

INTRODUÇÃO À FÍSICA COMPUTACIONAL

Projeto 2

Luiz Fernando S E Santos
N. Usp: 10892680

1. INTRODUÇÃO

Nesta prática tivemos contato com diferentes métodos utilizados em cálculo numérico para derivação, integração e aproximação de raízes de polinômios. Grande parte dos métodos foram possíveis graças à propriedades da expansão do polinômio de Taylor para diferentes pontos e intervalos.

Foram então analisados os erros das medidas em relação ao seu valor “real” para os diferentes métodos e sua eficiência. Os resultados foram analisados e discutidos logo após.

2. METODOLOGIA

2.1 Derivação numérica

A)

Desejando calcular o erro cometido ao utilizarmos a fórmula:

$$f'_{5s}(x) = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h}$$

Abrimos o polinômio de Taylor para cada ponto que aparece na fórmula e o multiplicamos pelo termo também mostrado:

$$-f(x+2h) = -f(x) - 2f'(x)h - \frac{4}{2!}f''(x)h^2 - \dots$$

$$8f(x+h) = 8f(x) + 8f'(x)h + \frac{8}{2!}f''(x)h^2 + \dots$$

$$-8f(x-h) = -8f(x) + 8f'(x)h - \frac{8}{2!}f''(x)h^2 + \dots$$

$$f(x-2h) = f(x) - 2f'(x)h + \frac{4}{2!}f''(x)h^2 + \dots$$

Somando todos os termos, diversos deles se cancelam e ao final sobra:

$$12f'(x)h + Af^{(5)}(x)h^5 + \dots = -f(x-2h) + 8f(x-h) - 8f(x+h) + f(x+2h)$$

Onde A é uma constante. Como o termo que multiplica A constitui o h de menor índice (e, portanto, maior valor para $h < 1$), este termo é usado como o erro. E então:

$$f'(x) + \frac{A}{12}f^{(5)}(x)h^4 = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h}$$

Portanto o erro é da ordem da quarta potência de h.

B)

Sendo a função que desejamos derivar:

$$f(x) = e^{2x} \cos\left(\frac{x}{4}\right)$$

Calculei manualmente sua primeira e segunda derivadas, obtendo assim respectivamente:

$$f'(x) = \frac{-e^{2x}(\sin(x/4) - 8\cos(x/4))}{4}$$

$$f'''(x) = \frac{-e^{2x}(16\sin(x/4) - 63\cos(x/4))}{16}$$

Então, utilizando os valores de h mostrados na tabela, calculo a derivada no ponto $x = 1$ numericamente ao passo que vario o h. O erro do meu cálculo será determinado pelo módulo da diferença entre a derivada calculada numericamente e da calculada diretamente pela fórmula. Estes erros serão então analisados.

C)

De forma similar ao último método, dessa vez vario h de 10^{-15} até 1. Faço então uma tabela com $\log(h) \times \log(E)$ que será usada para a elaboração de um gráfico.

Sendo o erro dado por:

$$|e| = Ah^\alpha$$

Isso quer dizer que:

$$\log(|e|) = \alpha \log(h) + \log(a)$$

Portanto, espera-se que o gráfico resultante seja uma reta, onde seu coeficiente angular é a ordem de h, α .

2.2 Integração numérica

A)

Pelo teorema fundamental do cálculo, uma integral entre dois pontos a e b é dada por:

$$\int_a^b f(x) dx = F(b) - F(a)$$

Sendo assim, temos que:

$$\int_{n-2h}^{n+2h} f(x) dx = F(n+2h) - F(n-2h)$$

Abrindo os polinômios de Taylor de $F(n+2h)$ e $F(n-2h)$, somando os termos e substituindo as derivadas por meio de outras séries, ao final chegamos que a ordem do erro local desse método deve ser da ordem de $O(h^7)$. Sendo assim, o erro global será de $O(h^6)$.

B)

Para calcular a integral de um intervalo utilizando o método do trapézio, usamos a fórmula:

$$I_t = h[0.5f(a) + f(a+h) + \dots + f(b-h) + 0.5f(b)]$$

Para fazer isso computacionalmente, defino um loop externo, que varia o h, e outro interno, em que varia um inteiro i e multiplica h por ele obtendo, $[h, 2h, 3h, \dots, (n-1)h]$, onde $a + (n-1)h = b - h$, assim percorrendo todo o intervalo desejado de integração. O motivo do -1 vem de o primeiro e o último termos da integral ($f(a)$ e $f(b)$) serem multiplicados por um número diferente dos demais, o que torna mais fácil somar esses dois termos previamente e, após isso, somarmos os termos do meio. Isso nos poupa memória que seria gasta com a utilização de uma estrutura de controle de fluxo dentro do loop, que em vez de i variar de 1 até n-1, variaria de 0 até n.

O mesmo vale para os demais casos, de Simpson e de Boole. A fórmula para esses dois é, respectivamente:

$$I_s = \frac{h}{3} [f(a) + 4f(a+h) + 2f(a+2h) + \dots + 4f(b-2h) + 2f(b-h) + f(b)]$$

$$I_B = \frac{2h}{45} [7f(a) + 32f(a+h) + 12f(a+2h) + 32f(a+3h) + 14f(a+4h) + \dots + 7f(b)]$$

A fórmula para a regra de Boole é obtida somando os intervalos $[x-2h, x+2h]$. Em ambas temos algo em comum nos termos intermediários que é a variação de um número que multiplica nosso f entre certo k e $k+2$ (2 e 4 ou 12 e 14). Podemos resolver isso apenas decompondo o valor em: $(k + \text{auxiliar} * 2)$ onde a variável auxiliar varia entre 0 e 1. O caso da regra de Boole também tem outro complicador: esta variação só vale para os termos pares, ao passo que os ímpares são todos multiplicados por 32. Nesse caso, para poder utilizar a solução da multiplicação do auxiliar por 2 descrita acima nesse caso também, desta vez utilizo uma estrutura de controle de fluxo para que possa alterar o valor do auxiliar APENAS após passar um índice par, caso contrário ele não se altera.

Primeiramente, vario o auxiliar da seguinte forma:

```
auxiliar = 0
x = 0
DO
    auxiliar += (-1)^x
    x ++
FIM DO
```

Como x varia entre par e ímpar, o auxiliar, que inicia em 0, é primariamente acrescido de 1, para logo na próxima iteração ser decrescido do mesmo valor, voltando à zero.

Combinando isso com o IF/ELSE:

```
IF (ímpar):
    (...)
ELSE:
    auxiliar += (-1)^x
    x ++
FIM IF
```

Agora só altero o auxiliar ao passar por um valor par.

Do valor obtido ao final das iterações, multiplico todas por seus respectivos $\text{NÚMERO} * h$, mostrado no início de cada equação, e calculo o desvio de cada método, para cada h , em relação ao valor verdadeiro, calculado diretamente pela integral da função entre a e b .

C)

Variamos agora h de 10^{-15} até 1 e realizamos o mesmo procedimento citado em b) para calcular cada integral. Desta vez, faço minha tabela utilizando o logaritmo de erro obtido e também o log do h utilizado. Plotamos o gráfico e é esperado o mesmo comportamento do descrito e explicado em 1.c).

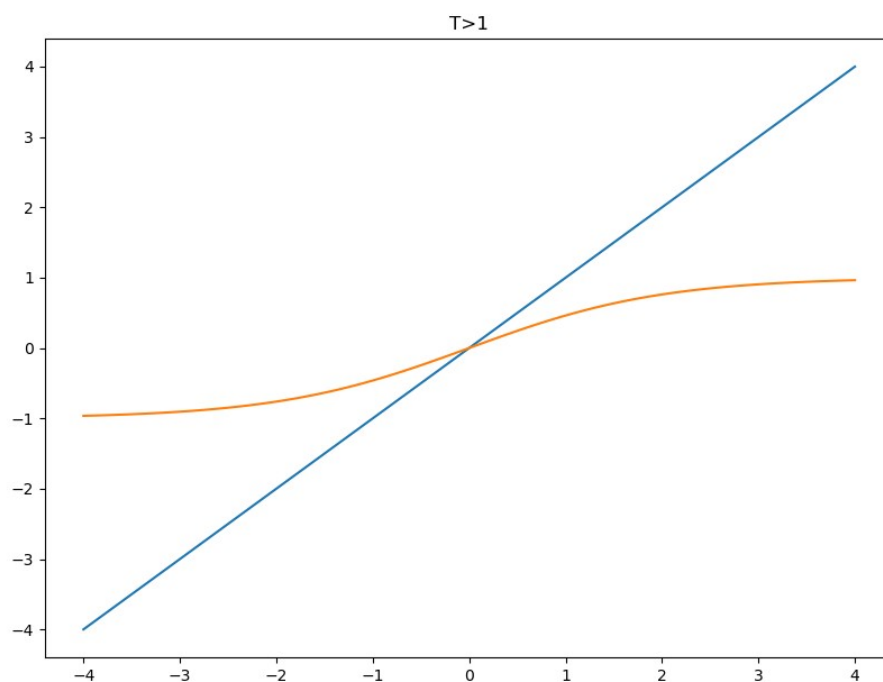
2.3 Equações algébricas não lineares

A)

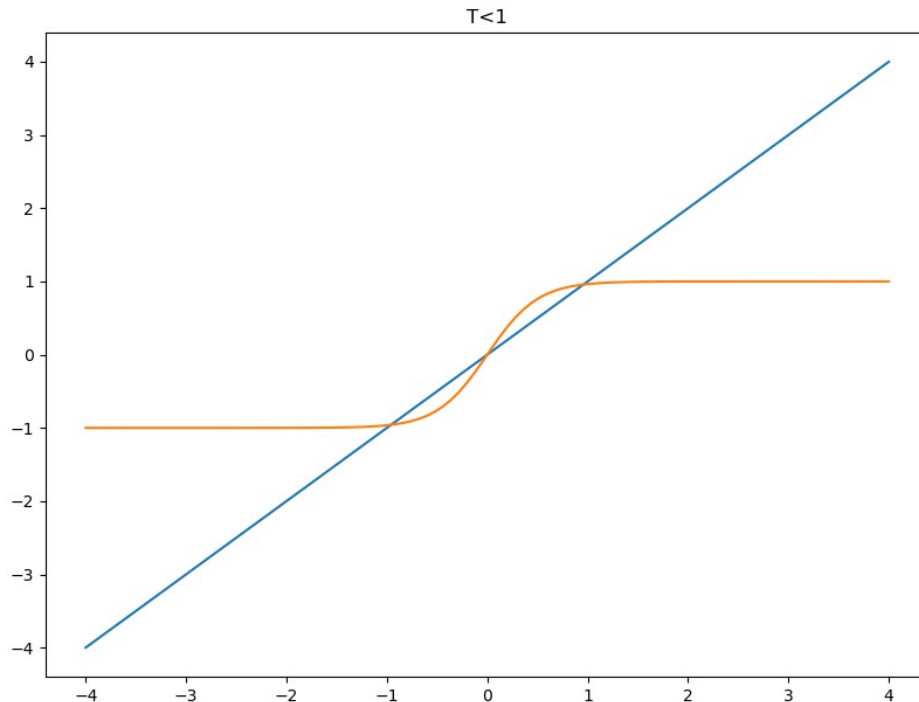
Conhecendo a função transcendental

$$m = \tanh\left(\frac{m}{T}\right)$$

Considerando T constante, os valores de m que fazem a igualdade valer são aqueles em que $g(m) = m$ e $h(m) = \tanh(m/t)$ se interceptam. O número de pontos onde isso acontece pode mudar, dependendo do valor de T , como pode ser observado:



Para $T > 1$ (a única solução é $m = 0$). E



Para $T < 1$. Agora podemos observar 3 soluções.

Um modo de determinar quais são essas soluções é simplesmente definindo a função

$$f(m) = \tanh\left(\frac{m}{T}\right) - m$$

Suas raízes serão os valores desejados de m .

C)

Agora, com a função em mãos, aplico ambos os métodos discutidos para determinar a maior raiz M , solução da equação, utilizando tanto o método da bisseção quanto o de Newton-Raphson.

Começo definindo como funções do programa, $f(x)$ terá como valor de entrada um certo número m e calculará a função descrita em a) usando ele, e $fder(x)$, que é a primeira derivada dessa função, calculada manualmente. Essas duas funções serão usadas no método de Newton-Raphson, em que

$$m_{n+1} = m - \frac{f(m)}{f'(m)}$$

é o próximo valor de M (com maior precisão) calculado. Como desejo fazer esse procedimento até atingir a precisão desejada, preciso comparar o valor “novo”, obtido após calculada a expressão acima, com seu valor “antigo” (o m usado para calcular a expressão). Para isso, criei uma variável m , para representar o valor antigo, e mpp (“ $m++$ ”), a que será calculada usando m . Defino dentro do meu loop em que faço a operação, uma estrutura de controle de fluxo em que, caso a precisão seja atingida, interrompe o ciclo. Após calculada a diferença entre mpp e m , posso fazer $m = mpp$, para o próximo loop.

Para o método da bisseção, começo com o intervalo $a = 0$ e $b = 1$. Dentro do meu loop, calculo o ponto médio, c , entre a e b . Como sei, por definição, que se a raiz está contida em um intervalo, digamos $[c, b]$, o produto $f(c) \cdot f(b)$ será menor que zero, e caso contrário será maior (uma vez que não passa por $y = 0$ em nenhum ponto (isso considerando a existência/ausência de apenas

uma raiz no intervalo), o que implica uma multiplicação de termos de mesmo sinal, resultando em um valor maior que zero. Calculo, assim, meu c, para que possa calcular o produto de $f(c)$ com $f(a)$ e $f(b)$. Se tenho certeza que há uma raiz no meu intervalo inicial, sei que, caso $f(b)*f(c) > 0$, com certeza a raiz está no intervalo $[a, c]$. Desta forma, faço $b = c$ e consigo um novo intervalo, de metade de tamanho do anterior e mais próximo da raiz.

Também defino a variável lc (“last c”), para ser usado de modo análogo á m e mpp e determinar a precisão do meu resultado.

Para completar a tabela atingir a precisão de 10^{-6} , conforme pedido, sabendo que o método da bisseção requer um número maior de loops para atingir a mesma precisão que o de Newton-Raphson, defino seu erro como a condição para quebrar o loop e coloco ambos os métodos dentro de uma mesma estrutura de repetição, assim reduzindo a complexidade.

D e E)

Utilizo o mesmo método descrito em C), mas dessa vez adiciono uma outra estrutura de repetição externa, que varie minha precisão E. Escrevo o $\log(E)$ e nmax, o log do número de iterações necessárias para se atingir E. Faço isso tanto para o método da bisseção quanto para o de Newton-Raphson, e para os valores de T 0.9 e 0.3. Os resultados são comparados e discutidos em seguida.

F)

Agora, adiciono nas funções do meu programa ($f(x)$ e $fder(x)$), um segundo parâmetro T, que significa uma temperatura também variável. Calculo como a raiz encontrada pelo programa, m, se comporta conforme vario T. O erro da raiz encontrada é tomado como o erro da máquina para o tipo de variável utilizado (REAL*8). Os resultados de M para cada T são calculados e posteriormente analisados.

3.RESULTADOS E DISCUSSÕES

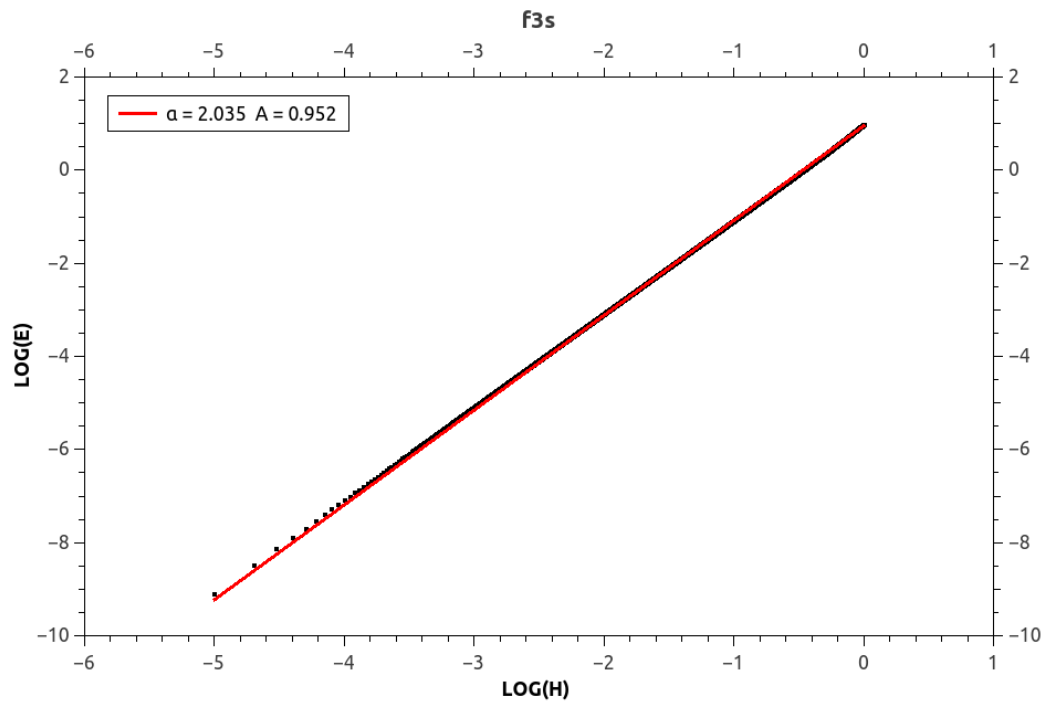
3.1 Derivação numérica

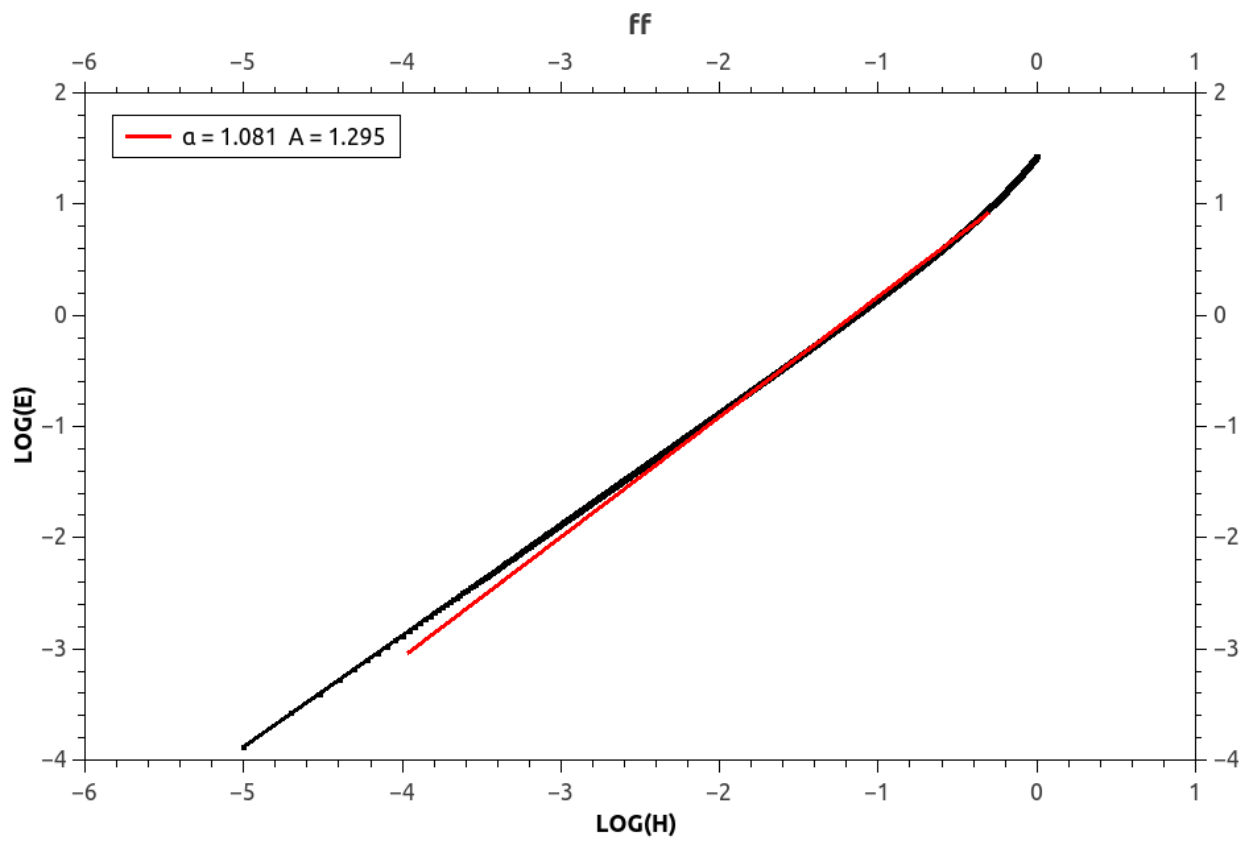
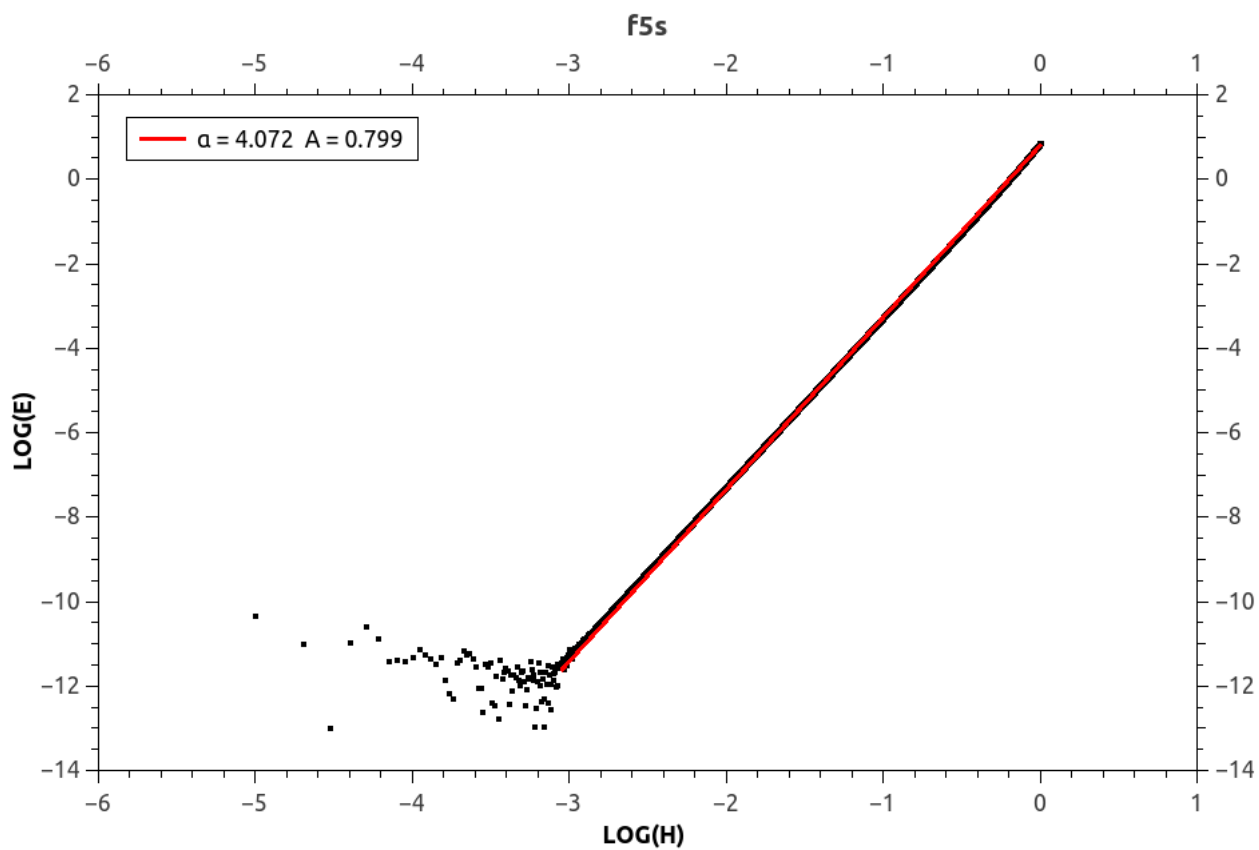
A tabela obtida pelo método descrito foi:

ff'	ft'	f'3s	f'5s	f''
9.1991178720791602	4.9371253446305730	2.1309962637242936	0.35717703031651560	1.9106344765015351
1.4038455737628368	1.2398006362701981	8.2022468746318467E-002	5.2893051089419885E-004	7.4610143412414232E-002
0.67999314578382197	0.63903137721667846	2.0480884283571754E-002	3.2977204014628114E-005	1.8638503092073933E-002
0.13263190558268789	0.13099406759237553	8.1891899515618150E-004	5.2722093357715494E-008	7.4536058841090380E-004
6.6109823016276437E-002	6.5700368461733660E-002	2.0472727727138817E-004	3.2953124673440470E-009	1.8633868408812759E-004
1.3189118762475971E-002	1.3172740646709258E-002	8.1890578833565542E-006	7.5424111400934635E-012	7.4522672974808302E-006
6.5925107177804421E-003	6.5884161921587747E-003	2.0472628108336721E-006	2.3607782395629329E-012	1.8629605023079421E-006
1.3181744931216599E-003	1.3180107059014290E-003	8.1893610115457705E-008	5.0395243533785106E-012	3.3312957725684100E-008
6.5906680025307196E-004	6.5902582146293298E-004	2.0489395069489547E-008	2.5764279598661233E-011	4.7740216757574672E-007
1.3181010861096354E-004	1.3180852941907517E-004	7.8959594418392953E-010	4.6771475581408595E-011	1.1846085936184636E-005
6.5928498287348702E-006	6.5930469901331890E-006	9.8580699159356300E-011	6.4156857604302786E-010	9.9416808181871374E-003
1.3259517999131276E-006	1.3297016749902468E-006	1.8749375385596068E-009	1.8749375385596068E-009	0.19468279211581674
6.3317263254702993E-007	6.6356786021515290E-007	1.5197613834061485E-008	2.5559694805110666E-008	0.42704210167427092
1.1802914912095730E-007	5.9606534819067747E-008	2.9211307150944776E-008	3.6612794573898100E-008	8.5982835629154302

Podemos ver que, para $f'3s$, $f'5s$ e f'' , em certo ponto a incerteza para de diminuir, aumenta e passa a se comportar de forma estranha. Isso é por conta do erro da máquina, que sendo do tipo de variável REAL*8, chega até 10^{-8} . Assim, o h ótimo para cada caso é aquele que dexia o resultado da imprecisão mais próximo desse valor, sem ultrapassálo (EX: para $f'3s$ o nono h , para $f'5s$ o quarto, para f'' o oitavo...).

Foi então feito um gráfico $\log(h) \times \log(\text{erro})$, obtendo assim os seguintes resultados:





Pelo que foi explicado anteriormente, a relação entre o coeficiente angular e a ordem do erro prevista para h são equivalentes com uma precisão maior que 5% em todos os casos.

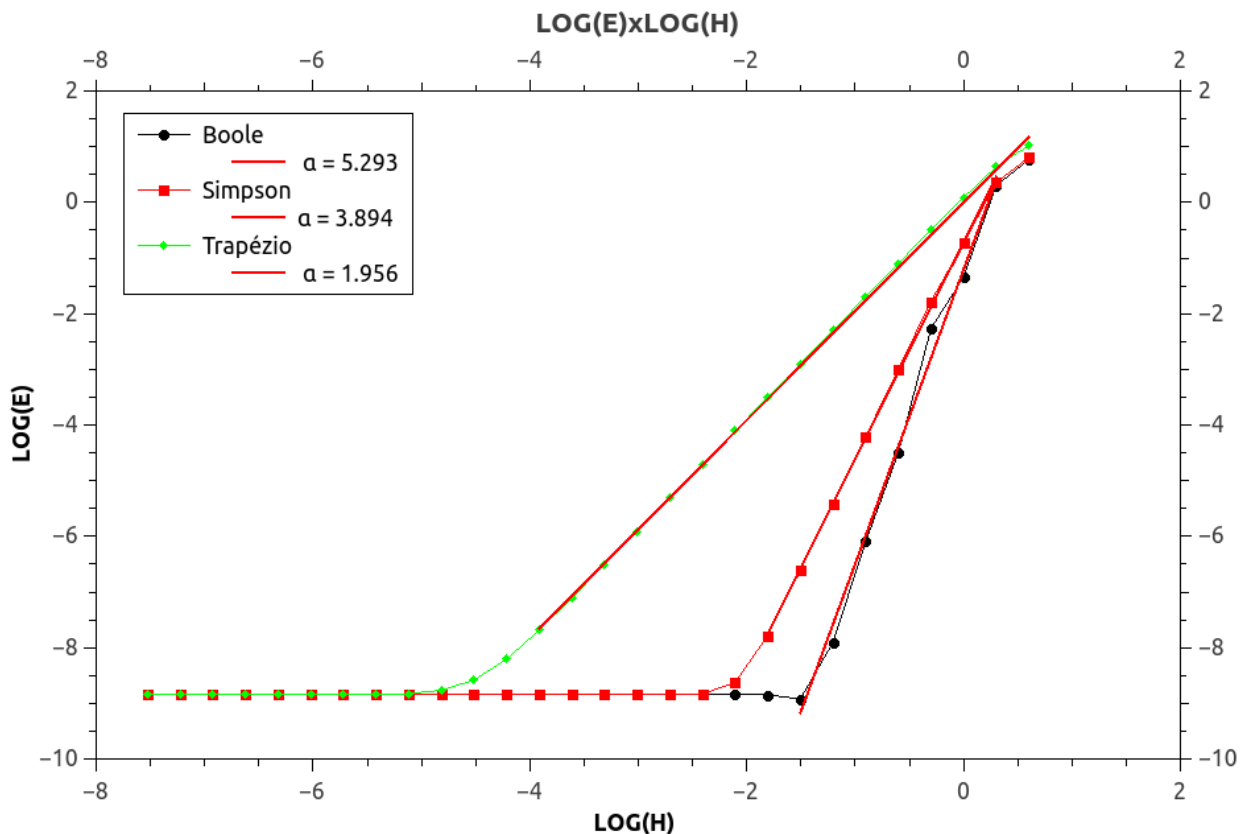
3.2 Integração numérica

A tabela obtida foi:

h	It	Is	Ib
0.5000000000000000	0.31742022909119028	1.6482824675025709E-002	5.3181425236923374E-003
0.2500000000000000	8.0105545321877303E-002	1.0006507321065339E-003	3.1494197421544357E-005
0.1250000000000000	2.0072730063970923E-002	6.1791644668351964E-005	7.9896116078614909E-007
6.2500000000000000E-002	5.0210702469279944E-003	3.8503079133889173E-006	1.2447870290088758E-008
3.1250000000000000E-002	1.2554488834479738E-003	2.4176228818895140E-007	1.1925795906364556E-009
1.5625000000000000E-002	3.1387454567899375E-004	1.6433089555789593E-008	1.4111423141116575E-009
7.8125000000000000E-003	7.8470401347008334E-005	2.3532360504674443E-009	1.4145788984620822E-009
3.9062500000000000E-003	1.9618705308310780E-005	1.4732950415208279E-009	1.4146330773456839E-009
1.9531250000000000E-003	4.9057400519636474E-006	1.4183008101298356E-009	1.4146344096133134E-009
9.7656250000000000E-004	1.2274961611602464E-006	1.4148640037348059E-009	1.4146352977917331E-009
4.8828125000000000E-004	3.0793502547865614E-007	1.4146477322896089E-009	1.4146315230334494E-009
2.4414062500000000E-004	7.8044733120563592E-008	1.4146321891672642E-009	1.4146330773456839E-009
1.2207031250000000E-004	2.0572151315789711E-008	1.4146277482751657E-009	1.4146324112118691E-009

É visível como a ordem do erro cai mais rápido com h com um método do que com o outro. Pela mesma lógica do erro da máquina, explicada anteriormente, que agora pode ser visualizada como com o erro convergindo para um valor, mesmo alterando h , o h ótimo para It é o décimo terceiro, para Is é o sétimo e para Ib o quinto.

Do gráfico $\log(h) \times \log(E)$, obtemos:



Dessa vez, os valores estão menos precisos que no caso anterior, principalmente no caso do método de Boole (que devia nos dar uma linha com coef angular 6). Mesmo assim, eles mostram a convergência dos valores a partir de certo h , por conta do erro da máquina, conforme o esperado. Também assumem valores equivalentes em zero. Isso também está de acordo com o esperado para o comportamento do erro com função da ordem de h .

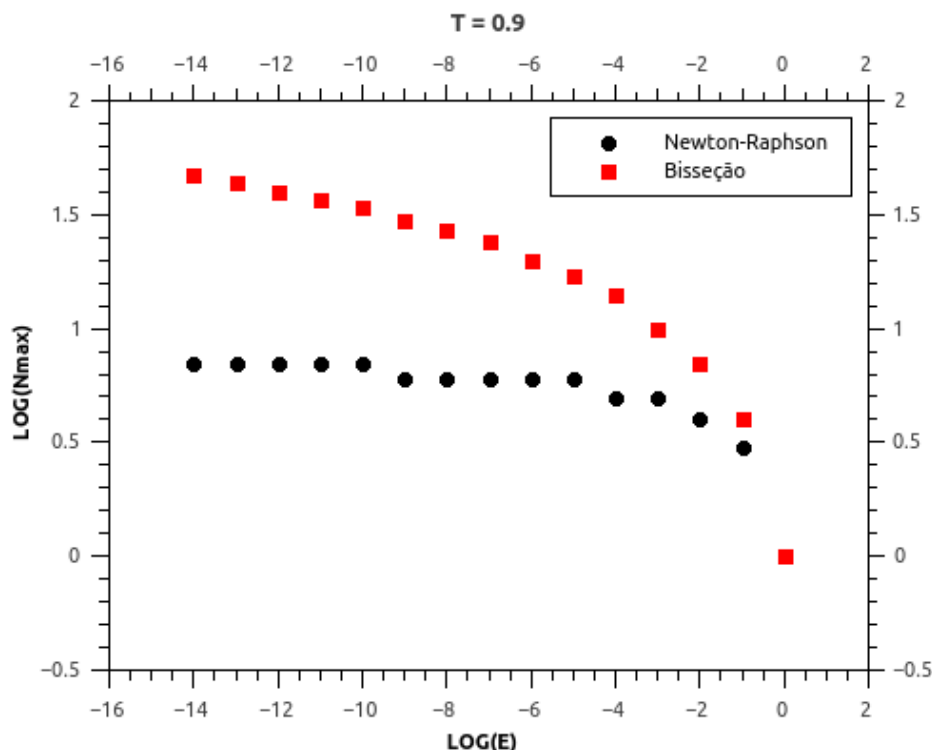
3.3 Equações algébricas não lineares

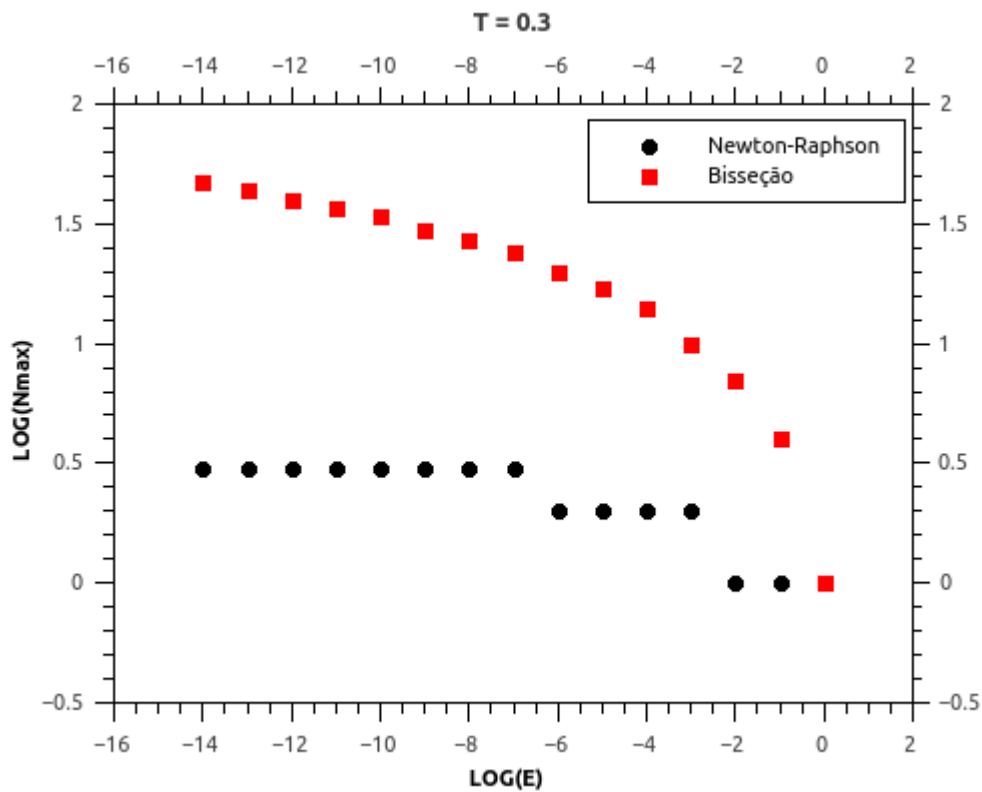
A tabela obtida foi:

a	b	$ b - a $	M	$ m - mpp $
0.0000000000000000	1.0000000000000000	1.0000000000000000	1.0000000000000000	4.1891800434786797E-002
0.5000000000000000	1.0000000000000000	0.5000000000000000	0.95810819956521320	6.0403624757887631E-004
0.7500000000000000	1.0000000000000000	0.2500000000000000	0.95750416331763433	1.3924035813239755E-007
0.8750000000000000	1.0000000000000000	0.1250000000000000		
0.9375000000000000	1.0000000000000000	6.2500000000000000E-002		
0.9375000000000000	0.9687500000000000	3.1250000000000000E-002		
0.9531250000000000	0.9687500000000000	1.5625000000000000E-002		
0.9531250000000000	0.9609375000000000	7.8125000000000000E-003		
0.9570312500000000	0.9609375000000000	3.9062500000000000E-003		
0.9570312500000000	0.9589843750000000	1.9531250000000000E-003		
0.9570312500000000	0.9580078125000000	9.7656250000000000E-004		
0.9570312500000000	0.9575195312500000	4.8828125000000000E-004		
0.9572753906250000	0.9575195312500000	2.4414062500000000E-004		
0.9573974609375000	0.9575195312500000	1.2207031250000000E-004		
0.9574584960937500	0.9575195312500000	6.1035156250000000E-005		
0.9574890136718750	0.9575195312500000	3.0517578125000000E-005		
0.9574890136718750	0.9575042724609375	1.5258789062500000E-005		
0.95749664306640625	0.9575042724609375	7.6293945312500000E-006		
0.95750045776367188	0.9575042724609375	3.8146972656250000E-006		
0.95750236511230469	0.9575042724609375	1.9073486328125000E-006		

Podemos ver claramente uma grande diferença no número de iterações necessárias para se atingir a mesma precisão entre o método da Bissecção e o de Newton-Raphson.

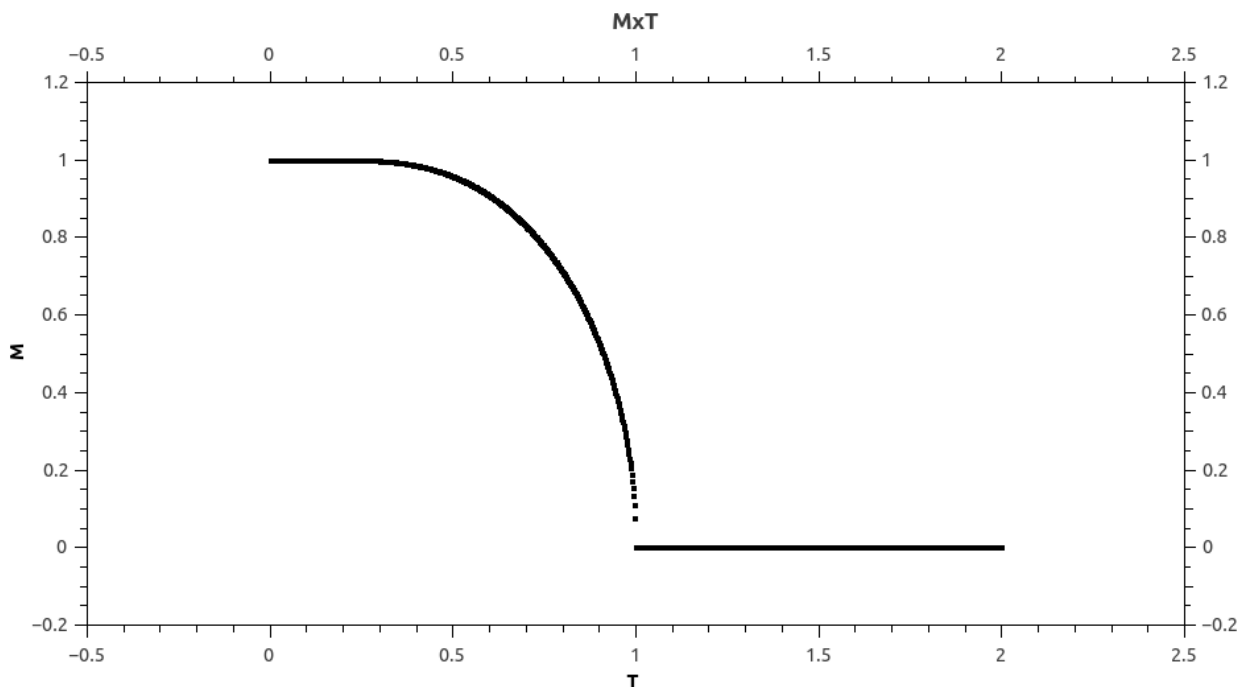
Dos gráficos de log, obtivemos:



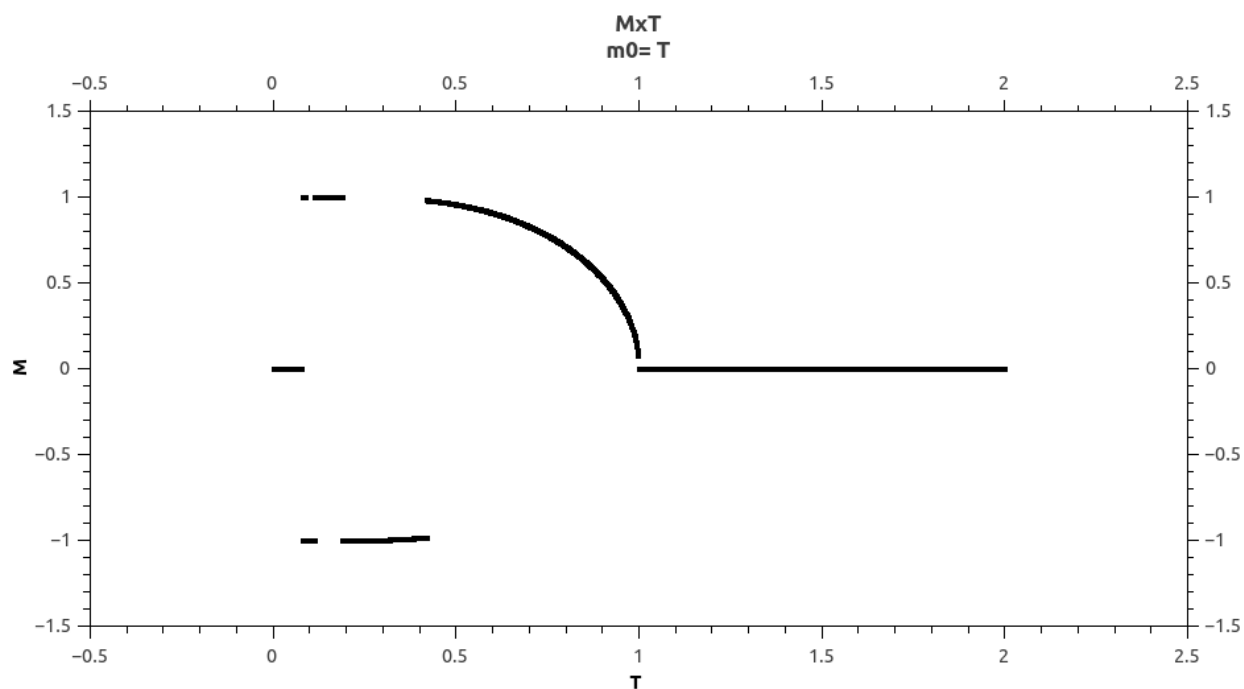


Podemos ver que, ao passo que o formato do gráfico para o método de Newton-Raphson muda, o para o método da bisseção permanece exatamente igual.

Agora, analisando a variação do M ao varirmos T:



Para $m_0 = 1$ temos este resultado. Dele podemos tirar que a maior raiz, M, para $T \ll 1$, tende à 1, ao passo que para $T > 1$ é zero. O seguinte gráfico foi feito tomando $m_0 = T$:



Nele podemos ver, além de uma estrutura similar ao caso de $m_0 = 1$ mostrado anteriormente, traços do que parem ser as outras raízes da função. Elas seguem muito bem o esperado para as três raízes 0 , $+M$ e $-M$.