Quasilinear Heat Equation in Three Dimensions and Stefan Problem in Permafrost Soils in the Frame of Alternating Directions Finite Difference Scheme

Taras A. Dauzhenka and Igor A. Gishkeluk

Abstract—The quasilinear heat equation with thermal conductivity and heat capacity depending on the temperature field in three spatial dimensions is studied in application to the phase transition problem in permafrost soils. The conditions under which the alternating directions Douglas – Rachford finite difference scheme retains numerical stability are explicitly formulated. The comparison with the known analytical similarity solution to the Stefan problem in one spatial dimension is performed.

Index Terms—quasilinear heat equation, Stefan problem, finite differences, alternating directions scheme, numerical stability.

I. INTRODUCTION

SINCE the early formulations of the alternating directions implicit methods (ADI methods) [1], [2], they have been tremendously developed and found a vast number of applications [3], [4]. Nevertheless, serious difficulties are encountered with the use of these methods in application to problems with complex geometries [5] and/or nonlinear equations of mathematical physics [6].

Although the schemes of the ADI methods are proved to be efficient and economic with respect to time consumption and, in most cases, unconditionally stable, they exhibit some disadvantages:

1) Their finite-difference formulations permit to consider only rectangular spatial domains (due to commutativity requirements imposed on the factorized and split operators) [7];

2) The application of the ADI schemes to the problems with Neumann and Robin boundary conditions that are varying in time encounters serious problems due to the necessity of evaluation or approximation of these boundary conditions at the intermediate steps of the scheme [8];

3) When applied to the solution of nonlinear heat equations, the operators constituting an ADI scheme do not commute, thus leading to the loss of unconditional stability of the scheme [6].

The first of the above disadvantages can be overcome either by the use of finite elements methods in conjunction with operator splitting techniques, or by domain decomposition techniques. The second and the third disadvantages

present an important problem for the successful application of the ADI scheme. To the best of our knowledge, no complete stability analysis for an ADI scheme applied to the nonlinear heat equation in a three-dimensional spatial domain is available in the literature, thus motivating this work.

Another motivation for the present work is the application of ADI scheme to the modeling of heat transfer in large scale environmental systems (e.g., large areas of permafrost soils) which, in the case of purely explicit finite-differences schemes, imposes stiff constraints on the time-step in order to guarantee the numerical stability. At the same time, implementation of implicit schemes can often lead to much greater computing expenses than that of explicit schemes, especially for the problems with rapidly changing coefficients in complex geometries and substantially nonhomogeneous meshes. Thus, in modeling of heat transfer in large scale systems the necessity of making an optimal choice between explicit and implicit schemes arises. In case of finite-elements method, applied to modeling of processes in permafrost soils, the analysis of numerical stability appears to be so complex, that the stability criterion is often established empirically [9].

In the present paper we discuss the application of the ADI Douglas – Rachford scheme to the solution of Stefan problem in porous permafrost soils. The paper is organized as follows: next section contains the problem formulation and some assumptions that will be used in the proof of numerical stability of the ADI scheme while section 3 exposes the proof itself. Section 4 presents some numerical results and is followed by Conclusions.

II. PROBLEM SETUP AND FORMALISM

A. Quasilinear heat equation

The equation describing heat transfer in a system with phase transition has the following form:

$$C(u)\frac{\partial u}{\partial t} = \nabla(k(u)\nabla(u)), \tag{1}$$

where $u = u(x, y, z, t)$ is the unknown function (temperature: $u(t_\mathrm{f}) \in \mathcal{C}_+^2(\bar{\Omega})$ at every fixed t_f, $\mathcal{C}_+^2(\bar{\Omega})$ being a real linear space of positive valued functions), $C(u)$ is the heat capacity, $k(u)$ is thermal conductivity. By Ω we denote a bounded, open subset of an Euclidian space E^3 with boundary $\dot{\Omega}$, the closure of Ω being denoted by $\bar{\Omega}$. Following the classification given in [10], we call Eq.(1) a quasilinear heat equation. The initial and boundary conditions are taken to be:

$$u(x, y, z, t = 0) = f_0(x, y, z)|(x, y, z) \in \Omega, \qquad (2)$$

$$u(x, y, z, t) = f(x, y, z, t)|(x, y, z) \in \dot{\Omega}, t \in R_+. \qquad (3)$$

We consider the cases when $C(u) \geq \mu > 0$ and $k(u) \geq \nu > 0$, thus Eq. (1) is uniformly parabolic and has a unique solution to the initial-boundary-value problem Eq. (3)–(2). [11].

B. Stefan problem in permafrost soils

Basing on results of [12], we formulate the model for the Stefan problem without explicitly invoking the front-tracking condition. This approach is justified by the observations that in case of explicit front-tracking models applied to Stefan problem in salted permafrost soils, there appears an over-cooled region (frozen fringe zone) [13] (the analysis of this zone and related frost heave and cryogenic suction processes [9] are beyond the scope of the present work). Thus, we take into account the phase transition by introducing effective heat capacity $C_{\text{eff}}(u)$ which incorporates the latent heat per unit mass L_0:

$$\begin{cases} C_{\text{eff}}(u) = C_0(u) - (1 - w(u)) \cdot \rho_w \cdot L_0 \cdot \frac{dw(u-0)}{du}, \\ C_0(u) = C_{\text{th}} + (C_{\text{f}} - C_{\text{th}}) \cdot w(u) \end{cases}$$
$$(4)$$

$$w(u) = \begin{cases} 1 - \frac{1}{1+S \cdot (T_{\text{ph}} - u)}, & u \leq T_{\text{ph}}; \\ 0, & u > T_{\text{ph}}, \end{cases} \qquad (5)$$

where C_{th} and C_{f} are the values of volumetric heat capacities of the soil in thawed and frozen phases respectively, $w(u)$ is the fraction of frozen water, S is the smoothing parameter, T_{ph} is the phase transition temperature, ρ_w is the density of water. The thermal conductivity is taken to be

$$k(u) = k_{\text{th}} + (k_{\text{f}} - k_{\text{th}}) \cdot w(u), \qquad (6)$$

where k_{th} and k_{f} are the values of thermal conductivities in thawed and frozen phases respectively.

C. The linear space $\mathcal{C}(\Omega_h)$

For the formulation of finite difference scheme we introduce the following discretization procedure. Let the vector $h = (h_x, h_y, h_z) \in E^3$ have positive coordinates and G_h be the set of all points $(i \cdot h_x, j \cdot h_y, k \cdot h_z) \in E^3$, the indices i, j, k being integers. Two points $x_1 = (i_1 \cdot h_x, j_1 \cdot h_y, k_1 \cdot h_z)$ and $x_2 = (i_2 \cdot h_x, j_2 \cdot h_y, k_2 \cdot h_z)$ belonging to G_h are called neighbors if $\sqrt{(i_1 - i_2)^2 \cdot h_x^2 + (j_1 - j_2)^2 \cdot h_y^2 + (k_1 - k_2)^2 \cdot h_z^2} = h_{x,y,z}$. The points $x \in G_h \bigcap \Omega$ all of whose neighbor belong to $\bar{\Omega}$ are denoted by Ω_h. Following [14], we denote the points $x \in G_h \backslash \Omega_h$ with the property that at least one neighbor belongs to Ω_h by $\dot{\Omega}_h$. Thus, the full spatial mesh is $\bar{\Omega}_h = \Omega_h \bigcup \dot{\Omega}_h$, $\dot{\Omega}_h$ being the set of boundary points (outside of Ω_h). For the following we assume that $\bar{\Omega}_h$ is a homogeneous (i.e., h_x, h_y, h_z are constants) cubic domain.

Let $T = [0, \infty)$ and $T^0 = T \backslash \{0\}$. Then the time mesh is defined as follows:

$$T_\tau = \{t \in T | t = m \cdot \tau, n = 0, 1...\}. \qquad (7)$$

The approximate solutions of Eq. (1), $U(t)$, are defined on T_τ and take their values in a real finite-dimensional linear space $\mathcal{C}(\Omega_h)$, the dimension of which is equal to the number of points in Ω_h. A function $U(t) \in \mathcal{C}(\bar{\Omega}_h)$ is called admissible if $U(t = 0) = f_0(x, y, z)$ and $U = u(x, y, z, t)$ on $\dot{\Omega}_h \times T_\tau^0$. As was pointed out in [14], $u(x, y, z, t)$ is, in general, not known in $\dot{\Omega}_h \times T_\tau^0$ and thus the following assumption should be made: there exists a null sequence (that converges to zero) $\{h_{x,y,z}^\beta\}$ of mesh spacings such that $\dot{\Omega}_{h^\beta} \subset \dot{\Omega}$ and $h_{x,y,z}$ always belongs to the sequence $\{h_{x,y,z}^\beta\}$. This assumption guarantees that an admissible function is uniquely specified in $(G_h \times T_\tau^0) \backslash (\Omega_h \times T_\tau^0)$ in terms of the initial and boundary values of $u(x, y, z, t)$.

For every $U(t) \in \mathcal{C}(\bar{\Omega}_h)$ we define the first-order forward and backward difference operators:

$$\begin{cases} h_\alpha \cdot D_\alpha^+ U(t) = U_{i_\alpha+1}(t) - U_{i_\alpha}(t), \\ h_\alpha \cdot D_\alpha^- U(t) = U_{i_\alpha-1}(t) - U_{i_\alpha}(t). \end{cases} \qquad (8)$$

where $\alpha = \{x, y, z\}$, $2 \leq i_\alpha \leq (\dim_\alpha(\bar{\Omega}_h) - 1)$, $\dim_\alpha(\bar{\Omega}_h)$ being the number of points in $\bar{\Omega}_h$ along α-direction in the Cartesian coordinates. For any suitably defined function $U(t) \in \mathcal{C}(\bar{\Omega}_h)$ we set

$$k_{i_\alpha \pm 1/2, n} = k(\frac{U_{i_\alpha \pm 1} + U_{i_\alpha}}{2}). \qquad (9)$$

Thus, $k_{i_\alpha \pm 1/2, n}$ are the values of $k(U(t))$ at the fictitious intermediate nodes of the mesh Ω_h at $t = \tau \cdot n \in T_\tau, U(t) \in \mathcal{C}(\bar{\Omega}_h)$.

From the Eqs. (8)–(9) one obtains:

$$-D_{i_\alpha}^+(k_n(U) \cdot D_{i_\alpha}^- U) = \frac{1}{h_{i_\alpha}}(\Lambda_{i_\alpha, n}^+ + \Lambda_{i_\alpha, n}^-)U, \qquad (10)$$

where

$$\begin{cases} \Lambda_{i_\alpha, n}^- U = k_{i_\alpha-1/2, n} \cdot \frac{U_{i_\alpha-1} - U_{i_\alpha}}{h_{i_\alpha}}, \\ \Lambda_{i_\alpha, n}^+ U = k_{i_\alpha+1/2, n} \cdot \frac{U_{i_\alpha+1} - U_{i_\alpha}}{h_{i_\alpha}}, \end{cases} \qquad (11)$$

and $k_n(U)$ corresponds the value of $k(U(t))$ at $t = \tau \cdot n \in T_\tau, U(t) \in \mathcal{C}(\bar{\Omega}_h)$, i_α and α being defined as $\alpha = \{x, y, z\}$; $1 \leq i_\alpha \leq \dim_\alpha(\bar{\Omega}_h)$. We also provide $\mathcal{C}(\bar{\Omega}_h)$ with the l_2-inner product (U, V) and induced norm $\|U\| = (U, U)^{1/2}$ on $\bar{\Omega}_h$:

$$(U, V) = h_x \cdot h_y \cdot h_z \sum_{p \in G_h} U(p)V(p). \qquad (12)$$

The maximum norm is defined as following:

$$\|U\|_\infty = \max(U). \qquad (13)$$

We define on $\bar{\Omega}_h$ a modified A-inner product $(U, V)_{n,A}$ and induced norm $\|U\|_{n,A} = (U, U)_{n,A}^{1/2}$ by

$$(U, V)_{n,A} = h_x \cdot h_y \cdot h_z \sum_{\alpha=x,y,z} \sum_{p \in G_h} U(p)A_{n,\alpha}V(p), \qquad (14)$$

$$A_{n,\alpha}(V(p)) = -C_n^{-1}(V(p)) \cdot D_\alpha^+(k_n(V(p))D_\alpha^- V(p)), \qquad (15)$$

where $C_n^{-1}(V(p))$ and $k_n(V(p))$ correspond the values of $C^{-1}(v(x,y,z,t))$ and $k(v(x,y,z,t))$ at $p = (x,y,z) \in G_h, t = \tau \cdot n \in T_\tau, v(t) \in \mathcal{C}(\bar{\Omega}))$ respectively. Note, that the so defined operator $A_{n,\alpha}$ implies that it is generally not true that $A_{n,\alpha} : \mathcal{C}(\bar{\Omega}_h) \to \mathcal{C}(\bar{\Omega}_h)$ (see a discussion in the next subsection).

D. Alternating direction finite difference scheme

Let us now put in correspondence to Eq. (1) the following finite differences scheme:

$$(1+\tau\cdot A_{n,x})(1+\tau\cdot A_{n,y})(1+\tau\cdot A_{n,z})U_\tau^n = \sum_\alpha A_{n,\alpha}(U^n),$$
(16)

where $A_{n,\alpha}$ are given by Eq. (15), $\alpha = \{x,y,z\}$, $U_\tau^n = (U^{n+1} - U^n)/\tau$, $U^n = U(t)$ and $U^{n+1} = U(t+\tau)$ at $t = n \cdot \tau \in T_\tau$, $U(t) \in \mathcal{C}(\bar{\Omega}_h)$. Taking into account Eq. (8)–(11) and (15), Eq. (16) can be rewritten in the following form:

$$\begin{cases} C(U^n)\frac{U^*-U^n}{\tau} = \frac{1}{h_x}\Lambda_{i_x,n}U^* + \frac{1}{h_y}\Lambda_{i_y,n}U^n + \frac{1}{h_z}\Lambda_{i_z,n}U^n, \\ C(U^n)\frac{U^{**}-U^*}{\tau} = \frac{1}{h_y}\Lambda_{i_y,n}(U^{**} - U^n), \\ C(U^n)\frac{U^{n+1}-U^{**}}{\tau} = \frac{1}{h_z}\Lambda_{i_z,n}(U^{n+1} - U^n), \end{cases}$$
(17)

where $\Lambda_{i_x,n} = \Lambda_{i_x,n}^+ + \Lambda_{i_x,n}^-$ (same for i_y, i_z), U^* and U^{**} are the auxiliary functions that are used at the intermediate steps. Note that the indices i_x, i_y, i_z in the first equation of system (17) change in the following ranges $2 \leq i_x \leq (\dim_x(\bar{\Omega}_h) - 1)$, $2 \leq i_y \leq (\dim_y(\bar{\Omega}_h) - 1)$, $2 \leq i_z \leq (\dim_z(\bar{\Omega}_h) - 1)$, while for the last two equations in (17) $1 \leq i_x \leq \dim_x(\bar{\Omega}_h)$, $2 \leq i_y \leq (\dim_y(\bar{\Omega}_h) - 1)$, $1 \leq i_z \leq \dim_z(\bar{\Omega}_h)$ and $1 \leq i_x \leq \dim_x(\bar{\Omega}_h)$, $1 \leq i_y \leq \dim_y(\bar{\Omega}_h)$, $2 \leq i_z \leq (\dim_z(\bar{\Omega}_h) - 1)$ respectively. The scheme (17) has been described in [1] and is called the Douglas-Rachford scheme. It has been pointed out by Douglas in [1] that the scheme (17) with $\Lambda_{i_x,n}$ defined by Eq. (11) is locally second-order correct in space, unless the boundary conditions lower the accuracy. From Eq. (16) one can see that Eq. (1) is approximated with the accuracy $O(\tau + h_x^2 + h_y^2 + h_z^2)$ in Ω_h.

The boundary and initial conditions (3)–(2) are formulated on the mesh $\bar{\Omega}_h$ in the following way:

$$\begin{cases} U(p,t) = f_h(p,t)|p \in \dot{\Omega}_h, t = n \cdot \tau \in T_\tau, \\ U(p) = f_h^0(p)|p \in \Omega_h, t = 0 \in T_\tau. \end{cases}$$
(18)

where $f_h, f_h^0 \in \bar{\Omega}_h$.

It is well known that the scheme (16) is unconditionally stable when C_n^{-1} and k_n in Eq. (15) are constants (i.e., $C_n^{-1} = C^{-1}$ and $k_n = k$) and $A_{n,\alpha}$ are the finite-difference forms of the Laplace operators. This is readily seen from Eq.(16) and Eq. (14) if noticed that $A_{n,\alpha} = A_\alpha$ are self-adjoint positive definite operators, $A_\alpha = A_\alpha^* > 0$. In this case one can use the following Theorem proved in [15]:

Theorem 1: If operators $[(1+\tau\cdot A_x)(1+\tau\cdot A_y)(1+\tau\cdot A_z)]$ and $\sum_\alpha A_\alpha$ in Eq.(16) are self-adjoint, do not change in time, and $\sum_\alpha A_\alpha$ is positive definite, then the condition

$$(1+\tau\cdot A_x)(1+\tau\cdot A_y)(1+\tau\cdot A_z) \geq \frac{\tau}{2}\sum_\alpha A_\alpha, \quad (19)$$

is necessary and sufficient for the following estimate to hold:

$$\|U^n\|_A \leq \|U^{n=0}\|_A, \quad (20)$$

where $\|U^n\|_A$ is the norm induced by the A-inner product (14) in the case of constant C_n^{-1} and k_n, $U^{n=0}$ is the initial condition, and U^n is the numerical solution obtained at the n-th iteration.

Indeed, if $C_n^{-1} = C^{-1}$ and $k_n = k$ are constants, the operators A_α defined by Eq. (15) are pairwise commutative and thus operator relation (19) holds for any small h_x, h_y, h_z, τ. Relation (20) means that the scheme (16) is numerically stable when C_n^{-1} and k_n are constants.

The introduction of C_n^{-1} and k_n that depend on U^n makes the operators $A_{n,\alpha}$ defined by (15) nonlinear, i.e. it is generally not true that $A_{n,\alpha} : \mathcal{C}(\bar{\Omega}_h) \to \mathcal{C}(\bar{\Omega}_h)$. This fact makes it impossible to prove numerical stability in a single norm, as it is done in Eq.(20). In order to proceed, we slightly modify the definition of operator $A_{n,\alpha}$ and make the following assumptions:

Assumption 1: In the ADI scheme (16) the operators $A_{n,\alpha}$ are acting on the vectors U^{n+1} in the following way:

$$A_{n,\alpha}(U^{n+1}) = -C^{-1}(U^n) \cdot D_\alpha^+(k(U^n)D_\alpha^- U^{n+1}). \quad (21)$$

This assumption leads to consideration of a linear finite-difference scheme of type (16) for the heat equation with variable coefficients C_n and k_n, the forms of which depend on the temperature field U^n calculated at the previous time level. This kind of assumption is commonly used when considering nonlinear time dependent problems (e.g., [16]).

Let us rewrite (21) in the following equivalent form:

$$A_{n,\alpha}(V(p)) = \bar{A}_{n,\alpha}(V(p)) + \delta_{n,\alpha}(V(p)), \quad (22)$$

where

$$\begin{cases} \bar{A}_{n,\alpha}(V(p)) = \frac{-K_{i_\alpha,n}^+ D_\alpha^+(D_\alpha^- V(p))}{C_n}, \\ \delta_{n,\alpha}(V(p)) = \frac{K_{i_\alpha,n}^-(D_\alpha^+ - D_\alpha^-)V(p)}{C_n}. \end{cases}$$
(23)

$$\begin{cases} K_{i_\alpha,n}^+ = (k_{i_\alpha+1/2,n} + k_{i_\alpha-1/2,n})/2, \\ K_{i_\alpha,n}^- = (k_{i_\alpha+1/2,n} - k_{i_\alpha-1/2,n})/(2\cdot h_\alpha). \end{cases}$$
(24)

Assumption 2: The functions $C^{-1} : \mathcal{C}(\bar{\Omega}_h) \to \ell^\infty(\bar{\Omega}_h)$, $K^+ : \mathcal{C}(\bar{\Omega}_h) \to \ell^\infty(\bar{\Omega}_h)$ and $K^- : \mathcal{C}(\bar{\Omega}_h) \to \ell^\infty(\bar{\Omega}_h)$ (defined by (4)–(6), (9) and (24)) are mapping the elements $U^n(p) \in \mathcal{C}(\bar{\Omega}_h)$ into a sequence space $\ell^\infty(\bar{\Omega}_h)$, the elements of which $(K_n^+(p), K_n^-(p), C_n^{-1}(p) \in \ell^\infty(\bar{\Omega}_h))$ satisfy the following relations:

$$\begin{cases} C^{-1}(U^n(p)) \equiv C_n^{-1}(p), \\ |D_\alpha^+(C_n^{-1}(p))| \leq c_1 \cdot C_n^{-1}(p), \\ |D_\alpha^-(C_n^{-1}(p))| \leq c_1 \cdot C_n^{-1}(p), \end{cases}$$
(25)

$$\begin{cases} K^+(U^n(p)) \equiv K_n^+(p) = \frac{k_{i_\alpha+1/2,n}(p)+k_{i_\alpha-1/2,n}(p)}{2}, \\ |D_\alpha^+(K_n^+(p))| \leq c_2 \cdot K_n^+(p), \\ |D_\alpha^-(K_n^+(p))| \leq c_2 \cdot K_n^+(p), \end{cases}$$
(26)

$$\begin{cases} K^-(U^n(p)) \equiv K_n^-(p) = \frac{k_{i_\alpha+1/2,n}(p)-k_{i_\alpha-1/2,n}(p)}{2 \cdot h_\alpha}, \\ |D_\alpha^+(K_n^-(p))| \leq c_3 \cdot K_n^-(p), \\ |D_\alpha^-(K_n^-(p))| \leq c_3 \cdot K_n^-(p), \end{cases} \tag{27}$$

where the operators D_α^+, D_α^- are defined similarly to (8); c_1, c_2, c_3 are real positive constants. $\ell^\infty(\bar{\Omega}_h)$ is the Banach space with the maximum norm.

III. NUMERICAL STABILITY OF ADI SCHEME APPLIED TO THE NONLINEAR HEAT EQUATION

In this section we make use of Assumptions 1 and 2 and obtain conditions under which the scheme (16) with Eq.(21) can be successfully applied to the solution of nonlinear heat equation (1).

The main problem one has to consider is the following: obtain some relation between energy norms of operators $A_{n,x}, A_{n,y}, A_{n,z}$ that could guarantee numerical stability of (16). In order to proceed, let us accept the following definition of numerical stability given in [15]:

Definition 1: Let $D : H \to H$ be a positive definite self-adjoint operator acting in a (real) Hilbert space H. Let $\rho > 0$ be a constant, such that $\rho^n \leq M_1$ for any n from (7), $M_1 > 0$ being a constant that does not depend on h_α, n, τ. The scheme (16) is said to be stable in a Hilbert space H_A (with scalar product and norm defined by (14)) if for any $U^n \in H$ and for any sufficiently small $|h_\alpha|$ and τ, for the solution of equation $U^{n+1} = (E + \tau \cdot ((1 + \tau \cdot A_{n,x})(1 + \tau \cdot A_{n,y})(1 + \tau \cdot A_{n,z}))^{-1} \sum_\alpha A_{n,\alpha})U^n$ the following relation holds:

$$(AU^{n+1}, U^{n+1}) \leq \rho^2 \cdot (AU^n, U^n), \tag{28}$$

where E is the identity operator.

In order to prove the that the scheme (16) is stable in the above sense, we will need some auxiliary results proved in [15]:

Lemma 1: Let H be a Hilbert space (real or complex). Let $L : H \to H$ be a linear operator on H, such that $L_0 = 0.5 \cdot (L + L^*) > 0$. Then the relation

$$(L_0)^{-1} = \mathrm{Re}(L^{-1}) \geq \frac{\tau}{1 + \rho}E, \tag{29}$$

(with $\rho \geq 1$) is sufficient for

$$\|E + \tau \cdot L\| \leq \rho, \tag{30}$$

to hold.

Lemma 2: Let $L_n : H \to H$ be a time-dependent (for n from (7)), positive definite and self-adjoint operator. Then, relation

$$L_n \leq (1 + c_4 \cdot \tau)L_{n-1}, \tag{31}$$

is equivalent to relation

$$\|L_n^{1/2}L_{n-1}^{-1/2}\| \leq \sqrt{1 + c_4 \cdot \tau}, \tag{32}$$

where $c_4 > 0$ is a constant that does not depend on τ, h_α, n.

Now, based on the results of [15], we can formulate the main theorem:

Theorem 2: Let the operators $A_{n,\alpha}$ (defined by (21)) satisfy the following condition for any $n \in T_\tau$

$$(1 - c_4 \cdot \epsilon^n) \sum_\alpha A_{n-1,\alpha} \leq \sum_\alpha A_{n,\alpha} \leq (1 + c_4 \cdot \epsilon^n) \sum_\alpha A_{n-1,\alpha}, \tag{33}$$

where $\epsilon^n = \|(U^n - U^{n-1})\|_\infty$, $c_4 > 0$ is a constant. Then the relation

$$(1 + \tau \cdot A_{n,x})(1 + \tau \cdot A_{n,y})(1 + \tau \cdot A_{n,z}) \geq \frac{\tau}{1 + \rho} \sum_\alpha A_{n,\alpha}. \tag{34}$$

(with n from (7), $\rho \geq 1$) is sufficient for the following estimate to hold:

$$(\sum_\alpha A_{n,\alpha}U^{n+1}, U^{n+1}) \leq \rho^2(1 + c_4 \cdot \epsilon^n)(\sum_\alpha A_{n-1,\alpha}U^n, U^n), \tag{35}$$

where $U^n, U^{n+1} \in \mathcal{C}(\bar{\Omega}_h)$ are the solutions of Eq.(16).

Proof: Operators $A_{n,\alpha}$ (defined by (21)) are positive definite and self-adjoint, thus permitting to define $(\sum_\alpha A_{n,\alpha})^{\pm 1/2}$ and $(1 + \tau \cdot A_{n,\alpha})^{-1}$. Let us rewrite Eq.(16) in the following form:

$$\bar{U}^{n+1} = (E + \tau \cdot L_n)\hat{U}^n, \tag{36}$$

where

$$\begin{cases} L_n = (\sum_\alpha A_{n,\alpha})^{1/2}(B_n)^{-1}(\sum_\alpha A_{n,\alpha})^{1/2}, \\ B_n = (1 + \tau \cdot A_{n,x})(1 + \tau \cdot A_{n,y})(1 + \tau \cdot A_{n,z}), \end{cases} \tag{37}$$

$$\bar{U}^{n+1} = (\sum_\alpha A_{n,\alpha})^{1/2}U^{n+1}; \hat{U}^n = (\sum_\alpha A_{n,\alpha})^{1/2}U^n. \tag{38}$$

Now, using Lemma 1 and (34) we obtain:

$$\|(E + \tau \cdot L_n)\| \leq \rho. \tag{39}$$

Thus, for $\bar{U}^{n+1}$ and $\hat{U}^n$ one has:

$$\|\bar{U}^{n+1}\| \leq \|(E + \tau \cdot L_n)\| \cdot \|\hat{U}^n\| \leq \rho \cdot \|\hat{U}^n\|. \tag{40}$$

Taking into account Lemma 2 and the first equation of (38), we obtain:

$$\|\hat{U}^n\| = \|(\sum_\alpha A_{n,\alpha})^{1/2}(\sum_\alpha A_{n-1,\alpha})^{-1/2}\bar{U}^n\|, \tag{41}$$

$$\|(\sum_\alpha A_{n,\alpha})^{1/2}(\sum_\alpha A_{n-1,\alpha})^{-1/2}\bar{U}^n\| \leq \sqrt{1 + c_4 \cdot \epsilon^n}\|\bar{U}^n\|. \tag{42}$$

Hence,

$$\|\bar{U}^{n+1}\| \leq \rho\|\hat{U}^n\| \leq \rho \cdot \sqrt{1 + c_4 \cdot \epsilon^n}\|\bar{U}^n\|. \tag{43}$$

And, finally, returning to the scalar products, one obtains (35).

Let us now examine the conditions under which relation (34) holds. One can see that the l.h.s of (34) contains the terms $\tau^2 \cdot (A_{n,x}A_{n,y} + A_{n,x}A_{n,z} + A_{n,y}A_{n,z})$ and $\tau^3 A_{n,x}A_{n,y}A_{n,z}$ (in addition to $\tau \sum_\alpha A_{n,\alpha}$) which are required for the inequality to hold. Taking into account

Assumption 2, a straightforward calculation of a quantity $\tau^2 \cdot A_{n,\alpha} A_{n,\beta}$ gives us:

$$\tau^2 \cdot A_{n,\alpha} A_{n,\beta} \leq \frac{\tau^2}{h_\alpha^4} \cdot ((\frac{\max(K_n^+)}{\min(C_n)})^2 \cdot N_1 +$$

$$\frac{\max(h_\alpha \cdot K_n^-) \cdot \max(K_n^+)}{(\min(C_n))^2} \cdot N_2 +$$

$$(\frac{\max(h_\alpha \cdot K_n^-)}{\min(C_n)})^2 \cdot N_3), \qquad (44)$$

where

$$N_1 = (c_1^2 + 2 \cdot c_1 c_2 + c_2^2) \cdot P_1 +$$
$$(c_1 + c_2) \cdot (P_2 + P_3) + P_4, \qquad (45)$$
$$N_2 = (c_1^2 + 2 \cdot c_1 c_3 + c_2^3) \cdot P_5 +$$
$$(c_1 + c_3) \cdot (P_6 + P_7) + P_8 + P_9 - P_{10}, \qquad (46)$$
$$N_3 = P_{11} + P_{12}, \qquad (47)$$
$$P_1^2 = \max(|(U, h_\alpha^2 \cdot D_\beta^+ D_\beta^- U)|),$$
$$P_2^2 = \max(|(U, h_\alpha^3 \cdot D_\alpha^+ D_\beta^+ D_\beta^- U)|),$$
$$P_3^2 = \max(|(U, h_\alpha^3 \cdot D_\alpha^- D_\beta^+ D_\beta^- U)|),$$
$$P_4^2 = \max(|(U, h_\alpha^4 \cdot D_\alpha^+ D_\alpha^- D_\beta^+ D_\beta^- U)|),$$
$$P_5^2 = \max(|(U, h_\alpha \cdot (D_\beta^+ - D_\beta^-)U)|),$$
$$P_6^2 = \max(|(U, h_\alpha^2 \cdot D_\alpha^+ (D_\beta^+ - D_\beta^-)U)|),$$
$$P_7^2 = \max(|(U, h_\alpha^2 \cdot D_\alpha^- (D_\beta^+ - D_\beta^-)U)|),$$
$$P_8^2 = \max(|(U, h_\alpha^3 \cdot D_\alpha^+ D_\alpha^- (D_\beta^+ - D_\beta^-)U)|),$$
$$P_9^2 = \max(|(U, h_\alpha^3 D_\alpha^+ D_\beta^+ D_\beta^- U)|),$$
$$P_{10}^2 = \max(|(U, h_\alpha^3 D_\alpha^- D_\beta^+ D_\beta^- U)|),$$
$$P_{11}^2 = \max(|(U, h_\alpha^2 D_\alpha^+ (D_\beta^+ - D_\beta^-)U)|),$$
$$P_{12}^2 = \max(|(U, h_\alpha^2 D_\alpha^- (D_\beta^+ - D_\beta^-)U)|), \qquad (48)$$

the $\max()$ and $\min()$ being taken over all $p \in \bar{\Omega}_h$.

From (44)–(48) it follows that for the inequality (34) to hold, the sufficient condition is:

$$\frac{\rho \cdot \tau}{1 + \rho} \sum_\alpha A_{n,\alpha} \geq 3 \cdot \frac{\tau^2}{h_\alpha^4} \cdot ((\frac{\max(K_n^+)}{\min(C_n)})^2 \cdot (N_1 - P_4) +$$

$$\frac{\max(h_\alpha \cdot K_n^-) \cdot \max(K_n^+)}{(\min(C_n))^2} \cdot N_2 +$$

$$(\frac{\max(h_\alpha \cdot K_n^-)}{\min(C_n)})^2 \cdot N_3), \quad (49)$$

or, more rigorously:

$$\frac{\rho \cdot \tau}{(1 + \rho)h_\alpha^2} \frac{\min(K_n^+)}{\max(C_n)} \sum_\alpha \min(|(U, h_\alpha^2 D_\alpha^+ D_\alpha^- U)|) \geq$$

$$3 \cdot \frac{\tau^2}{h_\alpha^4} \cdot ((\frac{\max(K_n^+)}{\min(C_n)})^2 \cdot (N_1 - P_4) +$$

$$\frac{\max(h_\alpha \cdot K_n^-) \cdot \max(K_n^+)}{(\min(C_n))^2} \cdot N_2 +$$

$$(\frac{\max(h_\alpha \cdot K_n^-)}{\min(C_n)})^2 \cdot N_3). \quad (50)$$

From (50), the sufficient stability criterion (that relates $\tau, h_\alpha, \rho, C_n, K_n^+$ and K_n^- and guarantees that (34) holds on every time level $n \in T_\tau$) can be directly expressed as follows:

$$\frac{\tau}{h_\alpha^2} \leq \frac{\rho}{1 + \rho} \frac{\min(K_n^+)}{\max(C_n)} (\frac{\min(C_n)}{\max(K_n^+)})^2 \cdot c_0, \qquad (51)$$

where c_0 is a constant that can be deduced from (50). Note that (51) resembles the textbook expression for the stability criterion of a finite-differences scheme for a one-dimensional heat equation with variable coefficients: $\tau/h_\alpha^2 \leq \min(C_n)/(2 \cdot \max(K_n))$.

IV. NUMERICAL EXPERIMENTS

In this section we demonstrate the results of application of the ADI scheme (16) to the solution of a Stefan problem in a rectangular parallelepiped. We compare the obtained numerical results with the known analytical expressions of the similarity solutions for infinite domains [17].

Consider the following problem: a one-dimensional semi-infinite domain $\Omega = [0, \infty)$ contains the material in its liquid state in $\Omega_{liq}(t) = [0, s(t))$ and in its solid state in $\Omega_{sol}(t) = \Omega \backslash \bar{\Omega}_{liq}(t) = (s(t), \infty)$, $s(t)$ being the point that separates the liquid and the solid phases. The temperature at the point $z \in \Omega$ at time moment t is denoted by $u(z,t)$. The phase transition temperature is assumed to be $u(s(t), t) \equiv T_{ph} = 0$ degrees Celsius. We also assume zero heat fluxes at the ends of Ω and a piecewise constant initial condition:

$$u(z,0) = \begin{cases} u_{liq} = 10°\text{C}, & 0 \leq z \leq s(t=0); \\ T_{ph} = 0°\text{C}, & z = s(t=0); \\ u_{sol} = -1°\text{C}, & s(t=0) \leq z \leq \infty. \end{cases} \quad (52)$$

The formulated problem is known to have the following solution [17]:

$$u(z,t) = \begin{cases} -\frac{u_{liq} \cdot \text{erfc}(\alpha/\sqrt{K_{liq}})}{2 - \text{erfc}(\alpha/\sqrt{K_{liq}})} + \\ \frac{u_{liq} \cdot \text{erfc}(z - s(0))/(2\sqrt{K_{liq}t})}{2 - \text{erfc}(\alpha/\sqrt{K_{liq}})}, & z < s(t); \\ u_{sol} - \frac{u_{sol} \cdot \text{erfc}((z - s(0))/(2\sqrt{K_{sol}t}))}{\text{erfc}(\alpha/\sqrt{K_{sol}})}, & z \geq s(t). \end{cases} \quad (53)$$

where $s(t) = s(0) + 2 \cdot \alpha \sqrt{t}$ and the constant α is obtained by solving the equation:

$$\alpha = \frac{\sqrt{K_{sol}}}{\sqrt{\pi}L} \frac{u_{sol}}{\text{erfc}(\alpha/\sqrt{K_{sol}})} \exp(-\frac{\alpha^2}{K_{sol}}) +$$
$$\frac{\sqrt{K_{liq}}}{\sqrt{\pi}L} \frac{u_{liq}}{2 - \text{erfc}(\alpha/\sqrt{K_{liq}})} \exp(-\frac{\alpha^2}{K_{liq}}), \qquad (54)$$

where $\text{erfc}()$ is the complementary error function, L is the latent heat of phase transition, K_{sol} and K_{liq} are the thermal diffusivities in the solid and the liquid phases respectively which are assumed to be constant in time and position.

In our numerical experiment we have considered the spatial domain with dimensions: $0 \leq x \leq 15$, $0 \leq y \leq 15$, $2.9 \leq z \leq 20$ meters. The spatial mesh $\bar{\Omega}_h$ has the following nodes numbers: $4 \times 4 \times 172$ nodes. On all the faces of the parallelepiped the Neumann boundary conditions with zero heat fluxes have been imposed. The initial condition was taken to be (52) with $s(t = 0) = 10$ meters. Thermal conductivities and heat capacities of thawed and frozen phases that appear in (4)–(6) are taken to be: $C_{th} = 2240000$ $J/(\text{m}^3\text{K})$, $C_f = 2160000$ J/(m^{3}K), $k_{th} = 1.8$ W/(m · K),
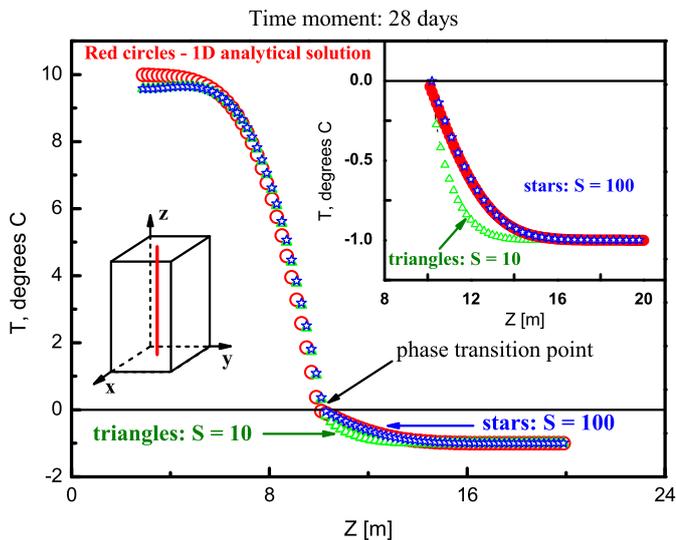
Fig. 1. The dependence of temperature on the Z-coordinate: comparison of numerical results (blue stars and green triangles), obtained with the scheme (16), and analytical expression of the similarity solution (53) (red circles). Time moment 28 days. Inset shows the same curves at temperatures below $T_{\rm ph} = 0$ Celsius. Red vertical line inside the parallelepiped indicates the set of points, for which the results are presented.

$k_{\rm f} = 2$ W/(m · K), $L_0 = 334$ kJ/(kg), $\rho_w = 1000$ kg/m^3, $S = 100$ K^{-1}. Numerical stability for the problem with these parameters has been achieved when the time step τ and spacial step $h \equiv \min(h_{x,y,z})$ were subject to the following condition: $\tau/h^2 \leq \min(C_{\rm eff}(u))/(20 \cdot \max(k(u)))$. This gave us the number of iterations: 6486.

Fig. 1 shows the numerical results along the line ($x = 5; y = 5; 2.9 \leq z \leq 20$) meters (i.e., along the central line of the parallelepiped) together with the corresponding analytical result obtained with (53). One can see both a qualitative and quantitative agreement between numerical and analytical results, the accuracy being within 0.5 degree Celsius. As expected, a better agreement is achieved at greater values of the smoothing parameter S in (5), which is responsible for the account of sharpness of the phase transition.

A good agreement between numerical and analytical results is also observed for the time evolution of temperature at a given point. Thus, Fig. 2 shows the time dependence of temperature at the point with coordinates ($x = 5, y = 5, z = 10.4$) meters.

V. CONCLUSIONS

We have shown that under the Assumptions 1 and 2, the uniform numerical stability (Definition 1) can be established for the Douglas – Rachford scheme (17) (which approximates the equation (1)), the general form of the sufficient stability criterion being (51). The used approach consists in reduction of a quasilinear equation to an equation with variable coefficients (subject to the constraints (25)–(27)), the forms of which change after each iteration. The comparison of obtained numerical results and an analytical solution (Fig. (1) - (2)) suggests that the ADI method can be successfully applied to the solution of Stefan problems in permafrost soils.
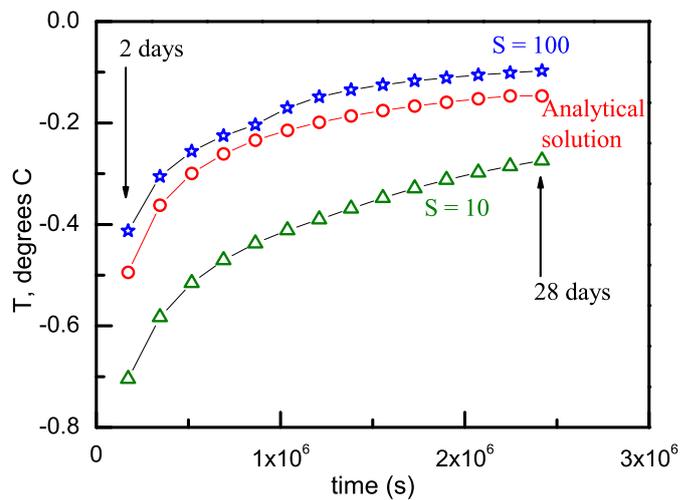
ACKNOWLEDGMENT

Fig. 2. The dependence of temperature on time t at the point ($x = 5, y = 5, z = 10.4$) meters: comparison of numerical results, obtained with the scheme (16), (blue stars and green triangles) and analytical expression of the similarity solution (53) (red circles). $2 \leq t \leq 28$ days.

REFERENCES

[1] J. Douglas, "Alternating direction methods for three space variables," *Numerische Mathematik*, vol. 4, no. 1, pp. 41–63, Dec. 1962.
[2] N. N. Yanenko, *The method of fractional steps;: The solution of problems of mathematical physics in several variables*, 1st ed. Springer-Verlag, 1971.
[3] G. I. Marchuk, *Splitting and Alternating Direction Methods*, ser. Finite Difference Methods. North Holland: Elsevier Science Publishers B.V., 1990, vol. 1, pp. 199–462.
[4] K. J. in 't Hout and B. D. Welfert, "Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms," *Applied Numerical Mathematics*, vol. 59, no. 3-4, pp. 677–692, Mar. 2009.
[5] P. Angot, J. Keating, and P. D. Minev, "A direction splitting algorithm for incompressible flow in complex geometries," *Computer Methods in Applied Mechanics and Engineering*, vol. 217-220, pp. 111–120, 2012.
[6] E. Hansen and A. Ostermann, "Dimension splitting for quasilinear parabolic equations," *IMA Journal of Numerical Analysis*, vol. 30, no. 3, pp. 857–869, 2010.
[7] G. Birkhoff and R. S. Varga, "Implicit alternating direction methods," *Trans. Amer. Math. Soc.*, vol. 92, pp. 13–24, 1959.
[8] B. Bialecki and R. Fernandes, "An alternating-direction implicit orthogonal spline collocation scheme for nonlinear parabolic problems on rectangular polygons," *SIAM Journal on Scientific Computing*, vol. 28, no. 3, pp. 1054–1077, 2006.
[9] H. R. Thomas, P. Cleall, Y. C. Li, C. Harris, and M. Kern-Luetschg, "Modelling of cryogenic processes in permafrost and seasonally frozen soils," *Géotechnique*, vol. 59, no. 3, pp. 173–184, 2009.
[10] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and quasi-linear equations of parabolic type*, ser. Translations of Mathematical Monographs. Providence, RI: American Mathematical Society, 1968, vol. 23.
[11] S. Kamenomostskaya, "On the Stefan problem." *Mat. Sb., N. Ser.*, vol. 53(95), no. 4, pp. 489 –514, 1961.
[12] A. A. Samarskii and P. N. Vabishchevich, *Mathematical Modelling, Volume 1, Computational Heat Transfer*, volume 1 ed. Wiley, 1996.
[13] L. Bronfenbrener and E. Korin, "Two-phase zone formation conditions under freezing of porous media," *Journal of Crystal Growth*, vol. 198-199, pp. 89–95, 1999.
[14] M. Lees, "Alternating direction and semi-explicit difference methods for parabolic partial differential equations," *Numerische Mathematik*, vol. 3, no. 1, pp. 398–412, 1961.
[15] A. A. Samarskii and A. V. Gulin, *Stability of Difference Schemes*. Moskow: Nauka, 1973.
[16] J. E. Dendy, "Alternating direction methods for nonlinear Time-Dependent problems," *SIAM J. Numer. Anal.*, vol. 14, no. 2, pp. 313–326, 1977.
[17] E. Javierre, C. Vuik, F. J. Vermolen, and S. van der Zwaag, "A comparison of numerical models for one-dimensional stefan problems," *Journal of Computational and Applied Mathematics*, vol. 192, no. 2, pp. 445–459, 2006.