# EXAM OF ADVANCED NUMERICAL METHODS, 5/22/2014.
# FULL RESOLUTION

Below is a full, mostly handwritten, answer to the exam. It is much more complete than what can be expected from the student in a time-constrained real exam situation. Far less complete answers are enough for the student to obtain all the points in each section.

The main aim of this detailed resolution is for future students to learn from it.

# ANM, EXAM OF 5/22/2014

## Part 1, Exercise 1

February = month 2 $\Rightarrow f(2) = 112,\ f'(2) = 0$ (max.)

August = month 8 $\Rightarrow f(8) = 49,\ f'(8) = 0$ (min.)

May = month 5 $\Rightarrow f(5) = 76$

With results rounded off to 6 decimals:

A) We need a table of divided differences (not finite differences), regardless of whether the nodes are equally spaced or not, because we have data derivatives, so the polynomial will be an osculating one):

Table of divided differences with repetitions.

| $i$ | $z_i$ | $f_{i,0}$ | $f_{i,1}$ | $f_{i,2}$ | $f_{i,3}$ | $f_{i,4}$ |
|---|---|---|---|---|---|---|
| 0 | 2 | 112 | — | — | — | — |
| 1 | 2 | 112 | $\underline{\underline{0}}$ | — | — | — |
| 2 | 5 | 76 | −12 | −4 | — | — |
| 3 | 8 | 49 | −9 | 0.5 | 0.75 | — |
| 4 | 8 | 49 | $\underline{\underline{0}}$ | 3 | 5/12 | $-0.0\hat{5}$ |
| 5 | 3 | 104 | −11 | 2.2 | 0.4 | $-0.01\hat{6}$ |

only for section B)

$0.03\hat{8} = \frac{7}{180}$

$\frac{76-112}{5-2} = -12 \quad \frac{49-76}{8-5} = -9$ ; $\underline{\underline{0}}$ and $\underline{\underline{0}}$ because $f'(2) = f'(8) = 0$.

$\frac{-12-0}{5-2} = -4 \quad \frac{-9--12}{8-2} = 0.5 \quad \frac{0--9}{8-5} = 3 \quad \frac{0.5--4}{8-2} = 0.75$

$\frac{3-0.5}{8-2} = 5/12 \quad \frac{5/12-0.75}{8-2} = -0.0\hat{5}$

Interpolation polynomial (Newton representation):

$$\boxed{P_4(x) = 112 - 4(x-2)^2 + 0.75(x-2)^2(x-5) - 0.0\hat{5}(x-2)^2(x-5)(x-8)}$$

$-0.055556 = -1/18$

Evaluate it optimally for April, i.e. for $x=4$:

To evaluate it "optimally" (from all 3 points of view, namely computational cost, roundoff error propagation and memory storage) we have to use the Hörner-like algorithm:

$$p_4(x) = \left\{\left[-0.0\hat{5}(x-8)+0.75\right](x-5)-4\right\}(x-2)^2+112$$

$$p_4(4) = \underbrace{\left\{\underbrace{\left[\underbrace{\underbrace{-0.0\hat{5}(4-8)}_{0.\hat{2}}+0.75}_{0.97\hat{2}}\right](4-5)-4}_{-4.97\hat{2}}\right\}(4-2)^2}_{-19.\hat{8}}+112 =$$

$92.\hat{1}$ = $\underline{92.111111 \text{ kW}}$ (again with 6 decimals)

B) The easiest way to estimate the error made in A) is to use the new datum, $f(3)=104$, to calculate a better estimation of $f(4)$, i.e. $p_5(4)$. Since this is our best estimation of $f(4)$, our best estimation of the error made in A) must be $p_5(4) - p_4(4)$, which is the term $h_5(x)$ that must be added to $p_4(x)$ in order to obtain $p_5(x)$, evaluated at $x=4$.

To write $h_5(x)$, we add a new row to our previous table of divided differences (with pencil on p. 1; the numbers are calculated exactly like the previous rows):

$$h_5(x) = 0.03\hat{8}\,(x-2)^2(x-5)(x-8)^2$$

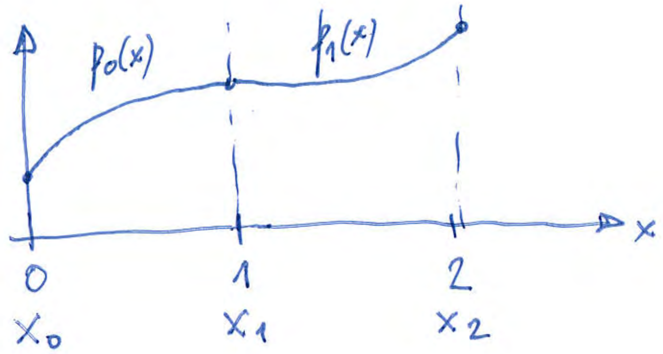
$$e_4(4) = f(4) - p_4(4) \simeq p_5(4) - p_4(4) = h_4(4) =$$

$$= \frac{7}{180}\,4\,(-1)\,(-4)^2 = -2.4\hat{8} = \boxed{-2.488889\ \mathrm{kW}}$$

(with 6 decimals)

It is important to know that, the error being negative, A is in excess, i.e. we now estimate that $92.\hat{1}$ was too large by about 2.5 kW.

$$\boxed{\text{Part 1, Exercise 2}}$$




Continuity at $x=1$:

$$\alpha + \gamma = -\alpha + \beta - 5\alpha + 1 \Rightarrow \underline{7\alpha - \beta + \gamma = 1}$$

Continuity of $s'(x)$ at $x=1$:

$$\left[3\alpha x^2 + \gamma = -3\alpha x^2 + 2\beta x - 5\alpha\right]_{x=1}$$

$$3\alpha + \gamma = -3\alpha + 2\beta - 5\alpha \Rightarrow \underline{11\alpha - 2\beta + \gamma = 0}$$

Continuity of $s''(x)$ at $x=1$:

$$\left[6\alpha x = -6\alpha x + 2\beta\right]_{x=1}$$

$$6\alpha = -6\alpha + 2\beta \quad \Rightarrow \quad \underline{6\alpha - \beta = 0}$$

We have a 3×3 linear system. To solve "by hand" it is often convenient to use Gauss's method (or a variant):

$$\begin{pmatrix} 7 & -1 & 1 & 1 \\ 11 & -2 & 1 & 0 \\ 6 & -1 & 0 & 0 \end{pmatrix} \langle F_2 - F_1 \rangle \begin{pmatrix} 7 & -1 & 1 & 1 \\ 4 & -1 & 0 & -1 \\ 6 & -1 & 0 & 0 \end{pmatrix} \langle F_3 - F_2 \rangle \begin{pmatrix} 7 & -1 & 1 & 1 \\ 4 & -1 & 0 & -1 \\ 2 & 0 & 0 & 1 \end{pmatrix}$$

Backward substitution:

$$\alpha = 1/2 \rightarrow \beta = 4\alpha + 1 = 3 \rightarrow \gamma = 1 + \beta - 7\alpha = \frac{2+6-7}{2} = 1/2$$

$$\boxed{\alpha = \frac{1}{2}, \ \beta = 3, \ \gamma = \frac{1}{2}}$$

The justification is obviously that a cubic spline $s(x)$ must be $s \in C^2([a,b])$ as part of its definition ("class 2").

If the spline is natural, it will satisfy $s''(0) = s''(2) = 0$:

$$6\alpha x = 3x\Big]_{x=0} = 0 \quad \underline{\underline{ok}}$$

$$-6\alpha x + 2\beta = -3x + 6\Big]_{x=2} = -6+6=0 \quad \underline{\underline{ok}}$$

Therefore <u>it can be a natural spline.</u>

Can it be a spline with boundary conditions?

Yes, if they coincide with $s'(0)$ and $s'(2)$ respectively:

$$s'(0) = 3\alpha x^2 + \gamma\Big]_{x=0} = \frac{1}{2}$$

$$s'(2) = -3\alpha x^2 + 2\beta x - 5\alpha\Big]_{x=2} = -3\,\frac{1}{2}\,4 + 2\cdot 3\cdot 2 - 5\,\frac{1}{2} =$$

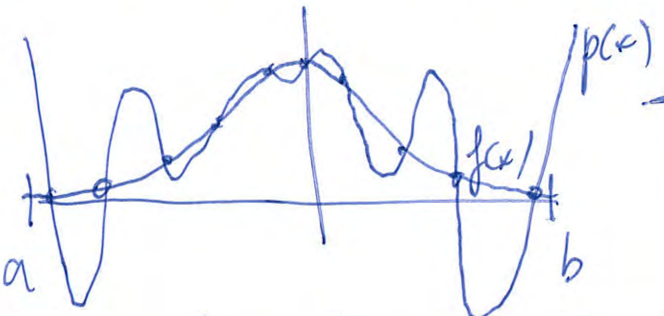
$$= -6+12-\frac{5}{2} = \frac{12-5}{2} = \frac{7}{2}$$

Therefore <u>it can also be a spline with boundary conditions $f'(0) = 1/2$, $f'(2) = 7/2$.</u>

## Part 1, Exercise 3

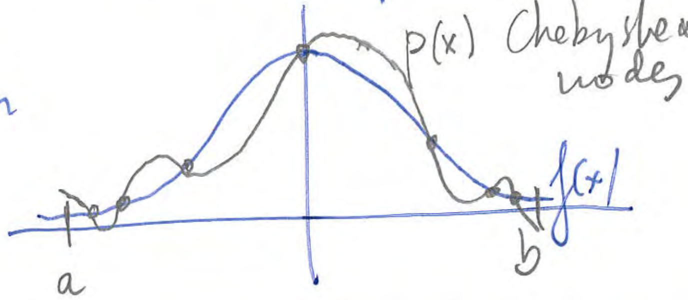
The Runge effect is the spurious (not representing the true nature of $f(x)$) oscillations that polynomials of high degree (even moderate) typically exhibit with respect to the function they are interpolating.

A well known example happens when $f(x)$ has the shape of a bell, like $e^{-x^2}$ or maybe $\frac{1}{1+x^2}$. The aspect would be like:




← see Classroom Notes for better figures.

This would be a typical aspect with equally spaced nodes. The effect can be (almost) minimized by choosing Chebyshev nodes rather than equally-spaced ones. That is because the Runge effect tends to be stronger near the endpoints of the interval of interpolation $[a, b]$, and the Chebyshev nodes are closer to one another near a and b; they are more spaced towards the center, but there the Runge effect tends to be weak. The aspect with Chebyshev nodes could be:



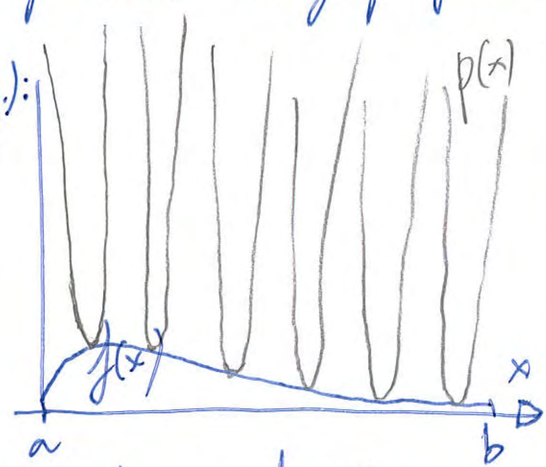

Observe that the maximum error between nodes tends to be similar (and it would be the same, i.e. achieved several times, if a certain derivative of $f(x)$ were constant).

It is important to note 3 things:

- The Runge effect is not the result of roundoff errors.

It is truncation errors $e(x) = f(x) - p(x)$ we are talking about, i.e. with exact arithmetic.

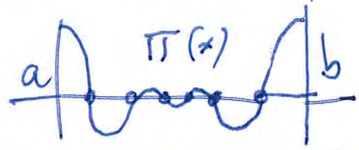
- Osculating polynomials do <u>not</u> present a better behavior. The Hermite polynomial of a plane's wing profile could look something like this (e.g.):




(See better fig. in Classroom Notes).

- The Runge effect is one of the main reasons why <u>polynomials of high degree are very rarely used</u>. Even moderate degrees, like 15 or 20, should be avoided. This is also a good <u>motivation to use splines</u>, which minimize oscillations under certain conditions (and way to measure them).

P.S. I forgot to add that $e(x) = f[x_0, x_1, \ldots, x_n, x]\, \Pi(x)$ closely resembles the aspect of $\pi(x)$ if $f[\ldots]$, related with $f^{(n+1)}(\xi)$ by a factor $\frac{1}{(n+1)!}$, is approx. constant; so the aspect of $\Pi(x)$ is the reason why oscillations are stronger near $a$ and $b$:




$\Pi(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$

**Part 1, Exercise 4**

N.B. L = natural log = log(x) in Matlab.

$I = \int_2^3 \frac{Lx}{\sqrt{(x-2)(3-x)}}\,dx$ looks like, after the linear change of variable mapping $[-1,1]$ onto $[2,3]$, we will have an integral of the Gauss-Chebyshev type, i.e. with some weight function $w(x) = \frac{1}{\sqrt{1-x^2}}$. (Even if we do not, we still have to make the change of variable, because both limits of integration are finite; the "worst" that can happen is that we have to integrate using Gauss-Legendre rules, which require either the use of tables or the calculation of Legendre polynomials → their roots (= nodes) → → the weights).

Hence, $x = 2.5 + \frac{\xi}{2}$ $\begin{cases} x(0) = 2.5 \text{ ok} \\ x(-1) = 2 \text{ ok} \\ x(1) = 3 \text{ ok} \\ \text{linear ok} \end{cases}$

$dx = \frac{d\xi}{2}$

$$I = \int_{-1}^{1} \frac{L(2.5+\xi/2)\,\frac{d\xi}{2}}{\sqrt{(0.5+\xi/2)(0.5-\xi/2)}} = \int_{-1}^{1} \frac{L(2.5+\xi/2)}{\sqrt{0.25-\xi^2/4}}\,\frac{d\xi}{2}$$

$$= \int_{-1}^{1} \frac{\overbrace{L(2.5+\xi/2)}^{f(\xi)}}{\sqrt{1-\xi^2}}\,d\xi = \int_{-1}^{1} \frac{f(x)\,dx}{\sqrt{1-x^2}}$$

$\xi$, $x$ are dummy variables.

So yes, it's Gauss-Chebyshev, which we can do without tables because the roots of the Chebyshev polynomials (i.e. the nodes) are easy to obtain, and the sum of the weights (coefficients of the quadrature rule) is always $\pi$ (and they are all equal, which reminds of ...)

The plan is to use 1 node, 2 nodes, 3, ... etc. until the distance (difference in absolute value) between the last two values obtained does not exceed 0.01% of the last one (which is our best estimation of $I$ at that moment). This does not guarantee that the precision is better than 0.01%, but it is very likely it will. We are therefore talking about a termination criterion of iterations, rather than a precision (strictly speaking).

With one node $x_0 = 0$, $w_0 = \pi$ $(n=0)$: $\quad f(x) = L\left(2.5 + \frac{x}{2}\right)$

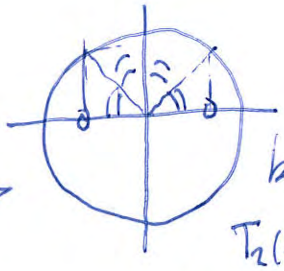
$$I \simeq Q_0 = w_0 f(x_0) = \pi\, L(2.5 + 0/2) = 2.878612$$

N.B. I will write all numbers rounded off to 6 decimals, but the internal operations will be with double precision arithmetic.

With 2 nodes $(n=1)$:

Weights $w_0 = w_1 = \pi/2$

Nodes $\pm \cos\frac{\pi}{4} = \pm \sqrt{2}/2$



because $T_n(t) = \cos(n \arccos t)$

$T_2(t) = \cos(2 \arccos t)$;

for $t = \cos\frac{\pi}{4}$, $\cos(2\frac{\pi}{4}) = 0$ ok

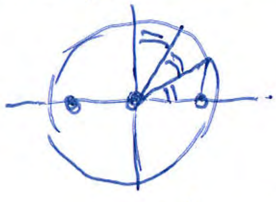
$$I \simeq Q_1 = w_0 f(x_0) + w_1 f(x_1) = \frac{\pi}{2} L\left(2.5 + \frac{-\sqrt{2}/2}{2}\right) + \frac{\pi}{2} L\left(2.5 + \frac{\sqrt{2}/2}{2}\right) =$$

$$= 2.846878$$

Termination criterion:

$$\left|\frac{Q_1 - Q_0}{Q_1}\right| \times 100 = 1.1147\% > 0.01\% \Rightarrow \text{show must go on.}$$

With 3 nodes $(n=2)$:



Weights $w_i = \pi/3$

Nodes $\left\{\cos\left(\frac{\pi/2}{3}\right), 0, -\cos\left(\frac{\pi/2}{3}\right)\right\} = \{-0.866025, 0, 0.866025\}$

$$Q_2 = \frac{\pi}{3} L\left(2.5 + \frac{-0.866}{2}\right) + \frac{\pi}{3} L\left(2.5 + \frac{0}{2}\right) + \frac{\pi}{3} L\left(2.5 + \frac{0.866}{2}\right) =$$

$= Q_2 = 2.8467154243$

Termination criterion: $\left|\frac{Q_2 - Q_1}{Q_2}\right| \times 100\% = 0.005707\% < 0.01\%$

$\Rightarrow \boxed{I \simeq Q_2 = 2.846715}$

N.B. The exact value turns out to be $I = \pi L\left(\frac{5}{4} + \sqrt{\frac{3}{2}}\right) =$

$= 2.8467143113428\ldots$ so the error made with $Q_2$ is

$\simeq -1.113 \times 10^{-6}$ (i.e. in excess).

## Part 1, Exercise 5

This is basically exercise 35 of the Classroom Notes, thoroughly discussed there, with slightly different phrasing of the questions. See lengthy discussion there.

A) Since $E = K \cdot f^{(4)}(\xi)$, it must be exact in $\mathbb{P}_3$ ($1, x, x^2, x^3$ all have zero fourth derivative) but not in $\mathbb{P}_4$ ($f = x^4 \Rightarrow f^{(4)} \equiv 24 \neq 0 \Rightarrow E \neq 0 \Rightarrow Q$ inexact) $\Rightarrow$ the polynomial degree is exactly $N = 3$.

Exact in $\mathbb{P}_2$ with 3 nodes $\Rightarrow$ formula of interpolatory kind.

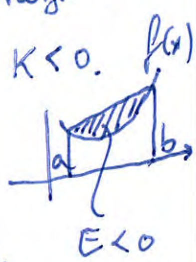
Inexact in $\mathbb{P}_5$ w/ 3 nodes $\Rightarrow$ not a Gauss one.

The extra unit in polynomial degree (exact not only in $\mathbb{P}_2$, like any interpolatory rule of 3 nodes "by construction", but also in $\mathbb{P}_3$) is typical, but not exclusive, of Newton-Cotes rules of an odd number of nodes. Not exclusive because any rule with 3 nodes symmetrically located on both sides of midpoint $m = (a+b)/2$ will be exact in $\mathbb{P}_3$ (and even if the nodes are not symmetrical, but such that $\int_a^b \pi(x)\,dx = 0$). There are infinitely many rules of 3 nodes that are exact in $\mathbb{P}_3$ that are not Newton-Cotes ones.

Another clue is that $h$ appears in the term $E$, and $h$ is typically the distance between nodes in Newton-Cotes rules. So it looks like a N-C rule. However, we are told to "reason out" according to the "properties" (fundamental ones) of the rule. Saying "h appears, so it is N-C" is reasoning according to a notation. And not even a universal one at that, since for many non-N-C rules, like Gauss ones, $h$ can be the total width $h=b-a$.

Therefore we can only <u>be sure</u> that it is a <u>non-Gauss interpolatory rule</u>; <u>possibly (but not necessarily) a Newton-Cotes one</u>.

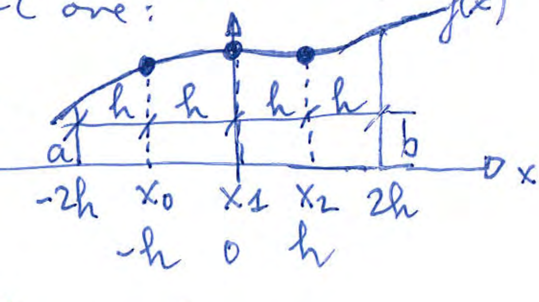
If the rule is indeed a N-C one, it must be the open one, because the constant K in the error term ($14h^5/45$) is positive; closed rules have a negative constant in the error term. For instance, the trapezoidal rule (closed) has error in excess (i.e. negative) for functions that are concave upwards (positive $f''$), hence $K<0$.




Therefore <u>it looks like the open Newton-Cotes rule of 3 nodes</u>.

But I cannot assure it, at least without calculating its error term E and seeing if it is precisely $14h^5/45 \cdot f^{(4)}(\xi)$. (And, strictly speaking, not even then, because the meaning of $h$ is not specified.)

B) To obtain the coefficients I need the positions of the nodes, so I will assume the rule is the open N-C one:




Exact in 1 $\rightarrow \int_{-2h}^{2h} 1\,dx = 4h = A_0 1 + A_1 1 + A_2 1$

$\Rightarrow \underline{A_0 + A_1 + A_2 = 4h}$

Exact in x $\rightarrow \int_{-2h}^{2h} x\,dx = 0 = A_0(-h) + \cancel{A_1 0} + A_2 h \Rightarrow \underline{A_0 = A_2}$

Exact in $x^2 \rightarrow \int_{-2h}^{2h} x^2\,dx = 2\frac{(2h)^3}{3} = A_0(-h)^2 + \cancel{A_1 0^2} + A_2 h^2 = 2A_0 h^2 = \frac{16h^3}{3}$

$\Rightarrow \boxed{A_0 = \frac{8h}{3} = A_2}$ ; $A_1 = 4h - 2A_0 = \frac{12h}{3} - \frac{16h}{3} = \frac{-4h}{3}$

$\boxed{A_1 = \frac{-4h}{3}}$

Even if it is not asked, I will calculate the error term E with a twofold purpose: check that I got my coefficients right and that the rule was indeed the open N-C one (possibly, anyway):

$$\int_{-2h}^{2h} x^4 dx = 2\frac{(2h)^5}{5} = \frac{64h^5}{5} = \frac{8h}{3}(-h)^4 + \cancel{\frac{-4h}{3}0^4} + \frac{8h}{3}h^4 + K \cdot f^{4)}(\xi)$$

$$f = x^4 \to f^{4)} = 4! = 24 \Rightarrow \frac{64h^5}{5} = \frac{16h^5}{3} + K \cdot 24 \Rightarrow K = \frac{h^5}{24}\left(\frac{64}{5} - \frac{16}{3}\right) =$$

$$= \frac{h^5}{24} \times \frac{192-80}{15} = \frac{h^5}{\cancel{24}_3} \cdot \frac{\cancel{112}\,14}{15} = \frac{14h^5}{45} \ \underline{\underline{\text{OK}}}$$

P.S. In principle one could look for a different rule that looks alike ($E = 14h^5/45 \times f^{4)}(\xi)$) even if $h = b-a$, for instance.

$\boxed{\text{Part 2, Exercise 1}}$

I'll call $x = y_1$, $y = y_2$ so I can write $\underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

Then: $\underline{y}' = \underbrace{\begin{pmatrix} t & 1 \\ 1 & t \end{pmatrix}}_{J = J(t) \text{ (linear system)}} \underline{y} = \underline{f}(t, \underline{y})$ (system of ODEs)

Initial Conditions: $\underline{y}(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Rightarrow \begin{cases} t_0 = 0 \\ \underline{y}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{cases}$

Modified Euler $\equiv$ Midpoint, with advance formula:

$$\begin{cases} \underline{K}_1 = \underline{f}(t_k, \underline{y}_k)\, h_k \\ \underline{K}_2 = \underline{f}\left(t_k + \frac{h_k}{2}, \underline{y}_k + \frac{\underline{K}_1}{2}\right) h_k \\ \underline{y}_{k+1} = \underline{y}_k + \underline{K}_2 \end{cases} \qquad \text{with } h = 0.1$$

First step: $\underline{K}_1 = \underline{f}\left(0, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) h = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} 0.1 = \begin{pmatrix} -0.1 \\ 0.1 \end{pmatrix}$

$$\underline{K}_2 = \underline{f}\left(0.05, \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -0.05 \\ 0.05 \end{pmatrix}\right) \cdot 0.1 = \begin{pmatrix} 0.05 & 1 \\ 1 & 0.05 \end{pmatrix} \begin{pmatrix} 0.95 \\ -0.95 \end{pmatrix} 0.1 = \begin{pmatrix} -0.09025 \\ 0.09025 \end{pmatrix}$$

$$\underline{y}_1 = \underline{y}_0 + \underline{K}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -0.09025 \\ 0.09025 \end{pmatrix} = \begin{pmatrix} 0.90975 \\ -0.90975 \end{pmatrix}$$

$$\Rightarrow \boxed{\begin{array}{l} x(0.1) \simeq 0.90975 \\ y(0.1) \simeq -0.90975 \end{array}}$$

Second step: abusing notation by "overwriting" $\underline{K}_1$, $\underline{K}_2$:

$$\underline{K}_1 = \underline{f}\left(0.1, \begin{pmatrix} 0.90975 \\ -0.90975 \end{pmatrix}\right) \cdot 0.1 = \begin{pmatrix} 0.1 & 1 \\ 1 & 0.1 \end{pmatrix} \begin{pmatrix} 0.90975 \\ -0.90975 \end{pmatrix} \cdot 0.1 = \begin{pmatrix} -0.0818775 \\ 0.0818775 \end{pmatrix}$$

$$\underline{k_2} = f\left(0.15, \begin{pmatrix} 0.90975 \\ -0.90975 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -0.0818755 \\ 0.0818775 \end{pmatrix}\right) \cdot 0.1 =$$

$$= \begin{pmatrix} 0.15 & 1 \\ 1 & 0.15 \end{pmatrix} \begin{pmatrix} 0.86881125 \\ -0.8688125 \end{pmatrix} \cdot 0.1 = \begin{pmatrix} -0.0738490 \\ 0.0738490 \end{pmatrix}$$

$$\underline{y_2} = \underline{y_1} + \underline{k_2} = \begin{pmatrix} 0.90975 \\ -0.90975 \end{pmatrix} + \begin{pmatrix} -0.073849 \\ 0.073849 \end{pmatrix} = \begin{pmatrix} 0.83590104375 \\ -0.83590104375 \end{pmatrix}$$

With 6 significant digits: $\boxed{\begin{array}{l} x(0.2) \simeq 0.835901 \\ y(0.2) \simeq -0.835901 \end{array}}$

**Part 2, Exercise 2**

$$y_{n+1} = y_{n-1} + \frac{h}{3}\left(7 f_n - 2 f_{n-1} + f_{n-2}\right)$$

... already passed to the computer.

## Part 2, Exercise 2

This is again the advance formula: $y_{n+1} - y_{n-1} = \frac{h}{3}\left(7 f_n - 2 f_{n-1} + f_{n-2}\right)$ (1)

To study the convergence and the order of convergence, we will apply the following "recipe":

First write the multistep linear method in its general form:

$$y_{n+k} = -\sum_{j=0}^{k-1} \alpha_j y_{n+j} + h\sum_{j=0}^{k} \beta_j f_{n+j} \qquad (n = 0, ..., N-k) \tag{2}$$

Then calculate its first characteristic polynomial (with $\alpha_k = 1$):

$$\rho(z) = \sum_{j=0}^{k} \alpha_j z^j$$

and its 2nd characteristic polynomial: $\sigma(z) = \sum_{j=0}^{k} \beta_j z^j$

The method is *consistent* iff[1] $\rho(1) = 0$ and $\rho'(1) = \sigma(1)$.

The method is *stable*[2] iff the roots $z_i$ of $\rho(z)$ verify $\|z_i\| \le 1$ on the complex plane, and the ones with modulus 1 are simple roots.

The method is *convergent* iff it is *consistent* and *stable*; and its order of convergence is $p$ iff $\rho(1) = 0$ (already checked for consistency) and:

$$\frac{1}{m}\sum_{j=0}^{k} j^m \alpha_j = \sum_{j=0}^{k} j^{m-1}\beta_j \qquad (m = 1, 2, ..., p)$$

We must first identify (1) and (2) so as to obtain $k$ and the coefficients $\alpha_j$, $\beta_j$.

Let us start with $k$. The points used run from $t_{n-2}$ to $t_{n+1}$, so they are 4 points, hence $k = 3$. You can also see that in that the maximum difference of indices in (1) is $(n+1)-(n-2) = 3$, while in (2) it is $(n+k)-(n+0) = k$; identifying both we get again $k = 3$.

Substituting $k = 3$ in (2) leaves $y_{n+3}$ on the left-hand side, so I will add two units to every index in (1) so that it is easier to identify coefficients. I will also isolate $y_{n+3}$:

$$y_{n+3} = y_{n+1} + \frac{h}{3}\left(7 f_{n+2} - 2 f_{n+1} + f_n\right)$$

If we now expand (2) for $k = 3$:

$$y_{n+3} = \left(-\alpha_0 y_n - \alpha_1 y_{n+1} - \alpha_2 y_{n+2}\right) + h\left(\beta_0 f_n + \beta_1 f_{n+1} + \beta_2 f_{n+2} + \beta_3 f_{n+3}\right)$$

Identifying coefficients between these two expressions it is now immediate to obtain:

$$\alpha_0 = 0;\quad \alpha_1 = -1;\quad \alpha_2 = 0;\quad (\alpha_3 = 1);\quad \beta_0 = \frac{1}{3},\quad \beta_1 = \frac{-2}{3},\quad \beta_2 = \frac{7}{3},\quad \beta_3 = 0$$

The characteristic polynomials are then:

$$\rho(z) = \sum_{j=0}^{k} \alpha_j z^j = \underline{-z + z^3}\,; \qquad \sigma(z) = \sum_{j=0}^{k} \beta_j z^j = \underline{\frac{1}{3} + \frac{-2}{3} z + \frac{7}{3} z^2}$$

---

[1] We use "iff" meaning "if and only if".

[2] Stability of the method itself, regardless of the problem it is solving. A stable method can work unstably (i.e. it can propagate local errors in an amplified manner, typically rendering outrageous results) if the step size $h$ is larger than some threshold $h_t$. The methods that never become unstable when applied on stable ODEs or systems thereof are called *unconditionally stable* or *A-stable*, and are the methods whose absolute stability region covers the entire left complex half plane.

Check for consistency: $\rho(z) = -z + z^3 \quad \Rightarrow \quad \rho(1) = -1 + 1^3 = 0$ ok

$$\rho'(z) = -1 + 3z^2 \quad \Rightarrow \quad \rho'(1) = -1 + 3 = 2\,; \qquad \sigma(1) = \frac{1}{3} + \frac{-2}{3} + \frac{7}{3} = \frac{6}{3} = 2 = \rho'(1) \quad \text{ok}$$

Check for stability:

$$\rho(z) = -z + z^3 = z(-1 + z^2) = 0 \quad \Rightarrow \quad \begin{cases} z = 0 \text{ (with modulus <1)} \\ z = \pm 1 \text{ (with modulus 1, simple roots)} \end{cases} \qquad \text{ok}$$

So the method is also stable. Since it is consistent and stable, it is convergent[3]. Let us calculate the order of convergence. We already checked that $\rho(1) = 0$, so:

For $m = 1$:

$$\left.\begin{aligned} &\frac{1}{1}\left(0^1\alpha_0 + 1^1\alpha_1 + 2^1\alpha_2 + 3^1\alpha_3\right) = -1 + 3 = 2 \\ &\beta_0 + \beta_1 + \beta_2 + \beta_3 = \frac{1}{3} + \frac{-2}{3} + \frac{7}{3} + 0 = \frac{6}{3} = 2 \end{aligned}\right\} \quad \text{equal} \quad \Rightarrow \quad \text{ok}$$

Note that the coefficient of $\beta_0$ would be $0^0$ in the "recipe", which is indeterminate; but the coefficients of $\beta_1$, $\beta_2$ and $\beta_3$ in the same expression are all 1, so it looks natural that the coefficient of $\beta_0$ should also be 1; besides, we are already familiar with that expression, which is $\sigma(1)$, calculated above. Making the coefficient of $\beta_0$ equal to 1 is actually the right way to apply the "recipe" (and the only way for both values to be equal).

For $m = 2$:

$$\left.\begin{aligned} &\frac{1}{2}\left(0^2\alpha_0 + 1^2\alpha_1 + 2^2\alpha_2 + 3^2\alpha_3\right) = \frac{1}{2}(-1 + 9) = 4 \\ &0\beta_0 + 1\beta_1 + 2\beta_2 + 3\beta_3 = 1\frac{-2}{3} + 2\frac{7}{3} = \frac{-2 + 14}{3} = 4 \end{aligned}\right\} \quad \text{equal} \quad \Rightarrow \quad \text{ok}$$

For $m = 3$:

$$\left.\begin{aligned} &\frac{1}{3}\left(0^3\alpha_0 + 1^3\alpha_1 + 2^3\alpha_2 + 3^3\alpha_3\right) = \frac{1}{3}(-1 + 27) = \frac{26}{3} \\ &0^2\beta_0 + 1^2\beta_1 + 2^2\beta_2 + 3^2\beta_3 = \frac{-2}{3} + 4\frac{7}{3} = \frac{-2 + 28}{3} = \frac{26}{3} \end{aligned}\right\} \quad \text{equal} \quad \Rightarrow \quad \text{ok}$$

For $m = 4$:

$$\left.\begin{aligned} &\frac{1}{4}\left(0^4\alpha_0 + 1^4\alpha_1 + 2^4\alpha_2 + 3^4\alpha_3\right) = \frac{1}{4}(-1 + 81) = 20 \\ &0^3\beta_0 + 1^3\beta_1 + 2^3\beta_2 + 3^3\beta_3 = \frac{-2}{3} + 8\frac{7}{3} = \frac{-2 + 56}{3} = \frac{54}{3} \end{aligned}\right\} \quad \text{different}$$

Therefore <u>the order of the method is $p = 3$</u>.

The method is <u>explicit</u>, because in (1) the unknown $y_{n+1}$, when isolated, is written explicitly in terms of already-known values (the points at $t_n$, $t_{n-1}$ and $t_{n-2}$ are all known by the time one tries to calculate the one at $t_{n+1}$). In other words, $y_{n+1}$ does not appear on the right-hand side of (1), so one does not need to iterate in order to apply the advance formula. The computational cost is therefore small (basically one evaluation of $f(t,y)$ per step).

[3] Consistency guarantees that each local error (the error attributable to each single step) tends to zero fast enough as the step size $h$ tends to zero (namely, faster than $h$). Stability guarantees that the propagation of those local errors can occur in a damped, rather than amplified, manner, i.e., that the global error $E$ at the end of a time interval can be less than the sum of all the local errors made in the interval. When actually applying a method to solve a specific problem, there exists an "amplification factor" that governs the propagation, or composition, of the local errors to form the global error; and for the method to work stably, that amplification factor must be less than 1 in absolute value. That depends not only on the method, but also on the problem being solved and on the step size $h$ used to solve it with that method. Here we are talking all the time about truncation errors, i.e., with exact arithmetic.

The method is not of Adams-Moulton because AM methods are implicit; but neither is it of Adams-Bashforth, which are explicit, because in those methods you have $y_n$, rather than $y_{n-1}$, on the right-hand side of the advance formula (Adams-Bashforth methods integrate the interpolation polynomial $p(t)$ of $f_i = f(t_i, y_i)$ at $i = n, n-1, n-2, \ldots$ between $t_n$ and $t_{n+1}$, and add the value of that integral to $y_n$). So it is <u>an explicit multistep linear method of some other type</u> than AB. Specifically, this method integrates $p(t)$ between $t_{n-p}$ and $t_{n+1}$ for $p = 1$ (and that's why it starts $y_{n+1} = y_{n-1} + \ldots$); these methods are called *Nyström* methods.

Finally, by definition, <u>the number of steps is 3</u> because that is the number of known points used (namely the points at $t_{n-2}$, $t_{n-1}$ and $t_n$; even if it were an implicit method, by definition, they number that determines the number of steps of the method is the number of "already calculated" points used on the right-hand side of the advance formula).

# Part 2, Exercise 3

First I write the Taylor series expansions of interest. With a little bit of foresight I realize Taylor remainders of order 3 will work just fine (otherwise write suspension points and truncate later):

$$f(z+2h) = f(z) + f'(z)\,2h + \frac{f''(z)}{2!}(2h)^2 + \frac{f'''(\xi_1)}{3!}(2h)^3$$

$$f(z-h) = f(z) - f'(z)\,h + \frac{f''(z)}{2!}h^2 - \frac{f'''(\xi_2)}{3!}h^3$$

for some $\xi_1 \in (z, z+2h)$, $\xi_2 \in (z-h, z)$, and assuming that $f \in C^3([z-h, z+2h])$ (or at least that $f^{3)}$ exists there and is bounded).

We are requested to achieve the highest possible order of convergence, so I will try to get rid of the terms of order 2. I can achieve just that by multiplying the second by $(-4)$ and adding term to term:

$$f(z+2h) - 4f(z-h) = -3f(z) + f'(z)\,6h + \emptyset + \frac{8f'''(\xi_1) + 4f'''(\xi_2)}{3!}h^3$$

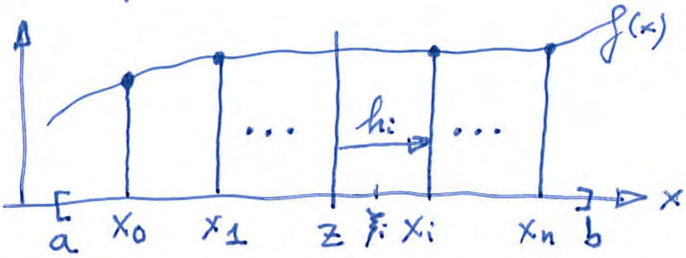
Isolating $f'(z)$:

$$f'(z) = \frac{f(z+2h) + 3f(z) - 4f(z-h)}{6h} - \frac{2f'''(\xi_1) + f'''(\xi_2)}{9}h^2$$

As $h \to 0^+$, $\xi_1 \to 0^+$ and $\xi_2 \to 0^-$; if $f \in C^3$ near $z$, the last numerator will tend to $3f'''(z)$ and actually be equal to $3f'''(\xi)$ for some $\xi$ near $z$. That's essentially correct, but for a formal proof consider the function $g(x) = 3f'''(x) \in C([\xi_1, \xi_2])$. The value $2f'''(\xi_1) + f'''(\xi_2)$ is obviously intermediate between $g(\xi_1)$ and $g(\xi_2)$; therefore, by virtue of Weierstrass's Intermediate Value Theorem, there must exist at least one $\xi \in [\xi_1, \xi_2]$ such that $g(\xi) = 3f'''(\xi) = 2f'''(\xi_1) + f'''(\xi_2)$. Substituting:

$$\boxed{f'(z) = \frac{f(z+2h) + 3f(z) - 4f(z-z)}{6h} + \underbrace{\frac{-f'''(\xi)}{3}h^2}_{E = O(h^2)}}$$

That was done with an ad-hoc manipulation of the Taylor series expansions of interest. One can also obtain the same result in a more systematic way, i.e., applying the "theory". Let's remember it:




Taylor expansion of $f$ at $x_i = z + h_i$ around $z$:

$$f(x_i) = f(z) + f'(z) h_i + \frac{f''(z)}{2!} h_i^2 + \dots + \frac{f^{k)}(z)}{k!} h_i^k + \dots + \frac{f^{n)}(z)}{n!} h_i^n + \frac{f^{n+1)}(z)}{(n+1)!} h_i^{n+1} + \dots$$

$$\dots + \frac{f^{m)}(z)}{m!} h_i^m + \frac{f^{m+1)}(\xi_i)}{(m+1)!} h_i^{m+1}$$

That's assuming:

- $f \in C^{m+1}([a,b])$ (or at least $f^{m+1)}$ $\exists$ and bounded in $(a,b]$) for the expansion to be valid;
- $n \geq k$ (so we have enough points to approximate $f^{k)}(z)$; in our case, $k=1$, $n=2$);
- $m \geq n$ (for the valid series expansion to be long enough).

Like always, $f^{k)}(z) = D + E = \sum_{i=0}^{n} A_i f(x_i) + E$. Substituting:

$$f^{k)}(z) = \sum_{i=0}^{n} A_i \left[ f(z) + f'(z) h_i + \frac{f''(z)}{2!} h_i^2 + \dots \right] + E =$$

$$= \underbrace{\left( \sum_{i=0}^{n} A_i \right)}_{=0} f(z) + \underbrace{\left( \sum_{i=0}^{n} A_i h_i \right)}_{=0} f'(z) + \underbrace{\left( \frac{1}{2!} \sum_{i=0}^{n} A_i h_i^2 \right)}_{=0} f''(z) + \dots + \underbrace{\left( \frac{1}{k!} \sum_{i=0}^{n} A_i h_i^k \right)}_{=1} f^{k)}(z) + \dots$$

$$+ \underbrace{\left( \frac{1}{n!} \sum_{i=0}^{n} A_i h_i^n \right)}_{=0} f^{n)}(z) + \left( \frac{1}{(n+1)!} \sum_{i=0}^{n} A_i h_i^{n+1} \right) f^{n+1)}(z) + \dots + \left( \frac{1}{m!} \sum_{i=0}^{n} A_i h_i^m \right) f^{m)}(z) +$$

$$+ \frac{1}{(m+1)!} \left( \sum_{i=0}^{n} A_i h_i^{m+1} f^{m+1)}(\xi_i) \right) + E \quad (*)$$

If we impose the conditions indicated $=0, =0, \dots, =1, \dots, =0$ above, then everything cancels out except the last terms $(*)$, which must be then $0$. From $(*) = 0$ we can isolate $E$ and get the expression of the error term (applying Weierstrass's Intermediate Value Theorem in the process). The conditions $=0, \dots, =1, \dots, =0$ represent an $(n+1) \times (n+1)$ set of linear equations in $A_i$; solving it we obtain the coefficients $A_0, A_1, \dots, A_n$.

In our case: $K=1$, $n=2$, $m=2$ (we assume $f\in C^3([a,b])$),
and $h_0=-h$, $h_1=0$, $h_2=2h$. Substituting:

$$\begin{cases} A_0+A_1+A_2=0 \\ A_0(-h)+A_1\cdot 0+A_2(2h)=1 \Rightarrow -A_0+2A_2=1/h \\ A_0(-h)^2+A_1\cdot 0^2+A_2(2h)^2=0 \Rightarrow A_0+4A_2=0 \end{cases} \overset{(+)}{\Rightarrow} 6A_2=1/h \quad A_2=\frac{1}{6h} \rightarrow$$

$$A_0=-4A_2=\frac{-4}{6h} \rightarrow A_1=-A_0-A_2=\frac{4}{6h}-\frac{1}{6h}=\frac{3}{6h}$$

$$\Rightarrow \boxed{D=\frac{-4f(z-h)+3f(z)+f(z+2h)}{6h}}$$ like before.

As for the error term $E$, let's substitute into $(*)=0$. Only the last term(s) remain:

$$\frac{1}{3!}\left(A_0(-h)^3 f'''(\xi_0)+0+A_2(2h)^3 f'''(\xi_2)\right)+E=0 \Rightarrow$$

$$E=\frac{-1}{3!}\left(\frac{+4}{6h}h^3 f'''(\xi_0)+\frac{1}{6h}(2h)^3 f'''(\xi_2)\right)$$ where $\xi_0\in(z-h,z)$, $\xi_2\in(z,z+2h)$.

Therefore $E=\dfrac{-f'''(\xi_0)-2f'''(\xi_2)}{9}h^2$ and, by a reasoning similar to the one above, $E=\dfrac{-f'''(\xi)}{3}h^2$ for some $\xi$ near $z$.

P.S. If it was not obvious enough before:

If $f'''(\xi_0)=f'''(\xi_2)$: $\quad 3f'''(\xi_0)=f'''(\xi_0)+2f'''(\xi_2)=3f'''(\xi_2)$

If $f'''(\xi_0)<f'''(\xi_2)$: $\quad 3f'''(\xi_0)<f'''(\xi_0)+2f'''(\xi_2)<3f'''(\xi_2)$

If $f'''(\xi_0)>f'''(\xi_2)$: $\quad 3f'''(\xi_0)>f'''(\xi_0)+2f'''(\xi_2)>3f'''(\xi_2)$

In all three cases, $f'''(\xi_0)+2f'''(\xi_2)$ is an intermediate value of $3f'''\in C([\xi_0,\xi_2])$ between its values on $\xi_0$ and on $\xi_2$, so one can apply Weierstrass's Intermediate Value Theorem and conclude that $\exists\, \xi\in[\xi_0,\xi_2]\subset[z-h,z+2h] \,/\, 3f'''(\xi)=f'''(\xi_0)+2f'''(\xi_2)$.

P.S.2 - Of course there's also the possibility to develop the "systematic theory" just for the problem at hand, namely use $K=1$, $n=m=2$ from the very beginning, truncating the Taylor expansions at order 3. You should be able to do that yourself.