Błędy Numeryczne

Metody Numeryczne

dr inż. Grzegorz Fotyga

Politechnika Gdańska
Wydział Elektroniki, Telekomunikacji i Informatyki
Katedra Inżynierii Mikrofalowej i Antenowej

18 maja 2018

Plan prezentacji

Porażka systemu Patriot (1)

- 25 II 1991, Wojna w zatoce Perskiej
- Dharan, Arabia Saudyjska
- System Patriot chroniący amerykańską bazę nie wykrył i nie przechwycił rakiety typu Scud
- Zginęło 28 amerykańskich żołnierzy, 100 zostało rannych.

http://www-users.math.umn.edu/~arnold/disasters/patriot.html
https://hownot2code.com/2016/11/09/r-17-vs-patriot-a-rounding-issue-bugs-in-a-missile-defense-system/

Porażka systemu Patriot (2) - baza

https://www.reddit.com/r/todayilearned/comments/24s9wi/til_that_in_february_1991_28_american_soldiers/

Porażka systemu Patriot (3) - przyczyna

- Czas jako zmienna *int* przechowywana w **24-bitowym rejestrze**,
- inkrementacja co 0.1s
- 0.1 reprezentowane za pomocą 24 bitów: 0.00011001100110011001100
- Błąd zaokrąglenia: 0.0000000000000000000000011001100 (binarnie), 0.000000095s (dziesiętne).
- 100 godzin $\Rightarrow$ błąd **narasta** do **0.34s.**
- Prędkość R-17: ok. 1600m/s, co daje **błąd ok. 500m** w algorytmie lokalizacji systemu Patriot.
- Następnego dnia (26 II 1991) ***bug fix*** dotarł do Dharan.

https://hownot2code.com/2016/11/09/r-17-vs-patriot-a-rounding-issue-bugs-in-a-missile-defense-system/

Eksplozja Ariane 5

- Wybuch rakiety Ariane 5 w 1996r
- Strata: 7 mld$, 10 lat przygotowań
- Przyczyna: konwersja z 64 -bitowej zmiennej typu *float* do 16-bitowej zmiennej typu *int*
- **Nadmiar** (ang. overflow)

https://prosesosis.wordpress.com/2015/05/13/numeros-simples-hacen-colapsar-a-muchas-computadoras/
http://www-users.math.umn.edu/~arnold/disasters/ariane.html

Błąd względny i bezwzględny

- **Błąd bezwzględny** (ang. absolute error):

$$\Delta x = x - \widetilde{x} \tag{1}$$

gdzie x jest wartością dokładną, a $\widetilde{x}$ przybliżoną

- **Błąd względny** (ang. relative error):

$$\delta_x = (\Delta x)/x \tag{2}$$

- często wyrażany w %:

$$\delta_x = (\Delta x)/x \cdot 100\% \tag{3}$$

dla $x \neq 0$

Błąd względny i bezwzględny - przykład

Prędkość światła: $c = 299798458 m/s$, $\widetilde{c} = 3e8 m/s$.

- $\Delta c = -201542$
- $\delta_c = -0.067226\%$

Wzór na prędkość światła: $c = \dfrac{1}{\sqrt{\mu_0 \epsilon_0}}$, gdzie:

Przenikalność elektryczna: $\epsilon_0 = 8.854187817e\text{-}12\ F/m$, $\quad \widetilde{\epsilon_0} = \dfrac{1}{36\pi}e\text{-}9\ F/m$.

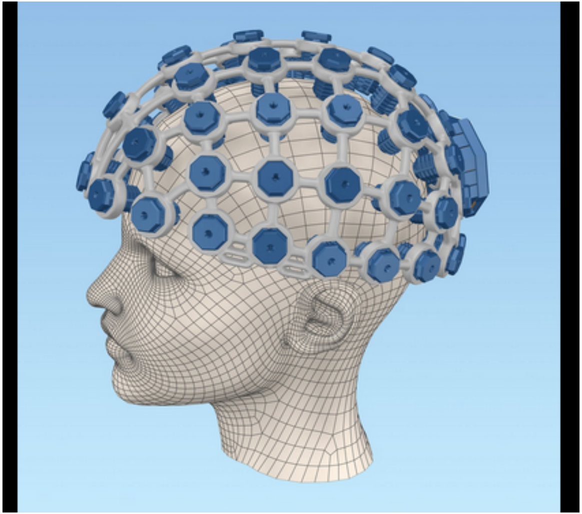
- $\Delta\epsilon_0 = -1.2247e\text{-}14$
- $\delta_\epsilon = -0.13831\%$

Przenikalność magnetyczna: $\mu_0 = 12.566370614e\text{-}7\ H/m$, $\quad \widetilde{\mu_0} = 4\pi e\text{-}7\ H/m$.

- $\Delta\mu_0 = -3.5917e\text{-}17$
- $\delta_{\mu_0} = -2.8582e\text{-}09\%$

Błąd względny i bezwzględny - przykład 2

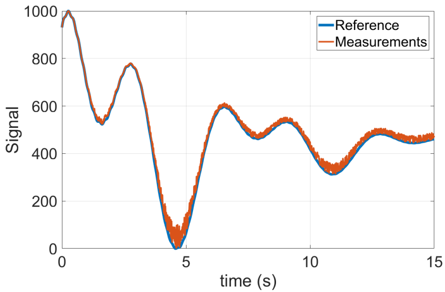
Test kasku z elektrodami do EEG-BCI

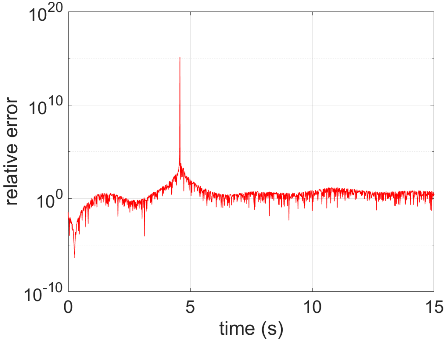


https://www.fastcompany.com/3053063/new-kickstarter-eeg-kit-promises-cheap-mind-reading

Błąd względny i bezwzględny - przykład 2

Wykresy mocy sygnału odbiornika referencyjnego (o dużej czułości) i odbiornika testowanego. Pomiar trwa 15 s. **Błąd względny.**

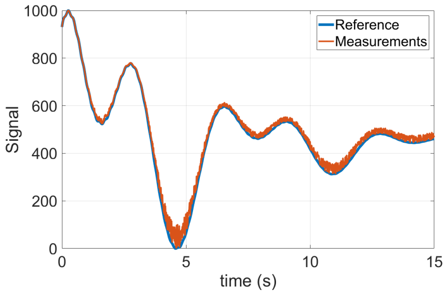


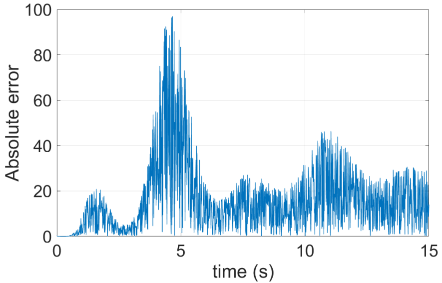


Kask jest **bezużyteczny**.

Błąd względny i bezwzględny - przykład 2

Wykresy mocy sygnału odbiornika referencyjnego (o dużej czułości) i odbiornika testowanego. Pomiar trwa 15 s. **Błąd bezwzględny.**





Kask jest **OK**.

Zapis zmiennopozycyjny (1)

- Liczby rzeczywiste (zbiór $\mathbb{R}$) są zapisywane w większości współczesnych maszyn cyfrowych w tzw. postaci *zmiennopozycyjnej*
- Tworzą one podzbiór zbioru $\mathbb{R}$, oznaczony jako **F**.
- Są definiowane jako:

$$x = M \times N^W \tag{4}$$

- M - **mantysa** (część ułamkowa)
- W - **wykładnik** - liczba całkowita, decyduje o zakresie liczb dopuszczalnych w danej maszynie
- $N = 2$
- Liczba rzeczywista jest więc reprezentowana za pomocą dwóch grup bitów.
- Dwa ważne ograniczenia reprezentacji zmiennoprzecinkowej:
  - reprezentowane liczby nie mogą być dowolnie **wielkie**/**małe**,
  - są między nimi **odstępy**.

Zapis zmiennopozycyjny (2) - Przykład

Przykład

- Mantysa określona za pomocą 5 bitów, wykładnik za pomocą 3, gdzie pierwszy bit oznacza znak (1 – liczba ujemna):

$$x = \underset{M}{(1)1101} \quad \underset{W}{(0)10} \tag{5}$$

- (5) w zapisie dwójkowym:

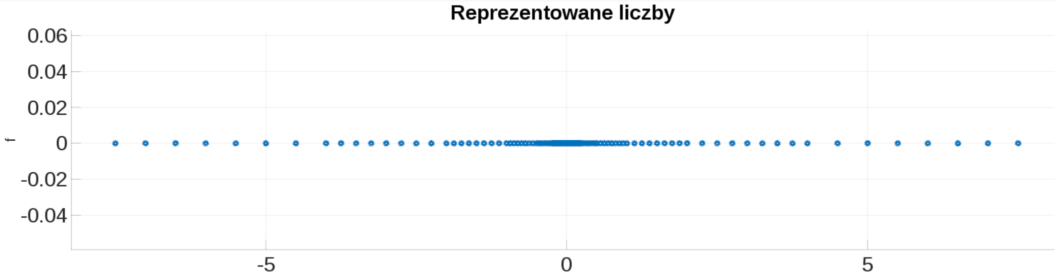
$$-0.1101 \times 2^{+10} = -\left(\frac{1}{2} + \frac{1}{4} + \frac{0}{8} + \frac{1}{16}\right) \times 2^{+(1\cdot 2+0\cdot 1)} \tag{6}$$

- (5) w zapisie dziesiętnym:

$$-3.25 \tag{7}$$

Zapis zmiennopozycyjny (3)

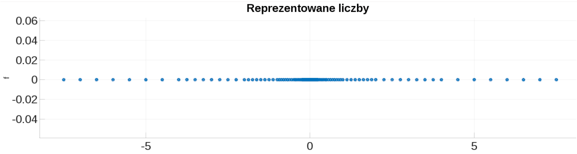
- W powyższym przykładzie możemy wykonywać obliczenia za pomocą tylko niektórych liczb dodatnich i ujemnych w zakresie $\langle -7.5, 7.5 \rangle$:



Zapis zmiennopozycyjny (4) - reprezentacja liczb rzeczywistych w **F**

Operacje na liczbach rzeczywistych:

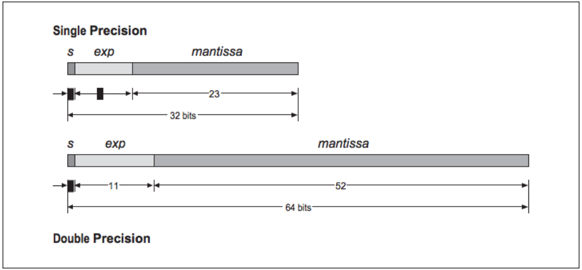
- Reprezentacja liczby $x_R \in \mathbb{R} = 6.26$ w **F**: $x_F = 6.5$ ($x_F = (0)1101(0)11$), błąd względny: $\delta = 0.038339$,
- Reprezentacja liczby $x_R \in \mathbb{R} = 0.2$ w **F**: $x_F = 0.1875$ ($x_F = (0)1100(1)10$), błąd względny: $\delta = 0.0628$,



Zapis zmiennopozycyjny (5) - IEEE

Standard **IEEE** definiowania liczb zmiennoprzecinkowych

- Pojedyncza precyzja:
  - 32 bity
  - 23 mantysa
  - 8 wykładnik
- Podwójna precyzja:
  - 64 bity
  - 52 mantysa
  - 11 wykładnik



Zapis zmiennopozycyjny (6) - Podwójna precyzja

Specyfika definiowania liczb zmiennoprzecinkowych w podwójnej precyzji

- Największa liczba: $1.79 \times 10e308$
- Najmniejsza liczba: $2.23 \times 10e\text{-}308$
- *Rozdzielczość* podzbioru **F** jest reprezentowana za pomocą **epsilonu maszynowego**:

$$\epsilon_M = 2^{-52} \approx 2.22 \times 10^{-16} \tag{8}$$

- Rozpatrzmy funkcję *fl*: $\mathbb{R} \to \mathbf{F}$, która przyporządkowuje każdej liczbie rzeczywistej najbliższą wartość z podzbioru **F**.
- dla każdego $x \in \mathbb{R}$ istnieje $\widetilde{x} \in \mathbf{F}$ taki, że:

$$|x - \widetilde{x}| \leqslant \epsilon_M |x| \tag{9}$$

- dla każdego $x \in \mathbb{R}$ istnieje ϵ (gdzie $|\epsilon| \leqslant \epsilon_M$) taki, że:

$$fl(x) = x(1 + \epsilon) \tag{10}$$

Zapis zmiennopozycyjny (7)

Specyfika **działań** na liczbach zmiennoprzecinkowych

- Wszystkie operacje matematyczne wykonywane na komputerze można sprowadzić do **elementarnych** operacji arytmetycznych: $+, -, \times$ i $\div$.
- Jednak w arytmetyce zmiennoprzecinkowej korzystamy z przestrzeni **F**, w której istnieją analogiczne działania: $\oplus, \ominus, \otimes$ i $\oslash$.
- Dla $x, y \in \mathbf{F}$:
  - $x \oplus y = fl(x + y)$
  - $x \ominus y = fl(x - y)$
  - $x \otimes y = fl(x \times y)$
  - $x \oslash y = fl(x \div y)$

Zapis zmiennopozycyjny (8) - Lemat Wilkinsona

Specyfika **działań** na liczbach zmiennoprzecinkowych

- $x, y \in \mathbf{F}$. Błąd wprowadzany przez działania $\oplus, \ominus, \otimes$ i $\oslash$, względem $+, -, \times$ i $\div$ jest następujący (**lemat Wilkinsona**):
  - $x \oplus y = x(1+\epsilon) + y(1+\epsilon)$
  - $x \ominus y = x(1+\epsilon) - y(1+\epsilon)$
  - $x \otimes y = (x \times y)(1+\epsilon)$
  - $x \oslash y = (x \div y)(1+\epsilon)$
- gdzie $|\epsilon| \leqslant \epsilon_M$.
- **Wniosek**: Każda operacja arytmetyczna w przestrzeni **F** jest wykonywana z dokładnością do błędu względnego nie większego niż ϵ_M.
- Fakt ten powoduje powstawanie tzw. błędów zaokrągleń (ang. round-off errors).

Błąd zaokrągleń (1)

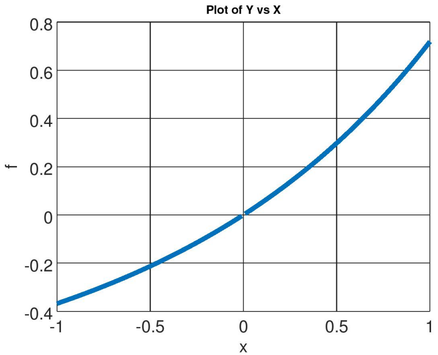
```
% round off errors - przyklad 1
sum = 0;
for i = 1:1000000
   sum = sum + 0.001;
end
sum = 999.999999983265

% round off errors - przyklad 2
x = 1e3;
x - (x + (1e-17)) = 0
```

Błąd zaokrągleń (2) - przykład

- Funkcja: $f(x) = \dfrac{(e^x - 1)}{x} - 1$
- Granica funkcji $f(x)$ przy $x \to 0$ daje wyrażenie nieoznaczone: $0/0$
- Korzystając z reguły de l'Hospitala:
- $\lim\limits_{x\to 0} \dfrac{g(x)}{h(x)} = \lim\limits_{x\to 0} \dfrac{g'(x)}{h'(x)} = \lim\limits_{x\to 0} \dfrac{e^x - 1}{1} = 0$

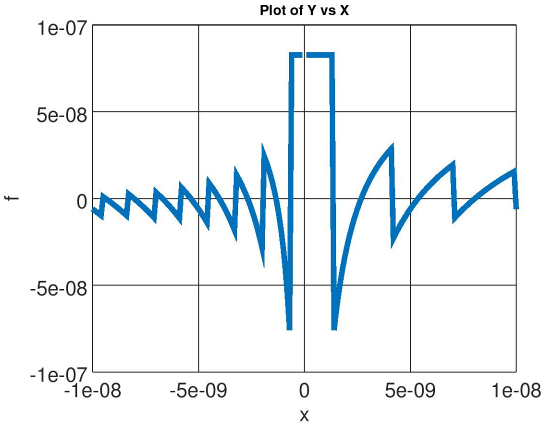
```
x = -1:(1e-2):1;
f = (exp(x) - 1)./x-1;
plot(x,f,'linewidth',5);
xlabel('x','fontsize',20)
ylabel('f','fontsize',20)
title('Plot of Y vs X','fontsize',14)
```



Błąd zaokrągleń (3) - przykład

Dziwne zjawisko spowodowane **błędem zaokrąglenia** w pobliżu 0.

```
x = (-1e-8):(1e-10):(1e-8);
f = (exp(x) - 1)./x-1;
plot(x,f,'linewidth',5);
xlabel('x','fontsize',20)
ylabel('f','fontsize',20)
title('Plot of Y vs X','fontsize',14)
```



Dygresja o kropce

Czym się różnią w Matlabie (Octave) działania:

- ^ i /
- .^ i ./ ?

```
>> x = rand(1,3)
       x = 0.088235    0.862485    0.325445
>> 2/x
       error: operator / (dzielenie przez wektor!)
>> 2./x
       ans = 22.6667     2.3189     6.1454
>> 2^x
       error: for x^A, Use .^ for elementwise power.
>> 2.^x
       ans = 1.0631    1.8182    1.2531
```

Błąd danych wejściowych / błąd zaokrągleń (1)

1. **Błąd danych** - błąd, którym obarczone są dane wejściowe zadań numerycznych.
2. **Źródła błędów** danych wejściowych:
   - błędy pomiarów,
   - błędy z wcześniejszej fazy obliczeń,
   - zastępowanie wartości stałych bedących liczbami niewymiernymi ich przybliżeniami
3. Błędy te są zwykle **niewielkie**, jednak nawet małe względne błędy danych wejściowych zadania numerycznego mogą skutkować **dużymi** błędami względnymi wyników.

Błąd danych wejściowych / błąd zaokrągleń (2)

Układ trzech równań

Przypadek a)

$$\begin{cases} 1x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 1 \\ \frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = 2 \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = 3 \end{cases}$$

Przypadek b)

$$\begin{cases} 1.00x_1 + 0.50x_2 + 0.33x_3 = 1 \\ 0.50x_1 + 0.33x_2 + 0.25x_3 = 2 \\ 0.33x_1 + 0.25x_2 + 0.20x_3 = 3 \end{cases}$$

Błąd danych wejściowych / błąd zaokrągleń (3)

Wyniki:

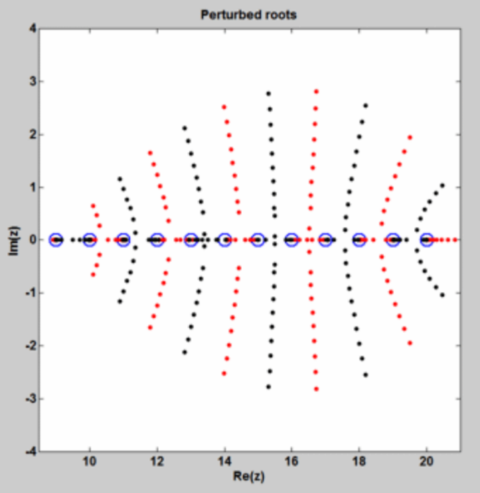
- **Błąd względny między** 0.33 **i** $1/3$**:**
  $\delta = 1\%$
- Przypadek a) $x = [x_1 \;\; x_2 \;\; x_3] = [27.0 \;\; -192.0 \;\; 210.0]^T$
- Przypadek b) $x = [266.66 \;\; -1433.33 \;\; 1366.67]^T$
- **Błąd względny:**
  $\delta = [-887.65\% \;\; -646.53\% \;\; -550.79\%]^T$
- Mały błąd względny danych wejściowych powoduje ogromny błąd wyniku działania algorytmu. Oznacza to, że problem jest **źle uwarunkowany**.

Błąd danych wejściowych / błąd zaokrągleń (4)

- **Rozpatrzmy wielomian:** $P20 = (x-1)\cdot(x-2)\cdot(x-3)\cdot(x-4)\cdot(x-5)\cdot(x-6)\cdot(x-7)\cdot(x-8)\cdot(x-9)\cdot(x-10)\cdot(x-11)\cdot(x-12)\cdot(x-13)\cdot(x-14)\cdot(x-15)\cdot(x-16)\cdot(x-17)\cdot(x-18)\cdot(x-19)\cdot(x-20)$.
- **Miejsca zerowe** $P20$: $[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]$
- Wprowadzamy **zaburzenie**: $P20 - 10^{-8}x^{19}$
- **Miejsca zerowe** $P20 - 10^{-8}x^{19}$: $[1.0000, 2.00000, 3.00000, 4.00000, 5.00000, 6.00000]$, $[6.99997, 8.00060, 8.99177, 10.0925, 10.6779, 20.2403]$, $[12.14447 - 0.90522i, 12.14447 + 0.90522i]$

Błąd danych wejściowych / błąd zaokrągleń (5)

Wykres miejsc zerowych wielomianu Wilkinsona dla zaburzeń mniejszych niż 1.1921e-07 (2^{-23}) (płaszczyzna zespolona!)



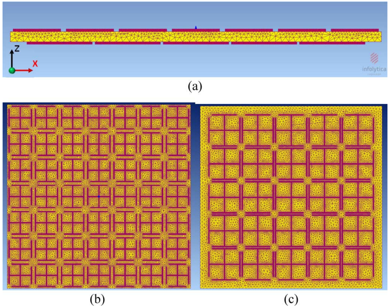
Błędy odcięcia

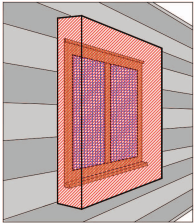
Błędy odcięcia (ang. truncation error), np.:

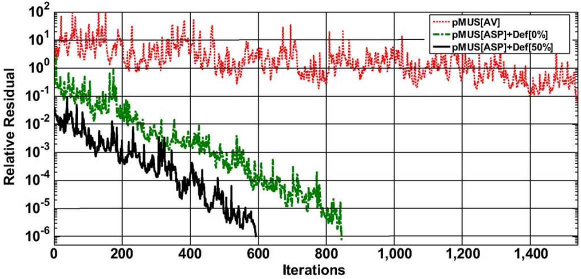
1. przybliżenie nieskończonej sumy szeregu sumą skończonej liczby jego elementów.
2. przybliżenie pochodnej ilorazem różnic skończonych.
3. przybliżenie całki sumą skończoną.

Jeszcze do tego wrócimy

Błędna metoda (1)







ieeexplore.ieee.org/document/7366073/
ieeexplore.ieee.org/document/777134/

Wybory w Niemczech, partia Zielonych (1)

- Niedziela. Wyniki wyborów: **partia Zielonych: 5%,** wchodzi do sejmu
- Poniedziałek. Błąd programu komisji wyborczej, **jednak 4.97%**.
- Błędne zaokrąglenie do 1 miejsca po przecinku.

https://catless.ncl.ac.uk/Risks/13.37.html https://www.stern.de/politik/deutschland/die-gruenen--vier-gruende--warum-man-sie-einfach-nicht-mehr-waehlen-kann-7418740.html