ESAIM: M2AN
Vol. 39, N° 4, 2005, pp. 755–780
DOI: 10.1051/m2an:2005033

ESAIM: Mathematical Modelling and Numerical Analysis

FINITE ELEMENT APPROXIMATION FOR DEGENERATE PARABOLIC EQUATIONS. AN APPLICATION OF NONLINEAR SEMIGROUP THEORY

Akira Mizutani[1], Norikazu Saito[2] and Takashi Suzuki[3]

Abstract. Finite element approximation for degenerate parabolic equations is considered. We propose a semidiscrete scheme provided with order-preserving and L^1 contraction properties, making use of piecewise linear trial functions and the lumping mass technique. Those properties allow us to apply nonlinear semigroup theory, and the wellposedness and stability in L^1 and L^∞, respectively, of the scheme are established. Under certain hypotheses on the data, we also derive L^1 convergence without any convergence rate. The validity of theoretical results is confirmed by numerical examples.

Mathematics Subject Classification. 35K65, 47H20, 65M60.

Received: June 1, 2004. Revised: March 2, 2005.

1. Introduction

This paper is concerned with the finite element method applied to the initial-boundary value problem for degenerate parabolic equation,

$$\begin{cases} u_t - \Delta f(u) = 0 & \text{in } \Omega \times (0,T), \\ f(u) = 0 & \text{on } \partial\Omega \times (0,T), \\ u|_{t=0} = u_0(x) & \text{on } \Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, denotes a bounded domain with the Lipschitz boundary $\partial\Omega$, T an arbitrary positive constant, and f a non-decreasing continuous function defined on $\mathbb{R}$ satisfying $f(0) = 0$.

Problem (1.1) describes, for instance, the flow of homogeneous fluid in porous media if

$$f(u) = u\,|u|^{\gamma-1} \tag{1.2}$$

with $\gamma > 1$, the fast diffusion if (1.2) with $0 < \gamma < 1$, and the two phase Stefan problem in enthalpy formulation if

$$f(u) = \begin{cases} \alpha(u+1) & (u \le -1) \\ 0 & (-1 < u < 1) \\ \beta(u-1) & (u \ge 1) \end{cases} \tag{1.3}$$

with $\alpha, \beta > 0$. See, for more detail [14, 15, 32].

Keywords and phrases. Finite element method, degenerate parabolic equation, nonlinear semigroup.

[1] Department of Mathematics, Faculty of Science, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan.

[2] Faculty of Education, Toyama University, 3190 Gofuku, Toyama 930-8555, Japan. saito@edu.toyama-u.ac.jp

[3] Division of Mathematical Science, Department of System Innovation, Graduate School of Engineering Science, Osaka University, 1-1 Machikaneyama, Toyonaka, 560-0043, Japan.

L^1 theory to (1.1) was developed in early 1970's in use of nonlinear semigroups. To summarise it, we set $X = L^1(\Omega)$ and introduce operators L and A in X by $Lv = -\Delta v$ for $v \in D(L) = \{v \in W_0^{1,1}(\Omega) \mid Lv \in X\}$ and $Av = Lf(v)$ for $v \in D(A) = \{v \in X \mid f(v) \in D(L)\}$, respectively. Then, problem (1.1) is reduced to the nonlinear evolution equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} + Au = 0 \qquad \text{with} \qquad u(0) = u_0 \tag{1.4}$$

in X. Brezis-Strauss [6] proved that the operator $-A$ is *m-dissipative* in X. This means that

$$\|v - \hat{v}\|_{L^1(\Omega)} \le \|v - \hat{v} + \lambda Av - \lambda A\hat{v}\|_{L^1(\Omega)} \tag{1.5}$$

holds for $v, \hat{v} \in D(A)$ and $\lambda > 0$, and also $R(I + \lambda A) = L^1(\Omega) = \overline{D(A)}$. Then, theory of Crandall-Liggett [12] guarantees the generation of semigroup $\{S(t)\}_{t\ge 0}$ on X by

$$S(t) = \underset{m\to\infty}{s\text{-}\lim} \left(I + \frac{t}{m}A\right)^{-m}, \tag{1.6}$$

and $u(t) = S(t)u_0$ is regarded as the solution to (1.4). Another important property of A is the order-preserving, that is,

$$g \ge \hat{g} \qquad \Rightarrow \qquad (I + \lambda A)^{-1} g \ge (I + \lambda A)^{-1}\hat{g}. \tag{1.7}$$

Relations (1.5) and (1.7) are summarised as

$$\left\|[v - \hat{v}]_+\right\|_{L^1(\Omega)} \le \left\|[v - \hat{v} + \lambda Av - \lambda A\hat{v}]_+\right\|_{L^1(\Omega)} \tag{1.8}$$

for $v, \hat{v} \in D(A)$, where $[v]_+ = \max\{0, v\}$. This implies

$$\left\|[S(t)u_0 - S(t)\hat{u}_0]_+\right\|_{L^1(\Omega)} \le \left\|[u_0 - \hat{u}_0]_+\right\|_{L^1(\Omega)} \tag{1.9}$$

by (1.6), where $u_0, \hat{u}_0 \in X$ and $t \in [0, T]$. Inequality (1.9) means that $\{S(t)\}_{t\ge 0}$ is an order-preserving and L^1 contraction semigroup on X.

L^∞ stability of the resolvent,

$$\left\|(I + \lambda A)^{-1} g\right\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}, \tag{1.10}$$

is also proven in [6], where $g \in X \cap L^\infty(\Omega)$ and $\lambda > 0$. This implies L^∞ stability of the semigroup

$$\|S(t)u_0\|_{L^\infty(\Omega)} \le \|u_0\|_{L^\infty(\Omega)}, \tag{1.11}$$

where $u_0 \in X \cap L^\infty(\Omega)$ and $t \in [0, T]$.

So far, several schemes of time discretization have been examined. In fact, those structures of the problem, particularly (1.6), justify the backward difference approximation to (1.1), which was studied by [28]. Another scheme was obtained by the use of the nonlinear Chernoff formula of [5], where solution at each discrete time level is approximated by a linear elliptic equation. This approach was taken first by [3]. Whereas L^1 framework was employed in [3,28], L^2 error estimates were obtained by [20,26,27] for modified schemes of [3]. Those works were done in the literature of porous media or that of Stefan problems. For fast diffusion problems, we refer to [23,24].

On the other hand, for porous media and Stefan problems, fully discrete schemes where the space variable was discretized by finite element methods were also studied by many authors; [10,13,21,30,31,34,36]. Some of them gave error analysis in the H^{-1} framework. We will mention a few remarks on such schemes in the next section, after having presented our scheme.

The present paper deals with a spatial discretization for (1.1), that is,

$$\frac{\mathrm{d}u_h}{\mathrm{d}t} + A_h u_h = 0 \qquad \text{with} \qquad u_h|_{t=0} = u_{0h},$$

where A_h, u_h, and u_{0h} stand for the finite element approximations of A, u, and u_0, respectively.

Our purposes are twofold. Firstly, we introduce the spatial discretization A_h of A which preserves above mentioned properties. It can be done by making use of piecewise linear trial functions and the lumping mass technique, if a family of the triangulation $\{\mathcal{T}_h\}$ of Ω, $h > 0$ being the discretization parameter, is of acute type (the definition will be recalled in Sect. 3). Actually, in Sections 2 and 3, we introduce A_h and prove that A_h satisfies the discrete analogue of (1.8) in a suitable Banach space X_h, respectively. From this, we immediately obtain the nonlinear semigroup $\{S_h(t)\}_{t\geq 0}$ on X_h which is generated by $-A_h$ and satisfies the discrete analogue of (1.9). Moreover, as will be mentioned in Section 4, A_h and $S_h(t)$ are L^∞ stable as well as A and $S(t)$ are so.

The second purpose of this paper is to make error analysis. The goal of this end is to derive

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} \|u_h(t) - u(t)\|_{L^1(\Omega)} = 0. \tag{1.12}$$

Our main theorem (Th. 7.1) shows that (1.12) is valid, for example, if $\Omega \subset \mathbb{R}^2$ is convex, u_0 is continuous on $\overline{\Omega}$ with the value zero on $\partial\Omega$, f is strictly increasing, and $\{\mathcal{T}_h\}$ is provided with acuteness and quasi-uniformity. Further an extension of Theorem 7.1 to the case where f is nondecreasing is also discussed (Prop. 7.1 and Lem. 7.2). However we have no error estimates and they will be studied in subsequent works. Proof of (1.12) follows the principle that the convergence of semigroup is a consequence of that of resolvents. Thus, Sections 5–7 are devoted to the proof of the convergence of the resolvent, the Yosida approximation, and the semigroup, respectively.

Finally, in Section 8, we present results of numerical experiments for the porous media nonlinearity. The time discretization makes use of the forward difference formula. We observe that L^1 convergence of numerical solutions really takes place.

At this stage, we clarify our motivation of this work. As was mentioned above, several physical phenomena are modelled by (1.1), and therefore order-preserving and L^1 contraction properties are essential requirements from not only mathematical but also physical points of view. Consequently we are interested in discrete schemes which preserve such properties of the original problem. However it seems that little effort has been made in this direction. The first contribution of this paper is to give an discrete scheme enjoying a discrete version of order-preserving and L^1 contraction properties for a general nondecreasing f. Moreover our presented scheme is well suited for an actual computation. The second contribution is a convergence result of the form (1.12). Our result can be applied to porous media and fast diffusion nonlinearities (1.2). Especially we do not know any convergence results for a spatial discretization to the fast diffusion problem at present. On the other hand, for the Stefan nonlinearity (1.3), our convergence result may be restrictive, since f and u_0 are assumed to be strictly increasing and continuous, respectively. The main interest here is to reveal a general nature of convergence rather than to go into details under specific assumptions on f. The convergence result itself is to be expected from semigroup theory. But, as is well-known, fundamental theorems in nonlinear semigroup theory were established by quite technical and somewhat tricky arguments. Therefore, it is not obvious that the similar argument works for discrete problems. For example, the effect of perturbation on f causes a new issue which have not appeared in the continuous problem (see Rem. 7.1). Thus, in the present paper, we will develop a discrete nonlinear semigroup theory. Also we note that, concerning the regularity of solutions, we only have $u(t) \in X$ and $f(u(t)) \in W_0^{1,1}(\Omega)$, even if u_0 is continuous. Our argument does not require any redundant assumptions on the regularity of solutions.

Recently some of the authors and their colleague published the monograph [17], where finite element approximation to (1.1) on flat torus with uniform triangulation is studied. Some lemmas and theorems described below are proven similarly, but we shall give them for completeness. Furthermore, the method of [17] for the convergence of resolvent is restrictive, and we shall provide new arguments here.

We follow the standard notation of [1]. We put $\|\cdot\|_p = \|\cdot\|_{L^p(\Omega)}$ for $p \in [1,\infty]$. The space $W_0^{m,p}(\Omega)$ stands for the closure in $W^{m,p}(\Omega)$ of $C_0^\infty(\Omega)$, the set of C^∞ functions with compact supports in Ω. We write $H^m(\Omega)$ and $H_0^m(\Omega)$ instead of $W^{m,2}(\Omega)$ and $W_0^{m,2}(\Omega)$, respectively, for $m \geq 0$. The standard inner product of $L^2(\Omega)$ is denoted by $(\cdot,\cdot)$. Furthermore we set

$$W = \{v \in C(\overline{\Omega}) |\ v|_{\partial\Omega} = 0\}. \tag{1.13}$$

Generic positive constants depending on Ω are denoted by C, C_1, and so forth. If it is necessary to specify the dependence on other parameters, say γ_1, γ_2, $\cdots$, then we write $C(\Omega, \gamma_1, \gamma_2, \cdots)$. We shall use the same symbol I to indicate the identity operator on any space.

2. Finite element approximation

For the sake of simplicity, in what follows, we suppose that Ω is an n-dimensional polyhedron. We take a family of triangulations $\{\mathcal{T}_h\} = \{\mathcal{T}_h\}_{h\downarrow 0}$ defined on $\overline{\Omega}$, where each element $\sigma \in \mathcal{T}_h$ is a closed simplex. The maximum side length of all elements in $\mathcal{T}_h$ is denoted by h. We take the piecewise linear approximation, putting

$$X_h = \{\chi \in W \mid\ \chi \text{ is linear on } \sigma \text{ for each } \sigma \in \mathcal{T}_h\}.$$

Let I_h be the set of vertices of $\sigma \in \mathcal{T}_h$ belonging to Ω. For $a \in I_h$, the function $w_a \in X_h$ is defined by

$$w_a = \begin{cases} 1 & (\text{at } a) \\ 0 & (\text{at } b \in I_h \setminus \{a\}). \end{cases}$$

Then, $\{w_a \mid a \in I_h\}$ forms a basis of X_h and the interpolation operator $\pi_h : W \to X_h$ is defined by

$$\pi_h v = \sum_{a \in I_h} v(a) w_a.$$

Each $a \in I_h$ takes barycentric domain D_a. See [17], p. 203 for its precise definition. Let

$$\overline{w}_a(x) = \begin{cases} 1\ (x \in D_a) \\ 0\ (x \in \overline{\Omega} \setminus D_a), \end{cases}$$

and denote by $\overline{X}_h$ the vector space spanned by $\{\overline{w}_a \mid a \in I_h\}$. The linear transformation $M_h : X_h \to \overline{X}_h$, referred to as the lumping operator, is defined through $w_a \mapsto \overline{w}_a$. Sometimes, we shall write $\overline{\chi}_h$ for $M_h \chi_h$, where $\chi_h \in X_h$.

The semidiscrete scheme studied in this paper is to solve $u_h \in C^1([0,T]; X_h)$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}(\overline{u}_h, \overline{v_h}) + (\nabla \pi_h f(u_h), \nabla v_h) = 0 \qquad \text{with} \qquad (u_h(0), v_h) = (u_{0h}, v_h) \tag{2.1}$$

for any $v_h \in X_h$, where $u_{0h} \in X_h$ is an appropriate approximation of $u_0 \in X$. In order to convert (2.1) to the operator theoretic form, we introduce the following operators. Let $L_h : X_h \to X_h$ be the finite element approximation L defined by

$$(L_h \chi_h, v_h) = (\nabla \chi_h, \nabla v_h)$$

for $\chi_h, v_h \in X_h$. Let $M_h^* : \overline{X}_h \to X_h$ be the adjoint operator of M_h associated with the L^2 inner product, and set

$$K_h = M_h^* M_h : X_h \to X_h.$$

Then (2.1) is expressed as

$$K_h \frac{\mathrm{d}u_h}{\mathrm{d}t} + L_h \pi_h f(u_h) = 0 \qquad \text{with} \qquad u_h(0) = u_{0h}. \tag{2.2}$$

The operator M_h is invertible in X_h and hence $K_h^{-1} = M_h^{-1}(M_h^*)^{-1}$ is well-defined. Therefore, scheme (2.2) is equivalent to

$$\frac{\mathrm{d}u_h}{\mathrm{d}t} + A_h u_h = 0 \qquad \text{with} \qquad u_h(0) = u_{0h} \tag{2.3}$$

in X_h, where

$$A_h v = K_h^{-1} L_h \pi_h f(v)$$

is defined for $v \in W$.

We here describe some examples of $u_{0h} \in X_h$:

$$\begin{aligned} u_{0h} &= P_h u_0 \qquad (\text{if } u_0 \in L^2(\Omega)); \\ u_{0h} &= R_h u_0 \qquad (\text{if } u_0 \in H_0^1(\Omega)); \\ u_{0h} &= \pi_h u_0 \qquad (\text{if } u_0 \in W), \end{aligned}$$

where $R_h : H_0^1(\Omega) \to X_h$ and $P_h : L^2(\Omega) \to X_h$ denote the Ritz and the orthogonal projection operators. They are defined by

$$(\nabla(v - R_h v), \nabla \chi_h) = 0 \qquad (\chi_h \in X_h) \tag{2.4}$$

and

$$(v - P_h v, \chi_h) = 0 \qquad (\chi_h \in X_h), \tag{2.5}$$

respectively. Further, if $u_0 \in W_0^{1,1}(\Omega)$, we can apply Scott and Zhang's interpolation operator $\Pi_h : W_0^{1,1}(\Omega) \to X_h$ and take $u_{0h} = \Pi_h u_0$. (For the precise definition of Π_h, see [35]. A version of such interpolation is described in [4].) In convergence analysis presented below, we assume $u_0 \in W$ and take $u_{0h} = \pi_h u_0$.

Before concluding this section, we state a remark on another finite element scheme to (1.4). From the L^2 theoretical point of view, it may be natural to take

$$\frac{\mathrm{d}}{\mathrm{d}t}(u_h, v_h) + (\nabla f(u_h), \nabla v_h) = 0 \qquad \text{with} \qquad (u_h(0), v_h) = (u_{0h}, v_h) \tag{2.6}$$

for $v_h \in X_h$. In this case, the operator theoretic representation reads

$$\frac{\mathrm{d}u_h}{\mathrm{d}t} + L_h R_h f(u_h) = 0 \qquad \text{with} \qquad u_h(0) = u_{0h}.$$

If f is locally Lipschitz continuous, scheme (2.6) is well-defined, because then $v_h \in X_h$ implies $f(v_h) \in H_0^1(\Omega)$. Namely, in this case (2.6) is conforming and was studied by [21, 30, 31, 34] including its time discretizations. Based on the energy method, they discussed the stability, convergence, and error estimate in the L^2 norm for the porous media and the Stefan nonlinearities.

However, the linear part $L_h R_h$ does not have such properties as (3.8), (3.9) and (3.10) given below. Thus, in general, $v_h \in X_h \mapsto -L_h R_h f(v_h) \in X_h$ is not m-dissipative. This means that, even if an approximate solution converges to the original one, it is not certain that the approximate solution has order-preserving and L^1 contraction properties. On the contrary, $-A_h$ is m-dissipative as will be shown in the next section.

3. Wellposedness

We pose on $\{\mathcal{T}_h\}$ that

(H1) *Acuteness.* Given $\sigma \in \mathcal{T}_h$, a vertex $P_0 \subset \sigma$, and the opposite face $F \subset \sigma$ to P_0, let S be a plane including F. Then the foot of the perpendicular from P_0 to S is always included in $\overline{F}$.

Remark 3.1. If $n = 1$, (H1) always holds. If $n = 2$, it is equivalent to saying that each $\sigma \in \mathcal{T}_h$ is a right or an acute triangle. Generally, it corresponds to the non-negative type of Ciarlet and Raviart [9] or acuteness of Fujii [16].

This section is devoted to the proof of the following theorem, which is a discrete analogue of (1.8).

Theorem 3.1. *Assume that* (H1) *holds. Then we have*

$$\|M_h\pi_h[v_h - \hat{v}_h]_+\|_1 \le \left\|M_h\pi_h\left[v_h - \hat{v}_h + \lambda A_h v_h - \lambda A_h \hat{v}_h\right]_+\right\|_1, \tag{3.1}$$

for $v_h, \hat{v}_h \in X_h$ *and* $\lambda > 0$. *Furthermore, it holds that* $R(I + \lambda A_h) = X_h$.

This assures the unique solvability of (2.3). In fact, X_h forms a Banach space equipped with the norm

$$\|\chi_h\|_{1,h} = \int_\Omega M_h\pi_h\,|\chi_h| \tag{3.2}$$

for $\chi_h \in X_h$. Theorem 3.1 means that $-A_h$ is m-dissipative in X_h with respect to this norm. Therefore, from the generation theorem of [12], scheme (2.3) is uniquely solvable globally in time and the solution is given as $u_h(t) = S_h(t)u_{0h}$, where

$$S_h(t) = \lim_{m\to\infty}\left(I + \frac{t}{m}A_h\right)^{-m}. \tag{3.3}$$

Combining (3.1) with (3.3), we deduce

$$\|[S_h(t)u_{0h} - S_h(t)\hat{u}_{0h}]_+\|_{1,h} \le \|[u_{0h} - \hat{u}_{0h}]_+\|_{1,h} \tag{3.4}$$

for $u_{0h}, \hat{u}_{0h} \in X_h$ and $t \in [0, T]$. Therefore, it holds that

$$u_{0h} \ge \hat{u}_{0h} \quad\Rightarrow\quad S_h(t)u_{0h} \ge S_h(t)\hat{u}_{0h}$$

In particular, $S_h(t)u_{0h} \ge 0$ follows from $u_{0h} \ge 0$ and it holds that

$$\|S_h(t)u_{0h}\|_{1,h} \le \|u_{0h}\|_{1,h}.$$

At this stage, we assume that

(H2) *Regularity.* There is a positive constant ν_1 independent of h such that

$$\rho(\sigma) \ge \nu_1 d(\sigma)$$

for any $\sigma \in \mathcal{T}_h$, where $\rho(\sigma)$ and $d(\sigma)$ indicate diameters of the inscribing and the circumscribing balls of σ, respectively.

Under such a reasonable assumption, we have a constant $C > 0$ independent of h satisfying

$$C^{-1}\|\chi_h\|_1 \le \|\chi_h\|_{1,h} \le C\|\chi_h\|_1 \tag{3.5}$$

for $\chi_h \in X_h$. Hence, by (3.4), we obtain

$$\left\|[S_h(t)v_h - S_h(t)\hat{v}_h]_+\right\|_1 \le C\left\|[v_h - \hat{v}_h]_+\right\|_1.$$

Remark 3.2. Inequalities (3.5) follows from

$$C^{-1}\|\chi_h\|_{L^1(\sigma)} \le \|M_h\pi_h|\chi_h|\|_{L^1(\sigma)} \le C\|\chi_h\|_{L^1(\sigma)} \quad (\chi_h \in X_h). \tag{3.6}$$

for any $\sigma \in \mathcal{T}_h$. Because $\{\mathcal{T}_h\}$ is regular, inequality (3.6) is reduced to the case $\sigma = \hat{\sigma}$, where $\hat{\sigma}$ denotes the canonical reference element. The linear functions on $\hat{\sigma}$ form a finite dimensional vector space Y. The desired estimate holds because Y is isometric to an Euclidean space, and any two norms on Y are equivalent to each other. We also have

$$\|K_h\chi_h\|_p + \|K_h^{-1}\chi_h\|_p \le C\|\chi_h\|_p \quad (\chi_h \in X_h,\ 1 \le p \le \infty). \tag{3.7}$$

under (H2). See [17], p.174.

Before stating the proof of Theorem 3.1, we collect some inequalities concerning linear part $K_h^{-1}L_h$, which hold under (H1). They are shown in [17], Sect. 5.1 and the proof is omitted here. First, discrete maximum principle

$$\max_{\overline{\Omega}}(I+\lambda K_h^{-1}L_h)^{-1}v_h \le \max_{\overline{\Omega}} \pi_h[v_h]_+ \tag{3.8}$$

holds, where $v_h \in X_h$ and $\lambda > 0$. Here, well-definedness of $(I+\lambda K_h^{-1}L_h)^{-1} : X_h \to X_h$ is included. It follows from (3.8) that

$$0 \le v_h \in X_h,\ \lambda > 0 \qquad \Rightarrow \qquad (I+\lambda K_h^{-1}L_h)^{-1}v_h \ge 0. \tag{3.9}$$

Next, discrete L^1 contraction property is expressed as

$$0 \le v_h \in X_h,\ \lambda > 0\ \Rightarrow\ \int_\Omega M_h(1+\lambda K_h^{-1}L_h)^{-1}v_h \le \int_\Omega M_h v_h. \tag{3.10}$$

The proof of (3.8) and (3.10) is explicitly mentioned for the case $n = 2$ in [17]. However, the other cases $n = 1, 3$ can be done similarly under the assumption (H1). In (3.9) and (3.10), contribution of mass lumping is essential for $\lambda > 0$. If the consistent mass is employed, then (3.9) is restricted to the range $0 < h^2/\lambda \ll 1$, while property (3.10) is not certain to hold. See Ciarlet-Raviart [9] and Fujii [16] for the former fact.
Thanks to (3.9) and (3.10), we can prove the following inequality, which is comparable to Kato's one of [22].

Lemma 3.1. *Assume that* (H1) *holds. Then we have*

$$\int_\Omega M_h\pi_h\left[(K_h^{-1}L_h\pi_h v)\ \mathrm{sgn}^+ v\right] \ge 0 \tag{3.11}$$

for $v \in W$, where

$$\mathrm{sgn}^+ v = \begin{cases} 1 & (v \ge 0) \\ 0 & (v < 0). \end{cases}$$

Proof. First, we show that

$$\int_\Omega M_h\pi_h\left[\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h\right]_+ \le \int_\Omega M_h\pi_h\left[v_h\right]_+ \tag{3.12}$$

holds for $v_h \in X_h$ and $\lambda > 0$. In fact, taking

$$v_h^\pm \equiv \pi_h\left[v_h\right]_\pm = \pm\sum_{a\in I_h^\pm} v_h(a)w_a, \tag{3.13}$$

we have $0 \le v_h^\pm \in X_h$ and $v_h = v_h^+ - v_h^-$, where $I_h^\pm = \{a \in I_h \mid \pm v_h(a) \ge 0\}$ and $[\ \cdot\ ]_\pm = \max\{0, \pm\ \cdot\ \}$. This implies $\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h^\pm \ge 0$ by (3.9), and hence

$$\left[\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h\right]_+ \le \left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h^+.$$

Because π_h and M_h are order-preserving, we have

$$M_h\pi_h\left[\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h\right]_+ \le M_h\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h^+,$$

which implies

$$\int_\Omega M_h\pi_h\left[\left(I+\lambda K_h^{-1}L_h\right)^{-1}v_h\right]_+ \le \int_\Omega M_h v_h^+$$

by (3.10). This means (3.12).

Given $v \in W$, we take $v_h = \pi_h v$ and $u_h = \left(I + \varepsilon K_h^{-1} L_h\right)^{-1} v_h$ for $\varepsilon > 0$. Because of

$$v_h - u_h = \varepsilon \left(I + \varepsilon K_h^{-1} L_h\right)^{-1} K_h^{-1} L_h v_h,$$

we have

$$\begin{aligned}\varepsilon \int_\Omega M_h \pi_h \left[\left(\left(I + \varepsilon K_h^{-1} L_h\right)^{-1} K_h^{-1} L_h v_h\right) \cdot \operatorname{sgn}^+ v_h\right] &= \int_\Omega M_h \pi_h \left[(v_h - u_h) \cdot \operatorname{sgn}^+ v_h\right] \\ &= \int_\Omega M_h \pi_h \left[v_h\right]_+ - \int_\Omega M_h \pi_h \left[u_h \cdot \operatorname{sgn}^+ v_h\right] \\ &\geq \int_\Omega M_h \pi_h \left[v_h\right]_+ - \int_\Omega M_h \pi_h \left[u_h\right]_+ .\end{aligned}$$

The right-hand side is non-negative by (3.12), and hence

$$\int_\Omega M_h \pi_h \left[\left(\left(1 + \varepsilon K_h^{-1} L_h\right)^{-1} K_h^{-1} L_h v_h\right) \cdot \operatorname{sgn}^+ v_h\right] \geq 0.$$

Making $\varepsilon \downarrow 0$, we have

$$\int_\Omega M_h \pi_h \left[\left(K_h^{-1} L_h v_h\right) \cdot \operatorname{sgn}^+ v_h\right] \geq 0.$$

Hence noting

$$\pi_h \left(\eta \cdot \operatorname{sgn}^+ v\right) = \sum_{a \in I_h \cap \{\pi_h v \geq 0\}} \eta(a) w_a = \pi_h \left(\eta \cdot \operatorname{sgn}^+ \pi_h v\right)$$

for any $\eta \in W$, we obtain (3.11). The proof is complete. □

Now we give the following.

Proof of Theorem 3.1. To prove (3.1), we show more generally that

$$\left\|M_h \pi_h \left[v - \hat{v}\right]_+\right\|_1 \leq \left\|M_h \pi_h \left[v - \hat{v} + \lambda A_h v - \lambda A_h \hat{v}\right]_+\right\|_1, \tag{3.14}$$

where $v, \hat{v} \in W$ and $\lambda > 0$. To this end, we suppose that f is strictly increasing. Otherwise, we replace f by $f_\varepsilon(u) = f(u) + \varepsilon u$ and make $\varepsilon \downarrow 0$. Putting $g = v + \lambda K_h^{-1} L_h \pi_h f(v)$ and $\hat{g} = \hat{v} + \lambda K_h^{-1} L_h \pi_h f(\hat{v})$, we get that

$$\begin{aligned}\left\|M_h \pi_h \left[v - \hat{v}\right]_+\right\|_1 &= \int_\Omega M_h \pi_h \left[(v - \hat{v}) \cdot \operatorname{sgn}^+ (v - \hat{v})\right] \\ &= \int_\Omega M_h \pi_h \left[(g - \hat{g}) \cdot \operatorname{sgn}^+ (v - \hat{v})\right] \\ &\quad - \lambda \int_\Omega M_h \pi_h \left[\left(K_h^{-1} L_h \pi_h \left(f(v) - f(\hat{v})\right)\right) \cdot \operatorname{sgn}^+ (v - \hat{v})\right].\end{aligned}$$

Here, we have $\operatorname{sgn}^+ w = \operatorname{sgn}^+ (v - \hat{v})$ holds for $w = f(v) - f(\hat{v}) \in W$, because f is strictly increasing. Therefore, (3.11) guarantees that

$$\int_\Omega M_h \pi_h \left[\left(K_h^{-1} L_h \pi_h \left(f(v) - f(\hat{v})\right)\right) \cdot \operatorname{sgn}^+ (v - \hat{v})\right] = \int_\Omega K_h^{-1} \pi_h \left[K_h^{-1} L_h \pi_h w \cdot \operatorname{sgn}^+ w\right] \geq 0.$$

This leads to

$$\begin{aligned}\|M_h \pi_h [v - \hat{v}]_+\|_1 &\leq \int_\Omega M_h \pi_h \left[(g - \hat{g}) \cdot \operatorname{sgn}^+ (v - \hat{v})\right] \\ &\leq \int_\Omega M_h \pi_h \left[g - \hat{g}\right]_+ = \left\|M_h \pi_h \left[v - \hat{v} + \lambda A_h v - \lambda A_h \hat{v}\right]_+\right\|_1,\end{aligned}$$

and hence (3.1) follows.

Now we prove the maximality $(I+\lambda A_h)X_h = X_h$ for $\lambda > 0$. Namely, given $g_h \in X_h$, we show the existence of $v_h \in X_h$ satisfying $v_h + \lambda A_h v_h = g_h$. In fact, $T_\lambda = I + \lambda A_h$ is a continuous mapping on X_h, a finite dimensional vector space provided with the norm $\|\cdot\|_{1,h}$. In use of (3.1) we can take an open ball $\mathcal{O} \subset X_h$ sufficiently large such that $g_h \notin T_\lambda(\partial\mathcal{O})$ for any $\lambda > 0$. We may suppose that $g_h \in \mathcal{O}$. Then the topological degree deg $(T_\lambda, g_h, \mathcal{O})$ is well-defined and its homotopy invariance implies

$$\deg\,(T_\lambda, g_h, \mathcal{O}) = \deg\,(I, g_h, \mathcal{O}) = 1.$$

This means that $g_h \in T_\lambda(\mathcal{O})$, and the proof is complete. □

4. L^∞ STABILITY

This section is devoted to the L^∞ stability of approximate solutions. Precisely, we show the following.

Theorem 4.1. *Under the assumption* (H1), *it holds that*

$$\|S_h(t)u_{0h}\|_\infty \le \|u_{0h}\|_\infty\,, \tag{4.1}$$

where $u_{0h} \in X_h$ *and* $t \in [0,T]$.

Note that, as will be verified at the end of this section, (4.1) gives

$$\left\|(I+\lambda A_h)^{-1} g_h\right\|_\infty \le \|g_h\|_\infty\,, \tag{4.2}$$

for $g_h \in X_h$, $\lambda > 0$.

To prove Theorem 4.1, we make use of the nonlinear Chernoff formula, taking a finite element analogue of the time-discretization scheme of [3]. For the moment, we suppose that f is locally Lipschitz continuous. Let $\mu > 0$ be the Lipschitz constant of f on $[-M, M]$, where $M = \|u_{0h}\|_\infty$ for $u_{0h} \in X_h$. We take $\tau = T/N$ for $N \in \mathbb{N}$ and put $t_m = m\tau$ for $0 \le m \le N$. Then, we introduce the regularizing parameter $s_\tau > 0$ satisfying

$$\lim_{\tau\downarrow 0} s_\tau = 0 \qquad \text{and} \qquad \frac{\mu\tau}{s_\tau} \le 1, \tag{4.3}$$

and take $\{w_h^\tau(t_m)\}_{m=0}^N \subset X_h$ by

$$\frac{w_h^\tau(t_{m+1}) - w_h^\tau(t_m)}{\tau} + \left(\frac{1 - \mathrm{e}^{-s_\tau K_h^{-1} L_h}}{s_\tau}\right)\pi_h f(w_h^\tau(t_m)) = 0$$

with $w_h^\tau(0) = u_{0h}$, where $\{\mathrm{e}^{-sK_h^{-1}L_h}\}_{s\ge 0}$ denotes the linear semigroup in X_h generated by $K_h^{-1}L_h$. We extend $w_h^\tau(t_m)$ to all $t \in [0,T]$ as

$$w_h^\tau(t) = \begin{cases} w_h^\tau(0) & (t=0) \\ w_h^\tau(t_m) & (t_{m-1} < t \le t_m,\ 1 \le m \le N) \end{cases}. \tag{4.4}$$

The following lemma is proven similarly to [3].

Lemma 4.1. *In addition to the basic assumption on f, suppose that f is locally Lipschitz continuous on $\mathbb{R}$. Then $w_h^\tau(t) \in X_h$ is well-defined for all $t \in [0,T]$, and moreover*

$$\lim_{\tau\downarrow 0} \sup_{t\in[0,T]} \|w_h^\tau(t) - S_h(t)u_{0h}\|_{1,h} = 0 \tag{4.5}$$

for $u_{0h} \in X_h$.

Proof. We have the formula

$$w_h^\tau(t_m) = F_h(\tau)^m u_{0h},$$

where

$$F_h(\tau)\phi_h = \phi_h + \frac{\tau}{s_\tau}\left[\mathrm{e}^{-s_\tau K_h^{-1}L_h}\pi_h f(\phi_h) - \pi_h f(\phi_h)\right]. \tag{4.6}$$

Since, by $\mu\tau/s_\tau \le 1$, the mapping $r \mapsto r - (\tau/s_\tau)f(r)$ is non-increasing, we have

$$-M - \frac{\tau}{s_\tau}f(-M) \le u_{0h} - \frac{\tau}{s_\tau}\pi_h f(u_{0h}) \le M - \frac{\tau}{s_\tau}f(M). \tag{4.7}$$

On the other hand, (3.8) implies $0 \le \left(I + \lambda K_h^{-1}L_h\right)^{-1} v_h^\pm \le \max_{\overline{\Omega}} v_h^\pm$ for $v_h \in X_h$ and $\lambda > 0$ with $v_h^\pm \in X_h$ defined by (3.13). In particular,

$$\max_{\overline{\Omega}} \left(1 + \lambda K_h^{-1}L_h\right)^{-1} \pi_h[v]_\pm \le \max_{\overline{\Omega}} \pi_h[v]_\pm$$

holds for any $v \in W$ and $\lambda > 0$. Then, the linear semigroup theory guarantees that

$$\max_{\overline{\Omega}} \mathrm{e}^{-sK_h^{-1}L_h}\pi_h[v]_\pm \le \max_{\overline{\Omega}} \pi_h[v]_\pm$$

for any $s > 0$ and $v \in W$. Therefore, noting that $f(-M) \le \pi_h f(u_{0h}) \le f(M)$, we can deduce

$$f(-M) \le \mathrm{e}^{-s_\tau K_h^{-1}L_h}\pi_h f(u_{0h}) \le f(M). \tag{4.8}$$

Inequalities (4.7) and (4.8) imply

$$-M \le w_h^\tau(t_1) = u_{0h} + \frac{\tau}{s_\tau}\left[\mathrm{e}^{-s_\tau K_h^{-1}L_h}\pi_h f(u_{0h}) - \pi_h f(u_{0h})\right] \le M,$$

which means $\|F_h(\tau)u_{0h}\|_\infty \le M$. Therefore, we get by an induction that

$$\|w_h^\tau(t_m)\|_\infty \le \|u_{0h}\|_\infty. \tag{4.9}$$

This allows us to assume that f is Lipschitz continuous with Lipschitz constant μ in $\mathbb{R}$ by replacing $f(u)$ by $f(\pm M)$ for $\pm u \ge M$. Then, $r \mapsto f(r)$ and $r \mapsto r - (\tau/s_\tau)f(r)$ are non-decreasing on $\mathbb{R}$, and it follows that

$$\frac{\tau}{s_\tau}|f(r) - f(s)| + \left|(r - s) - \frac{\tau}{s_\tau}(f(r) - f(s))\right| = |r - s| \tag{4.10}$$

for $r, s \in \mathbb{R}$. On the other hand, from (3.1) and (3.3) applied to $f(u) = u$, we have

$$\left\|\mathrm{e}^{-sK_h^{-1}L_h}\pi_h[v]_+\right\|_{1,h} \le \|\pi_h[v]_+\|_{1,h}$$

for $v \in W$. This, together with (4.10), gives that

$$\begin{aligned}\|F_h(\tau)\phi_h - F_h(\tau)\psi_h\|_{1,h} &\le \frac{\tau}{s_\tau}\|f(\phi_h) - f(\psi_h)\|_{1,h} + \left\|(\phi_h - \psi_h) - \frac{\tau}{s_\tau}(f(\phi_h) - f(\psi_h))\right\|_{1,h} \\ &= \|\phi_h - \psi_h\|_{1,h}\end{aligned} \tag{4.11}$$

for $\phi_h, \psi_h \in X_h$.

Now we shall show (4.5). It is a consequence of the Chernoff formula, Theorem 3.1 of [5]. Namely, it suffices to prove that

$$\lim_{\tau\downarrow 0}\left[I+\frac{\lambda}{\tau}\left(I-F_h(\tau)\right)\right]^{-1}\phi_h=\left(I+\lambda A_h\right)^{-1}\phi_h \tag{4.12}$$

for $\phi_h\in X_h$ and $\lambda>0$. For this purpose, we put

$$\psi_h=\left(I+\lambda A_h\right)^{-1}\phi_h,\quad \psi_h^\tau=\left[I+\frac{\lambda}{\tau}\left(I-F_h(\tau)\right)\right]^{-1}\phi_h,\quad \phi_h^\tau=\psi_h+\frac{\lambda}{\tau}\left(I-F_h(\tau)\right)\psi_h.$$

Then, we have

$$\phi_h=\psi_h^\tau+\frac{\lambda}{\tau}\left(I-F_h(\tau)\right)\psi_h^\tau$$

and

$$\phi_h-\phi_h^\tau=\left(1+\frac{\lambda}{\tau}\right)\left(\psi_h^\tau-\psi_h\right)+\frac{\lambda}{\tau}\left(F_h(\tau)\psi_h-F_h(\tau)\psi_h^\tau\right).$$

Therefore, inequality (4.11) gives that

$$\left(1+\frac{\lambda}{\tau}\right)\left\|\psi_h^\tau-\psi_h\right\|_{1,h}\le\left\|\phi_h-\phi_h^\tau\right\|_{1,h}+\frac{\lambda}{\tau}\left\|\psi_h^\tau-\psi_h\right\|_{1,h},$$

and hence

$$\left\|\psi_h^\tau-\psi_h\right\|_{1,h}\le\left\|\phi_h-\phi_h^\tau\right\|_{1,h}. \tag{4.13}$$

Inequality (4.13) provides an *a priori* estimate and hence the existence of ψ_h^τ follows similarly to the proof of Theorem 3.1.

Finally, by (4.6), we have

$$\phi_h^\tau=\psi_h-\frac{\lambda}{s_\tau}\left[\mathrm{e}^{-s_\tau K_h^{-1}L_h}\pi_h f(\psi_h)-\pi_h f(\psi_h)\right]$$

and hence

$$\lim_{\tau\downarrow 0}\phi_h^\tau=\psi_h+\lambda K_h^{-1}L_h\pi_h f\left(\psi_h\right)=\psi_h+\lambda A_h\psi_h=\phi_h.$$

Thus, we get (4.12) by (4.13) and the proof is complete. □

Now, we give the following.

Proof of Theorem 4.1. If f is locally Lipschitz continuous, then we have (4.5) and (4.9), which implies (4.1) by dim $X_h<+\infty$.

If this is not the case, we take the Yosida approximation, a family $\{f_\lambda\}$ converging to f locally uniformly as $\lambda\downarrow 0$. Namely, in use of the maximal monotone graph $\beta=f^{-1}$, we define the inverse function of f_λ as

$$f_\lambda^{-1}\equiv\beta_\lambda=\frac{1}{\lambda}\left[1-\left(1+\lambda\beta\right)^{-1}\right] \tag{4.14}$$

which is non-decreasing, $f_\lambda(0)=0$, and locally Lipschitz continuous. Let $A_h^\lambda v=L_h\pi_h f_\lambda(v)$. Then it generates the semigroup $\{S_h^\lambda(t)\}_{t\ge 0}$ in X_h satisfying

$$\left\|S_h^\lambda(t)u_{0h}\right\|_\infty\le\left\|u_{0h}\right\|_\infty$$

for $u_{0h}\in X_h$ and $t\in[0,T]$. Making $\lambda\downarrow 0$, we obtain (4.1) by dim $X_h<+\infty$. □

We proceed to the proof of (4.2). For this end, we take the duality map $F : X_h \to X_h^*$, regarding X_h as a closed subspace of $L^\infty(\Omega)$. Namely, for $v_h, \chi_h \in X_h$ it holds that

$$\chi_h \in F(v_h) \iff \langle v_h, \chi_h \rangle = \|v_h\|_\infty^2 = \|\chi_h\|_*^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing between X_h and X_h^*, and $\|\cdot\|_*$ the operator norm. See Miyadera [29], *e.g.*, for the existence of such an operator. Then, by making use of (4.1), it holds that

$$\begin{aligned}\left\langle \left(\frac{S_h(\tau) - 1}{\tau} \right) v_h, \chi_h \right\rangle &= \frac{1}{\tau} \left\{ \langle S_h(\tau) v_h, \chi_h \rangle - \langle v_h, \chi_h \rangle \right\} \\ &= \frac{1}{\tau} \left\{ \langle S_h(\tau) v_h, \chi_h \rangle - \|v_h\|_\infty^2 \right\} \\ &\leq \frac{1}{\tau} \left\{ \|S_h(\tau) v_h\|_\infty - \|v_h\|_\infty \right\} \|\chi_h\|_* \leq 0\end{aligned}$$

for $v_h \in X_h$, $\tau > 0$, and $\chi_h \in F(v_h)$. Hence, by making $\tau \downarrow 0$, we obtain $\langle A_h v_h, \chi_h \rangle \leq 0$ for $v_h \in X_h$ and $\chi_h \in F(v_h)$. The general theory of the duality map, say Corollary 2.7 of [29], guarantees that

$$\|g_h\|_\infty \leq \|(1 + \lambda A_h) g_h\|_\infty$$

for any $g_h \in X_h$ and $\lambda > 0$. Thus we establish (4.2).

5. Convergence of resolvent

Convergence of semigroup follows from that of resolvent. We assume the following condition concerning the domain $\Omega \subset \mathbb{R}^n$:

(D) If $n = 3$ the Dirichlet problem

$$-\Delta w = g \quad \text{in} \quad \Omega, \qquad w = 0 \quad \text{on} \quad \partial\Omega$$

admits the elliptic estimate

$$\|w\|_{W^{2,p}(\Omega)} \leq C_p \|g\|_p$$

for $p \in (1, \mu)$, where $\mu > n = 3$.

As for the triangulation, we suppose

(H3) *Inverse inequality.* There is a positive constant ν_2 independent of h such that

$$d(\sigma) \geq \nu_2 h$$

for any $\sigma \in \mathcal{T}_h$.

This section is devoted to the

Theorem 5.1. *If Ω is convex and provided with the property (D) (if $n = 3$), $\{\mathcal{T}_h\}$ satisfies* (H1), (H2) *and* (H3), *and f is strictly increasing, then it holds that*

$$\lim_{h \downarrow 0} \left\| (I + \lambda A)^{-1} g - (I + \lambda A_h)^{-1} \pi_h g \right\|_\infty = 0, \tag{5.1}$$

where $g \in W$ and $\lambda > 0$.

Several remarks are in order.

Remark 5.1. The family of triangulation $\{\mathcal{T}_h\}$ satisfying (H2) and (H3) is often called *quasi-uniform*.

Remark 5.2. Convexity of $\Omega \subset \mathbb{R}^n$ assures

$$L^{-1}(L^2(\Omega)) \subset H_0^1(\Omega) \cap H^2(\Omega) \subset W. \tag{5.2}$$

In fact, the second inclusion is a consequence of Sobolev's embedding theorem by $n = 1, 2, 3$. On the other hand, the first inclusion follows from the elliptic regularity of the Green operator of L. See [19].

Remark 5.3. Rannacher and Scott [33] showed that if $n = 2$, Ω is convex, and $\{\mathcal{T}_h\}$ satisfies (H2) and (H3), then the following estimate holds for the Ritz operator R_h defined by (2.4). That is, there is $h_0 > 0$ such that

$$\|R_h w\|_{W^{1,p}(\Omega)} \leq C \|w\|_{W^{1,p}(\Omega)} \tag{5.3}$$

for any $w \in H_0^1(\Omega) \cap W^{1,p}(\Omega)$, $0 < h \leq h_0$, and $p \in [2, \infty]$. (See also [17] for the proof.) By virtue of Theorem 7.5.3 of Brenner and Scott [4], on the other hand, the same conclusion follows if $n = 3$, Ω is provided with (D), and $\{\mathcal{T}_h\}$ satisfies (H2) and (H3).

For later use, it is sufficient for (5.3) to hold with some $p > n$. It is obvious for $n = 1$, because we can take $p = 2$ then. Namely, assumptions on Ω are reduced to (5.2) and (5.3) with some $p > n$.

Remark 5.4. Condition (D) is fulfilled, when all edges and all vertices of a polyhedron $\Omega \subset \mathbb{R}^3$ are small enough not to produce singularities. See, for a more complete description, Theorems 8.2.1.2 and 8.2.2.8 of Grisvard [19].

Remark 5.5. Given $g \in W$ and $u = (I + \lambda A)^{-1} g$, we have

$$f(u) = \lambda^{-1} L^{-1} (g - u) \in W$$

by (1.10) and (5.2). Therefore, if f is strictly increasing, then $u \in W$ follows.

First, we show the following.

Lemma 5.1. *Let $\lambda > 0$, $g \in W$, and $u_h = (I + \lambda A_h)^{-1} \pi_h g$. Then, under the assumptions of the previous theorem, the family $\{u_h\}$ is relatively compact as $h \downarrow 0$ in W.*

Proof. Recall $u_h = (I + \lambda A_h)^{-1} \pi_h g$ with $\lambda > 0$ and $g \in W$. We shall show that any $\varepsilon > 0$ admits $\delta > 0$ and $h_1 > 0$ such that

$$0 < h \leq h_1,\ x, y \in \overline{\Omega},\ |x - y| < \delta \ \Rightarrow\ |u_h(x) - u_h(y)| < \varepsilon. \tag{5.4}$$

Then, Ascoli-Arzela's theorem assures that any $\left\{u_{h_j}\right\}$ with $h_j \downarrow 0$ admits a subsequence, uniformly converging on $\overline{\Omega}$. Thus, the lemma is proven.

In fact, we have

$$L_h \pi_h f(u_h) = \frac{1}{\lambda} K_h(\pi_h g - u_h). \tag{5.5}$$

Putting $\phi_h = \lambda^{-1} K_h(\pi_h g - u_h) \in X_h \subset W$, we take w satisfying

$$-\Delta w = \phi_h \quad \text{in} \quad \Omega \qquad \text{with} \qquad w = 0 \quad \text{on} \quad \partial\Omega.$$

Because $\phi_h = L_h\pi_h f(u_h)$ holds, we obtain $R_h w = \pi_h f(u_h)$. By virtue of (5.3), (5.2), (3.7), and (4.2), it follows for $p > n$ that

$$\begin{aligned}
\|\pi_h f(u_h)\|_{W^{1,p}} &\le C_p \|w\|_{W^{1,p}(\Omega)} \le C \cdot C_p \|w\|_{H^2(\Omega)} \\
&\le C \cdot C_p \cdot C \|\phi_h\|_2 = \lambda^{-1} C_p' \|K_h(\pi_h g - u_h)\|_2 \\
&\le \lambda^{-1} C_p'' \|\pi_h g - u_h\|_2 \\
&\le \lambda^{-1} C_p'' |\Omega|^{1/2} \|\pi_h g - u_h\|_\infty \\
&\le \lambda^{-1} C_p'' |\Omega|^{1/2} (\|\pi_h g\|_\infty + \|u_h\|_\infty) \\
&\le 2\lambda^{-1} C_p'' |\Omega|^{1/2} \|g\|_\infty
\end{aligned}$$

for $0 < h \le h_0$. Here and henceforth, $|\Omega|$ denotes the n dimensional volume of Ω. Therefore, by Morrey's inequality, there is a constant $\tilde{C} = \tilde{C}(\lambda, g, \Omega, h_0)$ such that

$$|\pi_h f(u_h(x)) - \pi_h f(u_h(y))| \le \tilde{C} |x - y|^\alpha$$

for $0 < h \le h_0$ and $x, y \in \overline{\Omega}$, where $\alpha = 1 - n/p > 0$.

Let B_h be the set of nodal points of $\mathcal{T}_h$ belonging to $\partial\Omega$, and put $\overline{I}_h = I_h \cup B_h$. Since $\pi_h f(u_h(x)) = f(u_h(x))$ for $x \in \overline{I}_h$, we have

$$|f(u_h(x_1)) - f(u_h(x_2))| \le \tilde{C} |x_1 - x_2|^\alpha \tag{5.6}$$

for $x_1, x_2 \in \overline{I}_h$.

Let $\sigma \in \mathcal{T}_h$ and $V(\sigma)$ be the set of vertices of σ. Because $u_h \in X_h$, we have $\overline{x}, \underline{x} \in V(\sigma)$ such that

$$u_h(\overline{x}) = \max_\sigma u_h \qquad \text{and} \qquad u_h(\underline{x}) = \min_\sigma u_h.$$

This implies

$$\max_\sigma f(u_h) = f(u_h(\overline{x})) \qquad \text{and} \qquad \min_\sigma f(u_h) = f(u_h(\underline{x}))$$

because f is non-decreasing. Therefore, if $x, y \in \sigma$ we have by (5.6) that

$$|f(u_h(x)) - f(u_h(y))| \le f(u_h(\overline{x})) - f(u_h(\underline{x})) \le \tilde{C} |\overline{x} - \underline{x}|^\alpha \le \tilde{C} h^\alpha. \tag{5.7}$$

We shall combine (5.6) and (5.7) in the following way. Namely, given $x, y \in \overline{\Omega}$ in $|x - y| \le h$, we take $\sigma_1, \sigma_2 \in \mathcal{T}_h$ satisfying $x \in \sigma_1$ and $y \in \sigma_2$. We also take $x_1 \in V(\sigma_1)$ and $x_2 \in V(\sigma_2)$. Then, we get from those inequalities that

$$\begin{aligned}
|f(u_h(x)) - f(u_h(y))| &\le |f(u_h(x)) - f(u_h(x_1))| \\
&\quad + |f(u_h(x_1)) - f(u_h(x_2))| + |f(u_h(x_2)) - f(u_h(y))| \\
&\le \tilde{C} h^\alpha + \tilde{C} |x_1 - x_2|^\alpha + \tilde{C} h^\alpha \\
&\le 2\tilde{C} h^\alpha + \tilde{C} (|x_1 - x| + |x - y| + |y - x_2|)^\alpha \le 5\tilde{C} h^\alpha.
\end{aligned}$$

Finally, f is continuous and strictly increasing, the inverse function f^{-1} is uniformly continuous on $[-\|g\|_\infty, \|g\|_\infty]$. Because of $\|u_h\|_\infty \le \|g\|_\infty$, each $\varepsilon > 0$ admits $\delta_1 > 0$ such that

$$|f(u_h(x)) - f(u_h(y))| < \delta_1 \quad \Rightarrow \quad |u_h(x) - u_h(y)| < \varepsilon \tag{5.8}$$

for $x, y \in \overline{\Omega}$. Those relations (5.8) and (5.8) imply (5.4) and the proof is complete.

We also make use of the following lemma, where P_h denotes the L^2 orthogonal projection defined by (2.5). For the proof, see that of Theorems 1.12 and 5.4 in [17]. We note that those results hold even if Ω is not convex, or (D) does not hold.

Lemma 5.2. *Suppose* (H2) *and* (H3), *and take*

$$q \in \left[1, \min\left(2, \frac{n}{n-1}\right)\right). \tag{5.9}$$

Then it holds that

$$\|\chi_h\|_{W^{1,q}(\Omega)} \le C \|L_h \chi_h\|_{L^1(\Omega)}, \tag{5.10}$$

$$\left\|L_h^{-1} K_h P_h\right\|_{L^1(\Omega), W^{1,q}(\Omega)} \le C, \tag{5.11}$$

and

$$\left\|L_h^{-1} K_h P_h v - L^{-1} v\right\|_{W^{1,q}(\Omega)} = 0, \tag{5.12}$$

where $\chi_h \in X_h$ *and* $v \in L^1(\Omega)$.

Now we can give the

Proof of Theorem 5.1. Given $\lambda > 0$ and $g \in W$, we put $g_h = \pi_h g$ and

$$u_h = (1 + \lambda A_h)^{-1} g_h. \tag{5.13}$$

In use of (5.5), (3.7), and (4.2), we get

$$\begin{aligned} \|L_h \pi_h f(u_h)\|_1 &= \lambda^{-1} \|K_h(\pi_h g - u_h)\|_1 \le \lambda^{-1} C \|\pi_h g - u_h\|_1 \\ &\le \lambda^{-1} C |\Omega| \left(\|\pi_h g\|_\infty + \|u_h\|_\infty\right) \\ &\le 2\lambda^{-1} C |\Omega| \|g\|_\infty . \end{aligned}$$

Therefore, we have $\|\pi_h f(u_h)\|_{W^{1,q}} \le C$ by (5.10), where q is taken from (5.9).

From this inequality and Lemma 5.1, any $h_j \downarrow 0$ admits $\{h'_j\} \subset \{h_j\}$, $w \in W^{1,q}(\Omega)$, and $u \in C_0(\overline{\Omega})$ satisfying

$$\begin{array}{ll} \pi_h f(u_h) \to w & \text{weakly in } W^{1,q}(\Omega) \\ \pi_h f(u_h) \to w & \textit{a.e.} \text{ in } \Omega \\ u_h \to u & \text{uniformly on } \overline{\Omega} \end{array}$$

as $h = h'_j \downarrow 0$.

Here, we show that

$$w = f(u) \quad \textit{a.e.} \tag{5.14}$$

holds by Egorov's theorem. In fact, given $\varepsilon > 0$, we have a measurable set $\Omega_\varepsilon \subset \Omega$ satisfying $|\Omega \setminus \Omega_\varepsilon| < \varepsilon$ and $u_h \to u$ uniformly on Ω_ε. This implies $\pi_h f(u_h) \to f(u)$ uniformly on Ω_ε, because $f(u)$ is continuous on $\overline{\Omega}$ and

$$\|\pi_h f(u_h) - f(u)\|_{L^\infty(\Omega_\varepsilon)} \le 3 \|f(u_h) - f(u)\|_{L^\infty(\Omega_\varepsilon)} + \|(\pi_h - 1) f(u)\|_{L^\infty(\Omega_\varepsilon)}$$

follows from

$$\pi_h f(u_h) - f(u) = (\pi_h - I)\left(f(u_h) - f(u)\right) + \left(f(u_h) - f(u)\right) + (\pi_h - I) f(u).$$

Hence we deduce $w = f(u)$ *a.e.* on Ω_ε, and therefore (5.14) follows. Thus, we have

$$\pi_h f(u_h) \;\to\; f(u) \qquad \text{weakly in } W^{1,q}(\Omega) \tag{5.15}$$

as $h = h'_j \downarrow 0$.

On the other hand, we have by (5.11) and (5.12) that

$$\begin{aligned}\left\|L_h^{-1}K_hu_h - L^{-1}u\right\|_{W^{1,q}} &\le \left\|L_h^{-1}K_hP_h\left(u_h-u\right)\right\|_{W^{1,q}} + \left\|L_h^{-1}K_hP_hu - L^{-1}u\right\|_{W^{1,q}}\\ &\le C\left\|u_h-u\right\|_1 + \left\|L_h^{-1}K_hP_hu - L^{-1}u\right\|_{W^{1,q}}\\ &\le C\left|\Omega\right|\cdot\left\|u_h-u\right\|_\infty + \left\|L_h^{-1}K_hP_hu - L^{-1}u\right\|_{W^{1,q}} \to 0 \end{aligned} \tag{5.16}$$

and similarly,

$$\left\|L_h^{-1}K_hg_h - L^{-1}g\right\|_{W^{1,q}} \to 0 \tag{5.17}$$

as $h = h_j' \downarrow 0$. Writing (5.13) as $L_h^{-1}K_hu_h + \lambda\pi_hf(u_h) = L_h^{-1}K_hg_h$, we obtain $L^{-1}u + \lambda f(u) = L^{-1}g$ by (5.15), (5.16), and (5.17). This means $u = (I+\lambda A)^{-1}g$ and the proof is complete. □

6. Convergence of Yosida approximation

Throughout this and the following sections, supposing

$$u_0 \in W,$$

we take

$$u_{0h} = \pi_h u_0.$$

Relation (1.12) is referred to as a convergence of the semigroup. To show this result, we make use of the Yosida approximation. Since $-A$ is an m-dissipative operator in $X = L^1(\Omega)$, we can apply the abstract theory. See Miyadera [29] for the proof of the following facts.

First, the Yosida approximation of A is defined by $A_\lambda = \lambda^{-1}(I - J_\lambda)$, where $\lambda > 0$ and $J_\lambda = (I+\lambda A)^{-1}$. It is $(2/\lambda)$-Lipschitz continuous in X, because J_λ is a contraction. Furthermore $-A_\lambda$ is m-dissipative in X. Hence, it generates a contraction semigroup and, for $u_0 \in X$, we have a unique solution $u_\lambda \in C^1([0,T];X)$ to

$$\frac{\mathrm{d}u_\lambda}{\mathrm{d}t} + A_\lambda u_\lambda = 0 \qquad \text{with} \qquad u_\lambda(0) = u_0. \tag{6.1}$$

The Yosida approximation $A_{h,\lambda}$ of the approximate operator A_h is also defined similarly in X_h. We have $A_{h,\lambda} = \lambda^{-1}(I - J_{h,\lambda})$ with $J_{h,\lambda} = (I+\lambda A_h)^{-1}$. We note that $A_{h,\lambda}$ has the same properties on X_h equipped with the norm $\|\cdot\|_{1,h}$ as those for A_λ on X with $\|\cdot\|_1$. It is $(2/\lambda)$-Lipschitz continuous and $-A_{h,\lambda}$ is m-dissipative. We have a unique $u_{h,\lambda} \in C^1([0,T];X_h)$ satisfying

$$\frac{\mathrm{d}u_{h,\lambda}}{\mathrm{d}t} + A_{h,\lambda}u_{h,\lambda} = 0 \qquad \text{with} \qquad u_{h,\lambda}(0) = \pi_h u_0. \tag{6.2}$$

This section is devoted to the following.

Lemma 6.1. *Suppose that Ω is convex and is provided with the property (D) if $n = 3$, that $\{\mathcal{T}_h\}$ satisfies* (H1), (H2) *and* (H3), *and that f is strictly increasing. Given $\lambda > 0$ and $u_0 \in W$, let u_λ and $u_{h,\lambda}$ be the solutions to* (6.1) *and* (6.2), *respectively. Then, it holds that*

$$\lim_{h\downarrow 0}\sup_{t\in[0,T]}\left\|u_{h,\lambda}(t) - u_\lambda(t)\right\|_1 = 0. \tag{6.3}$$

There is a technical difficulty to prove the above lemma. That is, it is not obvious that $u_\lambda(t) \in W$ follows from $u_0 \in W$ in spite that $u_\lambda(t) \in L^\infty(\Omega)$ actually follows from $u_0 \in L^\infty(\Omega)$. This causes a problem because the interpolation operator π_h works only to continuous functions. To avoid such an issue, we take time discretizations and derive an analogous result first.

Taking $\tau = T/N$ with $N \in \mathbb{N}$, we introduce the backward difference approximation to (6.1):

$$\begin{cases} \dfrac{u_\lambda^\tau(t_{m+1}) - u_\lambda^\tau(t_m)}{\tau} + A_\lambda u_\lambda^\tau(t_{m+1}) = 0 \quad (0 \le m \le N) \\ u_\lambda^\tau(0) = u_0, \end{cases} \tag{6.4}$$

where $t_m = m\tau$. It is defined only at a discrete time level t_m, and the extension to the continuous time interval $[0,T]$ is given by

$$u_\lambda^\tau(t) = \begin{cases} u_\lambda^\tau(0) & (t = 0) \\ u_\lambda^\tau(t_m) & (t_{m-1} < t \le t_m,\ 1 \le m \le N). \end{cases}$$

The backward difference approximation is also taken to (6.2):

$$\begin{cases} \dfrac{u_{h,\lambda}^\tau(t_{m+1}) - u_{h,\lambda}^\tau(t_m)}{\tau} + A_{h,\lambda} u_{h,\lambda}^\tau(t_{m+1}) = 0 \quad (0 \le m \le N) \\ u_{h,\lambda}^\tau(0) = \pi_h u_0 \end{cases} \tag{6.5}$$

and the extension $u_{h,\lambda}^\tau(t)$ to the continuous time interval is defined similarly.

Remark 6.1. The relation $u = (I + \tau A_\lambda)^{-1} g$ is equivalent to

$$\left(\frac{\lambda}{\tau} + 1\right) u = \frac{\lambda}{\tau} g + J_\lambda u \tag{6.6}$$

and hence

$$\left(1 + \frac{\lambda}{\tau}\right) \|u\|_\infty - \frac{\lambda}{\tau} \|g\|_\infty \le \left\| u + \frac{\lambda}{\tau}(u - g) \right\|_\infty = \|J_\lambda u\|_\infty \le \|u\|_\infty$$

follows from (4.2). This implies $\|u\|_\infty \le \|g\|_\infty$, that is, L^∞ stability of A_λ described as

$$\left\| (I + \tau A_\lambda)^{-1} g \right\|_\infty \le \|g\|_\infty .$$

On the other hand, relation (6.6) reads

$$u = u + \frac{\lambda}{\tau}(u - g) + \lambda L f\left(u + \frac{\lambda}{\tau}(u - g)\right)$$

and hence $Lf\left(u + \frac{\lambda}{\tau}(u-g)\right) \in L^\infty(\Omega)$ follows from $u, g \in L^\infty(\Omega)$. This relation implies $u \in W$ if $g \in W$ and f is strictly increasing, in the similar way as Remark 5.5. In particular, $u_\lambda^\tau(t_m) = (I + \tau A_\lambda)^{-m} u_0 \in W$ follows from $u_0 \in W$. On the other hand, it is obvious that

$$u_{h,\lambda}^\tau(t_m) = (I + \tau A_{h,\lambda})^{-m} \pi_h u_0 \in X_h.$$

Under those preparations, first we show the following

Lemma 6.2. *Let $u_\lambda^\tau(t)$ and $u_{h,\lambda}^\tau(t)$ be the solutions to* (6.4) *and* (6.5)*, respectively, where $\tau > 0$, $\lambda > 0$, and $u_0 \in W$. Then, under the same assumptions of Lemma 6.1, it holds that*

$$\lim_{h \downarrow 0} \sup_{t \in [0,T]} \left\| u_{h,\lambda}^\tau(t) - u_\lambda^\tau(t) \right\|_1 = 0. \tag{6.7}$$

Proof. By the associative law of operators, we calculate as

$$\begin{aligned} u^{\tau}_{h,\lambda}(t_m) - u^{\tau}_{\lambda}(t_m) &= \left(I + \tau A_{h,\lambda}\right)^{-m} \pi_h u_0 - \left(I + \tau A_{\lambda}\right)^{-m} u_0 \\ &= (I + \tau A_{\lambda})^{-m} \pi_h u_0 - \left(I + \tau A_{\lambda}\right)^{-m} u_0 \\ &\quad + \sum_{l=1}^{m} \Big[(I + \tau A_{\lambda})^{-(m-l)} (I + \tau A_{h,\lambda})^{-l} \\ &\quad - (I + \tau A_{\lambda})^{-(m-l+1)} (I + \tau A_{h,\lambda})^{-(l-1)} \Big] \pi_h u_0. \end{aligned} \tag{6.8}$$

Because $(I + \tau A_{\lambda})^{-1}$ is a contraction in X, it holds that

$$\left\| (I + \tau A_{\lambda})^{-1} v - (I + \tau A_{\lambda})^{-1} \hat{v} \right\|_1 \leq \| v - \hat{v} \|_1 \tag{6.9}$$

for $v, \hat{v} \in X$. This implies

$$\| (I + \tau A_{\lambda})^{m} \pi_h u_0 - (I + \tau A_{\lambda})^{m} u_0 \|_1 \leq \| (\pi_h - 1) u_0 \|_1 .$$

On the other hand, the L^1 norm of the third term of the right-hand side of (6.8) is estimated from above by

$$\begin{aligned} &\sum_{l=1}^{m} \left\| \left[(I + \tau A_{\lambda}) \left(I + \tau A_{h,\lambda}\right)^{-l} - \left(I + \tau A_{h,\lambda}\right)^{-(l-1)} \right] \pi_h u_0 \right\|_1 \\ &\quad = \sum_{l=1}^{m} \left\| \left[(I + \tau A_{\lambda}) - (I + \tau A_{h,\lambda}) \right] \left(I + \tau A_{h,\lambda}\right)^{-l} \pi_h u_0 \right\|_1 \\ &\quad = \tau \sum_{l=1}^{m} \left\| \left[A_{\lambda} \left(1 + \tau A_{h,\lambda}\right)^{-l} - A_{h,\lambda} (1 + \tau A_{h,\lambda})^{-l} \right] \pi_h u_0 \right\|_1 \\ &\quad \leq \tau \sum_{l=1}^{m} \left(I_1 + I_2 \right), \end{aligned}$$

where

$$I_1 = \left\| \left[A_{\lambda} (I + \tau A_{h,\lambda})^{-l} - A_{\lambda} (I + \tau A_{\lambda})^{-l} \right] \pi_h u_0 \right\|_1, \quad I_2 = \left\| \left[A_{\lambda} (I + \tau A_{\lambda})^{-l} - A_{h,\lambda} (I + \tau A_{h,\lambda})^{-l} \right] \pi_h u_0 \right\|_1 .$$

In use of the $(2/\lambda)$-Lipschitz continuity of A_{λ}, we get

$$\begin{aligned} I_1 &\leq \frac{2}{\lambda} \left\| \left[(1 + \tau A_{h,\lambda})^{-l} - (I + \tau A_{\lambda})^{-l} \right] \pi_h u_0 \right\|_1 \\ &\leq \frac{2}{\lambda} \left(\left\| (I + \tau A_{h,\lambda})^{-l} \pi_h u_0 - (I + \tau A_{\lambda})^{-l} u_0 \right\|_1 + \left\| (I + \tau A_{\lambda})^{-l} u_0 - (1 + \tau A_{\lambda})^{-l} \pi_h u_0 \right\|_1 \right), \end{aligned}$$

which, together with (6.9), leads to

$$I_1 \leq \frac{2}{\lambda} \left(\left\| u^{\tau}_{h,\lambda}(t_l) - u^{\tau}_{\lambda}(t_l) \right\|_1 + \| (\pi_h - 1) u_0 \|_1 \right).$$

To estimate I_2, we note that

$$\begin{aligned}
\left[A_\lambda(I+\tau A_\lambda)^{-l}-A_{h,\lambda}(I+\tau A_{h,\lambda})^{-l}\right]\pi_h u_0 &= A_\lambda\left(I+\tau A_\lambda\right)^{-l}\pi_h u_0 - A_\lambda\left(I+\tau A_\lambda\right)^{-l}u_0\\
&\quad+\left\{A_\lambda(I+\tau A_\lambda)^{-l}u_0 - A_{h,\lambda}\left(I+\tau A_{h,\lambda}\right)^{-l}\pi_h u_0\right\}\\
&= A_\lambda\left(I+\tau A_\lambda\right)^{-l}\pi_h u_0 - A_\lambda\left(I+\tau A_\lambda\right)^{-l}u_0\\
&\quad+\frac{1}{\lambda}\left\{\left(I+\tau A_\lambda\right)^{-l}u_0-\left(I+\tau A_{h,\lambda}\right)^{-l}\pi_h u_0\right\}\\
&\quad-\frac{1}{\lambda}\left\{J_\lambda\left(I+\tau A_\lambda\right)^{-l}u_0 - J_{h,\lambda}\left(I+\tau A_{h,\lambda}\right)^{-l}\pi_h u_0\right\}\\
&= A_\lambda(I+\tau A_\lambda)^{-l}\pi_h u_0 - A_\lambda\left(I+\tau A_\lambda\right)^{-l}u_0\\
&\quad+\frac{1}{\lambda}\left\{(I+\tau A_\lambda)^{-l}u_0-(I+\tau A_{h,\lambda})^{-l}\pi_h u_0\right\}\\
&\quad-\frac{1}{\lambda}\left[J_\lambda - J_{h,\lambda}\pi_h\right](I+\tau A_\lambda)^{-l}u_0\\
&\quad-\frac{1}{\lambda}\left\{J_{h,\lambda}\pi_h(1+\tau A_\lambda)^{-l}u_0 - J_{h,\lambda}(I+\tau A_{h,\lambda})^{-l}\pi_h u_0\right\}.
\end{aligned}$$

We have

$$\| A_\lambda(I+\tau A_\lambda)^{-l}\pi_h u_0 - A_\lambda\left(I+\tau A_\lambda\right)^{-l}u_0\|_1 \le \frac{2}{\lambda}\|(\pi_h-1)u_0\|_1$$

as before. Moreover we obtain by (3.5)

$$\left\|\frac{1}{\lambda}\left\{J_{h,\lambda}\pi_h(1+\tau A_\lambda)^{-l}u_0 - J_{h,\lambda}(I+\tau A_{h,\lambda})^{-l}\pi_h u_0\right\}\right\|_1 \le \frac{C}{\lambda}\|\pi_h u_\lambda^\tau(t_l)-u_{h,\lambda}^\tau(t_l)\|_1$$

because $J_{h,\lambda}$ is a contraction in X_h with respect to $\|\cdot\|_{1,h}$. Those relations yield

$$I_2 \le \frac{C}{\lambda}\left(\|(\pi_h-1)u_0\|_1 + \left\|u_{h,\lambda}^\tau(t_l)-u_\lambda^\tau(t_l)\right\|_1 + \left\|\pi_h u_\lambda^\tau(t_l)-u_{h,\lambda}^\tau(t_l)\right\|_1 + \|(J_\lambda - J_{h,\lambda}\pi_h)u_\lambda^\tau(t_l)\|_1\right).$$

We can summarise the above relations as

$$\begin{aligned}
\left\|u_{h,\lambda}^\tau(t_m)-u_\lambda^\tau(t_m)\right\|_1 &\le \frac{C\tau}{\lambda}\sum_{l=1}^{m}\left\|u_{h,\lambda}^\tau(t_l)-u_\lambda^\tau(t_l)\right\|_1 + C\left(1+\frac{t}{\lambda}\right)\|(\pi_h-1)u_0\|_\infty\\
&\quad+\frac{C\tau}{\lambda}\sum_{l=1}^{m}\|(\pi_h-1)u_\lambda^\tau(t_l)\|_1 + \frac{\tau}{\lambda}\sum_{l=1}^{m}\|(J_\lambda-J_{h,\lambda}\pi_h)u_\lambda^\tau(t_l)\|_1.
\end{aligned}$$

Now applying the discrete Gronwall's lemma, we obtain

$$\begin{aligned}
\sup_{t\in[0,T]}\left\|u_{h,\lambda}^\tau(t)-u_\lambda^\tau(t)\right\|_1 &\le \exp(CT/\lambda)\left[C\left(1+\frac{T}{\lambda}\right)\|(\pi_h-1)u_0\|_\infty\right.\\
&\quad\left.+\frac{C\tau}{\lambda}\sum_{l=1}^{N}\|(\pi_h-1)u_\lambda^\tau(t_l)\|_1 + \frac{\tau}{\lambda}\sum_{l=1}^{N}\|(J_\lambda-J_{h,\lambda}\pi_h)u_\lambda^\tau(t_l)\|_1\right].
\end{aligned}$$

As is noted, $u_0\in W$ implies $u_\lambda^\tau(t_l)\in W$. Therefore, the right-hand side tends to 0 as $h\downarrow 0$ by (5.1), and the proof is complete. □

Now we are able to state the

Proof of Lemma 6.1. Since the semigroup generated by A_λ is a contraction on X, we have $\|u_\lambda(t)\|_1 \le \|u_0\|_1$ and hence

$$\|A_\lambda u_\lambda(t) - A_\lambda u_\lambda(s)\|_1 \le \frac{2}{\lambda}\|u_\lambda(t) - u_\lambda(s)\|_1 \le \frac{2}{\lambda}\left|\int_s^t \|A_\lambda u_\lambda(s)\|_1 \,\mathrm{d}s\right| \le \frac{4}{\lambda^2}|t-s|\,\|u_0\|_1\,. \tag{6.10}$$

We shall show that

$$\sup_{t\in[0,T]}\|u_\lambda^\tau(t) - u_\lambda(t)\|_1 \le \frac{2T\tau}{\lambda^2}\|u_0\|_1 \tag{6.11}$$

holds. In fact, we have

$$u_\lambda^\tau(t_{m+1}) - u_\lambda^\tau(t_m) + \tau A_\lambda u_\lambda^\tau(t_{m+1}) = 0$$

and

$$u_\lambda(t_{m+1}) - u_\lambda(t_m) + \int_{t_m}^{t_{m+1}} A_\lambda u_\lambda(s)\ \mathrm{d}s = 0$$

so that the error function $\mathrm{e}^\tau(t_m) = u_\lambda^\tau(t_m) - u_\lambda(t_m)$ satisfies

$$(I+\tau A_\lambda)u_\lambda^\tau(t_{m+1}) - (I+\tau A_\lambda)u_\lambda(t_{m+1}) = \mathrm{e}^\tau(t_m) + \int_{t_m}^{t_{m+1}}[A_\lambda u_\lambda(s) - A_\lambda u_\lambda(t_{m+1})]\ \mathrm{d}s.$$

This, together with (6.9), implies

$$\begin{aligned}\|\mathrm{e}^\tau(t_{m+1})\|_1 &\le \|(1+\tau A_\lambda)\,u_\lambda^\tau(t_{m+1}) - (1+\tau A_\lambda)\,u_\lambda(t_{m+1})\|_1\\ &\le \|\mathrm{e}^\tau(t_m)\|_1 + \int_{t_m}^{t_{m+1}}\|A_\lambda u_\lambda(t_{m+1}) - A_\lambda u_\lambda(s)\|_1\,\mathrm{d}s\end{aligned}$$

and hence

$$\|\mathrm{e}^\tau(t)\|_1 \le \sum_{l=1}^m \int_{t_{l-1}}^{t_l}\|A_\lambda u_\lambda(t_l) - A_\lambda u_\lambda(s)\|_1\,\mathrm{d}s.$$

In use of (6.10) we obtain

$$\|\mathrm{e}^\tau(t_m)\|_1 \le \sum_{l=1}^m \frac{4}{\lambda^2}\cdot\frac{\tau^2}{2}\|u_0\|_1 = \frac{2t_m\tau}{\lambda^2}\|u_0\|_1\,,$$

which yields (6.11).

Similarly, we have

$$\sup_{t\in[0,T]}\left\|u_{h,\lambda}^\tau(t) - u_{h,\lambda}(t)\right\|_{1,h} \le \frac{2T\tau}{\lambda^2}\|\pi_h u_0\|_{1,h}\,.$$

Therefore, it follows from (3.5) that

$$\sup_{t\in[0,T]}\|u_{h,\lambda}(t) - u_\lambda(t)\|_1 \le \frac{CT\tau}{\lambda^2}\|u_0\|_\infty + \sup_{t\in[0,T]}\left\|u_{h,\lambda}^\tau(t) - u_\lambda^\tau(t)\right\|_1.$$

Now, send $h\downarrow 0$ and then $\tau\downarrow 0$. Then, relation (6.3) follows from (6.7). $\square$

7. Convergence of semigroup

We complete the proof of (1.12), one of the main result of the present paper.

Theorem 7.1. *If Ω is convex and is provided with* (D) *(in the case of $n = 3$), $\{\mathcal{T}_h\}$ satisfies* (H1), (H2), *and* (H3), *and f is strictly increasing, then it holds that*

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} \|S_h(t)\pi_h u_0 - S(t)u_0\|_1 = 0, \tag{7.1}$$

where $u_0 \in W$.

We begin with the following.

Lemma 7.1. *We have*

$$\left\|K_h^{-1} L_h \pi_h L^{-1}\right\|_{L^2(\Omega),L^2(\Omega)} \le C. \tag{7.2}$$

Proof. Let R_h and P_h be the Ritz and the orthogonal projection operators defined as (2.4) and (2.5), respectively. Then it holds that $L_h^{-1}P_h = R_h L^{-1}$ and hence

$$L_h\pi_h L^{-1} = L_h\left(\pi_h - R_h\right)L^{-1} + P_h = L_h R_h\left(\pi_h - 1\right)L^{-1} + P_h. \tag{7.3}$$

We also have

$$\|\nabla R_h v\|_2 \le C\,\|\nabla v\|_2 \tag{7.4}$$

for $v \in H_0^1(\Omega)$ and

$$\|(\pi_h - 1)\,v\|_2 + h\,\|\nabla(\pi_h - 1)\,v\|_2 \le Ch^2\,\|v\|_{H^2(\Omega)} \tag{7.5}$$

for $v \in H_0^1(\Omega)\cap H^2(\Omega)$. Furthermore, (H3) leads to

$$\|\nabla\chi_h\|_2 \le Ch^{-1}\,\|\chi_h\|_2$$

for $\chi_h \in X_h$. See [17], Sect. 1.4 or [7] for those fundamental facts.

From the last fact, we have

$$|(L_h\chi_h, \psi_h)| = |(\nabla\chi_h, \nabla\psi_h)| \le Ch^{-1}\,\|\nabla\chi_h\|_2\,\|\psi_h\|_2$$

for $\chi_h, \psi_h \in X_h$, and hence

$$\|L_h\chi_h\|_2 \le Ch^{-1}\,\|\nabla\chi_h\|_2$$

for $\chi_h \in X_h$. This, together with (7.4) and (7.5), implies that

$$\begin{aligned}\left\|L_h R_h\left(\pi_h - 1\right)L^{-1}v\right\|_2 &\le Ch^{-1}\left\|\nabla R_h\left(\pi_h - 1\right)L^{-1}v\right\|_2\\ &\le Ch^{-1}\left\|\nabla\left(\pi_h - 1\right)L^{-1}v\right\|_2\\ &\le C\left\|L^{-1}v\right\|_{H^2(\Omega)}\\ &\le C\,\|v\|_2\end{aligned}$$

for $v \in L^2(\Omega)$. As a result, by (7.3), we obtain

$$\|L_h\pi_h L^{-1}v\|_2 \le C\|v\|_2 + \|v\|_2$$

for $v \in L^2(\Omega)$. Therefore, (7.2) follows from $\left\|K_h^{-1}\right\|_{L^2(\Omega),L^2(\Omega)} \le C$. $\square$

Now we can state the

Proof of Theorem 7.1. It is made of two steps. Let $u_0 \in W$.

Step 1. We show that the theorem is true under the additional assumption

$$f(u_0) \in H^2(\Omega). \tag{7.6}$$

In doing so, for $\lambda > 0$, we introduce solutions $u_\lambda(t)$ and $u_{h,\lambda}(t)$ of (6.1) and (6.2), respectively, and show

$$\sup_{t\in[0,T]} \|u_\lambda(t) - u(t)\|_1 \le 3(\sqrt{\lambda T} + \lambda)\, \|Lf(u_0)\|_1\,, \tag{7.7}$$

$$\sup_{0\le t\le T} \|u_{h,\lambda}(t) - u_h(t)\|_{1,h} \le 3C(\sqrt{\lambda T} + \lambda)\, \|Lf(u_0)\|_2\,. \tag{7.8}$$

In fact, a formula below (4.5) of [29] assures

$$\left\| S(t)u_0 - J_\lambda^{[t/\lambda]} u_0 \right\|_1 \le 2(\lambda^2 + \lambda t)^{1/2}\, \|Au_0\|_1 \le 2(\sqrt{\lambda t} + \lambda)\, \|Au_0\|_1$$

for $J_\lambda = (I + \lambda A)^{-1}$. Similarly a formula above (3.49) of [29] reads as

$$\left\| S_\lambda(t)u_0 - J_\lambda^{[t/\lambda]} u_0 \right\|_1 \le (\sqrt{\lambda t} + \lambda)\, \|Au_0\|_1\,.$$

Therefore, for $u_0 \in D(A)$,

$$\|S(t)u_0 - S_\lambda(t)u_0\|_1 \le 3(\sqrt{\lambda t} + \lambda)\, \|Au_0\|_1\,,$$

which implies (7.7).

Similarly, we obtain

$$\|u_{h,\lambda}(t) - u_h(t)\|_{1,h} \le 3(\sqrt{\lambda t} + \lambda)\, \|A_h\pi_h u_0\|_{1,h}\,. \tag{7.9}$$

Here, we have by (7.2) that

$$\begin{aligned} \|A_h\pi_h u_0\|_{1,h} &\le \left\|K_h^{-1} L_h \pi_h f(u_0)\right\|_{1,h} \\ &\le C \left\|K_h^{-1} L_h \pi_h L^{-1} \cdot Lf(u_0)\right\|_2 \\ &\le C\, \|Lf(u_0)\|_2 \end{aligned}$$

from the assumption. Combining this with (7.9), we get (7.8).

In use of (7.7) and (7.8), we have

$$\begin{aligned} \sup_{t\in[0,T]} \|u_h(t) - u(t)\|_1 &\le \sup_{0\le t\le T} \|u_h(t) - u_{h,\lambda}(t)\|_1 + \sup_{0\le t\le T} \|u_{h,\lambda}(t) - u_\lambda(t)\|_1 + \sup_{t\in[0,T]} \|u_\lambda(t) - u(t)\|_1 \\ &\le \sup_{t\in[0,T]} \|u_{h,\lambda}(t) - u_\lambda(t)\|_1 + \ C(\sqrt{\lambda T} + \lambda)\, \|Lf(u_0)\|_2\,. \end{aligned}$$

Hence by (6.3)

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} \|u_h(t) - u(t)\|_1 \le C\left(\sqrt{\lambda T} + \lambda\right) \|Lf(u_0)\|_2\,. \tag{7.10}$$

Then (7.1) follows by sending $\lambda \downarrow 0$.

Step 2. We deal with general $u_0 \in W$. For this purpose, we recall that f is strictly increasing and set $v_0 = f(u_0)$. We take a sequence $\{v_j\}_{j=1}^{\infty} \subset H^2(\Omega) \cap W$ satisfying $\|v_j - v_0\|_\infty \to 0$ as $j \to \infty$. Then, $u_j = f^{-1}(v_j)$ satisfies that

$$f(u_j) \in H^2(\Omega) \cap W \qquad \text{and} \qquad \lim_{j\to\infty} \|u_j - u_0\|_\infty = 0,$$

and, as saw in Step 1, we know

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} \|S_h(t)\pi_h u_j - S(t)u_j\|_1 = 0 \quad (j = 1, 2, \ldots).$$

On the other hand, by (1.9), (3.4) and (3.5), we have

$$\begin{aligned}
\|S(t)u_0 - S_h(t)\pi_h u_0\|_1 &\le \|S(t)u_0 - S(t)u_j\|_1 + \|S(t)u_j - S_h(t)\pi_h u_j\|_1 + \|S_h(t)\pi_h u_j - S_h(t)\pi_h u_0\|_1 \\
&\le \|u_0 - u_j\|_1 + \|S(t)u_j - S_h(t)\pi_h u_j\|_1 + C\|\pi_h u_j - \pi_h u_0\|_1 \\
&\le (|\Omega| + C)\|u_j - u_0\|_\infty + \|S(t)u_j - S_h(t)\pi_h u_j\|_1 .
\end{aligned}$$

This leads to

$$\sup_{t\in[0,T]} \|S_h(t)\pi_h u_0 - S(t)u_0\|_1 \le (|\Omega| + C)\|u_j - u_0\|_\infty + \sup_{0\le t\le T} \|S_h(t)\pi_h u_j - S(t)u_j\|_1 .$$

Making $h \downarrow 0$ and then $j \to \infty$, we obtain (7.1) and the proof is complete. □

We describe some observations on a generalization of Theorem 7.1.
Let $\lambda > 0$ and β_λ be the Yosida regularization of $\beta = f^{-1}$. Putting $f_\lambda = \beta_\lambda^{-1}$, we introduce the semigroup $\{S_h^\lambda(t)\}$ generated by $A_h^\lambda v = L_h \pi_h f_\lambda(v)$ $(v \in W)$.

Proposition 7.1. *Suppose that the same assumptions on Ω and $\{\mathcal{T}_h\}$ as that of Theorem 7.1 hold. Let $u_0 \in W$ and suppose that there is $\{u_j\}_{j=1}^{\infty} \subset W$ such that*

$$f(u_j) \in H^2(\Omega) \quad \textit{and} \quad \lim_{j\to\infty} \|u_j - u_0\|_\infty = 0. \tag{7.11}$$

Furthermore, assume that there is a positive function $\varepsilon_T(\lambda)$ of $\lambda > 0$ such that $\varepsilon_T(\lambda) \to 0$ as $\lambda \downarrow 0$ which is independent of h and that

$$\sup_{t\in[0,T]} \|S_h(t)\pi_h v - S_h^\lambda(t)\pi_h v\|_1 \le \varepsilon_T(\lambda)\|Lf(v)\|_2 \tag{7.12}$$

for $v \in W$ with $f(v) \in H^2(\Omega)$. (We note that (7.12) *is comparable to* (7.8)*.) Then we have* (7.1)*, even if f is not strictly increasing.*

Remark 7.1. Unfortunately, it is not obvious that $\varepsilon_T(\lambda)$ in (7.12) really exists or not. In [11], Cockburn and Gripenberg considered the case of $\Omega = \mathbb{R}^n$ and derived an explicit continuous-dependence on f of solutions to (1.4). However, the case of a bounded Ω is open, and it seems to be difficult to derive a corresponding estimate for solutions to the discrete problem (2.3). It is an important and interesting open problem.

Before stating the proof of Proposition 7.1, we give a class (NI) of nonlineality f and inital data u_0 which ensures the condition (7.11). Actually, (NI) contains porous media, fast diffusion and Stefan nonlinearlities.

(NI) For $f \in C(\mathbb{R})$ with $f(0) = 0$ and $u_0 \in W$, the following conditions are satisfied:

(i) There are $\{a_i^\pm\}_{i=1}^m$ with $\cdots < a_i^- < a_i^+ < a_{i+1}^- < a_{i+1}^+ < \cdots$ and $\{b_i\}_{i=1}^m$ such that $f(s) = b_i$ for all $s \in Q_i = (a_i^-, a_i^+)$ for $i = 1, \ldots, m$;

(ii) f is strictly increasing on $\mathbb{R}\backslash\overline{Q}$, where $Q = \displaystyle\bigcup_{i=1}^m Q_i$;

(iii) $\displaystyle\lim_{s\to a_i^\pm \pm 0} \frac{f(a_i^\pm)-f(s)}{a_i^\pm - s} < \infty$ for $i = 1,\dots,m$.

(iv) Every $\partial D_i \backslash \partial\Omega$ is a finite number of hypersurfaces of class C^1, where

$$D_i = D_i(u_0) = \left\{x \in \Omega |\ a_{i-1}^+ < u_0(x) < a_i^-\right\},$$

for $i = 1,\dots,m+1$ with $a_{-1}^+ = -\infty$ and $a_{m+1}^- = \infty$. (We note that $\partial D_i \cap \partial\Omega = \emptyset$ if $a_{i-1}^+ \neq 0$ and $a_i^- \neq 0$.)

Then we have the

Lemma 7.2. *Let $f \in C(\mathbb{R})$ with $f(0) = 0$ and $u_0 \in W$. If (NI) is satisfied, then there is $\{u_j\}_{j=1}^\infty \subset W$ satisfying* (7.11).

Proof. By (iv), we can take $\{u_j\}_{j=1}^\infty \subset W$ such that

$$f(u_j)|_{\overline{D}_i} \in C^1(\overline{D}_i) \cap H^2(D_i) \quad \text{and} \quad \|u_j - u_0\|_\infty \to 0 \text{ as } j \to \infty.$$

Because of $f(u_j) = b_i$ on $\overline{\{x \in \Omega|\ a_i^- < u_0(x) < a_i^+\}}$, we have $f(u_j) \in H^2(\Omega)$. □

We finally state the

Proof of Proposition 7.1. Let $\lambda > 0$. We introduce the semigroup $\{S^\lambda(t)\}$ generated by $A^\lambda v = Lf_\lambda(v)$ $(v \in D(A^\lambda))$. In [2], Bénilan *et al.* proved

$$\lim_{\lambda\to 0} \sup_{t\in[0,T]} \|S(t)u_0 - S^\lambda(t)u_0\|_1 = 0 \tag{7.13}$$

for all $u_0 \in X$.

Since f_λ is strictly increasing with $f_\lambda(0) = 0$, we can apply Theorem 7.1 and obtain

$$\lim_{h\to 0} \sup_{0\le t\le T} \|S^\lambda(t)u_j - S_h^\lambda(t)\pi_h u_j\|_1 = 0, \quad (\lambda > 0,\ j = 1,2,\dots). \tag{7.14}$$

We observe that

$$\begin{aligned}\sup_{t\in[0,T]} \|S_h(t)\pi_h u_j - S(t)u_j\|_1 \le{}& \sup_{t\in[0,T]} \left\|S_h(t)\pi_h u_j - S_h^\lambda(t)\pi_h u_j\right\|_1 \\ &+ \sup_{t\in[0,T]} \left\|S_h^\lambda(t)\pi_h u_j - S^\lambda(t)u_j\right\|_1 + \sup_{t\in[0,T]} \left\|S^\lambda(t)u_j - S(t)u_j\right\|_1.\end{aligned}$$

This, together with (7.12) and (7.14), implies

$$\lim_{h\to 0} \sup_{t\in[0,T]} \|S_h(t)\pi_h u_j - S(t)u_j\|_1 \le \varepsilon_T(\lambda)\|Lf(u_j)\|_2 + \sup_{t\in[0,T]} \left\|S^\lambda(t)u_j - S(t)u_j\right\|_1.$$

Hence, from (7.13), we obtain

$$\lim_{h\to 0} \sup_{t\in[0,T]} \|S_h(t)\pi_h u_j - S(t)u_j\|_1 = 0, \quad (j = 1,2,\dots)$$

by sending $\lambda \downarrow 0$.

Then, in virtue of (7.11), we can repeat the argument of Step 2 in the proof of Theorem 7.1 and establish (7.1). □

TABLE 1. Relative L^1 error $E_1(N)$ and convergence rate α_N.

	$\gamma = 1.5$		$\gamma = 3.0$		$\gamma = 6.0$	
N	E_1	α_N	E_1	α_N	E_1	α_N
8	0.0529		0.0320		0.1780	
16	0.0153	1.78	0.0147	1.12	0.1542	0.21
32	0.0041	1.90	0.0074	0.99	0.1290	0.26
64	0.0010	1.99	0.0025	1.57	0.1090	0.24

8. Numerical examples

We assume that Ω is a unit square: $\Omega = \{0 < x_1 < 1,\ 0 < x_2 < 1\}$. We take $\mathcal{T}_h$ as a uniform mesh composed of $2N^2$ equal right triangles for $N \in \mathbb{N}$; each sides of Ω is divided into N intervals of same length, and then each small-square is decomposed into two equal triangles by a diagonal. Put $h = 1/N$. The time discretization makes use of the forward difference formula. Namely we find $\{u_h^\tau(t_m)\}_{m=0}^{\tilde{N}} \subset X_h$ satisfying

$$\begin{cases} \dfrac{u_h^\tau(t_{m+1}) - u_h^\tau(t_m)}{\tau} + K_h^{-1} L_h f\left(u_h^\tau(t_m)\right) = 0 \quad (0 \le m \le \tilde{N}) \\ u_h^\tau(0) = \pi_h u_0, \end{cases}$$

where $\tilde{N} \in \mathbb{N}$ and $\tau = T/\tilde{N}$. We choose a sufficiently small τ relative to h, (specifically we take $\tau = h^2/100$,) since we are interested in the effect of the space discretization on the accuracy of the scheme.

We recall that Barenblatt's self-similar solution [18]

$$u^*(x_1, x_2, t) = (t + T_0)^{-1/\gamma} \left[a^2 - \frac{(\gamma - 1)\{(x_1 - 1/2)^2 + (x_2 - 1/2)^2\}}{4\gamma^2 (t + T_0)^{1/\gamma}} \right]_+^{\frac{1}{\gamma - 1}}$$

solves $u_t - \Delta u^\gamma = 0$ and $u|_{\partial\Omega} = 0$ with the initial data $u_0(x_1, x_2) = u^*(x_1, x_2, 0)$ in a generalised sense. Here $a > 0$, $T_0 > 0$, and $\gamma > 1$ are given constants. We compute the discrete relative L^1 error:

$$E_1(N) = \left(\sum_{a \in I_h} |U_a| \right)^{-1} \sum_{a \in I_h} |U_a - u^*(a, T)|,$$

where we have put $u_h^\tau(T) = \sum U_a w_a$. We may suppose that the effect of the time discretization is relatively negligible because of $\tau = h^2/100$, and assume that $E_1(N) = Ch^\alpha = CN^{-\alpha}$.
We estimate the rate of convergence α by

$$\alpha = \alpha_N = \frac{\log E_1(N/2) - \log E_1(N)}{\log 2}.$$

In Table 1, we compare the result taking $\gamma = 3/2, 3$, and 6. These results show that the L^1 convergence really takes place. The shape of f affects the accuracy of the scheme. Especially, if the shape of f is like to a linear function, our scheme has a high accuracy. We also observe that the rate of convergence continuity depends on f.

Acknowledgements. The authors thank Professor B. Cockburn for his suggestions.

References

[1] R.A. Adams, *Sobolev Spaces*. Academic Press, New York, London (1975).

[2] P. Bénilan, M.G. Crandall and P. Sacks, Some L^1 existence and dependence results for semilinear elliptic equations under nonlinear boundary conditions. *Appl. Math. Optim.* **17** (1988) 203–224.

[3] A.E. Berger, H. Brezis and J.C.W Rogers, A numerical method for solving the problem $u_t - \Delta f(u) = 0$. *RAIRO Anal. Numer.* **13** (1979) 297–312.

[4] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*. Springer (1994).

[5] H. Brezis and A. Pazy, Convergence and approximation of semigroups of nonlinear operators in Banach spaces. *J. Funct. Anal.* **9** (1972) 63–74.

[6] H. Brezis and W. Strauss, Semi-linear second-order elliptic equations in L^1. *J. Math. Soc. Japan* **25** (1973) 565–590.

[7] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam (1978).

[8] P.G. Ciarlet, Basic Error Estimates for Elliptic Problems, in *Finite Element Methods (Part 1)*, P.G. Ciarlet and J.L. Lions Eds., Handbook of Numerical Analysis, 17–351, Elsevier Science Publishers B.V., Amsterdam (1991).

[9] P.G. Ciarlet and P.A. Raviart, Maximum principle and uniform convergence for the finite element method. *Comput. Methods Appl. Mech. Engrg.* **2** (1973) 17–31.

[10] J.F. Ciavaldini, Analyse numérique d'un problème de Stefan à deux phases par une méthode d'éléments finis. *SIAM J. Numer. Anal.* **12** (1975) 464–487.

[11] B. Cockburn and G. Gripenberg, Continuous dependence on the nonlinearities of solutions of degenerate parabolic equations. *J. Differential Equations* **151** (1999) 231–251.

[12] M.G. Crandall and T. Liggett, Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.* **93** (1971) 265–293.

[13] C.M. Elliott, Error analysis of the enthalpy method for the Stefan problem. *IMA J. Numer. Anal.* **7** (1987) 61–71.

[14] C.M. Elliott and J.R. Ockendon, *Weak and Variational Methods for Moving Boundary Problems*. Pitman, Boston. *Res. Notes Math.* **59** (1982).

[15] A. Friedman, *Variational Principles and Free-Boundary Problems*. Wiley, New York (1982).

[16] H. Fujii, Some remarks on finite element analysis of time-dependent field problems, in *Theory and Practice in Finite Element Structural Analysis*, University of Tokyo Press, Tokyo (1973) 91–106.

[17] H. Fujita, N. Saito and T. Suzuki, *Operator Theory and Numerical Methods*. North-Holland, Amsterdam (2001).

[18] B.H. Gilding and L.A. Peletier, On a class of similarity solutions of the porous media equation. *J. Math. Anal. Appl.* **55** (1976) 351–364.

[19] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston (1985).

[20] W. Jäger and J. Kačur, Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes. *RAIRO Modél. Math. Anal. Numér.* **29** (1995) 605–627.

[21] J. Kačur, A. Handlovicová and M. Kacurová, Solution of nonlinear diffusion problems by linear approximation schemes. *SIAM J. Numer. Anal.* **30** 1703-1722 (1993).

[22] T. Kato, Schrödinger operators with singular potentials. *Israel J. Math.* **13** (1972) 135–148.

[23] M.N. Le Roux, Semi-discretization in time for a fast diffusion equation. *J. Math. Anal. Appl.* **137** (1989) 354–370.

[24] M.N. Le Roux and P.E. Mainge, Numerical solution of a fast diffusion equation. *Math. Comp.* **68** (1999) 461–485.

[25] P. Lesaint and J. Pousin, Error estimates for a nonlinear degenerate parabolic equation. *Math. Comp.* **59** (1992) 339–358.

[26] E. Magenes, R.H. Nochetto and C. Verdi, Energy error estimates for a linear scheme to approximate nonlinear parabolic problems. *RAIRO Modél. Math. Anal. Numér.* **21** (1987) 655–678.

[27] E. Magenes, C. Verdi and A. Visintin, Semigroup approach to the Stefan problem with non-linear flux. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) **75** (1983) 24–33.

[28] E. Magenes, C. Verdi, and A. Visintin, Theoretical and numerical results on the two-phase Stefan problem. *SIAM J. Numer. Anal.* **26** (1989) 1425–1438.

[29] I. Miyadera, Nonlinear Semigroups. *Amer. Math. Soc. Colloq. Publ.* (1992).

[30] R.H. Nochetto, Error estimates for two-phase Stefan problems in several space variables. I. Linear boundary conditions. *Calcolo* **22** (1985) 457–499.

[31] P.H. Nochetto, and C. Verdi, Approximation of degenerate parabolic problems using numerical integration. *SIAM J. Numer. Anal.* **25** (1988) 784–814.

[32] L.A. Peletier, The porous media equation, in *Applications of Nonlinear Analysis in the Physical Sciences* (Bielefeld, 1979), Surveys Reference Works Math., 6, Pitman, Boston, Mass.-London (1981) 229–241.

[33] R. Rannacher and R. Scott, Some optimal error estimates for piecewise linear finite element approximation. *Math. Comp.* **38** (1982) 437–445.

[34] M. Rose, Numerical methods for flows through porous media, I. *Math. Comp.* **40** (1983) 435–467.

[35] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* **54** (1990) 483–493.

[36] R.E. White, An enthalpy formulation of the Stefan problem. *SIAM J. Numer. Anal.* **19** (1982) 1129–1157.