NUMBER APPROXIMATION - EXERCISES

1. Numbers

Exercise 1. Define one-to-one correspondences between the following sets of numbers:

a) $\mathbb{E} = \{n | n \in \mathbb{N}, n \bmod 2 = 0\}$, $\mathbb{O} = \{n | n \in \mathbb{N}, n \bmod 2 = 1\}$
Solution. $f: \mathbb{E} \to \mathbb{O}$, $f(n) = n + 1$ is one-to-one.

b) $\mathbb{N}$, $\mathbb{Z}$
Solution. $f: \mathbb{N} \to \mathbb{Z}$, with f defined by $\{0 \to 0, 1 \to -1, 2 \to 1, 3 \to -2, 4 \to 2, \ldots\}$ is one possibility. Introducing $[x]$ as the integer part of $x \in \mathbb{Q}$, i.e. $[x] = n$ with $n \leqslant x < n + 1$, f can be expressed as

$$f(n) = (-1)^{n \bmod 2}([n/2] + n \bmod 2)$$

In Julia $x \div y$ is integer division, so for $n \in \mathbb{N}$ $[n/2] = n \div 2$, and % is the modulo operator

```
∴ [4÷5 4÷4 4÷3 4÷2; 4%5 4%4 4%3 4%2]
```

$$\begin{bmatrix} 0 & 1 & 1 & 2 \\ 4 & 0 & 1 & 0 \end{bmatrix} \tag{1}$$

```
∴ function f(n)
    q = n÷2; r = n%2; s = 1-2*r;
    s*(q+r)
  end
```

```
f
```

```
∴ N=10; [collect(0:N)'; f.(0:N)']
```

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 0 & -1 & 1 & -2 & 2 & -3 & 3 & -4 & 4 & -5 & 5 \end{bmatrix} \tag{2}$$

```
∴
```

c) $\mathbb{Z}$, $\mathbb{Q}$
Solution. Construct a table and introduce diagonal traversal to obtain the positive rationals p/q.

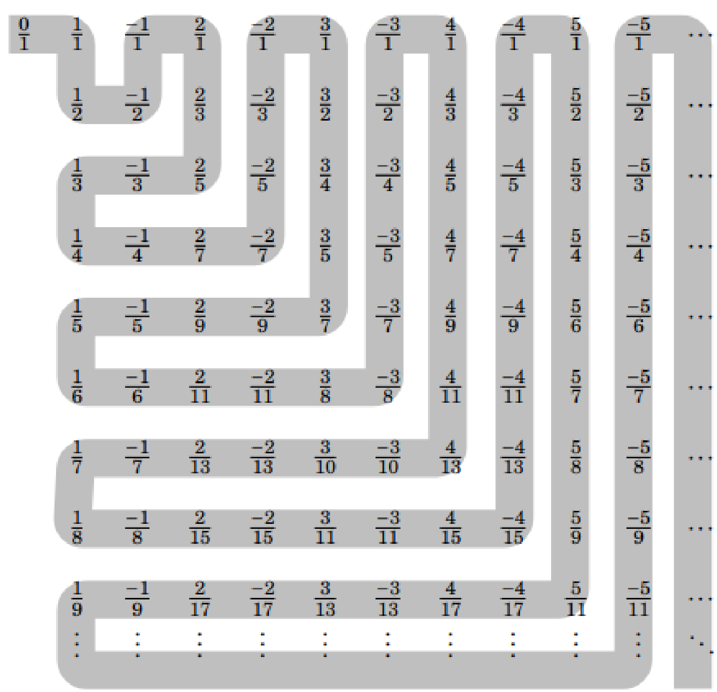

Figure 1. One-to-one mapping showing that $|\mathbb{Q}| = |\mathbb{N}|$

From above deduce $|\mathbb{N}|=|\mathbb{E}|=|\mathbb{O}|=|\mathbb{Z}|=|\mathbb{Q}|=\aleph_0$.

Exercise 2. Provide an example to show $|\mathbb{R}|>|\mathbb{N}|$.

Exercise 3. Let $\mathbb{N}_q=\{n|n\in\mathbb{N}, n<2^q\}$. Answer the following questions analytically. Also provide a Julia implementation.

a) Define a one-to-one correspondence $f:\mathbb{N}_{2q}\to\mathbb{N}_q\times\mathbb{N}_q$

b) Let $f(m_1)=(n_1,p_1)$, $f(m_2)=(n_2,p_2)$. Assume $m_1+m_2\in\mathbb{N}_{2q}$. Express $f(m_1+m_2)$ in terms of n_1,n_2,p_1,p_2 .

c) Assume $m_1\cdot m_2\in\mathbb{N}_{2q}$. Express $f(m_1\cdot m_2)$ in terms of n_1,n_2,p_1,p_2 .

Exercise 4. Construct a one-to-one representation of the positions of atoms within an hexagonal lattice, $f:\mathbb{Z}^2\to\mathbb{R}^2$. Implement f and f^{-1} as Julia functions. Use f to construct a graphical representation of a two-dimensional hexagonal lattice.

Exercise 5. Construct a one-to-one representation of the positions of atoms within an hexagonal lattice, $f:\mathbb{Z}^3\to\mathbb{R}^3$. Implement f and f^{-1} as Julia functions. Use f to construct a graphical representation of a three-dimensional hexagonal lattice.

2. Approximation

Exercise 6. Write Julia code to compute machine epsilon ϵ for `Float32` and `Float64`.

Solution.

```
∴ function MachEps(type)
    one=type(1.0); half=type(0.5); eps=one;
    while (one+half*eps != one)
      eps=half*eps;
    end
    return eps;
  end;
```

```
∴ [MachEps(Float32) eps(Float32) MachEps(Float64) eps(Float64)]
```

$$[1.1920928955078125e-7, 1.1920928955078125e-7, 2.220446049250313e-16, 2.220446049250313e-16] \tag{3}$$

```
∴
```

Exercise 7. Carry out a numerical experiment to verify the Axiom of floating point arithmetic within `Float32`, by computing $\pi+r$ in `Float32` and comparing to the result in `Float64`. Construct a scatter plot of (r,ε) with ε the error in computing $\pi+r$ in `Float32`.

Solution. The floating point axiom states $\mathrm{fl}(x)\circledast\mathrm{fl}(y)=(x*y)(1+\varepsilon)$, with $|\varepsilon|\leqslant\epsilon$, leading to

$$\varepsilon=\frac{\mathrm{fl}(x)\circledast\mathrm{fl}(y)}{x*y}-1,$$

or in this case

$$\varepsilon=\frac{\mathrm{fl}(\pi)\oplus\mathrm{fl}(r)}{\pi+r}-1.$$

The operations in $\mathbb{R}$ are computed in `Float64` in the following, with r randomly chosen.

```
∴ function ErrPlot(n)
    pi32=Float32(pi); pi64=Float64(pi); half=Float64(0.5); one=Float64(1.0);
    rscale = 1.0e3*half; e32 = eps(Float32);
    r64=Float64.( rscale*(rand(n) .- half) );
    r32=Float32.(r64);
    result32 = pi32 .+ r32; result64 = pi64 .+ r64;
    ε = result32 ./ result64 .- one;
    rmin=minimum(r64); rmax=maximum(r64);
   clf(); plot(r32,ε,".",[rmin rmax],[e32 e32],"dg",[rmin rmax],[-e32 -e32],"dg");
    xlabel("r"); ylabel("ε"); title("Float32␣addition␣error");
  end;
```

```
∴ ErrPlot(1000); savefig(homedir()*"/courses/MATH661/images/E01Fig02.eps")
```

```
∴
```

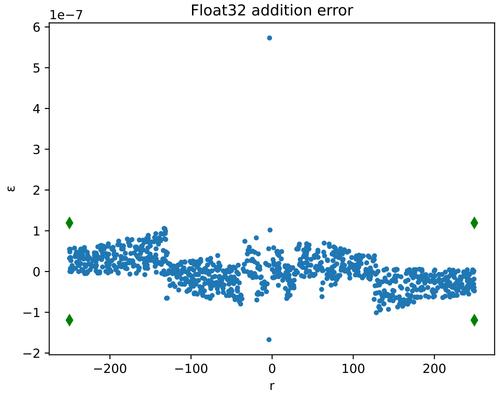

Figure 2. Numerical experiment on verification of floating point axiom. While for most numbers within this random sample the axiom is verified, there are a few cases when $r \cong 0$ the error is larger than ϵ.

Exercise 8. Consider the approximations of e

$$S_n = 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}, T_n = \frac{1}{n!} + \frac{1}{(n-1)!} + \cdots + 1.$$

a) Write Julia functions to compute S_n, T_n.

Solution.

```
∴ function S(n)
    fact=1.0; sum=1.0;
    for k=2:n
      fact = k*fact;
      sum = sum + 1/fact;
    end
    return sum;
  end;
```

```
∴ function T(n)
    fact=1.0;
    for k=n:-1:2
      fact=k*fact;
    end
    sum=0.0;
    for k=n:-1:1
      sum = sum + 1/fact;
      fact = fact/k;
    end
    return sum;
  end;
```

```
∴
```

b) Determine if $S_n = T_n$ for all $n \in \mathbb{N}$.

Solution. In $\mathbb{R}$, indeed $S_n = T_n$ by commutativity (proof by induction). In $\mathbb{F}$ there must exist some N such that for $n > N$, $S_n \neq T_n$ as a consequence of the existence of machine epsilon. Verify by computation (note organization of computations to use Julia broadcasting and presentation of results in a single table)

```
∴ r=1:8; s=S.(r); t=T.(r); chk = s.==t; [r s t chk]
```

$$\begin{bmatrix} 1.0 & 1.0 & 1.0 & 1.0 \\ 2.0 & 1.5 & 1.5 & 1.0 \\ 3.0 & 1.6666666666666667 & 1.6666666666666665 & 0.0 \\ 4.0 & 1.7083333333333335 & 1.7083333333333333 & 0.0 \\ 5.0 & 1.7166666666666668 & 1.7166666666666668 & 1.0 \\ 6.0 & 1.7180555555555557 & 1.7180555555555554 & 0.0 \\ 7.0 & 1.7182539682539684 & 1.7182539682539684 & 1.0 \\ 8.0 & 1.71827876984127 & 1.7182787698412698 & 0.0 \end{bmatrix} \tag{4}$$

$\therefore$

c) Determine if $|S_n - T_n| < \epsilon$ (ϵ is machine epsilon). Is the floating point axiom verified?

d) Determine if $|S_n - T_n| < (n-1)\epsilon$. Is the floating point axiom verified?

Exercise 9. Consider the approximations of $\pi/2$ given by Wallis's product

$$S_n = \left(\frac{2}{1}\right)\cdot\left(\frac{2}{3}\cdot\frac{4}{3}\right)\cdot\left(\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\frac{8}{7}\right)\ldots p_n,$$

$$T_n = p_n\cdot\ldots\cdot\left(\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\frac{8}{7}\right)\cdot\left(\frac{2}{3}\cdot\frac{4}{3}\right)\cdot\left(\frac{2}{1}\right)$$

a) Find the general term p_n.

b) Determine if $S_n = T_n$ for all $n \in \mathbb{N}$.

c) Determine if $|S_n - T_n| < \epsilon$ (ϵ is machine epsilon). Is the floating point axiom verified?

d) Determine if $|S_n - T_n| < (n-1)\epsilon$. Is the floating point axiom verified?

Exercise 10. Consider the approximations of $e/2$ given by Pippenger's product

$$S_n = \left(\frac{2}{1}\right)^{1/2}\cdot\left(\frac{2}{3}\cdot\frac{4}{3}\right)^{1/4}\cdot\left(\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\frac{8}{7}\right)^{1/8}\ldots p_n,$$

$$T_n = p_n\cdot\ldots\cdot\left(\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\frac{8}{7}\right)^{1/8}\cdot\left(\frac{2}{3}\cdot\frac{4}{3}\right)^{1/4}\cdot\left(\frac{2}{1}\right)^{1/2}$$

a) Find the general term p_n.

b) Determine if $S_n = T_n$ for all $n \in \mathbb{N}$.

c) Determine if $|S_n - T_n| < \epsilon$ (ϵ is machine epsilon). Is the floating point axiom verified?

d) Determine if $|S_n - T_n| < (n-1)\epsilon$. Is the floating point axiom verified?

3. Successive approximations

Exercise 11. Assume errors in successive numerical approximation of $a \in \mathbb{R}$, finite, are given by $e_n = a_n - a_{n-1}$, with $\{a_n\}_{n\in\mathbb{N}}$, $a_n = n^{1/3}$.

a) Construct a scatter plot of (n, e_n). Does the plot indicate convergence of the numerical approximation?

b) Compute $\lim_{n\to\infty} e_n$.

c) Suppose $|e_n| < \varepsilon$. What is an upper bound for $|a_n - a|$?

Exercise 12. Assume errors in successive numerical approximation of $a \in \mathbb{R}$, finite, are given by $e_n = a_n - a_{n-1}$, with $\{a_n\}_{n\in\mathbb{N}}$, $a_n = n^{-1/2}$.

a) Construct a scatter plot of (n, e_n). Does the plot indicate convergence of the numerical approximation?

b) Compute $\lim_{n\to\infty} e_n$.

c) Suppose $|e_n| < \varepsilon$. What is an upper bound for $|a_n - a|$?

Exercise 13. Consider a sequence of successive approximations of the derivative $f'(x_0)$

$$d_n = \frac{f(x_0 + 1/n) - f(x_0)}{1/n}, n \in \mathbb{N}.$$

a) Is $\{d_n\}_{n\in\mathbb{N}}$ a convergent sequence?

b) Is $\{d_n\}_{n\in\mathbb{N}}$ a Cauchy sequence?

c) Construct a scatter plot of (n, d_n) for $f(x) = \sin x$, $x_0 = \pi/4$. Does the plot indicate convergence of $\{d_n\}_{n\in\mathbb{N}}$?

d) Construct a scatter plot of (n, e_n), $e_n = d_n - d_{n-1}$, for $f(x) = \sin x$, $x_0 = \pi/4$. Does the plot indicate convergence of $\{d_n\}_{n\in\mathbb{N}}$?

Exercise 14. Consider errors in successive approximations $\{a_n\}_{n\in\mathbb{N}}$ given by $e_n = a_n - a_{n-1} = e_{n-1} + e_{n-2}$, i.e., errors at each step accumulate errors in previous two steps, with $e_1 = e_0 = 1$. Is this a convergent approximation? Present both an analytical solution, and a numerical experiment.

Exercise 15. Consider errors in successive approximations $\{a_n\}_{n\in\mathbb{N}}$ given by $e_n = a_n - a_{n-1} = \frac{5}{6}e_{n-1} + \frac{1}{6}e_{n-2}$, i.e., errors at each step are a weighted average of those in previous two steps, with $e_1 = e_0 = 1$. Is this a convergent approximation? Present both an analytical solution, and a numerical experiment.

Exercise 16. Consider errors in successive approximations $\{a_n\}_{n\in\mathbb{N}}$ given by $e_n = a_n - a_{n-1} = \frac{1}{2}e_{n-1} + \frac{1}{4}e_{n-2}$, i.e., errors at each step are less than a weighted average of those in previous two steps, with $e_1 = e_0 = 1$. Is this a convergent approximation? Present both an analytical solution, and a numerical experiment.