Numerical Algorithms 9(1995)223–244

A universal constant for the convergence of Newton's method and an application to the classical homotopy method

Jean-Claude Yakoubsohn

Laboratoire d'Approximation et d'Optimisation, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse cedex, France

Communicated by J. Della Dora

We give a new theorem concerning the convergence of Newton's method to compute an approximate zero of a system of equations. In this result, the constant $h_0 = 0.162434\ldots$ appears, which plays a fundamental role in the localization of "good" initial points for the Newton iteration. We apply it to the determination of an appropriate discretization of the time interval in the classical homotopy method.

Keywords: Newton method, homotopy, approximate zero.

AMS subject classification: 65H10.

0. Introduction

In this paper we consider the algebraic system

$$P(x) = 0$$

with $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $P(x) = (P_1(x), \ldots, P_n(x))$ where the $P_i(x)$ are polynomials in $\mathbb{R}^n[x]$ of degree d_i, $1 \leqslant i \leqslant n$, with $d = \max_{1 \leqslant d \leqslant n} d_i$.

We denote by $\alpha = (\alpha_1, \ldots, \alpha_n)$ a multi-index and by $|\alpha| = \alpha_1 + \ldots + \alpha_n$. For $k \in \mathbb{N}$, we use the classical notation:

$$\binom{k}{\alpha} = \frac{k!}{\alpha_1! \ldots \alpha_n!} \quad \text{and} \quad x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}.$$

For x and y in $\mathbb{R}^n$, we write the Taylor formula

$$P(y) = P(x) + \sum_{k=1}^{d} \frac{1}{k!} D^k P(x)(y - x)^k,$$

where $D^kP(x)(y-x)^k$ is the vector

$$\left(\sum_{|\alpha|=k}\binom{k}{\alpha}\frac{\partial^k P_1(x)}{\partial x^\alpha}(y-x)^\alpha,\ldots,\sum_{|\alpha|=k}\binom{k}{\alpha}\frac{\partial^k P_n(x)}{\partial x^\alpha}(y-x)^\alpha\right).$$

Furthermore, $D^k_\alpha P(x)$ is the vector

$$D^k_\alpha P(x)=\left(\frac{\partial^k P_1(x)}{\partial x^\alpha},\ldots,\frac{\partial^k P_n(x)}{\partial x^\alpha}\right).$$

We also shall denote by $D^{k+1}P(x)(y-x)^k$ the matrix with the coefficients

$$\begin{aligned}&(D^{k+1}P(x)(y-x)^k)_{ij}\\&=\sum_{\substack{|\alpha|=k+1\\ \alpha_1=\ldots=\alpha_{j-1}=0,\alpha_j\geqslant 1}}\binom{k+1}{\alpha}\frac{\partial^{k+1}P_i(x)}{\partial x^\alpha}\frac{(y-x)^\alpha}{y_j-x_j},\quad 1\leqslant i,j\leqslant n,\end{aligned}$$

which verifies $D^{k+1}P(x)(y-x)^k(y-x)=D^{k+1}P(x)(y-x)^{k+1}$. We use the max-norm i.e. $\|x\|=\max_{1\leqslant i\leqslant n}|x_i|$ for $x\in\mathbb{R}^n$ and $|||A|||=\max_{1\leqslant i\leqslant n}\sum_{j=1}^n|A_{ij}|$ for a matrix A. The open ball centered at x and of radius r associated to the max-norm is $B_{max}(x,r)$.

The max-norm of the matrix $D^{k+1}P(x)(y-x)^k$ is bounded by

$$|||D^{k+1}P(x)(y-x)^k|||\leqslant\max_{1\leqslant i\leqslant n}\sum_{|\alpha|=k+1}\binom{k+1}{\alpha}\left|\frac{\partial^{k+1}P_i(x)}{\partial x^\alpha}\right|\|y-x\|^k.$$

For this reason we introduce the quantities:

$$|||D^kP(x)|||\stackrel{\text{def}}{=}\max_{1\leqslant i\leqslant n}\sum_{|\alpha|=k}\binom{k}{\alpha}\left|\frac{\partial^k P_i(x)}{\partial x^\alpha}\right|.$$

These quantities, different from the max-norm of the operator $D^kP(x)$, occur naturally in the estimations established from the Taylor formula: see lemma 2.1.

The main goal of this paper is to prove the following theorem.

Main theorem

Let us consider an algebraic system $P(x)=0$ defined as previously. Let $h_0=1.62434\ldots$ be the smallest root of polynomial $4h^3-12h^2+8h-1$. Let $x^0\in\mathbb{R}^n$ and $h\in[0,h_0]$ be such that the inequalities

$$\frac{1}{k!}|||D^kP(x^0)|||\,|||DP(x^0)^{-1}|||^k\|P(x^0)\|^{k-1}\leqslant h^{k-1},\quad 2\leqslant k\leqslant d,\tag{1}$$

are satisfied.

(1) Then the sequence of vectors in $\mathbb{R}^n$ defined by

$$x^{p+1}=x^p-DP(x^p)^{-1}P(x^p)$$

converges to a simple solution x^* of the algebraic system $P(x)=0$.

(2) Let $h \in [0, h_0[$. The convergence of the sequence x_p is super-quadratic, i.e.

$$||x^{p+1} - x^p|| \leqslant a^p \left(\frac{h}{a^2}\right)^{2^p - 1} ||x^1 - x^0||,$$

where $a = 2h_0^2 - 4h_0 + 1 = 0.403031\ldots$ and $h/a^2 < 1$.

(3) For $x \in \mathbb{R}^n$ let us define the polynomial

$$\bar{L}(x,t) = -||DP(x)^{-1}P(x)|| + tL(x,t)$$

where

$$L(x,t) = 1 - \sum_{k=2}^{l} \frac{1}{k!} \left(\sum_{|\alpha|=k} \binom{k}{\alpha} ||DP(x)^{-1} D_\alpha^k P(x)|| \right) t^{k-1}.$$

We denote by $l^+(x)$ and $l^-(x)$ the positive roots (when they exist) of the polynomial $\bar{L}(x,t)$ and by $l(x)$ the positive root of the polynomial $L(x,t)$.

Then the union of the balls $B_{max}(x^p, l^+(x^p))$ for indices p such that $l(x^*) \geqslant l^-(x^p)$ contains only one solution of the algebraic system $P(x) = 0$ which is x^*.

We now can explain the considerations on which this paper is based. For this, we begin by a short digression. In [2], Dedieu and the author have studied exclusion algorithms in view of localizing all the real solutions of a system of algebraic equations which admits a finite number of roots. These algorithms consist in computing at a given point $x \in \mathbb{R}^n$ a ball $B_{max}(x, m(x))$ in which there is no solution of the system. The radius of this ball is named the exclusion function at x. The current exclusion function for the system $P(x) = 0$ considered here is $m(x) = \max_{1 \leqslant i \leqslant n} m_i(x)$, where each of $m_i(x)$ is the positive root of a concave polynomial in $\mathbb{R}[t]$ defined by:

$$M_i(x,t) = ||P_i(x)|| - \sum_{k=1}^{d_i} \sum_{|\alpha|=k} \binom{k}{\alpha} \left| \frac{\partial^k P_i(x)}{\partial x^\alpha} \right| t^k.$$

This exclusion function is in some sense equivalent to the distance of x to the solution set.

An initial ball and an accuracy ϵ being given, an exclusion algorithm consists in choosing a point x in an initial ball in which we want to localize the solutions of the system, to exclude the ball $B_{max}(x, m(x))$ of the initial ball and start again with another point x while the remainder is a non-empty set. This algorithm returns all the balls $B_{max}(x, m(x))$ such that $m(x) \leqslant \epsilon$: these balls are susceptible to contain solutions of the system. The complexity of this algorithm is the number of computations of the exclusion function: it is proportional to $\mathrm{Log}\, 1/\epsilon$.

In these exclusion algorithms, the number of exclusion tests is large near a solution. It is a natural idea to test if the hypotheses of the convergence of the Newton method are satisfied as soon as $m(x) \leqslant \epsilon$. If that is the case we must compute two quantities: an approximation of the solution x^* and a ball in which there is only the solution x^*. Then we prove in [2] that the complexity of the exclusion algorithm conjugated with the Newton method is in Log|Log ϵ| exclusion tests. The two parts of this program can be realized with a classic Newton–Kantorovitch Theorem, see [3,5] or [6]. For example, the translation in the polynomial case of the result which appears in [3, p. 263] gives:

Newton–Kantorovitch Theorem
Let $x^0 \in \mathbb{R}^n$, $P(x) \in \mathbb{R}^n[x]$ and the ball $B = \bar{B}(x^0, 2\|DP(x_0)^{-1}P(x_0)\|)$ be such that the condition:

$$2n|||DP(x^0)^{-1}|||\,\|DP(x^0)^{-1}P(x^0)\| \max_{i,j,z\in B} \sum_k \left| \frac{\partial^2 P_i(z)}{\partial x_j \partial x_k} \right| \leqslant 1$$

is satisfied.

Then the sequence $x^{p+1} = x^p - DP(x^p)^{-1}P(x^p)$ converges to the unique solution of the system $P(x) = 0$ in the ball $\bar{B}(x^0, 2\|DP(x_0)^{-1}P(x_0)\|)$.

In the context of exclusion algorithms there are two reasons for applying the main theorem instead of this theorem. The first is that the quantities $|||DP^k(x)|||$ are determined for computing the function $m(x)$. In other words, this strategy avoids the computation of the quantity $\max \sum |\partial^2 P_i(z)/\partial x_j \partial x_k|$ in the Newton–Kantorovitch theorem.

Next, the radius $2\|DP(x_0)^{-1}P(x_0)\|$ is in general small and the introduction of the polynomial $\bar{L}(x, t)$ in the main theorem improves the radius of the ball of unicity of the solution.

A modern approach of the convergence of the Newton method must be attributed to Smale in [9] in which the case of roots of a complex polynomial $P(z)$ is studied. In this paper the inequalities equivalent to (1) are:

$$\left| \frac{P^{(k)}(z)P(z)^{k-1}}{k!P'(z)^k} \right| \leqslant \min_{\theta: P'(\theta)=0} \left| \frac{1}{P(\theta) - P(z)} \right|.$$

Shub and Smale in [7] give the historical origin of these ratios which go back to Newton and Euler.

More recently, in [8], these two authors generalize the results concerning complex polynomials to systems of analytic equations. The technical background and the results developed here are different from this last paper.

In section 1 we give preliminary lemmas and we introduce polynomials which appear in the study of the convergence. The goal of sections 2 (convergence), 3 (complexity) and 4 (set of unicity) is to prove the main theorem. The stability is the object of section 5 and section 6 gives an application to an efficient discretization of the time interval [0, 1] in the classical homotopy method.

Technical lemmas used in the various sections are collected in section 1 and the reader eager for knowledge can start with section 2.

1. Preliminary lemmas

In the proof of lemma 2.1 we shall apply the following

Lemma 1.1
Let k and j be two integers. Let α and β be two multi-indices such that $|\alpha| = k$ and $|\beta| = j$. Then

$$\binom{\alpha_1+\beta_1}{\alpha_1}\cdots\binom{\alpha_n+\beta_n}{\alpha_n} \leqslant \binom{k+j}{j}.$$

Proof
By induction on the sum $k+j$. The inequality holds trivially for $k+j=1$. Suppose the inequality of all α and β such that $|\alpha|+|\beta| = k+j$. We have

$$\begin{aligned}\binom{\alpha_1+\beta_1+1}{\alpha_1+1}\binom{\alpha_2+\beta_2}{\alpha_2}\cdots\binom{\alpha_n+\beta_n}{\alpha_n} &= \binom{\alpha_1+\beta_1}{\alpha_1}\binom{\alpha_2+\beta_2}{\alpha_2}\\ \cdots\binom{\alpha_n+\beta_n}{\alpha_n} &+ \binom{\alpha_1+\beta_1}{\alpha_1+1}\binom{\alpha_2+\beta_2}{\alpha_2}\cdots\binom{\alpha_n+\beta_n}{\alpha_n}\\ &\leqslant \binom{k+j}{j}+\binom{k+j}{j-1} \quad \text{by the induction hypothesis}\\ &\leqslant \binom{k+j+1}{j}.\end{aligned}$$

Hence the inequality holds for all α and β such that $|\alpha|+|\beta| = k+j+1$. And the lemma is proved. □

We now introduce the following polynomials which play an important part in the study of the convergence:

$$T_d(h) = 1-\sum_{i=1}^{d-1}(i+1)h^i,$$

$$S_{kd}(h) = \sum_{i=0}^{d-k}\binom{k+i}{i}h^i,$$

$$Y_{kd}(h) = h^{k-1}S_{kd}(h)-(1-h)^{k-1}T_d^k(h).$$

Concerning these polynomials, we shall use the following

Lemma 1.2
Let $k \geqslant 2$, $d \geqslant 2$ and h be in the interval $[0, (5-\sqrt{13})/6]$. Then

$$S_{kd}^2(h) \leqslant S_{2d}^k(h).$$

Before proving this lemma we establish the following:

Lemma 1.3
Let $d > 2$ and h be in the interval $[0, (5-\sqrt{13})/6]$. We have

$$1 \leqslant (1-h)^2 S_{2d}(h).$$

Proof
By definition of the polynomial $S_{2d}(h)$ we have $S_{2d}(h) \geqslant 1 + 3h$. Hence $(1-h)^2 S_{2d}(h) \geqslant (1-h)^2(1+3h)$. And it is easy to verify the polynomial $(1-h)^2(1+3h) = 1 + h(3h^2 - 5h + 1)$ is greater than 1 on the interval $[0, (5-\sqrt{13})/6]$. □

Proof of lemma 1.2
We proceed by induction. The inequality is verified for $k = 2$. Suppose the inequality holds for k. We have

$$\begin{aligned} S_{k+1d}(h) &= \sum_{i=0}^{d-k-1} \binom{k+i+1}{i} h^i \\ &= \sum_{i=1}^{d-k-1} \binom{k+i}{i-1} h^i + \sum_{i=0}^{d-k-1} \binom{k+i}{i} h^i \\ &= h \sum_{i=0}^{d-k-2} \binom{k+i+1}{i} h^i + S_{kd}(h) - \binom{d}{d-k} h^{d-k} \\ &\leqslant h S_{k+1d}(h) + S_{kd}(h). \end{aligned}$$

Thus we have the inequality

$$S_{k+1d}(h) \leqslant \frac{1}{1-h} S_{kd}(h).$$

And by lemma 1.3 it follows that

$$S_{k+1d}^2(h) \leqslant S_{kd}^2(h) S_{2d}(h)$$

on the interval $[0, (5-\sqrt{13})/6]$. By the induction hypothesis we have $S_{kd}^2(h) \leqslant S_{2d}^k(h)$. Thus $S_{k+1d}^2(h) \leqslant S_{2d}^{k+1}(h)$ and the lemma holds. □

Lemma 1.4
(1) The positive roots of polynomials $T_d(h)$ form a strictly decreasing sequence of real numbers which converges to $(2-\sqrt{2})/2$.

(2) The polynomials Y_{kd} have only one positive root denoted by y_{kd} on the interval $]0,(2-\sqrt{2})/2[$.
(3) The sequence of roots $(y_{2d})_{d\geqslant 2}$ of the polynomials $Y_{2d}(h)$ is a strictly decreasing sequence which converges to the smallest root h_0 of the polynomial $4h^3-12h^2+8h-1$. An approximate value of this root is $h_0 = 0.162434565\ldots$.
(4) $Y_{kd}(y_{2d})<0$ for all k, $3\leqslant k\leqslant d$.
(5) We give a table of first values of y_{2d}:

d	y_{2d}	d	y_{2d}
2	0.228155	5	0.162916
3	0.173091	6	0.162537
4	0.164665		

Proof

(1) By Descartes' rule the polynomial $T_d(h)$ has only one positive root. More precisely since $T_d(0)=1$ and $T_d(1)=2-d-(d(d-1))/2<0$ this root is in the interval $[0,1]$. Furthermore, $T_{d+1}(h)=T_d(h)-h^{d+1}<T_d(h)$ on $]0,\infty[$. We deduce that the sequence of the positive roots of the polynomials $T_d(h)$ is a strictly decreasing sequence. We now prove the convergence of this sequence. The polynomial $1-T_d(h)$ is the derivative of the polynomial $\sum_{i=1}^{d-1}h^{i+1}$. We have

$$\left(\sum_{i=1}^{d-1}h^{i+1}\right)'=\left(\frac{h^2-h^{d+1}}{1-h}\right)'=\frac{2h-h^2-(d+1)h^d+dh^{d+1}}{(1-h)^2}.$$

Hence

$$T_d(h)=1-\frac{2h-h^2-(d+1)h^d+dh^{d+1}}{(1-h)^2}$$
$$=\frac{1-4h+2h^2+h^d(d+1-dh)}{(1-h)^2}.$$

We now fix $h\in]0,1[$. Then we have

$$\lim_{d\to\infty}T_d(h)=\frac{1-4h+2h^2}{(1-h)^2},$$

with

$$T_d(h)>\frac{1-4h+2h^2}{(1-h)^2}, \tag{2}$$

since $d+1-dh>0$. Consequently, the sequence of the positive roots of polynomials $T_d(h)$ converges to the root of the polynomial $1-4h+2h^2$ which lies inside $]0,1[$, i.e. $(2-\sqrt{2})/2$.

(2) Let t_d be the positive root of the polynomial $T_d(h)$ and $2\leqslant k\leqslant d$. We have $Y_{kd}(0)=-1$ and $Y_{kd}(t_d)=t_d^{k-1}S_{kd}(t_d)>0$. To prove the unicity of the root y_{kd} on

the interval $]0, t_d[$, we claim that $Y'_{kd}(h)$ is positive on this interval. In fact,

$$Y'_{kd}(h) = (k-1)h^{k-2}S_{kd}(h) + h^{k-1}S'_{kd}(h) + (k-1)(1-h)^{k-2}T_d^k(h) - k(1-h)^{k-1}T'_d(h)T_d^{k-1}(h).$$

On the interval $]0, t_d[$ we have $T'_d(h) < 0$ and then by the previous expression $Y'_{kd}(h) > 0$ for all $h \in]0, t_d[$.

(3) We have

$$Y_{2d+1}(h) - Y_{2d}(h) = h(S_{2d+1}(h) - S_{2d}(h)) + (1-h)(T_d^2(h) - T_{d+1}^2(h)) = \binom{d+1}{d-1}h^d + (1-h)(T_d(h) + T_{d+1}(h))(d+1)h^d.$$

Hence $Y_{2d+1}(h) - Y_{2d}(h) > 0$ on the interval $]0, (2-\sqrt{2})/2[$. Thus, applying the results of part (2) we have $y_{2d+1} < y_{2d}$.

To prove the convergence of this sequence we shall establish

Lemma 1.5

Let $h \in]0, (2-\sqrt{2})/2[$. Then we have

$$\lim_{d\to\infty} Y_{2d}(h) = \frac{4h^3 - 12h^2 + 8h - 1}{(1-h)^2},$$

with

$$Y_{2d}(h) < \frac{4h^3 - 12h^2 + 8h - 1}{(1-h)^2}.$$

Consequently the polynomial $4h^3 - 12h^2 + 8h - 1$ is increasing on the interval $]0, (2-\sqrt{2})/2[$ and $h_0 = 0.162434565$ is the root of this polynomial in this interval. Furthermore, $\lim_{d\to\infty} y_{2d} = h_0$ with $h_0 < y_{2d}$ and part (3) of lemma 1.4 is proved. □

Proof of lemma 1.5

We observe

$$S_{2d}(h) = 1 + \frac{1}{2}\left(\sum_{i=1}^{d-2} h^{i+2}\right)'' = 1 + \frac{1}{2}\left(\frac{h^3 + h^{d+1}}{1-h}\right)'' = 1 + \frac{1}{2}\left(\frac{6h - 6h^2 + 2h^3}{(1-h)^3}\right) + \frac{h^{d-1}}{(1-h)^3}(-d(d-1)h^2 + 2(d^2-1)h - d(d+1)).$$

We have $-d(d-1)h^2+2(d^2-1)h-d(d+1)<0$ and $\lim_{d\to\infty} h^{d-1}(-d(d-1)h^2+2(d^2-1)h-d(d+1))=0$ since $h\in]0,(2-\sqrt{2})/2[$. Hence

$$\lim_{d\to\infty} S_{2d}(h)=1+\frac{3h-3h^2+h^3}{(1-h)^3},$$

with

$$S_{2d}(h)<1+\frac{3h-3h^2+h^3}{(1-h)^3}.$$

Using inequality (2), we obtain

$$\begin{aligned}\lim_{d\to\infty} Y_{2d}(h)&=\left(1+\frac{3h-3h^2+h^3}{(1-h)^3}\right)-(1-h)\left(\frac{1-4h+2h^2}{(1-h)^2}\right)^2\\&=\frac{4h^3-12h^2+8h-1}{(1-h)^2},\end{aligned}$$

with

$$Y_{2d}(h)<\frac{4h^3-12h^2+8h-1}{(1-h)^2}.$$

Hence lemma 1.5 holds. □

(4) The positive root y_{2d} of the polynomial $Y_{2d}(h)$ verifies the relation

$$y_{2d}S_{2d}(y_{2d})=(1-y_{2d})T_d^2(y_{2d}). \tag{3}$$

On the other hand, we have $y_{2d}<\ldots<y_{22}$. As $y_{22}\sim 0.228<(5-\sqrt{13})/6\sim 0.232$ the inequality of lemma 1.2, $S_{kd}^2(h)\leqslant S_{2d}^k(h)$, holds on the interval $[0,y_{22}]$ which contains all the roots y_{kd}. By the identity (3) and using $y_{kd}<1-y_{kd}$ on this interval, we have successively

$$y_{2d}^{k-2}S_{kd}^2(y_{2d})\leqslant(1-y_{2d})^{k-2}S_{2d}^k(y_{2d}),$$

$$y_{2d}^{2k-2}S_{kd}^2(y_{2d})\leqslant(1-y_{2d})^{k-2}y_{2d}^kS_{2d}^k(y_{2d})=(1-y_{2d})^{2k-2}T_d^{2k}(y_{2d}),$$

$$y_{2d}^{k-1}S_{kd}(y_{2d})\leqslant(1-y_{2d})^{k-1}T_d^k(y_{2d}).$$

This previous inequality is $Y_{kd}(y_{2d})\leqslant 0$. □

The following lemma shall be used in the study of the complexity.

Lemma 1.6
Let us introduce the real function

$$\phi(h)=\frac{h}{2h^2-4h+1}\quad\text{for}\in[0,h_0].$$

Then

$$h^{k-1}\left(\frac{h}{1-h}\right)^{k-1}\frac{S_{kd}(h)}{T_d^k(h)} \leqslant (\phi(h)^2)^{k-1}, \quad 2 \leqslant k \leqslant d.$$

Proof

We remark that the polynomial $S_{kd}(h)$ is the Taylor series expansion of order $d-k+1$ at 0 of $1/(1-h)^{k+1}$. Hence $S_{kd}(h) \leqslant 1/(1-h)^{k+1}$. From (2) we obtain

$$h^{k-1}\left(\frac{h}{1-h}\right)^{k-1}\frac{S_{kd}(h)}{T_d^k(h)} \leqslant \frac{h^{2k-2}}{(2h^2-4h+1)^k} \leqslant (\phi(h)^2)^{k-1}, \quad 2 \leqslant k \leqslant d.$$

The last inequality is satisfied since $2h^2-4h+1 \leqslant 1$ for $h \in [0, h_0]$. Hence the conclusion of this lemma is proved. □

2. Convergence

To establish the first part of the main theorem, we prove the following:

Lemma 2.1

Let $x \in \mathbb{R}^n$ be such that $DP(x)^{-1}$ exists. Let us consider

$$y = x - DP(x)^{-1}P(x).$$

We have the following estimations:

(1)

$$\frac{||P(y)||}{||P(x)||} \leqslant \sum_{k=2}^{d}\frac{1}{k!}|||D^kP(x)|||\ |||DP(x)^{-1}|||^k||P(x)||^{k-1}. \tag{4}$$

(2)

$$\frac{1}{k!}|||D^kP(x)||| \leqslant \sum_{i=0}^{d-k}\binom{k+i}{i}\frac{1}{(k+i)!}|||D^{k+i}P(x)|||\ |||DP(x)^{-1}|||^i||P(x)||^i. \tag{5}$$

(3) If

$$\sum_{i=1}^{d-1}(i+1)\frac{1}{(i+1)!}|||D^{i+1}P(x)|||\ |||DP(x)^{-1}|||^{i+1}||P(x)||^i < 1, \tag{6}$$

then $DP(y)^{-1}$ exists and we have

$$|||DP(y)^{-1}||| \leqslant \frac{|||DP(x)^{-1}|||}{1-\sum_{i=1}^{d-1}(i+1)\frac{1}{(i+1)!}|||D^{i+1}P(x)|||\ |||DP(x)^{-1}|||^{i+1}||P(x)||^i}. \tag{7}$$

Proof
By definition of y it follows that $P(x) + DP(x)(y - x) = 0$. Hence

$$P(y) = \sum_{k=2}^{d} \frac{1}{k!} D^k P(x)(-DP(x)^{-1}P(x))^k.$$

(1) The estimation on $||P(y)||$ follows easily from the previous formula.
(2) To prove the estimation on $(1/k!)|||D^k P(x)|||$ we write for $1 \leqslant i \leqslant n$ and $2 \leqslant k \leqslant d$,

$$\frac{\partial^k P_i(x)}{\partial x^\alpha} = \sum_{j=0}^{d-k} \frac{1}{j!} \sum_{|\beta|=j} \binom{j}{\beta} \frac{\partial^{k+j} P_i(x)}{\partial x^{\alpha+\beta}} (-DP(x)^{-1}P(x))^\beta.$$

Consequently,

$$\sum_{|\alpha|=k} \frac{1}{k!} \binom{k}{\alpha} \left| \frac{\partial^k P_i(x)}{\partial x^\alpha} \right|$$

$$\leqslant \sum_{j=0}^{d-k} \binom{k+j}{j} \frac{1}{(k+j)!} \sum_{\substack{|\alpha|=k \\ |\beta|=j}} \binom{k}{\alpha} \binom{j}{\beta} \left| \frac{\partial^{k+j} P_i(x)}{\partial x^{\alpha+\beta}} \right| |(DP(x)^{-1}P(x))^\beta|$$

$$\leqslant \sum_{j=0}^{d-k} \binom{k+j}{j} \frac{1}{(k+j)!} \sum_{|\gamma|=k+j} \binom{k+j}{\gamma} \left| \frac{\partial^{k+j} P_i(x)}{\partial x^\gamma} \right| |||(DP(x)^{-1}|||^j \, ||P(x)||^j.$$

This previous inequality holds since $\sum_{|\alpha|=k, |\beta|=j} \leqslant \sum_{|\gamma|=k+j}$ and by lemma 1.1. Thus the estimation (5) follows easily.
(3) Estimation on $|||DP(y)^{-1}|||$. We have successively

$$|||I - DP(x)^{-1}DP(y)||| \leqslant \sum_{i=1}^{d-1} \frac{1}{i!} |||DP(x)^{-1} D^{i+1} P(x)(DP(x)^{-1}P(x))^i|||$$

$$\leqslant \sum_{i=1}^{d-1} \frac{1}{i!} |||DP(x)^{-1}||| \; |||D^{i+1} P(x)(DP(x)^{-1}P(x))^i|||$$

$$\leqslant \sum_{i=1}^{d-1} (i+1) \frac{1}{(i+1)!} |||D^{i+1} P(x)||| \; |||DP(x)^{-1}|||^{i+1} ||P(x)||^i.$$

By hypothesis we deduce $|||I - DP(x)^{-1}DP(y)||| < 1$. Then by [3, p. 264],

$$DP(y)^{-1} = (DP(x)(I - (I - DP(x)^{-1}DP(y))))^{-1}$$

exists and we have

$$|||DP(y)^{-1}||| \leqslant \frac{|||DP(x)^{-1}|||}{1 - |||I - DP(x)^{-1}DP(y)|||}.$$

Thus the estimation (7) holds. □

We now prove the convergence of the Newton sequence where the inequalities (1) are satisfied at the point x^0. In fact, we prove that we can take $h \in [0, y_{2d}]$ instead of $h \in [0, h_0]$ as we have claimed in the statement of the main theorem where y_{2d} is defined in lemma 1.4.

Using (1) and lemma 2.1 we have successively

$$||P(x^1)|| \leqslant \sum_{k=2}^{d} h^{k-1} ||P(x^0)|| \leqslant \frac{h}{1-h} ||P(x^0)||,$$

$$\frac{1}{k!} |||D^k P(x^1)||| \leqslant \left(\sum_{i=0}^{d-k} \binom{k+i}{i} h^i \right) \frac{h^{k-1}}{|||DP(x^0)^{-1}|||^k ||P(x^0)||^{k-1}},$$

$$|||DP(x^1)^{-1}||| \leqslant \frac{|||DP(x^0)^{-1}|||}{1 - \sum_{i=1}^{d-1} (i+1) h^i}.$$

Using the definition of the polynomials $S_{kd}(h)$ and $T_d(h)$ we finally obtain

$$\frac{1}{k!} |||D^k P(x^1)||| \, |||DP(x^1)^{-1}|||^k ||P(x^1)||^{k-1} \leqslant h^{k-1} \left(\frac{h}{1-h} \right)^{k-1} \frac{S_{kd}(h)}{T_d^k(h)}, \quad 2 \leqslant k \leqslant d,$$

since, by lemma 1.2, $T_d(h) > 0$ for $h \in [0, y_{2d}]$. We also have on this interval $Y_{kd}(h) \leqslant 0$. It is equivalent to

$$\left(\frac{h}{1-h} \right)^{k-1} \frac{S_{kd}(h)}{T_d^k(h)} \leqslant 1.$$

Hence the inequalities (1) hold at x^1.

By induction we deduce that the inequalities are satisfied for all points x^p and we have

$$||P(x^p)|| \leqslant \left(\frac{h}{1-h} \right)^p ||P(x^0)||.$$

Hence $\lim_{p \to \infty} |P(x^p)| = 0$ since $h/(1-h) < 1$ and by continuity the sequence $(x^p)_{p \geqslant 0}$ converges to a solution x^* of the algebraic system $P(x) = 0$.

We now prove that x^* is a simple root. Let us suppose x^* is a double solution. For p sufficiently large we have

$$\begin{aligned} &\tfrac{1}{2} |||D^2 P(x^p)||| \, |||DP(x^p)^{-1}|||^2 ||P(x^p)|| \\ &\sim \tfrac{1}{2} |||D^2 P(\alpha)||| \, |||(DP^2(\alpha)(x^p - \alpha))^{-1}|||^2 || \tfrac{1}{2} D^2 P(x)(x-\alpha)^2 || \\ &\geqslant \tfrac{1}{4} |||D^2 P(\alpha)||| \frac{||(DP^2(\alpha)(x^p - \alpha))^{-1} D^2 P(x)(x-\alpha)(x-\alpha)||}{||DP^2 P(x)(x-\alpha)||} \\ &\geqslant \tfrac{1}{4} |||D^2 P(\alpha)||| \frac{1}{||DP^2 P(x)|| \, ||(x-\alpha)||} ||(x-\alpha)|| = \tfrac{1}{4}. \end{aligned}$$

Then, if x^* is a double solution, the inequalities (1) are not satisfied for $k = 2$. Hence the convergence from a simple root of Newton's method under the hypothesis (1) is established. □

3. Complexity

We prove the following

Lemma 3.1
Let $x \in \mathbb{R}^n$ be such that $DP(x)^{-1}$ exists and the inequality (6) holds. Let us consider

$$y = x - DP(x)^{-1}P(x).$$

We have the following estimations:
(1)

$$|||DP(y)^{-1}DP(x)||| \leqslant \frac{1}{1 - \sum_{i=1}^{d-1}(i+1)\frac{1}{(i+1)!}|||D^{i+1}P(x)|||\,|||DP(x)^{-1}|||^{i+1}||P(x)||^i}; \tag{8}$$

(2)

$$\frac{||DP(x)^{-1}P(y)||}{||DP(x)^{-1}P(x)||} \leqslant \sum_{k=2}^{d}\frac{1}{k!}|||D^kP(x)|||\,|||DP(x)^{-1}|||^k||P(x)||^{k-1}. \tag{9}$$

Proof
(1) The estimation (8) results directly from $|||DP(y)^{-1}DP(x)||| \leqslant |||DP(y)^{-1}|||\,|||DP(x)^{-1}|||$ and from inequality (7).
(2) Let us prove the estimation (9). We have successively

$$\begin{aligned}
&||DP(x)^{-1}P(y)||\\
&\leqslant \left\|\sum_{k=2}^{d}\frac{1}{k!}DP(x)^{-1}D^kP(x)(-DP(x)^{-1}P(x))^{k-1}(-DP(x)^{-1}P(x))\right\|\\
&\leqslant \sum_{k=2}^{d}\frac{1}{k!}|||DP(x)^{-1}|||\,|||D^kP(x)(-DP(x)^{-1}P(x))^{k-1}|||\,||DP(x)^{-1}P(x)||.
\end{aligned}$$

The reader may easily conclude the proof. □

We now prove part (2) of the main theorem. If the equalities are satisfied at x, then from lemma 1.6 we have for $y = x - DP(x)^{-1}P(x)$:

$$\frac{1}{k!}|||D^kP(y)|||\,|||DP(y)^{-1}|||^k||P(y)||^{k-1} \leqslant (\phi(h)^2)^{k-1}.$$

We recall $\phi(h) = h/(2h^2 - 4h + 1)$. For this reason we introduce the sequence $\beta_0 = h \in [0, h_0[$, $\beta_p = \phi(\beta_{p-1})^2$. Let $a = 2h_0^2 - 4h_0 + 1 = 0.403031\ldots$. We have $\phi(h) \leqslant h/a$ and $\beta_1 \leqslant a^2(h/a^2)^{2^1}$ where $h/a^2 < h_0/(2h_0^2 - 4h_0 + 1)^2 = 1$. We easily prove by induction:

$$\beta_p \leqslant a^2\left(\frac{h}{a^2}\right)^{2^p}. \tag{10}$$

We now return to the Newton sequence (x^p). From inequalities (8) and (9) we have

$$\begin{aligned}||x^{p+1} - x^p|| &= ||DP(x^p)^{-1}P(x^p)|| \\ &\leqslant |||DP(x^p)^{-1}DP(x^{p-1})|||\, ||DP(x^{p-1})^{-1}P(x^p)|| \\ &\leqslant \frac{1}{T_d(\beta_{p-1})}\frac{\beta_{p-1}||DP(x^{p-1})^{-1}P(x^{p-1})||}{1-\beta_{p-1}}.\end{aligned}$$

From inequalities (2) and (10) we obtain

$$\begin{aligned}||x^{p+1} - x^p|| &\leqslant \phi(\beta_{p-1})(1-\beta_{p-1})||x^p - x^{p-1}|| \leqslant \frac{\beta_{p-1}}{a}||x^p - x^{p-1}|| \\ &\leqslant a\left(\frac{h}{a^2}\right)^{2^{p-1}}||x^p - x^{p-1}||.\end{aligned}$$

Finally,

$$||x^{p+1} - x^p|| \leqslant a^p\left(\frac{h}{a^2}\right)^{2^p - 1}||x^1 - x^0||$$

and part (2) of the main theorem follows. □

Corollary 3.2
Let $a = 2h_0^2 - 4h_0 + 1$ and $c = \frac{1}{2}(\log a \log(h/a^2)\log 2)^{1/2}$.

(1) For all $p \geqslant 0$ we have $||x^* - x^p|| \leqslant c(a/\sqrt{2})^{p-1}(h/a^2)^{2^{p-1}-1}||x^1 - x^0||$.
(2) If $h = 0.162$ then $c = 12.693384$.

Proof
Using the estimation on $||x^p - x^{p-1}||$ we write

$$\begin{aligned}||x^{p+q} - x^p|| &\leqslant \sum_{k=0}^{q-1}||x^{p+k+1} - x^{p+k}|| \\ &\leqslant \sum_{k=0}^{q-1} a^{p+k}\left(\frac{h}{a_2}\right)^{2^{p+k}-1}||x^1 - x^0||.\end{aligned}$$

We now give an upper bound for the series $\sum_{k=p}^{\infty} a^k b^{2^k - 1}$, where $b = h/a^2$. We have successively

$$\begin{aligned}\sum_{k=p}^{\infty} a^k b^{2^k-1} &\leqslant \left(\sum_{k=p}^{\infty} a^{2k}\right)^{1/2}\left(\sum_{k=p}^{\infty} b^{2(2^k-1)}\right)^{1/2}\\ &\leqslant \left(\int_{p-1}^{\infty} a^{2s}ds\right)^{1/2}\left(\int_{p-1}^{\infty} b^{2(2^s-1)}ds\right)^{1/2}\\ &\leqslant \left(\int_{p-1}^{\infty} a^{2s}ds\right)^{1/2}\left(\int_{p-1}^{\infty} \frac{2^s}{2^{p-1}} b^{2(2^s-1)}ds\right)^{1/2}\\ &\leqslant \frac{1}{2}\left(\log a \log\frac{h}{a^2}\log 2\right)^{-1/2}\left(\frac{a}{\sqrt{2}}\right)^{p-1}\left(\frac{h}{a^2}\right)^{2^{p-1}-1}.\end{aligned}$$

And the corollary follows. □

4. Set of unicity of a solution

Part (3) of the main theorem results from two lemmas

Lemma 4.1
Let $x \in \mathbb{R}^n$ such that $DP(x)^{-1}$ exists.

(1) The polynomial $L(x,t)$ is a strictly concave polynomial on the interval $[0,+\infty[$ which possesses only one positive root denoted by $l(x)$.
(2) The polynomial $\bar{L}(x,t)$ is a concave polynomial on $\mathbb{R}$ which possesses either no real root or two nonnegative roots $l^+(x)$ and $l^-(x)$ such that $0 \leqslant l^-(x) \leqslant l^+(x)$.
(3) Let us consider a simple root x^* of the algebraic system.
   (3.1) The functions $l^+(x)$ and $l^-(x)$ are well defined in the neighbourhood of x^* and are continuous.
   (3.2) $\lim_{x\to x^*} l^+(x) = l(x)$ and $\lim_{x\to x^*} l^-(x) = 0$.

Proof
The first part is left to the reader and the second follows directly from Descartes' rule. For the third part we note that $\lim_{x\to x^*} \bar{L}(x,t) = tL(x^*,t)$. The roots of this polynomial are 0 and $l(x)$. By continuity of the roots of a monic polynomial we conclude the proof of this lemma. □

We now give a lower bound for the distance between two solutions of the algebraic system.

Lemma 4.2
Let x^* be a simple solution of the algebraic system $P(x)=0$. Then $\|y^* - x^*\| \geqslant l(x^*)$ for all solutions $y^* \neq x^*$.

Proof
By Taylor's formula and the fact that $P(x^*) = P(y^*) = 0$ we have

$$DP(x^*)(y^* - x^*) + \sum_{k=2}^{d} \frac{1}{k!} D^k P(x^*)(y^* - x^*)^k = 0.$$

Since $DP(x^*)^{-1}$ exists, we have

$$\begin{aligned} 0 &= \left\| y^* - x^* + \sum_{k=2}^{d} \frac{1}{k!} DP(x^*)^{-1} D^k P(x^*)(y^* - x^*)^k \right\| \\ &= \left\| y^* - x^* + \sum_{k=2}^{d} \frac{1}{k!} \sum_{|\alpha|=k} \binom{k}{\alpha} DP(x^*)^{-1} D_\alpha^k P(x)(y^* - x^*)^\alpha \right\| \\ &\geqslant \| y^* - x^* \| \left(1 - \sum_{k=2}^{d} \frac{1}{k!} \sum_{|\alpha|=k} \binom{k}{\alpha} \| DP(x^*)^{-1} D_\alpha^k P(x^*) \| \, \| y^* - x^* \|^{k-1} \right) \\ &\geqslant \| y^* - x^* \| \, L(x^*, \| y^* - x^* \|). \end{aligned}$$

Hence $L(x^*, \| y^* - x^* \|) \leqslant 0$, which implies $\| y^* - x^* \| \geqslant l(x^*)$. □

It follows from the two previous lemmas that for each index p such that $l^-(x^p) \leqslant l(x^*)$ the ball $B_{max}(x^p, l^+(x^p))$ contains only the root x^* and the third part of the theorem holds. □

Example 4.3

(1) Let us consider the expression of the Wilkinson polynomial of degree 20: $P(x) = (x-1)\ldots(x-20)$. The radii $l(i)$, $1 \leqslant i \leqslant 20$, of the disk of unicity of roots are:

$$\begin{array}{llll} l(1) = 0.206876, & l(2) = 0.320212, & l(3) = 0.406108, & l(4) = 0.467400, \\ l(5) = 0.491714, & l(6) = 0.522499, & l(7) = 0.549990, & l(8) = 0.586759, \\ l(9) = 0.627686, & l(10) = 0.684167, & l(11) = 0.684953, & l(12) = 0.625967, \\ l(13) = 0.583502, & l(14) = 0.548625, & l(15) = 0.524204, & l(16) = 0.486883, \\ l(17) = 0.467892, & l(18) = 0.398490, & l(19) = 0.312693, & l(20) = 0.203898. \end{array}$$

(2) Let us consider the system [4]:

$$\begin{aligned} &5x_1^9 - 6x_1^5 x_2 + x_1 x_2^4 + 2x_1 x_3 = 0, \\ &-2x_1^6 x_2 + 2x_1^2 x_2^3 + 2x_2 x_3 + 0, \\ &x_1^2 + x_2^2 - 0.265625 = 0. \end{aligned}$$

There are eight real solutions in the box $[-0.6, 0.6]^2 \times [-0.02, 0.02]$. Table 1 gives the solutions with respectively the separation radius and the Jacobian norm.

Table 1

Solution	Separation	$l(x)$	Jacobian norm
$(-0.515388, 0, -0.012445)$	0.121872	0.043952	0.000308
$(0.515388, 0, -0.012445)$	0.121872	0.024878	0.000308
$(0.501577, 0.118513, 0.012389)$	0.121872	0.0520798	0.000359
$(-0.501577, 0.118513, 0.012389)$	0.121872	0.121496	0.000359
$(0, 0.515388, 0)$	0.271846	0.027213	0.000237
$(0, -0.515388, 0)$	0.728974	0.027213	0.000237
$(-0.261936, 0.443862 - 0.013194)$	0.271846	0.20187	0.038792
$(0.261936, 0.443862 - 0.013194)$	0.271846	0.119856	0.003879

5. Stability

The study of the stability is the computation of a set in which all the elements verify the inequalities (1).

We first introduce some notation. For $h > 0$ and $h\omega^2 \geqslant \eta \geqslant 0$, we introduce the polynomials in $\mathbb{R}[t]$

$$T_d(\omega, t) = \omega - \sum_{i=1}^{d-1} t^i,$$

$$R(h, \eta, \omega, t) = \eta(1-t) + t - ((\omega+1)t^2 - 2(\omega+1)t + \omega)^2 h.$$

First we give a technical lemma.

Lemma 5.1

(1) If $\omega \leqslant 1$ then $(\omega+1)t^2 - 2(\omega+1)t + \omega \leqslant 1$.

(2) $T_d(\omega, t) > \dfrac{(\omega+1)t^2 - 2(\omega+1)t + \omega}{(1-t)^2}$.

(3) The smallest positive root of the polynomial $(\omega+1)t^2 - 2(\omega+1)t + \omega$ is $1 - (\sqrt{\omega+1})/(\omega+1)$.

(4) The polynomial $R(h, \eta, \omega, t)$ possesses one positive root denoted by $r(h, \eta, \omega)$ which verifies $r(h, \eta, \omega) < 1 - (\sqrt{\eta+1})/(\eta+1)$.

Proof

The proof is easy and left to the reader. □

We now give a new set of stability for the Newton method.

Proposition 5.2

Let us consider $h > 0$, $P(x) \in \mathbb{R}^n[x]$, and $x \in \mathbb{R}^n$ such that $|||DP(x)^{-1}|||$ exists. We introduce the quantities:

$$\omega(x)^{-1} = \max_{k \geqslant 1} \frac{1}{k!} |||D^k P(x)||| \, |||DP(x)^{-1}|||, \; \eta(x) = |||DP(x)^{-1}||| \, ||P(x)|| \omega(x)$$

such that $h\omega(x)^2 \geqslant \eta(x)$. We also suppose

$$\frac{1}{k!} |||D^k P(x)||| \, |||DP(x)^{-1}|||^k ||P(x)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2.$$

Then, for all $y \in \bar{B}_{max}(x, r(h, \eta(x), \omega(x)))$ we have

$$\frac{1}{k!} |||D^k P(y)||| \, |||DP(y)^{-1}|||^k ||P(y)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2.$$

Proof

Let $y \in \mathbb{R}^n$. In the same way as in lemma 2.1 we prove the following estimations:

$$||P(y)|| \leqslant ||P(x)|| + |||DP(x)||| \, ||y - x|| + \sum_{i=2}^{d} \frac{1}{i!} |||D^i P(x)||| \, ||y - x||^i,$$

$$\frac{1}{k!} |||D^k P(y)||| \leqslant \sum_{i=0}^{d-k} \binom{k+i}{i} \frac{1}{(k+i)!} |||D^{k+i} P(x)||| \, ||y - x||^i,$$

$$|||DP(y)^{-1}||| \leqslant \frac{|||DP(x)^{-1}|||}{1 - \sum_{i=1}^{d-1} (i+1) \frac{1}{(i+1)!} |||D^{1+i} P(x)||| \, |||DP(x)^{-1}||| \, ||y - x||^i}.$$

This last estimation holds if the denominator is positive. Using lemma 5.1, this condition is equivalent to

$$||y - x|| < 1 - \frac{\sqrt{\omega(x) + 1}}{\omega(x) + 1}.$$

Since $\omega(x) \leqslant 1$, we obtain the following estimations from lemma 5.1:

$$\frac{1}{k!} |||D^k P(y)||| \, |||DP(y)^{-1}|||^k ||P(y)||^{k-1}$$

$$\leqslant \frac{\left(\sum_{i=0}^{d-k} \binom{k+i}{i} ||y - x||^i \right) \left(|||DP(x)^{-1}||| \, ||P(x)|| \, \omega(x) + \sum_{i=1}^{d} ||y - x||^i \right)^{k-1}}{\left(\omega(x) - \sum_{i=1}^{d-1} (i+1) ||y - x||^i \right)^k}$$

$$\leqslant \left(\frac{\eta(x)(1 - ||y - x||) + ||y - x||}{((\omega(x) + 1)||y - x||^2 - 2(\omega(x) + 1)||y - x|| + \omega(x))^2} \right)^{k-1}.$$

Hence the inequalities of the proposition will be satisfied if $R(h, \eta(x,) \omega(x), ||y - x||) \leqslant 0$, i.e. if $||y - x|| < r(h, \eta(x), \omega(x))$. And the proposition follows. □

Corollary 5.3
Let $h \in]0, h_0]$, where h_0 is the smallest root of the polynomial $4h^3 - 12h^2 + 8h - 1$. Let us consider x satisfying the hypotheses of the previous proposition.

(1) Then, for all $y \in \bar{B}_{max}(x, r(h, \eta(x), \omega(x))$, the Newton sequence

$$x^0 = y, \quad x^{p+1} = x^p - DP(x^p)^{-1}P(x^p)$$

converges to a root of $P(x)$.

(2) If $P(x) = 0$, $\omega(x) = 1$ and $h = h_0$, we have $r(h_0, 1, 0) = 0.07877298446\ldots.$

Proof
This is a consequence of the previous proposition and of the main theorem. □

6. Application to the classical homotopy method

We deal with the polynomials $P(x) = (P_1(x), \ldots, P_n(x)) \in \mathbb{R}^n[x]$ and $Q(x) = (Q_1(x), \ldots, Q_n(x)) \in \mathbb{R}^n[x]$. Let us consider the following linear homotopy:

$$H(x, t) = tP(x) + (1 - t)Q(x)$$

for $t \in [0, 1]$. Denote by $DH(x, t)$ the Jacobian matrix of the map $(x, t) \in \mathbb{R}^{n+1} \to H(x, t) \in \mathbb{R}^n$. The meaning of the notation $D_xH(x, t)$ and $D_tH(x, t)$ is clear. We also use the following notation:

$$|||D_tD_x^kH(x)||| \overset{\text{def}}{=} \max_{1 \leqslant i \leqslant n} \sum_{|\alpha|=k} \binom{k}{\alpha} \left| \frac{\partial^k P_i(x)}{\partial x^\alpha} - \frac{\partial^k Q_i(x)}{\partial x^\alpha} \right|.$$

Let $x^0 \in \mathbb{R}^n$ be such that $Q(x^0) = 0$ and $rank(DH(x^0, 0)) = n$. From [1, lemma 2.1.3] we know that there exists a continuously differentiable curve $t \in [-1, 1] \to c(t) \in \mathbb{R}^{n+1}$ which verifies for all $t \in [0, 1]$

(1) $c(0) = 0$,
(2) $H(c(t)) = 0$,
(3) $rank(DH(c(t)) = n$,
(4) $c'(t) \neq 0$.

In this study we shall assume that

(1) $c(t) = (x(t), t)$,
(2) $D_xH(x(t), t)^{-1}$ *exists* for all $t \in [0, 1]$.

For t_0 fixed in $[0, 1]$ and $0 < h \leqslant h_0$, we apply corollary 5.3 to the map $H(x(t_0), t_0)$. We obtain for all $y \in B_{max}(x(t_0), r(h, 0, \omega(x(t_0))))$ that the Newton sequence

$$x^0 = y, \quad x^{p+1} = x^p - D_xH(x^p, t_0)^{-1}H(x^p, t_0)$$

converges to $x(t_0)$.

Let $y \in \mathbb{R}^n$ be given. If we suppose

$$\frac{1}{k!}|||D_x^k H(y,t_0)|||\,|||D_x H(y,t_0)^{-1}|||^k ||H(y,t_0)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2,$$

the question is now to compute an interval $[t_0, \bar{t}]$ for which

$$\forall t \in [t_0, \bar{t}] \quad \frac{1}{k!}|||D_x^k H(y,t)|||\,|||D_x H(y,t)^{-1}|||^k ||H(y,t)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2.$$

For this we introduce the polynomial

$$U(h,\eta,\omega,t) = \eta + t - h(\omega - t)^2.$$

We have

Lemma 6.1

(1) If $h\omega^2 \geqslant \eta \geqslant 0$ and $0 < h \leqslant h_0$, the polynomial $U(h,\eta,\omega,t)$ has one positive root in the interval $]0,1[$. Write

$$u(h,\eta,\omega) = \frac{2\omega h + 1 - \sqrt{4h(\omega+\eta)+1}}{2h}$$

for this root. Furthermore $u(h,\eta,\omega) < \omega$.

(2) $u(h_0,0,1) = 0.124504\ldots$.

Proof

A simple computation gives this lemma. □

The interval $[t_0, \bar{t}]$ is given by

Proposition 6.2

Let $t_0 \in \mathbb{R}$, $y \in \mathbb{R}^n$, $0 < h \leqslant h_0$. Let us suppose that the quantities

$$\omega(y,t_0)^{-1} = \max\Bigg(\max_{k\geqslant 1}\frac{1}{k!}||D_x^k H(y,t_0)|||\,|||D_x H(y,t_0)^{-1}|||,$$

$$\max_{k\geqslant 1}\frac{1}{k!}|||D_t D_x^k H(y,t_0)|||\,|||D_x H(y,t_0)^{-1}|||, \quad |||D_t H(y,t_0)|||\,|||D_x H(y,t_0)^{-1}|||\Bigg),$$

$$\eta(y,t_0) = |||D_x H(y,t_0)^{-1}|||\,||H(y,t_0)||\,\omega(y,t_0),$$

verify $h\omega(y,t_0)^2 \geqslant \eta(y,t_0)$. If

$$\frac{1}{k!}|||D_x^k H(y,t_0)|||\,|||D_x H(y,t_0)^{-1}|||^k ||H(y,t_0)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2, \tag{11}$$

then for all $t \in [t_0, t_0 + u(h,\eta(y,t_0),\omega(y,t_0))]$ we have

$$\frac{1}{k!}|||D_x^k H(y,t)|||\,|||D_x H(y,t)^{-1}|||^k ||H(y,t)||^{k-1} \leqslant h^{k-1}, \quad k \geqslant 2.$$

Proof
Using Taylor's formula at t_0, we have the following estimations:

$$||H(y,t)|| \leqslant ||H(y,t_0)|| + |t-t_0|\,||D_tH(y,t_0)||,$$

$$\frac{1}{k!}|||D_x^kH(y,t)||| \leqslant \frac{1}{k!}(|||D_x^kH(y,t_0)||| + |t-t_0|\,|||D_tD_x^kH(y,t)|||),$$

$$|||D_xH(y,t)^{-1}||| \leqslant \frac{D_xH(y,t_0)^{-1}}{1-|t-t_0|\,|||D_xH(y,t_0)^{-1}|||\,|||D_tD_xH(y,t_0)|||}.$$

This previous inequality holds if $|t-t_0| < \omega(y,t_0)$. We also have $\omega(y,t_0) \leqslant 1$. We deduce for $k \geqslant 2$,

$$\frac{1}{k!}|||D_x^kH(y,t)|||\,|||D_xH(y,t)^{-1}|||^k||H(y,t)||^{k-1}$$

$$\leqslant \frac{(1+|t-t_0|)(\eta(y,t_0)+|t-t_0|)^{k-1}}{(\omega(y,t_0)-|t-t_0|)^k},$$

$$\leqslant \left(\frac{\eta(y,t_0)+|t-t_0|}{(\omega(y,t_0)-|t-t_0|)^2}\right)^{k-1},$$

since

$$1+|t-t_0| \leqslant \frac{1}{1-|t-t_0|} \leqslant \frac{1}{\omega(y,t_0)-|t-t_0|}.$$

The inequalities of the proposition are satisfied if $U(h,\eta(y,t_0),\omega(y,t_0), |t-t_0|) \leqslant 0$, i.e. if $|t-t_0| \leqslant u(h,\eta(y,t_0),\omega(y,t_0))$. And the proposition follows. □

We now consider the following algorithm. Let $h \in]0,h_0[$, $a = 2h_0^2 - 4h_0 + 1$ and $c = 1/2(\log a \log(h/a^2)\log 2)^{-1/2}$ as in corollary 3.2. Let us denote $\eta_i = \eta(x^{ip_i}, t_i)$ and $\omega_i = \omega(x^{ip_i}, t_i)$.

Inputs: $\epsilon > 0$, $p_0 = 0, x^{00} = x(0), t_0 = 0$.
$i := 1$
$t_1 = u(h, \eta_0, \omega_0)$
while $t_i < 1$ **do**
  begin
  $x^{i0} = x^{i-1,p_{i-1}}$
  $x^{i1} = x^{i0} - D_xH(x^{i0}, t_i)^{-1}H(x^{i0}, t_i)$
  $k_i = \min\{k : c(a/\sqrt{2})^{k-1}(h/a^2)^{2^{k-1}-1}||x^{i1} - x^{i0}|| \leqslant \epsilon\}$
  $k := 1$
  while $k \leqslant k_i$ **and** $h\omega_i^2 < \eta_i$ **do**
    begin
    $x^{i,k+1} = x^{ik} - D_xH(x^{ik}, t_i)^{-1}H(x^{ik}, t_i)$
    $k := k+1$
    end

$p_i := k$
$t_{i+1} = \min(t_i + u(h, \eta_i, \omega_i), 1)$
$i := i + 1$
end
$x^{i,0} = x^{i-1,p_{i-1}}$
Output $x^{i,0}$

Proposition 6.3
The sequence (x^{ik}) defined in the previous algorithm converges. More precisely, there exists some index i such that $t_{i-1} < 1 \leqslant t_i$ and $\lim_k x^{ik} = x(1)$, i.e. the limit of the sequence x^{ik} is a zero of the polynomial $P(x)$.

Proof
The algorithm starts with a root x^{00} of $Q(x)$. The condition $h\omega_0^2 > \eta_0 = 0$ is satisified and we can compute t_1. From proposition 6.2, the inequalities (11) hold in (x^{00}, t_1) and the Newton sequence $x^{10} = x^{00}$, $x^{1,k+1} = x^{1k} - D_xH(x^{1k}, t_1)^{-1}H(x^{1k}, t_1)$ converges to $x(t_1)$. The algorithm consists in computing x^{1,p_1} such that $\|x^{1,p_1} - x(t_1)\| \leqslant \epsilon$ and $h\omega_1^2 \geqslant \eta_1$ using the test of corollary 3.2. At this step we can compute $t_2 = \min(t_1 + u(h, \eta_1, \omega_1), 1)$. The inequalities (11) are satisfied at $(x^{1,p_1} - t_2) = (x^{20}, t_2)$. And so on in this way: at each step of the algorithm we construct a point (x^{i,p_i}, t_{i+1}) which verifies the inequalities (11). Since the sequence (t_i) is increasing, there exists some i such that $t_{i-1} < 1 \leqslant t_i$. The algorithm returns the point $x^{i-1,p_{i-1}} = x^{i,0}$ and the proposition follows. □

References

[1] E.L. Allgower and K. Georg, *Numerical Continuation Method: an Introduction* (Springer, Berlin, 1990).
[2] J.P. Dedieu and J.C. Yakoubsohn, Two seminumerical algorithms for solving polynomial systems, preprint, Toulouse (1993).
[3] B. Demidovitch and I. Maron, *Eléments de Calcul Numérique* (MIR, Moscow, 1979).
[4] R.B. Kearfott, Some tests of generalized bisection, ACM Trans. Math. Software 13 (1987) 197–220.
[5] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
[6] A. Ostrowski, *Solutions of Equations in Euclidean and Banach Spaces* (Academic Press, New York, 1973).
[7] M. Shub and S. Smale, Computational complexity: on the geometry of polynomials and theory of cost, part 1. Ann. Sci. Éc. Norm. Sup. 4 série, 18 (1985) 107–142.
[8] M. Shub and S. Smale, Complexity of Bezout's theorem I: geometric aspects, J. Amer. Soc. 6 (1993) 459–501.
[9] S. Smale, The fundamental theorem of algebra and complexity theory, Bull. Amer. Soc. 13 (1981) 1–36.