

MOWNIT

Laboratorium 1 Ćwiczenie 13

Jakub Karbowski

12 marca 2022

Przybliżoną wartość pochodnej funkcji $f(x)$ w punkcie x można obliczyć ze wzoru

$$f'(x) \approx f_h'(x) = \frac{f(x+h) - f(x)}{h}$$

Mając zadaną funkcję $f(x) = \sin(x) + \cos(3x)$, należało przeprowadzić analizę wpływu:

- precyzji liczb zmiennoprzecinkowych,
- oraz wartości h

na dokładność obliczonej pochodnej.

Parametry doświadczenia

Język programowania:

- Julia

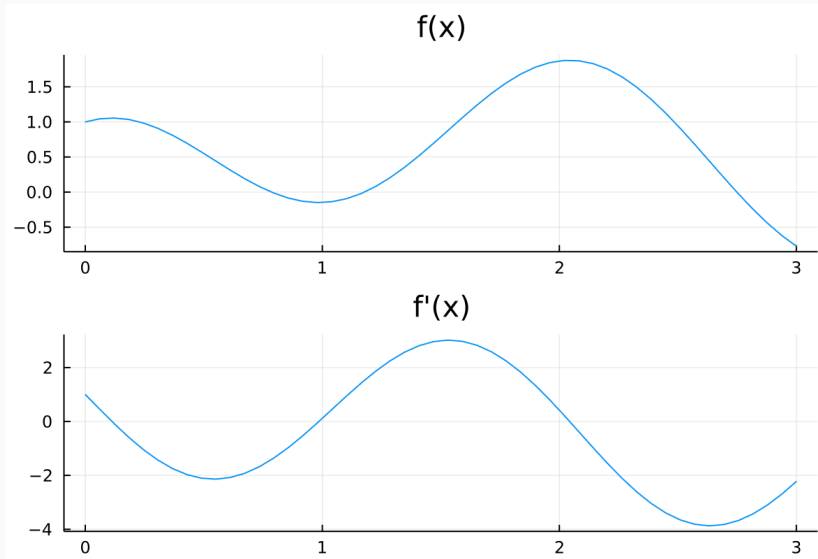
Sprawdzane wartości:

- $x = 1$
- $h = 2^{-n}$ ($n = 0, 1, 2, \dots, 40$)
- $h = 1 + 2^{-n}$

Sprawdzane typy:

- Float16
- Float32
- Float64

Badana funkcja



Rysunek 1: Badana funkcja

Liczmy ręcznie pochodną:

$$f(x) = \sin(x) + \cos(3x)$$

$$f'(x) = \cos(x) - 3 \sin(3x)$$

$$f'(1) = \cos(1) - 3 \sin(3)$$

$$\approx 0.11694228 \dots$$

Dokładna wartość nie jest liczbą wymierną.

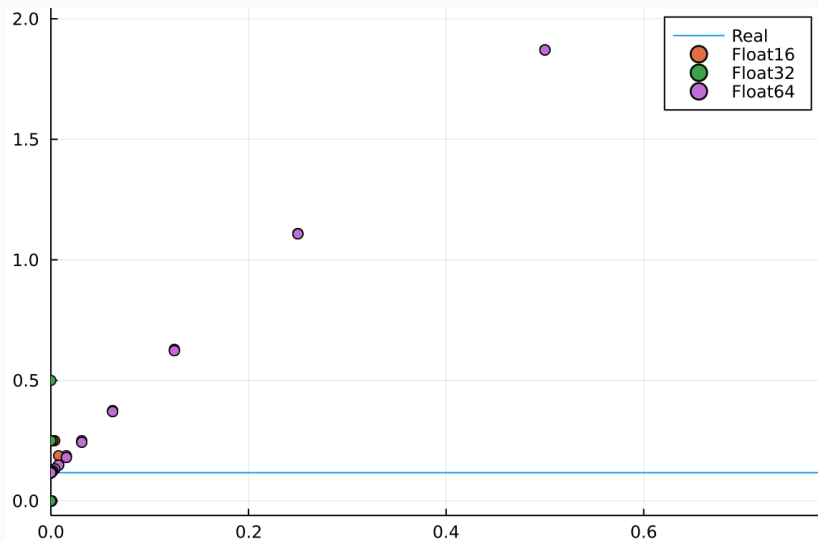
Epsilon maszynowy w okolicy $f'(1)$:

1. Float16 – $6.104 \cdot 10^{-5}$
2. Float32 – $7.450581 \cdot 10^{-9}$
3. Float64 – $1.3877787807814457 \cdot 10^{-17}$

Wartość $f'(1)$:

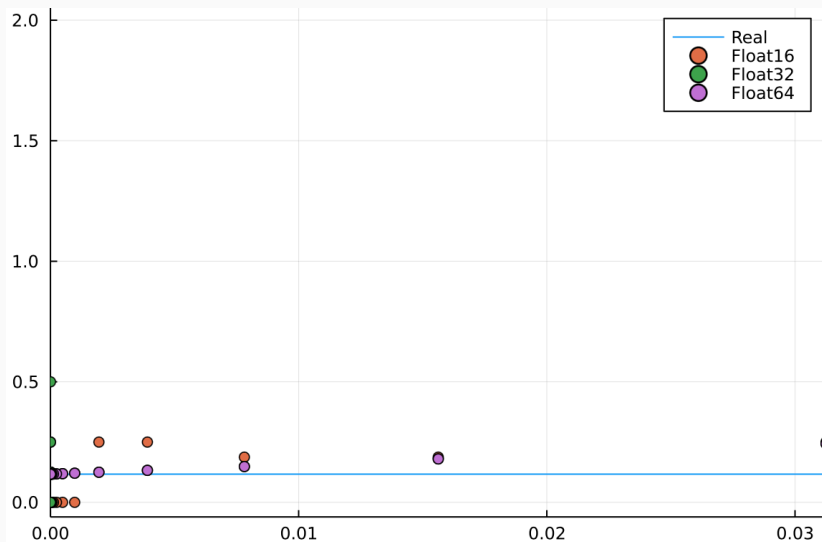
1. Float16 – 0.1172
2. Float32 – 0.11694229
3. Float64 – 0.11694228168853815

Zmniejszanie h



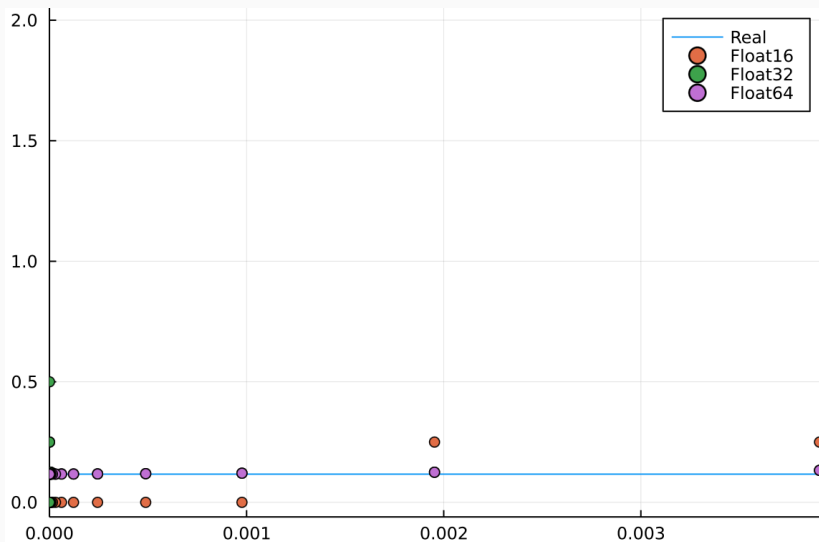
Rysunek 2: $h \in [0, 0.5]$

Float16 underflow



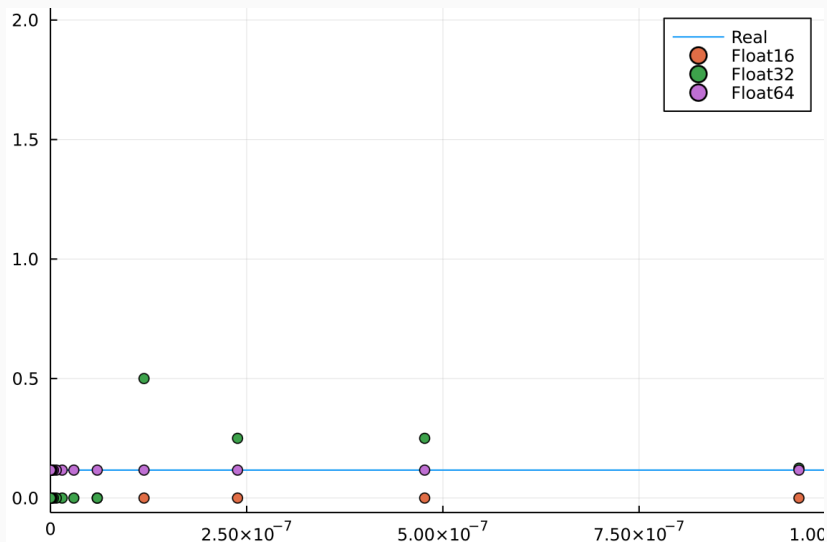
Rysunek 3: $h \in [0, 0.03]$

Float16 underflow



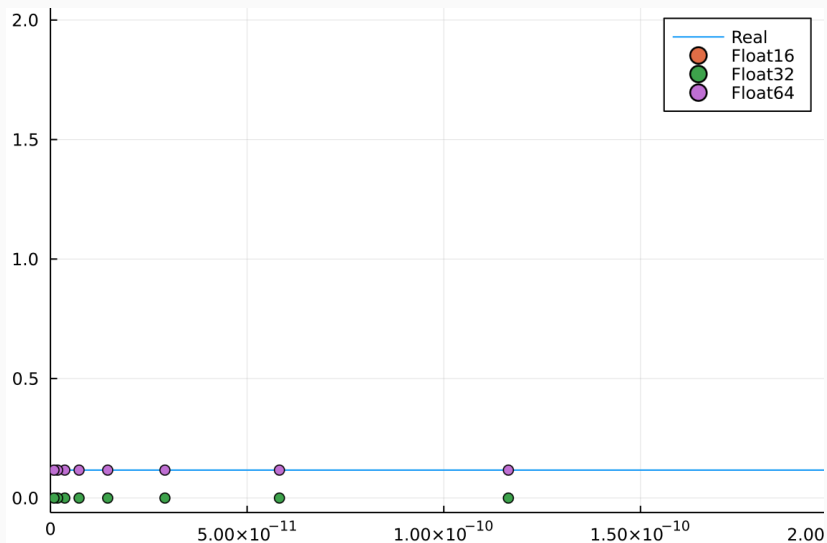
Rysunek 4: $h \in [0, 0.004]$

Float32 underflow



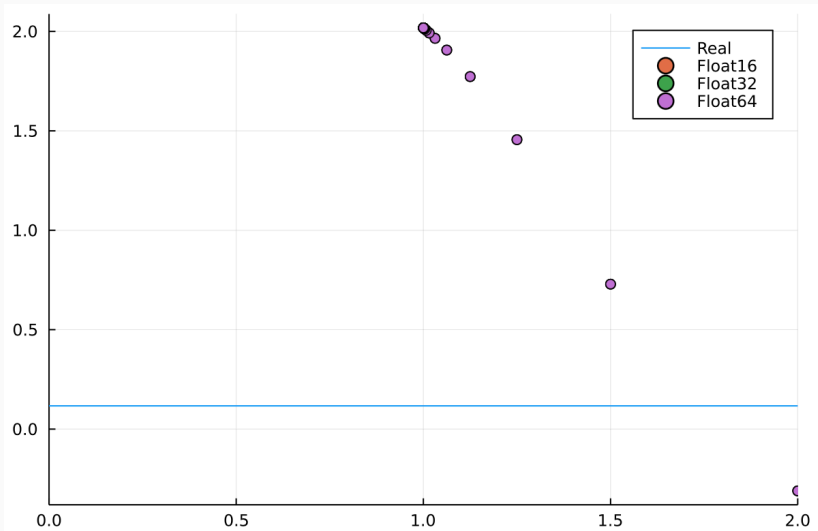
Rysunek 5: $h \in [0, 10^{-6}]$

Ustabilizowanie Float64



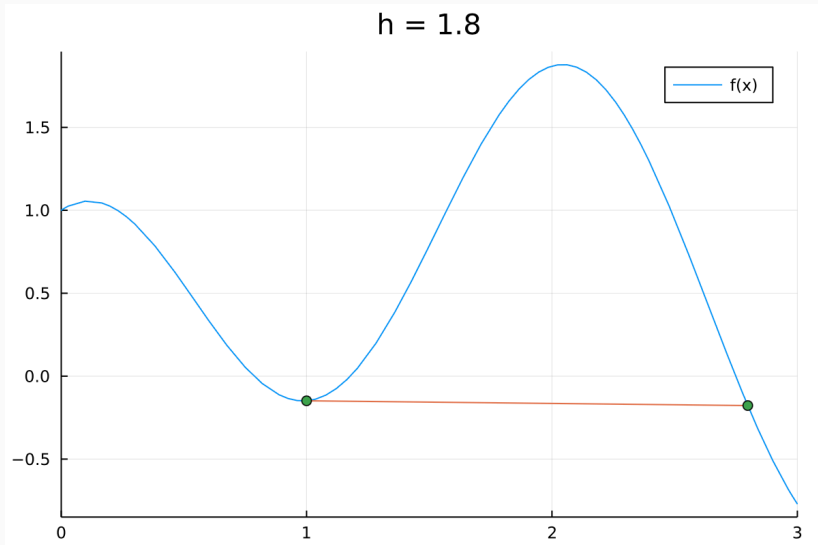
Rysunek 6: $h \in [0, 2 \times 10^{-10}]$

$$h + 1$$



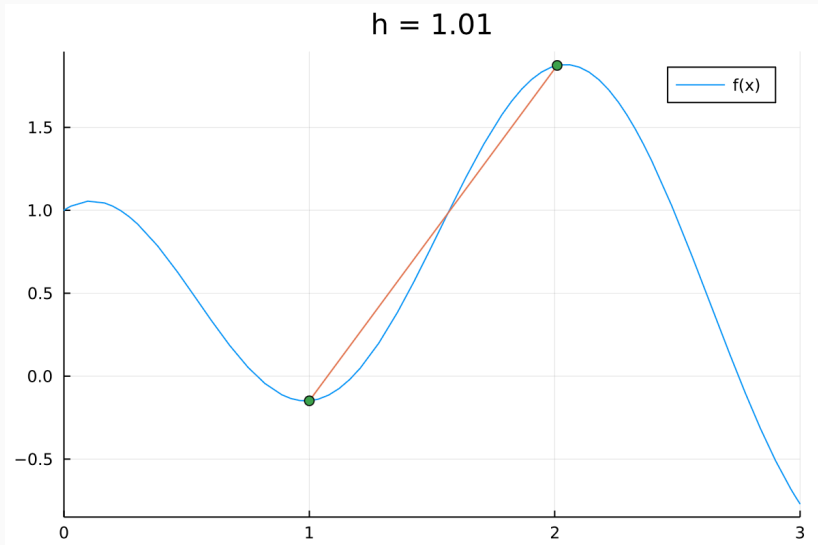
Rysunek 7: $h \in [1, 2]$

Dlaczego?



Rysunek 8: Wizualizacja pochodnej $h = 1.8$

Dlaczego?



Rysunek 9: Wizualizacja pochodnej $h = 1.01$

Tabela 1: h vs $f'(1)$

h	$f_h'(1)$ [Float64]	$f_h'(1)$ [Float32]	$f_h'(1)$ [Float16]
3.05e-5	0.11706539714577957	0.1171875	0.0
1.52e-5	0.11700383928837255	0.12109375	0.0
7.62e-6	0.11697306045971345	0.125	0.0
3.81e-6	0.11695767106721178	0.125	0.0
1.90e-6	0.11694997636368498	0.125	0.0
9.53e-7	0.11694612901192158	0.125	0.0
4.76e-7	0.1169442052487284	0.25	0.0
2.38e-7	0.11694324295967817	0.25	0.0
1.19e-7	0.11694276239722967	0.5	0.0
5.96e-8	0.11694252118468285	0.0	0.0
2.98e-8	0.116942398250103	0.0	NaN
1.49e-8	0.11694233864545822	0.0	NaN
7.45e-9	0.11694231629371643	0.0	NaN

Tabela 2: h vs $f'(1)$ (c.d.)

h	$f_h'(1)$ [Float64]	$f_h'(1)$ [Float32]	$f_h'(1)$ [Float16]
3.72e-9	0.11694228649139404	0.0	NaN
1.86e-9	0.11694222688674927	0.0	NaN
9.31e-10	0.11694216728210449	0.0	NaN
4.65e-10	0.11694216728210449	0.0	NaN
2.32e-10	0.11694192886352539	0.0	NaN
1.16e-10	0.11694145202636719	0.0	NaN
5.82e-11	0.11694145202636719	0.0	NaN
2.91e-11	0.11693954467773438	0.0	NaN
1.45e-11	0.116943359375	0.0	NaN
7.27e-12	0.1169281005859375	0.0	NaN
3.63e-12	0.116943359375	0.0	NaN
1.81e-12	0.11688232421875	0.0	NaN
9.09e-13	0.1168212890625	0.0	NaN

Tabela 3: Float64 underflow

h	$f_h'(1)$ [Float64]
7.27e-12	0.1169281005859375
3.63e-12	0.116943359375
1.81e-12	0.11688232421875
9.09e-13	0.1168212890625
$f'(1)$	0.11694228168853815

Dla najmniejszych sprawdzanych wartości h widać oddalenie się od wartości dokładnej $f'(1)$, mimo zmniejszania h . Sugeruje to wystąpienie underflow.

1. Nie należy stosować bardzo małych h bez uzasadnienia.
2. Należy zwrócić uwagę na kształt funkcji przy doborze h .
3. Szersze typy są bezpieczniejsze, ale nie niezawodne.