ANALYSIS OF MULTIGRID PRECONDITIONING FOR IMPLICIT PDE SOLVERS FOR DEGENERATE PARABOLIC EQUATIONS[*]

MARCO DONATELLI[†], MATTEO SEMPLICE [‡], AND STEFANO SERRA-CAPIZZANO [§]

Abstract. In this paper an implicit numerical method designed for nonlinear degenerate parabolic equations is proposed. A convergence analysis and the study of the related computational cost are provided. In fact, due to the nonlinear nature of the underlying mathematical model, the use of a fixed point scheme is required. The chosen scheme is the Newton method and its convergence is proven under mild assumptions. Every step of the Newton method implies the solution of large, locally structured, linear systems. A special effort is devoted to the spectral analysis of the relevant matrices and to the design of appropriate multigrid preconditioned Krylov methods. Numerical experiments for the validation of our analysis complement this contribution.

Key words. nonlinear degenerate parabolic equations, locally Toeplitz matrices, preconditioning and multigrid methods.

AMS subject classifications. 65N12, 65F10

1. Introduction. As a prototype we first consider a single equation of the form

$$\frac{\partial u}{\partial t} = \nabla \cdot (D(u)\nabla u), \tag{1.1}$$

in both one and two dimensional domains. $D(u)$ is a non-negative function and the equation is called degenerate whenever $D(u)$ vanishes. For the convergence analysis of our numerical methods, we will require that $D(u)$ is at least differentiable and $D'(u)$ is Lipschitz continuous, while the existence of solutions is guaranteed under the milder assumption of continuity (see [24]).

In the literature, degenerate parabolic equations have been discretized mainly using explicit or semi-implicit methods, thus avoiding to solve the nonlinear equations. A remarkable class of methods arises directly from the so-called non-linear Chernoff formula [6] for time advancement, coupling it with a spatial discretization: for finite differences this was started in [3] and for finite elements by [15]. The schemes in this class share a timestep restriction of the form $\Delta t = O(h^2)$ for stability, but, especially in the finite element setting, there is a mature theory that includes also error estimators and h-adaptivity. More recently, another class related to the relaxation approximation emerged: such numerical procedures exploit high order non-oscillatory methods typical of the approximation of conservation laws and their convergence can be proved making use of semigroup arguments similar to those relevant for proving the Chernoff formula [7]. Finally, we wish to point out the paper [10], where the authors derive conservative schemes for degenerate convection-diffusion equations with the elliptic operator in the form $\partial_{xx}^2(K(u))$, which is obviously equivalent to (1.1) in one space dimension, setting $K(u) = \int_0^u D(\xi) \mathrm{d}\xi$.

The present investigation is part of the search for suitable numerical techniques to integrate for long times nonlinear, possibly degenerate, parabolic equations like

[†]Dipartimento di Fisica e Matematica, Università dell'Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy (`marco.donatelli@uninsubria.it`)

[‡](`matteo.semplice@uninsubria.it`)

[§](`stefano.serrac@uninsubria.it`)

those appearing in models for monument degradation (see [2]) when chemical/microbiological pollutants are taken into consideration. We wish to point out that the techniques developed here have applications that go beyond the aforementioned models. For example, again in the area of conservation of the cultural heritage, they could be adapted to numerically investigate the more complete sulfation model described in [1] and the consolidation model presented in [8]. Some applications in the field of monument conservation have been presented in [18], where the mathematical tools developed in the present paper are employed for forecasting marble deterioration.

In this paper we consider a time discretization based on the Crandall-Liggett formula (Implicit Euler)

$$U(t^n, x) - \Delta t L_D(U(t^n, x)) = U(t^{n-1}, x), \qquad (1.2)$$

where time has been discretized, $\Delta t = t^n - t^{n-1}$, and $-L_D(\cdot)$ denotes the elliptic operator $u \mapsto -\nabla \cdot (D(u)\nabla u)$. We point out that using the Crank-Nicholson scheme instead of (1.2) would give a higher order scheme, but somewhat complicate the notation of this paper without affecting the line of the proofs. For such a generalization, at least in a specific case, see [18].

The computation of the numerical solution $U(t^n, x)$ with (1.2) requires to solve a nonlinear equation whose form is determined by the elliptic operator and the nonlinear function $D(u)$, but the convergence is guaranteed without restrictions on the time step Δt [9]. Furthermore, due to the nonlinear nature of the underlying mathematical model, the use of a fixed point scheme is required and the choice of the Newton-like methods implies the solution at every step of large, locally structured (in the sense of [22]) linear systems. A special effort is devoted to the spectral analysis of the relevant matrices and to the design of appropriate iterative solvers, with special attention to preconditioned Krylov methods and to multigrid procedures (see [12, 17, 13, 23] and references therein for a general treatment of iterative solvers).

The paper is organized as follows. In Section 2 we couple the time discretization (1.2) with a spatial discretization based on finite differences and set up a Newton method for the resulting system of nonlinear equations. In Section 3 we prove the convergence of the Newton methods under a mild restriction on Δt. In Section 4 we consider various iterative methods for the solution of the inner linear systems involved in the Newton method. A spectral analysis of the related matrix structures is provided in order to give an appropriate motivation for the good behaviour of the proposed iterative solvers. In Section 5 we perform some numerical tests. Finally, a conclusion section with a short plan for future investigations completes the paper.

2. Space and time discretization. In order to discretize equations like (1.1), we will employ a time semi-discretization given by the Crandall-Liggett formula and a space discretization based on finite differences, explained in the following subsection. The latter numerical choice leads to a system of coupled nonlinear equations that need to be solved at each discrete timestep in order to compute the solution of the PDE: this is achieved using the Newton method, as detailed in Subsection 2.3, where we also prove and comment convergence results.

2.1. One space dimension. We take into consideration a standard discretization in space using finite differences. Denoting $x_\xi = a + \xi h$, we consider $N + 2$ points with equal spacing $h = (b - a)/(N + 1)$ in the interval $[a, b]$ and we denote by u_k^n the approximate solution at time t^n and location x_k, where $k = 0, \ldots, N + 1$. When considering Dirichlet boundary conditions, the values u_0 and u_{N+1} are known and can

be eliminated by the equations. Let thus $\mathbf{u}^n$ be the vector of size N containing the collection of the values u_k^n, for $k = 1, \ldots, N$. When no potential confusion arises, we sometimes drop in both notations the superscript indicating the time level. Boundary conditions of Neumann or Robin type can be treated in similar ways.

In the following, we denote by $\mathsf{tridiag}_k[\beta_k, \alpha_k, \gamma_k]$ a square tridiagonal matrix where the kth row has entries β_k on the lower diagonal, α_k on the main diagonal, and γ_k on the upper diagonal. We also denote with $\mathrm{diag}(\alpha_k)$ the square diagonal matrix with α_k on the kth row. The order of such matrices is clear from the context and usually it is equal to N with k ranging from 1 to N.

We choose a standard 3-points second order approximation of the differential operator $(D(u)u_x)_x$. Namely, the action of the elliptic differential operator on $\mathbf{u}$ is described by $-\frac{1}{h^2}L_{D(\mathbf{u})}\mathbf{u}$, where the tridiagonal matrix

$$-L_{D(\mathbf{u})} = \mathsf{tridiag}_k[-D_{k-1/2}, D_{k-1/2} + D_{k+1/2}, -D_{k+1/2}] \qquad (2.1)$$

contains the values $D_{j+1/2} = (D(u_{j+1}) + D(u_j))/2$, $j = 0, \ldots, N$, and thus depends nonlinearly on the u_j's. It should be noticed that the latter is a second order approximation of $D(u(x_{j+1/2}))$ since u_k^n differs from $u(t^n, x_k)$ by $O(h^2)$ thanks to the second order scheme and since, by standard Taylor expansions, we have

$$D_{j+1/2} = \frac{D(u(x_{j+1})) + D(u(x_j))}{2} = D(u(x_{j+1/2})) + O(h^2),$$

under the mild assumption that $D(u)_x(\cdot)$ is Lipschitz continuous.

REMARK 2.1. *$L_{D(\mathbf{u})}$ is a symmetric real tridiagonal matrix. If $D(u(\cdot))$ is a nonnegative function, then the matrix $-L_{D(\mathbf{u})}$ is always positive semidefinite, since the matrix is weakly diagonally dominant by row. When $D(u(\cdot))$ is positive the matrix becomes irreducible and therefore we observe also invertibility.*

2.2. Two space dimensions. We now describe a straightforward generalization to a two-dimensional case. To this end, we consider a rectangular domain $\Omega = [a_0, a_1] \times [b_0, b_1] \subset \mathbb{R}^2$ and the grid points $x_{i,j} = (a_0 + ih, b_0 + jk)$. For simplicity and without loss of generality, we also assume that the region Ω is square and choose identical discretization steps in the two directions (i.e. $h = k$), so that using $N + 2$ points per direction we have $h = k = (a_1 - a_0)/(N + 1) = (b_1 - b_0)/(N + 1)$. The grid is thus composed of the $(N + 2)^2$ points $x_{i,j}$ for i and j ranging from 0 to $N + 1$. We denote with $u_{i,j}$ the numerical value approximating $u(x_{i,j})$. Of course, as in the one-dimensional case the use of Dirichlet boundary conditions reduces the gridding to the N^2 internal points; also in this case other boundary conditions can be considered in a similar way.

In order to write in matrix form the approximated differential operator above, we must choose an ordering of the unknowns $u_{i,j}$, arranging them into a vector $\mathbf{u}$ and approximate

$$(\nabla \cdot (D(u)\nabla u)(x_j))_{i,j=1}^N \simeq \frac{1}{h^2}L_{D(\mathbf{u})}\,\mathbf{u}.$$

The positions of the nonzero entries of the matrix $L_{D(\mathbf{u})}$ of course depend on the chosen ordering, so here we keep a double-index notation for the elements of $\mathbf{u}$ and

of the matrix entries. Therefore, $L_{D(\mathbf{u})}$ has entries

$$\left[L_{D(\mathbf{u})}\right]_{i,j}^{l,m} = \delta_{i,l}\delta_{j,m}\left(-D_{i+1/2,j} - D_{i-1/2,j} - D_{i,j+1/2} - D_{i,j-1/2}\right)$$
$$+ \delta_{l,i+1}\delta_{m,j}D_{i+1/2,j} + \delta_{l,i-1}\delta_{m,j}D_{i-1/2,j}$$
$$+ \delta_{l,i}\delta_{m,j+1}D_{i,j+1/2} + \delta_{l,i}\delta_{m,j-1}D_{i,j-1/2} \quad (2.2)$$

on the $(i,j)^{\text{th}}$ row and $(l,m)^{\text{th}}$ column, where we denoted $D_{i+1/2,j} = (D_{i+1,j}+D_{i,j})/2$ and $D_{i,j+1/2} = (D_{i,j+1} + D_{i,j})/2$.

REMARK 2.2. *Remark 2.1 has a natural counterpart in this two-dimensional setting, where the word tridiagonal is replaced by block tridiagonal with tridiagonal blocks.*

2.3. The nonlinear system and the Newton iteration. Following the Crandall-Liggett formula (1.2), in order to compute $\mathbf{u}^n$ from $\mathbf{u}^{n-1}$, we need to solve the nonlinear vector equation

$$\mathbf{u}^n = \mathbf{u}^{n-1} + \frac{\Delta t}{h^2}L_{D(\mathbf{u}^n)}\mathbf{u}^n.$$

Thus we set up Newton iterations for the vector function

$$F(\mathbf{u}) = \mathbf{u} - \frac{\Delta t}{h^2}L_{D(\mathbf{u})}\mathbf{u} - \mathbf{u}^{n-1}. \quad (2.3)$$

In the following, we denote $\mathbf{u}^{n,s}$ the sth Newton iterate for the computation of $\mathbf{u}^n$.

In the one-dimensional case, the generic partial derivative of $F(\mathbf{u})$ is

$$\frac{\partial F_k}{\partial u_j} = \delta_{jk} - \frac{\Delta t}{h^2}\left.L_{D(\mathbf{u})}\right|_{j,k} - \frac{\Delta t}{2h^2}\left[\begin{array}{c}\delta_{k-1,j}D'_{k-1}(u_{k-1} - u_k)+ \\ + \delta_{k,j}D'_k(u_{k-1} - 2u_k + u_{k+1})+ \\ + \delta_{k+1,j}D'_{k+1}(u_{k+1} - u_k).\end{array}\right], \quad (2.4)$$

so that the Jacobian is

$$F'(\mathbf{u}) = X_N(\mathbf{u}) + Y_N(\mathbf{u}), \quad (2.5a)$$

$$X_N(\mathbf{u}) = I_N - \frac{\Delta t}{h^2}L_{D(\mathbf{u})}, \quad (2.5b)$$

$$Y_N(\mathbf{u}) = -\frac{\Delta t}{2h^2}T_N(\mathbf{u})\operatorname{diag}_k[D'_k], \quad (2.5c)$$

$$T_N(\mathbf{u}) = \operatorname{tridiag}_k[u_{k-1} - u_k, u_{k-1} - 2u_k + u_{k+1}, u_{k+1} - u_k]. \quad (2.5d)$$

According to Remark 2.1, the matrix $X_N(\mathbf{u})$ is a symmetric positive definite having minimum eigenvalue $\lambda_{\min}(X_N(\mathbf{u})) \geq 1$.

In the two-dimensional case, the Jacobian of $F(\mathbf{u})$ has again a form similar to (2.5):

$$F'(\mathbf{u}) = I_{N^2} - \frac{\Delta t}{h^2}L_{D(\mathbf{u})} - \frac{1}{2}\frac{\Delta t}{h^2}Y(\mathbf{u}) \quad (2.6a)$$

where $L_{D(\mathbf{u})}$ is defined in (2.2) and

$$Y_{i,j}^{l,m}(\mathbf{u}) = \sum_{l,m}\frac{\partial\left[L_{D((\mathbf{u}))}\right]_{i,j}^{l,m}}{\partial u_{l,m}}u_{l,m}$$

with the double-index notation used in the previous section. A tedious but straight-forward computation yields

$$
\begin{aligned}
Y_{i,j}^{l,m}(\mathbf{u}) = {} & D'_{i,j}\delta_{i,l}\delta_{j,m}\left(-4u_{i,j}+u_{i+1,j}+u_{i-1,j}+u_{i,j+1}+u_{i,j-1}\right) \\
& + D'_{i+1,j}\delta_{l,i+1}\delta_{m,j}\left(u_{i+1,j}-u_{i,j}\right) + D'_{i-1,j}\delta_{l,i-1}\delta_{m,j}\left(u_{i-1,j}-u_{i,j}\right) \\
& + D'_{i,j+1}\delta_{l,i}\delta_{m,j+1}\left(u_{i,j+1}-u_{i,j}\right) + D'_{i,j-1}\delta_{l,i}\delta_{m,j-1}\left(u_{i,j-1}-u_{i,j}\right). \quad \text{(2.6b)}
\end{aligned}
$$

3. Convergence of the Newton method. We denote by $\|\cdot\|_2$ the Euclidean norm for vectors and the induced spectral norm for matrices. For a square matrix A we denote the real and the imaginary part of A with

$$
\mathrm{Re}(A) = \frac{A + A^\mathsf{T}}{2} \qquad \text{and} \qquad \mathrm{Im}(A) = \frac{A - A^\mathsf{T}}{2\mathrm{i}}, \qquad \mathrm{i}^2 = -1.
$$

Moreover, $\Sigma(A)$ will be the set of the eigenvalues of A and if A is also symmetric, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ will denote the minimum and the maximum eigenvalue of A, respectively.

In order to prove the convergence of the Newton method, we first consider some auxiliary results.

LEMMA 3.1 ([14]). *For a real $N \times N$ matrix A, the minimum singular value is*

$$
\sigma_N(A) \geq \lambda_{\min}\left(\mathrm{Re}(A)\right). \tag{3.1}
$$

In general it can be proved that for any complex-valued matrix A the minimal singular value is not less than the distance d_r of any straight line r separating the numerical range of A from the complex zero. Therefore a better estimate can be obtained by making the sup that we call d of d_r, for all straight lines that induce the separation. In our case we used the fact that $\mathrm{Re}(A)$ is positive definite and so our straight line becomes the set of all complex numbers having real part equal to $\lambda_{\min}(\mathrm{Re}(A))$. The estimate could be poor since the latter straight line is not necessarily tangent to the numerical range (a convex set by the Toeplitz-Hausdorff theorem, see [4]): thus d could be much larger than d_r. However in our setting such an estimate is already very satisfactory, as also stressed by the numerical experiments.

PROPOSITION 3.2. *Consider $F(\mathbf{u})$ as defined in (2.3) and let F' be assigned by equations (2.5) (one space dimension). Assuming that $\mathbf{u}$ is a sampling of a solution u of (1.1) with D differentiable and having Lipschitz continuous first derivative. If, in addition, u is differentiable with Lipschitz continuous first derivative, then*

$$
\left\|F'(\mathbf{u})^{-1}\right\|_2 \leq 1 + O(\Delta t). \tag{3.2}
$$

When using the induced l^∞ norm, we have a positive constant C_1 independent of h such that

$$
\left\|F'(\mathbf{u})^{-1}\right\|_\infty \leq C_1 \tag{3.3}
$$

for h sufficiently small and under the additional assumption that $\Delta t \leq C_\infty h$ for some $C_\infty > 0$.

Proof. We write

$$
\mathrm{Re}(F'(\mathbf{u})) = X_N(\mathbf{u}) + V_N(\mathbf{u}),
$$

where

$$V_N(\mathbf{u}) = -\frac{\Delta t}{4h^2}\, \mathsf{tridiag}_k \begin{bmatrix} (D'_{k-1} - D'_k)(u_{k-1} - u_k) \\ 2D'_k(u_{k-1} - 2u_k + u_{k+1}) \\ (D'_{k+1} - D'_k)(u_{k+1} - u_k) \end{bmatrix}. \tag{3.4}$$

By the regularity of D and u, we deduce that every entry of $V_N(\mathbf{u})$ is of order $\frac{\Delta t}{h^2}h^2$ that is $O(\Delta t)$ and hence $\|V_N(\mathbf{u})\|_2 = O(\Delta t)$. Thus, recalling that $\lambda_{\min}(X_N) \geq 1$, it holds

$$\lambda_{\min}\left(\mathrm{Re}(F'(\mathbf{u}))\right) \geq 1 - \widetilde{C}\Delta t \tag{3.5}$$

for some $\tilde{C} > 0$, that contains the infinity norms of the first derivatives of u and D and their Lipschitz constants. Using Lemma 3.1

$$\|F'(\mathbf{u})^{-1}\|_2 = \frac{1}{\sigma_{\min}(F'(\mathbf{u}))} \leq \lambda_{\min}\left(\mathrm{Re}(F'(\mathbf{u}))\right)^{-1}. \tag{3.6}$$

Inequality (3.2) now follows, combining (3.6) and (3.5).

For the proof of the estimate in l^∞ norm, we note that

$$F'(\mathbf{u}) = X_N(\mathbf{u}) - \frac{\Delta t}{2h}\mathsf{tridiag}_k \left[u'(\hat{\xi}_k), O(h), u'(\tilde{\xi}_k)\right] \mathsf{diag}_k \left[D'_k\right],$$

where $\hat{\xi}_k \in [x_{k-1}x_k]$, $\tilde{\xi}_k \in [x_k, x_{k+1}]$ and the constant in the $O(h)$ contains the Lipschitz constant of u'. We split $F'(\mathbf{u})$ as

$$F'(\mathbf{u}) = \frac{\Delta t}{h^2}\left(Z - W\right), \tag{3.7}$$

where

$$Z = \mathsf{diag}_k[z_k] = \mathsf{diag}_k\left[\frac{h^2}{\Delta t} + D_{k-1/2} + D_{k+1/2} + O(h^2)D'_k\right],$$

$$W = \mathsf{tridiag}_k\left[D_{k-1/2} + \frac{h}{2}u'(\hat{\xi}_k)D'_{k-1},\ 0,\ D_{k+1/2} + \frac{h}{2}u'(\tilde{\xi}_k)D'_{k+1}\right].$$

From (3.7), we have

$$[F'(\mathbf{u})]^{-1} = \frac{h^2}{\Delta t}\left(I - Z^{-1}W\right)^{-1}Z^{-1}. \tag{3.8}$$

For the factor Z^{-1}, recalling that $D(\cdot) \geq 0$, it holds

$$\|Z^{-1}\|_\infty = \max_k \frac{1}{|D_{k-1/2} + D_{k+1/2} + \frac{h^2}{\Delta t}(1 + O(\Delta t)D'_k)|} \leq c\frac{\Delta t}{h^2} \tag{3.9}$$

with c independent of h, for h sufficiently small, and assuming that $\Delta t \leq C_\infty h$. For the factor $(I - Z^{-1}W)^{-1}$, we note that

$$Z^{-1}W = \mathsf{tridiag}_k\left[\frac{D_{k-1/2} + \frac{h}{2}u'_k D'_k + O(h^2)}{z_k},\ 0,\ \frac{D_{k+1/2} + \frac{h}{2}u'_k D'_k + O(h^2)}{z_k}\right]$$

and hence

$$\|Z^{-1}W\|_\infty \le \max_k \frac{D_{k-1/2} + D_{k+1/2} + r_h}{\left| D_{k-1/2} + D_{k+1/2} + \frac{1}{C_\infty} h \left(1 + O(\Delta t)D_k'\right)\right|} \le \alpha < 1, \quad (3.10)$$

for $C_\infty > 0$ sufficiently small, since $r_h = O(h)$ with constant hidden in the big O not depending on C_∞. Thus the spectral radius of $Z^{-1}W$ is $\rho(Z^{-1}W) < 1$ and we have

$$\left(I - Z^{-1}W\right)^{-1} = \sum_{j=0}^\infty \left(Z^{-1}W\right)^j \quad \Rightarrow \quad \|\left(I - Z^{-1}W\right)^{-1}\|_\infty \le \frac{1}{1-\alpha}. \quad (3.11)$$

Finally, combining (3.11) and (3.9) with (3.8), estimate (3.3) holds with $C_1 = \frac{c}{1-\alpha}$. $\blacksquare$

PROPOSITION 3.3. *The estimates of Proposition 3.2 can be extended to the two-dimensional setting, where $F(\mathbf{u})$ is assigned by (2.3) and (2.2), while F' is given by equations (2.6).*

Proof. For the spectral norm, an analysis similar to the one made for $V_N(\mathbf{u})$ in the first part of the proof of Proposition 3.2, shows that $\|\mathrm{Re}(Y(\mathbf{u}))\|_2 = O(h^2)$. From equation (2.3), recalling that $\lambda_{\min}(I_{N^2} - \frac{\Delta t}{h^2} L_{D(\mathbf{u})}) \ge 1$ (see Remark 2.2), we infer that $\lambda_{\min}(\mathrm{Re}(F'(\mathbf{u}))) \ge 1 - \widetilde{C}\Delta t$ and hence $\|F'(\mathbf{u})\|_2 \le 1 + O(\Delta t)$.

For the proof of the estimate in l^∞ norm, we use again a splitting of the form (3.7) where Z is the diagonal part of $F'(\mathbf{u})$ up to the scaling factor $\Delta t/h^2$. Without reporting all details, we note that similarly to (3.9) it holds that $\|Z^{-1}\|_\infty \le c\Delta t/h^2$. In (3.10) the only term that is not of order h both in numerator and denominator is $D_{i+1/2,j} + D_{i-1/2,j} + D_{i,j+1/2} + D_{i,j-1/2}$ instead of $D_{k-1/2} + D_{k+1/2}$, but the same estimate is satisfied. Therefore inequality (3.11) can be obtained also here. $\blacksquare$

Both Propositions 3.2 and 3.3 can be proved with minor changes, under weaker assumptions. Indeed if both $u(\cdot)$ and $D(\cdot)$ are continuously differentiable, then every entry of $V_N(\mathbf{u})$ in (3.4) is of order

$$O\left(\max\left\{\frac{\Delta t}{h}\omega_{u'}(h), \frac{\Delta t}{h}\omega_{D'}(h)\|u'\|_\infty\right\}\right),$$

with $\omega_v(\cdot)$ denoting the modulus of continuity of a given function v. Therefore with the choice of Δt proportional to h and setting $\alpha(h) = \max\{\omega_{u'}(h), \omega_{D'}(h)\} = o(1)$, we find

$$\lambda_{\min}\left(\mathrm{Re}(F'(\mathbf{u}))\right) \ge 1 - \widetilde{C}\alpha(h)$$

and by Lemma 3.1 $\|F'(\mathbf{u})^{-1}\|_2 \le 1 + C\alpha(h)$. Furthermore, if we require that u is only Lipschitz continuous then the inequality regarding the norm of $F'(\mathbf{u})^{-1}$ reads as $\|F'(\mathbf{u})^{-1}\|_2 \le C$, where C linearly depends on the Lipschitz constant of u. Finally, the same results can be obtained with minor changes, when using the induced l^∞ norm.

In general, the solution u of (1.1) is not smooth, but only piecewise smooth with a finite number of cusps. For instance with $D(u) = u^m$ and continuous data with piecewise continuous derivative, the derivative of u is not defined in a finite number of points in one space dimension and in a finite number of smooth curves in two space dimensions; see [24]. The latter implies that the related matrices have the same features up to low rank correction terms whose cumulative rank is $O(N^{d-1})$ if the equation is in d dimensions.

Since the Crandall-Liggett formula does not induce any restriction on the timestep Δt [9], we have only to prove the convergence of the Newton method. We are interested in the choice $\Delta t = Ch$ for a constant C independent of h, which gives a method which is overall first order convergent. This is no restriction due to the presence of singularities at degenerate points: higher order methods would be computationally more intensive, without reaching their convergence rate, even if in practice a certain reduction of the error is expected.

The following result is a classical tool (see [16]) for handling the global convergence of the Newton procedure.

THEOREM 3.4 (Kantorovich). *Consider the Newton method for approximating the zero of a vector function $F(\mathbf{u})$, starting from the initial approximation $\mathbf{u}^{(0)}$. Under the assumptions that*

$$\| \left[F'(\mathbf{u}^{(0)}) \right]^{-1} \| \leq \beta \,, \tag{3.12a}$$

$$\| \left[F'(\mathbf{u}^{(0)}) \right]^{-1} F(\mathbf{u}^{(0)}) \| \leq \eta \,, \tag{3.12b}$$

$$\| F'(\mathbf{u}) - F'(\mathbf{v}) \| \leq \gamma \| \mathbf{u} - \mathbf{v} \| \,, \tag{3.12c}$$

and that

$$\beta \eta \gamma < \frac{1}{2} \,, \tag{3.13}$$

the method is convergent and, in addition, the stationary point of the iterations lies in the ball with centre $\mathbf{u}^{(0)}$ and radius

$$\frac{1 - \sqrt{1 - 2\beta\eta\gamma}}{\beta\gamma}.$$

For the choice $\Delta t = Ch$ we can prove the following result.

THEOREM 3.5. *The Newton method for $F(\mathbf{u})$ defined in (2.3) for computing $\mathbf{u}^n$ is convergent, both in the one-dimensional and in the two-dimensional setting, when initialized with the solution at the previous timestep (i.e. $\mathbf{u}^{n,0} = \mathbf{u}^{n-1}$) and for $\Delta t \leq Ch$, for a positive constant C independent of h.*

Proof. We will make use of the Kantorovich Theorem 3.4, so we need the estimates (3.12) and to show that the relation in (3.13) is satisfied. We will use the l^∞ vector norm and the induced matrix norm.

Concerning (3.12a), Proposition 3.2 in the one-dimensional case and Proposition 3.3 in the two-dimensional case imply that

$$\beta \leq C_1 \tag{3.14}$$

under the assumption $\Delta t \leq C_\infty h$.

Regarding (3.12b), we find

$$\left\| \left[F'(\mathbf{u}^{n-1}) \right]^{-1} F(\mathbf{u}^{n-1}) \right\|_\infty \leq \beta \left\| F(\mathbf{u}^{n-1}) \right\|_\infty = \beta \left\| \frac{\Delta t}{h^2} L_{D(\mathbf{u}^{n-1})} \mathbf{u}^{n-1} \right\|_\infty =$$
$$\beta \left\| \mathbf{u}^{n-2} - \mathbf{u}^{n-1} \right\|_\infty \leq \beta C_2 \Delta t \quad (3.15)$$

for a constant C_2 independent of h. The first equality in the previous calculation follows from (2.3), while the second one is a consequence of the fact that $\mathbf{u}^{n-1}$ is

the stationary point of the Newton iteration for the previous time step and thus it satisfies

$$\mathbf{u}^{n-1} + \frac{\Delta t}{h^2} L_{D(\mathbf{u}^{n-1})} \mathbf{u}^{n-1} = \mathbf{u}^{n-2}. \tag{3.16}$$

Therefore we obtain

$$\eta = C_2 \beta \Delta t. \tag{3.17}$$

For the Lipschitz constant of F', i.e., for estimating (3.12c), first we study in detail the one-dimensional case and then we give few hints concerning the two-dimensional case. Observe that in the one-dimensional case $F'(\mathbf{u}) - F'(\mathbf{v})$ is a tridiagonal matrix with two contributions:

$$F'(\mathbf{u}) - F'(\mathbf{v}) = \frac{\Delta t}{h^2}(L_{D(\mathbf{u})} - L_{D(\mathbf{v})}) + (Y_N(\mathbf{u}) - Y_N(\mathbf{v})), \tag{3.18}$$

with $L_{D(\cdot)}$ and $Y_N(\cdot)$ as in (2.1) and (2.5c), respectively. The first term can be estimated as follows:

$$\|L_{D(\mathbf{u})} - L_{D(\mathbf{v})}\|_\infty \le 4\|D'\|_\infty \|\mathbf{u} - \mathbf{v}\|_\infty. \tag{3.19}$$

In order to check that the last inequality is satisfied, one observes that the sum of the absolute values of the entries in each row of $L_{D(\mathbf{u})} - L_{D(\mathbf{v})}$ is smaller than the sum of 4 terms of the form

$$\left| D_{k\pm1/2}(\mathbf{u}) - D_{k\pm1/2}(\mathbf{v}) \right| = \left| D\left(\frac{u_{k\pm1} + u_k}{2}\right) - D\left(\frac{v_{k\pm1} + v_k}{2}\right) \right|$$

$$= |D'(\zeta)| \frac{|u_{k\pm1} + u_k - v_{k\pm1} - v_k|}{2} \le \|D'\|_\infty \|\mathbf{u} - \mathbf{v}\|_\infty.$$

For the second term in (3.18), we have

$$\|Y_N(\mathbf{u}) - Y_N(\mathbf{v})\|_\infty \le \frac{\Delta t}{2h^2}\|D'\|_\infty \|M\|_\infty, \tag{3.20}$$

where

$$M = \mathsf{tridiag}_k \begin{bmatrix} (u_{k-1} - u_k) - (v_{k-1} - v_k) \\ (u_{k-1} - 2u_k + u_{k+1}) - (v_{k-1} - 2v_k + v_{k+1}) \\ (u_{k+1} - u_k) - (v_{k+1} - v_k) \end{bmatrix}$$

and hence

$$\|M\|_\infty \le 8\|\mathbf{u} - \mathbf{v}\|_\infty. \tag{3.21}$$

Replacing equation (3.21) in (3.20) and combining (3.20) and (3.19) with (3.18), we obtain

$$\gamma \le 8\|D'\|_\infty \frac{\Delta t}{h^2}. \tag{3.22}$$

For the two-dimensional case, $L_{D(\cdot)}$ is defined in (2.2) and

$$\|L_{D(\mathbf{u})} - L_{D(\mathbf{v})}\|_\infty \le 8\|D'\|_\infty \|\mathbf{u} - \mathbf{v}\|_\infty.$$

Moreover, the matrix M is replaced by a pentadiagonal matrix $\widetilde{M}$ such that $\|\widetilde{M}\|_\infty \leq 16\|\mathbf{u} - \mathbf{v}\|_\infty$. As a conclusion, in the two-dimensional case when estimating γ, we have a further factor two, i.e., relation (3.22) is replaced by

$$\gamma \leq 8\,d\,\|D'\|_\infty \frac{\Delta t}{h^2}, \tag{3.23}$$

where $d = 1$ in the one-dimensional case and $d = 2$ in two dimensions.

Finally, putting together equations (3.14), (3.17), and (3.23), Theorem 3.4 implies that the Newton method converges, provided that

$$\frac{1}{2} \geq C_1^2 C_2 8\,d\,\|D'\|_\infty \frac{(\Delta t)^2}{h^2} \geq \beta\eta\gamma,$$

i.e., $\Delta t \leq Ch$, for h sufficiently small and for $C = \min\{C_\infty, 1/(4C_1\sqrt{C_2\,d\,\|D'\|_\infty}\,)\}$ (essentially) independent on h. $\square$

REMARK 3.6. *If one considers the stopping criterion for the Newton iterations, equation (3.16) has an extra ε on the right-hand-side and consequently (3.15) changes into*

$$\left\| \left[F'(\mathbf{u}^{n-1})\right]^{-1} F(\mathbf{u}^{n-1}) \right\|_2 \leq \beta \left\| \mathbf{u}^{n-2} - \mathbf{u}^{n-1} + \varepsilon \right\|_2 \leq \beta\,(C_2\Delta t + \varepsilon)$$

Thus if one chooses $\varepsilon = c \cdot \Delta t$ where c is a moderately small constant independent of Δt, estimate (3.17) is satisfied, only with a slightly higher constant and the above result is still true. Note also that more precision will be useless in practice and would make the Newton process more expensive, by increasing the iteration count.

4. Algorithms for the resulting linear systems. At each Newton iteration, we need to solve a linear system whose coefficient matrix is represented by the Jacobian $F'(\mathbf{u})$ defined by equations (2.5) or (2.6) in one and two space dimensions, respectively. In principle, the Jacobian is recomputed at each Newton iteration, so we are interested in efficient iterative methods for solving the related linear system. In order to choose appropriate iterative methods for solving the Jacobian linear system, we first analyse the spectral properties of the matrix $F'(\mathbf{u})$. This will lead us to consider preconditioned Krylov methods, multigrid methods and their combination.

4.1. Spectral analysis for the resulting matrix-sequences and preconditioning. In this subsection we give a detailed spectral analysis of the matrices involved in the definition of the Jacobian. This study will lead us to define an optimal preconditioner and to motivate the use of multigrid preconditioners.

Like in the previous section, we give all the details in the one-dimensional case and we extend the results to the two-dimensional case at the end. First, we study the term $Y_N(\mathbf{u})$ in the definition of $F'(\mathbf{u})$ in equation (2.5a) showing that it is negligible with respect to $X_N(\mathbf{u})$.

REMARK 4.1. *If $\mathbf{u}$ is a sampling of a solution $u(\cdot)$ of (1.1), if we assume that $u(\cdot)$ is at least continuous, and if $\omega_u(\cdot)$ denotes its modulus of continuity, then*

$$\|Y_N(\mathbf{u})\|_2 \leq e_u(\Delta t, h)$$

with

$$e_u(\Delta t, h) = 4\frac{\Delta t}{h^2}\|D'(u)\|_\infty \omega_u(h).$$

If we are interested in evaluating $\|Y_N(\tilde{\mathbf{u}})\|_2$ where $\tilde{u}$ is an approximation to the true solution u (this happens naturally in the numerical process discussed in the present section), then

$$\|Y_N(\tilde{\mathbf{u}})\|_2 \leq e_{\tilde{u}}(\Delta t, h) \leq e_u(\Delta t, h) + 4\frac{\Delta t}{h^2}\|D'(u)\|_\infty \|u - \tilde{u}\|_\infty$$

$$= 4\frac{\Delta t}{h^2}\|D'(u)\|_\infty \left(\omega_u(h) + \|u - \tilde{u}\|_\infty\right).$$

Hence, since we are using second order formulae, the error $\|u - \tilde{u}\|_\infty = O(h^2)$ and therefore $\|Y_N(\tilde{\mathbf{u}})\|_2$ is dominated by $\omega_u(h)$, which is of order h if the solution is Lipschitz continuous, that is

$$\|Y_N(\tilde{\mathbf{u}})\|_2 \leq 4M\frac{\Delta t}{h}\|D'(u)\|_\infty + O(\Delta t).$$

In conclusion, we can safely claim that the global spectrum of the Jacobian $F'(\tilde{\mathbf{u}})$ is decided, up to small perturbations, by the matrix $X_N(\tilde{\mathbf{u}})$. For making more explicit the latter statement, if we assume that $\Delta t = Ch$, where $C > 0$ is independent of h, then $\lambda_{min}(X_N(\tilde{\mathbf{u}})) \geq 1$, $\|X_N(\tilde{\mathbf{u}})\|_2 = O(h^{-1})$ while $\|Y_N(\tilde{\mathbf{u}})\|_2 = O(1)$. A similar estimate holds in the two-dimensional case as well, working in the same manner as in the first part of the proof reported in Proposition 3.3.

DEFINITION 4.2. *Let $\mathcal{C}_0(\mathbb{R}_0^+)$ be the set of continuous functions with bounded support defined over the nonnegative real numbers, d a positive integer, and θ a complex-valued measurable function defined on a set $G \subset \mathbb{R}^d$ of finite and positive Lebesgue measure $\mu(G)$. Here G will be often equal to $(0, 2\pi)^d$ so that $e^{iG} = \mathbb{T}^d$ with $i^2 = -1$ and $\mathbb{T}$ denoting the complex unit circle. A matrix sequence $\{A_N\}$ is said to be distributed (in the sense of the eigenvalues) as the pair (θ, G), or to have the eigenvalue distribution function θ ($\{A_N\} \sim_\lambda (\theta, G)$), if, $\forall F \in \mathcal{C}_0(\mathbb{C})$, the following limit relation holds*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} F(\lambda_j(A_N)) = \frac{1}{\mu(G)} \int_G F(\theta(t))\, dt, \qquad t = (t_1, \ldots, t_d). \qquad (4.1)$$

Along with the distribution in the sense of eigenvalues (weak*-convergence), for the practical convergence analysis of iterative solvers we are also interested in a further asymptotic property called here *clustering*.

DEFINITION 4.3. *A matrix sequence $\{A_N\}$ is strongly clustered at $s \in \mathbb{C}$ (in the eigenvalue sense), if for any $\varepsilon > 0$ the number of the eigenvalues of A_N off the disk*

$$D(s, \varepsilon) := \{z : |z - s| < \varepsilon\}$$

can be bounded by a pure constant q_ε possibly depending on ε, but not on n. In other words

$$q_\varepsilon(n, s) := \#\{\lambda_j(A_N) : \lambda_j \notin D(s, \varepsilon)\} = O(1), \quad n \to \infty.$$

If every A_N has only real eigenvalues (at least for all n large enough), then s is real and the disk $D(s, \varepsilon)$ reduces to the interval $(s - \varepsilon, s + \varepsilon)$. Finally, the term "strongly" is replaced by "weakly", if

$$q_\varepsilon(n, s) = o(n), \qquad n \to \infty.$$

REMARK 4.4. *It is clear that $\{A_N\} \sim_\lambda (\theta, G)$ with $\theta \equiv s$ a constant function is equivalent to $\{A_N\}$ being weakly clustered at $s \in \mathbb{C}$.*

Now we briefly use the above concepts in our specific setting. Given the linear restriction on Δt imposed by the convergence of the Newton method (Theorem 3.5), we are interested in the choice $\Delta t = Ch$ for $C > 0$ independent of h. However, for notational simplicity, here we assume $\Delta t = h$ and note that analogous results can be stated and obtained for a generic positive constant C. Moreover, we scale the Jacobian as $A_N = hF'(\mathbf{u})$ and we thus consider the sequence $\{A_N\}$. More in detail A_N is

$$A_N = -L_{D(\mathbf{u})} + R_N(\mathbf{u}) \tag{4.2}$$

where in the one-dimensional case

$$R_N(\mathbf{u}) = hI_N - \frac{1}{2}T_N(\mathbf{u})\operatorname{diag}_k(D'_k) \tag{4.3}$$

with $T_N(\mathbf{u})$ defined as in (2.5d), while in the two-dimensional case

$$R_N(\mathbf{u}) = hI_{N^2} - \frac{1}{2}Y(\mathbf{u}) \tag{4.4}$$

with $Y(\mathbf{u})$ defined as in (2.6b).

We have the following results, which are of crucial interest in the choice, design, and analysis of efficient solvers for the involved linear systems. Since we consider the case where the Newton method is convergent, in the following results we implicitly assume that $\mathbf{u}$ and D have the regularity required in Proposition 3.2.

PROPOSITION 4.5. *Given the sequence $\{A_N\}$ with A_N defined in (4.2), we obtain that*

$$\{A_N\} \sim_\lambda (\theta, G),$$

where in the one-dimensional case

$$\theta(x, s) = D(u(x))(2 - 2\cos(s)) \quad and \quad G = [a, b] \times [0, 2\pi],$$

while in the two-dimensional case

$$\theta(x, y, s, t) = D(u(x, y))(4 - 2\cos(s) - 2\cos(t)) \; and \; G = [a_0, b_0] \times [a_1, b_1] \times [0, 2\pi]^2.$$

Proof. If we assume that $\mathbf{u}$ is a sampling of a given function over a uniform grid, then $\{L_{D(\mathbf{u})}\} \sim_\lambda (-\theta, G)$ (see [22] for one-dimensional problems and [19] for two-dimensional problems). In our case the entries of $\mathbf{u}$ represent an approximation in infinity norm of the true solution, the latter being implied by the convergence of the method, and therefore by standard perturbation arguments we deduce $\{L_{D(\mathbf{u})}\} \sim_\lambda (-\theta, G)$. The remaining term $R_N(\mathbf{u})$ has trace norm (sum of all singular values i.e. Schatten p norm with $p = 1$; see [4]) bounded by a pure constant C independent of N, when assuming that D' is bounded and u is at least Lipschitz continuous. The latter implies that the distribution of $\{A_N\}$ is decided only by that of $\{-L_{D(\mathbf{u})}\}$; see Theorem 3.4 in [11]. □

The Proposition 4.5 gives the distribution of the zero order main term. Concerning the negligible term, we have $\{R_N(\mathbf{u})\} \sim_\lambda (0, G)$ and $\{R_N(\mathbf{u})/h\} \sim_\lambda (\psi, G)$

with $\psi(x, s) = 1 - D'(u(x))2\mathrm{i}\sin(s))$ in the one-dimensional case and $\psi(x, y, s, t) = 1 - D'(u(x, y))2\mathrm{i}(\sin(s) + \sin(t)))$ (distribution of the first order term).

PROPOSITION 4.6. *Let A_N be defined in (4.2), then*

$$\kappa_2(A_N) \,=\, \|A_N\|_2 \|A_N^{-1}\|_2 \,=\, O(N). \tag{4.5}$$

Moreover, for the spectrum of A_N it holds

$$\Sigma(A_N) \,\subset\, [\,ch,\, C\,] \,\times\, \mathrm{i}\,[\,-dh,\, dh\,], \tag{4.6}$$

for some positive constants c, d, and C independent of N.

Proof. From Propositions 3.2 and 3.3 we deduce $\|A_N^{-1}\|_2 = O(h^{-1})$. Moreover $\|A_N\|_2 = O(1)$ since $\|L_{D(\mathbf{u})}\|_2 = O(1)$, while $\|R_N(\mathbf{u})\|_2 = O(h)$ according to Remark 4.1. Therefore, combining the previous results, it follows that $\kappa_2(A_N) = O(N)$ and $\Sigma(\mathrm{Re}(A_N)) \subset [\,ch,\, C\,]$. The estimate (4.6) follows from Bendixson Theorem (see Theorem 3.6.1 of [21]) proving that $\Sigma(\mathrm{Im}(A_N)) \subset \mathrm{i}\,[\,-dh,\, dh\,]$. We note that in the one-dimensional case

$$\mathrm{Im}(A_N) = -\frac{1}{2}\,\mathrm{Im}\left(\widetilde{T}_N(\mathbf{u})\;\mathsf{diag}_k[D_k']\right),$$

where $\widetilde{T}_N(\mathbf{u}) = \mathsf{tridiag}_k\,[u_{k-1} - u_k\,,\; 0\,,\; u_{k+1} - u_k]$, while in the two-dimensional case $\widetilde{T}_N(\mathbf{u})$ is replaced by a matrix that has only 4 non zero entries of order h in each row. Therefore, $\mathrm{Im}(A_N) \subset \mathrm{i}\,[\,-dh,\, dh\,]$, where d depends on $\|D'(u)\|_\infty$ and on the Lipschitz constant of the solution. $\square$

From Proposition 4.5, A_N and $-L_{D(u)}$ have the same eigenvalue distribution, but they differ in the first order term in h. From Proposition 4.6, $\kappa_2(A_N) = O(N)$ while $\kappa_2(-L_{D(u)}) = O(N^2)$. This happens because a perturbation of the spectrum of $O(h)$ does not affect the eigenvalue distribution as stated in Definition 4.2. Therefore, for instance a shift of the spectrum of $O(h)$ does not change the eigenvalue distribution, but it can change the spectral condition number from $o(N)$ to $O(N)$ (assuming that the largest eigenvalue is bounded).

REMARK 4.7. *Since $F'(\mathbf{u}) = A_N/h$, the previous proposition can be rewritten as*

$$\kappa_2(F'(\mathbf{u})) \,=\, O(N) \qquad and \qquad \Sigma(F'(\mathbf{u})) \,\subset\, [\,c,\, CN\,] \,\times\, \mathrm{i}\,[\,-d,\, d\,].$$

PROPOSITION 4.8. *Let A_N be defined in (4.2), then setting*

$$P_N = -L_{D(\mathbf{u})} + hI,$$

where $I = I_N$ in one space dimension and $I = I_{N^2}$ in two dimensions, it holds

$$\Sigma(P_N^{-1}A_N) \,\subset\, [\,1 - c_1 h,\, 1 + c_2 h\,] \,\times\, \mathrm{i}\,[\,-d,\, d\,].$$

for some positive constants c_1, c_2, and d independent of N.

Proof.

Note that $\Sigma(P_N^{-1}A_N)$ belongs to the field of value of $P_N^{-1/2}A_N P_N^{-1/2}$. Fix $\alpha = \mathbf{x}^H P_N^{-1/2}A_N P_N^{-1/2}\mathbf{x}$, for all $\mathbf{x} \in \mathbb{C}^n$, $\|\mathbf{x}\|_2 = 1$. In order to apply the Bendixson Theorem, note that $\mathrm{Re}(\alpha) = \mathbf{x}^H P_N^{-1/2}\mathrm{Re}(A_N)P_N^{-1/2}\mathbf{x}$ which belongs to $[1 - c_1 h,\, 1 + c_2 h]$ by the analysis provided in Propositions 3.2 and 3.3. A similar analysis stands for $\mathrm{Im}(\alpha) = \mathbf{x}^H P_N^{-1/2}\mathrm{Im}(A_N)P_N^{-1/2}\mathbf{x}$ as in Proposition 4.6. In fact

$$|\mathrm{Im}(\alpha)| \leq \|P_N^{-1/2}\mathrm{Im}(A_N)P_N^{-1/2}\|_2 = \rho(P_N^{-1}\mathrm{Im}(A_N)) \leq \|P_N^{-1}\|_2\|\mathrm{Im}(A_N)\|_2$$

and the thesis follows from $\|P_N^{-1}\|_2 = 1/\lambda_{\min}(P_N) \leq 1/h$, since $-L_{D(\mathbf{u})}$ is positive semidefinite, and $\|\text{Im}(A_N)\|_2 \leq dh$ due to Proposition 4.6. $\square$

From Proposition 4.8 the preconditioned sequence $\{P_N^{-1}A_N\}$ has eigenvalues with real part strongly clustered at 1. Furthermore, a weak eigenvalue value clustering

$$\{P_N^{-1}A_N\} \sim_\lambda (1, G), \tag{4.7}$$

follows from the property of algebra of the Generalised Locally Toeplitz (GLT) sequences; see [19]. Therefore, the eigenvalues of $\{P_N^{-1}A_N\}$ have imaginary part weakly clustered at 1.

REMARK 4.9. *Proposition 4.8 holds unchanged replacing P_N and A_N with $X_N(\mathbf{u})$ and $F'(\mathbf{u})$, respectively, because $X_N(\mathbf{u})^{-1}F'(\mathbf{u}) = P_N^{-1}A_N$. In practice, Proposition 4.8 implies that $X_N(\mathbf{u})$ is an optimal preconditioner for $F'(\mathbf{u})$, i.e., a preconditioned Krylov method applied to $F'(\mathbf{u})$ with $X_N(\mathbf{u})$ as preconditioner converges in a number of iterations bounded by a constant depending on the precision, but not on the mesh that is on h (optimality of the method).*

4.2. Iterative methods for the linear system. In this section we consider some iterative methods for solving the linear system at each Newton step and study their convergence properties on the matrix sequence $\{A_N\}$. The same analysis holds for $\{F'(\mathbf{u})\}$ since the scaling factor h does not affect the convergence behaviour. A classical reference for the results quoted below is [17].

4.2.1. GMRES. We first consider the GMRES algorithm, since the antisymmetric part of A_N is negligible but not zero.

Assume that A_N is diagonalisable and let $A_N = W\Lambda W^{-1}$, where $\Lambda = \text{diag}_k(\lambda_k)$ is the diagonal matrix of the eigenvalues. Define

$$\epsilon^{(m)} = \min_{p \in \mathbb{P}_m : p(0)=1} \max_{k=1,\ldots,N} |p(\lambda_k)|.$$

Denoting with $r^{(m)}$ the residual at the mth step of GMRES, it is a classical result that

$$\|r^{(m)}\|_2 \leq \kappa_2(W)\epsilon^{(m)}\|r^{(0)}\|_2.$$

Thanks to Proposition 4.6, $\kappa_2(W) \approx 1$ since $\text{Im}(A_N) = O(h)$. Thus the GMRES convergence is determined by the factor $\epsilon^{(m)}$.

Thanks to Proposition 4.6, it is possible to construct an ellipse properly containing the spectrum of A_N and avoiding the complex 0, so that when one applies GMRES to the matrix A_N, it holds that

$$\epsilon^{(m)} \leq \left(1 - C\sqrt{h}\right)^m \tag{4.8}$$

for a positive constant C that is independent of the problem size N. This means that the number of iterations required to reach a preassigned tolerance could grow as $\sqrt{N}$.

Similarly, using P_N as preconditioner, Proposition 4.8 implies that

$$\epsilon^{(m)} \leq \widetilde{C}^m \tag{4.9}$$

for some $\widetilde{C} \in (0,1)$, independent of the problem size N. Even if the solution u is not enough regular to assure that the spectrum of $P_N^{-1}A_N$ belongs to $[1 - c_1h, 1 + c_2h] \times$ i$[-d, d]$, the strong cluster at 1 leads in practice to the superlinear convergence.

4.2.2. Conjugate gradient (CG). Let S_N be a symmetric definite matrix and define $\|\mathbf{x}\|_{S_N} = \|(S_N)^{1/2}\mathbf{x}\|_2$. We recall the following classical result about the convergence of the CG applied to a linear system with coefficient matrix S_N:

$$\|\mathbf{x}_m - \mathbf{x}_*\|_{S_N} \leq 2 \left(\frac{\sqrt{\kappa_2(S_N)} - 1}{\sqrt{\kappa_2(S_N)} + 1} \right)^m \|\mathbf{x}_0 - \mathbf{x}_*\|_{S_N}, \tag{4.10}$$

where $\mathbf{x}_m$ is the approximate solution obtained at the mth step of the CG algorithm and $\mathbf{x}_*$ the exact solution.

Combining (4.10) with Proposition 4.6, we expect the CG algorithm to converge in $O(\sqrt{N})$ iterations when applied to $\text{Re}(A_N)$. On the other hand, using P_N as preconditioner, similarly to Proposition 4.8 implies that the eigenvalues of $P_N^{-1}\text{Re}(A_N)$ are localised in the interval $[1 - c_1 h, 1 + c_2 h]$ and thus that CG converges in a constant number of iterations, independently on the size N of the problem.

Finally, according to Proposition 4.6 and 4.5, $\text{Im}(A_N)$ is negligible. Thus in practice one may apply the CG algorithm to the matrix A_N, expecting a convergence behaviour similar to that for $\text{Re}(A_N)$, in both the unpreconditioned and preconditioned cases.

4.2.3. Multigrid method (MGM). Proposition 4.8 and equation (4.7) implies that A_N has the same spectral behaviour as P_N. Hence, if an iterative method is effective for P_N and robust, it should be effective also for A_N. This is the case of MGM largely used when dealing with elliptic PDEs ([23]). The matrix P_N can be viewed, after re-scaling, as a regularized weighted Laplacian since in the coefficient matrix one adds h times the identity to $-L_{D(\mathbf{u})}$. In this way the conditioning is not growing as N^2 as in the standard Laplacian but grows only linearly with N. Therefore the standard V-cycle that is already optimal for $-L_{D(\mathbf{u})}$, i.e., the number of iterations is independent of the system size, continues to be optimal also for P_N.

MGM has essentially two degrees of indetermination: the choice of the grid transfer operators and the choice of the smoother (pre- and post-smoother, if necessary). In particular, let P_{i+1}^i be the prolongation operator from a coarse grid $i+1$ to a finer grid i. We consider a Galerkin strategy: the restriction operator is $(P_{i+1}^i)^{\mathsf{T}}$ and the coefficient matrix of the coarse problem is $A_{i+1} = (P_{i+1}^i)^{\mathsf{T}} A_i P_{i+1}^i$, where A_i is the coefficient matrix on the ith grid.

For the prolongation we consider the classical linear interpolation in one space dimension and the bilinear interpolation in two dimensions that are enough to obtain an optimal convergence for $-L_{D(\mathbf{u})}$. Therefore, since A_N has the same spectral distribution (Proposition 4.5) and is better conditioned (Proposition 4.6), it is not necessary to resort to more sophisticated grid transfer operators. The restriction is the full-weight since, according to the Galerkin approach, it is the transpose of the linear (bilinear) interpolation. Regarding the choice of the smoother, damped Jacobi and red-black Gauss-Seidel are considered.

Finally, we stress that a robust and effective strategy is to use a multigrid iteration as preconditioner for Krylov methods as confirmed in the numerical experiments. In fact we showed that P_N is an optimal preconditioner for A_N and the MGM is an optimal solver for a linear system with coefficient matrix P_N. However, the numerical experimentation in Section 5 confirms that the MGM can be efficiently applied also as a stand-alone solver.

5. Numerical tests. First, we consider in detail a one-dimensional problem. The two-dimensional setting is analysed at the end of the section.
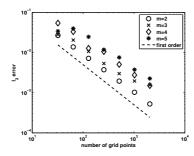
FIG. 5.1. l_2 error at final time for $N = 32, 64, \ldots, 2048$, final time $t = 20/32$, $\Delta t = h$.

The one-dimensional problem is a test case, the porous medium equation written in the form

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(mu^{m-1}\frac{\partial u}{\partial x}\right) \tag{5.1}$$

with homogeneous Dirichlet boundary conditions. Here $m \geq 1$, with $m = 1$ corresponding to the heat equation. In particular we consider the exact self-similar solution

$$u(t, x) = t^{-\alpha}\left[1 - \alpha\frac{m-1}{2m}\left(|x|(t+1)^{-\alpha}\right)^2\right]_+^{\frac{1}{m-1}}, \qquad \alpha = \frac{1}{m+1} \tag{5.2}$$

due to Barenblatt and Pattle [24]. (The subscript + denotes the positive part). The experiments are carried out in Matlab 7.0.

In this section we do not use the rescaling by h that was useful to lighten the notation of the previous section. Obviously, the analysis in Section 4 is not affected by any positive rescaling.

5.1. Convergence of the global method and of Newton's method. First we check the convergence of the method. We perform test for m ranging from 2 to 5, observing no appreciable difference in the convergence properties of the algorithm. In all tests we choose $\Delta t = h$.

Figure 5.1 plots the l_2 errors between the numerical solution at time $t = 20/32$ and the exact solution (5.2) and shows that the method is first order convergent, as expected for this choice of time stepping procedure and also due to the presence of the singularity in the first derivative of the exact solution. The dashed line is a reference slope for first order schemes. We observe that the convergence is not significantly affected by the parameter m.

Figure 5.2 plots the number of Newton iterations employed by the algorithm during the integration from $t = 0$ to $t = 20/32$. We plot the average (circles), minimum and maximum (solid lines) number of Newton iterations per timestep. Taking $\Delta t = h$ (Figure 5.2a), we observe that the number of Newton iterations slowly decreases when N increases and that, for any given N it increases only very moderately when m increases. In the case $m = 2$ we also tried to vary the step size from $\Delta t = h/10$ to $\Delta t = 5h$. The results are reported in Figure 5.2b, showing that the number of Newton iterations grows when taking larger Δt in (1.2). The larger variability (for fixed N) and the irregular behaviour of the mean value when increasing N in the case $\Delta t = 5h$ preludes to the loss of convergence that we observe if Δt is taken even larger.
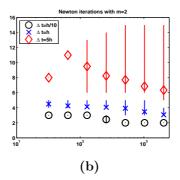
(a)　　　　　　　　　　　　(b)

FIG. 5.2. *Average, minimum and maximum number of Newton iterations performed during the integration until final time. The horizontal axis represents the number of points in the discretization. In (a) Δt was kept fixed and m varied, in (b) $m = 2$ was kept fixed and Δt varied.*
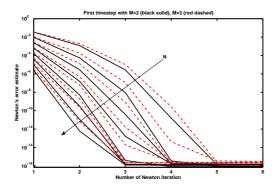

FIG. 5.3. *History of the convergence of the Newton iterations during the first timestep. Black solid lines correspond to $m = 2$ and red dashed ones to $m = 3$. We show the results for N ranging from 32 to 4096, with $\Delta t = h$: the behaviour under grid refinement is indicated by the thin arrow.*

Next we verify the convergence of the Newton's method. In Figure 5.3 we plot the Newton's error estimate $\|\mathbf{u}^{1,k+1} - \mathbf{u}^{1,k}\|_2 / \|\mathbf{u}^{1,k}\|_2$ obtained when computing the first timestep $\mathbf{u}^1$. We compare different number of grid points ($N = 32, 64, \ldots, 4096$) as indicated by the thin arrow and two values for the exponent m appearing in (5.1).

We emphasise that as prescribed in Proposition 3.5 the choice of $\Delta t = h$ is acceptable for the convergence both of the global numerical scheme and for the convergence of the Newton procedure.

5.2. Solution of the linear system. This section is devoted to computational proposals for the solution of a linear system where the coefficient matrix is the Jacobian in (2.5a), which is required at every step of the Newton procedure. For all the tests, we set $m = 2$, final time $t = 20/32$, $\Delta t = h$, and we let N be equal to $32, 64, \ldots, 1024$ for checking the optimality of the proposed best solvers.

As already stressed in Proposition 4.6, the matrix is (weakly) non-symmetric so we start by considering the use of preconditioned GMRES (PGMRES).

5.2.1. GMRES. In Figure 5.4a we plot the average (circles), minimum and maximum (vertical lines) number of GMRES iterations performed during the integration until final time, at different spatial resolutions. A least square fit (dashed
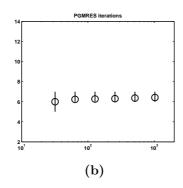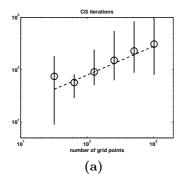
FIG. 5.4. *Average, minimum and maximum number of GMRES iterations (a) and precondi-tioned GMRES iterations (b) performed during the integration until final time. The dashed line in panel (a) is the least square fit.*
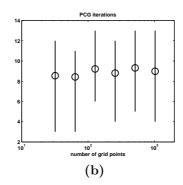

FIG. 5.5. *Average, minimum and maximum number of CG iterations (a) and PCG iterations (b), performed during the integration until final time. The dashed line in panel (a) is the least square fit.*
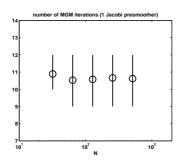
line) shows that the number of iterations grows as $N^{0.5320}$. This fact is in complete accordance with the analysis of Subsection 4.2 and in particular with equation (4.8).

In Proposition 4.8 and Remark 4.9 the use of $X_N(\mathbf{u})$ as preconditioner for $F'(\mathbf{u})$ was analysed and a strong spectral clustering of the preconditioned matrix at 1 was shown. As a consequence we expect a number of iterations not depending on the size N of the matrix as in (4.9): this fact is observed in practice and indeed the iteration count of the PGMRES is almost constant, with average value equal to 6 iterations (see Figure 5.4b).

The previous tests were conducted with a direct solver for the preconditioning step, in order to confirm the theoretical results of Section 4. According to the discussion in Section 4.2.3, the linear system with $X_N(\mathbf{u})$ can be solve by multigrid methods without change the GMRES iteration count (not shown). Later on we will test preconditioning with a single V-cycle for the whole matrix $F'(\mathbf{u})$ (instead of $X_N(\mathbf{u})$).

5.2.2. CG. Since the non-symmetric part of $F'(\mathbf{u})$ is negligible, we can try directly the solution of the whole system by using techniques such as the preconditioned CG (PCG) or the multigrid method which in theory should suffer from the loss of symmetry in the linear system.

In Figure 5.5a we plot the average (circles), minimum and maximum (vertical lines) number of CG iterations performed during the integration until final time, at

FIG. 5.6. *Average, minimum and maximum number of MGM iterations performed during the integration until final time.*

different spatial resolutions. A least square fit shows that the number of iterations grows as $N^{0.5491}$, which agrees with the discussion in Subsection 4.2. However, the number of iterations for fine grids is higher that the ones with GMRES (up to 950 instead of 160 with a grid of 1024 points). On the other hand, PCG does not need to memorize any further vector at each iteration.

In a similar way, we consider $X_N(\mathbf{u})$ as preconditioner in the PCG method. Results are shown in Figure 5.5b. The number of iteration is again essentially constant with respect to N, but it requires a slightly higher number of iterations to converge with respect to PGMRES. Furthermore we have a higher variance in the number of iterations, due to the weak non-symmetry of whole matrix.
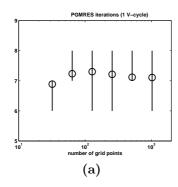
5.2.3. MGM. We test the optimality of MGM, as discussed in Section 4.2. We apply a single recursive call, that is the classical V-cycle procedure. As smoother, we use a single Jacobi step with damping factor equal to 2/3. We observe mesh independent behaviour with 10 or 11 iterations (see Figure 5.6).

We also tried other more sophisticated approaches by adding one step of post-smoother with Gauss-Seidel, CG or GMRES but without any substantial reduction of the number of iterations.

5.2.4. Krylov methods with MGM as preconditioner. The previous experiments confirm that the MGM is an excellent solver for our linear system. Often this method is also applied as preconditioner in a Krylov method instead of employing it as a solver. In other words, as preconditioning step, we perform a single V-cycle iteration, with the coefficient matrix $F'(\mathbf{u})$ and where the datum is the residual vector at the current iteration.

With the use of such very cheap MGM preconditioning, the PGMRES converges within about 7 iterations, independently of the size of the involved matrices (see Figure 5.7a). Comparing with the GMRES method preconditioned with $X_N(\mathbf{u})$ considered in 5.2.1 and Figure 5.4b, the present preconditioning strategy computationally cheaper since the solution of the preconditioned system is replaced by a V-cycle iteration without increasing the number of GMRES iterations. This must be partly ascribed to the fact that the preconditioner now uses the full matrix $F'(\mathbf{u}) = X_N(\mathbf{u}) + Y_N(\mathbf{u})$ and not only $X_N(\mathbf{u})$.

The application of the V-cycle preconditioned to PCG requires a further post-smoothing step (again damped Jacobi) to obtain a robust enough convergence. With one step of pre- and post-smoother the PCG method leads to a convergence within 7 or 8 iterations, again independently of the system sizes (see Figure 5.7b).
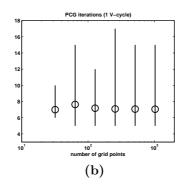
(a)   (b)

FIG. 5.7. *Average, minimum and maximum number of PGMRES (a) and PCG (b) iterations performed during the integration until final time, when using one MGM V-cycle as preconditioner.*

N	MGM(1,0)	MGM(1,1)	PGMRES(1,0)	PGMRES(1,1)	PCG(1,1)
32	0.07	0.05	0.11	0.09	0.08
64	0.17	0.11	0.25	0.20	0.22
128	0.45	0.31	0.60	0.49	0.53
256	1.23	0.84	1.50	1.27	1.34
512	3.46	2.39	3.95	3.40	3.56
1024	8.95	6.35	10.20	8.80	9.22

TABLE 5.1

Average CPU time for solving the inner linear system for the one-dimensional example. The preconditioner is one multigrid iteration where between bracket are denoted the steps of pre- and post-smoother.

In conclusion, V-cycle preconditioning in connection with GMRES has to be preferred, in terms of the robustness (less variance in the iteration count) and the number of iterations. Indeed, due to the small iteration count, also the memory requirement does not pose any difficulty, since the number of vectors that have to be stored in the GMRES process is very reasonable. However, since the computational cost for iteration is different for each solver, we compare also the CPU time. Table 5.1 shows the CPU time for the V-cycle method used as stand-alone solver and as preconditioner for GMRES and CG. All the different strategies give comparable results and the minimum CPU time is obtained using the multigrid as stand-alone solver with one iteration of pre- and post-smoother.

5.3. Two space dimensions. We performed our tests with the two-dimensional Barenblatt solution [24] with exponent $m = 4$ on grids of size $N \times N$ for N ranging from 32 to 1024. First of all we note that the number of Newton iterations required at each timestep is almost independent of N and is (on average) 4 when $\Delta t = 0.5h$, 4.5 when $\Delta t = h$ and 6.5 when $\Delta t = 2h$ (see Figure 5.8).

Here, for brevity, we discuss in detail only the PGMRES with the multigrid preconditioner since similar considerations hold for PCG like in the 1D case.

For the multigrid preconditioner, we employ a single V-cycle iteration, with a Galerkin approach using the bilinear interpolation as prolongation operator and one step of red-black Gauss-Seidel as pre-smoother. In Figure 5.9 we plot the mean (symbols) and minimum-maximum (solid lines) number of GMRES iterations needed at different spatial resolutions. Different colors correspond to different choices of Δt, namely $\Delta t = h/2$ (blue crosses), $\Delta t = h$ (black circles) and $\Delta t = 2h$ (red diamonds). The left panel shows that, without preconditioning, the number of GMRES iterations
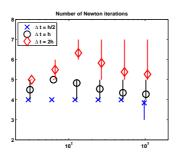
FIG. 5.8. *Average, minimum and maximum number of Newton iterations performed during the integration until final time. The labels on the horizontal axis refer to the number N of grid points per direction. The 3 data series for each N have been slightly shifted for clarity.*
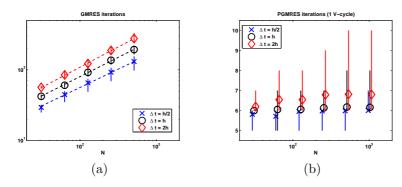

(a) (b)

FIG. 5.9. *Number of GMRES iterations at different grid sizes, in two dimensions. (a) without preconditioning. (b) with V-cycle preconditioner. On the right, the 3 data series for each N have been slightly shifted for clarity.*

grows with the grid size: least square fits yield the approximations $N^{0.5165}$, $N^{0.5435}$ and $N^{0.5702}$ respectively for the number of GMRES iterations on an $N \times N$ grid with the three choices of Δt mentioned above. For homogeneity, the results for $N = 1024$ are not reported in the graph, since they require the restarted GMRES method or a parallel implementation, due to memory limitations when run on a PC with 8Mb of RAM.

Figure 5.9b clearly demonstrates the optimality of the preconditioning strategy adopted, with the number of iterations being in the narrow range 5–10, when N ranges from 32 to 1024, with all the three choices of the time step, and with the average number of iterations being always between 5 and 7. As a side observation, we note that we also employed damped Jacobi as a smoother with analogous results on the optimality, but observing a slightly higher number of iterations (8–11 on average).

Finally, like in the one-dimensional case, Table 5.2 shows a comparison in terms of CPU time for the V-cycle method used as stand-alone solver and as preconditioner for GMRES and CG. We consider the two-dimensional Barenblatt solution with exponent $m = 2$, $\Delta t = h$, and red-black Gauss-Seidel as smoother. All the different strategies give comparable results and the minimum CPU time is again obtained using the multigrid as stand-alone solver with one iteration of pre- and post-smoother. We observe that even if the PGMRES converges in fewer iterations than PCG, when the size of the problem becomes large the PCG becomes faster than PGMRES since the

$N \times N$	MGM(1,0)	MGM(1,1)	PGMRES(1,0)	PGMRES(1,1)	PCG(1,1)
32×32	0.37	0.31	0.48	0.41	0.58
64×64	2.03	1.57	2.63	2.32	1.81
128×128	13.80	10.20	17.30	15.36	11.90
256×256	118.00	85.20	157.00	125.00	104.00
512×512	986,00	734.00	1278.00	1072.00	1001.00

TABLE 5.2

Average CPU time for solving the inner linear system for the two-dimensional example ($m = 2$). The preconditioner is one multigrid iteration where between bracket are denoted the steps of pre- and post-smoother.

vectors that PGMRES requires to memorize at each iteration are large. For PCG a multigrid preconditioner with a larger number of smoothing iterations could be usefull to improve the robustness of the method and to reduce the variance in the PCG iterations.

6. Conclusions and future developments. The novel contribution of this paper relies in the proposal of a fully implicit numerical method for dealing with nonlinear degenerate parabolic equations, in its convergence and stability analysis, and in the study of the related computational cost. Indeed the nonlinear nature of the underlying mathematical model requires the application of a fixed point scheme. We identified the classical Newton method in which, at every step, the solution of a large, locally structured, linear system has been handled by using specialised iterative solvers. In particular, we provided a spectral analysis of the relevant matrices which has been crucial for identifying appropriate preconditioned Krylov methods with efficient V-cycle preconditioners. Numerical experiments for the validation of our analysis complement this contribution.

The research related in this paper was prompted by the applications described in [18], whose natural evolution leads to consider more complicated geometries, in two and three spatial dimensions, and finite element techniques. We observe that the methods considered here can be to some extent generalised to finite elements approximations. In particular, when considering real two-dimensional and three-dimensional cases, the structure of the relevant matrices will depend heavily on the geometry of the domain, on the triangulation/gridding (often generated automatically), and on the type of finite elements (higher order or non Lagrangian etc.). In any case there exists a kind of information depending only on the continuous operator and which is inherited virtually unchanged in both finite differences and finite elements, provided that the grids are quasi-uniform in finite differences and the angles are not degenerating in finite elements. Such information consists in the locally Toeplitz structure (see [22, 19]) used in Section 4 and in the related spectral features (conditioning, subspaces related to small eigenvalues, global spectral behaviour etc.).

Concerning multigrid methods for nonlinear problems in a more general setting and in particular for finite elements approximations, the ideas in [20] and related works should be considered. Moreover, the FAS methods in [5] could be useful when the Jacobian is not easily computable. These nonlinear multigrid strategies will be investigated in the future in particular in connection with the applications in [18].

REFERENCES

[1] G. ALÌ, V. FURUHOLT, R. NATALINI, AND I. TORCICOLLO, *A mathematical model of sulphite chemical aggression of limestones with high permeability. I. Modeling and qualitative analysis.*, Transp. Porous Media, 69 (2007), pp. 109–122.

[2] D. AREGBA DRIOLLET, F. DIELE, AND R. NATALINI, *A mathematical model for the SO_2 aggression to calcium carbonate stones: numerical approximation and asymptotic analysis.*, SIAM J. Appl. Math., 64 (2004), pp. 1636–1667.

[3] A. BERGER, H. BREZIS, AND J. ROGERS, *A numerical method for solving the problem $u_t - \Delta f(u) = 0$*, RAIRO Numer. Anal., 13 (1979), pp. 297–312.

[4] R. BHATIA, *Matrix analysis*, vol. 169 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[5] A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[6] H. BRÉZIS AND A. PAZY, *Convergence and approximation of semigroups of nonlinear operators in Banach spaces*, J. Funct. Anal., 9 (1972), pp. 63–74.

[7] F. CAVALLI, G. NALDI, G. PUPPO, AND M. SEMPLICE, *High-order relaxation schemes for non linear degenerate diffusion problems*, SIAM J. Numer. Anal., 45 (2007), pp. 2098–2119.

[8] F. CLARELLI, C. GIAVARINI, R. NATALINI, C. NITSCH, AND M. SANTARELLI, *Mathematical models for the consolidation processes in stones*, in Proc. of "International Symposium: Stone Consolidation in Cultural Heritage - research and practice". Lisbona, May 2008., 2009.

[9] M. CRANDALL AND T. LIGGETT, *Generation of Semi-Groups of non linear transformations on general Banach spaces*, Amer. J. Math., 93 (1971), pp. 265–298.

[10] S. EVJE AND K. H. KARLSEN, *Monotone difference approximations of BV solutions to degenerate convection-diffusion equations*, SIAM J. Numer. Anal., 37 (2000), pp. 1838–1860.

[11] L. GOLINSKII AND S. SERRA-CAPIZZANO, *The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences*, J. Approx. Theory, 144 (2007), pp. 84–102.

[12] A. GREENBAUM, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[13] W. HACKBUSCH, *Multigrid methods and applications*, vol. 4 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1985.

[14] R. A. HORN AND C. R. JOHNSON, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.

[15] 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), pp. 655–678.

[16] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970.

[17] Y. SAAD, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2003.

[18] M. SEMPLICE, *Preconditioned implicit solvers for nonlinear PDEs in monument conservation*, SIAM J. Sci. Comput., 32 (2010), pp. 3071–3091.

[19] S. SERRA-CAPIZZANO, *The GLT class as a generalized Fourier analysis and applications*, Linear Algebra Appl., 419 (2006), pp. 180–233.

[20] G. STARKE, *Gauss-Newton multilevel methods for least-squares finite element computations of variably saturated subsurface flow*, Computing, 64 (2000), pp. 323–338. International GAMM-Workshop on Multigrid Methods (Bonn, 1998).

[21] J. STOER AND R. BULIRSCH, *Introduction to numerical analysis*, vol. 12 of Texts in Applied Mathematics, Springer-Verlag, New York, third ed., 2002.

[22] P. TILLI, *Locally Toeplitz sequences: spectral properties and applications*, Linear Algebra Appl., 278 (1998), pp. 91–120.

[23] U. TROTTENBERG, C. W. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press Inc., San Diego, CA, 2001.

[24] J. L. VÁZQUEZ, *The porous medium equation*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, Oxford, 2007. Mathematical theory.