Math 4513
Solutions to Homework Set 1

1.1. Given that

$$\frac{d^n}{dx^n}(\ln|x|) = (-1)^{n-1}(n-1)!x^{-n+1}$$

(a) Use the Taylor Theorem with Integral Remainder to find the magnitude of the error term $R_{100}(1.99)$ when one approximates $\ln[1.99]$ using the first 101 terms of the Taylor expansion about 1 of $\ln|x|$.

- We have

$$\ln|x| = \sum_{n=0}^{N} \frac{1}{n!}\frac{d^n}{dx^n}(\ln|x|)\bigg|_{x=1}(x-1)^n + R_N(x)$$

When $N = 100$, the error term is $R_{100}(x)$ given by

$$\begin{aligned} R_{100}(x) &= \frac{1}{100!}\int_1^x \frac{d^{101}}{dt^{101}}(\ln|t|)\bigg|_t (x-t)^{100}dt \\ &= \frac{1}{100!}\int_1^x (-1)^{100}(100!)t^{-100}(x-t)^{100}dt \end{aligned}$$

For $x = 1.99$ we have

$$\begin{aligned} R_{100}(1.99) &= \int_1^{1.99} t^{-100}(1.99-t)^{100}dt \\ &= -.4357523477\times 10^{-35} \end{aligned}$$

□

(b) Use the Taylor Theorem with Lagrange Remainder to obtain an upper bound on the error term $E_{100}(x)$ when x ranges from 1.985 to 1.995 for the Taylor expansion of $\ln|x|$ about 1.

- We have

$$\ln|x| = \sum_{n=0}^{N} \frac{1}{n!}\frac{d^n}{dx^n}(\ln|x|)\bigg|_{x=1}(x-1)^n + E_N(x)$$

When $N = 100$ the error term $E_{100}(x)$ is given by

$$\begin{aligned} E_{100}(x) &= \frac{1}{(101)!}\frac{d^{101}}{dt^{101}}(\ln|t|)\bigg|_{t=\xi}(x-1)^{101} \\ &= \frac{1}{101!}\left((-1)^{100}(100!)\xi^{-100}\right)(x-1)^{101} \\ &= \frac{1}{101}\xi^{-100}(x-1)^{101} \end{aligned}$$

for some $\xi \in [1, x]$. When $x = 1.99$ we have

$$E_{100}(1.99) = \frac{1}{101}(0.99)^{101}\xi^{-100}$$

Clearly, the maximal value of this error term on the interval $[1, 1.99]$ occurs when $\xi = 1$. so

$$\max E_{100}(1.99) = \frac{1}{101}(0.99)^{101} = 3.5878\times 10^{-3}$$

□

1.2. Show that if $b_n = \mathcal{O}(a_n)$ then $b_n/\ln|n| = \mathfrak{o}(a_n)$.

- Since $b_n = \mathcal{O}(a_n)$

$$b_n \leq C a_n \text{ for sufficiently large } n$$

Let $c_n = b_n / \ln |n|$. To see if $c_n = \mathfrak{o}(a_n)$ we look at

$$\lim_{n\to\infty} \frac{|c_n|}{|a_n|}$$

Now

$$0 \leq \frac{|c_n|}{|a_n|} = \frac{|b_n|}{|a_n| \ln |n|} \leq \frac{C}{\ln |n|} \quad \text{for sufficient large } n$$

But now

$$\lim_{n\to\infty} \frac{C}{\ln |n|} = 0 \quad \Rightarrow \quad \lim_{n\to\infty} \frac{|c_n|}{|a_n|} = 0$$

so $c_n = \mathfrak{o}(a_n)$. $\square$

1.3. Show that if $b_n = \mathfrak{o}(a_n)$ then $b_n = \mathcal{O}(a_n)$, but that the converse is not true.

- Since $b_n = \mathfrak{o}(a_n)$

$$\lim_{n\to\infty} \frac{|b_n|}{|a_n|} = 0$$

By definition, this means that for any number C there exists a number N such that $n > N$ implies $|b_n| / |a_n| < C$. Or

$$|b_n| < C |a_n| \quad \text{for sufficiently large } n$$

Hence, $b_n = \mathfrak{o}(a_n)$ implies $b_n = \mathcal{O}(a_n)$.

To see that the converse is not true, take $b_n = a_n$. Then $b_n = \mathcal{O}(a_n)$ (since $b_n \leq (1)a_n$ for all $n > 1$). But

$$\lim_{n\to\infty} \frac{|b_n|}{|a_n|} = \lim_{n\to\infty} \frac{|a_n|}{|a_n|} = 1 \neq 0$$

so $b_n \neq \mathfrak{o}(a_n)$. $\square$

1.4. Show that if $b_n = \mathcal{O}(a_n)$ and $c_n = \mathcal{O}(a_n)$, then $b_n + c_n = \mathcal{O}(a_n)$.

- If $b_n = \mathcal{O}(a_n)$ and $c_n = \mathcal{O}(a_n)$, then there exist constants C_1, C_2 and numbers N_1, N_2 such that

$$\begin{aligned} |b_n| &\leq C_1 |a_n| \quad , \quad n > N_1 \\ |c_n| &\leq C_2 |a_n| \quad , \quad n > N_2 \end{aligned}$$

But then

$$|b_n + c_n| \leq (C_1 + C_2) |a_n| \quad , \quad \text{for all } n \text{ greater than } \max(N_1, N_2)$$

Hence, by taking $C = C_1 + C_2$, $N = \max(N_1, N_1)$ we see that $b_n + c_n = \mathcal{O}(a_n)$. $\square$

1.5. Show that if $b_n = \mathfrak{o}(a_n)$ and $c_n = \mathfrak{o}(a_n)$, then $b_n + c_n = \mathfrak{o}(a_n)$.

- If $b_n = \mathfrak{o}(a_n)$ and $c_n = \mathfrak{o}(a_n)$,

$$0 = \lim \frac{|b_n|}{|a_n|} = \lim \frac{|c_n|}{|a_n|}$$

and so, using the triangle inequality and the fact that the limit of a sum of convergent sequences is equal to the sum of their limits we have

$$\lim_{n\to\infty} \frac{|b_n + c_n|}{|a_n|} \leq \lim_{n\to\infty} \left(\frac{|b_n|}{|a_n|} + \frac{|c_n|}{|a_n|} \right) = \lim_{n\to\infty} \frac{|b_n|}{|a_n|} + \lim_{n\to\infty} \frac{|c_n|}{|a_n|} = 0 + 0 = 0$$

so $b_n + c_n = \mathfrak{o}\,(a_n)$. $\square$

1.6. Show that for any $r > 0$, $x^r = \mathcal{O}(e^x)$ as $x \to \infty$.

- We need to show taht we can always find a constant C and a number N such that

$$|x^r| \leq C\,|e^x| \quad , \quad \text{for all } x > N$$

Without loss of generality we can assume x is positive and greater than 1. If we take the log both sides of this inequality (which maintains the inequality since $\ln|x|$ is monotonically increasing for $x > 1$) we get

$$r \ln|x| \leq \ln|C| x$$

or

$$C \geq r \frac{\ln|x|}{x}$$

Applying l'Hospital's rule to the right hand side one can see that

$$\lim_{x\to\infty} \frac{\ln|x|}{x} = \lim_{x\to\infty} \frac{\frac{1}{x}}{1} = \lim_{x\to\infty} \frac{1}{x} = 0$$

So for large enough x we can guarantee that $C \geq r\frac{\ln|x|}{x}$ for any constant $C > 0$. Hence we can maintain

$$|x^r| \leq C\,|e^x| \quad , \quad \text{for all sufficiently large } x$$

so $x^r = \mathcal{O}(e^x)$. $\square$

1.7. Show that for any $r > 0, \ln|x| = \mathcal{O}(x^r)$ as $x \to \infty$.

- We need to show that

$$\ln|x| \leq C\,|x^r|$$

for all suficiently large x. Consider

$$\lim_{x\to\infty} \frac{\ln|x|}{x^r}$$

applying l'Hospital's rule

$$\lim_{x\to\infty} \frac{\ln|x|}{x^r} = \lim_{x\to\infty} \frac{\frac{1}{x}}{rx^{r-1}} = \lim_{x\to\infty} \frac{1}{r} x^{-r} = 0 \quad \text{if } r > 0$$

Therefore, we can make the ratio

$$\frac{\ln|x|}{x^r} < C$$

for any positive number C by taking x sufficiently large. Hence,

$$\ln|x| < Cx^r \quad \text{for } x \text{ sufficiently large} \quad \Rightarrow \quad \ln|x| = \mathcal{O}(x^r)$$

$\square$

1.8. If $1/10$ is correctly rounded to the normalized binary number $(1.a_1a_2\ldots a_{23})_2 \times 2^m$, what is the roundoff error? What is the relative roundoff error?

- As demonstrated in class

$$\begin{aligned}\left(\frac{1}{10}\right)_{10} &= 0.0001100110011001100110011001100110011\overline{0011} \\ &= (1.10011001100110011001100\ldots) \times 2^{-4}\end{aligned}$$

Keeping only 23 bits past the decimal point we obtain the machine number that's obtained by rounding down is

$$\begin{aligned} q_- &= (1 + 2^{-1} + 2^{-4} + 2^{-5} + 2^{-8} + 2^{-9} + 2^{-12} + 2^{-13} + 2^{-16} + 2^{-17} + 2^{-20} + 2^{-21}) * 2^{-4} \\ &= (0.09999999406)_{10}\end{aligned}$$

The next machine number up will be

$$\begin{aligned} q_+ &= q_- + (0.00000000000000000000001) \times 2^{-4} \\ &= q_- + 2^{-27} \\ &= (0.1000000015)_{10}\end{aligned}$$

We have

$$\begin{aligned}\left|\frac{1}{10} - q_-\right| &= 5.94 \times 10^{-9} \\ \left|\frac{1}{10} - q_+\right| &= 1.5 \times 10^{-9}\end{aligned}$$

And so q_+ is the closest machine number, and the roundoff error is 1.5×10^{-9}. The relative roundoff error is

$$\frac{\left|\frac{1}{10} - q_+\right|}{\frac{1}{10}} = 1.5 \times 10^{-8}$$

□

1.9. Give examples of real numbers for which

$$fl\,(x \odot y) \neq fl\,(fl(x) \odot fl(y))$$

Illustrate this for all four arithmetic operators $(+, -, \times, \div)$ using a hypothetical machine with 5 decimal (base t0) digits.

- Addition: Let $x = 1.000014$ and $y = 1.000014$. Then

$$x + y = 1.000028$$

and

$$\begin{aligned} fl(x) &= fl(y) = 1.00001 \\ fl(x) + fl(y) &= 2.00002 \\ fl\,(fl(x) + fl(y)) &= 2.00002\end{aligned}$$

but

$$fl(x+y) = 1.00003 \neq 2.00002 = fl\,(fl(x) + fl(y))$$

- Subtraction:

$$\begin{aligned} x &= 1.000046 \\ y &= 1.000054 \\ fl(x) &= fl(y) = 1.00005 \\ fl(x) - fl(y) &= 0.00000 \\ fl\,(fl(x) - fl(y)) &= 0.00000 \\ x - y &= -0.000008 \\ fl(x-y) &= -0.00001\end{aligned}$$

$$fl(x-y) = -0.00001 \neq 0.00000 = fl\left(fl(x) - fl(y)\right)$$

- Multiplication

$$\begin{aligned}
x &= 1.000004 \\
y &= 1.000004 \\
fl(x) &= 1.00000 \\
fl(y) &= 1.00000 \\
fl(x) \times fl(y) &= 1.00000 \\
fl\left(fl(x) \times fl(y)\right) &= 1.00000 \\
x \times y &= 1.0000800016 \\
fl(x \times y) &= 1.00008
\end{aligned}$$

$$fl(x \times y) = 1.00008 \neq 1.00000 = fl\left(fl(x) \times fl(y)\right)$$

- Division

$$\begin{aligned}
x &= 0.0000125 \\
y &= 1.000000 \\
f(x) &= 0.00001 \\
fl(y) &= 1.00000 \\
fl(y)/fl(x) &= 1.00000 \times 10^5 \\
fl\left(fl(y)/fl(x)\right) &= 1.00000 \times 10^5 \\
y/x &= 80,0000 \\
fl(y/x) &= 8.0 \times 10^4
\end{aligned}$$

$$fl(y/x) = 8.0 \times 10^4 \neq 1.00000 \times 10^5 = fl\left(fl(y)/fl(x)\right)$$

$\square$