Infinity and Finite Arithmetic [a]

Walter Gander

gander@inf.ethz.ch

ETH Fachdidaktik Informatik

14. October 2020

[a] Chapter in: Adventures Between Lower Bounds and Higher Altitudes: Essays Dedicated to Juraj Hromkovič on the Occasion of His 60th Birthday, Springer, 2018

Infinity Is Man Made

- Infinity: something that never ends: sky, # stars, time
- However, everything is finite:
  1 liter water: 3.343×10^{25} molecules H_2O (Wolfram Alpha)
- Every human composed of finite number of molecules
- In principle one could also count the finite number of atoms which form our planet.

  > *"Two things are infinite: the universe and human stupidity; and I'm not sure about the universe." (Albert Einstein)*

- Scientists believe Einstein: universe has 10^{80} atoms.

Infinity does not exists in nature – it is man made.

Infinity in Mathematics

- Encyclopedia Britannica: 3 types of infinity: mathematical (endless sequence of numbers), physical (spatial and temporal), and metaphysical (God or the Absolute)
- Several hierarchies of ∞: countable like $\mathbb{N}$ or $\mathbb{Q}$, uncountable $\mathbb{R}$
- Infinite series $\sum_{k=1}^{\infty} a_k$ occur frequently in mathematics and one is interested if the partial sums

$$s_n = \sum_{k=1}^{n} a_k, \quad \lim_{n\to\infty} s_n = ?$$

converge to a limit or not.

Harmonic Sum $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$

- divergent, since

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \cdots, \quad |t| < 1$$

$$\int_0^z \frac{dt}{1-t} = -\log(1-z) = z + \frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \cdots, \qquad z < 1$$

$$z \to 1 \implies -\log(1-z) \to \infty$$

- dividing by z and integrate (get dilogarithm)

$$\frac{-\log(1-z)}{z} = 1 + \frac{z}{2} + \frac{z^2}{3} + \frac{z^3}{4} + \cdots$$

$$Li2(z) = \int_0^z \frac{-\log(1-u)}{u}\, du = z + \frac{z^2}{4} + \frac{z^3}{9} + \frac{z^4}{16} + \cdots.$$

For $z = 1$ we get the well known series of the reciprocal squares

ζ-Function

$$\zeta(z) = 1 + \frac{1}{2^z} + \frac{1}{3^z} + \frac{1}{4^z} + \cdots$$

- $$\int_0^1 \frac{-\log(1-u)}{u}\, du = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6}.$$
  ζ-function for $z = 2$
- divide $Li2(z)$ by z and integrate
  $$\int_0^1 \frac{Li2(z)}{z}\, dz = 1 + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{4^3} + \cdots = \zeta(3).$$
- can compute the numerical value for $\zeta(3)$
  nice result as for $\zeta(2)$ still not known

Infinity and Numerical Computations

- $s_n = \sum_{k=1}^{n} \frac{1}{k}$

```
s=1; term=1; k=1;
while s+term ~= s
  k=k+1; term=1/k; s=s+term;
end
```

- on the computer the harmonic series converges!
  For $n = 10^{15}$ we get $s_n \approx 35.116$ and `s+term = s`

- $\zeta(2)$: series of the inverse squares

```
s=1; term=1; n=1;
while s+term ~= s
  n=n+1; term=1/n^2; s=s+term;
end
```

terminates with $n = 94'906'266$ and $s = 1.644934057834575$

Series for $\zeta(2)$

- forward summation:

$$\sum_{k=1}^{94'906'266} \frac{1}{k^2} = 1.64493405{\color{red}7834575}$$

- reverse summation

$$\sum_{k=94'906'266}^{1} \frac{1}{k^2} = 1.64493405631151 4$$

- Maple with `Digits:=30`: $s = 1.64493405631151440597651536455$

  reverse summation correct

- However

$$\frac{\pi^2}{6} - s = 1.0537 \times 10^{-8}$$

Conclusion: straightforward numerical summation fails

Aitkens Δ^2-Acceleration

- Let $\{x_k\} \to s$ with linear convergence

$$\iff \lim_{k\to\infty} \frac{x_{k+1}-s}{x_k - s} = \rho, \quad \rho \neq 0, \quad |\rho| < 1$$
$$\iff x_n \sim s + C\rho^n, \quad C = x_0 - s.$$

- replace "$\sim$" by "$=$", solve for ρ, C and s

  $x_{n-2} = s + C\rho^{n-2}$, $x_{n-1} = s + C\rho^{n-1}$, $x_n = s + C\rho^n$

- We obtain $(s = x'_n)$

$$x'_n = x_{n-2} - \frac{(x_{n-1} - x_{n-2})^2}{x_n - 2\,x_{n-1} + x_{n-2}} = x_{n-2} - \frac{(\Delta x_{n-2})^2}{\Delta^2 x_{n-2}}$$

- If x'_n also linearly convergent, can iterate this transformation

Aitken Triangular Scheme

x_k is original sequence

$x'_k,\ x''_k,\ \ldots$
are extrapolated sequences

x_k	x'_k	x''_k	$\cdots$
x_1			
x_2			
x_3	x'_1		
x_4	x'_2		
x_5	x'_3	x''_1	
$\vdots$	$\vdots$	$\vdots$	$\ddots$

Computing the Euler-Mascheroni Constant γ

$$\sum_{k=1}^{n} \frac{1}{k} = \log(n) + \gamma + \frac{1}{2n} + O\left(\frac{1}{n^2}\right)$$

Consider $x_k = \sum_{j=1}^{2^k} \frac{1}{j} - \log(2^k), \quad k = 0, 1, \ldots, K$

Computing γ

Generate x_k

```
function x=HarmonicPartial(K);
y=[1:2^K]; y=1./y;
s=1; x=s;
for k=1:K
   s=s+sum(y(2^(k-1)+1:2^k));
   x=[x s-log(2^k)];
end
x=x';
```

Main Program

```
K=8
x=HarmonicPartial(K);
A=AitkenAcc(x)
```

Iterated Aitken-Scheme

```
function T=AitkenAcc(x)
n=length(x); m=floor((n+1)/2);
T=zeros(n,m);
T(:,1)=x;
for j=2:m
  for k=2*j-1:n
    Delta2=T(k,j-1)-2*T(k-1,j-1)+T(k-2,j-1);
    if Delta2==0, break, end
    T(k,j)=T(k-2,j-1)-(T(k-1,j-1)-T(k-2,j-1))^2/Delta2;
  end
end
```

Results

1.000000000000000				
0.806852819440055				
0.697038972213442	0.552329999700925			
0.638415601177307	0.571280073489448			
0.608140270989212	0.575806621446670	0.577227192023427		
0.592759292636793	0.576875788763135	0.577206420206234		
0.585007820346097	0.577132432059184	0.577213495239782	0.577211697690028	
0.581116828669555	0.577195084788389	0.577215319609806	0.577215953496545	
0.579167518337717	0.577210549043978	0.577215616874797	0.577215674740153	0.577215691876342

Using 256 terms of series we obtain 7 correct decimal digits of γ !

Extrapolation, Richardson

- $T(h)$ with $T(0) = a_0$ difficult to compute
- Idea: compute values $T(h_i)$ for $h_i > 0$, $i = 0, 1, \dots, n$

  Construct the interpolation polynomial $P_n(h) \implies P_n(0) \approx a_0$
- For $a_0 = \lim_{m\to\infty} s_m$ use transformation $h = 1/m$

  then $T(h) = s_m$ thus $\lim_{h=0} T(h) = a_0$
- Thm: If $\exists$ asymptotic expansion

  $$T(h) = a_0 + a_1 h + \cdots + a_k h^k + R_k(h) \quad \text{with} \quad |R_k(h)| < C_k h^{k+1}$$

  and if $h_{i+1} < c h_i$ with some $0 < c < 1$ then $P_n(0) \to a_0$ faster than $T(h_n)$

Aitken-Neville Interpolation

- Let $T_{ij}(x)$ be the polynomial of degree $\leq j$ that interpolates the data

x	x_{i-j}	x_{i-j+1}	$\cdots$	x_i
y	y_{i-j}	y_{i-j+1}	$\cdots$	y_i

- Aitken-Neville scheme

x	y				
x_0	$y_0 = T_{00}$				
x_1	$y_1 = T_{10}$	T_{11}			
$\vdots$	$\vdots$		$\ddots$		
x_i	$y_i = T_{i0}$	T_{i1}	$\cdots$	T_{ii}	
$\cdots$	$\cdots$	$\cdots$		$\cdots$	$\ddots$

$$
\begin{aligned}
T_{i0} &= y_i \\
T_{ij} &= \frac{(x_i - x)T_{i-1,j-1} + (x - x_{i-j})T_{i,j-1}}{x_i - x_{i-j}} \\
j &= 1, 2, \ldots, i \\
i &= 0, 1, 2, \ldots
\end{aligned}
$$

Simplifications

- For $x = 0$ the Aitken-Neville-Scheme recursion is:

$$T_{ij} = \frac{h_i T_{i-1,j-1} - h_{i-j} T_{i,j-1}}{h_i - h_{i-j}}.$$

- For $h_i = h_0 2^{-i}$

$$T_{ij} = \frac{2^j T_{i,j-1} - T_{i-1,j-1}}{2^j - 1}.$$

- If $T(h) = a_0 + a_2 h^2 + a_4 h^4 + \cdots$ use $P_n(h^2)$ and $h_i = h_0 2^{-i}$

$$T_{ij} = \frac{4^j T_{i,j-1} - T_{i-1,j-1}}{4^j - 1} \quad \text{Romberg}$$

ANS Program

```
function A=ANS(x,factor);
% ANS  Aitken-Neville-Scheme for x,  factor is 2 or 4
K=length(x);
A(1,1)=x(1);
for i=2:K
  A(i,1)=x(i); vhj=1;
  for j=2:i
    vhj=vhj*factor;
    A(i,j)=(vhj*A(i,j-1)-A(i-1,j-1))/(vhj-1);
  end;
end
```

Example $\zeta(2)$ compute partial sums

$$s_m = \sum_{k=1}^{2^m} \frac{1}{k^2}$$

```
K=8;
y=[1:2^K]; y=1./y.^2;
for j=0:K-1
  s=sum(y(1:2^j)); x(j+1)=s;
end
x=x';
A=ANS(x,2)
```

1.0000							
1.2500	1.5000						
1.4236	1.5972	1.629629					
1.5274	1.6312	1.642569	1.644418529				
1.5843	1.6412	1.644617	1.644909465	1.644942194			
1.6141	1.6439	1.644893	1.644933169	1.644934750	1.6449345100		
1.6294	1.6446	1.644928	1.644934037	1.644934095	1.6449340742	1.644934067322	
1.6371	1.6448	1.644933	1.644934065	1.644934067	1.6449340669	1.644934066809	1.64493406680539

- $2^7 = 128$ terms of series
- extrapolated value $A_{8,8} = 1.64493406680539$
- error $\pi^2/6 - A_{8,8} = 4.28 \cdot 10^{-11}$ – quite impressive!

Asymptotic Expansion of ζ-function

- $s_{m-1} = \sum_{k=1}^{m-1} \frac{1}{k^z} = \zeta(z) - \sum_{k=m}^{\infty} \frac{1}{k^z}$
- Euler-MacLaurin Summation Formula, asymptotic expansion:

$$\sum_{k=m}^{\infty} \frac{1}{k^z} \sim \frac{1}{z-1} \frac{1}{m^{z-1}} + \frac{1}{2} \frac{1}{m^z} + \frac{1}{z-1} \sum_{j=1} \binom{1-z}{2j} \frac{B_{2j}}{m^{z-1+2j}}$$

Bernoulli numbers:

$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}, B_6 = \frac{1}{42}, B_8 = -\frac{1}{30}, \ldots$ and $B_3 = B_5 = B_7 = \ldots = 0$.

- Thus

$$\sum_{k=1}^{m-1} \frac{1}{k^z} + \frac{1}{2} \frac{1}{m^z} \sim \zeta(z) - \frac{1}{z-1} \sum_{j=0} \binom{1-z}{2j} \frac{B_{2j}}{m^{z-1+2j}}$$

Examples

- For $z = 3$: asymptotic expansion has only even exponents

$$\sum_{k=1}^{m-1} \frac{1}{k^3} + \frac{1}{2}\frac{1}{m^3} \sim \zeta(3) - \frac{1}{2m^2} - \frac{1}{4m^4} + \frac{1}{12m^6} - \frac{1}{12m^8} \pm \cdots$$

- For $z = 2$: expansion has only odd exponents:

$$\sum_{k=1}^{m-1} \frac{1}{k^2} + \frac{1}{2}\frac{1}{m^2} \sim \zeta(2) - \frac{B_0}{m} - \frac{B_2}{m^3} - \frac{B_4}{m^5} - \cdots$$

- Knowing asymptotic expansions, no need to extrapolate. E.g. for $m = 1000$ we get machine precision with

$$\sum_{k=1}^{m-1} \frac{1}{k^3} + \frac{1}{2}\frac{1}{m^3} + \frac{1}{2m^2} + \frac{1}{4m^4} - \frac{1}{12m^6} = 1.202056903159593$$

Extrapolate ζ-Function

Only even exponents in asymptotic expansion for $\zeta(3)$
Extrapolate from 128 terms
$\zeta(3) = 1.202056903159594$
to machine precision!

```
K=8; m=1;
for j=1:K
  s=0;
  for k=1:m-1
    s=s+1/k^3;
  end
  x(j)=s+1/2/m^3;
  m=2*m;
end
A=ANS(x',4)
```

How about, if

$$T(h) = a_0 + a_1 h + a_2 h^3 + a_3 h^5 + a_4 h^7 + \cdots$$

has only odd powers?

Richardson Extrapolation

- Idea: eliminate lower order terms in

$$T(h) = a_0 + a_1 h + a_2 h^3 + a_3 h^5 + a_4 h^7 + \cdots$$

- Extrapolation scheme

$$\begin{array}{lllllllll}
T_{11} & = & T(h) & & & & & & \\
T_{12} & = & T(h/2) & T_{22} & = & 2T_{12} - T_{11} & & & \\
T_{31} & = & T(h/4) & T_{32} & = & 2T_{32} - T_{12} & T_{33} & = & \dfrac{2^3 T_{32} - T_{22}}{2^3 - 1} \\
 & \vdots & & & \vdots & & & \vdots & \ddots
\end{array}$$

- Recurrence relation

$$T_{ij} = \frac{2^{2j-3} T_{ij-1} - T_{i-1j-1}}{2^{2j-3} - 1}, \quad i = 2, 3, \ldots, \; j = 2, 3, \ldots, i$$

Richardson (cont.)

$$T_{11} = T(h) = a_0 + a_1 h + a_2 h^3 + a_3 h^5 + a_4 h^7$$

$$T_{21} = T(h/2) = a_0 + 1/2 a_1 h + 1/8 a_2 h^3 + 1/32 a_3 h^5 + 1/128 a_4 h^7$$

$$T_{22} = 2T_{21} - T_{11} = a_0 - 3/4 a_2 h^3 - 15/16 a_3 h^5 - 63/64 a_4 h^7$$

$$T_{32} = 2T(h/4) - T_{21} = a_0 - 3/32 a_2 h^3 - 15/512 a_3 h^5 - 63/8192 a_4 h^7$$

$$T_{33} = (2^3 T_{32} - T_{22})/(2^3 - 1) = a_0 + 45/448 a_3 h^5 + 135/1024 a_4 h^7$$

The terms with h and h^3 are eliminated in T_{33}

Again $\zeta(2)$

```
function A=ANSodd(x);
% ANSodd extrapolation for x with only
% odd exponents
K=length(x);
A(1,1)=x(1);
for i=2:K
  A(i,1)=x(i); vhj=2;
  for j=2:i
    A(i,j)=(vhj*A(i,j-1)-A(i-1,j-1))/(vhj-1);
    vhj=vhj*4;
  end;
end
```

```
K=8; m=1;
for j=1:8
  s=0;
  for k=1:m-1
    s=s+1/k^2;
  end
  x(j)=s+1/2/m^2;
  m=2*m;
end
A=ANSodd(x)
```

0.500				
1.125	1.750000			
1.392	1.659722	1.64682539		
1.519	1.646857	1.64502024	1.64496201552	
1.582	1.645177	1.64493716	1.64493448049	1.64493426368
1.613	1.644964	1.64493416	1.64493407101	1.64493406778
1.629	1.644937	1.64493407	1.64493406688	1.64493406685
1.637	1.644934	1.64493406	1.64493406684	1.64493406684

1.644934067405063		
1.644934066849075	1.644934066848803	
1.644934066848226	1.644934066848226	1.644934066848226

- this time we obtain machine precision by computing 128 terms
- Error= $1.644934066848226 - \pi^2/6 = -4.4409e{-16}$

The ε-Algorithm

- Model for Aitken's Δ^2-Acceleration $x_n \sim s + C\rho^n$
- Generalization by Shanks, asymptotic error model

$$x_n \sim s + \sum_{i=1}^{k} a_i \rho_i^n, \quad \text{for } k > 1.$$

  Replacing again "$\sim$" with "$=$" and using $2k+1$ consecutive iterations we get a system of nonlinear equations

$$x_{n+j} = s_{n,k} + \sum_{i=1}^{k} a_i \rho_i^{n+j}, \quad j = 0, 1, \ldots, 2k.$$

- Assuming we can solve this system, we obtain a new sequence $x'_n = s_{n,k}$. This is called a Shanks Transformation.

- Solving this nonlinear system is not easy (rather impossible!)
- different characterization for the Shanks Transformation:
- Let $P_k(x) = c_0 + c_1 x + \cdots + c_k x^k$ be the polynomial with zeros $\rho_1, \ldots, \rho_k$, normalized such that $\sum c_i = 1$
- consider the equations

$$\begin{aligned}
c_0(x_n - s_{n,k}) &= c_0 \sum_{i=1}^{k} a_i \rho_i^n \\
c_1(x_{n+1} - s_{n,k}) &= c_1 \sum_{i=1}^{k} a_i \rho_i^{n+1} \\
\vdots \quad &= \quad \vdots \\
c_k(x_{n+k} - s_{n,k}) &= c_k \sum_{i=1}^{k} a_i \rho_i^{n+k}.
\end{aligned}$$

- Adding all these equations we obtain the sum

$$\sum_{j=0}^{k} c_j(x_{n+j} - s_{n,k}) = \sum_{i=1}^{k} a_i \rho_i^n \underbrace{\sum_{j=0}^{k} c_j \rho_i^j}_{P_k(\rho_i)=0},$$

and since $\sum c_i = 1$, the extrapolated value becomes

$$s_{n,k} = \sum_{j=0}^{k} c_j x_{n+j}. \tag{1}$$

- Thus $s_{n,k}$ is a linear combination of successive iterates, a weighted average. If we knew the coefficients c_j of the polynomial, we could directly compute $s_{n,k}$.

Peter Wynn's Algorithm

- Wynn established in 1956 a remarkable recursion for $s_{n,k}$
- $\varepsilon_{-1}^{(n)} = 0$ and $\varepsilon_0^{(n)} = x_n$ for $n = 0, 1, 2, \ldots$
- $\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}}$
- ε-Table

$$
\begin{array}{cccccc}
\varepsilon_{-1}^{(0)} & & & & & \\
 & \varepsilon_{0}^{(0)} & & & & \\
\varepsilon_{-1}^{(1)} & & \varepsilon_{1}^{(0)} & & & \\
 & \varepsilon_{0}^{(1)} & & \varepsilon_{2}^{(0)} & & \\
\varepsilon_{-1}^{(2)} & & \varepsilon_{1}^{(1)} & & \varepsilon_{3}^{(0)} & \\
 & \varepsilon_{0}^{(2)} & & \varepsilon_{2}^{(1)} & & \ldots \\
\varepsilon_{-1}^{(3)} & & \varepsilon_{1}^{(2)} & & \ldots & \\
 & \varepsilon_{0}^{(3)} & & \ldots & & \\
\varepsilon_{-1}^{(4)} & & \ldots & & &
\end{array}
$$

- Wynn showed $\varepsilon_{2k}^{(n)} = s_{n,k}$ and $\varepsilon_{2k+1}^{(n)} = \frac{1}{S_k(\Delta x_n)}$
  where $S_k(\Delta x_n)$ is Shanks transformation of $\Delta x_n = x_{n+1} - x_n$
- Thus every second column in the ε-table is of interest
- MATLAB implementation:
  Write ε-table in lower triangular part of matrix E. Shift indices

$$
\begin{array}{llll}
0 = \varepsilon_{-1}^{(0)} = E_{11}, & & & \\
0 = \varepsilon_{-1}^{(1)} = E_{21} & x_1 = \varepsilon_0^{(0)} = E_{22}, & & \\
0 = \varepsilon_{-1}^{(2)} = E_{31} & x_2 = \varepsilon_0^{(1)} = E_{32} & \varepsilon_1^{(0)} = E_{33}, & \\
0 = \varepsilon_{-1}^{(3)} = E_{41} & x_3 = \varepsilon_0^{(2)} = E_{42} & \varepsilon_1^{(1)} = E_{43} & \varepsilon_2^{(0)} = E_{44}.
\end{array}
$$

Algorithm

```
function [s,Er]=EpsilonAlgorithm(x);
% EPSILONALGORITHM computes the eps-scheme E for sequence x.
%   Output is the reduced scheme Er (only even columns) and
%   diagonal element s.
n=length(x);
E=zeros(n+1,n+1);
for i=1:n
  E(i+1,2)=x(i);
end
for i=3:n+1
  for j=3:i
    D=E(i,j-1)-E(i-1,j-1);
    if D==0, s=E(i,j-1); return, end
    E(i,j)=E(i-1,j-2)+1/D;
  end
end
Er=E(2:end,2:2:end); s=E(end,end);
```

Example

- Use ε-algorithm to evaluate the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \cdots = \ln 2.$$

```
format short e
k=4;
v=1;
for j=1:2*k+1
  y(j)=v/j; v=-v;
end
x=cumsum(y);
[s,Er]=EpsilonAlgorithm(x);
Er
log(2)-s
```

```
Er =
   1.0000e+00            0            0            0            0
   5.0000e-01            0            0            0            0
   8.3333e-01   7.0000e-01            0            0            0
   5.8333e-01   6.9048e-01            0            0            0
   7.8333e-01   6.9444e-01   6.9333e-01            0            0
   6.1667e-01   6.9242e-01   6.9309e-01            0            0
   7.5952e-01   6.9359e-01   6.9317e-01   6.9315e-01            0
   6.3452e-01   6.9286e-01   6.9314e-01   6.9315e-01            0
   7.4563e-01   6.9335e-01   6.9315e-01   6.9315e-01   6.9315e-01
>> log(2)-s
ans =  -1.5179e-07
```

It is quite remarkable that we can obtain a result with about 7 decimal digits of accuracy by extrapolation using only partial sums of the first 9 terms, especially since the last partial sum still has no correct digit!

References

1. Walter Gander, Martin J. Gander, Felix Kwok, *Scientific Computing, an Introduction Using* MAPLE *and* MATLAB. Springer Verlag, 2014.
2. Juraj Hromkovič, *Algorithmic Adventures*. Springer Berlin Heidelberg, 2009.
3. Konrad Knopp, *Theorie und Anwendungen der unendlichen Reihen*. Springer Verlag, 1947.
4. Daniel Shanks, *Non-linear transformation of divergent and slowly convergent sequences*. Journal of Mathematics and Physics, 1955, Vol. 34, pp. 1–42.
5. P. Wynn, *On a device for computing the* $e_m(S_n)$*-transformation*. MTAC, 1956, Vol. 10, pp. 91-96.