JOURNAL OF INTEGRAL EQUATIONS
AND APPLICATIONS
Volume 26, Number 3, Fall 2014

FOR NONLINEAR INFINITE DIMENSIONAL EQUATIONS, WHICH TO BEGIN WITH: LINEARIZATION OR DISCRETIZATION?

LAURENCE GRAMMONT, MARIO AHUES AND FILOMENA D. D'ALMEIDA

Dedicated to the memory of Alain Largillier

ABSTRACT. To tackle a nonlinear equation in a functional space, two numerical processes are involved: discretization and linearization. In this paper we study the differences between applying them in one or in the other order. Linearize first and discretize the linear problem will be in the sequel called option (A). Discretize first and linearize the discrete problem will be called option (B). As a linearization scheme, we consider the Newton method. It will be shown that, under certain assumptions on the discretization method, option (A) converges to the exact solution, contrarily to option (B) which converges to a finite dimensional solution. These assumptions are not satisfied by the classical Galerkin, Petrov-Galerkin and collocation methods, but they are fulfilled by the Kantorovich projection method. The problem to be solved is a nonlinear Fredholm equation of the second kind involving a compact operator. Numerical evidence is provided with a nonlinear integral equation.

1. Introduction. We consider a complex Banach space $\mathcal{X}$ and a nonlinear Fréchet differentiable operator $F : \mathcal{O} \subseteq \mathcal{X} \to \mathcal{X}$ defined on a nonempty open set $\mathcal{O}$ of $\mathcal{X}$. The problem is set as

$$\text{Find } \varphi \in \mathcal{O} : \quad F(\varphi) = 0, \tag{1}$$

where 0 is the null vector of $\mathcal{X}$.

2010 AMS *Mathematics subject classification.* Primary 65J15, 45G10, 35P05.

Keywords and phrases. Nonlinear equations, Newton-like methods, Kantorovich projection approximation, integral equations.

This research has been partially supported by the project IFCPAR-CEFIPRA 4101-1 and received financial support provided through CMUP by the European Regional Development Fund through the programme COMPETE and by the Portuguese Government through the FCT Fundação para a Ciência e Tecnologia under the project PEst-C/MAT/UI0144/2011.

Received by the editors on September 23, 2013, and in revised form on February 16, 2013.

DOI:10.1216/JIE-2014-26-3-413

The exact Newton method in function spaces leads to the Newton sequence $(\varphi^{(k)})_{k\geq 0}$ defined through the relation:

$$F'(\varphi^{(k)})(\varphi^{(k+1)} - \varphi^{(k)}) = -F(\varphi^{(k)}), \quad \varphi^{(0)} \in \mathcal{O}. \tag{2}$$

(For convergence results on Newton method, see the slide of the conference of Villani **[15]** or the book of Argyros **[4]**).

Three options are to be considered:

(NK) Solve (2) in exact arithmetic (i.e., with so-called analytical methods).

(A) Discretize (2) and solve a finite-dimensional linear problem.

(B) Discretize (1), apply Newton's method to the discrete nonlinear problem, and solve the corresponding finite dimensional linear problem.

The first option is often impossible to perform, so one has to choose between option (A) and option (B).

The treatment of (1) depends on the kind of operator involved in the definition of F.

If F involves compact operators as integral operators, projection and iterated projection methods or numerical quadrature rules are well known techniques to compute approximate solutions to such a nonlinear operator equation: in **[13]**, the author explains how to build a sequence of approximate solutions φ_n of the operator equation $\varphi = K\varphi$, where K is a nonlinear operator in $\mathcal{X}$ and φ_n is the solution of an approximate operator equation $\varphi_n = K_n\varphi_n$. He analyses the error estimates for $K_n = \pi_n K$ where $(\pi_n)_{n\geq 1}$ is a sequence of linear projections onto finite dimensional subspaces. In **[12,** Chapter 4, page 244] and **[14]**, the authors approach the solution with that of a perturbed Galerkin equation $\varphi_n = \pi_n K\varphi_n + S_n\varphi_n$ where S_n is some other nonlinear operator. We can also find in **[7]** a comparison between the Galerkin approximation, obtained as a solution of the equation $\varphi_n^G = \pi_n K(\varphi_n^G)$, and the iterated Galerkin approximation defined as $\varphi_n^S := K(\varphi_n^G)$. The authors prove, under appropriate assumptions, that both $\|\varphi_n^G - \varphi\|$ and $\|\varphi_n^S - \varphi\|$ tend to 0 as $n \to \infty$ and they give corresponding error estimates. In **[9]**, the authors propose accelerated projection and iterated projection methods. They consist in decomposing the equation into two components, a finite dimensional one and an infinite dimensional one. In all of the above-mentioned papers, one notices that methods start with a discretization procedure which leads to a nonlinear approximate equation in a finite-dimensional linear space. Next, in these papers, the authors apply a numerical scheme to treat that nonlinear equation, such as Newton method or some of its many variants. This scheme involves at each step the resolution of an $n \times n$ updated linear system. This method corresponds to option (B). If the nonlinear operator F is sufficiently smooth, we suggest to proceed in the opposite sense. First linearize the original equation in the infinite dimensional context, and then solve numerically the underlying sequence of linear equations using some discretization scheme such as a numerical quadrature rule, if the function F involves an integral operator, or a projection method. This method corresponds to option (A).

In the domain of numerical PDEs, operators are not compact. Many papers have been published in the last decades on that issue (see **[3]**, **[8]**, **[10]** and **[16]**). In them the discretization procedure is a *Galerkin* or a *Petrov-Galerkin* method in which case option (A) and option (B) coincide (**[16]**). The notion of mesh independence is not linked to option (A) or (B) as it could seem at first glance:

> Mesh independence of Newton's method means that Newton's method applied to a family of finite dimensional discretizations of a Banach space nonlinear operator equation behaves essentially the same for all sufficiently fine discretizations **[10]**.

So in this domain, with this particular discretization, the linearization and discretization commute and our question has no interest.

For Fredholm equations of the second kind involving a compact operator, there exist numerical methods which are more efficient than the Galerkin or Petrov-Galerkin method. The aim of our paper is to exhibit a class of discretization scheme for which linearization and discretization do not commute. Under certain conditions on the discretization scheme and on the nonlinear operator F, we prove that option (A) behaves better than option (B) for a same and fixed discretization level.

The paper is organized as follows. In Section 2, we present the main theoretical result which is that, *under suitable assumptions on the operator and on the discretization process*, if a Newton type method is applied first and the discretization process of order n is used at each step of the Newton like method, then the sequence of iterates converges to φ, the exact solution of (1), for any fixed integer n large enough. This means that one can attain any desired accuracy by employing a suitable discretization method of such a fixed order n. In Section 3, we illustrate this result with an application to a nonlinear Fredholm equation of the second kind involving a compact operator K and a given function f :

$$F(x) := x - K(x) - f.$$

We discretize (2) with the so-called Kantorovich scheme (see **[2,** page 186]):

$$(I - \pi_n K'(\varphi_n^{(k)}))(\varphi_n^{(k+1)} - \varphi_n^{(k)}) = -\varphi_n^{(k)} + K(\varphi_n^{(k)}) + f,$$

where $(\pi_n)_{n\geq 1}$ is a sequence of linear bounded projections. We remark that π_n acts only on K', which makes a significant difference with respect to Galerkin, Petrov-Galerkin or collocation schemes in which π_n acts on the whole equation:

$$\pi_n[(I - K'(\varphi_n^{(k)}))(\varphi_n^{(k+1)} - \varphi_n^{(k)})] = \pi_n[-\varphi_n^{(k)} + K(\varphi_n^{(k)}) + f].$$

The scheme, which we propose in our paper, converges when K is compact (and hence K' too) and the sequence $(\pi_n)_{n\geq 1}$ is pointwise convergent to the identity operator I. Our work shows that options (A) and (B) are not equivalent and that (A) should be preferred to (B). Then we exhibit some numerical examples involving an integral operator, confirming in practice our theoretical results. In section 4, we give the discrete version of option (A). In section 5, we list some concluding remarks and provisory conclusions.

In this paper, $\mathcal{L}(\mathcal{X})$ denotes the real Banach algebra of all bounded linear operators from $\mathcal{X}$ into itself, $B_\rho(u)$ denotes the closed ball with center $u \in \mathcal{X}$ and radius $\rho > 0$ and $F'(x)$ denotes the Fréchet derivative of F at x.

2. Option (A) vs option (B).

2.1. Theoretical analysis of option (A). The solution of problem (1) is characterized as the limit of the sequence $(\varphi^{(k)})_{k\geq 0}$ defined through the relation

$$F'(\varphi^{(k)})(\varphi^{(k+1)} - \varphi^{(k)}) = -F(\varphi^{(k)}), \quad \varphi^{(0)} \in \mathcal{O}. \tag{3}$$

These iterations are functional linear equations. To solve them numerically, in most cases, we apply a discretization scheme. For example, if F is a Fréchet differentiable nonlinear integral operator then, for each $x \in \mathcal{O}$, $F'(x)$ is a linear integral operator, and the discretization process could be a numerical quadrature such as the Nyström scheme or a projection method such as Kantorovich scheme.

If $F'(\varphi^{(k)})$ is invertible, equation (3) can be rewritten as

$$\varphi^{(k+1)} = \varphi^{(k)} - F'(\varphi^{(k)})^{-1} F(\varphi^{(k)}).$$

Let $\Sigma_n : \mathcal{O} \to \mathcal{L}(\mathcal{X})$ be such that, in some discretized sense, for each $x \in \mathcal{O}$, $\Sigma_n(x)$ is an approximation to $F'(x)^{-1}$. Then corresponding discretized iterates satisfy:

$$\varphi_n^{(0)} \in \mathcal{O}, \quad \varphi_n^{(k+1)} = \varphi_n^{(k)} - \Sigma_n(\varphi_n^{(k)}) F(\varphi_n^{(k)}). \tag{4}$$

We give sufficient conditions on the approximate operator Σ_n to ensure the cited conditions. For this purpose, we will interpret equation (4) as a Newton-like method.

Theorem 2.1. A priori convergence theorem. *Suppose that F, $\mathcal{O}$, $\varphi \in \mathcal{O}$, $\mu > 0$, $R > 0$, $\ell > 0$ and $\alpha \in]0, 1]$ are such that:*

(i) *$F(\varphi) = 0$, $F'(\varphi)$ is invertible and $\|F'(\varphi)^{-1}\| \leq \mu$.*

(ii) *The closed ball $B_R(\varphi)$ is included in the open set $\mathcal{O}$, and $F' : \mathcal{O} \to \mathcal{L}(\mathcal{X})$ is $(\ell, \alpha)-$Hölder continuous on $B_R(\varphi)$.*

(iii) *For*

$$r := \min\left\{R, \frac{1}{(2\mu\ell)^{1/\alpha}}\right\},$$

there exists $\gamma_n \in]0, 1[$ such that

$$\sup_{x \in B_r(\varphi)} \|I - \Sigma_n(x) F'(x)\| \leq \gamma_n.$$

(iv) *The starting approximation $\varphi_n^{(0)}$ is chosen in the closed ball $B_{\rho_n}(\varphi)$, where*

$$\rho_n := \min\left\{r, \left(\frac{1 - \gamma_n}{4\ell\mu(1 + \gamma_n)}\right)^{1/\alpha}\right\}.$$

Then, for all k, $\varphi_n^{(k)} \in B_{\rho_n}(\varphi)$, and

$$\|\varphi_n^{(k)} - \varphi\| \leq \rho_n \left(\frac{1+\gamma_n}{2}\right)^k \to 0 \qquad \text{as } k \to \infty.$$

Proof. We first prove that, for all $x \in B_r(\varphi)$, $F'(x)$ is invertible and

$$\|F'(x)^{-1}\| \leq 2\mu.$$

In fact, for all $x \in B_r(\varphi)$,

$$F'(x) = F'(\varphi) + F'(x) - F'(\varphi) = F'(\varphi)[I + F'(\varphi)^{-1}(F'(x) - F'(\varphi))].$$

Since

$$\|F'(\varphi)^{-1}(F'(x) - F'(\varphi))\| \leq \|F'(\varphi)^{-1}\| \, \|F'(x) - F'(\varphi)\| \leq \mu \ell r^\alpha \leq \frac{1}{2},$$

we conclude that $F'(x)$ is invertible and that its inverse is uniformly bounded on $B_r(\varphi)$:

$$F'(x)^{-1} = \left[I + F'(\varphi)^{-1}(F'(x) - F'(\varphi))\right]^{-1} F'(\varphi)^{-1},$$

so

$$\|F'(x)^{-1}\| \leq \mu \sum_{k=0}^{\infty} \|F'(\varphi)^{-1}(F'(x) - F'(\varphi))\|^k \leq 2\mu.$$

Concerning $\Sigma_n(x)$, we remark that, for all $x \in B_r(\varphi)$,

$$\Sigma_n(x) = F'(x)^{-1} - (I - \Sigma_n(x)F'(x))F'(x)^{-1};$$

hence,

$$\|\Sigma_n(x)\| \leq 2\mu(1 + \gamma_n).$$

Since

$$\varphi_n^{(k+1)} - \varphi = \varphi_n^{(k)} - \varphi - \Sigma_n(\varphi_n^{(k)}) \left(F(\varphi_n^{(k)}) - F(\varphi)\right),$$

and

$$F(\varphi_n^{(k)}) - F(\varphi) = \int_0^1 F'\left((1-t)\varphi_n^{(k)} + t\varphi\right)(\varphi_n^{(k)} - \varphi)\, dt,$$

then

$$\varphi_n^{(k+1)} - \varphi = \int_0^1 \left[I - \Sigma_n(\varphi_n^{(k)})F'((1-t)\varphi_n^{(k)} + t\varphi)\right](\varphi_n^{(k)} - \varphi)\, dt.$$

Let $F'(\varphi_n^{(k)})$ be added to and subtracted from $F'((1-t)\varphi_n^{(k)} + t\varphi)$. We get

$$\begin{aligned}\varphi_n^{(k+1)} - \varphi \\ &= \int_0^1 \left[I - \Sigma_n(\varphi_n^{(k)})F'(\varphi_n^{(k)})\right](\varphi_n^{(k)} - \varphi)\, dt \\ &\quad + \int_0^1 \Sigma_n(\varphi_n^{(k)})\left(F'((1-t)\varphi_n^{(k)} + t\varphi) - F'(\varphi_n^{(k)})\right)(\varphi - \varphi_n^{(k)})dt,\end{aligned}$$

and

$$\begin{aligned}\|\varphi_n^{(k+1)} - \varphi\| \le \|I - \Sigma_n(\varphi_n^{(k)})F'(\varphi_n^{(k)})\|\, \|\varphi_n^{(k)} - \varphi\| \\ + \|\Sigma_n(\varphi_n^{(k)})\|\, \|\varphi_n^{(k)} - \varphi\| \\ \times \int_0^1 \|F'((1-t)\varphi_n^{(k)} + t\varphi) - F'(\varphi_n^{(k)})\|\, dt.\end{aligned}$$

Let $\varphi_n^{(k)} \in B_{\rho_n}(\varphi)$. Then $\|I - \Sigma_n(\varphi_n^{(k)})F'(\varphi_n^{(k)})\| \le \gamma_n$, and since $B_r(\varphi)$ is convex, for $t \in [0,1]$, $(1-t)\varphi_n^{(k)} + t\varphi \in B_r(\varphi)$ and

$$\|F'((1-t)\varphi_n^{(k)} + t\varphi) - F'(\varphi_n^{(k)})\| \le \ell\|\varphi_n^{(k)} - \varphi\|^\alpha.$$

Hence,

$$\begin{aligned}\|\varphi_n^{(k+1)} - \varphi\| &\le \|\varphi_n^{(k)} - \varphi\|\left(\gamma_n + 2\mu\ell(1+\gamma_n)\|\varphi_n^{(k)} - \varphi\|^\alpha\right) \\ &\le \frac{1+\gamma_n}{2}\|\varphi_n^{(k)} - \varphi\|.\end{aligned}$$

Since $1 + \gamma_n < 2$, the previous inequality implies that $\varphi_n^{(k+1)} \in B_{\rho_n}(\varphi)$ and that

$$\|\varphi_n^{(k)} - \varphi\| \le \rho_n\left(\frac{1+\gamma_n}{2}\right)^k \to 0 \quad \text{as } k \to \infty.$$

The proof is complete. $\square$

Remark 2.2. Notice that the sequence $(\varphi_n^{(k)})_{k\geq 0}$ tends to φ which does not depend on n. This means that we are not constrained by the value of n, provided the assumptions are satisfied. We will see in Section 3 that the numerical computation of $\varphi_n^{(k)}$ may be required to solve an $n\times n$ linear system. That is, by solving k linear systems of order n, we get an approximation $\varphi_n^{(k)}$ of any desired accuracy, if k is large enough.

In **[11]**, option (A) is applied to a Fredholm equation of the second kind $\varphi - K(\varphi) = f$, involving an integral operator K, with the Nyström method as the discretization process. In Section 3, we propose to discretize with a projection method–the Kantorovich projection method–fulfilling the assumptions of the theorem.

If the classical Galerkin method or collocation method, built upon the finite rank projection π_n, is applied to equation (3), then option (A) leads to a sequence $(\varphi_n^{(k)})_{k\geq 0}$ in $\mathcal{X}_n$, the range of π_n, satisfying

$$(I - \pi_n K'(\varphi_n^{(k)}))(\varphi_n^{(k+1)} - \varphi_n^{(k)}) = -\pi_n F(\varphi^{(k)}), \quad \varphi^{(0)} \in \mathcal{O}.$$

Then $\Sigma_n(\varphi_n^{(k)}) := (I - \pi_n K'(\varphi_n^{(k)}))^{-1}\pi_n$ is not invertible so that the condition $\|I - \Sigma_n(x)F'(x)\| \leq \gamma_n$ will not be satisfied.

2.2. Option (B). There is no possible general analysis of option (B). We can only state that, starting with some discretization scheme in an n-dimensional subspace of $\mathcal{X}$ applied to problem (1), we are led at some stage of computations to solve a nonlinear system of equations

$$\text{Find } \mathsf{x}_n^{(\infty)} \in \mathbf{C}^{n\times 1} : \mathsf{F}_n(\mathsf{x}_n^{(\infty)}) = 0.$$

This problem will be linearized and become a sequence of linear updated systems. For instance, under suitable conditions, the Newton method can be applied. With such a method, a sequence $(\mathsf{x}_n^{(k)})_{k\geq 0}$ in $\mathbf{C}^{n\times 1}$ is built with the recursion formula

$$\mathsf{F}_n'(\mathsf{x}_n^{(k)})(\mathsf{x}_n^{(k+1)} - \mathsf{x}_n^{(k)}) = -\mathsf{F}_n(\mathsf{x}_n^{(k)}).$$

Sufficient conditions on the starting point $\mathsf{x}_n^{(0)}$ for the convergence of this sequence may be found, for instance, in **[1]**. Next, the vector $\mathsf{x}_n^{(k)}$ should allow us to compute a function ψ_n which will be called an n-order approximation of φ.

3. Kantorovich projection approximation for equations of the second kind. Let $K : \mathcal{O} \to \mathcal{X}$ be a Fréchet-differentiable nonlinear compact operator and $T := K'$ denote the Fréchet dérivative of K.

Remark 3.1. We have chosen T to denote K' because the methods presented in this paper involve an approximation of K' built through an operator of finite rank n. A symbol such as K'_n is ambiguous: it may denote both the derivative of some operator K_n or an n-order approximation of K'. The advantage of writing T for K' is that T_n will always denote some n-order approximation of T.

The problem is:

$$\text{Given} \quad f \in \mathcal{X}, \text{ find } \varphi \in \mathcal{O} : \varphi - K(\varphi) = f. \tag{5}$$

Problem (5) will be handled as: Find $\varphi \in \mathcal{O} : F(\varphi) = 0$, where, for all $x \in \mathcal{O}$, $F(x) := x - K(x) - f$, and hence $F'(x) = I - T(x)$.

Let π_n be a projection onto an n-dimensional subspace of $\mathcal{X}$ spanned by an ordered basis $e_n := [e_{n,1}, \ldots, e_{n,n}] \in \mathcal{X}^{1\times n}$. Let $e_n^* := [e_{n,1}^*, \ldots, e_{n,n}^*] \in (\mathcal{X}^*)^{1\times n}$ be an ordered basis of the annihilator of the null space of π_n which is adjoint to $[e_{n,1}, \ldots, e_{n,n}]$. Each $e_{n,j}^*$ is a bounded semi-linear functional and π_n is characterized by:

$$\pi_n x = \sum_{j=1}^{n} \langle x, e_{n,j}^* \rangle e_{n,j}, \quad x \in \mathcal{X},$$

where

$$\langle x, e_{n,j}^* \rangle := \overline{e_{n,j}^*(x)}.$$

We remark that, for all $j \in [[1, n]]$,

$$\pi_n e_{n,j} = e_{n,j}, \qquad \pi_n^* e_{n,j}^* = e_{n,j}^*.$$

We shall use the following notation which allows us to simplify the description of matrices and linear combinations:

$$e_n \mathsf{x} := \sum_{j=1}^{n} \mathsf{x}(j) e_{n,j}$$

for all $\mathsf{x} \in \mathbf{C}^{n\times 1}$,

$$\langle\!\langle\, h, e_n^* \,\rangle\!\rangle(i,j) := \langle h_j, e_{n,i}^* \rangle$$

for all $h := [h_1, \ldots, h_p] \in \mathcal{X}^{1\times p}$,

$$Lx := [Lx_1, \ldots, Lx_m]$$

for all $x := [x_1, \ldots, x_m] \in \mathcal{X}^{1\times m}$ and all $L : \mathcal{X} \to \mathcal{X}$. For example, with such notation,

$$\pi_n x = e_n \langle\!\langle\, x, e_n^* \,\rangle\!\rangle, \quad x \in \mathcal{X}.$$

3.1. Option (A). With option (A), we apply the Kantorovich projection discretization to the linear operator equation issued from the Newton scheme:

$$\text{(6)} \qquad \text{Find} \quad \varphi_n^{(k+1)} \in \mathcal{X} : \varphi_n^{(k+1)} - \pi_n T(\varphi_n^{(k)})\varphi_n^{(k+1)} = g_n^{(k)},$$

where

$$g_n^{(k)} := K(\varphi_n^{(k)}) - \pi_n T(\varphi_n^{(k)})\varphi_n^{(k)} + f.$$

We suppose that

$$\text{(7)} \qquad \begin{cases} \text{(i)} & \text{Equation (5) has a unique solution } \varphi \in \mathcal{O}, \\ \text{(ii)} & I - T(\varphi) \text{ is invertible,} \\ \text{(iii)} & T : \mathcal{O} \to \mathcal{L}(\mathcal{X}) \text{ is } \ell\text{-Lipschitz.} \end{cases}$$

In the following, we prove that the assumptions of Theorem 2.1 are fulfilled.

Let $\mu > 0$ and $R > 0$ be such that

$$\|(I - T(\varphi))^{-1}\| \le \mu, \quad B_R(\varphi) \subset \mathcal{O}.$$

Concerning the constants of Theorem 2.1, fix

$$\alpha = 1, \quad r = \min\left\{R, \frac{1}{2\mu\ell}\right\}.$$

The discretization process is based upon the approximation

$$T_n(x) := \pi_n T(x), \quad x \in B_R(\varphi).$$

As in Theorem 2.1, we can prove that, for all $x \in B_r(\varphi)$, $I - T(x)$ is invertible and

$$\|(I - T(x))^{-1}\| \leq 2\mu.$$

Proposition 3.2. *Suppose that* (7) *holds, and also that*

(i) *For all* $x \in \mathcal{X}$, $\pi_n x \to x$ *as* $n \to \infty$.
(ii) *The set*

$$W := \{T(x)h \ : \ x \in B_R(\varphi),\, h \in \mathcal{X},\, \|h\| = 1\}$$

is relatively compact.

Then

$$\lim_{n\to\infty} \sup_{x \in B_R(\varphi)} \|T_n(x) - T(x)\| = 0.$$

Proof.

$$W := \{T(x)h : x \in B_R(\varphi), \quad h \in \mathcal{X},\ \|h\| = 1\}$$

is relatively compact, π_n tends to I pointwise and pointwise convergence is uniform on relatively compact sets. □

Proposition 3.3. *Under the assumptions of Proposition* 3.2, *for* n *large enough, the approximate inverse defined by*

$$\Sigma_n(x) := (I - T_n(x))^{-1}, \tag{8}$$

exists and is uniformly bounded for $x \in B_r(\varphi)$.

Proof. Set

$$\delta_n := \sup_{x \in B_r(\varphi)} \|T_n(x) - T(x)\|. \tag{9}$$

Since $\delta_n \to 0$ as $n \to \infty$, for all n large enough,

$$\delta_n < \frac{1}{2\mu}.$$

As

$$I - T_n(x) = (I - T(x))[I - (I - T(x))^{-1}(T_n(x) - T(x))]$$

and by the Neumann series theorem, $I - T_n(x)$ is invertible for all x in $B_r(\varphi)$, and the following uniform bound holds for all $x \in B_r(\varphi)$:

$$\|\Sigma_n(x)\| \leq \frac{2\mu}{1 - 2\mu\delta_n}. \tag{10}$$

Hence, for all n large enough, and all $x \in B_r(\varphi)$, the operator $\Sigma_n(x)$ defined by (8) is well defined, belongs to $\mathcal{L}(\mathcal{X})$ and is uniformly bounded over $B_r(\varphi)$. $\square$

Proposition 3.4. *Under the assumptions of Proposition* 3.2*, there exists* $\gamma_n < 1$ *such that*

$$\sup_{x \in B_r(\varphi)} \|I - \Sigma_n(x)(I - T(x))\| \leq \gamma_n.$$

Proof. For all $x \in B_r(\varphi)$,

$$I - \Sigma_n(x)F'(x) = I - (I - T_n(x))^{-1}(I - T(x)) = (I - T_n(x))^{-1}(T(x) - T_n(x)).$$

Hence,

$$\|I - \Sigma_n(x)(I - T(x))\| \leq \|(I - T_n(x))^{-1}\| \, \|T(x) - T_n(x)\| \leq \frac{2\mu\delta_n}{1 - 2\mu\delta_n}.$$

Define

$$\gamma_n := \frac{2\mu\delta_n}{1 - 2\mu\delta_n},$$

and choose n large enough to have $\gamma_n < 1$. $\square$

Then, assuming the conditions of Proposition 3.2, we proved that the hypotheses of Theorem 2.1 are satisfied and Theorem 2.1 can be applied. We can state that

$$\varphi_n^{(k)} \to \varphi \quad \text{as } k \to +\infty.$$

Let us focus on the implementation of option (A): for computational purposes, we remark that, as $g_n^{(k)} := K(\varphi_n^{(k)}) - \pi_n T(\varphi_n^{(k)})\varphi_n^{(k)} + f$,

$$(I - \pi_n)\varphi_n^{(k+1)} = (I - \pi_n)g_n^{(k)} = (I - \pi_n)(K(\varphi_n^{(k)}) + f),$$

and hence,

$$\varphi_n^{(k+1)} = (I - \pi_n)(K(\varphi_n^{(k)}) + f) + e_n \mathsf{x}_n^{(k+1)} \tag{11}$$

for a column vector $\mathsf{x}_n^{(k+1)} \in \mathbf{C}^{n\times 1}$ solving the linear system

$$(\mathsf{I}_n - \mathsf{A}_n^{(k)})\mathsf{x}_n^{(k+1)} = \mathsf{b}_n^{(k)},$$

where

$$\begin{aligned}
\mathsf{A}_n^{(k)} &:= \langle\, T(\varphi_n^{(k)})e_n, e_n^* \,\rangle, \\
\mathsf{b}_n^{(k)} &:= \langle\, K(\varphi_n^{(k)}), e_n^* \,\rangle - \langle\, T(\varphi_n^{(k)})\varphi_n^{(k)}, e_n^* \,\rangle \\
&\quad + \langle\, f, e_n^* \,\rangle + \langle\, T(\varphi_n^{(k)})(I - \pi_n)(K(\varphi_n^{(k)}) + f), e_n^* \,\rangle.
\end{aligned}$$

3.2. Option (B). With option (B), we define the Kantorovich projection approximation $\psi_n \in \mathcal{X}$ to be such that

$$\psi_n - \pi_n K(\psi_n) = f. \tag{12}$$

Hence,

$$(I - \pi_n)\psi_n = (I - \pi_n)f,$$

and setting

$$\mathsf{x}_n^{(\infty)} := \langle\, \psi_n, e_n^* \,\rangle,$$

we see that the function ψ_n is of the form

$$\psi_n = (I - \pi_n)f + e_n \mathsf{x}_n^{(\infty)}.$$

Equation (12) is equivalent to

$$\mathsf{F}_n(\mathsf{x}_n^{(\infty)}) = 0,$$

where F_n is the nonlinear operator defined from some open subset $\mathcal{O}_n$ of $\mathbf{C}^{n\times 1}$ into $\mathbf{C}^{n\times 1}$, by

$$\mathsf{F}_n(\mathsf{x}) := \mathsf{x} - (\, K\big((I-\pi_n)f + e_n\mathsf{x}\big), e_n^* \,) - (\, f, e_n^* \,).$$

The properties of K imply that F_n is Fréchet-differentiable, and its Jacobian matrix at x is

$$\mathsf{F}_n'(\mathsf{x}) = \mathsf{I}_n - (\, T((I-\pi_n)f + e_n\mathsf{x})e_n, e_n^* \,).$$

Suppose we approximate the vector $x_n^{(\infty)}$ through the Newton sequence $(\mathsf{x}_n^{(k)})_{k\geq 0}$. Then

$$\psi_n^{(k+1)} := (I-\pi_n)f + e_n\mathsf{x}_n^{(k+1)}, \tag{13}$$

and the linear system to be solved for $\mathsf{x}_n^{(k+1)}$ reads as

$$(\mathsf{I}_n - \mathsf{C}_n^{(k)})\mathsf{x}_n^{(k+1)} = \mathsf{d}_n^{(k)},$$

where

$$\mathsf{C}_n^{(k)} := (\, T(\psi_n^{(k)})e_n, e_n^* \,),$$
$$\mathsf{d}_n^{(k)} := (\, K(\psi_n^{(k)}), e_n^* \,) - (\, T(\psi_n^{(k)})e_n, e_n^* \,)\mathsf{x}_n^{(k)} + (\, f, e_n^* \,).$$

Comparing option (A) with option (B) through the linear systems to be solved, we notice that the right hand side of option (A) is richer than the one of option (B). Also the reconstruction formula (11) for option (A) is richer than (13) for option (B).

3.3. Numerical example. The aim of this section is to give an academic example illustrating the behavior of option (A) and option (B) in the case of Kantorovich projection approximations with an interpolatory projection. It corresponds to a particular choice of the sequence of projections and reads as follows:

We consider the Banach space $\mathcal{X} := C^0([0,1],\mathbf{R})$ of all continuous real-valued functions defined on $[0,1]$. We have chosen the integral operator

$$K(\varphi)(s) := \int_0^1 \kappa(s,t,\varphi(t))t, \quad \varphi\in\mathcal{X},\ s\in[0,1],$$

defined by the kernel κ given by

$$\kappa(s,t,u) := \sin(4\pi\, s)\, t\, u^2, \quad (s,t) \in [0,1]\times[0,1],\ u \in \mathbf{R},$$

and the function f is defined by

$$f(s) := \frac{3}{4}\sin(4\pi\, s), \quad s \in [0,1],$$

so that the exact solution φ is given by

$$\varphi(s) := \sin(4\pi\, s), \quad s \in [0,1].$$

The projection π_n is built upon a uniform grid in $[0,1]$. Let $t_{n,i} := (i-1)/(n-1)$ for all $i \in [[1,n]]$. Associated with this grid, we define the approximating space to be the subspace of all piecewise linear continuous functions. The canonical basis of this subspace is formed by the so-called hat functions $[e_{n,1},\ldots,e_{n,n}] \in \mathcal{X}^{1\times n}$. For $j =\in [[2,n-1]]$,

$$e_{n,j}(t) := \begin{cases} \frac{t-t_{n,j-1}}{t_{n,j}-t_{n,j-1}} & \text{for } t \in [t_{n,j-1},t_{n,j}],\\ \frac{t_{n,j+1}-t}{t_{n,j+1}-t_{n,j}} & \text{for } t \in [t_{n,j},t_{n,j+1}],\\ 0 & \text{otherwise.}\end{cases}$$

$$e_{n,1}(t) := \begin{cases} \frac{t_{n,2}-t}{t_{n,2}-t_{n,1}} & \text{for } t \in [t_{n,1},t_{n,2}],\\ 0 & \text{otherwise,}\end{cases}$$

$$e_{n,n}(t) := \begin{cases} \frac{t-t_{n,n-1}}{t_{n,n}-t_{n,n-1}} & \text{for } t \in [t_{n,n-1},t_{n,n}],\\ 0 & \text{otherwise.}\end{cases}$$

Let $e_n^* := [e_{n,1}^*,\ldots,e_{n,n}^*] \in (\mathcal{X}^*)^{1\times n}$ be the adjoint basis given by: $\langle x, e_{n,j}^*\rangle = x(t_{n,j})$ for all $x \in \mathcal{X}$. The corresponding projection π_n corresponds to the piecewise linear interpolation and is given by $\pi_n x = \sum_{j=1}^n x(t_{n,j})e_{n,j}$ for all $x \in \mathcal{X}$.

It is easy to check that the assumptions of Proposition 3.2 are satisfied so that the convergence result of option (A) can be applied. Concerning computations, the integrals needed for building matrices involved can be calculated by hand in this example.

Table 1 illustrates in practice that option (B) does not converge to the exact solution. On the contrary, independently of the value of n, option (A) converges to the exact solution. Although the rate of convergence of option (B) is quadratic, its limit is the discretized solution ψ_n.

n	option (A)	option (B)
10	0.999874127673875	0.984807753012208
	0.249968531918469	0.21809485857788
	0.049592000737629	0.087702839415384
	0.007542157161705	0.081124614006456
	0.001026726630759	0.081107606627030
	0.000136287926264	0.081107606513346
	0.000018020624652	0.081107606513346
	0.000002381507787	0.081107606513346
	0.000000314704968	0.081107606513346
	0.000000041586383	0.081107606513346
	0.000000005495385	0.081107606513346
	0.000000000726182	0.081107606513346
	0.000000000095961	0.081107606513346
1000	0.999999987660527	0.999998262219395
	0.249999996915132	0.249996289179079
	0.025002302500949	0.025005485898316
	0.000305308116624	0.000313039885821
	0.000000051593932	0.000008284241186
	0.000000000000836	0.000008237803859
		0.000008237803860
		0.000008237803861
		0.000008237803861
		0.000008237803861
		0.000008237803861

TABLE 1. Norm of the error relative to the exact solution.

Although theory predicts a linear convergence for option (A), practice shows that, as n increases, this rate may become superlinear and almost quadratic. The reason for this behavior is obviously the fact that the greater is n, the better the Fréchet derivative of F is approximated by the discretization of each linear step of Newton method. These aspects are also illustrated in Figures 1, 2 and 3.

4. Discrete version of option (A). If $F(\varphi_n^{(k)})$, in (4), cannot be computed exactly, then the proposed algorithm issued from option (A) needs further discretization. It is not compulsory to apply the discretization process used to discretize Newton iterations, nor the same level n which gives the order of the system to be solved. Let F_N be a numerical evaluation of F, with $N \gg n$. If F involves an integral operator, F_N can be built from F, replacing the integrals by numerical

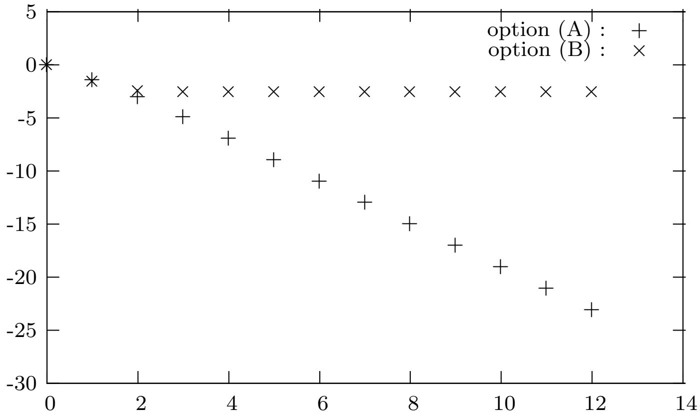

FIGURE 1. Logarithm of the errors of option (A) and option (B) for $n = 10$.

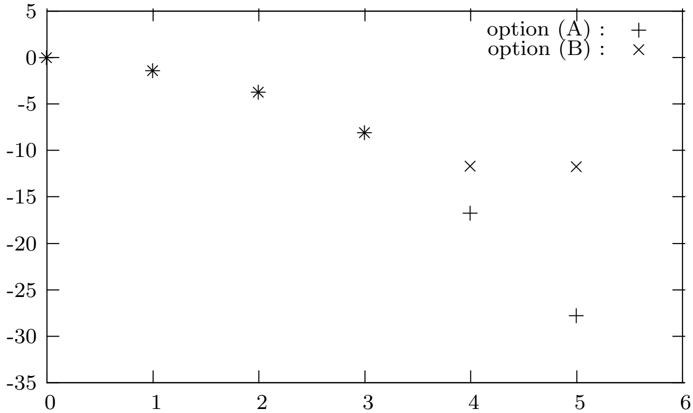

FIGURE 2. Logarithm of the errors of option (A) and option (B) for $n = 1000$.

quadratures. Then a discrete version of Theorem 2.1 can be written replacing the operator F by F_N.

Let us define the sequence of discrete Newton iterates as

$$\varphi_{N,n}^{(0)} \in \mathcal{O}_N, \qquad \varphi_{N,n}^{(k+1)} = \varphi_{N,n}^{(k)} - \Sigma_{N,n}(\varphi_{N,n}^{(k)}) F_N(\varphi_{N,n}^{(k)}). \tag{14}$$

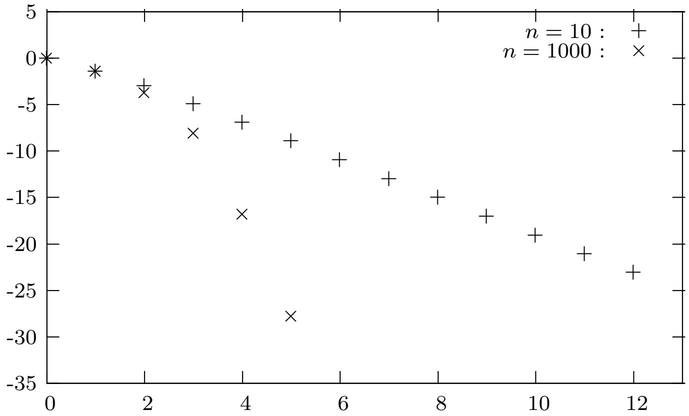

FIGURE 3. Logarithm of the errors of option (A).

Theorem 4.1. A priori convergence of the discrete version. *Suppose that F_N, $\mathcal{O}_N$, $\varphi_N \in \mathcal{O}$, $\mu_N > 0$, $R_N > 0$, $\ell_N > 0$ and $\alpha_N \in]0,1]$ are such that:*

(i) $F_N(\varphi_N) = 0$, *$F'_N(\varphi_N)$ is invertible and* $\|F'_N(\varphi_N)^{-1}\| \leq \mu_N$,
(ii) *The closed ball $B_{R_N}(\varphi_N)$ is included in the open set $\mathcal{O}_N$, and $F'_N : \mathcal{O}_N \to \mathcal{L}(\mathcal{X})$ is (ℓ_N, α_N)-Hölder continuous on $B_{R_N}(\varphi_N)$,*
(iii) *For*

$$r_N := \min\left\{R_N, \frac{1}{(2\mu_N \ell_N)^{1/\alpha_N}}\right\},$$

there exists $\gamma_{N,n} \in]0,1[$ such that

$$\sup_{x \in B_{r_N}(\varphi_N)} \|I - \Sigma_{N,n}(x) F'_N(x)\| \leq \gamma_{N,n},$$

(iv) *The starting approximation $\varphi_n^{(0)}$ is chosen in the closed ball $B_{\rho_{N,n}}(\varphi_N)$, where*

$$\rho_{N,n} := \min\left\{r_N, \left(\frac{1-\gamma_{N,n}}{4\ell_N \mu_N (1+\gamma_{N,n})}\right)^{1/\alpha_N}\right\}.$$

Then, for all k, $\varphi_{N,n}^{(k)} \in B_{\rho_{N,n}}(\varphi_N)$, and

$$\|\varphi_{N,n}^{(k)} - \varphi_N\| \leq \rho_n \left(\frac{1+\gamma_{N,n}}{2}\right)^k \longrightarrow 0 \quad \text{as } k \to \infty.$$

Proof. This is the same proof as for Theorem 2.1, replacing F by F_N. □

Remark 4.2. The discrete version of option A is not equivalent to the application of option B, except in the case where $n := N$ and $\Sigma_{N,N}(\varphi_N^{(k)}) := (F'_N(\varphi_N))^{-1}$. The philosophy of the discrete version of option A is to have the solution φ_N of $F_N(x) = 0$ at a cheaper cost.

The application of option B leads to linear systems of *size* N whose solutions $\psi_N^{(k)}$ tend to φ_N whereas the discrete version of option A leads to the resolution of linear systems of *size* $n \ll N$ whose solutions $\varphi_{N,n}^{(k)}$ tend also to φ_N but at a cheaper computational cost.

This discrete version seems close to Axelsson [**8**]. Given a coarse grid (mesh size H) and a fine grid (mesh size h) Axelsson's idea is to compute a coarse approximation ϕ_H of the exact solution, then correct it once using one Newton iteration on the fine grid, obtaining in this way an approximation ϕ_h^0. Next, the author properly relates the parameters H and h so that the error on ϕ_h^0 be the same as the error on the fine approximation ϕ_h.

Our philosophy is different. We intend to solve several Newton iterations on a coarse grid (for a computational cost reason). While Axelsson performs his Newton iteration on the fine grid, we perform our Newton iterations on the coarse grid. Our Newton sequence tends to the fine approximation φ_N and we need to solve several small systems. If the parameters H and h are chosen properly, the Axelsson approximation has the same order as the fine approximation and he needs to solve only one system but with high dimension.

Let us apply Theorem 4.1 to a particular application belonging to the case considered in Section 3, $F(x) = x - K(x) - f$ where K is a Urysohn integral operator defined by

$$K(\varphi)(s) := \int_0^1 \kappa(s,t,\varphi(t))\,dt, \quad \varphi \in \mathcal{X},\ s \in [0,1],$$

whose kernel κ is smooth. As Atkinson and Flores in **[6]**, F_N and F'_N are obtained by replacing all the integrals by numerical integrals. Then

$$F_N(x) := x - K_N(x) - f, \quad \text{and hence} \quad F'_N(x) = I - T_N(x),$$

where the Nyström approximation K_N of order N of the nonlinear operator K is given by

$$K_N(x)(s) = \sum_{j=1}^{N} \omega_{N,j} \kappa(s, t_{N,j}, x(t_{N,j})), \quad x \in \mathcal{O}, \; s \in [0,1].$$

and the Nyström approximation of order N of the linear operator $T(x) = K'(x)$ is denoted by $T_N(x)$ and is given by

$$T_N(x)h(s) = \sum_{j=1}^{N} \omega_{N,j} \frac{\partial \kappa}{\partial u}(s, t_{N,j}, x(t_{N,j}))h(t_{N,j}),$$
$$x \in \mathcal{O}, \quad h \in \mathcal{X}, \quad s \in [0,1].$$

Then φ_N is the Nyström approximation of φ and if the kernel κ is smooth enough, then the properties (i) and (ii) of Theorem 4.1 are fulfilled.

In the discrete version, we have

$$\Sigma_{N,n}(x) := (I - \pi_n T_N(x))^{-1},$$

where π_n is a projection such that for all $x \in \mathcal{X}$, $\pi_n x \to x$ as $n \to \infty$.

As in Theorem 2.1, we can prove that, for all $x \in B_{r_N}(\varphi_N)$, $I - T_N(x)$ is invertible and

$$\|(I - T_N(x))^{-1}\| \leq 2\mu_N.$$

The set

$$W_N := \{T_N(x)h \; : \; x \in B_{R_N}(\varphi_N), \quad h \in \mathcal{X}, \; \|h\| = 1\}$$

is relatively compact by Arzela-Ascoli, so that

$$\lim_{n\to\infty} \sup_{x \in B_{R_N}(\varphi_N)} \|\pi_n T_N(x) - T_N(x)\| = 0.$$

As in Section 3, we prove that, for all $x \in B_{r_N}(\varphi_N)$,

$$\|\Sigma_{N,n}(x)\| \leq \frac{2\mu_N}{1 - 2\mu_N \delta_{N,n}}, \tag{15}$$

where

$$\delta_{N,n} := \sup_{x \in B_{r_N}(\varphi_N)} \|\pi_n T_N(x) - T_N(x)\|.$$

As in Proposition 3.4, we prove that $\gamma_{N,n} < 1$ exists such that

$$\sup_{x \in B_{r_N}(\varphi_N)} \|I - \Sigma_{N,n}(x)(I - T_N(x))\| \leq \gamma_{N,n}.$$

Hence, we proved that assumption (iii) of Theorem 4.1 is fulfilled.

5. Final comments and conclusions.

(1) In what concerns option (B), the starting vector $\mathsf{x}_n^{(0)} \in \mathbf{C}^{n\times 1}$ should be chosen such that, for instance, the following Newton assumptions are satisfied: $\mathsf{I}_n - \mathsf{C}_n^{(0)}$ is invertible, and there exists m_n such that

$$\|(\mathsf{I}_n - \mathsf{C}_n^{(0)})^{-1}\| \leq m_n, \quad \nu_n := m_n^2 L_n \|\mathsf{F}_n(\mathsf{x}_n^{(0)})\| \leq \frac{1}{2},$$

L_n being a Lispchitz constant for the Jacobian matrix $\mathsf{x} \mapsto \mathsf{F}_n'(\mathsf{x})$ in some closed ball centered at $\mathsf{x}_n^{(0)}$. In that case, the convergence of option (B) is quadratic:

$$\|\mathsf{x}_n^{(k+1)} - \mathsf{x}_n^{(\infty)}\| \leq \frac{m_n L_n}{2(1 - 2\nu_n)} \|\mathsf{x}_n^{(k)} - \mathsf{x}_n^{(\infty)}\|^2,$$

but the limit of the process, as $k \to \infty$, leads to the n-order approximate solution ψ_n and not to the exact solution φ.

(2) Option (A) has a linear convergence but its limit function is φ, the exact solution.

(3) Since the convergence of option (A) is linear, why not conceive a cheaper fixed slope style iteration:

$$\xi_n^{(k+1)} = \xi_n^{(k)} - \Sigma_n(\xi_n^{(0)})F(\xi_n^{(k)}), \quad \xi_n^{(0)} \in B_{\varrho_n}(\varphi),$$

for some conveniently chosen radius ϱ_n? The following result goes in that direction.

Theorem 5.1. *Suppose that F, $\mathcal{O}$, $\varphi \in \mathcal{O}$, $\xi_n^{(0)} \in \mathcal{O}$ $\mu_n > 0$, $r > 0$, $\ell > 0$ and $\alpha \in]0,1]$ are such that:*

(a) $F(\varphi) = 0$.

(b) $\|\Sigma_n(\xi_n^{(0)})\| \leq \mu_n$.

(c) *The closed ball* $B_r(\varphi)$ *is included in the open set* $\mathcal{O}$, *and* $F' : \mathcal{O} \to \mathcal{L}(\mathcal{X})$ *is* (ℓ, α)*-Hölder continuous on* $B_r(\varphi)$.

(d) *The starting approximation* $\xi_n^{(0)}$ *and* $\gamma_n \in]0, 1[$ *satisfy*

$$\xi_n^{(0)} \in B_{\varrho_n}(\varphi), \quad \|I - \Sigma_n(\xi_n^{(0)})F'(\xi_n^{(0)})\| \leq \gamma_n,$$

where

$$\varrho_n := \min\left\{r, \left(\frac{1-\gamma_n}{3\ell\mu_n}\right)^{1/\alpha}\right\}.$$

Then, for all k, $\xi_n^{(k)} \in B_{\varrho_n}(\varphi)$, *and*

$$\|\xi_n^{(k)} - \varphi\| \leq (3\mu_n\ell\varrho_n^\alpha + \gamma_n)^k \longrightarrow 0 \quad \text{as } k \to \infty.$$

Proof. The keys are the identity

$$\begin{aligned}\xi_n^{(k+1)} - \varphi = {} & [I - \Sigma_n(\xi_n^{(0)})F'(\xi_n^{(0)})](\xi_n^{(k)} - \varphi) \\ & + \Sigma_n(\xi_n^{(0)})[F'(\xi_n^{(0)}) - F'(\xi_n^{(k)})](\xi_n^{(k)} - \varphi) \\ & - \Sigma_n(\xi_n^{(0)}) \int_0^1 [F'(\varphi + t(\xi_n^{(k)} - \varphi)) \\ & - F'(\xi_n^{(k)})](\xi_n^{(k)} - \varphi)\, dt,\end{aligned}$$

and the fact that $3\mu_n\ell\varrho_n^\alpha + \gamma_n < 1$. $\square$

(4) To conclude, let us remark that the classical Galerkin approximation to the equation $(I - T(x))v = g$ built upon the projection π_n does not enter the framework of this paper. Indeed, this approximation is defined as the solution of the zero projected residual:

$$\pi_n((I - T(x))v_n - g) = 0.$$

This means that, for n large enough and provided $I - \pi_n T(x)\pi_n$ is invertible, the Galerkin solution is given by

$$v_n = (I - \pi_n T(x)\pi_n)^{-1}\pi_n g.$$

In other words, in the context of this paper, the role of the operator $\Sigma_n(x)$ is played by $(I - \pi_n T(x)\pi_n)^{-1}\pi_n$ which is certainly not invertible. Hence, hypothesis (iii) of Theorem 2.1 and (d) of Theorem 5.1 will never be satisfied.

6. Acknowledgments. Part of this work was done when the third author was on leave at the University of Saint Étienne as Visiting Professor during May 2011.

REFERENCES

1. M. Ahues, *Newton methods with Hölder derivative*, Numer. Func. Anal. Optim. **25** (2004), 1–17.

2. M. Ahues, A. Largillier and B.V. Limaye, *Spectral computations for bounded operators*, Appl. Math. **18**, Chapman and Hall/CRC, Boca Raton, 2001.

3. E.L. Allgower, K. Böhmer, F.A. Potra and W.C. Rheinboldt, *A mesh independence application of the mesh-independence principle for operator equation and their discretization*, SIAM J. Numer. Anal.**23** (1986), 160–169.

4. I.K. Argyros, *Convergence and applications of Newton-type iterations*, Springer Science+Business Media, LLC, New York, 2008.

5. ______, *Some methods for finding errors bounds for Newton-like methods under mild differentiability conditions*, Acta Math. Hung. **61** (1993), 183–194.

6. K.E. Atkinson and J. Flores, *The discrete collocation method for nonlinear integral equation*, IMA J. Numer. Anal. **13** (1993), 195–213.

7. K.E. Atkinson and F.A. Potra, *Projection and iterated projection methods for nonlinear integral equations*, SIAM J. Numer. Anal. **24** (1987), 1352–1373.

8. O. Axelsson, *On mesh independence and Newton type methods*, Appl. Math. **38** (1993), 249–265.

9. D.R. Dellwo and M.B. Friedman, *Accelerated projection and iterated projection methods with applications to nonlinear integral equations*, SIAM J. Numer. Anal. **28** (1991), 236–250.

10. P. Deuflhard and F.A. Potra, *Asymptotic mesh independence of Newton-Garlerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal. **29** (1992), 1395–1412.

11. L. Grammont, *Nonlinear integral equations of the second kind: a new version of Nyström method*, Numer. Funct. Anal. Optim. **34** (2013), 496–515.

12. M.A. Krasnoselskii, G. Vainikko, P.P. Zabreiko, Ya.B. Rutitskii and V.Ya. Stetsenko, *Approximate solution of operator equations*, Noordhoff, Groningen, The Netherlands, 1972.

13. M.A. Krasnoselskii and P.P. Zabreiko, *Geometrical methods of nonlinear analysis*, Springer Verlag, Berlin, 1984.

14. G.M. Vainikko, *Galerkin's perturbation method and general theory of approximate methods for nonlinear equations*, USSR Comp. Math. Math. Phys. **7** (1967), 1–41.

15. C. Villani, *Le fabuleux destin de la méthode de Newton*, Inauguration at The Hadamard Foundation, May 17, 2011.

16. M. Weiser, A. Schiela and P. Deuflhard, *Asymptotic mesh independence of Newton's method revivited*, SIAM J. Numer. Anal. **42** (2005), 1830–1845.

Université de Lyon, Institut Camille Jordan, UMR 5208, 23 rue du Dr Paul Michelon, 42023 Saint-Étienne Cedex 2, France
Email address: **laurence.grammont@univ-st-etienne.fr**

Université de Lyon, Institut Camille Jordan, UMR 5208, 23 rue du Dr Paul Michelon, 42023 Saint-Étienne Cedex 2, France
Email address: **mario.ahues@univ-st-etienne.fr**

CMUP – Centro de Matemática and Faculdade Engenharia da Universidade Porto, Rua Roberto Frias, 4200-465 Porto, Portugal
Email address: **falmeida@fe.up.pt**