CHAPTER 4

Differences

4.1. Newton Series

On the contents of the lecture. The formula with the binomial series was engraved on Newton's gravestone in 1727 at Westminster Abbey.

Interpolation problem. Suppose we know the values of a function f at some points called *interpolation nodes* and we would like to interpolate the value of f at some point, not contained in the data. This is the so-called *interpolation problem.* Interpolation was applied in the computation of logarithms, maritime navigation, astronomical observations and in a lot of other things.

A natural idea is to construct a polynomial which takes given values at the interpolation nodes and consider its value at the point of interest as the interpolation. Values at $n+1$ points define a unique polynomial of degree n, which takes just these values at these points. In 1676 Newton discovered a formula for this polynomial, which is now called *Newton's interpolation formula.*

Consider the case, when interpolation nodes are natural numbers. Recall that the difference of a function f is the function denoted δf and defined by $\delta f(x) = f(x+1) - f(x)$. Define iterated differences $\delta^k f$ by induction: $\delta^0 f = f$, $\delta^{k+1} f = \delta(\delta^k f)$. Recall that $x^{\underline{k}}$ denotes the k-th factorial power $x^{\underline{k}} = x(x-1)\dots(x-k+1)$.

LEMMA 4.1.1. *For any polynomial $P(x)$, its difference $\Delta P(x)$ is a polynomial of degree one less.*

PROOF. The proof is by induction on the degree of $P(x)$. The difference is constant for any polynomial of degree 1. Indeed, $\delta(ax+b) = a$. Suppose the lemma is proved for polynomials of degree $\leq n$ and let $P(x) = \sum_{k=0}^{n+1} a_k x^k$ be a polynomial of degree $n+1$. Then $P(x) - a_{n+1}x^{\underline{n+1}} = Q(x)$ is a polynomial of degree $\leq n$. $\Delta P(x) = \Delta a x^{\underline{n+1}} + \Delta Q(x)$. By the induction hypothesis, $\Delta Q(x)$ has degree $\leq n-1$ and, as we know, $\Delta x^{\underline{n+1}} = (n+1)x^{\underline{n}}$ has degree n. □

LEMMA 4.1.2. *If $\Delta P(x) = 0$, and $P(x)$ is a polynomial, then $P(x)$ is constant.*

PROOF. If $\Delta P(x) = 0$, then degree of $P(x)$ cannot be positive by Lemma 4.1.1, hence $P(x)$ is constant. □

LEMMA 4.1.3 (Newton Polynomial Interpolation Formula). *For any polynomial $P(x)$*

$$P(x) = \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}}. \tag{4.1.1}$$

PROOF. If $P(x) = ax+b$, then $\Delta^0 P(0) = b$, $\Delta^1 P(0) = a$ and $\delta^k P(x) = 0$ for $k > 1$. Hence the Newton series (4.1.1) turns into $b + ax$. This proves our assertion for polynomials of degree ≤ 1. Suppose it is proved for polynomials of degree n. Consider $P(x)$ of degree $n+1$. Then $\Delta P(x) = \sum_{k=1}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}}$ by the induction hypothesis. Denote by $Q(x)$ the Newton series $\sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}}$ for $P(x)$.

Then

$$\begin{aligned}
\Delta Q(x) &= \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!}(x+1)^{\underline{k}} - \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}} \\
&= \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} \Delta x^{\underline{k}} \\
&= \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} k x^{\underline{k-1}} \\
&= \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{(k-1)!} x^{\underline{k-1}} \\
&= \sum_{k=0}^{\infty} \frac{\delta^k(\delta P(0))}{k!} x^{\underline{k}} \\
&= \Delta P(x).
\end{aligned}$$

Hence $\Delta(P(x) - Q(x)) = 0$ and $P(x) = Q(x) + c$. Since $P(0) = Q(0)$, one gets $c = 0$. This proves $P(x) = Q(x)$. $\square$

LEMMA 4.1.4 (Lagrange Formula). *For any sequence* $\{y_k\}_{k=0}^n$, *the polynomial* $L_n(x) = \sum_{k=0}^n (-1)^{n-k} \frac{y_k}{k!(n-k)!} \frac{x^{\underline{n+1}}}{x-k}$ *has the property* $L_n(k) = y_k$ *for* $0 \le k \le n$.

PROOF. For $x = k$, all terms of the sum $\sum_{k=0}^n (-1)^{n-k} \frac{y_k}{k!(n-k)!} \frac{x^{\underline{n+1}}}{x-k}$ but the k-th vanish, and $\frac{x^{\underline{k}}}{x-k}$ is equal to $k!(n-k)!(-1)^{n-k}$. $\square$

LEMMA 4.1.5. *For any function* f *and for any natural number* $m \le n$ *one has* $f(m) = \sum_{k=0}^n \frac{\delta^k f(0)}{k!} m^{\underline{k}}$.

PROOF. Consider the Lagrange polynomial L_n such that $L_n(k) = f(k)$ for $k \le n$. Then $\delta^k L_n(0) = \delta^k f(0)$ for all $k \le n$ and $\delta^k L_n(0) = 0$ for $k > n$, because the degree of L_n is n. Hence, $L_n(x) = \sum_{k=0}^{\infty} \frac{\delta^k f(0)}{k!} x^{\underline{k}} = \sum_{k=0}^{n} \frac{\delta^k f(0)}{k!} x^{\underline{k}}$ by Lemma 4.1.3. Putting $x = m$ in the latter equality, one gets $f(m) = L_n(m) = \sum_{k=0}^n \frac{\delta^k f(0)}{k!} m^{\underline{k}}$. $\square$

We see that the Newton polynomial gives a solution for the interpolation problem and our next goal is to estimate the interpolation error.

Theorem on extremal values. The least upper bound of a set of numbers A is called the *supremum* of A and denoted by $\sup A$. In particular, the ultimate sum of a positive series is the supremum of its partial sums. And the variation of a function on an interval is the supremum of its partial variations.

Dually, the greatest lower bound of a set A is called the *infinum* and denoted by $\inf A$.

THEOREM 4.1.6 (Weierstrass). *If a function* f *is continuous on an interval* $[a, b]$, *then it takes maximal and minimal values on* $[a, b]$.

PROOF. The function f is bounded by Lemma 3.6.3. Denote by B the supremum of the set of values of f on $[a, b]$. If f does not take the maximum value, then $f(x) \ne B$ for all $x \in [a, b]$. In this case $\frac{1}{B - f(x)}$ is a continuous function on $[a, b]$.

Hence it is bounded by Lemma 3.6.3. But the difference $B - f(x)$ takes arbitrarily small values, because $B - \varepsilon$ does not bound $f(x)$. Therefore $\frac{1}{B-f(x)}$ is not bounded. This is in contradiction to Lemma 3.6.3, which states that a locally bounded function is bounded. The same arguments prove that $f(x)$ takes its minimal value on $[a, b]$. $\square$

THEOREM 4.1.7 (Rolle). *If a function f is continuous on the interval $[a, b]$, differentiable in interval (a, b) and $f(a) = f(b)$, then $f'(c) = 0$ for some $c \in (a, b)$.*

PROOF. If the function f is not constant on $[a, b]$ then either its maximal value or its minimal value differs from $f(a) = f(b)$. Hence at least one of its extremal values is taken in some point $c \in (a, b)$. Then $f'(c) = 0$ by Lemma 3.2.1. $\square$

LEMMA 4.1.8. *If an n-times differentiable function $f(x)$ has $n + 1$ roots in the interval $[a, b]$, then $f^{(n)}(\xi) = 0$ for some $\xi \in (a, b)$.*

PROOF. The proof is by induction. For $n = 1$ this is Rolle's theorem. Let $\{x_k\}_{k=0}^{n}$ be a sequence of roots of f. By Rolle's theorem any interval (x_i, x_{i+1}) contains a root of f'. Hence f' has $n - 1$ roots, and its $(n - 1)$-th derivative has a root. But the $(n - 1)$-th derivative of f' is the n-th derivative of f. $\square$

THEOREM 4.1.9 (Newton Interpolation Formula). *Let f be an $n + 1$ times differentiable function on $I \supset [0, n]$. Then for any $x \in I$ there is $\xi \in I$ such that*

$$f(x) = \sum_{k=0}^{n} \frac{\delta^k f(0)}{k!} x^{\underline{k}} + \frac{f^{(k+1)}(\xi)}{(k+1)!} x^{\underline{k+1}}.$$

PROOF. The formula holds for $x \in \{0, 1, \dots n\}$ and any ξ, due to Lemma 4.1.5, because $x^{\underline{n+1}} = 0$ for such x. For other x one has $x^{\underline{n+1}} \neq 0$, hence there is C such that $f(x) = \sum_{k=0}^{n} \frac{\delta^k f(0)}{k!} x^{\underline{k}} + C x^{\underline{k+1}}$. The function $F(y) = f(y) - \sum_{k=0}^{n} \frac{\delta^k f(0)}{k!} x^{\underline{k}} - C y^{\underline{k+1}}$ has roots $0, 1, \dots, n, x$. Hence its $(n + 1)$-th derivative has a root $\xi \in I$. Since $\sum_{k=0}^{n} \frac{\delta^k f(0)}{k!} x^{\underline{k}}$ is a polynomial of degree n its $(n + 1)$-th derivative is 0. And the $(n + 1)$-th derivatives of $C x^{\underline{n+1}}$ and $C x^{n+1}$ coincide, because their difference is a polynomial of degree n. Hence $0 = F^{(n+1)}(\xi) = f^{(n+1)}(\xi) - C(n + 1)!$ and $C = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. $\square$

Binomial series. The series $\sum_{k=0}^{\infty} \frac{\delta^k f(0)}{k!} x^{\underline{k}}$ is called the *Newton series* of a function f. The Newton series coincides with the function at all natural points. And sometimes it converges to the function. The most important example of such convergence is given by the so-called *binomial series*.

Consider the function $(1 + x)^y$. This is a function of two variables. Fix x and evaluate its difference with respect to y. One has $\delta_y(1+x)^y = (1+x)^{y+1} - (1+x)^y = (1+x)^y(1+x-1) = x(1+x)^y$. This simple formula allows us immediately to evaluate $\delta_y^k (1 + x)^y = x^k (1 + x)^y$. Hence the Newton series for $(1 + x)^y$ as function of y is

$$(1 + x)^y = \sum_{k=0}^{\infty} \frac{x^k y^{\underline{k}}}{k!}. \tag{4.1.2}$$

For fixed y and variable x, the formula (4.1.2) represents the famous Newton binomial series. Our proof is not correct. We applied Newton's interpolation formula, proved only for polynomials, to an exponential function. But Newton's original proof was essentially of the same nature. Instead of interpolation of the whole function, he interpolated coefficients of its power series expansion. Newton considered the discovery of the binomial series as one of his greatest discoveries. And the role of the binomial series in further developments is very important.

For example, Newton expands into a power series $\arcsin x$ in the following way. One finds the derivative of $\arcsin x$ by differentiating identity $\sin \arcsin x = x$. This differentiation gives $\cos(\arcsin x) \arcsin' x = 1$. Hence $\arcsin' x = \frac{1}{\cos \arcsin x} = (1-x^2)^{-\frac{1}{2}}$. Since

$$(1-x^2)^{-\frac{1}{2}} = \sum_{k=0}^{\infty} \frac{(-x^2)^k (-\frac{1}{2})^{\underline{k}}}{k!} = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{2k!!} x^{2k}, \tag{4.1.3}$$

one gets the series for arcsin by termwise integration of (4.1.3). The result is

$$\arcsin x = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{2k!!} \frac{x^{2k+1}}{2k+1}.$$

It was more than a hundred years after the discovery Newton's Binomial Theorem that it was first completely proved by Abel.

THEOREM 4.1.10. *For any complex z and ζ such that $|z| < 1$, the series $\sum_{k=0}^{\infty} \frac{z^k \zeta^{\underline{k}}}{k!}$ absolutely converges to* $(1+z)^\zeta = \exp(\zeta \ln(1+z))$.

PROOF. The analytic function $\exp \zeta \ln(1+z)$ of variable z has no singular points in the disk $|z| < 1$, hence its Taylor series converges to it. The derivative of $(1+z)^\zeta$ by z is $\zeta(1+z)^{\zeta-1}$. The k-th derivative is $\zeta^{\underline{k}}(1+z)^{\zeta-k}$. In particular, the value of k-th derivative for $z=0$ is equal to $\zeta^{\underline{k}}$. Hence the Taylor series of the function is $\sum_{k=0}^{\infty} \frac{\zeta^{\underline{k}} z^k}{k!}$. □

On the boundary of convergence. Since $(1+z)^\zeta$ has its only singular point on the circle $|z| = 1$, and this point is -1, the binomial series for all z on the circle has $(1+z)^\zeta$ as its Abel's sum. In particular, for $z=1$ one gets

$$\sum_{k=0}^{\infty} \frac{x^{\underline{k}}}{k!} = 2^x.$$

The series on the left-hand side converges for $x > 0$. Indeed, the series becomes alternating starting with $k > x$. The ratio $\frac{k-x}{k+1}$ of modules of terms next to each other is less then one. Hence the moduli of the terms form a monotone decreasing sequence onward $k > x$. And to apply the Leibniz Theorem 2.4.3, one needs only to prove that $\lim_{n\to\infty} \frac{x^{\underline{n}}}{n!} = 0$. Transform this limit into $\lim_{n\to\infty} \frac{x}{n} \prod_{k=1}^{n-1}(\frac{x}{k} - 1)$. The product $\prod_{k=1}^{n-1}(\frac{x}{k} - 1)$ contains at most x terms which have moduli greater than 1, and all terms of the product do not exceed x. Hence the absolute value of this product does not exceed x^x. And our sequence $\{\frac{x^{\underline{n}}}{n!}\}$ is majorized by an infinitesimally small $\{\frac{x^{x+1}}{n}\}$. Hence it is infinitesimally small.

Plain binomial theorem. For a natural exponent the binomial series contains only finitely many nonzero terms. In this case it turns into $(1+x)^n = \sum_{k=0}^{n} \frac{n^{\underline{k}} x^k}{k!}$.

Because $(a+b)^n = a^n(1+\frac{b}{a})^n$, one gets the following famous formula

$$(a+b)^n = \sum_{k=0}^{n+1} \frac{n^{\underline{k}}}{k!} a^k b^{n-k}.$$

This is the formula that is usually called Newton's Binomial Theorem. But this simple formula was known before Newton. In Europe it was proved by Pascal in 1654. Newton's discovery concerns the case of non integer exponents.

Symbolic calculus. One defines the *shift operation* $\mathbf{S}^a$ for a function f by the formula $\mathbf{S}^a f(x) = f(x+a)$. Denote by $\mathbf{1}$ the identity operation and by $\mathbf{S} = \mathbf{S}^1$. Hence $\mathbf{S}^0 = \mathbf{1}$. The composition of two operations is written as a product. So, for any a and b one has the following sum formula $\mathbf{S}^a\mathbf{S}^b = \mathbf{S}^{a+b}$.

We will consider only so-called *linear* operations. An operation O is called linear if $O(f+g) = O(f)+O(g)$ for all f, g and $O(kf) = kO(f)$ for any constant k. Define the sum $A+B$ of operations A and B by the formula $(A+B)f = Af+Bf$. Further, define the product of an operation A by a number k as $(kA)f = k(Af)$. For linear operations O, U, V one has the distributivity law $O(U+V) = OU+OV$. If the operations under consideration commute $UV = VU$, (for example, all iterations of the same operation commute) then they obey all usual numeric laws, and all identities which hold for numbers extend to operations. For example, $U^2 - V^2 = (U-V)(U+V)$, or the plain binomial theorem.

Let us say that an operation O is *decreasing* if for any polynomial P the degree of $O(P)$ is less than the degree of P. For example, the operation of difference $\delta = \mathbf{S} - \mathbf{1}$ and the operation $\mathbf{D}$ of differentiation $\mathbf{D}f(x) = f'(x)$ are decreasing. For a decreasing operation O, any power series $\sum_{k=0}^{\infty} a_k O^k$ defines an operation at least on polynomials, because this series applied to a polynomial contains only finitely many terms. Thus we can apply analytic functions to operations.

For example, the binomial series $(1+\delta)^y = \sum_{k=0}^{\infty} \frac{\delta^k y^{\underline{k}}}{k!}$ represents $\mathbf{S}^y$. And the equality $\mathbf{S}^y = \sum_{k=0}^{\infty} \frac{\delta^k y^{\underline{k}}}{k!}$, which is in fact the Newton Polynomial Interpolation Formula, is a direct consequence of binomial theorem. Another example, consider $\delta_n = \mathbf{S}^{\frac{x}{n}} - \mathbf{1}$. Then $\mathbf{S}^{\frac{x}{n}} = \mathbf{1} + \delta_n$ and $\mathbf{S}^x = (\mathbf{1}+\delta_n)^n$. Further, $\mathbf{S}^x = \sum_{k=0}^{n} \frac{n^{\underline{k}}\delta_n^k}{k!} = \sum_{k=0}^{\infty} \frac{n^{\underline{k}}}{n^k}\frac{(n\delta_n)^k}{k!}$. Now we follow Euler's method to "substitute $n = \infty$". Then $n\delta_n$ converts into $x\mathbf{D}$, and $\frac{n^{\underline{k}}}{n^k}$ turns into 1. As result we get the Taylor formula $\mathbf{S}^x = \sum_{k=0}^{\infty} \frac{x^k \mathbf{D}^k}{k!}$. Our proof is copied from the Euler proof in his *Introductio* of $\lim_{n\to\infty}(1+\frac{x}{n})^n = \sum_{k=0}^{\infty}\frac{x^k}{k!}$. This substitution of infinity means passing to the limit. This proof is sufficient for decreasing operations on polynomials because the series contains only finitely many nonzero terms. In the general case problems of convergence arise.

The binomial theorem was the main tool for the expansion of functions into power series in Euler's times. Euler also applied it to get power expansions for trigonometric functions.

The Taylor expansion for $x = 1$ gives a symbolic equality $\mathbf{S} = \exp \mathbf{D}$. Hence $\mathbf{D} = \ln \mathbf{S} = \ln(\mathbf{1}+\delta) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{\delta^k}{k}$. We get a formula for numerical differentiation. Symbolic calculations produce formulas which hold at least for polynomials.

Problems.

1. Prove $(x+y)^{\underline{n}} = \sum_{k=0}^{n} \frac{n^{\underline{k}} x^{\underline{k}} y^{\underline{n-k}}}{k!}$.

2. Evaluate $\sum_{k=0}^{n} \frac{n^{\underline{k}}}{k!} 2^{n-k}$.

3. Prove: If p is prime, then $\frac{p^{\underline{k}}}{k!}$ is divisible by p.

4. Prove: $\frac{n^{\underline{k}}}{k!} = \frac{n^{\underline{n-k}}}{(n-k)!}$.

5. Deduce the plain binomial theorem from multiplication of series for exponenta.

6. One defines the Catalan number c_n as the number of correct placement of brackets in the sum $a_1 + a_2 + \cdots + a_n$. Prove that Catalan numbers satisfy the following recursion equation $c_n = \sum_{k=0}^{n-1} c_k c_{n-k}$ and deduce a formula for Catalan numbers.

7. Prove that $\Delta^k x^{\underline{n}} x^{\underline{m}} = 0$ for $x = 0$ and $k < n$.

8. Prove that $\sum_{k=0}^{n} (-1)^k \frac{n^{\underline{k}}}{k!} = 0$.

9. Get a differential equation for the binomial series and solve it.

10. Prove $(a+b)^{\underline{n}} = \sum_{k=0}^{n} \frac{n^{\underline{k}}}{k!} a^{\underline{k}} b^{\underline{n-k}}$.

11. Prove: A sequence $\{a_k\}$ such that $\Delta^2 a_k \geq 0$ satisfies the inequality $\max\{a_1, \dots, a_n\} \geq a_k$ for any k between 1 and n.

12. Prove $\sum_{k=0}^{\infty} (-1)^k \frac{x^{\underline{2k}}}{2k!} = 2^{x/2} \cos \frac{x\pi}{4}$.

13. Prove $\sum_{k=0}^{\infty} (-1)^k \frac{x^{\underline{2k+1}}}{(2k+1)!} = 2^{x/2} \sin \frac{x\pi}{4}$.

14. Prove $\Delta^n 0^p$ is divisible by $p!$.

*__15.__ Prove that $\Delta^n 0^p = \sum_{k=0}^{n-1} (-1)^{n-k} \frac{n^{\underline{k}}}{k!} k^p$.

16. Prove $\cos^2 x + \sin^2 x = 1$ via power series.

4.2. Bernoulli Numbers

On the contents of the lecture. In this lecture we give explicit formulas for telescoping powers. These formulas involve a remarkable sequence of numbers, which were discovered by Jacob Bernoulli. They will appear in formulas for sums of series of reciprocal powers. In particular, we will see that $\frac{\pi^2}{6}$, the sum of Euler series, contains the second Bernoulli number $\frac{1}{6}$.

Summation Polynomials. Jacob Bernoulli found a general formula for the sum $\sum_{k=1}^n k^q$. To be precise he discovered that there is a sequence of numbers $B_0, B_1, B_2, \dots, B_n, \dots$ such that

$$\sum_{k=1}^{n} k^q = \sum_{k=0}^{q+1} B_k \frac{q^{\underline{k-1}} n^{q+1-k}}{k!}. \tag{4.2.1}$$

The first 11 of the *Bernoulli numbers* are $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, -\frac{1}{30}, 0, \frac{5}{66}$. The right-hand side of (4.2.1) is a polynomial of degree $q+1$ in n. Let us denote this polynomial by $\psi_{q+1}(n)$. It has the following remarkable property: $\delta\psi_{q+1}(x) = (1+x)^q$. Indeed the latter equality holds for any natural value n of the variable, hence it holds for all x, because two polynomials coinciding in infinitely many points coincide. Replacing in (4.2.1) $q+1$ by m, n by x and reversing the order of summation, one gets the following:

$$\begin{aligned}
\psi_m(x) &= \sum_{k=0}^{m} B_{m-k} \frac{(m-1)^{\underline{m-k-1}}}{(m-k)!} x^k \\
&= \sum_{k=0}^{m} B_{m-k} \frac{(m-1)!}{k!(m-k)!} x^k \\
&= \sum_{k=0}^{m} B_{m-k} \frac{(m-1)^{\underline{k-1}}}{k!} x^k.
\end{aligned}$$

Today's lecture is devoted to the proof of this Bernoulli theorem.

Telescoping powers. Newton's Formula represents x^m as a factorial polynomial $\sum_{k=0}^n \frac{\delta^k 0^m}{k!} x^{\underline{k}}$, where $\Delta^k 0^m$ denotes the value of $\delta^k x^m$ at $x = 0$. Since $\delta x^{\underline{k}} = k x^{\underline{k-1}}$, one immediately gets a formula for a polynomial $\phi_{m+1}(x)$ which telescopes x^m in the form

$$\phi_{m+1}(x) = \sum_{k=0}^{\infty} \frac{\Delta^k 0^m}{(k+1)!} x^{\underline{k+1}}$$

This polynomial has the property $\phi_{m+1}(n) = \sum_{k=0}^{n-1} k^m$ for all n.

The polynomials $\phi_m(x)$, as follows from Lemma 4.1.2, are characterized by two conditions:

$$\Delta\phi_m(x) = x^{m-1}, \qquad \phi_m(1) = 0.$$

LEMMA 4.2.1 (on differentiation). $\phi'_{m+1}(x) = \phi'_{m+1}(0) + m\phi_m(x)$.

PROOF. Differentiation of $\Delta\phi_{m+1}(x) = x^m$ gives $\Delta\phi'_{m+1}(x) = mx^{m-1}$. The polynomial $m\phi_m$ has the same differences, hence $\Delta(\phi'_{m+1}(x) - m\phi_m(x)) = 0$. By Lemma 4.1.2 this implies that $\phi'_{m+1}(x) - m\phi_m(x)$ is constant. Therefore, $\phi'_{m+1}(x) -$

$m\phi_m(x) = \phi'_{m+1}(0) - m\phi_m(0)$. But $\phi_m(1) = 0$ and $\phi_m(0) = \phi_m(1) - \delta\phi_m(0) = 0 - 0^{m-1} = 0$. □

Bernoulli polynomials. Let us introduce the m-th *Bernoulli number* B_m as $\phi'_{m+1}(0)$, and define the *Bernoulli polynomial* of degree $m > 0$ as $B_m(x) = m\phi_m(x) + B_m$. The Bernoulli polynomial $B_0(x)$ of degree 0 is defined as identically equal to 1. Consequently $B_m(0) = B_m$ and $B'_{m+1}(0) = (m+1)B_m$.

The Bernoulli polynomials satisfy the following condition:

$$\Delta B_m(x) = mx^{m-1} \quad (m > 0).$$

In particular, $\Delta B_m(0) = 0$ for $m > 1$, and therefore we get the following *boundary conditions* for Bernoulli polynomials:

$$\begin{aligned} B_m(0) = B_m(1) = B_m &\quad \text{for } m > 1, \text{ and} \\ B_1(0) = -B_1(1) = B_1. & \end{aligned}$$

The Bernoulli polynomials, in contrast to $\phi_m(x)$, are *normed*: their leading coefficient is equal to 1 and they have a simpler rule for differentiation:

$$B'_m(x) = mB_{m-1}(x)$$

Indeed, $B'_m(x) = m\phi'_m(x) = m((m-1)\phi_{m-1}(x) + \phi'_m(0)) = mB_{m-1}(x)$, by Lemma 4.2.1.

Differentiating $B_m(x)$ at 0, k times, we get $B_m^{(k)}(0) = m^{\underline{k-1}}B'_{m-k+1}(0) = m^{\underline{k-1}}(m-k+1)B_{m-k} = m^{\underline{k}}B_{m-k}$. Hence the Taylor formula gives the following representation of the Bernoulli polynomial:

$$B_m(x) = \sum_{k=0}^{m} \frac{m^{\underline{k}}B_{m-k}}{k!}x^k.$$

Characterization theorem. The following important property of Bernoulli polynomials will be called the *Balance property*:

$$\int_0^1 B_m(x)\,dx = 0 \quad (m > 0). \tag{4.2.2}$$

Indeed, $\int_0^1 B_m(x)\,dx = \int_0^1 (m+1)B'_{m+1}(x)\,dx = \Delta B_{m+1}(0) = 0$.

The Balance property and the Differentiation rule allow us to evaluate Bernoulli polynomials recursively. Thus, $B_1(x)$ has 1 as leading coefficient and zero integral on $[0,1]$; this allows us to identify $B_1(x)$ with $x - 1/2$. Integration of $B_1(x)$ gives $B_2(x) = x^2 - x + C$, where C is defined by (4.2.2) as $-\int_0^1 x^{\underline{2}}\,dx = \frac{1}{6}$. Integrating $B_2(x)$ we get $B_3(x)$ modulo a constant which we find by (4.2.2) and so on. Thus we obtain the following theorem:

THEOREM 4.2.2 (characterization). *If a sequence of polynomials* $\{P_n(x)\}$ *satisfies the following conditions:*

- $P_0(x) = 1$,
- $\int_0^1 P_n(x)\,dx = 0$ *for* $n > 0$,
- $P'_n(x) = nP_{n-1}(x)$ *for* $n > 0$,

then $P_n(x) = B_n(x)$ *for all* n.

Analytic properties.

LEMMA 4.2.3 (on reflection). *$B_n(x) = (-1)^n B_n(1-x)$ for any n.*

PROOF. We prove that the sequence $T_n(x) = (-1)^n B_n(1-x)$ satisfies all the conditions of Theorem 4.2.2. Indeed, $T_0 = B_0 = 1$,

$$\int_0^1 T_n(x)\,dx = (-1)^n \int_1^0 B_n(x)\,dx = 0$$

and

$$\begin{aligned} T_n(x)' &= (-1)^n B_n'(1-x) \\ &= (-1)^n n B_{n-1}(1-x)(1-x)' \\ &= (-1)^{n+1} n B_{n-1}(x) \\ &= n T_{n-1}(x). \end{aligned}$$

□

LEMMA 4.2.4 (on roots). *For any odd $n > 1$ the polynomial $B_n(x)$ has on $[0,1]$ just three roots: $0, \frac{1}{2}, 1$.*

PROOF. For odd n, the reflection Lemma 4.2.3 implies that $B_n(\frac{1}{2}) = -B_n(\frac{1}{2})$, that is $B_n(\frac{1}{2}) = 0$. Furthermore, for $n > 1$ one has $B_n(1) - B_n(0) = n0^{n-1} = 0$. Hence $B_n(1) = B_n(0)$ for any Bernoulli polynomial of degree $n > 1$. By the reflection formula for an odd n one obtains $B_n(0) = -B_n(1)$. Thus any Bernoulli polynomial of odd degree greater than 1 has roots $0, \frac{1}{2}, 1$.

The proof that there are no more roots is by contradiction. In the opposite case consider $B_n(x)$, of the least odd degree > 1 which has a root α different from the above mentioned numbers. Say $\alpha < \frac{1}{2}$. By Rolle's Theorem 4.1.7 $B_n'(x)$ has at least three roots $\beta_1 < \beta_2 < \beta_3$ in $(0,1)$. To be precise, $\beta_1 \in (0,\alpha)$, $\beta_2 \in (\alpha, \frac{1}{2})$, $\beta_3 \in (\frac{1}{2}, 1)$. Then $B_{n-1}(x)$ has the same roots. By Rolle's Theorem $B_{n-1}'(x)$ has at least two roots in $(0,1)$. Then at least one of them differs from $\frac{1}{2}$ and is a root of $B_{n-2}(x)$. By the minimality of n one concludes $n-2 = 1$. However, $B_1(x)$ has the only root $\frac{1}{2}$. This is a contradiction. □

THEOREM 4.2.5. *$B_n = 0$ for any odd $n > 1$. For $n = 2k$, the sign of B_n is $(-1)^{k+1}$. For any even n one has either $B_n = \max_{x\in[0,1]} B_n(x)$ or $B_n = \min_{x\in[0,1]} B_n(x)$. The first occurs for positive B_n, the second for negative.*

PROOF. $B_{2k+1} = B_{2k+1}(0) = 0$ for $k > 0$ by Lemma 4.2.4. Above we have found that $B_2 = \frac{1}{6}$. Suppose we have established that $B_{2k} > 0$ and that this is the maximal value for $B_{2k}(x)$ on $[0,1]$. Let us prove that $B_{2k+2} < 0$ and it is the minimal value for $B_{2k+2}(x)$ on $[0,1]$. The derivative of B_{2k+1} in this case is positive at the ends of $[0,1]$, hence $B_{2k+1}(x)$ is positive for $0 < x < \frac{1}{2}$ and negative for $\frac{1}{2} < x < 1$, by Lemma 4.2.4 on roots and the Theorem on Intermediate Values. Hence, $B_{2k+2}'(x) > 0$ for $x < \frac{1}{2}$ and $B_{2k+2}'(x) < 0$ for $x > \frac{1}{2}$. Therefore, $B_{2k+2}(x)$ takes the maximal value in the middle of $[0,1]$ and takes the minimal values at the ends of $[0,1]$. Since the integral of the polynomial along $[0,1]$ is zero and the polynomial is not constant, its minimal value has to be negative. The same arguments prove that if B_{2k} is negative and minimal, then B_{2k+2} is positive and maximal. □

LEMMA 4.2.6 (Lagrange Formula). *If f is a differentiable function on $[a,b]$, then there is a $\xi \in (a,b)$, such that*

$$f(b) = f(a) + f'(\xi)\frac{f(b)-f(a)}{b-a}. \tag{4.2.3}$$

PROOF. The function $g(x) = f(x) - (x-a)\frac{f(b)-f(a)}{b-a}$ is differentiable on $[a,b]$ and $g(b) = g(a) = 0$. By Rolle's Theorem $g'(\xi) = 0$ for some $\xi \in [a,b]$. Hence $f'(\xi) = \frac{f(b)-f(a)}{b-a}$. Substitution of this value of $f'(\xi)$ in (4.2.3) gives the equality. □

Generating function. The following function of two variables is called the *generating function of Bernoulli polynomials.*

$$B(x,t) = \sum_{k=0}^{\infty} B_k(x)\frac{t^k}{k!} \tag{4.2.4}$$

Since $B_k \leq \frac{k!}{2^k}$, the series on the right-hand side converges for $t < 2$ for any x. Let us differentiate it termwise as a function of x, for a fixed t. We get $\sum_{k=0}^{\infty} kB_{k-1}(x)\frac{t^k}{k!} = tB(x,t)$. Consequently $(\ln B(x,t))'_x = \frac{B'_x(x,t)}{B(x,t)} = t$ and $\ln B(x,t) = xt + c(t)$, where the constant $c(t)$ depends on t. It follows that $B(x,t) = \exp(xt)k(t)$, where $k(t) = \exp(c(t))$. For $x = 0$ we get $B(0,t) = k(t) = \sum_{k=0}^{\infty} B_k\frac{t^k}{k!}$. To find $k(t)$ consider the difference $B(x+1,t) - B(x,t)$. It is equal to $\exp(xt+t)k(t) - \exp(xt)$. On the other hand the difference is $\sum_{k=0}^{\infty} \Delta B_k(x)\frac{t^k}{k!} = \sum_{k=0}^{\infty} kB_{k-1}(x)\frac{t^k}{k!} = tB(x,t)$. Comparing these expressions we get explicit formulas for the generating functions of Bernoulli numbers:

$$k(t) = \frac{t}{\exp t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!}t^k,$$

and Bernoulli polynomials:

$$B(x,t) = \sum_{k=+}^{0-1} B_k(x)\frac{t^k}{k!} = \frac{t\exp(tx)}{\exp t - 1}.$$

From (4.2.4) one gets $t = (\exp t - 1)\sum_{k=0}^{\infty} B_k\frac{t^k}{k!}$. Substituting $\exp t - 1 = \sum_{k=1}^{\infty} \frac{t^k}{k!}$ in this equality, by the Uniqueness Theorem 3.6.9, one gets the equalities for the coefficients of the power series

$$\sum_{k=1}^{n} \frac{B_{n-k}}{(n-k)!k!} = 0 \quad \text{for } n > 1.$$

Add $\frac{B_n}{n!}$ to both sides of this equality and multiply both sides by $n!$ to get

$$B_n = \sum_{k=0}^{n} \frac{B_k n^{\underline{k}}}{k!} \qquad \text{for } n > 1. \tag{4.2.5}$$

The latter equality one memorizes via the formula $B^n = (B+1)^n$, where after expansion of the right hand side, one should move down all the exponents at B turning the powers of B into Bernoulli numbers.

Now we are ready to prove that

$$\phi_m(1+x) = \frac{B_m(x+1)}{m} - \frac{B_m}{m} = \sum_{k=0}^{m} B_{m-k}\frac{(m-1)^{\underline{k-1}}}{k!}x^k = \psi_m(x). \tag{4.2.6}$$

Putting $x=0$ in the right hand side one gets $\psi_m(0) = B_m(m-1)^{\underline{-1}} = \frac{B_m}{m}$. The left-hand side takes the same value at $x=0$, because $B_m(1) = B_m(0) = B_m$. It remains to prove the equality of the coefficients in (4.2.6) for positive degrees.

$$\begin{aligned}\frac{B_m(x+1)}{m} &= \frac{1}{m}\sum_{k=0}^{m}\frac{m^{\underline{k}}B_{m-k}}{k!}(1+x)^k\\ &= \frac{1}{m}\sum_{k=0}^{m}\frac{m^{\underline{k}}B_{m-k}}{k!}\sum_{j=0}^{k}\frac{k^{\underline{j}}x^j}{j!}\end{aligned}$$

Now let us change the summation order and change the summation index of the interior sum by $i = m-k$.

$$\begin{aligned}&= \frac{1}{m}\sum_{j=0}^{m}\frac{x^j}{j!}\sum_{k=j}^{m}\frac{m^{\underline{k}}B_{m-k}}{k!}k^{\underline{j}}\\ &= \frac{1}{m}\sum_{j=0}^{m}\frac{x^j}{j!}\sum_{i=0}^{m-j}\frac{m^{\underline{m-i}}B_i}{(m-i)!}(m-i)^{\underline{j}}\end{aligned}$$

Now we change $\frac{m^{\underline{m-i}}(m-i)^{\underline{j}}}{(m-i)!}$ by $\frac{(m-j)^{\underline{i}}m^{\underline{j}}}{i!}$ and apply the identity (4.2.5).

$$\begin{aligned}&= \sum_{j=0}^{m}\frac{x^j m^{\underline{j}}}{mj!}\sum_{i=0}^{m-j}\frac{B_i(m-j)^{\underline{i}}}{i!}\\ &= \sum_{j=0}^{m}\frac{(m-1)^{\underline{j-1}}x^j}{j!}B_{m-j}.\end{aligned}$$

Problems.

1. Evaluate $\int_0^1 B_n(x)\sin 2\pi x\,dx$.
2. Expand $x^4 - 3x^2 + 2x - 1$ as a polynomial in $(x-2)$.
3. Calculate the first 20 Bernoulli numbers.
4. Prove the inequality $|B_n(x)| \le |B_n|$ for even n.
5. Prove the inequality $|B_n(x)| \le \frac{n}{4}|B_{n-1}|$ for odd n.
6. Prove that $\frac{f(0)+f(1)}{2} = \int_0^1 f(x)\,dx + \int_0^1 f'(x)B_1(x)\,dx$.
7. Prove that $\frac{f(0)+f(1)}{2} = \int_0^1 f(x)\,dx + \frac{\Delta f'(0)}{2} - \int_0^1 f''(x)B_2(x)\,dx$.
8. Deduce $\Delta B_n(x) = nx^{n-1}$ from the balance property and the differentiation rule.
9. Prove that $B_n(x) = B_n(1-x)$, using the generating function.
10. Prove that $B_{2n+1} = 0$, using the generating function.
11. Prove that $B_m(nx) = n^{m-1}\sum_{k=0}^{n-1} B_m\left(x+\frac{k}{n}\right)$.
12. Evaluate $B_n(\frac{1}{2})$.
13. Prove that $B_{2k}(x) = P(B_2(x))$, where $P(x)$ is a polynomial with positive coefficient (Jacobi Theorem).
14. Prove that $B_n = \sum_{k=0}^{\infty}(-1)^k\frac{\Delta^k 0^n}{k+1}$.
***15.** Prove that $B_m + \sum \frac{1}{k+1}[k+1$ is prime and k is divisor of $m]$ is an integer (Staudt Theorem).

4.3. Euler-Maclaurin Formula

On the contents of the lecture. From this lecture we will learn how Euler managed to calculate eighteen digit places of the sum $\sum_{k=0}^{\infty}\frac{1}{k^2}$.

Symbolic derivation. Taylor expansion of a function f at point x gives

$$f(x+1)=\sum_{k=0}^{\infty}\frac{f^{(k)}(x)}{k!}.$$

Hence

$$\delta f(x)=\sum_{k=1}^{\infty}\frac{\mathbf{D}^k f(x)}{k!},$$

where $\mathbf{D}$ is the operation of differentiation. One expresses this equality symbolically as

$$\delta=\exp\mathbf{D}-\mathbf{1}. \tag{4.3.1}$$

We are searching for F such that $F(n)=\sum_{k=1}^{n-1}f(k)$ for all n. Then $\delta F(x)=f(x)$, or symbolically $F=\delta^{-1}f$. So we have to invert the operation of the difference. From (4.3.1), the inversion is given formally by the formula $(\exp\mathbf{D}-\mathbf{1})^{-1}$. This function has a singularity at 0 and cannot be expanded into a power series in $\mathbf{D}$. However we know the expansion

$$\frac{t}{\exp t-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}t^k.$$

This allows us to give a symbolic solution of our problem in the form

$$\delta^{-1}=\mathbf{D}^{-1}\frac{\mathbf{D}}{\exp\mathbf{D}-\mathbf{1}}=\sum_{k=0}^{\infty}\frac{B_k}{k!}\mathbf{D}^{k-1}=\mathbf{D}^{-1}-\frac{1}{2}\mathbf{1}+\sum_{k=1}^{\infty}\frac{B_{2k}}{2k!}\mathbf{D}^{2k-1}.$$

Here we take into account that $B_0=1$, $B_1=-\frac{1}{2}$ and $B_{2k+1}=0$ for $k>0$. Since $\sum_{k=1}^{n-1}f(k)=F(n)-F(1)$, the latter symbolic formula gives the following summation formula:

$$\sum_{k=1}^{n-1}f(k)=\int_1^n f(x)\,dx-\frac{f(n)-f(1)}{2}+\sum_{k=1}^{\infty}\frac{B_{2k}}{(2k)!}(f^{(2k-1)}(n)-f^{(2k-1)}(1)). \tag{4.3.2}$$

For $f(x)=x^m$ this formula gives the Bernoulli polynomial ϕ_{m+1}.

Euler's estimate. Euler applied this formula to $f(x)=\frac{1}{(x+9)^2}$ and estimated the sum $\sum_{k=10}^{\infty}\frac{1}{k^2}$. In this case the k-th derivative of $\frac{1}{(x+9)^2}$ at 1 has absolute value $\frac{(k+1)!}{10^{k+2}}$. Hence the module of the k-th term of the summation formula does not exceed $\frac{B_k}{k10^{k+2}}$. For an accuracy of eighteen digit places it is sufficient to sum up the first fourteen terms of the series, only eight of them do not vanish. Euler conjectured, and we will prove, that the value of error does not exceed of the value of the first rejected term, which is $\frac{B_{16}}{16\cdot 10^{18}}$. Since $B_{16}=-\frac{3617}{510}$ this gives the promised accuracy.

B_1	B_2	B_4	B_6	B_8	B_{10}	B_{12}	B_{14}	B_{16}	B_{18}	B_{20}
$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{30}$	$\frac{1}{42}$	$-\frac{1}{30}$	$\frac{5}{66}$	$-\frac{691}{2730}$	$\frac{7}{6}$	$-\frac{3617}{510}$	$\frac{43867}{798}$	$-\frac{174611}{330}$

FIGURE 4.3.1. Bernoulli numbers

We see from the table (Figure 4.3.1) that increasing of the number of considered terms does not improve accuracy noticeably.

Summation formula with remainder. In this lecture we assume that all functions under consideration are differentiable as many times as needed.

LEMMA 4.3.1. *For any function $f(x)$ on $[0,1]$ one has*

$$\frac{1}{2}(f(1)+f(0)) = \int_0^1 f(x)\,dx - \int_0^1 f'(x)B_1(x)\,dx.$$

PROOF. Recall that $B_1(x) = x - \frac{1}{2}$, hence $\int_0^1 f'(x)B_1(x)\,dx = \int_0^1 (x-\frac{1}{2})\,df(x)$. Now, integration by parts gives

$$\int_0^1 (x-\frac{1}{2})\,df(x) = \frac{1}{2}(f(1)+f(0)) - \int_0^1 f(x)\,dx.$$

□

Consider the *periodic Bernoulli polynomials* $B_m\{x\} = B_m(x-[x])$. Then $B'_m\{x\} = mB_{m-1}\{x\}$ for non integer x.

Let us denote by $\sum_m^n a_k$ the sum $\frac{1}{2}a_m + \sum_{k=m+1}^{n-1} a_k + \frac{1}{2}a_n$.

LEMMA 4.3.2. *For any natural p, q and any function $f(x)$ one has*

$$\sum_p^q f(k) = \int_p^q f(x)\,dx - \int_p^q f'(x)B_1\{x\}\,dx.$$

PROOF. Applying Lemma 4.3.1 to $f(x+k)$ one gets

$$\begin{aligned}\frac{1}{2}(f(k+1)+f(k)) &= \int_0^1 f(x+k)\,dx + \int_0^1 f'(x+k)B_1(x)\,dx\\ &= \int_k^{k+1} f(x)\,dx + \int_k^{k+1} f'(x)B_1\{x\}\,dx.\end{aligned}$$

Summing up these equalities for k from p to q, one proves the lemma. □

LEMMA 4.3.3. *For $m > 0$ and a function f one has*

$$\int_p^q f(x)B_m\{x\}\,dx = \frac{B_{m+1}}{m+1}(f(q)-f(p)) - \int_p^q f'(x)B_{m+1}\{x\}\,dx. \tag{4.3.3}$$

PROOF. Since $B_m\{x\}dx = d\frac{B_{m+1}\{x\}}{m+1}$ and $B_{m+1}\{k\} = B_{m+1}$ for any natural k, the formula (4.3.3) is obtained by a simple integration by parts. □

THEOREM 4.3.4. *For any function f and natural numbers n and m one has:*

$$\sum_1^n f(k) = \int_1^n f(x)\,dx + \sum_{k=1}^{m-1} \frac{B_{k+1}}{(k+1)!}\Big(f^{(k)}(n) - f^{(k)}(1)\Big) + \frac{(-1)^{m+1}}{m!}\int_1^n f^{(m)}(x)B_m\{x\}\,dx. \tag{4.3.4}$$

PROOF. The proof is by induction on m. For $m = 1$, formula (4.3.4) is just given by Lemma 4.3.2. Suppose (4.3.4) is proved for m. The *remainder*

$$\frac{(-1)^{m+1}}{m!}\int_1^n f^{(m)}(x)B_m\{x\}\,dx$$

can be transformed by virtue of Lemma 4.3.3 into

$$\frac{(-1)^{m+1}B_{m+1}}{(m+1)!}(f^{(m)}(n) - f^{(m)}(1)) + \frac{(-1)^{m+2}}{(m+1)!}\int_1^n B_{m+1}\{x\}f^{(m+1)}(x)\,dx.$$

Since odd Bernoulli numbers vanish, $(-1)^{m+1}B_{m+1} = B_{m+1}$ for $m > 0$. □

Estimation of the remainder. For $m = \infty$, (4.3.4) turns into (4.3.2). Denote

$$R_m = \frac{(-1)^{m+1}}{m!}\int_1^n f^{(m)}(x)B_m\{x\}\,dx.$$

This is the so-called *remainder* of Euler-Maclaurin formula.

LEMMA 4.3.5. *$R_{2m} = R_{2m+1}$ for any $m > 1$.*

PROOF. Because $B_{2m+1} = 0$, the only thing which changes in (4.3.4) when one passes from $2m$ to $2m+1$ is the remainder. Hence its value does not change either. □

LEMMA 4.3.6. *If $f(x)$ is monotone on $[0,1]$ then*

$$\operatorname{sgn}\int_0^1 f(x)B_{2m+1}(x)\,dx = \operatorname{sgn}(f(1) - f(0))\operatorname{sgn} B_{2m}.$$

PROOF. Since $B_{2m+1}(x) = -B_{2m+1}(1-x)$, the change $x \to 1-x$ transforms the integral $\int_{0.5}^1 f(x)B_{2m+1}(x)\,dx$ to $-\int_0^{0.5} f(1-x)B_{2m+1}(x)\,dx$:

$$\begin{aligned}\int_0^1 f(x)B_{2m+1}(x)\,dx &= \int_0^{0.5} f(x)B_{2m+1}(x)\,dx + \int_{0.5}^1 f(x)B_{2m+1}(x)\,dx \\ &= \int_0^{0.5} (f(x) - f(1-x))B_{2m+1}(x)\,dx.\end{aligned}$$

$B_{2m+1}(x)$ is equal to 0 at the end-points of $[0, 0.5]$ and has constant sign on $(0, 0.5)$, hence its sign on the interval coincides with the sign of its derivative at 0, that is, it is equal to $\operatorname{sgn} B_{2m}$. The difference $f(x) - f(1-x)$ also has constant sign as $x < 1-x$ on $(0, 0.5)$ and its sign is $\operatorname{sgn}(f(1) - f(0))$. Hence the integrand has constant sign. Consequently the integral itself has the same sign as the integrand has. □

LEMMA 4.3.7. *If $f^{(2m+1)}(x)$ and $f^{(2m+3)}(x)$ are comonotone for $x \geq 1$ then*

$$R_{2m} = \theta_m \frac{B_{2m+2}}{(2m+2)!}(f^{(2m+1)}(n) - f^{(2m+1)}(1)), \quad 0 \leq \theta_m \leq 1.$$

PROOF. The signs of R_{2m+1} and R_{2m+3} are opposite. Indeed, by Lemma 4.2.5 $\operatorname{sgn} B_{2m} = -\operatorname{sgn} B_{2m+2}$, and $\operatorname{sgn}(f^{(2m+1)}(n) - f^{(2m+1)}(1)) = \operatorname{sgn}(f^{(2m+3)}(n) - f^{(2m+3)}(1)$ due to the comonotonity condition. Hence $\operatorname{sgn} R_{2m+1} = -\operatorname{sgn} R_{2m+3}$ by Lemma 4.3.6.

Set

$$T_{2m+2} = \frac{B_{2m+2}}{(2m+2)!}(f^{(2m+1)}(n) - f^{(2m+1)}(1)).$$

Then $T_{2m+2} = R_{2m+1} - R_{2m+2}$. By Lemma 4.3.5, $T_{2m+2} = R_{2m+1} - R_{2m+3}$. Since R_{2m+3} and R_{2m+1} have opposite signs, it follows that $\operatorname{sgn} T_{2m+2} = \operatorname{sgn} R_{2m+1}$ and $|T_{2m+2}| \geq |R_{2m+1}|$. Hence $\theta_m = \frac{R_{2m+1}}{T_{2m+2}} = \frac{R_{2m}}{T_{2m+2}}$ belongs to $[0,1]$. □

THEOREM 4.3.8. *If $f^{(k)}$ and $f^{(k+2)}$ are comonotone for any $k > 1$, then*

$$\left|\sum_1^n f(k) - \int_1^n f(x)\,dx - \sum_{k=1}^m \frac{B_{2k}}{(2k)!}\left(f^{(2k-1)}(n) - f^{(2k-1)}(1)\right)\right| \leq \left|\frac{B_{2m+2}}{(2m+2)!}\left(f^{(2m+1)}(n) - f^{(2m+1)}(1)\right)\right|.$$

Hence the value of the error which gives the summation formula (4.3.2) with m terms has the same sign as the first rejected term, and its absolute value does not exceed the absolute value of the term.

THEOREM 4.3.9. *Suppose that $\int_1^\infty |f^{(k)}(x)|\,dx < \infty$, $\lim\limits_{x\to\infty} f^{(k)}(x) = 0$ and $f^{(k)}$ is comonotone with $f^{(k+2)}$ for all $k \geq K$ for some K. Then there is a constant C such that for any $m > K$ for some $\theta_m \in [0,1]$*

$$\sum_{k=1}^n f(k) = C + \frac{f(n)}{2} + \int_1^n f(x)\,dx + \sum_{k=1}^m \frac{B_{2k}}{(2k)!}f^{(2k-1)}(n) + \theta_m \frac{B_{2m+2}}{(2m+2)!}f^{(2m+1)}(n). \tag{4.3.5}$$

LEMMA 4.3.10. *Under the condition of the theorem, for any $m \geq K$,*

$$\frac{(-1)^m}{m!}\int_p^\infty f^{(m)}(x)B_m\{x\}\,dx = -\theta_m \frac{B_{2m+2}}{(2m+2)!}f^{(2m+1)}(p). \tag{4.3.6}$$

PROOF. By Lemma 4.3.7,

$$\frac{(-1)^{m+1}}{m!}\int_p^q f^{(m)}(x)B_m\{x\}\,dx = \theta_m \frac{B_{2m+2}}{(2m+2)!}(f^{(2m+1)}(q) - f^{(2m+1)}(p)).$$

To get (4.3.6), pass to the limit as q tends to infinity. □

PROOF OF THEOREM 4.3.9. To get (4.3.5) we change the form of the remainder R_K for (4.3.4). Since

$$\int_1^n B_K\{x\}f^{(K)}\,dx = \int_1^\infty B_K\{x\}f^{(K)}(x)\,dx - \int_n^\infty B_K\{x\}f^{(K)}(x)\,dx,$$

applying the equality (4.3.3) to the interval $[n,\infty)$, one gets

$$-\frac{(-1)^{k+1}B_k}{k!}\int_n^\infty B_k\{x\}f^{(k)}(x)\,dx = \frac{(-1)^{k+1}B_{k+1}}{(k+1)!}f^{(k)}(n) - \frac{(-1)^{k+2}B_{k+1}}{(k+1)!}\int_n^\infty B_{k+1}\{x\}f^{(k+1)}(x)\,dx.$$

Iterating this formula one gets

$$R_K = \int_1^\infty B_K\{x\}f^{(K)}\,dx + \sum_{k=K}^{m}\frac{B_{k+1}}{(k+1)!}f^{(k)}(n) + \frac{(-1)^m}{m!}\int_n^\infty B_m\{x\}f^{(m)}(x)\,dx.$$

Here we take into account the equalities $(-1)^kB_k = B_k$ and $(-1)^{m+2} = (-1)^m$. Now we substitute this expression of R_K into (4.3.4) and set

$$C = (-1)^{K+1}\int_1^\infty B_K\{x\}f^{(K)}(x)\,dx - \frac{f(1)}{2} - \sum_{k=1}^{K-1}\frac{B_{k+1}}{(k+1)!}f^{(k)}(1). \tag{4.3.7}$$

□

Stirling formula. The logarithm satisfies all the conditions of Theorem 4.3.9 with $K = 2$. Its k-th derivative at n is equal to $\frac{(-1)^{k+1}(k-1)!}{n^k}$. Thus (4.3.5) for the logarithm turns into

$$\sum_{k=1}^{n}\ln k = n\ln n - n + \sigma + \frac{\ln n}{2} + \sum_{k=1}^{m}\frac{B_{2k}}{2k(2k-1)n^{2k-1}} + \frac{\theta_m B_{2m+2}}{(2m+2)(2m+1)n^{2k-1}}.$$

By (4.3.7), the constant is

$$\sigma = \int_1^\infty \frac{B_2\{x\}}{x^2}\,dx - \frac{B_2}{2}.$$

But we already know this constant as $\sigma = \frac{1}{2}\ln 2\pi$. For $m = 0$, the above formula gives the most common form of Stirling formula:

$$n! = \sqrt{2\pi n}\,n^n e^{-n+\frac{\Theta}{12n}}.$$

Problems.

1. Write the Euler-Maclaurin series telescoping $\frac{1}{x}$.
2. Prove the uniqueness of the constant in Euler-Maclaurin formula.
3. Calculate ten digit places of $\sum_{k=1}^{\infty}\frac{1}{n^3}$.
4. Calculate eight digit places of $\sum_{k=1}^{1000000}\frac{1}{k}$.
5. Evaluate $\ln 1000!$ with accuracy 10^{-4}.

4.4. Gamma Function

On the contents of the lecture. Euler's Gamma-function is the function responsible for infinite products. An infinite product whose terms are values of a rational function at integers is expressed in terms of the Gamma-function. In particular it will help us prove Euler's factorization of sin.

Telescoping problem. Given a function $f(x)$, find a function $F(x)$ such that $\delta F = f$. This is the *telescoping problem* for functions. In particular, for $f = 0$ any periodic function of period 1 is a solution. In the general case, to any solution of the problem we can add a 1-periodic function and get another solution. The general solution has the form $F(x) + k(t)$ where $F(x)$ is a particular solution and $k(t)$ is a 1-periodic function, called the periodic constant.

The Euler-Maclaurin formula gives a formal solution of the problem, but the Euler-Maclaurin series rarely converges. Another formal solution is

$$F(x) = -\sum_{k=0}^{\infty} f(x+k). \tag{4.4.1}$$

Trigamma. Now let us try to telescope the Euler series. The series (4.4.1) converges for $f(x) = \frac{1}{x^m}$ provided $m \geq 2$ and $x \neq -n$ for natural $n > 1$. In particular, the function

$$\Gamma(x) = \sum_{k=1}^{\infty} \frac{1}{(x+k)^2} \tag{4.4.2}$$

is analytic; it is called the *trigamma* function and it telescopes $-\frac{1}{(1+x)^2}$. Its value $\Gamma(0)$ is just the sum of the Euler series.

This function is distinguished among others functions telescoping $-\frac{1}{(1+x)^2}$ by its finite variation.

Theorem 4.4.1. *There is a unique function $\Gamma(x)$ such that $\delta\Gamma(x) = -\frac{1}{(1+x)^2}$, $\operatorname{var}_\Gamma[0,\infty] < \infty$ and $\Gamma(0) = \sum_{k=1}^{\infty} \frac{1}{k^2}$.*

Proof. Since Γ is monotone, one has $\operatorname{var}_\Gamma[0,\infty] = \sum_{k=0}^{\infty} |\delta\Gamma| = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$. Suppose $f(x)$ is another function of finite variation telescoping $\frac{1}{(1+x)^2}$. Then $f(x) - \Gamma(x)$ is a periodic function of finite variation. It is obvious that such a function is constant, and this constant is 0 if $f(1) = \Gamma(1)$. □

Digamma. The series $-\sum_{k=0}^{\infty} \frac{1}{x+k}$, which formally telescopes $\frac{1}{x}$, is divergent. However the series $-\sum_{k=0}^{\infty} \left(\frac{1}{x+k} - \frac{1}{k}[k \neq 0]\right)$ is convergent and it telescopes $\frac{1}{x}$, because adding a constant does not affect the differences. Indeed,

$$-\sum_{k=0}^{\infty} \left(\tfrac{1}{x+1+k} - \tfrac{1}{k}[k \neq 0]\right) + \sum_{k=0}^{\infty} \left(\tfrac{1}{x+k} - \tfrac{1}{k}[k \neq 0]\right) = -\sum_{k=0}^{\infty} \delta \tfrac{1}{x+k} = \tfrac{1}{x}.$$

The function

$$\digamma(x) = -\gamma + \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{x+k}\right) \tag{4.4.3}$$

is called the*digamma* function. Here γ is the Euler constant. The digamma function is an analytic function, whose derivative is the trigamma function, and whose difference is $\frac{1}{1+x}$.

Monotonicity distinguishes $\digamma$ among others function telescoping $\frac{1}{1+x}$.

THEOREM 4.4.2. *There is a unique monotone function $\digamma(x)$ such that $\delta\digamma(x) = \frac{1}{1+x}$ and $\digamma(0) = -\gamma$.*

PROOF. Suppose $f(x)$ is a monotone function telescoping $\frac{1}{1+x}$. Denote by v the variation of $f-\digamma$ on $[0,1]$. Then the variation of $f-\digamma$ over $[1,n]$ is nv. On the other hand, $\operatorname{var}_f[1,n] = \sum_{k=1}^{n} \frac{1}{k} < \ln n + \gamma$. Hence the variation of $f(x) - \digamma(x)$ on $[1,n]$ is less than $2(\gamma + \ln n)$. Hence v for any n satisfies the inequality $nv \leq 2(\gamma + \ln n)$. Since $\lim_{n\to\infty} \frac{\ln n}{n} = 0$, we get $v = 0$. Hence $f - \digamma$ is constant, and it is zero if $f(1) = \digamma(1)$. □

LEMMA 4.4.3. *$\digamma' = \Gamma$.*

PROOF. To prove that $\digamma'(x) = \Gamma(x)$, consider $F(x) = \int_1^x \Gamma(t)\,dt$. This function is monotone, because $F'(x) = \Gamma(x) \geq 0$. Further $(\delta F)' = \delta F' = \delta\Gamma(x) = -\frac{1}{(1+x)^2}$. It follows that $\delta F = \frac{1}{1+x} + c$, where c is a constant. By Theorem 4.4.2 it follows that $F(x+1) - cx - \gamma = \digamma(x)$. Hence $\digamma(x)' = F'(x+1) + c = \Gamma(x)$. This proves that $\digamma'$ is differentiable and has finite variation. As $\delta\digamma(x) = \frac{1}{1+x}$ it follows that $\delta\digamma'(x) = -\frac{1}{(1+x)^2}$. We get that $\digamma'(x) = \Gamma(x)$ by Theorem 4.4.1. □

Telescoping the logarithm. To telescope the logarithm, we start with the formal solution $-\sum_{k=0}^{\infty} \ln(x+k)$. To decrease the divergence, add $\sum_{k=1}^{\infty} \ln k$ termwise. We get $-\ln x - \sum_{k=1}^{\infty}(\ln(x+k) - \ln k) = -\ln x - \sum_{k=1}^{\infty} \ln(1+\frac{x}{k})$. We know that $\ln(1+x)$ is close to x, but the series still diverges. Now convergence can be reached by the subtraction of $\frac{x}{k}$ from the k-th term of the series. This substraction changes the difference. Let us evaluate the difference of $F(x) = -\ln x - \sum_{k=1}^{\infty}(\ln(1+\frac{x}{k}) - \frac{x}{k})$. The difference of the n-th term of the series is

$$\begin{aligned}
&\left(\ln\left(1+\tfrac{x+1}{k}\right) - \tfrac{x+1}{k}\right) - \left(\ln\left(1+\tfrac{x}{k}\right) - \tfrac{x}{k}\right)\\
&\quad= \left(\ln(x+k+1) - \ln k - \tfrac{x+1}{k}\right) - \left(\ln(x+k) - \ln k - \tfrac{x}{k}\right)\\
&\quad= \delta\ln(x+k) - \tfrac{1}{k}.
\end{aligned}$$

Hence

$$\begin{aligned}
\delta F(x) &= -\delta\ln x - \textstyle\sum_{k=1}^{\infty}\left(\delta\ln(x+k) - \tfrac{1}{k}\right)\\
&= \lim_{n\to\infty}\left(-\delta\ln x - \textstyle\sum_{k=1}^{n-1}\left(\delta\ln(x+k) - \tfrac{1}{k}\right)\right)\\
&= \lim_{n\to\infty}\left(\ln x - \ln(n+x) + \textstyle\sum_{k=1}^{n-1}\tfrac{1}{k}\right)\\
&= \ln x + \lim_{n\to\infty}(\ln(n) - \ln(n+x)) + \lim_{n\to\infty}\left(\textstyle\sum_{k=1}^{n-1}\tfrac{1}{k} - \ln n\right)\\
&= \ln x + \gamma.
\end{aligned}$$

As a result, we get the following formula for a function, which telescopes the logarithm:

$$\Theta(x) = -\gamma x - \ln x - \sum_{k=1}^{\infty}\left(\ln\left(1+\frac{x}{k}\right) - \frac{x}{k}\right). \tag{4.4.4}$$

THEOREM 4.4.4. *The series* (4.4.4) *converges absolutely for all* x *except negative integers. It presents a function* $\Theta(x)$ *such that* $\Theta(1) = 0$ *and* $\delta\Theta(x) = \ln x$.

PROOF. The inequality $\frac{x}{1+x} \le \ln(1+x) \le x$ implies

$$|\ln(1+x) - x| \le \left|\frac{x}{1+x} - x\right| = \left|\frac{x^2}{1+x}\right|. \tag{4.4.5}$$

Denote by ε the distance from x to the closest negative integer. Then due to (4.4.5), the series $\sum_{k=1}^{\infty} \ln\left(\left(1+\frac{y}{k}\right) - \frac{y}{k}\right)$ is termwise majorized by the convergent series $\sum_{k=1}^{\infty} \frac{x^2}{\varepsilon k^2}$. This proves the absolute convergence of (4.4.4).

Since $\lim_{n\to\infty} \sum_{k=1}^{n-1} (\ln(1+\frac{1}{k}) - \frac{1}{k}) = \lim_{n\to\infty}(\ln n - \sum_{k=1}^{n-1}\frac{1}{k}) = -\gamma$, one gets $\Theta(1) = 0$. □

Convexity. There are a lot of functions that telescope the logarithm. The property which distinguishes Θ among others is convexity.

Throughout the lecture θ and $\overline{\theta}$ are nonnegative and *complementary* to each other, that is $\theta + \overline{\theta} = 1$. The function f is called *convex* if, for any x, y, it satisfies the inequality:

$$f(\theta x + \overline{\theta} y) \le \theta f(x) + \overline{\theta} f(y) \quad \forall \theta \in [0,1]. \tag{4.4.6}$$

Immediately from the definition it follows that

LEMMA 4.4.5. *Any linear function* $ax + b$ *is convex.*

LEMMA 4.4.6. *Any sum (even infinite) of convex functions is a convex function. The product of a convex function by a positive constant is a convex function.*

LEMMA 4.4.7. *If* $f(p) = f(q) = 0$ *and* f *is convex, then* $f(x) \ge 0$ *for all* $x \notin [p,q]$.

PROOF. If $x > q$ then $q = x\theta + p\overline{\theta}$ for $\theta = \frac{q-p}{x-p}$. Hence $f(q) \le f(x)\theta + f(p)\overline{\theta} = f(x)$, and it follows that $f(x) \ge f(q) = 0$. For $x < p$ one has $p = x\theta + q\overline{\theta}$ for $\theta = \frac{q-p}{q-x}$. Hence $0 = f(p) \le f(x)\theta + f(q)\overline{\theta} = f(x)$. □

LEMMA 4.4.8. *If* f'' *is nonnegative then* f *is convex.*

PROOF. Consider the function $F(t) = f(l(t))$, where $l(t) = x\overline{\theta} + y\theta$. Newton's formula for $F(t)$ with nodes 0, 1 gives $F(t) = F(0) + \delta F(0)t + \frac{1}{2}F''(\xi)t^{\underline{2}}$. Since $F''(\xi) = (y-x)^2 f''(\xi) > 0$, and $t^{\underline{2}} = t(t-1) < 0$ we get the inequality $F(t) \le F(0)\overline{\theta} + tF(1)$. Since $F(\theta) = f(x\overline{\theta} + y\theta)$ this is just the inequality of convexity. □

LEMMA 4.4.9. *If* f *is convex, then* $0 \le f(a) + \theta\delta f(a) - f(a+\theta) \le \delta^2 f(a-1)$ *for any* a *and any* $\theta \in [0,1]$

PROOF. Since $a + \theta = \overline{\theta}a + \theta(a+1)$ we get $f(a+\theta) \le f(a)\overline{\theta} + f(a+1)\theta = f(a) + \theta\delta f(a)$. On the other hand, the convex function $f(a+x) - f(a) - x\delta f(a-1)$ has roots -1 and 0. By Lemma 4.4.7 it is nonnegative for $x > 0$. Hence $f(a+\theta) \ge f(a) + \theta\delta f(a-1)$. It follows that $f(a) + \theta\delta f(a) - f(a+\theta) \ge f(a) + \theta\delta f(a) - f(a) - \theta\delta f(a-1) = \theta\delta^2 f(a-1)$. □

THEOREM 4.4.10. $\Theta(x)$ *is the unique convex function that telescopes* $\ln x$ *and satisfies* $\Theta(1) = 1$.

PROOF. Convexity of Θ follows from the convexity of the summands of its series. The summands are convex because their second derivatives are nonnegative.

Suppose there is another convex function $f(x)$ which telescopes the logarithm too. Then $\phi(x) = f(x) - \Theta(x)$ is a periodic function, $\delta\phi = 0$. Let us prove that $\phi(x)$ is convex. Consider a pair c, d, such that $|c-d| \leq 1$. Since $f(c\theta + d\overline{\theta}) - \theta f(c) - \overline{\theta} f(d) \leq 0$, as f is convex, one has

$$\begin{aligned}\phi(c\theta + d\overline{\theta}) - \theta\phi(c) - \overline{\theta}\phi(d) &= (f(c\theta + d\overline{\theta}) - \theta f(c) - \overline{\theta} f(d)) \\ &\quad - (\Theta(c\theta + d\overline{\theta}) - \theta\Theta(c) - \overline{\theta}\Theta(d)) \\ &\leq \theta\Theta(c) + \overline{\theta}\Theta(d) - \Theta(c\theta + d\overline{\theta}).\end{aligned}$$

First, prove that ϕ satisfies the following ε-relaxed inequality of convexity:

$$\phi(c\theta + d\overline{\theta}) \leq \theta\phi(c) + \overline{\theta}\phi(d) + \varepsilon. \tag{4.4.7}$$

Increasing c and d by 1, we do not change the inequality as $\delta\phi = 0$. Due to this fact, we can increase c and d to satisfy $\frac{1}{c-1} < \frac{\varepsilon}{3}$. Set $L(x) = \Theta(c) + (x-c)\ln c$. By Lemma 4.4.9 for $x \in [c, c+1]$ one has $|\Theta x - L(x)| \leq \delta^2\Theta(c-1) = \ln c - \ln(c-1) = \ln(1 + \frac{1}{c-1}) \leq \frac{1}{c-1} < \frac{\varepsilon}{3}$. Since $|\Theta(x) - L(x)| < \frac{\varepsilon}{3}$ for $x = c, d, \frac{c+d}{2}$, it follows that $\theta\Theta(c) + \overline{\theta}\Theta(d) - \Theta(c\theta + d\overline{\theta})$ differs from $\theta L(c) + \overline{\theta} L(d) - L(c\theta + d\overline{\theta}) = 0$ by less than by ε. The inequality (4.4.7) is proved. Passing to the limit as ε tends to 0, one eliminates ε.

Hence $\phi(x)$ is convex on any interval of length 1 and has period 1. Then $\phi(x)$ is constant. Indeed, consider a pair a, b with condition $b - 1 < a < b$. Then $a = (b-1)\theta + b\overline{\theta}$ for $\theta = b - a$. Hence $f(a) \leq f(b)\theta + f(b-1)\overline{\theta} = f(b)$. □

LEMMA 4.4.11. $\Theta''(1+x) = \mathit{\Gamma}(x)$.

PROOF. The function $F(x) = \int_1^x \mathit{\Gamma}(t)\,dt$ is convex because its second derivative is $\mathit{\Gamma}$. The difference of $F' = \mathit{\Gamma}$ is $\frac{1}{1+x}$. Hence $\delta F(x) = \ln(x+1) + c$, where c is some constant. It follows that $F(x-1) - cx + c = \Theta(x)$. Hence Θ is twice differentiable and its second derivative is $\mathit{\Gamma}$. □

Gamma function. Now we define Euler's *gamma function* $\Gamma(x)$ as $\exp(\Theta(x))$, where $\Theta(x)$ is the function telescoping the logarithm. Exponentiating (4.4.4) gives a representation of the Gamma function in so-called *canonical Weierstrass form*:

$$\Gamma(x) = \frac{e^{-\gamma x}}{x}\prod_{k=1}^{\infty}\left(1 + \frac{x}{k}\right)^{-1} e^{\frac{x}{k}}. \tag{4.4.8}$$

Since $\delta \ln \Gamma(x) = \ln x$, one gets the following *characteristic equation* of the Gamma function

$$\Gamma(x+1) = x\Gamma(x). \tag{4.4.9}$$

Since $\Theta(1) = 0$, according to (4.4.4), one proves by induction that $\Gamma(n) = (n-1)!$ using (4.4.9).

A nonnegative function f is called *logarithmically convex* if $\ln f(x)$ is convex.

THEOREM 4.4.12 (characterization). $\Gamma(x)$ *is the unique logarithmically convex function defined for* $x > 0$, *which satisfies equation* (4.4.9) *for all* $x > 0$ *and takes the value* 1 *at* 1.

PROOF. Logarithmical convexity of $\Gamma(x)$ follows from the convexity of $\Theta(x)$. Further $\Gamma(1) = \exp \Theta(1) = 1$. If f is a logarithmically convex function satisfying the gamma-equation, then $\ln f$ satisfies all the conditions of Theorem 4.4.4. Hence, $\ln f(x) = \Theta(x)$ and $f(x) = \Gamma(x)$. $\square$

THEOREM 4.4.13 (Euler). *For any $x \geq 0$ one has $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$.*

Let us check that the integral satisfies all the conditions of Theorem 4.4.12. For $x = 1$ the integral gives $\int_0^\infty e^{-t}\,dt = -e^{-t}\,|_0^\infty = 1$. The integration by parts $\int_0^\infty t^x e^{-t}\,dt = -\int_0^\infty t^x\,de^{-t} = -t^x e^{-t}\,|_0^\infty + \int_0^\infty e^{-t}xt^{x-1}\,dx$ proves that it satisfies the gamma-equation (4.4.9). It remains to prove logarithmic convexity of the integral.

LEMMA 4.4.14 (mean criterium). *If f is a monotone function which satisfies the following* mean inequality $2f(\frac{x+y}{2}) \leq f(x) + f(y)$ *for all x, y then f is convex.*

PROOF. We have to prove the inequality $f(x\theta + y\overline{\theta}) \leq \theta f(x) + \overline{\theta} f(y) = L(\overline{\theta})$ for all θ, x and y. Set $F(t) = f(x + (y-x)t)$; than F also satisfies the mean inequality. And to prove our lemma it is sufficient to prove that $F(t) \leq L(t)$ for all $t \in [0,1]$.

First we prove this inequality only for all *binary rational* numbers t, that is for numbers of the type $\frac{m}{2^n}$, $m \leq 2^n$. The proof is by induction on n, the degree of the denominator. If $n = 0$, the statement is true. Suppose the inequality $F(t) \leq L(t)$ is already proved for fractions with denominators of degree $\leq n$. Consider $r = \frac{m}{2^{n+1}}$, with odd $m = 2k+1$. Set $r^- = \frac{k}{2^n}$, $r^+ = \frac{k+1}{2^n}$. By the induction hypothesis $F(r^\pm) \leq L(r^\pm)$. Since $r = \frac{r^+ + r^-}{2}$, by the mean inequality one has $F(r) \leq \frac{f(r^+)+f(r^-)}{2} \leq \frac{L(r^+)+L(r^-)}{2} = L(\frac{r^+ + r^-}{2}) = L(r)$.

Thus our inequality is proved for all binary rational t. Suppose $F(t) > L(t)$ for some t. Consider two binary rational numbers p, q such that $t \in [p, q]$ and $|q - p| < \frac{F(t)-L(t)}{|f(y)-f(x)|}$. In this case $|L(p) - L(t)| \leq |p-t||f(y) - f(x)| < |F(t) - L(t)|$. Therefore $F(p) \leq L(p) < F(t)$. The same arguments give $F(q) < F(t)$. This is a contradiction, because t is between p and q and its image under a monotone mapping has to be between images of p and q. $\square$

LEMMA 4.4.15 (Cauchy-Bunyakovski-Schwarz).

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \leq \int_a^b f^2(x)\,dx \int_a^b g^2(x)\,dx. \tag{4.4.10}$$

PROOF. Since $\int_a^b (f(x) + tg(x))^2\,dx \geq 0$ for all t, the discriminant of the following quadratic equation is non-negative:

$$t^2 \int_a^b g^2(x)\,dx + 2t \int_a^b f(x)g(x)\,dx + \int_a^b f^2(x)\,dx = 0. \tag{4.4.11}$$

This discriminant is $4\left(\int_a^b f(x)g(x)\,dx\right)^2 - 4\int_a^b f^2(x)\,dx \int_a^b g^2(x)\,dx$. $\square$

Now we are ready to prove the logarithmic convexity of the Euler integral. The integral is obviously an increasing function, hence by the mean criterion it is sufficient to prove the following inequality:

$$\left(\int_0^\infty t^{\frac{x+y}{2}-1}e^{-t}\,dt\right)^2 \leq \int_0^\infty t^{x-1}e^{-t}\,dt \int_0^\infty t^{y-1}e^{-t}\,dt. \tag{4.4.12}$$

This inequality turns into the Cauchy-Bunyakovski-Schwarz inequality (4.4.10) for $f(x) = t^{\frac{x-1}{2}} e^{-t/2}$ and $g(t) = t^{\frac{y-1}{2}} e^{-t/2}$.

Evaluation of products. From the canonical Weierstrass form it follows that

$$\prod_{n=1}^{\infty} \{(1 - x/n)\exp(x/n)\} = \frac{-e^{\gamma x}}{x\Gamma(-x)}, \tag{4.4.13}$$

$$\prod_{n=1}^{\infty} \{(1 + x/n)\exp(-x/n)\} = \frac{e^{-\gamma x}}{x\Gamma(x)}.$$

One can evaluate a lot of products by splitting them into parts which have this canonical form (4.4.13). For example, consider the product $\prod_{k=1}^{\infty}\left(1 - \frac{x^2}{k^2}\right)$. Division by n^2 transforms it into $\prod_{k=1}^{\infty}(1 - \frac{1}{2n})^{-1}(1 + \frac{1}{2n})^{-1}$. Introducing multipliers $e^{\frac{1}{2n}}$ and $e^{-\frac{1}{2n}}$, one gets a canonical form

$$\prod_{n=1}^{\infty}\left\{\left(1 - \frac{1}{2n}\right)e^{\frac{1}{2n}}\right\}^{-1} \prod_{n=1}^{\infty}\left\{\left(1 + \frac{1}{2n}\right)e^{-\frac{1}{2n}}\right\}^{-1}. \tag{4.4.14}$$

Now we can apply (4.4.13) for $x = \frac{1}{2}$. The first product of (4.4.14) is equal to $-\frac{1}{2}\Gamma(-1/2)e^{-\gamma/2}$, and the second one is $\frac{1}{2}\Gamma(1/2)e^{\gamma/2}$. Since according to the characteristic equation for Γ-function, $\Gamma(1/2) = -\frac{1}{2}\Gamma(1/2)$, one gets $\Gamma(1/2)^2/2$ as the value of Wallis product. Since the Wallis product is $\frac{\pi}{2}$, we get $\Gamma(1/2) = \sqrt{\pi}$.

Problems.

1. Evaluate the product $\prod_{n=1}^{\infty}\left(1 + \frac{x}{n}\right)\left(1 + \frac{2x}{n}\right)\left(1 - \frac{3x}{n}\right)$.
2. Evaluate the product $\prod_{k=1}^{\infty}\frac{k(5+k)}{(3+k)(2+k)}$.
3. Prove: The sum of logarithmically convex functions is logarithmically convex.
4. Prove $\Gamma(x) = \lim_{n\to\infty}\frac{n!n^x x^{-n}}{x}$.
5. Prove $\prod_{k=1}^{\infty}\frac{k}{x+k}\left(\frac{k+1}{k}\right)^x = \Gamma(x+1)$.
6. Prove Legendre's doubling formula $\Gamma(2x)\Gamma(0.5) = 2^{2x-1}\Gamma(x+0.5)\Gamma(x)$.

4.5. The Cotangent

On the contents of the lecture. In this lecture we perform what was promised at the beginning: we sum up the Euler series and expand $\sin x$ into the product. We will see that sums of series of reciprocal powers are expressed via Bernoulli numbers. And we will see that the function responsible for the summation of the series is the cotangent.

An ingenious idea, which led Euler to finding the sum $\sum_{k=1}^{\infty}\frac{1}{k^2}$, is the following. One can consider $\sin x$ as a polynomial of infinite degree. This polynomial has as roots all points of the type $k\pi$. Any ordinary polynomial can be expanded into a product $\prod(x-x_k)$ where x_k are its roots. By analogy, Euler conjectured that $\sin x$ can be expanded into the product

$$\sin x = \prod_{k=-\infty}^{\infty}(x-k\pi).$$

This product diverges, but can be modified to a convergent one by division of the n-th term by $-n\pi$. The division does not change the roots. The modified product is

$$\prod_{k=-\infty}^{\infty}\left(1-\frac{x}{k\pi}\right) = x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2\pi^2}\right). \tag{4.5.1}$$

Two polynomials with the same roots can differ by a multiplicative constant. To find the constant, consider $x=\frac{\pi}{2}$. In this case we get the inverse to the Wallis product in (4.5.1) multiplied by $x=\frac{\pi}{2}$. Hence the value of (4.5.1) is 1, which coincides with $\sin\frac{\pi}{2}$. Thus it is natural to expect that $\sin x$ coincides with the product (4.5.1).

There is another way to tame $\prod_{k=-\infty}^{\infty}(x-k\pi)$. Taking the logarithm, we get a divergent series $\sum_{k=-\infty}^{\infty}\ln(x-k\pi)$, but achieve convergence by termwise differentiation. Since the derivative of $\ln\sin x$ is $\cot x$, it is natural to expect that $\cot x$ coincides with the following function

$$\operatorname{ctg}(x) = \sum_{k=-\infty}^{\infty}\frac{1}{x-k\pi} = \frac{1}{x}+\sum_{k=1}^{\infty}\frac{2x}{x^2-k^2\pi^2}. \tag{4.5.2}$$

Cotangent expansion. The expansion $\frac{z}{e^z-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}z^k$ allows us to get a power expansion for $\cot z$. Indeed, representing $\cot z$ by Euler's formula one gets

$$i\frac{e^{iz}+e^{-iz}}{e^{iz}-e^{-iz}} = i\frac{e^{2iz}+1}{e^{2iz}-1} = i+\frac{2i}{e^{2iz}-1} = i+\frac{1}{z}\frac{2iz}{e^{2iz}-1} = i+\frac{1}{z}\sum_{k=0}^{\infty}\frac{B_k}{k!}(2iz)^k.$$

The term of the last series corresponding to $k=1$ is $2izB_1=-iz$. Multiplied by $\frac{1}{z}$, it turns into $-i$, which eliminates the first i. The summand corresponding to $k=0$ is 1. Taking into account that $B_{2k+1}=0$ for $k>0$, we get

$$\cot z = \frac{1}{z}+\sum_{k=1}^{\infty}(-1)^k\frac{4^kB_{2k}}{(2k)!}z^{2k-1}.$$

Power expansion of $\operatorname{ctg}(z)$**.** Substituting

$$\frac{1}{z^2-n^2\pi^2} = -\frac{1}{n^2\pi^2}\frac{1}{1-\frac{z^2}{n^2\pi^2}} = -\sum_{k=0}^{\infty}\frac{z^{2k}}{(n\pi)^{2k+2}}$$

into (4.5.2) and changing the order of summation, one gets:

$$\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\frac{z^{2k}}{(n\pi)^{2k+2}} = \sum_{k=0}^{\infty}\frac{z^{2k}}{\pi^{2k+2}}\sum_{n=1}^{\infty}\frac{1}{n^{2k+2}}.$$

The change of summation order is legitimate in the disk $|z|<1$, because the series absolutely converges there. This proves the following:

LEMMA 4.5.1. $\operatorname{ctg}(z)-\frac{1}{z}$ *is an analytic function in the disk* $|z|<1$*. The* n*-th coefficient of the Taylor series of* $\operatorname{ctg}(z)-\frac{1}{z}$ *at* 0 *is equal to* 0 *for even* n *and is equal to* $\frac{1}{\pi^{n+1}}\sum_{k=1}^{\infty}\frac{1}{k^{n+1}}$ *for any odd* n*.*

Thus the equality $\cot z = \operatorname{ctg}(z)$ would imply the following remarkable equality:

$$\boxed{(-1)^n\frac{4^n B_{2n}}{2n!} = -\frac{1}{\pi^{2n}}\sum_{k=1}^{\infty}\frac{1}{k^{2n}}}$$

In particular, for $n=1$ it gives the sum of Euler series as $\frac{\pi^2}{6}$.

Exploring the cotangent.

LEMMA 4.5.2. $|\cot z|\le 2$ *provided* $|\operatorname{Im} z|\ge 1$*.*

PROOF. Set $z=x+iy$. Then $|e^{iz}|=|e^{ix-y}|=e^{-y}$. Therefore if $y\ge 1$, then $|e^{2iz}|=e^{-2y}\le\frac{1}{e^2}<\frac{1}{3}$. Hence $|e^{2iz}+1|\le\frac{1}{e^2}+1<\frac{4}{3}$ and $|e^{2iz}-1|\ge 1-\frac{1}{e^2}>\frac{2}{3}$. Thus the absolute value of

$$\cot z = i\frac{e^{iz}+e^{-iz}}{e^{iz}-e^{-iz}} = i\frac{e^{2iz}+1}{e^{2iz}-1}$$

is less than 2. For $y\ge 1$ the same arguments work for the representation of $\cot z$ as $i\frac{1+e^{-2iz}}{1-e^{-2iz}}$. □

LEMMA 4.5.3. $|\cot(\pi/2+iy)|\le 4$ *for all* y*.*

PROOF. $\cot(\pi/2+iy) = \frac{\cos(\pi/2+iy)}{\sin(\pi/2+iy)} = \frac{-\sin iy}{\cos iy} = \frac{e^t-e^{-t}}{e^t+e^{-t}}$. The module of the numerator of this fraction does not exceed $e-e^{-1}$ for $t\in[-1,1]$ and the denominator is greater than 1. This proves the inequality for $y\in[-1,1]$. For other y this is the previous lemma. □

Let us denote by $\pi\mathbb{Z}$ the set $\{k\pi \mid k\in\mathbb{Z}\}$ of π*-integers.*

LEMMA 4.5.4. *The set of singular points of* $\cot z$ *is* $\pi\mathbb{Z}$*. All these points are simple poles with residue* 1*.*

PROOF. The singular points of $\cot z$ coincide with the roots of $\sin z$. The roots of $\sin z$ are roots of the equation $e^{iz}=e^{-iz}$ which is equivalent to $e^{2iz}=1$. Since $|e^{2iz}|=|e^{-2\operatorname{Im} z}|$ one gets $\operatorname{Im} z=0$. Hence $\sin z$ has no roots beyond the real line. And all its real roots as we know have the form $\{k\pi\}$. Since $\lim_{z\to 0} z\cot z = \lim_{z\to 0}\frac{z\cos z}{\sin z} = \lim_{z\to 0}\frac{z}{\sin z} = \frac{1}{\sin' 0} = 1$, we get that 0 is a simple pole of $\cot z$

with residue 1 and the other poles have the same residue because of periodicity of $\cot z$. □

LEMMA 4.5.5. *Let $f(z)$ be an analytic function on a domain D. Suppose that f has in D finitely many singular points, they are not π-integers and D has no π-integer point on its boundary. Then*

$$\oint_{\partial D} f(\zeta)\cot\zeta d\zeta = 2\pi i \sum_{k=-\infty}^{\infty} f(k\pi)[k\pi \in D] + 2\pi i \sum_{z\in D} \operatorname{res}_z(f(z)\cot z)[z \notin \pi\mathbb{Z}].$$

PROOF. In our situation every singular point of $f(z)\cot z$ in D is either a π-integer or a singular point of $f(z)$. Since $\operatorname{res}_{z=k\pi}\cot z = 1$, it follows that $\operatorname{res}_{z=k\pi} f(z)\cot z = f(k\pi)$. Hence the conclusion of the lemma is a direct consequence of Residue Theory. □

Exploring $\operatorname{ctg}(z)$**.**

LEMMA 4.5.6. $\operatorname{ctg}(z+\pi) = \operatorname{ctg}(z)$ *for any z.*

PROOF.

$$\begin{aligned}
\operatorname{ctg}(z+\pi) &= \lim_{n\to\infty}\sum_{k=-n}^{n}\frac{1}{z+\pi-k\pi}\\
&= \lim_{n\to\infty}\sum_{k=-n-1}^{n-1}\frac{1}{z+k\pi}\\
&= \lim_{n\to\infty}\frac{1}{z-(n+1)\pi} + \lim_{n\to\infty}\frac{1}{z-n\pi} + \lim_{n\to\infty}\sum_{k=-(n-1)}^{(n-1)}\frac{1}{z+\pi-k\pi}\\
&= 0+0+\operatorname{ctg}(z).
\end{aligned}$$

□

LEMMA 4.5.7. *The series representing* $\operatorname{ctg}(z)$ *converges for any z which is not a π-integer.* $|\operatorname{ctg}(z)| \le 2$ *for all z such that* $|\operatorname{Im} z| > \pi$.

PROOF. For any z one has $|z^2 - k^2\pi^2| \ge k^2$ for $k > |z|$. This provides the convergence of the series. Since $\operatorname{ctg}(z)$ has period π, it is sufficient to prove the inequality of the lemma in the case $x \in [0,\pi]$, where $z = x+iy$. In this case $|y| \ge |x|$ and $\operatorname{Re} z^2 = x^2 - y^2 \le 0$. Then $\operatorname{Re}(z^2 - k^2\pi^2) \le -k^2\pi^2$. It follows that $|z^2 - k^2\pi^2| \ge k^2\pi^2$. Hence $|\operatorname{ctg}(z)|$ is termwise majorized by $\frac{1}{\pi} + \sum_{k=1}^{\infty}\frac{1}{k^2\pi^2} < 2$. □

LEMMA 4.5.8. $|\operatorname{ctg}(z)| \le 3$ *for any z with* $\operatorname{Re} z = \frac{\pi}{2}$.

PROOF. In this case $\operatorname{Re}(z^2 - k^2\pi^2) = \frac{\pi^2}{4} - y^2 - k^2\pi^2 \le -k^2$ for all $k \ge 1$. Hence $|C(z)| \le \frac{2}{\pi} + \sum_{k=1}^{\infty}\frac{1}{k^2} \le 1 + 2 = 3$. □

LEMMA 4.5.9. *For any $z \ne k\pi$ and domain D which contains z and whose boundary does not contain π-integers, one has*

$$\oint_{\partial D}\frac{\operatorname{ctg}(\zeta)}{\zeta - z}d\zeta = 2\pi i\operatorname{ctg}(z) + 2\pi i\sum_{k=-\infty}^{\infty}\frac{1}{k\pi - z}[k\pi \in D]. \tag{4.5.3}$$

PROOF. As was proved in Lecture 3.6, the series $\sum_{k=-\infty}^{\infty} \frac{1}{(\zeta-z)(\zeta-k\pi)}$ admits termwise integration. The residues of $\frac{1}{(\zeta-z)(\zeta-k\pi)}$ are $\frac{1}{k\pi-z}$ at $k\pi$ and $\frac{1}{z-k\pi}$ at z. Hence

$$\oint_{\partial D} \frac{1}{(\zeta-z)(\zeta-k\pi)} d\zeta = \begin{cases} 2\pi i \frac{1}{z-k\pi} & \text{for } k\pi \notin D, \\ 0 & \text{if } k\pi \in D. \end{cases}$$

It follows that

$$\begin{aligned} \oint_{\partial D} \frac{\operatorname{ctg}(\zeta)}{\zeta - z} d\zeta &= 2\pi i \sum_{k=-\infty}^{\infty} \frac{1}{z-k\pi}[k\pi \notin D] \\ &= 2\pi i \operatorname{ctg}(z) - \sum_{k=-\infty}^{\infty} \frac{1}{z-k\pi}[k\pi \in D]. \end{aligned}$$

□

LEMMA 4.5.10. $\operatorname{ctg}(z)$ *is an analytic function defined on the whole plane, having all* π*-integers as its singular points, where it has residues* 1.

PROOF. Consider a point $z \notin \pi\mathbb{Z}$. Consider a disk D, not containing π-integers with center at z. Then formula (4.5.3) transforms to the Cauchy Integral Formula. And our assertion is proved by termwise integration of the power expansion of $\frac{1}{\zeta-z}$ just with the same arguments as was applied there. The same formula (4.5.3) allows us to evaluate the residues. □

THEOREM 4.5.11. $\cot z = \frac{1}{z} + \sum_{k=1}^{\infty} \frac{2z}{z^2-k^2\pi^2}$.

PROOF. Consider the difference $R(z) = \cot z - \operatorname{ctg}(z)$. This is an analytic function which has π-integers as singular points and has residues 0 in all of these. Hence $R(z) = \frac{1}{2\pi i}\oint_{\partial D} \frac{R(\zeta)}{\zeta-z} d\zeta$ for any $z \notin \pi\mathbb{Z}$. We will prove that $R(z)$ is constant. Let z_0 and ζ be a pair of different points not belonging to $\pi\mathbb{Z}$. Then for any D such that $\partial D \cap \pi\mathbb{Z} = \varnothing$ one has

$$\begin{aligned} R(z) - R(z_0) &= \frac{1}{2\pi i} \oint_{\partial D} R(\zeta) \left(\frac{1}{\zeta - z} - \frac{1}{\zeta - z_0} \right) d\zeta \\ &= \frac{1}{2\pi i} \oint_{\partial D} \frac{R(z)(z-z_0)}{(\zeta-z)(\zeta-z_0)}. \end{aligned} \tag{4.5.4}$$

Let us define D_n for a natural $n > 3$ as the rectangle bounded by the lines $\operatorname{Re} z = \pm(\pi/2 - n\pi)$, $\operatorname{Im} z = \pm n\pi$. Since $|R(z)| \leq 7$ by Lemmas 4.5.2, 4.5.3, 4.5.7, and 4.5.8 the integrand of (4.5.4) eventually is bounded by $\frac{7|z-z_0|}{n^2}$. The contour of integration consists of four monotone curves of diameter $< 2n\pi$. By the Estimation Lemma 3.5.4, the integral can be estimated from above by $\frac{32\pi n 7|z-z_0|}{n^2}$. Hence the limit of our integral as n tends to infinity is 0. This implies $R(z) = R(z_0)$. Hence $R(z)$ is constant and the value of the constant we find by putting $z = \pi/2$. As $\cot \pi/2 = 0$, the value of the constant is

$$\operatorname{ctg}(\pi/2) = \lim_{n\to\infty} \sum_{k=-n}^{n} \frac{1}{\pi/2 - k\pi} = \frac{2}{\pi} \lim_{n\to\infty} \sum_{k=-n}^{n} \frac{1}{1-2k}.$$

This limit is zero because

$$\sum_{k=-n}^{n}\frac{1}{1-2k}=\sum_{k=-n}^{0}\frac{1}{1-2k}+\sum_{k=1}^{n}\frac{1}{1-2k}=\sum_{k=0}^{n}\frac{1}{2k+1}+\sum_{k=1}^{n}-\frac{1}{2k-1}=\frac{1}{2n+1}.$$

□

Summation of series by $\cot z$**.**

THEOREM 4.5.12. *For any rational function $R(z)$, which is not singular in integers and has degree ≤ -2, one has $\sum_{k=-\infty}^{\infty} R(n) = -\sum_z \operatorname{res} \pi\cot(\pi z)R(z)$.*

PROOF. In this case the integral $\lim_{n\to\infty}\oint_{\partial D_n/pi} R(z)\pi\cot\pi z = 0$. Hence the sum of all residues of $R(z)\pi\cot\pi z$ is zero. The residues at π-integers gives $\sum_{k=-\infty}^{\infty} R(k)$. The rest gives $-\sum_z \operatorname{res}\pi\cot(\pi z)R(z)$. □

Factorization of $\sin x$**.** Theorem 4.5.11 with πz substituted for z gives the series $\pi\cot\pi z = \sum_{k=-\infty}^{\infty}\frac{1}{z-k}$. The half of the series on the right-hand side consisting of terms with nonnegative indices represents a function, which formally telescopes $-\frac{1}{z}$. The negative half telescopes $\frac{1}{z}$. Let us bisect the series into nonnegative and negative halves and add $\sum_{k=-\infty}^{\infty}\frac{1}{k}[k\neq 0]$ to provide convergence:

$$\sum_{k=-\infty}^{-1}\left(\frac{1}{z-k}+\frac{1}{k}\right)+\sum_{k=0}^{\infty}\left(\frac{1}{z-k}+\frac{1}{k+1}\right)$$
$$=\sum_{k=1}^{\infty}\left(-\frac{1}{k}+\frac{1}{z+k}\right)+\sum_{k=1}^{\infty}\left(\frac{1}{z+1-k}+\frac{1}{k}\right).$$

The first of the series on the right-hand side represents $-\digamma(z)-\gamma$, the second is $\digamma(-z+1)+\gamma$. We get the following *complement formula* for the digamma function:

$$-\digamma(z)+\digamma(1-z)=\pi\cot\pi z.$$

Since $\Theta''(z+1)=\digamma'(z)=\mathit{\Gamma}(z)$ (Lemma 4.4.11) it follows that $\Theta'(1+z)=\digamma(z)+c$ and $\Theta'(-z)=-(\digamma(1-z)+c)$. Therefore $\Theta'(1+z)+\Theta'(-z)=\pi\cot\pi z$. Integration of the latter equality gives $-\Theta(1+z)-\Theta(-z)=\ln\sin\pi z+c$. Changing z by $-z$ we get $\Theta(1-z)+\Theta(z)=-\ln\sin\pi z+c$. Exponentiating gives $\Gamma(1-z)\Gamma(-z)=\frac{1}{\sin\pi z}c$. One defines the constant by putting $z=\frac{1}{2}$. On the left-hand side one gets $\Gamma(\frac{1}{2})^2=\pi$, on the right-hand side, c. Finally we get the *complement formula for the Gamma-function*:

$$\Gamma(1-z)\Gamma(z)=\frac{\pi}{\sin\pi z}. \tag{4.5.5}$$

Now consider the product $\prod_{k=1}^{\infty}(1-\frac{x^2}{k^2})$. Its canonical form is

$$\prod_{n=1}^{\infty}\left\{\left(1-\frac{x}{n}\right)e^{\frac{x}{n}}\right\}^{-1}\prod_{n=1}^{\infty}\left\{\left(1+\frac{x}{n}\right)e^{-\frac{x}{n}}\right\}^{-1}. \tag{4.5.6}$$

The first product of (4.5.6) is equal to $-\frac{e^{\gamma x}}{x\Gamma(-x)}$, and the second one is $\frac{e^{-\gamma x}}{x\Gamma(-x)}$. Therefore the whole product is $-\frac{1}{x^2\Gamma(x)\Gamma(-x)}$. Since $\Gamma(1-x)=-x\Gamma(-x)$ we get the following result

$$\frac{1}{\Gamma(x)\Gamma(1-x)}=x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2}\right).$$

Comparing this to (4.5.5) and substituting πx for x we get the Euler formula:

$$\sin x = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{\pi^2 k^2} \right).$$

Problems.

1. Expand $\tan z$ into a power series.

2. Evaluate $\sum_{k=1}^{\infty} \frac{1}{1+k^2}$.

3. Evaluate $\sum_{k=1}^{\infty} \frac{1}{1+k^4}$.

4.6. Divergent Series

On the contents of the lecture. "Divergent series is a pure handiwork of Diable. It is a full nonsense to say that $1^{2n} - 2^{2n} + 3^{2n} - \cdots = 0$. Do you keep to die laughing about this?" (N.H. Abel letter to ...). The twist of fate: now one says that that the above mentioned equality holds in *Abel's sense*.

The earliest analysts thought that any series, convergent or divergent, has a sum given by God and the only problem is to find it correctly. Sometimes they disagreed what is the correct answer. In the nineteenth century divergent series were expelled from mathematics as a "handiwork of Diable" (N.H. Abel). Later they were rehabilitated (see G.H. Hardy's book *Divergent Series*[1]). Euler remains the unsurpassed master of divergent series. For example, with the help of divergent series he discovered Riemann's functional equation of the ζ-function a hundred years before Riemann.

Evaluations with divergent series. Euler wrote: "My pen is clever than myself." Before we develop a theory let us simply follow to Euler's pen. The fundamental equality is

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}. \tag{4.6.1}$$

Now we, following Euler, suppose that this equality holds for all $x \neq 1$. In the second lecture we were confused by some unexpected properties of divergent series. But now in contrast with the second lecture we do not hurry up to land. Let us look around.

Substituting $x = -e^y$ in (4.6.1) one gets

$$1 - e^y + e^{2y} - e^{3y} + \cdots = \frac{1}{1+e^y}.$$

On the other hand

$$\frac{1}{1+e^y} = \frac{1}{e^y - 1} - \frac{2}{e^{2y} - 1}. \tag{4.6.2}$$

Since

$$\frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k. \tag{4.6.3}$$

One derives from (4.6.2) via (4.6.3)

$$\frac{1}{e^y + 1} = \sum_{k=1}^{\infty} \frac{B_k(1-2^k)}{k!} y^{k-1}. \tag{4.6.4}$$

Let us differentiate repeatedly n-times the equality (4.6) by y. The left-hand side gives $\sum_{k=0}^{\infty}(-1)^k k^n e^{ky}$. In particular for $y = 0$ we get $\sum_{k=0}^{\infty}(-1)^k k^n$. We get on the right-hand side by virtue of (4.6.4) the following

$$\left(\frac{d}{dy}\right)^n \frac{1}{1+e^y} = \frac{B_{n+1}(1-2^{n+1})}{n+1}.$$

Combining these results we get the following equality

$$1^n - 2^n + 3^n - 4^n + \cdots = \frac{B_{n+1}(2^{n+1}-1)}{n+1}. \tag{4.6.5}$$

[1] G.H. Hardy, *Divergent Series*, Oxford University Press, 1949.

Since odd Bernoulli numbers vanish, we get

$$1^{2n} - 2^{2n} + 3^{2n} - 4^{2n} + \dots = 0.$$

Consider an even analytic function $f(x)$, such that $f(0) = 0$. In this case $f(x)$ is presented by a power series $a_1x^2 + a_2x^4 + a_3x^6 + \dots$, then

$$\begin{aligned}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{f(kx)}{k^2} &= \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^2}\sum_{n=1}^{\infty}a_nx^{2n}k^{2n}\\ &= \sum_{n=1}^{\infty}a_nx^{2n}\sum_{k=1}^{\infty}(-1)^{k-1}k^{2n-2}\\ &= a_1x^2(1-1+1-1+\dots)\\ &= \frac{a_1x^2}{2}.\end{aligned}$$

In particular, for $f(x) = 1 - \cos x$ this equality turns into

$$\sum_{k=1}^{\infty}(-1)^{k-1}\frac{1-\cos kx}{k^2} = \frac{x^2}{4}. \tag{4.6.6}$$

For $x = \pi$ the equality (4.6.6) gives

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots = \frac{\pi^2}{8}.$$

Since

$$\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2} = \sum_{k=1}^{\infty}\frac{1}{k^2} - \sum_{k=1}^{\infty}\frac{1}{(2k)^2} = \left(1-\frac{1}{4}\right)\sum_{k=1}^{\infty}\frac{1}{k^2}$$

one derives the sum of the Euler series:

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}.$$

We see that calculations with divergent series sometimes give brilliant results. But sometimes they give the wrong result. Indeed the equality (4.6.6) generally is untrue, because on the left-hand side we have a periodic function and on the right-hand side a non-periodic one. But it is true for $x \in [-\pi, \pi]$. Termwise differentiation of (4.6.6) gives the true equality (3.4.2), which we know from Lecture 3.4.

Euler's sum of a divergent series. Now we develop a theory justifying the above evaluations. Euler writes that the value of an infinite expression (in particular the sum of a divergent series) is equal to the value of a finite expression whose expansion gives this infinite expression. Hence, numerical equalities arise by substituting a numerical value for a variable in a generating functional identity. To evaluate the sum of a series $\sum_{k=0}^{\infty} a_k$ Euler usually considers its *power generating function* $g(z)$ represented by the power series $\sum_{k=0}^{\infty} a_kz^k$, and supposes that the sum of the series is equal to $g(1)$.

To be precise suppose that the power series $\sum_{k=0}^{\infty} a_kz^k$ converges in a neighborhood of 0 and there is an analytic function $g(z)$ defined in a domain U containing a path p from 0 to 1 and such that $g(z) = \sum_{k=0}^{\infty} a_kz^k$ for z sufficiently close to 0 and 1 is a regular point of g. Then the series $\sum_{k=0}^{\infty} a_k$ is called *Euler summable* and the value $g(1)$ is called its *analytic Euler sum* with respect to p. And we will use a special sign $\simeq$ to denote the analytical sum.

By the Uniqueness Theorem 3.6.9 the value of analytic sum of a series is uniquely defined for a fixed p. But this value generally speaking depends on the path. For example, let us consider the function $\sqrt{1+x}$. Its binomial series for $x=-2$ turns into

$$-1+1-\frac{1}{2!}-\frac{1\cdot 3}{3!}-\frac{1\cdot 3\cdot 5}{4!}-\cdots-\frac{(2k-1)!!}{(k+1)!}-\cdots.$$

For $p(t)=e^{i\pi t}$ one sums up this series to i, because it is generated by the function $\exp\frac{\ln(1+z)}{2}$ defined in the upper half-plane. And along $p(t)=e^{-i\pi t}$ this series is summable to $-i$ by $\exp\frac{-\ln(1+z)}{2}$ defined in the lower half-plane.

For a fixed path the analytic Euler sum evidently satisfies the Shift, Multiplication and Addition Formulas of the first lecture. But we see that the analytic sum of a real series may be purely imaginary. Hence the rule $\operatorname{Im}\sum_{k=0}^{\infty}a_k \simeq \sum_{k=0}^{\infty}\operatorname{Im}a_k$ fails for the analytic sum. The Euler sum along $[0,1]$ coincides with the Abel sum of the series in the case when both of them exist.

In above evaluations we apply termwise differentiation to functional series. If the Euler sum $\sum_{k=1}^{\infty}f_k(z)$ is equal to $F(z)$ for all z in a domain this does not guarantee the possibility of termwise differentiation. To guarantee it we suppose that the function generating the equality $\sum_{k=1}^{\infty}f_k(z)\simeq F(z)$ analytically depends on z. To formalize the last condition we have to introduce analytic functions of two variables.

Power series of two variables. A power series of two variables z,w is defined as a formal unordered sum $\sum_{k,m}a_{km}z^kw^m$, over $\mathbb{N}\times\mathbb{N}$ — the set of all pairs of nonnegative integers.

For a function of two variables $f(z,w)$ one defines its *partial derivative* $\frac{\partial f(z_0,w_0)}{\partial z}$ with respect to z at the point (z_0,w_0) as the limit of $\frac{f(z_0+\Delta z,w_0)-f(z_0,w_0)}{\Delta z}$ as Δz tends to 0.

LEMMA 4.6.1. *If $\sum a_{km}z_1^kw_1^m$ absolutely converges, then both $\sum a_{km}z^kw^m$ and $\sum ma_{km}z^kw^{m-1}$ absolutely converge provided $|z|<|z_1|$, $|w|<|w_1|$. And for any fixed z, such that $|z|<|z_1|$ the function $\sum ma_kz^kw^{m-1}$ is the partial derivative of $\sum a_{km}z^kw^m$ with respect to w.*

PROOF. Since $\sum|a_{km}||z_1|^k|w_1|^m<\infty$ the same is true for $\sum|a_{km}||z|^k|w|^m$ for $|z|<|z_1|$, $|w|<|w_1|$. By the Sum Partition Theorem we get the equality

$$\sum a_{km}z^kw^m=\sum_{m=0}^{\infty}w^m\sum_{k=0}^{\infty}a_{km}z^k.$$

For any fixed z the right-hand side of this equality is a power series with respect to w as the variable. By Theorem 3.3.9 its derivative by w, which coincides with the partial derivative of the left-hand side, is equal to

$$\sum_{m=0}^{\infty}mw^{m-1}\sum_{k=0}^{\infty}a_{km}z^k=\sum ma_{km}w^{m-1}z^k.$$

□

Analytic functions of two variables. A function of two variables $F(z,w)$ is called analytic at the point (z_0,w_0) if for (z,w) sufficiently close to (z_0,w_0) it can be presented as a sum of a power series of two variables.

THEOREM 4.6.2.

(1) *If $f(z,w)$ and $g(z,w)$ are analytic functions, then $f+g$ and fg are analytic functions.*
(2) *If $f_1(z), f_2(z)$ and $g(z,w)$ are analytic functions, then $g(f_1(z), f_2(w))$ and $f_1(g(z,w))$ are analytic functions.*
(3) *The partial derivative of any analytic function is an analytic function.*

PROOF. The third statement follows from Lemma 4.6.1. The proofs of the first and the second statements are straightforward and we leave them to the reader. □

Functional analytical sum. Let us say that a series $\sum_{k=1}^{\infty} f_k(z)$ of analytic functions is *analytically summable* to a function $F(z)$ in a domain $U \subset \mathbb{C}$ along a path p in $\mathbb{C} \times \mathbb{C}$, such that $p(0) \in U \times 0$ and $p(1) \in U \times 1$, if there exists an analytic function of two variables $F(z,w)$, defined on a domain W containing p, $U \times 0$, $U \times 1$, such that for any $z_0 \in U$ the following two conditions are satisfied:

(1) $F(z_0, 1) = F(z_0)$.
(2) $F(z,w) = \sum \frac{f_m^{(k)}(z_0)}{k!}(z-z_0)^k w^m$ for sufficiently small $|w|$ and $|z-z_0|$.

Let us remark that the analytic sum does not change if we change p keeping it inside W. That is why one says that the sum is evaluated along the domain W.

To denote the functional analytical sum we use the sign $\cong$. And we will write also $\cong_W$ and $\cong_p$ to specify the domain or the path of summation.

The function $F(z,w)$ will be called the *generating function* for the analytical equality $\sum_{k=1}^{\infty} f_k(z) \cong F(z)$.

LEMMA 4.6.3. *If $f(z)$ is an analytic function in a domain U containing 0, such that $f(z) = \sum_{k=0}^{\infty} a_k z^k$ for sufficiently small $|z|$, then $f(z) \cong_W \sum_{k=0}^{\infty} a_k z^k$ in U for $W = \{(z,w) \mid wz \in U\}$.*

PROOF. The generating function of this analytical equality is $f((z-z_0)w)$. □

LEMMA 4.6.4 (on substitution). *If $F(z) \cong_p \sum_{k=0}^{\infty} f_k(z)$ in U and $g(z)$ is an analytic function, then $F(g(z)) \cong_{g(p)} \sum_{k=0}^{\infty} f_k(g(z))$ in $g^{-1}(U)$.*

PROOF. Indeed, if $F(z,w)$ generates $F(z) \cong_p \sum_{k=0}^{\infty} f_k(z)$, then $F(g(z),w))$ generates $F(g(z)) \cong_{g(p)} \sum_{k=0}^{\infty} f_k(g(z))$. □

N. H. Abel was the first to have some doubts about the legality of termwise differentiation of functional series. The following theorem justifies this operation for analytic functions.

THEOREM 4.6.5. *If $\sum_{k=1}^{\infty} f_k(z) \cong_p F(z)$ in U then $\sum_{k=1}^{\infty} f_k'(z) \cong_p F'(z)$ in U.*

PROOF. Let $F(z,w)$ be a generating function for $\sum_{k=1}^{\infty} f_k(z) \cong_p F(z)$. We demonstrate that its partial derivative by z (denoted $F'(z,w)$) is the generating function for $\sum_{k=1}^{\infty} f_k'(z) \cong_p F'(z)$. Indeed, locally in a neighborhood of $(z_0, 0)$ one has $F(z,w) = \sum \frac{f_m^{(k)}(z_0)}{k!} w^m (z-z_0)^k$. By virtue of Lemma 4.6.1 its derivative by z is $F'(z,w) = \sum \frac{f_m^{(k)}(z_0)}{(k-1)!} w^m (z-z_0)^{k-1} = \sum \frac{f_m'^{(k)}(z_0)}{k!} w^m (z-z_0)^k$. □

The dual theorem on termwise integration is the following one.

THEOREM 4.6.6. *Let $\sum_{k=1}^{\infty} f_k \cong F$ be generated by $F(z,w)$ defined on $W = U \times V$. Then for any path q in U one has $\int_q F(z)\,dz \simeq \sum_{k=1}^{\infty} \int_q f_k(z)\,dz$.*

PROOF. The generating function for integrals is defined as $\int_q F(z,w)\,dz$. □

The proof of the following theorem is left to the reader.

THEOREM 4.6.7. *If* $\sum_{k=0}^{\infty} f_k \cong_p F$ *and* $\sum_{k=0}^{\infty} g_k \cong_p G$ *then* $\sum_{k=0}^{\infty}(f_k+g_k) \cong_p F+G$, $\sum_{k=1}^{\infty} f_k \cong_p F - f_0$, $\sum_{k=0}^{\infty} cf_k \cong_p cF$

Revision of evaluations. Now we are ready to revise the above evaluation equipped with the theory of analytic sums. Since all considered generating functions in this paragraph are single valued, the results do not depend on the choice of the path of summation. That is why we drop the indications of path below.

The equality (4.6.1) is the analytical equivalence generated by $\frac{1}{1-tx}$. The next equality (4.6.7) is the analytical equivalence by Lemma 4.6.4. The equality (4.6.3) is analytical equivalence due to Lemma 4.6.3. Termwise differentiation of (4.6.7) is correct by virtue of Theorem 4.6.5. Therefore the equality (4.6.5) is obtained by the restriction of an analytical equivalence. Hence the Euler sum of $\sum_{k=1}^{\infty}(-1)^k k^{2n}$ is equal to 0. Since the series $\sum_{k=1}^{\infty}(-1)^k k^{2n} z^k$ converges for $|z|<1$ its value coincides with the value of the generating function. And the limit $\lim_{z\to 1-0}\sum_{k=1}^{\infty}(-1)^k k^{2n} z^k$ gives the Euler sum, which is zero. Hence as a result of our calculations we have found *Abel's sum* $\sum_{k=1}^{\infty}(-1)^k k^{2n} = 0$.

Now we choose another way to evaluate the Euler series. Substituting $x = e^{\pm i\theta}$ in (4.6.1) for $0<\theta<2\pi$ one gets

$$\begin{aligned} 1 + e^{i\theta} + e^{2i\theta} + e^{3i\theta} + \ldots &\cong \frac{1}{1-e^{i\theta}}, \\ 1 + e^{-i\theta} + e^{-2i\theta} + e^{-3i\theta} + \ldots &\cong \frac{1}{1-e^{-i\theta}}. \end{aligned} \tag{4.6.7}$$

Termwise addition of the above lines gives for $\theta \in (0,2\pi)$ the following equality

$$\cos\theta + \cos 2\theta + \cos 3\theta + \cdots \cong -\frac{1}{2}. \tag{4.6.8}$$

Integration of (4.6.8) from π to x with subsequent replacement of x by θ gives by Theorem 4.6.6:

$$\sum_{k=1}^{\infty} \frac{\sin k\theta}{k} \cong \frac{\pi-\theta}{2} \quad (0<\theta<2\pi).$$

A second integration of the same type gives

$$\sum_{k=1}^{\infty} \frac{\cos k\theta - (-1)^k}{k^2} \cong \frac{(\pi-\theta)^2}{4}.$$

Putting $\theta = \frac{\pi}{2}$ we get

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} - \frac{1}{4}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \simeq \frac{\pi^2}{16}.$$

Therefore

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{12}.$$

Since

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} + 2\sum_{k=1}^{\infty} \frac{1}{(2k)^2}$$

one gets

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{1}{2}\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{6}.$$

Problems.

1. Prove that the analytic sum of convolution of two series is equal to the product of analytic sums of the series.

2. Suppose that for all $n \in \mathbb{N}$ one has $A_n \simeq \sum_{k=0}^{\infty} a_{n,k}$ and $B_n \simeq \sum_{k=0}^{\infty} a_{k,n}$. Prove that the equality $\sum_{k=0}^{\infty} A_k = \sum_{k=0}^{\infty} B_k$ holds provided there is an analytic function $F(z,w)$ coinciding with $\sum a_{k,n} z^k w^n$ for sufficiently small $|w|$,$|z|$ which is defined on a domain containing a path joining $(0,0)$ with $(1,1)$ analytically extended to $(1,1)$ (i.e., $(1,1)$ is a regular point of $F(z,w)$).