Issue 1 July–December 1998 ISSN 1083-0464

SOUTHWEST JOURNAL OF PURE AND APPLIED MATHEMATICS

CAMERON UNIVERSITY
DEPARTMENT OF MATHEMATICAL SCIENCES
LAWTON, OK 73505, U.S.A

Southwest Journal
of
Pure and Applied Mathematics

Editor in Chief: Ioannis Argyros
Managing Editor: Mohammad Tabatabai
Technical Editor: Gary Huckabay

Editorial Office

SWJPAM
Cameron University
Department of Mathematical Sciences
Lawton, OK 73505, U.S.A.

Publication Program

2 issues per year
July, December

Information for Authors

Scope. SWJPAM is an electronic journal devoted to all aspects of Pure and Applied mathematics, and related topics. Authoritative expository and survey articles on subjects of special interest are also welcomed.

Since SWJPAM aims to serve not only the specialists, but also the user of mathematics (promoting, in such a way, a strong interaction between them), particular emphasis is placed on applications. Thus, SWJPAM publishes papers which, although not directly contributing to mathematics, contain, however, significant results (from any field) obtained by making an essential and well stressed use of the methods and results of mathematics.

Submission of Manuscripts. Since SWJPAM is a completely electronic journal, all manuscripts must be submitted in either AMSTEX, LaTEX, or TEXform. Papers for publication should be submitted to ftp://rattler.cameron.edu/SWJPAM_incoming by anonymous ftp. The articles must be placed in the 'SWJPAM_incoming' directory. The submission should include the title of the paper, the author's name and affiliation, and a short abstract which summarizes the paper; a list of keywords and a classification according to the Mathematical Reviews' Subject Classification (1991). Authors are informed by e-mail of referee communication.

Key Words. Key words should be comprehensive and subject specific. It is not necessary to list the subject area of the journal's coverage as a key word. Four key words should be sufficient to cover the major subjects off a given paper, although more can always be supplied if the author deems it necessary.

 (continued on inside back cover)

SOUTHWEST JOURNAL
OF
PURE AND APPLIED MATHEMATICS
ISSUE 1, JULY–DECEMBER 1998

CONTENTS

ISSN 1083-0464

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISBN 1083-0464
Issue 1 (July) 1998, pp. 1 - 8
Submitted: October 2, 1997

On Multivariate Approximation by Mejer-Konig-Zeller Type Polynomials

Tianzi Jiang

Abstract

Multivariate positive linear operators have many good properties, such as convergence, convexity preserving, simultaneous approximation, and so on. They can been used in the field of computer graphics. The typical example is multivariate Bernstein polynomial. Recently, much effort was devoted to study the properties of multivariate positive linear operators. In this paper, we give a multivariate version of Mejer-Konig-Zeller operator and find out the approximation properties of this operator. Our results are theoretical, but they will provide the benchmark for the use of such operators in practice.

AMS subject classification: 41A36, 41A63.
Key words and phrases: Multivariate approximation, Mejer-Konig-Zeller polynomials, positive linear operator.

School of Mathematics, The University of New South Wales, Sydney 2052, Australia and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, P.R. China
E-mail: jiangtz@maths.unsw.EDU.AU and jiangtz@prlsun6.ia.ac.cn
This work was partially supported by the Chinese National Science Fundation, the Chinese National Key Projects for Basic Science Researchers, and the UNSW Vice-Chancellor's Post Doctoral Fellowship in Australia.

1 Introduction

Multivariate positive linear operators have many good properties, such as convergence, convexity preserving, simultaneous approximation, and so on. They can been used in the field of computer graphics. The typical example is multivariate Bernstein polynomial. Recently, much effort was devoted to study the properties of multivariate positive linear operators. We only metion references [1-6]. In this paper, we give a multivariate version of Mejer-Konig-Zeller type operator and find out the approximation properties of this operator. Our results are theoretical, but they will provide the benchmark for the use of such operators in practice.

Let $\Delta = \{(x,y) : x, y \geq 0, x+y \leq 1\}$. For positive integer r, we define $C(\Delta)$ and $C^r(\Delta)$ as follows

$$C(\Delta) = \{f : f \quad \text{is continuous on} \quad \Delta \quad \text{and with real value}\}.$$

$$C^r(\Delta) = \left\{ f \in C(\Delta) : \frac{\partial^k f}{\partial^i \partial^{k-i}} \in C(\Delta), 0 \leq i \leq k, 0 \leq k \leq r \right\}.$$

For $f \in C(\Delta)$ and $\delta > 0$, we define a various modules of continuity of f as follows.

$$\begin{aligned} \omega(f;\delta) = \sup\{f(x',y') - f(x,y)| : (x',y'), (x,y) \in \Delta \\ |x'-x| \leq \delta, |y'-y| \leq \delta\}. \end{aligned}$$

$$\omega_1(f;\delta) = \sup\{|f(x_1,y) - f(x_2,y)| : (x_1,y), (x_2,y) \in \Delta |x_1 - x_2| \leq \delta\}.$$

$$\begin{aligned} \omega_2(f;\delta) = \sup\{f(x,y_1) - f(x,y_2)| : (x,y_1), (x,y_2) \in \Delta \\ |y_1 - y_2| \leq \delta\}. \end{aligned}$$

For $f \in C^1(\Delta)$, we define $\omega_{10}(f;\delta), \omega_{01}(f;\delta)$ as

$$\omega_{10}(f;\delta) = \omega(f'_x;\delta); \omega_{01}(f;\delta) = \omega(f'_y;\delta),$$

where f'_x and f'_y are the first order partial derivatives of f with respect to variables x and y, respectively. Now we define $L(\alpha,\delta)$ as

$$L(\alpha,\delta) = \{f \in C(\Delta) : \omega(f,\delta) = O(\delta^\alpha), \text{as} \quad \delta \to 0+\}.$$

Let $f \in L(\Delta)$. We define the multivariate Mejer-Konig-Zeller type operator $L_{mn}(f;x,y)$ as

$$L_{mn}(f;x,y) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} f(\frac{k}{m+k}, \frac{l}{n+l}) M_{mnkl}(x,y),$$

where

$$M_{mnkl} = \left(\begin{array}{c} m+k \\ k \end{array} \right) x^k (1-x)^{m+1} \left(\begin{array}{c} n+l \\ l \end{array} \right) y^l (1-y)^{n+1}.$$

The rest of this is organized as follows. Section 2 provides some lemmas which will be used in the course of poving the main results. Section 3 includes all approximate properties of the multivariate Mejer-Konig-Zeller type and their proofs.

2 Lemmas

In order to prove our main results we need the following lemmas.

Lemma 2.1

$$L_{mn}(1; x, y) = 1 \tag{1}$$

$$L_{mn}(s; x, y) = x \tag{2}$$

$$L_{mn}(t; x, y) = y. \tag{3}$$

$$L_{mn}(s^2; x, y) = x^2 + \varepsilon_m(x), \tag{4}$$

where

$$\varepsilon_m(x) \leq \frac{x(1-x)}{m}. \tag{5}$$

$$L_{mn}(t^2; x, y) = x^2 + \varepsilon_n(y), \tag{6}$$

where

$$\varepsilon_n(y) \leq \frac{y(1-y)}{n}. \tag{7}$$

Proof: (1)-(3) are obvious.
Now we prove (4) and (5).

$$L_{mn}(s^2; x, y) = \sum_{k=1}^{\infty} \frac{(m+k-1)!}{(k-1)!m!} \frac{k}{m+k} x^k (1-x)^{m+1}$$
$$= \sum_{k=2}^{\infty} \frac{(m+k-2)!}{(k-2)!m!} x^k (1-x)^{m+1}$$

$$
\begin{aligned}
&+\sum_{k=1}^{\infty}\frac{(m+k-1)!}{(k-1)!m!}\left(\frac{k}{m+k}-\frac{k-1}{m+k-1}\right)x^k(1-x)^{m+1}\\
&=x^2+\sum_{k=1}^{\infty}\frac{(m+k-1)!}{(k-1)!m!}\frac{m}{(m+k)(m+k-1)}x^k(1-x)^{m+1}\\
&=x^2+x(1-x)\sum_{k=1}^{\infty}\frac{(m-1+k-1)!}{(k-1)!(m-1)!}\frac{1}{m+k}x^{k-1}(1-x)^m\\
&=x^2+x(1-x)\sum_{k=0}^{\infty}\frac{(m-1+k)!}{(k)!(m-1)!}\frac{1}{m+k+1}x^k(1-x)^m\\
&:=x^2+\varepsilon_m(x).
\end{aligned}
$$

Since

$$\frac{1}{m+k+1}\le\frac{1}{m}$$

we have

$$\varepsilon_m(x)\le\frac{x(1-x)}{m}.$$

Similarly, we can prove (6) and (7). This completes the proof of Lemma 2.1.

Lemma 2.2

$$L_{mn}((s-x)^2;x,y)\le\frac{x(1-x)}{m} \tag{8}$$

$$L_{mn}((t-y)^2;x,y)\le\frac{y(1-y)}{n}, \tag{9}$$

$$L_{mn}(|s-x|;x,y)\le\sqrt{\frac{x(1-x)}{m}}, \tag{10}$$

$$L_{mn}(|t-y|;x,y)\le\sqrt{\frac{y(1-y)}{n}}, \tag{11}$$

Proof: By Lemma 2.1 we have

$$
\begin{aligned}
L_{mn}((s-x)^2;x,y)&=L_{mn}(x^2;x,y)-2xL_{mn}(s;x,y)+L_{mn}(s^2;x,y)\\
&=\varepsilon_m(x)\le\frac{x(1-x)}{m}.
\end{aligned}
$$

Similarly, we can prove (9). By Holder inequality, we have

$$L_{mn}(|s-x|;x,y)\le\left(\sqrt{L_{mn}((s-x)^2;x,y)}\right)^{1/2}.$$

From this and (5) we can obtain (10). Similarly we can obtain (11). This completes the proof of Lemma 2.2.

3 Main Results and Their Proofs

Theorem 3.1 *If function $f(s,t)$ is continuous at $(x,y) \in \Delta$, then*

$$\lim_{m,n\to\infty} L_{mn}(f;x,y) = f(x,y). \tag{12}$$

Moreover, if $f \in C(\Delta)$, $L_{mn}(f;x,y)$ is uniformly continuous on Δ.

Proof: From Lemma 2.1, as $m,n \to \infty$ we have

$$L_{mn}(1;x,y) \Rightarrow 1,$$

$$L_{mn}(s;x,y) \Rightarrow x,$$

$$L_{mn}(t;x,y) \Rightarrow y,$$

$$L_{mn}(s^2+t^2;x,y) \Rightarrow x^2+y^2,$$

By Korovkin's theorem [7], this yields Theorem 3.1.

Theorem 3.2 *For any $f \in C(\Delta)$, as $m,n \to \infty$, we have*

$$\|L_{mn}(f) - f\|_{C(\Delta)} \le 2\omega(f, 1/\sqrt{\frac{1}{m}+\frac{1}{n}}). \tag{13}$$

Proof: $\forall \delta \in (0,1)$, by Holder's inequality and Lemma 2.2 we have

$$\begin{aligned}
|L_{mn}(f;x,y) - f(x,y)| &\le \omega(f,\delta)\left(1 + \frac{1}{\delta}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} M_{mnkl}(x,y)\|(\frac{k}{m+k},\frac{l}{n+l}) - (x,y)\|\right) \\
&\le \omega(f,\delta)\left(1 + \frac{1}{\delta}\{\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} M_{mnkl}(x,y)[(\frac{k}{m+k}-x)^2 + (\frac{l}{n+l}-y)^2]\}^{1/2}\right) \\
&\le \omega(f,\delta)\left(1 + \frac{1}{\delta}\{L_{mn}((s-x)^2+(t-y)^2;x,y)\}^{1/2}\right) \\
&\le \omega(f,\delta)\left(1 + \frac{1}{\delta}\sqrt{\frac{x(1-x)}{m}+\frac{y(1-y)}{n}}\right) \\
&\le \omega(f,\delta)\left(1 + \frac{1}{\delta}\sqrt{\frac{1}{m}+\frac{1}{n}}\right).
\end{aligned}$$

Let $\delta = 1/\sqrt{\frac{1}{m}+\frac{1}{n}}$, we can obtain the desired result. This completes the proof of Theorem 3.2.

Corollary 3.1 *If $f \in Lip(\alpha;\Delta)(0<\alpha\le 1)$, then*

$$\|L_{mn}f - f\| \le K_f(\frac{1}{m}+\frac{1}{n})^{\alpha/2},$$

where K_f is a constant only depending on f.

Theorem 3.3 *For any $f \in C^1(\Delta)$, as $m,n\to\infty$, we have*

$$\|L_{mn}(f;x,y)-f(x,y)\|_{C(\Delta)} \le 2\omega_{10}\left(f, 1/\sqrt{x(1-x)/m+y(1-y)/n}\right)$$
$$+2\omega_{01}(f,1/\sqrt{x(1-x)/m+y(1-y)/n}).$$

Proof: Let $f\in C^1(\Delta), \forall (x,y)\in\Delta$, by Taylor's formula [8, pp.129-139] we have

$$\begin{aligned} f(s,t)-f(x,y) &= f'_x(x,y)(s-x)+f'_y(x,y)(t-y) \\ &+ \left[f'_x(x+\theta_1(s-x),y+)(s-x),y+\theta_2(t-y))-f'_x(x,y)\right](s-x) \\ &+ \left[f'_y(x+\theta_1(s-x),y+)(s-x),y+\theta_2(t-y))-f'_y(x,y)\right](t-y), \end{aligned}$$

where $0<\theta_1,\theta_2<1$.

Let $\delta>0$, by Lemma 2.2 we have

$$\begin{aligned} &|L_{mn}(f;x,y)-f(x,y)| \\ &\le \omega_{10}(f,\delta)\left\{C_{mn}(x,y)+\frac{1}{\delta}[C^2_{mn}(x,y)+C_{mn}(x,y)d_{mn}(x,y)]\right\} \\ &+\omega_{01}(f,\delta)\left\{C_{mn}(x,y)+\frac{1}{\delta}[C_{mn}(x,y)d_{mn}(x,y)+C^2_{mn}(x,y)]\right\} \end{aligned}$$

where

$$C_{mn}(x,y)=\sqrt{L_{mn}((s-x)^2;x,y)}\le\sqrt{\frac{x(1-x)}{m}}.$$

$$d_{mn}(x,y)=\sqrt{L_{mn}((t-y)^2;x,y)}\le\sqrt{\frac{y(1-y)}{m}}.$$

It is clear that

$$C^2_{mn}(x,y)+C_{mn}(x,y)d_{mn}(x,y)\le\sqrt{\frac{x(1-x)}{m}+\frac{y(1-y)}{n}}.$$

$$C_{mn}(x,y)d_{mn}(x,y)+d^2_{mn}(x,y)\le\sqrt{\frac{x(1-x)}{m}+\frac{y(1-y)}{n}}.$$

Let $\delta=\sqrt{x(1-x)/m+y(1-y)/n}$, we can obtain the requred result. This completes the proof of Theorem 3.3.

Corollary 3.2 *For any $f \in C^1(\Delta)$, if $f'_x, f'_y \in Lip(1, \Delta)$, as $m, n \to \infty$, we have*

$$\|L_{mn}(f) - f\|_{C(\Delta)} = O(\frac{1}{m} + \frac{1}{n}).$$

Theorem 3.4 *Denote $L_n(f; x, y) = L_{nn}(f; x, , y)$. For any $f \in C^2(\Delta)$, we have*

$$\begin{aligned}
&\lim_{m,n\to\infty} n\left\{L_n(f; x, y) - f(x, y)\right\} \\
&= (1 - 3x)f'_x(x, y) + (1 - 3y)f'_y(x, y) \\
&+x(1 - x)f''_{xx}(x, y) + (x + y - xy)f''_{xy}(x, y) + y(1 - y)f''_{yy}(x, y),
\end{aligned}$$

where f''_{xx} and f''_{yy} are the second order partial derivatives of f with respect to variables x and y, respectively, and f''_{xy} is the second order mixed partial derivative of f with respect to variables x and y.

Proof: Let $f \in C^2(\Delta), \forall (x, y) \in \Delta$, by Taylor's formula [8, pp.129-139] we have

$$\begin{aligned}
f(s, t) &= f(x, y) + f'_x(x, y)(s - x) + f'_y(x, y)(t - y) \\
&+\frac{1}{2}\left\{f''_{xx}(x, y)(s - x)^2 + 2f''_{xy}(x, y)(s - x)(t - y)\right. \\
&\left.+f''_{yy}(x, y)(t - y)^2\right\} + R_2(s, t; x, y),
\end{aligned} \tag{14}$$

where

$$R_2(s, t; x, y) = \sum_{i=0}^{2} \alpha_i(s, t)(s - t)^{2-i}(t - y)^i.$$

Here $\alpha_i(s, t)$ is bounded on Δ and

$$\lim_{(s,t)\to(x,y)} \alpha_i(s, t) = 0. \tag{15}$$

From the properties of $\alpha_i(s, t)$ and Lemma 2.1 we have

$$L_n(R_2; x, y) = o(\frac{1}{n}), \quad (n \to \infty). \tag{16}$$

By (14) and Lemma 2.1 we have

$$\begin{aligned}
L_n(f; x, y) &= f(x, y) + f'_x(x, y)L_n(s - x; x, y) + f'_y(x, y)L_n(t - y; x, y) \\
&+\frac{1}{2}\left\{f''_{xx}(x, y)L_n((s - x)^2; x, y) + 2f''_{xy}(x, y)L_n((s - x)(t - y); x, y)\right. \\
&\left.+f''_{yy}(x, y)L_n((t - y)^2; x, y)\right\} + L_n(R_2; x, y).
\end{aligned} \tag{17}$$

From Lemma 2.1 and (15)-(17) we can obtain the required result. This completes the proof of Theorem 3.4.

Acknowledgement

The author is highly grateful to the referees for their comments, which improve the presentation of the paper.

References

[1] M.M.Derriennic, On Multivariate Approximation by Bernstein-Type Polynomials, J. Approx. Theory,**45**(1985), 155-166.

[2] Z.Ditzian, X.Zhou, Kantorovich-Bernstein polynomials. Constr. Approx., 6 (1990), 421–435.

[3] Z.Ditzian, X.Zhou, Optimal approximation class for multivariate Bernstein operators, Pacific J. Math. 158 (1993), 93–120. Pacific J. Math.,**158**(1993), 93-120.

[4] R.Jia, Z.Wu, Bernstein polynomials on simplices, (Chinese)Acta Math. Sinica 31 (1988), no. 4, 510–522.

[5] J.Wu, Some properties of Bernstein-Durrmeyer operators for two variables, Chinese Ann. Math. Ser. A, 8:3 (1987), 384–393.

[6] Z.Wu, Sharp error bounds for approximation by Bernstein polynomials on n-dimensional simplices, (Chinese) Chinese Ann. Math. Ser. A 10 (1989), no. 1, 23–30.

[7] P.P. Korovkin, Linear Operators and Approximation Theory, Hindustan Pub. Corp., Delhi, 1960.

[8] C.H. Edwards, Jr., Advanced Calculus of Several Variables, Academic Press, New York, 1973.

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Issue No. 01, 1998, pp. 9–23.
Submitted: September 15, 1997. Published: July 31, 1998

ON THE TRANSLATED EXPONENTIAL MODEL WITH CENSORING

A. A. Bartolucci[1], J. M. Dickey[2], K. P. Singh[1], M. A. Tabatabai[3]

Abstract. Two distinct Bayesian methodologies are given for "exact"inference on translated exponential survival models in one and two population problems with censoring. The first approach, based on the use of natural conjugate families of prior distributions, differs from the traditional frequentist analysis in that the derived posterior distributions are exact under the model. The second method is based on noncontinuous-type extensions of natural conjugate priors, and involves the use of Bayes factors for sharply defined hypotheses. The two methods are intended for different states of prior information and would yield different inferences for similar data. Inferences from the two methods are presented comparing data in two breeds of mice to determine if there are genetic differences with respect to surviving the bacterial strain salmonella typhimurium.

A.M.S. (MOS) Subject classification Codes. 62,60,92.

Keywords and Phrases. Bayesian Analysis, Conjugate Prior, Bayes Factor, Mean Survival, Censoring, Exponential Distribution.

1. INTRODUCTION.

An exponential survivor function is often indicated when sample data is plotted by the Kaplan-Meier (1958) technique. At the cellular immunobiology laboratory of the University of Alabama at Birmingham (U.A.B.) Cancer Center recent survival analyses of mice exposed to a bacterial strain have suggested exponential fits with a possible non zero location parameter. We thus consider a time to failure T measured from the time of exposure, as having the probability density

$$\mu^{-1}\exp[-(t-\delta)/\mu], \quad t>\delta, \tag{1.1}$$

that is,

$$T|(\mu,\delta) \sim \text{Exponential}(\mu,\delta), \quad \mu>0, \quad \delta>0. \tag{1.2}$$

Zelen (1966) discussed the role of δ in clinical trial experience as the "guarantee time,"or minimal time until the first possible failure. The parameter μ is then the mean time to failure from time δ. We address here the important issue of inference on the mean total time to failure

$$\tau = E(t) = \mu + \delta$$

[1]Department of Biostatistics, University of Alabama at Birmingham, E-mail address: abartolucci@crl.soph.uab.deu
[2]School of Statistics, University of Minnesota
[3]Department of Mathematical Sciences, Cameron University, E-mail address: mohammadt@cameron.edu

for μ and δ both unknown. In the two-sample case, i.e., comparing samples from (1.2) and the distribution $T^*|(\mu^*, \delta^*) \sim$ Exponential (μ^*, δ^*), we also extend our inferential procedures to the quotient of means, $\tau/\tau^* = (\mu + \delta)/(\mu^* + \delta^*)$.

Our first approach to the problem of inference on τ and τ/τ^*, given in Sections 2 and 3, respectively, is a Bayesian prior-to-posterior analysis from smooth prior densities in a new conjugate prior family. The posterior distributions are employed in straightforward fashion to yield probability regions for the parameters of interest.

A second approach, Section 4, offers a genuine alternative to these methods for inference on point or sharp hypotheses. This is intended for situations where there is appreciable prior density mounted up over a narrow neighborhood surrounding a simplified sampling model, that is , situations where the experimenter seriously suspects that the simplified model may be a good approximation to the unknown actual sampling model. This will involve the use, as approximations, of positive probabilities assigned to the lower dimensional hypothesis sets in the parameter space, and bayes factors expressed in terms of density ordinates rather than tail areas (Dickey and Lientz 1970). This is a further refinement and extension of the approach discussed in the stochastic curtailing procedure of Halpern et al. (1982), Lan et al. (1983), and Lachin and Lan (1992).

In Section 5, we give numerical illustrations of the methods using the bacterial infection example.

2. A SINGLE POPULATION.

Under the model (1.1) those observations which are censored have likelihood contributions in the form of right tail area probabilities. Namely, a clinical trial patient known only to survive beyond time t contributes a factor to the likelihood,

$$\exp[-(t-\delta)/\mu \tag{2.1}$$

A sample of n failure times and $m - n \geq 0$ censored times for a single treatment regimen,

$$\text{Data}: \; t_i, \; i = 1, \ldots, n, \ldots, m \tag{2.2}$$

yields the likelihood function of μ and δ,

$$\mu^{-n}\exp\{-n[\bar{t} - (m/n)\delta]/\mu\}, \quad 0 \leq \delta \leq d, \, \mu \geq 0, \tag{2.3}$$

where

$$d = \min(t_1, \ldots, t_n), \bar{t} = \sum_{i=1}^{m} t_1/n. \tag{2.4}$$

For notational convenience, we have assigned the n initial indices to the observed failures. Note that the total number of patients is m. Under type-II censoring for the case $\delta = 0$, the reciprocal observed failure rate $\bar{t}$ is distributed as $\mu\chi^2_{2n}/2n$, and hence it would estimate μ unbiasedly (Bain 1978).

The family of distributions which is natural-conjugate to the family of likelihoods (2.3) will have the form with density proportional to (2.3.),

$$p(\mu, \delta) = k' \mu^{-n'-1}\exp\{-n'[\bar{t}' - (m'/n')\delta]/\mu\} \tag{2.5}$$

$$0 < \delta < d' \quad \mu > o, \quad d' > 0,$$

where the normalizing constant can be seen to be

$$\begin{aligned} k' &= \{m'(n'-1)n^{n'-1}/\Gamma(n')\} \\ &/\{[\bar{t}' - (m'/n')d']^{-(n'-1)} - [\bar{t}']^{-(n'-1)}\}. \end{aligned}$$

This family is parameterized (or "hyper parameterized") by $n', \bar{t}', m', d' \geq 0$. Assuming $n'\bar{t}' - m'd' \geq 0$ and $n', m' > 0$, we have that with prior probability one, $n't' - m'\delta > 0$.

Defining

$$S'(\delta) = \bar{t}' - (m'/n')\delta,$$

for which $S'(\delta) > 0$, we see that the conditional density of μ given δ takes the recognizable chi-square form,

$$S'(\delta)/\mu|\delta \sim \chi^2_{2n'}/2n', \tag{2.6}$$

whereas under the reverse conditioning

$$p(\delta|\mu) = (m'/\mu)e^{m'\delta/\mu}/(e^{m'd'/\mu} - 1), \tag{2.7}$$

In both cases, (2.6) and (2.7), note that the ranges of the conditional distributions do not depend on the conditioning variables.

The marginal prior density for μ is

$$p(\mu) = k'm^{-1}\mu^{-n'}e^{-n'\bar{t}'/\mu}(e^{m'd'/\mu} - 1), \quad \mu > 0. \tag{2.8}$$

i.e., a linear combination (with one coefficient negative) of two reciprocal chi-square densities each with $2(n' - 1)$ degrees of freedom (variables $2(n'\bar{t}' - m'd')/\mu$ and $2n'\bar{t}'/\mu$). The marginal density of δ takes the simple form,

$$p(\delta) = k'\Gamma(n')\{n'S'(\delta)\}^{-n'}, \quad 0 \leq \delta \leq d. \tag{2.9}$$

We have found this family to be large enough in practice to allow realistic modeling of prior opinions. This model, like many models (Dickey 1979), can bring order into expert-opinion processes. The elicitation of a distribution within the family, to be described here in Section 5, involves a statement of choice of survival quartiles by the expert investigator (e.g. clinician). For further general theory on elicitation through prior percentiles, see Kadane et al. (1980) and Dickey (1979).

In this family, the prior-to-posterior transformation of hyper parameters for updating to observational data by Bayes Theorem takes a pooled-sample form, as is usual with conjugate families for general-exponential-form sampling models,

$$\begin{aligned} n'' &= n' + n \\ \bar{t}'' &= (n'\bar{t}' + n\bar{t})/(n' + n) \\ m'' &= m' + m \\ d'' &= \min(d', d). \end{aligned} \tag{2.10}$$

If the prior density is parameterized by $n', \bar{t}', m', d',$, then the corresponding posterior density will have parameters $n'', \bar{t}'', m'', d''$. (Note that there are two sample-size type hyper parameters, n and m.) Again, $\bar{t}'' > (m''/n'')/d''$. The marginal posterior distributions of μ and of δ, respectively, are then given by (2.8) and (2.9) with n', t', m', d', replaced by $n'', \bar{t}'', m'', d''$. These are not recognized as traditional distributions.

Note that for such linear functions of μ and δ as the overall mean survival,

$$\tau = \mu + \delta, \tag{2.11}$$

an estimate can be obtained as a linear function of the two posterior means, as $E(\mu) + E(\delta)$. The actual prior density of this parameter τ has the form,

$$p(\tau) \propto \begin{cases} \Gamma \quad (n^{'}, a/\tau) \quad \text{if max} \quad (0, \tau - d^{'}) = 0 \\ \Gamma \quad (n^{'}, a/\tau) - \Gamma(n^{'}, a/(\tau - d^{'})) \quad \text{if max} \quad (0, \tau - d^{'}) = \tau - d^{'} \end{cases} \tag{2.12}$$

where $a = n^{'}\bar{t}^{'} - m^{'}\tau$ and in a standard notation (Abramowitz and Stegun 1964),

$$\Gamma(x, y) = \int_{y}^{\infty} z^{x-1} e^{-z} dz. \tag{2.13}$$

The corresponding posterior density is

$$p(\tau|\text{Data}) \propto \begin{cases} \Gamma \quad (n^{''}, a^{''}/\tau) \quad \text{if max} \quad (0, \tau - d^{''}) = 0 \\ \Gamma \quad (n^{''}, a^{''}/\tau) - \Gamma(n^{''}, a^{''}/\tau - d^{''})) \quad \text{if max} \quad (0, \tau - d^{''}) = \tau - d^{''} \end{cases} \tag{2.14}$$

and $a^{''} = n^{''}\bar{t}^{''} - m^{''}\tau$. Numerical computation of posterior credible regions for τ are feasible by use of equation (2.14).

3. COMPARING TWO POPULATIONS.

Consider also an independent second sample of patient survival and censored times,

$$\text{Data}^* : t_i^*, i = 1, \ldots, n^*, \ldots, m^*, \tag{3.1}$$

of which n^* are actual failures ($n^* \leq m^*$). This data may represent the observed survival experience for a second therapy with underlying mean, $\tau^* = \mu^* = \delta^*$. We establish in this section inference procedures on the relation of τ to τ^* in either situation of the unknown locations δ, δ^* being assumed equal or unequal.

In the second of these two cases, that is $\delta \neq \delta^*$, assuming the natural conjugate prior distribution for (μ^*, δ^*) independent of (μ, δ), with parameters $n^*, \bar{t}^{*'}, m^{*'}$ and $d^{\star'}$, one obtains a posterior distribution for $(\mu^\star, \delta^\star)$ independent of (μ, δ) and identical in form to that given in Section 2. (All the prior and posterior derived distributions there are taken to apply again to the second population parameters $\mu^\star, \delta^\star, \tau^\star$, but with appropriately starred hyper parameters.)

To facilitate the comparison of the population means $\tau = \mu + \delta$ and $\tau^\star = \mu^\star + \delta^\star$, we perform the change of variables to (η, τ), where

$$\eta = \tau/\tau^*. \tag{3.2}$$

Inference can now be focused on suitable null values, $\eta = \eta_0$, where the natural choice of $\eta_0 = 1$ implies equality of the two population means. From (3.2), the marginal posterior density of η, in proportionality form, is

$$p(\eta|\text{Data, Data}^*) \propto I_1 + I_2 + I_3, \tag{3.3}$$

where, assuming $d^{''} \leq d^{\star''}$,

$$
\begin{aligned}
I_1 =& \frac{1}{\eta^2}\int_0^{d''}\left\{\tau^2(b-c\tau)^{1-g}\left[\frac{b^*\eta-c^*\tau}{\eta}\right]^{i-g^*}\right.\\
&\left.\cdot\Gamma\left(g-1,\frac{b-c\tau}{\tau}\right)\Gamma\left(g^*-1,\frac{b^*\eta-c^*\tau}{\tau}\right)\right\}d\tau\\
I_1 =& \frac{1}{\eta^2}\int_0^{d''}\left\{\tau^2(b-c\tau)^{1-g}\left[\frac{b^*\eta-c^*\tau}{\eta}\right]^{1-g^*}\right.\\
&\cdot\left[\Gamma\left(g-1,\frac{b-c\tau}{\tau}\right)-\Gamma\left(g-1,\frac{b-c\tau}{\tau-d''}\right)\right]\\
&\left.\cdot\Gamma\left(g^*-1,\frac{b^*\eta-c^*\tau}{\tau}\right)\right\}.d\tau\\
I_3 =& \frac{1}{\eta^2}\int_{d^{*\prime\prime}}^{\infty}\left\{\tau^2[b-c\tau]^{1-g}\left[\frac{b^*\eta-c^*\tau}{\eta}\right]^{1-g^*}\right.\\
&\cdot\left[\Gamma\left(g-1,\frac{b-c\tau}{\tau}\right)-\Gamma\left(g-1,\frac{b-c\tau}{\tau-d''}\right)\right]\\
&\left.\cdot\left[\Gamma\left(g^*-1,\frac{b^*\eta-c^*\tau}{\tau}\right)-\Gamma\left(g^*-1,\frac{b^*\eta-c^*\tau}{\eta(\eta-d^{*\prime\prime})}\right)\right]\right\}d\tau
\end{aligned}
$$

and where

$$
\begin{aligned}
&P(\eta\le\eta_0|Data,Data^*).\\
&b=n''\bar{t}'', \qquad && b^*=n^{*\prime\prime}\bar{t}^{*\prime\prime}\\
&c=m'', && c^*=m^{*\prime\prime}\\
&g=n'', && g^*=n^*.
\end{aligned}
\tag{3.4}
$$

For given fixed values, η_0, expression (3.3) can be computed numerically to yield posterior credible regions.

For the remaining case of equal but unknown location, $\delta=\delta^\star$, our results are much simpler. A consideration of $\tau\le\tau^\star$ reduces to $\mu\le\mu^\star$. Thus the mean survivals τ and $\tau^\star$ can be compared in terms of μ and $\mu^\star$. It is well known that the conjugate prior family for a simplified sampling model can be obtained by conditioning in the conjugate prior family of a more complicated sampling model, provided one uses a conditioning variable which is a linear combination of the joint variables (Gunel and Dickey 1974). We condition on the variable and value $\delta-\delta^\star=0$. The preceding natural conjugate prior assumption on the parameters (μ,δ) and $(\mu^\star,\delta^\star)$ taken independently yields the joint posterior density $p(\mu,\delta|Data)\cdot p(\mu^\star,\delta^\star|Data^\star)$, each factor in the form (2.5). Then taking $\delta=\delta^\star$ and integrating with respect to δ yields the density $p(\mu,\mu^\star|\delta-\delta^\star=0,Data,Data^\star)$. Finally, by changing variable to (ρ,μ),

$$\rho=\mu/\mu^*,$$

and integrating with respect to μ, we obtain

$$
\begin{aligned}
&p(\rho|\delta-\delta^*=0,Data,Data^*)\\
&=k''k^{*\prime\prime}\Gamma(n''+n^{*\prime\prime}-1)\left\{\rho^{n^{*\prime\prime}-1}(m''+m^{*\prime\prime}\rho)\right\}\\
&\cdot\left\{[n''\bar{t}''+n^{*\prime\prime}\bar{t}^{*\prime\prime}\rho-min(d'',d^{*\prime\prime})(m''+m^{*\prime\prime}\rho)]^{-(n''+^{*\prime\prime}-1}\right.\\
&\left.-[n''\bar{t}+n^{*\prime\prime}\bar{t}^{*\prime\prime}\rho]^{-(n''+n^{*\prime\prime}-1)}\right\}.
\end{aligned}
\tag{3.5}
$$

Expression (3.5) can be used to compute posterior credible regions of the form

$$p\{\rho \leq \rho_0 | \delta - \delta^* = 0, Data, Data^*\} \tag{3.6}$$

for fixed values of ρ_0.

The marginal posterior density of the common value δ in the case of $\delta = \delta^\star$ has the form,

$$\begin{aligned} p(\delta|\delta - \delta^* = 0, Data, Data^*) \propto & p(\delta|Data) \cdot p(\delta^* = \delta|Data^*) \\ & \propto S''(\delta)^{-n''} S^{*\prime\prime}(\delta)^{-n^{*\prime\prime}} \\ & 0 < \delta < min(d'', d^{*\prime\prime}) \end{aligned} \tag{3.7}$$

where S'' and $S^{\star\prime\prime}$ are the posterior version of S' and $S^{\star\prime}$ defined in Section 2.

Sometimes, it may even be of interest to infer whether $\mu \leq \mu^\star$ in the case of $\delta \neq \delta^\star$. Given our preceding prior distributional assumptions on (μ, δ) and $(\mu^\star, \delta^\star)$, the marginal posterior density of $\rho = \mu/\mu^\star$ is

$$\begin{aligned} p(\rho|Data, Data^*) =& k''k^{*\prime\prime}\Gamma(Q)(m''m^{*\prime\prime})^1 \rho^{n^{*\prime\prime}-2} \\ \cdot\{&[n''\bar{t}'' - m''d'' + (n^{*\prime\prime}\bar{t}^{*\prime\prime} - m^{*\prime\prime}d^{*\prime\prime})\rho]^{-Q} \\ & - [(n''\bar{t}'' - m''d'') + n^{*\prime\prime}\bar{t}^{*\prime\prime}\rho]^{-Q} \\ & - [n''\bar{t}'' + (n^{*\prime\prime}\bar{t}^{*\prime\prime} - m^{*\prime\prime}d^{*\prime\prime})\rho]^{-Q} \\ & + [n''\bar{t}'' + n^{*\prime\prime}\bar{t}^{*\prime\prime}\rho]^{-Q}\} \\ & forQ = n'' + n^{*\prime\prime} - 2. \end{aligned} \tag{3.8}$$

4 THE BAYES FACTOR.

A lack of realism in parameterized families of ordinary continuous-type joint prior distributions appears during attempts to assign appreciable prior probability to small neighborhood hypotheses. Such an attempt is obviously desirable, to model situations where the experimenter's prior belief includes appreciable suspicion (but not near certainty) that the unknown true sampling model is contained in such a specified neighborhood. If prior distributions express prior uncertainty, they should be able to express such a nontrivial suspicion. If the experimenter has prior uncertainty of this kind, he should use the following kind of inference method. If he does not, he should use other kinds of methods, as in the preceding sections. Consider an interval or "strip "hypothesis of the form

$$H_\varepsilon : |\eta - \eta_0| < \varepsilon. \tag{4.1}$$

The preceding joint prior densities give nearly zero prior probability to H_ε for small ε. Dickey (1976) corrected for this deficiency by considering a mixture of ordinary continuous distributions, which then has a mound of local high density over the region H_ε. However, working with a mixture is mathematically cumbersome. We approximate such a distribution by using a positive prior probability for the sharp hypothesis H,

$$H : \eta = \eta_0 vs. H^c : \eta \neq \eta_0. \tag{4.2}$$

The approximation of appreciable probability for H_ε (4.1) by positive probability for H (4.2) leads to inferences in terms of the Bayes factor or weighted likelihood ratio, as discussed by I. J. Good (1950), Savage et al. (1963) and Dickey and Lientz (1970). This ratio has the basic form,

$$\begin{aligned} B(H) = & \int lik(\tau, \eta = \eta_0|Data) p(\tau|H) d\tau \\ & / \int lik(\eta, \tau|Data) p(\eta, \tau|H^c) d\eta d\tau \end{aligned} \tag{4.3}$$

and was described by I. J. Good (1950) as the "factor in favor of the hypothesis, H, in virtue of the result of the experiment." Dickey and Lientz (1970) showed that the Bayes factor can be expressed as Savage's Density ratio:

$$B(H) = \lim_{\eta \to \eta_0}[p(\eta|H^c, Data)/p(\eta|H^c] \tag{4.4}$$

For our present sampling model, the Bayes factor can be calculated using our continuous-case prior densities in the form,

$$p(\eta_0|Data)/p(\eta_0). \tag{4.5}$$

Bartolucci and Dickey (1977) developed $B(H)$ for survival data problems in which the guarantee time is known.

One observes that if the prior degrees of freedom for $p(\eta)$ are taken to approach zero, i.e., $n^{'} = n^{\star'} \to 0$, then the posterior density corresponding to the non-integrable, "ignorance" prior density $p(\eta)$ proportional to $1/\eta$ is obtained immediately from our conjugate posterior density (3.3). But then the data becomes less and less likely under H^c, and $B(H) \to \infty$. However, if the prior density ordinate is not sent to zero but is rather assumed to be non-trivial and to locally approximately satisfy the same proportionality as for $n^{'} = n^{\star'} = 0$, namely $p(\eta|H^c) = constant/\eta$ (implying log η has a locally uniform prior density), the $B(H)$ takes the simple form,

$$B(H) = p(\eta_0|Data)/(constant/\eta_0). \tag{4.6}$$

In addition to the formal decision-theory criterion based on a loss function (Dickey and Leinz 1970), inference from the Bayes factor can also be carried out in an informal way by noting its rough order of magnitude. An extreme value will tend to indicate a strong inference common to a large range of opinions.

Of course, the Bayes factor gives the ratio of the posterior odds to the prior odds for $H(or H_e)$,

$$B(H) = [P(H|Data)/P(H^c|Data)]/[P(H)/P(H^c)],$$

and this ratio makes good inferential sense even without actually assigning a specific value to the prior probability $P(H)$.

The explanations and results of this section also apply verbatim to the hypothesis $H : \rho = \rho_0$ for our parameterization $\rho = \mu/\mu^\star$. Examples demonstrating the methods are given in the following section.

5 EXAMPLE.

Scientists at UAB initiated a study to compare the survival of two different breeds of mice exposed to salmonella typhimurium bacteria. This bacterium is most often the strain found in humans suffering from "food poisoning." Table 1 is a summary example of survival results at the time of analysis. Notice that the data follow a Type II censoring pattern.

To demonstrate the methods developed here, we have listed in Tables 3 and 4 Bayesian inference for the hypotheses $\eta_0 = 1(\eta = \tau/\tau^*$ and $\rho_0 = 1(\rho = \mu/\mu^*)$, respectively. We take sampling models, $T_i|(\mu, \delta) \sim$ Exponential $(\mu, \delta)(i = l, \ldots, 29)$ in Breed A and $T_j^*|(\mu^*, \delta^*) \sim Exponential(\mu^*, \delta^*)(j = l, \ldots, 35)$ for Breed B.

It is not uncommon in planning clinical trials for investigators to quote survival percentiles based on their previous experience in treating the particular disease type. The medical literature contains numerous examples of quotations of median survival and other percentiles, especially in articles on investigative therapies. (See, for example, Presant, C.A., et al. 1978 and Smalley, R.V. et al. 1976.)

One method of incorporating such information in prior probability assessments is seen in Birch and Bartolucci (1983). Expert opinions can be characterized by the predictive distribution of time to failure conditional on the hyper parameters. That is, the distribution after averaging out the sampling parameters according to their prior distribution. In the case of the exponential distribution (1.2) and the prior density (2.5) the predictive density takes the form

$$p(t|\bar{t}^{'}, n^{'}, m^{'}, d^{'}) \propto (n^{'}\bar{t}^{'} + t - (m^{'} + 1)d^{'})^{-n^{'}} - (n^{'}\bar{t}^{'} + t)^{-n^{'}} 0 < t < \infty \tag{5.1}$$

One can simply elicit percentiles of this predictive distribution from the experimenter and solve for the prior parameters. This basically involves eliciting four percentiles and solving interactively four equations with four unknown parameters. To do this we use a Newton-Raphson iterative technique. Let the f_i^{th} percentile of (5.1) be $t_i i = 1, \ldots, 4$. Define

$$\int_0^{t_i} p(t|\bar{t}^{'}, n^{'}, m^{'}, d^{'})dt - f_i = \phi_i(\bar{t}^{'}, n^{'}, m^{'}, d^{'}) i = 1, \ldots, 4. \tag{5.2}$$

The ususal methodology would be to expand (5.2) in Taylor Series of first order terms, i.e.

$$\begin{aligned} \phi_i(\bar{t}^{'}, n^{'}, m^{'}, d^{'}) &= \phi_i(\bar{t}_0, n_0, m_0, d_0) \\ &= \frac{\partial \phi_i}{\partial \bar{t}^{'}} \partial_{\bar{t}^{'}} + \frac{\partial \phi_i}{\partial n^{'}} \partial_{n^{'}} + \frac{\partial \phi_i}{\partial m^{'}} \partial_{m^{'}} + \frac{\partial \phi_i}{\partial d^{'}} \partial_{d^{'}} = 0 \\ i &=, \ldots, 4 \end{aligned}$$

where $\bar{t}_0, n_0, m_0$ and $d_0^{'}$ are initial values of $t^{'}, n^{'}, m^{'}$, and $d^{'}$, respectively and $\delta_\chi = \chi - \chi_0$. In our case we choose to condition on $d^{'}$ and $m^{'}$ to keep the number of equations and elicited percentiles to a manageable number. To expect to elicit more than two or three percentiles from an investigator may be unrealistic. In our example, we condition on $d^{'}$ and $m^{'}$ since they are most easily determined for our problem. That is to say, the choice $d^{'} = 5$ or $d^{'} = 8$ is most consistent with the minimum expected survival for the salmonella experiment and $m^{'}$ is chosen initially to be 30 since the value of $m^{'}$ will usually vary from 20 to 50. Thus, we can solve iteratively for $n^{'}$ and $\bar{t}^{'}$ given $d^{'}$ and $m^{'}$. We have an interative program that will perform this task. Kadane et al. (1980) investigated the assessment of the parameters of a matrix t distribution by a similar elicitation procedure.

In Table 2 we have the elicitation results for two sets of percentiles, (12, 62) and (10, 50), respectively. Primarily to demonstrate our analysis technique, in Tables 3 and 4 we held $n^{'}$ fixed for each set of percentiles and give results varying $m^{'}$. the lower and upper percentiles are given in parentheses in the $n^{'}$ column. The resulting prior densities form models of reasonable opinions for the experimental situation we are considering.

The assumption in table 3 is that of unequal and unknown locations $\delta \neq \delta^\star$, with independent prior distributions for (μ, δ) and $(\mu^\star, \delta^\star)$,namely (2.5) and the starred version of (2.5), respectively. The hypothesis of interest is $H : \eta = 1$. Expression (3.3) is the density kernel used for calculating the posterior tail areas (column 3 for $\eta_0 = 1$). These tail probabilities and Bayes factors cohere with the various types of prior input as listed and with $d^{'} = 5, \quad d^{''} = min \quad (5, 8) = 5 \quad \text{and} \quad d^{\star'} = 8, \quad d^{\star''} = min(8, 8) = 8$.

For illustration we have taken equality of the prior parameters, $m^{'} = m^{\star'}, n = n^{\star'}$, and $\quad \bar{t}^{'} = \bar{t}^{\star'}$ although such an assumption is not necessary.

The prior probabilities for H are zero in the calculation of the posterior tail area, and arbitrary in the calculation of the Bayes factor, where they do not enter. The posterior tail probabilities in Table 3 range from 0.000006 to 0.048608. The Bayes factors, $B(H)$, were computed as in expression (4.4) and are roughly consistent with the interpretation of the tail areas, in that we have moderate evidence against $\eta_0 = 1$ as $B(H)$ ranges from 0.00010 to 0.36098.

In table 4, equations (3.5) and (4.4), with $\rho = \mu/\mu^\star$ in place of η, were used in the computations of columns 3 and 4 respectively. The assumption here is that of equal but unknown locations, $\delta = \delta^\star$, and the hypothesis of interest is $H : \rho = 1$. The realistic choice of $d^{'} = d^{\star'} = 8, d^{''} = 8$ leads to results fairly consistent with Table 3 in that there is moderate evidence against $\rho_0 = 1$. The results of both tables are fairly robust when other values in the neighborhood of $\eta_0 = 1 \quad or \quad \rho_0 = 1$ considered.

Upon further examination of the Bayes factor in each table for the case $m^{'} = m^{\star'} = 25$, one notes that as $n^{'}$ and $n^{\star'}$ decreases, $B(H)$ increases. For the case $n^{'} = n^{\star'} = 0$, our effective prior uncertainty range in either case of $\eta \quad or \quad \rho$ is chosen to be the closed interval [1/5, 5], implying $p(log_{10}\eta|H^c) = 0.716$ or equivalently, $p(\eta|H^c) = 0.716/\eta$ or $p(\rho|H^c) = 0.716/\rho$.

For traditional inference for testing $\rho_0 = 1.0$, we have available the "approximate"F distribution of Gehan (1965) or Cox (1953) which, based on our data, yields the statistic

$$(33/)[n\bar{t} - 29t_{(1)}]/[n^*\bar{t}^* - 35t^*_{(1)}] = 1.62 \tag{5.3}$$

for $t(1) = 10$ and $t^{\star}_{(1)} = 8$. The tabled 5 percent critical value on 54 and 66 degrees of freedom is about 1.56. Thus, classically, we also have rejection of $\rho_0 = 1$ at the 0.05 level, which harmonizes with our "exact "Bayesian results.

The inference problem for which the traditional expression (5.3) may be considered is one of the "dividing "type, $H : \rho \leq \rho_0$ (Cox 1977). The posterior tail area could be considered as a Bayesian analogue for this situation (Box and Tiao 1973). However, it does not involve positive prior probability on H as does the Bayes factor. Different prior distributions for different stages of prior uncertainty are involved in the two Bayesian methods, as explained at the beginning of Section 4. Examining Table 4, one also sees an instance of the general tendency for traditional tests to reject more readily than do reasonable thresholds on Bayes factors (Lindley 1957; Dickey 1976). Figures 1 to 4 give the posterior densities for the cases in Tables 3 and 4 for which $m^{'} = m^{\star'} = 30.0$. For the sake of standardization each density point was divided by the maximum value. The overall conclusion from the analysis is that the Breed A mice appear to survive the salmonella exposure for longer periods of time than do the Breed B mice.

We have given here a methodology for Bayesian inference on the survival mean parameters $\tau = \mu + \delta$ and $\tau^{\star} = \mu^{\star} + \delta^{\star}$ for $\delta \neq \delta^{\star}$, whereas most traditional comparisons of exponential samples either impose the restriction $\delta = \delta^{\star}$ or impose appropriate first-order-statistic estimates on δ and $\delta^{\star}$ and then proceed with approximate tests of $H : \rho_0 = 1$, especially in the cases involving censoring. In this paper we have overcome these constraints and provided "exact "Bayesian inference procedures for the parameters of interest. Our methods help answer the need for a cautious coherent approach to analyzing data with proper consideration of the influence of relevant predata uncertainty for reasonable differences in prior opinions.

TABLE 1.

Survival Time in Days.

Breed A			Breed B		
10	16	29	8	12	15
10	16	34	9	12	15
11	16	34	9	12	16
11	18	37+	9	12	18
11	18	50+	9	13	18
11	19		10	13	19
11	19		10	13	19
14	23		11	13	23
15	23		11	14	23
15	23		11	14	23+
16	23		12	14	27+
16	23		12	14	

$$m = 29 \qquad \bar{t} = 21.185 \qquad m^\star = 35 \qquad t^\star = 14.939$$
$$n = 27 \qquad t_{(1)} = 10.0 \qquad n^\star = 33 \qquad t^\star_{(1)} = 8.0$$

+ indicates a censored observation

TABLE 2.

Percentiles		Prior Shape	Prior Mean
0.25	0.75	n'	$\bar{t}'$
12	62	8.018	19.687
10	50	15.023	11.763

TABLE 3.

Unknown Unequal Locations.

$$d'' = 8.0 \qquad d^{\star''} = 5.0 \qquad \eta = \tau/\tau^\star$$

$n' = n^{\star'}$	$m' = m^{\star'}$	$\bar{t}' = \bar{t}^{*'}$	$P[\eta < 1 \mid Data, Data^\star]$	$B(H)$ for $H : \eta = 1$
0	25		0.048608	0.36098
8.018	25	19.687	0.02555	0.22527
$(12, 62)$				
	30	22.805	0.01593	0.15085
	40	29.041	0.00508	0.05541
15.023	25	11.763	0.000069	0.00105
$(10, 50)$				
	30	13.4276	0.000032	0.00052
	40	16.7558	0.000006	0.00010

TABLE 4.

Unknown Equal Locations.

$$d^{''} = d^{\star''} = 8.0 \qquad \rho = \mu/\mu^{\star}$$

$n^{'} = n^{\star'}$	$m^{'} = m^{\star'}$	$\bar{t}^{'} = \bar{t}^{*'}$	$P[\rho < 1 \mid \delta = \delta^{\star}, Data, Data^{\star}]$	$B(H)$ for $H : \rho = 1$
0	14		.005324	0.77740
8.018	25	29.0411	0.04196	0.49886
$(12, 62)$				
	30	34.0299	0.04095	0.48878
	40	44.0075	0.03923	0.47168
15.023	25	16.7558	0.01088	0.20527
$(10, 50)$				
	30	19.41784	0.01066	0.20168
	40	24.7435	0.00997	0.18995

FIGURE 1
POSTERIOR DENSITY

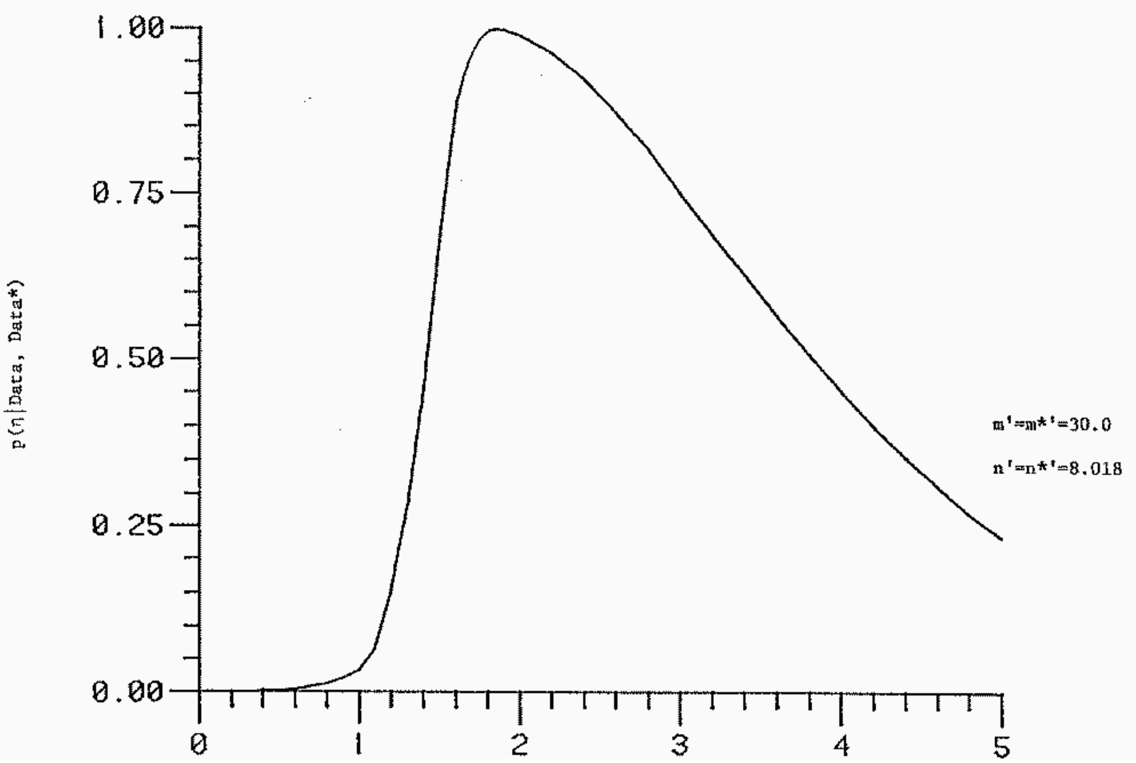


FIGURE 2
POSTERIOR DENSITY

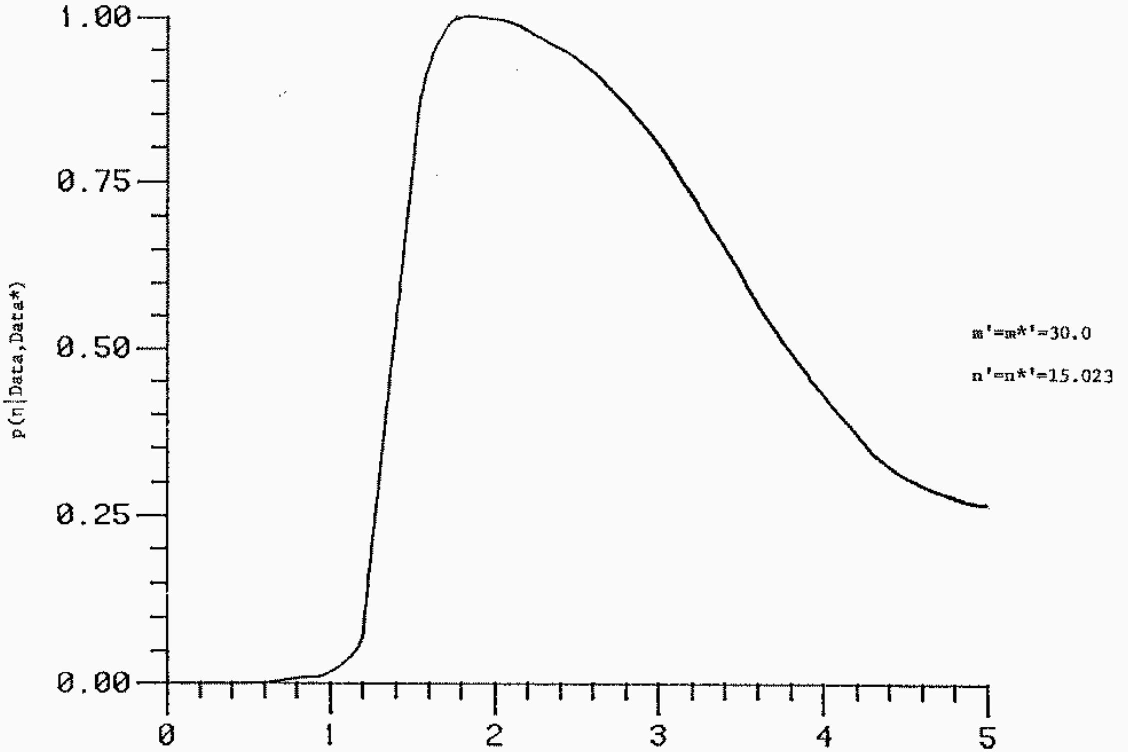


FIGURE 3
POSTERIOR DENSITY

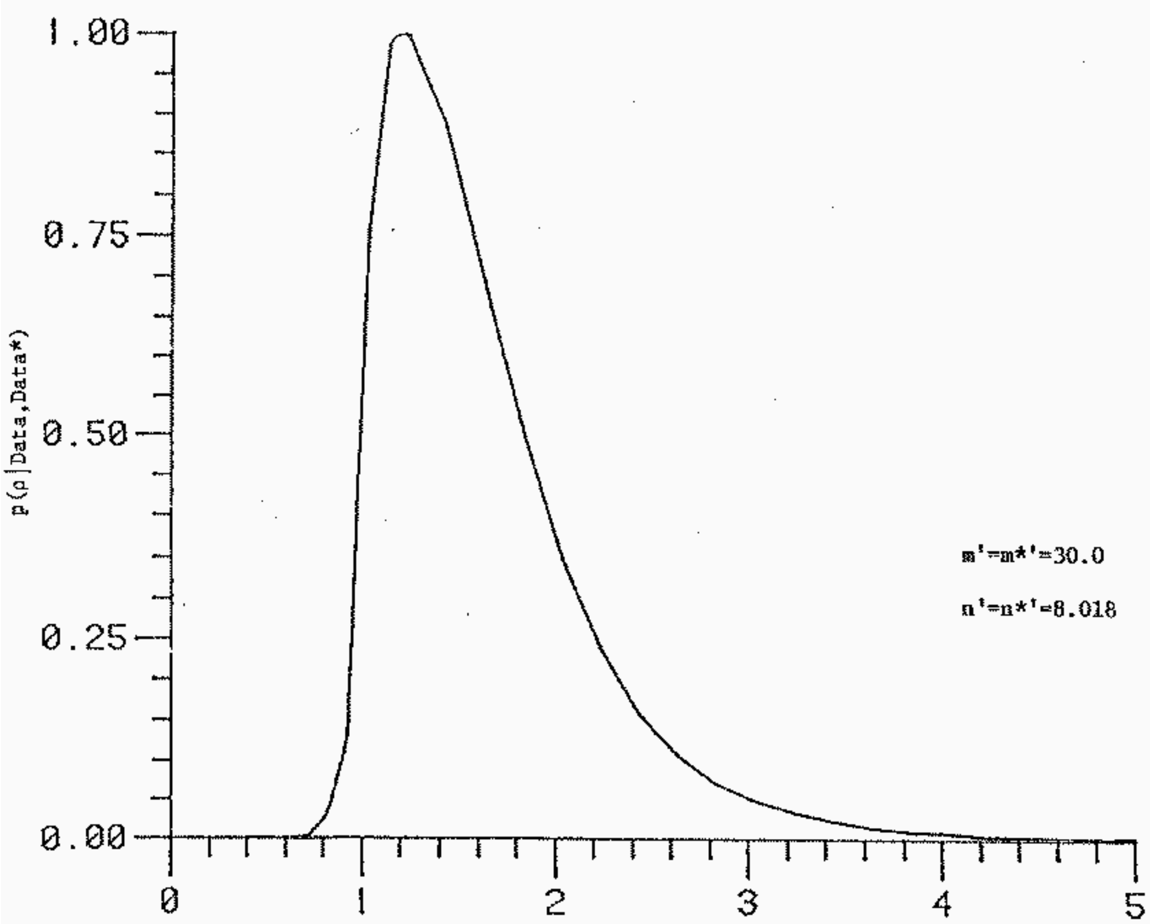
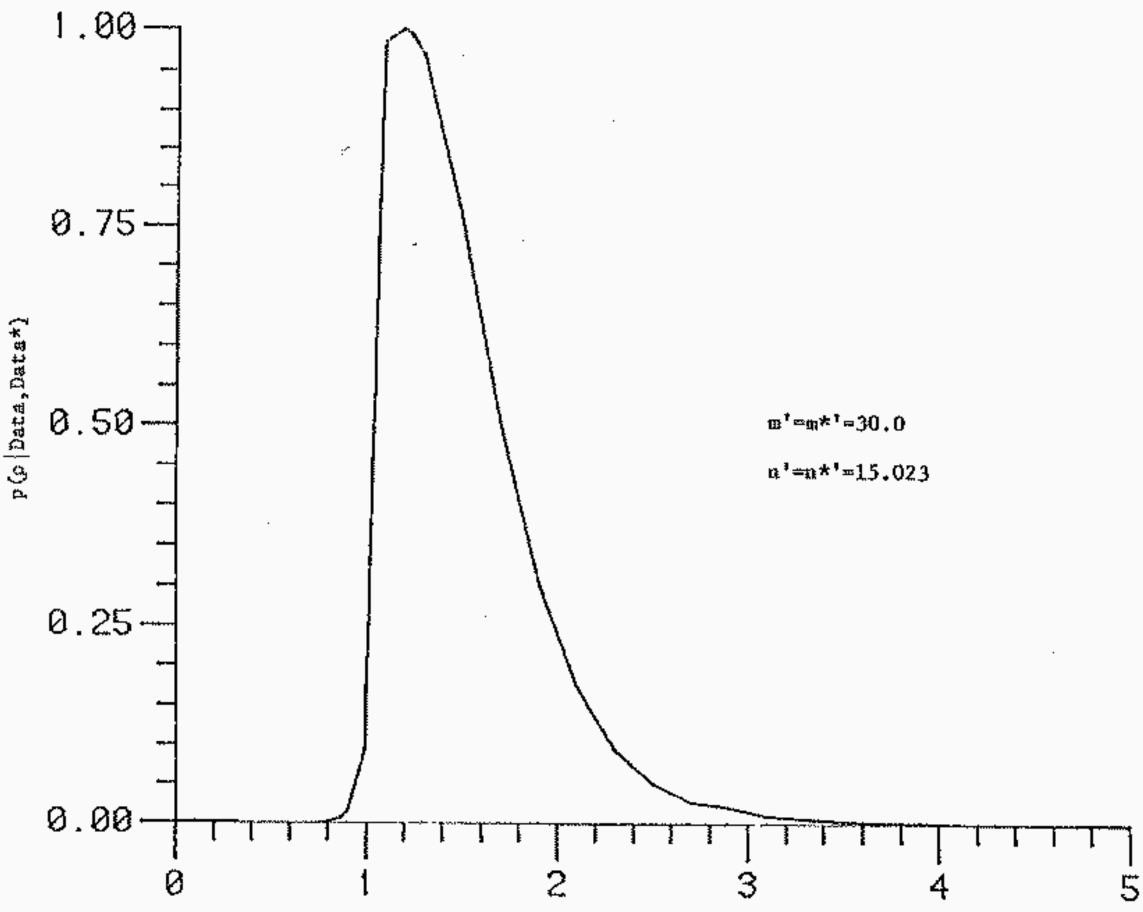


FIGURE 4
POSTERIOR DENSITY



References

Abramowitz, M., Stegun, I.A. (editors), Handbook of Mathematical Functions (1964), Dover Publications Inc., New York..

Bain, L.J., *Theory and Methods*, Statistical Analysis of Reliability and Life-Testing Models **24** (1978), Marcel Dekker Inc., 136-142.

Bartolucci, A.A. and Dickey, J.M., *Comparative Bayesian and Traditional Inference for Gamma-modled Survival Data*, Biometrics **33** (1977), 343-354.

Birch, R. and Bartolucci, A.A., *Determination of the Hyperparameters of a Prior Probability Model in Survival Analysis*, Computer Programs in Biomedicine **17** (1983), 98-94.

Box, G.E.P. and Tiao, George C., Bayesian Inference in Statistical Analysis (1973), Addison-Wesley Publ. Co., Reading, Massachusetts.

Cox, D.R., *Some Simple Approximate Tests for Poisson Variates*, Biometrika **40** (1953), 354-360.

Cox, D.R., *Regression Models and Life Tables (with Discussion)*, Journal of the Royal Statistical Society, Series B **34** (1972), 187-220.

Cox, D.R., *The Role of Significance Tests*, Scand. J. of Statistics **4** (1977), 49-70.

Dickey, J.M., *Scientific Reporting and Personal Probabilities: Student's Hypothesis*, Journal of the Royal Statistical Society, Series B **35, No.2** (1973), 285-305.

Dickey, J.M., *Approximate Posterior Distributions*, Journ. of the Amer. Statist. Assoc. **71, No.355** (1976), 680-689.

Dickey, J.M., *" Beliefs about Beliefs, A Theory for Stochastic Assessment of Subjective Probabilities." Invited paper given at the International Meeting on Bayesian Statistics, Valencia, Spain, May 28 - June 2, 1979* (1979), Published proceedings of Trabajos Estadistica.

Dickey, J.M. and Lientz, B.P., *The Weighted Likelihood Ratio, Sharp Hypothesis about Chances, The Order of a Markov Chain*, Annals of Mathematical Statistics **41** (1970), 214-226.

Gehan, E.A., *A Generalized Wilcoxon Test for Comparing Arbitrarily Singly Censored Samples*, Biometrika **52** (1965), 203-223.

Good, I.J., Probability and the Weighing of Evidence (1950), Charles Giffin, London.

Gunel, E., and Dickey, J.M., *Bayes Factors for Independence in Contingency Tables*, Biometrika **61, NO. 3** (1974), 545-557.

Halpern, M., Lan K.K.G., Ware, J.H., Johnson, N.J., DeMats, D.L., *An Aid to Data Monitoring in Long Term Clinical Trials*, Controlled Clinical Trials **3** (1982), 311-323.

Kadane, J.B., Dickey, J.M., Winkler, R.L., Smith, W.S., and Peters, S.C., *Interactive Elicitation of Opinion for a Normal Linear Model*, Journ. Amer. Statist. Assoc. **75** (1980), 845-854.

Kaplan, E.L. and Meier, P., *Non-parametric Estimation from Incomplete Observations*, Journ. Amer. Statist. Assoc. **53** (1958), 457-481.

Lachin, J.M., Lan Shu-Ping, and the Lupus Nephritis Collaborative Study Group, *Termination of a Clinical Trial with No Treatment Group Difference: The Lupus Nephritis Collaborative Study*, Controlled Clinical Trials **13** (1992), 62-79.

Lan, K.K.G., Simon, R., Halpern, M., *Stochastically Curtailed Tests in Long Term Clinical Trials*, Communications in Statistics **1** (1983), 207-219.

Lindley, D.V., *A Statistical Paradox*, Biometrika **44** (1957), 187-192.

Presant, C.A., Batolucci, A.A., Smalley, R.V. and Vogler, W.R., Edited by W.D. Terry and D. Windhorst, *Effects of Corynebacterium Parvum on Combination Chemotherapy of Disseminated Malignant Melanoma*, Immunotherapy of Cancer: Present Status of Trials in Man, Raven Press, New York, 113-121.

Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory (1961), M.I.T. Press, Cambrige, Massachusetts.

Savage, L.J. et al., *Bayesian Statistical Inference for Psychological Research*, Psychological Review **70** (1963), 193-242.

Smalley, R.V. et al., *Combination versus Sequential Five Drug Chemotherapy in Metastatic Carcinoma of the Breast*, Cancer Research **36** (1976), 3911-3916.

Zelen, M., *Applications of Exponential Models to Problems in Cancer Research*, Journ. Royal Statistical Soc. A. **129** (1966), 368-398.

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Issue No. 02, 1997, pp. 24–28.
Submitted: October 5, 1997. Published: July 31, 1998.

COMPARING THE RADII OF SOME BALLS APPEARING IN CONNECTION TO THREE LOCAL CONVERGENCE THEOREMS FOR NEWTON'S METHOD

Ioannis K. Argyros I.

Abstract. We provide three local convergence theorems for Newton's method in Banach space. These theorems use information on the first, second, and first-second Fréchet derivatives of the operator involved, respectively. In all cases, we show that for any starting point inside a ball centered at the solution of the nonlinear equation, Newton's method converges to that solution. Our goal is to obtain the smallest and the largest convergence region for method (2). Our results apply to solve autonomous differential equations. A numerical example is also provided.

1991 A.M.S. (MOS) Subject Classification Codes. 65J15, 65B05, 47H17, 49D15.

Key Words and Phrases. Newton's method, Banach space, local convergence theorem.

1. INTRODUCTION.

In this study we are concerned with the problem of approximating a locally unique solution x^* of the nonlinear equation

$$F(x) = 3D0, \tag{1}$$

where F is a twice Fréchet-differentiable operator defined on a convex subset D of a Banach space E_1 with values in a Banach space E_2.

The most popular method for generating a sequence $\{x_n\}$ $(n \geq 0)$ converging to x^* is undoubtedly Newton's method

$$x_{n+1} = 3Dx_n - F'(x_n)^{-1}F(x_n) \quad (n \geq 0)\ (x_0 \in D). \tag{2}$$

Department of Mathematics, Cameron University, Lawton, OK 73505 USA
E-mail Address: ioannisa@cameron.edu

Here $F'(x) \in L(E_1, E_2)$ $(x \in D)$ the space of bounded linear operators from E_1 into E_2. Sufficient convergence conditions for the convergence of Newton's method (2) to x^* have been given by many authors [1], [3], [5].

In this study, we assume the existence of a solution $x^* \in D$ of equation (1), and using Lipschitz conditions on the first and second Fréchet-derivative we provide three local convergence theorems for method (2). Each theorem guarantees the convergence of Newton's method (2) to x^* as long as the starting point x_0 is inside a ball centered at x^* and of a certain radius. Our goal is to obtain the smallest and the largest convergence region for method (2). Our results have theoretical value, but also practical use, since they can be applied to solve autonomous differential equations [1], [2], [4].

Finally, we complete this study with a numerical example.

2. CONVERGENCE ANALYSIS.

Let $x^* \in D$ be a simple solution of equation $F(x) = 3D0$ in the sense that $F'(x^*)$ has a bounded inverse. We denote by $U(x^*, r)$ the open ball with center x^* and radius $r > 0$. Let also α, β, γ be constants such that

$$\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq \alpha\|x - y\|, \tag{3}$$

$$\|F'(x^*)^{-1}F''(x^*)\| \leq \beta, \tag{4}$$

and

$$\|F'(x^*)^{-1}(F''(x) - F''(y))\| \leq \gamma\|x - y\| \tag{5}$$

for all $x, y \in D$. Define parameters r_1, r_2, r_3 by

$$r_1 = 3D\frac{2}{3\alpha}, \tag{6}$$

$$r_2 = 3D\frac{12}{9\beta + \sqrt{81\beta^2 + 96\gamma}}, \tag{7}$$

$$r_3 = 3D\frac{12}{3(\beta + 2\alpha) + \sqrt{9(\beta + 2\alpha)^2 + 24\gamma}}, \tag{8}$$

and functions f_1, f_2, f_3 on $[0, +\infty)$ by

$$f_1(r) = 3D3ar - 2, \tag{9}$$

$$f_2(r) = 3D4\gamma r^2 + 9\beta - 6 \tag{10}$$

$$f_3(r) = 3D\gamma r^2 + 3(\beta + 2\alpha)r - 6. \tag{11}$$

Using (2) we can have the approximations:

(12)

$$\begin{aligned} x_{n+1} - x^* = 3D[F'(x_n)^{-1}F'(x^*)]\Big\{\int_0^1 F'(x^*)^{-1}[F'(x^* + t(x_n - x^*)) \\ - F'(x_n)]\,dt(x_n - x^*)\Big\}, \end{aligned}$$

(13)
$$x_{n+1} - x^* = 3D[F'(x_n)^{-1}F'(x^*)]\Big\{\int_0^1 [F'(x^*)^{-1}(F''(x_n + t(x^* - x_n))$$
$$- \; F''(x_0))(1-t)\,dt(x^* - x_n)^2 + \frac{1}{2}F'(x^*)^{-1}F''(x_0)(x^* - x_n)^2\Big\}.$$

Consider also the identities

$$F'(x^*)^{-1}F'(x) = 3DI + F'(x^*)^{-1}(F'(x) - F'(x^*)) \tag{14}$$

and

$$\int_{x^*}^{x} F'(x^*)^{-1}[F''(y) - F''(x^*)]\,dy \tag{15}$$
$$= 3D\; F'(x^*)^{-1}F'(x) - I - F'(x^*)^{-1}F''(x^*)(x - x^*).$$

Then we can obtain the following local convergence theorems for Newton's method (2).

Theorem 1. *Let $F\colon D \subseteq E_1 \to E_2$ be Fréchet differentiable, and x^* be a simple solution of equation $F(x) = 3D0$. Assume Lipschitz condition* (3) *holds, and $U(x^*, r_1) \subseteq D$, where r_1 is given by* (6)*. Then, for any $x_0 \in U(x^*, r_1)$, Newton's method* (2) *converges to x^*. Moreover the following error bounds hold for all $n \geq 0$*

$$\|x_{n+1} - x^*\| \leq \frac{\alpha}{2(1 - \alpha\|x_n - x^*\|)}\|x_n - x^*\|^2. \tag{16}$$

Proof. The proof was essentially given in [6] and uses (3), (12), (14). However it was not given in affine invariant form as it is here. The advantages of such a form have been explained in some detail in [1], [3].

Theorem 2. *Let $F\colon D \subseteq E_1 \to E_2$ be twice Fréchet-differentiable, and x^* be a simple solution of equation $F(x) = 3D0$. Assume conditions* (4), (5) *hold, and $U(x_0, r_2) \subseteq D$, where r_2 is given by* (7)*. Then, for any $x_0 \in U(x_0, r_2)$, Newton's method* (2) *converges to x^*. Moreover the following error bounds hold for all $n \geq 0$*

$$\|x_{n+1} - x^*\| \leq \frac{\frac{\gamma}{3}\|x_n - x^*\| + \beta}{2\big[1 - \big(\frac{1}{2}\gamma\|x_n - x^*\|^2 + \beta\|x_n - x^*\|\big)\big]}\|x_n - x^*\|^2. \tag{17}$$

Proof. For $x_0 \in U(x_0, r_2)$ using induction on $n \geq 0$, (14), and the Banach lemma on invertible operators [2], [5], we obtain $F'(x_n)^{-1}$ exists for all $n \geq 0$,

$$\|F'(x_n)^{-1}F'(x^*)\| \leq \left[1 - \left(\frac{1}{2}\gamma\|x_n - x^*\| + \beta\right)\|x_n - x^*\|\right]^{-1} \tag{18}$$

and

$$\left\|\int_0^1 F'(x^*)^{-1}[F''(x_n + t(x^* - x_n))(1-t)\,dt(x^* - x_n)^2\right\| \tag{19}$$
$$\leq \; \frac{1}{2}\left(\frac{\gamma}{3}\|x_n - x^*\| + \beta\right)\|x_n - x^*\|^2.$$

By (13), (18) and (19) we obtain (17). The choice of r_2 guarantees that there exists $c \in [0,1]$ such that

$$\|x_{n+1} - x^*\| \leq c\|x_n - x^*\| \leq c^{n+1} r_2 \to 0 \quad \text{as } n \to \infty.$$

Hence, we get $\lim\limits_{n\to\infty} x_{n+1} = 3Dx^*$.

That completes the proof of the theorem. ♠

Theorem 3. *Let $F: D \subseteq E_1 \to E_2$ be twice Fréchet-differentiable, and x^* be a simple solution of equation $F(x) = 3D0$. Assume conditions (3)–(5) hold, and $U(x_0, r_3) \subseteq D$, where r_3 is given by (18). Then for any $x_0 \in U(x_0, r_3)$, Newton's method (2) converges to x^*. Moreover the following error bounds hold for all $n \geq 0$:*

$$\|x_{n+1} - x^*\| \leq \frac{\frac{\gamma}{3}\|x_n - x^*\| + \beta}{2(1 - a\|x_n - x^*\|)}\|x_n - x^*\|^2. \tag{20}$$

Proof. As in the proof of Theorem 2 but we use (14) instead of (15).

That completes the proof of Theorem 3. ♠

Remark 1. It follows from (3)–(8) that when $\gamma = 3D0$, then $\beta = 3D\alpha$ and $r_1 = 3Dr_2 = 3Dr_3$. Many other possibilities exist. For example $f_2(r_1) < 0$, $f_2(r_3) < 0$ and $f_3(r_1) < 0$ imply $r_1 < r_3 < r_2$ (see the example that follows).

Remark 2. The results obtained here have theoretical and practical value. Indeed, in practice they can be used to solve autonomous differential equations of the form

$$F'(x) = 3DT(F(x)) \quad (x \in D), \tag{21}$$

where T is a known operator [2], [4].

We note that (21) gives $F'(x^*) = 3DT(F(x^*)) = 3DT(0)$, which is known. So we can use Theorems 1–3, even when the solution x^* is not known.

We conclude this study with an example involving Theorems 1–3 and Remarks 1 and 2.

Example. Let $E_1 = 3DE_2 = 3D\mathbf{R}$, $D = 3D[-1,1]$. Define the function F by

$$F(x) = 3De^x - 1, \quad (x \in D). \tag{22}$$

Choose $T(x) = 3Dx + 1$ $(x \in D)$. Then F so defined satisfies (21). Using (3)–(8), (22), we get $F'(x^*) = 3DT(0) = 3D1$, $\alpha = 3D\gamma = 3De$, $\beta = 3D1$,

$$r_1 = 3D.245253, \quad r_2 = 3D.4364902 \quad \text{and} \quad r_3 = 3D.2982063.$$

Therefore the conclusions of Theorems 1–3 hold. In particular, for $x_0 \in U(x_0, s)$ with $s = 3Dr_1$ or r_2 or r_3, Newton's method converges to the solution $x^* = 3D0$ of equation $F(x) = 3D0$, where F is given by (22).

REFERENCES

[1] Argyros, I.K. and Szidarovszky, F., *The Theory and Applications of Iteration Methods*, CRC Press, Inc. Boca Raton, Florida, U.S.A., 1993.

[2] Davis, H.T., *Introduction to Nonlinear Differential and Integral Equations*, Dover Publ., New York, NY, 1962.

[3] Deuflhard, P. and Heindl, G.A., Affine invariant convergence theorems for Newton's method and extension to related methods, *SIAM J. Num. Anal.* **16** 1, (1979), 1–10.

[4] Dickson, L.E., *New First Course in the Theory of Equations*, John Wiley & Sons, New York, NY, 1939.

[5] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis*, Pergamon Press, New York, NY, 1982.

[6] Rheinboldt, W.C. An adaptive continuation process for solving systems of nonlinear equations, *Polish Academy of Science, Banach Ctr. Publ.* **3** (1977), 129–142.

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISBN 1083-0464
Issue No. 1 July 1998, pp. 29 - 43
Submitted: June 3, 1997. Published: July 31, 1998

Exact Internal Controllability of the Elasticity System

V. N. Domingos Cavalcanti, J. S. Prates Filho

Abstract

We study the exact internal controllability of the system:

$$\varphi'' - \mu\Delta\varphi - (\lambda+\mu)\nabla(\alpha(x) div\, \varphi) = h \quad \text{in } \Omega\times]0,T[,$$

where Ω is a bounded domain in $\mathbf{R}^n$, λ and μ are the Lamè's constants, and α is a positive real function defined in $\overline{\Omega}$. The result is obtained by applying HUM, Hilbert Uniqueness Method, due to J. L. Lions.

AMS Subject Classification: 93C20, 35L15
Key words: Exact internal controllability, HUM, Elasticity System.

1. Introduction

Let Ω be a bounded domain in $\mathbf{R}^n$ with a C^2-boundary Γ and let ω be a nonempty open subset of Ω. For $T > 0$, we denote by Q the finite cylinder $\Omega\times]0,T[$ and by Σ its lateral boundary $\Gamma\times]0,T[$.

Departamento de Matemática - Universidade Estadual de Maringá 87020-900
Maringá - PR, Brazil

In this paper we are concerned with the exact internal controllability of the elasticity system:

$$(1.1)\qquad \left| \begin{array}{l} \varphi'' - \mu\Delta\varphi - (\lambda+\mu)\nabla(\alpha(x)div\varphi) = h\chi_\omega \quad \text{in} \quad Q \\ \varphi = 0 \quad \text{on} \quad \Sigma \\ \varphi(0) = \varphi^0, \quad \varphi'(0) = \varphi^1 \quad \text{in} \quad \Omega, \end{array} \right.$$

where χ_ω denotes the characteristic function of ω, $\lambda > 0$ and $\mu > 0$ are the Lamè's constants, and α is a positive real function defined in $\overline{\Omega}$.

The exact controllability problem for the above system is formulated as follows:

" Given $T > 0$, find a Hilbert space $\mathcal{H}$ such that for each pair $\{\varphi_0, \varphi_1\} \in \mathcal{H}$ there exists a control function $h \in L^2(0,T;(L^2(\omega))^n)$ such that the solution φ of the system (1.1) satisfies:

$$\varphi(T) = \varphi'(T) = 0."$$

Exact controllability of the elasticity system was firstly studied by J. L. Lions [10]. Later, R. Fuentes Apolaya [5] considered the problem in noncylindrical domains and E. Zuazua [16] attacked the question of the approximate controllability. However, all of them assume $\alpha(x) \equiv$ constant. The main difficulty found in the resolution of the problem lies in the fact that we cannot apply the Holmgren's Theorem to differential operators with non constant coefficients. To bypass this difficulty, we establish a new argument based on regularization techniques. We refer the reader to [2], [3], [4], [6], [7] and [9] for other works on exact controllability.

In Section 2 we establish our main result Theorem 2.1. By assuming an "inverse inequality", we prove Theorem 2.1 with application of HUM. The arguments are standard, yet nontrivial. The proof of the "inverse inequality" is given in Section 3 by means of an argument which is in part related to the one used in Prates Filho-Soriano [13].

2. The Main Result and Application of HUM

For a fixed $x^0 \in \mathbf{R}^n$ we define

$$m(x) = x - x^0, \quad \forall\, x \in \overline{\Omega}$$

and

$$R(x^0) = \sup_{x\in\overline{\Omega}} \|m(x)\| = \sup_{x\in\overline{\Omega}} (\sum_{i=1}^{n} |x_i - x_i^0|^2)^{\frac{1}{2}}.$$

We also define

$$\Gamma_0 = \Gamma(x^0) = \{x \in \Gamma;\, m(x)\cdot\nu(x) > 0\}$$

and

$$\Gamma_*(x^0) = \Gamma \backslash \Gamma_0,$$

where $\nu(x)$ denotes the unit exterior normal at the point x of Γ.

We assume that $\alpha \in C^1(\overline{\Omega})$ and there exists some $\alpha_0 > 0$ such that

$$\alpha(x) \geq \alpha_0 \quad \forall\, x \in \overline{\Omega} \quad \text{and} \quad \|\nabla\alpha\|_{L^\infty(\Omega)} < \frac{\alpha_0}{R(x^0)}. \tag{2.1}$$

The following functional spaces will be considered:

$$H = [L^2(\Omega)]^n, \quad (u, v)_H = \sum_{i=1}^{n} (u_i, v_i) \quad \forall\, u, v \in H,$$

$$V = [H_0^1(\Omega)]^n, \quad ((u, v))_V = \sum_{i=1}^{n} (\nabla u_i, \nabla v_i) \quad \forall\, u, v \in V,$$

where $(\cdot, \cdot)$ denotes the standard inner-product of $L^2(\Omega)$. If we denote by V' the topological dual of V, we then have

$$V' = [H^{-1}(\Omega)]^n.$$

Now we are ready to state our main result.

Theorem 2.1 *Assume that condition* (2.1) *holds for a fixed* $x_0 \in \mathbf{R}^n$ *and let* ω *be a neighbourhood of* $\overline{\Gamma}_0$*, that is, there exists an open subset* $\mathcal{O}$ *of* $\mathbf{R}^n$ *such that* $\overline{\Gamma}_0 \subset \mathcal{O}$ *and* $\omega = \Omega \cap \mathcal{O}$.

Then given $T > T_0$ *where*

$$T_0 = \frac{2R(x^0)}{\sqrt{\mu}(1 - ||\nabla\alpha||_{L^\infty(\Omega)} R(x^0)\alpha_0^{-1})}, \tag{2.2}$$

there exists, for each pair $\{\varphi^0, \varphi^1\} \in H \times V$*, a control* $h \in L^2(0, T; (L^2(\omega))^n)$ *such that the solution* φ *of the problem* (1.1) *satisfies*

$$\varphi(T) = \varphi'(T) = 0.$$

Proof. First we consider a pair $\{\phi^0, \phi^1\} \in [D(\Omega)]^n \times [D(\Omega)]^n$. Then from the smoothness of the inital data, the homogeneous system

$$\left| \begin{array}{l} \phi'' - \mu\Delta\phi - (\lambda + \mu)\nabla(\alpha(x) div \phi) = 0 \quad \text{in} \ \ Q \\ \phi = 0 \quad \text{on} \ \ \Sigma \\ \phi(0) = \phi^0, \quad \phi'(0) = \phi^1 \quad \text{in} \ \ \Omega, \end{array} \right. \tag{2.3}$$

has a unique regular solution. Next we consider the backward problem

$$(2.4)\qquad \left|\begin{array}{l} \psi'' - \mu\Delta\psi - (\lambda+\mu)\nabla(\alpha(x)div\psi) = -\phi\chi_\omega \quad \text{in } Q \\ \psi = 0 \quad \text{on } \Sigma \\ \psi(0) = \psi^0, \quad \psi'(0) = \psi^1 \quad \text{in } \Omega, \end{array}\right.$$

where ϕ is the unique solution of (2.3).

Choosing $\{\psi^0, \psi^1\} \in V \times H$ and taking into account the regularity of ϕ, problem (2.4) has a unique solution ψ satisfying

$$\psi \in C^0([0,T];V) \cap C^1([0,T];H).$$

Therefore the mapping $\Lambda : [D(\Omega)]^n \times [D(\Omega)]^n \to H \times V$ given by

$$(2.5)\qquad \Lambda\{\phi^0, \phi^1\} = \{\psi'(0), -\psi(0)\}$$

is well defined.

On the other hand, if we choose $\{\phi^0, \phi^1\} \in H \times V'$, we get the so called ultra-weak solution ϕ of (2.3) by a transposition argument (cf. Lions [10]). More precisely, $\phi \in L^\infty(0,T;H)$ and satisfies

$$\sum_{i=1}^n \int_Q \phi_i f_i dxdt = \langle \phi^1, \zeta(0)\rangle_{V',V} - (\phi^0, \zeta'(0))_H,$$

for every weak solution of

$$\left|\begin{array}{l} \zeta'' - \mu\Delta\zeta - (\lambda+\mu)\nabla(\alpha(x)div\zeta) = f \quad \text{in } Q \\ \zeta = 0 \quad \text{on } \Sigma \\ \zeta(T) = \zeta'(T) = 0 \quad \text{in } \Omega, \end{array}\right.$$

where $\zeta \in L^1(0,T;H)$. Then, in this case, the ultra-weak solution ϕ of (2.3) satisfies

$$(2.6)\qquad \sum_{i=1}^n \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt = (\phi^0, \psi'(0))_H + \langle \phi^1, \psi(0)\rangle_{V',V}.$$

This is done by multiplying ψ and the equation in (2.3), and noting that $[D(\Omega)]^n \times [D(\Omega)]^n$ is dense in $H \times V'$.

Next, we define the bilinear form $(\cdot,\cdot)_* : ([D(\Omega)]^n)^2 \times ([D(\Omega)]^n)^2 \to \mathbf{R}$ by

$$(2.7)\qquad (\{\phi^0, \phi^1\}, \{\zeta^0, \zeta^1\})_* = \sum_{i=1}^n \int_0^T \int_\omega \phi_i \zeta_i dxdt$$

where ϕ and ζ are the solutions of (2.3) associated to the initial data $\{\phi^0, \phi^1\}$ and $\{\zeta^0, \zeta^1\}$, respectively. The functional given in (2.7) defines a inner product in $[(D(\Omega))^n]^2 \times [(D(\Omega))^n]^2$ and, consequently, we can consider the following norm in $[D(\Omega)]^n \times [D(\Omega)]^n$:

$$\|\{\phi^0, \phi^1\}\|_* = (\sum_{i=1}^n \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt)^{\frac{1}{2}}.$$

In order to extend Λ to the whole space $H \times V'$, we consider the Hilbert space $\mathcal{F}$ obtained by taking the closure of $[D(\Omega)]^n \times [D(\Omega)]^n$ with respect to $\|\cdot\|_*$. Then observing that if ϕ is the ultra-weak solution of (2.3), there exists a constant $k_2 > 0$ such that

$$\sum_{i=1}^{n} \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt \leq k_2(\|\phi^0\|_H^2 + |\phi^1|_{V'}^2).$$

Symmetrically, we claim that if $T > T_0$, then there exists a constant $k_1 > 0$ such that

$$k_1(|\phi^0|_H^2 + |\phi^1|_{V'}^2) \leq \sum_{i=1}^{n} \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt. \tag{2.8}$$

The proof of this "inverse inequality" is rather technical and it is postponed to Section 3. It follows that

$$\mathcal{F} = H \times V',$$

and therefore Λ has a unique continuous extension to the whole $\mathcal{F}$ which satisfies, for all $\{\phi^0, \phi^1\}, \{\zeta^0, \zeta^1\} \in \mathcal{F}$,

$$\langle \Lambda\{\phi^0, \phi^1\}, \{\zeta^0, \zeta^1\}\rangle_{\mathcal{F},\mathcal{F}'} = (\{\phi^0, \phi^1\}, \{\zeta^0, \zeta^1\})_*.$$

Moreover, since $\mathcal{F}$ is a Hilbert space when equiped with the inner product $(\cdot,\cdot)_*$, from the Lax-Milgran's Lemma for each pair $\{y^0, y^1\} \in \mathcal{F}'$ there exists only one pair $\{\phi^0, \phi^1\} \in \mathcal{F}$ such that

$$\Lambda\{\phi^0, \phi^1\} = \{y^0, y^1\}.$$

Therefore, assuming $\{\varphi^0, \varphi^1\} \in V \times H$, there exists only one pair $\{\phi^0, \phi^1\} \in H \times V'$ such that $\Lambda\{\phi^0, \phi^1\} = \{\varphi^1, -\varphi^0\}$, and from (2.5) we obtain that

$$\{\varphi^1, -\varphi^0\} = \{\psi'(0), -\psi(0)\}, \tag{2.9}$$

where ψ is the solution of (2.4) and ϕ is the solution of (2.3) with initial data $\{\phi^0, \phi^1\}$.

Finally we summarize our arguments. For $T > T_0$ and $\{\varphi^0, \varphi^1\} \in H \times V$ there exists only one pair $\{\phi^0, \phi^1\} \in H \times V'$ such that if ϕ is the solution of (2.3) and ψ is the solution of (2.4), we obtain from (2.9) that φ and ψ are solutions of the original problem (1.1) associated to the initial data $\{\varphi^0, \varphi^1\}$ when we choose

$$h = -\phi \in L^2(0,T;(L^2(\omega))^n).$$

¿From the unicity of solution of the elasticity system (1.1) we have that

$$\varphi(T) = \varphi'(T) = 0.$$

This ends the proof. □

3. Proof of the Inverse Inequality

This section is dedicated to the proof of the inequality (2.8), which is a consequence of the following theorem.

Theorem 3.1 (Inverse Inequality) *Let ω be a neighbourhood of $\overline{\Gamma}_0$. If $T > T_0$, T_0 given by* (2.2)*, there exists a constant $c = c(T) > 0$ such that:*

$$|\phi^0|_H^2 + |\phi^1|_{V'}^2 \leq c \sum_{i=1}^{n} \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt \tag{3.1}$$

holds for every $\{\phi^0, \phi^1\} \in H \times V'$, where $\phi = (\phi_1, \cdots \phi_n)$ is the corresponding ultra-weak solution of (2.3).

Proof. Assume that if θ is the weak solution of (2.3) associated to the initial data $\{\theta^0, \theta^1\} \in V \times H$, we have for some $c_1 = c_1(T)$ the following estimate:

$$\|\theta^0\|_V^2 + |\theta^1|_H^2 + |div\, \theta^0|_{L^2(\Omega)}^2 \leq c_1 \sum_{i=1}^{n} \int_0^T \int_\omega |\theta_i^{'}(x,t)|^2 dxdt. \tag{3.2}$$

Consider now $\{\phi^0, \phi^1\} \in H \times V^{'}$ and let ϕ be the ultra-weak solution of (2.3). Defining on $V \times V$ a coercive, continuous and bilinear form $b(u,v)$ given by

$$b(u,v) = \mu \sum_{i,j=1}^{n} \int_\Omega \frac{\partial u_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} dx + (\lambda + \mu) \int_\Omega \alpha(x)(div\, u)(div\, v) dx, \tag{3.3}$$

we obtain from the Lax-Milgran's Lemma that there exists only one $\eta \in V$ such that

$$b(\eta, v) = \langle -\phi^1, v \rangle_{V',V}, \quad \forall v \in V.$$

We observe that, provided (3.3) holds, it comes from the above identity that:

$$\mu \Delta \eta + (\lambda + \mu) \nabla(\alpha(x) div \eta) = \phi^1 \quad \text{in} \quad V'. \tag{3.4}$$

Now, let us define

$$\Psi(t) = \int_0^t \phi(s) ds + \eta \tag{3.5}$$

Then, Ψ is the weak solution of

$$\left| \begin{array}{l} \Psi'' - \mu \Delta \Psi - (\lambda + \mu) \nabla(\alpha(x) div \Psi) = 0 \;\text{ in } \; Q \\ \Psi = 0 \quad \text{on} \;\; \Sigma \\ \Psi(0) = \eta, \quad \Psi'(0) = \phi^0 \quad \text{in} \;\; \Omega, \end{array} \right.$$

and consequently Ψ satisfies (3.2), that is,

$$\|\eta\|_V^2 + |\phi^0|_H^2 + |div\,\eta|_{L^2(\Omega)}^2 \leq c_1 \sum_{i=1}^{n} \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt. \tag{3.6}$$

Since $-\Delta : H_o^1(\Omega) \to H^{-1}(\Omega)$ is a isometry, it follows from (3.4) and (3.6) that

$$|\phi^0|_H^2 + |\phi^1|_{V'}^2 \leq c \sum_{i=1}^{n} \int_0^T \int_\omega |\phi_i(x,t)|^2 dxdt,$$

which is the desired inequality. □

Then, from the above explanation, we conclude that to prove Theorem 3.1 we have only to prove that every weak solution of (2.3) satisfies the inequality given in (3.2). This is our next objective.

We recall that the energy of the system (2.3), when weak solutions are concerned, is given by:

$$E(t) = \frac{1}{2}|\phi'(t)|_H^2 + \frac{\mu}{2}\|\phi(t)\|_V^2 + \frac{(\lambda+\mu)}{2}|\alpha^{\frac{1}{2}}(x) div\,\phi(t)|_{L^2(\Omega)}^2. \tag{3.7}$$

So, if we denote by E_0 the initial energy, we have:

$$E_0 = \frac{1}{2}|\phi^1|_H^2 + \frac{\mu}{2}||\phi^0||_V^2 + \frac{(\lambda+\mu)}{2}|\alpha^{\frac{1}{2}}(x) div\,\phi^0|_{L^2(\Omega)}^2. \tag{3.8}$$

Lemma 3.2 *For every $T > T_0$, with T_0 defined in* (2.2)*, and every weak solution θ of (2.3), we have the following inequality:*

$$(T - T_0)E_0 \leq c_3 R(x^0) \int_0^T \int_{\Gamma_0} \{\mu \sum_{i=1}^{n} (\frac{\partial \theta_i}{\partial \nu})^2 + (\lambda+\mu)\alpha(x)(div\theta)^2\} d\Gamma dt \tag{3.9}$$

where E_0 is given in (3.8).

Proof. We are going to use the multiplier method as in J. L. Lions [10]. Consider $m = x - x^0$ and let m_k be its k-component. If θ is the weak solution of (2.3) and we multiply the i component of the equation

$$\theta'' - \mu\Delta\theta - (\lambda+\mu)(\alpha(x) div\theta) = 0$$

by $m_k \frac{\partial \theta_i}{\partial x_k}$, we obtain after using Green and Gauss' Theorems and integration by parts the following expression:

$$\frac{1}{2}\int_\Sigma \{\mu \sum_{i=1}^{n} (\frac{\partial \theta_i}{\partial \nu})^2 + (\lambda+\mu)\alpha(x)(div\theta)^2\} m \cdot \nu d\Sigma =$$

$$\sum_{i,k=1}^{n}(\theta_i^{'}(t),(x_k-x_k^0)\frac{\partial\theta_i(t)}{\partial x_k})|_{t=0}^{t=T}+$$

$$\frac{(n-1)}{2}\int_Q\{\sum_{i=1}^{n}|\theta_i^{'}(t)|^2-\mu\sum_{i=1}^{n}|\nabla\theta_i|^2-(\lambda+\mu)\alpha(x)(div\theta)^2\}dxdt+$$

$$\int_0^T E(t)dt-\frac{(\lambda+\mu)}{2}\int_Q(div\theta)^2\nabla\alpha\cdot mdxdt.$$

¿From the above expression we have that:

$$\frac{R(x^0)}{2}\int_0^T\int_{\Gamma_0}\{\mu\sum_{i=1}^{n}(\frac{\partial\theta_i}{\partial\nu})^2+(\lambda+\mu)\alpha(x)(div\theta)^2\}d\Gamma dt\geq$$

$$\sum_{i=1}^{n}(\theta_i^{'}(t),\sum_{k=1}^{n}(x_k-x_k^0)\frac{\partial\theta_i(t)}{\partial x_k}+\frac{(n-1)}{2}\theta_i(t))|_{t=0}^{t=T}+$$

$$(1-||\nabla\alpha||_{L^\infty(\Omega)}R(x^0)\alpha_0^{-1})\int_0^T E(t)dt\geq$$

$$\frac{-2\mathrm{R}(x^0)}{\sqrt{\mu}}E_0+(1-||\nabla\alpha||_{L^\infty(\Omega)}R(x^0)\alpha_0^{-1})TE_0,$$

which proves the lemma. □

Lemma 3.3 *Let us consider $T>T_0$ and $\varepsilon>0$ such that $T-2\varepsilon>T_0$. Then, for every weak solution θ of (2.3) we have:*

$$||\theta^0||_V^2+|\theta^1|_H^2+|div\theta^0|_{L^2(\Omega)}^2$$

$$\leq c_4(T)\int_\varepsilon^{T-\varepsilon}\int_{\Gamma_0}\{\mu\sum_{i=1}^{n}(\frac{\partial\theta_i}{\partial\nu})^2+(\lambda+\mu)\alpha(x)(div\theta)^2\}d\Sigma \tag{3.10}$$

Proof. Consider $T>T_0$ and $\varepsilon>0$ such that $T-T_0>2\varepsilon$.

Let θ be the weak solution of (2.3) associated to the initial data $\{\theta^0,\theta^1\}\in V\times H$ and let $E(t)$ be the energy relation defined in (3.7).

If we define

$$\zeta(t)=\theta(t+\varepsilon),\quad t\in[0,T-\varepsilon]$$

we have that ζ is the weak solution of

$$\left|\begin{array}{l}\zeta''-\mu\Delta\zeta-(\lambda+\mu)\nabla(\alpha(x)div\zeta)=0\ \text{ in }\ C^0([0,T-2\varepsilon];V')\\ \zeta=0\quad on\ \ \Sigma\\ \zeta(0)=\theta(\varepsilon),\quad \zeta'(0)=\theta'(\varepsilon)\quad\text{in }\ \Omega,\end{array}\right.$$

and consequently from (3.9) it comes that

$$E(\varepsilon) \le c_5 \int_0^{T-2\varepsilon} \int_{\Gamma_0} \{\mu \sum_{i=1}^{n} (\frac{\partial \zeta_i}{\partial \nu})^2 + (\lambda+\mu)\alpha(x)(div\zeta)^2\} d\Gamma dt.$$

If we make the change of variables $s = t + \varepsilon$ in the above integral and taking into account that the energy is conservative we obtain the desired result. □

Consider $T > T_0$ and let $\varepsilon > 0$ be such that $T - 2\varepsilon > 0$. Let us define a vector field

$$q = q(x,t) \in [W^{1,\infty}(Q)]^n \tag{3.11}$$

which satisfies the following properties:
(i) $q(x,t) = \nu(x), \quad \forall (x,t) \in \Gamma_0 \times]\varepsilon, T-\varepsilon[$.
(ii) $q(x,t)\cdot \nu(x) \ge 0, \quad \forall (x,t) \in \Gamma \times]0,T[$.
(iii) $q(x,0) = q(x,T) = 0, \quad \forall x \in \Omega$.
(iv) $q(x,t) = 0, \quad \forall (x,t) \in (\Omega \backslash \omega) \times]0,T[$.

Let us consider at a first time the strong solution θ of (2.3). If we multiply the i-component of the equation given in (2.3) by $\nabla\theta_i \cdot q$, if we observe the properties of the field q and if we use Lemma 3.3; we obtain from a density argument the following result:

Lemma 3.4 *Consider $T > T_0$ and $\varepsilon > 0$ such that $T - 2\varepsilon > 0$. If θ is the weak solution of (2.3), there exists a constant $c_6 = c_6(T) > 0$ which satisfies:*

$$||\theta^0||_V^2 + |\theta^1|_H^2 + |div\theta^0|^2_{L^2(\Omega)}$$

$$\le c_6 \int_\varepsilon^{T-\varepsilon} \int_\omega \{\sum_{i=1}^{n} |\theta_i'(x,t)|^2 + \sum_{i=1}^{n} |\nabla\theta_i(x,t)|^2 + |div\theta|^2\} dxdt. \tag{3.12}$$

Remark 3.5 Consider $\hat{\omega}$ a neighbourhood of $\overline{\Gamma}_0$ such that $\Omega \cap \overline{\hat{\omega}} \subset \omega$. So, if we construct to $\hat{\omega}$ a field $\hat{q}$ similar to the field q constructed to ω in (3.11) we obtain:

$$||\theta^0||_V^2 + |\theta^1|_H^2 + |div\theta^0|^2_{L^2(\Omega)} \le$$

$$c_7 \int_\varepsilon^{T-\varepsilon} \int_{\hat{\omega}} \{\sum_{i=1}^{n} |\theta_i'(x,t)|^2 + \sum_{i=1}^{n} |\nabla\theta_i(x,t)|^2 + |div\theta|^2\} dxdt. \tag{3.13}$$

Let us consider now a function

$$p = p(x,t) \in W^{1,\infty}(Q)$$

which satisfies:
(i) $p(x,t) \geq 0$, $\forall (x,t) \in \Omega \times]0,T[$
(ii) $p(x,t) = 1$, $\forall (x,t) \in \hat{\omega} \times]\varepsilon,\, T-\varepsilon[$
(iii) $p(x,t) = 0$, $\forall (x,t) \in (\Omega \backslash \omega) \times]0,T[$
(iv) $0 < p(x,t) < 1$, $\forall (x,t) \in \{\omega \backslash \hat{\omega} \times]0,T[\} \cup \{\hat{\omega} \times]o,\varepsilon] \cup [T-\varepsilon, T[)\}$
(v) $p(x,0) = p(x,T) = 0$, $\forall x \in \Omega$
(vi) $\frac{|\nabla p|^2}{p} \in L^\infty(\omega \times]0,T[)$.

Lemma 3.6 *Consider $T > T_0$ and $\varepsilon > 0$ such that $T - 2\varepsilon > T_0$. Let ω and $\hat{\omega}$ be the neighbourhoods of $\overline{\Gamma_0}$ mentioned above and let θ be the weak solution of (2.3). Then, we have*

$$\int_\varepsilon^{T-\varepsilon} \int_{\hat{\omega}} \{\sum_{i=1}^n |\nabla \theta_i(x,t)|^2 + |div\theta|^2\} dxdt \leq c_8 \sum_{i=1}^n \int_0^T \int_\omega (|\theta_i(x,t)|^2 + |\theta_i^{'}(x,t)|^2) dxdt.$$

Proof. Suppose that θ is the strong solution of (2.3). If we multiply the equation

$$\theta^{''} - \mu\Delta\theta - (\lambda+\mu)\nabla(\alpha(x) div\theta) = 0$$

by $p\theta$ we obtain, by using Green and Gauss' Theorems and integration by parts, that:

$$\mu \sum_{i=1}^n \int_0^T \int_\omega |\nabla\theta_i|^2 p dxdt = \sum_{i=1}^n \int_0^T \int_\omega p(\theta_i^{'})^2 dxdt + \sum_{i=1}^n \int_0^T \int_\omega p^{'} \theta_i \theta_i^{'} dxdt -$$

$$\mu \sum_{i=1}^n \int_0^T \int_\omega (\nabla\theta_i \cdot \nabla p)\theta_i dxdt - (\lambda+\mu) \int_0^T \int_\omega \alpha(x) div\theta(\nabla p \cdot \theta) dxdt -$$

$$(\lambda+\mu) \int_0^T \int_\omega \alpha(x)(div\theta)^2 p dxdt.$$

¿From the properties of the function p and provided that $0 < \alpha_0 \leq \alpha(x) \leq \alpha_1$ on $\overline{\Omega}$, we obtain the desired result by using a density argument in the above inequality. □

Lemma 3.6 combined with Remark 3.5 allow us to prove that:

$$||\theta^0||_V^2 + |\theta^1|_H^2 + |div\theta^0|_{L^2(\Omega)}^2 \leq c_9 \sum_{i=1}^n \int_0^T \int_\omega (|\theta_i(x,t)|^2 + |\theta_i^{'}(x,t)|^2) dxdt, \tag{3.14}$$

holds for every weak solution of (2.3).

¿From (3.14) we conclude that in order to prove inequality (3.2) for every weak solution of (2.3), we have only to prove that there exists $c_{10} > 0$ which verifies:

$$\sum_{i=1}^n \int_0^T \int_\omega |\theta_i(x,t)|^2 dxdt \leq c_{10} \sum_{i=1}^n \int_0^T \int_\omega |\theta_i^{'}(x,t)|^2 dxdt, \tag{3.15}$$

for every weak solution θ of (2.3). Indeed, if we suppose that (3.15) is not true, we obtain for each $k \in N$ a pair of initial data $\{\theta_k^0, \theta_k^1\} \in V \times H$ such that if θ_k is the weak solution of (2.3) associated to these data, we have that:

$$|\psi_k|_{L^2(0,T;(L^2(\omega))^n)} = 1\,, \quad k \in N \tag{3.16}$$

and

$$|\psi_k^{'}|_{L^2(0,T;(L^2(\omega))^n)} \leq \frac{1}{\sqrt{k}}\,, \quad k \in N \tag{3.17}$$

where

$$\psi_k = \frac{\theta_k}{|\theta_k|_{L^2(0,T;(L^2(\omega))^n)}}.$$

Taking into account that θ_k is the weak solution of (2.3) associated to $\{\theta_k^0,\, \theta_k^1\}$, we obtain, from the linearity of the system, that ψ_k is the weak solution of (2.3) associated to $\{\psi_k^0, \psi_k^1\}$ where

$$\psi_k^0 = \frac{\theta_k^0}{|\theta_k|_{L^2(0,T;(L^2(\omega))^n)}}$$

and

$$\psi_k^1 = \frac{\theta_k^1}{|\theta_k|_{L^2(0,T;(L^2(\omega))^n)}}.$$

Then, we obtain from (3.14), (3.16) and (3.17) that there exists $c_{11} > 0$ such that:

$$||\psi_k^0||_V^2 + |\psi_k^1|_H^2 + |div \psi_k^0|_{L^2(\Omega)}^2 \leq c_{11} \;\;, \quad \forall k \in \mathbf{N}. \tag{3.18}$$

So, we are able to extract subsequences from $\{\psi_k^0\}_{k\in N}$ and $\{\psi_k^1\}_{k\in N}$, which we are going to denote by the same notation, such that:

$$\psi_k^0 \rightharpoonup \psi^0 \quad \text{in } V \quad \text{and} \quad \psi_k^1 \rightharpoonup \psi^1 \quad \text{in } H. \tag{3.19}$$

In addition, considering that the energy is conservative and observing the expression given in (3.18) we have that

$$\{\psi_k\}_{k\in N} \;\text{ is bounded in }\; L^\infty(0,T;V)$$

and

$$\{\psi_k'\}_{k\in N} \;\text{ is bounded in }\; L^\infty(0,T;H). \tag{3.20}$$

Consequently, from Aubin-Lions' Lemma we can conclude that there exists a subsequence which satisfies

$$\psi_k \to \psi \quad in \;\; L^2(0,T;H).$$

We also obtain from (3.20) that there exists a subsequence such that

$$\psi_k' \rightharpoonup \psi' \quad in \;\; L^2(0,T;H)$$

which implies, in combination with (3.17), that

$$\psi'(x,t) = 0 \;\; a.e. \;\; in \;\; \omega\times]0,T[. \tag{3.21}$$

On the other hand, as ψ_k is the weak solution of (2.3) associated to $\{\psi_k^0, \psi_k^1\}$, we obtain that ψ is the weak solution of (2.3) associated to the inital data $\{\psi^0, \psi^1\}$. Consequently we have that $u = \psi'$ is the ultra-weak solution of

$$\left| \begin{array}{l} u'' - \mu\Delta u - (\lambda+\mu)\nabla(\alpha(x) div\, u) = 0 \quad \text{in} \;\; Q \\ u = 0 \quad \text{on} \;\; \Sigma \\ u(0) = u^0, \quad u'(0) = u^1 \quad \text{in} \;\; \Omega, \end{array} \right. \tag{3.22}$$

where $u^0 = \psi^1$ and $u^1 = \psi''(0)$.

Moreover, from (3.21) we obtain that:

$$u = 0 \quad \text{a.e. in} \quad \omega\times]0,T[\tag{3.23}$$

$$u' = 0 \quad \text{a.e. in} \quad \omega\times]0,T[. \tag{3.24}$$

Lemma 3.7 *Let u be the ultra-weak solution of (3.22) which satisfies (3.23) and (3.24). Then, we have that*

$$\{u^0, u^1\} = \{0,0\}.$$

Proof. Let us define the function $U : [-T,T] \to H$ given by:

$$U(t) = \left| \begin{array}{l} -3u(-t) + 4u(-\frac{t}{2}) \;\; \text{if} \;\; t \in [-T,0] \\ \\ u(t) \;\; \text{if} \;\; t \in [0,T], \end{array} \right.$$

and let us consider the mollifier sequence $\{\rho_\nu\}_{\nu\in N}$ such that:

$$\rho_\nu \in C_0^\infty(R), \;\; \rho_\nu \geq 0, \;\; supp(\rho_\nu) \subset]\frac{-1}{\nu}, 0[, \;\; \int_{\frac{-1}{\nu}}^0 \rho_\nu(t)dt = 1, \qquad \forall \nu \in \mathbf{N}.$$

Given $\varepsilon > 0$, we are going to define for large enough ν, the following sequence:

$$u_\nu(t) = (U * \rho_\nu)(t) = \int_{\frac{-1}{\nu}}^0 U(t-s)\rho_\nu(s)ds, \quad \forall t \in [0, T-\varepsilon].$$

As we have that $u \in C^0([0,T];\text{H})$ it comes that

$$u_\nu\,, u_\nu^{'}\,, u_\nu^{''}\,, ... \in C^0([0,T-\varepsilon];H), \quad \forall \nu \in \mathbf{N} \tag{3.25}$$

and u_ν satisfies

$$u_\nu^{''} - \mu\Delta u_\nu - (\lambda+\mu)\nabla(\alpha(x)divu_\nu) = 0 \quad in \;\; C^0([0,T-\varepsilon];(H^{-2}(\Omega))^n).$$

Then for each $t \in [0, T-\varepsilon]$ we obtain that:

$$-\mu\Delta u_\nu(t) - (\lambda+\mu)\nabla(\alpha(x)divu_\nu(t)) = -u_\nu''(t) \in H.$$

So, observing the bilinear form given in (3.3), we obtain by Lax-Milgram's Lemma and by (3.25) that:

$$u_\nu(t) \in V, \;\; \forall t \in [0, T-\varepsilon] \tag{3.26}$$

and

$$u_\nu \in C^0([0,T-\varepsilon];V), \;\; \forall \nu \in \mathbf{N}.$$

¿From (3.25) and (3.26) we obtain that:

$$u_\nu(0) \in V \quad and \quad u_\nu^{'}(0) \in H$$

and consequently u_ν is the weak solution of

$$\left| \begin{array}{l} u_\nu'' - \mu\Delta u_\nu - (\lambda+\mu)\nabla(\alpha(x)divu_\nu) = 0 \;\; \text{ in } \; C^0([0,T-\varepsilon];H) \\ u_\nu = 0 \;\; \text{ on } \; \Gamma\times]0,T-\varepsilon[\\ u_\nu(0) = u_\nu^0, \;\; u_\nu'(0) = u_\nu^1 \;\; \text{ in } \; \Omega. \end{array} \right.$$

So, u_ν satisfies the inequality given in (3.14) for $]0,T-\varepsilon[\times\omega$, that is

$$||u_\nu^0||_V^2 + |u_\nu^1|_H^2 + |divu_\nu^0|_{L^2(\Omega)}^2 \leq c_{12}\int_0^{T-\varepsilon}\int_\omega (\sum_{i=1}^n |u_{\nu_i}^{'}|^2 + \sum_{i=1}^n |u_{\nu_i}|^2)dxdt;$$

which in combination with (3.23) and (3.24) implies that:

$$u_\nu^0 = 0 \quad and \quad u_\nu^1 = 0 \;\; \forall\, \nu \in \mathbf{N}. \tag{3.27}$$

On the other hand, it comes by the definition of u_ν that

$$u_\nu^0 \to u^0 \;\; in \;\; H \quad and \quad u_\nu^1 \to u^1 \;\; in \;\; V'$$

and by (3.27) we have that:

$$u^0 = 0 \quad \text{and} \quad u^1 = 0$$

which proves Lemma 3.7. □

¿From the above lemma and from the fact that u is the ultra-weak solution of (3.22), we conclude that

$$u = 0 \quad in \quad C^0([0,T]; H).$$

In other words, we have that

$$\psi'(x,t) = 0 \quad \forall t \in [0,T] \quad and \quad a.e. \;\; in \quad \Omega$$

or

$$\psi(x,t) = \psi(x) \quad \forall t \in [0,T] \quad and \quad a.e. \;\; in \quad \Omega.$$

Then, we are able to conclude that ψ^0 =0 and $\psi^1 = 0$, which combined with the fact that ψ is the solution of (2.3), implies that $\psi = 0$. This is a contradition with (3.16). So, we have proved inequality (3.15) and consequently inequality (3.2), which ends the proof of Theorem 3.1.

Acknowledgement: The authors are grateful to Professor Manuel Milla Miranda and Professor Luiz Adauto Medeiros for their suggestions and stimulating discussions. They are also grateful to Marcelo Moreira Cavalcanti and Juan Amadeo Soriano for their helpful comments. They should always thank God for showing the way to go on.

References

[1] Brézis H., *Analyse Fonctionelle: Théorie et Applications*, Masson, Paris, 1983.

[2] Cavalcanti M. M., Exact controllability of the wave equation with Neumann boundary condition, preprint (1996).

[3] Cioranescu D., Donato P. & Zuazua E., Exact boundary controllability for the wave equation in domains with small holes, *J. Math. Pures Appl.* **71**, 343-377 (1992).

[4] Coron J. M., Contrôlabilité exacte frontière de l'équation d'Euler des fluides parfaits incompressibles bidimensionnels, *C. R. Acad. Sci. Paris* **317**, 271-276 (1993).

[5] Fuentes Apolaya R., HUM para um modelo elástico, *Proc. 40th Brazilian Seminar of Analysis*, (1994).

[6] Grisvard P., Contrôlabilité exacte des solutions de l'équation des ondes en présence de singularités, *J. Math. Pures Appl.* **68**, 215-259 (1989).

[7] Kim J. U., Control of a second-order integro differential equation, *SIAM J. Control Optim.* **31**, 101-110 (1993).

[8] Lagnese J. E. & Lions J. L., *Modelling Analysis and Control of Thin Plates*, Masson, Paris, 1989.

[9] Lasiecka I. & Triggiani R., Exact controllability of semilinear abstract systems with application to waves and plates boundary control problems, *Appl. Math. Optim.* **23**, 109-154 (1991).

[10] Lions J. L., *Contrôlabilité Exacte, Perturbations et Stabilization de Systèmes Distribués*, Tome 1, Masson, Paris, 1988.

[11] Medeiros L. A., Exact Controllabity for a Timoshenko model of vibrations of Beams, *Advances in Math. Sci. Appl.* **1**, 47-61 (1993).

[12] Milla Miranda M. & Medeiros L. A., Exact controllability for Schroendinger equations in non cylindrical domains, *C .R. Acad. Sci. Paris* **319**, 685-689 (1994).

[13] Prates Filho J. S. & Soriano J. A., Exact internal controllability of a Timoshenko system, preprint (1996).

[14] Soriano J. A., Controlabilidad exacta de la equación del telégrafo generalizada, *Rev. Mat. Univ. Complutense de Madrid* **8**, 459-493 (1995).

[15] Zuazua E., *Controlabilidad Exacta y Estabilización de la Ecuación de Ondas*, Lectures Notes, IM-UFRJ, Rio de Janeiro, 1990.

[16] Zuazua E., Contrôlabilité approchée du système de l'élasticitée avec contraintes sur les contrôles, *C. R. Acad. Sci. Paris* **319**, 61-66 (1994).

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISBN 1083-0464
Issue 1 July 1998, pp. 44 – 55
Submitted: January 2, 1998. Published: July 31, 1998

Newton-type Iterations as Approximations of Chebyshev's Method

J. A. Ezquerro and M. A. Hernández

Abstract

This paper examines two convergence theorems, under the same assumptions, for a modification of the Chebyshev method which are based upon majorant principle and recurrence relations. It is shown that from the last we obtain better results.

Keywords: Nonlinear equations in Banach spaces, second-order iterations, convergence theorems, Kantorovich type assumptions, recurrence relations, a priori error bounds.
Classification A.M.S. 1991: 47H17, 65J15.
Supported in part by a grant of the University of La Rioja: 96PYA17JEF.

University of La Rioja. Department of Mathematics and Computation.
Luis de Ulloa s/n. 26004 Logroño. Spain.
E-mail: jezquer@siur.unirioja.es , mahernan@siur.unirioja.es

1 Introduction

Let X and Y be Banach spaces and $F : \Omega \subseteq X \to Y$ be once Fréchet differentiable in an open convex domain Ω. Then the Chebyshev method ([1], [3], [9]), one of the most famous third-order iterations for solving the equation

$$F(x) = 0, \tag{1}$$

which starts from the initial approximation $x_0 \in \Omega$, is defined by

$$x_{n+1} = x_n - \left[I + \frac{1}{2}\Gamma_n F''(x_n)\Gamma_n F(x_n)\right]\Gamma_n F(x_n), \quad n \geq 0,$$

where $\Gamma_n = F'(x_n)^{-1}$, provided that $F'(x_n)$ is invertible at each step.

In [4], the goal of the paper is to reduce operational cost and at the same time to drop the conditions on the second Fréchet-derivative by replacing it with a fixed bilinear operator. So it was introduced an iteration, which was called a modification of Chebyshev's method, for solving (1) and defined by

$$\begin{aligned} x_{n+1} &= x_n - \Gamma_n F(x_n) - \tfrac{1}{2}\Gamma_n A(\Gamma_n F(x_n))^2 \\ &= x_n - \left[I + \tfrac{1}{2}\Gamma_n A \Gamma_n F(x_n)\right]\Gamma_n F(x_n), \quad n \geq 0, \end{aligned} \tag{2}$$

where $A : X \times X \to Y$ is a general bounded bilinear operator such that $\|A\| = \alpha$.

The "penalty" for this approach is the order of convergence drops from three to two. On the other hand, notice that (2) can be written in the form

$$\begin{aligned} F'(x_n)(y_n - x_n) &= -F(x_n), \\ F'(x_n)(x_{n+1} - y_n) &= -A(y_n - x_n)^2/2, \quad n \geq 0. \end{aligned}$$

We have therefore to solve twice the same system appearing in Newton's method with the only modification of the independent terms. As a result, under a similar computational cost, (2) remains faster than Newton's method. Observe that (2) is reduced to the Newton method if $A \equiv 0$.

To develop a convergence theorem for (2), using Kantorovich type assumptions ([5], [6]), two different analyses were done in [4]. The convergence assumptions were relaxed in the second analysis and the following was assumed:

(c$_1$) there exists a continous linear operator $\Gamma_0 = F'(x_0)^{-1}$, $x_0 \in \Omega$,

(c$_2$) $\|\Gamma_0(F'(x) - F'(y))\| \leq k\|x - y\|$, $x, y \in \Omega$, $\quad k \geq 0$,

(c$_3$) $\|A\| = \alpha$, $\quad \|\Gamma_0 A\| \leq \alpha/b$, $\quad \|\Gamma_0 F(x_0)\| = a/b$,

(c$_4$) $b - 2ak \geq 0$.

It was also employed the equation

$$p(t) \equiv \frac{kN}{2}t^2 - bt + a = 0 \tag{3}$$

has two positive roots r_1 and r_2 such that $r_1 \leq r_2$, where N is a parameter satisfying $N \in \left[b + \frac{2\alpha}{k}, \frac{b^2}{2ak}\right]$.

After it was defined the scalar sequence $\{t_n\}$ by

$$t_0 = 0, \quad t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}\left(1 + \frac{\alpha}{2}\frac{p(t_n)}{p'(t_n)^2}\right), \quad n \geq 0, \tag{4}$$

and proved that $\{t_n\}$ is well-defined for $n \geq 0$, monotonously increases and converges quadratically to r_1 for all $0 \leq \alpha \leq \frac{b}{4a}(b - 2ak)$.

We now state the convergence theorem for (2) which is based upon majorant principle. We denote $\overline{B(x,r)} = \{y \in X; \|y - x\| \leq r\}$ and $B(x,r) = \{y \in X; \|y - x\| < r\}$.

Theorem 1.1 *Let us assume that conditions* $(\mathbf{c}_1)$–$(\mathbf{c}_4)$ *hold and* $0 \leq \alpha \leq \dfrac{b}{4a}(b - 2ak) = \alpha_1$. *Then the sequence* $\{x_n\}$ *defined by (2) converges to a solution* x^* *of (1) in* $\overline{B(x_0, r_1)} \cap \Omega$ *for* $N \in \left[b + \frac{2\alpha}{k}, \frac{b^2}{2ak}\right]$. *The limit* x^* *is the unique solution of (1) in* $B(x_0, r) \cap \Omega$ *where* $r = r_2 + \dfrac{2(N-b)}{kN}$. *Moreover* $\|x^* - x_n\| \leq r_1 - t_n$, $n \geq 0$.

Following that and taking into account "test" function (3), we observe, by analysing the asymptotic error constant $C_\alpha = \left|\dfrac{p''(r_1) - \alpha}{2p'(r_1)}\right|$, sequence (4) converges faster to r_1 for increasing values of the parameter $\alpha \in [0, kN]$.

Notice that the speed of convergence of sequence (2) is majorized by scalar sequence (4). Then this fact leads to extend the values of the parameter α so that the number of the general bounded bilinear operator A considered in (2) is increased. To reach that, we establish another result on the convergence of method (2) under the same general assumpptions $(\mathbf{c}_1)$–$(\mathbf{c}_4)$. A new technique consisting of a system of recurrence relations is used and we can also obtain better error estimates in practice as we can see in example 2.

In section 2, we derive several basic results on a scalar sequence and present the recurrence relations. Next, in section 3, we apply the results to establish a convergence theorem for process (2) and see the upper bound for α appearing in theorem 1.1 is improved, since the interval of valid α is extended. Finally, in section 4, we yield two examples where theses improvements are shown.

2 Preliminaires

Let us assume $(\mathbf{c}_1)$–$(\mathbf{c}_4)$ and define a sequence of positive real numbers depending on two parameters by the following recurrence relation. Denote $\delta = ak/b$ and $a_0 = a\alpha/b^2$. Set

$$a_{n+1} = a_n f(a_n)^2 g(a_n), \quad n \geq 0, \tag{5}$$

where

$$f(x) = \frac{2}{2(1-\delta) - \delta x} \tag{6}$$

and

$$g(x) = \frac{1}{2}\left[x + \delta\left(1 + \frac{x}{2}\right)^2\right]. \tag{7}$$

We first need a lemma whose proof is trivial.

Lemma 2.1 *Let f and g the two real functions defined in (6) and (7) respectively. Then*

(i) *f and g are increasing for $\delta \geq 0$,*

(ii) *$f(\gamma x) \leq f(x)$ and $g(\gamma x) \leq \gamma g(x)$ for $\gamma \in (0,1)$.*

After that some properties for the sequence $\{a_n\}$ given by (5) are provided.

Theorem 2.2 *Let f and g the two real functions given by (6) and (7) respectively. Let $0 \leq \alpha < \dfrac{b^2}{a} h\left(\dfrac{ak}{b}\right) = \alpha_2$ where*

$$h(x) = \frac{2(2x^2 - 3x - 1 + \sqrt{1 + 8x - 4x^2})}{x(1-2x)}. \tag{8}$$

Then

(i) *$f(a_0)^2 g(a_0) < 1$,*

$(\mathbf{ii}_n)$ *$a_{n+1} \leq a_n$ for all $n \geq 0$,*

$(\mathbf{iii}_n)$ *$\delta(2 + a_n) < 2$ for all $n \geq 0$.*

Proof. ¿From the hypotheses it is easy to check **(i)** and $(\mathbf{iii}_0)$. Consequently, $a_1 \leq a_0$. We now suppose that $(\mathbf{ii}_j)$–$(\mathbf{iii}_j)$ are true for $j = 1, 2, \ldots, n-1$. Taking into account lemma 2.1 (i), we deduce $(\mathbf{ii}_n)$ and $(\mathbf{iii}_n)$. ∎

Theorem 2.3 *Assume the hypotheses of theorem 2.2 are satisfied. Let us define* $\gamma = a_1/a_0$. *Then*

(i) $a_n \leq \gamma^{2^n-1} a_0$ *for all* $n \geq 2$,

(ii) $f(a_n)g(a_n) \leq \dfrac{\gamma^{2^n}}{f(a_0)}$ *for all* $n \geq 1$.

Proof. Notice that $\gamma = a_1/a_0 = f(a_0)^2 g(a_0) < 1$ by theorem 2.2 (i), and $a_2 = a_1 f(a_1)^2 g(a_1) \leq \gamma^2 a_0 f(a_0)^2 g(a_0) = \gamma^2 a_1$ by lemma 2.1 (ii). If $a_j \leq \gamma^{2^{j-1}} a_{j-1}$ is supposed for $j = 2, 3, \ldots, n-1$, then

$$a_n = a_{n-1} f(a_{n-1})^2 g(a_{n-1}) \leq \gamma^{2^{n-1}} a_{n-2} f(a_{n-2})^2 g(a_{n-2}) = \gamma^{2^{n-1}} a_{n-1}$$

and, as a result, $a_n \leq \gamma^{2^n-1} a_0$.

Finally, for all $n \geq 1$, we have

$$f(a_n)g(a_n) \leq f(\gamma^{2^n-1} a_0) g(\gamma^{2^n-1} a_0) \leq \gamma^{2^n-1} f(a_0) g(a_0) = \frac{\gamma^{2^n}}{f(a_0)}$$

and **(ii)** also holds. ■

Firstly, we observe that taking into account initial hypotheses **(c**$_1$**)**–**(c**$_4$**)** and assuming that $y_0 \in \Omega$ we have

$$\|\Gamma_0 A \Gamma_0 F(x_0)\| \leq \|\Gamma_0 A\| \|\Gamma_0 F(x_0)\| \leq a_0.$$

Then the iterate x_1 is well defined and

$$\|x_1 - x_0\| \leq \left\| \left[I + \frac{1}{2} \Gamma_0 A \Gamma_0 F(x_0) \right] \right\| \|\Gamma_0 F(x_0)\| \leq \left(1 + \frac{a_0}{2}\right) \|\Gamma_0 F(x_0)\|.$$

Next we prove the following items are true for all $n \geq 1$ by mathematical induction:

[I$_n$**]** $\|\Gamma_n\| = \|F'(x_n)^{-1}\| \leq f(a_{n-1}) \|\Gamma_{n-1}\|$,

[II$_n$**]** $\|\Gamma_n F(x_n)\| \leq f(a_{n-1}) g(a_{n-1}) \|\Gamma_{n-1} F(x_{n-1})\|$,

[III$_n$**]** $\|\Gamma_n A\| \|\Gamma_n F(x_n)\| \leq a_n$,

[IV$_n$**]** $\|x_{n+1} - x_n\| \leq \left(1 + \dfrac{a_n}{2}\right) \|\Gamma_n F(x_n)\|$.

We now assume that $[\mathbf{I}_n]$–$[\mathbf{IV}_n]$ are true for some n and $x_n \in \Omega$. ¿From lemma 2.1 (i) and theorem 2.2 if $x_{n+1} \in \Omega$, it is follows that the operator $\Gamma_n F'(x_{n+1})$ is invertible, since

$$\|I - \Gamma_n F'(x_{n+1})\| \leq \prod_{j=0}^{n-1} \|\Gamma_{j+1} F'(x_j)\| \|\Gamma_0 (F'(x_n) - F'(x_{n+1}))\|$$

$$\leq k\|x_{n+1} - x_n\| \prod_{j=0}^{n-1} f(a_j) \leq k\left(1 + \frac{a_n}{2}\right) \|\Gamma_0 F(x_0)\| \prod_{j=0}^{n-1} f(a_j)^2 g(a_j) \leq \delta\left(1 + \frac{a_n}{2}\right) < 1,$$

and, by the Banach lemma, Γ_{n+1} is defined and

$$\|\Gamma_{n+1}\| = \|\Gamma_{n+1} F'(x_n) \Gamma_n\| \leq \frac{\|\Gamma_n\|}{1 - \|I - \Gamma_n F'(x_{n+1})\|} \leq \frac{\|\Gamma_n\|}{1 - \delta\left(1 + \frac{a_n}{2}\right)} = f(a_n)\|\Gamma_n\|.$$

Therefore $[\mathbf{I}_{n+1}]$ holds.

Using from (2) the expression for $x_{n+1} - x_n$, we have

$$F(x_n) + F'(x_n)(x_{n+1} - x_n) + \int_0^1 [F'(x_n + t(x_{n+1} - x_n)) - F'(x_n)](x_{n+1} - x_n)\, dt$$

$$= -\frac{A}{2}[\Gamma_n F(x_n)]^2 + \int_0^1 \prod_{j=0}^{n-1} [\Gamma_{j+1} F'(x_j)] \Gamma_0 [F'(x_n + t(x_{n+1} - x_n)) - F'(x_n)](x_{n+1} - x_n)\, dt,$$

so that

$$\|\Gamma_n F(x_{n+1})\| \leq \frac{1}{2}\|\Gamma_n A\| \|\Gamma_n F(x_n)\|^2 + \frac{k}{2}\|x_{n+1} - x_n\|^2 \prod_{j=0}^{n-1} \|\Gamma_{j+1} F'(x_j)\|$$

$$\leq \frac{1}{2}\left[\|\Gamma_n A\| \|\Gamma_n F(x_n)\| + k\left(1 + \frac{a_n}{2}\right)^2 \|\Gamma_n F(x_n)\| \prod_{j=0}^{n-1} f(a_j)\right] \|\Gamma_n F(x_n)\|$$

$$\leq \frac{1}{2}\left[a_n + \delta\left(1 + \frac{a_n}{2}\right)^2 \prod_{j=0}^{n-1} f(a_j)^2 g(a_j)\right] \|\Gamma_n F(x_n)\|$$

$$\leq \frac{1}{2}\left[a_n + \delta\left(1 + \frac{a_n}{2}\right)^2\right] \|\Gamma_n F(x_n)\| = g(a_n)\|\Gamma_n F(x_n)\|.$$

follows as a consequence of theorem 2.2. Hence $[\mathbf{II}_{n+1}]$ is true, since

$$\|\Gamma_{n+1} F(x_{n+1})\| \leq \|\Gamma_{n+1} F'(x_n)\| \|\Gamma_n F(x_{n+1})\| \leq f(a_n) g(a_n) \|\Gamma_n F(x_n)\|.$$

Now it is easy to see $[\mathbf{III}_{n+1}]$ and $[\mathbf{IV}_{n+1}]$:

$$\|\Gamma_{n+1} A\| \|\Gamma_{n+1} F(x_{n+1})\| \leq a_n f(a_n)^2 g(a_n) = a_{n+1}$$

and

$$\begin{aligned}\|x_{n+2}-x_{n+1}\| &\le \left(1+\frac{1}{2}\|\Gamma_{n+1}A\|\|\Gamma_{n+1}F(x_{n+1})\|\right)\|\Gamma_{n+1}F(x_{n+1})\| \\ &\le \left(1+\frac{a_{n+1}}{2}\right)\|\Gamma_{n+1}F(x_{n+1})\|.\end{aligned}$$

Then the induction is complete.

3 Main convergence theorem

After last section we are ready to prove the following convergence theorem.

Theorem 3.1 *In addition to assumptions* $(\mathbf{c}_1)$–$(\mathbf{c}_4)$ *and definition (5), assume that* $B(x_0, 1/k) \subset \Omega$ *and* $0 \le \alpha < \dfrac{b^2}{a} h\left(\dfrac{ak}{b}\right) = \alpha_2$ *where* h *is given by (8). Then iterates (2) are well-defined for all* $n \ge 0$, *lie in* $B(x_0, 1/k)$ *for* $n \ge 0$ *and converge R-quadratically at least to a unique solution* $x^* \in B(x_0, 1/k)$ *of equation (1). Furthermore the error bounds*

$$\|x^*-x_n\| \le \left(1+\frac{a_0}{2}\gamma^{2^n-1}\right)\gamma^{2^n-1}\frac{\Delta^n}{1-\Delta}\frac{a}{b}, \quad n \ge 0, \tag{9}$$

hold, where $\gamma = a_1/a_0 < 1$ *and* $\Delta = 1/f(a_0) < 1$.

Proof. Firstly, we note

$$\begin{aligned}\left(1+\frac{a_n}{2}\right)\|\Gamma_n F(x_n)\| &\le \left(1+\frac{a_n}{2}\right) f(a_{n-1})g(a_{n-1})\|\Gamma_{n-1}F(x_{n-1})\| \\ &\le \cdots \le \left(1+\frac{a_n}{2}\right)\|\Gamma_0 F(x_0)\| \prod_{j=0}^{n-1} f(a_j)g(a_j)\end{aligned}$$

as a consequence of estimate $[\mathbf{II}_n]$. Next, from theorem 2.3, we have

$$\prod_{j=0}^{n-1} f(a_j)g(a_j) \le \prod_{j=0}^{n-1}\left(\gamma^{2^j}\Delta\right) = \gamma^{2^n-1}\Delta^n$$

where $\gamma = a_1/a_0 < 1$ and $\Delta = 1/f(a_0) < 1$. Thus

$$\begin{aligned}\|x_{n+m}-x_n\| &\le \|x_{n+m}-x_{n+m-1}\| + \|x_{n+m-1}-x_{n+m-2}\| + \cdots + \|x_{n+1}-x_n\| \\ &\le \left(1+\frac{a_n}{2}\right)\sum_{i=n-1}^{n+m-2}\left[\prod_{j=0}^{i} f(a_j)g(a_j)\right]\|\Gamma_0 F(x_0)\|\end{aligned}$$

$$\leq \left(1+\frac{a_0}{2}\gamma^{2^n-1}\right)\gamma^{2^n-1}\frac{1-\Delta^m}{1-\Delta}\Delta^n\|\Gamma_0 F(x_0)\|. \tag{10}$$

As a result, we have for $n = 0$

$$\|x_m - x_0\| < \left(1+\frac{a_0}{2}\right)\frac{1-\Delta^m}{1-\Delta}\|\Gamma_0 F(x_0)\| = 1/k$$

and $x_m \in B(x_0, 1/k)$ for all $m \geq 0$. Moreover, by letting $m \to \infty$ in (10), we get

$$\|x^* - x_n\| \leq \left(1+\frac{a_0}{2}\gamma^{2^n-1}\right)\gamma^{2^n-1}\frac{\Delta^n}{1-\Delta}\frac{a}{b}.$$

This proves that $\{x_n\}$ converges to a limit $x^* \in B(x_0, 1/k)$ and bounds (9) hold.

To see that x^* is a solution of (1), we have $\|\Gamma_n F(x_n)\| \to 0$ as $n \to \infty$. Taking into account that $\|F(x_n)\| \leq \|F'(x_n)\|\|\Gamma_n F(x_n)\|$ and the sequence $\{\|F'(x_n)\|\}$ is bounded, we infer that $\|F(x_n)\| \to 0$ as $n \to \infty$. Consequently, we obtain $F(x^*) = 0$ by the continuity of F.

To prove the uniqueness, assume a some other solution y^* of (1) in $B(x_0, 1/k)$. According to Argyros and Chen ([1], [2]), from the approximation,

$$0 = F(y^*) - F(x^*) = \int_0^1 F'(x^* + t(y^* - x^*))\,dt\,(y^* - x^*),$$

we have to prove that the operator $\int_0^1 F'(x^* + t(y^* - x^*))\,dt$ is invertible and then $y^* = x^*$. Indeed, from

$$\int_0^1 \|\Gamma_0[F'(x^* + t(y^* - x^*)) - F'(x_0)]\|\,dt \leq k\int_0^1 \|x^* + t(y^* - x^*) - x_0\|\,dt$$

$$\leq k\int_0^1 \left[(1-t)\|x^* - x_0\| + t\|y^* - x_0\|\right]\,dt < 1,$$

it follows that $\left[\int_0^1 F'(x^* + t(y^* - x^*))\,dt\right]^{-1}$ exists.

It remains to be shown that sequence (2) has R-order of convergence two at least. We deduce from (9) that

$$\|x^* - x_n\| \leq \left(1+\frac{a_0}{2}\right)\frac{a/b}{\gamma(1-\Delta)}\gamma^{2^n}, \quad n \geq 0.$$

Hence the convergence of (2) is R-quadratical at least (see [8]). ■

Remark 1. Observe that the interval of valid α has been improved under the same general assumptions $(\mathbf{c}_1)$–$(\mathbf{c}_4)$. Note that $\alpha_1 = \frac{b}{4a}(b - 2ak)$ and $\alpha_2 = \frac{b^2}{a}h(\frac{ak}{b})$ where h is defined in (8) are the upper bounds for α appearing in theorem 1.1 and theorem 3.1 respectively. It is easy to prove that $\alpha_2 = \max\{\alpha_1, \alpha_2\}$, and the values of α are therefore extended.

4 Examples

In this section we provide two examples where we compare the results obtained in this paper with those from theorem 1.1.

Example 1. We first apply iterative method (2) to the cubic function, introduced by Ostrowski [7], $F : [-4, 4] \to \mathbb{R}$ where $F(x) = x^3 - 2x - 5$. We choose the initial value $x_0 = 2.4$. Then we can easily find the parameters appearing in theorem 3.1.:

$$a = 4.024, \quad b = 15.28 \quad \text{and} \quad k = 2.63916.$$

In consequence $\alpha < \alpha_2 = 8.36148$. Notice the obtained upper bound α_2 for α is much bigger than the upper bound $\alpha_1 = 2.50537$ appearing in theorem 1.1.

As the hypotheses of theorem 3.1 are satisfied, sequence (2), where $A = \alpha$, converges to $x^* = 2.09455148154$ with 12 significant digits for any $\alpha \in [0, \alpha_2)$. Moreover this solution is unique in $B(x_0, 0.636667)$.

On the other hand, observe that the higher the value of $\alpha \in [0, 8.36148)$ is, the smaller the asymptotic error constant $C_{e,\alpha} = \dfrac{12.5673 - \alpha}{22.3229}$ is. As a result, sequence (2) converges faster to x^* for increasing values of parameter $\alpha \in [0, 8.36148)$ (see table I).

n	x_n ($\alpha = 5$)	x_n ($\alpha = 8$)
0	2.40000000000	2.40000000000
1	2.11395496200	2.10033841040
2	2.09459866332	2.09454655415
3	2.09455148180	2.09455148154

Table I

Example 2. A more interesting application is given by the following example. We take the differential equation quoted in [4]

$$\begin{cases} y'' + y^2 = 0, \\ y(0) = 0 = y(1). \end{cases} \tag{11}$$

Divide the interval $[0, 1]$ into n subintervals and we set $h = 1/n$. We denote the points of subdivision by $0 = w_0 < w_1 < \ldots < w_n = 1$ with the corresponding values of the function $y_0 = y(w_0), y_1 = y(w_1), \ldots, y_n = y(w_n)$. A standard approximation for the second derivative at these points is

$$y_i'' = \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2}, \quad i = 1, 2, \ldots, n-1.$$

Noting that $y_0 = 0 = y_n$, define the operator $F : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ by

$$F(y) = Gy + h^2 g(y)$$

where

$$G = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix}, \quad g(y) = \begin{pmatrix} y_1^2 \\ y_2^2 \\ \vdots \\ y_{n-1}^2 \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix}.$$

Then

$$F'(y) = G - 2h^2 \begin{pmatrix} y_1 & 0 & \dots & 0 \\ 0 & y_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & y_{n-1} \end{pmatrix}.$$

Let $y \in \mathbb{R}^{n-1}$, then the norm will be $\|y\| = \max_{1 \le i \le n-1} |y_i|$. The corresponding norm on $G \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ is

$$\|G\| = \max_{1 \le i \le n-1} \sum_{k=1}^{n-1} |g_{ik}| = 4.$$

We note that for all $x, y \in \mathbb{R}^{n-1}$:

$$\|F'(x) - F'(y)\| = \|\mathrm{diag}\{2h^2(x_p - y_i)\}\| = 2h^2 \max_{1 \le i \le n-1} |x_i - y_i| \le 2h^2 \|x - y\|.$$

To show the convergence of (2) to a solution y^*, set $n = 10$ and as the solution would vanish at the endpoints and be positive in the interior, a reasonable choice of initial approximation seems to be $\frac{1}{4} \sin \pi x$. This gives us the following vector:

$$x_0 = \begin{pmatrix} 0.0772542 \\ 0.1469460 \\ 0.2022540 \\ 0.2377640 \\ 0.2500000 \\ 0.2377640 \\ 0.2022540 \\ 0.1469460 \\ 0.0772542 \end{pmatrix}.$$

For this initial approximation we obtain the following parameters:

$$a = 0.0195386, \quad b = 0.0763837 \quad \text{and} \quad k = 0.261836.$$

Then, under the same hypotheses, we obtain $\alpha < \alpha_2 = 0.412466$ (theorem 3.1) versus $\alpha \leq \alpha_1 = 0.0646531$ (theorem 1.1). In consequence the valid interval of α has been quite increased.

By theorem 3.1, iteration (2) converges to a unique solution $y^* = (y_1^*, y_2^*, \ldots, y_9^*)$ in $B(x_0, 3.81919)$.

On the other hand, the error estimates obtained by theorem 1.1 and theorem 3.1, under the same value of $\alpha = 0.06$ (see [4]), are compared in table II. It is shown that better error bounds are obtained from the last one.

n	error estimates from th. 1.1	error estimates from th. 3.1
1	0.1535885817753795	0.0466942344593571
2	0.0406992460289781	0.0032037407239804
3	0.0046814846994008	0.0002411379471532
4	0.0000777752985228	0.0000275164672344

Table II

Remark 2. We indicate the choice of the matrix A is an open study for future papers. In this paper, we only choose those with a little operational cost taking into account the influence of $\|A\|$ in the speed of convergence of iteration (2).

References

1. Argyros, I.K. and Chen, D.: Results on the Chebyshev Method in Banach Spaces, Proyecciones 12, (1993), 119–128.
2. Argyros, I.K. and Chen, D.: Parameter-Based Algorithms for Approximating Local Solution of Nonlinear Complex Equations, Proyecciones 33, (1994), 53–61.
3. Candela, V. and Marquina, A.: Recurrence Relations for Rational Cubic Methods II: The Chebyshev Method, Computing 45, (1990), 355–367.
4. Ezquerro, J.A.: A Modification of the Chebyshev Method, IMA J. Numer. Anal. 17, (1997), 511–525.
5. Kantorovich, L.V. and Akilov, G.P.: Functional Analysis, Pergamon Press, Oxford, 1984.
6. Ortega, J.M. and Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
7. Ostrowski, A.M.: Solution of Equations in Euclidean and Banach Spaces, Academic Press, New York, 1973.

8. Potra, F.A. and Pták, V.: Nondiscrete Induction and Iterative Processes, Pitman, New York, 1984.

9. Safiev, R.A.: On some iterative processes, Ž. Vyčcisl. Mat. Fiz. 4, (1964), 139-143 (Translated into English by L.B. Rall as MRC Technical Summary Report, No. 649, Univ. Wisconsin-Madison, 1966).

Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISBN 1083-0464
Issue 1 July 1998, pp. 56 – 63
Submitted: November 10, 1998. Published: July 31, 1998

Helly property associated with the discrete planes.

A. Bretto and B. Laget

Abstract

Discrete planes are basic concepts in discrete geometry and digital topology. In this note we introduce a discrete version of the Helly theorem, and we show that only 8-connected grids have the Helly property.

AMS subject classification: A. M. S. (MOS) Subject Classification Codes 05, 94.
Keywords: Digital topology; discrete geometry; hypergraph; Helly property; grids graph.

Université Jean Monnet Equipe d'ingénierie de la vision (UMR 5512)
Groupe de Combinatoire et Traitement d'Images.
Site G.I.A.T Industries
3, rue Javelin Pagnon BP 505
42007 Saint-Etienne cedex 1 FRANCE
E-mail: bretto@univ-st-etienne.fr

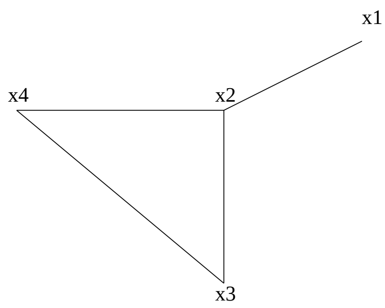


Figure 1: Example of a graph

1 Introduction

The growth of computer sciences has highlighted discrete mathematics. Discrete geometry and digital topology which are derived from classical mathematical fields, are the focus of much research works. These fields are crucial to image analysis. One major concept of discrete geometry and digital topology is the notion of grid graph or discrete plane. In this note we associate the helly property with these planes.

2 Preliminaries

2.1 Graph

Recall that a *graph* is a couple $G = (V; E)$, where V is a set of elements called *vertices*, and E is a set of unordered pairs of members of V called *edges.*
Two vertices are *adjacent* if they are linked by an edge.
Given a graph G, we denote by $\Gamma(x)$ the *neighborhood* of a vertex x, i. e. the set formed by all vertices adjacent to x is defined by:

$$\Gamma(x) = \{y \in V, \{x, y\} \in E\}$$

Figure 1 is an example of a graph.

The number of neighborhood vertices x is the *degree of* x , (denoted by dx). If dx is constant, the graph is called *regular.* A graph is *loopless* if it does not contain an edge of type (x, x).
A *chain* in a graph is a sequence of distinct vertices $(v_1, v_2, \dots v_p)$ linked by an edge. A *cycle* is a chain such as the first vertex and the last vertex are the same. We note a cycle with the length equal to n, C_n.
A graph is *connected* if any pair of vertices is linked by a chain.
A cycle is *centered* if there exists a vertex adjacent to all the vertices from the cycle.
A *sun* S_3 is a graph with six vertices $a_1, a_2, a_3, b_1, b_2, b_3$ such that :

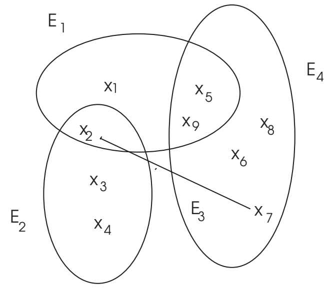


Figure 2: Representation of the hypergraph. In this example we have S the set of vertices with $\mathbf{S} = \{x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ x_9\}$. The set of hyperedges is E with $E =\{\ E_1 = \{x_1, x_2, x_5, x_9\},\ E_2 = \{x_2, x_3, x_4\},\ E_3 = \{x_2, x_7\},\ E_4 = \{x_5, x_6, x_7, x_8, x_9\}\}$, the star centered on x_2 is $H(x_2) = \{E_1,\ E_2,\ E_3\}$, with degree $dx_2 = 3$ and the star centered on x_4 is $H(x_4) = \{E_2\}$, with degree $dx_4 = 1$.

- a_1, a_2, a_3, a_1 is a cycle.
- b_i has only two neighbors a_i et a_{i+1} (i mod 3)

Figure 4 shows a cycle C_4 and a sun S_3.

2.2 Hypergraph

A *hypergraph* H on a set $\mathbf{S}$ is a family $(E_i)_{i\in I}$ of non-empty subsets of $\mathbf{S}$ called hyperedges with:

$$\bigcup_{i\in I} E_i = \mathbf{S}$$

$$\forall i \in I \quad I = \{1, 2, \ldots, n\} \quad n \in \mathbb{N}$$

Let us note :

$$H = (\mathbf{S}; (E_i)_{i\in I})$$

One calls *n-family* any sub-family from $(E_i)_{i\in I}$ with n edges.
For $x \in \mathbf{S}$, a *star* of H (with center x) is the set of hyperedges which contains x, and is called $H(x)$. The *degree of* x is the cardinality of the star $H(x)$ denoted by :

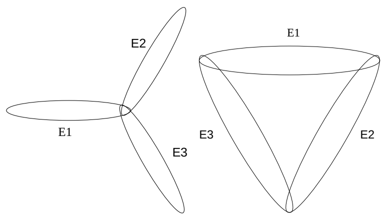
$$dx = card(H(x)).$$

We can represent a hypergraph by the figure 2.

A hypergraph has the *Helly property* if each family of hyperedges intersecting two by two (intersecting family) has a non-empty intersection (belonging to a star).

Figure 3 shows two examples of intersecting hyperedges :

- Fig. 3(a) : A family of three hyperedges intersecting two by two with a non-empty intersection
- Fig. 3(b) : A family of three hyperedges intersecting two by two with an empty intersection



(a) star (b) triangle

Figure 3: Suppose that the hypergraphs have only three hyperedges, then the first hypergraph has the Helly property.

Remark. *A loopless graph with no isolated vertex is therefore a particular hypergraph with every hyperedge having cardinality 2.*

A hypergraph is *regular* if any edge has the same cardinality.

2.2.1 Neighborhood hypergraph

Given a graph G, the hypergraph having the vertices of G as vertices and the neighborhood of these vertices as hyperedges (including these vertices), is called the *neighborhood hypergraph* of G. To each graph we can associate a neighborhood hypergraph:

$$H_G = (\mathbf{S}, (E_x = \{x\} \cup \Gamma(x)))$$

The neighborhood hypergraph associated with the example in figure 1 is:

$$\mathbf{S} = \{x_1, x_2, x_3, x_4, \}$$

$$E_{x_1} = \{x_1, x_2\},\ E_{x_2} = \{x_1, x_2, x_3, x_4\},$$

$$E_{x_3} = \{x_2, x_3, x_4\},\ E_{x_4} = \{x_2, x_3, x_4\}$$

A *discrete plane* (resp a *grid graph*) is a loopless regular connected graph such that :
The vertices are the points of a regular lattice $\mathbf{X}$ from $\mathbb{R}^2$ which can be either square or triangular or hexagonal.
The set of edges is generated thanks to classical distances :

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|$$

$$d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$$

and

$$d_2(x, y) = ((x_1 - y_1)^2 + (x_2 - y_2)^2)^{\frac{1}{2}}$$

On a square lattice the grid associated with d_1 (resp d_∞) is the 4-connected grid (resp 8-connected grid), one notes it Γ^4 (resp Γ^8). On a triangular lattice the grid associated with d_2 is the 6-connected grid, one note it Γ^6. On an hexagonal lattice one defines the 3-connected grid thanks to d_2, which is noted by Γ^3.

Generally the grid is defined by :

$$\Gamma(x) = \{y \in \mathbf{X}, \mathbf{d}(\mathbf{x}, \mathbf{y}) = \mathbf{1}\}$$

On a grid, one defines the neighborhood graph (or neighborhood relation) in this way:

$$\Gamma_p(x) = \{y \in \mathbf{X}, \mathbf{1} \leq \mathbf{d}(\mathbf{x}, \mathbf{y}) \leq \mathbf{p}(\mathbf{x})\}$$

where p is a mapping from $\mathbf{X}$ in $\mathbf{N}^*$ which gives the size of the neighborhood of x. This graph leads to the following neighborhood hypergraph:

$$H_{\Gamma_p} = (\mathbf{X}, (\{x\} \cup \Gamma_p(x)))$$

If $p(x)$ is a constant n, the hypergraph is uniform, it is noted by:

$$H_n = (X, (E_x = \{x\} \cup \Gamma_n(x)))$$

3 Grid graph and Helly property

Theorem 1. *Let H_{Γ_p} be a neighborhood hypergraph associated with a discrete plane.*
H_{Γ_p} has the Helly property if and only if the discrete plane is an 8-connected grid.

Proof. The condition is sufficient. Indeed we will prove the following lemma:

Lemma 1. *The neighborhood hypergraphs associated with the grid Γ^8 have the Helly property.*

Proof. Let $(E_i)_{i \in I}$ be an intersecting n-family, and (x_i, y_i) the coordinates of the vertex i which is the generator of E_i. The relation

$$E_i \cap E_j \neq \emptyset$$

can be expressed in the following way:

$$|x_i - x_j| \leq p_i + p_j \text{ and } |y_i - y_j| \leq p_i + p_j$$

with $p(x_i, y_i) = p_i$ et $p(x_j, y_j) = p_j$. The first inequality implies that :

$$x_i - p_i \leq x_j + p_j \quad \forall\, i, j \in I$$

consequently

$$x^- = Max(x_i - p_i) \leq Min(x_j + p_j) = x^+$$

In the same way, we have:

$$y^- = Max(y_i - p_i) \leq Min(y_j + p_j) = y^+$$

The result comes from the equality:

$$\bigcap_{i \in I} E_i = \{(x, y), x^- \leq x \leq x^+ \text{ et } y^- \leq x \leq y^+\}$$

□

To prove that the condition is necessary , it is sufficient to show that the uniform hypergraphs associated with the grid different from Γ^8 do not have the Helly property. Indeed it is sufficient to find a configuration of edges which does not verify the Helly property and to vary the size of the edges from the configuration.

Let H be a neighborhood hypergraph associated with a graph. If this graph contains either a no centered cycle with a length equal to four, five, six, or a non-centered sun S_3, obviously the neighborhood hypergraph does not have the Helly property. Figure 4 gives two examples from C_4 and S_3.

Lemma 2. *The hypergraphs associated with the grid Γ^4 do not have the Helly property.*

Proof. Now we will show that there exists an intersecting 4-family which is not a star. To do this, one builts four rectangular triangles on the grid by the following algorithm

- Let (x, a, u) be a rectangular triangle such that $d_1(x, a) = n-1$, $d_1(a, u) = n$, from the Montanari theorem [6] $d_1(x, u) = 2n - 1$

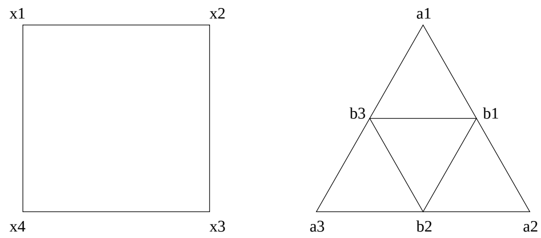


Figure 4: In the first figure we have $H_{C_4} = \{E_{x1} = \{x1, x2, x4\}; E_{x2} = \{x1, x2, x3\}; E_{x3} = \{x3, x2, x4\}; E_{x4} = \{x4, x1, x3\}$. These edges intersecte two by two but $\bigcap_{i=1}^{i=4} E_{xi} = \emptyset$. Second figure shows a S_3, we have $\bigcap_{i=1}^{i=3} E_{ai} = \emptyset$ but these edges intersecte two by two.

- Let (u, b, v) be a rectangular triangle such that $d_1(u, b) = n-1$, $d_1(b, v) = n$, one has $d_1(u, v) = 2n - 1$
- Let (v, c, w) be a rectangular triangle such that $d_1(v, c) = n-1$, $d_1(c, w) = n$, one has $d_1(v, w) = 2n - 1$
- Let (w, d, x) be a rectangular triangle such that $d_1(w, d) = n-1$, $d_1(d, x) = n$, one has $d_1(w, x) = 2n - 1$

Consequently, one obtains a cycle (x, u, v, w, x) with a length equal to four which is not centered, so the family has an empty intersection. □

Lemma 3. *The neighborhood relation Γ_n associated with the grid Γ^6 contains a sun S_3 which is not centered. So, H_n the neighborhood hypergraph associated with this grid does not verify the Helly property.*

Proof. Let $a_1,\ a_2,\ a_3$ be three vertices with: $d_2(a_i, a_{i+1}) = 2n$ $(i = 0, 1, 2\ mod\ 3)$, such that these vertices form an equilateral triangle (this triangle exists because it is the union of equilateral unity triangles). $2n$ being even, there exists a vertex b_i between a_i and b_i such that: $d_2(a_i, b_i) = d_2(b_i, a_{i+1}) = n$, so a_i is adjacent to b_i which is adjacent to a_{i+1}. Moreover, since $d_2(a_i, b_i) = d_2(b_{i-1}, a_i) = n$, and since the b_i are the middle of the sides of the equilateral triangle , one has : $d_2(b_i b_{i+1}) = n$ $(i = 0,\ 1,\ 2\ mod\ 3)$. From this algorithm, one obtains a no centered sun S_3, thanks to : $a_1,\ a_2,\ a_3,\ b_1,\ b_2,\ b_3$. □

Lemma 4. *Let H_n be a hypergraph associated with the grid Γ^3:*

- *$n = 1$, Γ_n contains a non-centered cycle with a length equal to six. We note C_6.*
- *$n \geq 2$, Γ_n a non-centered sun S_3.*

The hypergraphs H_n do not have the Helly property.

Proof. If $n = 1$: it is obvious.

Suppose $n \geq 2$: Let x be a vertex, $a_1, \ a_2, \ a_3$, the vertices such that $d_2(x, a_1) = 1$, $d_2(x, a_2) = 1$, $d_2(x, a_3) = 1$. The three vertices $a_1, \ a_2, \ a_3$ form a cycle C_3 centered in x. Let a_i, a_{i+1} $(mod 3)$, there exists a vertex b_i such that $d_2(b_i, a_i) = n$ and $d_2(b_i, a_{i+1}) = n$.

One obtains a cycle $(a_1 \ a_2 \ a_3 \ b_1 \ b_2 \ b_3)$ with : $(a_1 \ a_2 \ a_3)$ which is a cycle C_3, b_i is only adjacent to a_i and a_{i+1}. So $(a_1 \ a_2 \ a_3 \ b_1 \ b_2 \ b_3)$ form a non-centered sun S_3 □

4 Conclusion

The main applications of discrete geometry and digital topology are in the image analysis and the pattern recognition. The introduction of the Helly property in image processing allows both to generalize the region growing segmentation [3], [1] and to give a new class of image filters [2].

References

[1] A. Bretto, J. Azema, H. Cherifi and B. Laget *Combinatoric and image processing.* CVGIP:Graphical Model and Image Processing. Vol. 59, No. 5,(1997), September.

[2] A. Bretto and H. Cherifi *A noise cancellation algorithm based on hypergraph modelling* IEEE Workshop on Digital Signal Processing Loen, Norway Sept 1-4, (1996) 5-8.

[3] Bretto A. Ubéda S. *Hypergraph model of digital topology for grey level images* D. G. C. I. in Lecture Notes in Computer Sciences 1176 Springer-Verlag, (1996) 211-223.

[4] Didier M. *Personal Communications* C. E. L. Université de Montpellier II.

[5] Kovalevsky V. A. "Finite Topology as Applied to Image Processing." *Computer Vision Graphics, and Image Processing 46, (1989) 141-161.*

[6] Montanari U. *A note on minimal length polygonal approximation to a digitalized contour* Comm. ACM, 13(1), (1970), 41-47

[7] A. Rosenfeld *Digital Image Processing.* Academic press (1982).

[8] Voss K. *Discrete Images, Object and Functions in* $\mathbb{Z}^n$ Algorithms and Combinatorics Springer Verlag, (1990).