1. Additions to Chapter 10 book

1.1. Construction of Gauss quadrature. This subsection replaces [SM03, §10.1, §10.2 until Thm. 10.2].

For \( a < b, \ w(x) > 0 \) on \([a, b] \) almost everywhere (a.e.) with \( \int_a^b w(x)dx < \infty \), let for \( f \in C[a,b] \),

\[
I(f) := \int_a^b w(x)f(x) \, dx.
\]

For \( \{x_0, \ldots, x_n\} \in [a, b], \ w_0, \ldots, w_n \in \mathbb{R} \), let

\[
Q_n(f) := \sum_{i=0}^n w_i f(x_i).
\]

**Theorem 1.1.** Let \( Q_n \) be exact on \( \mathcal{P}_n \) (so \( w_i = \int_a^b w(x)L_i^{(n)}(x) \, dx \)). Then it is exact on \( \mathcal{P}_m \) for an \( m > n \) if and only if

\[
\pi_{n+1} \perp_{(\cdot,\cdot)_w} \mathcal{P}_{m-n-1}, \]

where \( \pi_{n+1}(x) := \prod_{i=0}^n(x-x_i) \) and \( (g,k)_w := \int_a^b w(x)g(x)k(x) \, dx \).

**Proof.** Suppose \( Q_n \) is exact on \( \mathcal{P}_m \). Then for all \( p \in \mathcal{P}_{m-n-1} \), it holds that

\[
\int_a^b w(x)p(x)\pi_{n+1}(x) \, dx = \sum_{i=0}^n w_ip(x_i)\pi_{n+1}(x_i) = 0.
\]

Conversely, let \( \pi_{n+1} \perp_{(\cdot,\cdot)_w} \mathcal{P}_{m-n-1} \). Each \( p \in \mathcal{P}_m \) can be written as \( p = q\pi_{n+1} + r \) where \( q \in \mathcal{P}_{m-n-1} \) and \( r \in \mathcal{P}_n \). We have

\[
\int_a^b w(x)p(x) \, dx = \int_a^b w(x)q(x)\pi_{n+1}(x) \, dx + \int_a^b w(x)r(x) \, dx = \int_a^b w(x)r(x) \, dx,
\]

and

\[
\sum_{i=0}^n w_ip(x_i) = \sum_{i=0}^n w_iq(x_i)\pi_{n+1}(x_i) + \sum_{i=0}^n w_ir(x_i) = \sum_{i=0}^n w_ir(x_i).
\]

From the fact that \( Q_n \) is exact on \( \mathcal{P}_n \), we infer that the expressions on both right-hand sides are equal, and so \( I(p) = Q_n(p) \), meaning that \( Q_n \) is exact on \( \mathcal{P}_m \). \( \Box \)

By selecting \( \pi_{n+1} \) such that \( \pi_{n+1} \perp_{(\cdot,\cdot)_w} \mathcal{P}_n \), that is, by taking \( \{x_0, \ldots, x_n\} \) the roots of the orthogonal polynomial of degree \( n+1 \), and by determining the weights \( w_0, \ldots, w_n \) such that \( Q_n \) is exact on \( \mathcal{P}_n \), the resulting \((n+1)\)-point formula, known as the Gauss formula, is thus exact on \( \mathcal{P}_{2n+1} \).

An example of a Gauss formula for \( n = 1, [a, b] = [0, 1] \) and \( w \equiv 1 \) is given in [SM03, Example 10.1]. There the weights are computed as \( w_i = \int_a^b w(x)L_i^{(n)}(x)^2 \, dx \).

---

*Date: April 5, 2020.*
Apparently, with this choice of quadrature points it holds that
\[ \int_a^b w(x)L_i^{(n)}(x)^2 \, dx = \int_a^b w(x)L_i^{(n)}(x) \, dx, \] cf. [SM03, Exer. 10.2].

**Proposition 1.2.** For the \((n+1)\)-point Gauss formula \(Q_n\), and \(f \in C^{(2n+2)}[a,b]\), it holds that
\[ I(f) - Q_n(f) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} \int_a^b w(x)\pi_{n+1}^2(x) \, dx, \]
for some \(\xi \in [a,b]\).

**Proof.** With \(p \in P_{2n+1}\) being the Hermite interpolation polynomial of \(f\) on \(\{x_0, \ldots, x_n\}\), it holds that \(Q_n(f) = Q_n(p) = I(p)\), and so
\[ I(f) - Q_n(f) = I(f - p) = \int_a^b w(x)\pi_{n+1}^2(x) \frac{f^{(2n+2)}(\xi(x))}{(2n+2)!} \, dx = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} \int_a^b w(x)\pi_{n+1}^2(x) \, dx, \]
for some \(\xi \in [a,b]\), where we have used that \(w(x)\pi_{n+1}^2(x) \geq 0\) on \([a,b]\). □
2. Additions to Chapter 11 book

2.1. A upper bound for the error of best approximation from a spline space. For \( a = x_0 < x_1 < \ldots < x_m = b \), and \( n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), let

\[
S_n^{(-1)} := \{ f : [a, b] \to \mathbb{R} : f|_{[x_i, x_{i+1}]} \in \mathcal{P}_n, i = 0, \ldots, m-1 \}.
\]

For \( k \in \{0, \ldots, n\} \), we set

\[
S_n^{(k)} := C^k[a, b] \cap S_n^{(-1)}.
\]

Obviously \( \mathcal{P}_n(a, b) \subseteq S_n^{(n)} \subseteq \cdots \subseteq S_n^{(-1)} \). It holds that

\[
(2.1) \quad \dim S_n^{(k)} = (n+1)m - (k+1)(m-1) = m(n-k) + k + 1,
\]

where the subtraction of \((k+1)(m-1)\) corresponds to the loss of \(k+1\) degrees of freedom at each of the points \(x_1, \ldots, x_{m-1}\) as a consequence of the \(C^k\) constraint.

So in particular \( \dim S_n^{(n)} = n+1 \), meaning that \( S_n^{(n)} \) is simply equal to \( \mathcal{P}_n(a, b) \). From here on we exclude this non-interesting case, and consider \( k \in \{0, \ldots, n-1\} \).

In some books, any of these spaces \( S_n^{(k)} \) are called \textit{spline spaces}, and in other books the name \textit{spline} is used exclusively for functions from \( S_1^{(1)} \).

Remarkably, a bound as in \((2.2)\) is also valid for approximation from the much smaller ‘true’ spline space \( S_n^{(n-1)} \):

\[
(2.3) \quad \inf_{s \in S_n^{(n-1)}} \| f - s \|_\infty = O(h^{n+1}\|f^{(n+1)}\|_\infty)
\]

(assuming \( S_n^{(n-1)} := \mathbb{R} \)), \( n \in \mathbb{N} \) w.r.t. equidistant interpolation points \( x_i = y_0 < y_1 < \cdots < y_{n+1} = x_{i+1} \), then \( \| (f-s)|_{[x_i, x_{i+1}]} \|_\infty = O(h^{n+1}\|f^{(n+1)}\|_\infty) \).

Remarkably, a bound as in \((2.2)\) is also valid for approximation from the much smaller ‘true’ spline space \( S_n^{(n-1)} \):

\[
(2.3) \quad \inf_{s \in S_n^{(n-1)}} \| f - s \|_\infty = O(h^{n+1}\|f^{(n+1)}\|_\infty)
\]

(and thus also for approximation from any ‘intermediate space’ \( S_n^{(k)} \)). Note that \( \dim S_n^{(n-1)} = m + n \), which thinking of \( m \) ‘large’ and say \( n = 10 \), is \textit{only slightly} larger than \( \dim S_n^{(0)} = m + 1 \).

We do not provide a proof for \((2.3)\) in the general case. Instead, in Exer. 6 (first statement of part (f)), the statement is proven for \( n = 3 \). In this exercise, an interpolant \( s \in S_3^{(2)} \) is constructed that realizes this upper bound on the approximation error. Other than the continuous piecewise Lagrange interpolant in \( S_n^{(0)} \), or the \( C^1 \) piecewise cubic Hermite interpolant in \( S_3^{(1)} \), the construction of the interpolant in \( S_3^{(2)} \) is \textit{not} local, i.e., \( s|_{[x_i, x_{i+1}]} \) does \textit{not} depend exclusively on \( f|_{[x_i, x_{i+1}]} \).
Exer. 6. Let $I_1 : C[a,b] \rightarrow S_1^{(0)}$ the continuous piecewise linear interpolator (i.e., the mapping from a continuous function to its continuous piecewise linear interpolant), and let $I_3 : C^1[a,b] \rightarrow S_3^{(2)}$ the “complete cubic spline interpolator” defined by

$$(2.4) \quad s(x_i) = f(x_i) \quad (i \in \{0,\ldots,m\}),$$

$$(2.5) \quad s'(x_0) = f'(x_0), \quad s'(x_m) = f'(x_m),$$

where $s$ is here a shorthand notation for $I_3(f)$. Note that the number of conditions equals $\text{dim} S_3^{(2)}$.

The aim of this exercise is to show that $\|f - I_3(f)\|_{\infty} = \mathcal{O}(h^4\|f'''\|_{\infty})$ assuming $f \in C^4[a,b]$.

For $s \in S_3^{(2)}$, it holds that $s'' \in S_1^{(0)}$, so that for $i = 1, \ldots, m$,

$$s''_{|x_{i-1},x_i}(x) = \frac{(x - x_i)^3}{6h^3} \sigma_{i-1} + \frac{(x - x_{i-1})^3}{6h^3} \sigma_i + \alpha_i (x - x_{i-1}) + \beta_i (x_i - x),$$

where, for $i \in \{0, \ldots, m\}$, $\sigma_i := s''(x_i)$. By integrating this relation twice, we obtain

$$(2.6) \quad s_{|x_{i-1},x_i}(x) = \frac{(x - x_i)^4}{24h^3} \sigma_{i-1} + \frac{(x - x_{i-1})^4}{24h^3} \sigma_i + \frac{(x - x_{i-1})^2}{12h^2} \alpha_i + \frac{(x_i - x)^2}{12h^2} \beta_i,$$

for some scalars $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$.

By imposing (2.4) we obtain

$$\alpha_i = \frac{f(x_i)}{h} - \frac{h}{6} \sigma_i, \quad \beta_i = \frac{f(x_{i-1})}{h} - \frac{h}{6} \sigma_{i-1}.$$

(a) By using the continuity of $s'$ in $x_1, \ldots, x_{m-1}$ and (2.5), show that

$$A [\sigma_0 \ldots \sigma_m]^\top = b,$$

where $A \in \mathbb{R}^{(m+1) \times (m+1)}$ is defined by

$$A = \begin{bmatrix} 4 & 2 & \cdots & \cdots & \cdots & \cdots & 1 & 4 & 1 \\ 1 & 4 & 1 & & & & 2 & 4 \end{bmatrix},$$

and $b \in \mathbb{R}^{m+1}$ by

$$b_i = \left\{ \begin{array}{ll} 12 \left[ \frac{f(x_{i-1})-f(x_i)}{h^3} - \frac{f'(x_i)}{h^2} \right] & \text{when } i = 0, \\
6 \left[ \frac{f(x_{i+1})-2f(x_i)+f(x_{i-1})}{h^3} \right] & \text{when } i \in \{1, \ldots, m-1\}, \\
12 \left[ \frac{f'(x_m)}{h^2} - \frac{f'(x_{m-1})}{h^3} \right] & \text{when } i = m. \end{array} \right.$$  

Elementary linear algebra shows that $A$ is invertible, and that $\|A^{-1}\|_{\infty} \leq \frac{1}{2}$, i.e., that $\max_i |(A^{-1}x)_i| \leq \frac{1}{2} \max_i |x_i|$. (Indeed, writing $A = 4(I - (I - \frac{1}{4}A))$ and using that $\|I - \frac{1}{4}A\|_{\infty} = \frac{1}{2}$, shows that $\|A^{-1}\|_{\infty} \leq \frac{1}{2}$).

(b) Show that $\|I_3(f)'''\|_{\infty} \leq 3\|f'''\|_{\infty}$. (Hint: Show that for $s \in S_3^{(2)}$, $\|s''\|_{\infty} = \max_{0 \leq i \leq m} |s''(x_i)|$, and that $\max_{0 \leq i \leq m} |b_i| \leq 6\|f''\|_{\infty}$.)

(c) Show that $I_3$ is a projector, i.e., that $I_3(s) = s$ for any $s \in S_3^{(2)}$.

(d) Show that for any $p \in S_1^{(0)}$ there exists an $\bar{s} \in S_3^{(2)}$ with $\bar{s}'' = p$. 

(e) Let \( \tilde{s} \in S_3^{(2)} \) be such that \( \tilde{s}'' = I_1(f'') \). Show that \( f - I_3(f) = f - \tilde{s} - I_3(f - \tilde{s}) \), and with that, show that
\[
\|f'' - I_3(f)''\|_\infty \leq 4\|f'' - \tilde{s}''\|_\infty \leq \frac{1}{2} h^2 \|f^{(4)}\|_\infty.
\]
(f) Show that \( I_1(f - I_3(f)) = 0 \), and with that show that
\[
\|f - I_3(f)\|_\infty \leq \frac{1}{16} h^4 \|f^{(4)}\|_\infty,
\]
as well as
\[
\|f' - I_3(f')\|_\infty \leq \frac{1}{2} h^3 \|f^{(4)}\|_\infty.
\]
2.2. **Construction of a local basis for \( S_{n}^{(n-1)} \).** For storing functions from \( S_{n}^{(n-1)} \), or for computing the best approximation w.r.t. the (weighted) \( L_{2}(a,b) \)-norm of some function by an element of \( S_{n}^{(n-1)} \), one needs a basis of \( S_{n}^{(n-1)} \). Preferably this basis is local, meaning that the number of basis functions that are non-zero at a point \( x \in [a,b] \) is bounded uniformly in \( m \) (and \( x \)).

**Remark 2.1.** Other than for \( n = 1 \) (cf. [SM03, §11.3]), for \( n > 1 \), such a local basis that additionally is interpolating does not exist (recall: A basis is (Lagrange) interpolating, when for each basis function there exists a point in \( [a,b] \) in which it doesn’t vanish, but in which all other basis functions do vanish.)

A local basis for \( S_{n}^{(n-1)} \) is constructed in Exer. 7. It generalizes the construction of the basis from [SM03, §11.3] for \( n = 1 \) to \( n \in \mathbb{N}_{0} \).

**Exer. 7.** For convenience, let \( a = 0 \). With

\[
S_{(n)}(x) := \sum_{k=0}^{n+1} (-1)^{k} \binom{n+1}{k} (x - kh)^{n}_{+},
\]

see Figure 1, we define \( S_{(n,\ell)}(x) := S_{(n)}(x - \ell h) \) for \( \ell \in \mathbb{Z} \).

![Figure 1. The functions \( S_{(n)} \) for \( n = 0,\ldots,4 \).](image)

(a) Show that \( S_{(n,\ell)}|_{[0,b]} \in S_{n}^{(n-1)} \).

(b) Show that \( \text{supp} S_{(n,\ell)} \subseteq [\ell h, (\ell + n + 1)h] \).

(c) Show that \( \dim S_{n}^{(n-1)} = m + n = \# \{ \ell \in \mathbb{Z} : S_{(n,\ell)}|_{[a,b]} \neq 0 \} \).

(d) Show that \( S_{(n+1,\ell)}(x) = (n+1)(S_{(n,\ell)}(x) - S_{(n,\ell+1)}(x)) \) (when \( n = 0 \) only for \( x \notin h\mathbb{Z} \)).

(e) From [SM03, Exer. 11.6], we know that

\[
S_{(n+1,\ell)}(x) = (x - \ell h)S_{(n,\ell)}(x) + ((n + 2 + \ell)h - x)S_{(n,\ell+1)}(x).
\]

Using induction to \( n \), from this show that

\[
\sum_{\ell \in \mathbb{Z}} S_{(n,\ell)}(x) = h^n n!.
\]

Now we are going to show that for all \( p \in \mathbb{Z} \),

\[
(2.7) \quad \sum_{\ell \in \mathbb{Z}} c_{\ell} S_{(n,\ell)}|_{[\ell h, (\ell+1)h]} = 0 \Rightarrow c_{\ell} = 0 \text{ for } p - n - 1 < \ell < p + 1.
\]

(f) Show that (2.7) holds for \( n = 0 \).
(g) Now let (2.7) be valid for some \( n \in \mathbb{N}_0 \). Let \( \sum_{\ell \in \mathbb{Z}} c_{\ell} S_{(n+1,\ell)} \) and so \( \sum_{\ell \in \mathbb{Z}} c_{\ell} S'_{(n+1,\ell)} \) vanish on \((ph, (p + 1)h)\). Using (d), show that this implies that for some constant \( c \in \mathbb{R} \), \( c_{\ell} = c \) for all \( p - n - 2 < \ell < p + 1 \), and with that, that

\[
\sum_{\ell \in \mathbb{Z}} c_{\ell} S_{(n+1,\ell)} |_{(ph, (p + 1)h)} = c \sum_{\ell \in \mathbb{Z}} S_{(n+1,\ell)} |_{(ph, (p + 1)h)} = ch^{n+1}(n + 1)!
\]

Conclude that (2.7) is valid for \( n + 1 \), and so for any \( n \in \mathbb{N}_0 \).

(h) Using (a), (c), and (2.7), show that

\[
\{ S_{(n,\ell)} |_{[0,b]} : \ell \in \{-n, \ldots, m - 1\} \}
\]

is a basis for \( S_n^{(n-1)} \).
3. Additions to Chapter 12 book

From [SM03] we skip

- the proof of Picard’s Theorem, Thm 12.1.
- §12.3 with the exception of Definition 12.2
- §12.4

The reason to skip §12.4 is that an implicit (1-step) ODE solver cannot be written as $y_{n+1} = y_n + h\Phi(x_n, y_n; h)$ since $\Phi$ doesn’t have $y_{n+1}$ as one of its arguments (in the book they try to solve this by defining $\Phi$ in an implicit way, but this doesn’t lead to an analysis that is correct). Therefore, we replace §12.4 by the analysis of the trapezium rule given below in §3.1.

3.1. Trapeziumrule. Writing

$$y(x_{n+1}) - y(x_n) = \int_{x_n}^{x_{n+1}} f(x, y(x)) \, dx,$$

and approximating the integral by the trapezium rule leads to the following implicit one-step method

$$y_{n+1} = y_n + \frac{1}{2}h[f(x_n, y_n) + f(x_{n+1}, y_{n+1})].$$

For each $n = 0, \ldots, N - 1$, $y_{n+1}$ is given implicitly as the solution of an equation. The first question is whether this equation has a solution:

3.1.1. Existence.

**Lemma 3.1** (Banach’s fixed point theorem). Let $(X, d)$ be a non-empty complete metric space, and let $F : X \to X$ be a contraction, meaning that for some $K < 1$,

$$d(F(x), F(y)) \leq Kd(x, y) \quad (x, y \in X).$$

Then $\exists x \in X$ with $F(x) = x$.

**Proof.** Select $x_0 \in X$ arbitrarily, and define $(x_n)_{n \in \mathbb{N}} \subset X$ by $x_{n+1} = F(x_n)$. Then $d(x_{n+1}, x_n) \leq Kd(x_n, x_{n-1}) \leq \cdots \leq K^n d(x_1, x_0)$, and so for $m \geq n$,

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + \cdots + d(x_{n+1}, x_n) \leq (K^{m-1} + \cdots + K^n)d(x_1, x_0) \leq \frac{K^n}{1 - K} d(x_1, x_0).$$

So $(x_n)_n$ is a Cauchy sequence, and because $X$ is complete, therefore convergent, say with limit $x$. A consequence of $F$ being a contraction is that $F$ is continuous. So $x = \lim_{n \to \infty} x_{n+1} = \lim_{n \to \infty} F(x_n) = F(\lim_{n \to \infty} x_n) = F(x)$. Now let $z$ be another fixed point of $F$. Then $d(x, z) = d(F(x), F(z)) \leq Kd(x, z)$, and so $d(x, z) = 0$, or $x = z$. \hfill $\square$

Returning to the trapeziumrule, let $f$ be continuous on

$$D := [x_0, x_M] \times [y_0 - C, y_0 + C],$$

and Lipschitz continuous w.r.t. its second variable, i.e., for some constant $L > 0$,

$$|f(x, u) - f(x, v)| \leq L|u - v| \quad ((x, u), (x, v) \in D).$$

Set $Q := \max_{(x, u) \in D} |f(x, u)|$. Fixing $n$, let

$$y_n \in [y_0 - C + hQ, y_0 + C - hQ],$$
which interval is non-empty when $h \leq \frac{C}{Q}$. Then $F$ defined by $F(y) := y_n + \frac{1}{2}h[f(x_n, y_n) + f(x_{n+1}, y)]$ maps $[y_0 - C, y_0 + C]$ into $[y_0 - C, y_0 + C]$. Furthermore it holds that for $y, z \in [y_0 - C, y_0 + C]$, $|F(y) - F(z)| = \frac{1}{2}h|f(x_{n+1}, y) - f(x_{n+1}, z)| \leq \frac{1}{2}hL|y - z|$,\footnote{Truncation error.}

Applying the fixed point theorem, with $M := [y_0 - C, y_0 + C]$, $d(x, y) := |x - y|$, we conclude that (3.1) has a unique solution whenever $h$ is sufficiently small such that additionally $hL < 2$.

### 3.1.2. Approximation of $y_{n+1}$

Generally the solution $y_{n+1}$ of (3.1) cannot be determined exactly. An obvious way to approximate it is to apply a number of iterations $y^{(i+1)} = F(y^{(i)})$ starting with say $y^{(0)} = y_n$, or even better, $y^{(0)} = y_n + h f(x_n, y_n)$. It holds that

$$|y_{n+1} - y^{(i+1)}| = |F(y_{n+1}) - F(y^{(i)})| \leq \frac{1}{2} hL |y_{n+1} - y^{(i)}|, \tag{3.2}$$

Assuming that $f$ is $2 \times$ continuous differentiable as function of its second variable a much faster converging iteration is given by the Newton iteration, described in [SM03, §1.4] starting from Definition 1.6, applied to the equation $G(y) := y - y_n - \frac{1}{2}[f(x_n, y_n) + f(x_{n+1}, y)] = 0$. If $G''(y_{n+1}) = 1 - \frac{1}{2} \frac{df}{dy}(x_{n+1}, y) \neq 0$, which is the case when $h$ is sufficiently small, then for $y^{(0)}$ being sufficiently close to the solution $y_{n+1}$, the Newton iteration converges, and

$$\lim_{i \to \infty} \frac{y_{n+1} - y^{(i+1)}}{(y_{n+1} - y^{(i)})^2} = -\frac{G''(y_{n+1})}{2G'(y_{n+1})}, \tag{3.3}$$

i.e., quadratic convergence, being way more favourable than the ‘R–linear’ convergence (3.2).

### 3.1.3. Truncation error

Knowing that the recursion (3.1) is well-defined, and that we can approximate its solution at any desired accuracy, we study the accuracy of the approximations for $y(x_n)$ that are produced by it. First we define a truncation error:

$$T_n := \frac{y(x_{n+1}) - y(x_n)}{h} - \frac{1}{2} (f(x_{n+1}, y(x_{n+1})) + f(x_n, y(x_n)))$$

$$= \frac{y(x_{n+1}) - y(x_n)}{h} - \frac{1}{2}(y'(x_{n+1}) + y'(x_n)).$$

To determine the order of the trapezium rule, the usual (and recommended) way is to Taylor all terms in the right-hand side around some suitable point:

$$y(x_{n+1}) = y(x_n + \frac{1}{2}h) + \frac{1}{2} hy'(x_n + \frac{1}{2}h) + \frac{1}{2}(\frac{1}{2}h)^2 y''(x_n + \frac{1}{2}h) + \frac{1}{6}(\frac{1}{2}h)^3 y'''(\xi_1)$$

$$y(x_n) = y(x_n + \frac{1}{2}h) - \frac{1}{2} hy'(x_n + \frac{1}{2}h) + \frac{1}{2}(\frac{1}{2}h)^2 y''(x_n + \frac{1}{2}h) - \frac{1}{6}(\frac{1}{2}h)^3 y'''(\xi_2)$$

yielding

$$\frac{y(x_{n+1}) - y(x_n)}{h} = y'(x_n + \frac{1}{2}h) + \frac{1}{24} h^2 y'''(\xi_1) + y'''(\xi_2)$$

$$= y'(x_n + \frac{1}{2}h) + \frac{1}{24} h^2 y'''(\xi).$$
for some \( \xi \in [\xi_1, \xi_2] \subset [x_n, x_{n+1}] \) using the intermediate value theorem. From
\[
y'(x_{n+1}) = y'(x_n + \frac{1}{2}h) + \frac{1}{2}h y''(x_n + \frac{1}{2}h) + \frac{1}{2}(\frac{1}{2}h)^2 y'''(\eta_1)
y'(x_n) = y'(x_n + \frac{1}{2}h) - \frac{1}{2}h y''(x_n + \frac{1}{2}h) + \frac{1}{2}(\frac{1}{2}h)^2 y'''(\eta_2)
\]
we have
\[
\frac{1}{2}(y'(x_{n+1}) + y'(x_n)) = y'(x_n + \frac{1}{2}h) + \frac{1}{8}h^2 y'''(\eta)
\]
for some \( \eta \in [\eta_1, \eta_2] \subset [x_n, x_{n+1}] \) again using the intermediate value theorem. This yields
\[
(3.4) \quad T_n = \frac{1}{2} h^2 y'''(\xi) - \frac{1}{8} h^2 y'''(\eta) = -\frac{1}{12} h^2 y'''(\xi) + o(h^2),
\]
where we have assumed that \( y \in C^2 \). Consequently, the trapeziumrule has order 2.

An analysis that avoids the \( o(h^2) \)-term exploits the following trick:
\[
\int_{x_n}^{x_{n+1}} (x - x_n)(x - x_n)y'''(x) \, dx = (x - x_n)(x - x_n)y'''(x) \bigg|_{x_n}^{x_{n+1}} + \int_{x_n}^{x_{n+1}} (x_{n+1} + x_n - 2x)y''(x) \, dx
\]
\[
= (x_{n+1} + x_n - 2x)y'(x) \bigg|_{x_n}^{x_{n+1}} + \int_{x_n}^{x_{n+1}} 2y'(x) \, dx = 2h T_n.
\]
On the other hand, thanks to \( (x - x_{n+1})(x - x_n) \leq 0 \) on \( [x_n, x_{n+1}] \), it holds that
\[
\int_{x_n}^{x_{n+1}} (x - x_{n+1})(x - x_n)y'''(x) \, dx = y'''(\xi) \int_{x_n}^{x_{n+1}} (x - x_{n+1})(x - x_n) \, dx = -\frac{1}{6} h^3 y'''(\xi)
\]
for some \( \xi \in [x_n, x_{n+1}] \), showing that
\[
T_n = -\frac{1}{12} h^2 y'''(\xi).
\]

3.1.4. **Global error.** Subtracting
\[
y_{n+1} = y_n + \frac{1}{2} h (f(x_{n+1}, y_{n+1}) + f(x_n, y_n))
\]
from
\[
y(x_{n+1}) = y(x_n) + \frac{1}{2} h (f(x_{n+1}, y(x_{n+1})) + f(x_n, y(x_n))) + h T_n
\]
yields the following recursion for the **global error** \( e_n = y(x_n) - y_n \):
\[
e_{n+1} = e_n + \frac{1}{2} h (f(x_{n+1}, y(x_{n+1})) - f(x_{n+1}, y_{n+1}) + f(x_n, y(x_n)) - f(x_n, y_n)) + h T_n
\]
and so using the Lipschitz continuity of \( f \) w.r.t. the second variable with constant \( L \),
\[
|e_{n+1}| \leq |e_n| + \frac{1}{2} h L (|e_{n+1}| + |e_n|) + h |T_n|
\]
and so for \( h L < 2 \),
\[
(1 - \frac{1}{2} h L) |e_{n+1}| \leq (1 + \frac{1}{2} h L) |e_n| + h |T_n|
\]
giving
\[
|e_{n+1}| \leq \frac{1 + \frac{1}{2} h L}{1 - \frac{1}{2} h L} |e_n| + \frac{h |T_n|}{1 - \frac{1}{2} h L}.
\]
Using that \(e_0 = 0\), with \(M := \sup_{x \in [x_0, x_M]} |y''(x)|\) one infers that
\[
|e_{n+1}| \leq \sum_{j=0}^{n} \left(1 + \frac{1}{2}hL\right)^{n-j} \frac{h^jT_{j}}{1 - \frac{1}{2}hL} \leq \frac{h^3M}{1 - \frac{1}{2}hL} \sum_{j=0}^{n} \left(1 + \frac{1}{2}hL\right)^{n-j} \frac{1}{j!}.
\]

Now from \(\frac{1}{1 - \frac{1}{2}hL} = 1 + \frac{1}{2}hL + O(h^2)\), thus \(\frac{h^3M}{1 - \frac{1}{2}hL} = 1 + hL + O(h^2) \leq e^{hL + O(h^2)}\), and so \((1 + \frac{1}{2}hL)^n \leq e^{hL + O(h^2)}n = e^{nhL + O(h)} = e^{nhL}e^{O(h)} = e^{nhL}(1 + O(h))\) we infer that \(\sup_{nh \leq X_M - x_0} \left(1 + \frac{1}{2}hL\right)^n < \infty\). We conclude the following result:

**Theorem 3.2** (trapezium rule). Let \(f\) be continuous on \(D := [x_0, X_M] \times [y_0 - C, y_0 + C]\), and Lipschitz continuous w.r.t. its second variable with constant \(L > 0\), i.e.,
\[
|f(x, u) - f(x, v)| \leq L|u - v| \quad ((x, u), (x, v) \in D).
\]
Let the exact solution \(y \in C^3([x_0, X_M])\). Then \(e_n = O(h^2)\) uniform in \(n\) and \(h\) with \(nh \leq X_M - x_0\), and \(hL < 2\).

### 3.2. Homogeneous recursions.

Together with the next two subsections §3.3 and §3.4, this subsection replaces [SM03, §12.7-9]. In particular we provide a proof for Dahlquist’s Equivalence Theorem ([SM03, Thm. 12.5]).

As a preparation for the next subsection, we study solutions of homogeneous recursions of the form
\[
\alpha_k v_{n+k} + \ldots + \alpha_0 v_n = 0 \quad n = 0, 1, \ldots
\]
where \(\alpha_k \neq 0, \alpha_0 \neq 0\). In particular, we will be interested in finding conditions on the coefficients \(\alpha_i\) so that \(v_n\) remains bounded when \(n \to \infty\).

**Theorem 3.3.** Let \(z_1, \ldots, z_l (\neq 0)\) be the roots of the characteristic polynomial
\[
\rho(z) := \alpha_k z^k + \ldots + \alpha_0,
\]
where \(z_r\) has multiplicity \(m_r\), so that \(m_1 + \ldots + m_l = k\). Then \((v_n)_{n \geq 0}\) is a linear combination of
\[
\left\{\left(\frac{d^r}{dz^r} z^n \right)_{n \geq 0} : 1 \leq r \leq l, 0 \leq q \leq m_r - 1\right\}.\]

**Proof.** Because each solution of (3.5) is uniquely determined by \(v_0, \ldots, v_{k-1}\), these solutions span a linear space of dimension \(\leq k\). It remains to show that the \(k\) sequences given in (3.6) are solutions, and that they are linearly independent.

\[1\]In case of multiple roots, the basis used in [SM03, (12.41)] is slightly different.
For each root $z_r$, and $0 \leq q \leq m_r - 1$, it holds that
\[
\alpha_k \left( \frac{d^q}{dz^q} z^{n+k} \right)_{z=z_r} = \left( \frac{d^q}{dz^q} (z^n \rho(z)) \right)_{z=z_r} = \left( \sum_{p=0}^q \frac{q!}{p!} \left( \frac{d^p}{dz^p} z^n \right) \left( \frac{d^{q-p}}{dz^{q-p}} \rho(z) \right) \right)_{z=z_r} = 0,
\]
so that indeed $\left( \frac{d^q}{dz^q} z^n \right)_{z=z_r}$ satisfies the recursion.

For $z_1, \ldots, z_k$ being simple roots, using induction one can show that
\[
\begin{vmatrix}
  1 & 1 & \cdots & 1 \\
  z_1 & z_2 & \cdots & z_k \\
  \vdots & \vdots & \ddots & \vdots \\
  z_1^{k-1} & z_2^{k-1} & \cdots & z_k^{k-1}
\end{vmatrix} = \prod_{1 \leq r < s \leq k} (z_s - z_r) \neq 0.
\]
This means that \{$(z^n_i)^0_{0 \leq n \leq k-1}$, $(z^n_k)^0_{0 \leq n \leq k-1}$\} are linearly independent, and so are the infinite sequences. The matrix in the last displayed formula is known as the Vandermonde matrix.

For the general case, one can show, see e.g. [Kal84], that the determinant of the $k \times k$ generalised Vandermonde matrix matrix with columns given by the first $k$ elements of the sequences from (3.6) is equal to
\[
\prod_{1 \leq r < s \leq l} (z_s - z_r)^{m_r} \neq 0.
\]
so that again the infinite sequences are linearly independent. \qed

Theorem 3.3 shows that the solution $(v_n)_{n \geq 0}$ is a (unique) linear combination of the $k$ special solutions given in (3.6). The coefficients in this linear combination can be found by solving a $k \times k$ linear system obtained by equating the linear combination of the $k$ special solutions, restricted to their first $k$ entries, to $(v_0, \ldots, v_{k-1})$. In other words, denoting the $k$ special solutions as $(z^{(i)}_n)_{n \geq 0}$ for $1 \leq i \leq k$ (where thus $z^{(i)}_n = z^n_i$ in case all roots are simple), it holds that $(v_n)_{n \geq 0} = \sum_{i=1}^{k} \gamma_i (z^{(i)}_n)_{n \geq 0}$, where
\[
\begin{pmatrix}
  z^{(1)}_0 & \cdots & z^{(k)}_0 \\
  \vdots & \ddots & \vdots \\
  z^{(1)}_{k-1} & \cdots & z^{(k)}_{k-1}
\end{pmatrix}
\]
\[
\begin{pmatrix}
  \gamma_1 \\
  \vdots \\
  \gamma_k
\end{pmatrix}
\]
\[
B
\]
\[
\begin{pmatrix}
  v_0 \\
  \vdots \\
  v_{k-1}
\end{pmatrix}.
\]

**Corollary 3.4.** If, and only if, all roots $z$ of the characteristic polynomial $\rho$ satisfy (3.7) $|z| \leq 1$, with any of them with modulus 1 being simple (root condition), then
\[
\sup_{0 \neq (v_0, \ldots, v_{k-1}) \in \mathbb{R}^k} \frac{\sup_{n \geq k} |v_n|}{\max\{|v_0|, \ldots, |v_{k-1}|\}} < \infty.
\]

**Proof.** From $(v_n)_{n \geq 0} = \sum_{i=1}^{k} \gamma_i (z^{(i)}_n)_{n \geq 0}$ the sufficiency follows from $\max_{1 \leq n \leq k} |\gamma_n| \leq \|B^{-1}\| \max_{0 \leq n \leq k-1} |v_n|$, and the necessity follows by letting $(v_n)_{0 \leq n \leq k-1}$ run over the special solutions $(z^{(i)}_n)_{0 \leq n \leq k-1}$. \qed
3.3. Multi-step methods. We consider
\[
\begin{align*}
\left\{ \begin{array}{l}
y'(x) = f(x, y(x)) \quad x \in [x_0, X_M], \\
y(x_0) = y_0,
\end{array} \right.
\end{align*}
\]
and assume that the conditions of the Picard theorem are fulfilled on \( D = [x_0, X_M] \times [y_0 - C, y_0 + C] \).

For a given stepsize \( h \), we set \( x_n := x_0 + nh, n = 0, \ldots, N := \frac{X_M - x_0}{h} \).

Given starting values \( y_0, \ldots, y_{k-1} \subseteq [y_0 - C, y_0 + C] \), we study the multi-step
\[
\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f(x_{n+j}, y_{n+j}) \quad n = 0, 1, \ldots, N - k,
\]
where \( \alpha_k \neq 0, \alpha_0^2 + \beta_0^2 \neq 0 \). For \( k > 1 \), a common way to provide suitable \( y_1, \ldots, y_{k-1} \) is to start with \((k - 1)\)-steps of a Runge-Kutta method of sufficiently high order (cf. the forthcoming Thm. 3.10).

**Example 3.5.** An obvious to arrive at an approximation scheme of type (3.8) is to apply a \((k + 1)\)-point Newton-Cotes formula (e.g. the Simpson rule where \( k = 2 \)) to the right-hand side of
\[
y(x_{n+k}) - y(x_n) = \int_{x_n}^{x_{n+k}} y'(s) \, ds = \int_{x_n}^{x_{n+k}} f(s, y(s)) \, ds.
\]

**Remark 3.6.** One step methods as Forward Euler and the Trapeziumrule obviously are of type (3.8) where \( k = 1 \). This holds not true for Runge-Kutta methods. (Explicit) Runge-Kutta methods can be analysed using [SM03, Thm. 12.3].

Possibly after rescaling, for notational convenience in the following we assume that
\[
\alpha_k = 1,
\]
and define the first and second characteristic polynomial of the multi-step by
\[
\rho(z) := \sum_{j=0}^{k} \alpha_j z^j \quad \text{(cf. [SM03, Thm. 3.3])}, \quad \sigma(z) := \sum_{j=0}^{k} \beta_j z^j.
\]
For the moment (cf. however Sect. 3.4), we will assume existence and uniqueness of the solution \((y_n)_{0 \leq n \leq N}\), which isn’t obvious for the implicit case \( \beta_k \neq 0 \), and furthermore that \((y_n)_{0 \leq n \leq N} \subseteq [y_0 - C, y_0 + C]\).

3.3.1. Truncation error. In correspondence with earlier definitions for one-step methods, we define the truncation error \( T_n \) by
\[
T_n := \frac{\sum_{j=0}^{k} \alpha_j y(x_{n+j}) - \sum_{j=0}^{k} \beta_j f(x_{n+j}, y(x_{n+j}))}{\sigma(1)}.
\]
The division by \( \sigma(1) \) is made to make the definition independent of arbitrary scalings of (3.8) as the one we applied to make \( \alpha_k = 1 \). The current definition of \( T_n \) coincides with ones given earlier for the Forward Euler and trapezoidal rule. Later, cf. footnote 2, we will see that for a valid multi-step \( \sigma(1) \neq 0 \).

**Definition 3.7.** The multi-step is said to have order \( p \) when, for sufficiently smooth solutions \( x \mapsto y(x) \) on \([x_0, X_M]\), for \( n = 0, \ldots, N - k \), it holds that \(|T_n| = O(h^p)\).

(When \( p > 0 \), the method is called consistent).
The determination of the order of a multi-step can follow the same lines as in Sect. 3.1.3 for the trapezium rule. That is, using that
\[ T_n = \frac{h^{-1} \sum_{j=0}^{k} \alpha_j y(x_{n+j}) - \sum_{j=0}^{k} \beta_j y'(x_{n+j})}{\sigma(1)}, \]
the order can be determined by replacing each of the terms in the numerator by a Taylor expansion of sufficiently high order around some arbitrary point \( \bar{x} \in [x_n, x_{n+h}] \). When doing so, the leading \( h^{-1} \)– and \( h^0 \)–terms disappear iff
\[ \rho(1) = 0 \text{ and } h^{-1} \sum_{j=0}^{k} \alpha_j (x_{n+j} - \bar{x}) = \sigma(1), \]
respectively. Writing
\[ h^{-1} \sum_{j=0}^{k} \alpha_j (x_{n+j} - \bar{x}) = \sum_{j=0}^{k} \alpha_j + h^{-1}(x_n - \bar{x}) \sum_{j=0}^{k} \alpha_j = \rho'(1) - h^{-1}(x_n - \bar{x}) \rho(1), \]
we conclude that the method is consistent if and only if \(^2\)
\[ \rho(1) = 0 \text{ & } \rho'(1) = \sigma(1). \]

**Example 3.8.** Some examples of multi-step methods, where \( f(x_n, y_n) \) is abbreviated by \( f_n \), are the (Forward) Euler, implicit (or Backward) Euler, trapezium rule, Adams-Bashforth, Adams-Moulton methods, given by
\[
\begin{align*}
  y_{n+1} &= y_n + hf_n, & \rho(z) &= z - 1, & \sigma(z) &= 1, \\
  y_{n+1} &= y_n + hf_{n+1}, & \rho(z) &= z - 1, & \sigma(z) &= z, \\
  y_{n+1} &= y_n + \frac{1}{2}h(f_{n+1} + f_n), & \rho(z) &= z - 1, & \sigma(z) &= \frac{1}{2}(z + 1), \\
  y_{n+4} &= y_{n+3} + \frac{1}{24}h(55f_{n+3} - 59f_{n+2} + 37f_{n+1} - 9f_n), & \rho(z) &= z^3 - z, & \sigma(z) &= \frac{55z^3 - 59z^2 + 37z - 9}{24}, \\
  y_{n+3} &= y_{n+2} + \frac{1}{24}h(9f_{n+3} + 19f_{n+2} - 5f_{n+1} + f_n), & \rho(z) &= z^3 - z^2, & \sigma(z) &= \frac{9z^3 + 19z^2 - 5z + 1}{24},
\end{align*}
\]
respectively.

### 3.3.2. Global error
Subtracting (3.8) from (3.9), the latter written as
\[
\sum_{j=0}^{k} \alpha_j y(x_{n+j}) = \sum_{j=0}^{k} \beta_j f(x_{n+j}, y(x_{n+j})) + h \sigma(1) T_n,
\]
yields for the global error
\[ e_n := y(x_n) - y_n \]
the recursion
\[
\sum_{j=0}^{k} \alpha_j e_{n+j} = \sum_{j=0}^{k} \beta_j (f(x_{n+j}, y(x_{n+j})) - f(x_{n+j}, y_{n+j})) + h \sigma(1) T_n \quad n = 0, 1, \ldots, N-k.
\]
The Lipschitz continuity of \( f \) on \( D \) w.r.t. the second variable means that
\[
|f(x_{n+j}, y(x_{n+j}))-f(x_{n+j}, y_{n+j})| \leq L |e_{n+j}|,
\]
\(^2\) Later we will see that a multi-step can only be convergent when \( \rho \) has no multiple roots with modulus 1. In addition to consistency it means that \( \rho'(1) \neq 0 \), and so indeed \( \sigma(1) \neq 0 \).
or, in other words, that $f(x_{n+j}, y(x_{n+j})) - f(x_{n+j}, y_{n+j}) = \ell_{n+j}e_{n+j}$ for some $|\ell_{n+j}| \leq L$. By substituting this we obtain the recursion

$$
(3.10) \quad \sum_{j=0}^{k} \alpha_j e_{n+j} = h \sum_{j=0}^{k} \beta_j \ell_{n+j} + h \sigma(1) T_n \quad n = 0, 1, \ldots, N - k.
$$

We write this scalar $k$-step recursion as an 1-step recursion for $k$-vectors. With the $k$-vectors and the $k \times k$-matrices

$$
\begin{align*}
\hat{z}_n := & \begin{bmatrix} e_n \\ e_{n+1} \\ \vdots \\ e_{n+k-1} \end{bmatrix}, \quad \hat{t}_n := \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \quad E_n := \text{diag} [0 \cdots 0 \beta_k \ell_{n+k}], \\
D_n := & \begin{bmatrix} 0 \\ \vdots \\ [\beta_0 \ell_n \cdots \beta_{k-1} \ell_{n+k-1}] \\ 0 \end{bmatrix}, \quad A := \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\
& \begin{bmatrix} -\alpha_0 & -\alpha_1 & \cdots & -\alpha_{k-1} \end{bmatrix}
\end{align*}
$$

(the latter known as the companion matrix), (3.10) reads as

$$
(3.11) \quad (I - hE_n)\hat{z}_{n+1} = (A + hD_n)\hat{z}_n + h\hat{t}_n \quad n = 0, 1, \ldots, N - k.
$$

where we have used that $\alpha_k = 1$.

We consider stepsizes with

$$
(3.12) \quad hL|\beta_k| \leq \frac{1}{2}.
$$

Then $\|hE_n\| \leq \frac{1}{2}$, showing that $I - hE_n$ is invertible, $(I - hE_n)^{-1} = \sum_{i=0}^{\infty} (hE_n)^i$, $\|(I - hE_n)^{-1}\| \leq (1 - \|hE_n\|)^{-1} \leq 2$, and $\|I - (I - hE_n)^{-1}\| \leq \frac{\|hE_n\|}{1 - \|hE_n\|} = O(h)$.

Writing

$$
\hat{z}_{n+1} = (I - hE_n)^{-1}(A + hD_n)\hat{z}_n + h (I - hE_n)^{-1}\hat{t}_n,
$$

we infer

$$
\hat{z}_{n+1} = A_n \cdots A_0 \hat{z}_0 + h \sum_{j=0}^{n} A_n \cdots A_{j+1} \delta_j.
$$

We have $\|\delta_j\| \leq 2|\sigma(1)| |T_j|$ and from $(n + 1)h \leq X_M - x_0$, we conclude that

$$
\|\hat{z}_{n+1}\| \leq \left[\|\hat{z}_0\| + (X_M - x_0)2|\sigma(1)| \max_{0 \leq j \leq n} |T_j|\right] \max_{0 \leq j \leq n} \|A_n \cdots A_{j}\|.
$$

Writing $A_i$ as $A + F_i$, $F_i = hD_i + (I - hE_i)^{-1} - I)(A + hD_i)$ so that $\|F_i\| \leq \gamma h$ for some constant $\gamma$. We expand $(A + F_n) \cdots (A + F_j)$ into a sum of $(n - j + 1)$-fold products of matrices $A$ and $F_i$. For each $0 \leq \ell \leq n - j + 1$ there are $(n-j+1)^{\ell}$ terms in this sum with $\ell$ factors $F_i$, and so each of these terms contains at most $\ell + 1$ factors of the form $A^p$ for some $1 \leq p \leq n - j + 1$. We conclude that, with

$$
R_N := \max_{1 \leq p \leq N-k+1} \|A^p\|,
$$

we have

$$
\|\hat{z}_{n+1}\| \leq \left[\|\hat{z}_0\| + (X_M - x_0)2|\sigma(1)| \max_{0 \leq j \leq n} |T_j|\right] \max_{0 \leq j \leq n} \|A_n \cdots A_{j}\|.
$$
for \(0 \leq j \leq n \leq N - k\) it holds that
\[
\|(A + F_n) \cdots (A + F_j)\| \leq \sum_{\ell=0}^{n-j+1} \binom{n-j+1}{\ell} (h\gamma)^\ell R_N^\ell = R_N (1 + h\gamma R_N)^n = R_N e_h \gamma R_N, 
\]

\((3.14)\)

It remains to find conditions under which \(R_N\) is bounded uniformly in \(N \to \infty\), i.e., in \(h \downarrow 0\). Noting that \(\|A^n\| = \sup_{\nu \in \mathbb{R}} \|A^n \nu\|\), given \(\tilde{v}_0 = [v_0 \cdots v_{k-1}]^\top\), let \(\tilde{v}_p := A^p \tilde{v}_0\). The equivalence of (3.10) and (3.11) in the case \(\beta_0 = \ldots = \beta_k = 0\), and \(T_n = 0 \ (\forall n)\) shows that \(\tilde{v}_p = [v_p \cdots v_{p+k-1}]^\top\) where \((v_n)_{n \geq k}\) is defined by
\[(3.14)\]
\[\alpha_k v_{n+k} + \ldots + \alpha_0 v_n = 0 \quad n = 0, 1, \ldots \]

From Corollary 3.4 we conclude that if, and only if all roots \(z\) of \(\rho(z) = \sum_{j=0}^k \alpha_j z^j\) satisfy the root criterion (3.7), then \(R_N\) from (3.13) is bounded uniformly in \(N\), so that the following theorem is valid.

**Remark 3.9.** Despite being excluded in §3.2, this conclusion also holds true when for some \(\ell \geq 1\), \(\alpha_0 = \ldots = \alpha_{\ell-1} = 0\), as with the Adams-Bashforth and Adams-Moulton methods. In that case \(\rho\) has a root \(z = 0\) with multiplicity \(\ell\), and instead of being a \(k\)-step-recursion (3.14) is actually a \((k - \ell)\)-step recursion.

**Theorem 3.10.** Let \((y_n)_{0 \leq n \leq N} \subset \lbrack y_0 - C, y_0 + C \rbrack\) be a solution of the multi-step (3.8), and let (3.12) be valid. Then if, and generally only if the first characteristic polynomial \(\rho\) satisfies the root condition (3.7)\(^3\), there exists a constant \(R > 0\), such that
\[
|e_n| \leq R \left( \max\{|e_0|, \ldots, |e_{k-1}|\} + (X_M - x_0)|\sigma(1)| \max_{0 \leq j \leq n-k} |T_j| \right), 
\]

\((n = k, \ldots, N := \frac{X_M - x_0}{h})\), with the truncation errors \(T_j\) as defined in (3.9).

In particular, if in that case the method is of order \(p\), i.e. \(|T_j| = \mathcal{O}(h^p)\), and \(\max\{|e_0|, \ldots, |e_{k-1}|\} = \mathcal{O}(h^p)\), then \(\max_{0 \leq n \leq N} |e_n| = \mathcal{O}(h^p)\).

The statement of this theorem is often abbreviated by saying that a multi-step is **convergent** if and only if it is **consistent** and **stable**.

To show that the root condition in Theorem 3.10 is generally necessary, consider a consistent multi-step applied to the simple initial value problem \(\left\{ \begin{array}{ll} y'(x) = 0, & \ y(x_0) = y_0, \end{array} \right. \)

which has solution \(y(x) \equiv y_0\). Given \(y_1, \ldots, y_{k-1}\), the solution of the multi-step is the solution of the recursion \(\sum_{j=0}^k \alpha_j y_{n+j} = 0\) for \(n = 0, \ldots, N - k\). From \(T_n \equiv 0\), and \(\sum_{j=0}^k \alpha_j y_{n+j} = y_0 \sum_{j=0}^k \alpha_j = y_0 \rho(1) = 0\) we have \(\sum_{j=0}^k \alpha_j e_{n+j} = 0\) for \(n = 0, \ldots, N - k\), and Theorem 3.3 shows that the root condition is generally needed to guarantee boundedness of \((e_n)_{n \geq 0}\).

**3.4. Existence and uniqueness of discrete solutions in the implicit case.**

What is left to discuss is existence and uniqueness of a solution \((y_n)_{0 \leq n \leq N} \subset \lbrack y_0 - C, y_0 + C \rbrack\) of (3.8). Possibly by decreasing \(X_M\), we may assume that for some \(\varepsilon \in (0, C)\), \(|y(x) - y_0| \leq C - \varepsilon\) for \(x \in [x_0, X_M]\).

We proceed by induction. For some \(n \geq 0\), let \(y_0, \ldots, y_{n+k}\) exist uniquely in \([y_0 - C, y_0 + C]\). If the root condition is fulfilled, and for some \(p > 0\), \(\max\{|e_0|, \ldots, |e_{k-1}|\} = \mathcal{O}(h^p)\), then...
\(O(h^p)\), and \(|T_j| = O(h^p)\), then Thm. 3.10 shows that for some constant \(Q\), independent of \(n\), it holds that \(|e_j| \leq Qh^p, 1 \leq j \leq n + k\).

By subtracting (3.8) from (3.9) (the latter multiplied by \(\sigma(1)h\)), and by using that for \(j \leq k - 1\), \(f(x_{n+1+j}, y(x_{n+1+j})) - f(x_{n+1+j}, y_{n+j+1}) = \ell_{n+1+j}e_{n+1+j}\) all as before, existence and uniqueness of \(y_{n+k+1} \in [y(x_{n+k+1}) - \varepsilon, y(x_{n+k+1}) + \varepsilon]\) is equivalent to existence and uniqueness of a solution \(e_{n+k+1} \in [-\varepsilon, \varepsilon]\) of the fixed point equation

\[
e_{n+k+1} = h\beta_k(f(x_{n+k+1}, y(x_{n+k+1})) - f(x_{n+k+1}, y(x_{n+k+1}) - e_{n+k+1})) + \sum_{j=0}^{k-1} h\beta_j \ell_{n+1+j}e_{n+1+j} + h\sigma(1)T_{n+1} - \sum_{j=0}^{k-1} \alpha_j e_{n+1+j}.
\]

By the induction hypothesis, it holds that \(\Phi(0) = \sum_{j=0}^{k-1} (h\beta_j \ell_{n+1+j} - \alpha_j) e_{n+1+j} + h\sigma(1)T_{n+1} = O(h^p)\), so for \(h\) sufficiently small we have \(|\Phi(0)| < \varepsilon/2\). Furthermore, for \(\xi, \nu \in [-\varepsilon, \varepsilon]\), it holds that \(|\Phi(\xi) - \Phi(\nu)| \leq h|\beta_k|L|\xi - \nu| \leq \frac{1}{2}\varepsilon|\xi - \nu|\). From \(|\Phi(\xi)| \leq |\Phi(\xi)| + |\Phi(0)| \leq \frac{1}{2}\varepsilon|\xi| + \frac{\varepsilon}{2}\), it follows that \(\Phi: [-\varepsilon, \varepsilon] \to [-\varepsilon, \varepsilon]\). An application of Banach’s fixed point theorem (Lemma 3.1) shows that there exists a unique \(e_{n+k+1} \in [-\varepsilon, \varepsilon]\) with \(e_{n+k+1} = \Phi(e_{n+k+1})\), and thus that there exists a unique \(y_{n+k+1} \in [y(x_{n+k+1}) - \varepsilon, y(x_{n+k+1}) + \varepsilon] \subset [y_0 - C, y_0 + C]\) that solves (3.8).

**References**
