Numerieke Methodes in de Algebra

Prof. Dr. Guido Vanden Berghe

Chapter 3

Wortels van een vergelijking Stelsels niet–lineaire vergelijkingen

---

Doelstelling

Dit hoofdstuk behandelt het probleem van de wortelbepaling van vergelijkingen van het type $f(x) = 0$. Dit is een probleem dat regelmatig voorkomt in wetenschappelijk werk. Bijvoorbeeld komt in de theorie van de lichtbreking de vraag naar voren om de wortels te bepalen van de vergelijking $x - \mathrm{tg}x = 0$; of, bij de berekening van planeetbanen, heeft men de wortels van de Kepler vergelijking $x - a \sin x = b$ nodig. In dit hoofdstuk zullen we ons in eerste instantie beperken tot het bepalen van reële wortels van reële functies. Klassieke methoden, zoals "de halveringsmethode", de "regula falsi methode", de "Newton-Raphson methode" en de "secans methode" worden in detail behandeld. Nadien wordt een algemene theorie voor zgn. één-punt iteratiemethodes ontwikkeld. Het convergentiegedrag van dergelijke methoden wordt geanalyseerd. Het numeriek bepalen van meervoudige wortels en van oplossingen van stelsels niet-lineaire vergelijkingen wordt eventjes aangeraakt. Het hoofdstuk wordt afgesloten met een gedetailleerde studie van de wortelbepaling van veeltermen. Zowel analytische als numerieke methoden komen aan bod. Het vinden van zowel reële als complex toegevoegde wortels wordt besproken.

3.1 Inleiding

Het vinden van één of meer wortels van een vergelijking

$$f(z) = 0 \tag{3.1}$$

met f een éénwaardige complexe functie van één complexe onafhankelijke veranderlijke, is één van de frequent optredende problemen in de toegepaste wiskunde. De numerieke methoden ter bepaling van dergelijke wortels zijn praktisch alle iteratief van aard. Alvorens op deze methoden in detail in te gaan, wensen we vooreerst enkele algemene opmerkingen te maken.

Opmerking 3.1.1

Vraagt men zowel naar de reële als naar de imaginaire wortels van de vergelijking $f(z) = 0$ dan is het vraagstuk steeds herleidbaar tot een gelijkwaardig stelsel van twee vergelijkingen, met twee reële onbekenden, elk met een reële operator. Immers, uit de theorie van de éénwaardige complexe functies en complexe veranderlijken, weten we dat $f(z)$ steeds te schrijven is als (zie cursus Algebra) :

$$f(z) = f(x + iy) = u(x, y) + v(x, y)i$$

waarbij $u(x, y)$ en $v(x, y)$ respectievelijk het reëel deel en de coëfficiënt van het imaginair deel van $f(x + iy)$ zijn. Dan is uiteraard $f(z) = 0$ equivalent met

$$\left\{ \begin{array}{rcl} u(x,y) & = & 0 \\ v(x,y) & = & 0 \,, \end{array} \right. \tag{3.2}$$

een stelsel van twee reële vergelijkingen in twee reële onbekenden.

Voorbeeld 3.1.1

Gevraagd de wortels van $ze^z - 2 = 0$.

Oplossing

Substitueer $z \to x + iy$:

$$\begin{aligned} f(z) &= (x+iy)e^{x+iy} - 2 = (x+iy)e^x(\cos y + i \sin y) - 2 \\ &= [e^x(x \cos y - y \sin y) - 2] + ie^x[x \sin y + y \cos y] \,. \end{aligned}$$

Dus $f(z) = 0$ wordt hier equivalent met

$$\left\{ \begin{array}{l} e^x(x \cos y - y \sin y) - 2 = u(x,y) = 0 \\ e^x(x \sin y + y \cos y) = v(x,y) = 0 \,. \end{array} \right.$$

◇

Opmerking 3.1.2

Wanneer de vergelijking uitgedrukt is met een complexe operator f bestaat de mogelijkheid om over te gaan op een aanverwant vraagstuk van wortelbepaling, waarbij de operator reëel is. Als in dit geval gegeven is $f(z) = 0$ met f complex, beschouw dan de vergelijking $f^*(z)f(z) = 0$ met f^* de complex toegevoegde operator van f. De verwantschap $z \to f^*(z)f(z)$ is er één van een rekenvoorschrift dat reëel is, zodat

$$f^*(z)f(z) = 0 \qquad \text{resulteert in} \qquad F(z) = 0$$

met F een reële operator.
Is α een reële wortel van $F(z) = 0$, dan is α een reële wortel van $f(z)$ (maar ook van $f^*(z) = 0$) met gehalveerde multipliciteit. Zijn α en α^* een paar complex toegevoegde wortels van $F(z) = 0$, dan is één onder hen een wortel van $f(z) = 0$ en de andere van $f^*(z) = 0$. Eenvoudige substitutie van α in $f(z)$ en berekening van de resulterende waarde leert of α al dan niet een wortel is van $f(z) = 0$. Hier is er behoud van multipliciteit.

Voorbeeld 3.1.2

Gegeven de hogere machtsvergelijking

$$A_0 z^n + A_1 z^{n-1} + \ldots + A_n = 0 \qquad A_0 \neq 0 \qquad n \geq 3$$

met $A_0, A_1, \ldots, A_n$ zulkdanige complexe coëfficiënten dat

$$\frac{A_1}{A_0}, \frac{A_2}{A_0}, \ldots, \frac{A_n}{A_0} \quad \text{niet alle tegelijk reëel zijn, dan is}$$

$f = A_0(\ )^n + A_1(\ )^{n-1} + \ldots + A_n$ een complexe operator.
Beschouw $f^*(z)f(z) = 0$
of $(A_0^* z^n + A_1^* z^{n-1} + \ldots + A_n^*)(A_0 z^n + A_1 z^{n-1} + \ldots + A_n) = 0$,
of $A_0^* A_0 z^{2n} + (A_1^* A_0 + A_0^* A_1) z^{2n-1} + (A_0^* A_2 + A_1^* A_1 + A_2^* A_0) z^{2n-2} + \ldots +$
$A_n^* A_n = 0$.
Hierin is $A_0^* A_0$ een reëel positief getal, $(A_1^* A_0 + A_0^* A_1)$ de som van twee complex toegevoegde grootheden en dus reëel, enz …; m.a.w. alle coëfficiënten in deze vergelijking zijn reëel. ◇

Opmerking 3.1.3

Uit voorgaande opmerkingen blijkt dat bij alle problemen de vergelijkingen kunnen herleid worden tot vergelijkingen met reële operatoren. Daarom zullen we in het vervolg slechts vergelijkingen van dit type beschouwen, nl.

$$f(z) = 0 \qquad \text{met} \qquad f \text{ reëel}.$$

We onderscheiden essentieel :

– polynomiale vergelijkingen (zgn. hogere machtsvergelijkingen)

$$A_0 z^n + A_1 z^{n-1} + \ldots + A_n = 0$$

of na deling door A_0

$$z^n + a_1 z^{n-1} + \ldots + a_n = 0 \quad \text{met} \quad \{a_i \mid 1 \leq i \leq n\} \quad \text{reële coëfficiënten.}$$

Voor dit soort vergelijkingen zullen we typische numerieke methoden aanhalen;

– rationale of transcendente vergelijkingen waarin functies van z optreden die niet uitdrukbaar zijn in eindige sommen van gehele niet–negatieve machten van z.

3.2 Eenvoudige methodes voor de bepaling van reële wortels

3.2.1 De halveringsmethode (bisection method)

Stel dat men door tabellatie erin geslaagd is, een reële wortel α van $f(x) = 0$ min of meer scherp te localiseren, m.a.w. men is erin geslaagd een interval $[a, b]$ te bepalen waarvoor

$$f(a)f(b) < 0 \, , \tag{3.3}$$

d.w.z. als de functie $f(x)$ continu is moet ze tenminste één wortel bezitten in $[a, b]$. Volgende reeks bewerkingen levert een benaderde waarde voor α binnen een tolerantie ϵ :

1. definieer $c = (a + b)/2$
2. als $b - c \leq \epsilon$ accepteer dan $\alpha \cong c$ en beëindig algoritme
3. als [teken $(f(b)) \times$ teken $(f(c))$] ≤ 0 dan $a = c$, anders $b = c$
4. start opnieuw bij stap 1.

Het interval $[a, b]$ wordt bij iedere doorgang door het algoritme in grootte gehalveerd; omwille van stap 3 zal het interval $[a, b]$ steeds een wortel van $f(x)$ bevatten.

In algoritme taal kan bovenstaande als volgt vertaald worden :

Algoritme 3.2.1

> **input** $a, b, M, \delta, \epsilon$
> $u \leftarrow f(a)$
> $v \leftarrow f(b)$
> $e \leftarrow b - a$
> **output** a, b, u, v
> **if** $teken(u) = teken(v)$ **then stop end if**
> **for** $k = 1, 2, \ldots, M$ **do**
> $e \leftarrow e/2$
> $c \leftarrow a + e$
> $w \leftarrow f(c)$
>   **output** k, c, w, e
>   **if** $|e| < \epsilon$ **or** $|w| < \delta$ **then stop end if**
>   **if** $teken(w) \neq teken(u)$ **then**
> $b \leftarrow c$
> $v \leftarrow w$
>   **else**
> $a \leftarrow c$
> $u \leftarrow w$
>   **end if**
> **end for**
> **end**

Verschillende delen van die pseudo-code vragen een bijkomende uitleg. Merk op dat we het middelste punt berekenen als $c \leftarrow a+(b-a)/2$ eerder dan als $c \leftarrow (a+b)/2$. Dit is verbonden met een strategie dat het in numerieke berekeningen altijd aangeraden is een kleine correctieterm toe te voegen aan een voorgaande benadering. Vervolgens is het beter te testen dat $teken(w) \neq teken(u)$ dan $wv < 0$, om na te gaan dat een functie van teken verandert in het interval. Merk ook op dat we drie stop criteria inlassen. Vooreerst levert M het maximum aantal stappen dat de gebruiker wil doorlopen. Zo een stop mogelijkheid moet altijd aanwezig zijn om het creëren van een oneindige lus te voorkomen. Vervolgens kan de berekening ook stoppen wanneer hetzij de fout e klein genoeg is, hetzij de waarde van $f(c)$ nul voldoende dicht nadert.

Om de verdere analyse goed te begrijpen, laat ons de verschillende intervallen die in het proces optreden als volgt noteren : $[a_0, b_0]$,$[a_1, b_1], \ldots$. Volgende eigenschappen kunnen aan die grootheden toegeschreven worden :

$$a_0 \leq a_1 \leq a_2 \leq \ldots \leq b_0,$$

$$b_0 \geq b_1 \geq b_2 \geq \ldots \geq a_0,$$

en

$$b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)\,. \tag{3.4}$$

Vermits de rij $\{a_n\}$ stijgend is en naar boven begrensd, convergeert ze. Op dezelfde wijze convergeert $\{b_n\}$. Als we (3.4) recursief toepassen vinden we dat

$$b_n - a_n = 2^{-n}(b_0 - a_0).$$

Dus

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty} 2^{-n}(b_0 - a_0) = 0.$$

Als we nu de volgende notatie invoeren :

$$r = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$$

dan kunnen we door de limiet te nemen in de ongelijkheid $0 \geq f(a_n)f(b_n)$ bekomen dat $0 \geq |f(r)|^2$, waaruit $f(r) = 0$. Veronderstel dat in een zekere stap van het proces het interval $[a_n, b_n]$ gedefinieerd is. Als het proces dan stopt, is het zeker dat de wortel in dit interval ligt. De beste schatting voor die wortel is op dit moment niet a_n of b_n maar het middelste punt van dit interval, d.i.

$$c_n = \frac{(a_n + b_n)}{2}.$$

Derhalve kunnen we dan vooropstellen dat de fout begrensd is door

$$|r - c_n| \leq \frac{1}{2}|b_n - a_n| \leq 2^{-(n+1)}(b_0 - a_0)\,.$$

Samenvattend kunnen we het volgend theorema verwoorden.

Theorema 3.2.1

Als $[a_0, b_0], [a_1, b_1], \ldots, [a_n, b_n], \ldots$ de intervallen voorstellen in de halveringsmethode, dan bestaan de limieten $\lim_{n\to\infty} a_n$ en $\lim_{n\to\infty} b_n$; daarenboven hebben die twee limieten eenzelfde waarde en stellen ze een wortelpunt van f voor. Als $r = \lim_{n\to\infty} c_n$ en $c_n = \frac{1}{2}(a_n + b_n)$ dan is

$$|r - c_n| \leq 2^{-(n+1)}(b_0 - a_0)\,.$$

Steunend op (1.17) kunnen we vooropstellen dat deze halveringsmethode lineair convergent is met een maat gelijk aan $1/2$.

3.2.2 De Regula Falsi methode

Omdat de halveringsmethode relatief traag convergeert heeft men gepoogd sneller convergerende methodes af te leiden. De regula falsi methode behoort tot die algoritmes. We vertrekken hier opnieuw van de veronderstelling (3.3). Zij A en B de punten met coördinaten $(a, f(a))$ en $(b, f(b))$ (zie figuren 1 en 2) dan wordt in de regula falsi methode de koorde AB geconstrueerd en haar snijpunt c met de x–as bepaald. De absciswaarde c wordt gebezigd als benadering voor de wortel α van $f(x)$.

figuur 1 figuur 2

De absciswaarde volgt rechtstreeks uit de analytische voorstelling van de rechten, nl.

$$c = b - f(b)\left[\frac{b-a}{f(b)-f(a)}\right] = \frac{af(b)-bf(a)}{f(b)-f(a)}\,. \tag{3.5}$$

Hierop steunend kan het proces als volgt omschreven worden :

1. laatste $c = 2b - a$ (om de eerste doorgang door de fouttest in stap 4 op een behoorlijke wijze te laten gebeuren)
2. $c = b - f(b)\left[\dfrac{(b-a)}{f(b)-f(a)}\right]$
3. als [teken $(f(b))$ $\times$ teken $(f(c))$] ≤ 0 dan $a = c$, anders $b = c$
4. als $\mid c -$ laatste $c \mid \leq \epsilon$ dan $\alpha \cong c$ en beëindig algoritme
5. anders laatste $c = c$ en keer terug naar stap 2.

In een algoritme kunnen we dit als volgt verwoorden :

Algoritme 3.2.2

> **input** $a, b, M, \delta, \epsilon$
> $u \leftarrow f(a)$
> $v \leftarrow f(b)$
> $laatstec \leftarrow 2 \times b - a$
> **output** a, b, u, v
> **if** $teken(u) = teken(v)$ **then stop end if**
> **for** $k = 1, 2, \ldots, M$ **do**
> $c \leftarrow b - f(b)\dfrac{b-a}{f(b)-f(a)}$
> $w \leftarrow f(c)$
>   **output** k, c, w
>   **if** $|w| < \delta$ **then stop end if**
>   **if** $teken(w) \neq teken(u)$ **then**
> $b \leftarrow c$
> $v \leftarrow w$
>   **else**
> $a \leftarrow c$
> $u \leftarrow w$
>   **end if**
>   **if** $|c - laatste\,c| \leq \epsilon$ **then stop**
>   **else**
> $laatste\,c = c$
>   **end if**
> **end for**
> **end**

Om het convergentiegedrag van de regula falsi methode te begrijpen is er een foutenformule nodig. Uit (3.5) volgt

$$\alpha - c = \alpha - b + f(b)\,\frac{b-a}{f(b)-f(a)}\ ,$$

waarbij α de wortel van $f(x)$ voorstelt, i.e. $f(\alpha) = 0$.
Door in het rechterlid $(\alpha - b)(\alpha - a)$ te factorizeren en enkele algebraïsche bewerkingen uit te voeren (trek dit zelf na als oefening), is bovenstaande uitdrukking te herleiden tot

$$\alpha - c = -(\alpha - b)(\alpha - a)\,\frac{f[a,b,\alpha]}{f[a,b]}\ . \tag{3.6}$$

De grootheden $f[a,b]$ en $f[a,b,\alpha]$ zijn differentiequotiënten, respectievelijk van eerste en tweede orde en zijn algemeen als volgt gedefinieerd :

$$\begin{aligned} f[x_0,x_1] &= \frac{f(x_1)-f(x_0)}{x_1-x_0} & (3.7)\\ f[x_0,x_1,x_2] &= \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}\,. & (3.8)\end{aligned}$$

Deze grootheden zijn uitdrukbaar in termen van afgeleiden van $f(x)$ (zie hoofdstuk 5), i.e.

$$f[x_0,x_1] = f'(\beta)\ , \quad f[x_0,x_1,x_2] = \frac{1}{2}\,f''(\delta) \tag{3.9}$$

met β tussen x_0 en x_1 en δ tussen het minimum en maximum van x_0, x_1 en x_2. Hiermee rekening houdend is (3.6) noteerbaar als :

$$\alpha - c = -(\alpha-b)(\alpha-a)\,\frac{f''(\delta)}{2f'(\beta)}\ . \tag{3.10}$$

Bij het gebruik van de regula falsi methode is doorgaans $f''(\alpha) \neq 0$ en is α de enige wortel in $[a,b]$. Om de verdere analyse te vereenvoudigen nemen we aan dat $f''(x) \geq 0$ voor $a \leq x \leq b$; een compleet analoge discussie kan gevoerd worden in het geval $f''(x) < 0$. Bij de aanname $f'' \geq 0$ ligt de koorde AB (figuur 1), die de punten $(x_1, f(x_1))$ en $(x_2, f(x_2))$ verbindt, boven de grafiek van $y = f(x)$, voor alle $a \leq x_1 < x_2 \leq b$.

Geval 1 : $f'(\alpha) > 0$

Rekening houdend met bovenstaande keuze, ligt c altijd tussen a en α (figuur 1), d.w.z. in stap 3 van het regula falsi algoritme wordt a steeds vervangen door c, m.a.w. de factor $\alpha - b$ blijft constant. Noemen we a_n de n^{de} waarde van a in het algoritme, dan is volgens (3.10)

$$\alpha - a_{n+1} = -(\alpha-a_n)(\alpha-b)\,\frac{f''(\delta_n)}{2f'(\beta_n)}\ , \qquad\qquad n \geq 0\ ,$$

met δ_n en β_n tussen a_n en b. Laat ons definiëren :

$$\mu = \max_{x\in[a,b]}\left|(\alpha-b)\frac{f''(x)}{2f'(x)}\right| \leq |a-b| \max_{x\in[a,b]}\left|\frac{f''(x)}{2f'(x)}\right|\ ,$$

zodat uit voorgaande uitdrukking volgt

$$|\alpha - a_{n+1}| \leq \mu|\alpha - a_n| \qquad\qquad n \geq 0\,.$$

We stellen dus lineaire convergentie vast (vergelijk met (1.16)) met convergentiemaat μ.

Waar deze maat bij de halveringsmethode constant was (nl. 1/2) stellen we vast dat ze hier afhankelijk wordt van de afgeleiden van de functie f en van de grenzen a en b van het oorspronkelijk interval. Dit wil zeggen dat de regula falsi methode soms sneller kan convergeren dan de halveringsmethode en soms niet.

Geval 2 : $f'(\alpha) < 0$

In dit geval zal c steeds liggen tussen α en b. Het punt a blijft onveranderd, terwijl c bij elke iteratiestap b zal vervangen. De redenering onder geval 1 kan hier hernomen worden en de bereikte conclusie is identiek als hierboven.

3.2.3 De methode van Newton–Raphson

Zoals bij de regula falsi–methode, benaderen we $y = f(x)$ in de omgeving van de wortel α door een rechte lijn, maar nu gebruiken we geen koorde maar een raaklijn. Startend van een initiale schatting voor α, nl. x_0 construeren we een loodrechte ophaallijn (zie figuur 3); in het snijpunt van deze ophaallijn en de curve $y = f(x)$ construeren we de raaklijn; het snijpunt van deze raaklijn en de x–as levert een nieuwe benadering voor α.

In de driehoek ABC is

$$\begin{aligned} BA &= CB \operatorname{tg} \hat{C} \qquad \text{of} \\ f(x_0) &= (x_0 - x_1) f'(x_0) \end{aligned}$$

waaruit

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \,.$$

figuur 3

Na n maal bovenstaande constructie te hebben doorgevoerd, bekomen we de volgende formule

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \,, \qquad n \geq 0 \,. \tag{3.11}$$

D.i. het *Newton–Raphson* voorschrift, welke de best bekende procedure is voor het bepalen van wortels van vergelijkingen. Dit Newton–Raphson voorschrift (3.11) kan ook afgeleid worden uit een linearizatieproces. Beschikt men over een benadering voor α nl. x_n dan is $\alpha = x_n + h$ waarbij h de exacte correctie is, m.a.w. $f(x_n + h) = 0$. Ontwikkelen we het linkerlid in Taylorreeks rond x_n (met sluitterm van Lagrange)

$$f(x_n + h) = f(x_n) + (\alpha - x_n) f'(x_n) + \frac{(\alpha - x_n)^2}{2} f''(\xi_n) = 0$$

met ξ_n gelegen tussen α en x_n.

Opgelost naar α krijgen we

$$\alpha = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{(\alpha - x_n)^2}{2} \frac{f''(\xi_n)}{f'(x_n)} \,. \tag{3.12}$$

Linearizeren (= het weglaten van hogere–orde termen in $h = \alpha - x_n$) levert ons opnieuw een benadering voor α van de vorm (3.11). Door (3.11) lid aan lid af te trekken van (3.12) volgt er

$$\alpha - x_{n+1} = -(\alpha - x_n)^2 \frac{f''(\xi_n)}{2f'(x_n)} \,. \tag{3.13}$$

Deze formule zal aangewend worden om aan te tonen dat deze methode een kwadratische orde van convergentie $p = 2$ bezit.

Het rekenproces kan als volgt in een algoritme voorgesteld worden :

Algoritme 3.2.3

> **input** x_0, M, δ, ϵ
> $v \leftarrow f(x_0)$
> **output** $0, x_0, v$
> **if** $|v| < \delta$ **then stop end if**
> **for** $k = 1, 2, \ldots, M$ **do**
> $x_1 \leftarrow x_0 - v/f'(x_0)$
> $v \leftarrow f(x_1)$
>   **output** k, x_1, v
>   **if** $|x_1 - x_0| \leq \epsilon$ **or** $|v| \leq \delta$ **then stop end if**
> $x_0 \leftarrow x_1$
> **end for**
> **end**

Voorbeeld 3.2.1

Vind een wortel α van

$$f(x) \equiv x^6 - x - 1 = 0$$

in [1,2]. De exacte oplossing tot op 9 decimale cijfers is $\alpha = 1.134724138$.

Oplossing

Dit leidt tot het iteratief doorlopen van

$$x_{n+1} = x_n - \frac{x_n^6 - x_n - 1}{6x_n^5 - 1} \qquad\qquad n \geq 0\,.$$

De resultaten bekomen met startwaarde $x_0 = 2.0$ vind je in de eerstvolgende tabel.

n	x_n	$f(x_n)$	$\alpha - x_n$
0	2.0	61.0	–8.652E–1
1	1.680628273	19.85	–5.459E–1
2	1.430738989	6.147	–2.960E–1
3	1.254970957	1.652	–1.202E–1
4	1.161538433	2.943E–1	–2.681E–2
5	1.136353274	1.683E–2	–1.629E–3
6	1.134730528	6.574E–5	–6.390E–6
7	1.134724138	–4.0E–9	< 1.0E–9

We merken op dat de Newton–Raphson methode zeer snel convergeert eenmaal een x_{n+1}–waarde bereikt wordt in de omgeving van α. Dit wordt geïllustreerd door x_4, x_5, x_6, x_7. De waarden x_0, x_1, x_2, x_3 vertonen de trage convergentie die ontstaat door een slechte initiale keuze x_0. Hadden we $x_0 = 1$ gekozen, dan was x_4 reeds correct tot op 7 significante cijfers en x_5 tot op 10. ◇

De convergentie zal hieronder bestudeerd worden. De snelheid van convergentie en het interval waarin de initale waarde x_0 kan gekozen worden, zullen aangegeven worden.

Theorema 3.2.2 - Convergentie theorema

Laat $f(x)$, $f'(x)$ en $f''(x)$ continu zijn voor alle x in de omgeving van α en veronderstel $f(\alpha) = 0$, $f'(\alpha) \neq 0$. Wanneer x_0 voldoende dicht bij α gekozen is, dan convergeren de x_n $(n \geq 0)$ uit (3.11) naar α. Bovendien is

$$\lim_{n\to\infty} \frac{(\alpha - x_{n+1})}{(\alpha - x_n)^2} = -\frac{f''(\alpha)}{2f'(\alpha)} \tag{3.14}$$

aantonende dat, als $f''(\alpha) \neq 0$, er een orde van convergentie $p = 2$ is.

Bewijs

Beschouw een interval $I = [\alpha - \epsilon, \alpha + \epsilon]$ en laat

$$M = \frac{\max\limits_{x\in I} \mid f''(x) \mid}{2\min\limits_{x\in I} \mid f'(x) \mid}\,.$$

Uit (3.13) volgt dan

$$\begin{aligned} \mid \alpha - x_1 \mid &\leq M \mid \alpha - x_0 \mid^2 \qquad \text{en} \\ M \mid \alpha - x_1 \mid &\leq (M \mid \alpha - x_0 \mid)^2 \, . \end{aligned}$$

Kies $\mid \alpha - x_0 \mid \leq \epsilon$ en $M \mid \alpha - x_0 \mid < 1$, dan is $M \mid \alpha - x_1 \mid < 1$ en $M \mid \alpha - x_1 \mid \leq M \mid \alpha - x_0 \mid$ wat ook wil zeggen $\mid \alpha - x_1 \mid \leq \epsilon$. Wij kunnen deze redenering inductief toepassen op $x_1, x_2, \ldots$, aantonende dat $\mid \alpha - x_n \mid \leq \epsilon$ en $M \mid \alpha - x_n \mid < 1$ voor alle $n \geq 1$.
Om nu de convergentie aan te tonen vertrekken we opnieuw uit (3.13), i.e.

$$\begin{aligned} \mid \alpha - x_{n+1} \mid &\leq M \mid \alpha - x_n \mid^2 \qquad \text{of} \\ M \mid \alpha - x_{n+1} \mid &\leq (M \mid \alpha - x_n \mid)^2 \end{aligned}$$

en door inductie

$$\begin{aligned} M \mid \alpha - x_n \mid &\leq (M \mid \alpha - x_0 \mid)^{2^n} \qquad \text{of} \\ \mid \alpha - x_n \mid &\leq \frac{1}{M}(M \mid \alpha - x_0 \mid)^{2^n} \, . \end{aligned}$$

Vermits $M \mid \alpha - x_0 \mid < 1$ toont dit aan dat $x_n \to \alpha$ als $n \to \infty$. In (3.13) ligt het onbekende punt ξ_n tussen x_n en α en dus $\xi_n \to \alpha$ als $n \to \infty$. Dus

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^2} = - \lim_{n\to\infty} \frac{f''(\xi_n)}{2f'(x_n)} = - \frac{f''(\alpha)}{2f'(\alpha)} \, .$$

Om convergentie te hebben van x_n naar α houdt de voorwaarde $M \mid \alpha - x_0 \mid < 1$ in dat we zouden moeten hebben

$$\mid \alpha - x_0 \mid < \frac{1}{M} \, .$$

Dus M is een maat aangevende hoe dicht x_0 bij α moet gekozen worden om de convergentie te waarborgen. ◇

3.2.4 De secans methode (secant method)

We vertrekken van het Newton iteratie schema :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \, . \tag{3.15}$$

Eén van de nadelen van deze methode is dat de afgeleide van de functie, waarvan men de wortel wil bepalen, moet gekend zijn. Een middel om dit te omzeilen bestaat erin $f'(x)$ te vervangen door een benadering, i.e.

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \, . \tag{3.16}$$

De benadering gegeven in (3.16) wordt rechtstreeks afgeleid uit de definitie van f' als een limiet, nl.

$$f'(x) = \lim_{u \to x} \frac{f(x) - f(u)}{x - u} \, .$$

Wanneer die substitutie gebeurd is, wordt het resulterende schema de *secans* methode genoemd. Ze ziet er als volgt uit :

$$x_{n+1} = x_n - f(x_n) \left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right] \qquad (n \geq 1) \, . \tag{3.17}$$

Vermits voor de berekening van x_{n+1} zowel x_n als x_{n-1} nodig zijn, moeten twee initiale punten gegeven worden. Nochtans vereist elke nieuwe x_{n+1} waarde slechts één enkele nieuwe evaluatie van f. De grafische interpretatie van de secans methode is vergelijkbaar met deze van Newton's methode. De raaklijn aan de curve wordt hier vervangen door de secans lijn (zie figuur 4).

figuur 4

In algoritme taal kunnen we deze methode omschrijven als

Algoritme 3.2.4

input $a, b, M, \delta, \epsilon$
$u \leftarrow f(a)$
$v \leftarrow f(b)$
output $0, a, u$
output $1, b, v$
for $k = 2, 3, \ldots, M$ **do**

```
    if |u| < |v| then
      a ↔ b
      u ↔ v
    end if
    s ← (b − a)/(v − u)
    a ← b
    u ← v
    b ← b − v × s
    v ← f(b)
    output k, b, v
    if |v| < δ or |b − a| < ϵ then stop end if
  end for
end
```

Merk op dat in elke stap de punten a en b kunnen gewisseld worden om ervoor te zorgen dat de rij $\{|f(x_k)| \mid k = 0, 1, \ldots\}$ niet stijgend is.

Voorbeeld 3.2.2

Gebruik de secans methode voor het vinden van een wortel van de functie

$$f(x) = x^3 - \mathrm{sh}x + 4x^2 + 6x + 9\,.$$

Oplossing

Een ruw tabelleren toont aan dat tussen 7 en 8 een wortel ligt. We nemen deze punten als x_0 en x_1 uit het algoritme. Deze methode levert als resultaat :

n	x_n	$f(x_n)$
0	7.00000	0.417×10^{2}
1	8.00000	-0.665×10^{3}
2	7.05895	0.208×10^{2}
3	7.11764	-0.183×10^{1}
4	7.11289	0.710×10^{-1}
5	7.11306	0.244×10^{-3}
6	7.11306	0.610×10^{-4} .

◇

Om het convergentiegedrag te onderzoeken, zullen we vooreerst een foutenanalyse in de secans methode doorvoeren. Uit de definitie (3.17) van de secans methode vinden we met $e_n = x_n - \alpha$ (met α terug een notatie voor de exacte wortel) :

$$\begin{aligned} e_{n+1} = x_{n+1} - \alpha &= \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})} - \alpha \\ &= \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})} \,. \end{aligned}$$

Door $e_n e_{n-1}$ te factorizeren en $\dfrac{x_n - x_{n-1}}{x_n - x_{n-1}}$ in te voeren bekomen we

$$e_{n+1} = \Big[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\Big]\Big[\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}}\Big] e_n e_{n-1} \,. \tag{3.18}$$

Uit het Taylor theorema volgt :

$$f(x_n) = f(\alpha + e_n) = f(\alpha) + e_n f'(\alpha) + \frac{1}{2} e_n^2 f''(\alpha) + \mathcal{O}(e_n^3) \,.$$

Vermits $f(\alpha) = 0$ levert dit

$$f(x_n)/e_n = f'(\alpha) + \frac{1}{2} e_n f''(\alpha) + \mathcal{O}(e_n^2) \,,$$

of ook $n \leftrightarrow n-1$

$$f(x_{n-1})/e_{n-1} = f'(\alpha) + \frac{1}{2} e_{n-1} f''(\alpha) + \mathcal{O}(e_{n-1}^2) \,.$$

Door aftrekking van die laatste twee betrekkingen bekomen we

$$f(x_n)/e_n - f(x_{n-1})/e_{n-1} = \frac{1}{2}(e_n - e_{n-1}) f''(\alpha) + \mathcal{O}(e_{n-1}^2) \,.$$

Vermits $x_n - x_{n-1} = e_n - e_{n-1}$, bereiken we de betrekking :

$$\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}} \approx \frac{1}{2} f''(\alpha) \,.$$

De eerste uitdrukking tussen haakjes in (3.18) kan benaderd worden door

$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \approx \frac{1}{f'(\alpha)} \,.$$

Derhalve hebben we aangetoond dat

$$e_{n+1} \approx \frac{f''(\alpha)}{2f'(\alpha)} e_n e_{n-1} = C e_n e_{n-1} \,. \tag{3.19}$$

Uit die gegevens kan men zich een idee vormen over de orde van convergentie. Hiertoe postuleren we de volgende asymptotische verwantschap (zie ook (1.16)) :

$$|e_{n+1}| \sim A|e_n|^p \,. \tag{3.20}$$

Dus ook

$$|e_n| \sim A|e_{n-1}|^p \text{ en } |e_{n-1}| \sim (A^{-1}|e_n|)^{1/p} \,. \tag{3.21}$$

In vergelijking (3.19) substitueren we de asymptotische waarden voor $|e_{n+1}|$ en $|e_{n-1}|$ uit relaties (3.20) en (3.21). Dit resulteert in

$$A|e_n|^p \sim C|e_n|A^{-1/p}|e_n|^{1/p} \,,$$

wat kan geschreven worden als

$$A^{1+1/p}C^{-1} \sim |e_n|^{1-p+1/p} \,. \tag{3.22}$$

Vermits de linkerzijde van die relatie een van nul verschillende constante is en $e_n \to 0$ als $n \to \infty$, moeten we concluderen we dat $1 - p + 1/p = 0$ of $p = (1 + \sqrt{5})/2 \approx 1.62$. Derhalve stellen we vast dat de secans methode *superlineair* is (d.w.z. beter dan lineair). We kunnen nu ook A bepalen vermits de rechterzijde van de relatie (3.22) 1 wordt. Dus, gebruik makend van $1 + 1/p = p$, hebben we

$$A = C^{1/(1+1/p)} = C^{1/p} = C^{p-1} = C^{0.62} = \Big[\frac{f''(\alpha)}{2f'(\alpha)}\Big]^{0.62} \,.$$

Met die waarde voor A krijgen we finaal voor de secans methode :

$$|e_{n+1}| \approx A|e_n|^{(1+\sqrt{5})/2} \,.$$

Vermits $(1 + \sqrt{5})/2 \approx 1.62 < 2$ is de convergentiesnelheid niet zo goed als bij de Newton–Raphson methode maar beter dan bij de halveringsmethode. Nochtans vereist elke stap in de secans methode slechts *één* nieuwe functie-evaluatie, terwijl elke stap in het Newton algoritme *twee* functie-evaluaties vergt, nl. $f(x)$ en $f'(x)$. Vermits functie-evaluaties de meeste rekentijd in beslag nemen in de beschouwde methoden, kunnen we vooropstellen dat twee stappen in de secans methode in dezelfde tijd worden uitgevoerd als één stap in de Newton–Raphson methode.

DEMO 5 (Maple)

Zie Claroline

3.3 Een algemene theorie voor één–punt iteratie methodes

3.3.1 Theoretische afleiding en theorema's

We beschouwen het bepalen van een wortel α van de vergelijking $x = g(x)$ door het definiëren van

$$x_{n+1} = g(x_n) , \qquad\qquad n \geq 0 ,$$

met x_0 een initiale schatting voor α. De Newton–Raphson methode behoort tot dit soort vergelijkingen met

$$g(x) = x - \frac{f(x)}{f'(x)} .$$

Ook de regula falsi methode behoort tot die klasse omdat één van de eindpunten onveranderd blijft (zie paragraaf 3.2.2).

De oplossingen van $x = g(x)$ worden vaste punten van g genoemd. We zijn meestal geïnteresseerd in de oplossingen van een vergelijking $f(x) = 0$ en er zijn doorgaans verschillende manieren om deze te herformuleren in de vorm van een vast–punt probleem. We willen dit hier vooreerst illustreren met een voorbeeld.

Voorbeeld 3.3.1

Beschouw $x^2 - a = 0$ voor $a > 0$ (één van de oplossingen is $+\sqrt{a}$). Mogelijke vaste–punt herformuleringen zijn :

(1) $x = x^2 + x - a$ of algemener $x = x + c(x^2 - a)$ voor $c \neq 0$

(2) $x = a/x$

(3) $x = \frac{1}{2}(x + \frac{a}{x})$ (d.i. in feite de Newton–Raphson formulering).

Voor bv. $a = 3$, $x_0 = 2$ geven we in de volgende tabel de numerieke resultaten, opgeleverd door de 3 bovenstaande gevallen

	geval (1)	geval (2)	geval (3)
n	x_n	x_n	x_n
0	2.0	2.0	2.0
1	3.0	1.5	1.75
2	9.0	2.0	1.732143
3	87.0	1.5	1.732051

Het is duidelijk dat enkel geval (3) tot goede resultaten leidt. Het is evident zich af te vragen waarom deze drie iteratieschema's zich gedragen zoals in dit voorbeeld. Hierna zullen we een algemene theorie ontwikkelen om dit gedrag te verklaren en om een methode te hebben om andere nieuwe iteratiemethodes te analyseren. $\diamond$

Lemma 3.3.1

Weze $g(x)$ continu in het interval $a \leq x \leq b$ en neem aan dat $a < g(x) < b$ voor elke $a \leq x \leq b$ (We zeggen g zendt $[a, b]$ in $[a, b]$ en noteren dit verder als $g([a, b]) \subset [a, b]$), dan bezit $x = g(x)$ minstens één oplossing in $[a, b]$.

Bewijs

Beschouw de continue functie $g(x) - x$. Binnen bovengenoemde aannamen is ze bij $x = a$ positief en bij $x = b$ negatief, d.w.z. steunend op de stelling van de tussenliggende waarden moet ze een wortel hebben in $[a, b]$. In figuur 5 zijn de wortels snijpunten van $y = x$ en $y = g(x)$. $\diamond$

figuur 5

Lemma 3.3.2

Weze $g(x)$ continu in $[a, b]$ en onderstel $g([a, b]) \subset [a, b]$. Onderstel bovendien dat er een constante $0 < \lambda < 1$ bestaat waarvoor

$$| g(x) - g(y) | \leq \lambda | x - y | \quad \text{voor alle} \quad x, y \in [a, b] , \tag{3.23}$$

dan bezit $x = g(x)$ een unieke oplossing α in $[a, b]$. De opeenvolgende x_n gegenereerd uit

$$x_n = g(x_{n-1}) \qquad\qquad n > 0$$

zullen naar α convergeren voor elke keuze x_0 in $[a, b]$.

Bewijs

Veronderstel dat $x = g(x)$ twee oplossingen α en β bezit in $[a, b]$. Dan is

$$\mid \alpha - \beta \mid = \mid g(\alpha) - g(\beta) \mid \leq \lambda \mid \alpha - \beta \mid$$

of $\quad (1 - \lambda) \mid \alpha - \beta \mid \leq 0$.

Vermits $0 < \lambda < 1$ impliceert dit $\alpha = \beta$. Bovendien weten we uit lemma 3.3.1 dat er tenminste een wortel α in $[a, b]$ is.
Om de convergentie van de rij $\{x_n\}$ te onderzoeken, kunnen we vooreerst opmerken dat alle x_n in $[a, b]$ gelegen zijn, want als

$$x_n \in [a, b] \qquad \text{dan} \qquad x_{n+1} = g(x_n) \in [a, b] \; .$$

Steunend op wiskundige inductie volgt hieruit $x_n \in [a, b]$ voor alle n. Voor de convergentie vertrekken we van

$$\mid \alpha - x_{n+1} \mid = \mid g(\alpha) - g(x_n) \mid \leq \lambda \mid \alpha - x_n \mid$$

en door inductie vinden we

$$\mid \alpha - x_n \mid \leq \lambda^n \mid \alpha - x_0 \mid , \qquad\qquad n \geq 0 \; . \tag{3.24}$$

Als $n \to \infty$, $\lambda^n \to 0$ en dus $x_n \to \alpha$. ◇

Noteer nu ook dat als $g(x)$ afleidbaar is in $[a, b]$ dan

$$g(x) - g(y) = g'(\xi)(x - y) \qquad\qquad \xi \text{ tussen } x \text{ en } y$$

voor alle $x, y \in [a, b]$. Definiëren we

$$\lambda = \max_{a \leq x \leq b} \mid g'(x) \mid$$

dan is $\mid g(x) - g(y) \mid \leq \lambda \mid x - y \mid$ voor $x, y \in [a, b]$.

Theorema 3.3.1

Veronderstel dat $g(x)$ continu afleidbaar is in $[a, b]$, (d.w.z. van de klasse $\mathcal{C}^1$), dat $g([a, b]) \subset [a, b]$ en dat

$$\lambda = \max_{a \leq x \leq b} \mid g'(x) \mid < 1 \, , \tag{3.25}$$

dan

(1) heeft $x = g(x)$ *een unieke oplossing* α *in* $[a,b]$ *;*

(2) voor elke keuze x_0 *in* $[a,b]$ *, met* $x_{n+1} = g(x_n)$ *,* $n \geq 0$ *is* $\lim_{n\to\infty} x_n = \alpha$ *;*

(3) $\mid \alpha - x_n \mid \leq \lambda^n \mid \alpha - x_0 \mid$
en

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha) \ . \tag{3.26}$$

Bewijs

Elk van de opgesomde punten volgt rechtstreeks uit voorgaande lemma's, behalve voor wat de maat van de convergentie (3.26) betreft. We weten echter dat

$$\alpha - x_{n+1} = g(\alpha) - g(x_n) = g'(\xi_n)(\alpha - x_n) \ , \qquad n \geq 0 \ , \tag{3.27}$$

met ξ_n een onbekend punt tussen α en x_n. Vermits $x_n \to \alpha$ moeten we ook hebben $\xi_n \to \alpha$ en dus

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = \lim_{n\to\infty} g'(\xi_n) = g'(\alpha) \, .$$

◇

Merk vooral de belangrijkheid op van de aanname (3.25) over de grootte van $g'(x)$. Was $\mid g'(\alpha) \mid > 1$ dan zou, wanneer x_n voldoende dicht bij α ligt, uit (3.26) volgen dat de fout $\mid \alpha - x_{n+1} \mid$ groter wordt dan $\mid \alpha - x_n \mid$, waaruit we moeten besluiten dat convergentie niet mogelijk is wanneer $\mid g'(\alpha) \mid \geq 1$.

Beschouwen we nu opnieuw ons voorbeeld van het begin van deze paragraaf. Berekenen we voor elk van de gevallen $g'(\alpha)$

(1) $g(x) = x^2 + x - 3, \ g'(\alpha) = g'(\sqrt{3}) = 2\sqrt{3} + 1 > 1$;
dus geen convergentie;

(2) $g(x) = 3/x, \ g'(\sqrt{3}) = -3/(\sqrt{3})^2 = -1$;
dus geen convergentie ;

(3) $g(x) = \frac{1}{2}(x + \frac{3}{x}), \ g'(x) = \frac{1}{2}(1 - \frac{3}{x^2}) ; \ g'(\sqrt{3}) = 0$;
dus wel convergentie.

Om de theorie over één–punt iteratiemethoden te vervolledigen, beschouwen we ook methodes met een orde van convergentie groter dan 1, zoals bv. de Newton-Raphson methode.

Theorema 3.3.2

Veronderstel dat α een wortel is van $x = g(x)$ en dat voor een waarde $p \geq 2$ $g(x)$ p maal continu afleidbaar is voor alle x in de omgeving van α (d.w.z. van de klasse $\mathcal{C}^p$). Neem bovendien aan dat

$$g'(\alpha) = \ldots = g^{(p-1)}(\alpha) = 0 \qquad \text{en} \qquad g^{(p)}(\alpha) \neq 0\,, \tag{3.28}$$

dan, als de initiale schatting x_0 voldoende dicht bij α gekozen is, zal de iteratie

$$x_{n+1} = g(x_n)\,, \qquad n \geq 0$$

een orde van convergentie p hebben en

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} = (-1)^{p-1}\,\frac{g^{(p)}(\alpha)}{p!}\;. \tag{3.29}$$

Bewijs

Ontwikkel $g(x_n)$ in Taylorreeks rond α :

$$\begin{aligned} x_{n+1} &= g(x_n) = g(\alpha) + (x_n - \alpha)g'(\alpha) + \ldots \\ &\qquad + \frac{(x_n-\alpha)^{p-1}}{(p-1)!}\,g^{(p-1)}(\alpha) + \frac{(x_n-\alpha)^p}{p!}\,g^{(p)}(\xi_n) \end{aligned}$$

met ξ_n tussen x_n en α. Omwille van (3.28) en het feit dat $\alpha = g(\alpha)$ volgt hieruit

$$\alpha - x_{n+1} = -\frac{(x_n-\alpha)^p}{p!}\,g^{(p)}(\xi_n)\,.$$

Neem de limiet $n \to \infty$ en steun op $x_n \to \alpha$ om het bewijs te vervolledigen. $\diamond$

De Newton–Raphson methode kan d.m.v. dit resultaat geanalyseerd worden, i.e.

$$g(x) = x - \frac{f(x)}{f'(x)}\;, \qquad g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$
$$g'(\alpha) = 0 \quad \text{en} \quad g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}\;.$$

Dit tezamen met (3.29) levert het vroeger reeds bekomen convergentieresultaat (3.14).

3.3.2 Fouttesten

Bij het gebruik van iteratieve methodes voor het oplossen van $f(x) = 0$ moet beslist worden hoe de iteratie te stoppen. Er wordt veel gebruik gemaakt van de test

$$| f(x_n) | \leq \delta \,, \tag{3.30}$$

met δ een vooraf ingebouwde tolerantie. We wensen erop te wijzen dat het gebruik hiervan alleen geen betrouwbare test is voor de nauwkeurigheid van x_n t.o.v. α. Veronderstel dat (3.30) is voldaan; beschouw dan het middelwaarde theorema

$$f(x_n) = f(x_n) - f(\alpha) = f'(\xi_n)(x_n - \alpha)$$

waaruit

$$\alpha - x_n = -\frac{f(x_n)}{f'(\xi_n)} \tag{3.31}$$

met ξ_n gelegen tussen x_n en α. Dus $| \alpha - x_n | \approx | f(x_n) |$ als $f'(\alpha) \approx 1$ en de waarde x_n zal met een nauwkeurigheid δ de wortel α benaderen. Wanneer echter $f'(\alpha) \ll 1$ dan is $| \alpha - x_n |$ veel groter dan de opgegeven tolerantie δ en wanneer $f'(\alpha) | \gg 1$ dan is $| \alpha - x_n |$ veel kleiner dan δ, wat onnodige bewerkingen veroorzaakt.

Voor de methode van Newton–Raphson volgt uit (3.31)

$$x_{n+1} - x_n = -\frac{f(x_n)}{f'(x_n)} \approx -\frac{f(x_n)}{f'(\xi_n)} = \alpha - x_n$$

en dus

$$| x_{n+1} - x_n | \leq \epsilon \quad \text{waarborgt} \quad | \alpha - x_n | \leq \epsilon \tag{3.32}$$

als x_n dicht genoeg ligt bij α opdat de benadering $f'(x_n) \approx f'(\xi_n)$ waar zou zijn.

Een tweede natuurlijke test om de iteratie te beëindigen kan steunen op

$$| x_{n+1} - x_n | \leq \epsilon \,.$$

Dit is een waardevolle test voor de methode van Newton–Raphson (zie (3.32)), maar ze kan onnauwkeurig zijn voor veel lineair convergente methodes. Beschouw bijvoorbeeld een lineair convergente één–punt methode $x_{n+1} = g(x_n)$ en veronderstel $g'(\alpha) \approx 1$. Voor waarden x_n, die voldoende dicht α benaderen hebben we

$$\alpha - x_{n+1} = g(\alpha) - g(x_n) = g'(\xi_n)(\alpha - x_n) \approx g'(\alpha)(\alpha - x_n)$$

en

$$x_{n+1} - x_n = (\alpha - x_n) - (\alpha - x_{n+1}) \approx [1 - g'(\alpha)](\alpha - x_n) \,,$$

waaruit

$$\alpha - x_n \approx \frac{1}{1-g'(\alpha)}\,(x_{n+1} - x_n)\ . \tag{3.33}$$

Dus $\mid \alpha - x_n \mid \gg \mid x_{n+1} - x_n \mid$ als $g'(\alpha) \approx 1$.
Als we echter het zwakker resultaat aannemen

$$\mid \alpha - x_{n+1} \mid \leq \lambda \mid \alpha - x_n \mid \qquad\qquad 0 < \lambda < 1 \tag{3.34}$$

voor alle n , krijgen we de volgende nauwkeurigheidstest :

$$\begin{aligned} \mid x_{n+1} - x_n \mid = \mid (\alpha - x_n) - (\alpha - x_{n+1}) \mid \ &\geq\ \mid \alpha - x_n \mid - \mid \alpha - x_{n+1} \mid \\ &\geq\ (1-\lambda) \mid \alpha - x_n \mid \end{aligned}$$

of

$$\mid \alpha - x_n \mid \leq \frac{1}{1-\lambda} \mid x_{n+1} - x_n \mid ,$$

waaruit

$$\mid \alpha - x_{n+1} \mid \leq \frac{\lambda}{1-\lambda} \mid x_{n+1} - x_n \mid\ . \tag{3.35}$$

Om dus $\mid \alpha - x_{n+1} \mid \leq \epsilon$ te hebben, itereer tot x_{n+1} voldoet aan

$$\mid x_{n+1} - x_n \mid \leq \Big(\frac{1-\lambda}{\lambda}\Big)\epsilon\ . \tag{3.36}$$

Dit is een algemeen bruikbare foutentest voor lineair convergente methodes. De te gebruiken waarde van λ kan als volgt bepaald worden. Uit (3.25) en (3.26) volgt

$$\frac{\alpha - x_{n+1}}{\alpha - x_n} \approx \lambda = \text{constante} \qquad\qquad (n \geq 0)\ .$$

Hieruit volgt

$$\frac{\alpha - x_{n+1}}{\alpha - x_n} \approx \frac{\alpha - x_{n+2}}{\alpha - x_{n+1}} \approx \lambda\ , \qquad n \geq 0 \tag{3.37}$$

en ook

$$\begin{aligned} \frac{x_{n+2} - x_{n+1}}{x_{n+1} - x_n} \ &=\ \frac{(\alpha - x_{n+1}) - (\alpha - x_{n+2})}{(\alpha - x_n) - (\alpha - x_{n+1})} \\ &=\ \frac{1 - \dfrac{\alpha - x_{n+2}}{\alpha - x_{n+1}}}{\dfrac{\alpha - x_n}{\alpha - x_{n+1}} - 1} \approx \frac{1-\lambda}{\dfrac{1}{\lambda} - 1} = \lambda\ . \end{aligned}$$

Opmerking 3.3.1

Uit deze beschouwingen kan men gemakkelijk de zgn. extrapolatieformule van Aitken afleiden. Als n voldoende groot is volgt uit (3.37) dat

$$(\alpha - x_{n+1})^2 \approx (\alpha - x_n)(\alpha - x_{n+2}),$$

waaruit

$$\alpha(x_n + x_{n+2} - 2x_{n+1}) \approx -x_{n+1}^2 + x_n x_{n+2},$$

of

$$\alpha \approx \frac{x_n x_{n+2} - x_{n+1}^2}{x_n - 2x_{n+1} + x_{n+2}} = x_{n+2} - \frac{(x_{n+2} - x_{n+1})^2}{x_n - 2x_{n+1} + x_{n+2}}. \tag{3.38}$$

Deze benaderingsformule voor de te zoeken wortel is bekend als *Aitken's extrapolatieformule*. Als (3.37) voor opeenvolgende n-waarden een quasi constante λ oplevert, volgt uit het rechterlid van deze extrapolatieformule een veel betere schatting voor de te zoeken wortel dan x_{n+2}. ◇

Voorbeeld 3.3.2

Beschouw de vergelijking

$$x = x + c(x^2 - 3).$$

Bepaal c zodat de convergentie voor een in te voeren iteratieschema verzekerd is.

Oplossing

Steunend op theorema 3.3.1 en het feit dat $\alpha = \sqrt{3}$ en $g'(x) = 1 + 2cx$, moet c zo bepaald worden dat

$$-1 < 1 + 2c\sqrt{3} < 1.$$

Als we $c = -1/4$ kiezen dan is $g'(\sqrt{3}) = 1 - \dfrac{\sqrt{3}}{2} = 0.134$. Dit levert het volgende één-punt iteratieschema :

$$x_{n+1} = x_n - \frac{1}{4}(x_n^2 - 3) \qquad\qquad n \geq 0$$

In de volgende tabel starten we met $x_0 = 2$ en vinden we iteratieresultaten t.e.m. $n = 7$. We geven tevens de verhouding $(\alpha - x_n)/(\alpha - x_{n-1})$ die volgens theorema 3.3.1 de theoretische waarde $g'(\sqrt{3})$ moet benaderen.

n	x_n	$\alpha - x_n$	$(\alpha - x_n)/(\alpha - x_{n-1})$
0	2.0	$-2.68\,10^{-1}$	–
1	1.75	$-1.79\,10^{-2}$	0.0668
2	1.7343750	$-2.32\,10^{-3}$	0.130
3	1.7323608	$-3.10\,10^{-4}$	0.134
4	1.7320923	$-4.15\,10^{-5}$	0.134
5	1.7320564	$-5.56\,10^{-6}$	0.134
6	1.7320516	$-7.45\,10^{-7}$	0.134
7	1.7320509	$-1.00\,10^{-7}$	0.134

Om nu de Aitken extrapolatieformule te illustreren passen we ze toe startend vanaf de punten $x_3,\ x_4$ en x_5 . We bekomen

$$\alpha \approx x_5 - \frac{(x_5 - x_4)^2}{(x_3 - 2x_4 + x_5)} = 1.73205086\,.$$

Deze waarde bezit een fout $-5.\,10^{-8}$ wat ongeveer 100 maal beter is dan de fout op x_5 en ook nog beter dan deze voor x_6 en x_7 . We moeten ons realizeren dat deze verbetering bekomen werd met enkele eenvoudige rekenkundige bewerkingen. ◇

Opmerking 3.3.2

De formule (3.38) wordt vaak de Aitken Δ^2 methode genoemd, omdat de grootheden in het rechterlid ook uitdrukbaar zijn met behulp van de voorwaartse differentieoperator Δ. (Voor een gedetailleerde studie van die operator verwijzen we naar hoofdstuk 5.)
Als we beschikken over een rij $\{x_n \mid (n = 0, 1, \ldots)\}$ van waarden, dan wordt de voorwaartse differentie Δx_n gedefinieerd als

$$\Delta x_n = x_{n+1} - x_n \qquad\qquad n \geq 0\,.$$

Hogere machten $\Delta^k x_n$ worden recursief als volgt ingevoerd :

$$\Delta^k x_n = \Delta^{k-1}(\Delta x_n) \qquad \text{voor} \quad k \geq 2\,.$$

Uit de definitie volgt

$$\begin{aligned}
\Delta^2 x_n &= \Delta(x_{n+1} - x_n) = \Delta x_{n+1} - \Delta x_n \\
&= (x_{n+2} - x_{n+1}) - (x_{n+1} - x_n) = x_{n+2} - 2x_{n+1} + x_n\,.
\end{aligned}$$

Hiermee rekening houdend is het evident dat (3.38) te schrijven is als

$$\alpha \approx x_{n+2} - \frac{(\Delta x_{n+1})^2}{\Delta^2 x_n} \qquad \text{voor alle } n \geq 0\,.$$

◇

3.3.3 Het numeriek bepalen van meervoudige wortels

Een functie heeft een meervoudige wortel α met multipliciteit $p > 1$ als

$$f(x) = (x - \alpha)^p h(x) \tag{3.39}$$

met $h(\alpha) \neq 0$ en $h(x)$ continu in $x = \alpha$. Als $h(x)$ voldoende afleidbaar is bij $x = \alpha$ dan is (3.39) uiteraard equivalent met

$$f(\alpha) = f'(\alpha) = \ldots = f^{(p-1)}(\alpha) = 0\,, \quad f^{(p)}(\alpha) \neq 0\,.$$

Laat ons het effect van (3.39) op de Newton–Raphson methode beschouwen, i.e.

$$x_{n+1} = g(x_n)\,, \qquad g(x) = x - \frac{f(x)}{f'(x)} \qquad\qquad x \neq \alpha\,.$$

Voor $f(x)$ van het type (3.39) is

$$f'(x) = (x - \alpha)^p\, h'(x) + p(x - \alpha)^{p-1}\, h(x)$$

en

$$g(x) = x - \frac{(x - \alpha)h(x)}{ph(x) + (x - \alpha)h'(x)}\,.$$

Dit afleiden resulteert in

$$g'(x) = 1 - \frac{h(x)}{ph(x) + (x - \alpha)h'(x)} - (x - \alpha)\frac{d}{dx}\Big[\frac{h(x)}{ph(x) + (x - \alpha)h'(x)}\Big]$$

en dus

$$g'(\alpha) = 1 - \frac{1}{p} \neq 0 \qquad \text{voor} \qquad p > 1\,. \tag{3.40}$$

M.a.w. de Newton–Raphson methode wordt een lineair convergente methode met maat $(p - 1)/p$.
Om dit te verbeteren construeren we een functie $g(x)$ waarvoor $g'(\alpha) = 0$. Kiezen we

$$g(x) = x - p\frac{f(x)}{f'(x)}$$

dan volgt uit de aanpassing van de afleiding van (3.40) inderdaad dat $g'(\alpha) = 0$ en ook

$$\alpha - x_{n+1} = -\frac{1}{2}(\alpha - x_n)^2\, g''(\xi_n)\,,$$

aantonende dat we aldus opnieuw een orde van convergentie gelijk aan twee bereiken, zoals dit het geval was bij de originele Newton–Raphson methode voor enkelvoudige wortels.

Terwijl die aanpassing van het Newton-Raphson schema het probleem van meervoudige wortels schijnt op te lossen, moeten we ons realizeren dat de multipliciteit eigenlijk onbekend is en de moeilijkheid blijft bestaan. Een andere manier, die succesvoller is om het probleem te behandelen, bestaat erin een nieuwe functie te definiëren

$$\mu = \frac{f(x)}{f'(x)} \, .$$

Als α een wortel met multipliciteit p ($p \geq 1$) is en $f(x) = (x-\alpha)^p h(x)$, dan bezit

$$\mu(x) = \frac{(x-\alpha)h(x)}{ph(x) + (x-\alpha)h'(x)}$$

ook een wortel α maar met multipliciteit 1. De Newton-Raphson methode kan dan toegepast worden op de functie μ, wat resulteert in

$$g(x) = x - \frac{\mu(x)}{\mu'(x)} = x - \frac{f(x)/f'(x)}{[(f'(x))^2 - f(x)f''(x)]/(f'(x))^2}$$

of

$$g(x) = x - \frac{f(x)f'(x)}{[f'(x)]^2 - f(x)f''(x)} \, . \tag{3.41}$$

Als g aan de vereiste continuïteitsvoorwaarden voldoet, zal de iteratie (3.41), zoals elke Newton-Raphson iteratie, kwadratisch convergent zijn, wat de multipliciteit van de wortel α ook is. Vanuit theoretisch standpunt is er één bezwaar tegen die methode : de bijkomende bepaling van $f''(x)$ kan soms zeer moeilijk zijn. Merken we tevens op dat in de praktijk zal blijken dat de aanwezigheid van een meervoudige wortel ernstige afrondingsproblemen kan meebrengen.

DEMO 8 (Matlab)

Beschouw de vergelijking $f(x) = x^4 - 4x^2 + 4 = 0$, die een wortel met multipliciteit 2 bezit bij $x = \sqrt{2}$. Bepaal die wortel numeriek met het eenvoudige Newton-Raphson schema (3.11) en met een aangepast schema

$$x_{n+1} = x_n - p\frac{f(x_n)}{f'(x_n)} \, ,$$

waarbij p eerst 2 gekozen wordt en nadien 3. Start steeds met $x_0 = 1.5$. Bepaal in elk geval het aantal iteraties om een nauwkeurigheid te bereiken van de orde 10^{-8}. Pas tevens het schema (3.41) toe en vergelijk je resultaat met de voorgaande.

3.4 Stelsels van niet–lineaire vergelijkingen Uitbreiding van de Newton–Raphson methode

We zoeken hier benaderingen voor reële oplossingen van een stelsel niet–lineaire vergelijkingen met reële operatoren. Het stelsel bevat evenveel vergelijkingen als onbekenden. Voor de eenvoud van presentatie beperken we ons tot een stelsel van twee vergelijkingen met twee onbekenden

$$\begin{cases} f_1(x,y) = 0 \\ f_2(x,y) = 0\,. \end{cases}$$

De veralgemening naar n vergelijkingen met n onbekenden zal voor de hand liggend zijn, eenmaal de algemene methode is toegelicht. We zullen hier verder steunen op de linearizatietechniek ook besproken in paragraaf (3.2.3); nl. zij $\eta = (\alpha, \beta)$ de voorstelling van de exacte reële oplossing en weze $w_0 = (x_0, y_0)$ een aanvangsapproximatie, dan $\eta = w_0 + \delta$ met $\delta = (h, k)$ exacte correcties en ook $f_i(x_0 + h,\, y_0 + k) \equiv 0$ $(i = 1, 2)$.
Maken we nu gebruik van de Taylor reeksontwikkeling voor functies van twee veranderlijken, dan is

$$f_i(x_0 + h,\, y_0 + k) = \quad 0 = \quad f_i(x_0,\, y_0) + h\Big(\frac{\partial f_i}{\partial x}\Big)_0 + k\Big(\frac{\partial f_i}{\partial y}\Big)_0 + \Theta(h^2, hk, k^2)\,, \qquad (i = 1, 2)\,,$$

waarbij de index 0 betekent “te nemen in $x_0,\, y_0$”.
Linearizatie levert een exact stelsel voor h_1 en k_1 , respectievelijk benaderingen van h en k , nl.

$$\begin{cases} 0 = f_1(x_0,\, y_0) + h_1\Big(\dfrac{\partial f_1}{\partial x}\Big)_0 + k_1\Big(\dfrac{\partial f_1}{\partial y}\Big)_0 \\ 0 = f_2(x_0,\, y_0) + h_1\Big(\dfrac{\partial f_2}{\partial x}\Big)_0 + k_1\Big(\dfrac{\partial f_2}{\partial y}\Big)_0. \end{cases} \tag{3.42}$$

Dit stelsel levert unieke waarden voor h_1 en k_1 als het een stelsel van Cramer is; als nieuwe approximatie hebben we dan

$$w_1 = w_0 + \delta_1 \qquad \text{met} \qquad \delta_1 = (h_1\,,\, k_1)\,.$$

We kunnen de procedure herbeginnen met

$$w_1 = (x_1, y_1) \longrightarrow w_2 = (x_2, y_2) \longrightarrow w_3 = (x_3, y_3) \longrightarrow \ldots$$

Vermits bij iedere overgang van $w_j \rightarrow w_{j+1}$ de berekening in (3.42) van $f(w_j)$ vereist is, kan men nagaan of er al dan niet voldaan is aan

$$\mid f(w_j) \mid < \epsilon$$

met ϵ een opgegeven tolerantie. Is aan die voorwaarde voldaan, dan kan de iteratie gestopt worden. Het is geenszins zeker of dit iteratieproces zal convergeren of niet. Dit kan op een correcte wijze bestudeerd worden, net zoals dit gebeurd is in paragraaf (3.3.1) voor de één–punt iteratie methode. Deze studie valt echter buiten het kader van deze cursus.

Uitbreiding naar een stelsel van N vergelijkingen met N onbekenden is eenvoudig : Zij

$$f_i(u_1, u_2, \ldots, u_N) = 0 \quad (i = 1, \ldots, N)$$

Weze V_i, $(i = 1, \ldots, N)$ een gekende benadering van de exacte oplossing α_i, $(i = 1, \ldots, N)$, dan is $\alpha_i = V_i + \epsilon_i$. Laten we die relatie substitueren in de gegeven vergelijkingen en het linkerlid in Taylorreeks ontwikkelen, i.e.

$$\begin{aligned} &f_i(V_1 + \epsilon_1\,,\ V_2 + \epsilon_2\,, \ldots,\ V_N + \epsilon_N) \\ &\quad = f_i(V_1, V_2, V_3, \ldots, V_N) + \Big[\frac{\partial f_i}{\partial u_1}\epsilon_1 + \frac{\partial f_i}{\partial u_2}\epsilon_2 + \ldots + \frac{\partial f_i}{\partial u_N}\epsilon_N\Big]_{u_j = V_j} = 0 \qquad (3.43) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (i = 1, \ldots, N) \end{aligned}$$

We hebben hier tevens gelinearizeerd, d.w.z. hogere orde termen in ϵ_i verwaarloosd. Vergelijking (3.43) stelt N lineaire vergelijkingen voor in de N onbekenden $\epsilon_1, \epsilon_2, \ldots, \epsilon_N$. Wanneer de ϵ waarden berekend zijn, kunnen de $(V_i + \epsilon_i)$, $(i = 1, \ldots, N)$ gebruikt worden als startwaarden voor een volgende iteratie. Dit proces kan verder gezet worden tot voor de α_i benaderingen met de gewenste graad van nauwkeurigheid gevonden zijn.

3.5 Wortels van veeltermen

De methoden besproken in de vorige paragrafen kunnen alle toegepast worden voor het bepalen van wortels van veeltermen. Maar voor wortelbepaling bij polynomen is het gewenst de speciale structuur van dergelijke functies te benutten, indien mogelijk. Vooral omdat wij bij veeltermen soms wel geïnteresseerd zijn in de *complexe* en *reële* wortels. Daarom verdient de wortelbepaling van veeltermen onze speciale aandacht, en ook omdat dit probleem gedurende een periode van bijna 400 jaar sterk in de belangstelling van wiskundigen gestaan heeft.

3.5.1 Klassieke methoden voor derde- en vierdegraadsveeltermen

Vooreerst wensen we op te merken dat voor veeltermen met graad kleiner of gelijk vier, algebraïsche methodes bestaan om de wortels exact te bepalen. Voor vierkantsvergelijkingen is deze opmerking triviaal. Voor de derdegraadsvergelijkingen kunnen we o.m.

verwijzen naar de oplossingsmethode van Cardano. Beschouwen we hiertoe de kubische vergelijking

$$x^3 + b_1x^2 + b_2x + b_3 = 0\,.$$

(Merk op dat we de coëfficiënt van de hoogste graadsterm reeds gelijk 1 hebben gekozen).
Door de substitutie

$$x = y - b_1/3$$

herleidt bovenstaande vergelijking zich tot :

$$y^3 + py + q = 0\ , \tag{3.44}$$

waarbij $p = \dfrac{3b_2 - b_1^2}{3}\ , \quad q = \dfrac{2b_1^3}{27} - \dfrac{b_1b_2}{3} + b_3\,.$
Voeren we een nieuwe substitutie door van de volgende vorm :

$$y = u + v$$

met u en v twee nieuwe onbekenden. Vorige kubische vergelijking is dan equivalent met :

$$\begin{aligned}&(u+v)^3 + p(u+v) + q = 0 \qquad \text{of}\\ &u^3 + v^3 + q + (3uv + p)(u+v) = 0\,.\end{aligned}$$

Aangezien aan u en v nog een tweede voorwaarde mag opgelegd worden, kiezen we u en v zo dat

$$3uv + p = 0\,, \tag{3.45}$$

waardoor de kubische vergelijking zich herleidt tot

$$u^3 + v^3 + q = 0\ . \tag{3.46}$$

Vergelijkingen (3.45), (3.46) kunnen aangewend worden ter bepaling van u en v; deze laatste zijn nl. de oplossingen van het stelsel

$$\begin{cases} u^3 + v^3 = -q \\ 3uv = -p \end{cases} \qquad \text{of} \qquad \begin{cases} u^3 + v^3 = -q \\ u^3v^3 = -p^3/27\,. \end{cases}$$

Hieruit volgt dat u^3 en v^3 de wortels zijn van de vierkantsvergelijking

$$z^2 + qz - p^3/27 = 0 \tag{3.47}$$

m.a.w.

$$\left.\begin{array}{l} u^3 \\ v^3 \end{array}\right\} = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \, .$$

De drie y–wortels worden dan gegeven door

$$y_i = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} = \alpha_i + \beta_i \quad (i = 1, 2, 3) \tag{3.48}$$

waarbij $\alpha_i\beta_i = -p/3$. De wortels van de oorspronkelijke kubische vergelijking zijn dan

$$x_i = y_i - b_1/3 \quad (i = 1, 2, 3).$$

Als in vergelijking (3.44) $p = 0$, dan zijn de wortels van die vergelijking $\sqrt[3]{-q}$; als $q = 0$ dan zijn de wortels van (3.44) $y_1 = 0$ en $y_{2,3} = \pm\sqrt{-p}$. Als $p \neq 0$ en $q \neq 0$ maar $(q/2)^2 + (p/3)^3 = 0$ dan zijn de wortels van (3.44)

$$y_1 = \frac{3q}{p} \qquad y_2 = y_3 = -\frac{3q}{2p}.$$

Dit kan gemakkelijk als volgt ingezien worden. Uit $(q/2)^2 + (p/3)^3 = 0$ kan afgeleid worden dat

$$-\frac{q}{2} = \left(\frac{3q}{2p}\right)^3 .$$

In dit geval is

$$\left.\begin{array}{l} u^3 \\ v^3 \end{array}\right\} = -\frac{q}{2} = \left(\frac{3q}{2p}\right)^3 .$$

Hieruit kunnen de volgende α_i , β_i $(i = 1, 2, 3)$ afgeleid worden.

$$\begin{aligned} \alpha_1 &= \frac{3q}{2p} & \beta_1 &= \frac{3q}{2p} \\ \alpha_2 &= \frac{3q}{2p}\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) & \beta_2 &= \frac{3q}{2p}\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) \\ \alpha_3 &= \frac{3q}{2p}\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) & \beta_3 &= \frac{3q}{2p}\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) . \end{aligned}$$

Rekening houdend met het feit dat $\alpha_i\beta_i = -p/3$ $(i = 1, 2, 3)$ krijgen we de vooropgestelde y_i waarden $(i = 1, 2, 3)$.

Merk wel op dat wanneer alle coëfficiënten van de kubische vergelijking reële getallen zijn, de wortels kunnen gevonden worden zonder expliciet gebruik te maken van Cardano's formule (3.48). Zeker in het geval waarbij

$$\frac{q^2}{4} + \frac{p^3}{27} < 0$$

is het nuttig dit even onder ogen te nemen. In dit bijzonder geval zijn de wortels van de kwadratische vergelijking elkaars complex toegevoegde. Die beide wortels u^3 en v^3 kunnen kortweg genoteerd worden als $Re^{\pm i\phi}$, waarbij $R^2 = -p^3/27$ en $R(e^{i\phi} + e^{-i\phi}) = -q$ of $\cos\phi = -\frac{q}{2}\sqrt{-(3/p)^3}$. Hieruit kan men dan gemakkelijk de y_k waarden afleiden, nl.

$$y_k = 2\sqrt{\frac{-p}{3}}\cos\frac{\phi + 2k\pi}{3}\,, \qquad (k = 1, 2, 3). \tag{3.49}$$

Merk op dat in dit geval de drie wortels van (3.44) reëel zijn.

Voorbeeld 3.5.1

Beschouw de vergelijking $x^3 + 3x^2 - 6x + 20 = 0$.
De substitutie $x = y - 1$ levert :

$$y^3 - 9y + 28 = 0\,,$$

waarvoor we onmiddellijk kunnen schrijven :

$$y = \sqrt[3]{-14 + \sqrt{196 - 27}} + \sqrt[3]{-14 - \sqrt{196 - 27}} = \sqrt[3]{-1} + \sqrt[3]{-27}\,.$$

Vermits $\alpha = \sqrt[3]{-1}$ volgende waarden bezit :

$$\alpha_1 = -1 \quad ; \alpha_2 = \frac{1}{2} + \frac{\sqrt{3}}{2}i \quad ; \alpha_3 = \frac{1}{2} - \frac{\sqrt{3}}{2}i$$

en $\beta = \sqrt[3]{-27}$ met de bijkomende voorwaarde $\alpha_i\beta_i = 3 \quad (i = 1, 2, 3)$ vinden we

$$\beta_1 = -3 \quad ; \quad \beta_2 = 3\Big(\frac{1}{2} - \frac{\sqrt{3}}{2}i\Big) \quad ; \quad \beta_3 = 3\Big(\frac{1}{2} + \frac{\sqrt{3}}{2}i\Big)$$

zodat

$$y_1 = \alpha_1 + \beta_1 = -4 \quad ; y_2 = \alpha_2 + \beta_2 = 2 - i\sqrt{3} \quad ; y_3 = \alpha_3 + \beta_3 = 2 + i\sqrt{3}$$

$$\text{en} \qquad x_1 = -5 \quad ; x_2 = 1 - i\sqrt{3} \qquad ; x_3 = 1 + i\sqrt{3}\,.$$

◇

Voorbeeld 3.5.2

Beschouw de vergelijking

$$y^3 - 3y + 1 = 0$$

waarvoor $(q/2)^2 + (p/3)^3 = \frac{1}{4} - 1 < 0$. We passen (3.49) toe en vinden $\cos\phi = -\frac{1}{2}\sqrt{1} = -\frac{1}{2}$, $\phi = 2\pi/3$; derhalve is

$$y_k = 2\cos(2\pi/3 + 2k\pi)/3, \qquad k = 1, 2, 3,$$

wat betekent $y_1 \approx -1.880$, $y_2 \approx 0.348$ en $y_3 \approx 1.532$. $\diamond$

Voor het oplossen van reële vierdegraadsvergelijkingen is bv. de methode van Descartes bruikbaar. Startend van de vergelijking $x^4 + ax^3 + bx^2 + cx + d = 0$ wordt de kubische term verwijderd door de transformatie $x = y - a/4$ resulterend in de vergelijking $y^4 + qy^2 + ry + s = 0$, waarbij

$$\begin{aligned} q &= b - 3a^2/8 \\ r &= c - ab/2 + a^3/8 \\ s &= d - ac/4 + a^2b/16 - 3a^4/256\,. \end{aligned}$$

Vermits we van een reële quartische vergelijking vertrokken zijn, kan deze vier, twee of geen reële wortels hebben. In elk geval is het mogelijk een opsplitsing te bewerkstelligen in een product van twee vierkantsvergelijkingen, nl.

$$(y^2 + 2ky + l)(y^2 - 2ky + m) = y^4 + qy^2 + ry + s$$

met reële k, l en m. Door de identificatie van de coëfficiënten van gelijke machten van y verkrijgen we

$$\begin{cases} l + m - 4k^2 = q \\ 2k(m - l) = r \\ lm = s\ . \end{cases}$$

Door uit die drie betrekkingen l en m te elimineren en $k^2 = z$ te stellen, bekomen we

$$z^3 + \alpha z^2 + \beta z + \gamma = 0$$

met $\alpha = q/2$; $\beta = (q^2 - 4s)/16$ en $\gamma = -r^2/64$. Vermits $\gamma \leq 0$ is er tenminste één positieve wortel z en kunnen k, l en m de één na de andere afgeleid worden. Op deze wijze is het oplossen van de vierdegraadsvergelijking gereduceerd tot het bepalen van de wortels van twee vierkantsvergelijkingen.

Wenst men de reële wortels van veeltermen numeriek te bepalen, dan kan men uiteraard beroep doen op één van de iteratieve methoden uit voorgaande paragrafen. Er bestaan echter typische methoden voor veeltermen, o.a. de methode van Bairstow voor het benaderen van zowel reële als complexe wortels van hogere machtsvergelijkingen met reële coëfficiënten.

DEMO 5 (Maple)

Zie Claroline

3.5.2 Methode van Bairstow

Schrijven we expliciet de Euclidische deling van een gegeven n^{de} graadsveelterm door een willekeurige tweede–graadsveelterm $z^2 + pz + q$:

$$\begin{aligned} & z^n + a_1 z^{n-1} + \ldots + a_n \\ & \quad = (z^2 + pz + q) \quad \underset{\text{quotiëntveelterm}}{(z^{n-2} + b_1 z^{n-3} + \ldots + b_{n-2})} \quad + \quad \underset{\text{restveelterm}}{(rz + s)} \end{aligned}$$

waaruit we bekomen

$$\begin{cases} a_1 = b_1 + p \\ a_2 = b_2 + pb_1 + q \\ a_3 = b_3 + pb_2 + qb_1 \\ \vdots \\ a_k = b_k + pb_{k-1} + qb_{k-2} \\ \vdots \\ a_{n-2} = b_{n-2} + pb_{n-3} + qb_{n-4} \\ a_{n-1} = r + pb_{n-2} + qb_{n-3} \\ a_n = s + qb_{n-2} \ . \end{cases} \tag{3.50}$$

Voor gekozen p en q en gegeven a_i $(i = 1, \ldots, n)$ is dit een stelsel lineaire vergelijkingen in b_i $(i = 1, \ldots, n-2)$, r en s. Deze laatste grootheden kunnen door het van boven naar onder doorlopen van (3.50) recursief gevonden worden. Voor verder gebruik introduceren we twee nieuwe grootheden b_{n-1} en b_n en definiëren

$$b_k = a_k - pb_{k-1} - qb_{k-2} \qquad (k = 1, 2, \ldots, n) \tag{3.51}$$

met $b_0 = 1$ en $b_{-1} = 0$ en waarbij

$$\begin{aligned} b_{n-1} &= a_{n-1} - pb_{n-2} - qb_{n-3} = r \\ b_n &= a_n - pb_{n-1} - qb_{n-2} = s - pb_{n-1} \ . \end{aligned} \tag{3.52}$$

We streven er uiteraard naar zulkdanige p en q te vinden dat de restveelterm nul wordt. Uit bovenstaande blijken r en s polynomiale functies R en S in p en q te zijn, m.a.w. we streven er naar oplossingen te vinden van

$$\begin{cases} R(p,q) = 0 \\ S(p,q) = 0 \; . \end{cases}$$

Om een reële oplossing van dit stelsel te bepalen, gebruiken we de Newton–Raphson methode uit paragraaf 3.4, vertrekkend van een aanvangsapproximatie p_0, q_0. Het stelsel (3.42) overgebracht in de hier gebruikte notatie luidt dan :

$$\begin{cases} (\frac{\partial R}{\partial p})_0 \, h_1 + (\frac{\partial R}{\partial q})_0 \, k_1 = -R(p_0 \, , \, q_0) \\ (\frac{\partial S}{\partial p})_0 \, h_1 + (\frac{\partial S}{\partial q})_0 \, k_1 = -S(p_0 \, , \, q_0) \end{cases} \tag{3.53}$$

met h_1 en k_1 benaderende correcties voor p_0 en q_0 en de index 0 wijzend op het feit dat de waarden moeten bepaald worden voor $p = p_0$ en $q = q_0$. Gebruik makend van (3.52) is dit ook te noteren als

$$\begin{cases} \Big(\frac{\partial b_{n-1}}{\partial p}\Big)_0 h_1 + \Big(\frac{\partial b_{n-1}}{\partial q}\Big)_0 k_1 = -(b_{n-1})_0 \\ \Big(\frac{\partial b_n}{\partial p} + p\frac{\partial b_{n-1}}{\partial p} + b_{n-1}\Big)_0 h_1 + \Big(\frac{\partial b_n}{\partial q} + p\frac{\partial b_{n-1}}{\partial q}\Big)_0 k_1 = -(b_n + pb_{n-1})_0 \; . \end{cases}$$

Als we de eerste vergelijking vermenigvuldigen met p_0 en van de tweede aftrekken, krijgen we volgend gelijkwaardig stelsel :

$$\begin{cases} \Big(\frac{\partial b_{n-1}}{\partial p}\Big)_0 h_1 + \Big(\frac{\partial b_{n-1}}{\partial q}\Big)_0 k_1 = -(b_{n-1})_0 \\ \Big(\frac{\partial b_n}{\partial p} + b_{n-1}\Big)_0 h_1 + \Big(\frac{\partial b_n}{\partial q}\Big)_0 k_1 = -(b_n)_0 \; . \end{cases} \tag{3.54}$$

Partieel afleiden naar p en q van (3.51) levert

$$\begin{aligned} -\frac{\partial b_k}{\partial p} &= b_{k-1} + p\frac{\partial b_{k-1}}{\partial p} + q\frac{\partial b_{k-2}}{\partial p} \; ; \; \frac{\partial b_0}{\partial p} = \frac{\partial b_{-1}}{\partial p} = 0 \; ; \\ -\frac{\partial b_k}{\partial q} &= b_{k-2} + p\frac{\partial b_{k-1}}{\partial q} + q\frac{\partial b_{k-2}}{\partial q} \; ; \; \frac{\partial b_0}{\partial q} = \frac{\partial b_{-1}}{\partial q} = 0 \; . \end{aligned} \tag{3.55}$$

Stel nu $\partial b_k/\partial p = -c_{k-1}$ $(k = 1, 2, \ldots, n)$ dan bekomen we door inductie uit (3.55) :

$$\frac{\partial b_k}{\partial q} = \frac{\partial b_{k-1}}{\partial p} = -c_{k-2}$$

en (3.55) is dan te schrijven als

$$\begin{aligned} c_{k-1} &= b_{k-1} - pc_{k-2} - qc_{k-3} \\ c_{k-2} &= b_{k-2} - pc_{k-3} - qc_{k-4} \end{aligned}$$

wat herleidbaar is tot één vergelijking

$$c_k = b_k - pc_{k-1} - qc_{k-2} \; ; \; c_0 = 1 \; ; \; c_{-1} = 0 \qquad (k = 1, 2, \ldots, n-1)\,. \tag{3.56}$$

Hieruit zien we dat de c_k op dezelfde wijze formeel berekenbaar zijn uit de b_k als de b_k uit de a_k (zie (3.51). Het stelsel (3.54) is dan herschrijfbaar als

$$\begin{aligned} &(c_{n-2})_0\, h_1 + (c_{n-3})_0\, k_1 = (b_{n-1})_0 \\ &[(c_{n-1})_0 - (b_{n-1})_0]\, h_1 + (c_{n-2})_0\, k_1 = (b_n)_0\,. \end{aligned} \tag{3.57}$$

Eenmaal h_1 en k_1 bepaald, kunnen we opnieuw vertrekken van een nieuwe approximatie

$$\begin{aligned} p_1 &= p_0 + h_1 \\ q_1 &= q_0 + k_1 \end{aligned}$$

en de procedure herbeginnen om de sequentie

$$\begin{aligned} p_0 \to p_1 \to p_2 &\longrightarrow p \\ q_0 \to q_1 \to q_2 &\longrightarrow q \end{aligned}$$

op te bouwen. De finaal gevonden waarden kunnen dan aangewend worden in de vierkantsvergelijking $z^2 + pz + q = 0$. Wortels van deze vierkantsvergelijking zijn dan uiteraard ook oplossingen van de oorspronkelijke $n^{\rm de}$ graadsvergelijking.

Als aanvangsapproximatie p_0 en q_0 kan men de volgende keuze maken :

(a) weet men iets over de orde van grootte van de te berekenen wortels, dan kan men $p_0 = -(\text{som})$ en $q_0 = \text{product}$ kiezen;

(b) weet men niets over de wortels, dan kan men bv. vertrekken van $p_0 = q_0 = 0$.

Het bovenstaande iteratief proces kan als volgt gestopt worden. Bij het doorlopen van (3.51)(3.52) in de berekening van de $j^{\rm de}$ approximatie (p_j , q_j), bepalen we uiteraard ook $(b_{n-1})_{j-1}$ en $(b_n)_{j-1}$ $(j = 0, 1, 2, \ldots)$. Wanneer deze beide getallen in absolute waarden terzelfdertijd kleiner worden dan een vooropgegeven tolerantie ϵ zullen uiteraard ook $\mid (r)_{j-1} \mid < \epsilon$ en $\mid (s)_{j-1} \mid < \epsilon$ en kan het iteratieproces stopgezet worden.

3.5.3 Polynomiale deflatie

Eenmaal twee wortels van een n^{de} graadsvergelijking bepaald d.m.v. de Bairstow methode, of m.a.w. eenmaal een Euclidische deler van de vorm $z^2 + pz + q$ afgeleid, is het uiteraard mogelijk de deling echt uit te voeren en een veelterm van de graad $(n-2)$ te bekomen. Dit reductieproces wordt deflatie genoemd. De Bairstow methode kan op die lagere–orde veelterm opnieuw toegepast worden, enz... Nochtans kunnen hier ernstige problemen rijzen. Vermits p en q niet exact gevonden worden, zal de lagere–graad polynoom fouten vertonen in elk van zijn coëfficiënten, zodat bij grote n–waarden de laatst berekende wortels nog weinig betrouwbaar kunnen zijn.

J.H. Wilkinson heeft ook de effecten van deflatie onderzocht en beveelt de volgende strategie aan.

(a) Start met het bepalen van de wortels met de kleinste waarde, en eindig met die welke het grootst zijn.

(b) Na benaderingen te hebben bekomen voor alle wortels, herbegin het Bairstow proces voor de originele veelterm, maar kies p_0 en q_0 waarden gebaseerd op die benaderingen.

Algemene nota

Methoden om veeltermvergelijkingen op te lossen behoren tot een zeer oud onderzoeksgebied, dat teruggaat op zijn minst tot in de Babylonische periode. Er zijn veel methoden hiervoor beschreven in de literatuur en veel nieuwe methoden zijn in de laatste decaden ontwikkeld.

De eerste publicatie, die de oplossing van de vergelijkingen van de derde graad bevat, vond plaats in het werk van *Hieronimo Cardano* “Ars magna de Regulis Algebraicis”, dat, gedrukt in Nürnberg, in 1545 verscheen. Cardano hield zich niet alleen bezig met wiskunde, die hij, nauwelijks 22 jaar oud, te Pavia doceerde, maar was na 1535 ook werkzaam als arts, nadat hij in 1526 in Padua de graad van dokter in de geneeskunde had behaald. Cardano noemt zichzelf niet de ontdekker van die formule (3.48), maar noemt op de eerste plaats een zekere Scipione del Ferro.

René Descartes of *Cartesius* (1596-1650) van wie we de methode voor de wortelbepaling van quartische vergelijkingen hebben besproken, is vooral bekend als Frans filosoof. Hij ontving zijn opleiding in het jezuïetencollege van La Flèche, dat hij in 1612 verliet. Zijn familie had hem voorbestemd voor een militaire loopbaan en zo diende hij onder prins Maurits van Nassau en de keurvorst van Beieren. Nadat hij het militair leven vaarwel zegde, vestigde hij zich eerst te Parijs en later in Holland. In 1637 publiceerde hij in het Frans zijn beroemde “Discours de la méthode”. Ook op zuiver wetenschappelijk gebied heeft Descartes vele sporen nagelaten; hij schreef o.a. over meteoren en over het licht. Bovenal dankt men aan hem de scherpe formulering van een nieuwe methode voor meetkundig onderzoek, nl. die m.b.v. coördinaten. Hij zette zijn ideeën uiteen in een geschrift “La géométrie” uitgegeven in Leiden in 1637.

Sir Isaac Newton (1642-1727) is een Engels wis-, natuur- en sterrenkundige. In 1660 slaagde hij, op 18-jarige leeftijd, met glans voor het toelatingsexamen voor Trinity College te

Cambridge. In 1668 verwierf hij de doctorstitel en werd hij hoogleraar aan hetzelfde College. In zijn beginperiode was hij vooral geïnteresseerd in opticaproblemen. In 1669 ontwikkelde Newton een methode voor het berekenen van wortels van een kubische vergelijking. Later publiceerde hij de methode die nu toegeschreven wordt aan Newton en Raphson, als een middel om de Kepler vergelijking voor de bepaling van de baan van een planeet op te lossen. In 1687 verscheen zijn "Philosophiae Naturalis Principia Mathematica", het fundament van de mechanica en de gehele klassieke natuurkunde. In 1701 trad hij af als hoogleraar te Cambridge; zijn ex-collega's kozen hem om hen in het parlement te vertegenwoordigen; in 1705 verhief koningin Anna hem in de adelstand. In de laatste decennia van zijn leven was de wetenschappelijke activiteit van Newton zeer beperkt. Hij verrichtte veel werk voor de voorbereiding van de tweede uitgave van de Principia, in samenwerking met de jonge wiskundige Cotes (zie ook cursus Numerieke Analyse en numerieke methoden ter benadering van bepaalde integralen). Met Leibnitz kan hij beschouwd worden als de grondlegger van de differentiaal- en integraalrekening (hijzelf sprak van fluxierekening). In zijn "Arithmetica universalis" behandelde hij het onderzoek van krommen, de algebraïsche vergelijkingen en vele andere wiskundige problemen.

In ongeveer 1690 formuleerde *Joseph Raphson* Newton's ideeën voor het speciale geval van een veelterm in een vorm die dichter aansluit bij de huidige formulering van de formule die gebruikt wordt als de Newton-Raphson methode.
