VOLUME XXVI (1993)

THE PUNJAB UNIVERSITY

JOURNAL OF MATHEMATICS

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF THE PUNJAB
LAHORE-54590
PAKISTAN

Notice to Contributors

1. The Journal is meant for publication of research papers and review articles covering state of the art in a particular area of mathematical science.

2. Manuscripts should be typewritten and in a form suitable for publication. As far as possible, the use of complicated notations should be avoided. Figures, drawn on separate sheets of white paper in Black Ink, should be of size suitable for inclusion in the Journal.

3. References should be given at the end of the paper and be referred to by numbers in serial order on square brackets, e.g. [3]. References should be typed as follows;

Reference to Paper:

Hoo, C.S.: BCI-algebras with conditions, *Math. Japonica* 32, No. 5 (1987) 749-756.

Reference to Book:

Mitchell, B.: Theory of categories, New York: Academic Press, 1965.

4. Contributions and other correspondence should be addressed to Dr. A. Majeed, Mathematics Department, Punjab University, Quaid-e-Azam Campus, Lahore-54590, Pakistan.

5. The decision to accept or reject a paper for publication in the Journal rests fully with the Editorial Board.

6. Authors, whose papers will be published in the Journal, will be supplied 30 free reprints of their papers and a copy of the issue containing their contributions.

7. The Journal which is published annually will be supplied free of cost in exchange with other Journals of Mathematics.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 1 – 26

EXISTENCE AND UNIQUENESS OF SOLUTIONS FOR PROBLEMS WITH A PARAMETER

Tadeusz Jankowski

Technical University of Gdansk,
Department of Numerical Methods,
Faculty of Applied Mathematics,
80-952 GDANSK, POLAND.

ABSTRACT

The purpose of this paper is to give, by the comparative method, sufficient conditions for the existence, uniqueness and convergence of successive approximations for some functional equation with a parameter. As a consequence corresponding conditions for differential equations with deviated arguments are obtained from the general considerations.

1. Introduction

We consider the equation with a parameter.

$$y(t) = f(t, y(\cdot), \lambda), \ t \in J = [a, b] \tag{1}$$

subject to the following condition

$$L(y(\cdot), \lambda) = \Theta, \tag{2}$$

where Θ is zero element in R^q. By Y we denote the space of bounded functions $y : J \to R^p$. The functions $f : J \times Y \times R^q \to R^p$, $L: Y \times R^q \to R^q$ are given. By a solution of (1–2) we mean a function $y \in Y$ and a parameter $\lambda \in R^q$, which satisfy the equation (1) and the condition (2). We say that the problem (1–2) is solved if such y and λ are found.

We can give some conditions which guarantee that there exists a function $\Lambda: Y \to R^q$ such that $\lambda = \Lambda(y(\cdot))$ is a solution of (2). For this case we have

$$y(t) = f(t, y(\cdot), \Lambda(y(\cdot))), \ t \in J \tag{3}$$

Note that although the problem (1–2) may be converted to (3), the function Λ appearing in (3) is not known explicitly and therefore (1–2) can not be solved numerically by sòlving (3).

We are concerned with the question of the existence, uniqueness and continuous dependence of solutions on the right-hand side of (1). Existence theorems can be prov?d by using the Banach or Schauder theorems on the fixed point and also by the method of successive approximations. Unfortunately, the Banach fixed point theorem yields some strong conditions imposed on the right-hand side of (1–2). These conditions can be weakened if we use the comparative method (for the abstract from of this method, see [12]). The comparative method will be applied in this paper to obtain results for existence and uniqueness of solutions of problem (1–2). For this reason two sequences $\{y_n\}$, $\{\lambda_n\}$ are introduced in the following way.

$$\left.\begin{array}{l} y_0 : J \to R^p \text{ is arbitrary,} \\ y_{n+1}(t) = f(t, y_n(\cdot), \lambda_n),\ t \in J,\ n = 0, 1, \ldots.. \end{array}\right\} \tag{4}$$

and

$$\left.\begin{array}{l} \lambda_0 \text{ is arbitrary in } R^q, \\ \lambda_{n+1} = \lambda_n - B^{-1}L(y_n(\cdot), \lambda_n),\ n = 0, 1, \ldots\ldots. \end{array}\right\} \tag{5}$$

where B is some nonsingular square matrix of order q. For given y_0 and λ_0 we determine y_1 from (4) and then λ_1 from (5), and so on. Sufficient conditions under which the sequences $\{y_n\}$. $\{\lambda_n\}$ converge to the solution $(\bar{y}, \bar{\lambda})$ of (1–2) are established. The estimates between the approximate solution (y_n, λ_n) and the exact solution $(\bar{y}, \bar{\lambda})$ are given too.

The functional problem (3) is a special case of (1–2). Volterra integral equations, Fredholm integral equations, integro-functional equations are also special cases of (1). Some equations can be reduced to (1): for example differential and functional-differential equations, differential equations with deviated arguments (see section 6). The Niccoletti condition is a special case of (2).

The paper is organized as follows. In sections 2–4 we investigate the general problem (1–2) giving sufficient conditions for the existence and uniqueness of solutions of (1–2). The question of the continuous dependence of solutions on the right-hand side of (1) with (2) is considered too. Existence and uniqueness theorems are formulated by using the nonlinear comparison operators Ω and Γ. The considerations for the functional problem (section 5) and for the differential equation with deviated arguments (section 6) are obtained, by applying the general results of sections 2–4 for our cases.

This paper contains a generalization of some results of [2,3,6,10].

2. Existence of Solutions of (1–2)

To formulate such theorem we first introduce the general assumptions for our problem (1–2). Let E denote the collection of all non-negative functions $u(\cdot)$ which are defined and bounded on J.

Assumption H $_1$: Suppose that

1° f:J×Y×$R^q \to R^p$, L:Y×$R^q \to R^q$ and if $x(t) = f(t, y(\cdot), \lambda)$ for $t \in Y$, $y \in Y$, $\lambda \in R^q$, then $x = Y$,

2° there exists a function Ω:J×E×$R_+ \to R_+ = [0, \infty)$ which is nonnegative, nondecreasing and montonicaly continuous with respect to the last two variables (*i.e.* if $u_n \searrow u$, $w_n \searrow w$, then $\Omega(t,u_n,w_n) \searrow \Omega(t,u,w)$ for t, $\Omega(t,0,0)=0$ and Ω is such that the condition

$$||f(t,x_1(\cdot),\mu_1)-f(t,x_2(\cdot),\mu_2)|| \le \Omega(t,||x_1(\cdot)-x_2(\cdot)||, ||\mu_1-\mu_2||$$

holds for $t \in J$, $x_1 . x_2 \in Y$, $\mu_1 . \mu_2 \in R^q$,

3° there exist a nonsingular square matrix B of order q and constants $h \ge 0$, $d > 0$ such that $hd < 1$. $d \ge ||B^{-1}||$ and

$$||L(x(\cdot), \mu) - L(x(\cdot), \bar{\mu}) - B(\mu-\bar{\mu})|| \le h||\mu-\bar{\mu}||. \qquad (6)$$

for $x = Y$, μ, $\bar{\mu} = R^q$, where the matrix norm is consistent with the vector norm,

4° there exists a nonnegative, nondecreasing and monotonicaly continuous function Γ:E–R_+. $\Gamma(0) = 0$ and such that for $x_1, x_2 \in Y$, $\mu \in R^q$ the inequality

$$||L(x_1(\cdot), \mu) - L(x_2(\cdot), \mu)|| \le \Gamma(||x_1(\cdot) - x_2(\cdot)||) \text{ holds.}$$

Assumption H_2: Suppose that

1° there exists a solution $w^* \in E$ of the inequality

$$\Omega(t.w(\cdot), d(1-dh)^{-1} [\Gamma(w(\cdot)) + ||L(y_0(\cdot), \lambda_0)||] + v(t) \leq w(t) \quad (7)$$

where $dh \in [0, 1)$ and

$$v(t) = ||f(t, y_0(\cdot), \lambda_0) - y_0(t)||, \; t \in J,$$

2° in the class of functions $w \in E$ satisfying the condition $w(t) \leq w^*(t)$, $t \in J$, the function $w(t)=0$, $t \in J$ and $u=0$ are only the solution of the system

$$\left.\begin{aligned} & w(t)=\Omega\;(t, w(\cdot), d(1-dh)^{-1}\,\Gamma(w(\cdot))), t\in J \\ & u=d(1-dh)^{-1}\,(w(\cdot)) \end{aligned}\right\} \quad (8)$$

Remark 1

It is known that the matrix norm is consistent with the vector norm if

$$||Cx|| \leq ||C||\;||x||,$$

where C is an $q \times q$ matrix and $x \in R^q$ (see for example [11]).

To prove the convergence of the sequences $\{y_n\}$, $\{\lambda_n\}$ to the solution $(\bar{y}, \bar{\lambda}) \in Y \times R^q$ of (1–2) we define two sequences by the following relations:

$$\left.\begin{aligned} & u_0=u^*=d(1-dh)^{-1}\,[\Gamma(w^*(\cdot)) + ||L(y_0(\cdot), \lambda_0)|| \\ & u_{n+1}-d[hu_n + \Gamma(w_n(\cdot))], n = 0, 1, \ldots\ldots \end{aligned}\right\} \quad (9)$$

and

$$\left.\begin{aligned} & w_0(t) = w^*(t), t \in J, \\ & w_{n+1}(t)=\Omega(t, w_n(\cdot), u_n), t\in J, n=0, 1, \ldots\ldots \end{aligned}\right\} \quad (10)$$

where $dh \in [0, 1)$

First of all we formulate the following.

Lemma 1

If Assumptions H_1 and H_2 are satisfied, then

(i) $u_{n+1} \leq u_n \leq u^*$, $n = 0, 1, \ldots..,$

$w_{n+1}(t) \leq w_n(t) \leq w^*(t)$, $t \in J$, $n = 0, 1, \ldots.,$

(ii) $u_n \to 0,$

$$w_n(t) \rightarrow 0, \ t \in J.$$

(the last sign $\rightarrow$ denotes the uniform convergence on J if w_n's are continuous),

(iii) $$u_n = (dh)^n u^* + d \sum_{i=0}^{n-1} (dh)^{n-1-i} \Gamma(w_i(\cdot)), n = 1, 2, ...,$$

(iv) $$\lim_{n\to\infty} \sum_{i=0}^{n-1} (dh)^{n-1-i} \Gamma(w_i(\cdot)) = 0.$$

Proof

The relations (i) we can prove by induction. In view of (9–10) and (ii), the sequences $\{u_n\}$, $\{w_n\}$ are convergent, so

$$\lim_{n\to\infty} u_n = \bar{n}, \ \lim_{n\to\infty} w_n(t) \rightarrow \bar{w}(t), \ t \in J,$$

where $\bar{u}$ and $\bar{w}$ satisfy the following equations

$$\bar{u} = d\,[h\bar{u} + \Gamma(\bar{w}(\cdot))],$$

$$\bar{w}(t) = \Omega(t, \bar{w}(\cdot), \bar{u}), \quad t \in J.$$

Hence and by Assumption H_2, the function $\bar{w}$ equals zero and $\bar{u}=0$ too. By induction, we can get (iii) and this leads to the last condition (iv).

Lemma 2

If Assumption H_1 and the condition 1° of Assumption H_2 are satisfied, then the estimates

$$||\lambda_n - \lambda_0|| \leq u^*, \quad n = 0, 1, ..., \tag{11}$$

$$||y_n(t) - y_0(t)|| \leq w^*(t), t \in J, n = 0, 1, ..., \tag{12}$$

hold true.

Proof

Indeed, the inequalities (11) and (12) are true for $n=0$. We assume that they are satisfied for $n \geq 0$. Using the assumptions we see that

$$\|\lambda_{n+1}-\lambda_0\| = \|B^{-1}[B(\lambda_n-\lambda_0)-L(y_n(\cdot), \lambda_n)+L(y_n(\cdot),\lambda_0)-L(y_n(\cdot), \lambda_0)$$
$$+ L(y_0(\cdot), \lambda_0) - L(y_0(\cdot), \lambda_0)]\|$$
$$\leq d\{h\|\lambda_n-\lambda_0\| + \Gamma(\|y_n(\cdot) - y_0(\cdot)\|) + \|L(y_0(\cdot), \lambda_0)\|\}$$
$$\leq d\{hu^* + \Gamma(w^*(\cdot)) + \|L(y_0(\cdot), \lambda_0)\|\} = u^*, \text{ and}$$

$$\|y_{n+1}(t) - y_0(t)\| \leq \|f(t, y_n(\cdot), \lambda_n) - f(t, y_0(\cdot), \lambda_0)\|$$
$$+ \|f(t, y_0(\cdot), \lambda_0) - y_0(t)\|$$
$$\leq \Omega(t, \|y_n(\cdot)-y_0(\cdot)\|, \|\lambda_n-\lambda_0\|) + \|f(t, y_0(\cdot),\lambda_0) - y_0(t)\|$$
$$\leq \Omega(t, w^*(\cdot), u^*) + \|f(t, y_0(\cdot), \lambda_0) - y_0(t)\|$$
$$\leq w^*(t), \ t \in J.$$

It is clear that (11) and (12) follow by induction.

We also need the following:

Lemma 3

If Assumption H_1 and the condition 1° of Assumption H_2 are satisfied, then the inequalities

$$\|\lambda_{n+j} - \lambda_n\| \leq u_n, \ n, j = 0, 1, ..., \tag{13}$$

$$\|y_{n+j}(t) - y_n(t)\| \leq w_n(t), \ t \in J, n, j = 0, 1, ... \tag{14}$$

hold.

Proof

The inequalities (13) and (14) are true for $n=0$ (see Lemma 2). Assume that they are true for $n, j \geq 0$. Then we obtain

$$\|\lambda_{n+j+1}-\lambda_{n+1}\| = \|B^{-1}[B(\lambda_{n+j}-\lambda_n) - L(y_n(\cdot), \lambda_{n+j})+L(y_n(\cdot), \lambda_n)$$
$$+ L(y_n(\cdot), \lambda_{n+j}) - L(y_{n+j}(\cdot), \lambda_{n+j})\|$$
$$\leq d\{h\|\lambda_{n+j}-\lambda_n\| + \Gamma(\|y_n(\cdot)-y_{n+j}(\cdot)\|)\}$$
$$\leq d\{hu_n + \Gamma(w_n(\cdot))\} = u_{n+1}, \text{ and}$$

$$\|y_{n+j+1}(t)-y_{n+1}(t)\| = \|f(t, y_{n+j}(\cdot), \lambda_{n+j})-f(t, y_n(\cdot), \lambda_n)\|$$
$$\leq \Omega(t, \|y_{n+j}(\cdot) - y_n(\cdot)\|, \|\lambda_{n+j}-\lambda_n\|)$$
$$\leq \Omega(t, w_n(\cdot), u_n) = w_{n+1}(t), \ t \in J.$$

Hence we have (13) and (14) by induction.

The main result of this section is the following existence theorem.

Theorem 1

If Assumptions H_1 and H_2 are satisfied, then there exists a solution $(\bar{\lambda}, \bar{y}) \in R^q \times Y$ of the problem (1–2). This solution $(\bar{\lambda}, \bar{y})$ is the limit of the sequences $\{\lambda_n\}$, $\{y_n\}$, respectively and the estimates

$$||\bar{\lambda} - \lambda_n|| \leq u_n, \quad n = 0, 1, ..., \tag{15}$$

$$||\bar{y}(t) - y_n(t)|| \leq w_n(t), \quad t \in J, \ n = 0, 1, ..., \tag{16}$$

hold true.

Moreover, this solution is unique in the class satisfying the relations;

$$||\bar{\lambda} - \lambda_0|| \leq u^*, \quad ||\bar{y}(t) - y_0(t)|| \leq w^*(t), \ t \in J, \tag{17}$$

Proof

Because of Lemmas 1 and 3 we see that the sequences $\{\lambda_n\}$, $\{y_n\}$ are convergent to $\bar{\lambda}$ and $\bar{y}$, respectively, and that the estimates (15) and (16) are satisfied.

In view of the inequalities:

$$\begin{aligned} ||f(t,\bar{y}(\cdot),\bar{\lambda}) - \bar{y}(t)|| &\leq ||f(t, \bar{y}(\cdot), \bar{\lambda}) - f(t, y_n(\cdot), \lambda_n)|| + ||y_{n+1}(t) - \bar{y}(t)|| \\ &\leq \Omega(t, ||\bar{y}(\cdot) - y_n(\cdot)||, ||\bar{\lambda} - \lambda_n||) + ||y_{n+1}(t) - \bar{y}(t)|| \\ &\leq \Omega(t, w_n(\cdot), u_n) + w_{n+1}(t) = 2w_{n+1}(t). \end{aligned}$$

and

$$\begin{aligned} ||L(\bar{y}(\cdot),\bar{\lambda})|| &= ||L(\bar{y}(\cdot),\bar{\lambda}) - L(\bar{y}(\cdot),\lambda_n) - B(\bar{\lambda} - \lambda_n) + L(\bar{y}(\cdot), \lambda_n) \\ &\quad - L(y_n(\cdot), \lambda_n) + B(\bar{\lambda} - \lambda_{n+1})|| \\ &\leq hu_n + \Gamma(w_n(\cdot)) + ||B|| \, u_{n+1}, \end{aligned}$$

and Lemma 1, we see that $(\bar{\lambda}, \bar{y})$ is a solution of (1–2). To prove that $(\bar{\lambda}, \bar{y})$ is unique we assume that the problem (1–2) has another solution $(\bar{\bar{\lambda}}, \bar{\bar{y}})$ in the class of functions satisfying relations (17), and that $(\bar{\lambda}, \bar{y}) \not\equiv (\bar{\bar{\lambda}}, \bar{\bar{y}})$. We can prove, by induction, that

$$||\bar{\bar{\lambda}} - \lambda_n|| \leq u_n, \ n = 0, 1, ...,$$

$$||\bar{\bar{y}}(t) - y_n(t)|| \le w_n(t),\ t \in J,\ n = 0, 1, \ldots.$$

It means that $\bar{\bar{\lambda}} = \bar{\lambda}$ and $\bar{\bar{y}}(t) = \bar{y}(t)$, $t \in J$. This contradiction proves the uniqueness of $(\bar{\lambda}, \bar{y})$ in the class pointed above.

Remark 2

Instead of (5) we may take the following process

$$\lambda_{n+1} = \lambda_n - B^{-1}(y_n(\cdot), \lambda_n)\, L(y_n(\cdot), \lambda_n),\ \ n = 0, 1, \ldots,$$

provided that the matrix $B_{q\times q}$ is nonsingular for each $y_n \in Y$ and $\lambda_n \in R^q$.

Remark 3

Assume that:

(i) Y is a set of continuous functions on J and the mapping $t \to f(t, x(\cdot), \lambda)$ is continuous for $t \in J$, $x \in Y$, $\lambda \in R^q$,

(ii) $L(x(\cdot), \lambda) = \widetilde{M} \int_a^b x(\tau)\, d\tau - \widetilde{P}$, where $\widetilde{M}_{q\times p}$ and $\widetilde{P}_{q\times 1}$ are given matrices,

(iii) $\widetilde{B}_{p\times q}$ is a continuous matrix function on J such that

$$\widetilde{M} \int_a^b \widetilde{B}(t)\, dt \text{ is nonsingular.}$$

The sequences $\{y_n\}$, $\{\lambda_n\}$ may be now defined by

$$\left.\begin{aligned} &y_{n+1}(t) = f(t, y_n(\cdot), \lambda_n) + \widetilde{B}(t)\,(\lambda_{n+1} - \lambda_n),\ t \in J,\ n = 0, 1, \ldots.. \\ &\widetilde{M} \int_a^b y_{n+1}(t)\,dt = \widetilde{P} \end{aligned}\right\} \quad (18)$$

By (18), it follows that

$$\lambda_{n+1} = \lambda_n - \left(\widetilde{M} \int_a^b B(t)dt\right)^{-1} \left[\widetilde{M} \int_a^b f(t, y_n(\cdot), \lambda_n)\, dt - \widetilde{P}\right],\ n = 0, 1, \ldots$$

3. Uniqueness Theorem

Now we will give the conditions under which the problem (1–2) has at most one solution. These conditions do not guarantee the existence of the solution of (1–2). We formulate the following

Theorem 2

If Assumption H_1 is satisfied and for $(c, e) \in E \times R_+$ the system

$$\left.\begin{aligned} & c(t) \leq \Omega(t, c(\cdot), d(1-dh)^{-1}\Gamma(c(\cdot))), \ t \in J \\ & e \leq d(1-dh)^{-1}\Gamma(c(\cdot)) \end{aligned}\right\} \tag{19}$$

has only zero solution *i.e.* $c(t)=0$, $t \in J$ and $e=0$, then the problem (1–2) has at most one solution.

Proof

Assume that the problem (1–2) has two solutions $(\mu_i, x_i) \in R^q \times Y$, $i=1,2$. Put

$$c(t) = \|x_1(t) - x_2(t)\|, \ t \in J, \ e = \|\mu_1 - \mu_2\|.$$

Now, in view of the conditions 3°, 4° of Assumption H_1, we have

$$\begin{aligned} \|\mu_1 - \mu_2\| &= B^{-1}[B(\mu_1-\mu_2) - L(x_1(\cdot), \mu_1) + L(x_1(\cdot), \mu_2) \\ &\quad - L(x_1(\cdot), \mu_2) + L(x_2(\cdot), \mu_2)]\| \\ &\leq d\{h\|\mu_1 - \mu_2\| + \Gamma(c(\cdot))\}, \text{ so} \end{aligned}$$

$$\|\mu_1 - \mu_2\| \leq d(1-dh)^{-1}\Gamma(c(\cdot)). \tag{20}$$

Similarly, using (20) and the condition 2° of Assumption H_1, we can write

$$\begin{aligned} c(t) &= \|f(t, x_1(\cdot), \mu_1) - f(t, x_2(\cdot), \mu_2)\| \\ &\leq \Omega(t, c(\cdot), d(1-dh)^{-1}\Gamma(c(\cdot)), \ t \in J. \end{aligned}$$

By the assumptions of Theorem 2, we get $c(t)=0$, $t \in J$ and $e=0$ *i.e.* $x_1(t)=x_2(t)$, $t \in J$, and $\mu_1=\mu_2$. This contradiction proves this theorem.

Remark 4

If Assumption H_2 is satisfied, then the functions $c(t)=0$, $t \in J$ and $e=0$ are only solution of (19) in the class of functions of E satisfying the condition $c(t) \leq w^*(t)$, $t \in J$.

By induction, we can prove

$$c(t) \leq w_n(t),\ \ t \in J \ \text{ and } \ e \leq u_n,\ \ n = 0, 1, ...,$$

and if $n \to \infty$, then $c(t) = 0, t \in J, e = 0$.

4. Continuous Dependence

Consider the equation

$$z(t) = f_0(t, z(\cdot), \gamma), \quad t \in J, \tag{21}$$

with the condition

$$L_0(z(\cdot), \gamma) = \Theta, \tag{22}$$

where the vector functions f_0 and L_0 have the same properties as f and L, respectively. We have

Theorem 3

If Assumption H_1 is satisfied, and

1° $(\bar{\lambda}, \bar{y}), (\bar{\gamma}, \bar{z}) \in R^q \times Y$ are solutions of problems (1–2) and (21–22) respectively,

2° the sequence $\{v_n\}$ defined by

$$\left.\begin{array}{l} v_0(t) = ||\bar{y}(t)|| + ||\bar{z}(t)||, t \in J \\ v_{n+1}(t) = \Omega(t, v_n(\cdot), d(1-dh)^{-1}[\Gamma(v_n(\cdot)) + v^{**}]) + v^*(t) \\ \qquad t \in J,\ n = 0, 1, ... \end{array}\right\} \tag{23}$$

$$v^*(t) = ||f(t, \bar{z}(\cdot), \bar{\gamma}) - \bar{z}(t)||,\ \ t \in J,$$

$$v^{**} = ||L(\bar{z}(\cdot), \bar{\gamma}) - L_0(\bar{z}(\cdot), \bar{\gamma})||$$

has a limit $\bar{v} \in E$ on J.

Then the realtions

$$||\bar{y}(t) - \bar{z}(t)|| \leq \bar{v}(t),\ \ t \in J, \tag{24}$$

$$||\bar{\lambda} - \bar{\gamma}|| \leq d(1 - dh)^{-1}\,[\Gamma(\bar{v}(\cdot)) + v^{**}] \tag{25}$$

hold true.

Proof

Put $\quad c = ||\bar{\lambda} - \bar{\gamma}||,\ e(t) = ||\bar{y}(t) - \bar{z}(t)||,\ t \in J.$

Now, in view of Assumption H_1, we can write

$$c = ||B^{-1}[B(\bar\lambda - \bar\gamma) - L(\bar y(\cdot), \bar\lambda) + L(\bar y(\cdot), \bar\gamma) - L(\bar y(\cdot), \bar\gamma) + L(\bar z(\cdot), \bar\gamma) - L(\bar z(\cdot), \bar\gamma) + L_0(\bar z(\cdot), \bar\gamma)]|| \le d\{h||\bar\lambda - \bar\gamma|| + \Gamma(||\bar y(\cdot) - \bar z(\cdot)|| + v^{**}\},$$ and

$$e(t) \le ||f(t, \bar y(\cdot), \bar\lambda) - f(t, \bar z(\cdot), \bar\gamma)|| + ||f(t, \bar z(\cdot), \bar\gamma) - \bar z(t)|| \le \Omega(t, ||\bar y(\cdot) - \bar z(\cdot)||.\ ||\bar\lambda - \bar\gamma||) + v^*(t) = \Omega(t, e(\cdot), d(1-dh)^{-1}[\Gamma(e(\cdot)) + v^{**}]) + v^*(t),\ t \in J$$

Indeed $\qquad e(t) \le v_0(t),\ t \in J,$

and hence, we see

$$e(t) \le \Omega(t, v_0(\cdot), d(1-dh)^{-1}[\Gamma(v_0(\cdot)) + v^{**}]) + v^*(t) = v_1(t),\ t \in J.$$

Furthermore by induction, we can get

$$e(t) \le v_n(t),\quad t \in J,\ n = 0, 1, \ldots\ldots$$

Because of $\lim_{n\to\infty} v_n(t) = \bar v(t)$, $t \in J$, we have the assertion of this theorem.

Remark 5

Assume that there exists a solution $\tilde v_0 \in E$ of the inequality
$$\Omega(t, \tilde v_0(\cdot), d(1-dh)^{-1}[\Gamma(\tilde v_0(\cdot)) + v^{**}] + \max(v_0(t), v^*(t)) \le \tilde v_0(t),\ t \in J.$$

It is easy to prove

$$v_n(t) \le \tilde v_n(t),\ t \in J,\ n = 0, 1, \ldots.,$$
$$\tilde v_{n+1}(t) \le \tilde v_n(t) \le \tilde v_0(t),\ t \in J,\ n = 0, 1, \ldots,$$

where

$$\tilde v_{n+1}(t) = \Omega(t, \tilde v_n(\cdot), d(1-dh)^{-1}[\Gamma(\tilde v_n(\cdot)) + v^{**}]) + v^*(t),\ t \in J,\ n = 0, 1, \ldots.$$

Hence $\lim_{n\to\infty} \tilde v_n(t) = \tilde v(t)$, $t \in J$. Now, Theorem 3 is true with $\tilde v$ instead of $\bar v$.

5. The Functional Equation

In this section we consider the following problem

$$\left.\begin{array}{l} y(t) = f(t, y(\beta_1(t)), y(\beta_2(t)), \ldots y(\beta_s(t)), \lambda)\ t \in J, \\ L(y(b), \lambda) = \Theta, \end{array}\right\} \qquad (26)$$

$$\text{with}\quad \left.\begin{array}{l} f: J\times(R^p)^s \times R^q \to R^p,\ L: R^p\times R^q\to R^q,\\ \beta_i: J\to J,\ \beta_i(a) = 0,\ i = 1, 2, \ldots, s. \end{array}\right\} \tag{27}$$

The problem (26) is a special case of (1–2), and

$$f(t, y(\cdot), \lambda) = f(t, y(\beta_1(t)), y(\beta_2(t)),\ldots,y(\beta_s(t)), \lambda),\ t\in J, \tag{28}$$

$$L(y(\cdot), \lambda) = L(y(b), \lambda). \tag{29}$$

Now we assume that Assumptions H_1 and H_2 are satisfied with

$$\Omega(t, w(\cdot), c) = \Omega(t, w(\beta_1(t)),\ldots,w(\beta_s(t)), c),\ t\in J, \tag{30}$$

$$\Gamma(w(\cdot)) = \Gamma(w(b)). \tag{31}$$

Theorem 1 on the existence of solutions of (1–2) implies.

Theorem 4

Suppose the condition (27) and Assumption H_1 (except 1°) and H_2 are satisfied with f,L,Ω and Γ defined by (28–31). Then the problem (26) has a solution $(\bar{\lambda}, \bar{y}) = R^q\times R^p$ which is a limit of the sequences $\{\lambda_n\}$, $\{y_n\}$ and the estimates (15–16) hold too.

We introduce the equation

$$z(t) = f_0(t, z(\delta_1(t), z(\delta_2(t)),\ldots,z(\delta_s(t), \gamma),\ \ t\in J, \tag{32}$$

with the condition

$$L_0(z(b), \lambda) = \Theta, \tag{33}$$

to formulate a theorem on the continuous dependence of solutions of (26). Indeed, the functions f_0, δ_i, L_0 have the same properties as f, β_i and L, respectively.

Theorem 3 implies

Theorem 5

Suppose the condition (27), Assumption H_1 (except for 1°) and the condition 2° of Theorem 3 are satisfied with f,L,Ω and Γ defined by (28–31). Further, let $(\bar{\lambda}, \bar{y})$, $(\bar{\gamma}, \bar{z})\in R^q\times R^p$ be solutions of (26) and (32–33), respectively. Then the inequalities (24–25) are satisfied.

Remark 6

Assume that there exist matrices $X_{q\times q}$, $Z_{q\times q}$ such that for all $u\in R^p$, $v\in R^q$ the matrix

$$T(u, v) = D_vL(u, v) + X(u, v) \text{ with } D_vL(u, v) = \left[\frac{\partial Li(u, v)}{\partial v_j}\right]$$

has a representation of the form

$$T(u, v) = P_0(I + Z(u, v)),$$

with a constant nonsingular matrix P_0. In addition, we assume that

$$||P_0Z(u, v)|| \le v_1,\ ||X(u, v)|| \le v_2 \text{ for all } u\in R^p,\ v\in R^q.$$

Now with a suitable choice of B, namely $B = P_0$, the condition (6) for our problem is satisfied with $h = v_1 + v_2$.

Of course we have

$$L(x(b),\lambda)-L(x(b),\mu)-P_0(\lambda-\mu) = [\int_0^1 D_vL(x(b),\mu+\tau(\lambda-\mu))d\tau-P_0](\lambda-\mu)$$

$$= \int_0^1 [P_0Z(x(b),\ \mu+\tau(\lambda-\mu))-X(x(b),$$

$$\mu+\tau(\lambda-\mu))]\ d\tau(\lambda-\mu),$$

and hence, we have the assertion.

Now if $p=q$ then we may take $X(u, v) = D_uL(u, v)$ (see [11]). Moreover, if the vector function L is linear of the form

$$L(u, v) = \widetilde{M}u + \widetilde{N}v - \widetilde{K}$$

with $\widetilde{M}_{q\times q}$, $\widetilde{N}_{q\times q}$, $\widetilde{K}_{q\times 1}$, then we may choose $B=\widetilde{M}+\widetilde{N}$ provided that it is nonsingular. Notice that

$$T(u, v) = \widetilde{M} + \widetilde{N},\ v_1 = 0,\ ||\widetilde{M}|| \le v_2 = h.$$

Now we consider the linear case of functions Ω and Γ. Assume that

$$\Omega(t, w_1, w_2,\dots,w_s, c) = \sum_{i=1}^{s} l_i(t)w_i + m(t)c,\ t\in J, \tag{34}$$

$$\Gamma(w) = Nw, \tag{35}$$

where $l_i : J \to R_+$, $m : J \to R_+$ and $N \geq 0$.

Put $\quad M(t) = m(t)\, d(1 - dh)^{-1} N,$

$$V(t) = m(t)\, d(1 - dh)^{-1} \,||L(y_0(b), \lambda_0)|| + v(t)$$

for $t \in J$ with $dh \in [0, 1)$, and

$$v(t) = ||f(t, y_0(\beta_1(t)),\ldots,y_0(\beta_s(t)), \lambda_0) - y_0(t)||, \quad t \in J.$$

For our linear case, (7) and (8) have the form

$$M(t)w(b) + \sum_{i=1}^{s} l_i(t)w(\beta_i(t)) + V(t) \leq w(t), \; t \in J, \tag{36}$$

$$w(t) = M(t)w(b) + \sum_{i=1}^{s} l_i(t)w(\beta_i(t)), \; t \in J, \tag{37a}$$

$$u = d(1 - dh)^{-1} Nw(b). \tag{37b}$$

Indeed, if

$$\widetilde{S} = \sup_{t \in J} \left[Mt + \sum_{i=1}^{s} l_i(t) \right] < 1, \tag{38}$$

then the function

$$w^*(t) = (1 - \widetilde{S})^{-1} \sup_{\tau \in J} V(\tau)$$

is a solution of (36).

Moreover, in the class of functions $w \in E$ satisfying the condition $w(t) \leq w^*(t)$, $t \in J$, the function $w(t) = 0$, $t \in J$ and $u = 0$ are the only solution of (37).

The last assertion follows from the fact that

$$w(t) \leq (\widetilde{S})^n\, w^*(t), \; t \in J, \; n = 0, 1, \ldots.,$$

which can be proved by induction.

The condition (38) guarantees that Assumption H_2 is satisfied for our case. Now we will give the conditions weaker than (38). Of course, as the function w^* of Assumption H_2, we can take the solution of the following equation

$$M(t)w(b) + \sum_{i=1}^{s} l_i(t)w(\beta_i(t) + V(t) = w(t), \; w \in J \tag{39}$$

instead of the inequality (36).

For $t \in J$, $i_n = 1, 2, \dots, s$, $n = 0, 1, \dots$, let

$$\beta_0^{i_0}(t) = t,\ \beta_0^{i_0,\dots,i_{n+1}}(t) = \beta_n^{i_1,\dots,i_{n+1}}(\beta_{i_0}(t)),$$

$$l_0^{i_0}(t) = \frac{1}{s},\ l_{n+1}^{i_0,\dots,i_{n+1}}(t) = l_{i_0}(t)\, l_n^{i_1,\dots,i_{n+1}}(\beta_{i_0}(t)).$$

Now we are able to formulate conditions by which Assumption H_2 is fulfilled for our linear case (34–35).

Lemma 4

1° $\quad P(t) = \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \dots \sum_{i_n=1}^{s} l_n^{i_0,\dots,i_n}(t) M(\beta_n^{i_0,\dots,i_n}(t)) < \infty,$

2° $\quad Q(t) = \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \dots \sum_{i_n=1}^{s} l_n^{i_0,\dots,i_n}(t)\, V(\beta_n^{i_0,\dots,i_n}(t)) < \infty,$

3° $\quad P(b) < 1$, then

(i) the function w^* defined by

$$w^*(t) = Q(t) + P(t)\, Q(b)(1-P(b))^{-1},\ t \in J \tag{40}$$

is a solution of (39), and there is no other solution of (39) in the class of functions $u \in E$ such that $u(t) \le w^*(t)$, $t \in J$,

(ii) in the class of functions $u \in E$ satisfying the condition $u(t) \le w^*(t)$, $t \in J$, the function $u(t) = 0$, $t \in J$ is the only solution of the equation (37a).

The functions M, V, l_i, β_i are defined in this section.

Proof

We prove the function w^* satisfies the equation (39). At first, we see that

$$\sum_{i=1}^{s} l_i(t) Q(\beta_i(t)) = \sum_{i=1}^{s} l_i(t) \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \dots \sum_{i_n=1}^{s} l_n^{i_0,\dots,i_n}(\beta_i(t)) V(\beta_n^{i_0,\dots,i_n}(\beta_i(t)))$$

$$= \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \dots \sum_{i_{n+1}=1}^{s} l_{n+1}^{i_0,\dots,i_{n+1}}(t) V(\beta_{n+1}^{i_0,\dots,i_{n+1}}(t)),\ t \in J,$$

and, by changing the sum index, we have

$$\sum_{i=1}^{s} l_i(t)\, Q(\beta_i(t)) = Q(t) - V(t),\ \ t \in J.$$

Similarly, we can get

$$\sum_{i=1}^{s} l_i(t)\, P(\beta_i(t)) = P(t) - M(t), t \in J.$$

Now, combining it and using 3°, we have

$$M(t)w^*(b) + \sum_{i=0}^{s} l_i(t)w^*(\beta_i(t)) + V(t)$$

$$= M(t)Q(b)\,[1 + P(b)(1-P(b))^{-1}] + Q(t) + Q(b)(1-P(b))^{-1}(P(t)-M(t))$$

$$= w^*(t),\ \ t \in J,$$

so w^* is the solution of (39).

Indeed, the function w^* is the unique solution of the equation

$$w(t) = P(t)\, w(b) + Q(t) = Sw(t),\ \ t \in J. \tag{41}$$

Let $u_0 \in E$ be a solution of (39) such that $u_0(t) \le w^*(t)$, $t \in J$. Put

$$\varphi_1(t) = M(t)\, u_0(b) + V(t), t \in J.$$

It is known that the equation

$$w(t) = \sum_{i=1}^{s} l_i(t)\, w(\beta_i(t)) + \varphi_1(t),\ \ t \in J \tag{42}$$

has a unique solution

$$q_0(t) = P(t)\, u_0(b) + Q(t) = Su_0(t),\ \ t \in J,$$

in the class $0 \le w(t) \le w^*(t)$, $t \in J$. Since u_0 is also a solution of (42) in this class, then $q_0 = u_0$. Hence u_0 is a solution of (41), and now $u_0 = w^*$.

The part (ii) may be proved by a similar argument.

Remark 5

Suppose that

1° $a \ge 0$, $\beta_i(t) \le \beta_i t$, $0 \le \beta_i \le 1$, $t \in J$, $i = 1, 2, \dots, s$,

2° $l_i(t) = l_i \ge 0$, $i = 1, 2, \dots, s$, 3° $\mathrm{M}(t) = \mathrm{M}_1 t^{\xi}, \ t \in \mathrm{J}, \mathrm{M}_1 \geq 0,$

4° $\mathrm{V}(t) = \mathrm{M}_2 t^{\rho}, \ t \in \mathrm{J}, \mathrm{M}_2 \geq 0,$

5° $0 \leq \sum_{i=1}^{s} l_i \, \beta_i^{\eta} < 1, \ \eta = \min(\xi, \rho),$

6° $\mathrm{M}_3 = 1 - \sum_{i=1}^{s} l_i \, \beta_i^{\xi} - \mathrm{M}_1 b^{\xi} > 0,$

then the assumptions of Lemma 4 are satisfied and

$$w^*(t) \leq \mathrm{M}_2 \, [t^{\rho} + \mathrm{M}_1 t^{\xi} \, b^{\rho}/\mathrm{M}_3] \left(1 - \sum_{i=1}^{s} l_i \beta_i^{\rho}\right)^{-1}, \ t \in \mathrm{J}$$

Theorems 1 and 4 and Lemma 4 imply.

Theorem 6

Suppose Assumption H_1 (except for 1°), the assumptions of Lemma 4 and conditions (27, 34–35) are satisfied with f, L, Ω and Γ defined by (28–31). Then there exists a unique solution $(\bar{\lambda}, \bar{y}) \in \mathrm{R}^q \times \mathrm{R}^p$ of (26) with the properties (15–17), where

$$\left.\begin{aligned} u_0 &= u^* = d(1-dh)^{-1} \, [\mathrm{N} w^*(b) + \|\mathrm{L}(y_0(b), \lambda_0)\|] \\ u_n &= (dh)^n u^* + d\mathrm{N} \sum_{i=1}^{n-1} (dh)^{n-1-i} w_i(b), \ n = 1, 2, \ldots\ldots \end{aligned}\right\}$$

$$\left.\begin{aligned} w_0(t) &= w^*(t) = \mathrm{Q}(t) + \mathrm{P}(t) \, \mathrm{Q}(b) \, (1-\mathrm{P}(b))^{-1} \\ w_{n+1}(t) &= \sum_{i=1}^{s} l_i(t) \, w_n(\beta_i(t)) + m(t) u_n, \ n = 0, 1, \ldots\ldots \end{aligned}\right\}$$

Now Theorems 3 and 5 and Remark 5 imply.

Theorem 7

If

1° Assumption H_1 (except for 1°) and the conditions (27, 34–35) are satisfied with f, L, Ω and Γ defined by (28–31),

2° the assumption of Lemma 4 are satisfied with V_0 instead of V, where

$$\mathrm{V}_0(t) = m(t) d v^{**} (1-dh)^{-1} + \max(v^*(t), \ \|\bar{y}(t)\| + \|\bar{z}(t)\|), \ t \in \mathrm{J}$$

$$v^{**} = ||L(\bar{z}(b), \bar{\gamma}) - L_0(\bar{z}(b), \bar{\gamma}),$$

$$v^*(t) = ||f(t, \bar{z}(\beta_1(t)), \bar{z}(\beta_2(t)),......, \bar{z}(\beta_s(t)), \bar{\gamma}) - \bar{z}(t)||\ t \in J,$$

then

$$||\bar{y}(t) - \bar{z}(t)|| \leq Q(t) + P(t)\ Q(b)\ (1 - P(b))^{-1},\ t \in J$$

$$||\bar{\lambda} - \bar{\gamma}|| \leq dN\ (1 - dh)^{-1}\ Q(b)\ (1 - P(b))^{-1} + d(1-dh)^{-1} v^{**}.$$

6. The Differential Equation with Deviated Arguments

Consider the problem

$$\left.\begin{aligned} &x'(t) = f(t, x(\alpha_1(t)),...x(\alpha_r(t)), x'(\beta_1(t)),...,x'(\beta_s(t)),\lambda),\ t \in J = [0,b] \\ &x(0) = x_p \in R^p,\ A(x(b)) = \Theta \in R^q. \end{aligned}\right\} \quad (43)$$

The functions $f: J \times (R^p)^{r+s} \times R^q \to R^p$, $\alpha_i: J \to J$, $\beta_j: J \to J$, $A: R^p \to R^q$ are continuous. We seek a parameter $\lambda \in R^q$ and a function $x \in C_1(J, R^p)$ such that (43) to be satisfied ($C^1(J, R^p)$ denotes the class of all continuous functions from J into R^p with a continuous first derivative).

By the substitution $y(t) = x'(t)$, the problem (43) is equivalent to the following one

$$\left.\begin{aligned} &y(t) = F(t, y, \lambda),\ t \in J \\ &A(x_p + \int_0^b y(\tau) d\tau) = \Theta \end{aligned}\right\} \quad (44)$$

with $F(t,y,\lambda) \equiv f(t, x_p + \int_0^{\alpha_1(t)} y(\tau)d\tau, ..., x_p + \int_0^{\alpha_r(t)} y(\tau)d\tau, y(\beta_1(t)),...,$

$y(\beta_s(t)), \lambda)$.

The problem (44) is a particular case of (1–2). A question of the existence and uniqueness and continuous dependence of solutions on the right-hand side of (44) was considered in [2,3,6–10]. Using the general considerations we give the results more general than obtained in [2,3,6,10]. The numerical methods for finding a numerical solution of (43) are considered in [4,5].

We introduce

Assumption H_3. Suppose that

1° the functions $f:J\times(R^p)^{r+s}\times R^q\to R^p$, $A:R^p\to R^q$ are continuous,

2° $\alpha_i, \beta_j \in C(J, J)$, $\alpha_1(0) = \beta_j(0)$, $i = 1, 2,\dots, r, j=1, 2,\dots, s$,

3° there exists a nonnegative function $\Omega:J\times(R_+)^{r+s+1}\to R_+$, which is nondecreasing and monotonicly continuous with respect to the last $r+s+1$ variables $\Omega(t, 0, \dots, 0)\equiv 0$, and

$$||f(t, x_1,\dots, x_r, g_1, \dots, g_s, \mu) - f(t, \bar{x}_1,\dots, \bar{x}_r, \bar{g}_1,\dots, \bar{g}_s, \mu)||$$
$$\leq\Omega(t, ||x_1-\bar{x}_1||,\dots, ||x_r-\bar{x}_r||, ||g_1-\bar{g}_1||,\dots, ||g_s-\bar{g}_s||, ||\mu-\bar{\mu}||)$$

for $t\in J$, $\mu, \bar{\mu} \in R^q$, $x_i, \bar{x}_i, g_j, \bar{g}_j\in R^p$, $i=1, 2, \dots, r$, $j=1,2,\dots, s$.

4° there exist a nonsingular square matrix B of order q and constants $h\geq 0$, $d>0$ such that $hd<1$, $d\geq||B^{-1}||$ and

$$||A(x_p+\int_0^b F(t,y,\bar{\mu})dt)-A(x_p+\int_0^b F(t,y,\mu)dt-B(\bar{\mu}-\mu)||\leq h||\bar{\mu}-\mu||$$

for μ, $\bar{\mu}\in R^q$, $y\in R^p$, where the matrix norm is consistent with the vector norm.

5° there exists a nonnegative, nondecreasing and mono-tonicly continuous function $\Gamma:R_+\to R_+$, $\Gamma(0)=0$, such that the inequality

$$||A(x) - A(g)|| \leq \Gamma(||x - g||)$$

holds for $x, g\in R^p$,

6° there exists a solution $u^*\in C(J, R_+)$ of the inequality

$$\Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t)), \dots, u(\beta_s(t)),$$

$$d(1-dh)^{-1}\,\Gamma^*(u)) + v(t) \leq u(t), t\in J, \tag{45}$$

where $\Gamma^*(u)=\Gamma(\int_0^b \Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau,\dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t)),\dots,$

$$u(\beta_s(t)),0)dt) + ||A(x_p + \int_0^b F(t, y_0, \lambda_0)\, dt)||,$$

$$v(t) = ||F(t, y_0, \lambda_0) - y_0(t)||, \; t\in J,$$

7° in the class of functions $u \in M(J, R_+)$ satisfying the condition $0 \le u(t) \le u^*(t)$, $t \in J$, the function $u(t) = 0$, $t \in J$ is the only solution of the equation

$$\Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t)), \dots, u(\beta_s(t)), d(1-dh)^{-1} \times$$

$$\Gamma^{**}(u)) = u(t), \ t \in J, \tag{46}$$

where $\Gamma^{**}(u) = \Gamma(\int_0^b \Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t))), \dots,$

$u(\beta_s(t)), 0)\, dt)$

($M(J, R_+)$ denotes the collection of all measurable and bounded functions in J with a range in R_+).

Now the functions Ω and Γ are of the form

$$\Omega(t, u(\cdot), c) = \Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t)), \dots, u(\beta_s(t)), c),$$

$$\Gamma(u(\cdot)) = \Gamma(\int_0^b \Omega(t, \int_0^{\alpha_1(t)} u(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} u(\tau)d\tau, u(\beta_1(t)), \dots, u(\beta_s(t)), 0)dt).$$

Indeed, Assumption H_3 guarantees Assumptions H_1 and H_2 to be satisfied. We define the sequences $\{y_n\}$ and $\{\lambda_n\}$ by the relations

$$y_{n+1}(t) = F(t, y_n, \lambda_n), \ t \in J, n = 0, 1, \dots\dots \tag{47}$$

$$\lambda_{n+1} = \lambda_n - B^{-1} D(y_n, \lambda_n), \ n = 0, 1, \dots.. \tag{48}$$

where $D(y, \lambda) = A(x_p + \int_0^b F(t, y, \lambda)\, dt)$,

and $y_0 \in C(J, R^p)$, $\lambda_0 \in R^q$ are arbitrary fixed elements.

Theorem 1 implies

Theorem 8

If Assumption H_3 is satisfied, then there exists a solution $(\bar{\lambda},\bar{y}) \in R^q \times R^p$ of (44) which is the limit of the sequences $\{\lambda_n\}$ and $\{y_n\}$ and the estimates (15–16) hold with the sequences $\{u_n\}$, $\{w_n\}$ defined by

$$\left.\begin{aligned} u_0 &= \bar{u} = d(1-dh)^{-1}\Gamma^*(u^*) \\ u_{n+1} &= d[hu_n + \Gamma^{**}(w_n)],\ n = 0, 1, \dots. \end{aligned}\right\}$$

$$\left.\begin{aligned} & w_0(t) = u^*(t),\ t \in J \\ & w_{n+1}(t) = \Omega\Big(t, \int_0^{\alpha_1(t)} w_n(\tau)d\tau, \dots, \int_0^{\alpha_r(t)} w_n(\tau)d\tau, w_n(\beta_1(t)), \dots, w_n(\beta_s(t)), u_n\Big) \\ & t \in J,\ n = 0,1,\dots \end{aligned}\right\}$$

Suppose we are given the problem

$$\left.\begin{aligned} & z(t) = F_0(t, z, \gamma),\ t \in J \\ & A_0(\bar{x}_p + \int_0^b z(\tau)d\tau) = \Theta \end{aligned}\right\} \qquad (49)$$

with

$$F_0(t,z,\gamma) = f_0(t, \bar{x}_p + \int_0^{\rho_1(t)} z(\tau)d\tau, \dots, \bar{x}_p + \int_0^{\rho_r(t)} z(\tau)d\tau, z(\delta_1(t)), \dots, z(\delta_s(t)), \gamma),$$

where f_0, A_0, $\bar{x}_p$, ρ_i, δ_j, γ have the same properties as $f, A, x_p, \alpha, \beta_j, \lambda$. Now we are in a position to formulate the theorem on the continuous dependence of solutions of (44) on the right-hand side of (44). Theorem 3 implies.

Theroem 9

If Assumption H_3 (except for 6°–7°) is satisfied, and

1° $(\bar{\lambda}, \bar{y})$, $(\bar{\gamma}, \bar{y}) \in R^q \times R^p$ are solutions of (44) and (49), respectively,

2° the sequence $\{v_n\}$ defined by

$$\left.\begin{aligned} &v_0(t) = ||\bar{y}(t)|| + ||\bar{z}(t)||, t\in J \\ &v_{n+1}(t) = \Omega(t, \int_0^{\alpha_1(t)} v_n(\tau)d\tau, \ldots, \int_0^{\alpha_r(t)} v_n(\tau)d\tau, v_n(\beta_1(t)), \ldots, v_n\beta_s(t)), \\ &d(1-dh)^{-1}[\Gamma^{**}(v_n)+v^{**}] + v^{**}]+v^*(t), t\in J, n=0,1,\ldots. \end{aligned}\right\}$$

$$v^*(t) = ||F(t, \bar{z}, \bar{\gamma}) - \bar{z}(t)||, t \in J,$$

$$v^{**} = ||A(x_p + \int_0^b F(t, \bar{z}, \bar{\gamma})\, dt) - A_0(\bar{x}_p + \int_0^b F_0(t, \bar{z}, \bar{\gamma})\, dt||,$$

has a limit $\bar{v}{:}J \to R_+$,

then the inequalities

$$||\bar{y}(t) - \bar{z}(t)|| \le \bar{v}(t), \quad t \in J,$$

$$||\bar{\lambda} - \bar{\gamma}|| \le d(1 - dh)^{-1} [\Gamma^{**}(\bar{v}) + v^{**}], \text{ hold.}$$

Now we consider the linear case for Ω and Γ. Let

$$\Omega(t, c_1, \ldots, c_r, d_1, \ldots, d_s, d_0) = \sum_{i=1}^{r} k_i(t)\, c_i + \sum_{i=1}^{s} l_i(t) d_i + m(t) d_0, \quad (50)$$

$$\Gamma(c) = Nc, N \ge 0, \quad (51)$$

where $k_i, l_j, m \in C(J, R_+)$, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, s$.

Put $M_1(t) = m(t)\, d(1 - dh)^{-1} N, \; t \in J,$

$$V_1(t) = m(t) d(1-dh)^{-1} ||A(x_p + \int_0^b F(t, y_0, \lambda_0) d\tau)|| + ||F(t, y_0, \lambda_0) - y_0(t)||, \quad t\in J.$$

Now (45) and (46) are of the form

$$Gu(t) + M_1(t) \int_0^b Gu(\tau)\, d\tau + V_1(t) \le u(t), t\in J, \quad (52)$$

$$Gu(t) + M_1(t) \int_0^b Gu(\tau) d\tau = u(t), t\in J, \quad (53)$$

where $Gu(t) = \sum_{i=1}^{r} k_i(t) \int_0^{\alpha_i(t)} u(\tau)d\tau + \sum_{i=1}^{s} l_i(t)u(\beta_i(t)),\ t\in J.$

Put $Q_0(t) = \sum_{i=1}^{r} k_i(t) + M_1(t) \int_0^b \sum_{i=1}^{r} k_i(\tau)d\tau,$

$$P_1(t) = \sum_{i=1}^{r} k_i(t)\,\alpha_i(t) + \sum_{i=1}^{s} l_i(t),\ \ t\in J,$$

$$Q_1(t) = M_1(t) \int_0^b P_1(\tau)d\tau + P_1(t),\ t\in J,$$

$$\overset{*}{u}_0(t) = V_1(t) + (1 - \int_0^b Q_0(\tau)d\tau)^{-1} Q_0(t) \int_0^b V_1(\tau)d\tau,\ t\in J,$$

$$\overset{*}{u}_1(t) = (1 - \sup_{t\in J} Q_1(t))^{-1} \sup_{\tau\in J} V_1(\tau),\ t\in J.$$

Theorem 8 implies.

Theorem 10

If

1° the conditions 1°–5° of Assumption H_3 and (50–51) are satisfied;

2° $\int_0^b Q_0(t)\,dt < 1$ and $l_j(t) = 0,\ t\in J,\ j = 1, 2, \ldots.$

or

2°° $\sup_{t\in J} Q_1(t) < 1,$

then the assertion of Theorem 8 is true with the function $u^* = \overset{*}{u}_0$ or $u^* = \overset{*}{u}_1$ for the case 2° or 2°°, respectively.

Now, we want to consider the problem (43) of the delay type when the functions α_i and β_j satisfy the conditions

$0\le\alpha_i(t)\le t,\ 0\le\beta_j(t)\le\beta_j t,\ 0\le\beta_j\le 1,\ t\in J,\ i=1, 2, \ldots, r, j=1,2,\ldots,s.$ (54)

Assume that (50) and (51) are satisfied with

$$\sum_{i=1}^{r} k_i(t) = k,\ l_j(t) = l_j,\ t \in \mathrm{J}, j = 1,2,\ldots, s. \qquad (55)$$

Now (45) and (46) are in shape of

$$\mathrm{H}u(t) + \mathrm{M}_1(t) \int_0^b \mathrm{H}u(\tau)d\tau + \mathrm{V}_1(t) = u(t),\ t \in \mathrm{J}, \qquad (56)$$

$$\mathrm{H}u(t) + \mathrm{M}_1(t) \int_0^b \mathrm{H}u(\tau)d\tau = u(t),\ t \in \mathrm{J}, \qquad (57)$$

where $\mathrm{H}u(t) = k\int_0^t u(\tau)\, d\tau + \sum_{i=1}^{s} l_i u(\beta_i t),\ t \in \mathrm{J}.$

Here we are in a position to formulate a lemma by which the conditons 6° and 7° of Assumption H_3 are satisfied for the linear case (50–51) of Ω and Γ when the problem (44) is of the delay type.

Lemma 5 (see[3]). If

1° $\mathrm{M}_1, \mathrm{V}_1 \in \mathrm{C}(\mathrm{J}, \mathrm{R}_+)$ and are nondecreasing,

2° $0 \le \beta_i \le 1,\ i = 1, 2, \ldots, s,$

3° $0 \le \sum_{i=1}^{s} l_i \beta_i < 1,\ l_i \ge 0,$

4° $\mathrm{T}_1(t) = \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \cdots \sum_{i_n=1}^{s} \left(\prod_{j=0}^{n-1} l_{i_j}\right) \mathrm{V}_1\left(t \prod_{j=0}^{n-1} \beta_{i_j}\right) < \infty,\ \prod_0^{-1} .. = 1 \,\&\, \mathrm{T}_1 \in \mathrm{C}(\mathrm{J}, \mathrm{R}_+),$

5° $\mathrm{T}_2(t) = \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \cdots \sum_{i_n=1}^{s} \left(\prod_{j=0}^{n-1} l_{i_j}\right) \mathrm{M}_1\left(t \prod_{j=0}^{n-1} \beta_{i_j}\right) < \infty$, and $\mathrm{T}_2 \in \mathrm{C}(\mathrm{J}, \mathrm{R}_+),$

6° there exists a unique nondecreasing solution $\tilde{u} \in \mathrm{C}(\mathrm{J}, \mathrm{R}_+)$ of the equation

$$u(t) = \frac{k}{s} \sum_{n=0}^{\infty} \sum_{i_0=1}^{s} \cdots \sum_{i_n=1}^{s} \left(\prod_{j=0}^{n-1} l_{i_j}\right) \int_0^{t\prod_{j=0}^{n-1}\beta_{i_j}} u(\tau)d\tau + \frac{1}{s}\mathrm{T}_2(t)\int_0^b \mathrm{H}u(\tau)d\tau + \frac{1}{s}\mathrm{T}_1(t),$$

$t \in \mathrm{J},\ k \ge 0,$ (58)

then

(i) in the class of functions $u \in M(J, R_+)$ satisfying the condition $0 \le u(t) \le \tilde{u}(t)$, $t \in J$, the function $\tilde{u}$ is the unique, continuous and nondecreasing solution of the equation (56),

(ii) in the class of functions $u \in M(J, R_+)$ satisfying the condition $0 \le u(t) \le \tilde{u}(t)$, $t \in J$, the function $u(t) = 0$, $t \in J$, is the unique solution of the equation (57).

Theorem 8 and Lemma 5 imply.

Theorem 11

If Assumption $H_3(1°-5°)$ for the linear case (50–51), and the conditions (54) and (55) and also the assumptions of Lemma 5 are satisfied, then the assertion of Theorem 8 is true with $\tilde{u}$ instead of u^*.

REFERENCES

1. R. Conti, Problems lineaires pour les equations differentielles ordinaires, *Math. Nachr.* 23(1961), 161–178.

2. L. A. Goma, The method of successive approximations in a two-point boundary value problem with a parameter (Russian), *Ukrain. Mat. Z.* 29(1977), 800–807.

3. T. Jankowski, Boundary value problems with a parameter of differential equations with deviated arguments, *Math. Nachr.* 125(1986), 7–28.

4. T. Jankowski, One-step methods for ordinary differential equations with parameters, *Apl. Mat.* 35(1990), 67–83.

5. T. Jankowski, One-step methods for retarded differential equations with parameters, *Computing* 43(1990), 343–359.

6. N.S. Kurpel and A.G. Marusjak, A multipoint boundary value problem for differential equations with parameters (Russian), Ukrain. *Mat. Z.* 32(1980), 223–226.

7. A. Pasquali, Un procedimento di calcolo connesso ad un noto problema ai limiti per l'equazione $x'=f(t,x,\lambda)$, *Le Mathematiche* 23(1968), 319–328.

8. T. Pomentale, A boundary-value problem for the equation $y'=\lambda f(x,y)+h(x,y)-g(x,y)y$, *ZAMM* 54(1974), 723–728.

9. T. Pomentale, A constructive theorem of existence and uniqueness for the problem $y'=f(x,y,\lambda)$, $y(a)=\alpha$, $y(b)=\beta$, *ZAMM* 56(1976), 387–388.

10. Z.B. Seidov, Boundary value problems with a parameter of differential equations in a Banach space (Russian), Sibirsk. *Mat. Z.* 9(1968), 223–228.

11. J. Stoer and R. Bulirsch, Introduction to numerical analysis, New York, Heidelberg, Berlin 1980.

12. T. Wazewski, Sur un procede de prouver la convergence des approximations successives sans utilisation des series de comparaison, *Bull. Acad. Sci. ser. sci. math. astr. et phys.*, 8(1960), 45–52.

13. L.A. Zyvotovskii, On the existence a of solutions of differential equations with deviated argument of neutral type (Russian), *Differen. Urav.* 8(1972), 1936–1942.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 27 - 40

A GEOMETRIC RATIONAL SPLINE WITH TENSION CONTROLS: AN ALTERNATIVE TO THE WEIGHTED NU-SPLINE

Muhammad Sarfraz

Department of Mathematics,
University of the Punjab,
Quaid-i-Azam Campus,
Lahore-54590, Pakistan.

ABSTRACT

A description and analysis of a rational cubic spline curve is made for use in CAGD (Computer Aided Geometric Design). This rational spline provides not only computationally simple alternative to the exponenetial based spline under tension [Barsky'84, Cline'74, Preuss'76, Schweikert'66] but also provides an alternative to the well known existing methods like cubic ν-spline of Nielson [Nielson'74], γ-splines of Boehm [Bohem'85] and weighted ν-splines [Foley'87]. The method also recovers the *rational spline with tension* [Gregory and Sarfraz'90] and the ν-spline of Nielson [Nielson'74] as a special case. Two shape parameters are introduced in each interval which provide a variety of shape controls like *point* and *interval* tension. The spline is presented in interpolatory form.

Keywords. Rational cubic spline, rational Bernstein-Bezier, interpolation, tension, shape control.

1. INTRODUCTION

A rational cubic spline with tension was described and analysed in [Gregory and Sarfraz'90] with a view to its application in CAGD (Computer Aided Geometric Design). It provides a C^2 computationally simpler alternative to the exponential spline-under-tension [Schweikert'66, Cline'74, Preuss'76]. Regarding shape characteristics, it has a shape control parameter associated with each interval which can be used to flatten or tighten the curve both locally and globally.

This paper generalizes the idea of the rational cubic spline with tension and introduces some further shape parameters in the description of continuity when the piecewise rational functions are stitched together. The shape parameters provide a variety of shape controls like *interval* and *point* tensions and they occure in the scheme in such a way that given n data points, the rational spline has $2n-1$ parameters that control the shape of the piecewise rational cubic curve. Thus the characteristics and the number of shape parameters occuring in this scheme are similar to those in weighted ν-spline [Foley'87] except with the difference that our interval tensions will play a more extensive role as compared to the interval weights in weighted ν-spline of Foley. In addition to this our scheme involves a rational piecewise function whereas in [Foley'87] piecewise cubics were used. Furthermore a GC^2 continuity will also be achieved as compared to only C^1 continuity.

The rational spline is not restricted to *homogeneous coordinate* form of having a GC^2 cubic spline numerator and denominator. Thus, in general, it is not a projection from a GC^2 cubic spline in $\mathbb{R}^4$ into $\mathbb{R}^3$ as, for example, in the case of non-uniform rational B-spline (NURBS). This, we believe, gives more freedom to develop shape control parameters for the curve, which behave in a well defined and well controlled way.

The shape parameters of the rational spline can be utilized to achieve a variety of shape controls like point and interval tensions. Since the spline is defined on a non-uniform knot partition, or by cummulative chord length, or by some other appropriate means.

The rational spline is based on a rational cubic Hermite interpolant which is introduced in Section 2 together with some preliminary analysis. Section 3 describes the geometric rational spline and analyses of its behaviour with respect to shape parameters in each interval is done in Section 4. Section 5 consists of some illustrative examples.

2. C^1 PIECEWISE RATIONAL CUBIC HERMITE INTERPOLANT

A piecewise rational cubic Hermite parametric function $P \in C^1[t_0, t_n]$, with parameters r_i, $i=0,\dots,n-1$, is defined for $t \in [t_i, t_{i+1}]$, $i=0,\dots,n-1$, by

$$(1) \quad P(t) = P_i(t_i\ r_i)$$

$$= \frac{(1-\theta)^3 X_i + \theta(1-\theta)^2(r_i X_i + h_i D_i) + \theta^2(1-\theta)(r_i X_{i+1} - h_i D_{i+1}) + \theta^3 X_{i+1}}{1 + (r_i - 3)\,\theta(1-\theta)}$$

where the notations X_i and $D_i \in \mathbb{R}^N$ are, respectively, the data values and the first derivative values at the knots t_i, $i=0,\ldots,n$ with $t_0<t_1<,\ldots,<t_n$, $h_i=t_{i+1}-t_i$, $\theta=(t-t_i)/h_i$ and $r_i \geq 0$.

The function $P(t)$ has the Hermite interpolation properties that

(2) $P(t_i) = X_i$ and $P^{(1)}(t_i) = D_i$, $i = 0, \ldots, n$.

The r_i, $i = 0,\ldots,n-1$, will be used as shape parameters to control and fine tune the shape of the curve. The denominator in (1) can be written as

$$(1-\theta)^3 + r_i\,\theta(1-\theta)^2 + r_i\,\theta^2(1-\theta) + \theta^3.$$

The case $r_i=3$, $i=0,\ldots,n-1$, is that of cubic Hermite interpolation and the restriction $r_i \geq -1$ ensures a positive denominator in (1).

For $r_i \neq 0$, (1) can be written in the form

(3) $P_i(t_i;r_i) = R_0(\theta;r_i)X_i + R_1(\theta;r_i)V_i + R_2(\theta;r_i)W_i + R_3(\theta;r_i)X_{i+1}$,

where

(4) $V_i = X_i + h_i D_i/r_i$, $W_i = X_{i+1} - h_i D_{i+1}/r_i$

and $R_j(\theta;r_i)$, $=0, 1, 2, 3$, are appropriately defined rational functions with

(5) $\sum_{j=0}^{3} R_j(\theta;r_i) = 1.$

Moreover, these functions are rational Bernstein-Bezier weight functions which are non-negative for $r_i > 0$. Thus in $\mathbb{R}^N$, $N>1$ and for $r_i > 0$, we have:

Proposition 1 (Convex hull property)

The curve segement P_i lies in the convex hull of the control points $\{X_i, V_i, W_i, X_{i+1}\}$.

We now consider the variation diminishing property of the rational cubic and for this we require some preliminary analysis. Let

(6) $p(\theta) = \sum_{i=0}^{3} a_i A_i \binom{3}{i} \theta^i (1-\theta)^{3-i}$, and

(7) $$q(\theta) = \sum_{i=0}^{3} a_i \binom{3}{i} \theta^i (1-\theta)^{3-i}$$

be scalar curves with $a_i > 0$, $\forall_i$. Since $p(\theta)$ is a Bezier curve and since $a_i > 0$, we have

$$V(p) \leq V(a_0 A_0, \ldots, a_3 A_3) = V(A_0, \ldots, A_3),$$

where $V(\cdot)$ denotes the number of sign change of a function or sequence. Also, since $q(\theta) > 0$, we have

(8) $$V\left(\frac{p}{q}\right) = V(p) \leq V(A_0, \ldots, A_3).$$

Let $p(\theta)$ now be considered as planar curve, say, $p(\theta) = (p_1(\theta), p_2(\theta))$ where $A_i = (x_i, y_i) \in \mathbb{R}^2$ and let $L \equiv ax+by+c=0$ be any line. Then the number of times the line L crosses the rational cubic curve $p(\theta)/q(\theta)$ is the same as it crosses the cubic Bezier curve $p(\theta)$, since $q(\theta) > 0$. This number is (using similar arguments as in [Goodman'89]).

$$V(ap_1+bp_2+c) = V(a\sum a_i x_i \binom{3}{i} \theta^i (1-\theta)^{3-i} + b\sum a_i y_i \binom{3}{i} \theta^i (1-\theta)^{3-i} + c)$$

$$= V(a\sum x_i \binom{3}{i} \theta^i (1-\theta)^{3-i} + b\sum y_i \binom{3}{i} \theta^i (1-\theta)^{3-i} + c\sum \binom{3}{i} \theta^i (1-\theta)^{3-i}$$

$$= V\left(\sum (ax_i + by_i + c) \binom{3}{i} \theta^i (1-\theta)^{3-i}\right)$$

$$\leq V(ax_0 + by_0 + c, \ldots, ax_3 + by_3 + c)$$

= the number of times the line L crosses the ploygon $A_0, \ldots, A_3$.

These arguments can be extended to a rational curve of any degree in $\mathbb{R}^N$ with any hyper plane of dimension $N-1$. Thus we have:

Proposition 2 (Variation diminishing property)

The curve segment P_i crosses any (hyper) plane of dimension $N-1$ no more times than it crosses the control polygon joining X_i, V_i, W_i, X_{i+1}.

The rational cubic (1) can, then, be expressed in the form

(9) $P_i(t;r_i) = l_i(t) + e_i(t;r_i)$, where

(10) $l_i(t) = (1-\theta)X_i + \theta X_{i+1}$,

(11) $$e_i(t;r_i) = \frac{h_i\theta(1-\theta)\{(\Delta_i - D_i)(\theta-1) + (\Delta_i - D_{i+1})\theta\}}{1 + (r_i - 3)\theta(1-\theta)}$$

This immediately leads to:

Proposition 3 (Interval tension property)

For given fixed (or bounded) D_i, D_{i+1}, the rational cubic Hermite interpolant (9) converges uniformly to the linear interpolant (10) on $[t_i, t_{i+1}]$ as $r_i \to \infty$ *i.e.*

(12) $$\lim_{r_i\to\infty} ||e_i|| = \lim_{r_i\to\infty} ||P_i - l_i|| = 0,$$

($||\cdot||$ denotes the uniform norm). Moreover, the component functions of e_i tend to zero monotonically, both uniformly and pointwise on $[t_i, t_{i+1}]$.

Remark 4

The interval tension property can also be observed from the behaviour of the control points V_i, W_i defined by (4), and hence of the Bernstein-Bezier convex hull, as $r_i \to \infty$.

3. GC^2 RATIONAL CUBIC INTERPOLATION

Now we construct a GC^2 rational spline interpolant. This requires knowledge of the second derivative which, after some simplification, is given by

(13a) $$P_i^{(2)}(t;r_i) = \frac{2\{\alpha_i(1-\theta)^3 + \beta_i\theta(1-\theta)^2 + \gamma_i\theta^2(1-\theta) + \delta_i\theta^3\}}{h_i\{1 + (r_i-3)\theta(1-\theta)\}^3}$$

where

(13b) $$\begin{cases} \alpha_i = r_i(\Delta_i - D_i) - D_{i+1} + D_i, \\ \beta_i = 3(\Delta_i - D_i), \\ \gamma_i = 3(D_{i+1} - \Delta_i), \\ \delta_i = r_i(D_{i+1} - \Delta_i) - D_{i+1} + D_i. \end{cases}$$

We now follow the familiar procedure of allowing the derivative parameters D_i, $i=0,\ldots,n$ to be degrees of freedom which are constrained by the imposition of GC^2 continuity conditions

$$(14)\quad \begin{bmatrix} P(t_{i+}) \\ P^{(1)}(t_{i+}) \\ P^{(2)}(t_{i+}) \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & \nu_i & 1 \end{bmatrix} = \begin{bmatrix} P(t_{i-}) \\ P^{(1)}(t_{i-}) \\ P^{(2)}(t_{i-}) \end{bmatrix}$$

These GC^2 conditions give, from (13a) and (13b), the linear system of *consistency equations*

$$(15)\quad h_i D_{i-1} + \{h_i h_{i-1} \nu_i/2 + h_i(r_{i-1}-1) + h_{i-1}(r_i-1)\}\, D_i + h_{i-1} D_{i+1}$$
$$= h_i r_{i-1}\, \Delta_{i-1} + h_{i-1}\, r_i \Delta_i,\ i = 1,\ldots, n-1.$$

With appropriate end conditions D_0 and D_n, (15) is a tridiagonal linear system in the unknowns D_i, $i=1,\ldots,n-1$. Assume that

$$(16)\quad r_i \geq r > 2,\ \ \nu_i \geq 0,$$

then the tridiagonal linear system is strictly diagonally dominant and hence has a unique solution which can be easily calculated by use of the tridiagonal *LU* decomposition algorithm. Thus a geometric rational cubic spline interpolant can be constructed with tension parameters r_i, $i=0,\ldots, n-1$, and ν_i, $i=1,\ldots,n-1$, where the special cases are such that

A. the case $r_i = 3$, $\nu_i = 0$, $\forall i$ corresponds to cubic spline interpolation.

B. the case $\nu_i = 0$, $\forall i$ corresponds to the rational cubic spline with tension [Gregory and Sarfraz'90].

C. the case $r_i = 3$, $\forall i$ corresponds to Nielson's ν-spline [Nielson'74].

4. SHAPE CONTROL

We now examine the behaviour of the geometric rational spline interpolant with respect to the tension parameters r_i and ν_i in the following propositions.

Proposition 5 (Global tension property)

Let $l \in C^0[t_0, t_n]$ denote the piecewise linear interpolant defined for $t \in [t_i, t_{i+1}]$ by $l(t) = l_i(t)$, see (10). Suppose that r_i are as in (16) and ν_i are assumed fixed. Then the rational spline interpolant converges uniformly to l as $r \to \infty$, *i.e.* on $[t_0, t_n]$

(17) $\lim_{r\to\infty} ||P - l|| = 0.$

Proof

Suppose $r_i = r$, $i = 0,\ldots, n-1$. Then from (15) it follows that

(18) $$\lim_{r\to\infty} D_i = \frac{(h_i\Delta_{i-1} + h_{i-1}\Delta_i)}{(h_i + h_{i-1})}, \quad i = 1,\ldots,n-1.$$

More generally, for r_i satisfying (16), it can be shown that

(19) $$\max_{1\le i\le n-1} ||D_i||_\infty \le \max\{||\Delta||_\infty r/((r-2)+h\nu_i/2), ||D_0||, ||D_n||\}$$

where

(20) $$||\Delta||_\infty = \max_{0\le i\le n-1} ||\Delta_i||_\infty \text{ and } h = \max_{0\le i\le n-1} h_i.$$

Hence the solution D_i, $i = 1,\ldots,n-1$, of the consistency equations (15) is bounded with respect to r. Now, from (11), the tension property (12) of Proposition 3 can clearly be extended to the case of bounded D_i. Thus applying (12) on each interval gives the desired result (17).□

Proposition 6 (Local tension property)

Let $\nu_i \ge 0$, and $r_i \ge r > 2$, $\forall$ i and consider an interval $[t_k, t_{k+1}]$ for a fixed $k\in\{0,\ldots,n-1\}$. Then, on $[t_k, t_{k+1}]$, the rational spline interpolant converges uniformly to the line segment l_k as $r_k\to\infty$ and ν_k is kept fixed, *i.e.*

(21) $\lim_{r_k\to\infty} ||P_k - l_k|| = 0.$

Proof

The boundedness property (19) holds as in Proposition 5 (where we can assume the additional constraint $r_k \ge r > 2$ to the hypotheses currently being imposed). Thus (12) applies for the case $i=k$.□

Remarks 7

We note that

(*i*) the tension properties in Propositions 5 and 6 can also be applied in the scalar case. We apply this to the curve segment $(t, P_i(t;r_i)) \in \mathbb{R}^2$, $t \in (t_i, t_{i+1})$, with control points

(22) $\{(t_i, X_i), (t_i + h_i/r_i, V_i), (t_{i+1} - h_i/r_i, W_i), (t_{i+1}, X_{i+1})\}$

This is a consequence of the identity

(23) $t \equiv R_0(\theta;r_i)t_i + R_1(\theta;r_i)(t_i + h_i/r_i) + R_2(\theta;r_i)(t_{i+1} - h_i/r_i) + R_3(\theta;r_i)t_{i+1}$

In fact, $(t, P(t))$ can be considered as an application of the interpolation scheme in $\mathbb{R}^2$ to the values $(t_i, X_i) \in \mathbb{R}^2$ and derivatives $(1, D_i) \in \mathbb{R}^2$, $i = 0, ..., n$.

(*ii*) increasing r_i tightens the curve both locally and globally (*c.f.* Propositions 5 and 6). For the range $2 < r_i < 3$ the rational spline produces a more flexible, *i.e. looser*, curve than the cubic spline curve, both locally and globally.

Now we look at the effects of the shape parameters ν_i and consider the curve as parametric one. It can be noted that

(T1) (Point tension) for fixed $i = k$ if we assume $\nu_k \to \infty$ and keep r_i, $i = k-1$, k fixed, then the kth equation of the system of equations (15) results as:

(24) $\lim\limits_{\nu_k \to \infty} D_k = 0.$

Thus the curve at the point P_k will appear to have a *corner*.

(T2) (Interval tension) Similarly as above large values of ν_k and ν_{k+1} (where r_i, $i = k-1$, k, $k+1$ are regarded as fixed) cause D_k and D_{k+1} to approach zero. This behaviour tightens the curve in the interval $[t_k, t_{k+1}]$.

Remark 8

The above interval tension property (T2) does not hold for the first and last intervals unless we assume the natural end conditions i.e. $D_i=0$, $i=0$, n. The spline curve can be globally tightened if the interval tension property is applied, in this fashion, in each of the interval.

5. EXAMPLES

The tension behaviour of the rational cubic spline interpolants is illustrated by the following simple examples for data sets in $\mathbb{R}^2$. Unless otherwise stated, in all the Figures the parameters v_i will be assumed as zero $\forall$ i and the parameters r_i as 3 for all i.

The Figure 1 shows the effect of a progressive increase in global tension with r = 3 (the cubic spline case), 5 and 50. The effect of the high tension parameter is clearly seen in that the resulting interpolant approaches piecewise linear form. Figure 2 illustrates the effect of progressively increasing the value of the tension parameter as r_4 = 3, 5 and 50 in one interval, whilst elsewhere the tension parameters are fixed equivalent to 3. Figures 3 demonstrate the result of Remark 7(*ii*) regarding the achievement of a *looser* curve than a cubic spline curve; the second curve of the figure is a cubic spline curve whereas the first and the last curves show the local and global behaviour against the value 2.1 of the corresponding shape parameters.

The Figure 4 illustrates the effect of progressivley increaisng the value of the point tension parameter ν_4 at the knot t_4 whilst the Figure 5 shows the interval tension effect due to progressive increases in ν_4 and ν_5.

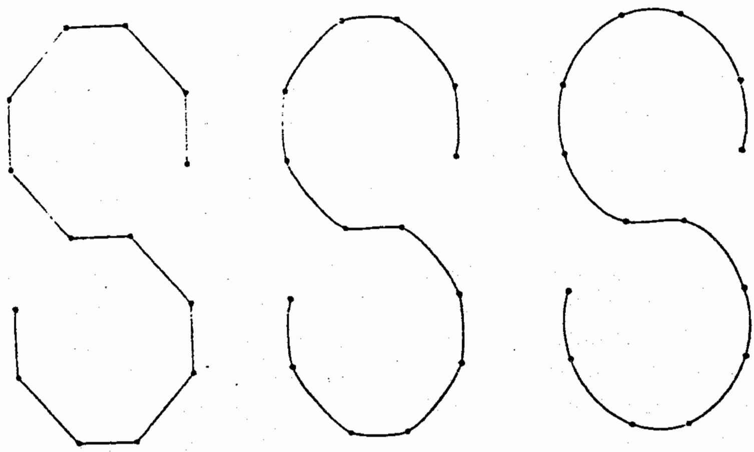


Figure 1. *Interpolatory rational splines with global tension* $r_i = r$.

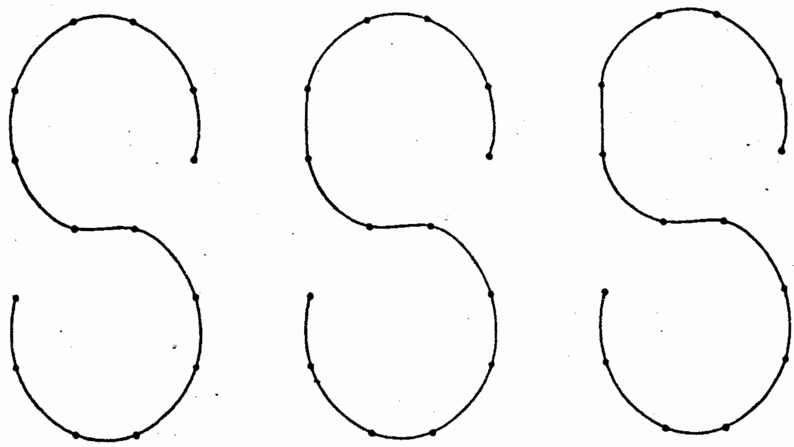


Figure 2. *Interpolatory rational splines with tension r_4 varying.*

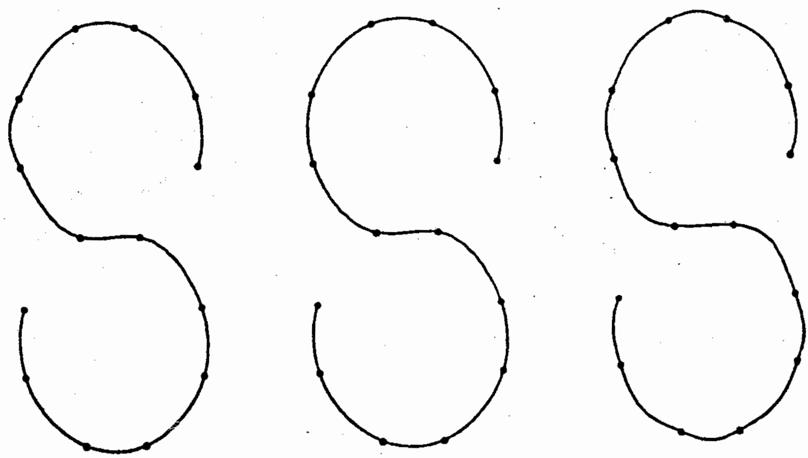


Figure 3. *Interpolatory rational splines can produce looser curves than cubic spline.*

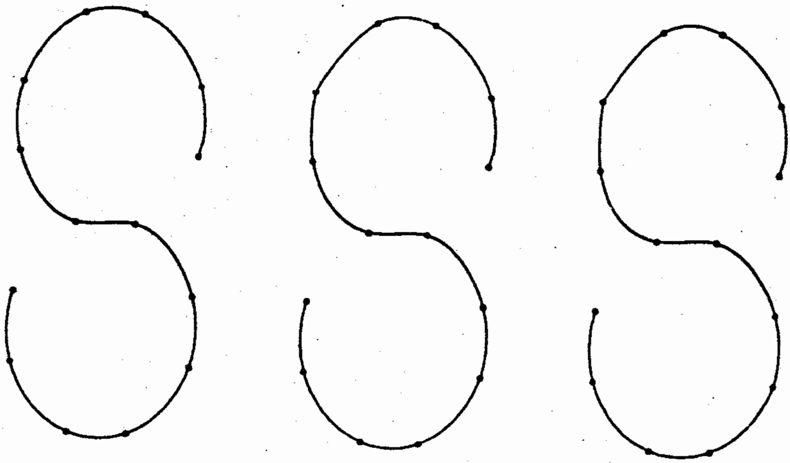


Figure 4. *Interpolatory rational splines with* ν_4 *varying for point tension.*

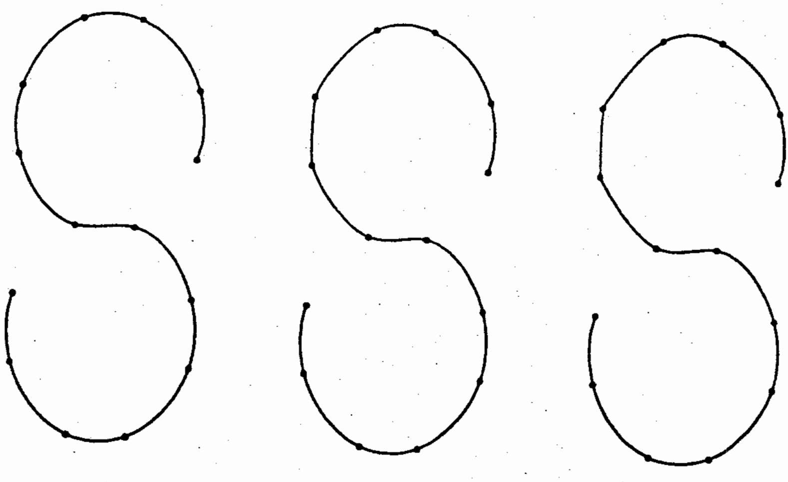


Figure 5. *Interpolatory rational splines with* ν_4 *and* ν_5 *varying for interval tension control.*

6. CONCLUDING REMARKS

An analysis of a GC^2 interpolatory rational cubic spline is developed with a view to its application in CAGD. It is quite reasonable to construct a rational form which involves two shape parameters per interval and provides a variety of local and global controls like interval and point shape effects. In particular, it has been found that only one shape parameter per interval is enough when local or global interval tension is required. The rational spline method can be applied to tensor product surfaces but unfortunately, in the context of interactive surface design, this tensor product surface is not that useful because any one of the tension parameters controls an entire corresponding interval strip of the surface. Thus as an application of GC^2 rational spline for the surfaces, Nielson's [Nielson'86] spline blended methods can be adopted. This will produce local shape control which is quite useful regarding computer graphics.

REFERENCES

1. Barsky, B.A. (1984), Exponential and polynomial methods for applying tension to an interpolatng spline curves, *Comput. Vision Graph. and Image Process*. 27, 1–18.

2. Boehm, W. (1985), Curvature continuous curves and surfaces, *Computer Aided Geometric Design*. 2(2), 313–323.

3. Cline, A. (1974), Curve fitting in one and two dimensions using splines under tension, *Comm. ACM* 17, 218–223.

4. Foley, T.A. (1987), Interpolation with interval and point tension controls using cubic weighted υ-splines, *ACM Trans. Math. Softw*. 13, 68–96..

5. Goodman, T.N.T. (1989), Shape preserving representations, in Lyche, T. and Schumaker, L., eds., *Mathematical Methods in Computer Aided Geometric Design*, Academic press.

6. Gordon, W.J. (1971), Blending function methods of bivariate and multivariate interpolation and approximation, *SIAM J. Num. Anal.* 8, 158–177.

7. Gregory, J.A. and Srafraz, M. (1990), A rational spline with tension, *Computer Aided Geometric Design*, 7, 1–13.

8. Nielson, G.M. (1974), Some piecewise polynomial alternatives to splines under tension, in Barnhill, R.F., eds., *Computer Aided Geometric Design*, Academic press, New York.

9. Preuss, S. (1976), Properties of splines in tension, *J. Approx. Thoery* 17, 86–96.

10. Schweikert, D. (1966), An interpolation curve using splines in tension, *J. Math. and Phys.* 45, 312–317.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 41 – 48

A GEOMETRIC CHARACTERISATION OF PARAMETRIC RATIONAL QUADRATIC CURVES

Muhammad Sarfraz

Department of Mathematics,
University of the Punjab,
Quaid-i-Azam Campus,
Lahore-54590, Pakistan.

ABSTRACT

A B-spline like basis has been constructed for rational quadratic splines. The design curve, thus formulated, produces a freeform curve with point tension control (both locally and globally) due to the presence of shape parameters in its description.

Keywords. B-spline, Bernstein-Bezier, freeform, tension, shape control.

1. INTRODUCTION

One of the common problem in CAGD (Computer Aided Geometric Design) is the designing of curves. Many people have worked in this area in the last couple of years. For example Boehm [1] used curvature continuous cubic splines, Dierecks and Tytgat [2] utilized beta splines, Nielson [3] has made a presentation of Nu-splines, Foley gave the description of weighted Nu-splines [4], Sarfraz [5] has made use of rational splines etc. These spline methods have the capability to design different shapes and then make changes in a different way according to the provision in the method. The above mentioned methods make use of cubic or rational cubic splines for their construction.

This paper presents a method, for the designing of curves, using the rational quadratic B-splines. This generation of freeform curves has parameters to control the shape. We first construct the freeform rational quadratic spline curve in the following section. Effects of the shape parameters are analysed and demonstrated by illustrative examples.

2. CONSTRUCTION OF FREEFORM RATIONAL QUADRATIC SPLINE CURVE

This section provides the structure of the rational quadratic spline (B-spline representation) with shape control.

2.1. Local Support Basis

For the purpose of the analysis, let additional knots be introduced outside the interval $[t_0, t_n]$, defined by $t_{-1}<t_0$ and $t_n<t_{n+1}<t_{n+2}<t_{n+3}$. Let $r_i \geq r > 0$, $i=0,\ldots,n+2$, be shape parameters defined on the extended partition $t_{-1}<\cdots<t_{n+2}<t_{n+3}$. Let $h_i=t_{i+1}-t_i$ and $\theta=(t-t_i)/h_i$. Rational quadratic spline functions ϕ_j, $j = 0,\ldots,n+2$, are constructed such that

$$\phi_j(t) = \begin{cases} 0 & \text{for} \quad t < t_{j-1}, \\ 1 & \text{for} \quad t \geq t_{j+1}. \end{cases} \tag{2.1}$$

On the two intervals $[t_i, t_{i+1}]$, $i = j-1, j$, ϕ_j will have the rational quadratic form

$$\phi_j(t) = R_0(\theta; r_i)\, \phi_j(t_i) + R_1(\theta; r_i)\, \hat{V}_{j,i} + R_2(\theta; r_i)\, \phi_j(t_{i+1}), \tag{2.2}$$

where $R_k(\theta; r_i)$, $k = 0, 1, 2$ are defined as:

$$R_0(\theta; r_i) = (1-\theta)^2/Q_0(0; r_i),$$

$$R_1(\theta; r_i) = r_i\theta(1-\theta)/Q_0(\theta; r_i),$$

$$R_2(\theta; r_i) = \theta^2/Q_0(\theta; r_i),$$

where $Q_0(\theta; r_i) = (1-\theta)^2 + r_i\theta(1-\theta) + \theta^2$.

The functions $R_k(\theta; r_i)$, $k=0, 1, 2$ are actually the Bernstein-Bézier weight functions.

The requirement that $\phi_j \in C^1(-\infty,\infty)$ (in particular at t_{j-1}, t_j and t_{j+1}) uniquely determines the following

$$\hat{V}_{j,i} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases} \tag{2.3a}$$

$$\phi_j(t_j) = \mu_j, \text{ where} \tag{2.3b}$$

$$\mu_j = h_{j-1}r_j/(h_{j-1}r_j + h_j r_{j-1}), \tag{2.4}$$

The local support rational quadratic B-spline basis is now defined by the difference functions

(2.5) $B_j(t) = \phi_j(t) - \phi_{j+1}(t),\ j = 0, ..., n + 1.$

Thus, there immediately follows:

Proposition 2.1 (Rational B-spline)

The rational spline functions $B_j(t),\ j = 0, ..., n + 1$, are such that

(2.6) (Local support) $B_j(t) = 0$, for $t \in (t_{j-1}, t_{j+2})$,

(2.7) (Partition of unity) $\sum_{j=0}^{n+1} B_j(t) = 1$ for $t \in [t_0, t_n]$.

An explicit representation of the rational quadratic B-spline B_j on any interval $[t_i, t_{i+1}]$ can be calculated from (2.3)-(2.5) as

(2.8) $B_j(t) = R_0(\theta; r_i)\, B_j(t_i) + R_1(\theta; r_i)\, V_{j,i} + R_2(\theta; r_i)\, B_j(t_{i+1})$,

where

(2.9) $B_j(t_i) = V_{j,i} = 0$, for $i \neq j, j + 1$, and

(2.10) $\begin{cases} B_j(t_j) = \mu_j, & V_{j,j} = 1, \\ B_j(t_{j+1}) = 1 - \mu_j, & V_{j,j+1} = 0. \end{cases}$

Careful examination of the Bernstein-Bézier vertices of $B_j(t)$ in (2.8) shows these to be non-negative for $r_i > 0$ and we thus have:

Proposition 2.2

The rational B-spline functions are such that

(2.11) (Positivity) $B_j(t) \geq 0$, for all t.

2.2 Design Curve

To apply the rational quadratic B-spline as a practical method for curve design, a convenient method for computing the curve representation

(2.12) $p(t) = \sum_{j=0}^{n+1} P_j\, B_j(t),\ t \in [t_0, t_n]$,

is required, where $P_j \in \mathbb{R}^2$ define the control points of the representation. Now, by the local support property,

$$(2.13)\quad p(t) = \sum_{j=i}^{i+2} P_j\, B_j(t),\ t \in [t_i, t_{i+1}],\ i = 0, \ldots, n-1.$$

Substitution of (2.8) then gives the piecewise defined rational Bernstein-Bézier representation

$$(2.14)\quad p(t) = R_0(\theta; r_i)\, F_i + R_1(\theta; r_i)\, V_i + R_2(\theta; r_i)\, F_{i+1}, \text{ where}$$

$$(2.15)\quad \begin{cases} F_i = (1 - \mu_i)\, P_{i-1} + \mu_i P_i, \\ V_i = P_i. \end{cases}$$

Let $X_i = [F_i\ V_i\ F_{i+1}]^T,\ Z_i = [P_{i-1}\ P_i\ P_{i+1}]^T$

and $$Y_i = \begin{bmatrix} 1-\mu_i & \mu_i & \\ & 1 & \\ & 1-\mu_{i+1} & \mu_{i+1} \end{bmatrix}$$

then the transformation (2.14) can also be represented in matrix notation as

$$(2.16)\quad X_i = Y_i\, Z_i.$$

The transformation to rational Bernstein-Bézier form is very convenient for computational purposes and also leads to the *Variation Diminishing* property [7]:

2.3 Shape Properties

The shape properties of the rational B-spline representation are examined in the following propositions.

Proposition 2.3 (Point tension property)

Let r_i be as assumed in Subsection 2.1 and $r_k \to \infty$ for some k, $1 \le k \le n$. Then following holds

$$(2.17)\quad \lim_{r_k \to \infty} p(t_k) = P_k$$

Proof

From (2.8) and (2.13),

$$p(t_k) - P_k = \sum_{j=0}^{n+1} (P_j - P_k)\, B_j(t_k)$$

$$= (P_{k+1} - P_k)\, B_k(t_k).$$

(by local support property)

$$= (P_{k+1} - P_k)\mu_k.$$

It can be simply shown that

$$\lim_{r_k \to \infty} \mu_k = 0,$$

and thus, (2.17) follows straightaway.□

Remark 2.4

Proposition 2.3 shows that if $r_k \to \infty$, then the part of the design curve is pulled towards the control point P_k. This could be proved directly by studying the behaviour of the Bernstein-Bézier control points in (2.15). This approach can be followed to prove the following:

Corollary 2.5 (Global tension property)

Let $r_i \geq r > 0$, $i = -1, \ldots, n+2$, and let P denote the rational B-spline control polygon, defined explicity on $[t_i, t_{i+1}]$, $i = 0, \ldots, n$ by

(2.8) $P(t) = (1 - \theta) P_i + \theta P_{i+1}, \theta(t) = (t - t_i)/h_i.$

Then the rational B-spline representation (2.13) converges uniformly to P on $[t_0, t_{n+1}]$ as $r \to \infty$.

3. EXAMPLES

For illustration of the tension results, consider a data set in $\mathbb{R}^2$ which define the control points of the rational B-spline representation. The curve in Figure A1 is the quadratic B-spline whereas the curves in Figures A2 and A3 demonstrate the shape effects at one and two adjacent data points respectively. Global shape effect is shown in Figure A4. It should be noted that, in all these figures, where ever the shape effect is achieved it is done by taking shape parameters equivalent to 50 and otherwise they are considered equivalent to 2 everywhere.

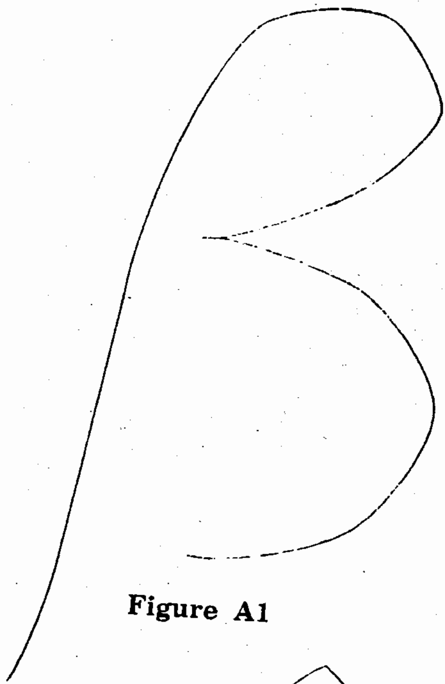


Figure A1

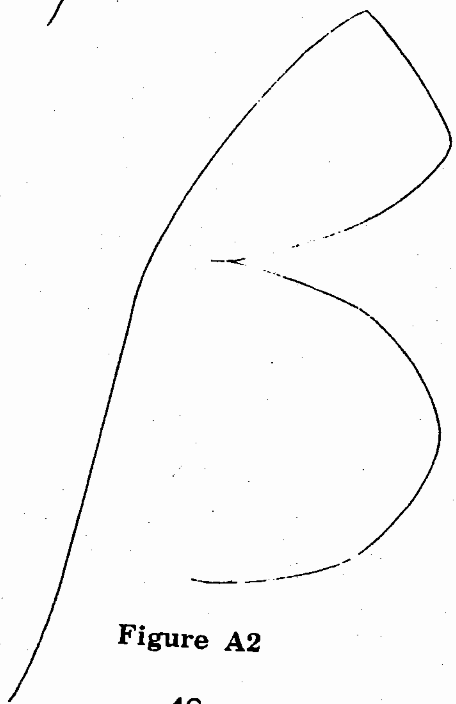


Figure A2

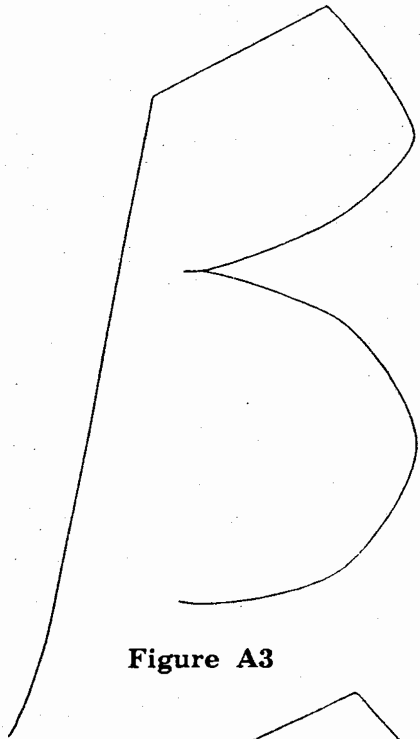


Figure A3

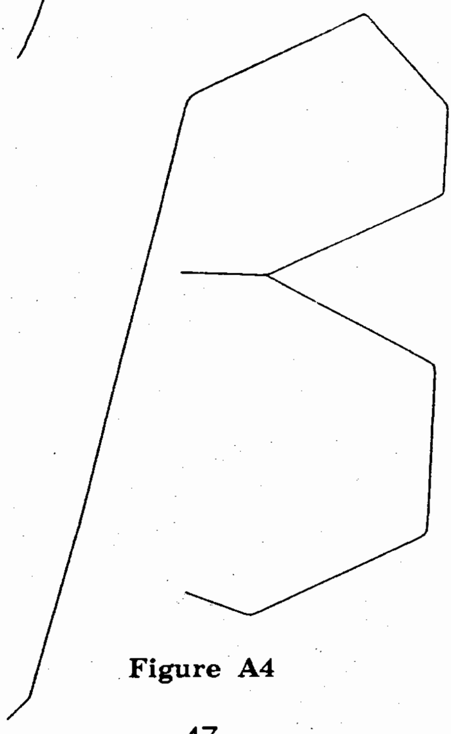


Figure A4

4. CONCLUDING REMARKS

A local B-spline kind of basis has been constructed to generate a freeform C^1 rational quadratic curve. The geometry of the curve has the provision in its description for the alteration of the shape of the curve. This alteration is due to the presence of the parameters in the rational functions. Alterations in the shape of the curve can be made any where according to the desire of the user. Being a rational quadratic scheme, it is computationally economical. Moreover computation of the curve has been suggested through the Bernstein-Bézier representation which is quite convenient for computational purposes.

REFERENCES

1. Boehm, W. (1985), Curvature continuous curves and surfaces, *Computer Aided Geometric Design.* 2(2), 313–323.

2. Dierckx, P. and Tytgat, B. (1989), Generating the Bézier points of β-spline curve, *Computer Aided Geometric Design,* 6, 279–291.

3. Neilson, G.M. (1986), Rectangular ν-splines, *IEEE Computer Graphics and Applications,* 6, 35–40.

4. Foley, T.A. and Ely, H.S. (1989), Surface interpolation with tension controls using cardinal bases, *Computer Aided Geometric Design,* 6, 97–109.

5. Sarfraz, M. (1992), Interpolatory rational cubic spline with biased, point and interval tension, *Comput. and Graphics,* 16, 427–430.

6. Farin, G.E. (1988), *Curves and surfaces for computer aided geometric design,* Academic press, New York.

7. Goodman, T.N.T. (1989), Shape preserving representations, in Lyche, T. and Schumaker, L., eds., *Mathematical Methods in Computer Aided Geometric Design,* Academic Press.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 49 – 53

COINCIDENCE POINTS OF MULTIVALUED MAPPINGS

Ismat Beg
Department of Mathematics,
Quaid-i-Azam University,
Islamabad, Pakistan.

and **Akbar Azam**
Department of Mathematics,
F.G. Post-Graduate College,
Islamabad, Pakistan.

ABSTRACT

A coincidence theorem in a metric space is proved for a multivalued mapping that commutes with two single valued mappings and satisfies a general multivalued contraction type condition.

1991 Mathematics subject classification: 47H10, 54H25, 47H04.

Kewords and Phrases: Coincidence point, fixed point, commuting mappings, multivalued contraction.

INTRODUCTION

Jungck [2] generalized the Banach contraction principle by introducing a contraction condition for a pair of commuting self mappings on a metric space. He also pointed out in [3] and [4] the potential of commuting mappings for generalized fixed point theorems. Subsequently a variety of extensions, generalizations and applications of this followed; e.g., see [5], [7] and [8]. This paper is a continuation of these investigations.

Let (X, d) be a metric space. We shall use the following notation and definitions.

CB(X) = {A: A is a nonempty bounded closed subset of X},
$N(\varepsilon, A) = \{x \in X: (\exists\, a \in A)\ (d(x, a) < \varepsilon)\}$, where $\varepsilon > 0$.

$$H(A,B) = \begin{cases} \inf \{\varepsilon > 0: A \subseteq N(\varepsilon, B),\ B \subseteq N(\varepsilon, A)\} & \text{if the infimum exists,} \\ \infty & \text{otherwise} \end{cases}$$

This function H is a metric for CB(X) and is called Hausdorff metric.

Let $T : X \to CB(X)$ be a mapping, then $C_T = \{f : X \to X : Tx \subseteq fX$ and $(\forall x \in X)(fTx \subseteq Tfx)\}$. T and f are said to be commuting mappings if for each $x \in X$, $f(Tx) = fTx \subseteq Tfx = T(fx)$. A point x is said to be a fixed point of a single valued mapping f (multivalued mapping T) provided $fx=x (x \in Tx)$. The point x is called a coincidence point of f and T if $fx \in Tx$. For details see Nadler [6] and Rhoades, Singh and Chitra [8].

MAIN RESULTS

Lemma 1

Let X be a metric space and $T : X \to CB(X)$ a continuous mapping. Let $f \in C_T$ and continuous such that f and T have a coincidence point z in X. If $\lim_{n\to\infty} f^n z = t < \infty$, then t is a common fixed point of f and T.

Proof

Obviously, $fz \in Tz$ implies that $f^2 z \in fTz \subseteq Tfz$. Therefore $f^{n+1} \in Tf^n z$. It follows that $t \in Tt$. Moreover $ft = f \lim_{n\to\infty} f^n z = \lim_{n\to\infty} f^{n+1} z = t$. Hence t is a common fixed point of f and T.

Theorem 2

Let X be a metric space and $T : X \to CB(X)$ a continuous mapping. Let $f, g \in C_T$ and continuous such that the following condition is satisfied:

$$H(Tx,Ty) \le Ad(fx, gy) + B\{d(fx, Tx) + d(gy, Ty)\} + C\{d(fx, Ty) + d(gy, Tx)\} + D\{1 + d(fx, gy)\}^{-1} d(fx, Tx)\, d(gy, Ty), \tag{1}$$

for all $x, y \in X$, A, B, C, D ≥ 0 and $0 < \frac{A + B + C}{1 - B - C - D} < 1$. Then there is a common coincidence point of f and T, and g and T.

Proof

Assume that $M = \frac{(A + B + C)}{(1 - B - C - D)}$. Let x_0 be an arbitrary but a fixed element of X. We shall construct two sequences $\{x_n\}$ and $\{y_n\}$ as follow. Let $x_1 \in X$ be such that $y_1 = fx_1 \in Tx_0$. Using the definition of

Hausdorff metric and the fact that $TX \subseteq gX$, we may choose $y_2 = gx_2 \in Tx_1$, such that

$$d(y_1, y_2) = d(fx_1, gx_2) \leq H(Tx_0, Tx_1) + A + B + C.$$

Since $Tx \subseteq fX$, we may choose $x_3 \in X$, such that $y_3 = fx_3 = Tx_2$ and $d(y_2, y_3) = d(gx_2, fx_3) \leq H(Tx_1, Tx_2) + \dfrac{(A + B + C)^2}{(1 - B - C - D)}$.

By induction we produce two sequences of points of X, such that

$$y_{2k+1} = fx_{2k+1} \in Tx_{2k}$$
$$y_{2k+2} = gx_{2k+2} \in Tx_{2k+1}, \qquad (2)$$

where k is any natural number. Furthermore,

$$d(y_{2k+1}, y_{2k+2}) = d(fx_{2k+1}, gx_{2k+2}) \leq H(Tx_{2k}, Tx_{2k+1}) + \frac{(A + B + C)^{2k+1}}{(1 - B - C - D)^{2k}}$$

and $$d(y_{2k+2}, y_{2k+3}) = d(gx_{2k+2}, fx_{2k+3}) \leq H(Tx_{2k+1}, Tx_{2k+2}) + \frac{(A + B + C)^{2k+2}}{(1 - B - C - D)^{2k+1}}$$

Hence $$d(fx_{2k+1}, gx_{2k+2}) \leq Ad(fx_{2k+1}, gx_{2k}) + B\{d(fx_{2k+1}, Tx_{2k+1}) + d(gx_{2k}, Tx_{2k})\} + C\{d(fx_{2k+1}, Tx_{2k}) + d(gx_{2k}, Tx_{2k+1})\} + D\{1 + d(fx_{2k+1}, gx_{2k})\}^{-1} d(fx_{2k+1}, Tx_{2k+1})\, d(gx_{2k}, Tx_{2k}) + \frac{(A + B + C)^{2k+1}}{(1 - B - C - D)^{2k}}$$
$$\leq (A + B + C)\, d(fx_{2k+1}, gx_{2k}) + (B + C + D)\, d(fx_{2k+1}, gx_{2k+2}) + \frac{(A + B + C)^{2k+1}}{(1 - B - C - D)^{2k}}$$
$$\leq Md(fx_{2k+1}, gx_{2k}) + M^{2k+1}.$$

Similarly,

$$d(gx_{2k}, fx_{2k+1}) \leq Md\,(fx_{2k-1}, gx_{2k}) + M^{2k}.$$

it further implies that

$$d(y_n, y_{n+1}) \leq M^{n-1}\, d(fx_1, gx_2) + (n - 1)\, M^n.$$

For $p \geq 1$ and $m = n + p$, we have

$$
\begin{aligned}
d(y_{n+1}, y_{m+1}) &\leq d(y_{n+1}, y_{n+2}) + d(y_{n+2}, y_{n+3}) + \ldots + \\
&\quad d(y_{n+p}, y_{n+p+1}) \leq \{M^n d(fx_1, gx_2) + nM^{n+1}\} \\
&\quad + \{M^{n+1} d(fx_1, gx_2) + (n+1)M^{n+2}\} + \ldots + \\
&\quad \{M^{n+p-1} d(fx_1, gx_2) + (n+p-1)\, M^{n+p}\} \\
&\leq \sum_{i=n}^{n+p-1} M^i d(fx_1, gx_2) + \sum_{i=n}^{n+p-1} iM^{i+1}
\end{aligned}
$$

It follows that the sequence $\{y_n\}$ is a Cauchy sequence. Hence there exists z in X such that $y_n \to z$. Therefore $fx_{2k+1} \to z$ and $gx_{2k+2} \to z$. The continuity of T implies that $Tfx_{2k+1} \to Tz$ and $Tgx_{2k+2} \to Tz$.

From (2), we have

$$gfx_{2k+1} \in gTx_{2k} \subseteq Tgx_{2k}$$

$$fgx_{2k+2} \in fTx_{2k+1} \subseteq Tfx_{2k+1}.$$

Since f and g are continuous, by letting $k \to \infty$, we obtain

$$gz \in Tz \text{ and } fz \in Tz$$

This completes the proof of the Theorem.

Corollary 3

Let X be a complete metric space and $T : X \to CB(X)$ a continuous mapping. Let $f, g \in C_T$ and continuous such that (1) is satisfied. Moreover, assume that

$$\{fz, gz\} \subset Tz \text{ implies } \lim_{n\to\infty} f^n z = \lim_{n\to\infty} g^n z = t < \infty.$$

Then t is a common fixed point of f, g and T.

Remark 4

Several other results may also be seen to follow as immediate corollaries to Theorem 2. Included among these are the following, Dube and Singh theorem 1 [1], Jungck [2] Kaneko [5] and Nadler theorem 5 [6].

REFERENCES

1. L.S. Dube and S.P. Singh, On multivalued contraction mappings, *BULL. MATH. de la Soc. Sci. Math. de la R.S. de Roumanie*, **14**(62), nr. 3(1970), 307–310.

2. G. Jungck, Commuting mappings and fixed points, *Amer. Math. Monthly*, **83**(1976), 261–263.

3. G. Jungck, Periodic and fixed points, and commuting mappings, *Proc. Amer. Math. Soc.*, **76**(1979), 333–338.

4. G. Jungck, Common fixed point theorem for semigroups of maps of L-spaces, *Math. Japonica*, **26**(1981), 625–631.

5. H. Kaneko, Single valued and multivalued f-contraction, *Bull. U.M.I.*, **4A**(1985), 29–33.

6. S.B. Nadler, Jr., Multivalued contraction mappings, *Pacific J. Math.*, **30**(1969), 475–488.

7. B.E. Rhoades, S. Sessa, M.S. Khan and M. Swaleh, On fixed points of asymptotically regular mappings, *J. Austral. Math. Soc.*, (Series A) **43**(1987), 328–346.

8. B.E. Rhoades, S.L. Singh and Kulshrestha Chitra, Coincidence theorem for some multivalued mappings, *Internat. J. Math. and Math. Sci.*, **7(3)** (1984), 429–434.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 54 – 62

SHARP ERROR BOUNDS FOR THE SECANT METHOD UNDER WEAK ASSUMPTIONS

Ioannis K. Argyros

Department of Mathematics,
Cameron University, Lawton,
OK 73505, U.S.A.

ABSTRACT

Using the Secant method we approximate a solution of a nonlinear operator equation in a Banach space. The nonlinear operator involved must satisfy a weak smoothness assumption. Our convergence results generalize and improve existing one's.

Key Words and Phrases: Banach space, Nonlinear operator, regularly continuous operator.

(1980) A.M.S. Classification Codes: 47D15, 47H17, 65J15, 65B05.

1. INTRODUCTION

Consider the equation

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

where F is a nonlinear operator mapping a subset E_3 of a Banach space E_1 into another Banach space E_2.

In this paper we are concerned with approximating a solution x^* of equation (1) using the Secant iterations

$$x_{n+1} = x_n - \delta F(x_n, x_{n-1})^{-1} F(x_n),\ n \geq 0 \tag{2}$$

where x_{-1} and $x_0 \in E_3$, and δF is a consistent approximation of the Frechet-derivative F' of F. Denote by N the class of all continuous nondecreasing functions $w : |R^+ \to |R^+$ with $w(0)=0$. Most authors (see, for example [1], [3], [6]) impose a condition of the form

$$||\delta G(y, z) - G'(x)|| \leq w(||x-y||^p + ||x-z||^p),\ 0 \leq p \leq 1$$

for all $x, y, z \in E_3$. However, this condition does not provide sharp error estimates for the Secant method when $0<p<1$ (see, for example [1], [3], [6]). In the elegant paper by Galperin and Waksman [2] sharp error bounds have been found for Newton's method using the notion of w-regularly continuous operator. Here, using a generalized notion of the above definition we provide sharp error bounds for the Secant method. Our results can be reduced to the ones in [2] for $\delta F = F'$ and are compared favorably with the ones in [1] – [7], for $\delta F = F'$ (or not).

2. CONVERGENCE RESULTS

Given an operator $G : E_3 \times E_3 \to E_2$, we say that G is w-continuous at a point $(y, z) \in E_3 \times E_3$ if the function w belongs to the class

$$K(G,(y,z); E_3{\times}E_3) := \{w{\in}N \mid \forall x{\in}E_3 \; ||G(y,z) - G'(x)|| \leq w\,(max\{||x-y||, ||x-z||\})$$

and that G is w-continuous on $E_3{\times}E_3$ if w belongs to

$$K(G, E_3{\times}E_3) := \{w{\in}N \mid \forall (y,z), (x,x){\in}E_3{\times}E_3 \; ||G(y,z) - G'(x)|| \leq w\,(max\{||x-y||, ||x-z||\})$$

All functions of the first set are called here local continuity moduli of G (at (y, z)), whereas those of the second set are called (global) continuity moduli of G (on $E_3 \times E_3$).

Let N^* denote the subclass of N consisting of all $w{\in}N$ that are concave. For an operator $G : E_3 \times E_3 \to E_2$, denote

$$H(x, y, z) = min\ \{||G'(x)||, ||G(y, z)||\}\ x, y, z \in E_3.$$

Given $w{\in}N^*$, we say that G is w-regularly continuous on $E_3{\times}E_3$, if, for all $x, y, z \in E_3$, the inequlaity

$$w^{-1}(H(x, y, z) + ||G(y, z) - G'(x)||) - w^{-1}(H(x, y, z)) \leq max\{||y-x||, ||z-x||\}. \tag{3}$$

Here $w^{-1}(s)$ stands for the least root of the equation $w(t)=s$. Clearly, w^{-1} is an increasing convex function defined on $[0, w(\infty))$. Because of w^{-1} convexity, the above inequality implies $w{\in}K(G,E_3{\times}E_3)$. As in [2] we can show that the converse is not true. For x_{-1}, x_0, x, $y \in E_3$ we define the numbers α, β, r, a, b, q by $||x_0-x_{-1}|| \leq \alpha$, $||\delta F(x_0, x_{-1})^{-1} F(x)|| \leq \beta$, $\gamma = ||x - y||$,

$a = w^{-1}(||G'(x)||)$, $b = w^{-1}(||G'(y)||) - r$, $q = a - b$, the functions $q(s, t)$, R^+, A, B, D by $q(s, t) = min\{t, s - t\}$, $R^+ = max\{R, 0\}$,

$$A(a,b,r) = \int_0^r [w(min\{a, (a-q(s,t))^+\}+t) - w(min\{a, (a-q(s,t))^+\})]\, dt,$$

$$B(r) = \frac{D(r) + 2w(r)\, r}{1 - (w(\alpha) + w(r))} + \alpha + \beta, \text{ with}$$

$$D(r) = A(a_0,b_0,r),\ a_0 = a_0(r) = w^{-1}(1 - w(\alpha) - w(r)), \text{ and } b_0 = b_0(r) = w^{-1}(1 - w(\alpha) - w(r)) - r.$$

Finally, define the iteration $\{t_n\}$, $n \geq -1$ by $t_{-1} = 0$, $t_0 = \alpha$, $t_1 = \alpha + \beta$ and for $n \geq 0$.

$$t_{n+2} = t_{n+1} + \frac{D(t_{n+1}-t_n) + w(t_{n+1}-t_n)(t_{n+1}-t_n) + w(max\{t_{n+1}-t_n, t_{n+1}-t_{n-1}\})(t_{n+1}-t_n)}{1-(w(t_0-t_{-1}) + w(max\{t_{n+1}-t_0, t_n-t_0\}))}$$

We can now state the main result:

Theorem

Let $F : E_3 \subseteq E_2 \rightarrow E_1$

Assume:

(i) there exist $x_0, x_{-1} \in E_3$ and positive numbers α, β such that $\delta F(x_0, x_{-1})$ is invertible, and $||x_0 - x_{-1}|| \leq \alpha$, $||\delta F(x_0, x_{-1})^{-1} F(x_0)|| \leq \beta$;

(ii) the number α is such that $1 - w(\alpha) > 0$ and there exists a minimum number $r^* \in (0, w^{-1}(1 - w(\alpha))$ such that

$$B(r) \leq r \text{ for all } r \in (0, r^*] \tag{4}$$

(iii) $U = U(x_1, r^*) = \{x \in E_3 |\ ||x - x_1|| \leq r^*\}\ E_3$.

(iv) The operator $\delta F(x_0, x^{-1})^{-1}\ \delta F : U \times U \rightarrow L(E_1, E_2)$ is w-regularly continuous on $U \times U$ and F is Frechet differentiable on U. Then

(1) the function A does not increase in each of its first arguments and increases in the third one;

(2) the iteration $\{t_n\}$, $n \geq -1$ is increasing and bounded above by r^* with $t^* = \lim_{n \to \infty} t_n \leq r^*$;

(3) the operator $\delta F(u, v)$ is invertible for all $u, v \in U$;

(4) the Secant iterations (2) are well defined, remain in $U(x_1, t^*)$ and converge to a solution x^* of equation (1);

(5) x^* is the unique solution of equation (1) in $U(x_1, r^*)$;

(6) the following estimates are true for all $n \geq -1$

$$||x_{n+1} - x_n|| \leq t_{n+1} - t_n \tag{5}$$

$$||x_n - x^*|| \leq t^* - t_n; \tag{6}$$

$$||x_n - x^*|| \leq \frac{D(||x_{n-1}-x_n||) + w(||x_{n-1}-x_n||)||x_{n-1}-x_n|| + w(max\{||x_{n-1}-x_n||, ||x_{n-1}-x_{n-2}||\})||x_{n-1}-x_n||}{1-(w(||x_0-x_{-1}) + w(max\{||x^*-x_0||, ||x_n-x_0||\}))}$$

and (7)

$$||x_{n+1}-x_n|| \leq ||x^*-x_n|| + \frac{D(||x^*-x_n||) + w(||x^*-x_n||) \, ||x^*-x_n|| + w(max\{||x^*-x_n||, ||x^*-x_{n-1}||\}||x^*-x_n||)}{1-(w(||x_0-x_{-1}||) + w(max\{||x_n-x_0||, ||x_{n-1}-x_0||\}))}. \tag{8}$$

(7) the convergence condition (4) and the estimates (5)–(7) are sharp.

Proof

(1) The proof of this part as identical to the corresponding one in [2, Lemma 2.1] is omitted.

(2) The first three members of the iteration $\{t_n\}$, $n \geq -1$ are such that $t_1 < t_0 < t_1 \leq r^*$. Therefore the denominator of the fraction appearing in the definition of the sequence is positive. That is $t_1 \leq t_2$ (since the numerator is obviously nonnegative). Let us assume $t_k \leq t_{k+1}$, $k=-1, 0, 1, 2, ..., n$. Then by the definition of the sequence $\{t_n\}$, $n \geq -1$ $t_{k+1} \leq t_{k+2}$. That is $t_{n+1} \leq t_{n+2}$ for $n = k+1$. So far we showed that the scalar sequence $\{t_n\}$, $n \geq -1$ is increasing for all $n \geq -1$. We will show that $t_n \leq r^*$ for all $n \geq -1$. For $n = -1, 0, 1$, this is true by hypothesis. For $n=2$, $t_2 \leq r^*$, since $t_2 \leq B(r^*) \leq r^*$. Let us assume that $t_k \leq r^*$, $k=-1,0,1,2,...,n$, then

$$D(t_1-t_0) + (t_2-t_1) + ... + D(t_{k+1}-t_k) \leq D(t_{k+1}-t_k) \leq D(t_{k+1}-t_0) \leq D(t_{k+1}) \leq D(r^*)$$

since the function w is increasing and

$(t_1-t_0) + (t_2-t_1) + ... + (t_k-t_{k-1}) + (t_{k+1}-t_k) = t_{k+1} - t_0$. Moreover $w(\max\{t_{k+1}-t_0, t_k-t_0\}) \leq w(r^*)$, $k=-1, 0, 1, 2, ..., n$. Hence $t_{k+1} \leq B(r^*) \leq r^*$, which completes the induction. Therefore the sequence $\{t_n\}$,

$n \geq -1$ is increasing and bounded above by r^* and as such it converges to some t^* such that $0 < t^* < r^*$.

(3) Let us observe that the linear operator $\delta F(u, v)$ is invertible for all u, v with $\max\{||u - x_0||, ||v - x_0||\} < w^{-1}(1-w(\alpha))$. Indeed $||\delta F(x_0, x_{-1})^{-1}(\delta F(u, v) - \delta F(x_0, x_{-1}))|| \leq ||\delta F(x_0, x_{-1})^{-1}(\delta F(u, v) - F'(x_0))|| + ||\delta F(x_0, x_{-1})^{-1}(F'(x_0) - \delta F(x_0, x_{-1}))|| \leq w(max\{||u - x_0||\}) + w(\alpha) < 1$, so that according to Banach's lemma $\delta F(u, v)$ is invertibel and

$$||\delta F(u,v)^{-1}\delta F(x_0,x_{-1})|| \leq [1-(w(||x_0-x_{-1}||)+w(max\{||u-x_0||, ||v-x_0||\}))]. \quad (10)$$

It now follows that if (2) is well defined for $n=0,1,2,\ldots,k$ and if (5) holds for $n \leq k$, then

$$||x_0 - x_n|| \leq t_n - t_0 \leq t^* - t_0 \text{ for } n \leq k.$$

This shows that (10) is satisfied for $u=x_i$ and $v=x_j$ with $i,j \leq k$. Thus (2) is well defined for $n=k+1$ too. Also from

$$||x_1 - x_k|| \leq t_k - t_1 \leq t_k \leq t^*$$

we obtain $x_k \in U(x_1, t^*)$

(4) –(6) We now choose $x, y, z \in U$ and bound each one of the three norms that appear at the right hand side of the following estimate

$$||T(F(x) - F(y) - F'(y)(x - y)|| \leq ||T(F(x) - F(y) - F'(y)(x - y)|| + ||T(F'(x) - F'(y))(x - y)|| + ||T(F'(x) - \delta F(y, z))(x - y)||,$$

where T denotes $\delta F(x_0, x_{-1})^{-1}$. Using (3) we obtain

$$||T(F(x)-F(y)-F'(y)(x-y))|| = \left|\left|\int_0^1 [TF'(x+t(y-x))-TF'(y)](x-y)dt\right|\right|$$

$$\leq \int_0^1 ||TF'(x + t(y - x)) - TF'(y)|| \, . \, ||x - y|| \, dt$$

$$\leq \int_0^1 [w(w^{-1}(H(y,x+t(y-x)))+t\,||y-x||)-w(w^{-1}(H(y,x+t(y-x))))]\,||x-y||\,dt$$

$$\leq \int_0^r [w(\min\{a,\ (a-q(s,\ t))^+\} + t) - w(\min\{a,\ (a-q(s,\ t)^+\})]\ dt$$

$$= A(a,\ b,\ r) \tag{11}$$

where we've used the result from [2, Lemma 2.1] that

$$w^{-1}(H(y,\ x+t(y-x))) = \min\ \{a,\ (a-q(s,\ t))^+\} \text{ for } x, y \in U,\ 0\leq t\leq \tau\leq r,\ 0\leq \tau\leq r.$$

Moreover

$$||TF'(x) - TF'(y))\ (x-y)|| \leq ||TF'(x) - TF'(y)||\ ||x-y|| \leq w(r)r.$$

Furthermore

$$||T\delta F(y,\ z) - TF'(x))\ (x-y)|| \leq w(\max\{||x-y||,\ ||x-z||\})\ ||x-y||$$
$$\leq w(\max\ \{r,\ ||x-z||\}\ r.$$

Since $||x_{-1}-x_0|| \leq t_0-t_{-1}$ and $||x_1-x_0|| \leq t_1-t_0$, let us assume that $||x_{k+1}-x_k||\leq t_{k+1}-t_k$, $k=-1,0,1,2,\ldots,n$, and apply the above estimates for $x=x_{k+1}$, $y=x_k$, $z=x_{k-1}$, $u=x_{k+1}$, and $v=x_k$.

Then, we obtain

$$||x_{k+2}-x_{k+1}|| = ||\ (\delta F(x_{k+1},x_k)^{-1}\ \delta F(x_0,x_{-1}))\ [\delta F(x_0,x_{-1})^{-1}\ (F(x_{k+1}) - F(x_k) - \delta F(x_k,x_{k-1}))\ (x_{k+1}-x_k))]||$$

$$\leq ||\delta F(x_{k+1},x_k)^{-1}T^{-1}||\ .\ ||T(F(x_{k+1})-F(x_k)-\delta F(x_k,x_{k-1}))(x_{k+1}-x_k)||$$

$$\leq \frac{A(a,b,||x_{k+1}-x_k||)+w(||x_{k+1}-x_k||)||x_{k+1}-x_k||+w(\max\{||x_{k+1}-x_k||,||x_{k+1}-x_{k-1}||\}\ ||x_{k+1}-x_k||)}{1-(w||x_0-x_{-1}||)+w(\max\{||x_{k+1}-x_0||,||x_k-x_0\})} \tag{12}$$

$$\leq \frac{A(a_0,b_0,t_{k+1}-t_k)+w(t_{k+1}-t_k)(t_{k+1}-t_k)+w(\max\{t_{k+1}-t_k,t_{k+1}-t_{k-1}\})\ (t_{j+1}-t_k)}{1-(w(t_0-t_1)+w(\max\{t_{k+1}-t_0,t_k-t_0\}))}$$

$= t_{k+2}-t_{k+1}$. That is the estimate $||x_{n+1}-x_n||\leq t_{n+1}-t_n$ is true for $n=k+1$ too. Hence, $\{x_n\}$, $n\geq -1$ is a Cauchy sequence in a Banach space and as such it converges to a point $x^* \in U$. Note that the numerator of the last inequality is an upper bound for $||TF(x_{k+1})||$ which tends to 0 as $k\to\infty$.

Hence, by continuity $F(x^*)=0$. The estimate (6) now follows easily from (5).

To show uniqueness, let us assume that there exist two solutions x^* and y^* in $U(x_1, r^*)$ and consider the estimate $F(x^*) - F(y^*) = L^*(x^*-y^*)$ with

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

Then as before we can show $||I - TL^*|| < 1$. That is L^* is invertible which shows that $x^*=y^*$.

$$\text{Set } L = \int_0^1 F'(x^* + t(x_n - x^*))\, dt$$

and use the estimate (12) for $k=n-1$, $x_{n+1}=x^*$ to obtain

$$||x_n-x^*|| = ||L^{-1}F(x_n)|| = ||(TL)^{-1}|| \cdot ||TF(x_n)||$$

$$\leq \frac{A(a_0,b_0,||x_{n-1}-x_n||) + w(||x_{n-1}-x_n||)\, ||x_{n-1}-x_n|| + w(max\{||x_{n-1}-x_n||,||x_{n-1}-x_{n-2}||\}\, ||x_{n-1}-x_n||)}{1-(w(||x_0-x_{-1}||) + w(max\{||x^*-x_0)||,||x_n-x_0||\})}$$

Moreover by taking norms in the identity

$$x_{n+1}-x_n=x^*-x_n+(T\delta F(x_n,x_{n-1}))^{-1}[T(F(x^*)-F(x_n)-\delta F(x_n,x_{n-1})(x^*-x_n))]$$

and using (11) we obtain

$$||x_{n+1}-x_n|| \leq ||x^*-x_n|| +$$

$$\frac{A(a_0,b_0,||x^*-x_n||)+w(||x^*-x_n||)\, ||x^*-x_n|| + w(max\{||x^*-x_n||,||x^*-x_{n-1}||\})\, ||x^*-x_n||}{1-(w||x_0-x_{-1}||) + w(max\{||x_n-x_0||,x_{n-1}-x_0||\})}$$

(7) This follows exactly as in part (5) in Theorem 2.1 in [2], which completes the proof of the theorem.

It can easily be seen that if w is linear and the sequence $||x^*-x_n||$ is monotone then (7) and (8) can provide an upper and lower bound on $||x^*-x_n||$ respectively expressed in terms of the rest of the norms.

Let $w(t)=ct^p$ for some $p\in[0,1]$ and define the real functions f, g by $F(r) = (4+3p)\, cr^{1+p} + (1 + p)\, (c\alpha^p-1)\, r - (1+p)\, (\alpha+\beta)\, cr^p + (1+p)\, (\alpha+\beta)\, (1-c\alpha^p)\, \lambda$ and $G(r) = 1 - c\alpha^p - cr^p$.

It can easily be seen that conditions (*ii*), (4) become equivalent to

$$F(r) \le 0 \text{ and } G(r) > 0.$$

Case 1: Let $p = 1$ and define the numbers r_1, r_2, r_3 and Δ by

$$r_1 = \frac{1 + c\beta - \sqrt{\Delta}}{7c}, r_2 = \frac{1 + c\beta + \sqrt{\Delta}}{7c}, r_3 = \frac{1 - c\alpha}{c} \text{ and}$$

$$\Delta = (1 + c\beta)^2 - 14(\alpha + \beta)(1 - c\alpha)c.$$

It is simple calculus to show that if $r \in [r_1, r_3)$ and $\Delta > 0$, then conditions (B) are satisfied and the conclusions of the theorem hold (if the rest of the hypotheses are satisfied).

Case 2: Let $p = \frac{1}{2}$ and define the number s, s_m, r_6 by

$$s = (1 + p)(c\alpha^{p-1} - (\alpha + \beta)c), s_m = \frac{1 + c\beta}{7c} \text{ and } r_6 = \sqrt{\frac{1 - c\sqrt{\alpha}}{c}}$$

Assume that

$$s < 0, s^2 > 4(4 + 3p)(1 + p)(\alpha + \beta)(1 - c\alpha^p)c \text{ and } f(s_m) < 0. \quad (14)$$

Then there exist two positive roots r_4 and r_5 with $r_4 \le r_5$ of the equation $f(t) = 0$. Let $I = (0, r_6)\ [r_4, r_5] \ne \emptyset$.

Claim: Then for all $r \in I$, conditions (13) are satisfied.

Indeed the function

$$h(t) = (4+3p)r^2 + (1+p)[(c\alpha^p - 1) - (\alpha+\beta)c]^r + (1+p)(\alpha+\beta)(1 - c\alpha^p)$$

has a minimum at s_m. The first two assumptions of (14) imply that this equation has two positive roots. Taking into account that $f(s_m) < 0$, since $f(t)$ is continuous, $f(0) > 0$ and $f(t) > 0$ for t sufficiently large we are assured that $f(t)$ has two positive roots r_4 and r_5. That completes the claim.

REFERENCES

1. Argyros, I.K., On Newton's method and nondiscrete mathematical induction, *Bull. Austral. Math. Soc.* V. **38**, (1988), 131–140.

2. Galperin, A. and Waksman, Z., Newton's method under a weak smoothness assumption, (to appear in the *J. Comp. and Appl. Math.*).

3. Miel, G.J., Semilocal analysis of equations with smooth operators, *Intern. J. Math. and Math. Sci.* V.**4**, (1981), 553–563.

4. Ortega, J.M. and Rheinboldt, W.C. Iterative Solution of Nonlinear Equations in Several Variables, Acad. Press, NY, 1970.

5. Potra, F.A. and Ptak, V., Nondiscrete induction and iterative processes, Pitman Publ. London, 1983.

6. Rockne, J., Newton's method under mild differentiability conditions with error analysis. *Numer. Math.* **18**, (1972), 412.

7. Zabrejko, P.P. and Nguen, D.F. The majorant method in the Theory of Newton-Kantorovich approximations and Ptak error estimates. *Numer. Funct. Anal. Optimiz.* V. **9**, (1987), 671–684.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 63 – 74

FIXED POINTS IN VECTOR LATTICES FOR GENERALIZED CONTRACTIONS

Ismat Beg

Department of Mathematics,
Quaid-i-Azam University,
Islamabad, Pakistan

and Faryad Ali

Department of Mathematics,
Lawrence College,
Murree, Pakistan.

ABSTRACT

It is shown that a discontinuous contractive mapping T on a σ-complete vector lattice must be (0)-continuous at a fixed point. The same is true for pair of maps. There are a number of theorems of the form, if the sequence $\{x_n\}$ of Mann iterates of a certain type (0)-converges to a point p, then p is a fixed point of the map T. Further, if (0)$-\lim x_n = p$, then p is a fixed point of T if and only if T is (0)-continuous at p.

1991 Mathematics subject classification. 47H10, 46A40, 47H89.

Key Words and Phrases. Fixed point, vector lattice.

The order theoretic fixed point theory has numerous important applications to the kinetics of chemical reactions, diffussion processes, the theory of nonlinear heat conduction and mathematical biology. The first contractive definition is that of Banach. It requires the mapping to be continuous in the whole space. In 1968 Kannan [3] gave an example of a contractive definition that does not require the continuity of the map. The purpose of the present paper is to study fixed points in vector lattices for contractive mappings. It is shown that these mappings are (0)-continuous at their fixed point. The same is true for pair of mappings. In the sequel we have obtained the analogue of results of Rhoades [5], Naimpally [4] and many others.

For notation and other facts regarding vector lattices, we refer to Cristeseu [1].

Theorem 1 [Voicu (6)]

Let X be a σ-complete vector lattice and T an operator from X into X. If for all $y, z \in X$,

$$|T(y) - T(z)| \leq \alpha|y - z| + \beta\left(|y - T(y)| + |z - T(z)|\right) + \gamma\left(|y - T(z)| + |z - T(y)|\right) \quad (1)$$

where $\alpha, \beta, \gamma \in R_+$ and $\alpha + 2\beta + 2\gamma < 1$. Then for any $x \in X$ the sequence $\{T^n(x)\}$ is (0)-convergent to an element p of X and p is the unique fixed point of T.

For sake of completeness we give outline of the proof of Theorem 1.

Outline of the Proof

Define $x_n = T^n(x)$. In (1) take $y = T^{n+p-1}(x)$ and $z = T^{n-1}(x)$ then

$$|x_{n+p} - x_n| \leq \frac{\lambda^{n-1}}{1-\lambda}|x_2 - x_1|,$$

where $\lambda = \dfrac{\alpha + \beta + \gamma}{1 - \beta - \gamma}$.

It implies that $\{x_n\}$ is a (0)-Cauchy sequence in X. There exists an element p in X such that

$$p = (0) - \lim_{n\to\infty} x_n = (0) - \lim_{n\to\infty} T^n(x).$$

Using inequality (1), we have,

$$\begin{aligned}|p - T(p)| &\leq (1+\gamma)\,|p - x_n| + \alpha|x_{n-1} - p| + \beta\lambda^{n-2}\,|x_2 - x_1| \\ &\quad + \gamma|x_{n-1} - T(p)| + \beta|p - T(p)| \\ &\leq \gamma|p - T(p)| + \beta|p - T(p)|.\end{aligned}$$

It follows that, $|p - T(p)| = 0$. Therefore p is a fixed point of T. Uniqueness of p can be verified easily.

Theorem 2

Let X be a σ-complete vector lattice and $T : X \to X$. If there exists $\alpha, \beta, \gamma \in R_+$ such that $\alpha + 2\beta + 2\gamma < 1$ and for all $y, z \in X$, (1) is satisfied, then T is (0)-continuous at the fixed point p.

Proof

From the Proof of Theroem 1, the sequence $\{x_n\}$, defined by $x_n = T^n(x)$, (0)-converges to the unique fixed point p. Thus $T(x_n) \xrightarrow{(0)} p$.

Let $\{y_n\}$ be any sequence in X, (0)-converging to p. Using (1)

$$\begin{aligned}|T(x_n)-T(y_n)| &\le \alpha|x_n - y_n| + \beta(|x_n - T(x_n)| + |y_n - T(y_n)|) \\ &\quad + \gamma(|x_n - T(y_n)| + |y_n - T(x_n)|) \\ &\le \alpha|x_n - y_n| + \beta(|T(x_n) - x_n| + |T(x_n) - y_n|) \\ &\quad + \gamma(|x_n - T(x_n)| + |T(x_n) - T(y_n)| + |T(x_n) - y_n|) \\ &= \alpha|x_n - y_n| + (\beta+\gamma)\,|T(x_n)-x_n| + (\beta+\gamma)\,|T(x_n)-y_n| \\ &\qquad + (\beta + \gamma)\,|T(x_n) - T(y_n)|.\end{aligned}$$

Therefore,

$$(0)-\lim_{n\to\infty} |T(x_n) - T(y_n)| \le (\beta + \gamma)\ (0) - \lim_{n\to\infty} |T(x_n) - T(y_n)|$$

It further implies,

$$(1 - \beta - \gamma)\ (0) - \lim_{n\to\infty} |T(x_n) - T(y_n)| \le 0.$$

Since $0 < \alpha + \beta + \gamma < 1 - \beta - \gamma$, therefore

$$|T(x_n) - T(y_n)| \xrightarrow{(0)} 0.$$

Hence $|T(y_n)-T(p)| \le |T(y_n)-T(x_n)| + |T(x_n)-p| \xrightarrow{(0)} 0$ and therefore

$$T(y_n) \xrightarrow{(0)} T(p) = p.$$

Hence T is (0)-continuous at p.

Theorem 3

Let X be a σ-complete vector lattice and $T : X \to X$. If there exists $\alpha, \beta, \gamma \in R_+$ such that $\alpha + 2\beta + 2\gamma < 1$ and for all x, $T(y) \in X$.

$$\begin{aligned}|T(x)-T(T(y))| &\le \alpha|x - T(y)| + \beta(|x - T(x)| + |T(y) - T(T(y))|) \\ &\qquad + \gamma(|x - T(T(y))| + |T(y) - T(x)|) \quad (2)\end{aligned}$$

Then T is (0)-continuous at the fixed point p.

Proof

In (1), if we put $y=x$ and $z=T(y)$, we get (2). Hence from Theorem 1, the sequence $\{x_n\}$ defined by $x_n = T^n(x)$ (0)-converges to a unique fixed point p.

Let $\{y_n\}$ be any other sequence in X such that $y_n \xrightarrow{(0)} p$. Using (2)

$$|T(y_n)-T(T(x_n))| \le \alpha|y_n-T(x_n)| + \beta(|y_n-T(y_n)| + |T(x_n)-T(x_n)|)$$
$$+ \gamma(|T(x_n) - T(y_n)| + |y_n - T(T(x_n))|)$$
$$\le \alpha|y_n-T(x_n)| + (\beta+\gamma)|y_n - T(T(x_n))| + (\beta+\gamma)$$
$$|T(T(x_n)) - T(y_n)| + (\beta+\gamma)\ |T(x_n) - T(T(x_n))|.$$

Therefore,

$$(1-\beta-\gamma)\ (0) - \lim_{n\to\infty}\ |T(y_n) - T(T(x_n))| \le 0.$$

Since $(1-\beta-\gamma) > 0$, thus

$$|T(y_n) - T(T(x_n))| \xrightarrow{(0)} 0.$$

Now, $|T(y_n)-T(p)| \le |T(y_n) - T(T(x_n))| + |T(T(x_n)) - p|$.

Letting $n \to \infty$,

$$|T(y_n) - T(p)| \xrightarrow{(0)} 0$$

Hence $T(y_n) \xrightarrow{(0)} T(p) = p$

Thus T is (0)-continuous at p.

Let X be a vector lattice, $T : X \to X$. The Mann iteration is defined by

$$x_0 \in X,\ x_{n+1} = (1-c_n)\,x_n + c_n T(x_n),\ n\ge 0, \qquad (3)$$

where $\{c_n\}$ satisfies $c_0=1$, $0\le c_n\le 1$, $n > 0$, and $\sum c_n = \infty$. We shall be interested in those iterations for which $\{c_n\}$ satisfies

$$c_0 = 1,\ \ 0 < c_n \le 1 \ \text{ for } \ n > 0, \qquad (4)$$

and $\{c_n\}$ is bounded away from zero.

Theorem 4

Let X be a σ-complete vector lattice, $T : X \to X$, satisfying one of the following contractive definitions at each pair of points $x,y\in X$:

$$|T(x) - T(y)| \le \sup\{c|x-y|,\ |x-T(x)| + |y-T(y)|,\ |x-T(y)| + |y-T(x)|\},\ c \ge 0,\ 0 \le k < 1, \qquad (5)$$

$$|T(x) - T(y)| \le \sup\{|x-y|,\ \frac{1}{2}[|x-T(x)| + |y-T(y)|],$$
$$\frac{1}{2}[|x-T(y)| + |y-T(x)|], \qquad (6)$$

At each pair of points x, y, T satisfies at least one of the following:

$$|x - T(x)| + |y - T(y)| \leq \beta[|x - T(y)| + |y - T(x)| + |x - y|,$$

$$\frac{1}{2} \leq \beta < \frac{2}{3}, \qquad (7a)$$

$$|x - T(x)| + |y - T(y)| + |T(x) - T(y)| \leq \gamma[|x - T(y)| +$$

$$|y - T(x)|], 1 \leq \gamma < \frac{2}{3}, \qquad (7b)$$

$$|T(x) - T(y)| \leq \delta \sup\{|x - y|, |x - T(x)|, |y - T(y)|,$$

$$\frac{1}{2}[|x - T(y) + |y-T(x)|]\}, 0 < \delta < 1 \quad (7c)$$

Let $\{x_n\}$ satisfy (3) and (4), and suppose that $x_n \xrightarrow{(0)} p$. Then p is a fixed point of T and T is (0)-continuous at p.

Proof

Since $x_{n+1} = (1 - c_n) x_n + c_n T(x_n)$.

Therefore,

$$(T(x_n) - p) + (p - x_n) = \frac{1}{c_n}(x_{n+1} - x_n)$$

Hence $|T(x_n) - p)| \leq |x_n - p| + \frac{1}{c_n}|x_{n+1} - x_n|$.

Letting $n \to \infty$, and from (4), we have $T(x_n) \xrightarrow{(0)} p$.

Substitute $x = x_n, y = p$ in each of (5)-(7), and then take the (0)-limit as $n \to \infty$, we obtain $T(p) = p$.

Let $\{y_n\}$ be any sequence in X with (0)-limit p.

Substituting $x = x_n, y = y_n$ in each of (5)-(7), and, using arguments similar to those from preceding theorems we obtain $|T(x_n) - T(y_n)| \xrightarrow{(0)} 0$. Therefore,

$$|T(y_n) - T(p)| \leq |T(x_n) - T(y_n)| + |T(x_n) - T(p)| \xrightarrow{(0)} 0$$

Hence T is (0)-continuous at p.

Ishikawa [2] developed the following iteration scheme to obtain a fixed point for a certain class of maps defined over a Hilbert space. Let $x_0 \in X$,

$$\left.\begin{aligned} y_n &= (1-\beta_n)\,x_n + \beta_n T(x_n),\ n \geq 0, \\ x_{n+1} &= (1-\alpha_n)\,x_n + \alpha_n T(y_n),\ n \geq 0, \end{aligned}\right\} \tag{8}$$

where $0 \leq \alpha_n \leq \beta_n \leq 1$, $\lim \beta_n = 0$, $\sum \alpha_n \beta_n = \infty$. In this paper we shall use the conditions

(i) $0 \leq \alpha_n, \beta_n \leq 1$, (ii) $\underline{\lim}\ \alpha_n > 0$, (iii) $\overline{\lim}\ \beta_n < 1$.

Theorem 5

Let X be a σ-complete vector lattice, T an operator defined on X with values in X and satisfying: for each $x, y \in X$,

$$|T(x) - T(y)| \leq k \sup\{|x-y|, |x - T(x)|, |y - T(y)|,$$

$$|x - T(y)| + |y - T(x)|\},\ 0 \leq k < 1. \tag{9}$$

Let $\{x_n\}$ be defined by (8) with $\{\alpha_n\}$, $\{\beta_n\}$ satisfying *(i)* and *(ii)*. If $(x_n\}$ (0)-converges to p, then p is a fixed point of T and T is (0)-continuous at p.

Proof

It follows from (8) that $x_{n+1} - x_n = \alpha_n(T(y_n) - x_n)$. Since $x_n \xrightarrow{(0)} p$ and $(\alpha_n\}$ is bounded away from zero, therefore $|T(y_n) - x_n| \xrightarrow{(0)} 0$. Hence $|p - T(y_n)| \xrightarrow{(0)} 0$.

Since T satisfies (9), we have

$$|T(y_n) - T(x_n) \leq k \sup\{|x_n - y_n|, |x_n - T(x_n)|, |y_n - T(y_n)|,$$

$$|x_n - T(y_n)| + |y_n - T(x_n)|\}, \tag{10}$$

where $|y_n - x_n| = |\beta_n T(x_n) + (1-\beta_n)x_n - x_n| \leq |x_n - T(y_n)| + (T(y_n) - T(x_n)|$,

$$\begin{aligned} |y_n - T(y_n)| &= |\beta_n T(x_n) + (1-\beta_n)\,x_n - T(y_n)| \\ &\leq |x_n - T(y_n)| + |T(x_n) - T(y_n)|, \end{aligned}$$

and

$$\begin{aligned} |y_n - T(x_n)| &= |\beta_n T(x_n) + (1-\beta_n)\,x_n - T(x_n)| \\ &\leq |x_n - T(y_n)| + |T(y_n) - T(x_n)|. \end{aligned}$$

Substituting these values in (10)

$$|T(x_n) - T(y_n)| \le k \sup\{|x_n - T(y_n)| + |T(x_n) - T(y_n)|, |x_n - T(y_n)| + |T(y_n) - T(x_n)|, |x_n - T(y_n)| + |T(x_n) - T(y_n)|, 2|x_n - T(y_n)| + |T(y_n) - T(x_n)|\}$$

$$\le \left(\frac{2k}{1-k}|x_n - T(y_n)|\right) \xrightarrow{(0)} 0.$$

Using triangle inequality, it follows that

$$|x_n - T(x_n)| \xrightarrow{(0)} 0 \text{ and } |p - T(x_n)| \xrightarrow{(0)} 0.$$

Again by using (9), we obtain

$$|T(x_n) - T(p)| \le k \sup\{|x_n - p|, |x_n - T(x_n)|, |p - T(p)|, |p - T(x_n)| + |x_n - T(p)|\}$$

$$\le k \sup\{|x_n - p|, |x_n - T(x_n)|, |p - x_n| + |x_n - T(x_n)| + |T(x_n) - T(p)|, |p - T(x_n)| + |x_n - T(x_n)| + |T(x_n) - T(p)|\}.$$

Letting $n \to \infty$, we obtain

$$|T(x_n) - T(p)| \xrightarrow{(0)} 0.$$

Hence $|p - T(p)| \le |p - T(x_n)| + |T(x_n) - T(p)| \xrightarrow{(0)} 0.$

It implies that p is a fixed point of T.

Let $\{y_n\}$ be any sequence in X such that $y_n \xrightarrow{(0)} p$. Since

$$|T(y_n) - T(p)| \le |T(y_n) - T(x_n)| + |T(x_n) - p| \xrightarrow{(0)} 0.$$

Hence $T(y_n) \xrightarrow{(0)} T(p)$, and so T is (0)-continuous at p.

Theorem 6

Let X be a σ-complete vector lattice, T an operator defined on X with values in X and satisfying:

At each pair of points $x, y \in X$, T satisfies at least one of the following:

$$|x - T(x)| + |y - T(y)| \le b[|x - T(y)| + |y - T(x)| + |x - y|],$$

$$\frac{1}{2} \le b < \frac{3}{2}, \tag{11A}$$

$$|x - T(x)| + |y - T(y)| + |T(x) - T(y)| \le c[|x - T(y)| + |y - T(x)|],$$

$$1 \le c < \frac{3}{2}, \tag{11B}$$

$$|T(x) - T(y)| \le k \sup\{|x - y|, |x - T(x)|, |y - T(y)|, \frac{1}{2}[|x - T(y)| + |y - T(x)|]\}, 0 < k < 1. \tag{11C}$$

Let $\{x_n\}$ be defined by (8) with $\{\alpha_n\}$, $\{\beta_n\}$ satisfying (i)–(iii). If $\{x_n\}$ (0)-converges to p, then p is a fixed point of T and T is (0)-continuous at p.

Proof

From (8), we get $x_{n+1} - x_n = \alpha_n(T(y_n) - x_n)$. Since $x_n \xrightarrow{(0)} p$, $|x_{n+1} - x_n| \xrightarrow{(0)} 0$. As $\underline{\lim}\ \alpha_n > 0$, therefore $|T(y_n) - x_n| \xrightarrow{(0)} 0$.

Thus $|p - T(y_n)| \le |p - x_n| + |x_n - T(y_n)| \xrightarrow{(0)} 0.$

Case I: If (11A) is satisfied, then

$$2|T(x_n) - T(y_n)| \le (1 + b)[|x_n - T(y_n)| + |y_n - T(x_n)|] + b|y_n - x_n|,$$

Since $|y_n - x_n| \le \beta_n [|x_n - T(y_n)| + |(Ty_n) - T(x_n)|]$ (12)

and $|y_n - T(x_n)| \le (1 - \beta_n)[|x_n - T(y_n)| + |T(y_n) - T(x_n)|]$ (13)

Therefore

$$|T(x_n) - T(y_n)| \le \frac{[2(1 + b) - \beta_n]}{[1 - b + \beta_n]} |x_n - T(y_n)|$$

Case II: If (11B) is satisfied, then

$$3|T(x_n) - T(y_n)| \le (c + 1)[|x_n - T(y_n)| + |y_n - T(x_n)|].$$

Using (13), we get

$$|T(x_n) - T(y_n)| \le \frac{(1 + c)(2 - \beta_n)}{[2 - c + (1 + c)\beta_n]} |x_n - T(y_n)|$$

Case III: If (11C) is satisfied, then

$$|T(x_n) - T(y_n)| \le k \sup\{|x_n - y_n|, |x_n - T(x_n)|, |y_n - T(y_n)|, \frac{1}{2}[|x_n - T(y_n)| + |y_n - T(x_n)|]\}$$

Also

$$\frac{1}{2}[|x_n - T(y_n)| + |y_n - T(x_n)|] \le \frac{1}{2}[|x_n - T(y_n)| + (1-\beta_n)\,|x_n - T(x_n)|]$$

$$\le |x_n - T(y_n)| + \frac{1}{2}|T(x_n) - T(y_n)| \tag{14}$$

Using (12), (13) and (14), we have

$$|T(x_n) - T(y_n)| \le \frac{k}{1-k}\;|x_n - T(y_n)|$$

Thus in any case (11A, 11B and 11C).

$$|T(x_n) - T(y_n)| \le \max\left\{\frac{2(1+b)-\beta_n}{1-b+\beta_n}, \frac{(1+c)(2-\beta_n)}{2-c+(1+c)\beta_n}, \frac{k}{1-k}\right\}$$

$$|x_n - T(y_n)|. \tag{15}$$

Since $\overline{\lim}\;\beta_n < 1$, (15) becomes $|T(x_n) - T(y_n)| \xrightarrow{(0)} 0$.

Therefore, $|x_n - T(x_n)| \xrightarrow{(0)} 0$ and $|p - T(x_n)| \xrightarrow{(0)} 0$.

If (11A) is satisfied and $x_n \xrightarrow{(0)} p$, then we have

$$2|T(x_n) - T(y_n)| \le (1+b)\,[|x_n - T(p)| + |p - T(x_n)|] + b\,|x_n - p|$$

$$\le (1+b)\,[|x_n - T(x_n)| + |T(x_n) - T(p)| + |p - T(x_n)|] + b\,|x_n - p|$$

If (11B) is satisfied and $x_n \xrightarrow{(0)} p$, then we have

$$|T(x_n) - T(p)| \le \frac{c+1}{2-c}[|x_n - T(x_n)| + |p - T(x_n)|].$$

If (11C) is satisfied and $x_n \xrightarrow{(0)} p$, then we have

$$|T(x_n) - T(p)| \le k\,\{|x_n - p|, |x_n - T(x_n)|, |p - T(p)|,$$

$$\frac{1}{2}[|x_n - T(p)| + |p - T(x_n)|]\}$$

$$\le k \sup\{|x_n - p|, |x_n - T(x_n)|, |p - x_n| + |x_n - T(x_n)| +$$

$$|T(x_n) - T(p)|, \frac{1}{2}[|x_n - T(x_n)| + |T(x_n) - T(p)| + |p - T(x_n)|]\}$$

$$\le \frac{k}{1-k}[|x_n - p| + |p - T(x_n)| + |x_n - T(x_n)|]$$

Therefore, $|T(x_n) - T(p)| \leq \max\left\{\frac{1+b}{1-b}, \frac{c+1}{2-c}, \frac{k}{1-k}\right\} |x_n - p| +$

$$|p - T(x_n)| + |T(p) - x_n|] + |x_n - T(x_n)|$$

Thus $|p-T(p)| \leq |p - x_n| + |x_n - T(x_n)| + \max\left\{\frac{1+b}{1-b}, \frac{c+1}{2-c}, \frac{k}{1-k}\right\}$

$[|x_n-p| + |p-T(x_n)| + |x_n-T(p)|] + |x_n - T(x_n)|$. It further implies that, $|p - T(p)| \xrightarrow{(0)} 0$.

Hence $T(p) = p$.

We now prove that T is (0)-continuous at p. Let $\{y_n\}$ be any sequence in X with (0)-lim $y_n = p$. Then

$$|T(y_n) - T(p)| \leq |T(y_n) - T(x_n)| + |T(x_n) - T(p)| \xrightarrow{(0)} 0$$

Hence $T(y_n) \xrightarrow{(0)} T(p)$. It further implies that T is (0)-continuous at p.

We now establish some results on Mann iterations for pair of maps.

Theorem 7

Let X be a σ-complete vector lattice, S, T be two operators defined on X with values in X and satisfying:

$$|S(x)-T(y)| \leq k\sup\{c|x-y|, |x-S(x)| + |y-T(y)|, |x-T(y)| + |y-S(x)|\} \quad (16)$$

where $c \geq 0$, $0 \leq k < 1$. If there exists a point x_0 such that the Mann iterates for S or T, defined by (3) and (4), (0)-converges to a point p, then p is a common fixed point of S and T. Moreover, S and T are (0)-continuous at p.

Proof

Suppose the Mann iterates using S, (0)-converges to p. Since $\{c_n\}$ is bounded away from zero. Therefore $S(x_n) \xrightarrow{(0)} p$. Using (16).

$$|S(x_n)-T(p)| \leq k \sup\{c|x_n-p|, |x_n-S(x_n)| + |p-T(p)|, |x_n-T(y)| + |p-S(x_n)|\},$$

Letting $n \to \infty$,

$$|p - T(p)| \leq k\, |p - T(p)|.$$

Since $0 \leq k < 1$, therefore $|p-T(p)| = 0$. Hence p is a fixed point of T. Using (16),

$$|S(p) - p| = |S(p) - T(p)| \leq k|p - S(p)|.$$

It further implies that $S(p)=p$. Hence p is a common fixed point of S and T.

Let $\{y_n\}$ be any sequence with (0)-lim $y_n=p$. Inequality (16) implies,

$$|S(y_n) - S(p)| = |S(y_n) - T(p)|$$

$$\leq k \sup \{c\, |y_n-p|,\, |y_n-S(y_n)| + |p-T(p)|,\, |y_n-T(p)| + |p-S(y_n)|\}$$

$$\leq k \sup \{c\, |y_n-p|,\, |y_n-p| + |T(p)-S(y_n)| + |p-T(p)|,\, |y_n-p| + |p-T(p)| + |T(p)-S(y_n)|\}$$

Letting $n \to \infty$, we obtain

$$|S(y_n) - T(p)| \leq \frac{k}{1-k} |y_n - p| \xrightarrow{(0)} 0.$$

Hence $S(y_n) \xrightarrow{(0)} T(p) = S(p)$.

Thus S is (0)-continuous at p. Similarly, it can be shown that T is (0)-continuous at p.

Proof for Mann iterates of T is similar.

Theorem 8

Let X be a σ-complete vector lattice, S and T be two operators defined on X with values in X satisfying; for each $x,y \in X$,

$$|S(x) - T(y)| \leq k \sup\{c\, |x-y|,\, |x-S(x)| + |y-T(y)|,\, |x-T(y)| + |y-S(x)|\}, \quad (17)$$

where $c \geq 0$, $0 \leq k < 1$. If $x_n \xrightarrow{(0)} p$, where $\{x_n\}$ is defined by

$$\left. \begin{aligned} & x_0 \in X,\ x_{2n+1} = (1-c_{2n})x_{2n} + c_{2n}\, S(x_{2n}), \\ & x_{2n+2} = (1-c_{2n+1})\, x_{2n+1} + c_{2n+1}\, T(x_{2n+1}) \end{aligned} \right\} \quad (18)$$

where $\{c_n\}$ satisfies (4), then S and T have p as a common fixed point, and S and T are (0)-continuous at p.

Proof

Since $\{c_n\}$ is bounded away from zero, and $x_n \xrightarrow{(0)} p$, it follows from (18) that $S(x_{2n}) \xrightarrow{(0)} p$ and $T(x_{2n+1}) \xrightarrow{(0)} p$. From (17),

$$|S(x_{2n})-T(p)| \le k \sup \{c\,|x_{2n}-p|,\ |x_{2n}-S(x_{2n})| + |p-T(p)|,\ |x_{2n}-p| + |p-S(x_{2n}))|\}$$

Letting $n \to \infty$, we obtain $|p - T(p)| \le k|p - T(p)|$. It implies that $T(p)=p$. Similarly, we can prove $S(p) = p$.

Let $\{y_n\}$ be any sequence in X with (0)-lim $y_n=p$. From (17),

$$|S(x_{2n})-T(y_n)| \le k \sup \{c\,|x_{2n} - y_n|,\ |x_{2n} - S(x_{2n})| + |y_n - T(y_n)|,\ |x_{2n}-T(y_n)| + |y_n - S(x_{2n})|\}$$

$$\le k \sup \{c\,|x_{2n} - y_n|,\ |x_{2n} - S(x_{2n})| + |y_n - S(x_{2n})| + |S(x_{2n}) - T(y_n)|,\ |x_{2n} - S(x_{2n})| + |S(x_{2n}) - T(y_n)| + |y_n - S(x_{2n})|\}$$

$$\le \frac{k}{1-k} \sup \{c\,|x_{2n} - y_n|,\ |x_{2n} - S(x_{2n})| + |y_n - S(x_{2n})|\}.$$

It implies taht $|S(x_{2n}) - T(y_n)| \xrightarrow{(0)} 0$. It further implies that $|T(y_n) - T(p)| \xrightarrow{(0)} 0$. Thus T is (0)-continuous at p. Similarly, we can prove that S is (0)-continuous at p.

Acknowledgements

One of the authors (I.B.) would like to thank Professor Abdus Salam, the International Atomic Energy Agency and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste, Italy.

REFERENCES

1. R. Cristesçu, Ordered vector spaces and linear operators, Editura Academiei, Bucuresti, Romania and Abacus Press, Tunbridge Wells, Kent England (1976).
2. S. Ishikawa, Fixed points by a new itertion method, *Proc. Amer. Math. Soc.*, 44(1974), 147–150.
3. R. Kannan, Some results on fixed points, *Bull. Calcutta Math. Soc.*, **60**(1968), 689–694.
4. S. A. Naimpally and K.L. Singh, Extensions of some fixed point theorems of Rhoades, *J. Math. Anal. Appl.*, **96**(1983), 437–446.
5. B.E. Rhoades, Contractive definitions and continuity, *Contemporary Math., Amer. Math. Soc.*, **72**(1988), 233–245.
6. F. Voicu, Teoreme de punct fix in spatii liniare σ-reticulate, *Bul. Stiint. Al Institutului de Constructii Bucuresti*, (Tome 37), Nr. 2 (1987), 87–91.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 75 – 83

A NOTE ON TRIANGULAR NUMBERS

Heakyung Lee & Muhammad Zafrullah
Department of Mathematics,
Winthrop University
Rock Hill, South Carolina 29733, U.S.A.

INTRODUCTION

For $r \in \mathrm{N}$, let t_r denote the r^{th} triangular number. It is well known that $t_r = \frac{r(r+1)}{2}$. In this note we pose and answer the question: If $s \in \mathrm{N}$ is fixed, for what $r \in \mathrm{N}$ is $t_{r+s} - t_r$ a triangular number? If there is an m such that $t_{r+s} - t_r = t_m$, let us call $(r+s, r, m)$ a triangular triplet. This question is of interest in that, on the one hand, the n^{th} triangular number $t_n = \frac{n(n+1)}{2}$ can be interpreted as the sum $t_n = \sum_{r=1}^{n} r$ and as the binomial coefficient $\binom{n+1}{2}$ on the other. So this work may lead to the question: If s is fixed, for what r is $\binom{r+s}{k} - \binom{r}{k}$ a binomial coefficient $\binom{m}{k}$ for $k \geq 3$? Moreover with very little effort the answers to the above questions can be interpreted as solution of the diophantine equation $t_m + t_n = t_k$ where $k, m, n \in \mathrm{N}$. We treat the question in two ways. In section 1, we find expressions for r that depend upon factorizations $s = ab$ where $a > 0$, $b > 0$ and $(a, b) = 1$. We also show that if $s \neq 1$ and if there is another factorization $s = a'b'$ but $a' \neq a$, then these expressions give totally different sets of values of r. In fact if $s > 1$, these (coprime) factorizations of s partition the set $\{r \in \mathrm{N} \mid t_{r+s} - t_r \text{ is triangular}\}$. In section 2, we include an alternative approach to the problem by producing an algorithm that can be used for computing a sequence of triangular triplets for any given $s \in \mathrm{N}$.

1. Evaluating r from factorizations of s.

Our tools are the following elementary results that should belong to undergraduate text books on number theory.

Lemma 1.1

Let $a, b \in$ N. If $(a, b) = 1$, then there exists $k \in$N such that $a|k$ and $b|k + 1$.

Proof

Since $(a, b) = 1$, there exist $x, y \in \mathbb{Z}$ such that $xa + yb = 1$. If $x < 0$, then we can put $-xa = k$ which gives $yb = k + 1$. If $x > 0$, we rewrite $xa + yb = 1$ as $x[b-(b-1)]\,a + yb = 1$ or $-x(b-1)a + (x+y)b = 1$. Now $k = x(b-1)a$ gives $(xa + y)b = k + 1$. As we can always choose non-zero x and y there is no need to consider the case of $x = 0$.

The following lemma can be proved using prime factorization but the proof included here, which is definitely well-known, can be introduced at a much earlier stage.

Lemma 1.2

If, in N, $a|bc$, then $a = a_1a_2$ where $a_1|b$ and $a_2|c$.

Proof

Let $a_1 = (a, b)$. Then $a = a_1a_2$ and $b = a_1b_2$ where $(a_2, b_2) = 1$. Now $a|bc$ implies $a_1a_2|a_1b_2c$, which imples $a_2|c$.

First of all let us consider the well-known case when $s = 1$. Then $t_{r+1} - t_r$ is triangular if and only if $r+1 = \dfrac{m(m+1)}{2}$ for some $m \in$N, if and only if $r = \dfrac{(m+2)(m-1)}{2}$ for some $m \in$N. Since r must be positive we must have $m > 1$. Thus for $m > 1$, $r = \dfrac{(m+2)(m-1)}{2}$ makes $t_{r+1} - t_r$ triangular and conversely. We note that, for $r \in$N, if $t_{r+s} - t_r = t_m$ for some $m \in m$ for some $m \in$N, then

$$rs + \frac{s^2 + s}{2} = \frac{m^2 + m}{2} \qquad \text{(i)}$$

So obviously $m > s$. We can rewrite (i) as

$$rs = \frac{m^2 + m - (s^2 + s)}{2} \text{ or as } r = \frac{(m-s)(m+s+1)}{2s} \qquad \text{(ii)}$$

Thus if we know m, we can find r. Now s must divide $m(m+1)$. So we factorize s into coprime factors by Lemma 1.2 and invoke Lemma 1.1. Now the fact that 2 is in the denominator of (*ii*) may cause a slight confusion. So for the moment, let us assume that s is odd. In that case obviously one of $m-s$ or $m+s+1$ is even and so 2 does not interfere with the divisibility by s.

Proposition 1.3

Let s be an odd number. Then $t_{r+s}-t_r$ is a triangular number if and only if

$$r = \frac{[\alpha_0 + (t+1)b]\,[-\beta_0 + (t-1)a]}{2}$$

where $s=ab$ is a factorization of s such that $(a, b) = 1$, (α_0, β_0) is a solution of $\alpha a + \beta b = 1$ with $\alpha_0 < 0$ ($\beta_0 > 0$) and t is an integer such that $t < \frac{-b-\alpha_0}{b}$.

Proof

($\Longrightarrow$) If $t_{r+s}-t_r=t_m$ for some $m\in N$, then by (ii) $r = \frac{(m-s)\,(m+1+s)}{2s}$.
Now $s\,|\,(m-s)\,(m+1+s)$ and so $s = ab$ where $a\,|\,m-s$ and $b\,|\,m+1+s$. Since $a\,|\,s$ and $b\,|\,s$, $a\,|\,m$ and $b\,|\,m+1$. Now as $(a, b) = 1$, there exist α_0, β_0 in $\mathbb{Z}$ such that $a\alpha_0+b\beta_0=1$. By Lemma 1 we can assume that $\alpha_0<0$ and $\beta_0>0$. This provides us with one possible choice of m, i.e. $m=|\alpha_0 a|$. For a more general solution we note that the general solution of $\alpha a + \beta b = 1$ is given by $\alpha=\alpha_0+bt$ and $\beta=\beta_0-at$ where t runs over integers. Again to be able to choose m we must have $\alpha=\alpha_0+bt<0$ (and $\beta_0-at>0$) so that $t < -\frac{\alpha_0}{b}$ (and $t < \frac{\beta_0}{b}$). Now as $m=|\alpha_0+bt|a$ and as by (*ii*) $m-s$ must be positive, we must have

$$|\alpha_0 + bt|\, a > s$$

or $$|\alpha_0 + bt| > b$$

or $$\alpha_0 + bt < -b$$

or $$t < \frac{-b-\alpha_0}{b} = -1 - \frac{\alpha_0}{b} < -\frac{\alpha_0}{b}$$

A simple calculation shows that $-\frac{\alpha_0}{b} \leq \frac{\beta_0}{a}$ and so $t < -1 - \frac{\alpha_0}{b}$ is the only restriction on t to give us the required $m = |\alpha_0 + bt|a$ or $m = -(\alpha_0 + bt)a$. This leads to

$$r = \frac{[-(\alpha_0 + bt)a - ab]\,[-(\alpha_0 + bt)a + ab + 1]}{2ab}$$

or, as $\alpha_0 a = 1 - \beta_0 b$,

$$r = \frac{[\alpha_0 + (t + 1)\,b]\,[-\beta_0 + (t - 1)a]}{2}$$

Indeed any factorization of s will give rise to a set of r's and a value of m for each r.

($\Leftarrow$) If $r = \frac{[\alpha_0 + (t + 1)\,b]\,[-\beta_0 + (t - 1)a]}{2}$

where α_0, β_0, a, b and t satisfy the conditions stated above, then

$$r = \frac{[\alpha_0 a + (t + 1)\,s]\,[-\beta_0 b + (t - 1)s]}{2s}$$

$$= \frac{[(\alpha_0 + bt)a + s]\,[\alpha_0 a - 1 + tab - s]}{2s}$$

$$= \frac{[(\alpha_0 + bt)a + s]\,[(\alpha_0 - bt)a - 1 - s]}{2s}$$

Let $m = -(\alpha_0 + bt)a$, then

$$r = \frac{(-m + s)\,(-m - 1 - s)}{2s}$$

or $2rs = m^2 + m - [s^2 + s]$ or $rs + \frac{s^2 + s}{2} = \frac{m^2 + m}{2}$.

Hence $t_{r+s} - t_r = \frac{m(m + 1)}{2}$.

If in the proof of 'necessity' in Proposition 1.3 we had proceeded without assuming that s was odd, we would find that

$$r = \frac{[\alpha_0 + (t + 1)b]\,[-\beta_0 + (t - 1)a]}{2}$$

where $\alpha_0 < 0$, t is such that r is a positive integer. Now if s is even and $s = ab$ such that $(a, b) = 1$ and a is even, then since $\alpha_0 a + \beta_0 b = 1$, β_0 must be odd. So $-\beta_0 + (t - 1)a$ is odd under all circumstances. Now we show that α_0 can be chosen to be even. If α_0 is even, we are done. If not, consider $1 = (b + 1) - b$ and hence $\alpha_0(b + 1 - b)a + \beta_0 b = 1$ or $\alpha_0(b + 1)a + (\beta_0 - \alpha_0 a)b = 1$. Letting $\alpha_0(b+1) = \alpha'_0$ and $(\beta_0 - \alpha_0 a) = \beta'_0$, renaming α'_0, β'_0 as α_0, β_0, we have the desired solution, *i.e.* with α_0 even. Now $\alpha_0 + (t + 1)b$ is even only if t is odd. If b is even, a similar argument can be used to get β_0 and t will again have to run over odd integers. This leads to the following statement.

Proposition 1.4

Let s be even. Then $t_{r+s} - t_r$ is a triangular number if and only if

$$r = \frac{[\alpha_0 + (t + 1)b]\,[-\beta_0 + (t - 1)a]}{2}$$

where $s = ab$ such that $(a, b) = 1$ and (α_0, β_0) is a solution of $\alpha a + \beta b = 1$ with $\alpha_0 < 0$. Moreover t runs over odd numbers less than $\frac{-b - \alpha_0}{b}$, and α_0 is even (β_0 is even) if and only if a is even (respectivley b is even).

From the computation angle: If r and s are given, then m may or may not exist and if it does then it is a function of the factorization of s... as we have already seen. So if r and s are given and m exists, then, for some factorization $s = ab$ with $\alpha_0 a + \beta_0 b = 1$ and for some $t < -1 - \frac{\alpha_0}{b}$, where t is fixed by the value of r, $m = |\alpha_0 + bt|a$. The fact that m is a function of the factorization $s = ab$ gives rise to the questions: Can two different factorizations give the same m? Can two different m's give the same r? (for a given s!) This provides a sufficient excuse for the last part of our project.

Now for the last part of our project let us introduce the following notation. Let $s = ab$ where $(a, b) = 1$. Define $S(a, b) = \{r \in \mathbb{N} \mid r = (m - s)(m + 1 + s)/2s \text{ where } a \mid m \text{ and } b \mid m + 1\}$. Now the following result answers the questions raised above.

Proposition 1.5

Let $s = ab = cd$ where $(a, b) = (c, d) = 1$. If $a \neq c$, then $S(a, b) \cap S(c, d) = \phi$.

Proof

Suppose that $S(a,b) \cap S(c,d) \neq \phi$. Then there exist $m, m' \in N$ such that

$$r = \frac{(m - s)(m + 1 + s)}{2s} = \frac{(m' - s)(m' + 1 + s)}{2s}$$

where $a|m$, $b|m + 1$, $c|m'$ and $d|m' + 1$. The above equation reduces to $m^2 - m'^2 = -(m - m')$. If $m \neq m'$, we have $m + m' = -1$ which is impossible. Hence, $m = m'$. But $m = -(\alpha_0 + bt)a$ and $m' = -(\alpha'_0 + dt')c$ for some $\alpha_0, \alpha'_0, t, t'$ given by Proposition 1.3 and 1.4. This gives, with the restrictions on t, t'

$$\alpha_0 a + tab = \alpha'_0 c + t'dc \qquad (*).$$

Since $a|cd$, then $a = a_1 h$ where $h = (a, c)$ and $a_1|d$. Now as a divides the left hand side of (*) and $a|dc$, we have $a|\alpha'_0 c$ which implies $a_1|\alpha'_0$. Since $\alpha'_0 c + \beta'_0 d = 1$ and $a_1|d$, we have $a_1 = 1$. Hence $a|c$. Similarly $c|a$. Thus $m = m'$ only if $a = c$, contrary to the assumption that $a \neq c$.

Corollary 1.6

For any factorization $s = ab$ where $(a,b) = 1$, $S(a,b) \cap S(b,a) = \phi$, unless $a = b = 1$.

2. Sequences of triangular triples

To this point our work has been of technical nature, that is, we have established existence of r and we have shown how the factorization of s dictates the generation of r such that $t_{r+s} - t_r$ is a triangular number t_m. But if one wants to find, for a given $s \in N$, all possible triangular triplets $(r + s, r, m)$ such that $t_{r+s} - t_r = t_m$, then the first section is of no help. One aim of this section is to provide a computer program for all required triplets of the kind mentioned above. To develop a computer program we need to produce an algorithmic procedure. For that we proceed as follows.

Suppose we have for s, r, m$\in$N such that $t_{r+s} - t_r = t_m$. Then

$$\frac{(r+s)(r+s+1)}{2} - \frac{r(r+1)}{2} = \frac{m(m+1)}{2}$$

or $\quad s^2 + (2r+1)s = m(m+1)$.

Obviously $m \geq s$. So let $m = s + k$ for some $k \in \mathrm{N}$. This gives $s^2 + (2r+1)s = s^2 + (2k+1)s + k(k+1)$ which gives

$(2r+1)s = (2k+1)s + k(k+1)$ or $2(r-k)s = k(k+1)$ or $(r-k)s = \frac{k(k+1)}{2} = t_k$.

Obviously if m exists then $s \mid t_k$ and we have $r = k + \frac{t_k}{s}$. On the other hand if k is such that $s \mid t_k$ then we can write $r = k + \frac{t_k}{s}$ or $(r-k)s = \frac{k(k+1)}{2}$ from which the equation can be easily constructed. Hence for a given k such that $s \mid t_k$ there is a unique triangular triplet $(r+s, r, s+k)$ that satisfies $t_{r+s} - t_r = t_m$. This lead to the following "pseudo code" which may help in writing the proper program in the language the reader is familiar with.

10 $\quad s = ?$

20 $\quad k : 1, 2, \ldots, 600$

30 $\quad t_k = \frac{k(k+1)}{2}$

40 $\quad$ If $\frac{t_k}{s}$ is an integer, go to 60

50 $\quad$ If $\frac{t_k}{s}$ is not an integer next k

60 $\quad$ Print $r = k + \frac{t_k}{s}$

70 $\quad$ Print $x = r + s, y = r, z = k + s$

80 $\quad$ If $t_x - t_y \neq t_z$ print : there is something wrong

90 $\quad$ If $t_x - t_y = t_z$ next k

The following is a program in Pascal that produces triangular triplets. (We have chosen Pascal for availability. A program in Basic will work equally well.)

```
Program Triangular (OUTPUT):
CONST OutFile = 'LPT1';
Columns = 4;
Number Printed = 28;
Space = '       ';
VAR f: TEXT;
Procedure Triangular Triples (s: INTEGER);
VAR  k, Count,
     t, r, x, z,
     tx, ty, tz:  LONGINT;
BEGIN
       WRITELN (f);
       WRITELN (f);
       WRITELN (f);
       WRITELN (f ;     S = ', s);
       WRITELN (f);
       WRITELN (f);
       k := 0;
       Count := 0;
       WHILE Count < NumberPrint DO
           BEGIN
               k := k + 1;
               t := (k*(k + 1)) div 2;
               IF (t MOD s = 0)
               THEN
                   BEGIN
                       r := k + (t div s);
                       x := r + s;
                       z := k + s;
                       Count := Count + 1;
                       WRITE (f,  '      (', x: ',', r:,',z:,)');
                       IF Count MOD Columns = 0
                       THEN
                           BEGIN
                               WRITELN (f);
                               WRITELN (f);
```

```
                END
                tx := x*(x + 1) div 2;
                ty := r*(r + 1) div 2;
                tz := z*(z + 1) div 2;
                F (tx – ty <> tz)
                THEN
                   BEGIN
                      WRITELN (f, 'ERROR');
                END;
          END;
     END;
END;
```

In the following we produce lists of triangular triplets for $s = 1, 15, 70$.

$s = 1$

(3, 2, 2)	(6, 5, 3)	(10, 9, 4)	(15, 14, 5)
(21, 20, 6)	(28, 27, 7)	(36, 35, 8)	(45, 44, 9)
(55, 54, 10)	(66, 65, 11)	(78, 77, 12)	(91, 90, 13)
(105, 104, 14)	(120, 119, 15)	(136, 135, 16)	(153, 152, 17)
(171, 170, 18)	(190, 189, 19)	(210, 209, 20)	(231, 230, 21)
(253, 252, 22)	(276, 275, 23)	(300, 299, 24)	(325, 324, 25)
(351, 350, 26)	(378, 377, 27)	(406, 405, 28)	(435, 434, 29)

$s = 15$

(21, 6, 20)	(27, 12, 24)	(36, 21, 29)	(38, 23, 30)
(49, 34, 35)	(59, 44, 39)	(73, 58, 44)	(76, 61, 45)
(92, 77, 50)	(106, 91, 54)	(125, 110, 59)	(129, 114, 60)
(150, 135, 65)	(168, 153, 69)	(192, 177, 74)	(197, 182, 75)
(223, 208, 80)	(245, 230, 84)	(274, 259, 89)	(280, 265, 90)
(311, 296, 95)	(337, 322, 99)	(371, 356, 104)	(378, 363, 105)
(414, 399, 110)	(444, 429, 114)	(483, 468, 119)	(491, 476, 120)

$s = 70$

(93, 23, 90)	(114, 44, 105)	(147, 77, 125)	(205, 135, 154)
(252, 182, 174)	(291, 221, 189)	(348, 278, 209)	(351, 281, 210)
(414, 344, 230)	(465, 395, 245)	(538, 468, 265)	(654, 584, 294)
(741, 671, 314)	(810, 740, 329)	(907, 837, 349)	(912, 842, 350)
(1015, 945, 370)	(1096, 1026, 385)	(1209, 1139, 405)	(1383, 1313, 434)
(1510, 1440, 454)	(1609, 1539, 469)	(1746, 1676, 489)	(1753, 1683, 490)
(1896, 1826, 510)	(2007, 1937, 525)	(2160, 2090, 545)	(2392, 2322, 574)

Acknowledgements

We gratefully acknowledge some interesting discussions about this note with Billy Hodges, Don Aplin and Tom Polaski. In fact Tom Polaski wrote the above program in Pascal.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 84 – 96

MONOTONICITY PRESERVING PIECEWISE RATIONAL INTERPOLATION WITH POINT TENSION CONTROL

Muhammad Sarfraz

Department of Mathematics,
Punjab University,
Quaid-i-Azam Campus,
Lahore-54590, Pakistan.

Farzana S. Chaudhry

Department of Mathematics,
Islamia University,
Bahawalpur, Pakistan.

ABSTRACT

A C^1 piecewise Rational cubic interpolant is utilized to solve the problem of shape preserving interpolation. It is shown that the interpolation method can be applied to monotonic sets of data. This monotonic preserving interpolant has also the characteristics to control the shape at data points i.e. it has point tension control. The scheme generalizes the monotoniticity preserving results of Sarfraz [8].

Keywords. Rational, interpolation, plane curve, shape preserving, monotonic, shape control.

1. INTRODUCTION

A number of authors have considered the problem of shape preserving and shape controlling interpolation. For brevity the reader being refered to ([1]–[10]). Some of the methods (e.g. [5]–[7]) are global and non-parametric whereas some of the methods (e.g. [2]–[4]) discuss local and parametric methods.

This paper uses a piecewise rational cubic interpolant to solve the problem of shape preserving interpolation together with the characteristic of shape control provided at the knot positions. Both, the shape preservation and the shape control effects, are obtained by introducing some parameters in the interpolant. The results derived here are actually the generalizations of the monotonicity results of Sarfraz [8] who developed a C^1 shape preserving interpolation scheme for parametric curves using a piecewise rational function with one parameter in one interval. Sarfraz derived the constraints, on the shape parameters occuring in the rational function to make the interpolant preserve the monotonic shape of the data.

The notion adopted here is that for parametric curves although some mention is made for scalar curves as well.

This paper introduces some extra degrees of freedom, in the form of parameters, in the rational interpolant of Sarfraz [8] such that the resulting interpolant remains C^1. The treatment adopted for monotonic shape preservation is similar to that in Sarfraz [8] except that the extra degrees of freedom give rise to a new analysis together with the benefit of shape control at the data points. Following section begins with some preliminaries about this rational cubic interpolant. The constraints with monotonic data, are derived in Section 3. These constraints are dependent on the tangent vectors. The description of the tangent vectors, which are consistent and dependent on the given data, is made in Section 4. The monotonicity preserving results are explained with examples in Section 5.

2. THE RATIONAL CUBIC INTERPOLANT

Let $F_i \in \mathbb{R}^2$, $i = 0,\dots, n$ be a given set of data points, where $t_0 < t_1 < \dots < t_n$ is the knot spacing. Also let $D_i \in \mathbb{R}^2$, denote the first derivative values defined at the knots. We consider the C^1 piecewise rational cubic Hermite function defined by

$$(2.1)\quad P|_{[t_i,t_{i+1}]}(t) = \frac{(1-\theta)^3\alpha_i F_i + \theta(1-\theta)^2(r_i+\alpha_i)V_i + \theta^2(1-\theta)(r_i+\alpha_{i+1})W_i + \theta^3\alpha_{i+1}F_{i+1}}{(1-\theta)^2\alpha_i + \theta(1-\theta)r_i + \theta^2\alpha_{i+1}}$$

$\theta(t) \equiv (t-t_i)/h_i$, $h_i = t_{i+1} - t_i$, and

$$(2.2)\quad V_i = F_i + \frac{\alpha_i}{r_i+\alpha_i}h_i D_i,\; W_i = F_{i+1} - \frac{\alpha_{i+1}}{r_i+\alpha_{i+1}}h_i D_{i+1}$$

We shall use this to generate an interpolatory planar curve which preserves not only the shape of the data but also controls the shape at the data points. Let

$$
(2.3) \quad \begin{cases} P(t) = (p_1(t), p_2(t)), \\ F_i = (x_i, y_i) \\ D_i = (D_i^x, D_i^y), \\ \Delta_i = (\Delta_i^x, \Delta_i^y), \end{cases}
$$

where,

$$
(2.4) \quad \Delta_i^x = \frac{(x_{i+1} - x_i)}{h_i}, \quad \Delta_i^y = \frac{(y_{i+1} - y_i)}{h_i}.
$$

and D_i denote the tangent vector to the curve at the knot t_i. It can be noted that $P(t)$ interpolates the points F_i and the tangent vectors D_i at the knots t_i.

The parameters α_i's and r_i's are to be chosen such that $\alpha_i \geq 0$, $\alpha_{i+1} \geq 0$ and $r_i \geq -\alpha_i, -\alpha_{i+1}$, $\forall i$, which ensure a strictly positive denominator in the rational cubic. The scalar weights in he numerator of (2.1) are those given by degree raising the denominator to cubic form, since

$$
(1-\theta)^2\alpha_i + \theta(1-\theta)r_i + \theta^2\alpha_{i+1} = (1-\theta)^3\alpha_i + \theta(1-\theta)^2(r_i + \alpha_i) + \theta^2(1-\theta)(r_i + \alpha_{i+1}) + \theta^3\alpha_{i+1}
$$

It follows that if

$$
\alpha_i, \alpha_{i+1} \geq 0 \text{ and } r_i > -\alpha_i, -\alpha_{i+1},
$$

then the denominator is positive, and from bernstein-Bezier theory (see [10]), the curve segment $P|_{[t_i, t_{i+1}]}$ lies in the convex hull of the control points $\{F_i, V_i, W_i, F_{i+1}\}$ and its variation diminishing property holds with respect to the *control polygon* joining these points. The case $\alpha_i = 1 = \alpha_{i+1}$ recovers the rational cubic interpolant in [8]. Thus from now onward we shall assume

$$
0 \leq \alpha_i, \; \alpha_{i+1} \leq 1
$$

and this will not loss any generality.

The following *tension* properties of the rational Hermite form are immediately apparent from (2.1) and (2.2).

2.1 Point Tension

The *point tension* behaviour can be immediately observed from the following:

$$\lim_{\alpha_i \to 0} V_i = F_i,$$

$$\lim_{\alpha_i \to 0} P|_{[t_i, t_{i+1}]}(t) = \frac{(1-\theta)^2 r_i F_i + \theta(1-\theta)(r_i + \alpha_{i+1}) W_i + \theta^2 \alpha_{i+1} F_{i+1}}{(1-\theta)\, r_i + \theta\, \alpha_{i+1}}$$

and

$$\lim_{\alpha_{i+1} \to 0} W_i = F_{i+1},$$

$$\lim_{\alpha_{i+1} \to 0} P|_{[t_i, t_{i+1}]}(t) = \frac{(1-\theta)^2 \alpha_i F_i + \theta(1-\theta)(r_i + \alpha_i) V_i + \theta^2 r_i F_{i+1}}{(1-\theta)\, \alpha_i + \theta r_i}$$

Thus if the pieces of the rational cubics are joined together with bounded derivatives then the curve has a corner at the joining point if the α associated with that point approaches infinity.

3. INTERPOLATION OF MONOTONIC DATA

For our purposes α_i's and r_i will be chosen to ensure that the interpolant preserves the shape of the data. This choice requires the knowledge of $P^{(1)}(t)$ which is as follows:

$$(3.1)\quad P^{(1)}(t) = \frac{b_{1,i}(1-\theta)^4 + b_{2,i}\theta(1-\theta)^3 + b_{3,i}\theta^2(1-\theta)^2 + b_{4,i}\theta^3(1-\theta) + b_{5,i}\theta^4}{\{(1-\theta)^2\alpha_i + \theta(1-\theta)r_i + \theta^2\alpha_{i+1}\}^2}$$

where

$$(3.2)\quad \begin{cases} b_{1,i} = \alpha_i^2 D_i, \\ b_{2,i} = 2\alpha_i\, [(r_i + \alpha_{i+1})\, \Delta_i - \alpha_{i+1} D_{i+1}], \\ b_{3,i} = [(\alpha_i + r_i)(\alpha_{i+1} + r_i) + 3\alpha_i\alpha_{i+1}]\Delta_i - (\alpha_i + r_i)\alpha_{i+1} D_{i+1} \\ \qquad -(\alpha_{i+1} + r_i)\, \alpha_i D_i, \\ b_{4,i} = 2\alpha_{i+1}\, [(r_i + \alpha_i)\, \Delta_i - \alpha_i D_i], \\ b_{5,i} = \alpha_{i+1}^2\, D_{i+1}, \end{cases}$$

and we denote

$$(3.3)\quad b_{j,i} = (b_{j,i}^x, b_{j,i}^y).$$

Let us assume for simplicity that

$$(3.4)\quad \Delta_i^x \neq 0, \; i = 0, \ldots, n-1,$$

and that the data is monotonic increasing and arises from a function. Then we must have

(3.5) $$\frac{\Delta_i^y}{\Delta_i^x} \geq 0, \ i = 0, \ldots, n-1,$$

i.e. Δ_i^x and Δ_i^y are of the same sign. (The case of monotonic decreasing set of data can be treated in a similar manner when the inequalities are reversed.) The necessary conditions for the interpolant $P(t)$ to be monotonic are, then, the following:

(3.6) $$\frac{D_i^y}{D_i^x} \geq 0, \ i = 0, \ldots, n-1,$$

i.e. D_i^x and D_i^y are of the same sign. We also note that D_i^x and D_i^y must have the same sign as Δ_i^x and Δ_i^y respectively. Thus we have the following:

Lemma 3.1

The monotonicity conditions imply that

(3.7) $$\begin{cases} \Delta_i^x \Delta_i^y, D_i^y D_i^x \geq 0, \\ \Delta_i^x D_i^y, \Delta_i^y D_i^x \geq 0, \\ \Delta_i^x D_{i+1}^y, \Delta_i^y D_{i+1}^x \geq 0, \end{cases}$$

for $i = 0, \ldots, n-1$.

Remark 3.1

Let

(3.8) $$\begin{cases} \beta_{1,i} = \Delta_i^x \Delta_i^y, \\ \beta_{2,i} = D_i^x \Delta_i^y + D_i^y \Delta_i^x, \\ \beta_{3,i} = \Delta_i^x D_{i+1}^y + \Delta_i^y D_{i+1}^x, \\ \beta_{4,i} = D_i^x D_{i+1}^y + D_i^y D_{i+1}^x, \\ \beta_{5,i} = D_i^x D_i^y. \end{cases}$$

Then it follows from (3.7) that

(3.9) $$\beta_{j,i} \geq 0, \ j = 1, \ldots, 5, \ i = 0, \ldots, n-1.$$

Now, $P(t)$ is monotonic increasing if and only if

(3.10) $\dfrac{p_2^{(1)}(t)}{p_1^{(1)}(t)} \geq 0 \ , \ \forall\, t \in [t_0, t_n].$

i.e. $p_1^{(1)}(t)$ and $p_2^{(1)}(t)$ are of the same sign. Thus (3.10) can be equivalently written as

(3.11) $p_1^{(1)}(t)\, p_2^{(1)}(t) \geq 0\,,\ \forall\, t \in [t_0, t_n].$

After some simplifications, using (2.3) – (2.4) and (3.1) – (3.3), it can be shown that for $t \in [t_i, t_{i+1}]$,

(3.12) $$p_1^{(1)}(t)\, p_2^{(1)}(t) = \frac{\sum_{j=1}^{9} \gamma_{j,i}\, (1-\theta)^{9-j}\, \theta^{j-1}}{\{(1-\theta)^2 \alpha_i + \theta(1-\theta) r_i + \theta^2 \alpha_{i+1}\}^4},$$

where $\gamma_{1,i} = \alpha_i^4\, D_i^x\, D_i^y$,

$$\gamma_{2,i} = b_{1,i}^x\, b_{2,i}^y + b_{2,i}^x\, b_{1,i}^y = 2\alpha_i^3\, [r_i \beta_{2,i} - \alpha_{i+1}(\beta_{4,i} - \beta_{2,i})]\,,$$

$$\gamma_{3,i} = b_{1,i}^x\, b_{3,i}^y + b_{3,i}^x\, b_{1,i}^y + b_{2,i}^x\, b_{2,i}^y$$

$$\gamma_{4,i} = b_{1,i}^x\, b_{4,i}^y + b_{4,i}^x\, b_{1,i}^y + b_{2,i}^x\, b_{3,i}^y + b_{3,i}^x\, b_{2,i}^y$$

$$\gamma_{5,i} = b_{1,i}^x\, b_{5,i}^y + b_{5,i}^x\, b_{1,i}^y + b_{2,i}^x\, b_{4,i}^y + b_{4,i}^x\, b_{2,i}^y + b_{3,i}^x\, b_{3,i}^y$$

$$\gamma_{6,i} = b_{2,i}^x\, b_{5,i}^y + b_{5,i}^x\, b_{2,i}^y + b_{3,i}^x\, b_{4,i}^y + b_{4,i}^x\, b_{3,i}^y$$

$$\gamma_{7,i} = b_{3,i}^x\, b_{5,i}^y + b_{5,i}^x\, b_{3,i}^y + b_{4,i}^x\, b_{4,i}^y$$

$$\gamma_{8,i} = b_{4,i}^x\, b_{5,i}^y + b_{5,i}^x\, b_{4,i}^y = 2\, \alpha_{i+1}^3\, [r_i \beta_{3,i} - \alpha_i(\beta_{4,i} - \beta_{3,i})],$$

$$\gamma_{9,i} = \alpha_{i+1}^4\, D_{i+1}^x\, D_{i+1}^y$$

The conditions

(3.14) $D_j^x\, D_j^y \geq 0, j = i, i+1$

are necessary for the interpolant to be monotonic increasing on $[t_i, t_{i+1}]$ (see Lemma 3.1) and, assuming these necessary conditions, sufficient conditions are

(3.15) $\gamma_{j,i} \geq 0$, $j = 2, ..., 8$.

It should be noted that if $\Delta_i^y = 0$, then $D_i^y = D_{i+1}^y = 0$ and hence $\beta_{2,i} = \beta_{3,i} = 0$. Moreover, $p_2(t) = y_i$, $t_i \leq t \leq t_{i+1}$. Therefore $P(t)$ is constant on $[t_i, t_{i+1}]$.

If $\Delta_i^y \neq 0$, then a sufficient condition for (3.15) is

$$(3.16) \quad r_i \geq max\left\{\alpha_{i+1}\left(\frac{\beta_{4,i}}{\beta_{2,i}}\right), \alpha_i\left(\frac{\beta_{4,i}}{\beta_{3,i}} - 1\right)\right\}.$$

Moreover, since

$$(3.17) \quad max\left\{\frac{\beta_{4,i}}{\beta_{2,i}}, \frac{\beta_{4,i}}{\beta_{3,i}}\right\} \geq max\left\{\alpha_{i+1}\left(\frac{\beta_{4,i}}{\beta_{2,i}} - 1\right), \alpha_i\left(\frac{\beta_{4,i}}{\beta_{3,i}} - 1\right)\right\},$$

the choice,

$$(3.18) \quad r_i = \frac{\beta_{4,i}(\beta_{2,i} + \beta_{3,i})}{\beta_{2,i}\,\beta_{3,i}},$$

satisfies (3.16) and provides nice graphical results. This is the same choice as discovered in [8].

Remarks 3.2

The scalar case can be considered as an application of interpolation scheme $(t, P(t))$ in $\mathbb{R}^2$ to the values $(t_i, F_i) \in \mathbb{R}^2$ and derivatives $(1, D_i^y) \in \mathbb{R}^2$, $i = 0, ..., n$. It can als be noted that $\Delta_i = (1, \Delta_i^y)$.

4. CHOICE OF TANGENT VECTORS

In most applications, the tangent vectors D_i will not be given and hence must be determined from the data $F_i \in \mathbb{R}^2$, $i = 0, ..., n$. We describe here the geometric mean choices of tangent vectors for our plane curves which satisfy the shape preserving conditions. The geometric mean choice of tangent vectors is

$$D_i = ((\Delta_{i-1}^x)^{\lambda_i} (\Delta_i^x)^{1-\lambda_i}, (\Delta_{i-1}^y)^{\lambda_i} (\Delta_i^y)^{1-\lambda_i}), \quad i = 1, ..., n-1,$$

with the end conditions

$$D_0 = ((\Delta_0^x)^{\lambda_0} (\Delta_{2,0}^x)^{1-\lambda_0}, (\Delta_0^y)^{\lambda_0} (\Delta_{2,0}^y)^{1-\lambda_0}),$$

$$D_n = ((\Delta^x_{n-1})^{\lambda_n} (\Delta^x_{n,n-2})^{1-\lambda_n}, (\Delta^y_{n-1})^{\lambda_n} (\Delta^y_{n,n-2})^{1-\lambda_n}),$$

where

$$\lambda_i = \frac{h_i}{(h_{i-1} + h_i)}, \quad i = 1,\ldots, n-1,$$

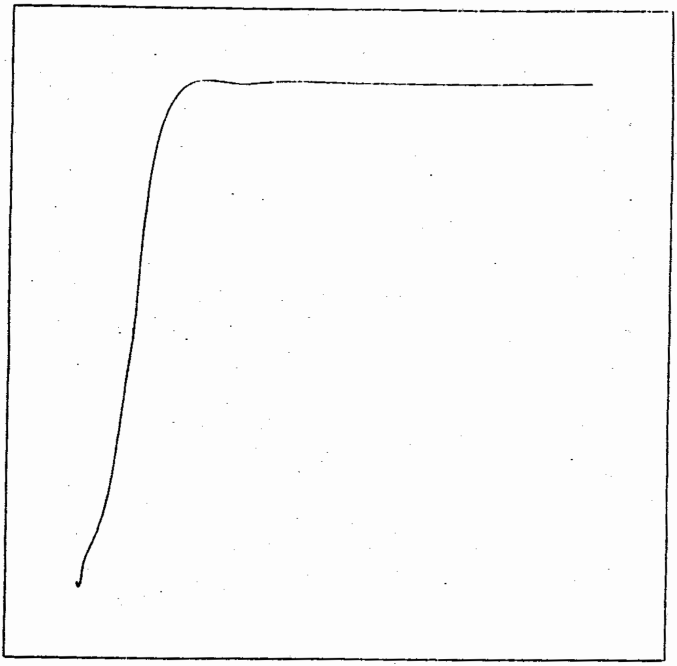
$$\lambda_0 = 1 + \frac{h_0}{h_1}, \quad \lambda_n = 1 + \frac{h_{n-1}}{h_{n-2}},$$

$$\Delta_{2,0} = \frac{F_2 - F_0}{t_2 - t_0}, \quad \Delta_{n,n-2} = \frac{F_n - F_{n-2}}{t_n - t_{n-2}}.$$

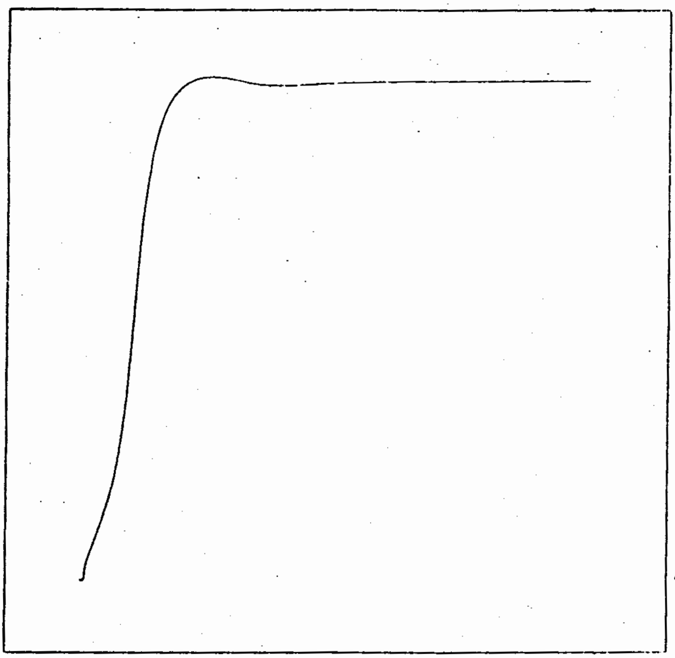
These geometric mean approximations are suitable for monotonic data since they satisfy the necessary conditions for monotonicity and produce pleasing graphical results.

5. EXAMPLES AND DISCUSSION OF RESULTS

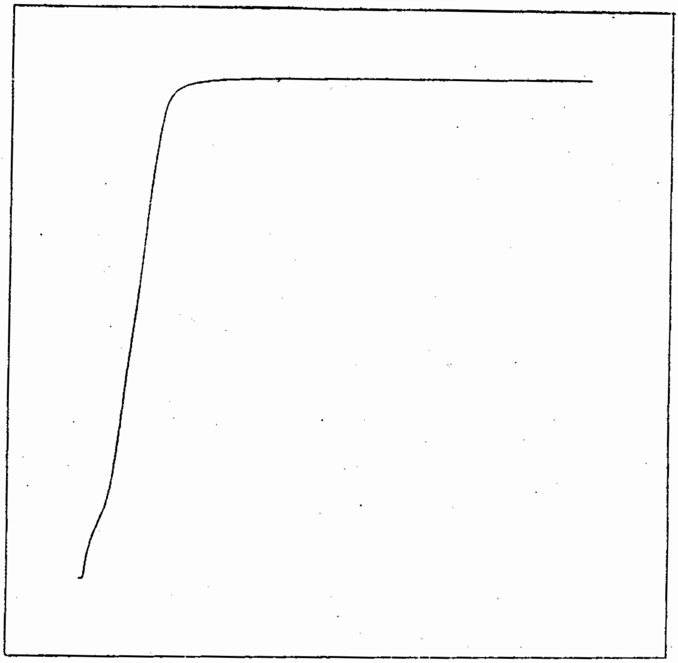
Demonstration of the results, of this research, is made onto the data due to Fritsch and Carlson [1]. All figures except Figures 1 and 2 (these are standard cubic spline scalar and parametric curves respectively) demonstrate the monotonicity preserving results corresponding to the geometric mean derivative values of Section 4. Unless and otherwise stated we shall take $\alpha_i = 1, \forall\, i$ (this is the case of monotonicity preserving curves in [8] where the user does not have freedom to play with the picture and make modifications in the desired regions of the curve) and the parametrization used here is the chord length parametrization: some other parametrization can also be used. The curves in Figures 3 and 4 are, respectively, the scalar and parametric monotonicity preserving interpolations. The rest of the figures 5–7 represent monotonicity preserving parametric curves with shape control: the curve in Figure 5 is tightened at the sixth data point with $\alpha_6 = .01$: the curve in Figure 6 is tightened in the sixth interval with $\alpha_6 = \alpha_7 = .01$: the curve in Figure 7 is tightend globally with $\alpha_i = .01, \forall\, i$, such that it approaches the control polygon in limit. Thus, in general, user can made interactions whereever the picture is not pleasing and hence possesses a reasonable fascility in hand to implement any monotonic data rising from some scientific phenomena.



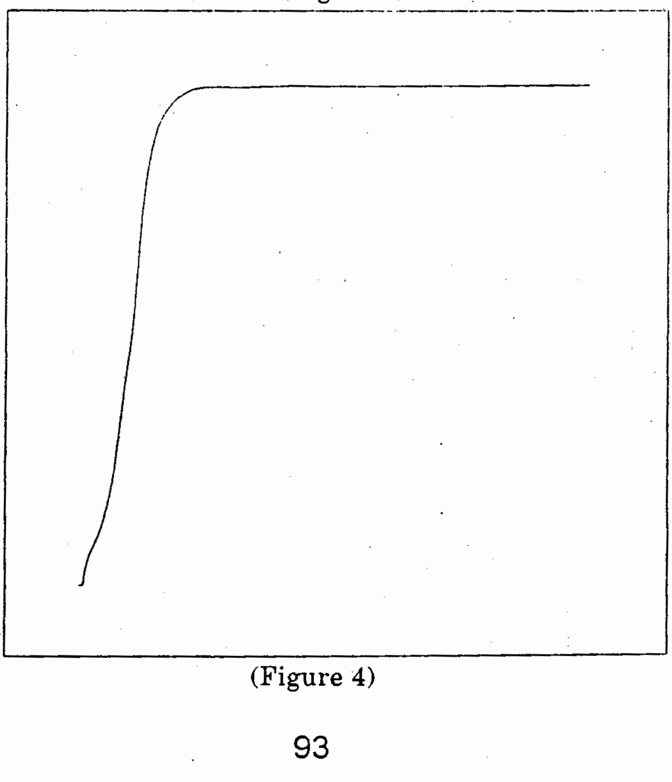
(Figure 1)



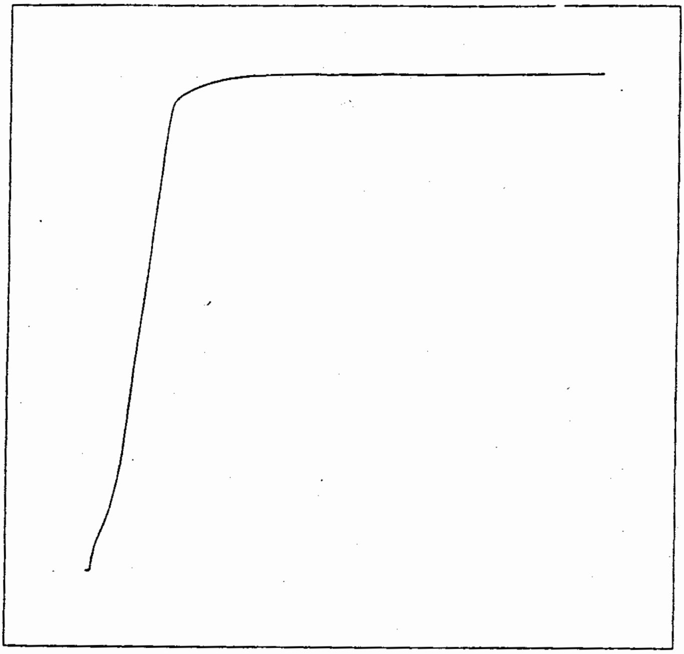
(Figure 2)



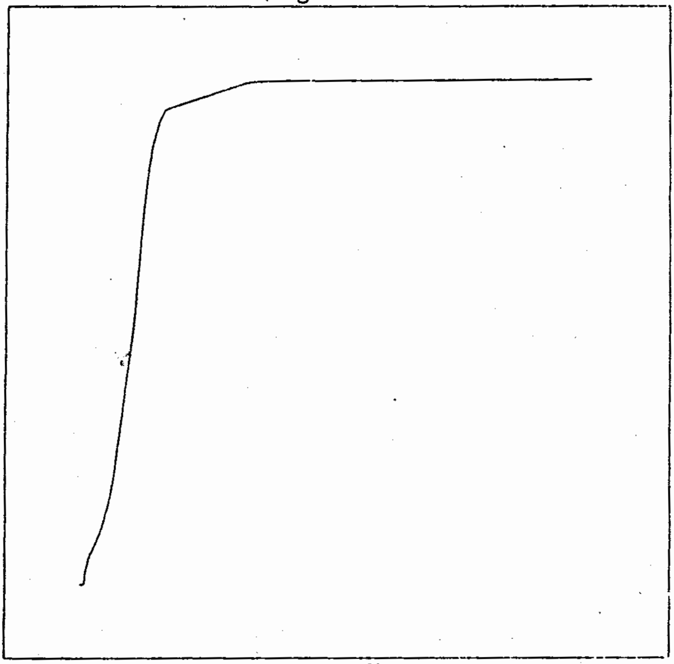
(Figure 3)



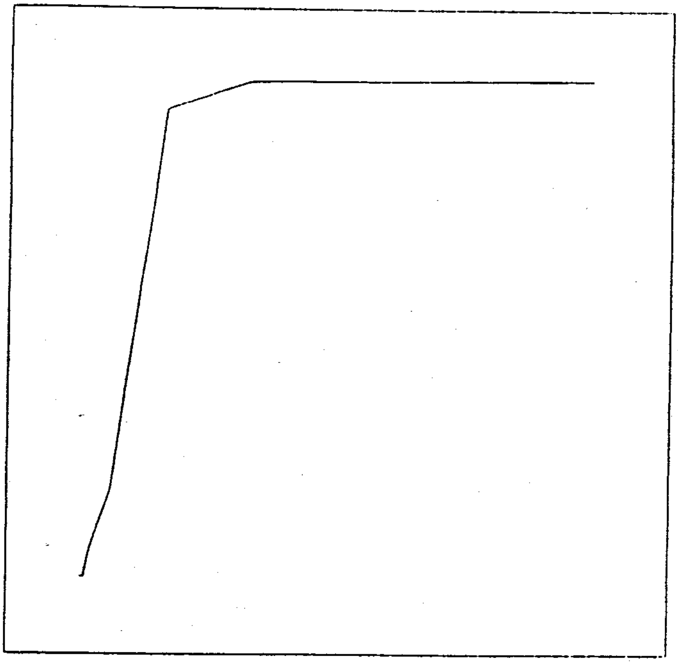
(Figure 4)



(Figure 5)



(Figure 6)



(Figure 7)

6. CONCLUDING REMARKS

C^1 rational cubic Hermite interpolant, with two families of shape parameters, has been utilized to obtain a C^1 monotonicity preserving plane curve method with the provision of point tension control. Data dependent shape constraints are derived on one family of shape parameters to assure the shape preservation of the data while the other family of parameters is left free to control and finetune the shape further in the desired areas of curves. Choice of the tangent vectors, which are consistent and dependent on the data, has also been made. This scheme can also be implemented in the scalar case.

REFERENCES

1. F.N. Fritsch and R.E. Carlson, Monotone piecewise cubic interpolation, *SIAM J. Num. Anal.* 17(1980), 238–246.
2. T.N.T Goodman, Shape preserving interpolation by parametric rational cubic splines, *Proc. Int. Conf. on Numerical Mathematics*, Int. series Num. Math. 86, Birkhauser Verlag, Basel, (1988).

3. T.N.T Goodman and K. Unsworth, Shape preserving interpolation by parametrically defined curves, *SIAM J. Num. Anal.* 25(1988), 1–13.

4. T.N.T Goodman and K. Unsworth, Shape preserving interpolation by curvature continuous parametric curves, *Comput. Aided Geom. Des.* 5(1988), 323–340.

5. J. A. Gregory, Shape preserving spline interpolation, Comput. Aided Des. 18(1986), 53–57.

6. R. Delbougo and J. A. Gregory, C2 rational quadratic spline interpolation to monotonic data, *IMA J. Num. Anal.* 3(1983), 141-152.

7. R. Delbougo and J. A. Greogry, Shape preserving piecewise rational interpolation. *SIAM J. Sci. Stat. Comput.* 6(1985), 967–976.

8. M. Sarfraz, Monotonicity preserving piecewise rational interpolation, *Punjab University Journal of Mathematics*, 26(1993).

9. J. A. Gregory and M. Sarfraz, A rational cubic spline with tension, Comput. Aided Geom. Des. 7(1990), 1–13.

10. W. Boehm, G. Farin and J. Kahmann, A survey of curve and surface methods in *CAGD, Comput. Aided Geom. Des.* 1(1984), 1–60.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 97 – 100

RANK OF MODULES OVER A LEFT ORE DOMAIN

M. A. Rauf Qureshi
Department of Mathematics,
University of Karachi, Pakistan.

1. INTRODUCTION

We assume familiarity with Ore domains. However, a reference may be made to [3], where it is shown that a left Ore domain R can be embeded in a left skew-field D of fractions of the form $T^{-1}R$, where $T=R-\{0\}$. In what follows R and D will be used in this sense except in 3.8.

The object of the paper is to carry over the concept of rank from the category of abelian groups to the category R-mod of unitary left R-modules, and study some of its features. Among other things, we shall determine a necessary and sufficient condition for $A\in R$-mod to have prescribed rank in terms of a property of A, and show that the rank of A is same as that of its injective envelope.

2. BASIC NOTIONS

Besides the idea of an Ore domain, we shall also need the definitions of torsion and torsion free objects given in [4], which may also be found in [3], where the words "T-torsion" and "T-torsion free" are used. In the following statments A,A′, ..., will stand for objects in R-mod unless stated otherwise.

2.1 Proposition

The set A_0 of all torsion elements of A is a submodule of A, and A/A_0 is torsion free.

We now state two more results, the proof of the first is given in [4], while for the other reference may be made to [2]. In fact, the proof of the second is similar to that of (1) in [1, ch. vii, p. 130].

2.2 Proposition

If A is torsion free, then its injective envelope is isomorphic to $D\otimes_R A$.

2.3 Proposition

D is flat as right R-module.

In the next section we shall suppose that all tensor products are taken over R, or to be precise $\otimes = \otimes_R$.

3. THEORY OF RANK

We define the rank of A, denoted by $r(A)$, by the dimension of the left Vector Space $D\otimes A$.

3.1 Proposition

We have

(i) $r(A) = 0$, whenever A is torsion,

(ii) $r(A) = \sum_{i\in J} r(A_i)$, if A is the direct sum $\bigcup_{i\in J} A_i$.

(iii) $r(A) = r(A') + r(A'')$, in case the sequence $0 \to A' \to A \to A'' \to 0$.

in R-mod is exact.

Proof

(i) follows, since $D\otimes A = 0$ in case A is torsion, and (ii) is imediate by observing that $D\otimes \bigcup_{i\in J} A_i$ is isomorphic to $\bigcup_{i\in J} (D\otimes A_i)$. For (iii) we use 2.3 to obtain the exact sequence $0\to D\otimes A' \to D\otimes A \to D\otimes A'' \to 0$, which splits. Hence we have (*iii*) in view of (*ii*).□

From (i) and (ii) of 3.1 we immediately get

3.2 Corollary

Let $0\to A'\to A\to A''\to 0$ be an exact sequence in R-mod. Then

(i) $r(A) = r(A'')$, if A' is torsion,

(ii) $r(A) = r(A')$, whenever A'' is torsion.□

3.3 Proposition

If E is the injective envelope of A, then

$r(A) = r(A/A_0) = r(E)$.

Proof

We can suppose that $A \subseteq E$, so that E/A is torsion, and the result follows using 3.2 on the exact sequences $0\to A_0\to A\to A/A_0\to 0$, and $0\to A\to E\to E/A\to 0$.□

3.4 Corollary

$r(R) = r(D) = 1$.

Proof

Since D $\in$ R-mod is injective envelope of R, the result follows from 3.3.□

3.5 Corollary

Let A be free with a basis B. Then $r(A) = O(B)$, where O(B) is the cardinal number of B.

Proof

Since A is isomorphic to $\bigcup_{b \in B} R_b$, where $R_b = R$, for each b, therefore, by 3.1 and 3.4 we obtain $r(A) = O(B)$.□

Although, 3.5 establishes invariance of basis number for any free object A, but this consequence is weaker in the presence of a more general result (see [3, p.18]), which states that, if R_1 is a subring of a ring R_2 with identity element $1 \in R$, then R_1 has 1BN whenever R_2 does. Hence R has 1BN.

3.6 Proposition

Suppose that A is torsion free. Then the injective envelope E of A is isomorphic to $\bigcup_{i \in J} D_i$, where each D_i is D and $r(A) = O(J)$.

Proof

By 2.2 we have $E \cong D \otimes A \cong \bigcup_{i \in J} D_i$.

so that $r(A) = O(J)$ by 3.1 and 3.4.□

3.7 Proposition

A has rank α if and only if the injective envelope of A/A_0 is isomorphic to $\bigcup_{i \in J} D_i$ with $D_i = D$ for each i and $O(J) = \alpha$.

Proof

In view of 2.1 and 3.3 we can suppose that A is torsion free. Let the injective envelope of A be isomorphic to $\bigcup_{i \in J} D_i$, where $\alpha = O(J)$.

Then by 3.1 and 3.3 $r(A) = r(\bigcup_{i \in J} D_i) = \sum_{i \in J} r(D_i) = O(J)$. The converse follows from 3.6.□

Observe that we have not made full use of the properties of R and D in the proofs of 3.1 and 3.2. Hence with obvious definitions 3.1, 3.2 and half of 3.3 may be stated, more generally, as

3.8 Proposition

Let R be a subring of a skew-field D, which is flat as right R-module, and R contains the identity of D. Then the statements of 3.1 and 3.2 hold, and furthermore $r(A) = r(E)$, where E is the injective envelope of $A \in$ R-mod.□

It is an open question, whether a ring satisfying the conditions of 3.8 is a left Ore domain.

REFERENCES

1. H. Cartan, and S. Eilenberg—Homological Algebra, Princeton University Press, 1956.

2. E.R. Gentile—On rings with one sided field of quotients, *Proc. AMS*, V.3, Part I, June 1960, 380–384.

3. D.S. Passman—A course in ring theory, Wardsworth & Brooks, California, 1991.

4. M. A. Rauf Qureshi—Left Ore domains, *Jour. Nat. Sci. and Math.*, V. 8(1968), 217–220.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 101 – 110

ON A CHARACTERIZATION OF $\hat{F}_4(2)$, THE REE-EXTENSION OF THE CHEVALLEY GROUP $F_4(2)$...I

Syed Muhammad Husnine
Department of Mathematics,
University of the Punjab,
Lahore, Pakistan.

ABSTRACT

In this paper we investigate a finite group G having a subgroup isomorphic to the centralizer of an involution in $F_4(2)$ under a simpler condition than that treated in Husnine [4].

1. INTRODUCTION

Let $F_4(2)$ denote the Chevalley group of type (F_4) over the field of two elements, say, $\Gamma=\{1, 0\}$. Then the centre of a Sylow_2 -subgroup S of $F_4(2)$ is elementary abelian of order 4. We denote the three involutions of the centre Z(S) of S by t_1, t_2, and $t_3=t_1 t_2$.

Now in $F_4(2)$, $C(t_1) \cong C(t_2)$ and $C(t_3) = C(t_1) \cap C(t_2)$. In Husnine [4] a characterization of $F_4(2)$ has been given by the following theorem:

Theroem 1. [4]

Let G be a finite group and y_1, a 2-central involution in G, so that $C = C_G(y_1) \cong C(t_1)$ and assume that G does not contain a normal complement to C of odd order. Then $G \cong F_4(2)$.

In the above theorem the condition for y_1 to be 2-central implies that a Sylow_2-subgroup of C is also a Sylow_2-subgroup of G. We intend to generalize this result by removing this condition. In fact we make the following conjecture which is more general.

Conjecture. Let G be a finite group having an involution y_1 such that $C=C(y_1)$ in G is isomorphic to $C(t_1)$. If y_1 is not a central involution then, $G \cong \hat{F}_4(2)$, an extension of $F_4(2)$ by the automorphism of order two used by Ree in [5].

In section 2, we describe the group $F_4(2)$ and $\hat{F}_4(2)$ following Carter [1], Guterman [3], and Ree [5], so that the paper becomes independent to a reasonable degree.

We state and prove our results in section 3. This involves the determination of a Sylow_2-subgroup of G and a 2-local subgroup of G, not contained in the centralizer of y_1 in G.

All notations are standard and follow Gorenstein [2]. However, we define the conjugate of y under x to be $x^{-1}y\,x$ for the elements x, y of a group G. We also write y^x for $x^{-1}y\,x$ and H^x for $x^{-1}H\,x$ for the elements x and y and the subset H of G. The set of all conjugates of an element y in G is denoted by $ccl_G(y)$. We write 'S_p-subgroup' for 'Sylow_p-subgroup' and $\sum_n$ for the symmetric group of degree n.

(2.1) The root system $\sum$ of type (F_4) consists of 48 roots; $\pm\xi_i$, $\pm\xi_i\pm\xi_j$, $\frac{1}{2}(\pm\xi_i \pm\xi_j \pm\xi_m \pm\xi_n)$, where $i, j, m, n = 1, 2, 3, 4$ and i, j, m, n are all distinct. We take $r_1 = \xi_4$, $r_2 = \xi_3 - \xi_4$, $r_5 = \xi_2 - \xi_3$, and $r_{10} = \frac{1}{2}(\xi_1 - \xi_2 - \xi_3 - \xi_4)$ as a system of fundamental roots. If we denote the root $ar_1 + br_2 + cr_5 + dr_{10}$ by $(abcd)$, then the positive roots are:

$r_1 = 1000$,	$r_2 = 0100$,	$r_3 = 1100$,	$r_4 = 2100$,
$r_5 = 0010$,	$r_6 = 0110$,	$r_7 = 1110$,	$r_8 = 2110$,
$r_9 = 2210$,	$r_{10} = 0001$,	$r_{11} = 1001$,	$r_{12} = 1101$,
$r_{13} = 2101$,	$r_{14} = 1111$,	$r_{15} = 2111$,	$r_{16} = 2211$,
$r_{17} = 3211$,	$r_{18} = 2102$,	$r_{19} = 2112$,	$r_{20} = 2212$,
$r_{21} = 3212$,	$r_{22} = 4212$,	$r_{23} = 4312$,	$r_{24} = 4322$.

Let Δ be the additive group generated by $\sum$. We define an inner product $<,>$ on $V = R \otimes \Delta$, *the vector space over the real numbers R, by $<\xi_i, \xi_j> = 0$ and $<\xi_i, \xi_i> = 1$ for $i, j = 1, 2, 3, 4; i \neq j$. For $r, s \in \sum$ let $\lambda(r) = <r, r>$ and $s(r) = 2<s, r>/<r, r>$. The values $\lambda(r_j)$ and $r_j(r_i)$ for $i = 1, 2, 5, 10$ and $1 \leq j \leq 24$ are given in table 1.

For each i, $1 \leq i \leq 24$ and each $s \in \sum$ let $\widetilde{w}_i(s) = s - s(r_i)r_i$. Then $\widetilde{w}_i$ is a permutation of $\sum$. The perutation group $\widetilde{W}$ generated by $\{\widetilde{w}_i \mid 1 \leq i \leq 24\}$ is the Weyl group of $\sum$.

(2.2) $\widetilde{W}$ is of order $2^7 3^2$ and is generated by $\widetilde{w}_1$, $\widetilde{w}_2$, $\widetilde{w}_5$ and $\widetilde{w}_{10}$. If $a_{ij} = |\widetilde{w}_i\,\widetilde{w}_j|$, then generators $\widetilde{w}_1$, $\widetilde{w}_2$, $\widetilde{w}_5$ and $\widetilde{w}_{10}$ together with

* We identify $I_R \otimes a$ with $a \in \Delta$.

the relations $(\widetilde{w}_i \widetilde{w}_j)^{a_{ij}} = 1$, $\{i, j\} \subseteq \{1, 2, 5, 10\}$, form a presentation of $\widetilde{W}$.

It will be convenient to think of the elements $\widetilde{w} \in \widetilde{W}$ as permutations of $\{\pm i \mid 1 \leq i \leq 24\}$ defined as follows:

$$\widetilde{w}(i) = \begin{cases} j \text{ if } \widetilde{w}(r_i) = r_j \ ; \quad \widetilde{w}(-i) = -\widetilde{w}(i) \\ -j \text{ if } \widetilde{w}(r_i) = -r_j \end{cases}$$

The values $\widetilde{w}_i(j)$ for $i = 1, 2, 5, 10$ and $1 \leq j \leq 24$ are also included in table 1.

Table - 1

i	$\widetilde{w}_1(i)$	$\widetilde{w}_2(i)$	$\widetilde{w}_5(i)$	$\widetilde{w}_{10}(i)$	$r_i(r_1)$	$r_i(r_2)$	$r_i(r_5)$	$r_i(r_{10})$	$\lambda(r_i)$	$\bar{i}$
1	−1	3	1	11	2	−1	0	−1	1	2
2	4	−2	6	2	−2	2	−1	0	2	1
3	3	1	7	12	0	1	−1	−1	1	4
4	2	4	8	18	2	0	−1	−2	2	3
5	5	6	−5	5	0	−1	2	0	2	10
6	8	5	2	6	−2	1	1	0	2	11
7	7	7	3	14	0	0	1	−1	1	18
8	6	9	4	19	2	−1	1	−2	2	12
9	9	8	9	20	0	1	0	−2	2	13
10	11	10	10	−10	−1	0	0	2	1	5
11	10	12	11	1	1	−1	0	1	1	6
12	13	11	14	3	−1	1	−1	1	1	8
13	12	13	15	13	1	0	−1	0	1	9
14	15	14	12	7	−1	0	1	1	1	19
15	14	16	13	15	1	−1	1	0	1	20
16	17	15	16	16	−1	1	0	0	1	22
17	16	17	17	21	1	0	0	−1	1	23
18	18	18	19	4	0	0	−1	2	2	7
19	19	20	18	8	0	−1	1	2	2	14
20	22	19	20	9	−2	1	0	2	2	15
21	21	21	21	17	0	0	0	1	1	24
22	20	23	22	22	2	−1	0	0	2	16
23	23	22	24	23	0	1	−1	0	2	17
24	24	24	23	24	0	0	1	0	2	21

(2.3) From table 1, we see that the values a_{ij} are as follows:

$a_{ii} = 1$ for $i = 1, 2, 5, 10$

$a_{ij} = 2$ for $\{i,j\} = \{1, 5\}, \{2, 10\}, \{5, 10\}$

$a_{ij} = 3$ for $\{i,j\} = \{1, 10\}, \{2, 5\}$

$a_{ij} = 4$ for $\{i,j\} = \{1, 2\}$

(2.4) $\widetilde{W}$ acts transitively on $\{r \in \Sigma \mid \lambda(r) = i\}$, i = 1, 2.

Let Γ be a field with two elements and let F be the Chevalley group of type (F_4) over Γ. Then F has the following properties:

(2.5) F is simple.

(2.6) For each i, $1 \le i \le 24$, there exists a homomorphism

$\phi_i : SL(2, 2) \to F$.

For each $\alpha \in \Gamma$, we define,

$$x_i(\alpha) = \phi_i\begin{pmatrix}1 & 0\\ \alpha & 1\end{pmatrix},\ x_{-i}(\alpha) = \phi_i\begin{pmatrix}1 & \alpha\\ 0 & 1\end{pmatrix},\ w_i = \phi_i\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

Note that $w_i^2 = x_i(\alpha)^2 = (w_i\, x_i(1))^3 = 1$.

(2.7) For each i, $1 \le i \le 24$, let $S_i = \{x_i(\alpha) \mid \alpha \in \Gamma\}$. Then each S_i is a group of order 2. The elements of S_i multiply according to the rule $x_i(\alpha)\, x_i(\beta) = x_i(\alpha + \beta)$, $\alpha, \beta \in \Gamma$.

(2.8) Let $S = \langle S_i \mid 1 \le i \le 24\rangle$. Then S is a Sylow$_2$-subgroup of F. Any element $x \in S$ can be expressed uniquely in the form $x = \prod_{i=1}^{24} x_i(\alpha_i)$ which we shall abbreviate as $x = \prod x_i(\alpha_i)$. Hence S has order 2^{24}. The product of any two elements of S may be obtained by use of the commutators $[x_i(1), x_j(1)]$, $1 \le i < j \le 24$. The nontrivial commutators are listed in table 2.

It is easily observed that $Z(S) = S_{21}S_{24}$.

(2.9) $N_F(S) = S$ and $S \cap W = 1$, where $W = \langle w_i \mid 1 \le i \le 24 \rangle$.

(2.10) For $w = w_{j_1} w_{j_2} .... w_{j_m}$, $1 \le j_1, ..., j_m \le 24$, let

$$\widetilde{w} = \widetilde{w}_{j_1}\ \widetilde{w}_{j_2} \ldots\ldots \widetilde{w}_{j_m}$$

Then $\quad w x_i(\alpha)\, w^{-1} = x_{\widetilde{w}(i)}(\alpha)$.

The map $w \to \widetilde{w}$ is a homomorphism from W onto $\widetilde{W}$. If $\widetilde{u} = 1$, Then $u \in C_W(S)$. It follows from (2.9) that $u = 1$. Therefore,

(2.11) The map $w \to \widetilde{w}$ is an isomorphism from W onto $\widetilde{W}$.

(2.12) For each $w \in W$ let $S_w = \langle S_i \mid 1 \le i \le 24;\ \widetilde{w}(i) < 0 \rangle$ and

$S'_w = \langle S_i \mid 1 \le i \le 24;\ \widetilde{w}(i) > 0 \rangle$. Then $S \cap S^w = S'_w$ and S_w is a complement to $S \cap S^w$ in S.

(2.13) Each element $g \in F$ can be expressed uniquely in the form

$g = sws'$ with $s \in S$, $w \in W$, and $s' \in S_w$.

For any root $r = ar_1 + br_2 + cr_5 + dr_{10}$ let

$$\bar{r} = (2b_1 + ar_2 + dr_5 + 2cr_{10})/\lambda(r)$$

Then $r \in \sum^+$ implies $\bar{r} \in \sum^+$. For each i, $1 \le i \le 24$ let $\bar{i}$ be defined by $r_{\bar{i}} = \bar{r}_i$. The values $\bar{i}$ are included in table 1.

(2.14) There exists an automorohism Φ of F such that

$\Phi(x_i(\alpha)) = x_{\bar{i}}(\alpha)$ for all $\alpha \in \Gamma$ and $\Phi(w_i) = w_{\bar{i}}$ for all i, $1 \le i \le 24$. This is the graph automorphism of F, used by Ree to obtain the Twisted Chevalley groups. The extension of F by the automorphism ϕ, which is of order 2, is denoted by $\hat{F}_4(2)$ and is called the Ree-Extension of F.

Table - 2

values $(i, j : m)$ for which $[x_i(1), x_j(1)] = x_m(1)$.			
(1,10 : 11)	(1, 12 : 13)	(1, 14 : 15)	(1, 16 : 17)
(2, 5 : 6)	(2, 8 : 9)	(2, 19 : 20)	(2, 22 : 23)
(3, 10 : 12)	(3, 11 : 13)	(3, 14 : 16)	(3, 15 : 17)
(4, 5 : 8)	(4, 6 : 9)	(4, 19 : 22)	(4, 20 : 23)
(5, 18 : 19)	(5, 23 : 24)	(6, 18 : 20)	(6, 22, 24)
(7, 10 : 14)	(7, 11 : 15)	(7, 12 : 16)	(7, 13 : 17)
(8, 18 : 22)	(8, 20 : 24)	(9, 18 : 23)	(9, 19 : 24)
(10, 17 : 21)	(11, 16 : 21)	(12, 15 : 21)	(13, 14 : 21)
values $(i, j : m, n)$ for which $[x_i(1), x_j(1)] = x_m(1)\, x_n(1)$.			
(1, 2 : 3, 4)	(1, 6 : 7, 8)	(1, 20 : 21, 22)	
(3, 5 : 7, 9)	(3, 19 : 21, 23)	(7, 18 : 21, 24)	
(2, 11 : 12, 18)	(2, 15 : 16, 24)	(4, 10 : 13, 18)	
(4, 14 : 17, 24)	(5, 12 : 14, 20)	(5, 13 : 15, 22)	
(6, 11 : 14, 19)	(6, 13 : 16, 23)	(8, 10 : 15, 19)	
(8, 12 : 17, 23)	(9, 10 : 16, 20)	(9, 11 : 17, 22)	

(2.15). [4]. Let $W_1 = \langle w_1, w_2, w_5 \rangle$ and $C_1 = \{sws' \mid s \in S,\ s' \in Sw,\ w \in W_1\}$.

Then $C_1 = C_F(x_{21}(1))$ (ii) $Z(C_1) = S_{21}$ (iii) $C_1 \cong O_2(C_1)$. Sp(6, 2)

(2.16). [4]. Let $W_2 = \langle w_1, w_2, w_{10} \rangle$ and $C_2 = \{sws' \mid s \in S,\ w \in W_2,\ s' \in S_w\}$

Then $C_2 = C_F(x_{24}(1))$; $Z(C_2) = S_{24}$ and $C_2 \cong O_2(C_2)$. Sp(6, 2)

(2.17). [4]. Let $W_3 = \langle w_1, w_2 \rangle$ and $C_3 = \{sws' \mid s \in S,\ w \in W_3,\ s' \in S_w\}$

Then $C_3 = C_F(x_{21}(1)\, x_{24}(1))$ and $C_3 \cong O_2(C_3)$, Σ_6, where Σ_6 is the symmetric roup on six letters.

(2.18). [3]. Let $R = \langle S_i \mid i > 0, i \neq 1, 2, 6, 11 \rangle$ and

$$C_4 = \{sws' \mid s \in R,\ w \in \langle w_5, w_{10} \rangle,\ s' \in S_w\}.$$

Then $C_4 = C(x_{16}(1)\, x_{22}(1))$

From now onwards, since the only involution in any root subgroup S_i of $F_4(2)$ is $x_i(1)$, we will write x_i for $x_i(1)$ except where there is ambiguity.

(2.19) We define $D_i = \Pi S_j, (j \neq i)$;

Then it is easy to check that $D_{10} = C_S(x_{17})$ and $D_5 = C_S(x_{23})$.

We write $D_{10} = M$ to correspond to [4]. Every involution of S whose centralizer in S is of order at least 2^{23} lies either in $Z(M) = S_{17}S_{21}S_{24}$ or in $Z(D_5) = S_{21}S_{23}S_{24}$. Hence M and D_5 are the only subgroups of S of order 2^{23} with centers of order 2^3.

(2.20). ([3]; 4.1). Let ν be an automorphism of S. Then

$$\{\nu(S_{21}), \nu(S_{24})\} = \{S_{21}, S_{24}\}.$$

(2.21). ([3] ; 4.6). The group C_1 admits an automorphism θ such that,

$$\theta(x_1) = x_1x_{22}(\in), \theta(x_3) = x_3x_{23}(\in), \theta(x_7) = x_7x_{24}(\in)$$

3. SLOW $_2$-STRUCTURE OF G AND THE $N_G(M)$.

Theroem A

Let G be a finite group with a non-central involution y_1 such that $C = C_G(y_1)$ is isomorphic to $C_{F_4(2)}(x_{21}) = C_1$. Identify C with C_1. Then the following hold:

(i) $N_G(S)$ is a Sylow$_2$-subgroup of G and $[N_G(S) : S] = 2$.

(ii) $N_G(M)/M \cong \Sigma_3$ and $N_G(M) \subseteq C_G(x_{24})$.

(iii) S is its own normalizer in $N_G(M)$. In particular, no two of the involutions x_{21}, x_{24} and $x_{21}x_{24}$ are conjugate in $N_G(M)$.

Lemma (3.1)

$N_G(S)$ is a 2-group and $[N_G(S) : S] = 2$.

Proof

Let $g \in N_G(S)$, then due to (2.23), $\{x_{21}^g, x_{24}^g\} = \{x_{21}, x_{24}\}$:

Thus $g^2 \in C$. But $N_C(S)$ is S due to (2.9). This implies $g^2 \in S$. This proves the Lemma.

Lemma (3.2)

For any $g \in N_G(S) \backslash S$, $M^g = D_5$, $x_{21}^g = x_{24}$ and $Z(N_G(S)) = \langle x_{21}x_{24} \rangle$

Proof

Let $U = N_G(S)$ and suppose $g \in U \backslash S$. Now M and D_5 are the only subgroups of S of order 2^{23} with centres of order 8. Thus $\{M^g, D_5^g\} = \{M, D_5\}$. Let $M^g = M$. Then $(Z(M))^g = Z(M)$. Now from table 2, $x_{21} = [S, Z(M)]$. Thus $x_{21}^g = [S^g, (Z(M))^g] = [S, Z(M)] = x_{21}$, a contradiction to the choice of g and the fact that S is an S_2-subgroup of C. Hence $M^g = D_5$ and $x_{21}^g = [S, Z(D_5)] = x_{24}$ by table 2. Now $x_{21}x_{24}$ is the only nontrivial element of Z(S) centralized by g. Hence the Lemma.

Lemma (3.3)

$N_G(S)$ is an S_2-subgroup of G.

Proof

Let $U = N_G(S)$. If $T \in Sylow_2(G)$ and U is properly contained in T, then $N_T(U)$ contains U properly. But $Z(U) = \langle x_{21}x_{24} \rangle$ and $Z_2(U) = Z(S)$ by (3.2). Thus $S = C_U(Z_2(U))$ is characteristic in U since $Z_2(U)$ is so. This implies S is normal in $N_T(U)$. Hence $N_T(U) = U$, a contradiction. The Lemma is thus proved.

Lemma (3.4)

$x_{21}x_{24}$ is not conjugate to any of x_{21} and x_{24} is G.

Proof

The centralizer of x_{21} has S as an S_2-subgroup whereas U, as defined above is an S_2-subgroup of $C_G(x_{21}x_{24})$ and $S \neq U$. Also x_{21} is conjugate to x_{24} in G by (3.2). This proves the Lemma.

Lemma (3.5)

$N_G(M) \cap C = S$.

Proof

This is directly verified by the structure of C and the tables.

Lemma (3.6)

S is an S_2-subgroup of $N_G(M)$.

Proof

If possible, let T be an S_2-subgroup of $N_G(M)$ such that S is properly contained in T. Then for $t \in T/S$, $M^t = D_5$ by (3.2). This contradicts the fact that $t \in N_G(M)$. Hence the Lemma.

Lemma (3.7)

$N_G(M)$ is not 2-closed (*i.e.*, S is not normal in $N_G(M)$).

Proof

From table 1 and (2.9), we find that w_5 normalizes D_5 but does not mormalize S which is an S_2-subgroup of $N(D_5)$ in G by similar arguments as in (3.6). Thus $N_G(D_5)$ is not 2-closed. Since M is conjugate to D_5 in G, the Lemma is evident.

Lemma (3.8)

$N_G(M)/M \cong \Sigma_3$, the symmetric group on three letters.

Proof

From (3.5) we have $S = C(x_{21})$ in $N_G(M)$. Thus $[N_G(M) : S] = |ccl_{N_G(M)}(x_{21})| \leq 6$, since Z(M) has 7 involutions including $x_{21}x_{24}$. Now $[N_G(M) : S]$ ust be odd due to (3.6). Also $N_G(M) \subseteq \text{Aut}\,(Z(M)) \cong GL(3, 2)$ whose order is $2^3.3.7$. Thus $[N_G(M) : S] = 3$ or 1. Due to (3.7) we are left with $[N_G(M) : S] = 3$. This forces $|N_G(M)/M| = 6$. Again due to (3.7), $N_G(M)/M$ cannot be abelian. Hence the only possibility is $N_G(M)/M$ is isomorphic to Σ_3. Hence the Lemma.

Lemma (3.9)

$N_G(M) \subseteq C_G(x_{24})$.

Proof

From tables 1 and 2 we find that $N_C(D_5) = \langle S, w_5 \rangle \cong \Sigma_3.D_5$. So $N_C(D_5)/D_5 \cong N_G(M)/M \cong N_G(D_5)/D_5$ due to (3.2), which, in turn, implies that $N_G(M)/M \cong N_{C_G(x_{24})}(M)/M$ and are in fact equal. The Lemma is proved.

Lemma (3.10)

S is its own normalizer in $N_G(M)$. In particular, no two of the involutions x_{21}, x_{24}, and $x_{21}x_{24}$ are conjugate in $N_G(M)$.

Proof

From (3.7) S is not normal in $N_G(M)$ and from (3.8) $[N_G(M):S]=3$. Hence N(S) in $N_G(M)$ must be S. Now, by Burnsides Theorem, no two elements of Z(S) are conjugate in $N_G(M)$. But Z(S) $= \{x_{21}, x_{24}, x_{21}x_{24}, I\}$. This proves the Lemma.

We have completed the proof of Theorem A: (3.1) and (3.3) prove (i); (2.8) and (2.9) prove (ii) ; while (2.10) takes care of (iii).

REFERENCES

1. Carter, R. W. Simple Groups of Lie Type. John Wiley & Sons. 1972.
2. Gorenstein, D. Finite Groups, Harper & Row, New York, 1968.
3. Guterman, M.A. A Characterization of $F_4(2^n)$, *J. Algebra.* **20**(1972), 1–23.
4. Husnine, S.M. A Characterization of Chevalley group $F_4(2)$ *J. Algebra* **44**(1977), 539–549.
5. Ree, R. A family of Simple Groups associated with the Simple Lie Algebra of Type (F_4), *Amer. J. Math.* **83**(1961), 401–420.

Punjab University
Journal of Mathematics
Vol. XXVI, (1993), pp 111 – 118

MONOTONICITY PRESERVING PIECEWISE RATIONAL INTERPOLATION

Muhammad Sarfraz

Department of Mathematics,
Punjab University,
Quaid-i-Azam Campus,
Lahore-54590, Pakistan.

ABSTRACT

A C^1 piecewise Rational cubic interpolant is utilized to solve the problem of shape preserving interpolation. It is shown that the interpolation method can be applied to monotonic sets of data. The scheme generalizes the monotoniticity preserving results of Delbourgo and Gregory [Delbourgo and Gregory'85].

Keywords. Rational, interpolation, plane curve, shape preserving, monotonic.

1. INTRODUCTION

A number of authors have considered the problem of shape preserving interpolation. For brevity the reader being refered to [Goodman'88, Gregory'86]. The methods in [Gregory'86] are global and non-parametric whereas [Goodman'88] discusses local and parametric shape preserving methods.

This paper uses a piecewise rational cubic interpolant to solve the problem of shape preserving interpolation. The results derived here are actually the extensions of the monotonicity results of Delbourgo and Gregory [Delbourgo and Gregory'85] who developed a C^1 shape preserving interpolation scheme for scalar curves using the same piecewise rational functions. They derived the constraints, on the shape parameters occuring in the rational function under discussion, to make the interpolant preserve the monotonic shape of the data.

This paper begins with some preliminaries about the rational cubic interpolant. The constraints with monotonic data are derived in Section 3. These constraints are dependent on the tangent vectors. The description of the tangent vectors, which are consistent and dependent on the given data, is made in Section 4. The monotonicity preserving results are explained with examples in Section 5.

2. THE RATIONAL CUBIC INTERPOLANT

Let $F_i \in \mathbb{R}^2$, $i = 0, \ldots, n$ be a given set of data points, where $t_0 < t_1 < \ldots < t_n$. We consider the C^1 piecewise rational cubic interpolant

(1) $$p(t) = \frac{(1-\theta)^3 F_i + \theta(1-\theta)^2 (r_i F_i + h_i D_i) + \theta^2 (1-\theta)(r_i F_{i+1} - h_i D_{i+1}) + \theta^3 F_{i+1}}{1 + (r_i - 3)\,\theta(1-\theta)}$$

$$\theta(t) = (t - t_i)/h_i, \quad h_i = t_{i+1} - t_i.$$

We will use this to generate an interpolatory planar curve which preserves the shape of the monotonic data. Let

(2) $$\begin{cases} p(t) = (p_1(t), p_2(t)), \\ F_i = (x_i, y_i), \\ D_i = (D_i^x, D_i^y), \\ \Delta_i = (\Delta_i^x, \Delta_i^y), \end{cases}$$

where

$$\Delta_i^x = \frac{(x_{i+1} - x_i)}{h_i}, \quad \Delta_i^y = \frac{(y_{i+1} - y_i)}{h_i},$$

and D_i denote the tangent vector to the curve at the knot t_i. It can be noted that $p(t)$ interpolates the points F_i and the tangent vectors D_i at the knots t_i.

The parameter r_i is to be chosen such that $r_i \geq -1$, which ensures a strictly positive denominator in the rational cubic. For our purposes r_i will be chosen to ensure that the interpolant preserves the shape of the data. This choice requires the knowledge of $p^{(1)}(t)$ which is as follows:

(3) $$p^{(1)}(t) = \frac{(1-\theta)^4 D_i + \alpha_{1,i}\theta(1-\theta)^3 + \alpha_{2,i}\theta^2(1-\theta)^2 + \alpha_{3,i}\theta^3(1-\theta) + D_{i+1}\theta^4}{\{1 + (r_i - 3)\,\theta(1-\theta)\}^2}$$

where

$$(5)\qquad \begin{cases} \alpha_{1,i} = 2(r_i\Delta_i - D_{i+1}), \\ \alpha_{2,i} = (r_i^2 + 3)\,\Delta_i - r_i(D_i + D_{i+1}), \\ \alpha_{3,i} = 2(r_i\Delta_i - D_i), \end{cases}$$

and we denote

$$(6)\qquad \alpha_{j,i} = (\alpha_{j,i}^x, \alpha_{j,i}^y).$$

3. INTERPOLATION OF MONOTONIC DATA

Let us assume for simplicity that

$$(7)\qquad \Delta_i^x \neq 0,\ i = 0,\dots,n-1,$$

and that the data is monotonic increasing and arises from a function. Then we must have

$$(8)\qquad \frac{\Delta_i^y}{\Delta_i^x} \geq 0,\ i = 0,\dots,n-1,$$

i.e. Δ_i^x and Δ_i^y are of the same sign. (The case of monotonic decreasing set of data can be treated in a similar manner when the inequalities are reversed.) The necessary conditions for the interpolant $p(t)$ to be monotonic are, then, the following:

$$(9)\qquad \frac{D_i^y}{D_i^x} \geq 0,\ i = 0,\dots,n-1,$$

i.e. D_i^x and D_i^y are of the same sign. We also note that D_i^x and D_i^y must have the same sign as Δ_i^x and Δ_i^y respectively. Thus we have the following:

Lemma 1

The monotonicity conditions imply that

$$(10)\qquad \begin{cases} \Delta_i^x\,\Delta_i^y,\ D_i^y\,D_i^x \geq 0, \\ \Delta_i^x\,D_i^y,\ \Delta_i^y\,D_i^x \geq 0, \\ \Delta_i^x\,D_{i+1}^y,\ \Delta_i^y\,D_{i+1}^x \geq 0, \end{cases}$$

for $i = 0, \dots, n-1$.

Remark 2

Let

$$(11)\quad \begin{cases} \beta_{1,i} = D_i^x \Delta_i^y + D_i^y \Delta_i^x , \\ \beta_{2,i} = \Delta_i^x D_{i+1}^y + \Delta_i^y D_{i+1}^x , \\ \beta_{3,i} = D_i^x D_{i+1}^y + D_i^y D_{i+1}^x , \\ \beta_{4,i} = D_i^x D_i^y . \end{cases}$$

Then it follows from (10) that

$$(12)\quad \beta_{j,i} \geq 0 , \; j = 1, \dots, 4, \; i = 0, \dots, n-1.$$

Now, $p(t)$ is monotonic increasing if and only if

$$(13)\quad \frac{p_2^{(1)}(t)}{p_1^{(1)}(t)} \geq 0 , \quad \forall\, t \in [t_0, t_n].$$

i.e. $p_1^{(1)}(t)$ and $p_2^{(1)}(t)$ are of the same sign. Thus (13) can be equivalently written as

$$(14)\quad p_1^{(1)}(t)\, p_2^{(1)}(t) \geq 0 , \quad \forall\, t \in [t_0, t_n].$$

After some simplifications, using (2) – (6), it can be shown that for $t \in [t_i, t_{i+1}]$,

$$(15)\quad p_1^{(1)}(t)\, p_2^{(1)}(t) = \frac{\sum_{j=1}^{9} \gamma_{j,i}\, (1-\theta)^{9-j}\, \theta^{j-1}}{\{1-(r_i-3)r_i\, \theta(1-\theta)\}^4},$$

where

$$\gamma_{1,i} = D_i^x D_i^y ,$$

$$\gamma_{2,i} = D_i^x \alpha_{1,i}^y + \alpha_{1,i}^x D_i^y = 2r_i\, \beta_{1,i} - 2\beta_{3,i} ,$$

$$\gamma_{3,i} = D_i^x \alpha_{2,i}^y + \alpha_{1,i}^x \alpha_{1,i}^y + \alpha_{2,i}^x D_i^y ,$$

$$\gamma_{4,i} = D_i^x \alpha_{3,i}^y + \alpha_{1,i}^x \alpha_{2,i}^y + \alpha_{2,i}^x \alpha_{1,i}^y + \alpha_{3,i}^x D_i^y$$

(16) $\gamma_{5,i} = D_i^x D_{i+1}^y + \alpha_{1,i}^x \alpha_{3,i}^y + \alpha_{2,i}^x \alpha_{2,i}^y + \alpha_{3,i}^x \alpha_{1,i}^y + D_{i+1}^x D_i^y$

$\gamma_{6,i} = \alpha_{1,i}^x D_{i+1}^y + \alpha_{2,i}^x \alpha_{3,i}^y + \alpha_{3,i}^x \alpha_{2,i}^y + D_{i+1}^x \alpha_{1,i}^y$

$\gamma_{7,i} = \alpha_{2,i}^x D_{i+1}^y + \alpha_{3,i}^x \alpha_{3,i}^y + D_{i+1}^x \alpha_{2,i}^x$

$\gamma_{8,i} = \alpha_{3,i}^x D_{i+1}^y + D_{i+1}^x \alpha_{3,i}^y = 2r_i \beta_{2,i} - 2\beta_{3,i}$,

$\gamma_{9,i} = D_{i+1}^x D_{i+1}^y$

The conditions

(17) $D_j^x D_j^y \geq 0, j = i, i+1$

are necessary for the interpolant to be monotonic increasing on $[t_i,t_{i+1}]$ and, assuming these necessary conditions, sufficient conditions are

(18) $\gamma_{j,i} \geq 0$, $j = 2, \ldots, 8$.

It should be noted that if $\Delta_i^y = 0$, then $D_i^y = D_{i+1}^y = 0$ and hence $\beta_{1,i} = \beta_{2,i} = 0$. Moreover, $p_2(t) = y_i$, $t_i \leq t \leq t_{i+1}$. Therefore $p(t)$ is constant on $[t_i,t_{i+1}]$.

If $\Delta_i^y \neq 0$, then a sufficient condition for (18) is

(19) $$r_i \geq max \left\{ \frac{D_i^x + D_{i+1}^x}{\Delta_i^x}, \frac{D_i^y + D_{i+1}^y}{\Delta_i^y} \right\}.$$

Moreover, since

(20) $$max \left\{ \frac{\beta_{3,i}}{\beta_{3,i}}, \frac{\beta_{1,i}}{\beta_{2,i}} \right\} \geq max \left\{ \frac{D_i^x + D_{i+1}^x}{\Delta_i^x}, \frac{D_i^y + D_{i+1}^y}{\Delta_i^y} \right\},$$

the choice,

(21) $$r_i = \frac{\beta_{3,i}(\beta_{1,i} + \beta_{2,i})}{\beta_{1,i} \beta_{2,i}},$$

satisfies (19) and provides nice graphical results.

Remarks 3

As the denominator in (1) has the form

$$(1-\theta)^3 + r_i\theta(1-\theta)^2 + r_i\theta^2(1-\theta) + \theta^3,$$

therefore (1) can be written as

$$p_i(t_i) = R_0(\theta;r_i)\, F_i + R_1(\theta;r_i)\, V_i + R_2(\theta;r_i)\, W_i + R_3(\theta;r_i)\, F_{i+1},$$

where

$$V_i = F_i + h_iD_i/r_i\,, \quad W_i = F_{i+1} - h_iD_{i+1}/r_i$$

and $R_j(\theta;\, r_i)$, $j = 0, 1, 2, 3$, are appropriately defined rational functions with

$$\sum_{j=1}^{3} R_j(\theta;\, r_i) = 1.$$

Now the scalar case can be dealt with as a consequence of the identity

$$t \equiv R_0(\theta;r_i)t_i + R_1(\theta;r_i)\,(t_i + h_i/r_i) + R_2(\theta;r_i)\,(t_{i+1} - h_i/r_i) + R_3(\theta;r_i)\, t_{i+1},$$

In fact this can be considered as an application of interpolation scheme $(t, p(t))$ in $\mathbb{R}^2$ to the values $(t_i, F_i) \in \mathbb{R}^2$ and derivatives $(1, D_i) \in \mathbb{R}^2$, $i = 0, \ldots., n$. It can also be noted that $\Delta_i = (1, \Delta_i)$. Therefore the monotonicity constraints, in this case, are

$$r_i \geq max\left\{2, \frac{D_{i+1} + D_i}{\Delta_i}\right\}, \tag{23}$$

which are san· as in [Delbourgo and Gregory'85].

4. CHOIC· OF TANGENT VECTORS

In most applications, the tangent vectors D_i will not be given and hence must be determined from the data $F_i \in \mathbb{R}^2$, $i = 0, \ldots, n$. We describe here the geometric mean choices of tangent vectors for our plane curves which satisfy the shape preserving conditions. The geometric mean choice of tangent vectors is

$$D_i = \left((\Delta_{i-1}^x)^{\lambda_i}\, (\Delta_i^x)^{1-\lambda_i},\, (\Delta_{i-1}^y)^{\lambda_i}\, (\Delta_i^y)^{1-\lambda_i}\right),\ i = 1, \ldots, n-1,$$

with the end conditions

$$D_0 = \left((\Delta_0^x)^{\lambda_0}\, (\Delta_{2,0}^x)^{1-\lambda_0},\, (\Delta_0^y)^{\lambda_0}\, (\Delta_{2,0}^y)^{1-\lambda_0}\right),$$

$$D_n = ((\Delta^x_{n-1})^{\lambda_n} (\Delta^x_{n,n-2})^{1-\lambda_n}, (\Delta^y_{n-1})^{\lambda_n} (\Delta^y_{n,n-2})^{1-\lambda_n}).$$

where

$$\lambda_i = \frac{h_i}{(h_{i-1} + h_i)}, \quad i = 1,\ldots, n-1,$$

and $$\lambda_0 = 1 + \frac{h_0}{h_1}, \quad \lambda_n = 1 + \frac{h_{n-1}}{h_{n-2}},$$

$$\Delta_{2,0} = \frac{F_2 - F_0}{t_2 - t_0}, \quad \Delta_{n,n-2} = \frac{F_n - F_{n-2}}{t_n - t_{n-2}}.$$

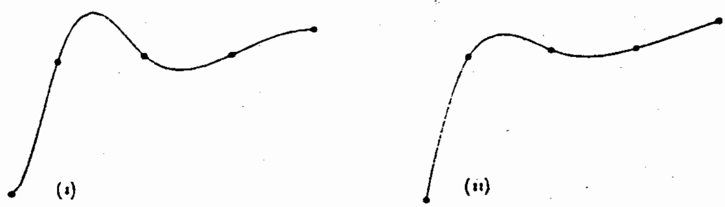
These geometric mean approximations are suitable for monotonic data since they satisfy the necessary conditions for monotonicity and produce pleasing graphical results.

5. EXAMPLES

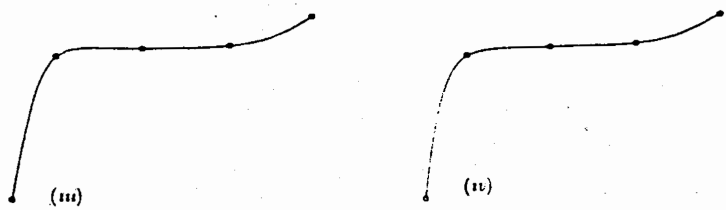
The Figure demonstrates the monotonicity preserving results corresponding to the geometric mean derivative values; the first and the second curves in this figure represent the scalar and parametric cubic spline interpolation respecitvely whereas the third and the fourth curves are respectively scalar and parametric montonicity preserving interpolations. The parametrization used here is the chord length parametrization. Some other parametrization may also be used.

6. CONCLUDING REMARKS

C^1 rational cubic Hermite interpolant with one shape parameter has been utilized to obtain a C^1 monotonicity preserving plane curve method. Data dependent shape constraints are derived on the shape parameters to assure the shape preservation of the data. Choice of the tangent vectors, which are consistent and dependent on the data, has also been made. This scheme can also be implemented in the scalar case.



(Figure)



(Figure)

REFERENCES

1. Goodman T.N.T, Shape preserving interpolation by parametric rational cubic splines, *Proc. Int. Conf. on Numerical Mathematics*, Int. series Num. Math. 86, Birkhauser Verlag, Basel, (1988).
2. Goodman T.N.T and Unsworth K., Shape preserving interpolation by parametrically defined curves, *SIAM J. Num. Anal.* 25, 1–13 (1988).
3. Gregory J. A., Shape preserving spline interpolation, *Comput. Aided Design*. 18, 53–57 (1986).
4. Delbougo R. and Greogry J. A., Shape preserving piecewise rational interpolation. *SIAM J. Sci. Stat. Comput.* 6, 967–976 (1985).

Published by Chairman Department of Mathematics for the University of the Punjab, Lahore-Pakistan.

Composed at
COMBINE GRAPHICS
Lawyer's Park, 1-Turner Road,
Lahore. Phone: 7244043

Printed at
AL-HIJAZ PRINTERS
Darbar Market, Lahore.
Phone: 7238009

CONTENTS