



CERN - DATA HANDLING DIVISION
DD/85/5
July 1985

ON AN ITERATIVE METHOD FOR A CLASS OF
INTEGRAL EQUATIONS OF THE FIRST KIND

H.N. Mülthei

Fachbereich Mathematik, Johannes Gutenberg-Universität Mainz,
Fed. Rep. Germany^{*)}

B. Schorr

CERN, Geneva, Switzerland

G E N E V A

1985

^{*)} Visitor at CERN, Geneva, Switzerland

ABSTRACT

In this paper, we investigate an iterative method which has been proposed [1] for the numerical solution of a special class of integral equations of the first kind, where one of the essential assumptions is the positivity of the kernel and the given right-hand side. Integral equations of this special type occur in experimental physics, astronomy, medical tomography and other fields where density functions cannot be measured directly, but are related to observable functions via integral equations. In order to take into account the non-negativity of density functions, the proposed iterative scheme was defined in such a way that only non-negative solutions can be approximated. The first part of the paper presents a motivation for the iterative method and discusses its convergence. In particular, it is shown that there is a connection between the iterative scheme and a certain concave functional associated with integral equations of this type. This functional can be interpreted as a generalization of the log-likelihood function of a model from emission tomography. The second part of the paper investigates the convergence properties of the discrete analogue of the iterative method associated with the discretized equation. Sufficient conditions for local convergence are given; and it is shown that, in general, convergence is very slow. Two numerical examples are presented.

1. INTRODUCTION

In experimental physics, astronomy, medical tomography, and other fields, density functions (of current, stars, radiation, etc.) are often measured by devices which cannot measure the unknown density f directly. In most cases the function g , which is read out from the device, is related to the unknown density f via an integral equation of the first kind. This is often a linear integral equation, which may always be reduced to the form

$$\int_0^1 k(x,y)f(y) dy = g(x), \quad x \in [0,1], \quad (1.1)$$

with given k and g , and unknown f . In certain applications (for instance in high-energy physics or in tomography) x and y can have more than one dimension, and the functions k and f can have singularities. Physicists often call the kernel k in (1.1) the apparatus function, while in engineering and tomography the function k is called the point response function. It is well known that the equation (1.1) is, in general, an ill-posed problem. This means that small changes in g may cause large changes in f .

There exists an extensive literature on the solution of general linear integral equations of the first kind. The case of density measurement, however, is special insofar as the functions f , g , and k in (1.1) are non-negative. An attempt to take this property into account was made by Kondor [1], who published an iterative method for the solution of (1.1). Kondor gave a purely formal description of the iterative method and presented a few examples to illustrate how the method works. A discrete version of this iterative method had already been used by Shepp and Vardi [2]. However, this was not mentioned in Kondor's paper. Shepp and Vardi used the iterative scheme in order to solve a constrained maximization problem, which they obtained from the maximum likelihood estimation of finitely many unknown parameters of Poisson-distributed random variables. The unknown parameters were discretized values of an unknown emission density which was the solution of a two-dimensional integral equation of the form (1.1). The underlying physical problem arises in emission tomography, where counts observed by a detector can be used to approximate the right-hand side of equation (1.1). Shepp and Vardi identified the (discrete) iterative method as a special case of a general class of algorithms which had been introduced and studied by Dempster, Laird and Rubin [3] when they investigated maximum likelihood estimators from incomplete data. Neither Kondor nor Shepp and Vardi studied the convergence of the iterative method. However, Shepp and Vardi provided a motivation for the (discrete) iterative scheme by showing that, in the case of convergence, the

limiting point maximizes the log-likelihood function which belongs to the problem. In order to show this, they refer to the Kuhn-Tucker theorem of concave programming and to results from Dempster et al. [3].

In this paper, we study the convergence properties of the iterative method in both the continuous and the discrete case. In section 2 we formulate the continuous iterative method as given by Kondor [1]. Then, in section 3, using ideas from Shepp and Vardi and from Dempster et al. (which they applied to the case of finitely many unknown parameters), we give a new motivation for the (continuous) iterative method. Convergence properties are investigated in sections 4 to 6. Finally, in section 7, some numerical tests are presented in order to illustrate how the method works.

2. THE ITERATIVE METHOD

In our own notation, the essential assumptions under which Kondor presented the iterative method are the following.

- (a) The kernel k is non-negative and continuous on $[0,1] \times [0,1]$ except possibly at isolated integrable singularities and discontinuities.
- (b) For given positive $g \in C[0,1]$ the equation (1.1) has a positive solution $f^* \in C[0,1]$.

For the theoretical treatment, in order to avoid unnecessary complications, we generally assume that:

- (α) $k \in C([0,1]^2)$ and positive,
- (β) $g \in C([0,1])$ and positive.

The requirements for f^* are always explicitly given in the following. Occasionally, for instance, we shall also admit that equation (1.1) has a solution $f^* \in C[0,1]$ which takes arbitrary values on $[0,1]$. The investigations which follow show that the general assumptions (α) and (β) can be weakened. However, in this paper, we do not aim at more generality.

The iterative method is then defined by

$$f_{n+1} = G(f_n) , \quad n \in \mathbb{N}_0 , \quad (2.1)$$

where

$$G(f)(y) := f(y)T(f)(y) ,$$

$$T(f)(y) := \frac{1}{a(y)} \int_0^1 \frac{g(x)k(x,y)}{F(x)} dx , \quad F(x) := \int_0^1 k(x,y)f(y) dy ,$$

$$a(y) := \int_0^1 k(x,y) dx, \quad y \in [0,1],$$

$$f, f_0 \in \mathcal{K} := \{h \in C[0,1] \setminus \{0\} : h(x) \geq 0 \text{ for all } x \in [0,1]\}.$$

Obviously we have $G : \mathcal{K} \rightarrow \mathcal{K}$. A function from \mathcal{K} can also vanish on sub-intervals of $[0,1]$.

Trivially, a solution $f^* \in \mathcal{K}$ of (1.1) is also a function from the set

$$\mathcal{C} := \{f \in \mathcal{K} : T(f)(y) \leq 1, y \in [0,1], \text{ equality holds for } y \text{ with } f(y) > 0\}.$$

The set \mathcal{C} corresponds to the Kuhn-Tucker conditions for the discrete case in [2], p. 119. It is easily shown that, for $y_0 \in [0,1]$ and $f \in \mathcal{C}$ with $f(y_0) = 0$ and $T(f)(y_0) < 1$, there exists $[a,b] \subset [0,1]$, $a < b$, $y_0 \in [a,b]$, such that f vanishes on $[a,b]$.

For $f \in \mathcal{C}$ the fixed point equation

$$f = G(f) \tag{2.2}$$

holds. Later, we shall also consider discrete versions of equations (1.1) and (2.1).

Kondor's [1] justification for using this iterative method seems to be essentially the success which he claims to have had in practical applications. In the following section we present a motivation for the method based on optimization principles, making use, to some extent, of ideas of Shepp and Vardi [2] and Dempster et al. [3].

3. NEW MOTIVATION FOR THE ITERATIVE METHOD

Later, in section 4, theorem 8, we shall show that the limiting function of the iterative method maximizes a certain constrained concave functional. In order to motivate the choice of the functional we use ideas which Shepp and Vardi [2] applied to a discrete model of emission tomography. In their model, photons are emitted from boxes and detected, with some probability, by a set of detectors. It is assumed that the photons are emitted according to a Poisson distribution with unknown mean value f_j for box j . The mean values g_i of the counts observed by the detectors are supposed to be related to the values f_j via the system of equations

$$g_i = \sum_j k_{ij} f_j ,$$

where the k_{ij} are known transition probabilities. The unknown values f_j are finally taken to be those values which maximize the log-likelihood function belonging to the observed counts in the different detectors.

In our case we may proceed as follows. Let

$$0 = x_1 < x_2 < \dots < x_{n-1} < x_n = 1 .$$

We set

$$g_i := g(x_i) = \int_0^1 k(x_i, y) f(y) dy , \quad i = 1(1)n, \quad (3.1)$$

$$r_n := \sum_{i=1}^n g_i = \sum_{i=1}^n \int_0^1 k(x_i, y) f(y) dy . \quad (3.2)$$

The relation (3.1), and hence (3.2), is a necessary condition for f if f is a solution of the integral equation (1.1). In the following, we will therefore consider (3.2) as a constraint on f .

Let P_i , $i = 1(1)n$, be n independent Poisson-distributed random variables with mathematical expectation $E(P_i) = g_i$. If p_i is a realization of P_i , $i = 1(1)n$, the corresponding likelihood function is given by

$$\ell_n(f) := \prod_{i=1}^n \exp(-g_i) g_i^{p_i} / p_i! , \quad (3.3)$$

where g_i is given by the right-hand side of (3.1). Taking the logarithm of both sides of (3.3), and leaving out unimportant constants, we obtain the log-likelihood function

$$\Lambda_n(f) := \sum_{i=1}^n p_i \ln \left(\int_0^1 k(x, y) f(y) dy \right) . \quad (3.4)$$

Maximizing Λ_n for all $f \in \mathcal{K}$ which satisfy the constraint (3.2) yields a maximum likelihood estimate of f .

The p_i in (3.4) are estimates of the values g_i , $i = 1(1)n$. We may therefore replace p_i in (3.4) by g_i . As $n \rightarrow \infty$ we may replace the sums in (3.2) and (3.4) by integrals to obtain formally

$$r := \int_0^1 g(x) dx = \int_0^1 \int_0^1 k(x,y)f(y) dy dx = \int_0^1 a(y)f(y) dy, \quad (3.5)$$

$$\Lambda(f) := \Lambda_{g,k}(f) := \int_0^1 g(x) \ln \left(\int_0^1 k(x,y)f(y) dy \right) dx. \quad (3.6)$$

We may consider Λ as the continuous version of the log-likelihood function Λ_n in (3.4). Λ is a functional for fixed k and g which is concave due to the concavity of the logarithm. The equation (3.5) is a constraint on $f \in \mathcal{K}$. In order to maximize (3.6) on \mathcal{K} under this constraint, one might think of using methods of concave programming as Shepp and Vardi did. They, of course, had a problem involving finitely many variables. However, maximizing concave functionals under equality and inequality constraints, by using methods of concave programming, implies the known complications of the choice of proper function spaces. We succeed in avoiding these difficulties by transferring results and ideas from Dempster et al. [3] for finite dimensional problems to our case of infinite dimensions. It turns out that all results connected with the maximization of $\Lambda(f)$ can be obtained essentially by using Jensen's inequality. We set

$$\mathcal{K}_r := \left\{ h \in \mathcal{K} : \int_0^1 a(y)h(y) dy = r \right\}.$$

Note that $G : \mathcal{K} \rightarrow \mathcal{K}_r$ (see theorem 5) and $\mathcal{C} \subset \mathcal{K}_r$ because of (2.2). We prove the following theorem:

Theorem 1:

Let $f^* \in \mathcal{C}$. Then, for all $f \in \mathcal{K}_r$, the inequalities

$$\Lambda(f) \leq \Lambda(f^*) \leq \int_0^1 g(x) \ln g(x) dx$$

hold; i.e. every $f^* \in \mathcal{C}$ is a global solution of the Maximization Problem MP: maximize $\Lambda(f)$ for $f \in \mathcal{K}_r$.

In particular, for a solution $f^* \in \mathcal{K}$ of (1.1) we have

$$\Lambda(f^*) = \int_0^1 g(x) \ln g(x) dx .$$

Proof: Let α and β be two non-negative integrable functions defined on $[0,1]$ with

$$\int_0^1 \alpha(t) dt = 1 .$$

Then, as a consequence of Jensen's inequality (Roberts and Varberg [4], p. 193), we have

$$\ln \left(\int_0^1 \alpha(t) \beta(t) dt \right) \geq \int_0^1 \alpha(t) \ln \beta(t) dt . \quad (3.7)$$

Using (3.7) we obtain with $F^*(x) := \int_0^1 k(x,y) f^*(y) dy$

$$\begin{aligned} \Lambda(f^*) - \Lambda(f) &= -\gamma \int_0^1 \frac{g(x)}{\gamma} \ln \frac{F(x)}{F^*(x)} dx \geq -\gamma \ln \left(\frac{1}{\gamma} \int_0^1 g(x) \frac{F(x)}{F^*(x)} dx \right) \\ &\geq -\gamma \ln \left(\frac{1}{\gamma} \int_0^1 \alpha(y) f(y) T(f^*)(y) dy \right) \geq -\gamma \ln 1 = 0 , \end{aligned}$$

due to $T(f^*)(y) \leq 1$ for $y \in [0,1]$, and

$$\begin{aligned} \int_0^1 g(x) \ln g(x) dx - \Lambda(f) &= -\gamma \int_0^1 \frac{g(x)}{\gamma} \ln \left(\frac{\int_0^1 k(x,y) f(y) dy}{g(x)} \right) dx \\ &\geq -\gamma \ln \left(\frac{1}{\gamma} \int_0^1 \alpha(y) f(y) dy \right) = -\gamma \ln 1 = 0 . \end{aligned}$$

This completes the proof.

Remark: The proof shows that theorem 1 is also valid if \mathcal{K}_γ is replaced by the larger class

$$\tilde{\mathcal{K}}_\gamma := \left\{ f \in \mathcal{K} : \int_0^1 \alpha(y) f(y) dy \leq \gamma \right\}$$

and if we only require that $f^* \in \tilde{\mathcal{K}}_\Gamma$ and $T(f^*)(y) \leq 1$ for $y \in [0,1]$. This apparently weaker requirement for f^* is, however, equivalent to $f^* \in \mathcal{C}$. In particular, f^* is on the 'boundary' of $\tilde{\mathcal{K}}_\Gamma$, namely in \mathcal{K}_Γ .

This statement may be proved as follows: assume that f^* satisfies the apparently weaker assumption. It then follows that

$$\Gamma = \int_0^1 a(y)f^*(y)T(f^*)(y) dy \leq \int_0^1 a(y)f^*(y) dy \leq \Gamma ,$$

which implies that $f^* \in \mathcal{K}_\Gamma$. Clearly, this inequality leads to the contradiction $\Gamma < \Gamma$ if we have $T(f^*)(y_0) < 1$ for a $y_0 \in [0,1]$ with $f^*(y_0) > 0$. Hence, $f^* \in \mathcal{C}$. In this case, however, the weaker assumption for f^* is trivially satisfied.

In order to motivate the iterative method of section 2 we proceed as follows. By analogy with Dempster et al. [3], we split the functional Λ into a difference of two functionals Q and H with $Q, H : \mathcal{K}^+ \times \mathcal{K} \rightarrow \mathbb{R}$. We denote by \mathcal{K}^+ and \mathcal{K}_Γ^+ the subset of all positive functions from \mathcal{K} and \mathcal{K}_Γ , respectively. We define, for $\psi \in \mathcal{K}^+$ and $f \in \mathcal{K}$:

$$Q(\psi, f) := \int_0^1 g(x) \int_0^1 \ln [k(x, y)\psi(y)]K(f; x, y) dy dx ,$$

$$H(\psi, f) := \int_0^1 g(x) \int_0^1 \ln [K(\psi; x, y)]K(f; x, y) dy dx ,$$

with

$$K(h; x, y) := \frac{k(x, y)h(y)}{\int_0^1 k(x, z)h(z)dz} , \quad x, y \in [0,1] , \quad h \in \mathcal{K} ,$$

for which

$$\int_0^1 K(h; x, y) dy = 1 ; \tag{3.8}$$

i.e. $K(h; x, \cdot)$ is a density function for $x \in [0,1]$ and $h \in \mathcal{K}$.

It is easy to verify that

$$\Lambda(\psi) = Q(\psi, f) - H(\psi, f) \quad \text{for} \quad \psi, f \in \mathcal{K}^+ . \tag{3.9}$$

Theorem 2:

For arbitrary ψ , $f \in \mathcal{K}^+$ the inequality

$$D_H(f, \psi) := H(f, f) - H(\psi, f) \geq 0$$

holds. This means that f maximizes $H(., f)$ on \mathcal{K}^+ .

Proof: We apply the inequality (3.7) to obtain with (3.8):

$$\begin{aligned} D_H(f, \psi) &= - \int_0^1 g(x) \int_0^1 \ln \left(\frac{K(\psi; x, y)}{K(f; x, y)} \right) K(f; x, y) dy dx \\ &\geq - \int_0^1 g(x) \ln \left(\int_0^1 K(\psi; x, y) dy \right) dx = 0 . \end{aligned} \quad (3.10)$$

Theorem 3:

For arbitrary $f \in \mathcal{K}^+$ and $\psi \in \mathcal{K}_r^+$ the inequality

$$D_Q(f, \psi) := Q(G(f), f) - Q(\psi, f) \geq 0$$

holds. This means that $G(f)$ maximizes $Q(., f)$ on \mathcal{K}_r^+ . In particular, the inequality

$$Q(f, f) \leq Q(G(f), f) \quad \text{for } f \in \mathcal{K}_r^+$$

holds.

Proof: Applying the inequality (3.7) we obtain:

$$\begin{aligned} D_Q(f, \psi) &= - \int_0^1 g(x) \int_0^1 \ln \left(\frac{\psi(y)}{G(f)(y)} \right) K(f; x, y) dy dx \\ &= - \int_0^1 \ln \left(\frac{\psi(y)}{G(f)(y)} \right) a(y) G(f)(y) dy \\ &\geq - r \ln \left(\frac{1}{r} \int_0^1 a(y) \psi(y) dy \right) = 0 . \end{aligned} \quad (3.11)$$

Remark: The iterative scheme (2.1) yields, at each iteration step, a function $f_{n+1} = G(f_n)$ which maximizes the functional $Q(., f_n)$ on \mathcal{K}_r^+ . The iterative method can, therefore, be interpreted as a generalized EM algorithm in the sense of Dempster et al. [3]. Each iteration consists of two substeps: an estimation step (E) and a maximization step (M). In our case these substeps coalesce. Shepp and Vardi [2] already found this in their specialized discrete case. This fact motivates the -- at first glance rather artificial -- iteration scheme introduced by Kondor.

The following theorem shows the essential connection between the iteration scheme (2.1) and the maximization problem MP.

Theorem 4:

For all $f \in \mathcal{K}_r$ the inequality

$$\Lambda(f) \leq \Lambda(G(f)) \quad (3.12)$$

holds. Equality is given in (3.12) if, and only if,

$$D_H(f, G(f)) = D_Q(f, f) = 0 ,$$

where $D_H(f, G(f))$ and $D_Q(f, f)$, for $f \in \mathcal{K}_r \setminus \mathcal{K}_r^+$, are given by (3.10) and (3.11), respectively.

Proof: From (3.9) we have for $f \in \mathcal{K}_r^+$:

$$\Lambda(G(f)) - \Lambda(f) = D_Q(f, f) + D_H(f, G(f)) \geq 0 , \quad (3.13)$$

where the inequality follows immediately from theorems 2 and 3. The proofs of theorems 2 and 3 show that (3.13) is also valid for $f \in \mathcal{K}_r$. The second part of theorem 4 is then trivial.

Corollary:

i) Suppose that f^* is a solution of the maximization problem MP. Then,

$$\Lambda(G(f^*)) = \Lambda(f^*) ,$$

$$D_H(f^*, G(f^*)) = D_Q(f^*, f^*) = 0 .$$

ii) Suppose that f^* is the only solution of the maximization problem MP. Then,

$$G(f^*) = f^* ;$$

in particular, for $f^* \in \mathcal{K}_r^+$, it follows that $f^* \in \mathcal{C}^+ := \mathcal{C} \cap \mathcal{K}^+$.

- iii) Suppose that Λ is strongly concave on \mathcal{K}_r . Then, $f^* \in \mathcal{K}_r^+$ is a solution of the maximization problem MP if, and only if, $f^* \in \mathcal{G}^+$.

Proof:

- i) From the assumption and from (3.12) it follows that

$$\Lambda(f^*) \leq \Lambda(G(f^*)) \leq \Lambda(f^*). \quad (3.14)$$

The second part of theorem 4 then yields the proof of point (i).

- ii) According to (3.14) the function $G(f^*)$ is also solution of MP from which the statement follows immediately.
- iii) If $f^* \in \mathcal{G}^+$, it then follows that $f^* \in \mathcal{K}_r^+$, and theorem 1 says that f^* maximizes Λ on \mathcal{K}_r . By assumption, the functional Λ is strongly concave on \mathcal{K}_r , which implies the uniqueness of f^* if f^* maximizes the problem MP. From point (ii) of this corollary therefore it follows that $f^* \in \mathcal{G}^+$.

Remark: It can be shown that Λ is strongly concave on \mathcal{K}_r as well as on \mathcal{K} if, and only if, the equation (1.1) has a unique solution in $C[0,1]$.

The results of this section clearly show that it is worthwhile to investigate further the properties of the iterative method (2.1). In the next two sections we shall investigate the convergence and some other properties of the method. The continuous case is considered in section 4, the discrete one in section 5.

4. ON THE CONVERGENCE OF THE CONTINUOUS CASE

We begin with some characteristics of the operator G . These are summarized in the following theorem:

Theorem 5:

Let $f^* \in \mathcal{G}$ and $f \in \mathcal{K}$. Then, G has the following properties:

- i) $G(cf) = G(f)$, for all $c \in \mathbb{R}_+$.
- ii) $G : \mathcal{K} \rightarrow \mathcal{K}_r$;
in particular, $f^* - G(f)$ changes sign at least once in $[0,1]$ for $f^* \neq G(f)$.
- iii) If, for $f \neq f^*$, either
 $f(y) \leq f^*(y)$ or $f(y) \geq f^*(y)$ for all $y \in [0,1]$,
then
 $f(y) \leq G(f)(y)$ or $f(y) \geq G(f)(y)$, resp., for all $y \in [0,1]$,
where equality only holds for all y with $f(y) = 0$.

Proof: Point (i) and the first part of point (ii) follow immediately from the structure of G . Owing to point (ii) and $\mathcal{G} \subset \mathcal{K}_r$ we have

$$\int_0^1 a(y) \left(f^*(y) - G(f)(y) \right) dy = 0 ,$$

which completes the proof of point (ii).

In order to show point (iii), assume that $0 < f(y) \leq f^*(y)$ for a $y \in [0,1]$. It then follows that

$$G(f)(y) = f(y)T(f)(y) > f(y)T(f^*)(y) = f(y) .$$

For $f(y) = 0$ the inequalities are trivial. The second part of (iii) is obtained analogously.

In order to investigate convergence of the iteration scheme (2.1), one would like to know whether the operator G is contracting. In this case, the first thing one usually does is to estimate its Fréchet derivative at a fixed point.

For any linear function mapping A we shall use the following notation: $Af(y) := (Af)(y)$, $y \in [0,1]$.

Lemma 1: The Fréchet derivative $G'(f)$, $f \in \mathcal{K}$, with respect to L_1 is given, for almost all $y \in [0,1]$, by

$$G'(f)h(y) = T(f)(y)h(y) - \frac{f(y)}{a(y)} \int_0^1 \left(\int_0^1 \frac{g(x)k(x,y)k(x,z)}{F^2(x)} dx \right) h(z) dz \quad (4.1)$$

for $h \in L_1[0,1]$.

In particular, for $f^* \in \mathcal{G}^+$,

$$G'(f^*)h = (I - L)h , \quad (4.2)$$

where I is the identity mapping and

$$Lh(y) := \frac{f^*(y)}{a(y)} \int_0^1 \ell(y,z)h(z) dz ,$$

$$\ell(y,z) := \int_0^1 \frac{g(x)k(x,y)k(x,z)}{(F^*(x))^2} dx .$$

In particular, for a solution $f^* \in \mathcal{K}$ of (1.1),

$$\ell(y, z) = \int_0^1 \frac{k(x, y)k(x, z)}{g(x)} dx.$$

Remark: Note that

i) $G'(f)f = 0$ for $f \in \mathcal{K}$,

ii) $Lf^* = f^*$ for $f^* \in \mathcal{G}$, i.e., f^* is an eigenfunction of L .

Proof: Consider

$$G(f+h)(y) - G(f)(y) = \frac{1}{a(y)} \int_0^1 g(x)k(x, y)q(x, y) dx$$

with

$$q(x, y) := \frac{f(y) + h(y)}{F(x) + H(x)} - \frac{f(y)}{F(x)},$$

$$H(x) := \int_0^1 k(x, y)h(y) dy.$$

For h with $\|h\|$, sufficiently small the identity

$$\begin{aligned} q(x, y) &= \frac{h(y)}{F(x)} - \frac{f(y)}{F^2(x)} H(x) + h(y) \left[\frac{1}{F(x) + H(x)} - \frac{1}{F(x)} \right] \\ &\quad + f(y) \left[\frac{1}{F(x) + H(x)} - \frac{1}{F(x)} + \frac{H(x)}{F^2(x)} \right] \\ &= \frac{h(y)}{F(x)} - \frac{f(y)}{F^2(x)} H(x) + h(y)O(H(x)) + f(y)O(H^2(x)) \end{aligned}$$

holds for almost all $y \in [0, 1]$. Denoting the right-hand side of (4.1) by $Ah(y)$, it can easily be shown that

$$|G(f+h)(y) - G(f)(y) - Ah(y)|$$

$$\leq c_1 \left(|h(y)| \frac{\int_0^1 k(x,z) |h(z)| dz}{F^2(x)} + |f(y)| \frac{\left(\int_0^1 k(x,z) |h(z)| dz \right)^2}{F^3(x)} \right) \\ \leq c_2 \|h\|_1 \left(|h(y)| + \|h\|_1 \right) \quad (4.3)$$

for almost all $y \in [0,1]$ and for $\|h\|_1$ sufficiently small, using the fact that the functions a , f , and k are continuous. It follows that

$$\|G(f+h) - G(f) - Ah\|_1 \leq c_3 \|h\|_1.$$

This implies that $A = G'(f)$. Equation (4.2) trivially follows from (4.1).

Remarks:

- i) The assumption $f^* \in \mathcal{G}^*$ is essential for the representation (4.2). In section 5, we shall also treat the discrete case which is analogous to the case of $f^* \in \mathcal{G}$.
- ii) The proof of lemma 1 shows that the representations (4.1) and (4.2), respectively, are also valid, for instance, for the Banach spaces $C[0,1]$ and $L_p[0,1]$, $p \in (1, \infty]$. This is easily seen by the following. Using the triangle inequality we obtain with (4.3)

$$\|G(f+h) - G(f) - Ah\|_p \leq \text{const.} \|h\|_1 \left(\|h\|_p + \|h\|_1 \right) \\ \leq \text{const.} \|h\|_p^2,$$

due to $\|h\|_1 \leq \text{const.} \|h\|_p$ for all $h \in L_p$, $p \in [1, \infty]$. The proof for $C[0,1]$ follows analogous lines.

Application of the triangle inequality to the right-hand side of equation (4.2) is of no help for showing that G is contracting in a neighbourhood of f^* . The question is whether generally there exists a Banach space \mathcal{B} with $\mathcal{K} \subset \mathcal{B}$ and $\|G'(f^*)\| < 1$, where $\|\cdot\|$ is the operator norm. Using the eigenvalues of the linear operator L , we shall show that this is not possible for $f^* \in \mathcal{G}^*$.

Theorem 6:

The eigenvalues of L lie in $[0,1]$. They are at most countable and only 0 can be a limiting point. Thereby 1 is always an eigenvalue, whereas 0 is an eigenvalue or at least a limiting point of eigenvalues.

Addendum: L has finitely many eigenvalues if, and only if, the kernel ℓ is degenerate; in particular, 0 is an eigenvalue in this case.

Corollary: For every Banach space B with $\mathcal{K} \subset B$ and $G'(f^*): B \rightarrow B$ the inequality $\|G'(f^*)\| \geq 1$ holds for $f^* \in \mathcal{G}^*$.

Remark: For instance, $C[0,1]$ and $L_p[0,1]$, $p \in [1, \infty]$, are trivially such Banach spaces.

Proof: If $Lh = \lambda h$, $\lambda \neq 0$, and $h \neq 0$, then $h \in C[0,1]$, and it follows, by changing the order of integration, that

$$\begin{aligned} |\lambda| \int_0^1 a(y) |h(y)| dy &= \int_0^1 a(y) |L(y)| dy \leq \int_0^1 a(z) T(f^*)(z) |h(z)| dz \\ &= \int_0^1 a(z) |h(z)| dz > 0, \end{aligned} \quad (4.4)$$

which implies that $|\lambda| \leq 1$, where 1 is an eigenvalue. The spectrum of L is identical to the spectrum of the symmetric linear operator

$$S := W L W^{-1},$$

with

$$W : h \mapsto \sqrt{\frac{a}{f^*}} h, \quad h \in C[0,1],$$

where the kernel of S is given by

$$s(x,y) := \sqrt{\frac{f^*(x)f^*(y)}{a(x)a(y)}} \ell(x,y), \quad x,y \in [0,1].$$

This kernel is positive semidefinite because it is symmetric and, for all $h \in C[0,1]$, the relation

$$\int_0^1 \int_0^1 s(x,y) h(x) h(y) dx dy = \int_0^1 \frac{g(x)}{(f^*(x))^2} \left(\int_0^1 \sqrt{\frac{f^*(x)}{a(x)}} k(z,x) h(x) dx \right)^2 dz \geq 0$$

holds. Therefore, all eigenvalues of S are in $[0, 1]$. According to Tricomi [5], p. 105, the eigenvalues of S are countable, and at most zero is a limiting point. In addition, S has finitely many eigenvalues if, and only if, its kernel is degenerate, which is equivalent to the degeneracy of the kernel ϱ . In the latter case 0 clearly is an eigenvalue.

The proof of the corollary follows from theorem 6 and the fact that the eigenfunctions h of L , which belong to the eigenvalues $\lambda \neq 0$, are elements of $C[0, 1] \subset \mathcal{B}$ such that with the norm $||\cdot||$ of \mathcal{B} we have

$$||G'(f^*)|| := \sup_{f \in \mathcal{B} \setminus \{0\}} \frac{||G'(f^*)f||}{||f||} \geq \frac{||(I - L)h||}{||h||} = 1 - \lambda.$$

If the kernel ϱ is degenerate, a continuous eigenfunction to the eigenvalue $\lambda = 0$ trivially exists. This completes the proof.

Remarks:

i) Let $||L||_{w,1}$ be the L_1 -norm of L with respect to the weight function w . Note that from (4.4) it easily follows that $||L||_{a,1} = 1$, since for $h := 1/\int_0^1 a(y)dy$ equality holds in (4.4).

ii) For the continuous case, the question whether the iterative method is convergent is still open. We had hoped to prove statements analogous to those given by Dempster et al. for EM algorithms (see [3], theorems 2 and 3). But careful checking shows that the proof of their theorem 2 is incorrect. (Inequality (3.13) in [3] is not true in general, as can be shown by using $\varphi^{(j)} := 1 + 1/2 + \dots + 1/j$, $j \in \mathbb{N}$, in the notation of Dempster et al. as a counter-example.) Therefore, it is not excluded that theorem 2 and theorem 3 (which is based on theorem 2) of Dempster et al. do not hold.

So far, we have for the continuous case the following result.

Theorem 7:

Let $f_{n+1} = G(f_n)$, $n \in \mathbb{N}_0$, $f_0 \in \mathcal{K}^+$. Then,

- i) $\Lambda(f_n)$, $n \in \mathbb{N}$, converges monotonically to some positive number for $n \rightarrow \infty$.
- ii) $\int_0^1 f_n(y) \left(1 - T(f_n)(y)\right)^2 dy \xrightarrow{n \rightarrow \infty} 0$.

Corollary: If, in addition to the assumptions of theorem 7,

$$f_n(y) \geq \varepsilon \quad \text{for } y \in J := [\alpha, \beta] \subset [0, 1], \quad \alpha < \beta, \quad n \in \mathbb{N}_0,$$

then $T(f_n)|_J$ [the restriction of $T(f_n)$ to J] converges uniformly to $1|_J$ for $n \rightarrow \infty$.

Remarks:

i) Note that in theorem 7 with respect to point (ii) we have

$$\int_0^1 f^*(y) \left(1 - T(f^*)(y)\right)^2 dy = 0$$

for all $f^* \in \mathcal{C}$.

ii) Point (ii) in theorem 7 allows that, for instance, $f_n|_J$ converges uniformly to zero as $n \rightarrow \infty$ even if $T(f_n)|_J$ does not converge to a non-negative function which is bounded by 1.

In order to prove theorem 7 we need the following

Lemma 2: Let $k_{\max} := \|k\|_{\infty}$, $k_{\min} := \min_{x,y \in [0,1]} k(x,y)$.

Then, for $f \in \mathcal{K}_r$,

$$i) \quad r/k_{\max} \leq \int_0^1 f(x) dx \leq r/k_{\min},$$

$$ii) \quad c_1 := (k_{\min}/k_{\max})^2 \leq T(f)(y) \leq c_2 := (k_{\max}/k_{\min})^2, \quad y \in [0,1].$$

iii) Let $f_{n+1} = G(f_n)$, $n \in \mathbb{N}_0$, $f_0 \in \mathcal{K}$. Then, the set $M_T := \{T(f_n), n \in \mathbb{N}\}$ is compact in $C[0,1]$.

Proof:

i) Using the mean value theorem for integrals, it follows that there exists a $\xi \in [0,1]$ such that

$$r = \int_0^1 a(y)f(y) dy = a(\xi) \int_0^1 f(y) dy.$$

From this equality the relation (i) follows easily.

ii) From point (i) it follows that

$$T(f)(y) \leq \left(k_{\max} / \left(k_{\min}^2 \int_0^1 f(z) dz \right) \right) \int_0^1 g(x) dx \leq c_2 ,$$

$$T(f)(y) \geq \left(k_{\min} / \left(k_{\max}^2 \int_0^1 f(z) dz \right) \right) \int_0^1 g(x) dx \geq c_1 .$$

iii) The Arzéla-Ascoli theorem yields point (iii) if we can show that

- a) M_T is uniformly bounded,
- b) the $T(f_n)$ are equicontinuous.

(a) follows from point (ii) and (b) is shown as follows. The function k is an element of $C([0,1]^2)$. This implies that for any $\epsilon > 0$ there exists a δ such that

$$\Delta(x; y_1, y_2) := \left| \frac{k(x, y_1)}{a(y_1)} - \frac{k(x, y_2)}{a(y_2)} \right| \leq \epsilon \quad \text{for } |y_1 - y_2| \leq \delta .$$

For $n \in \mathbb{N}$, it follows using (i) that

$$|T(f_n)(y_1) - T(f_n)(y_2)| \leq \int_0^1 \frac{g(x) \Delta(x; y_1, y_2)}{\int_0^1 k(x, z) f_n(z) dz} dx \leq \epsilon k_{\max} / k_{\min}$$

for $|y_1 - y_2| \leq \delta$. This completes the proof of point (iii).

We now prove theorem 7 and its corollary.

i) We have

$$0 < \Lambda(f_n) \leq \Lambda(f_{n+1}) \leq \int_0^1 g(x) \ln g(x) dx , \quad n \in \mathbb{N} ,$$

from theorems 4 and 1. The sequence $\{\Lambda(f_n)\}$ is therefore monotonically increasing and bounded, hence convergent.

ii) Using the relation

$$\ln z = \ln z_0 + (z - z_0)/z_0 - \int_0^1 \frac{(1-t) dt}{(z - t(z - z_0))^2} (z - z_0)^2 , \quad z_0, z \in \mathbb{R}_+ ,$$

we have, for $f \in \mathcal{K}_r^+$,

$$\begin{aligned}
 Q(f, f) &= \int_0^1 g(x) \int_0^1 \ln \left[k(x, y) \left(G(f)(y) + f(y) - G(f)(y) \right) \right] K(f; x, y) dy dx \\
 &= Q(G(f), f) + \int_0^1 g(x) \int_0^1 K(f; x, y) \left[\frac{f(y) - G(f)(y)}{G(f)(y)} \right. \\
 &\quad \left. - \int_0^1 \frac{(1-t) dt}{\left(f(y) - t(f(y) - G(f)(y)) \right)^2} \left(f(y) - G(f)(y) \right)^2 \right] dy dx \\
 &= Q(G(f), f) + \int_0^1 a(y) G(f)(y) \left[\frac{f(y) - G(f)(y)}{G(f)(y)} \right. \\
 &\quad \left. - \int_0^1 \frac{(1-t) dt}{\left(1 - t(1 - T(f)(y)) \right)^2} \left(1 - T(f)(y) \right)^2 \right] dy .
 \end{aligned}$$

Since $\int_0^1 a(y)f(y) dy = \int_0^1 a(y)G(f)(y) dy$, and using lemma 2, point (ii), we obtain

$$\begin{aligned}
 Q(G(f), f) - Q(f, f) &= \int_0^1 a(y) G(f)(y) \left(1 - T(f)(y) \right)^2 \int_0^1 \frac{(1-t) dt}{\left(1 - t(1 - T(f)(y)) \right)^2} dy \\
 &\geq \frac{1}{2 \max(1, \|T(f)\|_\infty^2)} \int_0^1 a(y) G(f)(y) \left(1 - T(f)(y) \right)^2 dy \\
 &\geq \frac{1}{2 \max(1, c_2^2)} \int_0^1 a(y) G(f)(y) \left(1 - T(f)(y) \right)^2 dy .
 \end{aligned}$$

Using theorem 4 and the relation (3.13) we obtain, for $n \in \mathbb{N}$:

$$\begin{aligned} \Lambda(f_{n+1}) - \Lambda(f_n) &\geq Q(f_{n+1}, f_n) - Q(f_n, f_n) \\ &\geq \frac{1}{2 \max(1, c_2^2)} \int_0^1 a(y) f_{n+1}(y) \left(1 - T(f_n)(y)\right)^2 dy . \end{aligned}$$

On the other hand, according to lemma 2, point (ii), we have

$$f_{n+1}(y) = f_n(y) T(f_n)(y) \geq c_1 f_n(y) ,$$

such that

$$\Lambda(f_{n+1}) - \Lambda(f_n) \geq \frac{c_1 k_{\min}}{2 \max(1, c_2^2)} \int_0^1 f_n(y) \left(1 - T(f_n)(y)\right)^2 dy > 0 .$$

The left-hand side of this inequality tends to zero as $n \rightarrow \infty$ according to point (i) of this theorem. This completes the proof of point (ii).

Proof of the corollary: From the assumptions, it follows that

$$\varepsilon \int_0^1 \left(1 - T(f_n)(y)\right)^2 dy \leq \int_0^1 f_n(y) \left(1 - T(f_n)(y)\right)^2 dy , \quad n \in \mathbb{N},$$

which implies that $T(f_n)|_J$ converges to $1|_J$ with respect to the L_2 -norm as $n \rightarrow \infty$. In order to show the uniform convergence of $T(f_n)|_J$ to $1|_J$, we proceed as follows. According to lemma 2, point (iii), $\{T(f_n)|_J\}$ is compact. It therefore contains an L_∞ -convergent subsequence whose limit can only be $1|_J$, since this subsequence converges to $1|_J$ with respect to the L_2 -norm. But the last conclusion also shows that the sequence $\{T(f_n)|_J\}$ has, with respect to the L_∞ -norm, only one limiting point which is $1|_J$. This completes the proof of the corollary.

Theorem 8:

For $f_0 \in \mathcal{K}^+$ let $f_{n+1} = G(f_n)$ converge to $f^* \in \mathcal{K}$ with respect to $L_1[0,1]$ as $n \rightarrow \infty$. Then, $f^* \in \mathcal{G}$; which means that f^* is a solution of the maximization problem MP.

Proof: Because of the inequality $||f^*||_1 - ||f_n||_1 \leq ||f^* - f_n||_1$, there exists, for $\varepsilon > 0$, an integer $n_0 = n_0(\varepsilon)$ such that

$$||f_n||_1 \geq \varepsilon \quad \text{for } n \geq n_0 .$$

Setting $F_n(x) := \int_0^1 k(x,z) f_n(z) dz$, we then obtain

$$\begin{aligned}
|T(f^*)(y) - T(f_n)(y)| &\leq \frac{1}{a(y)} \int_0^1 g(x)k(x,y) \frac{|F^*(x) - F_n(x)|}{F^*(x)F_n(x)} dx \\
&\leq \frac{c_2}{\varepsilon} \int_0^1 \frac{g(x)}{F^*(x)} dx \|f^* - f_n\|_1, \quad n \geq n_0.
\end{aligned}$$

The continuous functions $T(f_n)$ therefore uniformly converge to $T(f^*) \in C[0,1]$ as $n \rightarrow \infty$. Assume that there exists a $y_0 \in [0,1]$ with $T(f^*)(y_0) > 1$. Then, there exist a neighbourhood U of y_0 , a $\delta > 0$, and an integer m_0 such that for $n \geq m_0$ and $y \in U$ the inequality $T(f_n)(y) \geq 1 + \delta$ holds. Because of $f_{n+1} = T(f_n)f_n$ it then follows that

$$f_{m_0+k}(y) \geq (1 + \delta)^k f_{m_0}(y) > 0 \quad \text{for } y \in U, \quad k \in \mathbb{N}_0.$$

This contradicts the assumption that f_n converges as $n \rightarrow \infty$. Note that the positivity of f_{m_0} follows from the fact that $f_0 \in \mathcal{K}^+$. Hence, we have proved that

$$T(f^*)(y) \leq 1 \quad \text{for } y \in [0,1].$$

Now assume that there exists a $y_0 \in [0,1]$ with $T(f^*)(y_0) < 1$. Again, there exist a neighbourhood V of y_0 , a constant q with $0 < q < 1$, and an integer r_0 such that, for $n \geq r_0$ and $y \in V$, the inequalities

$$T(f_n)(y) \leq q < 1$$

hold. From this it can be deduced that

$$f_{r_0+k} \leq q^k f_{r_0}(y) \quad \text{for } y \in V, \quad k \in \mathbb{N}_0.$$

This result allows one to deduce that $f_{n|V}$ converges uniformly on V to $f^*|_V = 0$ as $n \rightarrow \infty$. Hence, if $f^*(y_0) > 0$, the equality $T(f^*)(y_0) = 1$ must hold. This completes the proof.

Remark: Theorem 8 and its proof are analogous to the discrete case considered in [2], p. 120. In particular, the theorem shows that the iterative method (2.1) is primarily oriented to the maximization problem MP; to the integral equation (1.1), however, only as a special case. See also the remarks in [2], p. 120.

Before we treat the discrete case we introduce the following simple example:

$$k(x,y) := \gamma(x-y) + \eta, \quad \eta > 0, \quad x, y \in [0,1],$$

$$\gamma(x) := \begin{cases} 1, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

$$g(x) := x + \eta.$$

For this example only the assumptions (a) and (b) of section 2 are satisfied, and $f^* = 1$ is the only solution. For convenience, we consider the limiting case $\eta = 0$ which yields the Volterra integral equation

$$\int_0^x f(y) dy = x, \quad x \in [0,1],$$

with the only solution $f^* = 1$. We shall refer to this example as our 'model' problem and shall use it to illustrate certain properties of the iterative method. For this example, the following expressions for the operators G , G' , and L introduced above are easily obtained:

$$\begin{aligned} G(f)(y) &= \frac{f(y)}{1-y} \int_y^1 \frac{x}{\int_0^x f(z) dz} dx, \quad f \in \mathcal{K}^+, \\ G'(1) &= I - L, \\ Lh(y) &= \frac{1}{1-y} \int_y^1 \frac{1}{x} \int_0^x h(z) dz dx \\ &= \frac{1}{1-y} \int_0^1 \ell(x,y) h(x) dx, \quad h \in L_1[0,1], \\ \ell(x,y) &:= \begin{cases} \ln y, & x \leq y, \\ \ln x, & x > y. \end{cases} \quad x, y \in [0,1], \end{aligned} \tag{4.5}$$

Remark: Here, G is only defined on \mathcal{K}^+ .

Theorem 9:

The operator L of the model problem defined by (4.5) has the Jacobi polynomials $p_n^{(0,1)}$ with respect to $[0,1]$ as eigenfunctions corresponding to the eigenvalues $\lambda_n = 1/(n+1)^2$, $n \in \mathbb{N}_0$.

Proof: Consider

$$(L - \lambda I)h = 0, \quad (4.6)$$

where $\lambda = 0$ is not an eigenvalue. Multiplying (4.6) by $1-x$ and differentiating, then multiplying the result by x and differentiating again yields, for $\lambda = \lambda_n$, the differential equation

$$x(1-x)h'' + (1-3x)h' + ((n+1)^2 - 1)h = 0,$$

which is the differential equation for the Jacobi polynomial $P_n^{(\alpha, \beta)}$ for $\alpha = 0$ and $\beta = 1$. This completes the proof of theorem 9, since by going backwards (integrating twice and dividing each time by x and $1-x$, respectively), the equation (4.6) for $\lambda = \lambda_n$ can be obtained from the differential equation for the $P_n^{(0,1)}$.

We shall return to the result of theorem 9 when we discuss a discrete version of our model problem.

5. ON THE CONVERGENCE OF THE DISCRETE CASE

In order to perform the iterations (2.1) numerically we discretize by using a quadrature formula. For convenience, we confine ourselves to quadrature formulae Q_m of the following kind:

$$\int_0^1 f(z) dz \approx Q_m(f) := \sum_{i=1}^m c_i f(z_i), \quad m \in \mathbb{N},$$

where

i) $z_i \in [0,1]$ are different for $i \neq j$, $i, j = 1(1)m$,

ii) $c_i > 0$, $i = 1(1)m$.

The dependence of the weights c_i and the knots z_i on m is not explicitly indicated. We use the notation $f_i := f(z_i)$ and $k_{ij} := k(z_i, z_j)$ for $i, j = 1(1)m$. The discretized form of the iteration scheme (2.1) then becomes

$$\varphi^{n+1} = G_m(\varphi^n), \quad n \in \mathbb{N}_0, \quad (5.1)$$

with

$$\varphi^n := (\varphi_i^n)_{i=1(1)m} \in \mathcal{K}_m,$$

$$G_m(\varphi) := \left(G_{m,i}(\varphi) \right)_{i=1(1)m}, \quad \varphi := (\varphi_i)_{i=1(1)m} \in \mathcal{K}_m,$$

$$G_{m,i}(\varphi) := \varphi_i T_{m,i}(\varphi), \quad T_{m,i}(\varphi) := \frac{1}{\sum_{r=1}^m c_r k_{ri}} \sum_{j=1}^m \frac{c_j k_{ji} g_j}{\sum_{s=1}^m c_s k_{js} \varphi_s}, \quad i = 1(1)m,$$

$$\mathcal{K}_m := \{ \varphi \in \mathbb{R}^m \setminus \{0\} : \varphi_i \geq 0, \quad i = 1(1)m \}.$$

Note that n is the iteration index and m the discretization index. By construction, we have $\varphi^n \in \mathcal{K}_m$ if $\varphi^0 \in \mathcal{K}_m$, for $n \in \mathbb{N}$.

The discretized form of \mathcal{G} is

$$\mathcal{G}_m := \{ \varphi \in \mathcal{K}_m : T_{m,i}(\varphi) \leq 1, \quad i = 1(1)m, \text{ equality holds for } i \text{ with } \varphi_i > 0 \}.$$

Let $C_m \varphi^* = \gamma$ be the discretized form of (1.1) with $C_m := (k_{ij} c_j) \in \mathbb{R}^{m,m}$, $\varphi^* := (\varphi_i^*) \in \mathcal{K}_m$, and $\gamma := (g_i) \in \mathbb{R}_+^m$ for $i, j = 1(1)m$. Then, $\varphi^* \in \mathcal{G}_m$ trivially follows, and for $\varphi^* \in \mathcal{G}_m$ we have $\varphi^* = G_m(\varphi^*)$. Note that $T_{m,i}(\varphi) > 0$, $i = 1(1)m$, for $\varphi \in \mathcal{K}_m$.

All statements in sections 3 and 4 concerning the continuous case can be carried over to the discrete case; therefore we do not formulate them explicitly. We only state some discrete analogues which we need in the following.

Theorem 10:

Let $\varphi^* \in \mathcal{G}_m$ and $\varphi \in \mathcal{K}_m$. Then, the mapping G_m has the following properties:

i) $G_m(c\varphi) = G_m(\varphi)$, for all $c \in \mathbb{R}_+$.

$$\text{ii) } \sum_{i=1}^m c_i \sum_{r=1}^m c_r k_{ri} G_{m,i}(\varphi) = \sum_{i=1}^m c_i g_i;$$

in particular, $\varphi^* - G_m(\varphi)$ has positive and negative components for $\varphi^* \neq G_m(\varphi)$.

iii) If, for $\varphi \neq \varphi^*$, either

$$\varphi \leq \varphi^* \quad \text{or} \quad \varphi \geq \varphi^* \quad (\text{by components}),$$

then

$$\varphi \leq G_m(\varphi) \quad \text{or} \quad \varphi \geq G_m(\varphi), \quad \text{respectively,}$$

where equality only holds for all components for which the corresponding components of φ are zero.

The proof follows the same lines as the proof of theorem 5.

The discretized analogue of lemma 1 is given by the following lemma:

Lemma 3: For $\varphi^* \in \mathcal{G}_m$,

$$G'_m(\varphi^*) = \tilde{I}_m - L_m,$$

holds, where

$$\tilde{I}_m := \text{diag} (t_i)_{i=1(1)m}, \quad t_i := T_{m,i}(\varphi^*) \in (0,1], \quad i = 1(1)m,$$

$$L_m := D_m C_m^T B_m C_m,$$

$$D_m := \text{diag} \left(\frac{\varphi_i^*}{c_i \sum_{r=1}^m c_r k_{ri}} \right)_{i=1(1)m}, \quad B_m := \text{diag} \left(\frac{c_i g_i}{\left(\sum_{r=1}^m c_r k_{ir} \varphi_r^* \right)^2} \right)_{i=1(1)m}.$$

At least one of the diagonal elements of \tilde{I}_m is 1, and for $\varphi^* \in \mathcal{G}_m^+ := \mathcal{G}_m \cap \mathbb{R}_+^m$ the matrix \tilde{I}_m is equal to the unit matrix I_m .

The proof of lemma 3 is trivial.

In order to investigate convergence of the iterative scheme (5.1) we are interested in the spectrum of $G'_m(\varphi^*)$, analogously to the continuous case.

Theorem 11:

The eigenvalues of L_m lie in $[0,1]$ with 1 being an eigenvalue.

Proof: Theorem 11 can be proved analogously to the proof of theorem 6 in section 4. We first assume that $\varphi^* \in \mathcal{G}_m^+$. The spectrum of L_m is the same as the spectrum of

$$S_m := W_m L_m W_m^{-1} = (C_m W_m^{-1})^T B_m C_m W_m^{-1} = S_m^T,$$

with $W_m := D_m^{-1/2}$. B_m is a diagonal matrix with positive diagonal elements. Hence, S_m is positive semidefinite, which implies that the eigenvalues of S_m are real and non-negative.

Assume now that $\varphi^* \in \mathcal{C}_m \setminus \mathcal{R}_+^m$. This implies that W_m does not exist. We then proceed as follows. For all $i \in A := \{j : \varphi_j^* = 0\}$ the i -th row of L_m consists of zeros; hence, zero is an eigenvalue. On the other hand, the i -th component of an eigenvector to an eigenvalue different from zero is zero. It is therefore sufficient to consider the matrix $L_{m,r}$ which is derived from L_m by leaving out the i -th row and the i -th column for all $i \in A$. In the same way as for $\varphi^* \in \mathcal{C}_m^*$ it can then be shown that $L_{m,r}$, and hence L_m , have eigenvalues which are non-negative.

For an eigenvalue $\lambda > 0$ of L_m and a corresponding eigenvector $\varphi = (\varphi_i)_{i=1(1)m}$ we have the relation

$$\begin{aligned} \lambda \sum_{i=1}^m \sum_{r=1}^m c_r k_{ri} c_i |\varphi_i| &= \sum_{i=1}^m \sum_{r=1}^m c_r k_{ri} c_i |(L_m \varphi)_i| \\ &\leq \sum_{j=1}^m \sum_{r=1}^m c_r k_{rj} T_{m,j}(\varphi^*) c_j |\varphi_j| \\ &\leq \sum_{j=1}^m \sum_{r=1}^m c_r k_{rj} c_j |\varphi_j| > 0. \end{aligned}$$

From this inequality, it immediately follows that $0 < \lambda \leq 1$. Finally, it is easy to verify that φ^* is an eigenvector corresponding to the eigenvalue 1. This completes the proof of theorem 11.

We set

$$G'_{m,r}(\varphi^*) := I_{m-\mu} - L_{m,r}, \quad \mu := |A|,$$

where A and $L_{m,r}$ were defined in the proof of theorem 11. A is empty if $\varphi^* \in \mathcal{C}_m^*$. In this case $\mu = 0$ and $L_{m,r} = L_m$. We always have $\mu < m$.

Theorem 12:

For the spectral radius $\varrho(G'_m(\varphi^*))$, $\varphi^* \in \mathcal{C}_m$, we have

$$\varrho(G'_m(\varphi^*)) = \max_{i \in A} \left(1 - \lambda_1^{(m)}, t_i \right), \quad \lambda_1^{(m)}, t_i \in [0, 1], \quad i \in A, \quad (5.2)$$

where $\lambda_1^{(m)}$ is the smallest eigenvalue of $L_{m,r}$. In particular, the iterative method (5.1) is (at least) locally linearly convergent to $\varphi^* = G_m(\varphi^*)$ with the

rate of convergence $\rho(G'_m(\varphi^*)) \in [0, 1)$, if $L_{m,r}$ is regular and $t_i < 1$ for $i \in A$. The latter is true, for instance, if K_m is regular and $\varphi^* \in \mathcal{C}_m^+$.

Proof: Trivially, all t_i , $i \in A$, are eigenvalues of $G'_m(\varphi^*)$ since the corresponding rows of L_m only contain zeros. Analogously to the proof of theorem 10, it follows that the remaining eigenvalues of $G'_m(\varphi^*)$ are the eigenvalues of $G'_{m,r}(\varphi^*)$. Since the spectrum of $L_{m,r}$ is in $[0, 1]$ as we know, equation (5.2) is proved. The statement of the convergence follows from Ostrowski's theorem (see Ortega and Rheinboldt [6], pp. 300-301). The set A is empty for $\varphi^* \in \mathcal{C}_m^+$, and, for regular K_m , the matrix S_m is positive definite. This implies that L_m is regular, which completes the proof of theorem 12.

It is expected, taking L_m as an appropriate approximation of L , that $1 - \lambda_1^{(m)}$ tends to 1 as $m \rightarrow \infty$, under certain assumptions. We shall not consider the general case, but treat only our model problem, for which we can compute the rate of convergence explicitly.

6. THE MODEL PROBLEM

For convenience, we use the extended right-side rectangular rule as a quadrature formula. This simple formula already reveals the general behaviour of the rate of convergence. For the model problem, we set

$$Q_m(f) := h \sum_{i=1}^m f(z_i), \quad h := 1/m, \quad z_i = ih, \quad i = 1(1)m.$$

In this case, we have

$$G_{m,i}(\varphi) = \frac{\varphi_i}{m+1-i} \sum_{j=i}^m \frac{j}{\sum_{r=1}^j \varphi_r}, \quad i = 1(1)m, \quad \varphi \in \mathcal{K}_m^+ := \mathbb{R}_+^m. \quad (6.1)$$

Note that from (6.1) we have

$$1 - z_i = \int_0^1 k(x, z_i) dx \simeq h \sum_{j=i}^m 1 = h(m+1-i) = 1 - z_{i-1}, \quad z_0 := 0,$$

$$\int_0^1 k(z_i, y) f(y) dy \simeq h \sum_{j=1}^i f_j, \quad i = 1(1)m.$$

In the first expression the quadrature error is $-h$, and it might therefore seem that the application of the rectangular rule is not appropriate. On the other hand, (6.1) yields for $i = m$ (i.e. for the critical point $y = 1$):

$$G_{m,m}(\varphi) = \frac{\varphi_m}{m} ,$$

$$h \sum_{r=1}^m \varphi_r$$

which is the natural approximation to the expression

$$\lim_{y \rightarrow 1} \frac{f(y)}{1-y} \int_y^1 \frac{x}{\int_0^x f(z) dz} dx = \frac{f(1)}{\int_0^1 f(z) dz} , \quad f \in \mathcal{K} .$$

We have

$$C_m = 1/m \begin{pmatrix} 1 & & 0 \\ \vdots & \ddots & \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{m,m} .$$

For $\varphi^* := (1, 1, \dots, 1)^T \in \mathbb{R}_+^m$ and $\gamma := (i/m)_{i=1(1)m}$, the equality $C_m \varphi^* = \gamma$ holds, and, in particular, $T_m(\varphi^*) = \varphi^* = G_m(\varphi^*)$. C_m is regular. It is easily seen that $\mathcal{G}_m = \{\varphi^*\}$ with A empty. Theorem 10 is still valid if \mathcal{K}_m is replaced by \mathcal{K}_m^+ , and if the last clause in point (iii) is dropped. A more specialized form of the theorems 11 and 12 for the model problem is given by theorem 13 below.

Theorem 13:

For the model problem we have:

i) The eigenvalues of L_m are

$$\lambda_i = 1/i^2 , \quad i = 1(1)m . \quad (6.2)$$

That is, they are equal to the first eigenvalues of L .

ii) The iteration scheme (5.1) is (at least) locally linearly convergent to φ^* with the rate of convergence

$$\rho(G_m(\varphi^*)) = 1 - 1/m^2 .$$

Remarks:

i) The convergence rate tends quadratically to one as the step size h decreases, which is certainly a big disadvantage for the application of the iterative scheme (5.1).

- ii) If some other quadrature formula is used, the exact relation (6.2) is not usually obtained, but only some approximation to the first m eigenvalues of L .

Proof of theorem 13:

For the model problem, we have

$$D_m = m^2 \operatorname{diag} \left(\frac{1}{m+1-i} \right)_{i=1(1)m}, \quad B_m = \operatorname{diag} \left(\frac{1}{i} \right)_{i=1(1)m}.$$

Using the permutation matrix

$$P := \begin{pmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{pmatrix} \in \mathbb{R}^{m,m}$$

we obtain

$$D_m = m^2 P B_m P, \quad C_m^T = P C_m P,$$

and since $P^2 = I$,

$$L_m = m^2 P B_m P^2 C_m P B_m C_m = A_m^2, \quad (6.3)$$

$$A_m := m P B_m C_m.$$

It is sufficient to determine the eigenvalues of A_m or A_m^{-1} . It can easily be verified that

$$C_m^{-1} = m \begin{pmatrix} 1 & & 0 \\ -1 & \ddots & \\ 0 & & -1 \end{pmatrix} \in \mathbb{R}^{m,m}. \quad (6.4)$$

From (6.4) we immediately obtain

$$A_m^{-1} = \frac{1}{m} C_m^{-1} B_m^{-1} P = \begin{pmatrix} & & & 1 \\ & 0 & & -2 & -1 \\ & & \ddots & \ddots & \\ & & & -2 & \\ m-1 & & & & 0 \\ m & -(m-1) & & & \end{pmatrix}.$$

The matrix A_m^{-1} is similar to the lower triangular matrix

$$E_m := \left((-1)^{i+1} j \frac{(m-j)!}{(m-i)!} \binom{i-1}{i-j} \right)_{i,j=1(1)m}$$

as will be proved in lemma 4. (Note that $1/(-n)! := 0$, $n \in \mathbb{N}$.)

The eigenvalues of E_m , and hence of A_m^{-1} , are $(-1)^{i+1} i$, $i = 1(1)m$, from which the relation (6.2) follows by using (6.3).

Lemma 4^{*}: With the upper triangular matrices

$$V_m := \left(\frac{(-1)^{m+1-i}}{(m+1-i)(i-1)!(j-i)!} \right), \quad V_m^{-1} = \left(\frac{(-1)^{m+1-i}(m+1-j)(j-1)!}{(j-i)!} \right),$$

where $i, j = 1(1)m$, the relation $V_m^{-1} A_m^{-1} V_m = E_m$ holds.

Proof: We have

$$\begin{aligned} (A_m^{-1} V_m)_{ij} &= \frac{(-1)^i i}{i(m-i)!(i+j-m-1)!} - \frac{(-1)^{i+1}(i-1)}{(i-1)(m+1-i)!(i+j-m-2)!} \\ &= \frac{(-1)^i j}{(m+1-i)!(i+j-m-1)!}, \quad i, j = 1(1)m. \end{aligned}$$

Setting

$$F_{ijm} := (-1)^{m+1} j \frac{(i-1)!}{(j-1)!},$$

we obtain

$$\begin{aligned} (V_m^{-1} A_m^{-1} V_m)_{ij} &= (-1)^{m+1-i} j \sum_{r=i}^m (-1)^r \frac{(m+1-r)(r-1)!}{(r-i)!(m+1-r)!(m+j-m-1)!} \\ &= (-1)^i F_{ijm} \sum_{r=i}^m (-1)^r \binom{r-1}{r-i} \binom{j-1}{m-r} = F_{ijm} \sum_{r=i}^m \binom{-i}{r-i} \binom{j-1}{m-r} \\ &= F_{ijm} \sum_{r=0}^{m-i} \binom{-i}{r} \binom{j-1}{m-i-r} = F_{ijm} \binom{j-i-1}{m-i} = (-1)^{i+1} j \frac{(m-j)!}{(m-i)!} \binom{i-1}{i-j} \end{aligned}$$

*) The authors are grateful to C. Schneider of the University of Mainz for providing what is essentially the matrix V_m .

for $i, j = 1(1)m$, where we used the identity

$$\sum_{r=0}^n \binom{p}{r} \binom{q}{n-r} = \binom{p+q}{n}, \quad n \in \mathbb{N}_0, \quad p, q \in \mathbb{R}.$$

Remark: Using (6.4) it can easily be seen that the condition number of C_m with respect to the maximum norm is equal to $2m$.

7. NUMERICAL TESTS

In order to illustrate how the iterative method works in practice, some numerical tests were carried out on a CDC CYBER 875 computer, with a single precision accuracy of approximately 14 digits. The method was applied to two different integral equations with different step sizes, different stopping criteria, and different starting functions. Let f^* be the solution of the integral equation (1.1), and let φ^* be the solution of the corresponding discretized system. For each problem, values of $f_d^* := (f_i^*)$, $i = 1(1)m$, and φ^* , of the starting vector φ^0 and of the iterates $\varphi^1, \varphi^2, \varphi^n$ (last iteration) are shown in the tables. In table 1, owing to space problems, not all components of the considered vectors are given for $m = 20$ and 40 . The vectors in tables 2 and 3 are symmetric; this means that $v_i = v_{m+1-i}$ for $i = 1(1)(m+1)/2$, if $v := (v_i)$ is any vector in table 2 or 3.

i) The model problem

For this problem the discretization was as described in section 6. Table 1A shows the results for the starting values

$$\varphi^0 := \left(5 + \cos(i/m) \right)_{i=1(1)m}.$$

Table 1B shows the results for the starting values

$$\varphi^0 := \left(1.5 + \cos(50i/m) \right)_{i=1(1)m}.$$

In both cases the iteration was stopped when $\|\varphi - \varphi^n\|_\infty \leq \epsilon$ with $\epsilon = 10^{-2}$.

We can see, particularly from table 1A that the iterations yield large improvements in the beginning. This phenomenon can also be observed for the other integral equation. It is partly explained by theorem 5. A comparison between the tables 1A and 1B shows that the number of iterations is strongly dependent on the starting values. Table 1B also shows a strong dependence of n on the step size. This can also be predicted theoretically. Decreasing ϵ slightly for the case considered in table 1B increased the number of iterations drastically. However, $\epsilon = 10^{-8}$, for instance, with starting values φ^0 of the first case

required $n = 217$, 239 , and 249 iterations for $m = 10$, 20 , and 40 , respectively, which corresponds to rather slow convergence.

ii) Probability problem

In this example, the function g is essentially proportional to the probability density function of a sum of two independent random variables under the condition that the sum has values in $[0,1]$. We have

$$k(x,y) := h(x - y) + \eta ,$$

$$h(x) := \begin{cases} 1 - x^2/\delta^2 , & |x| \leq \delta , \delta > 0 , \\ 0 , & \text{otherwise,} \end{cases}$$

$$f^*(y) := \alpha y^2 (1 - y)^2 + \beta ,$$

$$g(x) := g_0(x,b) - g_0(x,a) + \eta(\alpha/30 + \beta) ,$$

$$a := \max(0, x - \delta) ,$$

$$b := \min(1, x + \delta) ,$$

$$\begin{aligned} g_0(x,y) := & \alpha[y^3(1 - x^2/\delta^2)(1/3 - y/2 + y^2/5) \\ & + 2xy^4(1/4 - 2y/5 + y^2/6)/\delta^2 \\ & - y^5(1/5 - y/3 + y^2/7)/\delta^2] \\ & + \beta[y + (x - y)^3/(3\delta^2)] , \end{aligned}$$

where α , β , δ , and η are properly chosen constants. Note that, if all constants are positive, the functions k and g satisfy the assumptions (α) and (β) made for the theoretical treatment of the iterative method (2.1).

The extended Simpson's rule was used for discretization of the integral equation. The solutions φ^* of the discretized problems $C_m \varphi^* = \gamma$ for different m were computed by using double precision Gaussian elimination, yielding at least single precision machine accuracy in the residuals for the different step sizes considered. (The condition numbers with respect to the maximum norm ranged between 1.8×10^3 and 4.0×10^6 .)

Two different cases were considered whose results are shown in tables 2 and 3. In both cases, $\alpha = 100$, $\delta = 0.12$, and $\gamma = 0.1$ were taken, whereas $\beta = 1$ was used in the first case (table 2) and $\beta = -1$ in the second (table 3). This implies that in the latter case the solution f^* takes negative values in the

tails. For both cases the starting values were constant and equal to 1, but different stopping criteria were applied. The solutions φ^* are oscillating around f_d^* as can be seen from tables 2 and 3.

For $\beta = 1$ we first used $\|\varphi^* - \varphi^n\|_\infty \leq \epsilon$ for the stopping criterion. Requiring an error of the order of one per cent we had to iterate $n = 146$ times for $m = 11$ and $n = 43428$ times for $m = 21$. Both cases show the extremely slow convergence of the method. On the other hand, in the applications, owing to measurement errors, it is often preferred to use $\|C_m \varphi^n - \gamma\|_\infty \leq \epsilon$ for the stopping criterion. For the results in table 2 the iteration was therefore stopped in such a manner for $\epsilon = 10^{-2}$ in the case of $\beta = 1$.

In the second case ($\beta = -1$), where we expected φ^j to converge, as $j \rightarrow \infty$, to the solution of the discretized analogue to the maximization problem MP, the iteration was stopped when φ_i^n was less than or equal to $\epsilon = 10^{-14}$ for all i for which φ_i^* is negative.

The results of table 2, especially for $m = 21$ and 41, show that if we do not iterate too often, φ^n may be closer to the values of f_d^* than φ^* is.

Table 3 finally shows that in the case of $\beta = -1$ the iterates converge to a vector which has those components vanishing for which f_d^* (and φ^*) has negative values. According to theorem 8, interpreted for the discrete case, the limiting vector is a solution to the (discrete) maximization problem MP. Note the relatively fast convergence of the method in this case.

Summarizing we may state that the method converges very slowly. On the other hand, since the right-hand side of the integral equation is always disturbed in practical problems (discretization- and measurement-errors), determination of the solution with high precision must not be expected. Therefore, a few iterations with this iterative method may give at least a good idea of what the solution looks like.

REFERENCES

- [1] Kondor, A. Method of Convergent Weights - An Iterative Procedure of Solving Fredholm's Integral Equations of the First Kind. Nuclear Instrum. Methods 216 (1983) 177-181.
- [2] Shepp, L.A. and Vardi, Y. Maximum Likelihood Reconstruction for Emission Tomography. IEEE Trans. Medical Imaging, Vol. MI-1 (1982) 113-122.
- [3] Dempster, A.P., Laird, N.M. and Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. J. Roy. Stat. Soc., 39 (1977) 1-38.
- [4] Roberts, W. and Varberg, D.E. Convex Functions. Academic Press, New York and London (1973).
- [5] Tricomi, F.G. Integral Equations. Interscience, New York (1957).
- [6] Ortega, J.M. and Rheinbolt, W.C. Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York and London (1970).

Table 1

Model problem for two different starting functions, $\|\varphi^* - \varphi^n\|_\infty \leq 10^{-2}$

	x	$f_d^* = \varphi^*$	m = 10				m = 20				m = 40			
			φ^0	φ^1	φ^2	φ^9	φ^1	φ^2	φ^{10}	φ^1	φ^2	φ^{10}	φ^1	φ^2
A	0.1	1.0	5.9950	1.0121	1.0084	0.9997	1.0107	1.0075	0.9999	1.0100	1.0071	1.0000	1.0100	1.0071
	0.2	1.0	5.9801	1.0109	1.0081	1.0010	1.0094	1.0072	1.0009	1.0087	1.0067	1.0010	1.0087	1.0067
	0.3	1.0	5.9553	1.0082	1.0066	1.0016	1.0067	1.0055	1.0014	1.0059	1.0049	1.0014	1.0059	1.0049
	0.4	1.0	5.9211	1.0041	1.0037	1.0017	1.0025	1.0025	1.0014	1.0017	1.0019	1.0013	1.0017	1.0019
	0.5	1.0	5.8776	0.9986	0.9996	1.0012	0.9969	0.9982	1.0009	0.9960	0.9976	1.0007	0.9960	0.9976
	0.6	1.0	5.8253	0.9917	0.9942	1.0001	0.9899	0.9927	0.9999	0.9891	0.9920	0.9997	0.9891	0.9920
	0.7	1.0	5.7648	0.9835	0.9877	0.9985	0.9817	0.9361	0.9984	0.9808	0.9853	0.9981	0.9808	0.9853
	0.8	1.0	5.6967	0.9742	0.9801	0.9964	0.9723	0.9784	0.9965	0.9714	0.9776	0.9962	0.9714	0.9776
	0.9	1.0	5.6216	0.9637	0.9715	0.9939	0.9618	0.9697	0.9942	0.9609	0.9689	0.9938	0.9609	0.9689
	1.0	1.0	5.5403	0.9523	0.9620	0.9910	0.9503	0.9601	0.9916	0.9494	0.9592	0.9911	0.9494	0.9592
B						φ^{289}			φ^{1640}			φ^{3950}		
	0.1	1.0	1.7837	1.2923	1.3105	0.9999	1.2476	1.2611	1.0000	1.2175	1.2239	1.0001	1.2175	1.2239
	0.2	1.0	0.6609	0.4909	0.5109	1.0005	0.4584	0.4607	0.9993	0.4495	0.4507	0.9991	0.4495	0.4507
	0.3	1.0	0.7403	0.5429	0.5595	0.9989	0.5094	0.5089	0.9983	0.5016	0.5015	1.0001	0.5016	0.5015
	0.4	1.0	1.9081	1.3424	1.3339	1.0019	1.3047	1.2969	1.0040	1.2885	1.2846	1.0048	1.2885	1.2846
	0.5	1.0	2.4912	1.7186	1.6801	0.9959	1.6971	1.6825	1.0100	1.6792	1.6718	0.9923	1.6792	1.6718
	0.6	1.0	1.6543	1.1513	1.1369	1.0081	1.1263	1.1171	1.0028	1.1152	1.1109	1.0100	1.1152	1.1109
	0.7	1.0	0.5963	0.4220	0.4240	0.9900	0.4061	0.4033	0.9947	0.4023	0.4012	0.9944	0.4023	0.4012
	0.8	1.0	0.8331	0.5883	0.5911	1.0066	0.5663	0.5617	0.9971	0.5617	0.5600	1.0006	0.5617	0.5600
	0.9	1.0	2.0253	1.3861	1.3543	0.9980	1.3695	1.3526	0.9997	1.3622	1.3550	1.0004	1.3622	1.3550
	1.0	1.0	2.4650	1.6262	1.5398	1.0000	1.6467	1.6089	1.0000	1.6488	1.6318	1.0000	1.6488	1.6318

Table 2

Probability problem: $\beta = 1$, $\|C_{\varphi}^n - \gamma\|_{\infty} \leq 10^{-2}$

x	m = 11							m = 21							m = 41						
	f_d^*	φ^0	φ^1	φ^2	φ^{105}	φ^*	$\varphi^* - f_d^*$	φ^1	φ^2	φ^{31}	φ^*	$\varphi^* - f_d^*$	φ^1	φ^2	φ^{57}	φ^*	$\varphi^* - f_d^*$				
0.00	1.000	1.0	3.721	3.273	1.259	1.290	0.290	3.642	3.122	1.390	1.059	0.059	3.686	3.165	1.298	1.300	0.300				
0.03	1.059	1.0						3.589	3.027	1.234	0.807	-0.252	3.589	3.027	1.234	0.807	-0.252				
0.05	1.226	1.0						3.538	2.959	1.273	1.929	0.703	3.538	2.959	1.273	1.929	0.703				
0.08	1.481	1.0						3.531	2.956	1.413	1.117	-0.364	3.531	2.956	1.413	1.117	-0.364				
0.10	1.810	1.0	3.484	2.941	1.463	1.418	-0.392	3.537	2.992	1.647	2.394	0.584	3.564	3.012	1.657	2.700	0.890				
0.13	2.196	1.0						3.649	3.147	2.059	1.643	-0.553	3.649	3.147	2.059	1.643	-0.553				
0.15	2.626	1.0						3.769	3.337	2.605	3.942	1.316	3.769	3.337	2.605	3.942	1.316				
0.18	3.084	1.0						3.911	3.562	3.119	2.314	-0.771	3.911	3.562	3.119	2.314	-0.771				
0.20	3.560	1.0	4.058	3.845	5.063	5.242	1.682	4.068	3.817	3.672	5.342	1.782	4.068	3.817	3.672	5.342	1.782				
0.23	4.041	1.0						4.229	4.081	4.108	2.962	-1.079	4.229	4.081	4.108	2.962	-1.079				
0.25	4.516	1.0						4.393	4.353	4.553	7.026	2.510	4.393	4.353	4.553	7.026	2.510				
0.28	4.975	1.0						4.550	4.616	4.928	3.700	-1.275	4.550	4.616	4.928	3.700	-1.275				
0.30	5.410	1.0	4.607	4.758	4.168	4.063	-1.347	4.701	4.872	5.355	8.053	2.643	4.701	4.872	5.355	8.053	2.643				
0.33	5.813	1.0						4.841	5.110	5.746	4.360	-1.452	4.841	5.110	5.746	4.360	-1.452				
0.35	6.176	1.0						4.968	5.328	6.167	9.278	3.102	4.968	5.328	6.167	9.278	3.102				
0.38	6.493	1.0						5.079	5.521	6.489	4.873	-1.620	5.079	5.521	6.489	4.873	-1.620				
0.40	6.760	1.0	5.068	5.582	9.853	10.131	3.371	5.172	5.683	6.799	10.139	3.379	5.172	5.683	6.799	10.139	3.379				
0.43	6.972	1.0						5.248	5.815	6.976	5.230	-1.742	5.248	5.815	6.976	5.230	-1.742				
0.45	7.126	1.0						5.301	5.907	7.135	10.728	3.602	5.301	5.907	7.135	10.728	3.602				
0.48	7.219	1.0						5.335	5.967	7.184	5.469	-1.750	5.335	5.967	7.184	5.469	-1.750				
0.50	7.250	1.0	5.257	5.874	5.568	5.441	-1.809	5.344	5.984	7.230	10.623	3.373	5.344	5.984	7.230	10.623	3.373				

Table 3

Probability problem: $\beta = -1$, $\varphi_i^n \leq 10^{-14}$ for all i with $\varphi_i^* < 0$

m = 11										m = 21					m = 41				
x	f_d^*	φ^0	φ^1	φ^2	φ^{419}	φ^*	$\varphi^* - f_d^*$	φ^1	φ^2	φ^{401}	φ^*	$\varphi^* - f_d^*$	φ^0	φ^1	φ^{394}	φ^*	$\varphi^* - f_d^*$		
0.00	-1.000	1.0	1.735	1.279	0.000	-1.396	-0.396	1.659	1.160	0.000	-0.592	0.408	1.687	1.186	0.000	-1.318	-0.318		
0.03	-0.941	1.0						1.532	1.016	0.000	-0.859	-0.085	1.590	1.071	0.000	-0.717	0.223		
0.05	-0.774	1.0											1.540	1.017	0.000	-1.247	-0.472		
0.08	-0.519	1.0						1.554	1.059	0.000	-0.045	0.145	1.533	1.017	0.000	-0.393	0.126		
0.10	-0.190	1.0	1.511	1.017	0.000	-0.294	-0.104						1.566	1.065	0.000	-0.267	-0.077		
0.13	0.196	1.0						1.761	1.345	0.000	0.408	-0.218	1.651	1.182	0.000	0.152	-0.044		
0.15	0.626	1.0											1.771	1.351	0.000	0.936	0.311		
0.18	1.084	1.0						2.046	1.775	1.275	2.572	1.012	1.913	1.558	0.014	0.814	-0.271		
0.20	1.560	1.0	2.069	1.806	1.273	2.423	0.863						2.069	1.799	1.072	2.340	0.780		
0.23	2.041	1.0						2.368	2.292	3.153	1.694	-0.821	2.231	2.054	3.098	1.601	-0.439		
0.25	2.516	1.0											2.394	2.323	3.510	3.529	1.013		
0.28	2.975	1.0						2.678	2.819	2.494	5.227	1.817	2.552	2.586	2.920	2.266	-0.709		
0.30	3.410	1.0	2.652	2.787	2.863	2.551	-0.859						2.703	2.844	2.578	5.180	1.770		
0.33	3.813	1.0						2.937	3.269	3.587	3.119	-1.057	2.843	3.086	3.007	2.860	-0.952		
0.35	4.176	1.0											2.970	3.308	3.518	6.254	2.078		
0.38	4.493	1.0	3.087	3.568	6.328	7.179	2.419	3.149	3.649	5.785	6.962	2.202	3.081	3.507	4.515	3.369	-1.124		
0.40	4.760	1.0											3.174	3.672	5.230	7.149	2.389		
0.43	4.972	1.0						3.267	3.860	4.845	3.730	-1.396	3.250	3.809	5.573	3.732	-1.240		
0.45	5.126	1.0											3.302	3.903	5.074	7.658	2.532		
0.48	5.219	1.0						3.321	3.960	4.478	8.013	2.763	3.337	3.967	4.650	3.862	-1.357		
0.50	5.250	1.0	3.302	3.931	4.127	3.949	-1.301						3.346	3.983	4.212	8.130	2.880		