On the reformulation of the classical Stefan problem as a shape optimization problem

Rahel Brügger, Helmut Harbrecht

Departement Mathematik und Informatik
Fachbereich Mathematik
Universität Basel
CH-4051 Basel

Preprint No. 2021-08
April 2021

dmi.unibas.ch

ON THE REFORMULATION OF THE CLASSICAL STEFAN PROBLEM AS A SHAPE OPTIMIZATION PROBLEM

RAHEL C. BRÜGGER* AND HELMUT HARBRECHT*

Abstract. This article is concerned with the multi-dimensional one-phase Stefan problem, which belongs to the class of moving boundary problems. We suggest to reformulate the classical Stefan problem as a shape optimization problem, consisting of an objective functional for the moving boundary and a partial differential equation corresponding to a heat type equation. Minimizing the objective functional subject to the differential equation under consideration is equivalent to solving the Stefan problem. In order to apply gradient-based optimization algorithms, we analytically compute the shape gradient of the objective functional. A numerical example justifies our approach.

Key words. Shape optimization, Stefan problem, moving boundary problem, heat equation, space-time tube derivative

1. Introduction. The classical Stefan problem goes back to J. Stefan in 1889, who studied the formation of ice in the polar sea, see [36]. He stated that this growth problem is connected with heat type problems, where not only the temperature, which is the solution of a differential equation, is unknown, but also the position of the surface moves within time and thus is part of the problem.

This specific problem belongs to the class of the *moving boundary problems*. In general, such problems contain time-dependent boundaries which are unknown and depend on time and spatial variables. Moving boundary problems are also called Stefan problems. This is in contrast to *free boundary problems*. The latter also contain boundaries which are unknown beforehand, but these boundaries are in steady-state and are therefore not dependent on the time, cf. [5].

There exists a wide variety of literature on Stefan problems, see for example [16, 18, 28, 34, 37] and the references therein. This literature mostly treats the analysis of the Stefan problem. Stefan problems find their application, for example, in the modelling of phase transitions, chemical reactions, fluid flow in porous medium or melting of ice, compare [5].

To solve a Stefan problem numerically, one encounters the problem of treating the moving boundary. In [5], different numerical methods are explained: front-tracking methods, front-fixing methods and fixed-domain methods. Front-tracking methods compute the position of the moving boundary at every time step. Front-fixing methods try to fix the front by choosing a good spatial coordinate system. In fixed-domain methods, the problem is reformulated, for example, by the means of the enthalpy method such that the position of the boundary appears as a feature of the solution, see [5].

In the present article, we choose to solve the one-phase Stefan problem by using the tools of shape optimization introduced in [26] for the one-dimensional Stefan problem. More precisely, we reformulate the classical one-phase Stefan problem in multiple dimensions as a shape optimization

*Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel, Switzerland, {ra.bruegger,helmut.harbrecht}@unibas.ch

problem by introducing an objective functional and a parabolic state equation. The parabolic state equation is again the Stefan problem, but lacking the Stefan condition on the moving boundary. The objective functional is chosen such that it is minimal if the Stefan condition is satisfied. Therefore, the goal is to minimize this objective functional over all admissible surfaces. In order to apply gradient-based optimization algorithms, we compute the shape gradient of the objective functional by using the adjoint method, which is known to reduce the computational effort.

The remainder of this article is organized as follows. In Section 2, we introduce the classical one-phase Stefan problem, which we will then rewrite as a shape optimization problem. Section 3 is dedicated to the computation of the shape derivative of the objective functional. As the main focus of the article is the analytical computation of the shape gradient, the numerical example given in Section 4 serves as a proof of concept for our theoretical findings. In Section 5, we give some concluding remarks.

2. Problem formulation.

2.1. Classical one-phase Stefan problem. Let us consider the classical one-phase Stefan problem as described in [19]. This specific Stefan problem models the evolution of the solid-liquid phase interface. Thus, for every point of time $t \in [0, T]$, we have a time-dependent spatial domain which we denote by $\Omega_t \subset \mathbb{R}^d$, $d \geq 2$. This spatial domain has a time-dependent spatial boundary $\Gamma_t := \partial\Omega_t$. The setup is illustrated in Figure 2.1 for two spatial dimensions plus the temporal dimension. By setting

$$Q_T = \bigcup_{0<t<T} \left(\{t\} \times \Omega_t\right),$$

we obtain the space-time non-cylindrical domain (also called tube) with lateral boundary

$$\Sigma_T = \bigcup_{0<t<T} \left(\{t\} \times \Gamma_t\right).$$

We assume that the whole problem can be formulated in the hold-all domain $D \times (0, T)$.

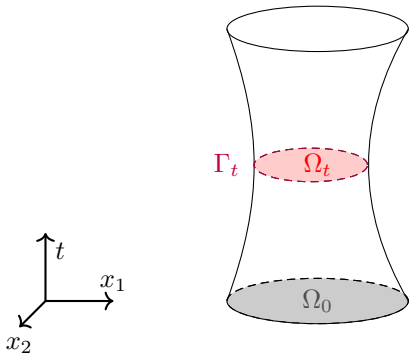


Figure 2.1: Setup of the Stefan problem.

For the formulation of the one-phase Stefan problem, we follow [19, 20]. The temperature $u(t, \mathbf{x})$

of the liquid in Ω_t is thus described by the partial differential equation

$$\partial_t u - \Delta u = 0 \qquad \text{in } \Omega_t, \tag{2.1}$$

$$\langle \mathbf{V}, \mathbf{n} \rangle = -\frac{\partial u}{\partial \mathbf{n}} \qquad \text{on } \Gamma_t, \tag{2.2}$$

$$u = 0 \qquad \text{on } \Gamma_t, \tag{2.3}$$

$$u(0, \cdot) = u_0, \qquad \text{in } \Omega_0 = \Omega. \tag{2.4}$$

The domain Ω in (2.4) is the initial shape of the liquid phase while condition (2.2) is called the Stefan condition [20]. It comes from the movement of the phase interface, see [40, pg. 387]. The Stefan condition expresses that the normal velocity $\langle \mathbf{V}, \mathbf{n} \rangle$ of the surface Γ_t equals minus the normal derivative of u at the boundary. We prescribe the initial position of the interface and the initial temperature distribution to make the problem meaningful. From this Stefan problem, we can see that the liquid freezes at zero temperature, cf. [19]. Notice that the one-phase Stefan problem is actually also a two-phase Stefan problem, but the temperature is only unknown in one region, while it is vanishing in the other region,compare [40].

The domain Ω_t, thus the region which contains the liquid phase, is characterized by $\{\mathbf{x} \in \mathbb{R}^d : u(t, \mathbf{x}) > 0\}$ if we choose $u_0 > 0$ in Ω. Therefore, u can be interpreted as a level set function. Due to (2.1), the parabolic Hopf lemma (see e.g. [15] for some remarks) implies $\partial u / \partial \mathbf{n} < 0$ on Γ_t for $t > 0$. Therefore, we obtain the so-called *Rayleigh-Taylor sign condition*

$$-\frac{\partial u_0}{\partial \mathbf{n}} \geq \lambda > 0 \text{ on } \Gamma_0,$$

which ensures the nondegeneracy in accordance with [19].

2.2. Notation. Since we will switch back and forth between spatial and space-time considerations depending on what is more useful for the task at hand, we introduce some notation in this section to clarify the difference between the two.

For every point of time t we denote the spatial unit normal by $\mathbf{n} = \mathbf{n}_t$, which is thus normal to Ω_t. The time-space unit normal is denoted by $\boldsymbol{\nu}$. Moreover, by ∇, we denote the spatial gradient

$$\nabla = \left[\tfrac{\partial}{\partial x_1}, \tfrac{\partial}{\partial x_2}, \dots, \tfrac{\partial}{\partial x_d} \right]^\mathsf{T},$$

while $\vec{\nabla}$ denotes the time-space nabla operator

$$\vec{\nabla} = \left[\tfrac{\partial}{\partial t}, \tfrac{\partial}{\partial x_1}, \tfrac{\partial}{\partial x_2}, \dots, \tfrac{\partial}{\partial x_d} \right]^\mathsf{T}.$$

Notice that for a time-space vector, the first entry always corresponds to the time component and the subsequent entries correspond to the spatial components. Thus, for the time-space normal, the time component is denoted by ν_1.

Moreover, we introduce the tangential gradient and denote it by ∇_Γ for space and $\vec{\nabla}_\Sigma$ for time-space. Accordingly we denote the tangential divergence (see Definition 3.4). The Jacobian matrix of a field $\mathbf{Z}$ is denoted by $\mathrm{D}\mathbf{Z}$ in space and for a time-space vector field $\vec{\mathbf{Z}}$ by $\vec{D}\vec{\mathbf{Z}}$ in time-space.

2.3. Generation of the tube. In order to generate a tube, we can adopt two different points of view. For both of them, let us assume that we have a spatial domain Ω_0. One can generate a tube Q_T by mapping this domain Ω_0 to a spatial domain Ω_t for every point of time t, see also Figure 2.1.

On the one hand, this can be done by considering a velocity field $\mathbf{V}$ and associate to it the solution $\mathbf{T}(t,\cdot) : \mathbf{x} \mapsto \mathbf{x}_t = \mathbf{T}(t,\mathbf{x})$ of the differential equation [42, pg. 6]

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{T}(t,\mathbf{x}) &= \mathbf{V}\big(t,\mathbf{T}(t,\mathbf{x})\big) \qquad \text{in } (0,T)\times\Omega_0, \\ \mathbf{T}(0,\mathbf{x}) &= \mathbf{x} \qquad\qquad\qquad\quad \text{in } \Omega_0. \end{aligned} \tag{2.5}$$

The map $\mathbf{T}(t,\cdot)$ thus describes the pathline of an individual particle being exposed to the velocity field $\mathbf{V}$. By setting $\Omega_t = \mathbf{T}(t,\Omega_0)$, we generate a tube Q_T (see Figure 2.2).

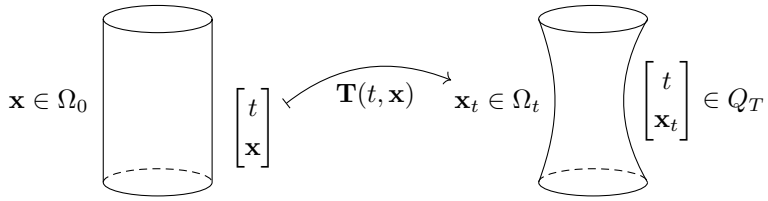


Figure 2.2: Generation of the tube by the mapping $\mathbf{T}(t,\cdot)$ induced by the velocity field $\mathbf{V}$.

On the other hand, we can assume to have a smooth C^2-diffeomorphism $\boldsymbol{\kappa}$ for every point of time t, which maps the initial domain Ω_0 onto the time-dependent domain Ω_t. In accordance with [29], we write

$$\boldsymbol{\kappa} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d, \quad (t,\mathbf{x}) \mapsto \boldsymbol{\kappa}(t,\mathbf{x}) \tag{2.6}$$

to emphasize the dependence of the mapping $\boldsymbol{\kappa}$ on the time, where we have $\boldsymbol{\kappa}(t,\Omega_0) = \Omega_t$. Here, $\boldsymbol{\kappa} \in C^2\big([0,T]\times\mathbb{R}^d\big)$ and, as in [23, pg. 826], we assume the uniformity condition

$$\|\boldsymbol{\kappa}(t,\mathbf{x})\|_{C^2([0,T]\times\mathbb{R}^d;\mathbb{R}^d)}, \|\boldsymbol{\kappa}(t,\mathbf{x})^{-1}\|_{C^2([0,T]\times\mathbb{R}^d;\mathbb{R}^d)} \leq C_{\boldsymbol{\kappa}} \tag{2.7}$$

for some constant $C_{\boldsymbol{\kappa}} \in (0,\infty)$.

[6, Theorem 2.1] and [8, Theorem 2.1] state that under certain assumptions it is equivalent to consider a family of velocity fields $\{\mathbf{V}(t)\}$ or a family of transformations $\{\mathbf{T}(t)\}$. Notice that the two points of view are connected by

$$\mathbf{V} = \partial_t\boldsymbol{\kappa} \circ \boldsymbol{\kappa}^{-1}.$$

In the following, we will adapt the latter parametrization point of view. Therefore, to reduce the technical level of the ensuing discussion, we assume that Ω_0 has C^2-smooth boundary which implies that the boundary of Ω_t has the same regularity.

REMARK 2.1. *Notice that, due to the uniformity condition* (2.7)*, we have as in [23]*

$$0 < \underline{\sigma} \leq \min\{\sigma(\mathrm{D}\boldsymbol{\kappa})\} \leq \max\{\sigma(\mathrm{D}\boldsymbol{\kappa})\} \leq \overline{\sigma} < \infty,$$

where $\sigma(.)$ *denote the singular values and* $\mathrm{D}\boldsymbol{\kappa}$ *denotes the Jacobian matrix of* $\boldsymbol{\kappa}$*. Moreover, as in [23, Remark 1, pg. 827], we assume* $\det(\mathrm{D}\boldsymbol{\kappa})$ *to be positive. The smoothness of the mapping also implies that the time derivative* $\partial_t\boldsymbol{\kappa}$ *is uniformly bounded.*

2.4. Rewriting the Stefan condition. Next, we intend to rewrite the Stefan condition (2.2) into a form, which will be more useful for the reformulation of the Stefan problem as a shape optimization problem. To this end, we first consider the spatial normal and the time-space normal.

Since u can be interpreted as a level set function, applying [32, Formula (1.2) pg. 9] implies that the outward pointing normal can be expressed as

$$\mathbf{n} = -\frac{\nabla u}{\|\nabla u\|} \tag{2.8}$$

provided that $\nabla u \neq \mathbf{0}$. Notice that a priori the normal could have a plus or a minus sign. Taking the scalar product of (2.8) with $\mathbf{n}$ yields

$$1 = -\frac{1}{\|\nabla u\|}\langle \nabla u, \mathbf{n}\rangle. \tag{2.9}$$

Due to the parabolic Hopf lemma, we have $\frac{\partial u}{\partial \mathbf{n}} < 0$ on Γ_t and, therefore, the minus sign is the correct sign. Hence, from (2.9), we can directly infer that the following lemma.

LEMMA 2.2. *It holds*

$$-\frac{\partial u}{\partial \mathbf{n}} = \|\nabla u\| \text{ on } \Gamma_t.$$

The time-space normal can be written as

$$\boldsymbol{\nu} = \frac{1}{\sqrt{1+v_\nu^2}}\begin{bmatrix} v_\nu \\ \mathbf{n} \end{bmatrix} \tag{2.10}$$

for some appropriate $v_\nu \in \mathbb{R}$, which will be determined in Lemma 2.4. Lemma 2.4 is a fundamental property of the tube Q_T, which then allows us to rewrite the Stefan condition in a form, which is computable in our numerical setting. For its proof, we need the following representation of the velocity field $\mathbf{V}$ in normal direction, which can be found in [28, Chapter II, pg. 37], see also [40, pg. 387].

LEMMA 2.3. *If $\nabla u \neq \mathbf{0}$, then it holds*

$$\langle \mathbf{V}, \mathbf{n}\rangle = \frac{\partial_t u}{\|\nabla u\|}.$$

Proof. Since u can be interpreted as a level set function, the interface is evolved according to [32, Formula (3.2), pg. 26] by the convection equation

$$\partial_t u + \langle \mathbf{V}, \nabla u\rangle = 0, \tag{2.11}$$

where $\mathbf{V}$ describes the velocity at every point of the implicit surface. In view of (2.8), we can rewrite this expression to arrive at the claim. ∎

According to [12, 13, 14], the vector field $\mathbf{V}$ belonging to the tube Q_T has to satisfy the following fundamental condition in order to generate the tube. Notice that for a given specific parametrization, this is also stated in [38].

LEMMA 2.4. *For the vector field $\mathbf{V}$, which generates the tube Q_T, it holds*

$$\langle \mathbf{V}, \mathbf{n}\rangle = -v_\nu,$$

where v_ν is the temporal component of the unnormalized time-space normal, see (2.10).

Proof. We prove this claim by assuming that the tube is represented by a level set function ϕ. Having a level set function, the time-space normal can be computed as

$$\frac{\vec{\nabla}\phi}{\|\vec{\nabla}\phi\|} = \frac{1}{\|\vec{\nabla}\phi\|}\begin{bmatrix}\partial_t\phi\\ \nabla\phi\end{bmatrix},$$

where $\vec{\nabla}$ denotes the time-space gradient. We can use the level set convection equation (2.11) to rewrite $\partial_t\phi$, because we can interpret $\phi(t,\cdot)$ for every point of time t as the level set function of the spatial domain. Therefore, we can compute the spatial normal as $\mathbf{n} = \frac{\nabla\phi}{\|\nabla\phi\|}$. Thus, we obtain

$$\boldsymbol{\nu} = \frac{\|\nabla\phi\|}{\|\vec{\nabla}\phi\|}\begin{bmatrix}-\langle\mathbf{V},\mathbf{n}\rangle\\ \mathbf{n}\end{bmatrix} = \frac{1}{\sqrt{\langle\mathbf{V},\mathbf{n}\rangle^2+1}}\begin{bmatrix}-\langle\mathbf{V},\mathbf{n}\rangle\\ \mathbf{n}\end{bmatrix}.$$

∎

We can now express the Stefan condition by means of the geometric quantity $\boldsymbol{\nu}$, which is stated in the following lemma.

LEMMA 2.5. *The left hand side of the Stefan condition* (2.2) *can be expressed as*

$$\langle\mathbf{V},\mathbf{n}\rangle = -\frac{\nu_1}{\sqrt{1-\nu_1^2}},$$

where ν_1 denotes the first entry of the normalized time-space normal $\boldsymbol{\nu}$.

Proof. From the representation (2.10), we infer that

$$\nu_1 = \frac{v_\nu}{\sqrt{1+v_\nu^2}}.$$

Taking the square and multiplying with the denominator gives

$$\nu_1^2(1+v_\nu^2) = v_\nu^2.$$

This expression can be solved for v_ν^2 by writing

$$\nu_1^2 = v_\nu^2(1-\nu_1^2),$$

and thus

$$v_\nu = \pm\frac{\nu_1}{\sqrt{1-\nu_1^2}}.$$

The correct sign is the plus sign, because ν_1 and v_ν have the same sign. Employing Lemma 2.4 yields finally the claim. ∎

REMARK 2.6. *Using Lemmata 2.2, 2.3, and 2.5, we immediately arrive at*

$$\partial_t u = \frac{\nu_1}{\sqrt{1-\nu_1^2}}\frac{\partial u}{\partial\mathbf{n}}. \tag{2.12}$$

Moreover, from (2.10)*, Lemma 2.4, and Lemma 2.5, we obtain*

$$\boldsymbol{\nu} = \begin{bmatrix}\dfrac{\nu_1}{\sqrt{1-\nu_1^2}}\,\mathbf{n}\end{bmatrix}. \tag{2.13}$$

2.5. Reformulation as a shape optimization problem. We are now in the position to reformulate the Stefan problem (2.1) to (2.4) as a shape optimization problem. To that end, we consider the reduced Stefan problem

$$\begin{aligned} \partial_t u - \Delta u &= 0 && \text{in } \Omega_t, \\ u &= 0 && \text{on } \Gamma_t, \\ u(0,\cdot) &= u_0, && \text{in } \Omega_0 = \Omega. \end{aligned} \tag{2.14}$$

This is a typical parabolic partial differential equation, where we assume that the boundary Γ_t is unknown. We would like to enforce the Stefan condition (2.2) by introducing a tracking-type functional for the Stefan condition. But instead of tracking (2.2), we will track the rewritten Stefan condition by using Lemma 2.5. Our choice of functional is hence

$$J(Q_T) = \frac{1}{2}\int_0^T \int_{\Gamma_t} \left(\frac{\partial u}{\partial \mathbf{n}} - \frac{\nu_1}{\sqrt{1-\nu_1^2}} \right)^2 \mathrm{d}\sigma \mathrm{d}t, \tag{2.15}$$

where u denotes the solution of (2.14) and ν_1 denotes the time component of the time-space normal $\boldsymbol{\nu}$, see (2.10).

Since the integrand in the objective functional (2.15) is non-negative, the objective functional is minimal if the Stefan condition (2.2) is satisfied. This amounts to the shape optimization problem

minimize $J(Q_T)$ from (2.15) over the class of admissible domains, where u satisfies (2.14).

Such problems can, for example, be numerically solved by applying a gradient-based method. Therefore, we shall compute the shape derivative of J in Section 3.

2.6. Solvability of the state equation. In order to state the solvability of (2.14), we introduce the following function spaces, which we already used in [1]. We set $Q_0 = (0,T) \times \Omega_0$, which has a time-independent boundary denoted by $\Sigma_0 := (0,T) \times \partial\Omega_0$. The anisotropic Sobolev spaces are defined by

$$H^{r,s}(Q_0) := L^2\big((0,T); H^r(\Omega_0)\big) \cap H^s\big((0,T); L^2(\Omega_0)\big)$$

for r, $s \in \mathbb{R}_{\geq 0}$, see, e.g., [2, 4, 25].

In order to include an initial condition, we can define

$$\hat{H}^{r,s}(Q_0) := \Big\{ u = U|_{Q_0} : U \in H^{r,s}\big((0,T)\times\Omega_0\big),\ U(t,\cdot) = 0,\ t < 0 \Big\}.$$

As in the elliptic case, we can include also (spatial) zero boundary conditions into the function spaces by setting

$$H_0^{r,s}(Q_0) := \big\{ u \in H^{r,s}(Q_0) : u|_{\Sigma_0} = 0 \big\}.$$

We are now in the position to introduce the non-cylindrical analogues of the above spaces by setting

$$H^{r,s}(Q_T) := \big\{ v \in L^2(Q_T) : v \circ \boldsymbol{\kappa} \in H^{r,s}(Q_0) \big\},$$

where the composition with $\boldsymbol{\kappa}$ only acts on the spatial component. Due to the chain rule, $v \circ \boldsymbol{\kappa}$ and v admit the same Sobolev regularity, provided that the mapping $\boldsymbol{\kappa}$ is smooth enough, see for example [27, Theorem 3.23] for the elliptic case. We define the norm of $H^{r,s}(Q_T)$ as

$$\|u\|_{H^{r,s}(Q_T)} = \|u \circ \boldsymbol{\kappa}\|_{H^{r,s}(Q_0)}$$

for $r, s \geq 0$.

THEOREM 2.7. *For every $u_0 \in L^2(\Omega_0)$, there exists a unique solution $u \in H_0^{1,\frac{1}{2}}(Q_T)$ satisfying* (2.14).

Proof. For the proof, we follow the strategy taken in [11, pg. 14–15]. Considering the problem with inhomogeneous initial condition $u_0 \in L^2(\Omega_0)$ but with homogeneous boundary condition and homogeneous source term, there exists a unique solution $u \in \mathcal{V}_0(Q_T)$ in accordance with [41, Theorem 26.1], since all the requested assumptions of this theorem are already shown in [1]. In here, the space $\mathcal{V}_0(Q_T)$ consists of all the functions v with $v \circ \boldsymbol{\kappa} \in \mathcal{V}_0(Q_0)$ and

$$\mathcal{V}_0(Q_0) := \big\{u \in L^2\big((0,T); H_0^1(\Omega_0)\big) : \partial_t u \in L^2\big((0,T); H^{-1}(\Omega_0)\big)\big\}.$$

Notice that this space is a dense subspace of $H_0^{1,\frac{1}{2}}(Q_0)$, thus we deduce the existence of a solution in $H_0^{1,\frac{1}{2}}(Q_T)$. The uniqueness follows by considering two different solutions, looking at their difference and using that, if $u_0 = 0$, then the unique solution of (2.14) in the space $\hat{H}_0^{1,\frac{1}{2}}(Q_T)$ is $u = 0$ according to [1]. ∎

REMARK 2.8. *If the initial data in* (2.14) *satisfy $u_0 \in H^1(\Omega_0)$, then the solution u of* (2.14) *lies in $H^{2,1}(Q_T)$. This is a consequence of [24, Chapter IV, Theorem 9.1].*

3. Shape calculus.

3.1. Perturbation of the tube. The shape calculus for time-dependent problems can be formulated under two points of view: the speed method (see e.g. [14, 29]) or the perturbation of identity (see [29]). In the speed method, the goal is to find the velocity field $\mathbf{V}$, which generates a mapping according to (2.5). This is an Eulerian setting. On the other hand, the perturbation of identity corresponds to a Lagrangian setting. This is the setting which we will use in the following. One considers the bijective mapping $\boldsymbol{\kappa}$ from (2.6), which maps the reference domain onto the tube. The mapping scheme is depicted in Figure 3.1.

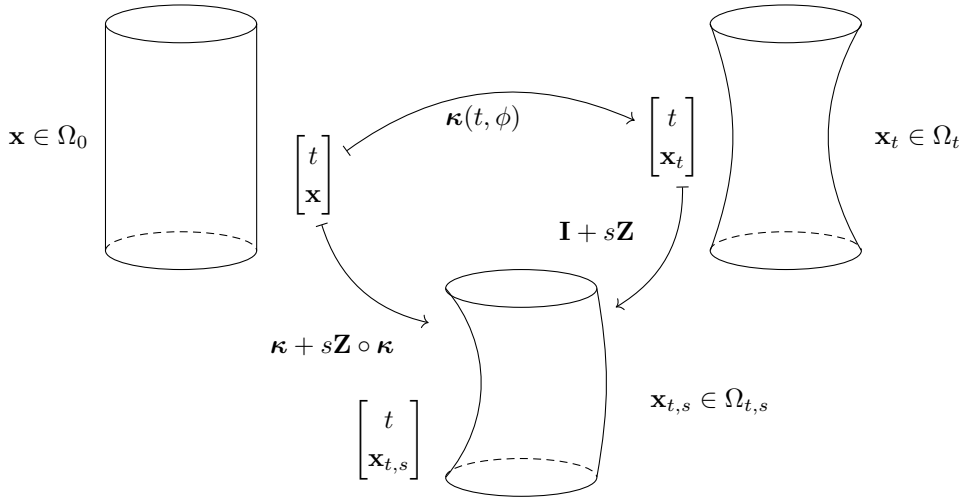


Figure 3.1: Perturbation of identity in the Lagrangian setting.

In order to compute a directional derivative of the objective functional, we perturb the tube Q_T using a vector field $\mathbf{Z}(t,\mathbf{x}) \in \mathbb{R}^d$ by applying a perturbation of identity $\mathbf{I} + s\mathbf{Z}$, satisfying again the

uniformity condition as in (2.7) when s is small enough. This yields a new tube

$$Q_T^s = \bigcup_{0<t<T} \big(\{t\} \times (\mathbf{I} + s\mathbf{Z})(\Omega_t)\big).$$

REMARK 3.1. *Our choice of the objective functional (see* (2.15)*) is more suitable for the Lagrangian approach of shape optimization than tracking the L^2-error of the Stefan condition* (2.2) *directly. This is due to the fact that the Stefan condition* (2.2) *is posed in an Eulerian form, since it explicitly contains the vector field $\mathbf{V}$, which generates the tube. In the speed method, we would perturb this vector field by considering $\mathbf{V} + s\mathbf{W}$ for an appropriate perturbation field $\mathbf{W}$.*

3.2. Local shape derivative. Let us define the space of admissible perturbation fields as

$$\mathcal{Z}_{\mathrm{ad}} := \Big\{ \mathbf{Z} \in C^2\big((0,T) \times D; \mathbb{R}^d\big) \Big\}$$

and consider a perturbation field $\mathbf{Z} \in \mathcal{Z}_{\mathrm{ad}}$. As in the time-independent case, we can define non-cylindrical material and local shape derivatives. The *material derivative* $\dot{v}[\mathbf{Z}]$ is defined as

$$\dot{v}[\mathbf{Z}] = \lim_{s \to 0} \frac{v_{t,s} \circ (\mathbf{I} + s\mathbf{Z}) - v_t}{s}, \tag{3.1}$$

where v_t is the state computed on Q_T and $v_{t,s}$ on the perturbed domain Q_T^s. By the relation

$$\delta v[\mathbf{Z}] = \dot{v}[\mathbf{Z}] - \nabla v \cdot \mathbf{Z}, \tag{3.2}$$

we can associate the so-called *local shape derivative* to a state, see [29]. The following theorem gives a characterization of the local shape derivative.

THEOREM 3.2. *The local shape derivative of the state u from* (2.14) *in the direction $\mathbf{Z} \in \mathcal{Z}_{\mathrm{ad}}$ can be computed as the solution of the partial differential equation*

$$\begin{aligned} \partial_t \delta u &= \Delta \delta u && \text{in } Q_T, \\ \delta u &= -\langle \mathbf{Z}, \mathbf{n} \rangle \frac{\partial u}{\partial \mathbf{n}} && \text{on } \Sigma_T, \\ \delta u(0, \cdot) &= 0 && \text{in } \Omega_0. \end{aligned} \tag{3.3}$$

For a proof, one can straightforwardly modify the proof given in [1], which is based on [2].

3.3. Tangential calculus. Before computing the directional derivative of the objective functional, we state some definitions and formulae of the tangential calculus which are used in the computations afterwards. To that end, we consider a general domain Ω with boundary Γ unless stated otherwise. Moreover, let $\mathbf{n}$ be the outward pointing normal.

Let us consider the *signed distance function* b (see [7, Chapter 7, Section 2] or [30, Section 2.5.6]), which can be used to describe surfaces. For smooth surfaces, it holds that the normal $\mathbf{n}$ to the surface corresponds to the gradient of b. Thus, we have a canonical extension $\boldsymbol{\mathcal{N}}$ of the normal from the surface into a tubular neighborhood of the surface. We denote the curvature operator $\nabla \boldsymbol{\mathcal{N}}$ by $\mathcal{R}$. According to [30, Theorem 2.5.18] and [17, Chapter 13.1], the curvature operator is a symmetric linear operator acting in the tangent plane. Thus, according to [30, Formula (2.5.162)], it holds

$$\mathcal{R}\mathbf{n} = 0. \tag{3.4}$$

We can define the curvature along the lines of [29, Definition 5.4] or [7, Chapter 9, Section 4.2]:

DEFINITION 3.3. *For a smooth surface* $\Gamma \subset \mathbb{R}^d$, $d \in \mathbb{N}$, *the* additive curvature $\mathcal{H}$ *of* Γ *is defined as*

$$\mathcal{H} = \operatorname{tr}(\mathrm{D}^2 b) = \Delta b = (d-1)\overline{H}.$$

Here, $\overline{H}$ *is the mean curvature and* $\mathrm{D}^2 b$ *is the second fundamental form (see [7, Chapter 9, Section 5]) and corresponds to the curvature operator* $\mathcal{R}$.

In [7, Chapter 9, Sections 5.1 & 5.2], we find the following definition of the tangential differential operators.

DEFINITION 3.4. *Let the boundary* Γ *be compact and let the level set function* b *be smooth enough in a tubular neighbourhood of* Γ. *Let us associate to* $f \in C^1(\Gamma)$ *a* C^1*-extension* F *in a tubular neighbourhood of* Γ. *Then, the* tangential gradient *of* f *is defined as*

$$\nabla_\Gamma f := \nabla F|_\Gamma - \frac{\partial F}{\partial \mathbf{n}}\mathbf{n}.$$

Moreover, we can define the tangential Jacobian matrix of a vector field $\mathbf{v} \in \left(C^1(\Gamma); \mathbb{R}^d\right)$, $d \geq 1$ *as*

$$\mathrm{D}_\Gamma \mathbf{v} = \mathrm{D}\mathbf{V}|_\Gamma - \mathrm{D}\mathbf{V}\mathbf{n}\mathbf{n}^\intercal, \tag{3.5}$$

where $\mathbf{V}$ *is a smooth extension of* $\mathbf{v}$ *in a tubular neighbourhood of* Γ. *The* tangential divergence *is defined as*

$$\operatorname{div}_\Gamma \mathbf{v} = \operatorname{div} \mathbf{V}|_\Gamma - \langle \mathrm{D}\mathbf{V}\mathbf{n}, \mathbf{n}\rangle.$$

The following lemma gives a formula of tangential calculus, which can be found in [7, Chapter 9, Section 5.4, pg. 497].

LEMMA 3.5. *Let* $\mathbf{W}$ *be a* C^1*-vector field in a tubular neighborhood of the boundary* Γ. *We set*

$$\mathbf{w} := \mathbf{W}|_\Gamma, \quad w_\mathbf{n} := \langle \mathbf{W}, \mathbf{n}\rangle, \quad \mathbf{w}_\Gamma := \mathbf{W} - w_\mathbf{n}\mathbf{n}.$$

Then, there holds

$$\nabla_\Gamma w_\mathbf{n} = (\mathrm{D}_\Gamma \mathbf{w})^\intercal \mathbf{n} + \mathrm{D}^2 b \mathbf{w}_\Gamma. \tag{3.6}$$

If we consider a surface containing a boundary, then we have the following variant of the tangential Stokes formula according to [9, Corollary 3.1] or [35, Proposition 2.58].

LEMMA 3.6. *Let* $S \subset \Gamma$ *be a* C^2*-manifold and* ∂S *be the boundary of* S. *For* $\mathbf{v} \in H^1(\Gamma; \mathbb{R}^d)$ *it holds*

$$\int_S \operatorname{div}_S \mathbf{v} \, \mathrm{d}S = \int_S \mathcal{H}\langle \mathbf{v}, \mathbf{n}\rangle \, \mathrm{d}S - \int_{\partial S} \langle \mathbf{v}, \boldsymbol{\tau}\rangle \, \mathrm{d}\partial S,$$

where $\mathcal{H}$ *denotes the additive curvature (see Definition 3.3) and* $\boldsymbol{\tau}$ *is the unit tangent vector to* S, *outward pointing from* S *and normal to the boundary* ∂S.

3.4. Ingredients for the shape derivative of the objective functional. To present the proof of the shape derivative of the objective functional in a clear manner, we shall provide some useful computations beforehand.

LEMMA 3.7. *There holds*

$$\delta\left(\frac{\partial u}{\partial \mathbf{n}}\right) = \frac{\partial \delta u}{\partial \mathbf{n}} \quad \text{on } \Gamma_t,$$

where δu *denotes the local shape derivative of* (2.14).

Proof. According to Lemma 2.2, we can compute

$$\delta\left(\frac{\partial u}{\partial \mathbf{n}}\right) = -\delta\left(\|\nabla u\|\right).$$

Thus, we have

$$\delta\left(\sqrt{\langle \nabla u, \nabla u\rangle}\right) = \frac{\langle \nabla u, \nabla \delta u\rangle}{\|\nabla u\|}$$

since spatial derivatives and the local shape derivative commute. Using finally (2.8) leads to the claim. ∎

LEMMA 3.8. *The local shape derivative of* $\nu_1/\sqrt{1-\nu_1^2}$ *is given by*

$$\delta\left(\frac{\nu_1}{\sqrt{1-\nu_1^2}}\right) = -\frac{1}{(1-\nu_1^2)^{\frac{3}{2}}}\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \vec{z}_{\boldsymbol{\nu}}\rangle,$$

where $\vec{\mathbf{e}}_1 \in \mathbb{R}^{1+d}$ *denotes the (first) canonical unit vector in* $\mathbb{R}^{1+d}$ *and* $\vec{\mathbf{Z}} = \left[\begin{smallmatrix} 0 \\ \mathbf{Z} \end{smallmatrix}\right]$.

Proof. With the chain rule and the quotient rule it follows

$$\delta\left(\frac{\nu_1}{\sqrt{1-\nu_1^2}}\right) = \frac{1}{(1-\nu_1^2)^{\frac{3}{2}}}\delta\nu_1.$$

It remains to compute $\delta\nu_1$. To this end, let us consider the whole time-space domain, which gets perturbed with the perturbation field $\vec{\mathbf{Z}}$ by applying the map $\vec{\mathbf{I}} + s\vec{\mathbf{Z}}$. Due to the choice of $\vec{\mathbf{Z}}$, this corresponds to a horizontal perturbation, thus, a perturbation of the spatial component in the direction $\mathbf{Z}$. For the normal $\boldsymbol{\nu}_s$ on the perturbed surface, it holds according to [7, Chapter 9, Section 4] that

$$\boldsymbol{\nu}_s \circ (\vec{\mathbf{I}} + s\vec{\mathbf{Z}}) = \frac{\left(\vec{\mathrm{D}}(\vec{\mathbf{I}} + s\vec{\mathbf{Z}})\right)^\intercal \boldsymbol{\nu}}{\left\|\left(\vec{\mathrm{D}}(\vec{\mathbf{I}} + s\vec{\mathbf{Z}})\right)^\intercal \boldsymbol{\nu}\right\|_{\mathbb{R}^{1+d}}}.$$

From this, we derive (see also [7])

$$\delta\boldsymbol{\nu} = \langle \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}, \boldsymbol{\nu}\rangle\boldsymbol{\nu} - (\vec{\mathrm{D}}\vec{\mathbf{Z}})^\intercal\boldsymbol{\nu} - (\vec{\mathrm{D}}^2\vec{b})\vec{\mathbf{Z}} \tag{3.7}$$

with the time-space signed distance function $\vec{b}$ as introduced in Section 3.3. Especially, $\delta\nu_1$ is the first coordinate of this (vector valued) equation.

Let us define $\vec{\mathbf{Z}}|_\Sigma := \vec{\mathbf{z}}$, $\vec{z}_{\boldsymbol{\nu}} := \langle \vec{\mathbf{Z}}, \boldsymbol{\nu}\rangle$, and $\vec{\mathbf{z}}_\Sigma := \vec{\mathbf{Z}} - \vec{z}_{\boldsymbol{\nu}}\boldsymbol{\nu}$ as in Lemma 3.5. Since (3.4) also holds with the time-space normal $\boldsymbol{\nu}$, we have on Σ_T

$$(\vec{\mathrm{D}}^2\vec{b})\vec{\mathbf{Z}} = (\vec{\mathrm{D}}^2\vec{b})\vec{\mathbf{z}}_\Sigma$$

As in [7, pg. 501], we can compute by using (3.5)

$$(\vec{\mathrm{D}}\vec{\mathbf{Z}})^\intercal\boldsymbol{\nu} = (\vec{\mathrm{D}}_\Sigma\vec{\mathbf{z}} + \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}\boldsymbol{\nu}^\intercal)^\intercal\boldsymbol{\nu} = (\vec{\mathrm{D}}_\Sigma\vec{\mathbf{z}})^\intercal\boldsymbol{\nu} + \boldsymbol{\nu}\boldsymbol{\nu}^\intercal\vec{\mathrm{D}}\vec{\mathbf{Z}}^\intercal\boldsymbol{\nu} = (\vec{\mathrm{D}}_\Sigma\vec{\mathbf{z}})^\intercal\boldsymbol{\nu} + \langle \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}, \boldsymbol{\nu}\rangle\boldsymbol{\nu}.$$

We can thus rewrite the first coordinate of the right hand side of (3.7) by

$$\begin{aligned}
&\left\langle \vec{\mathbf{e}}_1, \langle \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}, \boldsymbol{\nu}\rangle\boldsymbol{\nu}\right\rangle - \left\langle \vec{\mathbf{e}}_1, (\vec{\mathrm{D}}\vec{\mathbf{Z}})^\intercal\boldsymbol{\nu}\right\rangle - \langle \vec{\mathbf{e}}_1, \vec{\mathrm{D}}^2\vec{b}\vec{\mathbf{Z}}\rangle \\
&\quad = \left\langle \vec{\mathbf{e}}_1, \langle \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}, \boldsymbol{\nu}\rangle\boldsymbol{\nu}\right\rangle - \left\langle \vec{\mathbf{e}}_1, (\vec{\mathrm{D}}_\Sigma\vec{\mathbf{z}})^\intercal\boldsymbol{\nu} + \langle \vec{\mathrm{D}}\vec{\mathbf{Z}}\boldsymbol{\nu}, \boldsymbol{\nu}\rangle\boldsymbol{\nu}\right\rangle - \langle \vec{\mathbf{e}}_1, \vec{\mathrm{D}}^2\vec{b}\vec{\mathbf{z}}_\Sigma\rangle \\
&\quad = -\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma\vec{z}_{\boldsymbol{\nu}}\rangle,
\end{aligned}$$

where we used the time-space analogue of (3.6). ∎

LEMMA 3.9. *There holds the identity*

$$\frac{\partial}{\partial \mathbf{n}}\langle \nabla u, \mathbf{n}\rangle = \langle \mathrm{D}^2 u \mathbf{n}, \mathbf{n}\rangle =: \frac{\partial^2 u}{\partial \mathbf{n}^2} \quad \text{on } \Gamma_t.$$

Proof. Due to (3.4), the claim follows from

$$\frac{\partial}{\partial \mathbf{n}}\langle \nabla u, \mathbf{n}\rangle = \big\langle \nabla \langle \nabla u, \mathbf{n}\rangle, \mathbf{n}\big\rangle = \langle \mathrm{D}^2 u \mathbf{n} + \mathcal{R}\nabla u, \mathbf{n}\rangle.$$ ∎

To compute the second order normal derivative, we give the following lemma:

LEMMA 3.10. *The second normal derivative of u can be computed as*

$$\frac{\partial^2 u}{\partial \mathbf{n}^2} = \frac{\nu_1}{\sqrt{1-\nu_1^2}}\frac{\partial u}{\partial \mathbf{n}} - \overline{H}_{\mathbf{x}}\frac{\partial u}{\partial \mathbf{n}},$$

where $\overline{H}_{\mathbf{x}}$ denotes the spatial mean curvature, compare Definition 3.3.

Proof. Let us consider a fixed point of time t. According to [35, Proposition 2.68], for a smooth boundary Γ and function ϕ on Γ, it holds

$$\Delta \phi = \Delta_\Gamma \phi + \overline{H}_{\mathbf{x}}\frac{\partial \phi}{\partial \mathbf{n}} + \frac{\partial^2 \phi}{\partial \mathbf{n}^2},$$

where Δ_Γ denotes the Laplace-Beltrami operator, defined as $\Delta_\Gamma \phi := \mathrm{div}_\Gamma(\nabla_\Gamma \phi)$, compare [7, Chapter 9, Section 5.3]. Therefore, we can compute the second normal derivative of u as

$$\frac{\partial^2 u}{\partial \mathbf{n}^2} = \Delta u - \overline{H}_{\mathbf{x}}\frac{\partial u}{\partial \mathbf{n}},$$

where we used that u vanishes on the boundary Γ_t and, thus, the tangential derivative equals to zero. Due to the state equation, we have that $\partial_t u = \Delta u$ and, therefore, we arrive at the claim by using (2.12). ∎

The following lemma connects spatial and temporal derivatives of ν_1.

LEMMA 3.11. *It holds*

$$\langle \nabla \nu_1, \mathbf{n}\rangle = -\partial_t \nu_1 \frac{\nu_1}{\sqrt{1-\nu_1^2}}.$$

Proof. From (3.4) and the symmetry of the curvature operator, we have

$$\vec{\nabla}_{\boldsymbol{\nu}}\boldsymbol{\nu} = \vec{\mathcal{R}}\boldsymbol{\nu} = \mathbf{0},$$

where $\vec{\mathcal{R}}$ denotes the time-space curvature operator. Looking at the first entry of $\vec{\mathcal{R}}\boldsymbol{\nu}$ and using (2.13) thus gives

$$\partial_t \nu_1 \nu_1 + \langle \nabla \nu_1, \mathbf{n}\rangle \sqrt{1-\nu_1^2} = 0.$$ ∎

3.5. Shape derivative of the objective functional. With the previous preparations at hand, we can compute the shape derivative of the objective functional (2.15), which is defined by

$$\nabla J(Q_T)[\mathbf{Z}] = \lim_{s \to 0} \frac{J\left(Q_T^s\right) - J\left(Q_T\right)}{s}.$$

THEOREM 3.12. *The shape derivative of the objective functional* (2.15) *in the direction* $\mathbf{Z} \in \mathcal{Z}_{ad}$ *in Hadamard form reads*

$$\begin{aligned}\nabla J(Q_T)[\mathbf{Z}] = \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n} \rangle \Big\{ &- \frac{\partial p}{\partial \mathbf{n}} \frac{\partial u}{\partial \mathbf{n}} - \vec{\mathrm{div}}_\Sigma \left(p \frac{1}{1-\nu_1^2} \vec{\nabla}_\Sigma t \right) \\ &- p \overline{H}_{\mathbf{x}} \frac{\partial u}{\partial \mathbf{n}} + p \partial_t \nu_1 \frac{\nu_1}{(1-\nu_1^2)^2} + \frac{1}{2} \mathcal{H}_{\mathbf{x}} p^2 \Big\} \, \mathrm{d}\sigma \mathrm{d}t \\ &- \int_{\Gamma_0 \cup \Gamma_T} \frac{\tau_1}{\sqrt{1-\nu_1^2}} \langle \mathbf{Z}, \mathbf{n} \rangle p \, \mathrm{d}\sigma,\end{aligned}$$

Here, the adjoint state p *satisfies the following backward heat equation*

$$\begin{aligned} -\partial_t p - \Delta p &= 0 && \text{in } Q_T, \\ p &= \frac{\partial u}{\partial \mathbf{n}} - \frac{\nu_1}{\sqrt{1-\nu_1^2}} && \text{on } \Sigma_T, \\ p(T, \cdot) &= 0 && \text{on } \Omega_T, \end{aligned} \tag{3.8}$$

$\mathcal{H}_{\mathbf{x}}$ *denotes the spatial additive curvature of* Γ_t *(compare Definition 3.3), and* τ_1 *is the first entry of* $\boldsymbol{\tau}$ *described in Lemma 3.6.*

Proof. We shall first employ the formula from [29, Theorem 5.5], [29, Proposition 6.1] (which are stated for time-space) or [7, Chapter 9, Theorem 4.3] (only for the inner integral), which immediately yields

$$\begin{aligned}\nabla J(Q_T)[\mathbf{Z}] = A + B + C := &\frac{1}{2} \int_0^T \int_{\Gamma_t} \delta \left(\left(\frac{\partial u}{\partial \mathbf{n}} - \frac{\nu_1}{\sqrt{1-\nu_1^2}} \right)^2 \right) \mathrm{d}\sigma \mathrm{d}t \\ &+ \frac{1}{2} \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n} \rangle \frac{\partial}{\partial \mathbf{n}} \left(\left(\frac{\partial u}{\partial \mathbf{n}} - \frac{\nu_1}{\sqrt{1-\nu_1^2}} \right)^2 \right) \mathrm{d}\sigma \mathrm{d}t \\ &+ \frac{1}{2} \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n} \rangle \mathcal{H}_{\mathbf{x}} \left(\frac{\partial u}{\partial \mathbf{n}} - \frac{\nu_1}{\sqrt{1-\nu_1^2}} \right)^2 \mathrm{d}\sigma \mathrm{d}t\end{aligned} \tag{3.9}$$

While B and C are already in Hadamard form and only the normal derivative in B has to be treated, the integral A has to be brought into Hadamard form by using the adjoint problem. With the aid of the chain rule, Lemma 3.7, and Lemma 3.8, we compute A as

$$A = A_1 + A_2 = \int_0^T \int_{\Gamma_t} p \frac{\partial \delta u}{\partial \mathbf{n}} \, \mathrm{d}\sigma \mathrm{d}t + \int_0^T \int_{\Gamma_t} p \frac{1}{(1-\nu_1^2)^{\frac{3}{2}}} \langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \vec{z}_{\boldsymbol{\nu}} \rangle \, \mathrm{d}\sigma \mathrm{d}t. \tag{3.10}$$

The second term on the right-hand side is similar to [7, pg. 490 resp. 501], but since we only have the first component of the normal, we have the scalar product with $\vec{\mathbf{e}}_1$. To eliminate the Neumann derivative on δu, we derive the following integration by parts formula by applying Green's formula and Reynolds transport theorem:

$$\begin{aligned} 0 &= \int_0^T \int_{\Omega_t} (\partial_t \delta u - \Delta \delta u) p + \delta u (\partial_t p + \Delta p) \, \mathrm{d}\mathbf{x} \mathrm{d}t \\ &= \int_0^T \int_{\Omega_t} \partial_t (\delta u p) \, \mathrm{d}\mathbf{x} \mathrm{d}t + \int_0^T \int_{\Gamma_t} \frac{\partial p}{\partial \mathbf{n}} \delta u - \frac{\partial \delta u}{\partial \mathbf{n}} p \, \mathrm{d}\sigma \mathrm{d}t \\ &= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega_t} \delta u p \, \mathrm{d}\mathbf{x} \mathrm{d}t - \int_0^T \int_{\Gamma_t} \delta u p \langle \mathbf{V}, \mathbf{n} \rangle \, \mathrm{d}\sigma \mathrm{d}t + \int_0^T \int_{\Gamma_t} \frac{\partial p}{\partial \mathbf{n}} \delta u - \frac{\partial \delta u}{\partial \mathbf{n}} p \, \mathrm{d}\sigma \mathrm{d}t. \end{aligned}$$

By using the fundamental theorem of calculus together with the zero initial condition of δu and the zero end condition of p, we conclude that the domain integral vanishes. Lemma 2.5 yields

$$A_1 = \int_0^T \int_{\Gamma_t} \frac{\partial p}{\partial \mathbf{n}} \delta u \, \mathrm{d}\sigma \mathrm{d}t + \int_0^T \int_{\Gamma_t} \delta u p \left(\frac{\nu_1}{\sqrt{1-\nu_1^2}} \right) \mathrm{d}\sigma \mathrm{d}t.$$

Inserting the boundary condition of δu leads finally to

$$A_1 = \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n} \rangle \left\{ - \frac{\partial p}{\partial \mathbf{n}} \frac{\partial u}{\partial \mathbf{n}} - \frac{\partial u}{\partial \mathbf{n}} p \left(\frac{\nu_1}{\sqrt{1-\nu_1^2}} \right) \right\} \mathrm{d}\sigma \mathrm{d}t.$$

Let us next look at the term A_2. When inserting $\vec{\mathbf{e}}_1 = \vec{\nabla} t$, we get

$$A_2 = \int_0^T \int_{\Gamma_t} p \frac{1}{(1-\nu_1^2)^{\frac{3}{2}}} \langle \vec{\nabla} t, \vec{\nabla}_\Sigma \vec{z}_{\boldsymbol{\nu}} \rangle \, \mathrm{d}\sigma \mathrm{d}t.$$

We can split $\vec{\nabla} t$ in its tangential and normal component in time-space. The normal component vanishes within the scalar product due to [7, Chapter 9, Theorem 5.1]. Using

$$\mathrm{d}\Sigma = \frac{1}{\sqrt{1-\nu_1^2}} \mathrm{d}\sigma \mathrm{d}t \tag{3.11}$$

as it is stated in [29, Remark 6.3, pg. 167], we therefore obtain

$$A_2 = \int_{\Sigma_T} p \frac{1}{1-\nu_1^2} \langle \vec{\nabla}_\Sigma t, \vec{\nabla}_\Sigma \vec{z}_{\boldsymbol{\nu}} \rangle \, \mathrm{d}\Sigma.$$

Following the ideas of [7, Chapter 9, Section 5.7], we apply the product rule to get

$$A_2 = \int_{\Sigma_T} \vec{\mathrm{div}}_\Sigma \left(p \frac{1}{1-\nu_1^2} \vec{\nabla}_\Sigma t \vec{z}_{\boldsymbol{\nu}} \right) - \vec{\mathrm{div}}_\Sigma \left(p \frac{1}{1-\nu_1^2} \vec{\nabla}_\Sigma t \right) \vec{z}_{\boldsymbol{\nu}} \, \mathrm{d}\Sigma.$$

Using Lemma 3.6 then gives

$$\begin{aligned} A_2 = \int_{\Sigma_T} \mathcal{H}_{t,\mathbf{x}} p \frac{1}{1-\nu_1^2} \langle \vec{\nabla}_\Sigma t, \boldsymbol{\nu} \rangle \vec{z}_{\boldsymbol{\nu}} \, \mathrm{d}\Sigma - \int_{\Gamma_0 \cup \Gamma_T} p \frac{1}{1-\nu_1^2} \langle \vec{\nabla}_\Sigma t, \boldsymbol{\tau} \rangle \vec{z}_{\boldsymbol{\nu}} \, \mathrm{d}\sigma \\ - \int_{\Sigma_T} \vec{\mathrm{div}}_\Sigma \left(p \frac{1}{1-\nu_1^2} \vec{\nabla}_\Sigma t \right) \vec{z}_{\boldsymbol{\nu}} \, \mathrm{d}\Sigma. \end{aligned}$$

Herein, the first integral of the right-hand side vanishes due to $\langle \vec{\nabla}_\Sigma t, \boldsymbol{\nu} \rangle = 0$. In view of

$$\vec{\nabla}_\Sigma t = \vec{\mathbf{e}}_1 - \nu_1 \boldsymbol{\nu}, \tag{3.12}$$

and $\langle \boldsymbol{\nu}, \boldsymbol{\tau} \rangle = 0$, the second integral of the right-hand side reduces to $p\tau_1 \vec{z}_{\boldsymbol{\nu}} / (1-\nu_1^2)$, where τ_1 denotes the first coordinate of $\boldsymbol{\tau}$. By using (2.13), $\vec{z}_{\boldsymbol{\nu}} = \sqrt{1-\nu_1^2} \langle \mathbf{Z}, \mathbf{n} \rangle$, and (3.11), we thus have also A_2 in Hadamard form:

$$A_2 = - \int_{\Gamma_0 \cup \Gamma_T} p \frac{\tau_1}{\sqrt{1-\nu_1^2}} \langle \mathbf{Z}, \mathbf{n} \rangle \, \mathrm{d}\sigma - \int_0^T \int_{\Gamma_t} \vec{\mathrm{div}}_\Sigma \left(p \frac{1}{1-\nu_1^2} \vec{\nabla}_\Sigma t \right) \langle \mathbf{Z}, \mathbf{n} \rangle \, \mathrm{d}\sigma \mathrm{d}t.$$

Next, we shall treat the term B in (3.9). It can be computed by using Lemma 3.9:

$$B = \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n} \rangle p \left\{ \frac{\partial^2 u}{\partial \mathbf{n}^2} - \frac{\partial}{\partial \mathbf{n}} \left(\frac{\nu_1}{\sqrt{1-\nu_1^2}} \right) \right\} \mathrm{d}\sigma \mathrm{d}t.$$

In view of Lemma 3.10, we can eliminate the second order normal derivative. Moreover, the second term can be treated by using the quotient rule, resulting in

$$\frac{\partial}{\partial \mathbf{n}}\left(\frac{\nu_1}{\sqrt{1-\nu_1^2}}\right) = \frac{1}{(1-\nu_1^2)^{\frac{3}{2}}}\langle \nabla \nu_1, \mathbf{n}\rangle.$$

Application of Lemma 3.11 yields finally

$$B = \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n}\rangle p \left\{ \frac{\nu_1}{\sqrt{1-\nu_1^2}} \frac{\partial u}{\partial \mathbf{n}} - \overline{H}_{\mathbf{x}} \frac{\partial u}{\partial \mathbf{n}} + \partial_t \nu_1 \frac{\nu_1}{(1-\nu_1^2)^2} \right\} \mathrm{d}\sigma \mathrm{d}t.$$

The claim follows when taking finally into account that

$$C = \frac{1}{2}\int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n}\rangle \mathcal{H}_{\mathbf{x}} p^2 \, \mathrm{d}\sigma \mathrm{d}t$$

and computing $A_1 + A_2 + B + C$. ∎

3.6. Shape derivative for the numerical computations. The shape gradient of the objective functional given in Theorem 3.12 is in Hadamard form. Nevertheless, for numerical computations, we need to rewrite the term containing the surface divergence to make it computable. Two approaches can be chosen: On the one hand, we can compute the surface divergence directly by treating the three terms separately and try to reformulate them into computable terms. Especially, one would have to reformulate the surface gradient of the adjoint problem. On the other hand, we could stop the manipulations of the term A_2 in the proof of Theorem 3.12 at (3.10) in order to avoid the computation of the surface divergence of several other terms. We then have to compute the surface gradient of the perturbation field in normal direction. Since we choose a smooth setting for our numerical computations, we pursue this approach.

In view of the definition of A_2 in (3.10), we compute

$$\left\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \langle \vec{\mathbf{Z}}, \boldsymbol{\nu}\rangle \right\rangle = \sqrt{1-\nu_1^2}\left\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \langle \mathbf{Z}, \mathbf{n}\rangle \right\rangle - \langle \mathbf{Z}, \mathbf{n}\rangle \frac{\nu_1}{\sqrt{1-\nu_1^2}} \langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \nu_1\rangle.$$

The term $\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \nu_1\rangle$ corresponds to an entry in the time-space curvature operator, namely $\partial_t \nu_1$. Adding the terms A_1, B, and C from the proof of Theorem 3.12 to the so computed expression of A_2 yields

$$\begin{aligned}
\nabla J(Q_T)[\mathbf{Z}] = \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n}\rangle \Big\{ & - \frac{\partial u}{\partial \mathbf{n}} \frac{\partial p}{\partial \mathbf{n}} - p \frac{\partial u}{\partial \mathbf{n}} \frac{\nu_1}{\sqrt{1-\nu_1^2}} \Big\} \\
& + p \frac{1}{1-\nu_1^2} \left\langle \vec{\mathbf{e}}_1, \vec{\nabla}_\Sigma \langle \mathbf{Z}, \mathbf{n}\rangle \right\rangle - \langle \mathbf{Z}, \mathbf{n}\rangle p \frac{\nu_1}{(1-\nu_1^2)^2} \partial_t \nu_1 \\
& + \langle \mathbf{Z}, \mathbf{n}\rangle \left\{ p \left(\frac{\nu_1}{\sqrt{1-\nu_1^2}} \frac{\partial u}{\partial \mathbf{n}} - \overline{H}_{\mathbf{x}} \frac{\partial u}{\partial \mathbf{n}} + \partial_t \nu_1 \frac{\nu_1}{(1-\nu_1^2)^2} \right) + \frac{1}{2} \mathcal{H}_{\mathbf{x}} p^2 \right\} \mathrm{d}\sigma \mathrm{d}t.
\end{aligned}$$

Two terms cancel out and we therefore arrive at

$$\begin{aligned}
\nabla J(Q_T)[\mathbf{Z}] = & \int_0^T \int_{\Gamma_t} \langle \mathbf{Z}, \mathbf{n}\rangle \Big\{ - \frac{\partial u}{\partial \mathbf{n}} \frac{\partial p}{\partial \mathbf{n}} - \overline{H}_{\mathbf{x}} p \frac{\partial u}{\partial \mathbf{n}} + \frac{1}{2} \mathcal{H}_{\mathbf{x}} p^2 \Big\} \mathrm{d}\sigma \mathrm{d}t \\
& + \int_0^T \int_{\Gamma_t} p \frac{1}{1-\nu_1^2} \left\langle \vec{\mathbf{e}}_1, \vec{\nabla}_{\Sigma_T} \langle \mathbf{Z}, \mathbf{n}\rangle \right\rangle \mathrm{d}\sigma \mathrm{d}t.
\end{aligned} \tag{3.13}$$

We use this form of the shape gradient for our numerical computations.

4. Numerical experiment. In this section, we indicate how we solve the optimization problem. The example serves as a proof of concept, and is therefore intentionally kept simple.

4.1. Parametrization of the shape optimization problem. To numerically solve the Stefan problem, reformulated as a shape optimization problem, we restrict ourselves to a star-shaped spatial domain $\Omega_t \subset \mathbb{R}^2$ for every point of time t (compare also [1]). This choice can be parametrized in space by using a Fourier series for the unknown radial function, where the coefficients are time-dependent. Thus, the boundary Σ_T can be represented by

$$\Sigma_T = \left\{ \begin{bmatrix} t \\ \boldsymbol{\gamma}(t,\theta) \end{bmatrix} \in \mathbb{R}^3 : t \in [0,T],\ \theta \in [0,2\pi) \right\},$$

where the time-dependent parametrization $\boldsymbol{\gamma}(t,\cdot) : [0,2\pi) \to \Gamma_t$ employs polar coordinates

$$\boldsymbol{\gamma}(t,\theta) = w(t,\theta) \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}. \tag{4.1}$$

Here, $w(t,\theta)$ denotes the time- and angle-dependent radius, given by

$$w(t,\theta) := \sum_{\ell=0}^{N_L} L_\ell(t) \left(\alpha_{0,\ell} + \sum_{k=1}^{N_K} \left\{ \alpha_{k,\ell} \cos(k\theta) + \beta_{k,\ell} \sin(k\theta) \right\} \right),$$

with $L_\ell(t)$ being appropriate dilations and translations of the Legendre polynomials of degree ℓ.

Finding the optimal tube now corresponds to determining the unknown coefficients $\alpha_{k,\ell}$ and $\beta_{k,\ell}$ of the parametrization. Hence, we have the following finite dimensional problem:

$$\text{Seek } \boldsymbol{\gamma}^\star \in Z_N \text{ such that } \nabla J(\boldsymbol{\gamma}^\star)[\mathbf{Z}] = 0 \text{ for all } \mathbf{Z} \in Z_N.$$

Here, Z_N is the finite dimensional ansatz space of parametrizations. To compute the discrete shape gradient, we hence have to consider the directions

$$(\mathbf{Z} \circ \boldsymbol{\gamma})(t,\theta) = L_\ell(t) \cos(k\theta) \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \tag{4.2}$$

for all $\ell = 1, \dots, N_L$ and $k = 0, \dots, N_K$, and

$$(\mathbf{Z} \circ \boldsymbol{\gamma})(t,\theta) = L_\ell(t) \sin(k\theta) \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \tag{4.3}$$

for all $\ell = 1, \dots, N_L$ and $k = 1, \dots, N_K$. Notice that, since the initial domain Ω_0 has to remain fixed, we do not wobble at the shape parameters $\alpha_{k,0}$ and $\beta_{k,0}$.

4.2. Implementation of the shape gradient. With the parametrization of the boundary at hand, we shall next explain how to implement the shape gradient (3.13).

We first need to compute the solutions u of the state equation (2.14) and p of the adjoint equation (3.8). Although we only need the Neumann data on the boundary, we cannot apply the boundary element method easily as we have a non-trivial initial condition $u = u_0$ for $t = 0$. Therefore, we employ the finite element method in space and couple it with the theta-scheme to solve the parabolic equation.

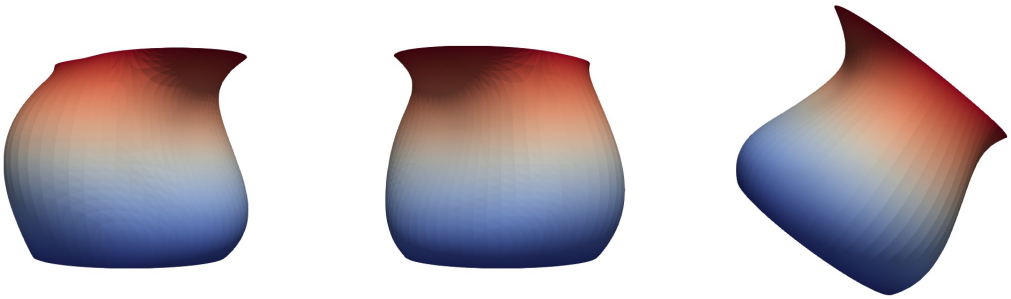
We use the space-time cylinder as reference domain, which means that we need to introduce a finite element mesh the unit circle. This mesh then gets mapped onto the spatial domain Ω_t described by the parametrization for every point of time t, similarly as in [22]. Then, standard piecewise linear finite elements can be used to solve the partial differential equation for every time step, when mapping the weak formulation back to the reference domain. For the time discretization, we use the theta-scheme on the reference domain, including the inhomogeneous Dirichlet data, see [39, Section 8.2.1] for example.

For the mapping of the domain onto the reference domain, we need to evaluate the Legendre polynomials in the parametrization, which can be done by using their three-term recurrence formula as described in [33].

From the finite element approximation of the state, the Neumann data can be computed, which are piecewise constant. The approximate Neumann data are then projected onto the space of piecewise linear functions as they enter the Dirichlet data of the adjoint state. Notice that the adjoint problem has a singularity at $t = T$. We do not treat this singularity specifically, but need to perform our computations on a fine level to resolve this singularity.

Another component for the shape gradient (3.13) is the additive curvature and the mean curvature in space. Since we consider a two-dimensional setting, both coincide and can be computed from the parametrization, see [10, pg. 21]. Finally, the surface gradient of $\langle \mathbf{Z}, \mathbf{n} \rangle$ can be computed as explained in [3] or [21, Section C.1] by using the parametrization at hand.

We have now all the components to compute the integrand of the shape gradient in (3.13). The integral is computed by using a trapezoidal rule in space and a trapezoidal rule in time on the reference cylindrical domain.



(a) View with the x_1-axis in front.

(b) View with the x_2-axis in front.

(c) Three-dimensional view.

Figure 4.1: Initial guess of the shape optimization problem. The colours correspond to the time slices.

4.3. Numerical results. For the parametrization of the boundary, we choose 15 Fourier coefficients in space ($N_K = 8$) and 10 Legendre polynomials in time ($N_L = 9$), leading to 150 design parameters in total, from which 135 are unknown as we let Ω_0 be fixed. We choose Ω_0 as the unit circle of radius 1. We set $u_0 = J_0\big(\|\mathbf{x}\|\lambda_0\big)$, where J_0 denotes the Bessel function of the first kind and λ_0 is its smallest positive root. In every time-step, we use 163'840 finite elements. We

choose $T = 0.2$ and a time step size of $\Delta_t = 0.0005$. We perform 50 iterations in the optimization procedure and use a quasi-Newton method updated by the limited memory inverse BFGS rule, where 10 updates are stored, see [31] for example. A second order line search is applied to find an appropriate step size in the quasi-Newton method. The optimization algorithm is started with the initial shape displayed in Figure 4.1.

In Figure 4.2 on the left, the evolution of the ℓ^∞-norm of the shape gradient is displayed during the course of the minimization algorithm, while on the right the evolution of the functional is shown.

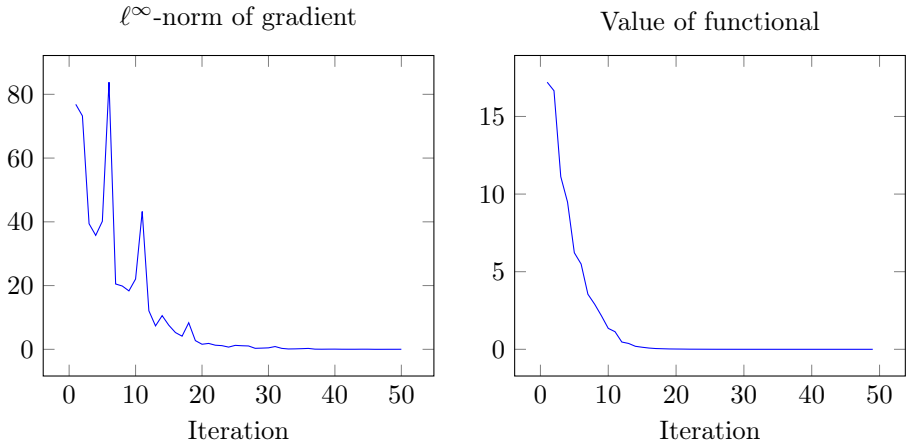


Figure 4.2: The histories of the shape gradient (*left*) and of the functional (*right*).

We clearly observe that these two values tend to zero, thus we have convergence of the minimization algorithm. Figure 4.3 shows the terminal shape at the end of the optimization process. It is a truncated cylinder. This solution corresponds to the intuition we have for the solution of the Stefan problem, as we would expect the initial circle to grow uniformly throughout time.

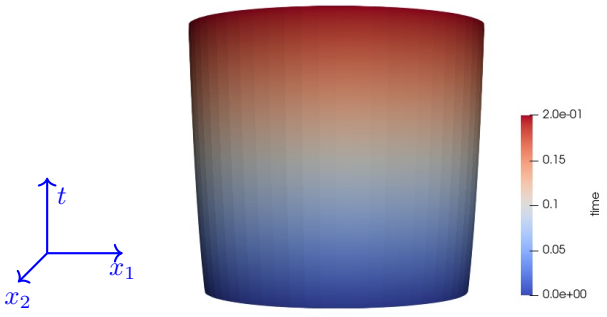


Figure 4.3: Terminal shape of the shape optimization problem.

5. Conclusion. In this article, we reformulated the Stefan problem as a shape optimization problem by introducing a shape functional subject to a differential equation. Bearing in mind that we would like to apply a gradient-based optimization algorithm, we computed the directional

derivative of the shape functional after rewriting the so-called Stefan condition in a suitable form. Using a parametrization of the boundary by means of a Fourier series allows for computing all terms of the discrete shape gradient. The theoretical results are supported by a numerical experiment, which serves as a proof of concept.

Acknowledgment. The authors would like to thank Enno Lenzmann for bringing their attention to the works of Mahir Hadžić and Steve Shkoller.

REFERENCES

[1] R. Brügger, H. Harbrecht, and J. Tausch. On the numerical solution of a time-dependent shape optimization problem for the heat equation. *SIAM J. Control Optim.*, 59(2):931–953, 2021.

[2] R. Chapko, R. Kress, and J.-R. Yoon. On the numerical solution of an inverse boundary value problem for the heat equation. *Inverse Problems*, 14(4):853–867, 1998.

[3] D. Colton and R. Kress. *Integral equation methods in scattering theory*. Krieger Publishing Company, Malabar, Florida, 1992.

[4] M. Costabel. Boundary integral operators for the heat equation. *Integral Equations and Operator Theory*, 13(4):498–552, 1990.

[5] J. Crank. *Free and moving boundary problems*. Clarendon Press Oxford, Oxford, 1984.

[6] M.C. Delfour and J.-P. Zolésio. Velocity method and lagrangian formulation for the computation of the shape hessian. *SIAM Journal on Control and Optimization*, 29(6):1414–1442, 1991.

[7] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metric, Analysis, Differential Calculus, and Optimization*. Society of Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2011.

[8] M.C. Delfour and J.P Zolésio. Structure of shape derivatives for nonsmooth domains. *Journal of Functional Analysis*, 104(1):1–33, 1992.

[9] F.R. Desaint and J.-P. Zolésio. Manifold derivative in the Laplace–Beltrami equation. *Journal of Functional Analysis*, 151(1):234–269, 1997.

[10] M.P. Do Carmo. *Differentialgeometrie von Kurven und Flächen*. Vieweg, Brauschweig/Wiesbaden, third edition, 1993.

[11] S. Dohr, K. Niino, and O. Steinbach. *Space-time methods*, volume 25, chapter 1: Space-time boundary element methods for the heat equation, pages 1–60. De Gruyter, Berlin/Boston, 2019.

[12] R. Dziri and J.-P. Zolésio. Dynamical shape control in non-cylindrical hydrodynamics. *Inverse Problems*, 15(1):113–122, 1999.

[13] R. Dziri and J.-P. Zolésio. Dynamical shape control in non-cylindrical Navier-Stokes equations. *Journal of Convex Analysis*, 6(2):293–318, 1999.

[14] R. Dziri and J.-P. Zolésio. Eulerian derivative for non-cylindrical functionals. In M. P. Polis J. Cagol and J.-P. Zolésio, editors, *Shape optimization and optimal design*, pages 87–107. Lecture notes in pure and applied mathematics, Marcel Dekker, Inc., New York, Basel, 2001.

[15] A. Friedman. Remarks on the maximum principle for parabolic equations and its applications. *Pacific Journal of Mathematics*, 8(2):201–211, 1958.

[16] A. Friedman and D. Kinderlehrer. A one phase Stefan problem. *Indiana University Mathematics Journal*, 24(11):1005–1035, 1975.

[17] A. Gray. *Modern differential Geometry of Curves and Surfaces with Mathematica*. CRC Press, Boca Raton, FI, 3rd edition, 1991.

[18] S.C. Gupta. *The classical Stefan problem: basic concepts, modelling and analysis*, volume 45 of *North-Holland Series in Applied Mathematics and Mechanics*. Elsevier, Amsterdam, 2003.

[19] M. Hadžić and S. Shkoller. Global stability and decay for the classical Stefan problem. *Communications on Pure and Applied Mathematics*, 68(5):689–757, 2014.

[20] M. Hadžić and S. Shkoller. Well-posedness for the classical Stefan problem and the zero surface tension limit. *Archive for Rational Mechanics and Analysis*, 223(1):213–264, 2016.

[21] H. Harbrecht. *Wavelet Galerkin schemes for the boundary element method in three dimensions*. PhD thesis, Technische Universität Chemnitz, 2001.

[22] H. Harbrecht. Analytical and numerical methods in shape optimization. *Mathematical Methods in the Applied Sciences*, 31(18):2095–2114, 2008.

[23] H. Harbrecht, M. Peters, and M. Siebenmorgen. Analysis of the domain mapping method for elliptic diffusion problems on random domains. *Numerische Mathematik*, 134(4):823–856, 2016.

[24] O.A. Ladyzenskaja, V.A. Solonnikov, and N.N. Ural'Ceva. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, Providence, RI, 1968.

[25] J.L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications II*. Springer, Berlin-Göttingen-Heidelberg, 1972.

[26] J. Lujano and J. Tausch. A shape optimization method for moving interface problems governed by the heat equation. *Journal of Computational and Applied Mathematics*, 390:113266, 2021.

[27] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, NY, 2000.

[28] A.M. Meirmanov. *The Stefan problem*. Walter de Gruyter, Berlin, 1992.

[29] M. Moubachir and J.-P. Zolésio. *Moving Shape Analysis and Control*. Chapman & Hall / CRC, Tayler & Francis Group, USA, 2006.

[30] J.-C. Nédélec. *Acoustic and electromagnetic equations: integral representations for harmonic problems*. Springer, New York, 2001.

[31] J. Nocedal and S.T. Wright. *Numerical Optimization*. Springer Science+Business Media, LLC, second edition, 2006.

[32] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, New York, 2003.

[33] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical recipes in FORTRAN 77. *Press Syndicate*, volume 1, 1992.

[34] L.I. Rubinstein. *The Stefan Problem*, volume 27. American Mathematical Society, Providence, Rhode Island, 1971. Translated from the Russian by A.D. Solomon.

[35] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization*. Springer, Berlin-Heidelberg, 1992.

[36] J. Stefan. Über die Theorie der Eisbildung, insbesondere über die Eisbildung im Polarmeere. *kais. Acad. d. Wiss. in Wien*, 98:269–286, 1889.

[37] D.A. Tarzia. A bibliography on moving-free boundary problems for the heat-diffusion equation. *MAT-Serie A*, 2, 1988.

[38] J. Tausch. Nyström method for BEM of the heat equation with moving boundaries. *Advances Compt. Math.*, 2019.

[39] J.A. Trangenstein. *Numerical solution of elliptic and parabolic partial differential equations*. Cambridge University Press, Cambridge, NY, 2013.

[40] A. Visintin. *Handbook of Differential Equations: Evolutionary Equations*, volume IV, chapter 8, pages 377–484. Elsevier, Amsterdam, 2008.

[41] J. Wloka. *Partial differential equations*. Cambridge University Press, Cambridge, NY, 1987.

[42] J.-P. Zolésio. *Identification de domaines par déformations*. PhD thesis, Université de Nice, 1979.