SIAM Review
Vol. 63, No. 3, pp. 585–624

How Many Steps Still Left to x^*?*

Emil Cătinaş†

Abstract. The high speed of $x_k \to x^* \in \mathbb{R}$ is usually measured using the C-, Q-, or R-orders:

$$\lim \frac{|x^* - x_{k+1}|}{|x^* - x_k|^{p_0}} \in (0, +\infty), \ \lim \frac{\ln |x^* - x_{k+1}|}{\ln |x^* - x_k|} = q_0, \ \text{or} \ \lim \left| \ln |x^* - x_k| \right|^{\frac{1}{k}} = r_0.$$

By connecting them to the natural, term-by-term comparison of the errors of two sequences, we find that the C-orders—including (sub)linear—are in agreement. Weird relations may appear though for the Q-orders: we expect $|x^* - x_k| = \mathcal{O}(|x^* - y_k|^\alpha)\ \forall \alpha > 1$ to imply "$\geq$" for the Q-orders of $\{x_k\}$ vs. $\{y_k\}$; the contrary is shown by an example providing no vs. infinite Q-orders. The R-orders appear to be even worse: an $\{x_k\}$ with infinite R-order may have unbounded nonmonotone errors: $|x^* - x_{k+1}|/|x^* - x_k| \to +\infty$.

Such aspects motivate the study of equivalent definitions, computational variants, and so on.

These orders are also the perspective from which we analyze the three basic iterative methods for nonlinear equations in $\mathbb{R}$. The Newton method, widely known for its quadratic convergence, may in fact attain any C-order from $[1, +\infty]$ (including sublinear); we briefly recall such convergence results, along with connected aspects (such as historical notes, known asymptotic constants, floating point arithmetic issues, and radius of attraction balls), and provide examples.

This approach leads to similar results for the successive approximations method, while the secant method exhibits different behavior: it may not have high C-orders, but only Q-orders.

uncaught typo: obviously, limsup .. = \infty instead of lim .. = \infty (see Remark 2.12c, where this is correctly treated.)

Key words. (computational) convergence orders, iterative methods, Newton method, secant method, successive approximations, asymptotic rates

AMS subject classifications. 65-02, 65-03, 41A25, 40-02, 97-02, 97N40, 65H05

DOI. 10.1137/19M1244858

Contents

*Received by the editors February 14, 2019; accepted for publication (in revised form) September 17, 2020; published electronically August 5, 2021.
https://doi.org/10.1137/19M1244858

†"Tiberiu Popoviciu" Institute of Numerical Analysis, Cluj-Napoca, Romania (ecatinas@ictp.acad.ro, http://www.ictp.acad.ro/catinas).

1. Introduction. The analysis of the convergence speed of sequences is an important task, since in numerical applications the aim is to use iterative methods with fast convergence, which provide good approximations in few steps.

The ideal setting for studying a given sequence $\{x_k\} := (x_k)_{k\geq 0} \subset \mathbb{R}$ is that of *error-based analysis*, where the finite limit x^* is assumed to be known and we analyze the absolute values of the errors, $e_k := |x_k - x^*|$. The errors allow the comparison of the speeds of two sequences in the most natural way, even if their limits are distinct (though here we write both as x^*). Let us first define notation, for later reference.

NOTATION. *Given $\dot{x}_k, \mathring{x}_k \to x^* \in \mathbb{R}$, denote their errors by $\{\dot{e}_k\}$, resp., $\{\mathring{e}_k\}$, etc.*

COMPARISON A. *$\{\dot{x}_k\}$ converges faster (not slower) than $\{\mathring{x}_k\}$ if*

$$(\dot{e}_k \leq \mathring{e}_k) \qquad |x^* - \dot{x}_k| \leq |x^* - \mathring{x}_k|, \quad k \geq k_0$$

(or, in brief, $\{\dot{x}_k\}$ is $(\dot{e}_k \leq \mathring{e}_k)$ faster than $\{\mathring{x}_k\}$) and strictly faster if

$$(\dot{e}_k < \mathring{e}_k) \qquad |x^* - \dot{x}_k| < |x^* - \mathring{x}_k|, \quad k \geq k_0.$$

Furthermore, increasingly faster convergence holds if
- for some $c \in (0,1)$ we have $|x^* - \dot{x}_k| \leq c|x^* - \mathring{x}_k|, k \geq k_0$;
- the constant c above is not fixed, but tends to zero (see, e.g., [9, p. 2], [53]*),*

$$(\dot{e}_k = o(\mathring{e}_k)) \qquad |x^* - \dot{x}_k| = o(|x^* - \mathring{x}_k|) \quad \text{as } k \to \infty;^{1}$$

[1] This means that $|x^* - \dot{x}_k| \leq c_k|x^* - \mathring{x}_k|$, $k \geq k_0$, with $c_k \to 0$, which allows a finite number of elements c_k to be greater than one.

- *given* $\alpha > 1$, *we have*

$$(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha})) \qquad |x^* - \dot{x}_k| = \mathcal{O}(|x^* - \mathring{x}_k|^{\alpha}) \text{ as } k \to \infty.^2$$

Obviously, $(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha})) \Rightarrow (\dot{e}_k = o(\mathring{e}_k)) \Rightarrow (\dot{e}_k < \mathring{e}_k) \Rightarrow (\dot{e}_k \le \mathring{e}_k)$.

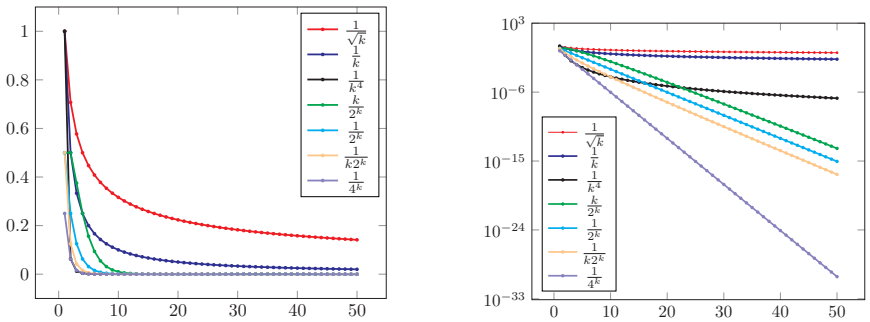
Let us first illustrate the convergence speed in an intuitive fashion, using some graphs.

EXAMPLE 1.1. *Consider the following two groups of sequences* ($\{x_k\} = \{e_k\}$)*:*
(a) $\{\frac{1}{\sqrt{k}}\}$, $\{\frac{1}{k}\}$, $\{\frac{1}{k^4}\}$;
(b) $\{\frac{k}{2^k}\}$, $\{\frac{1}{2^k}\} = \{(\frac{1}{2})^k\} = \{2^{-k}\}$, $\{\frac{1}{k2^k}\}$, $\{\frac{1}{4^k}\}$.

Handling the first terms of these sequences raises no difficulties, as all the common programming languages represent them as $\mathrm{fl}(x_k)$ *in standard double precision arithmetic (called* `digits64` *by the IEEE* 754-2008 *standard); see Exercise* 1.3 *below.*

uncaught typo: binary64 instead of digits64

The plotting of (k, e_k) *in Figure* 1.1(a) *does not help in analyzing the behavior of the errors, as for* $k \ge 10$ *we cannot see much, even if the graph is scaled.*[3] *All we can say is that* $\{\frac{1}{\sqrt{k}}\}$ *has the slowest convergence, followed by* $\{\frac{1}{k}\}$*, and also that the first terms of* $\{\frac{1}{k^4}\}$ *are smaller than those of* $\{\frac{1}{2^k}\}$.



(a) Plotting (k, e_k) does not help much.

(b) Plotting $(k, \lg e_k)$ is better.

Fig. 1.1 *(Scaled) Cartesian vs. semilog coordinates in visualizing the errors.*

In order to compare the last terms we must use the semilog coordinates $(k, \lg e_k)$ *in Figure* 1.1(b) *(see* [43, p. 211]*); for* $\{\frac{1}{2^k}\}$*, they are* $(k, k \lg \frac{1}{2})$*, belonging to the line* $y = ax$, $a = \lg \frac{1}{2}$. *As the graph is twice scaled,*[4] *it actually shows another line,* $y = bx$.

Figure 1.1(b) *shows that, in reality,* $\{\frac{1}{2^k}\}$ *is* $(\dot{e}_k < \mathring{e}_k)$ *faster than* $\{\frac{1}{k^4}\}$ $(k_0 = 17)$*. We will see later that the convergence is sublinear in group* (a) *and linear in group* (b).

EXERCISE 1.2. *Show that* (a) $\{\frac{1}{2^k}\}$ *is* $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha}))\ \forall \alpha > 1]$ *faster than* $\{\frac{1}{k^4}\}$*;*
(b) $\{\frac{1}{2^k}\}$ *is* $(\dot{e}_k = o(\mathring{e}_k))$ *but not* $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha}))$ *for some* $\alpha > 1]$ *faster than* $\{\frac{k}{2^k}\}$.

EXERCISE 1.3. (a) *The smallest positive* `digits64` *number, which we denote by* `realmin(digits64)`*, is* $2.225\,073\,858\,507\,201\,4 \cdot 10^{-308}$. *Analyzing its binary representation, express its value as a power of* 2 *(see* [69, p. 14], [28], *or* [45]*).*

uncaught typo: binary64 instead of digits64

[2] That is, $\exists K > 0$ such that $|x^* - \dot{x}_k| \le K|x^* - \mathring{x}_k|^{\alpha}$, $k \ge k_0$.

[3] The implicit fitting of the figures in the window usually results in different units for the two axes (scaled graph). In Figure 1.1(a) we would see even less if the graph were unscaled.

[4] The first scaling is by the use of an unequal number of units (50 on axis x vs. 36 on axis y), and the second one results from the fact that the final figure is a rectangle instead of a square.

(b) *For each sequence from* (a) *and* (b) *in the example above, compute the largest* k *such that* $\mathrm{fl}(x_k) \neq 0$.

(c) *Give a formula for the largest index* k *for which all the elements* $\mathrm{fl}(x_k)$ *of all the sequences from* (a) *and* (b) *in the example above are nonzero.*

Next we analyze some increasingly faster classes of sequences.

EXAMPLE 1.4. *Consider*

(c) $\left\{\frac{1}{2^{k^2}}\right\}$, $\left\{\frac{1}{2^{k^3}}\right\}$, $\left\{\frac{1}{k^k}\right\}$;

(d) $\left\{\frac{1}{2^{\frac{2^k}{k}}}\right\}$, $\left\{\frac{1}{2^{2^k}}\right\} = \left\{\left(\frac{1}{2}\right)^{2^k}\right\} = \{2^{-2^k}\}$, $\left\{\frac{1}{2^{k2^k}}\right\}$, $\left\{\frac{1}{3^{2^k}}\right\}$;

(e) $\left\{\frac{1}{2^{3^k}}\right\}$;

(f) $\left\{\frac{1}{2^{k2^{\sqrt{k}}}}\right\}$, $\left\{\frac{1}{2^{2^{k^2}}}\right\}$.

In Figure 1.2 *we plot* $\left\{\frac{1}{4^k}\right\}$, $\left\{\frac{1}{2^{k^2}}\right\}$, *and* $\left\{\frac{1}{k^k}\right\}$. *Their graphs are on a line (the fastest from Figure* 1.1(b)*), on a parabola* $y = cx^2$, *and, resp., in between.*

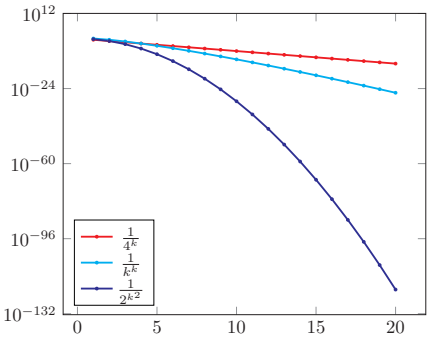


Fig. 1.2 *Linear vs. superlinear order.*

The computation with what appears at first sight to be a "reasonable" number of terms (say, 10*) becomes increasingly challenging as we successively consider* $\{x_k\}$ *from* (c)–(f).

We leave as an exercise the calculation/computation of the largest index k *for each* $\{x_k\}$ *in* (c)–(f) *such that* $\mathrm{fl}(x_k)$ *is a nonzero* `digits64` *number; we note though that all* $\mathrm{fl}(1/2^{2^{k^2}})$ *and* $\mathrm{fl}(1/2^{k2^{\sqrt{k}}})$ *are zero for* $k \geq 4$, *respectively,* $k \geq 5$, *and so* `digits64` *numbers are not enough here.*

uncaught typo: binary64 instead of digits64

There are plenty of choices for increased precision: the well-known MATLAB [58] *(but with the Advanpix toolbox* [1] *providing arbitrary precision), the recent, freely available Julia* [5], *and many others (not to mention the programming languages for symbolic computation). All the graphs from this paper were obtained using the tikz/pgf package* [92] *for LaTeX.*[5] *The figures are easy to customize and the instructions to generate them have a simple form, briefly,*

```
\addplot [domain=1:20, samples=20] {1/sqrt(x)};
```

Equally important, we will see that the tikz/pgf library allows the use of higher precision than `digits64` *(more precisely, smaller/larger numbers by increased representation of the exponent, e.g.,* `realmin` *around* 10^{-6500}*). The LaTeX sources for the figures are posted on https://github.com/ecatinas/conv-ord.*

[5]The initial TeX system was created by D. E. Knuth [54].

The sequence groups (c)–(f) *are presented in Figure* 1.3; $\left\{\frac{1}{2^{k^3}}\right\}$ *belongs to a cubic parabola, and* $\left\{\frac{1}{2^{2^k}}\right\}$ *to an exponential. The parabola of* $\left\{\frac{1}{2^{k^2}}\right\}$ *appears flat, which shows how much faster the other sequences converge. The descent of the three terms of* $\left\{1/2^{2^{k^2}}\right\}$ *is the steepest.*

As will be seen later, the sequence groups (c)–(f) *have increasing order: strict superlinear, quadratic, cubic, and infinite, respectively.*

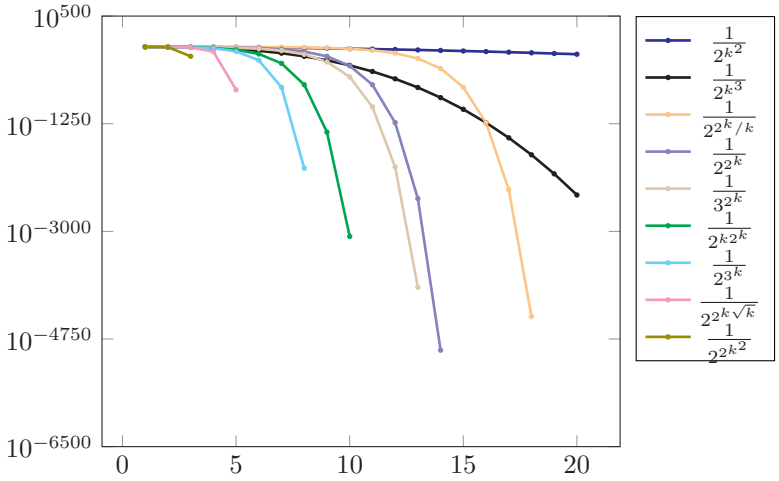


Fig. 1.3 *Sequences with superlinear, quadratic, cubic, and infinite orders.*

EXERCISE 1.5. *Show that* $\left\{\frac{1}{3^{2^k}}\right\}$ *is* $\left[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha})),\ \alpha \in (1, \frac{\ln 3}{\ln 2}]\right]$ *faster than* $\left\{\frac{1}{2^{2^k}}\right\}$.

Obviously, when comparing two sequences by $(\dot{e}_k < \mathring{e}_k)$, the faster one must have its graph below the other one $(k \geq k_0)$ (cf. Polak [75, p. 47]). While the more concave the graph the better (cf. Kelley [50, Ex. 5.7.15]), fast convergence does not in fact necessarily require smooth graphs (see Jay [47]).

EXERCISE 1.6. *Prove that any convergence curve between* $\left\{\frac{1}{2^{3^k}}\right\}$ *and* $\left\{\frac{1}{3^{2^k}}\right\}$ *is at least* $\left[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha})),\ \alpha \in (1, \frac{\ln 3}{\ln 2}]\right]$ *faster than* $\left\{\frac{1}{2^{2^k}}\right\}$ *(hint: use Exercise* 1.5*).*

Consider, for instance, the merging of certain sequences.

EXAMPLE 1.7. *Let*

$$x_k^a = \begin{cases} \frac{1}{3^{2^k}}, & k \text{ odd}, \\ \frac{1}{2^{3^k}}, & k \text{ even}, \end{cases} \qquad x_k^b = \begin{cases} \frac{1}{2^{2^k}}, & k \text{ even}, \\ \frac{1}{3^{2^k}}, & k \text{ odd}, \end{cases} \qquad x_k^c = \begin{cases} \frac{1}{2^{2^k}}, & k \text{ even}, \\ \frac{1}{5^{2^k}}, & k \text{ odd}. \end{cases}$$

$\left\{\frac{1}{2^{3^k}}\right\}, \{x_k^a\}, \left\{\frac{1}{3^{2^k}}\right\}, \{x_k^b\}, \left\{\frac{1}{2^{2^k}}\right\}$, *written in* $(\dot{e}_k \leq \mathring{e}_k)$ *order of decreasing speed, are plotted in Figure* 1.4(a). *The usual ranking by convergence orders is instead* $\left\{\frac{1}{2^{3^k}}\right\}$, $\left\{\frac{1}{3^{2^k}}\right\}$, $\left\{\frac{1}{2^{2^k}}\right\}$, $\{x_k^b\}$, $\{x_k^a\}$ *(C-orders* 3, 2, 2, *R-order* 2, *no order). Though* $\{x_k^a\}$ *has the second fastest convergence, the orders actually rank it as the slowest.* $\{x_k^b\}$ *does not have C- or Q-order (but just exact R-order* 2*), so is usually ranked below* $\left\{\frac{1}{2^{2^k}}\right\}$.

$\left\{\frac{1}{5^{2^k}}\right\}, \{x_k^c\}, \left\{\frac{1}{2^{2^k}}\right\}, \left\{\frac{1}{1.3^{2^k}}\right\}$ *from Figure* 1.4(b) *have ranking* $\left\{\frac{1}{5^{2^k}}\right\}$, $\left\{\frac{1}{2^{2^k}}\right\}$, $\left\{\frac{1}{1.3^{2^k}}\right\}$, $\{x_k^c\}$ *(C-orders* 2, 2, 2, *R-order* 2*). Though nonmonotone,* $\{x_k^c\}$ *has exact R-order* 2 *and is (at least)* $(\dot{e}_k < \mathring{e}_k)$ *faster than the C-quadratic* $\left\{\frac{1}{1.3^{2^k}}\right\}$.

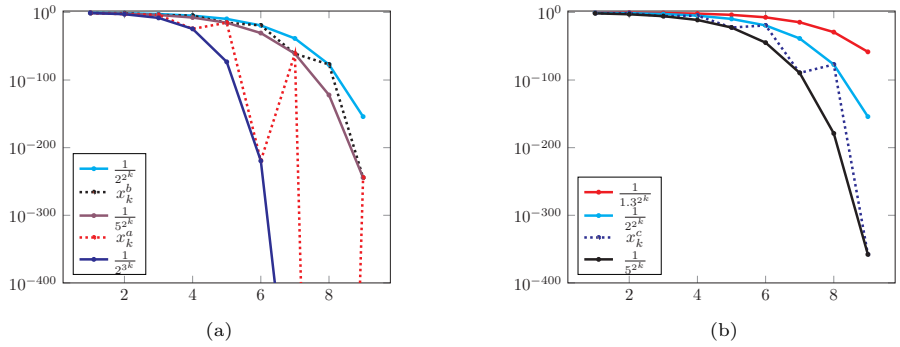


Fig. 1.4 *Ranking sequences visually by* $(\dot{e}_k \le \mathring{e}_k)$.

Quiz 1.8. Given $\{\dot{x}_k\}$, $\{\mathring{x}_k\}$ with errors shown in Figure 1.5, which one is faster?

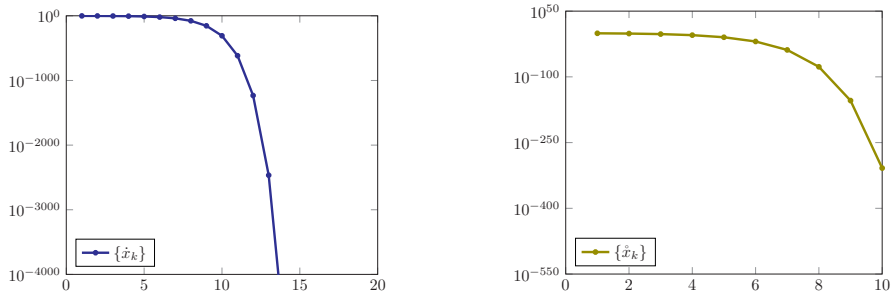


Fig. 1.5 *Quiz* 1.8 *(the answers to quizzes are given at the end of the paper).*

The study of errors using Comparison A seems[6] perfect in theory, but has a notable disadvantage: it requires two sequences. The alternative is to use a "scale" for measuring any single sequence—*the convergence orders*. The (classical) C-orders, though they demand some strong conditions, appear in the convergence of the most often encountered iterative methods for nonlinear equations.[7]

Still, some problems appear, for example, when one wants to check the usual predictions such as "in quadratic convergence, the error is squared at each step." Indeed, as a sequence is an infinite set, its order refers to the asymptotic range (all sufficiently large indices), where statements such as

$$\text{"if } \varepsilon > 0 \text{ is suff. small, } \exists k_0 \ge 0 \text{ s.t. [some property] holds } \forall k \ge k_0\text{"}$$

come to life. Clearly, the first k_0 terms ($k_0 = 2, 10$, or perhaps 10^6) are not necessarily connected to the order: a finite number of terms does not contribute to the order of an abstract sequence, while, most painful, the high order of an iterative method may not be reflected in its first steps if the initial approximation is not good enough.

[6] Of course, one cannot always compare all sequences by $(\dot{e}_k \le \mathring{e}_k)$: consider $\{x_k^a\}$ and $\{\frac{1}{6^{2^k}}\}$, for example.

[7] A notable exception: the higher orders of the secant method are Q- but not necessarily C-orders.

Therefore, such predictions, instead of being applicable to the very first term, are usually applied only starting from *the first terms from the asymptotic range*.

When the stringent conditions of C-orders are not fulfilled, the Q-orders are considered instead, but unfortunately they can be in disagreement with the term-by-term comparison of the errors of two sequences; indeed, we show that convergence even with no Q- (or C)-order may in fact hide a fast overall speed, only some of the consecutive errors do not decrease fast enough. The R-orders are instead known for their limited usefulness, requiring the weakest conditions and allowing nonmonotone errors.

Consider now a further problem: given some iterations for which the calculation of the order turns out to be difficult or even impossible, can the computer be used to approximate it? A positive answer is given by the computational convergence orders.

Despite the fact that convergence orders have a long history, it is only recently that some final connections were made and a comprehensive picture was formed of the C-, Q-, and R-orders, together with their computational variants [20].

The understanding of these notions is important from both the theoretical and the practical viewpoints, the comments of Tapia, Dennis, and Schäfermeyer [93] being most relevant:

> The distinction between Q- and R-convergence is quite meaningful and useful and is essentially sacred to workers in the area of computational optimization. However, for reasons not well understood, computational scientists who are not computational optimizers seem to be at best only tangentially aware of the distinction.

We devote section 2 to error-based analysis of the problem of measuring and comparing convergence speeds of abstract real sequences, using either the convergence orders we review, or the term-by-term comparison of errors.

In section 3 we deal with the computational versions of the convergence orders (based either on the corrections[8] $x_{k+1} - x_k$ or on the nonlinear residuals $f(x_k)$) and show that in $\mathbb{R}$ they are equivalent to the error-based ones.

In section 4 we deal with three main iterative methods (Newton, secant, and successive approximations), presenting results on their attainable convergence orders, as well as other connected aspects (asymptotic constants, estimation of the attraction balls, floating point arithmetic issues, brief history, etc.).

2. C-, Q-, and R-Orders $p_0 > 1$ vs. Asymptotic Rates. Even though the roots of iterative methods trace back to the Babylonians and Egyptians (ca. 1800 B.C.) [3], the first comment on convergence orders seems to have been made by Newton (ca. 1669) on the doubling of digits in quadratic convergence (quoted by Ypma in [101]): "But if I desire to continue working merely to twice as many figures, less one, . . . ," and then in 1675: "That is, the same Division, by wch you could finde the 6th decimal figure, if prosecuted, will give you all to the 11th decimal figure."

In the Journal Book of the Royal Society, it is recorded "17 December 1690: Mr Ralphson's Book was this day produced by E Halley, wherein he gives a Notable Improvemt of ye method [. . .], which doubles the known figures of the Root known by each Operation. . . ."

Halley noticed the tripling of digits in the method he introduced.

In his *Tracts*, Hutton showed that one of his schemes is of third order [3]. In 1818, Fourier [38] also noted that the iterates double the exact figures at each Newton step.

[8] Each x_{k+1} can be seen as being obtained from x_k by adding the correction.

In 1870, Schröder [87] implicitly defined the high C-orders for some iterative methods by considering conditions on the nonlinear mapping.

Three types of convergence orders are used at present: the classical C-order (notation adopted from [4]), the Q-order (which contains the C-order as a particular case), and the R-order. Outstanding contributions to their study were successively made by Ortega and Rheinboldt (1970) in their fundamental book [67], Potra and Pták (1984) [79], Potra (1989) [76], and finally by Beyer, Ebanks, and Qualls (1990), in the less known but essential paper [4]. In [20] we connected and completed these results (in $\mathbb{R}^N$).

We analyze here only the high orders, noting that the various equivalent definitions of the Q-orders lead to intricate implications for linear convergence [4].

2.1. C-Order $p_0 > 1$. The definition of C- and Q-orders is obtained by considering for $p \geq 1$ the *quotient convergence factors* [67, sect. 9.1]

$$Q_p(k) := \frac{e_{k+1}}{(e_k)^p} = \frac{|x^* - x_{k+1}|}{|x^* - x_k|^p}, \quad k = 0, 1, \ldots .3$$

It is assumed that $x_k \neq x^*$, $k \geq 0$. If, though, $x_{k_0} = x^*$ in an iterative method, then (hopefully) we get $x_k = x^*$, $k \geq k_0$ (this holds for the three methods in section 4).

Let us briefly list the four slowest types of C-order:

- no C-order, when $\nexists Q_1 := \lim_{k\to\infty} Q_1(k)$ (e.g., $x_k = \frac{1}{\sqrt{k}}$, k odd; $x_k = \frac{2}{\sqrt{k}}$, k even);
- C-sublinear, if $Q_1 = 1$ (e.g., $\{\frac{1}{k}\}$);
- C-linear, if $0 < Q_1 < 1$ (e.g., $\{\frac{1}{2^k}\}$);
- C-superlinear (defined later, usually called (strict) Q-superlinear);

notice that $Q_1 > 1$ cannot hold, and that $\{x_k^a\}, \{x_k^c\}$, with no C-order, in fact are fast.

REMARK 2.1. (a) *(see* [65, p. 620]*) C-linear order implies strict monotone errors:*

$$e_{k+1} < e_k, \quad k \geq k_0. \tag{2.1}$$

(b) *C-sublinear $\not\Rightarrow$ monotone errors (e.g., $x_k = \frac{1}{k-2}$, k even, $x_k = \frac{1}{k}$, k odd, $k \geq 3$).*

(c) *$\{x_k^a\}, \{x_k^c\}$, both nonmonotone, have no C-order.*

EXERCISE 2.2. *If $\{\mathring{x}_k\}$ has C-sublinear and $\{\dot{x}_k\}$ C-linear order, show that $\{\dot{x}_k\}$ is $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha}))\ \forall \alpha > 1]$ faster than $\{\mathring{x}_k\}$.*

The following definition of high orders is well known; $p_0 > 1$ is implicitly assumed throughout this paper even if not explicitly mentioned.

DEFINITION 2.3. *$\{x_k\}$ has C-order $p_0 > 1$ if*

$$Q_{p_0} := \lim_{k\to\infty} Q_{p_0}(k) \in (0, +\infty). \tag{C}$$

REMARK 2.4. *It is important to stress that Q_{p_0} above cannot be zero or infinite; these two cases will arise when analyzing the Q-orders.*

By denoting (see [67, sect. 9.1], [88, p. 83])

$$\underline{Q}_p = \liminf_{k\to\infty} Q_p(k), \quad \bar{Q}_p = \limsup_{k\to\infty} Q_p(k), \quad p \geq 1, \tag{2.2}$$

condition (C) is equivalent either to relation

$$0 < \underline{Q}_{p_0} = Q_{p_0} = \bar{Q}_{p_0} < \infty \tag{C_Q}$$

or to requiring that $\forall \varepsilon > 0$, $\exists k_0 \geq 0$ such that

$$(C_\varepsilon) \qquad (Q_{p_0} - \varepsilon)e_k^{p_0} \leq e_{k+1} \leq (Q_{p_0} + \varepsilon)e_k^{p_0} \qquad \forall k \geq k_0.$$

REMARK 2.5. $(C) \Rightarrow (2.1)$ *(by* (C_ε)*, or because* (C) *is stronger than C-linear).*

EXAMPLE 2.6. (a) $x_k = \theta^{p_0^k}$ *for some* $\theta \in (0,1)$, $p_0 > 1$, *is the standard example for C-order* p_0 *(*$Q_{p_0} = 1$*). The usual instances are the C-quadratic* $\{2^{-2^k}\}$, $\{3^{-2^k}\}$ *(*$\theta = \frac{1}{2}$, *resp.,* $\theta = \frac{1}{3}$, $p_0 = 2$*), the C-cubic* $\{2^{-3^k}\}$ *(*$\theta = \frac{1}{2}$, $p_0 = 3$*), etc.*

(b) $\{c \cdot 2^{-2^k}\}$ *(*$c \in \mathbb{R}$ *given,* $c \neq 0$*) also has C-quadratic convergence.*

(c) *The C-quadratic* $\{(-1)^k \cdot 2^{-2^k}\}$ *is nonmonotone, but with monotone* $\{e_k\}$.

(d) $x_k^d = \begin{cases} c(\frac{1}{3})^{2^k}, & k \text{ odd}, \\ \frac{1}{c}(\frac{1}{3})^{2^k}, & k \text{ even} \end{cases}$, $c > 1$, *does not have C-order 2:* $\underline{Q}_2 = \frac{1}{c^3}$, $\bar{Q}_2 = c^3$.

Quiz 2.7 ([42, Ex. 5.6 and p. 245]). What is the rate of $\{x_k\}$ if $\{e_k\}$ is
(a) $10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, \ldots$;
(b) $10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, \ldots$;
(c) $10^{-2}, 10^{-3}, 10^{-5}, 10^{-8}, 10^{-13}, 10^{-21}, \ldots$;
(d) $10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}, \ldots$?

REMARK 2.8 (see, e.g., [4]). *If it exists, the C-order* $p_0 > 1$ *of* $\{x_k\}$ *is unique. Indeed, if raising the power of the denominator in* $Q_{p_0}(k)$*, the quotient tends to infinity, since* $Q_{p_0+\varepsilon}(k) = Q_{p_0}(k)\frac{1}{(e_k)^\varepsilon}$ $\forall \varepsilon > 0$. *Similarly, if lowering the power, it tends to zero.*

EXAMPLE 2.9. *Let* $\{10^{-5} \cdot 2^{-2^k}\}$, $\{2^{-2^k/k}\}$, *and* $\{x_k^d\}$ *(*$c = 10$*); in Figure* 2.1 *we plot* $\{Q_{1.8}(k)\}$, $\{Q_2(k)\}$, *and* $\{Q_{2.2}(k)\}$ *for each of them.*

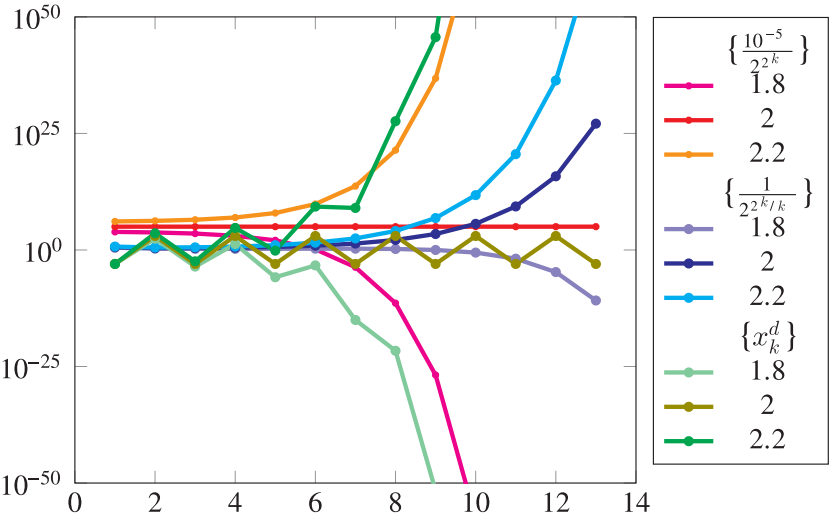


Fig. 2.1 $Q_p(k)$, $k = \overline{1,13}$ *(*$p = 1.8,\ 2,\ 2.2$*).*

The limiting values of $\underline{Q}_p$, $\bar{Q}_p$ *from* (2.2) *may be regarded as "functions" of* $p \geq 1$. *Their graphical illustration leads to the so-called Q-convergence profile of* $\{x_k\}$ [4] *(or Q-profile here, for short).*

For all C-orders $p_0 > 1$, *it holds that* $\bar{Q}_p = \underline{Q}_p = Q_p$ $\forall p \geq 1$, *and* Q_p *is a "function" given by* $Q_p = 0$, $p \in [1, p_0)$, $Q_{p_0} \in (0, +\infty)$, *and* $Q_p = +\infty$, $p > p_0$.

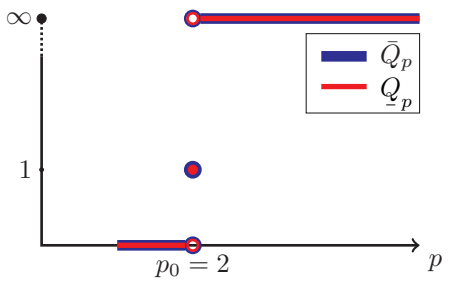


Fig. 2.2 *Q-convergence profile for* $\{\frac{1}{2^{2^k}}\}$*: the limit points of* $Q_p(k)$ *as "functions" of* p*. The vertical axis is not to scale.*

In Figure 2.2 *we plot the Q-profile of* $\{2^{-2^k}\}$.

EXERCISE 2.10 ([4]). *Show that* $x_k \to x^*$ *cannot have C-order* $p_0 < 1$.

EXERCISE 2.11. (a) *Show that if* $\{x_k\}$ *has C-order* $p_0 > 1$*, then so has* $\{\frac{e_{k+1}}{e_k}\}$ *and*

$$Q_{p_0}\left\{\frac{e_{k+1}}{e_k}\right\} = 1, \tag{2.3}$$

regardless of $Q_{p_0}\{e_k\} \neq 0$ [20, Rem. 3.9]*. Does the converse hold? (Hint: take* $\{\frac{1}{k}\}$*.)*

(b) [20, Rem. 3.9] *Find a similar statement for* $\{e_{k+1}e_k\}$.

(c) *(Kelley* [50, Ex. 5.7.15]*) If* $\{x_k\}$ *has C-order* $p_0 > 1$*, show that* $\lg e_k$ *is concave, i.e.,* $\lg e_{k+1} - \lg e_k$ *is decreasing.*

Not only convergence with any high order exists, but even the infinite C-order, defined either by $Q_p = 0\ \forall p > 1$ (take, e.g., $\{2^{-2^{2^k}}\}$ [67, E 9.1-3(g)] or $\{2^{-2^{k^2}}\}$) or by convention, when the convergence is in a finite number of steps.

"The higher the order, the better" is a well-known cliché, which we will express in subsection 2.5 in terms of the big Ohs.

2.2. Q-Order $p_0 > 1$. We propose the following definitions of Q-order convergence:

- no Q-order if $\bar{Q}_1 = \infty$ (e.g., $\{x_k^a\}, \{x_k^c\}$);
- Q-sublinear if $1 \leq \bar{Q}_1 < +\infty$ (e.g., $x_k = \frac{2}{k}$, k odd, $x_k = \frac{1}{k}$, k even);
- exact Q-sublinear if $0 < \underline{Q}_1 \leq 1 \leq \bar{Q}_1 < +\infty$;
- at least Q-linear if $\bar{Q}_1 < 1$;
- Q-linear if $0 < \bar{Q}_1 < 1$;
- exact Q-linear if $0 < \underline{Q}_1 \leq \bar{Q}_1 < 1$.

"with unbounded nonmonotone errors" instead of "no Q-order" (see Remark 2.12.c). One can find sequences with \bar{Q}_1=\infty, see [E. Catinas, The superlinear strict order can be faster than the infinite order, Numer. Algor, 2023, https://doi.org/10.1007/s11075-023-01604-y]

REMARK 2.12. (a) *Obviously, when* $x_k \to x^*$*, then* $\underline{Q}_1 \leq 1$ *and* $\underline{Q}_1 \leq \bar{Q}_1$.

(b) $\bar{Q}_1 \in [0, +\infty]$ *always exist, while* Q_1 *may not (i.e., when* $\underline{Q}_1 < \bar{Q}_1$*).*

(c) $\bar{Q}_1 < 1 \Rightarrow$ *monotone, while* $\bar{Q}_1 > 1 \Rightarrow$ *nonmonotone errors;* $\bar{Q}_1 = \infty$ *means* unbounded nonmonotone errors*. Unlike* $0 < Q_1 \leq 1$, $0 < \bar{Q}_1 \leq +\infty$ *alone does not necessarily imply slow speed (e.g.,* $\{x_k^a\}, \{x_k^c\}$ *with unbounded nonmonotone errors).*

The convergence may be fast when $\underline{Q}_1 = 0$*, even if* $\bar{Q}_1 = \infty$ *(e.g.,* $\{x_k^a\}, \{x_k^c\}$*), but is no longer fast when* $0 < \bar{Q}_1$.

Strict Q-superlinear is an intermediate order between linear and $p_0 > 1$.

DEFINITION 2.13 ([67, p. 285]). *$\{x_k\}$ has Q-superlinear order*[9] *if $\bar{Q}_1 = 0 (= Q_1)$:*

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = 0, \qquad (\Leftrightarrow \exists c_k \to 0 \ s.t. \ e_{k+1} = c_k e_k).$$

Strict Q-superlinear order holds when, moreover, $\bar{Q}_p = +\infty \ \forall p > 1$.

REMARK 2.14 (cf. [75, (37b)]). *Q-superlinear order holds if $\exists \theta_k > 0$ and $c > 0$ such that $\theta_k \to 0$ and $e_k = c\prod_{i=1}^{k} \theta_i$ (c may be taken to be 1).*

EXAMPLE 2.15 (strict Q-superlinear). (a) *Let* $\left\{\frac{1}{k^k}\right\}$ [6, p. 22], $\left\{\frac{1}{10^{k^2}}\right\}$ [11], [4], $\left\{\frac{1}{k!}\right\}$ [47], $\left\{\frac{1}{c^{k^2}}\right\}$, $c > 1$ [67, E 9.2-1(j)].

(b) [67, E 10.1-4], [79, p. 94] *Given $0 < c < \frac{1}{\mathrm{e}}$, let $x_{k+1} = -\frac{x_k}{\ln x_k}$, $k \geq 0$, $x_0 = c$.*

EXERCISE 2.16. *$\{x_k^b\}$ has Q-superlinear order, but not strict Q-superlinear order.*

The following formulation has been commonly used for half a century, since the 1970 book of Ortega and Rheinboldt: "$\{x_k\}$ converges with Q-order at least $p_0 > 1$" [67], [79], [76], [85]. Here we deal, as in [4] and [20], with a more restrictive notion, with the classic notion corresponding to *the lower Q-order q_l* from Definition 2.30.

DEFINITION 2.17 (see [20]; cf. [76], [4]). *$\{x_k\}$ has Q-order $p_0 > 1$ if any of the following equivalent conditions hold:*

$$(Q) \qquad \lim_{k\to\infty} Q_p(k) = \begin{cases} 0, & p \in [1, p_0), \\ +\infty, & p \in (p_0, +\infty); \end{cases}$$

$$(Q_L) \qquad \lim_{k\to\infty} Q_L(k) = p_0, \qquad Q_L(k) := \frac{\ln e_{k+1}}{\ln e_k};$$

$$(Q_\Lambda) \qquad \lim_{k\to\infty} Q_\Lambda(k) = p_0, \qquad Q_\Lambda(k) := \frac{\ln \frac{e_{k+2}}{e_{k+1}}}{\ln \frac{e_{k+1}}{e_k}};$$

or $\forall \varepsilon > 0$, $\exists A, B > 0$ such that

$$(Q_\varepsilon) \qquad A e_k^{p_0+\varepsilon} \leq e_{k+1} \leq B e_k^{p_0-\varepsilon} \quad \forall k \geq k_0.$$

REMARK 2.18. (a) *If the Q-order p_0 exists, it is unique (recall Remark 2.8).*

(b) *If $\{x_k\}$ has Q-order $p_0 > 1$, then the right-hand side inequalities in (Q_ε) imply that the errors are strictly monotone ($k \geq k_0$); see Remarks 2.1 and 2.5.*

In [20] we proved the equivalence of the four conditions above for $\{x_k\} \subset \mathbb{R}^N$, connecting and completing the independent, fundamental results of Potra [76] and Beyer, Ebanks, and Qualls [4]. This proof does not simplify for $\{x_k\} \subset \mathbb{R}$, but a natural connection between these conditions is found, e.g., by taking logarithms in (C_ε) to obtain (Q_L).[10] Also, (Q_L) and (Q_Λ) can be connected by Exercise 2.11 and as follows.

EXERCISE 2.19. *If $\{x_k\}$ has strict monotone errors for $k \geq k_0$, prove that $(Q_\Lambda) \Rightarrow (Q_L)$ (hint: use the Stolz–Cesàro Theorem).*

REMARK 2.20. (a) *One can always set $k_0 = 0$ in (Q_ε) (this is actually used in the definitions of Potra from [76]), by a proper choice of smaller A and larger B.*

[9] Or *at least Q-superlinear*, or even *superlinear*, as R-superlinear is seldom encountered.

[10] p_0 in (Q_L) roughly means that the number of figures is multiplied by p_0 at each step ($k \geq k_0$); see, e.g., [48] and [79, p. 91].

(b) *In* [20] *we have denoted by* (Q_ε) *a condition which is trivially equivalent to* (Q). *Here we use* (Q_ε) *instead of* $(Q_{I,\varepsilon})$ *from* [20] *("I" stood for "inequality" in that paper).*

(c) (Q) *implies that the Q-profile of* $\{x_k\}$ *has a single jump at* $p_0 > 1$, *but that the limit* Q_{p_0} *is not required to exist, so*

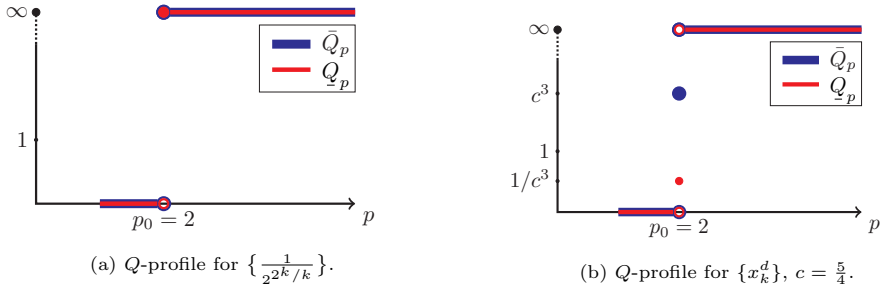
$$0 \le \underline{Q}_{p_0} \le \bar{Q}_{p_0} \le \infty, \tag{2.4}$$

and so six cases result for the possible values of $\underline{Q}_{p_0}$ *and* $\bar{Q}_{p_0}$ *(0, finite* > 0, $+\infty$*).*

(d) *Relations* $\underline{Q}_{p_0} = 0$ *or* $\bar{Q}_{p_0} = +\infty$ *might seem rather abstract, but they do actually occur (even simultaneously, as we shall see for the higher orders of the secant method; see also* [99, p. 252 and Chap. 7], *where* $\underline{Q}_2 = \bar{Q}_2 = +\infty$, *for a problem in* $\mathbb{R}^N$*).*

(e) Q*-order* $p_0 > 1$ *implies* $\bar{Q}_1 = 0$, *but the converse is not true.*

Figure 2.3 shows the Q-profiles of $\{2^{-2^k/k}\}$ and $\{x_k^d\}$ $(c = \frac{5}{4} = 1.25)$.



(a) Q-profile for $\{\frac{1}{2^{2^k/k}}\}$. (b) Q-profile for $\{x_k^d\}$, $c = \frac{5}{4}$.

Fig. 2.3 *Q-profiles:* (a) *Q-subquadratic;* (b) *exact Q-quadratic.*

Superlinearity may also be defined in conjunction with higher Q-orders, but this is tricky.

DEFINITION 2.21 (see [47], [20]). *The Q-order* $p_0 > 1$ *is called*

- *Q-superorder* p_0 *if* $\bar{Q}_{p_0} = 0 (= Q_{p_0})$;
- *Q-suborder* p_0 *if* $\underline{Q}_{p_0} = +\infty (= Q_{p_0})$.[11]

The Q-order 2 *with* $\bar{Q}_2 = 0$ *is Q-superquadratic (analogously, Q-supercubic, etc.).*

EXERCISE 2.22. (a) *The Q-superquadratic* $\{\dot{x}_k\} = \{2^{-k2^k}\}$, *the Q- (and C-)quadratic* $\{2^{-2^k}\}$, *and the Q-subquadratic* $\{\mathring{x}_k\} = \{2^{-\frac{2^k}{k}}\}$ *are in* $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha))\ \forall \alpha > 1]$ *decreasing order of speed. Study the Q-order of* $z_k = \begin{cases} \dot{x}_k, & k \text{ odd}, \\ \mathring{x}_k, & k \text{ even}, \end{cases}$ *(2 is erroneous in* [20]*).*

(b) *Although the C-quadratic* $\{2^{-2^k}\}$ *is* $(\dot{e}_k = o(\mathring{e}_k))$ *faster than* $\{k2^{-2^k}\}$, *the latter, considered in* [47], *is actually Q-superquadratic; similarly,* $\{2^{-2^k}/k\}$ *is Q-subquadratic. The terminology "Q-sub/superquadratic" appears here to be in disagreement with the real speed.*

(c) *Determine the Q-orders from Quiz* 2.7 *by using* (Q_L).

[11] In [20] this was defined by $\bar{Q}_{p_0} = +\infty$, which we believe is not strong enough.

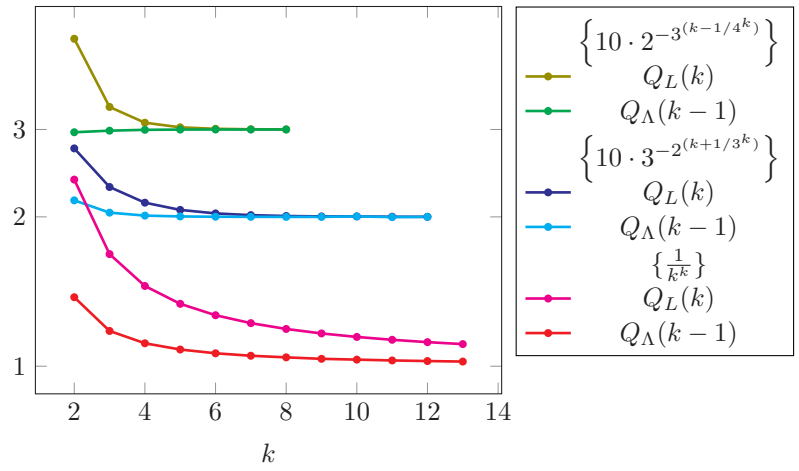


Fig. 2.4 *$Q_L(k)$ and $Q_\Lambda(k-1)$ computed for three sequences.*

EXAMPLE 2.23. *Let $\{10 \cdot 2^{-3^{(k-1/4^k)}}\}$ (C-cubic), $\{10 \cdot 3^{-2^{(k+1/3^k)}}\}$ (C-quadratic), and $\{\frac{1}{k^k}\}$ (Q-superlinear); in Figure 2.4 we see that $Q_L(k), Q_\Lambda(k-1)$ tend to 3, 2, and 1, respectively.*

REMARK 2.24. *Plotting/comparing $Q_L(k)$ and $Q_\Lambda(k-1)$ avoids conclusions based on $Q_\Lambda(k)$ using information ahead of that from $Q_L(k)$ (i.e., x_{k+2} vs. x_{k+1}).*

The calculation of the order may be made easier by using logarithms.

EXERCISE 2.25. *Let $x_k \to x^* \in \mathbb{R}$.*

(a) *If, for some given $q \geq 1$ and $A, B > 0$, one has*

$$Ae_k e_{k-1}^q \leq e_{k+1} \leq Be_k e_{k-1}^q, \quad k \geq k_0, \tag{2.5}$$

show that $\{x_k\}$ has Q-order

$$\lambda_q := \frac{1+\sqrt{1+4q}}{2}.$$

Thus, $\lambda_1 \approx 1.618$ (the golden ratio), $\lambda_2 = 2$, $\lambda_3 = \frac{1+\sqrt{13}}{2} \approx 2.3$, etc. Inequality (2.5) appears in the analysis of the secant method (see Theorems 4.18 and 4.22), and it does not necessarily attract C- or even exact Q-orders, even if $\exists \lim_{k\to\infty} \frac{e_{k+1}}{e_k e_{k-1}^q} = c \in (0, \infty)$.

(b) *Calculate the Q-order of $\{x_k\}$ if it satisfies (see* [94], [78]*)*

$$Ae_k^2 e_{k-1} \leq e_{k+1} \leq Be_k^2 e_{k-1}.$$

(c) *If $\{x_k\}$ verifies the general relation (see, e.g.,* [95, Chap. 3], [7], [76], [44], [72]*)*

$$Ae_k^{\alpha_0} \cdots e_{k-q}^{\alpha_q} \leq e_{k+1} \leq Be_k^{\alpha_0} \cdots e_{k-q}^{\alpha_q},$$

determine the condition verified by the Q-order (see [76] *for the exact Q-order).*

A special case holds when $\varepsilon = 0$ in (Q_ε) (see also [9], [88], [79], [11], [76]).

DEFINITION 2.26 (see, e.g., [8]). *$\{x_k\}$ has exact Q-order p_0 if $\exists A, B, k_0 \geq 0$, s.t.*

$$A \cdot e_k^{p_0} \leq e_{k+1} \leq B \cdot e_k^{p_0} \quad \forall k \geq k_0. \tag{$\bar{Q}$}$$

This case can be characterized in terms of the asymptotic constants $\underline{Q}_{p_0}, \bar{Q}_{p_0}$.

PROPOSITION 2.27. *$\{x_k\}$ has exact Q-order p_0 iff $\underline{Q}_{p_0}, \bar{Q}_{p_0}$ are finite and nonzero,*

$$(\bar{\underline{Q}}_Q) \qquad 0 < \underline{Q}_{p_0} \le \bar{Q}_{p_0} < \infty,$$

in which case the constants A, B in $(\bar{\underline{Q}})$ are bounded by

$$A \le \underline{Q}_{p_0} \le \bar{Q}_{p_0} \le B,$$

and these bounds are attained in some special restrictive circumstances.

EXAMPLE 2.28. (a) *$\{x_k^d\}$ in Example 2.6 has exact Q-order 2: $\underline{Q}_2 = \frac{1}{c^3} \neq \bar{Q}_2 = c^3$.*

(b) *$x_k = 2^{-2^k}$, k odd, $x_k = 3 \cdot 2^{-2^{k-\frac{1}{3^k}}}$, k even, has $Q(2k) < \underline{Q}_2 = \frac{1}{9}$, i.e., $A < \underline{Q}_2$.*

REMARK 2.29. *The sequences $\{\mathring{e}_k\}$ and $\{\underset{\circ}{e}_k\}$ are asymptotically similar when $\mathring{e}_k = \mathcal{O}(\underset{\circ}{e}_k)$ and $\underset{\circ}{e}_k = \mathcal{O}(\mathring{e}_k)$, as $k \to \infty$, denoted by $\underset{\circ}{e}_k = \Theta(\mathring{e}_k)$ (see, e.g., [53], [25, p. 50]).*

The exact Q-order may also be expressed as $e_{k+1} = \mathcal{O}(e_k^{p_0})$ and $e_k^{p_0} = \mathcal{O}(e_{k+1})$, as $k \to \infty$ (see [8], [9, p. 2], and also [71]), i.e., $e_{k+1} = \Theta(e_k^{p_0})$.

The classical Q-order of Ortega and Rheinboldt is the q_l below.

DEFINITION 2.30 ([4], [67], [88], [76]). *The lower/upper Q-orders of $\{x_k\}$ are*

$$(2.6) \quad q_l = \begin{cases} \infty, & \text{if } \bar{Q}_p = 0 \ \forall p \ge 1, \\ \inf\{p \in [1,\infty) \,:\, \bar{Q}_p = +\infty\}, \end{cases} \quad \text{resp.,} \quad q_u = \sup\big\{p \in [1,\infty) : \underline{Q}_p = 0\big\}.$$

When $\{x_k\}$ has Q-order $p_0 > 1$, the lower and upper orders coincide, $q_l = q_u = p_0$; otherwise, $q_l < q_u$ [4]; we will also see this in Theorem 2.48 (relation (2.13)).

Next we analyze $\{x_k^b\}$ from Example 1.7, which has no Q-order, despite it being $(\dot{e}_k \le \mathring{e}_k)$ faster than $\{\frac{1}{2^{2^k}}\}$. We keep in mind that the Q-order does not measure the overall speed, but just compares the consecutive errors.

EXAMPLE 2.31 ([20]). *Let $x_k^b = \begin{cases} 2^{-2^k}, & k \text{ even}, \\ 3^{-2^k}, & k \text{ odd}, \end{cases}$ with the Q-profile in Figure 2.5(a).*

$\{x_k^b\}$ does not have a Q-order, as $q_l = \log_3 4 = 1.2\ldots < q_u = 4\log_4 3 = 3.1, \ldots$, but still it has (exact) R-order 2 (see Definitions 2.35 and 2.38). We note the usual statement in this case that "$\{x_k^b\}$ converges with Q-order at least $q_l = 1.2\ldots$ and with R-order at least 2," that no longer holds in the setting of this paper.

EXERCISE 2.32. *Show that $\{x_k^c\}$ has $\underline{Q}_1 = 0, \bar{Q}_1 = +\infty$ and determine q_l, q_u (here we find $q_l < 1$, and the Q-profile can be extended, as in [4], for $p < 1$).*

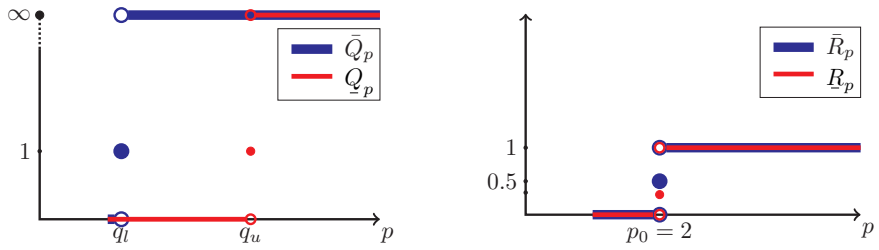
Determine the Q-profiles of $\{x_k^a\}$ and $\{x_k^c\}$.

The inequalities implied by the lower/upper Q-orders are as follows.

REMARK 2.33. *If $1 < q_l < \infty$, then $\forall \varepsilon > 0$, with $1 < q_l - \varepsilon$, $\exists B > 0, k_0 \ge 0$, s.t.* [80]

$$(2.7) \qquad e_{k+1} \le B \cdot e_k^{q_l - \varepsilon} \quad \forall k \ge k_0$$

(and relation (2.6) says that q_l is the largest value satisfying the above inequalities).



(a) Q-profile of $\{x_k^b\}$ (no Q-order). (b) R-profile of $\{x_k^b\}$ (exact R-order 2).

Fig. 2.5 (a) *A Q-profile;* (b) *an R-profile.*

When $1 < q_u < \infty$, *then* $\forall \varepsilon > 0$, $\exists A > 0, k_1 \geq 0$, *such that*

$$A \cdot e_k^{q_u+\varepsilon} \leq e_{k+1} \quad \forall k \geq k_1. \tag{2.8}$$

One can say that the lower Q-order is small when the maximum error reduction w.r.t. the previous step is small (i.e., small exponent $q_l + \varepsilon$ *for* e_k*). Analogously, the upper order is large when the minimum error reduction per step is small.*

The exact lower/upper Q-orders q_l, *respectively,* q_u *are defined if* $\varepsilon = 0$ *in* (2.7)–(2.8)*:*

$$A \cdot e_k^{q_u} \leq e_{k+1} \leq B \cdot e_k^{q_l} \quad \forall k \geq k_2.$$

The role of q_u will be seen in Example 2.46.

The lower and upper Q-orders verify the relations (see [76], [20])

$$q_l = \underline{Q}_L := \liminf_{k\to\infty} \frac{\ln e_{k+1}}{\ln e_k}, \tag{2.9}$$

$$q_u = \bar{Q}_L := \limsup_{k\to\infty} \frac{\ln e_{k+1}}{\ln e_k}, \tag{2.10}$$

which will be completed in formulae (3.5), respectively, (3.6) by some computational variants.

2.3. R-Order $p_0 > 1$. The root factors consider some averaged quantities instead of relating the consecutive terms to one another; they are defined for $p = 1$ by

$$R_1(k) = e_k^{\frac{1}{k}}, \quad k \geq 1.$$

We propose the following terminology (see also [67], [63], [75], [79], [100]):

- R-sublinear/with no R-order if $\bar{R}_1 = 1$ (e.g., $\{\frac{1}{k}\}, \{\frac{1}{\sqrt{k}}\}$; see [75, Rem. 1.2.40]);
- R-linear if $0 < \bar{R}_1 < 1$;
- at least R-linear if $\bar{R}_1 < 1$ (e.g., $x_k = \frac{1}{2^k}$, k odd, $x_k = \frac{1}{4^k}$, k even [75, Rem. 1.2.40], or $x_k = \frac{1}{2^k}$, k even, $x_k = \frac{1}{2^{2^k}}$, k odd);
- exact R-linear if $0 < \underline{R}_1 \leq \bar{R}_1 < 1$;
- (at least) R-superlinear if $\bar{R}_1 = 0 (= R_1)$ (e.g., $x_k = \frac{1}{k^k}$, k even, $x_k = x_{k-1}$, k odd).

The R-orders are alternatively defined by requiring the errors to be bounded by sequences converging to zero with corresponding Q-orders [32, p. 21], [51, Def. 2.1.3], [43, p. 218]:

- at least R-linear if $\exists\theta \in (0,1)$ and $c > 0$ s.t. $e_k \le c \cdot \theta^k$ $(k \ge 0)$ [75, (37a)];
- R-superlinear if $\exists\theta_k \to 0$ and $c > 0$ s.t. $e_k \le c \prod_{i=1}^{k} \theta_i$ [75, (37b)].

REMARK 2.34. (a) *At least Q-linear* $\Rightarrow$ *at least R-linear* [67, E 9.3-5], [75, Thm. 1.2.41], [77], [100], [43, p. 213] *(using, e.g.,* (2.14)*); the converse is false (see above or* [100]*).*

(b) *Q-superlinear* $\Rightarrow$ *R-superlinear, with the converse again being false (see, e.g., the above example).*

When $p > 1$, the root factors are defined by (see [67], [79], [76], [85])

$$R_p(k) = e_k^{\frac{1}{p^k}}, \quad k \ge 0.$$

DEFINITION 2.35 (see [20]; cf. [76], [7]). $\{x_k\}$ *has R-order* $p_0 > 1$ *if any of the equivalent relations hold:*[12]

$$(R) \qquad \lim_{k\to\infty} R_p(k) = \begin{cases} 0, & p \in [1, p_0), \\ 1, & p \in (p_0, +\infty); \end{cases}$$

$$(R_L) \qquad \lim_{k\to\infty} R_L(k) = p_0, \quad R_L := \left|\ln e_k\right|^{\frac{1}{k}};$$

$$(R_\Lambda) \qquad \lim_{k\to\infty} R_\Lambda(k) = p_0, \quad R_\Lambda := \left|\ln \frac{e_{k+1}}{e_k}\right|^{\frac{1}{k}};$$

or $\forall\varepsilon > 0$, $\exists A, B > 0$, $0 < \eta, \theta < 1$, *and* $k_0 \ge 0$ *such that*

$$(R_\varepsilon) \qquad A \cdot \eta^{(p_0+\varepsilon)^k} \le e_k \le B \cdot \theta^{(p_0-\varepsilon)^k} \quad \forall k \ge k_0.$$

REMARK 2.36. (a) *If the R-order* p_0 *exists, it is unique (see also* [67, 9.2.3, p. 289]*).*

(b) *We can always set* $k_0 = 0$ *in* (R_ε)*, by choosing smaller* A *and larger* B *suitably (cf.* [76]*, where* $k_0 = 0$ *was used in definitions).*

(c) *We prefer here the notation* (R_ε) *instead of* $(R_{I,\varepsilon})$ *from* [20].

For $p \ge 1$ one defines the asymptotic quantities (see [67, sect. 9.2] and [88, p. 85])

$$\underline{R}_p = \liminf_{k\to\infty} R_p(k), \quad \bar{R}_p = \limsup_{k\to\infty} R_p(k) \le 1. \tag{2.11}$$

An example of an R-profile is shown in Figure 2.5(b) for $\{x_k^b\}$, with (exact) R-order 2.

REMARK 2.37 ([67, p. 290], [20]). *If* $\bar{R}_{p_0} < 1$*, then* $\forall\varepsilon > 0$ *with* $\bar{R}_{p_0} + \varepsilon < 1$, $\exists k_0 \ge 0$ *s.t.*

$$e_k \le \left(\bar{R}_{p_0} + \varepsilon\right)^{p_0^k} \quad \forall k \ge k_0,$$

while if $\underline{R}_{p_0} > 0$*, then* $\forall\varepsilon > 0$ *with* $0 < \underline{R}_{p_0} - \varepsilon$, $\exists k_1 \ge 0$ *s.t.*

$$\left(\underline{R}_{p_0} - \varepsilon\right)^{p_0^k} \le e_k \quad \forall k \ge k_1.$$

The following notion is defined similarly to the exact Q-order.

[12] In order that relation (R_Λ) is properly defined and equivalent to the rest of the following definitions, an additional assumption is required (see [20]): $1 < q_l \le q_u < +\infty$.

DEFINITION 2.38 ([76]). *$\{x_k\}$ has exact R-order p_0 if $\exists A, B > 0, 0 < \eta, \theta < 1$, s.t.*

$$(\underline{\bar{R}}) \qquad A \cdot \eta^{p_0^k} \le e_k \le B \cdot \theta^{p_0^k} \quad \forall k \ge k_0.$$

EXAMPLE 2.39. *$\{x_k^b\}$, $\{x_k^c\}$ have exact R-orders 2: $\underline{R}_2 = \frac{1}{3}$, $\bar{R}_2 = \frac{1}{2}$, and $\underline{R}_2 = \frac{1}{5}$, $\bar{R}_2 = \frac{1}{2}$, respectively.*

EXERCISE 2.40. *The Q-sub-/superquadratic $\{2^{-2^k}/k\}$, $\{k2^{-2^k}\}$ have $R_2 = \frac{1}{2}$, but not exact R-order 2. The Q-sub-/superquadratic $\{2^{-2^k/k}\}$, $\{2^{-k2^k}\}$ have $R_2 = 1$, respectively, $R_2 = 0$, i.e., they are R-sub-/superquadratic, again without having exact R-order 2.*

REMARK 2.41. *Perhaps a proper definition of Q-superorder p_0 would be Q-order p_0 with $Q_{p_0} = 0 = R_{p_0}$, while Q-suborder p_0 would be Q-order p_0 with $Q_{p_0} = \infty$, $R_{p_0} = 1$.*

We characterize the exact R-order in the following result.

PROPOSITION 2.42. *The exact R-order p_0 of $\{x_k\}$ is equivalently defined by*

$$(\underline{\bar{R}}_R) \qquad 0 < \underline{R}_{p_0} \le \bar{R}_{p_0} < 1,$$

which implies the following bounds for η, θ from $(\underline{\bar{R}})$:

$$\eta \le \underline{R}_{p_0} \le \bar{R}_{p_0} \le \theta;$$

these bounds are attained in some special restrictive circumstances.

REMARK 2.43. *A particular instance from $(\underline{\bar{R}})$, i.e., $e_k \le B \cdot \theta^{2^k}$, was considered as a definition for (at least) R-quadratic convergence, and some computational scientists (who we suspect are not computational optimizers) have misled by simply calling it "quadratic convergence," leading to the confusion noted in* [93].

As the R-orders require weaker conditions than the Q-orders, they may allow nonmonotone errors. This aspect is perhaps not widely known; we have found it pointed out in [65, p. 620, and in the 1st ed.], [51, p. 14], [25, p. 51], and [47].

Clearly, iterative methods with nonmonotone errors are usually not desired.

We can easily find conditions for monotone errors in the case of exact R-order $p_0 > 1$.

THEOREM 2.44 (monotone errors in exact R-orders). *If $\{x_k\}$ obeys $(\underline{\bar{R}})$ and*

$$p_0 > \frac{\ln \eta}{\ln \theta} \left(\ge \frac{\ln \bar{R}_{p_0}}{\ln \underline{R}_{p_0}} \right) \quad \text{or} \quad \left(p_0 = \frac{\ln \eta}{\ln \theta} \text{ and } B < A \right),$$

then it has strict monotone errors ($k \ge k_0$).

The lower and upper R-orders, r_l, respectively, r_u, and further notations from [20] are

$$r_l = \inf\{p \in [1, \infty) : \bar{R}_p = 1\}, \qquad r_u = \sup\{p \in [1, \infty) : \underline{R}_p = 0\},$$

$$\underline{R}_L := \liminf_{k \to \infty} \left| \ln e_k \right|^{\frac{1}{k}}, \quad \bar{R}_L := \limsup_{k \to \infty} \left| \ln e_k \right|^{\frac{1}{k}},$$

$$\underline{R}_\Lambda := \liminf_{k \to \infty} \left| \ln \frac{e_{k+1}}{e_k} \right|^{\frac{1}{k}}, \quad \bar{R}_\Lambda := \limsup_{k \to \infty} \left| \ln \frac{e_{k+1}}{e_k} \right|^{\frac{1}{k}}.$$

REMARK 2.45. *The lower R-order r_l of $\{x_k\}$ was also defined as follows: $\exists\{\mathring{x}_k\}$ with C-order $\mathring{p}_0 = r_l$ s.t. $\{x_k\}$ is $(\dot{e}_k \le \mathring{e}_k)$ faster than $\{\mathring{x}_k\}$* [29], [76], [50, Def. 4.1.3].

We now consider some sequences similar to those in Example 1.7.

EXAMPLE 2.46. (a) *Let $\mathring{x}_k = 2^{-2^k}$ and take $\dot{x}_k = \begin{cases} 2^{-2^k}, & k \text{ even}, \\ 2^{-3^k}, & k \text{ odd} \end{cases}$ (Jay* [47]*).*

Then $\{\dot{x}_k\}$ converges $(\dot{e}_k \le \mathring{e}_k)$ faster than $\{\mathring{x}_k\}$ and though $\{\dot{x}_k\}$ has neither C-order, nor Q-order, nor R-order ($\dot{r}_l = 2$, $\dot{r}_u = 3$), $\{\mathring{x}_k\}$ has C-order 2.

(b) *Extending the above behavior, $\dot{x}_k = \begin{cases} 2^{-4^k}, & k \text{ even}, \\ 2^{-5^k}, & k \text{ odd} \end{cases}$ has no C-, Q-, or R-order, but it converges $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha}))\ \forall \alpha > 1]$ faster than $\mathring{x}_k = 2^{-2^k}$.*

(c) $\dot{x}_k = \begin{cases} 2^{-3^{2^k}}, & k \text{ even}, \\ 2^{-4^{2^k}}, & k \text{ odd} \end{cases}$ *is $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^{\alpha}))\ \forall \alpha > 1]$ faster than $\mathring{x}_k = 2^{-2^{2^k}}$; $\{\dot{x}_k\}$ has no C- or Q-order (but has infinite R-order), while $\{\mathring{x}_k\}$ has infinite C-order.*

Here $\{\dot{x}_k\}$ deserves a closer look, as it is a perfect candidate to support the comments from [93]*: indeed, it attains infinite R-order, but its errors are unbounded nonmonotone. Nocedal and Wright* [65, p. 620] *also noted such possible behavior.*

REMARK 2.47. *Clearly, statements such as those in the above example can appear only for sequences $\{\dot{x}_k\}$ with $\underline{Q}_1 = 0$, as this condition allows large upper orders q_u. When $\underline{Q}_1 > 0$, the corresponding sequence cannot converge quickly.*

2.4. Relations between the C-, Q-, and R-Orders of a Sequence. C-order $p_0 > 1$ of $\{x_k\}$ implies Q-order p_0 and in turn R-order p_0 (see Exercise 2.22 and Example 2.31 for converses). We state this below and, as in [4], we use curly braces for equivalent orders.

THEOREM 2.48 (see [20]; cf. [76], [4]). *Let $x_k \to x^*$ and $p_0 > 1$. Then (see footnote* 12*)*

$$\{C, C_Q, C_\varepsilon\} \underset{\not\Leftarrow}{\Rightarrow} \{Q, Q_L, Q_\Lambda, Q_\varepsilon\} \underset{\not\Leftarrow}{\Rightarrow} \{R, R_L, R_\Lambda, R_\varepsilon\}, \tag{2.12}$$

or, in a more generic fashion (i.e., $\{C\} := \{C, C_Q, C_\varepsilon\}$),

$$\{C\} \underset{\not\Leftarrow}{\Rightarrow} \{Q\} \underset{\not\Leftarrow}{\Rightarrow} \{R\}.$$

Moreover, the following relation holds for the lower and upper orders:

$$q_l = \underline{Q}_L \le \underline{R}_L = \underline{R}_\Lambda = r_l \le r_u = \bar{R}_L = \bar{R}_\Lambda \le \bar{Q}_L = q_u. \tag{2.13}$$

Relation (2.13) is obtained from the well-known inequalities for positive numbers,

$$\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} \le \liminf_{k\to\infty} |a_k|^{\frac{1}{k}} \le \limsup_{k\to\infty} |a_k|^{\frac{1}{k}} \le \limsup_{k\to\infty} \frac{a_{k+1}}{a_k}, \tag{2.14}$$

taking $a_k = \left|\ln e_k\right|$; see [76] and [20].

Any inequality from (2.13) may be strict. Now we see (e.g., in Examples 2.31 and 2.49) that the inner inequality may be an equality (i.e., obtain R-order), while one of the outer inequalities may be strict (i.e., no Q-order).

An $\{x_k\}$ with exact R-order $\tau > 1$ arbitrarily large and $1 < q_l$ (but arbitrarily close) is again suitable for justifying the comments from [93].

EXAMPLE 2.49 ([67, E 9.3-3], [76]). *Given any numbers $1 < s < \tau$, take $0 < \theta < 1$, $\eta = \theta^q$ with $q > 1$ such that $qs > \tau$. Then $x_k = \begin{cases} \theta^{\tau^k}, & k \text{ odd}, \\ \eta^{\tau^k}, & k \text{ even} \end{cases}$ has exact R-order τ, while $q_l = \underline{Q}_L = \frac{\tau}{q}$ and $q_u = \bar{Q}_L = \tau q (> \tau)$, and thus it has no Q-order (in the classical sense from* [67] *$\{x_k\}$ has Q-order at least $\frac{\tau}{q}$).*

2.5. Comparing the Speeds of Two Sequences. We consider only C-orders.

COMPARISON B (higher C-order, faster speed). *If $\{\mathring{x}_k\}, \{\dot{x}_k\}$ have C-orders $1 < \mathring{p}_0 < \dot{p}_0$, then $\{\dot{x}_k\}$ is $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha))\ \forall \alpha > 1]$ faster than $\{\mathring{x}_k\}$.*

For the proof, one may use (C_ε). In [21] we show some simplified proofs.

REMARK 2.50. (a) *Comparison* B *holds regardless of the magnitude of $\dot{p}_0 - \mathring{p}_0$.*

(b) *The C-orders (i.e., sublinear, linear, strict superlinear, $1 < \mathring{p}_0 < \dot{p}_0$, and infinite) form classes in $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha))\ \forall \alpha > 1]$ increasing speed (cf. Exercise 2.2).*

strict superlinear should be removed, as it is not a C-order (its def. requires \bar{Q}_p), and, moreover, [The superlinear strict order can be faster than the infinite order, E. Catinas, Numer. Algor, 2023, https://doi.org/10.1007/s11075-023-01604-y]

As seen in previous examples, comparing by $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha))\ \forall \alpha > 1]$ is much stronger than by $[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha)),\ \alpha \in (1, \alpha_0]]$ for some given $\alpha_0 > 1$.

How similarly do two sequences with the same C-order and identical Q_{p_0} behave?

COMPARISON C. *$\{\frac{1}{3^{2^k}}\}$ is $\left[(\dot{e}_k = \mathcal{O}(\mathring{e}_k^\alpha)),\ \alpha \in (1, \frac{\ln 3}{\ln 2}]\right]$ faster than $\{\frac{1}{2^{2^k}}\}$ (recall Exercise 1.5). This shows that if $\{\mathring{x}_k\}, \{\dot{x}_k\}$ have C-order p_0 with the same Q_{p_0}, this does not necessarily mean that (in the asymptotic range) one has $\{\dot{e}_k\} \approx \{\mathring{e}_k\}$.*

Brezinski [11] *noted that both $\{\frac{1}{k}\}$ and $\{\frac{1}{k^2}\}$ have the same $Q_1 = 1$ (i.e., C-sublinear order), but quite different speeds.*

REMARK 2.51. *Assessing the asymptotic constant Q_{p_0} may not make sense when $\{x_k\} \subset \mathbb{R}^N, N \geq 2$, as its definition is norm-dependent* [67, E 9.1-2], [20].

3. Computational Versions of the Convergence Orders. The error-based analysis of the orders above is similar in some ways to the local convergence theory for iterative methods from section 4: both assume the existence of the limit/solution x^* and infer essential properties, even if x^* is not known and in practice one needs to use quantities based solely on information available at each step k.

Next, we analyze two practical approaches, equivalent to error-based analysis: the replacing of $|x^* - x_k|$ by (the absolute value of) either the corrections $s_k := |x_{k+1} - x_k|$ or the nonlinear residuals $|f_k| := |f(x_k)|$.

We keep the notation from previous work and obtain a rather theoretical setting in this analysis (e.g., s_k requiring x_{k+1}); in numerical examples, however, we will use only information available at step k (i.e., $x_k, |f_k|, s_{k-1}, x_{k-1}, |f_{k-1}|, s_{k-2}, \ldots$, etc.).

3.1. Computational Convergence Orders Based on Corrections. When the corrections $\{s_k\}$ converge with lower Q-order q_l, then $\{x_k\}$ also converges and attains at least lower R-order q_l.

THEOREM 3.1 (see [75, Thm. 1.2.42], [43, Lem. 4.5.6]). *Let $\{x_k\} \subset \mathbb{R}$. If $\exists c \in (0,1)$ and $k_0 \in \mathbb{N}$ with*

$$|x_{k+1} - x_k| \leq c \cdot |x_k - x_{k-1}| \quad (\textit{i.e., } s_k \leq c \cdot s_{k-1}), \quad k \geq k_0 + 1,$$

then $\exists x^* \in \mathbb{R}$ such that $x_k \to x^*$ at least R-linearly.

If $\exists c > 0, p_0 > 1$, and $k_0 \in \mathbb{N}$ s.t.

$$c^{\frac{1}{p_0 - 1}} s_{k_0} < 1$$

and

$$s_k \leq c \cdot s_{k-1}^{p_0}, \quad k \geq k_0 + 1,$$

then $\exists x^* \in \mathbb{R}$ such that $x_k \to x^*$ with lower R-order at least p_0.

The errors and corrections are tightly connected when the convergence is fast.

LEMMA 3.2 (Potra–Pták–Walker Lemma; [79, Prop. 6.4], [98]). *$x_k \to x^*$ Q-superlinearly iff $x_{k+1} - x_k \to 0$ Q-superlinearly.*

In the case of Q-superlinear convergence, it holds (the Dennis–Moré Lemma [30]*) that*

$$\lim_{k\to\infty} \frac{|x_{k+1} - x_k|}{|x^* - x_k|} = 1. \tag{3.1}$$

REMARK 3.3. (a) *As pointed out by Dennis and Moré in* [30], (3.1) *alone does not imply Q-superlinear convergence: take $x_{2k-1} = \frac{1}{k!}$, $x_{2k} = 2x_{2k-1}$, $k \geq 1$, for example.*

(b) *While the statement and the proofs of this result consider multidimensional spaces and norms (or even metrics) in* (3.1)*, Brezinski* [10] *has noted that in $\mathbb{R}$ the sufficiency can also be proved by using* l'Hôpital's rule for sequences.

As in [4], we use an apostrophe for the resulting quotient factors ($p > 1$):

$$Q'_p(k) := \frac{s_{k+1}}{s_k^p} = \frac{|x_{k+2} - x_{k+1}|}{|x_{k+1} - x_k|^p}, \quad k \geq 0,$$

and the above lemma shows that C-order $p_0 \Rightarrow C'$-order p_0 with $Q_{p_0} = Q'_{p_0}$ (see also [40]). The same statement holds for the Q-orders and corresponding upper/lower limits of $Q_{p_0}(k)$. The equivalence of the error-based and computational orders ($\{C, C'\}$, etc.) is stated in subsection 3.3, which incorporates the nonlinear residuals as well.

The C'-, Q'-, and Q'_ε-orders are semicomputational: they do not use x^* but still require p_0. Instead, of much practical interest are the (full) computational expressions

$$\lim_{k\to\infty} Q'_L(k) = p_0, \qquad Q'_L(k) := \frac{\ln s_{k+1}}{\ln s_k}, \tag{Q'_L}$$

$$\lim_{k\to\infty} Q'_\Lambda(k) = p_0, \qquad Q'_\Lambda(k) := \frac{\ln \frac{s_{k+2}}{s_{k+1}}}{\ln \frac{s_{k+1}}{s_k}}, \tag{Q'_Λ}$$

as they are equivalent to the corresponding error-based orders (see Corollary 3.5).

3.2. Computational Convergence Orders Based on Nonlinear Residuals. In section 4 we study the speed of some iterative methods toward a zero x^* of f. The nonlinear mapping $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is assumed to be sufficiently many times differentiable and the solution is assumed simple:[13] $f'(x^*) \neq 0$ (this means that $f(x) = (x - x^*)g(x)$ and $g(x^*) \neq 0$). However, we will consider multiple solutions as well.

The use of the nonlinear residuals for controlling the convergence to x^* is natural, as they can be seen as "proportional" to the errors: by the Lagrange theorem,

$$f(x) = f'(x + \theta(x^* - x))(x - x^*) \quad \text{for some } \theta \in (0, 1), \tag{3.2}$$

which in applied sciences is usually written as

$$f(x) \approx f'(x^*)(x - x^*) \quad (x \approx x^*). \tag{3.3}$$

However, instead of an asymptotic relation, this is often regarded as an approximate equality. Such an approach is as precise as the idiomatic "the higher the order, the fewer the iterations (needed)": it may not refer to the asymptotic range.

[13] This terminology is used for equations in $\mathbb{R}$, while for systems in $\mathbb{R}^N$ the terminology is "nonsingular," as (the Jacobian) $F'(x^*)$ is a matrix.

Quiz 3.4. If f is a polynomial of degree 2 with $f'(x^*) = 1$ at the root $x^* = 0$, how large can $|f(x)|$ be when $|x - x^*| = 0.0001$ holds (and then for $|x - x^*| = 10^{-16}$)?

Returning to the quotient factors, we use, as in [20], double prime marks,

$$Q''_{p_0}(k) := \frac{|f(x_{k+1})|}{|f(x_k)|^{p_0}} =: \frac{|f_{k+1}|}{|f_k|^{p_0}},$$

and notice that if $Q_{p_0} \in (0, +\infty)$, then, by (3.2),

$$Q''_{p_0}(k) \to \frac{Q_{p_0}}{|f'(x^*)|^{p_0-1}}. \tag{3.4}$$

This leads us to the C''- and Q''-orders $p_0 > 1$. For instance,

$$\lim_{k\to\infty} Q''_L(k) = p_0, \qquad Q''_L(k) := \frac{\ln|f_{k+1}|}{\ln|f_k|}, \tag{Q''_L}$$

$$\lim_{k\to\infty} Q''_\Lambda(k) = p_0, \qquad Q''_\Lambda(k) := \frac{\ln|f_{k+2}/f_{k+1}|}{\ln|f_{k+1}/f_k|}. \tag{Q''_Λ}$$

3.3. The Equivalence of the Error-Based and Computational Orders. This equivalence was given in $\mathbb{R}^N$ [20] (see also [76] and [4]). Here we formulate it as a corollary, since in $\mathbb{R}$ the three C-orders are equivalent (see [20, Ex. 3.22] for $\{C''\} \not\Rightarrow \{C, C'\}$ in $\mathbb{R}^N$).

COROLLARY 3.5 (cf. [20]). *Let f be smooth, x^* simple, $x_k \to x^*$, $p_0 > 1$. Then (see footnote 12)*

$$\{C, C', C''\} \underset{\not\Leftarrow}{\Rightarrow} \{Q, Q', Q''\} \underset{\not\Leftarrow}{\Rightarrow} \{R, R', R''\}.$$

Moreover, the lower/upper orders and asymptotic constants obey

$$\underline{Q}_L = \underline{Q}'_L = \underline{Q}''_L = q_l \quad \textit{and} \quad \bar{Q}_L = \bar{Q}'_L = \bar{Q}''_L = q_u, \quad \textit{resp.}, \tag{3.5}$$

$$\underline{Q}_{q_u} = \underline{Q}'_{q_u} = |f'(x^*)|^{p_0-1}\underline{Q}''_{q_u} \quad \textit{and} \quad \bar{Q}_{q_l} = \bar{Q}'_{q_l} = |f'(x^*)|^{p_0-1}\bar{Q}''_{q_l}. \tag{3.6}$$

Consequently, one has C-order $p_0 > 1$ iff

$$0 < Q_{p_0} = Q'_{p_0} = |f'(x^*)|^{p_0-1}Q''_{p_0} < \infty$$

and Q-order $p_0 > 1$ iff

$$p_0 = \underline{Q}_L = \underline{Q}'_L = \underline{Q}''_L = q_l = \bar{Q}_L = \bar{Q}'_L = \bar{Q}''_L = q_u = Q_\Lambda = Q'_\Lambda = Q''_\Lambda.$$

Relation (3.5) shows equalities of orders, while (3.6) shows that of constants.

The equivalence of the computational orders based on the corrections follows from the Potra–Pták–Walker lemma, while (3.5) and (3.6) follow from the Dennis–Moré lemma.

The equivalence of the computational orders based on the nonlinear residuals is obtained using the Lagrange theorem, by (3.2).

4. Iterative Methods for Nonlinear Equations. Unless otherwise stated, the nonlinear mapping $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is sufficiently smooth and the solution x^* is simple.

Since an equation $f(x) = 0$ might have a unique, several (even infinite), or no solution at all, the iterative methods usually need an initial approximation close to an x^*.

DEFINITION 4.1 ([67, Def. 10.1.1]). *A method converges locally to x^* if $\exists V \subseteq D$ (neighborhood of x^*) s.t. $\forall x_0 \in V$, the generated iterations $\{x_k\} \subset D$ and $x_k \to x^*$.*

4.1. The Newton Method. The first result on the local convergence of a method was given for the Newton iterations (see [67, NR 10.2-1], [101], [23, sect. 6.4]),

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, \ldots, x_0 \in D, \tag{4.1}$$

and it is to due to Cauchy in 1829 ([22], available on the Internet). These iterates are also known as the tangent method, though their initial devising did not consider this geometric interpretation.[14]

4.1.1. History. This method can be traced back to ancient times (Babylon and Egypt, ca. 1800 B.C.), as it is conjectured that it was used in the computation (with five correct—equivalent—decimal digits) of $\sqrt{2}$ (see [3], [49, sect. 1.7], [74], with the references therein, and also [86, p. 39] for an annotated picture). While this is not certain, there is no doubt that Heron used these iterations for $\sqrt{a}$ [3].

"Heron of Alexandria (first century A.D.) seems to have been the first to explicitly propose an iterative method of approximation in which each new value is used to obtain the next value," as noted by Chabert et al. [23, p. 200]. Indeed, Heron has approximated $\sqrt{720}$ in an iterative fashion by elements corresponding to the Newton method for solving $x^2 - 720 = 0$ [23, sect. 7.1].

We can also recognize the Newton method for $x^2 - A = 0$ used by Theon of Alexandria (ca. 370 A.D.) in approximating $\sqrt{4500}$, based entirely on geometric considerations (originating from Ptolemy; see [74]). See [23, sect. 7.2].

So efficient is this method in computing two of the basic floating point arithmetic operations—the square root and division (see Exercise 4.4)—that it is still a choice even in current (2021) codes (see, e.g., [62, pp. 138, 124]).

Let us recall some general comments on the subsequent history of this method, thoroughly analyzed by Ypma in [101].

> A method algebraically equivalent to Newton's method was known to the 12th century algebraist Sharaf al-Dīn al-Ṭūsī [...], and the 15th century Arabic mathematician Al-Kāshī used a form of it in solving $x^p - N = 0$ to find roots of N. In western Europe a similar method was used by Henry Briggs in his Trigonometria Britannica, published in 1633, though Newton appears to have been unaware of this [...]. [101] (see also [82])

In solving nonlinear problems using such iterates, Newton (≈1669) and subsequently Raphson (1690) dealt only with polynomial equations,[15, 16] described in [90].

[14] The first such interpretation was given by Mourraille (1768) [23, sect. 6.3].

[15] $x^3 - 2x - 5 = 0$ is "the classical equation where the Newton method is applied" [90].

[16] Actually, as noted by Ypma, Newton considered such iterations in solving a transcendent equation, namely, the Kepler equation $x - e \sin x = M$. "However, given the geometrical obscurity of the argument, it seems unlikely that this passage exerted any influence on the historical development of the Newton-Raphson technique in general" [101].

Newton considered the process of successively computing the corrections, which are ultimately added together to form the final approximation (cf. [90]); see Algorithm 4.1.

Algorithm 4.1 The algorithm devised by Newton (left), and his first example (right)

Let $f(x)$, x_0	$[f(x) = x^3 - 2x - 5,\ x_0 = 2]$
Let $g_0(s) = f(x_0 + s)$	$[g_0(s) = (2+s)^3 - 2(2+s) - 5 = s^3 + 6s^2 + 10s - 1]$
For $k = 0 : m - 1$	
Compute $\tilde{g}_k(s)$ (the order-one approx. of $g_k(s)$)	$[\tilde{g}_0(s) = 10s - 1]$
Solve for s_k in $\tilde{g}_k(s) = 0$	$[s_0 = 0.1]$
Let $g_{k+1}(s) = g_k(s_k + s)$	$[g_1(s) = g_0(s_0 + s) = s^3 + 6.3s^2 + 11.23s + 0.061]$
$x_n := x_0 + \sum_{k=0}^{n-1} s_k$	$[x_n = 2 + 0.1 - \frac{0.061}{11.23} + \cdots + s_{n-1}]$

Raphson considered approximations updated at each step, a process equivalent to (4.1) (see [55], [101], [90]). However, the derivatives of f (which could be calculated with the "fluxions" of that time) do not appear in their formulae, only the first-order approximations from the finite series development of polynomials (using the binomial formula). For more than a century, the belief was that these two variants represented two different methods;[17] it was Stewart (1745) who pointed out that they are in fact equivalent (see [90]). As noted by Steihaug [90],

> It took an additional 50 years before it was generally accepted that the methods of Raphson and Newton were identical methods, but implemented differently. Joseph Lagrange in 1798 derives the Newton-Raphson method [...] and writes that the Newton's method and Raphson's method are the same but presented differently and Raphson's method is *plus simple que celle de Newton*.

Simpson (1740) was the first to apply the method to transcendent equations, using "fluxions."[18, 16] Even more important, he extended it to the solving of nonlinear systems of two equations (see [23], [101], [90]),[19] subsequently generalized to the usual form known today: $F(x) = 0$ with $F : \mathbb{R}^N \to \mathbb{R}^N$, for which, denoting the Jacobian of F at x_k by $F'(x_k)$, one has to solve for s_k at each step the linear system

$$F'(x_k)s = -F(x_k). \tag{4.2}$$

Cajori [13] noted (see also [101]):

> Then appear writers like Euler, Laplace, Lacroix and Legendre, who explain the Newton-Raphson process, but use no names. Finally, in a publication of Budan in 1807, in those of Fourier of 1818 and 1831, in that of Dandelin in 1826, the Newton-Raphson method is attributed to Newton.

[17] "Nearly all eighteenth century writers and most of the early writers of the nineteenth century carefully discriminated between the method of Newton and that of Raphson" (Cajori [13]). For a given input (n fixed), the result of Algorithm 4.1 may depend on how accurately the different fractions (e.g., $\frac{0.061}{11.23}$) are computed.

[18] The fluxions $\dot{x}$ did not represent f', but are "essentially equivalent to dx/dt; implicit differentiation is used to obtain dy/dt, subsequently dividing through by dx/dt as instructed produces the derivative $A = dy/dx$ of the function. [...] Thus Simpson's instructions closely resemble, and are mathematically equivalent to, the use of [(4.1)]. The formulation of the method using the now familiar $f'(x)$ calculus notation of [(4.1)] was published by Lagrange in 1798 [...]" [101] (see also [90]).

[19] The first nonlinear system considered was $y + \sqrt{y^2 - x^2} - 10 = 0$, $x + \sqrt{y^2 + x} - 12 = 0$, with $(x_0, y_0) = (5, 6)$, according to [101] and [90].

> The immense popularity of Fourier's writings led to the universal adoption of the misnomer "Newton's method" for the Newton-Raphson process.

While some authors use "the Newton–Raphson method,"[20] the name Simpson—though appropriate[21]—is only occasionally encountered (e.g., [59], [34], [41], [39]). In a few cases "Newton–Kantorovich" appears, but refers to the mid-1900 contributions of Kantorovich to semilocal convergence results when F is defined on normed spaces.

4.1.2. Attainable C-Orders. The method attains at least C-order 2 for simple solutions. Regarding rigorous convergence results, "Fourier appears to have been the first to approach this question in a note entitled *Question d'analyse algébrique* (1818) [...], but an error was found in his formula," as noted by Chabert et al. [23, sect. 6.4], who continue by noting that

> Cauchy studied the subject from 1821 onwards [...], but did not give a satisfactory formulation until 1829. [...] The concern for clarity and rigour, which is found in all of Cauchy's work, tidies up the question of the convergence of Newton's method for the present.

THEOREM 4.2 (cf. [22], [23, Thm. II, p. 184]). *If x^* is simple and $f''(x^*) \neq 0$, then the Newton method converges locally, with C-quadratic order:*

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^2} = \left|\frac{f''(x^*)}{2f'(x^*)}\right| =: Q_2. \tag{4.3}$$

REMARK 4.3. *The usual proofs for* (4.3) *consider the Taylor developments of $f(x_k)$ and $f'(x_k)$, assuming small errors and neglecting "higher-order terms":*

$$\ell_{k+1} = \ell_k - \frac{f(x^*) + f'(x^*)\ell_k + \cdots}{f'(x^*) + f''(x^*)\ell_k + \cdots} \approx \frac{f''(x^*)}{2f'(x^*)}\ell_k^2 \qquad (\ell_k := x_k - x^*,\ \textit{i.e., signed}). \tag{4.4}$$

EXERCISE 4.4. (a) *Write the Newton iterates for computing $\sqrt{2}$ (cf.* [62, p. 138]*).*
(b) *(Newton everywhere* [96, sect. 17.9]*) Apply* (4.1) *to compute the division a/b of real numbers in terms of only the operations $+, -, *$ (cf.* [62, p. 124]*).*

REMARK 4.5. (a) *The usual perception is that the smaller Q_2, the faster the asymptotic speed; expression* (4.3) *shows that for the Newton method this is realized with smaller $|f''|$ and larger $|f'|$, and this means that near x^*, the graph of f resembles a line ($|f''|$ small) having large angle with Ox (the line is not close to Ox).*[22] *While this interpretation holds locally for the asymptotic range of the Newton iterates, it is worth mentioning that the larger the region with such a property, the larger the attraction set (see also* (4.12)*), and the better conditioned the problem of solving $f(x) = 0$ (see subsection* 4.1.6*).*

We note that taking $2f(x)$ instead of $f(x)$, despite doubling $f'(x)$, does not reduce Q_2, leading in fact to the same iterates (this property is called affine invariance*; see* [34]*).*

[20] "Who was '–Raphson'?" (N. Bićanić and K. H. Johnson, 1979) has the popular answer "Raphson was Newton's programmer" (S. Nash, 1992).

[21] "[...] it would seem that the Newton–Raphson–Simpson method is a designation more nearly representing the facts of history in reference to this method" (Ypma [101]).

"I found no source which credited Simpson as being an inventor of the method. None the less, one is driven to conclude that neither Raphson, Halley nor anyone else prior to Simpson applied fluxions to an iterative approximation technique" (Kollerstrom [55]).

[22] In the opposite situation, when the angle is zero (x^* is multiple), the convergence is only linear.

(b) *Of course, it would be of major interest to state a conclusion like "the smaller Q_2, the fewer the iterates needed." However, even if supported by a number of numerical examples, such a statement should be treated with great care. The reason is simple: we do not address similar objects (Q_2 is an asymptotic constant, referring to $k \to \infty$, while the desired conclusion refers to a finite number of steps). Moreover, when we say "the smaller Q_2," this means that we consider different mappings f; even if they have the same x^*, the attraction set and the dynamics of the iterates may be different.*

The role of the asymptotic constants will be analyzed in the forthcoming work [21].

Let us analyze the expectation that this method behaves similarly for the same Q_2.

EXAMPLE 4.6. (a) *Let $a \in \mathbb{R}$ be given and*

$$f(x) := x + x^2 + ax^3 = 0, \qquad \text{with } x^* = 0.$$

We get $f'(0) = 1$ and $f''(0) = 2$, so by (4.3), $Q_2 = 1$ ($\forall a \in \mathbb{R}$ given).

The full Taylor series gives the exact relation for the signed errors from (4.4):

$$\ell_{k+1} = \frac{1 + 2a\ell_k}{1 + 2\ell_k + 3a\ell_k^2}\ell_k^2. \tag{4.5}$$

Take $x_0 = 0.0001 (= e_0)$ fixed. By (4.4), we should obtain $|\ell_1| = e_1 \approx e_0^2 = 10^{-8}$ for any $a > 0$. In Table 4.1 we see this does not hold when a has large values.

As $|a|$ grows, despite being "small," the error e_0 in fact is not "squared," as suggested by (4.4). Indeed, when $|a|$ is very large, in (4.4) the terms with larger errors (corresponding to $f'''(x) = 6a$) are actually neglected, despite their having higher powers.

(b) *Taking $x_0 = 0.0001$ for f as above and then for $f(x) = x + 2x^2$ shows that Q_2 and the number of iterates needed (for the same accuracy) are not in direct proportion.*

Table 4.1 *$e_1 = |\ell_1|$ computed in two equivalent ways, for different a's (`digits64`, Julia).*

uncaught typo: binary64 instead of digits64

a	$\ell_1 (= x_1)$	ℓ_1 by (4.5)
1	$9.999\,999\,700\,077\,396 \cdot 10^{-9}$	$9.999\,999\,700\,059\,996 \cdot 10^{-9}$
10	$1.001\,799\,339\,592\,549\,2 \cdot 10^{-8}$	$1.001\,799\,339\,592\,279\,7 \cdot 10^{-8}$
10^5	$2.093\,301\,435\,406\,632\,7 \cdot 10^{-7}$	$2.093\,301\,435\,406\,699 \cdot 10^{-7}$
10^6	$1.951\,077\,460\,687\,246\,8 \cdot 10^{-6}$	$1.951\,077\,460\,687\,245 \cdot 10^{-6}$
10^7	$1.538\,994\,000\,922\,936\,2 \cdot 10^{-5}$	$1.538\,994\,000\,922\,934\,8 \cdot 10^{-5}$

REMARK 4.7. *Igarashi and Ypma* [46] *noticed that a smaller Q_2 does not necessarily attract fewer iterates for attaining a given accuracy. However, the numerical examples used to support this claim were based not on revealing the individual asymptotic regime of the sequences, but on the number of iterates required to attain a given (`digits64`) accuracy.*

If $f''(x^*) = 0$, then (4.3) implies Q-superquadratic order, but actually the order is higher: it is given by the first index≥ 2 of the nonzero derivative at the solution.

THEOREM 4.8 (see, e.g., [37]). *Let x^* be simple. The Newton method converges locally with C-order $p_0 \in \mathbb{N}$, $p_0 \geq 2$, iff*

$$f^{(p_0)}(x^*) \neq 0 \quad \text{and} \quad f''(x^*) = \cdots = f^{(p_0-1)}(x^*) = 0,$$

in which case

$$Q_{p_0} = \frac{p_0 - 1}{p_0!} \left| \frac{f^{(p_0)}(x^*)}{f'(x^*)} \right|.$$

4.1.3. Examples for the Attained C-Orders. We consider decreasing orders.
(a) *Infinite C-order.* If $f(x) = ax + b$, $a \neq 0$, then $x_1 = x^*$ $\forall x_0 \in \mathbb{R}$.
(b) *Arbitrary natural C-order* $p_0 \geq 2$ (see Remark 4.13 for $p_0 \notin \mathbb{N}$).

EXAMPLE 4.9. *Let $p \geq 2$ be a natural given number and consider*

$$f(x) = x + x^p, \quad x^* = 0, \ p = 2, 3, 4 \ [83].$$

By Theorem 4.8*, we obtain local convergence with C-order p and $Q_p = p - 1$.*

Let us first deal with $p = 2$. In double precision, for $x_0 = 0.46$ we obtain six nonzero iterates (shown truncated in Table 4.2*). In analyzing $Q_L(k-1)$ (shown with the last six decimal digits truncated), we note that $Q_L(4)$ is a better approximation than the last one, $Q_L(5)$. This phenomenon may appear even if we increase the precision (see Figure* 4.2*).*

When $x_0 = 0.47$ and 0.45 we obtain five nonzero iterates (see Exercise 4.11*), and $Q_L(k)$ is increasingly better.*

Table 4.2 *Newton iterates for $x + x^2 = 0$ (*`digits64`*, Julia).*

uncaught typo: binary64 instead of digits64

k	x_k	$Q_L(k-1)$	x_k	$Q_L(k-1)$	x_k	$Q_L(k-1)$
0	0.47		0.46		0.45	
1	$1.1 \cdot 10^{-1}$	2.877 706 159	$1.1 \cdot 10^{-01}$	2.840 052 802	$1.0 \cdot 10^{-1}$	2.803 816 781
2	$1.0 \cdot 10^{-2}$	2.094 428 776	$9.9 \cdot 10^{-03}$	2.090 320 979	$9.3 \cdot 10^{-3}$	2.086 305 525
3	$1.0 \cdot 10^{-4}$	2.004 592 994	$9.7 \cdot 10^{-05}$	2.004 275 304	$8.6 \cdot 10^{-5}$	2.003 972 042
4	$1.1 \cdot 10^{-8}$	2.000 023 942	$9.4 \cdot 10^{-09}$	2.000 021 019	$7.4 \cdot 10^{-9}$	2.000 018 386
5	$1.4 \cdot 10^{-16}$	2.000 000 001	$8.8 \cdot 10^{-17}$	2.000 000 001	$5.4 \cdot 10^{-17}$	2.000 000 001
6	0		$1.2 \cdot 10^{-32}$	1.987 981 962	0	
7	0		0			

For $p = 3$ and 4, we obtain four nonzero iterations for the three choices of x_0, and therefore a single meaningful term in computing $\{Q'_\Lambda(k)\}$.

REMARK 4.10. *The higher the convergence order, the faster* $\mathrm{fl}(x^*)$ *may be reached, and we might end up with fewer distinct terms.*

Higher precision is needed in this study. We have chosen to use the Julia language here, which allows not only quadruple precision (`digits128`*, compliant with IEEE* 754-2008*) but arbitrary precision as well (by using the* `setprecision` *command). Some examples are posted at https://github.com/ecatinas/conv-ord and may be easily run online (e.g., by IJulia notebooks running on Binder).*[23]

In Figure 4.1 *we show the resulting $Q_L(k-1)$, $Q_\Lambda(k-2)$, $Q'_L(k-2)$, $Q'_\Lambda(k-3)$ (in order to use only the information available at step k).*

We see that all four (computational) convergence orders approach the corresponding p's, but we cannot clearly distinguish their speed.

In Figure 4.2 *we present $|Q_\Lambda(k-2) - p|$ and $|Q_L(k-1) - p|$.*

[23] Such examples may be also performed in MATLAB [58] (by choosing the Advanpix [1] toolbox, which handles arbitrary precision), or in other languages.

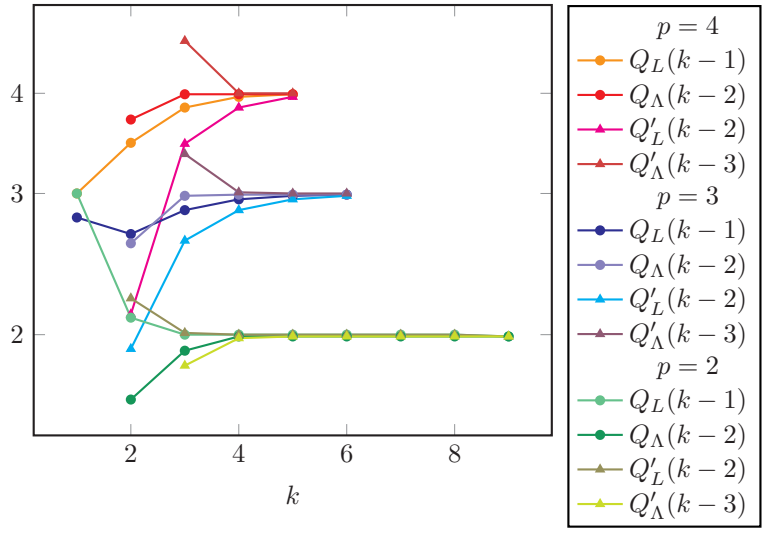


Fig. 4.1 *Convergence orders: the Newton iterates for* $f_{[p]}(x) = x + x^p = 0$, $p = 2, 3, 4$.

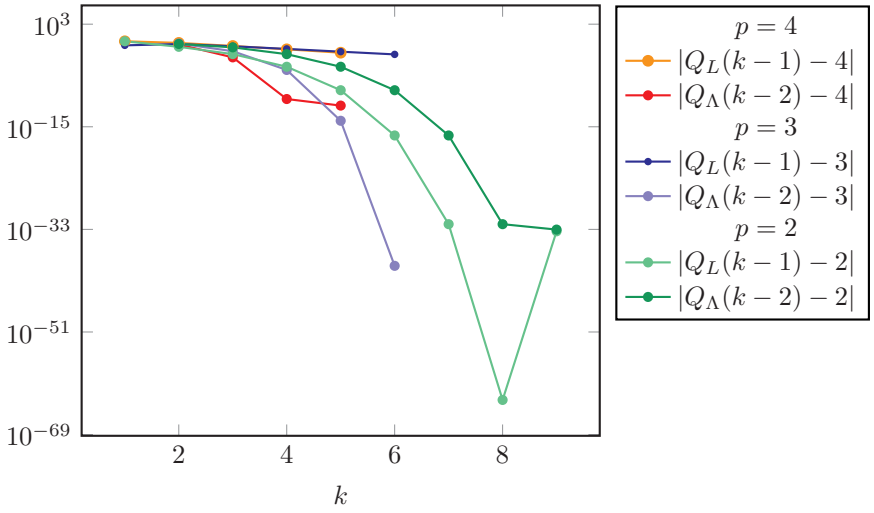


Fig. 4.2 $|Q_\Lambda(k-2) - p|$, $|Q_L(k-1) - p|$ *(Newton iterates,* $f_{[p]}(x) = x + x^p = 0$*).*

EXERCISE 4.11. *Examining the numerator and denominator in the Newton corrections, give the two different reasons why* x_7 *is* 0 *in Table* 4.2 *(when* $x_0 = 0.47$ *and* 0.45*). When is* `realmin` *or* `machine epsilon` *(or both) too large?*

(c) *C-order* $p \in (1, 2)$. The C-order 2 is lost with the smoothness of f (when $\nexists f''$).

EXAMPLE 4.12 ([67, E 10.2-4], [31]). *The Newton method has C-order* $1 + \alpha$ *for*

$$f(x) = x + |x|^{1+\alpha}, \quad \alpha \in (0, 1) \text{ given}, \quad x^* = 0.$$

REMARK 4.13. *Modifying the above example by taking* $f(x) = x + |x|^{p_0}$ *shows that the Newton method may in fact attain any real nonnatural order* $p_0 \in (2, +\infty)$.

(d) *C-linear order.* This appears for a multiple solution x^*, which is implicitly assumed to be isolated ($f'(x) \neq 0\ \forall x \neq x^*$ in a neighborhood of x^*).[24] Indeed, if the multiplicity of x^* is $q \in \mathbb{N}$,

$$f(x^*) = f'(x^*) = \cdots = f^{(q-1)}(x^*) = 0, \quad f^{(q)}(x^*) \neq 0 \tag{4.6}$$

(i.e., $f(x) = (x - x^*)^q g(x)$, $g(x^*) \neq 0$), then (see, e.g., [67, E 10.2-8], [36], [41])

$$Q_1 = 1 - \frac{1}{q}. \tag{4.7}$$

If q is known, one may use the Schröder formula [68] to restore the C-order 2:

$$x_{k+1} = x_k - q\frac{f(x_k)}{f'(x_k)}, \qquad \text{with } Q_2 = \frac{|f^{(q+1)}(x^*)|}{q(q+1)|f^{(q)}(x^*)|} \text{ (see, e.g., [68, (8.13)]).}$$

EXERCISE 4.14. (a) *Calculate the Newton and Schröder iterates for* $f(x) = x^2$.
(b) *The Schröder iterates are the Newton iterates for* $\sqrt[q]{f(x)} = (x - x^*)\sqrt[q]{g(x)}$.

Approximation of the unknown q is surveyed in [60, sect. 7.9]; see also [81, sect. 6.6.2].

(e) *C-sublinear order.*

EXERCISE 4.15. *Let* $f(x) = e^{-1/x^2}$, $x \neq 0$, $f(0) = 0$. *Show that the Newton method converges locally to* $x^* = 0$, *with* $Q_1(\{x_k\}) = 1$, *i.e., C-sublinearly.*

4.1.4. Convexity. The influence of x_0 and the convexity of f were first analyzed by Mourraille (1768) [23, sect. 6.3], in an intuitive fashion.

In 1818, Fourier ([38], available on the Internet) gave the condition

$$f(x_0)f''(x_0) > 0, \tag{4.8}$$

which ensures the convergence of the Newton method, provided f maintains the monotony and convexity in the interval determined by x^* and x_0. This condition can lead to sided convergence intervals (i.e., not necessarily centered at x^*).

Clearly, the method may converge even on the whole axis, a result which is more interesting to consider in $\mathbb{R}^n$ (cf. the Baluev theorem [66, Thm. 8.3.4]).

REMARK 4.16. (a) *By* (4.4), *the signs of the errors are periodic with period either* 1 *(i.e., monotone iterates) or* 2 *(*$k \geq k_0$*).*

(b) *The Fourier condition is sufficient for convergence but not also necessary; the iterates may converge locally (by Theorem* 4.2*) even when* (4.8) *does not hold: take, e.g.,* $f(x) = \sin(x)$, $x^* = 0$, *with the signs of errors alternating at each step (*$k \geq k_0$*).*

4.1.5. Attraction Balls. Estimations for $B_r(x^*) = \{x \in D : |x^* - x| < r\} = (x^* - r, x^* + r)$ as V in Definition 4.1 are usually given using the condition

$$|f'(x) - f'(y)| \leq L|x - y| \quad \forall x, y \in D. \tag{4.9}$$

The Lipschitz constant L of f' measures the nonlinearity of f (i.e., how much the graph of f is bent); its exact bound shows this is in accordance with Remark 4.5, since

$$L \geq \sup_{x \in D} |f''(x)|. \tag{4.10}$$

[24] Otherwise, an iterative method would not be able to converge to a specific solution we were interested in (take, e.g., the constant function $f(x) = 0$ with $\mathbb{R}$ as the solution set, or $f(x) = x \sin \frac{1}{x}$, $x \neq 0$, $f(0) = 0$).

The usual assumptions also require

$$|f'(x^*)| \geq \frac{1}{\beta}, \tag{4.11}$$

and the Dennis–Schnabel [32], Rheinboldt [84], and Deuflhard–Potra [35] estimations are, respectively,

$$r = \frac{1}{2\beta L}, \quad r = \frac{2}{3\beta L}, \quad \text{and } r = \frac{2}{\omega}, \tag{4.12}$$

where ω satisfies $|f'(x)^{-1}| \cdot |f'(x+t(y-x)) - f'(x)| \leq t\omega|y-x|$ $\forall t \in (0,1)$, $\forall x, y \in D$, i.e., it can be viewed as a combination of both L and β (see also [2]).

We note that (4.12) estimates a small r for Example 4.6, as L is large.

4.1.6. Floating Point Arithmetic. Curious behavior may arise in this setting.

The derivative f' measures the sensitivity in computing $f(x)$ using $\tilde{x} \approx x$ instead: $|f(x) - f(\tilde{x})| \approx |f'(x)| \cdot |x - \tilde{x}|$. This shows that for large $|f'(x)|$, the absolute error in x may be amplified. When the relative error in x is of interest, the *condition number* becomes $|f'(x)| \cdot |x|$: $|f(x) - f(\tilde{x})| \approx |f'(x)| \cdot |x| \frac{|x-\tilde{x}|}{|x|}$.

For the relative errors of f, the *condition numbers* are given by [69], [42, sects. 1.2.6 and 5.3]

$$\frac{|f(x) - f(\tilde{x})|}{|f(x)|} \approx \frac{|xf'(x)|}{|f(x)|} \frac{|x - \tilde{x}|}{|x|} \quad \text{or} \quad \frac{|f(x) - f(\tilde{x})|}{|f(x)|} \approx \frac{|f'(x)|}{|f(x)|} |x - \tilde{x}|,$$

which show that the relative errors in $f(x)$ may be large near a solution x^*.

However, as shown in [42, sects. 1.2.6 and 5.3] and [81, Ex. 2.4, sect. 6.1], the problem of solving an equation has the conditioning number $1/|f'(x^*)|$, which is equal to the inverse of the conditioning number for computing $f(x)$.

When $|f'(x)|$ (i.e., $|f'(x^*)|$) is not small, Shampine, Allen, and Pruess argue that the Newton method is "stable at limiting precision" [89, p. 147], as the Newton corrections $\frac{f(x)}{f'(x)}$ are small.

When $|f'(x^*)|$ is small (e.g., x^* is multiple), notable cancellations may appear in computing f and prevent the usual graphical interpretation. The (expanded) polynomial $(x-2)^9$ of Demmel [28, p. 8] computed in double precision is well known.

Even simpler examples exist: $f(x) = x^3 - 2x^2 + \frac{4}{3}x - \frac{8}{27} (= (x - \frac{2}{3})^3)$ (Sauer [86, p. 45]) and $f(x) = x^3 e^{3x} - 3x^2 e^{2x} + 3xe^x - 1 (= (e^x - 1)^3)$, $x^* \approx 0.56$ (Shampine, Allen, and Pruess [89, Fig. 1.2, p. 23]).

$|f'|$ not being small is no guarantee that things won't get weird due to other reasons.

Quiz 4.17 (Wilkinson; see [89, Ex. 4.3]). Let $f(x) = x^{20} - 1$, $x^* = 1$, and (here) $x_0 = 10$. Why, in double precision, are the (rounded) Newton iterates: $x_1 = 9.5, x_2 = 9.0, x_3 = 8.6, x_4 = 8.1$, etc.? Take $\ell_0 = \pm 0.001$ and compute x_1 and e_1.

4.1.7. Nonlinear Systems in $\mathbb{R}^N$. When Newton-type methods are used in practice to solve nonlinear systems in $\mathbb{R}^N$, obtaining high Q-orders becomes challenging. As the dimension may be huge ($N = 10^6$ and even higher) the main difficulty resides—having solved the representation of the data in memory[25]—in accurately solving the

[25] One can use Newton–Krylov methods, which do not require storage of the matrices $F'(x)$, but instead computation of matrix-vector products $F'(x)v$, further approximated by finite differences [12].

resulting linear systems (4.2) at each step. In many practical situations, superlinear convergence may be considered satisfactory.

The most common setting allows the systems (4.2) to be only approximately solved (the corrections verify $F'(x_k)s_k = -F(x_k) + r_k$, $r_k \neq 0$), leading to the *inexact Newton method*; its superlinear convergence and the orders $p \in (1, 2]$ were characterized by Dembo, Eisenstat, and Steihaug in [27], the approach being easily extended by us for further sources of perturbation in [14], [17].

The quasi-Newton methods form another important class of Newton-type iterations; they consider approximate Jacobians, in order to produce linear systems that are easier to solve [31]. Their superlinear convergence was characterized by Dennis and Moré [30].

The inexact and quasi-Newton methods are in fact tightly connected models [17].

4.2. The Secant Method. This is a two-step method, requiring $x_0 \neq x_1 \in D$,

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k)-f(x_{k-1})}{x_k - x_{k-1}}} = \frac{x_{k-1}f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad k = 1, 2, \ldots, \tag{4.13}$$

but none of the above formulae is recommended for programming (see subsection 4.2.5).

4.2.1. History. Though nowadays the secant method is usually seen as a Newton-type method with derivatives replaced by finite differences (i.e., a quasi-Newton method) or an inverse interpolatory method, this is not the way it first appeared. The roots of this method can be traced back to approximately the same time as that of the Newton method, the 18th century B.C., found on Babylonian clay tablets and on the Egyptian Rhind Papyrus [23], [70], when it was used as a single iteration for solving linear equations.[26] During these times and even later, the terminology for such a noniterated form was "the rule of double false position" (see, e.g., [70]); indeed, two initial "false positions" x_0, x_1 yield in one iteration the true solution of the linear equation.

Heron of Alexandria [33] approximated the cubic root of 100 by this method (as revealed by Luca and Păvăloiu [57], it turns out that it can be seen as applied equivalently to $x^2 - \frac{100}{x} = 0$).

While reinvented in several cultures (in China,[27] India,[28] Arab countries, and Africa,[29] sometimes even for 2-by-2 systems of equations), its use as an iterative method was first considered by Cardano (1545), who called it *Regula Liberae Positionis* [70] (or *regula aurea* [82, note 5, p. 200]).

Viète also used a secant-type method, and it appears that Newton (around 1665) independently rediscovered the secant method [70], [101].

4.2.2. Attainable Q-Orders. The standard convergence result shows the usual convergence with C-order given by the golden ratio $\lambda_1 := \frac{1+\sqrt{5}}{2} \approx 1.618$; as noted by Papakonstantinou and Tapia [70], this was first obtained by Jeeves in 1958 (in a technical report, published as [48]).

THEOREM 4.18 ([68, Chap. 12, sects. 11–12]). *If x^* is simple and $f''(x^*) \neq 0$,*

[26] Such equations appear trivial these days, but their solution required ingenuity back then: the decimal system and the zero symbol were used much later, in the 800s (see, e.g., [61, pp. 192–193]).

[27] Jiuzhang Suanshu (*Nine Chapters on the Mathematical Art*) [23, sect. 3.3].

[28] For example, in the work of Madhava's student Paramesvara [73], [74], [23, sect. 3.4].

[29] al-Ṭūsī, al-Kāshī, al-Khayyām (see [82], [101], [23, sect. 3.5]).

then the secant method has local convergence with C-order $\lambda_1 = \frac{1+\sqrt{5}}{2}$*:*

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^{\lambda_1}} = \left|\frac{f''(x^*)}{2f'(x^*)}\right|^{\lambda_1 - 1} =: Q_{\lambda_1}. \tag{4.14}$$

REMARK 4.19. *The Q-order* λ_1 *follows by considering signed errors (see, e.g.,* [48]*)*

$$\ell_{k+1} = \frac{[x_{k-1}, x_k, x^*; f]}{[x_{k-1}, x_k; f]}\ell_k\ell_{k-1}, \quad k \geq 1 \quad (\ell_k := x_k - x^*), \tag{4.15}$$

and then Exercise 2.25*. Indeed, the divided differences converge,* $[x_{k-1}, x_k; f] \to f'(x^*)$*,* $[x_{k-1}, x_k, x^*; f] \to \frac{f''(x^*)}{2}$*, and denoting* $l = \left|\frac{f''(x^*)}{2f'(x^*)}\right|$*, it follows that in* (2.5) *we may take* $A = l - \varepsilon$ *and* $B = l + \varepsilon$ *for some small* ε*.*

Quiz 4.20. Find the incompleteness in proving the C-order λ_1 in (4.14) by arguing that

$$(l-\varepsilon)\Big(\frac{e_{k-1}^{\lambda_1}}{e_k}\Big)^{\frac{1}{\lambda_1}} e_k^{\frac{1}{\lambda_1}+1-\lambda_1} \leq (l-\varepsilon)\frac{e_k e_{k-1}}{e_k^{\lambda_1}} \leq \frac{e_{k+1}}{e_k^{\lambda_1}} \leq (l+\varepsilon)\frac{e_k e_{k-1}}{e_k^{\lambda_1}} = (l+\varepsilon)\Big(\frac{e_{k-1}^{\lambda_1}}{e_k}\Big)^{\frac{1}{\lambda_1}},$$

$\frac{1}{\lambda_1} + 1 - \lambda_1 = 0$, and therefore Q_{λ_1} satisfies $Q_{\lambda_1} = lQ_{\lambda_1}^{-\frac{1}{\lambda_1}}$, i.e., is given by (4.14).

REMARK 4.21. *By* (4.15)*, the signs of the errors (and of* $f(x_k)$ *as well) are periodic with period at most* 3 *(*$k \geq k_0$*), as noted by Neumaier* [64, Cor. 5.1.3].

Higher Q- (but not necessarily C-)orders may be attained.

THEOREM 4.22 (Raydan [83]). *If* x^* *is simple and* $q \geq 1$ *is the first index such that* $f^{(q+1)}(x^*) \neq 0$*, then the secant method converges locally with Q-order given by*

$$\lambda_q = \frac{1 + \sqrt{1+4q}}{2}$$

and obeys

$$\lim \frac{e_{k+1}}{e_k e_{k-1}^q} = \frac{|f^{(q+1)}(x^*)|}{(q+1)!|f'(x^*)|}.$$

Moreover, specific results hold in the following cases, noting the attained orders:

- $q = 1$*: C-order* λ_1*;*
- $q = 2$*: exact Q-order* $\lambda_2 = 2$ *(but not necessarily C-order* 2*);*
- $q = 3$*: Q-order* $\lambda_3 \approx 2.3$ *(but not necessarily exact Q-order* λ_3*).*

REMARK 4.23. *When* $q \geq 2$*, the secant iterates may have only Q- but no C-order, despite* $\frac{e_{k+1}}{e_k e_{k-1}^q}$ *converging to a finite nonzero limit (see Exercise* 2.25*).*

EXAMPLE 4.24 (cf. [83]). *Let* $f_{[q+1]}(x) = x + x^{q+1}$*,* $x_0 = 0.5$*,* $x_1 = \frac{1}{(q+1)^{\lambda_q}}$*,* $q = 1, 2, 3$*, and compute the secant iterates (by* (4.16)*).*

The assertions of Theorem 4.22 *are verified in this case, since* $|Q_L(k-1) - \lambda_q|$ *and* $|Q_\Lambda(k-2) - \lambda_q|$*, computed using* `setprecision(500)`*, tend to* 0 *(see Figure* 4.3*).*

In Figure 4.4 *we plot* $Q_{\lambda_q}(k)$ *when* $q = 1, 2,$ *and* 3*.*

We see that $Q_{\lambda_1}(k)$ *tends to* 1*.*

For $q = 2$*,* $Q_{\lambda_2}(k)$ *oscillates (suggesting* $\underline{Q}_{\lambda_2}(k) \approx 0.57$*,* $\bar{Q}_{\lambda_2}(k) \approx 1.74$*).*

For $q = 3$*,* $\bar{Q}_{\lambda_3}(k)$ *tends to infinity (as rigorously proved by Raydan for* $q \geq 3$*). For this particular data it appears that* $\underline{Q}_{\lambda_3} = 0$ *(which remains to be rigorously*

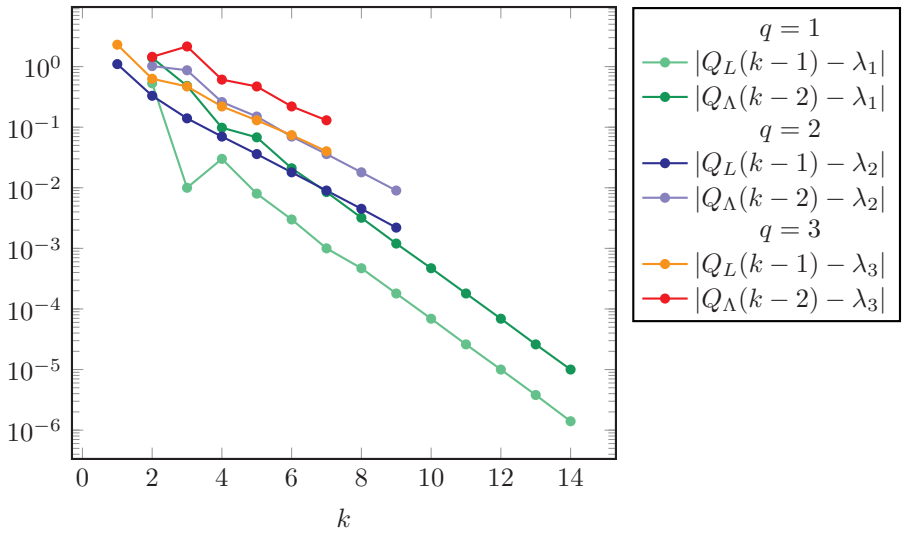


Fig. 4.3 $|Q_\Lambda(k-2) - \lambda_q|$, $|Q_L(k-1) - \lambda_q|$ *(secant iterates,* $f(x) = x + x^{1+q} = 0$*).*

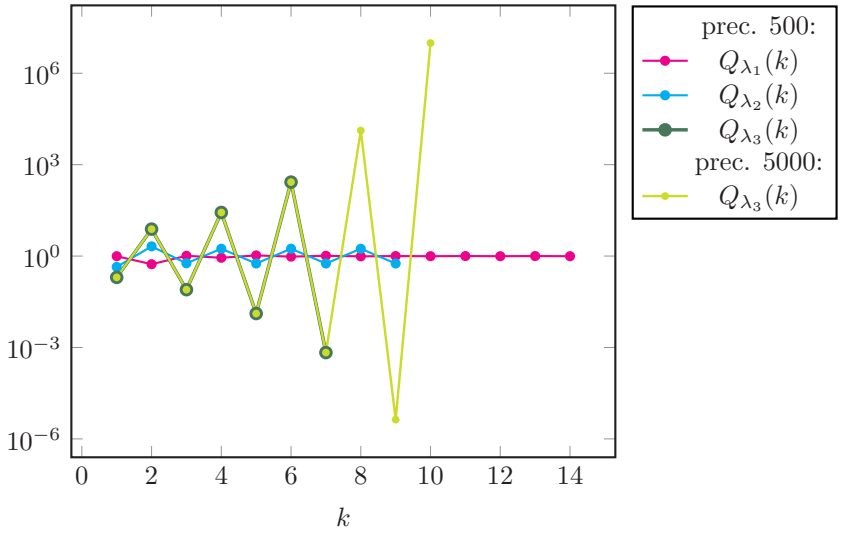


Fig. 4.4 $Q_{\lambda_q}(k)$ *(secant iterates,* $f_{[1+q]}(x)$*,* `setprecision(500)`*, resp.,* `(5000)`*).*

proved). In order to see it more clearly, we have also used further increased precision (`setprecision(5000)`*), providing confirmation.*

When $q = 3$*,* $\bar{Q}_{\lambda_3} = \infty$ *or* $\underline{Q}_{\lambda_3} = 0$ *do not necessarily hold for any* x_0, x_1*: they do not hold (numerically) for* $x_0 = 0.5, x_1 = 0.25$ *(but they do hold, e.g., for* $x_0 = 1, x_1 = 0.5$*).*

The higher the order, the faster the iterates attain $\mathrm{fl}(x^*)$*. We see in Figure* 4.4 *that* `setprecision(500)` *allows* 14 *iterates when* $q = 1$ *and only* 7 *iterates when* $q = 3$ *(increased to* 11 *for* `setprecision(5000)`*).*

4.2.3. Multiple Solutions. Not only the high orders, but even very local convergence may be lost if x^* is not simple.

EXAMPLE 4.25 (cf. [41]). *If* $f(x) = x^2$, $x_0 = -\varepsilon, x_1 = \varepsilon$, $\nexists x_2$ *in* (4.16) $(\forall \varepsilon > 0)$.

Assuming that the secant method converges, only a linear rate is obtained [95]; Díez [36] showed that if the multiplicity of x^* in (4.6) is

- $q = 2$, then $Q_1 = \frac{1}{\lambda_1} = \lambda_1 - 1 \approx 0.618$;
- $q \geq 3$, then Q_1 is the $0 < \lambda^* < 1$ solution of $x^q + x^{q-1} - 1 = 0$.

Also, if the iterates verify at each step $|f(x_k)| < |f(x_{k-1})|$ (by swapping), one obtains superlinear convergence for the root secant method of Neumaier [64, p. 241]:

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{1 - \sqrt{f(x_{k-1})/f(x_k)}}.$$

For other references on this topic, see [60, sect. 7.9] and [91, Ex. 1.9].

4.2.4. Attraction Balls. Estimates for the radius of an attraction ball were given by Liang [56]: if (4.9) and $\left|[x, y, z; f]\right| \leq K \ \forall x, y, z \in D$ hold, then

$$r = \frac{1}{3\beta K}.$$

When using a specific x_1 in the secant method ($x_0 \in B_r$ arbitrary, $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$), estimates similar to (4.12) were obtained: if (4.11) and (4.10), then

$$r = \frac{2}{3\beta L} \quad \text{(see [56])}.$$

4.2.5. Floating Point Arithmetic. The two expressions in (4.13) are not recommended for programming in an *ad literam* fashion, respectively, at all; indeed, the first one should be written instead as

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad k = 1, 2, \ldots, \tag{4.16}$$

while the second one can lead, close to x^*, to cancellations when $f(x_k)$ and $f(x_{k-1})$ have the same sign [26, p. 228], [60, p. 4].

Further equivalent formulae are

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{1 - \frac{f(x_{k-1})}{f(x_k)}}, \quad k = 1, 2, \ldots,$$

and (Vandergraft [97, p. 265])

$$x_{k+1} = x_k - \left[(x_{k-1} - x_k)\left(\frac{f(x_k)}{f(x_{k-1})}\right)\right] \Big/ \left(1 - \frac{f(x_k)}{f(x_{k-1})}\right),$$

which requires $|\frac{f(x_k)}{f(x_{k-1})}| < 1$ at each step (by swapping the iterates).

It is worth noting, however, that swapping the consecutive iterates in order to obtain $|f(x_k)| < |f(x_{k-1})|$ leads to iterates that may have a different dynamic than is standard.

The secant method may not be stable at limiting precision, even for large $|f'(x^*)|$ (Shampine, Allen, and Pruess [89, p. 147]), and higher precision may be needed.

4.3. Successive Approximations for Fixed Point Problems. We consider a sufficiently differentiable mapping $g : D \to D$, $D \subseteq \mathbb{R}$, and the fixed point problem

$$g(x) = x.$$

A solution x^* is called a fixed point of g; x^* is further called *an attraction fixed point* when one has local convergence for the successive approximations

$$x_{k+1} = g(x_k), \quad k \geq 0.$$

4.3.1. History. Assuming that the computation of $\sqrt{2}$ with five (decimal) digits is highly improbable in the absence of an iterative method, it is likely that successive approximations date back to at least 1800 B.C. (considering the Newton iterates as particular instances of successive approximations).

Regarding the first use of iterates not connected to the Newton or secant-type iterates, they seem have appeared in the fifth and sixth centuries in India (see Plofker [74]; in this paper the author describes as an example the *Brahmasphutasiddhanta* of Brahmagupta from the seventh century).

Later occurrences are mentioned for Ibn al-Banna as cited in [23, sect. 7.3], al-Ṭūsī [23, sect. 7.4], Viète [23, sect. 7.5], Parameśvara (aviśeṣa) [73], al-Kāshī [23, sect. 7.5], and Ḥabash al-Ḥāsib al-Marwazi [52], [3].

Further occurrences of such iterates appear for Kepler (1618–1622) in [23, sect. 7.6], Gregory (1672 [101]; 1674 [3]), Dary (1674) [3], and Newton (1674) [101].

4.3.2. Local Convergence, Attainable C-Orders. The first part of the following local convergence result was first obtained by Schröder in 1870 [87]. The extension to mappings in $\mathbb{R}^N$ implied the analysis of the eigenvalues of the Jacobian at the fixed point, which is an important topic in numerical analysis.[30] The standard formulation of the result was first given by Ostrowski in 1957, but Perron had essentially already obtained it in 1929 (see [67, NR 10.1-1]).

THEOREM 4.26.

(a) ([67, Thm. 10.1.3], [68]) *Let g be continuous on $D \subseteq \mathbb{R}$ and differentiable at the fixed point $x^* \in D$. If*

$$|g'(x^*)| < 1,$$

then x^* is an attraction fixed point.

(b) ([37]) *Conversely, if x^* is an attraction fixed point, then $|g'(x^*)| \leq 1$.*

REMARK 4.27 (see [67, E 10.1-2]). *Condition $|g'(x^*)| \leq 1$ is sharp: by considering $|g'(x^*)| = 1$ and taking $g(x) = x - x^3$, $x^* = 0$ is an attraction fixed point, while if $g(x) = x + x^3$, the same fixed point x^* is no longer of attraction.*

The higher orders attained are characterized as follows.

THEOREM 4.28 (see, e.g., [68, Chap. 4, sect. 10, p. 44], [95], [37]). *Let $p_0 \geq 2$ be an integer and x^* a fixed point of g. Then $x_{k+1} = g(x_k) \to x^*$ with C-order p_0 iff*

$$g'(x^*) = g''(x^*) = \cdots = g^{(p_0-1)}(x^*) = 0 \quad \textit{and} \quad g^{(p_0)}(x^*) \neq 0,$$

in which case

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^{p_0}} = \frac{|g^{(p_0)}(x^*)|}{p_0!} =: Q_{p_0}. \tag{4.17}$$

[30] Given $H \in \mathbb{R}^{N\times N}$ and the linear system $x = Hx + c$, the following result was called by Ortega [66, p. 118] the *fundamental theorem of linear iterative methods:* If $x = Hx + c$ has a unique solution x^*, then the successive approximations converge to x^* $\forall x_0 \in \mathbb{R}^N$ iff the spectral radius $\rho(H)$ is < 1.

EXAMPLE 4.29 ([67, E 10.1-10]). *Any C-order $p_0 > 1$ is possible: let $g(x) = x^{p_0}$.*

The C-sublinear order may be attained too.

EXERCISE 4.30. *Let $g(x) = x - x^3$ and $x_{k+1} = g(x_k) \to 0$; then $Q_1(\{x_k\}) = 1$.*

The successive approximations are usually regarded as the most general iterations, but actually, they may also be seen as instances of an inexact Newton method [15]; such an approach allows us to study acceleration techniques in a different setting.

4.3.3. Attraction Balls. The following result holds.

THEOREM 4.31 ([19]). *Let x^* be an attraction fixed point of g for which*

$$|g'(x^*)| \le q < 1.$$

Suppose there exist $r_1 > 0$, $L > 0$ such that g is differentiable on B_{r_1}, with g' Lipschitz continuous of constant L, and denote

$$r = \min\left\{r_1, \sqrt{\frac{2(1-q)}{L}}\right\}. \tag{4.18}$$

Then, $\forall x_0 \in B_r$, letting $t = \frac{L}{2}|x_0 - x^*| + q < 1$ we get

$$|x_{k+1} - x^*| \le t|x_k - x^*|, \quad k \ge 0,$$

which implies that $x_{k+1} = g(x_k) \in B_r$ and $x_{k+1} = g(x_k) \to x^*$.

REMARK 4.32. *It is important to note that this result does not require g to be a contraction on D (i.e., $L < 1$), in contrast to the classical Banach contraction theorem.*

The center-Lipschitz condition (i.e., x^* instead of y in (4.9)) is an improvement [24].

The estimate (4.18) may be sharp in certain cases.

EXAMPLE 4.33 ([19]). *Let $g(x) = x^2$ with $x^* = 0$, $g'(x^*) = 0$, $r_1 > 0$ arbitrary, and $L = 2$. Then (4.18) gives the sharp $r = r_2 = 1$, since $x^* = 1$ is another fixed point.*

We observe that $|g'(x)| = 2$ at $|x - x^*| = r = 1$, which upholds Remark 4.32.

Regarding the general case of several dimensions with $G : D \subseteq \mathbb{R}^N \to D$, we note that the trajectories with high convergence orders are characterized by the zero eigenvalue of $G'(x^*)$ and its corresponding eigenvectors (see [16]).

REMARK 4.34 ([18, Ex. 2.8]). *Definition 4.1 allows a finite number of $x_k \notin V$ (regardless of how "small" V is chosen to be): take, e.g., $G(x) = Ax$, $A = \left(\begin{smallmatrix} 0 & 2 \\ \frac{1}{8} & 0 \end{smallmatrix}\right)$, $V = B_r(0)$.*

5. Conclusions. Various aspects of convergence orders have been analyzed here, but rather than providing a final view, we see this work as a fresh start; in [21] we pursue the study of the following themes: the characterizations of the C-order, the connections between the Q- and R-orders and the big Oh rates, the asymptotic constants (with their immediate computational variants), the convergence orders of the discretization schemes (to mention a few).

6. Answers to Quizzes. Quiz 1.8: $\{\dot{x}_k\} = \{\frac{1}{2^{2^k}}\} = \{\mathring{x}_k\}$ in different windows.

Quiz 2.7: (a) $Q_1 = 10^{-1}$; (b) $Q_1 = 10^{-2}$; (c) 10^{-x_k}, $\{x_k\}$ the Fibonacci sequence; (d) $Q_2 = 1$.

Quiz 3.4: unbounded (take $f(x) = x + ax^2$).

Quiz 4.17: $\mathrm{fl}(10^{20} - 1) = \mathrm{fl}(10^{20})$, in double precision.

Quiz 4.20: the limit Q_{λ_1} is not proved to exist; see also Theorem 4.22.

Acknowledgments. I am grateful to two referees for constructive remarks which helped to improve the manuscript, particularly the one who brought the Julia language to my attention; also, I am grateful to my colleagues M. Nechita and I. Boros, as well as to S. Filip for useful suggestions regarding Julia.

REFERENCES

[1] Advanpix team, *Advanpix*, Version 4.7.0.13642, 2020, https://www.advanpix.com/ (accessed 2021/04/07). (Cited on pp. 588, 610)

[2] I. K. Argyros and S. George, *On a result by Dennis and Schnabel for Newton's method: Further improvements*, Appl. Math. Lett., 55 (2016), pp. 49–53, https://doi.org/10.1016/j.aml.2015.12.003. (Cited on p. 613)

[3] D. F. Bailey, *A historical survey of solution by functional iteration*, Math. Mag., 62 (1989), pp. 155–166, https://doi.org/10.1080/0025570X.1989.11977428. (Cited on pp. 591, 606, 618)

[4] W. A. Beyer, B. R. Ebanks, and C. R. Qualls, *Convergence rates and convergence-order profiles for sequences*, Acta Appl. Math., 20 (1990), pp. 267–284, https://doi.org/10.1007/bf00049571. (Cited on pp. 592, 593, 594, 595, 598, 602, 604, 605)

[5] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, SIAM Rev., 59 (2017), pp. 65–98, https://doi.org/10.1137/141000671. (Cited on p. 588)

[6] R. P. Brent, *Algorithms for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, NJ, 1973. (Cited on p. 595)

[7] R. P. Brent, S. Winograd, and P. Wolfe, *Optimal iterative processes for root-finding*, Numer. Math., 20 (1973), pp. 327–341, https://doi.org/10.1007/bf01402555. (Cited on pp. 597, 600)

[8] C. Brezinski, *Comparaison des suites convergentes*, Rev. Française Inform. Rech. Opér., 5 (1971), pp. 95–99. (Cited on pp. 597, 598)

[9] C. Brezinski, *Accélération de la Convergence en Analyse Numérique*, Springer-Verlag, Berlin, 1977. (Cited on pp. 586, 597, 598)

[10] C. Brezinski, *Limiting relationships and comparison theorems for sequences*, Rend. Circolo Mat. Palermo Ser. II, 28 (1979), pp. 273–280, https://doi.org/10.1007/bf02844100. (Cited on p. 604)

[11] C. Brezinski, *Vitesse de convergence d'une suite*, Rev. Roumaine Math. Pures Appl., 30 (1985), pp. 403–417. (Cited on pp. 595, 597, 603)

[12] P. N. Brown, *A local convergence theory for combined inexact-Newton/finite-difference projection methods*, SIAM J. Numer. Anal., 24 (1987), pp. 407–434, https://doi.org/10.1137/0724031. (Cited on p. 613)

[13] F. Cajori, *Historical note on the Newton-Raphson method of approximation*, Amer. Math. Monthly, 18 (1911), pp. 29–32, https://doi.org/10.2307/2973939. (Cited on p. 607)

[14] E. Cătinaş, *Inexact perturbed Newton methods and applications to a class of Krylov solvers*, J. Optim. Theory Appl., 108 (2001), pp. 543–570, https://doi.org/10.1023/a:1017583307974. (Cited on p. 614)

[15] E. Cătinaş, *On accelerating the convergence of the successive approximations method*, Rev. Anal. Numér. Théor. Approx., 30 (2001), pp. 3–8, https://ictp.acad.ro/accelerating-convergence-successive-approximations-method/ (accessed 2021/04/07). (Cited on p. 619)

[16] E. Cătinaş, *On the superlinear convergence of the successive approximations method*, J. Optim. Theory Appl., 113 (2002), pp. 473–485, https://doi.org/10.1023/a:1015304720071. (Cited on p. 619)

[17] E. Cătinaş, *The inexact, inexact perturbed and quasi-Newton methods are equivalent models*, Math. Comp., 74 (2005), pp. 291–301, https://doi.org/10.1090/s0025-5718-04-01646-1. (Cited on p. 614)

[18] E. Cătinaş, *Newton and Newton-type Methods in Solving Nonlinear Systems in* $\mathbb{R}^N$, Risoprint, Cluj-Napoca, Romania, 2007, https://ictp.acad.ro/methods-of-newton-and-newton-krylov-type/ (accessed on 2021/6/9). (Cited on p. 619)

[19] E. Cătinaş, *Estimating the radius of an attraction ball*, Appl. Math. Lett., 22 (2009), pp. 712–714, https://doi.org/10.1016/j.aml.2008.08.007. (Cited on p. 619)

[20] E. Cătinaş, *A survey on the high convergence orders and computational convergence orders of sequences*, Appl. Math. Comput., 343 (2019), pp. 1–20, https://doi.org/10.1016/j.amc.2018.08.006. (Cited on pp. 591, 592, 594, 595, 596, 598, 599, 600, 601, 602, 603, 605)

[21] E. Cătinaş, *How Many Steps Still Left to x*? Part* II. *Newton Iterates without f and f′*, manuscript, 2020. (Cited on pp. 603, 609, 619)

[22] A. Cauchy, *Sur la determination approximative des racines d'une equation algebrique ou transcendante*, Oeuvres Complete (II) 4, Gauthier-Villars, Paris, 1899, 23 (1829), pp. 573–609, http://iris.univ-lille.fr/pdfpreview/bitstream/handle/1908/4026/41077-2-4.pdf?sequence=1 (accessed 2021/04/07). (Cited on pp. 606, 608)

[23] J.-L. Chabert, ed., *A History of Algorithms. From the Pebble to the Microchip*, Springer-Verlag, Berlin, 1999. (Cited on pp. 606, 607, 608, 612, 614, 618)

[24] J. Chen and I. K. Argyros, *Improved results on estimating and extending the radius of an attraction ball*, Appl. Math. Lett., 23 (2010), pp. 404–408, https://doi.org/10.1016/j.aml.2009.11.007. (Cited on p. 619)

[25] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, SIAM, Philadelphia, PA, 2000, https://doi.org/10.1137/1.9780898719857. (Cited on pp. 598, 601)

[26] G. Dahlquist and Å. Björk, *Numerical Methods*, Dover, Mineola, NY, 1974. (Cited on p. 617)

[27] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408, https://doi.org/10.1137/0719025. (Cited on p. 614)

[28] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997, https://doi.org/10.1137/1.9781611971446. (Cited on pp. 587, 613)

[29] J. E. Dennis, *On Newton-like methods*, Numer. Math., 11 (1968), pp. 324–330, https://doi.org/10.1007/bf02166685. (Cited on p. 602)

[30] J. E. Dennis, Jr., and J. J. Moré, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp., 28 (1974), pp. 549–560, https://doi.org/10.1090/s0025-5718-1974-0343581-1. (Cited on pp. 604, 614)

[31] J. E. Dennis, Jr., and J. J. Moré, *Quasi-Newton methods, motivation and theory*, SIAM Rev., 19 (1977), pp. 46–89, https://doi.org/10.1137/1019005. (Cited on pp. 611, 614)

[32] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA, 1996, https://doi.org/10.1137/1.9781611971200. (Cited on pp. 599, 613)

[33] G. Deslauries and S. Dubuc, *Le calcul de la racine cubique selon Héron*, Elem. Math., 51 (1996), pp. 28–34, http://eudml.org/doc/141587 (accessed 2020/02/18). (Cited on p. 614)

[34] P. Deuflhard, *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*, Springer-Verlag, Heidelberg, 2004. (Cited on p. 608)

[35] P. Deuflhard and F. A. Potra, *Asymptotic mesh independence of Newton–Galerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal., 29 (1992), pp. 1395–1412, https://doi.org/10.1137/0729080. (Cited on p. 613)

[36] P. Díez, *A note on the convergence of the secant method for simple and multiple roots*, Appl. Math. Lett., 16 (2003), pp. 1211–1215, https://doi.org/10.1016/s0893-9659(03)90119-4. (Cited on pp. 612, 617)

[37] F. Dubeau and C. Gnang, *Fixed point and Newton's methods for solving a nonlinear equation: From linear to high-order convergence*, SIAM Rev., 56 (2014), pp. 691–708, https://doi.org/10.1137/130934799. (Cited on pp. 609, 618)

[38] J. Fourier, *Question d'analyse algébrique*, Bull. Sci. Soc. Philo., 67 (1818), pp. 61–67, http://gallica.bnf.fr/ark:/12148/bpt6k33707/f248.item (accessed 2021-04-07). (Cited on pp. 591, 612)

[39] W. Gautschi, *Numerical Analysis*, 2nd ed., Birkhäuser/Springer, New York, 2011, https://doi.org/10.1007/978-0-8176-8259-0. (Cited on p. 608)

[40] M. Grau-Sánchez, M. Noguera, and J. M. Gutiérrez, *On some computational orders of convergence*, Appl. Math. Lett., 23 (2010), pp. 472–478, https://doi.org/10.1016/j.aml.2009.12.006. (Cited on p. 604)

[41] A. Greenbaum and T. P. Chartier, *Numerical Methods. Design, Analysis, and Computer Implementation of Algorithms*, Princeton University Press, Princeton, NJ, 2012. (Cited on pp. 608, 612, 617)

[42] M. T. Heath, *Scientific Computing. An Introductory Survey*, 2nd ed., SIAM, Philadelphia, PA, 2018, https://doi.org/10.1137/1.9781611975581. (Cited on pp. 593, 613)
[43] M. Heinkenschloss, *Scientific Computing. An Introductory Survey*, Rice University, Houston, TX, 2018. (Cited on pp. 587, 599, 600, 603)
[44] J. Herzberger and L. Metzner, *On the Q-order of convergence for coupled sequences arising in iterative numerical processes*, Computing, 57 (1996), pp. 357–363, https://doi.org/10.1007/BF02252254. (Cited on p. 597)
[45] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, PA, 2002, https://doi.org/10.1137/1.9780898718027. (Cited on p. 587)
[46] M. Igarashi and T. J. Ypma, *Empirical versus asymptotic rate of convergence of a class of methods for solving a polynomial equation*, J. Comput. Appl. Math., 82 (1997), pp. 229–237, https://doi.org/10.1016/s0377-0427(97)00077-0. (Cited on p. 609)
[47] L. O. Jay, *A note on Q-order of convergence*, BIT, 41 (2001), pp. 422–429, https://doi.org/10.1023/A:1021902825707. (Cited on pp. 589, 595, 596, 601, 602)
[48] T. A. Jeeves, *Secant modification of Newton's method*, Commun. ACM, 1 (1958), pp. 9–10, https://doi.org/10.1145/368892.368913. (Cited on pp. 595, 614, 615)
[49] V. J. Katz, *A History of Mathematics. An Introduction*, 2nd ed., Addison-Wesley, 1998. (Cited on p. 606)
[50] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, PA, 1995, https://doi.org/10.1137/1.9781611970944. (Cited on pp. 589, 594, 602)
[51] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, PA, 1999, https://doi.org/10.1137/1.9781611970920. (Cited on pp. 599, 601)
[52] E. S. Kennedy and W. R. Transue, *A medieval iterative algorism*, Amer. Math. Monthly, 63 (1956), pp. 80–83, https://doi.org/10.1080/00029890.1956.11988762. (Cited on p. 618)
[53] D. E. Knuth, *Big Omicron and big Omega and big Theta*, ACM SIGACT News, 8 (1976), pp. 18–24, https://doi.org/10.1145/1008328.1008329. (Cited on pp. 586, 598)
[54] D. E. Knuth, *The TeXbook*, Addison-Wesley, 1984. (Cited on p. 588)
[55] N. Kollerstrom, *Thomas Simpson and Newton's method of approximation: An enduring myth*, British J. History Sci., 25 (1992), pp. 347–354, https://doi.org/10.1017/s0007087400029150. (Cited on pp. 607, 608)
[56] K. Liang, *Homocentric convergence ball of the secant method*, Appl. Math. Chin. Univ., 22 (2007), pp. 353–365, https://doi.org/10.1007/s11766-007-0313-3. (Cited on p. 617)
[57] D. Luca and I. Păvăloiu, *On the Heron's method for approximating the cubic root of a real number*, Rev. Anal. Numér. Théor. Approx., 26 (1997), pp. 103–108, https://ictp.acad.ro/jnaat/journal/article/view/1997-vol26-nos1-2-art15 (accessed 2021-04-07). (Cited on p. 614)
[58] Mathworks, *MATLAB*, Version 2019b, 2019, http://www.mathworks.com (accessed 2021/04/07). (Cited on pp. 588, 610)
[59] J. M. McNamee, *Numerical Methods for Roots of Polynomials, Part* I, Elsevier, Amsterdam, 2007. (Cited on p. 608)
[60] J. M. McNamee and V. Y. Pan, *Numerical Methods for Roots of Polynomials, Part* II, Elsevier, Amsterdam, 2013. (Cited on pp. 612, 617)
[61] U. C. Merzbach and C. B. Boyer, *A History of Mathematics*, 3rd ed., John Wiley & Sons, 2011. (Cited on p. 614)
[62] J.-M. Muller, N. Brunie, F. de Dinechin, C.-P. Jeannerod, M. Joldes, V. Lefèvre, G. Melquiond, N. Revol, and S. Torres, *Handbook of Floating-Point Arithmetic*, 2nd ed., Springer, New York, 2018, https://doi.org/10.1007/978-3-319-76526-6. (Cited on pp. 606, 608)
[63] A. S. Nemirovsky and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization*, Wiley, 1983. (Cited on p. 599)
[64] A. Neumaier, *Introduction to Numerical Analysis*, Cambridge University Press, 2001. (Cited on pp. 615, 617)
[65] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer, New York, 2006, https://doi.org/10.1007/978-0-387-40065-5. (Cited on pp. 592, 601, 602)
[66] J. M. Ortega, *Numerical Analysis. A Second Course*, SIAM, Philadelphia, PA, 1990, https://doi.org/10.1137/1.9781611971323. (Cited on pp. 612, 618)
[67] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia, PA, 2000, https://doi.org/10.1137/1.9780898719468. (Cited on pp. 592, 594, 595, 598, 599, 600, 602, 603, 606, 611, 612, 618, 619)
[68] A. M. Ostrowski, *Solution of Equations and Systems of Equations*, 2nd ed., Academic Press, New York, 1966; 3rd ed., 1972. (Cited on pp. 612, 614, 618)

[69] M. L. Overton, *Numerical Computing with IEEE Floating Point Arithmetic*, SIAM, Philadelphia, PA, 2001, https://doi.org/10.1137/1.9780898718072. (Cited on pp. 587, 613)

[70] J. M. Papakonstantinou and R. A. Tapia, *Origin and evolution of the secant method in one dimension*, Amer. Math. Monthly, 120 (2013), pp. 500–518, https://doi.org/10.4169/amer.math.monthly.120.06.500. (Cited on p. 614)

[71] I. Păvăloiu, *Sur l'ordre de convergence des méthodes d'itération*, Mathematica, 23 (46) (1981), pp. 261–265, https://ictp.acad.ro/convergence-order-iterative-methods/. (Cited on p. 598)

[72] I. Păvăloiu, *Optimal algorithms concerning the solving of equations by interpolation*, in Research on Theory of Allure, Approximation, Convexity and Optimization, Editura Srima, 1999, pp. 222–248, http://ictp.acad.ro/optimal-algorithms-concerning-solving-equations-interpolation/ (accessed 2021-04-07). (Cited on p. 597)

[73] K. Plofker, *An example of the secant method of iterative approximation in a fifteenth-century Sanskrit text*, Historia Math., 23 (1996), pp. 246–256, https://doi.org/10.1006/hmat.1996.0026. (Cited on pp. 614, 618)

[74] K. Plofker, *Use and transmission of iterative approximations in India and the Islamic World*, in From China to Paris: 2000 Years Transmission of Mathematical Ideas, Y. Dold-Samplonius, J. W. Dauben, M. Folkerts, and B. Van Dalen, eds., Franz Steiner Verlag Wiesbaden GmbH, Stuttgart, 2002, pp. 167–186. (Cited on pp. 606, 614, 618)

[75] E. Polak, *Optimization. Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997. (Cited on pp. 589, 595, 599, 600, 603)

[76] F. A. Potra, *On Q-order and R-order of convergence*, J. Optim. Theory Appl., 63 (1989), pp. 415–431, https://doi.org/10.1007/bf00939805. (Cited on pp. 592, 595, 597, 598, 599, 600, 601, 602, 605)

[77] F. A. Potra, *Q-superlinear convergence of the iterates in primal-dual interior-point methods*, Math. Program. Ser. A, 91 (2001), pp. 99–115, https://doi.org/10.1007/s101070100230. (Cited on p. 600)

[78] F. A. Potra, *A superquadratic version of Newton's method*, SIAM J. Numer. Anal., 55 (2017), pp. 2863–2884, https://doi.org/10.1137/17M1121056. (Cited on p. 597)

[79] F. A. Potra and V. Pták, *Nondiscrete Induction and Iterative Processes*, Pitman, Boston, MA, 1984. (Cited on pp. 592, 595, 597, 599, 600, 604)

[80] F. A. Potra and R. Sheng, *A path following method for LCP with superlinearly convergent iteration sequence*, Ann. Oper. Res., 81 (1998), pp. 97–114, https://doi.org/10.1023/A:1018942131812. (Cited on p. 598)

[81] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, Springer, Berlin, 2007. (Cited on pp. 612, 613)

[82] R. Rashed, *The Development of Arabic Mathematics: Between Arithmetic and Algebra*, Stud. Philos. Sci. 156, Springer, Boston, MA, 1994. (Cited on pp. 606, 614)

[83] M. Raydan, *Exact order of convergence of the secant method*, J. Optim. Theory Appl., 78 (1993), pp. 541–551, https://doi.org/10.1007/BF00939881. (Cited on pp. 610, 615)

[84] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, in Mathematical Models and Numerical Methods, Banach Ctr. Publ. 3, Polish Academy of Science, 1977, pp. 129–142, https://doi.org/10.4064/-3-1-129-142. (Cited on p. 613)

[85] W. C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, 2nd ed., SIAM, Philadelphia, PA, 1998, https://doi.org/10.1137/1.9781611970012. (Cited on pp. 595, 600)

[86] T. Sauer, *Numerical Analysis*, 2nd ed., Pearson, Boston, MA, 2012. (Cited on pp. 606, 613)

[87] E. Schröder, *Ueber unendlich viele Algorithmen zur Auflösung der Gleichungen*, Math. Ann., 2 (1870), pp. 317–365, https://doi.org/10.1007/bf01444024; available as *On Infinitely Many Algorithms for Solving Equations*, Tech reports TR–92–121 and TR–2990, Institute for Advanced Computer Studies, University of Maryland, 1992; translated by G. W. Stewart. (Cited on pp. 592, 618)

[88] H. Schwetlick, *Numerische Lösung Nichtlinearer Gleichungen*, VEB, Berlin, 1979. (Cited on pp. 592, 597, 598, 600)

[89] L. F. Shampine, R. C. Allen, Jr., and S. Pruess, *Fundamentals of Numerical Computing*, John Wiley and Sons, New York, 1997. (Cited on pp. 613, 617)

[90] T. Steihaug, *Computational science in the eighteenth century. Test cases for the methods of Newton, Raphson, and Halley:* 1685 *to* 1745, Numer. Algorithms, 83 (2020), pp. 1259–1275, https://doi.org/10.1007/s11075-019-00724-8. (Cited on pp. 606, 607)

[91] E. Süli and D. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, UK, 2003. (Cited on p. 617)
[92] T. Tantau, *The tikz and pgf Packages. Manual for Version* 3.1.5*b*-34-*gff*02*ccd*1, https://github.com/pgf-tikz/pgf. (Cited on p. 588)
[93] R. A. Tapia, J. E. Dennis, Jr., and J. P. Schäfermeyer, *Inverse, shifted inverse, and Rayleigh quotient iteration as Newton's method*, SIAM Rev., 60 (2018), pp. 3–55, https://doi.org/10.1137/15M1049956. (Cited on pp. 591, 601, 602)
[94] R. A. Tapia and D. L. Whitley, *The projected Newton method has order* $1+\sqrt{2}$ *for the symmetric eigenvalue problem*, SIAM J. Numer. Anal., 25 (1988), pp. 1376–1382, https://doi.org/10.1137/0725079. (Cited on p. 597)
[95] J. F. Traub, *Iterative Methods for the Solutions of Equations*, Prentice Hall, Englewood Cliffs, NJ, 1964. (Cited on pp. 597, 617, 618)
[96] E. E. Tyrtyshnikov, *A Brief Introduction to Numerical Analysis*, Birkhäuser/Springer, New York, 1997. (Cited on p. 608)
[97] J. S. Vandergraft, *Introduction to Numerical Computations*, 2nd ed., Academic Press, New York, 1983. (Cited on p. 617)
[98] H. F. Walker, *An approach to continuation using Krylov subspace methods*, in Computational Science in the 21st Century, J. Periaux, ed., John Wiley and Sons, 1997, pp. 72–81. (Cited on p. 604)
[99] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997, https://doi.org/10.1137/1.9781611971453. (Cited on p. 596)
[100] S. J. Wright, *Optimization algorithms for data analysis*, in The Mathematics of Data, IAS/Park City Math. Ser. 25, AMS, Providence, RI, 2018, pp. 49–97. (Cited on pp. 599, 600)
[101] T. J. Ypma, *Historical development of the Newton–Raphson method*, SIAM Rev., 37 (1995), pp. 531–551, https://doi.org/10.1137/1037125. (Cited on pp. 591, 606, 607, 608, 614, 618)