fixed step size $\tau = 1$.

• The initial step size $\tau_0^n$ is also bounded from above and below. We only use $\tau_0^n \in [\tau_{\text{min}}, \tau_{\text{max}}]$.

• To avoid unnecessary computations of $J_\alpha(\sigma^k)$ we use a history to get the maximum of the last $M$ functional values.

6.7. Implementation Step 4: Shrinkage operator $\tilde{S}_\alpha$

In general we assume an $f \in \tilde{H}^1(\Omega)$ with its numerical representation $\tilde{f}$. We can then define the discrete shrinkage operator $\tilde{S}_\mu$ with respect to the hat functions as follows

$$\tilde{S}_\mu(f) = \sum_{k=1}^{K} \tilde{f}_k^s \phi_k$$

with

$$\tilde{f}_k^s := \begin{cases} (	ilde{f}_k - \mu)\text{sign}(\tilde{f}_k), & \text{if } |\tilde{f}_k| > \mu \\ 0, & \text{else.} \end{cases} \quad (6.42)$$

As presented in Section 5.4 we must apply the shrinkage operator to $\delta \sigma^n - \tau \tilde{D}_s'(\sigma^n)$ with its corresponding numerical representations.

Remark. In this step it is necessary to include a restriction on the updated conductivity $\delta \sigma^{n+1}$ such that the ellipticity criterion $\lambda \leq \sigma_0 + \delta \sigma^{n+1} \leq \lambda^{-1}$ is satisfied.

Remark. The hat functions are not an orthonormal basis of $\tilde{H}^1(\Omega)$ with respect to the $H^1$ scalar product but it can be shown that they are a frame, i.e. there exist two constants $c_1$ and $c_2$ such that

$$c_1 \|f\|_{H^1(\Omega)}^2 \leq \sum_{k=1}^{K} |\langle f, \phi_k \rangle_{H^1(\Omega)}|^2 \leq c_2 \|f\|_{H^1(\Omega)}^2, \quad \forall f \in \tilde{H}^1(\Omega).$$

We therefore have $\sum_{k=1}^{K} |\langle f, \phi_k \rangle_{H^1(\Omega)}|^2 = |\tilde{f}^T K_s^T K_s \tilde{f}|$ and $\|f\|_{H^1(\Omega)}^2 = |\tilde{f}^T K_s \tilde{f}|$. $K_s$ is a positive definite symmetric matrix and it therefore induces a scalar product. The same holds for $K_s^T K_s$ so that the equivalence of the induced norms follows directly from the finite dimension.
6.8. Implementation Step 5: Stopping criterion

In Section 5.5 we presented the stopping criterion

$$\max_{n-\zeta+1 \leq k \leq n} \frac{\|\delta\sigma^{k+1} - \delta\sigma^k\|_{H^1(\Omega)}}{\|\delta\sigma^{k+1}\|_{H^1(\Omega)}} < s_{\text{stop}}.$$

With the discrete $H^1$ scalar product, all functions are elements in $\tilde{H}^1(\Omega)$ and the numerical representation $\delta c^i$ for $\delta \sigma^i$ we can rewrite the criterion as

$$\max_{n-\zeta+1 \leq k \leq n} \left( \frac{|(\delta c^{k+1} - \delta c^k)^T K_s (\delta c^{k+1} - \delta c^k)|}{|(\delta c^{k+1})^T K_s \delta c^{k+1}|} \right)^{\frac{1}{2}} < s_{\text{stop}}. \quad (6.43)$$

The algorithm is terminated if $s_{\text{stop}} < 1.0 \times 10^{-3}$. 
7. Application

In this section we apply the algorithm presented to simulated and real data. The first subsection deals with reconstructions from simulated data. The object considered imitates a water tank which was used to get the real data for the subsequent section. Additionally we investigate different electrode settings which were not covered by the data set for the real case.

7.1. Simulated data

Here we examine different inclusions and compare the reconstructions which were recorded by using different parameter choices for $\alpha$ and $\beta$. As will be shown later, the reconstructions are not really three-dimensional if the electrodes are continuous in one direction. We therefore consider different electrode sizes and positions.

To reflect the real case we use a cylindrical domain with electrodes attached to the curved surface area. The cylinder has a radius of 1 m and a height of 0.5 m. The first case is that we have 16 electrodes which are 0.5 m high and 0.2 m wide. The centre points are equally spaced. The second case has 32 electrodes subdivided into two rings each with 16 electrodes which are 0.2 m wide and 0.2 m high. The centre points in each ring are also equally spaced. The meshes for the two cases can be seen in Figure 7, where the first one is distributed in 83,098 tetrahedrons (16,456 vertices) and the second one in 81,539 tetrahedrons (16,198 vertices).

We assume that the background conductivity and the contact impedances are
1. Setting

2. Setting

Figure 8: Assumed test settings and the conductivity of the inclusions are as follows:
1. Setting: $\sigma = 0.2 \ \Omega^{-1}m^{-1}$, 2. Setting: $\sigma = 0.6 \ \Omega^{-1}m^{-1}$.

<table>
<thead>
<tr>
<th>Height no. $i$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_i [m]$</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.35</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 1: Heights of the cross sections for simulated data.

given by $\sigma_0 = 1 \ \Omega^{-1}m^{-1}$ and $z_i = 0.1 \ \Omega m^2$. The induced currents for the first case with 16 electrodes are dipole current patterns so that the current on one electrode is $1 \ A$ and $-1 \ A$ on the opposite electrode. For the second case with two rings of 16 electrodes we use the same current patterns for each ring. Current patterns which incorporate both rings were also tested but did not improve the results such that the higher costs are justified.

We consider two different test settings which we want to reconstruct. The first one is a sphere with a 0.15 m radius and the centre point in height 0.18 m and the second setting is two spheres with 0.15 m radius and centre points in height 0.25 m. The test settings, which are approximated in the mesh, and the assumed conductivity of the inclusions can be seen in Figure 8.

The visualization of three-dimensional results is more challenging so that we decided to show a three-dimensional shape of the reconstructed inclusion and the conductivity distribution in cross sections. The height of each cross section can be found in Table 1.

We first compare the two different electrode settings where the regularization parameter $\alpha$ and $\beta$ were selected by visual inspection. The results can be found in Figure 9 where the regularization parameters are $\alpha = 4.5 \times 10^{-2}$ and $\beta = 1.0 \times 10^{-3}$.
for the first mesh and $\alpha = 0.6$ and $\beta = 1.0$ for the second mesh. Only one ring of electrodes does not give us any information about how the conductivity changes as a function of the height. That means we cannot expect a differentiation of the reconstructed conductivity on the height scale, as can be seen in the middle column of Figure 9. Even if we use both rings of electrodes we can only observe a slight improvement in comparison to the predetermined conditions. The results are not satisfying, as can be seen in the right column of Figure 9, which can be explained by different causes. On the one hand the choice of the parameters $\alpha$ and $\beta$ could not have been optimal because it was based upon the principle of trial and error. On the other hand we could not have induced enough heights to achieve better results. To estimate the inclusion in $x$- and $y$-direction we used a lot more electrodes at different locations which gave us more detailed information about its position so that from this point of view the reconstruction gains in accuracy. Since the reconstruction is too big in size and the distribution of the conductivity is too smooth, the choice of parameters requires correction. Besides, the magnitude of the conductivity is underestimated which can be traced back to the fact that the $l_1$-penalty term is weighted too high, i.e. the difference between the background conductivity and estimated conductivity should be small.

In order to achieve a reconstruction which is close to the real case we make modifications which affect the parameters. At first we consider the reconstructions as a function of $\alpha$ with a constant $\beta$. Secondly we examine the same reconstruction with interchanged parameters, i.e. we have a constant $\alpha$ and a reconstruction as a function of $\beta$.

In the first case we observe a better reconstruction of the real case by increasing the parameter $\alpha$, cf. Figure 10. The higher the value we choose for $\alpha$ the sharper the reconstructions are. Yet the choice of $\alpha$ is not arbitrary because, as can be seen in the fourth column in Figure 10, a very high $\alpha$ leads to worse reconstructions of the inclusion. The problem of the incorrect magnitude persists and actually it can be said that the amplitude decreases as $\alpha$ increases.

In the other case we would expect the inclusion to get smoother when we increase $\beta$, which also distorts the magnitude of the conductivity. As Figure 11 shows, however, we cannot observe crucial differences for $\beta$ between zero and one. Using a higher value than one shows that the reconstruction becomes smoother and the amplitude of the conductivity decreases, as can be seen in the right column of Figure 11.
Figure 9: Setting 1: Cross sections from real conductivity (left), reconstructed conductivity on 1st mesh (middle) and reconstructed conductivity on 2nd mesh (right).
Figure 10: Setting 2: Real conductivity and reconstructed conductivity in the 1st mesh with different parameter $\alpha$ and $\beta = 1.0 \times 10^{-3}$ visualized in cross sections in heights $h_1, \ldots, h_5$ from top to bottom.
Figure 11: Setting 2: Real conductivity and reconstructed conductivity in the 1st mesh with different parameter \( \beta \) and \( \alpha = 0.015 \) visualized in cross sections in heights \( h_1, \ldots, h_5 \) from top to bottom.
In summary the results provide an insight into initial applications of the sparse approach to three-dimensional domains and give an idea of necessary future developments.

### 7.2. Real data

The experimental setup and the mesh which is divided into 16,742 tetrahedral elements and 3,911 interpolation nodes can be seen in Figure 12. In this case we used a rough mesh in order to accelerate the computation of the reconstructions so that a multiplicity of combinations of the parameters was possible in order to find the best one. The cylindrical tank was 0.28 m in diameter. Sixteen equally spaced metallic electrodes with 0.025 m in width and 0.07 m in height were attached to the inner surface of the tank. The tank was filled with tap water such that the water level was the same as the height of the electrodes. Objects with different shapes and made of different materials were placed in the tank. We used plastic cylinders, plastic cuboids and hollow steel bars in different combinations as can be seen in Figure 13. All objects were symmetric in height, which is not a real three-dimensional setup. Nevertheless we consider this case as it is comparable with the results from [12, 13]. The height of cross sections can be found in Table 2.

The measurements were conducted with the KIT 2 measurement system [30]. In these experiments different current injections between all adjacent electrode pairs were used. The applied sinusoidal current had an amplitude of $1.0 \times 10^{-3}$ A. One of the current carrying electrodes was the earth electrode at each injection and the potential differences to the other electrodes were measured. The contact impedances
and the background conductivity $\sigma_0$ were reconstructed from an earlier experiment in which the water tank was only filled with tap water. The estimation of these quantities based on a prior experiment with a homogeneous conductivity distribution is discussed in detail in [34]. The estimated background conductivity is $\sigma_0 = 0.02 \, \Omega^{-1}\text{m}^{-1}$ and the contact impedances can be found in Table 2. Figure 14 shows the reconstructions based on the real data and the corresponding parameter values for $\alpha$ and $\beta$ can be found in Table 2. In all cases we obtain a sharp separation of the inclusions from the background. If more than one inclusion appears, they are clearly separated as well. In the experimental setting we used plastic inclusions with a assumed conductivity close to zero and metal inclusions with a conductivity of about $1.0 \times 10^5 \, \Omega^{-1}\text{m}^{-1}$. The position of the inclusions is precisely identified. The conductivity of the plastic inclusions is reconstructed ten times smaller than the conductivity of the tap water. This does not match the real conductivity of plastic which is due to the choice of $\alpha$. The conductivity of the metal inclusions is greatly underestimated. We do not consider the reactive thin layer between the conductor and the tap water which could be the reason for the described instance. Finally, the three-dimensional shapes of the reconstructions can be found in Figure 13.
1. Experiment

2. Experiment

3. Experiment

4. Experiment

Figure 13: Experimental settings (left) and reconstructed features (right).
Figure 14: Reconstructions from real data visualized in cross sections in heights $h_1, \ldots, h_6$ from top to bottom.
Table 2: Parameter choices for $\alpha$ and $\beta$ (left-top), heights of the cross sections for real data (left-bottom) and estimated contact impedances (right).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>$1.0 \times 10^{-3}$</td>
</tr>
<tr>
<td>3</td>
<td>1.5</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>4</td>
<td>1.5</td>
<td>$1.0 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Electrode</th>
<th>$z_i[\Omega m^2]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.71 \times 10^{-2}$</td>
</tr>
<tr>
<td>2</td>
<td>$2.19 \times 10^{-2}$</td>
</tr>
<tr>
<td>3</td>
<td>$1.39 \times 10^{-2}$</td>
</tr>
<tr>
<td>4</td>
<td>$1.20 \times 10^{-2}$</td>
</tr>
<tr>
<td>5</td>
<td>$1.84 \times 10^{-2}$</td>
</tr>
<tr>
<td>6</td>
<td>$0.86 \times 10^{-2}$</td>
</tr>
<tr>
<td>7</td>
<td>$0.35 \times 10^{-2}$</td>
</tr>
<tr>
<td>8</td>
<td>$0.65 \times 10^{-2}$</td>
</tr>
<tr>
<td>9</td>
<td>$1.63 \times 10^{-2}$</td>
</tr>
<tr>
<td>10</td>
<td>$1.14 \times 10^{-2}$</td>
</tr>
<tr>
<td>11</td>
<td>$0.68 \times 10^{-2}$</td>
</tr>
<tr>
<td>12</td>
<td>$1.98 \times 10^{-2}$</td>
</tr>
<tr>
<td>13</td>
<td>$1.45 \times 10^{-2}$</td>
</tr>
<tr>
<td>14</td>
<td>$1.07 \times 10^{-2}$</td>
</tr>
<tr>
<td>15</td>
<td>$0.59 \times 10^{-2}$</td>
</tr>
<tr>
<td>16</td>
<td>$2.77 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 2: Parameter choices for $\alpha$ and $\beta$ (left-top), heights of the cross sections for real data (left-bottom) and estimated contact impedances (right).

8. Summary and outlook

The iterated soft shrinkage approach applied to the complete electrode model has been described in detail such that the three-dimensional implementation is easily understandable. The sparse reconstructions from simulated data show that real three-dimensional results can only be expected if the electrodes are positioned in all three dimensions. The results from real data are consistent with the results for the two-dimensional model presented in [13]. These results are promising but investigations in experimental settings with a three-dimensional electrode distribution need to be done to show the feasibility of the algorithm.

The parameters $\alpha$ and $\beta$ are chosen by visual inspection. This is not optimal and an automatic parameter choice would be preferable. The relations between $\alpha$ and $\beta$ need necessarily to be analysed in detail.

In this thesis the background conductivity and the contact impedances are estimated from a prior experiment with homogeneous conductivity distribution. Such a prior experiment cannot always be done, e.g. when imaging concrete, and we have to consider other solutions. The obvious one is to include the additional parameters.
Another solution would be an alternating algorithm.

Another problem occurs with the extension to the three-dimensional model which was not considered in this thesis. The computation time increases greatly because the degree of freedom in the finite element space is much higher than in two dimensions. The implementation has to be optimized in several ways. The forward operator can be accelerated by including the zero mean condition. We then have a symmetric and positively definite matrix and the linear system can therefore be solved more efficiently. The matrix $A$ has to be updated in each iteration several times. This can be optimized by a partial update only on the elements where the conductivity changed. The pre-computation of the sub-determinants requires more memory but can speed up the algorithm. Parallel computing could be a solution to accelerate the gradient calculation if the memory requirements are reduced. If they are not reduced, the parallelization has no effect.

The optimizations are partially implemented but can be improved in several ways. The experimental evaluation with more three-dimensional data, the automatic parameter choice and the estimation of contact impedances and background conductivity will be subject of future studies.
A. Proofs

A.1. Model

A.1.1. Uniform continuity

The proof techniques for uniform continuity follow [19].

**Theorem A.1.1** (Uniform continuity in $L^p$-norm). Let $\sigma \in \mathcal{A}$, where $\delta \sigma$ is compactly supported in $\Omega$ such that $\sigma + \delta \sigma \in \mathcal{A}$, $Q > 2$ according to Proposition 3.2.7 and $p > \frac{2Q}{Q-2}$. Then $F_\sigma$ is uniformly continuous with respect to $L^p$-norm, i.e.

$$
\|F(\sigma + \delta \sigma)I - F(\sigma)I\|_H \leq C\|\delta \sigma\|_{L^p(\Omega)}, \quad \forall I \in \Sigma \text{ with } \|I\|_2 < k, \ 0 < k < \infty. \quad (A.1)
$$

**Proof.** For easier notation we denote $F(\sigma) = F(\sigma)I$ for an arbitrary $I \in \Sigma$ with $\|I\|_2 < k, \ 0 < k < \infty$. We then obtain the weak formulations for $F(\sigma) = (u, U)$ and $F(\sigma + \delta \sigma) = (w, W)$ by

$$
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) dS = \langle V, I \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \quad \forall (v, V) \in H
$$

and

$$
\int_{\Omega} (\sigma + \delta \sigma) \nabla w \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (w - W_l)(v - V_l) dS = \langle V, I \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \quad \forall (v, V) \in H.
$$

By subtraction and converting we get

$$
\int_{\Omega} \sigma \nabla (u - w) \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} ((u - w) - (U_l - W_l))(v - V_l) dS = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla v \, dx.
$$

(A.2)

By taking $v = u - w \in H^1(\Omega)$ and $V = U - W \in \Sigma$ we get

$$
\int_{\Omega} \sigma (\nabla (u - w))^2 \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} ((u - w) - (U_l - W_l))^2 dS = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla (u - w) \, dx
$$

which is equivalent to

$$
B((u - w, U - W), (u - w, U - W)) = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla (u - w) \, dx.
$$
With the coercivity result for $B$, $\text{supp}(\delta \sigma)$ and applying Hölder’s inequality with $\frac{1}{p} + \frac{1}{q} + \frac{1}{2} = 1$ we get with a constant $c_0 > 0$

$$c_0 \| (u - w, U - W) \|^2_H \leq \int_{\Omega} \delta \sigma \nabla w \cdot \nabla (u - w) \, dx = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla (u - w) \, dx \leq \| \delta \sigma \|_{L^p(\Omega')} \| w \|_{L^q(\Omega')} \| \nabla (u - w) \|_{L^2(\Omega)} \leq \| \delta \sigma \|_{L^p(\Omega')} \| \nabla w \|_{L^q(\Omega')} \| u - w \|_{H^1(\Omega)} \leq \| \delta \sigma \|_{L^p(\Omega')} \| \nabla w \|_{L^q(\Omega')} \| (u - w, U - W) \|_H.$$

With the assumption $p > \frac{2Q}{Q-2}$ we get $2 < q < Q$ and can apply the variant of Meyer’s theorem given by Proposition 3.2.7.

$$\| \nabla w \|_{L^q(\Omega')} \leq \| w \|_{W^{1,q}(\Omega')} \leq C \| w \|_{H^1(\Omega)}.$$

With the coercivity of $B$ and the Cauchy Schwarz’s inequality we get

$$c_0 \|(u, U)\|^2_H \leq B((u, U), (u, U)) = \langle I, U \rangle_{\mathbb{R}^L \times \mathbb{R}^L} \leq \|I\|_2 \|U\|_2 \leq \|I\|_2 \|(u, U)\|_H$$

which implies

$$\|(u, U)\|_H \leq \frac{1}{c_0} \|I\|_2. \tag{A.3}$$

With this and with the assumption $\|I\|_2 < k$ we get

$$\|F(\sigma + \delta \sigma) - F(\sigma)\|_H \leq C \|\delta \sigma\|_{L^p(\Omega')}. \quad \square$$

### A.1.2. Differentiability in $L^p$ sense

The proof techniques for differentiability follow [19].

**Theorem A.1.2** (Differentiability in $L^p$-sense). Let $\sigma \in \mathcal{A}$, where $\delta \sigma$ is compactly supported in $\Omega$ such that $\sigma + \delta \sigma \in \mathcal{A}$, $Q > 2$ according to Proposition 3.2.7 and
\( p > \frac{2q}{q-2} \). With \((u, U) = F(\sigma)I, I \in \Sigma, \|I\|_2 < k, 0 < k < \infty\), we define a bounded linear operator \( F'(\sigma)[\delta\sigma]I = (\delta u, \delta U) \) as the solution to

\[
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{i=1}^{L} \frac{1}{z_i} \int_{e_i} (\delta u - \delta U_i)(v - V_i) \, dS = -\int_{\Omega} \delta \sigma \nabla u \cdot \nabla v \, dx, \quad \forall (v, V) \in H.
\]

Then \( F(\sigma) \) is Fréchet-differentiable in \( \sigma \) with respect to \( L^p \)-norm in the following sense

\[
\frac{\|F(\sigma + \delta\sigma)I - F(\sigma)I - F'(\sigma)[\delta\sigma]I\|_H}{\|\delta\sigma\|_{L^p(\Omega)}} \to 0 \quad \text{as} \quad \|\delta\sigma\|_{L^p(\Omega)} \to 0, \quad \forall I \in \Sigma.
\]

**Proof.** For easier notations we denote \( F(\sigma) = F(\sigma)I \) and \( F'(\sigma)[\delta\sigma] = F'(\sigma)[\delta\sigma]I \) for an arbitrary \( I \in \Sigma \). We first show linearity and boundedness of \( F'(\sigma)[\delta\sigma] \). We thus consider the weak formulation in Equation (3.30) with \((v, V) = (\delta u, \delta U)\), apply Hölder’s inequality with \( \frac{1}{p} + \frac{1}{q} + \frac{1}{2} = 1 \), i.e. \( 2 < q < Q \), and use the estimate \( \|u\|_{W^{1,q}(\Omega)} \leq C\|u\|_{H^1(\Omega)} \) from Proposition 3.2.7.

\[
B((\delta u, \delta U), (\delta u, \delta U)) = -\int_{\Omega} \delta \sigma \nabla u \cdot \nabla \delta u \, dx
\]

\[
\leq \int_{\Omega} \|\delta \sigma\|_{L^p(\Omega)} \|\nabla u\|_{L^q(\Omega)} \|\nabla \delta u\|_{L^2(\Omega)}
\]

\[
\leq \|\delta \sigma\|_{L^p(\Omega)} \|u\|_{W^{1,q}(\Omega)} \|\delta u\|_{H^1(\Omega)}
\]

\[
\leq C\|\delta \sigma\|_{L^p(\Omega)} \|(u, U)\|_H \|(\delta u, \delta U)\|_H
\]

\[
\leq C\|\delta \sigma\|_{L^p(\Omega)} \|I\|_2 \|(\delta u, \delta U)\|_H.
\]

For the last estimate compare Equation (A.3). With the coercivity result for the operator \( B \) we get the boundedness by

\[
\|F(\sigma)[\delta\sigma]\|_H \leq C\|\delta\sigma\|_{L^p(\Omega)}.
\]

The linearity

\[
F'(\sigma)[\delta\sigma_1 + \delta\sigma_2] = F'(\sigma)[\delta\sigma_1] + F'(\sigma)[\delta\sigma_2]
\]

follows directly from the weak formulation in Equation (3.30). To prove the main part of the Fréchet-differentiability we consider the weak formulations of \((u, U) = F(\sigma), (w, W) = F(\sigma + \delta\sigma) \) and \((\delta u, \delta U) = F'(\sigma)[\delta\sigma] \).

For \((u, U)\) and \((w, W)\) the weak formulations are given by Proposition 3.2.2

\[
\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dx + \sum_{i=1}^{L} \frac{1}{z_i} \int_{e_i} (u - U_i)(v - V_i) \, dS = \langle V, I \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \quad \forall (v, V) \in H
\]
and

\[ \int_{\Omega} (\sigma + \delta \sigma) \nabla w \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{\epsilon_l} (w - W_l)(v - V_l) \, dS = \langle V, I \rangle_{\mathbb{R}^L \times \mathbb{R}^L}, \quad \forall (v, V) \in H. \]

For \((\delta u, \delta U)\) the weak formulation is given by Equation (3.30)

\[ \int_{\Omega} \sigma \nabla \delta u \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{\epsilon_l} (\delta u - \delta U_l)(v - V_l) \, dS = -\int_{\Omega} \delta \sigma \nabla u \cdot \nabla v \, dx, \quad \forall (v, V) \in H. \]

Subtracting, converting and defining \((p, P) := (w - u - \delta u, W - U - \delta U)\) yields

\[ \int_{\Omega} \sigma \nabla p \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{\epsilon_l} (p - P_l)(v - V_l) \, dS = \int_{\Omega} \delta \sigma \nabla (u - w) \cdot \nabla v \, dx. \]

We now apply \((v, V) = (p, P)\) and get

\[ B((p, P), (p, P)) = \int_{\Omega} \delta \sigma \nabla (u - w) \cdot \nabla p \, dx. \]

With \(\frac{1}{p} + \frac{1}{q} + \frac{1}{2} = 1\) we apply Hölder’s inequality and restrict the integral to the support of \(\delta \sigma\)

\[ \int_{\Omega} \delta \sigma \nabla (u - w) \cdot \nabla p \, dx = \int_{\Omega} \delta \sigma \nabla (u - w) \cdot \nabla p \, dx \leq \|\delta \sigma\|_{L^p(\Omega)} \|\nabla (u - w)\|_{L^q(\Omega)} \|\nabla p\|_{L^2(\Omega)} \leq \|\delta \sigma\|_{L^p(\Omega)} \|\nabla (u - w)\|_{W^{1,q}(\Omega)} \|(p, P)\|_H. \]

Last step is to get an estimate for \(\|(u - w)\|_{W^{1,q}(\Omega')}\). We therefore consider the weak formulation given by Equation (A.2)

\[ \int_{\Omega} \sigma \nabla (u - w) \cdot \nabla v \, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{\epsilon_l} ((u - w) - (U_l - W_l))(v - V_l) \, dS = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla v \, dx, \]

for all \((v, V) \in H\). Assuming \(v \in C_0^\infty(\Omega)\) and \(V = 0\) we get

\[ \int_{\Omega} \sigma \nabla (u - w) \cdot \nabla v \, dx = \int_{\Omega} \delta \sigma \nabla w \cdot \nabla v \, dx, \quad \forall v \in C_0^\infty(\Omega), \]

which, by applying Greens formula, is equivalent to

\[ \int_{\Omega} \nabla \cdot (\sigma \nabla (u - w)) \, v \, dx = \int_{\Omega} \nabla \cdot (\delta \sigma \nabla w) \, v \, dx, \quad \forall v \in C_0^\infty(\Omega). \]
With the fundamental lemma of variational calculus \((u - w)\) is a weak solution of
\[
\nabla \cdot (\sigma \nabla (u - w)) = \nabla \cdot (\delta \sigma \nabla w) \text{ in } \Omega.
\]

Together with the assumption \(p > \frac{2Q}{Q-2}\) and \(\frac{1}{p} + \frac{1}{q} + \frac{1}{2} = 1\), which implies \(2 < q < Q\), we get with the variant of Meyer’s theorem by Proposition 3.2.7
\[
\| (u - w) \|_{W^{1,q}(\Omega')} \leq C(\| (u - w) \|_{H^1(\Omega)} + \| \delta \sigma \nabla w \|_{L^q(\Omega)}).
\]

We can estimate the first summand with the uniform continuity Theorem 3.2.8 by
\[
\| (u - w) \|_{H^1(\Omega)} \leq C\| \delta \sigma \|_{L^p(\Omega')}.
\]

The second summand needs more attention. We take an \(r\) with \(q < r < Q\) such that \(g \geq p\) with \(\frac{1}{r} + \frac{1}{g} = \frac{1}{q}\). Applying Hölder’s inequality and restricting to the support of \(\delta \sigma\) yields
\[
\| \delta \sigma \nabla w \|_{L^p(\Omega)} \leq \| \delta \sigma \|_{L^g(\Omega')} \| \nabla w \|_{L^r(\Omega')} \leq \| \delta \sigma \|_{L^g(\Omega')} \| w \|_{W^{1,r}(\Omega)}.
\]

Again with the variant of Meyer’s theorem we get
\[
\| w \|_{W^{1,r}(\Omega)} \leq C\| w \|_{H^1(\Omega)} \leq C\| I \|_2.
\]

All together with the coercivity of \(B\) we have
\[
\| (p, P) \|_H \leq C\| \delta \sigma \|_{L^p(\Omega')} \| (u - w) \|_{W^{1,q}(\Omega')} \\
\leq C\| \delta \sigma \|_{L^p(\Omega')} (\| \delta \sigma \|_{L^p(\Omega')} + \| \delta \sigma \|_{L^q(\Omega')}).
\]

Conversion yields
\[
\frac{\| F(\sigma + \delta \sigma) - F(\sigma) - F'(\sigma)[\delta \sigma] \|_H}{\| \delta \sigma \|_{L^p(\Omega')}} \leq C(\| \delta \sigma \|_{L^p(\Omega')} + \| \delta \sigma \|_{L^q(\Omega')}).
\]

With \(g \leq p\) the following holds and concludes the proof.
\[
\| \delta \sigma \|_{L^p(\Omega')} \to 0 \quad \text{as} \quad \| \delta \sigma \|_{L^p(\Omega')} \to 0.
\]
List of Figures

1. Definition of a metal electrode with thin layer. ........................................ 14
2. Examples for experimental settings. ......................................................... 16
3. Armijo’s criterion, acceptance area (gray area above the red line),
   acceptance interval \((0, \tau^\ast]\). .......................................................... 48
4. Consistency of a triangulation. ................................................................. 51
5. Interpolation nodes for polynomial functions of the first degree. .............. 53
6. Hat function on one facet \(\Gamma\). ............................................................. 54
7. Triangulations of a cylinder with 16 rectangular electrodes (left) and 32 rectangular electrodes arranged in two different heights (right). .... 71
8. Assumed test settings and the conductivity of the inclusions are as
   follows: 1. Setting: \(\sigma = 0.2 \ \Omega^{-1} \text{m}^{-1}\), 2. Setting: \(\sigma = 0.6 \ \Omega^{-1} \text{m}^{-1}\). . 72
9. Setting 1: Cross sections from real conductivity (left), reconstructed
   conductivity on 1st mesh (middle) and reconstructed conductivity on
   2nd mesh (right). .................................................................................. 74
10. Setting 2: Real conductivity and reconstructed conductivity in the
    1st mesh with different parameter \(\alpha\) and \(\beta = 1.0 \times 10^{-3}\) visualized in
    cross sections in heights \(h_1, \ldots, h_5\) from top to bottom. .................... 75
11. Setting 2: Real conductivity and reconstructed conductivity in the
    1st mesh with different parameter \(\beta\) and \(\alpha = 0.015\) visualized in
    cross sections in heights \(h_1, \ldots, h_5\) from top to bottom. .................... 76
12. Experimental setup (left) and finite element mesh (right). ...................... 77
13. Experimental settings (left) and reconstructed features (right). ............. 79
14. Reconstructions from real data visualized in cross sections in heights
    \(h_1, \ldots, h_6\) from top to bottom. ....................................................... 80
References


