High precision computation of number-theoretical constants

Henri Cohen

MSRI and Institut de Mathématiques de Bordeaux

March 10, 2011,   MSRI Berkeley

Introduction

In number theory, need to compute constants to hundreds of decimal places instead of 15 in numerical analysis : for instance, search for identities which one often finds numerically before giving a proof. For example, very active research on Mahler measures of rational functions of several variables lead to surprising links with special values of L-functions of elliptic curves and volumes of hyperbolic manifolds. Many of these identities have been shown to be true to thousands of decimal places but are still conjectures (I am not talking of BSD here). Many other uses for high precision computation of constants.

Goal of the talk : give some computational techniques, not all classical, for computing certain constants, and we want them to reasonably high accuracy, for instance several hundred decimal places.

Introduction

In number theory, need to compute constants to hundreds of decimal places instead of 15 in numerical analysis : for instance, search for identities which one often finds numerically before giving a proof. For example, very active research on Mahler measures of rational functions of several variables lead to surprising links with special values of L-functions of elliptic curves and volumes of hyperbolic manifolds. Many of these identities have been shown to be true to thousands of decimal places but are still conjectures (I am not talking of BSD here). Many other uses for high precision computation of constants.

Goal of the talk : give some computational techniques, not all classical, for computing certain constants, and we want them to reasonably high accuracy, for instance several hundred decimal places.

Part I : Series Summation I

(1). Naive Methods

Typical example : numerical computation of $\zeta(3) = \sum_{n \geq 1} 1/n^3$.

Direct computation : the Nth remainder is asymptotic to $1/(2N^2)$, so for $N = 10^8$ (reasonable limit), only 16 decimals, insufficient.

Use of rational functions : we note that :

$$\frac{1}{n-1} - 2\frac{1}{n} + \frac{1}{n+1} = \frac{2}{n^3 - n} = \frac{2}{n^3} + \frac{2}{n^5 - n^3} .$$

The LHS "telescopes" : summing for $n \geq 2$ we get

$$\zeta(3) = \frac{5}{4} - \sum_{n \geq 2} \frac{1}{n^5 - n^3} .$$

Here the Nth remainder is asymptotic to $1/(4N^4)$, so need only $N = 10^4$ terms to get 16 decimals, and for $N = 10^8$ we obtain 32. Of course, this method can be extended at will, obtaining even greater speedups.

Exist still other naive methods, no need to insist.

(1). Naive Methods

Typical example : numerical computation of $\zeta(3) = \sum_{n \geq 1} 1/n^3$.

Direct computation : the Nth remainder is asymptotic to $1/(2N^2)$, so for $N = 10^8$ (reasonable limit), only 16 decimals, insufficient.

Use of rational functions : we note that :

$$\frac{1}{n-1} - 2\frac{1}{n} + \frac{1}{n+1} = \frac{2}{n^3 - n} = \frac{2}{n^3} + \frac{2}{n^5 - n^3} \, .$$

The LHS “telescopes” : summing for $n \geq 2$ we get

$$\zeta(3) = \frac{5}{4} - \sum_{n \geq 2} \frac{1}{n^5 - n^3} \, .$$

Here the Nth remainder is asymptotic to $1/(4N^4)$, so need only $N = 10^4$ terms to get 16 decimals, and for $N = 10^8$ we obtain 32. Of course, this method can be extended at will, obtaining even greater speedups.

Exist still other naive methods, no need to insist.

Part I : Series Summation I

(1). Naive Methods

Typical example : numerical computation of $\zeta(3) = \sum_{n \geq 1} 1/n^3$.

Direct computation : the Nth remainder is asymptotic to $1/(2N^2)$, so for $N = 10^8$ (reasonable limit), only 16 decimals, insufficient.

Use of rational functions : we note that :

$$\frac{1}{n-1} - 2\frac{1}{n} + \frac{1}{n+1} = \frac{2}{n^3 - n} = \frac{2}{n^3} + \frac{2}{n^5 - n^3} \ .$$

The LHS "telescopes" : summing for $n \geq 2$ we get

$$\zeta(3) = \frac{5}{4} - \sum_{n \geq 2} \frac{1}{n^5 - n^3} \ .$$

Here the Nth remainder is asymptotic to $1/(4N^4)$, so need only $N = 10^4$ terms to get 16 decimals, and for $N = 10^8$ we obtain 32. Of course, this method can be extended at will, obtaining even greater speedups.

Exist still other naive methods, no need to insist.

(1). Naive Methods

Typical example : numerical computation of $\zeta(3) = \sum_{n\geq 1} 1/n^3$.

Direct computation : the Nth remainder is asymptotic to $1/(2N^2)$, so for $N = 10^8$ (reasonable limit), only 16 decimals, insufficient.

Use of rational functions : we note that :

$$\frac{1}{n-1} - 2\frac{1}{n} + \frac{1}{n+1} = \frac{2}{n^3-n} = \frac{2}{n^3} + \frac{2}{n^5-n^3}\ .$$

The LHS "telescopes" : summing for $n \geq 2$ we get

$$\zeta(3) = \frac{5}{4} - \sum_{n\geq 2} \frac{1}{n^5-n^3}\ .$$

Here the Nth remainder is asymptotic to $1/(4N^4)$, so need only $N = 10^4$ terms to get 16 decimals, and for $N = 10^8$ we obtain 32. Of course, this method can be extended at will, obtaining even greater speedups.

Exist still other naive methods, no need to insist.

(2). The Euler–MacLaurin Sum Formula

Much more powerful, although ancient. Review on Bernoulli numbers :

$$\frac{x}{e^x - 1} = \sum_{n\geq 0} B_n \frac{x^n}{n!} .$$

We have $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_{2n+1} = 0$ for $n \geq 1$.

One form of Euler–McLaurin, valid for "nice" functions f : for a suitable constant $C = C(f)$ we have :

$$\sum_{n=1}^{N} f(n) = C + \int_1^N f(t)\,dt + \frac{f(N)}{2} + \sum_{1\leq k\leq p} B_{2k} \frac{f^{(2k-1)}(N)}{(2k)!} + R_p(N) ,$$

where the remainder term $R_p(N)$ is "small" (in general smaller than the last term of the sum on k).

Series Summation II

(2). The Euler–MacLaurin Sum Formula

Much more powerful, although ancient. Review on Bernoulli numbers :

$$\frac{x}{e^x - 1} = \sum_{n \geq 0} B_n \frac{x^n}{n!} \ .$$

We have $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_{2n+1} = 0$ for $n \geq 1$.

One form of Euler–McLaurin, valid for “nice” functions f : for a suitable constant $C = C(f)$ we have :

$$\sum_{n=1}^{N} f(n) = C + \int_1^N f(t)\,dt + \frac{f(N)}{2} + \sum_{1 \leq k \leq p} B_{2k} \frac{f^{(2k-1)}(N)}{(2k)!} + R_p(N) \ ,$$

where the remainder term $R_p(N)$ is “small” (in general smaller than the last term of the sum on k).

Series Summation III

Not only does it allow to compute sums of infinite series under suitable assumptions :

$$\sum_{n\geq 1} f(n) = \sum_{n=1}^{N} f(n) + \int_N^\infty f(t)\,dt - \frac{f(N)}{2} - \sum_{1\leq k\leq p} B_{2k}\frac{f^{(2k-1)}(N)}{(2k)!} - R_p(N)\ ,$$

but also limits, such as Euler's constant

$$\gamma = \lim_{N\to\infty}\left(\sum_{1\leq n\leq N} 1/n - \log(N)\right) .$$

The convergence is exponential, essentially in $e^{-2\pi N}$. Thus for $N = 100$, we usually get more than 270 decimals !

Series Summation III

Not only does it allow to compute sums of infinite series under suitable assumptions :

$$\sum_{n\geq 1} f(n) = \sum_{n=1}^{N} f(n) + \int_N^{\infty} f(t)\,dt - \frac{f(N)}{2} - \sum_{1\leq k\leq p} B_{2k}\frac{f^{(2k-1)}(N)}{(2k)!} - R_p(N) ,$$

but also limits, such as Euler's constant

$$\gamma = \lim_{N\to\infty}\Big(\sum_{1\leq n\leq N} 1/n - \log(N)\Big) .$$

The convergence is exponential, essentially in $e^{-2\pi N}$. Thus for $N = 100$, we usually get more than 270 decimals !

Series Summation IV

Amusing consequence of Euler–MacLaurin : let

$$S = 4 \sum_{1 \leq n \leq 500000} \frac{(-1)^{n-1}}{2n-1} .$$

One computes that

$S = 3.14159065358979324046264338326950288419729139937510305$
$\pi = 3.14159265358979323846264338327950288419716939937510582$

Here the numbers which appear are the Euler numbers instead of the Bernoulli numbers, because we have an alternating sum : $E_0 = 1$, $E_1 = 1$, $E_2 = 5$, $E_3 = 61$.

Series Summation IV

Amusing consequence of Euler–MacLaurin : let

$$S = 4 \sum_{1 \leq n \leq 500000} \frac{(-1)^{n-1}}{2n-1} \, .$$

One computes that

$$S = 3.1415906535897932404626433832695028841972913993751030 5$$
$$\pi = 3.1415926535897932384626433832795028841971693993751058 2$$

Here the numbers which appear are the Euler numbers instead of the Bernoulli numbers, because we have an alternating sum : $E_0 = 1$, $E_1 = 1$, $E_2 = 5$, $E_3 = 61$.

(3). Summation of Alternating Series

$S = \sum_{n\geq 0}(-1)^n a(n)$ with $a(n) \geq 0$. Of course, difference of two series with positive terms, but specific method (note that for $\zeta(3)$ we have the trivial identity $\sum_{n\geq 1}(-1)^n/n^3 = -(3/4)\zeta(3)$, and similarly for all values).

Idea : write $a(n)$ as the moment of a measure :

$$a(n) = \int_0^1 x^n w(x)\,dx$$

for a certain positive weight w. We then have :

$$S = \sum_{n\geq 0}(-1)^n a(n) = \int_0^1 \frac{w(x)}{1+x}\,dx\ .$$

Possible if and only if $a(n)$ is totally monotone ($a(n)$ and all its successive forward differences are nonnegative). However method valid much more generally : can even compute $\zeta(s)$ for complex s ! ! !

(3). Summation of Alternating Series

$S = \sum_{n\ge0}(-1)^n a(n)$ with $a(n) \ge 0$. Of course, difference of two series with positive terms, but specific method (note that for $\zeta(3)$ we have the trivial identity $\sum_{n\ge1}(-1)^n/n^3 = -(3/4)\zeta(3)$, and similarly for all values).

Idea : write $a(n)$ as the moment of a measure :

$$a(n) = \int_0^1 x^n w(x)\,dx$$

for a certain positive weight w. We then have :

$$S = \sum_{n\ge0}(-1)^n a(n) = \int_0^1 \frac{w(x)}{1+x}\,dx\ .$$

Possible if and only if $a(n)$ is totally monotone ($a(n)$ and all its successive forward differences are nonnegative). However method valid much more generally : can even compute $\zeta(s)$ for complex s ! ! !

(3). Summation of Alternating Series

$S = \sum_{n\geq 0}(-1)^n a(n)$ with $a(n) \geq 0$. Of course, difference of two series with positive terms, but specific method (note that for $\zeta(3)$ we have the trivial identity $\sum_{n\geq 1}(-1)^n/n^3 = -(3/4)\zeta(3)$, and similarly for all values).

Idea : write $a(n)$ as the moment of a measure :

$$a(n) = \int_0^1 x^n w(x)\, dx$$

for a certain positive weight w. We then have :

$$S = \sum_{n\geq 0}(-1)^n a(n) = \int_0^1 \frac{w(x)}{1+x}\, dx \; .$$

Possible if and only if $a(n)$ is totally monotone ($a(n)$ and all its successive forward differences are nonnegative). However method valid much more generally : can even compute $\zeta(s)$ for complex s ! ! !

Series Summation VI

Let P_n be a polynomial of degree n such that $P_n(-1) \neq 0$. Then $(P_n(-1) - P_n(X))/(1 + X)$ is also a polynomial of degree $n - 1$, so if we write

$$\frac{P_n(-1) - P_n(X)}{1 + X} = \sum_{k=0}^{n-1} c_{n,k} X^k$$

we have

$$\frac{1}{P_n(-1)} \sum_{n=0}^{n-1} c_{n,k} a(k) = \frac{1}{P_n(-1)} \int_0^1 \frac{P_n(-1) - P_n(x)}{1 + x} w(x)\, dx = S - R_n\ ,$$

with

$$|R_n| \leq \frac{M_n}{|P_n(-1)|} \int_0^1 \frac{w(x)}{1 + x}\, dx = \frac{M_n}{|P_n(-1)|} S\ ,$$

where $M_n = \sup_{x \in [0,1]} |P_n(x)|$

Series Summation VII

Best choice of P_n to minimize $M_n/|P_n(-1)|$: $P_n(X) = T_n(1-2X)$, T_n Chebychev polynomial ($P_n(\sin^2 t) = \cos(2nt)$).

This leads to an incredibly simple algorithm :

$d \leftarrow (3+\sqrt{8})^n$; $d \leftarrow (d+1/d)/2$; $b \leftarrow -1$; $c \leftarrow -d$; $s \leftarrow 0$; For $k = 0, \ldots, n-1$ do :
$c \leftarrow b - c$; $s \leftarrow s + c \cdot a(k)$; $b \leftarrow (k+n)(k-n)b/((k+1/2)(k+1))$;
The result is s/d.

Convergence in 5.83^{-n}. This is the `sumalt` program of Pari/GP.

Series Summation VII

Best choice of P_n to minimize $M_n/|P_n(-1)|$: $P_n(X) = T_n(1-2X)$, T_n Chebychev polynomial ($P_n(\sin^2 t) = \cos(2nt)$).

This leads to an incredibly simple algorithm :

$d \leftarrow (3+\sqrt{8})^n$; $d \leftarrow (d+1/d)/2$; $b \leftarrow -1$; $c \leftarrow -d$; $s \leftarrow 0$; For $k = 0, \ldots, n-1$ do :
$c \leftarrow b - c$; $s \leftarrow s + c \cdot a(k)$; $b \leftarrow (k+n)(k-n)b/((k+1/2)(k+1))$;
The result is s/d.

Convergence in 5.83^{-n}. This is the `sumalt` program of Pari/GP.

(4). The Use of Complex Analysis

If f is holomorphic, $\pi\cot\text{an}(\pi t)f(t)$ has only simple poles for $t = n \in \mathbb{Z}$ with residue $f(n)$, so the residue theorem allows us to compute $\sum_{0\le n\le N} f(n)$ as a contour integral, and with suitable assumptions on f, even obtain ordinary integrals. For instance, under suitable assumptions which I hide on purpose, we have

$$\sum_{n\geq 0}{}' f(n) = \int_0^\infty f(t)\,dt + i\int_0^\infty \frac{f(iy)-f(-iy)}{e^{2\pi y}-1}\,dy\ ,$$

where $\sum'$ means that $f(0)$ must be understood as $\lim_{x\to 0} f(x)/2$. This is one of the forms of the Abel–Plana formula.

Amusing consequence : if in addition f is an even function, the second (complex) integral vanishes, hence we obtain equality of the sum and the integral (another sophomore's dream), but I repeat, **under suitable assumptions**.

Series Summation VIII

(4). The Use of Complex Analysis

If f is holomorphic, $\pi \operatorname{cotan}(\pi t) f(t)$ has only simple poles for $t = n \in \mathbb{Z}$ with residue $f(n)$, so the residue theorem allows us to compute $\sum_{0 \le n \le N} f(n)$ as a contour integral, and with suitable assumptions on f, even obtain ordinary integrals. For instance, under suitable assumptions which I hide on purpose, we have

$$\sum_{n \ge 0}{}' f(n) = \int_0^\infty f(t)\,dt + i \int_0^\infty \frac{f(iy) - f(-iy)}{e^{2\pi y} - 1}\,dy\ ,$$

where $\sum'$ means that $f(0)$ must be understood as $\lim_{x \to 0} f(x)/2$. This is one of the forms of the Abel–Plana formula.

Amusing consequence : if in addition f is an even function, the second (complex) integral vanishes, hence we obtain equality of the sum and the integral (another sophomore's dream), but I repeat, **under suitable assumptions**.

Series Summation IX

Example : we have

$$\sum_{n\geq 0}{}' \left(\frac{\sin(n)}{n}\right)^k = \int_0^\infty \left(\frac{\sin(t)}{t}\right)^k dt \;,$$

for $k = 1, 2, 3, 4, 5$, and 6, but not for $k = 7$.
(values $\pi(1/2, 1/2, 3/8, 1/3, 115/384, 11/40)$).

Even better :

$$\sum_{n\geq 0}{}' \left(\frac{\sin(n/100)}{n}\right)^k = \int_0^\infty \left(\frac{\sin(t/100)}{t}\right)^k dt \;,$$

for $1 \leq k \leq 628$, but not for $k = 629$. (**Hint :** $2\pi = 6.28318\cdots$.)

Series Summation IX

Example : we have

$$\sum_{n\geq 0}{}' \left(\frac{\sin(n)}{n}\right)^k = \int_0^\infty \left(\frac{\sin(t)}{t}\right)^k dt\ ,$$

for $k = 1$, 2, 3, 4, 5, and 6, but not for $k = 7$.
(values $\pi(1/2, 1/2, 3/8, 1/3, 115/384, 11/40)$).

Even better :

$$\sum_{n\geq 0}{}' \left(\frac{\sin(n/100)}{n}\right)^k = \int_0^\infty \left(\frac{\sin(t/100)}{t}\right)^k dt\ ,$$

for $1 \leq k \leq 628$, but not for $k = 629$. (**Hint :** $2\pi = 6.28318\cdots$.)

Part II : Euler Products and Sums I

Here want to compute sums or products over the set P of prime numbers, for example :

$$\sum_{p \in P} \frac{1}{p^2} \quad \text{and} \quad \prod_{p \in P} \left(1 - \frac{1}{p(p-1)}\right) .$$

Sums seen in Part I involve regular functions $f(n)$, for instance C^∞ functions multiplied by simple periodic functions such as $(-1)^n$ or more generally $\chi(n)$ for a character χ. A priori these are the only functions to which these methods can be applied, of course with small variants. However prime numbers are very irregularly distributed, and we want to compute values exact to hundreds of decimals.

Part II : Euler Products and Sums I

Here want to compute sums or products over the set P of prime numbers, for example :

$$\sum_{p\in P}\frac{1}{p^2} \quad\text{and}\quad \prod_{p\in P}\left(1-\frac{1}{p(p-1)}\right)\ .$$

Sums seen in Part I involve regular functions $f(n)$, for instance C^∞ functions multiplied by simple periodic functions such as $(-1)^n$ or more generally $\chi(n)$ for a character χ. A priori these are the only functions to which these methods can be applied, of course with small variants. However prime numbers are very irregularly distributed, and we want to compute values exact to hundreds of decimals.

Euler Products and Sums II

Fundamental Idea : known since Riemann of course : link prime numbers to regular functions, for instance the Riemann zeta function $\zeta(s)$ thanks to the usual

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = \prod_{p \in P} \left(1 - \frac{1}{p^s}\right)^{-1} .$$

Note that if we have Euler sums or products involving characters, we would of course use the corresponding L-function.
Thus, for $\Re(s) > 1$:

$$\log(\zeta(s)) = -\sum_{p \in P} \log(1 - 1/p^{-s}) = \sum_{k \geq 1} \frac{1}{k} S(ks) ,$$

with $S(z) = \sum_{p \in P} 1/p^z$.

Euler Products and Sums III

We then use the Möbius inversion formula, and we obtain for $\Re(z) > 1$

$$S(z) = \sum_{k \geq 1} \frac{\mu(k)}{k} \log(\zeta(kz)) .$$

The computation of $\zeta(kz)$ is easy by using for instance Euler–MacLaurin seen above.

Convergence becomes much faster if we modify slightly the formula :

$$S(z) = \sum_{p \leq N,\ p \in P} \frac{1}{p^z} + \sum_{k \geq 1} \frac{\mu(k)}{k} \log(\zeta_{>N}(kz)) ,$$

where

$$\zeta_{>N}(s) = \zeta(s) - \sum_{n=1}^{N} \frac{1}{n^s} .$$

A reasonable choice is $N = 30$.

Euler Products and Sums III

We then use the Möbius inversion formula, and we obtain for $\Re(z) > 1$

$$S(z) = \sum_{k \geq 1} \frac{\mu(k)}{k} \log(\zeta(kz)) \, .$$

The computation of $\zeta(kz)$ is easy by using for instance Euler–MacLaurin seen above.

Convergence becomes much faster if we modify slightly the formula :

$$S(z) = \sum_{p \leq N,\ p \in P} \frac{1}{p^z} + \sum_{k \geq 1} \frac{\mu(k)}{k} \log(\zeta_{>N}(kz)) \, ,$$

where

$$\zeta_{>N}(s) = \zeta(s) - \sum_{n=1}^{N} \frac{1}{n^s} \, .$$

A reasonable choice is $N = 30$.

Euler Products and Sums IV

For instance we can compute $\sum_{p \in P} 1/p^2$ to thousands of decimal places, but also $\sum_{p \in P} 1/(p \log(p))$ (although this converges incredibly slowly), or $\lim_{N \to \infty} \left(\sum_{p \le N} 1/p - \log(\log(N)) \right)$, or even $\sum_{p \in P} \log(\log(p))^k / (p \log(p))$.

Similarly, can compute Euler products by taking logarithms.

Sample numerical values :

$$\sum_{p \in P} \frac{1}{p^2} = 0.452247420041065498506543364832247934173$$

$$\sum_{p \in P} \frac{1}{p \log p} = 1.636616323351260868569658003921863671188$$

$$\prod_{p \in P} \left(1 - \frac{1}{p(p-1)}\right) = 0.373955813619202288054728054346416415111$$

Euler Products and Sums IV

For instance we can compute $\sum_{p\in P} 1/p^2$ to thousands of decimal places, but also $\sum_{p\in P} 1/(p\log(p))$ (although this converges incredibly slowly), or $\lim_{N\to\infty}\left(\sum_{p\le N} 1/p - \log(\log(N))\right)$, or even $\sum_{p\in P} \log(\log(p))^k/(p\log(p))$.

Similarly, can compute Euler products by taking logarithms.

Sample numerical values :

$$\sum_{p\in P}\frac{1}{p^2} = 0.45224742004106549850654336483224793417$$

$$\sum_{p\in P}\frac{1}{p\log p} = 1.63661632335126086856965800392186367118$$

$$\prod_{p\in P}\left(1-\frac{1}{p(p-1)}\right) = 0.37395581361920228805472805434641641511$$

Euler Products and Sums IV

For instance we can compute $\sum_{p\in P} 1/p^2$ to thousands of decimal places, but also $\sum_{p\in P} 1/(p\log(p))$ (although this converges incredibly slowly), or $\lim_{N\to\infty}\left(\sum_{p\le N} 1/p - \log(\log(N))\right)$, or even $\sum_{p\in P} \log(\log(p))^k/(p\log(p))$.

Similarly, can compute Euler products by taking logarithms.

Sample numerical values :

$$\sum_{p\in P}\frac{1}{p^2} = 0.452247420041065498506543364832247934173$$

$$\sum_{p\in P}\frac{1}{p\log p} = 1.63661632335126086856965800392186367118$$

$$\prod_{p\in P}\left(1-\frac{1}{p(p-1)}\right) = 0.37395581361920228805472805434641641511$$

Euler Products and Sums V

As mentioned above, can be used for Euler products (or sums) with character, such as

$$L(\chi,3)=\prod_{p\in P}(1-\chi(p)/p^3)^{-1}\ ,$$

but **warning !** the conductor D of the character must be reasonable (for instance $D<10^6$).

In fact, if for a positive fundamental discriminant we set

$$L(D,3)=\prod_{p\in P}\left(1-\left(\frac{D}{p}\right)/p^3\right)^{-1}=\sum_{n\geq 1}\frac{\left(\frac{D}{n}\right)}{n^3}\ ,$$

then for instance for $D>10^{50}$ nobody knows how to compute this to more than 20 decimals, say (which can be done by naive sum or product), neither by the above method, nor by Euler–MacLaurin.

Euler Products and Sums V

As mentioned above, can be used for Euler products (or sums) with character, such as

$$L(\chi,3)=\prod_{p\in P}(1-\chi(p)/p^3)^{-1}\ ,$$

but **warning !** the conductor D of the character must be reasonable (for instance $D<10^6$).

In fact, if for a positive fundamental discriminant we set

$$L(D,3)=\prod_{p\in P}\left(1-\left(\frac{D}{p}\right)/p^3\right)^{-1}=\sum_{n\geq 1}\frac{\left(\frac{D}{n}\right)}{n^3}\ ,$$

then for instance for $D>10^{50}$ nobody knows how to compute this to more than 20 decimals, say (which can be done by naive sum or product), neither by the above method, nor by Euler–MacLaurin.

Part III : Numerical Integration I

Certainly the field where the most spectacular progress has been made in the past thirty years, essentially due to Japanese mathematicians. Want to compute numerically (assuming convergence, and as usual to hundreds of decimal places) integrals such as

$$\int_a^b f(x)\,dx\ , \quad \int_0^\infty f(x)\,dx\ , \quad \int_{-\infty}^\infty f(x)\,dx\ .$$

If $f(x)$ is irregular, nothing really better than naive trapezoidal rule. But the more f is regular, the more there exist efficient methods.

Classical Methods : Simpson and generalizations such as Romberg, or Gaussian integration, or others. Almost unusable if we want more than 20 decimal places, which is sufficient in numerical analysis, but not in number theory.

Part III : Numerical Integration I

Certainly the field where the most spectacular progress has been made in the past thirty years, essentially due to Japanese mathematicians. Want to compute numerically (assuming convergence, and as usual to hundreds of decimal places) integrals such as

$$\int_a^b f(x)\,dx\ , \quad \int_0^\infty f(x)\,dx\ , \quad \int_{-\infty}^\infty f(x)\,dx\ .$$

If $f(x)$ is irregular, nothing really better than naive trapezoidal rule. But the more f is regular, the more there exist efficient methods.

Classical Methods : Simpson and generalizations such as Romberg, or Gaussian integration, or others. Almost unusable if we want more than 20 decimal places, which is sufficient in numerical analysis, but not in number theory.

Numerical Integration II

It is only around 1970 that Mori and Takahashi noticed that if f is holomorphic (or even meromorphic) in a neighborhood of the interval of integration one can do much better, and rapidly obtain thousands of decimals if desired. Note that we need this holomorphy condition even though we are integrating a real function on a real interval.

Two fundamental ideas :

• (1). If F is a holomorphic function which tends to 0 sufficiently fast when $x \to \pm\infty$, x real, then the most efficient method to compute $\int_{\mathbb{R}} F(t)\,dt$ is indeed the trapezoidal rule. Note that this is a theorem, not so difficult but a little surprising nonetheless.

In practice, "sufficiently fast" means at least like e^{-ax^2} ($e^{-a|x|}$ is too slow), but best results obtained with a function which tends to 0 doubly exponentially fast, for instance as $\exp(-\exp(a|x|))$. Note that it is slightly worse to choose functions which tend to 0 even faster.

Numerical Integration II

It is only around 1970 that Mori and Takahashi noticed that if f is holomorphic (or even meromorphic) in a neighborhood of the interval of integration one can do much better, and rapidly obtain thousands of decimals if desired. Note that we need this holomorphy condition even though we are integrating a real function on a real interval.

Two fundamental ideas :

• (1). If F is a holomorphic function which tends to 0 sufficiently fast when $x \to \pm\infty$, x real, then the most efficient method to compute $\int_{\mathbb{R}} F(t)\,dt$ is indeed the trapezoidal rule. Note that this is a theorem, not so difficult but a little surprising nonetheless.

In practice, "sufficiently fast" means at least like e^{-ax^2} ($e^{-a|x|}$ is too slow), but best results obtained with a function which tends to 0 doubly exponentially fast, for instance as $\exp(-\exp(a|x|))$. Note that it is slightly worse to choose functions which tend to 0 even faster.

Numerical Integration II

It is only around 1970 that Mori and Takahashi noticed that if f is holomorphic (or even meromorphic) in a neighborhood of the interval of integration one can do much better, and rapidly obtain thousands of decimals if desired. Note that we need this holomorphy condition even though we are integrating a real function on a real interval.

Two fundamental ideas :

• (1). If F is a holomorphic function which tends to 0 sufficiently fast when $x \to \pm\infty$, x real, then the most efficient method to compute $\int_{\mathbb{R}} F(t)\,dt$ is indeed the trapezoidal rule. Note that this is a theorem, not so difficult but a little surprising nonetheless.
In practice, "sufficiently fast" means at least like e^{-ax^2} ($e^{-a|x|}$ is too slow), but best results obtained with a function which tends to 0 doubly exponentially fast, for instance as $\exp(-\exp(a|x|))$. Note that it is slightly worse to choose functions which tend to 0 even faster.

Numerical Integration III

• (2). The second idea is to reduce to to the case of doubly exponential decrease by change of variable, which must be holomorphic. Choosing one of the simplest examples, to compute

$$\int_{-1}^{1} f(x)\,dx$$

one makes the change of variable

$$x = \phi(t) := \tanh(\sinh(t)) \ ,$$

leading to the formula

$$\int_{-1}^{1} f(x)\,dx = \int_{-\infty}^{\infty} F(t)\,dt \quad \text{with} \quad F(t) = f(\phi(t))\phi'(t) \ ,$$

and $\phi'(t)$ tends to 0 doubly exponentially when $|t| \to \infty$, one exponential coming from sinh, the second from tanh.

Numerical Integration IV

With the above notation :

$$\int_{-1}^{1} f(x)\,dx = h\sum_{n=-N}^{N} f(\phi(nh))\phi'(nh) + R_N(h)\ ,$$

and can show that under suitable holomorphy assumptions on f, and choosing for instance $h = a\log(N)/N$ for a constant a close to 1, we have $R_N(h) = O(e^{-bN/\log(N)})$ for some other constant b, hence essentially exponential convergence of the method.

Typically, to obtain a few hundred decimal places (absolutely impossible with classical methods, and usually more than enough even in number theory), one can choose for instance $h = 1/200$ and $N = 500$, hence only 1000 evaluations of f ! ! !. This is the `intnum` routine of Pari/GP.

Numerical Integration IV

With the above notation :

$$\int_{-1}^{1} f(x)\,dx = h \sum_{n=-N}^{N} f(\phi(nh))\phi'(nh) + R_N(h) ,$$

and can show that under suitable holomorphy assumptions on f, and choosing for instance $h = a\log(N)/N$ for a constant a close to 1, we have $R_N(h) = O(e^{-bN/\log(N)})$ for some other constant b, hence essentially exponential convergence of the method.

Typically, to obtain a few hundred decimal places (absolutely impossible with classical methods, and usually more than enough even in number theory), one can choose for instance $h = 1/200$ and $N = 500$, hence only 1000 evaluations of f !!!. This is the `intnum` routine of Pari/GP.

Numerical Integration V

Here is a proven and precise theorem (P. Molin) : assume that f is holomorphic on the disc D centered at 0 with radius 2, which in particular contains the real interval $[-1,1]$. Then for all $N \geq 1$ we have

$$\left|\int_{-1}^{1} f(x)\,dx - \sum_{k=-N}^{N} a_k f(x_k)\right| \leq \left(e^4 \sup_D |f|\right) \exp\left(-\frac{5N}{\log(5N)}\right)\ ,$$

where

$$h = \frac{\log(5N)}{N}\ , \quad a_k = \frac{h\cosh(kh)}{\cosh^2(\sinh(kh))}\ , \quad \text{and} \quad x_k = \tanh(\sinh(kh))\ .$$

Numerical Integration VI

For integrals on $[a, b]$ with a and b finite, we reduce to $[-1, 1]$ by linear changes of variable. When the function has an algebraic singularity, one uses polynomial changes of variable. When one or both of the limits are infinite, use of other changes of variable :

- For $\int_0^\infty f(x)\,dx$, where f does not tend to 0 exponentially fast (for example $f(x) \sim 1/x^k$), we use $x = \phi(t) := \exp(\sinh(t))$.
- For $\int_0^\infty f(x)\,dx$ where f does tend to 0 exponentially fast (for example $f(x) \sim e^{-ax}$), we use $x = \phi(t) := \exp(t - \exp(-t))$.
- For $\int_{-\infty}^\infty f(x)\,dx$, we use $x = \phi(t) := \sinh(\sinh(t))$ if f does not tend to 0 exponentially fast and $x = \phi(t) := \sinh(t)$ otherwise.

Numerical Integration VI

For integrals on $[a, b]$ with a and b finite, we reduce to $[-1, 1]$ by linear changes of variable. When the function has an algebraic singularity, one uses polynomial changes of variable. When one or both of the limits are infinite, use of other changes of variable :

- For $\int_0^\infty f(x)\,dx$, where f does not tend to 0 exponentially fast (for example $f(x) \sim 1/x^k$), we use $x = \phi(t) := \exp(\sinh(t))$.
- For $\int_0^\infty f(x)\,dx$ where f does tend to 0 exponentially fast (for example $f(x) \sim e^{-ax}$), we use $x = \phi(t) := \exp(t - \exp(-t))$.
- For $\int_{-\infty}^\infty f(x)\,dx$, we use $x = \phi(t) := \sinh(\sinh(t))$ if f does not tend to 0 exponentially fast and $x = \phi(t) := \sinh(t)$ otherwise.

Numerical Integration VI

For integrals on $[a, b]$ with a and b finite, we reduce to $[-1, 1]$ by linear changes of variable. When the function has an algebraic singularity, one uses polynomial changes of variable. When one or both of the limits are infinite, use of other changes of variable :

- For $\int_0^\infty f(x)\,dx$, where f does not tend to 0 exponentially fast (for example $f(x) \sim 1/x^k$), we use $x = \phi(t) := \exp(\sinh(t))$.
- For $\int_0^\infty f(x)\,dx$ where f does tend to 0 exponentially fast (for example $f(x) \sim e^{-ax}$), we use $x = \phi(t) := \exp(t - \exp(-t))$.
- For $\int_{-\infty}^\infty f(x)\,dx$, we use $x = \phi(t) := \sinh(\sinh(t))$ if f does not tend to 0 exponentially fast and $x = \phi(t) := \sinh(t)$ otherwise.

Numerical Integration VI

For integrals on $[a, b]$ with a and b finite, we reduce to $[-1, 1]$ by linear changes of variable. When the function has an algebraic singularity, one uses polynomial changes of variable. When one or both of the limits are infinite, use of other changes of variable :

• For $\int_0^\infty f(x)\,dx$, where f does not tend to 0 exponentially fast (for example $f(x) \sim 1/x^k$), we use $x = \phi(t) := \exp(\sinh(t))$.

• For $\int_0^\infty f(x)\,dx$ where f does tend to 0 exponentially fast (for example $f(x) \sim e^{-ax}$), we use $x = \phi(t) := \exp(t - \exp(-t))$.

• For $\int_{-\infty}^\infty f(x)\,dx$, we use $x = \phi(t) := \sinh(\sinh(t))$ if f does not tend to 0 exponentially fast and $x = \phi(t) := \sinh(t)$ otherwise.

Numerical Integration VII

• For oscillating integrals such as $\int_0^\infty f(x)\sin(x)\,dx$, more subtle but similar method exists, only if, as here, the oscillations are completely under control.

Must pay attention to the relative closeness of the poles to the path of integration. Numerical example with the function $f(t) = 1/(1+t^2)$: On $[0,\infty]$ perfect. On $[0,1000]$ perfect. On $[-\infty,\infty]$ perfect.

But on $[-1000,1000]$, totally wrong (not a single correct decimal) because poles $\pm i$ too "close" after the linear change of variable (does not happen on $[-\infty,\infty]$ because different change of variable).

Numerical Integration VII

- For oscillating integrals such as $\int_0^\infty f(x)\sin(x)\,dx$, more subtle but similar method exists, only if, as here, the oscillations are completely under control.

Must pay attention to the relative closeness of the poles to the path of integration. Numerical example with the function $f(t) = 1/(1+t^2)$: On $[0,\infty]$ perfect. On $[0,1000]$ perfect. On $[-\infty,\infty]$ perfect.

But on $[-1000,1000]$, totally wrong (not a single correct decimal) because poles $\pm i$ too “close” after the linear change of variable (does not happen on $[-\infty,\infty]$ because different change of variable).

Numerical Integration VII

• For oscillating integrals such as $\int_0^\infty f(x)\sin(x)\,dx$, more subtle but similar method exists, only if, as here, the oscillations are completely under control.

Must pay attention to the relative closeness of the poles to the path of integration. Numerical example with the function $f(t) = 1/(1+t^2)$: On $[0,\infty]$ perfect. On $[0,1000]$ perfect. On $[-\infty,\infty]$ perfect.

But on $[-1000,1000]$, totally wrong (not a single correct decimal) because poles $\pm i$ too “close” after the linear change of variable (does not happen on $[-\infty,\infty]$ because different change of variable).

Numerical Integration VIII

One of the important applications of the method : computation of inverse Mellin transforms of gamma products, essential for L-function computations. Comparable to other methods (M. Rubinstein), a bit slower in low accuracy, but better for higher accuracy. In addition, proven error term.

Reference : recent PhD thesis (in French) of Pascal Molin, Bordeaux, 2010.

Numerical Integration VIII

One of the important applications of the method : computation of inverse Mellin transforms of gamma products, essential for L-function computations. Comparable to other methods (M. Rubinstein), a bit slower in low accuracy, but better for higher accuracy. In addition, proven error term.

Reference : recent PhD thesis (in French) of Pascal Molin, Bordeaux, 2010.