<u>Consequence:</u> Given a set of collocation points, we have a RK method.

[redacted]

Order (accuracy) of a collocation method

<u>Recall:</u> (Interpolation error)

Let $g: [t_0, t_0+h] \to \mathbb{R}^d$ and $p(t)$ be the interpolation polynomial of degree $s-1$ at nodes

$$\{(t_i, g(t_i))\}_{i=1}^{s}, \quad t_i = t_0 + c_i h.$$

Then $\exists\ \xi \in [t_0, t_0+h]$ s.t.

$$g(t) - p(t) = \frac{1}{s!}(t-t_1)(t-t_2)\cdots(t-t_s)\, g^{(s)}(\xi).$$

(<u>Proof:</u> Rolle's thm)

<u>Lemma 1:</u> Consider $y'(t) = f(t, y(t))$, $y(t_0) = y_0$,

f sufficiently smooth and Lipschitz continuous wrt second argument.

Let $u(t)$ be a collocation method on s points.

Then

corresponds to s in the defn of the collocation method.

$$\| y(t) - u(t) \| \leq C h^{s+1}, \quad t \in [t_0, t_0+h]$$

h small enough

$C > 0$ indep. of h.

<u>Rem:</u>
- This result shows that a collocation method is a "good" approximation of $y(t)$ in the whole interval $[t_0, t_0+h]$.
- If $t = t_0+h$, recall $u(t_0+h) = y_1 \Rightarrow |y(t_0+h) - y_1| \leq C h^{s+1}$

$\Rightarrow$ a collocation method on s points is at <u>least</u> of order s.

Proof: Let $p(t)$ be the interpolation polynomial of degree $s-1$ through $(t_i, f(t_i, y(t_i)))$, $i=1,2,\dots,s$
$t_i = t_0 + c_i h$

By definition:

Lagrange polynomial

$$p(t_0 + rh) = \sum_{j=1}^{s} \ell_j(r) f(t_j, y(t_j)) \qquad \text{(matches on } s \text{ points)}$$

And $\exists\, \xi \in [t_0, t_0+h]$ s.t.

$$\underbrace{f(t, y(t))}_{= y'(t)} - p(t) = \frac{1}{s!}(t-t_1)(t-t_2)\cdots(t-t_s) f^{(s)}(\xi, y(\xi)).$$

For $r \in [0,1]$:

$$y'(t_0+rh) - p(t_0+rh) = \frac{1}{s!}\overbrace{(t_0+rh-t_1)}^{=-c_1h}\cdots\overbrace{(t_0+rh-t_s)}^{=-c_sh} f^{(s)}(\xi, y(\xi))$$

$$= \frac{h^s}{s!}(r-c_1)\cdots(r-c_s) f^{(s)}(\xi, y(\xi))$$

$$=: h^s g(r)$$

Recall $u'(t_0 + rh) = \sum_{j=1}^{s} \ell_j(r) K_j$

see above $= y'(t_0+rh) - p(t_0+rh)$

$$y'(t_0+rh) - u'(t_0+rh) = h^s g(r) + p(t_0+rh) - u'(t_0+rh)$$

$$= h^s g(r) + \sum_{j=1}^{s} \ell_j(r) \underbrace{\left(f(t_j, y(t_j)) - K_j\right)}_{=: \Delta f_j} \qquad (*)$$

Recall $K_j = f(t_j, u(t_j))$

$$|\Delta f_j| = |f(t_j, y(t_j)) - f(t_j, u(t_j))|$$

$$\overset{\text{Lipschitz}}{\leq} L\, |y(t_j) - u(t_j)|$$

$$\leq L \max_{t \in [t_0, t_0+h]} |y(t) - u(t)|$$

Integrate (*):

$$y(t_0+rh)-u(t_0+rh) = \int_{t_0}^{t_0+h} (y'(s)-u'(s))\,ds$$

$$= h\int_0^r (y'(t_0+\tau h)-u'(t_0+\tau h))\,d\tau$$

$$= h^{s+1}\underbrace{\int_0^r g(\tau)\,d\tau}_{\leq C} + h\int_0^r \sum_{j=1}^{s} l_j(\tau)\,\Delta f_j\,d\tau$$

Hence

$$\max_{0\leq r\leq 1} |y(t_0+rh)-u(t_0+rh)| \leq C_1 h^{s+1} + C_2 h \max_{j=1,\dots,s} |\Delta f_j|$$

$$\overset{\text{see above}}{\leq} C_1 h^{s+1} + C_2 L h \underbrace{\max_{0\leq r\leq 1} |y(t_0+rh)-u(t_0+rh)|}_{\text{take this to the LHS, then factorize and isolate.}}$$

For h small enough

$$\Rightarrow \max_{0\leq r\leq 1} |y(t_0+rh)-u(t_0+rh)| \leq \frac{C_1}{\underbrace{1-C_2Lh}_{h \text{ small enough so that this term is positive}}} h^{s+1} \leq C h^{s+1}$$

#

Second approximation result

Relation between quadrature formulas and accuracy of collocation methods.

Lemma 2: The relations $a_{ij} = \int_0^{c_i} l_j(t)\,dt$, $b_i = \int_0^1 l_i(t)\,dt$, $i,j = 1,2,\dots,s$

are equivalent to

(conditions to be accurate in num. integration)

$$\begin{cases} C(q): & \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \quad k=1,2,\dots,q \\ B(q): & \sum_{i=1}^{s} \underset{\uparrow\,\text{weight}}{b_i}\, \underset{\uparrow\,\text{node}}{c_i^{k-1}} = \frac{1}{k}, \quad k=1,2,\dots,q \end{cases}$$

for $q=s$.

Proof: Exercises.

Rem: 1) If we choose $\{c_i\}_{i=1}^{s}$ properly, then $C(q)$, $B(q)$ can hold for $q > s$ (up to $q = 2s$).

[redacted]

Theorem 2: (Superconvergence, Guillou & Soulé 1969)

Let $f : [t_0, t_0+h] \times \mathbb{R}^d \to \mathbb{R}^d$ smooth and Lipschitz wrt second argument, and consider $y' = f(t,y)$, $y(t_0) = y_0$ with a collocation method $u(t)$ (defined on s nodes $0 \le c_1 < c_2 < \cdots < c_s \le 1$)

Assume that $B(p) : \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k}$, $k = 1, 2, \ldots, p$

holds for some $s \le p \le 2s$.

Then $|y(t_0+h) - u(t_0+h)| \le Ch^{p+1}$

$(= \mathcal{O}(h^{p+1})$ for $h \to 0)$

[redacted]

Preparation 1: (Quadratur formulas QF)

let $0 \le c_1 < c_2 < \cdots < c_s \le 1$ and $b_i \in \mathbb{R}$, $i = 1, 2, \ldots, s$

A QF is given by

$$\int_0^1 g(t)\,dt \approx \sum_{i=1}^{s} b_i g(c_i) .$$

We say that QF has order p if

$$\int_0^1 p(t)\,dt = \sum_{i=1}^{s} b_i\, p(c_i) \quad \forall\, p \in \mathbb{P}_{p-1} .$$

In this case, we have

$$\left| \int_{t_0}^{t_0+h} g(t)\,dt - h \sum_{i=1}^{s} b_i g(t_0 + c_i h) \right| \le C h^{p+1}$$

Preparation 2: (Variation of constants)

$$\begin{cases} \dot{w}(t) = A(t)\, w(t) + r(t) \\ w(t_0) = w_0 \end{cases}, \quad A(t) \ d \times d\text{-matrix}$$

Case 1: $A(t) = A$ independent of time:

$$w(t) = e^{(t-t_0)A} w_0 + \int_{t_0}^{t} e^{(t-s)A} r(s)\,ds$$

(Recall: $e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!}$) [redacted]

Case 2: $A(t)$ depends on time:

$$\Rightarrow w(t) = R(t, t_0)\, w_0 + \int_{t_0}^{t} R(t,s)\, r(s)\,ds$$

where the resolvent solves $\begin{cases} R'(t, t_0) = A(t)\, R(t, t_0) \\ R(t_0, t_0) = I \end{cases}$ identity matrix

[redacted]

AB

NIDS

Preparation 3: If $B(p)$ holds for $s \le p \le 2s$, then

$$\sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k} = \int_0^1 x^{k-1}\,dx \quad , k = 1,2,\dots,p.$$

$\Rightarrow$ given a fct g: $\sum_{i=1}^{s} b_i g(c_i)$ is a QF of order p.

We have now set the stage to prove Thm 2.

Proof:

① **Defect**

Denote $\delta(t) = u'(t) - f(t,u(t))$, notice that $\delta(t_0 + c_i h) = 0$ (only on collocation points) $i = 1,2,\dots,s$.

② **Linearisation:**

$$\underbrace{u'(t) - y'(t)}_{=w'(t)} = f(t,u(t)) - f(t,y(t)) + \delta(t)$$

$$= \underbrace{\frac{\partial f}{\partial y}(t,y(t))}_{=A(t)} \underbrace{(u(t) - y(t))}_{=w(t)} + r(t) + \delta(t)$$

$$= A(t)\, w(t) + r(t) + \delta(t)$$

where $\|r(t)\| \le C \|u(t) - y(t)\|^2 \le C h^{2(s+1)}$ by lemma 1

③ **Variation of constant formula:**

Denote $w(t) = u(t) - y(t)$, by Prep. 2:

$$\Rightarrow w(t_0+h) = R(t_0+h, t_0) \underbrace{w_0}_{=u(t_0)-y(t_0)=0} + \int_{t_0}^{t_0+h} R(t_0+h, s)(r(s) + \delta(s))\,ds$$

$$= \underbrace{\int_{t_0}^{t_0+h} R(t_0+h,s)\, r(s)\,ds}_{=I_1} + \underbrace{\int_{t_0}^{t_0+h} R(t_0+h,s)\,\delta(s)\,ds}_{=I_2}$$

$$\Rightarrow \|u(t_0+h) - y(t_0+h)\| \le \|I_1\| + \|I_2\|$$

by $\triangle$ inequality

④ Bound the remainders:

$$\|I_1\| \leq \int_{t_0}^{t_0+h} \| R(t_0+h, s)\, r(s)\| \, ds$$

$$\leq h \underbrace{\max_{s\in[t_0,t_0+h]} \|R(t_0+h,s)\|}_{\leq C} \underbrace{\max_{s\in[t_0,t_0+h]} \|r(s)\|}_{\overset{②}{\leq} C h^{2(s+1)}}$$

(matrix with n solutions which are bounded on a finite interval)

$$\leq \hat{C} h \cdot h^{2(s+1)} = \underbrace{\hat{C} h^2}_{\leq 1} \cdot \underbrace{h^{2s+1}}_{\leq h^{p+1}} \leq \hat{C} h^{p+1}.$$

because $p \leq 2s$

$$\|I_2\| = \left\| \int_{t_0}^{t_0+h} R(t_0+h, s)\, \delta(s)\, ds \right\|$$

$$\leq \left\| h \sum_{i=1}^{s} b_i R(t_0+h, t_0+c_i h) \overbrace{\delta(t_0+c_i h)}^{=0} \right\|$$

(no defect on the collocation points)

Δ inequality

(± same term)

$$+ \left\| \int_{t_0}^{t_0+h} R(t_0+h, s)\, \delta(s)\, ds - h \sum_{i=1}^{s} b_i R(t_0+h, t_0+c_i h)\, \delta(t_0+c_i h) \right\|$$

$$\overset{\text{Prep. 1\&3 (QF hypothesis)}}{\leq} C h^{p+1}.$$

The result follows.

#

Methods exploiting superconvergence (Guillou & Soulé thm)

We will mainly pick four classes of collocation methods:

(Names borrowed from quadrature theory)

1) Gauss collocation method of order $2s$ — Geometric Integration
2) Radau $2s-1$ — stiff problems
3) Lobatto $\to$ III A, $\to$ III B $2s-2$ — Geometric Integration.

Rem: If we apply Gauss or Lobatto to stiff problems $\to$ bad choice
Idem Radau and geometric integration.

1) Gauss methods: Let c_i, $i=1,2,\dots,s$ be the zeros of the (shifted) Legendre polynomials

$$p_s^G(x) = \frac{d^s}{dx^s}\left(x^s(1-x)^s\right).$$

Define b_i, a_{ij}

$$(*)\quad \begin{cases} B(s) & \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k}, \quad k=1,\dots,s \quad \text{or } b_i = \int_0^1 l_i(\tau)\,d\tau \\ C(s) & \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \quad k=1,\dots,s,\ i=1,\dots,s \quad \text{or } a_{ij} = \int_0^{c_i} l_j(\tau)\,d\tau \end{cases}$$

Result: (see exercises)

$B(p)$ is satisfied for $p=2s$

$\Rightarrow$ collocation method is of order $2s$.

Ex: $s=1 \Rightarrow c_1 = \frac{1}{2}$, $b_1 = 1$, $a_{11} = \frac{1}{2}$

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array}$$

implicit midpoint rule
good for geometric integration.

2) <u>Radau methods:</u> let c_i, $i = 1, 2, \ldots, s$ be the zeros of

$$\phi_s^R(x) = \frac{d^{s-1}}{dx^{s-1}}\left(x^{s-1}(1-x)^s\right) \Rightarrow c_s = 1 \text{ is a zero}$$

b_i, a_{ij} given by (*)

<u>Result:</u> $B(p)$ is satisfied for $p = 2s-1$.

<u>Ex:</u> $s = 1 \Rightarrow c_1 = 1,\ a_{11} = 1,\ b_1 = 1$

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

Implicit Euler method
good for stiff problems.

3) a) <u>Lobatto III A methods:</u> let c_i, $i = 1, 2, \ldots, s$ be the zeros of

$$\phi_s^L(x) = \frac{d^{s-2}}{dx^{s-2}}\left(x^{s-1}(1-x)^{s-1}\right) \Rightarrow c_s = 1,\ c_1 = 0 \text{ are zeros}$$

b_i, a_{ij} given by (*)

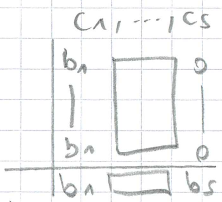
<u>Result:</u> $B(p)$ is satisfied for $p = 2s-2$.

<u>Ex:</u> $s = 2 \Rightarrow c_1 = 0,\ c_2 = 1 \Rightarrow$

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}$$

Implicit trapezoidal rule

b) <u>Lobatto III B methods:</u> $c_1, c_2, \ldots, c_s$ are defined as for Lobatto III A
b_1, b_s given as for Lobatto III A.



given, □ and ▭ to be found.

a_{ij}, b_i are given by $\begin{cases} a_{i1} = b_1 & \forall i \\ a_{is} = 0 & \forall i \end{cases}$

remaining a_{ij} given by $C(s-2)$, $b_2, \ldots, b_{s-1}$ are given by $B(s-2)$
(see exercises)

Unfortunately, Lobatto III B are <u>not</u> collocation methods (and thus the thm Guillou & Soulé cannot be applied).