On Brown's and Newton's methods with convexity hypotheses

J.P. Milaszewicz
Instituto Argentino de Matemática
Saavedra 15
1083 Buenos Aires, Argentina
Email: pmil@maths.tcd.ie

Abstract

It is proved that with convenient and easily verified hypotheses in the context of the monotone Newton theorem, analytic Brown iterations converge faster than Newton iterations. As a consequence, a similar result holds for discretised iterations. The same conclusions are established for the corresponding Fourier iterates[1].

1 Introduction

For a twice continuously differentiable function $F : D \subset R^n \to R^n$ consider the equation

$$F(x) = 0 \quad , \tag{1.1}$$

where $F(x) := (f_i(x)), 1 \leq i \leq n$. It will be assumed that D is an open convex set.

In order to find an approximate solution of (1.1) many methods related to Newton's method have been analysed when $n = 1$ (See for instance [19]). For $n > 1$, Brown proposed a recursion that combines the one dimensional Newton method with a Gauss-Seidel-like extension of the Gaussian elimination process to the nonlinear case; in fact, Brown's method turns to be Newton's method if $n = 1$, whereas it becomes Gaussian elimination if F is affine linear. This also holds for the discretised Brown method, where incremental quotients are the substitute for partial derivatives. Quadratic convergence for Brown's analytic method was

[1] Mathematics Subject Classification: 65H10

established in [5], but its early implementations showed a cost of $O(n^4)$ algebraic operations per iteration. Later on, Gay developed another algorithm to implement the method, following a suggestion by Brent, which reduced that figure to $O(n^3)$ (See [9] and [4]); at this point it is important to recall that for a function F, with full Jacobian matrix F', the a priori operational count yields that the discretised Brown method requires only $\frac{n^2+3n}{2}$ function evaluations at each iterative step, which is significantly less than the n^2+n function evaluations needed in Newton's discretised method; this makes the former a better option than the latter, as far as computational work per iteration is concerned; as a matter of fact, the number of function evaluations turns out to be the same for both discretised methods when F is almost linear, i.e. $F = A + \Phi$, where A is linear and Φ is a diagonal mapping; for instance, such situation arises in the discretisation of $\Delta u = e^u$, whereas the non sparse situation arises in the discretisation of nonlinear integral equations, such as Chandrasekhar's equation. It is thus particularly interesting to be able to compare the convergence properties of both methods. Frommer has pioneered such analysis in the context of the Monotone Newton Theorem (MNT) and has proved a monotone Brown theorem (see [6]), and that Brown's analytic method converges componentwise at least as fast as Newton's analytic method (see [7]); he also obtained partial results regarding the discretised versions of both methods (see [7]).

One of the objectives of the present paper is to prove that in the context of the MNT, when F is almost linear, discretised Brown iterates converge componentwise at least as fast as Newton's iterates, if the Brown differential increments are smaller or equal than their Newton counterparts; convergence becomes strictly faster with easily checked additional assumptions, whence the increments for the Brown iterations may vary in convenient neighbourhoods of the increments employed for the Newton iterations, while still yielding faster convergence. More interestingly, for a general F in the MNT context, the same assumptions also imply that convergence for the analytic Brown iterations is faster than for the analytic Newton iterations; as a consequence, similar conclusions do hold for the discretised iterations. On the other hand, if those assumptions are not fulfilled, discretised Newton iterates may converge faster than discretised Brown iterates, thus barring expectations of a general comparison theorem for discretised iterations similar to the one valid for analytic iterations. These results complement, and in some cases improve, those obtained by Frommer in [6] and [7]; they include the corresponding results for the Brown-Fourier iterates, whose role with respect to Brown iterations is analogous to the one played by the Newton-Fourier iterates with respect to Newton iterations.

To achieve the stated aims, the focus has been shifted back to the original algorithm implementing Brown's method; it will become apparent that it is very useful for analysis. The next section reviews the MNT as well as some of its implications, while in the third one some features of the two basic algorithms that implement Brown's method are examined; as a consequence, the results in [6] are improved and extended. The fourth section draws upon the previous two and contains the main results. These are numerically illustrated in the following section with both a one dimensional version of the differential equation mentioned above and Chandrasekhar's equation. In a final remarks section, developments are pointed out that could possibly be built around the present results.

2 The Monotone Newton Theorem

It will be assumed throughout the section that x^0 and y^0 are given in D satisfying $x^0 \le y^0$ (resp. $x^0 < y^0$) , i.e. $x_i^0 \le y_i^0$ (resp. $x_i^0 < y_i^0$), $1 \le i \le n$, and

$$F(x^0) \le 0 \le F(y^0) \quad .$$

For convex F, any first analytic Newton iterate produces such an y^0 whereas in many situations it is thereafter easy to obtain x^0 as well ([15]), as for instance when $F = A + \Phi$, with A a nonsingular M-matrix and Φ a convex (diagonal) mapping. We denote

$$\langle x^0, y^0 \rangle := \{y \in R^n : x^0 \le y \le y^0\} \subset D \quad .$$

F is supposed to be order convex, i.e. if x and y are in D and $x \le y$, then $F(\lambda x + (1-\lambda)y) \le \lambda F(x) + (1-\lambda)F(y)$ for all $0 \le \lambda \le 1$. Recall that the Jacobian matrix function F' is isotone, if

$$x \le y \quad \textit{implies} \quad F'(x) \le F'(y) \quad ,$$

and that if F' is isotone, then F is order convex.

Suppose that mappings $P_k, Q_k : \langle x^0, y^0 \rangle \to L(R^n)$ are given such that $P_k(y)$ and $Q_k(y)$ are nonnegative subinverses of $F'(y)$ whenever $y \in \langle x^0, y^0 \rangle$, and define

$$y^{k+1} := y^k - P_k(y^k)F(y^k) \quad , \quad k = 0, 1, \ldots \tag{2.1}$$

$$x^{k+1} := x^k - Q_k(y^k)F(x^k) \quad , \quad k = 0, 1, \ldots \tag{2.2}$$

Theorem 2.1 *The sequence (2.1) is well defined and it satisfies*

$$y^k \downarrow y^* \in \langle x^0, y^0 \rangle \quad .$$

If F' is isotone in $\langle x^0, y^0 \rangle$, then $x^k \uparrow x^* \le y^*$, as $k \to \infty$. The iterates satisfy $F(x^k) \le 0 \le F(y^k)$. If moreover, there exist nonsingular matrices P and Q such that

$$0 \le P \le P^k(y^k) \quad \textit{and} \quad 0 \le Q \le Q_k(x^k)$$

then

$$x^* = y^* \quad and \quad F(y^*) = 0$$

Finally, if it is also assumed that $F'(y)$ is nonsingular and $(F'(y))^{-1}$ is nonnegative whenever $y \in \langle x^0, y^0 \rangle$, and if it is set

$$Q_k(y^k) := P_k(y^k) := F'(y^k)^{-1} \quad , \tag{2.3}$$

then there exists a constant c such that

$$\left\| y^{k+1} - x^{k+1} \right\| \le c \left\| y^k - x^k \right\|^2 \quad , \quad k = 0, 1, \ldots .$$

Proof. See 13.3 in [15].

Remark. For $n = 1$ the iterates (2.2) with the choice (2.3) had been examined by Fourier; see [16] for a fairly complete account in this case. In the present context they have apparently been first analysed by Baluev ([2]) and then more thoroughly and leading to the theorem above in [15] (See also [14]). Theorem 2.1 will be referred to as the Monotone Newton Theorem (MNT), although this usually refers to the theorem with the choice (2.3). The hypotheses in this theorem are often satisfied when (1.1) is the system resulting from the discretisation of mildly nonlinear elliptic problems, nonlinear boundary value problems, and integral equations of Hammerstein type as well as Chandrasekhar's equation (see [15] and [1]).

It will be assumed throughout the section that $F'(y)$ is a nonsingular M-matrix, for each $y \in D$ (See [3] or [20] for standard properties of M-matrices).

Corollary 2.2 *Suppose that $F'(y)$ is irreducible whenever $y \in D$. Then, with the choice (2.3) it follows that*

(i) If $F(y^k) \neq 0$, then $y^{k+1} < y^k$.

(ii) If $F(x^k) \neq 0$, then $x^k < x^{k+1}$.

Consider $h_0 \in R,\ h_0 > 0$, such that

$$y + h_0\ e \in D \quad whenever \quad y \in \langle x^0, y^0 \rangle , \qquad e := (1, \ldots, 1)^t \quad .$$

Define, for $0 < h \le h_0$, $\delta F(y, h)$ in $L(R^n)$ by

$$\delta F(y,h)_{i,j} := \delta_j f_i(y,h) := \frac{1}{h}[f_i(y + h\ e^j) - f_i(y)] \quad , \quad 1 \le i, j \le n \quad ,$$

where e^j denotes the j-th unit coordinate vector.

Lemma 2.3 *The following hold:*

(i) $F'(y) \le \delta F(y,h) \le F'(y + h\ e) \quad for \quad y \in \langle x^0, y^0 \rangle$

(ii) $\delta F(y,h)$ is a nonsingular M-matrix.
If F' is isotone, then
(iii) $\delta F(x,h) \leq \delta F(y,h) \quad \forall \quad x,y \in \langle x^0, y^0 \rangle \quad , \quad x \leq y \quad .$
(iv) If $0 < h \leq h'$, then $\delta F(y,h) \leq \delta F(y,h') \quad \forall \quad y \in \langle x^0, y^0 \rangle$.
(v) If $F'(y^0)$ is irreducible, then there exists $h_0' > 0$ such that if $0 < h \leq h_0'$ then $\delta F(y,h)$ is irreducible for all $y \in \langle x^0, y^0 \rangle$.

Proof. (i) and (ii) are straightforward, by taking into account the order convexity of F and that $F'(y)$ is a nonsingular M-matrix.
(iii) The conclusion is implied by the isotonocity of F', when applied in

$$\delta_j f_i(y,h) = \int_0^1 \partial_j f_i(y + the^j)dt \quad . \tag{2.4}$$

(iv) Apply the same argument as in (iii).
(v) Since $0 < (F'(y^0))^{-1} \leq (F'(y))^{-1}$, whenever $y \in \langle x^0, y^0 \rangle$, the conclusion easily follows.

Set now in (2.1) and (2.2) for $0 < h_k \leq h_0$,

$$Q_k(y^k) := P_k(y^k) := \delta F(y^k, h_k) \tag{2.5}$$

and

$$y_N^{k+1} := y_N^k - F'(y_N^k)^{-1} \; F(y_N^k) \quad , \tag{2.6}$$

$$x_N^{k+1} := x_N^k - F'(y_N^k)^{-1} \; F(x_N^k) \quad . \tag{2.7}$$

Lemma 2.4 *With the notation in (2.6), (2.7), the following hold, where c_1 and C are convenient constants:*
(i) $x^k \leq x_N^k \leq y^* \leq y_N^k \leq y^k \qquad k = 0, 1, \ldots$
(ii) $\|y^{k+1} - x^{k+1}\| \leq c_1 \left[\|y^k - x^k\| h_k + \|y^k - x^k\|^2 \right] \qquad k = 0, 1, \ldots$
(iii) If for some d, $|h_k| \leq d \|F(y^k)\|$, then $\|y^{k+1} - x^{k+1}\| \leq C \|y^k - x^k\|^2$.

Proof. (i) follows easily from Lemma 2.3; (ii) and (iii) are standard facts (see [12]).
Remark. Note that even with $n = 1$, it is essential for the results above, to deal with forward differences. The iterates (2.2) with the choice of either (2.3) or (2.5) will be referred to as the Newton-Fourier iterates.

3 Brown's Method

While both Newton's analytic and discretised methods are one point stationary methods, Brown's method goes through the equations in (1.1) in a Gauss-Seidel-like manner in order to produce a new approximation to their solution. If, as before, y^0 denotes the starting point, the following algorithm produces one iteration, yielding the first analytic Brown iterate y^1.

Step 1. Set $\overline{y^0} := y^0$, $i := 1$ and $\overline{F}_1(y) := (f_{1,j}(y)) := (f_j(y))$.

Step 2. Consider a first order Taylor development of $f_{i,i}$ centered at $\overline{y^0}$, equate it to 0 and solve for y_i, the resulting identity being $y_i = g_i(y_{i+1}, \ldots, y_n)$.

Step 3. Define the $(i+1)$-th reduced system

$$\overline{F}_{i+1}(y_{i+1}, \ldots, y_n) := (f_{i+1,j}(y_{i+1}, \ldots, y_n)) = 0 \quad , \quad by$$

$$f_{i+1,j}(y_{i+1}, \ldots, y_n) := f_{i,j}(g_i(y_{i+1}, \ldots, y_n), y_{i+1}, \ldots, y_n) \quad ,$$
$$for \quad i+1 \leq j \leq n \quad .$$

Step 4. If $i+1 < n$, set $i := i+1$, $\overline{y^0} := (y^0_{i+1}, \ldots, y^0_n)$, and start over again with step 2.

Step 5. Consider a first order Taylor development of $\overline{F}_n = f_{n,n}$ centered at y^0_n, equate it to 0 and call its solution y^1_n.

Step 6. For $i = n-1$ to 1 define $y^1_i := g_i(y^1_{i+1}, \ldots, y^1_n)$.

Remark. It has been implicitly assumed that (1.1) is such that the coefficients of y_i in steps 2 and 5 do not vanish. If (1.1) has a solution y^*, F' is nonsingular at y^* and y^0 is sufficiently close to y^*, that assumption can be set aside with the introduction of pivoting at step 2 (See [5]). However, in the context of the MNT, no pivoting is needed, as will become apparent later on.

The discretised version of Brown's method is obtained by the substitution of difference quotients for derivatives in the Taylor developments; these difference quotients, as defined in the previous section, employ a fixed $h_k \neq 0$ throughout the k-th iteration; as with Newton's method, it is essential for the results presented here that $h_k > 0$, which will be assumed in the sequel. Notice also that there is another implicit assumption in Step 2, namely that the functions $f_{i,i}$ are defined in a convenient neighborhood of $\overline{y^0}$; for instance, when $i = 1$, the shifted point

$$(y^0_1 - \frac{f_1(y^0)}{\partial_1 f_1(y^0)}, y^0_2, \ldots, y^0_n) \tag{3.1}$$

must belong to D in order for the algorithm to carry on to its next step. It will also become evident that in the context of the MNT, these successive assumptions are automatically satisfied. Accordingly, it will be implicitly assumed, unless otherwise stated, that the algorithm can be carried out and it will be referred to as Algorithm

A. Finally, note that, as remarked in the introduction, if F is an affine function, one step of Brown's method becomes Gaussian elimination, and if $n = 1$, Brown's method coincides with Newton's method.

The early implementations of Algorithm A led to an operational count of $0(n^4)$ operations per iteration. Later on, a better implementation showed that only $0(n^3)$ operations per iteration were required ([9], [4]). The following algorithm corresponds to this later implementation and it will be referred to as Algorithm B; it follows the lines set in [17] to describe both Brown's and Brent's methods (See also [18]), and the notation is as in [6].

Step 0. Set $k = 1$

Step 1. Set $y^{n+1,k-1} = y^{k-1}$

Step 2. Set $R^{1,k} := I, \quad y^{1,k} := y^{n+1,k-1}, \quad i := 1$

Step 3. For $j = i$ to n, calculate

$$z_j^{i,k} := f_i'(y^{i,k}) R^{i,k} e^j$$

Step 4. If $i = n$, set $R^{i+1,k} := R^{i,k}$ and go to Step 6.

Else, for $j = i + 1$ to n, $m_j^{i,k} := -z_j^{i,k}/z_i^{i,k}$ and

$$E^{i,k} := \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & m_{i+1}^{i,k} & \dots & m_n^{i,k} \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

Step 5. $R^{i+1,k} := R^{i,k} E^{ik}$

Step 6. $y^{i+1,k} := y^{i,k} - [f_i(y^{i,k})/z_i^{i,k}] R^{i+1,k} e^i$

Step 7. If $i < n$ then $i := i + 1$ and go to Step 3.

Step 8. Set $y^k := y^{n+1,k}$, $k = k + 1$ and go to Step 1.

The discretised version of this algorithm is obtained by considering in its step 3

$$z_j^{i,k} := [f_i(y^{i,k} + h_k R^{i,k} e^j) - f_i(y^{i,k})]/h_k$$

with h_k fixed throughout the k-th iteration.

Remark. It is not difficult to verify that algorithms A and B produce the same iterates, in both their analytic and discretised versions. The crucial point is that Algorithm B performs as soon as possible the calculations involved in Step 6 of Algorithm A, as Gauss-Jordan elimination does with respect to Gaussian elimination when F is affine linear.

Note also that in step 6 of Algorithm B, $R^{i+1,k}e^i = R^{i,k}e^i$, which could lead to a simplification of this formulation of Brown's method; we chose to keep it as it is because Brent type algorithms, which extend Brown's method, do so by making different choices of $E^{i,k}$, whereby it could happen that $R^{i+1,k}e^i \neq R^{i,k}e^i$ (See [17] and [18]). Notice that for a general F the analytic version requires the calculation of n^2 derivative functions, besides n scalar function evaluations at each step, while the discretised version requires a total of $\frac{n^2+3n}{2}$ scalar function evaluations. Thus, if the determination of the required $\frac{n^2+n}{2}$ supplementary points in the discretised version is significantly less costly than that of the derivatives and its evaluations in the analytic version, then the former is to be preferred when considering the computational work involved in each iteration. In order for this preference to be completely justified, it is necessary to take into account that Brown's discretised version attains the quadratic behavior of the analytic version ([5]) with the choice $h_k \leq C\|F(y^k)\|$ (See [17] and [18]); recall that this choice also ensures quadratic convergence of Newton's discretised iterations.

Lemma 3.1 *In the analytic case, Algorithm A produces y^1 such that*

$$f'_{i,i}(\overline{y^0})(\overline{y^1} - \overline{y^0}) = -f_{i,i}(\overline{y^0}) \quad , \quad 1 \leq i \leq n \quad , \tag{3.2}$$

whereas in the discretised case, (3.2) holds with substitution of discretised gradients for the gradients.

Proof. It easily follows by mathematical induction, because on the one hand

$$y_1^1 = g_1(y_2^1, \ldots, y_n^1) = g_1(\overline{y^1}) \quad ,$$

while on the other $(y_2^1, \ldots, y_n^1)$ is the first Brown iterate when Algorithm A is applied to the reduced system

$$\overline{F}_2(\overline{y}) = 0 \quad ,$$

with starting point $\overline{y^0} = (y_2^0, \ldots, y_n^0)$.

Lemma 3.2 *Suppose one step of Algorithm A has been carried out.*
(i) In the discrete case $\overline{F}_2$ satisfies

$$\partial_j f_{2,i}(\overline{y},h) = \partial_j f_i(g_1(\overline{y}),\overline{y}) - \frac{\delta_j f_1(y^0)}{\delta_1 f_1(y^0)} * \partial_1 f_i(g_1(\overline{y}),\overline{y}) \quad ,$$

$$for \quad 2 \le i \le n \quad , \quad 2 \le j \le n \quad ,$$

while in the analytic case $\delta_j f_1(y^0)$ *and* $\delta_1 f_1(y^0)$ *are replaced by* $\partial_j f_1(y^0)$ *and* $\partial_1 f_1(y^0)$.
(ii) If F *is almost linear, then* $\overline{F}_2$ *is almost linear as well and in the discrete case, it also satisfies*

$$\delta_j f_{2,i}(\overline{y},h) = \delta_j f_i(g_1(\overline{y}),\overline{y}) - \frac{\delta_j f_1(y^0)}{\delta_1 f_1(y^0)} * \delta_1 f_i(g_1(\overline{y}),\overline{y}) \quad .$$

(iii) If F *is order convex, then* $\overline{F}_2$ *is order convex too.*

Proof.(i) By taking into account that in the discrete case

$$g_1(\overline{y}) = y_1^0 - \frac{1}{\delta_1 f_1(y^0)} \left[f_1(y^0) + \sum_{j=2}^{n} \delta_j f_1(y^0)(y_j - y_j^0) \right] \quad ,$$

the first conclusion easily follows.
(ii) The definition of F being almost linear can be applied to obtain the conclusion; it can also be easily obtained by applying (2.4).
(iii) If $\overline{x} \le \overline{y}$ and $0 \le \lambda \le 1$, then

$$\begin{aligned} \overline{F}_2(\lambda\overline{x} + (1-\lambda)\overline{y}) &= F(g_1(\lambda\overline{x} + (1-\lambda)\overline{y}), \lambda\overline{x} + (1-\lambda)\overline{y}) \\ &= F(\lambda g_1(\overline{x}) + (1-\lambda)g_1(\overline{y}), \lambda\overline{x} + (1-\lambda)\overline{y}) \\ &\le \lambda F(g_1(\overline{x}),\overline{x}) + (1-\lambda)F(g_1(\overline{y}),\overline{y}) \\ &= \lambda\overline{F}_2(\overline{x}) + (1-\lambda)\overline{F}_2(\overline{y}) \quad . \end{aligned}$$

In [6], Frommer introduces the class $\mathcal{C}$, namely $F \in \mathcal{C}$ if it is order convex and if, whenever $z^1 \ge z^2 \ge \ldots \ge z^n$ are in D, then the matrix of gradients

$$\begin{pmatrix} f_1'(z^1) \\ \vdots \\ f_n'(z^n) \end{pmatrix} \tag{3.3}$$

is a nonsingular M-matrix. We consider here the more, formally at least, general class $\mathcal{C}(y^0)$ of those functions F that are order convex and such that (3.3) is a nonsingular M-matrix whenever $y^0 \ge z^1 \ge z^2 \ge \ldots \ge z^n$, $z^i \in D$.

Lemma 3.3 *Suppose that $F \in \mathcal{C}(y^0)$. Then the following hold:*
(i) $\overline{F}_2$ is in the class $\mathcal{C}(\overline{y^0})$ on the convex domain $D_2 := \{z/(g_1(z), z) \in D\}$.
(ii) If $F(y^0) \geq 0$, then $\overline{F}_2(\overline{y^0}) \geq 0$ as well, $y^1 \leq y^0$ and $F(y^1) \geq 0$.
(iii) If moreover $F'(y)$ is always irreducible and $F(y^0) \neq 0$, then $y^1 < y^0$.
(iv) If F' is isotone, then $\overline{F}_2'$ is isotone as well.

Proof. Only the discretised version will be considered.
(i) Consider $\overline{y^0} \geq z^2 \geq \ldots \geq z^n$, $z^i \in D_2$; clearly the matrix

$$\begin{pmatrix} \delta f_1(y^0, h) \\ f_2'(g_1(z^2), z^2) \\ \vdots \\ f_n'(g_1(z^n), z^n) \end{pmatrix} \quad ,$$

where $\delta f_1(y^0, h) := (\delta_1 f_1(y^0, h), \ldots, \delta_n f_1(y^0, h))$ is the discretised gradient of f_1, is a nonsingular M-matrix. Thus, by defining

$$M := \begin{pmatrix} 1 & 0 & 0 & \ldots & \ldots & 0 \\ -m_{2,1} & 1 & 0 & \ldots & \ldots & 0 \\ -m_{3,1} & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 & 0 \\ -m_{n,1} & 0 & \ldots & \ldots & 0 & 1 \end{pmatrix} \quad ,$$

with

$$m_{i,1} := \frac{\partial_1 f_i(g_1(z^i), z^i)}{\delta_1 f_1(y^0, h)} \quad , \quad 2 \leq i \leq n \quad ,$$

it follows that

$$M * \begin{pmatrix} \delta f_1(y^0, h) \\ f_2'(g_1(z^2), z^2) \\ \vdots \\ f_n'(g_1(z^n), z^n) \end{pmatrix} = \begin{pmatrix} \delta_1 f_1(y^0, h) & \ldots & \ldots & \delta_n f_1(y^0, h) \\ 0 & \partial_2 f_{2,2}(z^2) & \ldots & \partial_n f_{2,2}(z^2) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \partial_2 f_{2,n}(z^n) & \ldots & \partial_n f_{2,n}(z^n) \end{pmatrix} \quad .$$

This last equality implies that $\overline{F}_2$ is in the class $\mathcal{C}(\overline{y^0})$ (See the proof of Lemma 3.3 in [10]).

(ii) If $F(y^0) \geq 0$, the order convexity implies for $2 \leq i \leq n$, that

$$\begin{aligned} f_{2,i}(\overline{y^0}) &= f_i(g_1(\overline{y^0}), \overline{y^0}) = f_i(g_1(\overline{y^0}), \overline{y^0}) - f_i(y^0) + f_i(y^0) \\ &\geq \partial_1 f_i(y^0)(g_1(\overline{y^0}) - y_1^0) + f_i(y^0) \geq f_i(y^0) \geq 0 \quad . \end{aligned}$$

Notice as above, that $f_{i,i}(\overline{y^0}) \geq 0$, and recall that

$$g_i(y_{i+1}, \ldots, y_n) = y_i^0 - \frac{1}{\delta_i f_{i,i}(\overline{y^0}, h)} \left[f_{i,i}(\overline{y^0}) + \sum_{j=2}^{n} \delta_j f_{i,j}(\overline{y^0}, h)(y_j - y_j^0) \right] \quad .$$

Since

$$y_n^1 = y_n^0 - \frac{f_{n,n}(y_n^0)}{\delta f_{n,n}(y_n^0, h)} \leq y_n^0 \quad ,$$

it follows inductively for $n-1 \geq i \geq 1$ that

$$y_i^1 = g_i(y_{i+1}^1, \ldots, y_n^1) \leq g_i(y_{i+1}^0, \ldots, y_n^0) \leq y_i^0 \quad .$$

Now, from Lemma 3.1

$$\begin{aligned} f_1(y^1) &= f_1(y^1) - f_1(y^0) + f_1(y^0) \geq f_1'(y^0)(y^1 - y^0) + f_1(y^0) \\ &\geq \delta f_1(y^0)(y^1 - y^0) + f_1(y^0) = 0 \quad . \end{aligned}$$

Analogously, isotonocity of $f_{i,i}$ and Lemma 3.1 imply that

$$f_i(y^1) = f_{i,i}(\overline{y^1}) \geq \delta f_{i,i}(\overline{y^0})(\overline{y^1} - \overline{y^0}) + f_{i,i}(\overline{y^0}) = 0 \quad .$$

(iii) Note first that, since $F'(y^0)$ is irreducible, then g_1 is strictly isotone, namely

$$g_1(z) < g_1(w) \quad , \quad if \quad z < w \quad .$$

Also, from the proof of (i), it follows that $\overline{F}_2'(\overline{y^0})$ is irreducible.
Now, from

$$f_{2,i}(\overline{y^0}) \geq \partial_1 f_i(y^0) * \left(-\frac{f_1(y^0)}{\delta_1 f_1(y^0)} \right) + f_i(y^0) \geq f_i(y^0) \geq 0 \quad , \quad 2 \leq i \leq n \quad ,$$

it follows that $\overline{F}_2(\overline{y^0}) = 0$ implies $F(y^0) = 0$, whence $\overline{F}_2(\overline{y^0}) \neq 0$.
Inductively, it is thus obtained that g_i is strictly isotone, that $\overline{F}_{i+1}'(\overline{y^0})$ is irreducible and that $\overline{F}_{i+1}(\overline{y^0}) \neq 0$, for $1 \leq i \leq n-1$.
Thus $f_{n,n}(y_n^0) > 0$, so that

$$y_n^1 < y_n^0 \quad ,$$

and

$$y_i^1 = g_i(\overline{y^1}) < g_i(\overline{y^0}) \leq y_i^0 \quad , \quad n-1 \geq i \geq 1 \quad .$$

Remark. Notice that if F' is isotone and $F'(y^0)$ is irreducible, then $F'(y)$ is irreducible whenever $y \leq y^0$.

Corollary 3.4 *Assume all the hypotheses in the preceding Lemma. The following conclusions hold:*
(i) $y^{k+1} \leq y^k$ *and* $F(y^k) \geq 0$ *for* $k \geq 0$.
(ii) If $F'(y)$ *is always irreducible, then, for each* k*, either* $F(y^k) = 0$ *or* $y^{k+1} < y^k$.
(iii) If there exists x^0 *such that* $F(x^0) \leq 0$*, then* $x^0 \leq y^k$.
(iv) Assume that $y^* = \lim y^k$ *exists and that there exist* b_i*,* $1 \leq i \leq n$ *, such that* $\partial_i f_{i,i}(\overline{y^k}) \leq b_i$ *or* $\delta_i f_{i,i}(\overline{y^k}, h_k) \leq b_i$ *in the case of discrete iterations. Then* $\lim F(y^k) = 0$*, i.e.* $F(y^*) = 0$ *if* $y^* \in D$.

Proof. (i) and (ii) are immediate consequences of the preceding Lemma, while (iii) follows from the results in [8]. As for (iv), the same arguments applied in [6] hold here as well; they can be simplified by taking into account the lemma above.
Remark. The conclusions in the Corollary above excluding (ii) have been proved by Frommer for $F \in \mathcal{C}$ based on Algorithm B (See [6]). Some of the features in the new proofs will be used to prove the main results in this paper.

From now on, all the hypotheses in the MNT will be assumed. In [6], a complementary sequence is introduced in Step 6 of Algorithm B, namely

$$x^{i+1,k} := x^{i,k} - [f_i(x^{i,k})/z_i^{i,k}]R^{i+1,k}e^i \quad , \tag{3.4}$$

while the corresponding one in Step 8 is

$$x^k := x^{n+1,k}$$

In order to obtain this sequence in the framework of Algorithm A, the following complementary steps are required for the analytic (resp. discretised) version

1'. Set $\overline{x^0} := x^0$ $(i = 1)$ and $\overline{F}_1^-(x) = (f_{1,j}^-(x)) := (f_j(x)) = F(x)$.

2'. Consider the affine aproximation $lf_{i,i}^-$ of $f_{i,i}^-$ centered at $\overline{x^0}$ with the (resp. discretised) gradient values of $f_{i,i}$ at $\overline{y^0}$ and solve for x_i , i.e.
$x_i = g_i^-(x_{i+1}, \ldots, x_n)$.

3'. Define the $(i+1)$-th reduced lower system

$$\overline{F}_{i+1}^-(x_{i+1}, \ldots, x_n) = 0 \quad , \quad by$$

$$f_{i+1,j}^-(x_{i+1}, \ldots, x_n) := f_{i,j}^-(g_i^-(x_{i+1}, \ldots, x_n), x_{i+1}, \ldots, x_n) \ , \ i+1 \leq j \leq n \ .$$

4'. Set $\overline{x^0} := (x^0_{i+1}, \ldots, x^0_n)$.

5'. Consider a first order aproximation of $f^-_{n,n}$ at x^0_n, with the slope given by $\partial_n f_{n,n}(\overline{y^0})$ (resp. by $\delta_n f_n(\overline{y^0})$) , equate it to 0 and call its solution x^1_n.

6'. For $i = n-1$ to 1 define $x^1_i := g^-_i(x^1_{i+1}, \ldots, x^1_n)$.

Remark. This sequence is similar to that of the Newton-Fourier iterates so that its terms will be called Brown-Fourier iterates. The compact description of the Brown-Fourier iterates given by (3.4) gives a further illustration of the usefulness of Algorithm B for computational purposes.

Lemma 3.5 *The first discretised Brown-Fourier iterate x^1 satisfies*

$$\delta f_{i,i}(\overline{y^0})(\overline{x^1} - \overline{x^0}) = -f^-_{i,i}(\overline{x^0}) \quad , \quad 1 \leq i \leq n \quad . \tag{3.5}$$

Proof. It follows as in Lemma 3.1.

Corollary 3.6 *The Brown-Fourier iterates satisfy*

$$x^k \leq x^{k+1} \leq y^* \quad , \quad \textit{and} \quad F(x^k) \leq 0 \quad .$$

If $F'(y)$ is always irreducible, then, for each k, either $x^k = y^*$ or $x^k < x^{k+1}$.

Proof. By noting that

$$g^-_1(\overline{x}) = x^0_1 - \frac{1}{\delta_1 f_1(y^0)} \left[f_1(x^0) + \sum_{j=2}^{n} \delta_j f_1(y^0)(x_j - x^0_j) \right] \quad ,$$

it follows that for $2 \leq i \leq n$,

$$\begin{aligned} f^-_{2,i}(\overline{x^0}) &= f_i(g^-_1(\overline{x^0}), \overline{x^0}) = f_i(g^-_1(\overline{x^0}), \overline{x^0}) - f_i(x^0) + f_i(x^0) \\ &\leq \partial_1 f_i(g^-_1(\overline{x^0}), \overline{x^0}) * (g^-_1(\overline{x^0}) - x^0_1) + f_i(x^0) \leq f_i(x^0) \leq 0 \quad . \end{aligned}$$

A simple induction argument and Lemma 3.5 now imply that $f^-_{i,i}(\overline{x^0}) \leq 0$, for $i \leq n$; in order to prove that $x^0 \leq x^1$, note that $x^1 - x^0$ is the solution of the upper triangular system (3.5), which has the same coefficients as (3.2); the proof of Lemma 3.3 yields that its matrix is a nonsingular M-matrix, i.e. its diagonal terms are positive and the off-diagonal terms are nonpositive while now the data are nonnegative. The last statement follows by recalling that the irreducibility hypothesis implies that (3.5) has in every row a nonvanishing term outside the diagonal.

Corollary 3.7 *If $F'(y)$ is always irreducible, then the following hold:*
(i) $F(y^0) = 0$ if and only if $\overline{F}_2(\overline{y^0}) = 0$.
(ii) $F(x^0) = 0$ if and only if $\overline{F}_2^-(\overline{x^0}) = 0$.

Proof. (i) From Lemma 3.3-(iii) and its proof it clearly follows that $F(y^0) \neq 0$ implies $\overline{F}_2(\overline{y^0}) \neq 0$. Now, if $\overline{F}_2(\overline{y^0}) \neq 0$, from the same proof one obtains that $y^1 \leq y^0$ and $y^1 \neq y^0$, and since $F(y^1) \geq 0$ yields $y^* \leq y^1$, the conclusion is that $F(y^0) \neq 0$.
(ii) It follows as (i) from the proof of the previous corollary.
Remark Note that the meaning of the previous corollary is interesting also when the function F is affine linear.

The results described so far provide a proof based on Algorithm A to the following theorem whose first part was proved in [6].

Theorem 3.8 *The analytic and discretised Brown and Brown-Fourier iterates (y^k) and (x^k) satisfy*
(i) $x^k \leq x^{k+1} \leq y^* \leq y^{k+1} \leq y^k$ and $\lim x^k = y^* = \lim y^k$.
(ii) If F' is isotone and $F'(y^0)$ is irreducible, $y^k \neq y^*$ implies $y^{k+1} < y^k$, while $x^k \neq y^*$ implies $x^k < x^{k+1}$.

In all that remains of the section we analyse some aspects of the Brown iterates with the following provisos; in case F is almost linear, then both analytic and discretised iterations are considered. For general F with the MNT hypotheses, only analytic iterations are considered. The notation corresponds to the discrete case.

Consider the modified system generated by Step 2, $i = 1$, in Algorithm A:

$$(LF)(y) = 0 \quad , \tag{3.6}$$

where it has been set

$$\begin{aligned} (LF)_1(y) &:= lf_1(y) := f_1(y^0) + \delta f_1(y^0)(y - y^0) \\ (LF)_i(y) &:= f_i(y) \quad , \quad 2 \leq i \leq n \quad . \end{aligned}$$

Lemma 3.9 *If $F(x) \leq 0$, then $(LF)(x) \leq 0$.*

Proof. Since necessarily $x \leq y^*$ (See [8]), the order convexity yields

$$lf_1(x) = f_1(y^0) + \delta f_1(y^0)(x - y^0) \leq f_1(y^0) + f_1'(y^0)(x - y^0) \leq f_1(x) \leq 0 \quad .$$

Corollary 3.10 *There exists a unique $y^{*,1}$ such that*

$$y^* \leq y^{*,1} \leq y^0 \, and \, (LF)(y^{*,1}) = 0 \, .$$

Proof. It follows by applying the MNT to (3.6) with starting interval $\langle y^*, y^0 \rangle$, and by also taking into account that for $y \leq y^0$

$$F'(y) \leq (LF)'(y) \quad ,$$

which implies that $(LF)'(y)$ is a nonsingular M-matrix as well.

Corollary 3.11 *If $(LF)(x) \leq 0$, then $x \leq y^{*,1}$, and if $(LF)(y) \geq 0$, then $y \geq y^{*,1}$.*

Recall that

$$g_1(\overline{y}) = y_1^0 - \frac{1}{\delta_1 f_1(y^0)} \left[f_1(y^0) + \sum_{j=2}^{n} \delta_j f_1(y^0)(y_j - y_j^0) \right] \quad .$$

Clearly it follows that

$$y_1^{*,1} = g_1(\overline{y^{*,1}}) \quad ,$$

and also that (3.6) is equivalent to

$$y_1 - g_1(\overline{y}) = 0 \quad ,$$

together with the reduced system

$$\overline{F}_2(\overline{y}) = 0 \quad . \tag{3.7}$$

The solution of the reduced system (3.7) is $\overline{y^{*,1}}$. The following theorem restates Lemma 3.3.

Theorem 3.12 *The reduced system (3.7) satisfies all the hypotheses of the MNT with respect to the starting interval $\left\langle \overline{y^*}, \overline{y^0} \right\rangle$. If $F'(y^0)$ is irreducible, then $\delta\overline{F}_2(\overline{y^0}, h)$ is irreducible too for h conveniently small, depending on $\langle x^0, y^0 \rangle$.*

Proof. Notice first that, g_1 being isotone,$\overline{F}_2'$ and $\delta\overline{F}_2$ are isotone as well. Note also that if $F'(y^0)$ is irreducible, then g_1 is strictly isotone, i.e.

$$if \quad \overline{x} < \overline{y} \quad , \quad then \quad g_1(\overline{x}) < g_1(\overline{y}) \quad .$$

Now $y_1^0 \geq g_1(\overline{y^0})$ yields

$$f_j(y^0) - f_{2,j}(\overline{y^0}) = \partial_1 f_j * (y_1^0 - g_1(\overline{y^0})) \leq 0 \quad ,$$

which implies that $f_{2,j}(\overline{y^0}) \geq 0$, for $2 \leq j \leq n$. Note now that by applying Lemma 3.8 in the identity

$$y_1 - g_1(\overline{y}) = \frac{1}{\delta_1 f_1(y^0)} l f_1(y) \quad ,$$

it follows that $y_1^* \le g_1(\overline{y^*})$; as a consequence, for $2 \le j \le n$,

$$\begin{aligned} f_{2,j}(\overline{y^*}) &= f_{1,j}(g_1(\overline{y^*}), \overline{y^*}) = f_{1,j}(g_1(\overline{y^*}), \overline{y^*}) - f_{1,j}(y^*) \\ &\le \partial_1 f_{1,j}(g_1(\overline{y^*}), \overline{y^*})(g_1(\overline{y^*}) - y_1^*) \le 0 \quad . \end{aligned}$$

In this way, Brown's method can be considered as a dimension reduction method that preserves the hypotheses in the MNT.

4 Brown and Newton iterations compared

Throughout the section, all the hypotheses in the MNT are assumed; also, unless stated otherwise, the increments in the finite difference quotients are the same at each discretised iteration of both Brown's and Newton's methods.

Lemma 4.1 *The following propositions hold for* $1 \le p \le n-1$*:*
(i) $g_p^-(\overline{y}) \le g_p(\overline{y}) \quad .$
(ii) $\delta_j f_{p+1,i}^-(\overline{y}) \le \delta_j f_{p+1,i}(\overline{y}) \quad , \quad p+1 \le i,j \le n \quad .$
(iii) $f_{p+1,i}(\overline{y}) \le f_{p+1,i}^-(\overline{y}) \quad , \quad p+1 \le i \le n.$

Proof. The proof proceeds inductively, and only its first step is shown.
(i) Notice that

$$f_1(y^0) - f_1(x^0) \le f_1'(y^0)(y^0 - x^0) \le \delta f_1(y^0)[(y - x^0) - (y - y^0)] \quad ,$$

whence

$$-[f_1(x^0) + \delta f_1(y^0)(y - x^0)] \le -[f_1(y^0) + \delta f_1(y^0)(y - y^0)] \quad .$$

Consequently,

$$g_1^-(x_2, \ldots, x_n) \le g_1(x_2, \ldots, x_n) \quad .$$

(ii) By taking into account (2.4), that $f_{2,i}^-(\overline{y}) = f_{1,i}(g_1^-(\overline{y}), \overline{y})$ and $f_{2,i}(\overline{y}) = f_{1,i}(g_1(\overline{y}), \overline{y})$, and that $\partial_j g_1^- = \partial_j g_1 \ge 0$, the conclusion is implied by (i) and isotonicity.
(iii) Order convexity and (i) imply that for $2 \le i \le n$,

$$\begin{aligned} f_{2,i}^-(\overline{y}) &= f_i(g_1^-(\overline{y}), \overline{y}) = f_i(g_1^-(\overline{y}), \overline{y}) - f_i(g_1(\overline{y}), \overline{y}) + f_i(g_1(\overline{y}), \overline{y}) \\ &= (\partial_1 f_i)(g_1^-(\overline{y}) - g_1(\overline{y})) + f_i(g_1(\overline{y}), \overline{y}) \\ &\ge f_i(g_1(\overline{y}), \overline{y}) = f_{2,i}(\overline{y}) \quad . \end{aligned}$$

(i) Let us finally notice that

$$\begin{aligned} f^-_{2,2}(\overline{y^0}) - f^-_{2,2}(\overline{x^0}) &\le \delta f^-_{2,2}(\overline{y^0})(\overline{y^0} - \overline{x^0}) \\ &\le \delta f_{2,2}(\overline{y^0})(\overline{y^0} - \overline{x^0}) \\ &= \delta f_{2,2}(\overline{y^0})[(\overline{y} - \overline{x^0}) - (\overline{y} - \overline{y^0})] \quad . \end{aligned}$$

Now, by applying (iii), and proceeding as in the proof of (i) above, it follows that

$$g^-_2(y_3, \dots, y_n) \le g_2(y_3, \dots, y_n) \quad .$$

Lemma 4.2 *The Newton-Fourier data satisfy the following inequalities:*

$$f^-_{p+1,i}(\overline{x^0}) \le f^-_{p,i}(\overline{x^0}) - \frac{\delta_p f_{p,i}(\overline{y^0})}{\delta_p f_{p,p}(\overline{y^0})} f^-_{p,p}(\overline{x^0})\,,\ 2 \le p+1 \le i \le n\,,$$

where it has been defined $f^-_{1,i} := f_i\,,\ 1 \le i \le n\,.$

Proof. Application of the previous lemma yields the conclusion in the following way:

$$\begin{aligned} f^-_{p+1,i}(\overline{x^0}) &= f^-_{p,i}(g^-_p(\overline{x^0}), \overline{x^0}) \\ &= f^-_{p,i}(g^-_p(\overline{x^0}), \overline{x^0}) - f^-_{p,i}(\overline{x^0}) + f^-_{p,i}(\overline{x^0}) \\ &\le \partial_p f^-_{p,i}(g^-_p(\overline{x^0}), \overline{x^0}) \left(-\frac{f^-_{p,p}(\overline{x^0})}{\delta_p f_{p,p}(\overline{y^0})} \right) + f^-_{p,i}(\overline{x^0}) \\ &\le \partial_p f_{p,i}(g^-_p(\overline{x^0}), \overline{x^0}) \left(-\frac{f^-_{p,p}(\overline{x^0})}{\delta_p f_{p,p}(\overline{y^0})} \right) + f^-_{p,i}(\overline{x^0}) \\ &\le \partial_p f_{p,i}(\overline{y^0}) \left(-\frac{f^-_{p,p}(\overline{x^0})}{\delta_p f_{p,p}(\overline{y^0})} \right) + f^-_{p,i}(\overline{x^0}) \quad . \end{aligned}$$

Starting with $\langle x^0, y^0 \rangle$, y^1_B and y^1_N denote, respectively, the first Brown and Newton (discretised) iterates, while x^1_B and x^1_N are their corresponding Fourier iterates. The proof of the following theorem follows the lines of Theorem 3.9 in [10].

Theorem 4.3 *Suppose that F is almost linear. The following inequalities hold for the first discretised iterates:*

$$x^1_N \le y^1_B \le y^* \le y^1_B \le y^1_N \quad . \tag{4.1}$$

Proof. Mathematical induction on n will first be applied to establish the comparison for the Brown and Newton iterates. Clearly (4.1) is trivially true for $n = 1$. Recall now that $\overline{y_B^1} = (y_2^1, \dots, y_n^1)$ is the first iterate generated by $\overline{y^0} = (y_2^0, \dots, y_n^0)$ for the reduced system (3.7). The inductive hypothesis yields

$$\overline{y^*} \leq \overline{y^{*,1}} \leq \overline{y_B^1} \leq (\overline{y^0})_N^1 := \overline{y^0} - (\delta \overline{F}_2(\overline{y^0}))^{-1} \overline{F}_2(\overline{y^0}) \quad .$$

Notice first that if M is defined as in Lemma 3.3 but with

$$m_{i,1} := \frac{\delta_1 f_i(g_1(\overline{y^0}), \overline{y^0})}{\delta_1 f_1(y^0)} \quad ,$$

then, taking into account Lemma 3.2, it follows that

$$M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}) = \begin{pmatrix} \delta_1 f_1(y^0) & \dots & \dots & \delta_n f_1(y^0) \\ 0 & & & \\ \vdots & & \delta \overline{F}_2(\overline{y^0}) & \\ 0 & & & \end{pmatrix} \quad . \tag{4.2}$$

Thus it follows for $i \geq 2$ that

$$\begin{aligned} ((\overline{y^0})_N^1)_i &= y_i^0 - \sum_{j=2}^{n} (\delta \overline{F}_2(y^0))_{i,j}^{-1} f_j(g_1(\overline{y^0}), \overline{y^0}) \\ &= y_i^0 - \sum_{j=2}^{n} (\delta(LF)(g_1(\overline{y^0}), \overline{y^0}))_{i,j}^{-1} f_j(g_1(\overline{y^0}), \overline{y^0}) \\ &= y_i^0 - \sum_{j=1}^{n} (\delta(LF)(g_1(\overline{y^0}), \overline{y^0}))_{i,j}^{-1} (LF)_j(g_1(\overline{y^0}), \overline{y^0}) \quad . \end{aligned}$$

On the other hand, LF being order convex, it yields

$$\begin{aligned} (LF)(y^0) - (LF)(g_1(\overline{y^0}), \overline{y^0})) &\leq (LF)'(y^0)(y^0 - (g_1(\overline{y^0}), \overline{y^0})) \\ &\leq \delta(LF)(y^0)(y^0 - (g_1(\overline{y^0}), \overline{y^0})) \quad , \end{aligned}$$

whence

$$(g_1(\overline{y^0}), \overline{y^0}) - (\delta(LF)(y^0))^{-1}(LF)(g_1(\overline{y^0}), \overline{y^0}) \leq l y_n^1 \quad , \tag{4.3}$$

where

$$ly^1_N := y^0 - (\delta(LF)(y^0))^{-1}(LF)(y^0) \quad .$$

Since $\delta(LF)(y^0) \leq \delta F(y^0)$, we obtain that

$$ly^1_N \leq y^0 - (\delta F(y^0))^{-1}(LF)(y^0) = y^1_N \quad .$$

Since from Lemma 2.3,

$$\delta(LF)(g_1(\overline{y^0}), \overline{y^0}) \leq \delta(LF)(y^0) \quad , \tag{4.4}$$

and from Theorem 3.12, $(LF)(g_1(\overline{y^0}), \overline{y^0}) \geq 0$, (4.4) applied in (4.3) implies

$$(g_1(\overline{y^0}), \overline{y^0}) - (\delta(LF)(g(\overline{y^0}), \overline{y^0}))^{-1}(LF)(g_1(\overline{y^0}), \overline{y^0}) \leq ly^1_N \leq y^1_N \quad . \tag{4.5}$$

From (4.5) now follows

$$(\overline{y^0})^1_N \leq \overline{(ly^1_N)} \leq \overline{y^1_N} \quad . \tag{4.6}$$

The inductive hypothesis and (4.6) imply

$$(\overline{y^0})^1_B = \overline{y^1_B} \leq \overline{y^1_N} \quad . \tag{4.7}$$

Now recall that by applying the MNT to (3.6) it follows that $lf_1(ly^1_N) \geq 0$, which implies

$$g_1(\overline{ly^1_N}) \leq (ly^1_N)_1 \quad . \tag{4.8}$$

It is apparent now that (4.7), (4.8) and (4.5) imply

$$(y^1_B)_1 = g_1(\overline{y^1_B}) \leq g_1(\overline{ly^1_N}) \leq (y^1_N)_1 \quad ,$$

which combined again with (4.7) finally gives

$$y^1_B \leq y^1_N \quad .$$

Consider the Fourier iterates. Mathematical induction is also applied here, but in a somewhat different way; its first step is shown whence it is easy to formulate the general step. Denote

$$x^1_{N(1)} := x^1_N = x^0 - \delta F(y^0)^{-1}F(x^0) \quad . \tag{4.9}$$

Recall now that

$$\delta(LF)(g_1(\overline{y^0}), \overline{y^0}) \leq \delta F(y^0) \quad , \tag{4.10}$$

and if M is as in (4.2) while

$$M_{N(1)} := \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ -m^1_{2,1} & 1 & 0 & \dots & \dots & 0 \\ -m^1_{3,1} & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 & 0 \\ -m^1_{n,1} & 0 & \dots & \dots & 0 & 1 \end{pmatrix} \quad ,$$

with

$$m^1_{i,1} := \frac{\delta_1 f_i(y^0)}{\delta_1 f_1(y^0)} \quad , \quad 2 \leq i \leq n \quad ,$$

then from (4.10) it follows that

$$M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}) \leq M_{N(1)} * \delta F(y^0) \quad , \tag{4.11}$$

because, for $2 \leq i, j \leq n$,

$$\delta_j f_i(g_1(\overline{y^0}), \overline{y^0}) - \frac{\delta_1 f_i(g_1(\overline{y^0}), \overline{y^0})}{\delta_1 f_1(y^0)} \delta_j f_1(y^0) \leq \delta_j f_i(y^0) - \frac{\delta_1 f_i(y^0)}{\delta_1 f_1(y^0)} \delta_j f_1(y^0) \quad .$$

From inequality (4.11) it also easily follows that

$$\delta F(y^0)^{-1} \leq (M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}))^{-1} * M_{N(1)} \quad , \tag{4.12}$$

which, when applied in (4.9), yields

$$x^1_{N(1)} \leq x^0 - (M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}))^{-1} * M_{N(1)} F(x^0) \quad ,$$

so that by applying (4.2) and Lemma 4.2, it results

$$\overline{x^1_{N(1)}} \leq x^1_{N(2)} := \overline{x^0} - \delta \overline{F}_2(y^0)^{-1} \overline{F}^-_2(\overline{x^0}) \quad . \tag{4.13}$$

The proposition to be proved along the same lines is that for $1 \leq p \leq n-1$,

$$\overline{x^1_{N(p)}} \leq x^1_{N(p+1)} := \overline{x^0} - \delta \overline{F}_{p+1}(y^0)^{-1} \overline{F}^-_{p+1}(\overline{x^0}) \quad . \tag{4.14}$$

The induction is carried on by taking the defining equality in (4.14) as the new starting point in (4.9). Recall now that

$$x^1_{N(n)} = (x^1_B)_n \quad , \tag{4.15}$$

and that for $1 \leq p \leq n-1$,

$$(x_B^1)_p = g_p^-(\overline{x_B^1}) = g_p^-((x_B^1)_{p+1}, \ldots, (x_B^1)_n) \quad . \tag{4.16}$$

Since from (4.14)

$$\overline{x_{N(n-1)}^1} \leq x_N^1(n) \quad , \tag{4.17}$$

(4.15) and (4.16) imply that

$$(x_{N(n-1)}^1)_{n-1} = g_{n-1}^-(\overline{x_{N(n-1)}^1}) \leq g_{n-1}^-(x_{N(n)}^1) = (x_B^1)_{n-1} \quad . \tag{4.18}$$

So that summing up (4.17) and (4.18)

$$x_{N(n-1)}^1 \leq \overline{x_B^1} \quad .$$

Thus, inductively backwards,

$$x_{N(p)}^1 \leq \overline{x_B^1} \quad , \quad n-1 \geq p \geq 1 \quad ,$$

so that the final result is

$$x_N^1 \leq x_B^1 \quad .$$

Remark. Note that the two inequalities following (4.3) are in fact equalities; however, if the Brown increments are allowed to be smaller or equal than their Newton counterparts, then the proof above requires only small additional modifications in order to hold with the extended hypothesis. Analogous minor modifications enable to extend all the results that follow. Most importantly, it is necessary to notice that the proof above also holds for analytic iterations for a general F in the MNT context, although in this case it can be simplified; this result has already been proved in [7], but the present proof enables to obtain strict inequalities as in the following theorem.

In the rest of the section it will be assumed that $n > 1$.

Theorem 4.4 *If F is almost linear, $F(y^0) \neq 0$ (resp. $F(x^0) \neq 0$), $F'(y)$ is always irreducible and*

$$\delta(LF)(g_1(\overline{y^0}), \overline{y^0}) \neq \delta(LF)(y^0) \quad , \tag{4.19}$$

then

$$y_B^1 < y_N^1 \quad (resp.\, x_N^1 < x_B^1) \quad .$$

Proof. Recall first from Corollary 3.7 that $F(y^0) \neq 0$ (resp. $F(x^0) \neq 0$) is equivalent to $\overline{F}_2(\overline{y^0}) \neq 0$ (resp. $\overline{F}_2^-(\overline{x^0}) \neq 0$), i.e. $(LF)(g_1(\overline{y^0}), \overline{y^0}) \neq 0$. Thus, the hypotheses imply that the first inequality in (4.5) is strict, because of (4.3), which yields the strict inequality for the Brown and Newton iterates. As for the Fourier iterates, note first that both terms in the inequality (4.11) are non singular M-matrices whose inverses have the same block structure, i.e.

$$(M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}))^{-1} = \begin{pmatrix} \frac{1}{\delta_1 f_1(y^0)} & v_B \\ 0 & (\delta\overline{F}_2(\overline{y^0}))^{-1} \end{pmatrix} \quad ,$$

and if

$$\begin{pmatrix} \delta_1 f_1(y^0) & \dots \\ 0 & F_{N(2)} \end{pmatrix} := M_{N(1)} * \delta F(y^0) \quad ,$$

then

$$(M_{N(1)} * \delta F(y^0))^{-1} = \begin{pmatrix} \frac{1}{\delta_1 f_1(y^0)} & v_N \\ 0 & (F_{N(2)})^{-1} \end{pmatrix} \quad .$$

From (4.10) now it follows, as (4.11), that $F_{N(2)}(\overline{y^0}) \neq \delta\overline{F}_2(\overline{y^0})$; hence, irreducibility implies that

$$0 < (F_{N(2)})^{-1} < (\delta\overline{F}_2(\overline{y^0}))^{-1} \quad . \tag{4.20}$$

Since

$$v_N = -\frac{1}{\delta_1 f_1(y^0)}(\delta_2 f_1(y^0), \dots, \delta_n f_1(y^0))F_{N(2)}^{-1} \ ,$$

while

$$v_B = -\frac{1}{\delta_1 f_1(y^0)}(\delta_2 f_1(y^0), \dots, \delta_n f_1(y^0))(\delta\overline{F}_2(y^0))^{-1} \ ,$$

and since for some $j, \delta_j f_1(y^0) < 0$, it follows from (4.20) that

$$0 < v_N < v_B \quad .$$

Thus, now

$$(\delta F(y^0))^{-1} = (M_{N(1)} * \delta F(y^0))^{-1} * M_{N(1)} < (M * \delta(LF)(g_1(\overline{y^0}), \overline{y^0}))^{-1} * M_{N(1)},$$

because for some $i, \delta_1 f_i(y^0) < 0$. Hence the inequality in (4.13) is strict, as $F(x^0) \neq 0$, and now it easily follows that

$$x_N^1 < x_B^1 \quad .$$

Remark. Note that (4.19) implies that $F(y^0) \neq 0$.

Theorem 4.5 *If $F'(y^0)$ is irreducible, $F(y^0) \neq 0$ (resp. $F(x^0) \neq 0$) and*

$$(LF)'(g_1(\overline{y^0}), \overline{y^0}) \neq (LF)'(y^0) \ , \tag{4.21}$$

then the analytic Brown and Newton (resp. Fourier) iterates satisfy

$$y_B^1 < y_N^1 \quad (resp.\, x_N^1 < x_B^1) \quad .$$

As a consequence, for h_1 sufficiently small, the first discretised iterates satisfy the same inequalities.

The following example shows that the assumption of F being almost linear can be essential for Theorem 4.3 in the discretised case while, perhaps more significantly, (4.19) or (4.21) can be so for the previous theorem. The example was provided by an unknown referee to show that it is not always true that in the MNT context, discretised Brown iterations converge faster than their Newton counterparts. It is good enough as to turn (4.21) into a necessary condition from a general point of view.

Consider $n = 2$ and F defined by

$$\begin{aligned} f_1(y_1, y_2) &:= y_1 - y_2 - 5 \\ f_2(y_1, y_2) &:= y_1 y_2 + 6 \end{aligned}$$

Clearly F' is isotone and it is an irreducibly diagonally dominant M-matrix whenever $y \in \langle x^0, y^0 \rangle$, where $x^0 := (3, -2)$ and $y^0 := (4, -1)$; here $y^* = (3, -2)$. If discretised Brown is applied, with the increment h in the difference quotients, it follows that

$$y_B^1 = (4 - \frac{2}{3+h}, -1 - \frac{2}{3+h}) \quad .$$

On the other hand, it is easy to check that

$$y_N^1 = (4 - \frac{2}{3}, -1 - \frac{2}{3}) \quad .$$

Theorem 4.6 *Suppose F is almost linear; then the discretised Brown and Newton iterates, as well as their Fourier counterparts satisfy*

$$x_N^k \leq x_B^k \leq y^* \leq y_B^k \leq y_N^k \quad , \quad 0 \leq k \quad .$$

These inequalities are also satisfied for the analytic iterates of a general F in the MNT context.

Proof. Only the proof for the Brown and Newton iterates is shown. Notice first that Theorem 4.3 yields

$$y_B^2 = (y_B^1)_B^1 \leq (y_B^1)_N^1 \quad . \tag{4.22}$$

The following comparison argument is similar to those in [13]. Let us denote $z := y_B^1$ and $w := y_N^1$. The order convexity implies that

$$z - (\delta F(w))^{-1} F(z) \leq w - (\delta F(w))^{-1} F(w) \quad .$$

But since $\delta F(z) \leq \delta F(w)$, it follows that

$$z_N^1 \leq w_N^1 \quad ,$$

which combined with (4.22) gives us

$$y_B^2 \leq y_n^2 \quad .$$

A simple induction argument can be used to complete the proof.

Corollary 4.7 *Assume the hypotheses in Theorem 4.4 and that F' is not constant on any open subset. Then the discretised iterates satisfy:*

$$y^* \leq y_B^k < y_N^k \ (resp.\, x_N^k < x_B^k \leq y^*) \ , \ 1 \leq k \quad .$$

Proof. Since $y_B^1 < y_N^1$, if $y^* \neq y_B^1$, then, with the notation in Theorem 4.6 it follows that $\delta F(z) \neq \delta F(w)$, so that

$$z_N^1 = z - (\delta F(z))^{-1} F(z) < z - (\delta F(w))^{-1} F(z) \leq w_N^1 \quad ,$$

and thus

$$y_B^2 < y_N^2 \quad .$$

If on the other hand $F(y_B^1) = 0$, recall from [12] that $y^* < y_N^{k+1} < y_N^k$ for all k.

Corollary 4.8 *Assume the hypotheses in Theorem 4.5, and that F' is not constant on any open subset. Then the analytic Brown and Newton iterates satisfy:*

$$y^* \leq y_B^k < y_N^k \ (resp.\ x_N^k < x_B^k \leq y^*) \ , \ 1 \leq k \quad .$$

As a consequence, the same inequalities hold for the discretised iterations, when the increments h_k are conveniently small.

Proof. It is only necessary to recall from [13] that $y_N^{k+1} < y_N^k$ for all k.

Remark. Note that the results above together with the remark following Theorem 4.3 imply that discretised Brown iterates will converge faster than their Newton counterparts provided that the increments for the former belong to convenient open neighborhoods of the latter, which must be conveniently small. Recall again that this condition is also necessary to ensure quadratic convergence of discretised Newton iterations.

5 Numerical Examples

The computations have been carried with the double precision of Fortran 77, and no substantial differences have been observed in the results obtained with algorithms A and B. Consider first the equation

$$u''(t) = e^{u(t)} \quad , \quad 0 \le t \le 1 \quad , \quad u(0) = u(1) = 0 \quad ,$$

and approximate it with the standard second divided difference with $h := \frac{1}{21}$. In the resulting nonlinear system a typical function is

$$F_i(u_1, \ldots, u_{20}) = h^2 e^{u_i} + 2u_i - u_{i-1} - u_{i+1} \quad , \quad 1 \le i \le 20 \quad ,$$

where u_i stands for $u(ih)$ and $u_0 := u_{21} := 0$. Consider y^0 as the vector whose components are all equal to 1. Clearly $F(y^0) \ge 0$. Only upper iterates are considered here, because their behaviour is demonstrative enough. The number of necessary analytic Newton iterations in order to satisfy

$$\|F(y^k)\|_\infty < .5 * 10^{-13} \tag{5.1}$$

has been 5, while for Brown's iterations their number has been 4. As to the discretised iterations, the following table shows that their behaviour is consistent with the theory. In each case the value of h has been held constant throughout all the iterations, i.e. until (5.1) has been satisfied. Only relevant values of h have been included in the first row, i.e. values for which the number of iterations in one of the methods change with respect to the previous or the following value of h. The second and third rows show the first value of k for which (5.1) is achieved.

$h :=$	10^{-1}	10^{-2}	10^{-3}	10^{-5}	10^{-6}	10^{-10}	10^{-11}
Newton	7	6	5	5	5	5	5
Brown	7	6	5	5	4	4	5

In the second example, the discretisation of Chandrasekhar's equation has been implemented as in [7]. More precisely, for

$$v(t) = 1 - \frac{1}{4}\int_0^1 \left(\frac{t}{s+t} * \frac{1}{v(s)} \right) ds \quad ,$$

the trapezoidal integration rule is applied at the points $ih, 0 \leq i \leq 64$, with $h := \frac{1}{64}$. By taking into account that $v(0) = 1$, the resulting nonlinear system is

$$F(x) = 0 \quad ,$$

where for $1 \leq i \leq 64$,

$$f_i(x) = x_i + \frac{1}{4}\left[w_0 + \sum_{j=1}^{64} w_j * \frac{i}{(i+j)} * \frac{1}{x_j} \right] - 1 \quad ,$$

with

$$w_0 := w_{64} := \frac{h}{2} \quad , \quad w_j := h, 1 \leq j \leq 63 \quad .$$

If now y^0 is taken as in the first example and $x^0 := \frac{1}{2}y^0$, it follows that the hypotheses in the MNT are satisfied. If (5.1) is used as stopping criterium, then 4 analytic Newton iterations are needed to satisfy it, while for Brown's analytic method 3 iterations suffice. In the following table the corresponding values for the discretised iterations are shown. As above only relevant values are included.

$h :=$	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-8}	10^{-9}
Newton	8	6	5	5	4	4	4	4
Brown	7	6	5	5	5	4	4	5

The results shown in this table go against the expectations, because the stopping criterium based on (5.1) is not good enough. In fact, for the relevant values corresponding to $v(1)$, with $h := 10^{-5}$ both the Brown and Fourier-Brown iterates yield the machine convergence value 0.799194702574 after 4 iterations; on the other hand, after 4 Newton iterations the same value is attained but the corresponding Newton-Fourier value differs in the last digit. If instead of (5.1), the stopping criterium is based on

$$\|y^k - x^k\|_\infty < .5 * 10^{-13} \quad , \tag{5.2}$$

then a significantly different table is obtained, consistent with both the theoretical prediction and the numerical results, namely

$h :=$	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-8}	10^{-9}
Newton	9	6	5	5	5	5	5	5
Brown	7	6	5	4	4	4	4	4

It is necessary to point out that with the stopping criterium given by (5.2), both analytic methods converge after 4 iterations. However, although for the relevant value of $v(1)$ all the four iterates coincide with the convergence value, only the Brown iterations do so after only 3 iterations.

The next table gives the values obtained as approximations to $v(1)$ by both the discretised Newton and Brown methods together with their Fourier counterparts with $h := -3 * 10^{-7}$ held constant. Convergence is attained after 4 iterations for both methods and with either (5.1) or (5.2) as stopping criteria. The subindexes N and B stand for Newton and Brown, respectively. Similar values have been obtained in Table 2 in [7] although not quite the same.

k	$(x_N^k)_{64}$	$(x_B^k)_{64}$	$(y_B^k)_{64}$	$(y_N^k)_{64}$
0	0.500000000000	0.500000000000	1.000000000000	1.000000000000
1	0.707150028325	0.793434228865	0.799636684959	0.803989531181
2	0.797361036475	0.799184364894	0.799194762877	0.799198386608
3	0.799194160116	0.799194702544	0.799194702574	0.799194702576
4	0.799194702574	0.799194702574	0.799194702574	0.799194702574

6 Final Remarks

The main conclusion to be retained from the paper is that in the context of the MNT the Brown method should often be a better option than the Newton method, be it in the analytic or in the discretised version; of course, in the latter case, it is necessary to check the sufficient conditions described here, especially if one is enticed by the possibility of significantly diminishing the computational cost. That conclusion can be viewed as the nonlinear equivalent that for linear systems states that with appropriate hypotheses, the Gauss-Seidel method converges faster than the Jacobi method ([20]). Since this relation is one of the basic elements in the analysis of linear iterative methods, the paper suggests the interest of further exploring this feature of Brown's method with respect to Newton's method, and also, to analise similar results for both block procedures and Brown versions of other methods, not necessarily Newton's; these possibilities will be examined elsewhere. Also, in the same way as in the multigrid approach to linear elliptic problems, smoothing of the error is usually accomplished by means of Gauss-Seidel iterations, for the equivalent procedure in nonlinear elliptic problems one could prefer Brown iterations rather than Newton iterations. Last but not least, other interesting questions arise that could be addressed with the present results in mind, as for instance those regarding effective computational implementation, taking into account the computational environment.

The author is warmly thankful both for the partial support received from the Institute for Numerical Computation and Analysis (INCA), Dublin, Ireland, and for the encouragement from its Director, Professor John J.H. Miller. The main results in the almost linear case have been announced in the third Hellenic-European Conference on Mathematics and Informatics, held in Athens in 1996.

References

[1] Anselone, P. (1964): Nonlinear Integral Equations. University of Wisconsin Press, Madison, Wisconsin.

[2] Baluev, A. (1952): On the method of Chaplygin (Russian). Dokl. Akad. Nauk SSSR **83**, 781-784.

[3] Berman, A., Plemmons, R. J. (1994): Nonnegative Matrices in the Mathematical Sciences. Society for Industrial and Applied Mathematics,SIAM, Philadelphia.

[4] Brent, R. (1973):Some efficient algorithms for solving systems of nonlinear equations. SIAM J. Numer. Anal. **10**, 327-344.

[5] Brown, K. (1969): A quadratically convergent Newton-like method based upon Gaussian elimination. SIAM J. Numer. Anal. **6**, 560-569.

[6] Frommer, A. (1988): Monotonicity of Brown's method. ZAMM **68**, 101-109.

[7] Frommer, A. (1988): Comparison of Brown's and Newton's method in the monotone case. Numer. Math. **52**, 511-521.

[8] Gale, D., Nikaido, H. (1965): The Jacobian matrix and global univalence of mappings. Math. Ann. **81**, 81-93.

[9] Gay, D.M. (1975): Brown's method and some generalizations with applications to minimization problems. Techn. Rep. 75-225, Dept. of Comput. Sci., Cornell Univ., Ithaca, New York.

[10] Milaszewicz, J.P., Abdel Masih, S. (1993): On elimination and fixed point iterations. Comput. Math. Appl. **25**, 43-53

[11] Milaszewicz, J.P., Abdel Masih, S.(1994): Errata to "On elimination and fixed point iterations". Comput. Math. Appl. **27**, 113-115.

[12] Milaszewicz, J.P. (1995): Monotone discrete Newton iterations and elimination. Comput. Math. Appl. **30**,79-90.

[13] Milaszewicz, J.P. (1995): Comparison theorems for monotone Newton-Fourier iterations and applications in functional elimination. Linear Algebra Appl. **220**, 343-357.

[14] Ortega, J.M., Rheinboldt, W.C. (1967): Monotone iterations for nonlinear equations with application to Gauss-Seidel methods. SIAM J. Numer. Anal. **4**, 171-190.

[15] Ortega, J.M., Rheinboldt, W.C. (1970): Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.

[16] Ostrowski, A.M. (1973): Solution of Equations in Euclidean and Banach Spaces. Academic Press, New York.

[17] Schmidt, J.W., Hoyer, W. (1977): Ein Konvergenzsatz für Verfahren vom Brown-Brent-Typ. ZAMM **57**, 397-405.

[18] Schwetlick, H. (1979): Numerische Lösung Nichtlinearer Gleichungssysteme. R. Oldenbourg Verlag, München.

[19] Traub, J.F. (1964): Methods for the Solution of Equations. Prentice-Hall, Englewood Cliffs.

[20] Varga, R.S. (1962): Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs.

Received in May 1999.