Uppsala Lectures on Calculus

E. V. Shchepin

September – November 2001

Contents

Chapter 1

Series

Legends

"One of the great mathematical challenges of the early 18th century was to find an expression for the sum of reciprocal squares

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \dots \tag{1.0.1}$$

Joh. Bernoulli eagerly sought for this expression for many decades." (E. Hairer, G. Wanner, *Analysis by Its Hystory*, Undergraduate Texts in Mathematics, Springer, 1997)

In 1689 Jac. Bernoulli proved convergence of the series. In 1728-1729 Goldbah and D. Bernoulli evaluated the series with accuracy 0.01. Stirling in 1730 found 8 digits of the sum.

L. Euler in 1734 calculated the first 18 digits (!) of the sum (1.0.1) after the decimal point and recognized $\pi^2/6$, which has the same 18 digits. He conjectured that the infinite sum is equal to $\pi^2/6$. In 1735 Euler discovered an expansion of the sine function into an infinite product of polynomials:

$$\frac{\sin x}{x} = \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{2^2\pi^2}\right)\left(1 - \frac{x^2}{3^2\pi^2}\right)\left(1 - \frac{x^2}{4^2\pi^2}\right)\dots \tag{1.0.2}$$

Comparing this presentation with the standard sine series expansion

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots \tag{1.0.3}$$

Euler not only proved that the sum (1.0.1) is equal to $\pi^2/6$, moreover he calculated all sums of the type

$$1 + \frac{1}{2^k} + \frac{1}{3^k} + \frac{1}{4^k} + \frac{1}{5^k} + \dots \tag{1.0.4}$$

for even k.

Putting $x = \frac{\pi}{2}$ in (1.0.2) he got the beautiful Wallis' Product:

$$\frac{\pi}{2} = \frac{2 \cdot 2}{1 \cdot 3}\frac{4 \cdot 4}{3 \cdot 5}\frac{6 \cdot 6}{5 \cdot 7}\frac{8 \cdot 8}{7 \cdot 9}\dots \tag{1.0.5}$$

which had been known since 1655. But Euler's first proof of (1.0.2) was not satisfactory. In 1748, in his famous "Introducio in Analysin Infinitorum", he presented a proof which was sufficiently rigorous for the 18th century. The series of reciprocal squares was named the *Euler series*.

If somebody wants to understand all the details of the above legend he had to study a lot of things, up to complex contour integrals. This is why the detailed mathematical exposition of Euler's series legend turns into an entire course of Calculus. The fascinating history of Euler's series is the guiding thread of the present course, "In Euler's footsteps".

1.1 Autorecursion of Infinite Expressions

On the contents of the lecture. The lecture presents a romantic style of early analytics. The motto of the lecture could be "infinity, equality and no definitions!". "Infinity" is the main personage we will play with today. We demonstrate how infinite expressions (i.e infinite sums, products, fractions) arise in solutions of simple equations, how it is possible to calculate them, and how the results of such calculations apply to finite mathematics. In particular, we will deduce the Euler-Binet formula for Fibonacci numbers, the first Euler's formula of the course. We become acquainted with the geometric series and the golden section.

Achilles and the turtle. The ancient Greek philosopher Zeno claimed that Achilles pursuing a turtle could never pass it by, in spite of the fact that his velocity was much greater than the velocity of the turtle. His arguments adopted to our purposes are the following.

First Zeno proposed a pursuing algorithm for Achilles:

Initialization. Assign to the variable "goal" the original position of the turtle.

Action. Reach the "goal".

Correction. If the current turtle's position is "goal", then STOP, else reassign to the variable "goal" the current position of the turtle and go to Action.

Secondly, Zeno remarks that this algorithms never stops if the turtle constantly moves in one direction.

And finally, he notes that Achilles has to follow his algorithm if he want pass the turtle by. He may be not aware of this algorithm, but unconsciously he must perform it. Because he cannot run the turtle down without reaching the original position of the turtle and then all positions of the turtle which the variable "goal" takes.

Zeno's algorithm generates a sequence of times $\{t_k\}$, where t_k is the time of execution of the k-th Action of the algorithm. And the whole time of work of the algorithm is the infinite sum $\sum\limits_{k=1}^{\infty} t_k$; and this sum expresses the time Achilles needs to run the turtle down. (The Corrections take zero time, because Achilles really does not think about them.) Let us name this sum *Zeno series*.

Assume that both Achilles and the turtle run with constant velocities v and w, respectively. Denote the initial distance between Achilles and the turtle by d_0. Then $t_1 = \frac{d_0}{v}$. The turtle in this time moves by the distance $d_1 = t_1 w = \frac{w}{v} d_0$. By his second Action Achilles overcomes this distance in time $t_2 = \frac{d_1}{v} = \frac{w}{v} t_1$, while the turtle moves away by the distance $d_2 = t_2 w = \frac{w}{v} d_1$. So we see that sequences of the times $\{t_k\}$ and distances $\{d_k\}$ satisfy the following *recurrence relations*: $t_k = \frac{w}{v} t_{k-1}$, $d_k = \frac{w}{v} d_{k-1}$.

Hence $\{t_k\}$ as well as $\{d_k\}$ are *geometric progressions* with ratio $\frac{w}{v}$. And the time t which Achilles needs to run the turtle down is

$$t = t_1 + t_2 + t_3 + \dots = t_1 + \frac{w}{v} t_1 + \frac{w^2}{v^2} t_1 + \dots = t_1 \left(1 + \frac{w}{v} + \frac{w^2}{v^2} + \dots \right)$$

In spite of Zeno, we know that Achilles does catch up with the turtle. And one easily gets the time t he needs to do it by the following argument: the distance between Achilles and the turtle permanently decreases with the velocity $v - w$. Consequently it becomes 0 in the time $t = \frac{d_0}{v-w} = t_1\frac{v}{v-w}$. Comparing the results we come to the following conclusion

$$\frac{v}{v-w} = 1 + \frac{w}{v} + \frac{w^2}{v^2} + \frac{w^3}{v^3} + \dots \tag{1.1.1}$$

Infinite substitution. We see that some infinite expressions represent finite values. The fraction in the left-hand side of (1.1.1) expands into the infinite series on the right-hand side. Infinite expressions play a key rôle in mathematics and physics. Solutions of equations quite often are presented as infinite expressions.

For example let us consider the following simple equation

$$t = 1 + qt \tag{1.1.2}$$

Substituting on the right-hand side $1 + qt$ instead of t, one gets a new equation $t = 1 + q(1 + qt) = 1 + q + q^2t$. Any solution of the original equation satisfies this one. Repeating this trick, one gets $t = 1+q(1+q(1+qt)) = 1+q+q^2+q^3t$. Repeating this infinitely many times, one eliminates t on the right hand side and gets a solution of (1.1.2) in an infinite form

$$t = 1 + q + q^2 + q^3 + \dots = \sum_{k=0}^{\infty} q^k \tag{1.1.3}$$

On the other hand, equation (1.1.2) solved in usual way gives $t = \frac{1}{1-q}$. As a result, we obtain the following formula

$$\frac{1}{1-q} = 1 + q + q^2 + q^3 + q^4 + \dots = \sum_{k=0}^{\infty} q^k \tag{1.1.4}$$

which represents a special case of (1.1.1) for $v = 1$, $w = q$.

Autorecursion. An infinite expression of the form $a_1 + a_2 + a_3 + \dots$ is called a *series* and is concisely denoted by $\sum_{k=1}^{\infty} a_k$. Now we consider a summation method for series which is inverse to the above method of infinite substitution. To find a sum of a series we shall construct an equation which is satisfied by its sum. We name this method *autorecursion*. Recursion means "return to something known". Autorecursion is "return to oneself".

The series $a_2 + a_3 + \dots = \sum_{k=2}^{\infty} a_k$ obtained from $\sum_{k=1}^{\infty} a_k$ by dropping its first term is called *the shift* of $\sum_{k=1}^{\infty} a_k$.

We will call the following equality the *shift formula*:

$$\sum_{k=1}^{\infty} a_k = a_1 + \sum_{k=2}^{\infty} a_k \tag{1.1.5}$$

Another basic formula we need is the following *multiplication formula*:

$$\lambda \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \lambda a_k \tag{1.1.6}$$

These two formulas are all one needs to find the sum of geometric series $\sum_{k=0}^{\infty} q^k$. To be exact, the multiplication formula gives the equality $\sum_{k=1}^{\infty} q^k = q\sum_{k=0}^{\infty} q^k$. Hence the shift formula turns into equation $x = 1 + qx$, where x is $\sum_{k=0}^{\infty} q^k$. The solution of this equation gives us the formula (1.1.4) for the sum of the geometric series again.

From this formula, one can deduce the formula for the sum of a finite geometric progression. By $\sum_{k=0}^{n} a_k$ is denoted the sum $a_0 + a_1 + a_2 + \cdots + a_n$. One has

$$\sum_{k=0}^{n-1} q^k = \sum_{k=0}^{\infty} q^k - \sum_{k=n}^{\infty} q^k = \frac{1}{1-q} - \frac{q^n}{1-q} = \frac{1-q^n}{1-q} \tag{1.1.7}$$

This is an important formula which was traditionally studied in school.

The series $\sum_{k=0}^{\infty} kx^k$. To find the sum of $\sum_{k=1}^{\infty} kx^k$ we have to apply additionally the following *addition formula*,

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k \tag{1.1.8}$$

which is the last general formula for series we introduce in the first lecture.

Reindexing the shift $\sum_{k=2}^{\infty} kx^k$ we give it the form $\sum_{k=1}^{\infty}(k+1)x^{k+1}$. Further it splits into two parts

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

by the addition formula. The first summand is the original sum multiplied by x. The second is a geometric series. We already know its sum. Now the shift formula for the sum $s(x)$ of the original series turns into equation $s(x) = x + x\frac{x}{1-x} + xs(x)$. Its solution is $s(x) = \frac{x}{(1-x)^2}$.

Fibonacci Numbers. Starting with $\phi_0 = 0$, $\phi_1 = 1$ and applying the recurrence relation

$$\phi_{n+1} = \phi_n + \phi_{n-1}, \tag{1.1.9}$$

one constructs an infinite sequence of numbers $0, 1, 2, 3, 5, 8, 13, 21, \ldots$, called *Fibonacci numbers*. We are going to get a formula for ϕ_n.

To do this let us consider the following function $\Phi(x) = \sum_{k=0}^{\infty} \phi_k x^k$. which is called the *generating function* for the sequence $\{\phi_k\}$. Since $\phi_0 = 0$, the sum $\Phi(x) + x\Phi(x)$ transforms in the following way:

$$\sum_{k=1}^{\infty} \phi_k x^k + \sum_{k=1}^{\infty} \phi_{k-1} x^k = \sum_{k=1}^{\infty} \phi_{k+1} x^k = \frac{\Phi(x) - x}{x} \tag{1.1.10}$$

Multiplying both sides of above equation by x and collecting all terms containing $\Phi(x)$ on the right-hand side, one gets $x = \Phi(x) - x\Phi(x) - x^2\Phi(x) = x$. It leads to

$$\Phi(x) = \frac{x}{1-x-x^2}$$

The roots of the equation $1-x-x^2 = 0$ are $\frac{-1\pm\sqrt{5}}{2}$. More famous is the pair of their inverses $\frac{1\pm\sqrt{5}}{2}$. The number $\phi = \frac{-1+\sqrt{5}}{2}$ is so called *golden section* or *golden mean*. It plays significant role in mathematics, architecture and biology. Its dual is $\hat{\phi} = \frac{-1-\sqrt{5}}{2}$. Then $\phi\hat{\phi} = -1$, and $\phi+\hat{\phi} = 1$. Hence $(1-x\phi)(1-x\hat{\phi}) = 1-x-x^2$, which in its turn leads to the following decomposition:

$$\frac{x}{x^2+x-1} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-\phi x} - \frac{1}{1-\hat{\phi}x}\right)$$

We expand both fractions on the right hand side into geometric series:

$$\frac{1}{1-\phi x} = \sum_{k=0}^{\infty} \phi^k x^k \qquad \frac{1}{1-\hat{\phi}x} = \sum_{k=0}^{\infty} \hat{\phi}^k x^k \tag{1.1.11}$$

This gives the following representation for the generating function

$$\Phi(x) = \frac{1}{\sqrt{5}}\sum_{k=0}^{\infty}(\phi^k - \hat{\phi}^k)x^k$$

On the other hand the coefficient at x^k in the original presentation of $\Phi(x)$ is ϕ_k. Hence

$$\phi_k = \frac{1}{\sqrt{5}}(\phi^k - \hat{\phi}^k) = \frac{(\sqrt{5}+1)^k + (-1)^k(\sqrt{5}-1)^k}{2^k\sqrt{5}} \tag{1.1.12}$$

This is called the *Euler–Binet* formula. It is possible to check it for small k and then prove it by induction using Fibonacci recurrence.

Continued fractions. Application of the method of infinite substitution to the solution of quadratic equation leads us to a new type of infinite expressions, so-called *continued fractions*. Let us consider the golden mean equation $x^2 - x - 1 = 0$. Rewrite it as $x = 1 + \frac{1}{x}$. Substituting $1 + \frac{1}{x}$ instead of x on the right-hand side we get $x = 1 + \frac{1}{1+\frac{1}{x}}$. Repeating the substitution infinitely many times we obtain a solution in the form of the *continued fraction*

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \dots}}} \tag{1.1.13}$$

As this fraction seems to represent a positive number and the golden mean is the unique positive root of the golden mean equation, it is natural to conclude that this fraction is equal to $\phi = \frac{1+\sqrt{5}}{2}$. This is true and this representation allows one to calculate the golden mean and $\sqrt{5}$ effectively with great precision.

To be precise, consider the sequence

$$1, \qquad 1+\frac{1}{1}, \qquad 1+\frac{1}{1+\frac{1}{1}}, \qquad 1+\frac{1}{1+\frac{1}{1+\frac{1}{1}}}\dots \tag{1.1.14}$$

of the so-called *convergents* of the continued fraction (1.1.13). Let us remark that all odd convergent are less than ϕ and all even are greater than ϕ. To see this, compare the nth convergent with the corresponding term of the following sequence of fractions:

$$1+\frac{1}{x}, \quad 1+\frac{1}{1+\frac{1}{x}}, \quad 1+\frac{1}{1+\frac{1}{1+\frac{1}{x}}}, \quad 1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{x}}}}\dots \tag{1.1.15}$$

We know that for $x=\phi$ all terms of the above sequence are equal to ϕ. Hence all we need is to observe how the removal of $\frac{1}{x}$ affects the value of the considered fraction. The value of the first fraction of the sequence decreases, the value of the second fraction increases. If we denote the value of nth fraction by f_n, then the value of the next fraction is given by the following recurrence relation:

$$f_{n+1}=1+\frac{1}{f_n} \tag{1.1.16}$$

Hence the increasing of f_n decreases f_{n+1} and decreasing of f_n increases f_{n+1}. Consequently in general all odd fractions of the sequence (1.1.15) are less than corresponding convergent, and all even are greater. The recurrence relation (1.1.16) is valid for the golden mean convergent. By this recurrence relation one can quickly calculate the first ten convergent

$$1, 2, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{13}{8}, \frac{21}{13}, \frac{34}{21}, \frac{55}{34}, \frac{89}{55}.$$

The golden mean lies between last two fractions, which have the difference $\frac{1}{34\cdot 55}$. This allows us to determine the first four decimal digits after the decimal point of it and of $\sqrt{5}$.

Problems.

1. Evaluate $\sum\limits_{k=0}^{\infty}\frac{2^{2k}}{3^{3k}}$.
2. Evaluate $1-1+1-1+\dots$
3. Evaluate $1+1-1-1+1+1-1-1+\dots$
4. Evaluate $\sum\limits_{k=1}^{\infty}\frac{k}{3^k}$
5. Evaluate $\sum\limits_{k=1}^{\infty}\frac{k^2}{2^k}$

6. Decompose into power series the fraction $\frac{1}{a+x}$
7. Find the generating function for the sequence $\{2^k\}$.
8. Find sum the $\sum\limits_{k=1}^{\infty} \phi_k 3^{-k}$
9. Prove by induction the Euler–Binet formula.
10. * Evaluate $1 - 2 + 1 + 1 - 2 + 1 \dots$
11. Approximate $\sqrt{2}$ by a rational with precision 0.0001
12. Find the value of $1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \dots}}}$
13. Find the value of $\sqrt{2 + \sqrt{2 + \sqrt{2 + \dots}}}$
14. By infinite substitution, solve the equation $x^2 - 2x - 1 = 0$, and represent $\sqrt{2}$ by a continued fraction.
15. Find the value of infinite product
$$2 \cdot 2^{\frac{1}{2}} \cdot 2^{\frac{1}{4}} \cdot 2^{\frac{1}{8}} \dots$$
16. Find a formula for n-th term of recurrent sequence $x_{n+1} = 2x_n + x_{n-1}$, $x_0 = x_1 = 1$.
17. Find the sum of the Fibonacci numbers $\sum\limits_{k=1}^{\infty} \phi_k$
18. Find sum $1 + 0 - 1 + 1 + 0 - 1 \dots$
19. Decompose into the sum of partial fractions $\frac{1}{x^2-3x+2}$

1.2 Positive Series

On the contents of the lecture. Infinity is pregnant with paradoxes. Paradoxes throw us down from heavens to the earth. We leave the poetry for prose, and rationalize the "infinity and equality" by working with "finiteness and inequality". We shall lay a solid foundation for a summation theory for positive series. And the reader will find out what $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$ precisely means.

Divergent series paradox. Let us consider the series $\sum_{k=0}^{\infty} 2^k$. This is a geometric series. We know how to sum it up by autorecursion. The autorecursion equation is $s = 1 + 2s$. The only number satisfying this equation is -1. The sum of positive numbers turns to be negative!? Something is wrong!

A way to save the situation is to admit infinity as a feasible solution. Infinity is an obvious solution of $s = 1 + 2s$. The sum of any geometric series $\sum_{k=0}^{\infty} q^k$ with denominator $q \geq 1$ is obviously infinite, isn't it?

Indeed, this sum is greater than $1 + 1 + 1 + 1 + \dots$, which symbolizes the infinity. (The autorecursion equation for $1 + 1 + 1 + \dots$ is $s = s + 1$. Infinity is the unique solution of this equation.)

The series $\sum_{k=0}^{\infty} 2^k$ represents Zeno's series in the case of *Mighty Turtle,* which is faster than Achilles. To be precise, this series arises if $v = d_0 = 1$ and $w = 2$. As the velocity of the turtle is greater than the velocity of Achilles he never reaches it. So the infinity is right answer for this problem. But the negative solution -1 also makes sense. One could interpret it as an event in the past. Just the point in time when *the turtle passed Achilles.*

Oscillating series paradoxes. Philosopher Gvido Grandy in 1703 attracted public attention to the series $1-1+1-1+\dots$. He claimed this series symbolized the Creation of Universe from Nothing. Namely, insertion of brackets in one way gives Nothing (that is 0), in another way, gives 1.

$$(1-1) + (1-1) + (1-1) + \cdots = 0 + 0 + 0 + \cdots = 0$$
$$1 - (1-1) - (1-1) - (1-1) \cdots = 1 - 0 - 0 - 0 \cdots = 1.$$

On the other hand, this series $1 - 1 + 1 - 1 + 1 - 1 + \dots$ is geometric with negative ratio $q = -1$. Its autorecursion equation $s = 1 - s$ has the unique solution $s = \frac{1}{2}$. Neither $+\infty$ nor $-\infty$ satisfy it. So $\frac{1}{2}$ seems to be its true sum.

Hence we see the Associativity Law dethroned by $1 - 1 + 1 - 1 + \dots$. The next victim is the Commutativity Law. The sum $-1+1-1+1-1+\dots$ is equal to $-\frac{1}{2}$. But the last series is obtained from $1 - 1 + 1 - 1 + \dots$ by transpositions of odd and even terms.

And the third amazing thing: diluting it by zeros changes its sum. The sum $1 + 0 - 1 + 1 + 0 - 1 + 1 + 0 - 1 + \dots$ by no means is $\frac{1}{2}$. It is $\frac{2}{3}$. Indeed, if we denote this sum by s then by shift formulas one gets

$$\begin{aligned} s &= 1+0-1+1+0-1+1+0-1+1+0-1+\dots \\ s-1 &= 0-1+1+0-1+1+0-1+1+0-1+1+\dots \\ s-1-0 &= -1+1+0-1+1+0-1+1+0-1+1+0+\dots \end{aligned}$$

Summing the numbers column-wise (i.e., by the Term-wise Addition Formula), we get

$$\begin{aligned} s+(s-1)+(s-1-0) &= (1+0-1)+(0-1+1)+(-1+1+0)+ \\ &\quad + (1+0-1)+(0-1+1)+(-1+1+0)+\dots \end{aligned}$$

The left-hand side is $3s-2$. The right-hand side is the zero series. That is why $s=\frac{2}{3}$.

The series $1-1+1-1+\dots$ arises as Zeno's series in the case of a blind Achilles directed by a cruel Zeno, who is interested, as always, only in proving his claim, and a foolish, but merciful turtle. The blind Achilles is not fast, his velocity equals the velocity of the turtle. At the first moment Zeno tells the blind Achilles where the turtle is. Achilles starts the rally. But the merciful turtle wishing to help him goes towards him instead of running away. Achilles meets the turtle half-way. But he misses it, being busy to perform the first step of the algorithm. When he accomplishes this step, Zeno orders: "Turn about!" and surprises Achilles by saying that the turtle is on Achilles initial position. The turtle discovers that Achilles turns about and does the same. The situation repeats ad infinitum. Now we see that assigning the sum $\frac{1}{2}$ to the series $1-1+1-1+\dots$ makes sense. It predicts accurately the time of the first meeting of Achilles and turtle.

Positivity. The paradoxes discussed above are discouraging. Our intuition based on handling finite sums fails when we turn to infinite ones. Observe that all paradoxes above involve negative numbers. And to eliminate the evil in its root, let us consider only nonnegative numbers.

We return to the ancient Greeks. They simply did not know what a negative number is. But in contrast to Greeks, we will retain zero. A series with nonnegative terms will be called a *positive* series. We will show that for positive series all familiar laws, including associativity and commutativity, hold true and zero terms do not affect the sum.

Definition of Infinite Sum. Let us consider what Euler's equality (1.2.1) could mean:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \tag{1.2.1}$$

The natural answer is: the *partial sums* $\sum\limits_{k=1}^{n} \frac{1}{k^2}$, which contain more and more reciprocal squares, approach closer and closer the value $\frac{\pi^2}{6}$. Consequently, all partial sums have to be less than $\frac{\pi^2}{6}$, its *ultimate sum*. Indeed, if some partial sum exceeds or coincides with $\frac{\pi^2}{6}$, then all subsequent sums will move away from $\frac{\pi^2}{6}$. Furthermore, any number c which is less than $\frac{\pi^2}{6}$ has to be surpassed by partial sums eventually, when they approach $\frac{\pi^2}{6}$ closer than by $\frac{\pi^2}{6} - c$. Hence the ultimate sum majorizes all partial ones, and any lesser number does not. This means that the ultimate sum is the smallest number which majorizes all partial sums.

Geometric motivation. Imagine a sequence of intervals of real line $[a_{i-1}, a_i]$. Denote by l_i the length of i-th interval. Let $a_0 = 0$ be the left end point of the first interval. Let $[0, A]$ be the smallest interval containing the whole sequence. Its length is naturally interpreted as the sum $\sum_{k=1}^{\infty} l_i$

This motivates the following definition.

Definition. *If the partial sums of the positive series $\sum_{k=1}^{\infty} a_k$ increase without bounds, its sum is defined to be ∞ and the series is called* divergent. *In the opposite case the series called* convergent, *and its sum is defined as the smallest number A such that $A \geq \sum_{k=1}^{n} a_k$ for all n.*

This Definition in equivalent to the following couple of principles. The first principle limits the ultimate sum from below:

Principle (One-for-All). *The ultimate sum of a positive series majorizes all partial sums.*

And the second principle limits the ultimate sum from above:

Principle (All-for-One). *If all partial sums of a positive series do not exceed a number, then the ultimate sum also does not exceed it.*

Term-Wise Addition Formula: $\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} (a_k + b_k)$

Proof. The inequality $\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k \leq \sum_{k=1}^{\infty} (a_k + b_k)$. is equivalent to $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} (a_k + b_k) - \sum_{k=1}^{\infty} b_k$. By All-for-One, the last is equivalent to the system of inequalities

$$\sum_{k=1}^{N} a_k \leq \sum_{k=1}^{\infty} (a_k + b_k) - \sum_{k=1}^{\infty} b_k \qquad N = 1, 2, \dots$$

This system is equivalent to the following system

$$\sum_{k=1}^{\infty} b_k \leq \sum_{k=1}^{\infty} (a_k + b_k) - \sum_{k=1}^{N} a_k \qquad N = 1, 2, \dots.$$

Each inequality of the last system, in its turn, is equivalent to the system of inequalities

$$\sum_{k=1}^{M} b_k \leq \sum_{k=1}^{\infty} (a_k + b_k) - \sum_{k=1}^{N} a_k \qquad M = 1, 2, \dots.$$

But these inequalities are true for all N and M, as the following computations show.

$$\sum_{k=1}^{M} b_k + \sum_{k=1}^{N} a_k \leq \sum_{k=1}^{M+N} b_k + \sum_{k=1}^{M+N} a_k = \sum_{k=1}^{M+N} (a_k + b_k) \leq \sum_{k=1}^{\infty} (a_k + b_k)$$

In the opposite direction, we see that any partial sum on the right-hand side $\sum_{k=1}^{n} (a_k + b_k)$ splits into $\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$. And by virtue of the One-for-All principle, this does not exceed $\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$. Now, the All-for-One principle provides the inequality in the opposite direction. $\square$

Shift Formula. The Shift Formula immediately follows from the Term-Wise Addition. To be precise, immediately from the Definition, one gets the following: $a_0 + 0 + 0 + 0 + 0 + \cdots = a_0$ and that $0 + a_1 + a_2 + a_3 + \cdots = \sum_{k=1}^{\infty} a_k$. Term-Wise Addition of these series gives

$$a_0 + \sum_{k=1}^{\infty} a_k = (a_0 + 0) + (0 + a_1) + (0 + a_2) + (0 + a_3) + \cdots = \sum_{k=0}^{\infty} a_k$$

Term-Wise Multiplication Formula. $\lambda \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \lambda a_k$. For any partial sum from the right-hand side one has $\sum_{k=1}^{n} \lambda a_k = \lambda \sum_{k=1}^{n} a_k \leq \lambda \sum_{k=1}^{\infty} a_k$ by the Distributivity Law for finite sums and One-for-All. This implies the inequality $\lambda \sum_{k=1}^{\infty} a_k \geq \sum_{k=1}^{\infty} \lambda a_k$ by All-for-One. The opposite inequality is equivalent to $\sum_{k=1}^{\infty} a_k \geq \frac{1}{\lambda} \sum_{k=1}^{\infty} \lambda a_k$. As any partial sum $\sum_{k=1}^{n} a_k$ is equal to $\frac{1}{\lambda} \sum_{k=1}^{n} \lambda a_k$, which does not exceed $\frac{1}{\lambda} \sum_{k=1}^{\infty} \lambda a_k$, one gets the opposite inequality.

Geometric series. We have to return to the geometric series, because the Autorecursion equation produced by shift and multiplication formulas says nothing about convergence. So we have to prove convergence for $\sum_{k=0}^{\infty} q^k$ with positive $q < 1$. It is sufficient to prove the following inequality for all n

$$1 + q + q^2 + q^3 + \cdots + q^n < \frac{1}{1-q}.$$

Multiplying both sides by $1 - q$ one gets on the left-hand side

$$\begin{aligned}(1-q) + (q - q^2) + (q^2 - q^3) + \cdots + (q^{n-1} - q^n) + (q^n - q^{n+1}) = \\ 1 - q + q - q^2 + q^2 - q^3 + q^3 - \cdots - q^n + q^n - q^{n+1} = 1 - q^{n+1}\end{aligned}$$

and 1 on the right-hand side. The inequality $1 - q^{n+1} < 1$ is obvious. Hence we have proved the convergence. Now the autorecursion equation $x = 1 + qx$

for $\sum_{k=0}^{\infty} q^k$ is constructed in usual way by Shift and Term-Wise multiplication. It leaves only two possibilities for $\sum_{k=0}^{\infty} q^k$, either $\frac{1}{q-1}$ or ∞. For $q < 1$ we have proved convergence, and for $q \geq 1$ infinity is the true answer.

Let us pay special attention to the case $q = 0$. We adopt a common convention:

$$0^0 = 1. \tag{1.2.2}$$

This means that the series $\sum_{k=0}^{\infty} 0^k$ satisfies the common formula for a convergent geometric series $\sum_{k=0}^{\infty} 0^k = \frac{1}{1-0} = 1$. Finally we state the theorem, which essentially due to Eudoxus, who proved the convergence of the geometric series with ratio $q < 1$.

Theorem 1 (Eudoxus). *For every nonnegative q one has*

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad \text{for } q < 1, \quad \text{and} \quad \sum_{k=0}^{\infty} q^k = \infty \quad \text{for } q \geq 1.$$

Comparison of series Quite often exact summation of series is too difficult, and for practical purposes it is enough to know the sum approximatively. In this case one usually compares the series with another one which sum is known. Such comparison is based on the following *Term-Wise Comparison Principle*, which immediately follows from the Definition of Sum.

Principle (Term-Wise Comparison). *If $a_k \leq b_k$ for k, then $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$*

The only series we have so far to compare with are the geometric ones. The following lemma is very useful for this purposes.

Lemma 2 (Ratio Test). *If $a_{k+1} \leq qa_k$ holds for some $q < 1$ and all terms of series $\sum_{k=0}^{\infty} a_k$ than $\sum_{k=0}^{\infty} a_k \leq \frac{a_0}{1-q}$*

Proof. By induction one proves inequality $a_k \leq a_0 q^k$. Now by Term-Wise Comparison one estimates $\sum_{k=0}^{\infty} a_k$ from above by the geometric series $\sum_{k=0}^{\infty} a_0 q^k = \frac{a_0}{1-q}$ □

If the series under consideration satisfies an autorecursion equation, to prove its convergence usually mean to evaluate it exactly. For proving of convergence the Term-Wise Comparison Principle can be strengthened. Let us say that the series $\sum_{k=1}^{\infty} a_k$ is *eventually* majorized by the series $\sum_{k=1}^{\infty} b_k$, if the inequality $b_k \geq a_k$ holds for each k starting from $k = n$ for some n. The following lemma is very useful to prove convergence.

Principle (Eventual Comparison). *A series $\sum_{k=1}^{\infty} a_k$, which is eventually majorized by a convergent series $\sum_{k=1}^{\infty} b_k$, is convergent.*

Proof. Consider a tail $\sum_{k=n}^{\infty} b_k$, which term-wise majorizes $\sum_{k=n}^{\infty} a_k$. Then $\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{n-1} a_k + \sum_{k=n}^{\infty} a_k \leq \sum_{k=1}^{n-1} a_k + \sum_{k=n}^{\infty} b_k \leq \sum_{k=1}^{n-1} a_k + \sum_{k=1}^{\infty} b_k < \infty$. $\square$

Example Consider series $\sum_{k=1}^{\infty} k2^{-k}$. The ratio of two successive terms $\frac{a_{k+1}}{a_k}$ of the series is $\frac{k+1}{2k}$. This ratio is less or equal to $\frac{2}{3}$ starting with $k = 3$. Hence this series is eventually majorities by geometric one $\sum_{k=0}^{\infty} a_3 \frac{2^k}{3^k}$, ($a_3 = \frac{2}{3}$). This proves its convergence. And now by autorecursion equation one gets its sum.

Harmonic series paradox. Now we have a solid background to evaluate positive series. Nevertheless, we must be careful about infinity! Consider the following calculation:

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

We get that the sum of $\sum_{k=1}^{\infty} \frac{1}{(2k-1)2k}$ satisfies the equation $s = \frac{s}{2}$. This equation has two roots 0 and ∞. But s satisfies inequalities $\frac{1}{2} < s < \frac{\pi^2}{6}$. What is wrong?

Problems.

1. Prove $\sum_{k=1}^{\infty} 0 = 0$
2. Prove $\sum_{k=1}^{\infty} 0^k = 1$
3. Prove $\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{2k} + \sum_{k=0}^{\infty} a_{2k+1}$
4. Prove $\sum_{k=1}^{\infty} (a_k - b_k) = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} b_k$ for convergent series
5. Evaluate $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$
6. Prove $(1 - \frac{1}{2}) + (\frac{1}{3} - \frac{1}{4}) + (\frac{1}{5} - \frac{1}{6}) + \cdots = 1 - [(\frac{1}{2} - \frac{1}{3}) + (\frac{1}{4} - \frac{1}{5}) + \dots$

7. Prove the convergence of $\sum\limits_{k=0}^{\infty} \frac{2^k}{k!}$

8. Prove the convergence of $\sum\limits_{k=1}^{\infty} \frac{1000^k}{k!}$

9. Prove the convergence of $\sum\limits_{k=1}^{\infty} \frac{k^{1000}}{2^k}$

10. Prove that $q^n < \frac{1}{n(1-q)}$ for $0 < q < 1$

11. Prove that for any positive $q < 1$ there is an n that $q^n < \frac{1}{2}$

12. Prove $\sum\limits_{k=1}^{\infty} \frac{1}{k!} \leq 2$

13. Evaluate $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+2)}$

14. Prove the convergence of the Euler series $\sum\limits_{k=1}^{\infty} \frac{1}{k^2}$

15. * Prove that $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} a_{ij} = \sum\limits_{j=1}^{\infty} \sum\limits_{i=1}^{\infty} a_{ij}$ for $a_{ij} \geq 0$

1.3 Unordered Sums

On the contents of the lecture Our summation theory culminates in the Sum Partition Theorem. This lecture will contribute towards evaluation of the Euler series in two ways: we prove its convergence, and even estimate its sum by 2. On the other hand, we will realize that evaluation of the Euler series with Euler's accuracy (10^{-18}) seems to be beyond human being's strength.

Consider a family $\{a_i\}_{i\in I}$ of nonnegative numbers indexed by elements of an arbitrary set I. An important special case of I is the set of pairs of natural numbers $\mathbb{N}\times\mathbb{N}$. Families indexed by $\mathbb{N}\times\mathbb{N}$ are called double series. They arise when one multiply one series by another one.

Any sum of the type $\sum\limits_{i\in K} a_i$, where K is a finite subset of I is called a *subsum* of $\{a_i\}_{i\in I}$ over K.

Definition. *The least number majorizing all subsums of $\{a_i\}_{i\in I}$ over finite subsets is called its (ultimate) sum and denoted by* $\sum\limits_{i\in I} a_i$

One-for-All and All-for-One principles for non-ordered sum are obtained from corresponding principles for ordered sum by replacing "partial sums" by "finite subsums".

Commutativity. In case $I=\mathbb{N}$ we have got a definition which apparently is new. But fortunately this definition is equivalent to the old one. Indeed, as any finite subsum of positive series does not exceed its ultimate (ordered) sum, the non-ordered sum also does not exceed it. On the other hand, any partial sum of the series is a finite subsum. This implies the opposite inequality. Therefore we have established the equality.

$$\sum_{k=1}^{\infty} a_k = \sum_{k\in\mathbb{N}} a_k$$

This means that positive series obeys the Commutativity Law. Because the non-ordered sum obviously does not depend on order of summands.

Partitions. A family of subsets $\{I_k\}_{k\in K}$ of a set I is called a *partition* of I and is written $\bigsqcup_{k\in K} I_k$ if $I=\bigcup_{k\in K} I_k$ and $I_k\cap I_j=\varnothing$ for all $k\neq j$.

Theorem 1 (Sum Partition Theorem). *For any partition $I=\bigsqcup_{j\in J} I_j$ of the indexing set and any family $\{a_i\}_{i\in I}$ of nonnegative numbers,*

$$\sum_{i\in I} a_i = \sum_{j\in J}\sum_{i\in I_j} a_i. \tag{1.3.1}$$

Iverson notation We will apply the following notation: a statement included into [] takes value 1, if the statement is true, and 0, if it is false. Prove the following simple lemmas to adjust to this notation. In both lemmas one has $K\subset I$.

Lemma 2. $\sum\limits_{i\in K} a_i = \sum\limits_{i\in I} a_i[i\in K]$

In particular, for $K = I$, Lemma **2** turns into

Lemma 3. $\sum\limits_{i\in I} a_i = \sum\limits_{i\in I} a_i[i \in I]$

Lemma 4. $\sum\limits_{k\in K} [i \in I_k] = [i \in I_K]$ *for all* $i \in I$ *iff* $I_K = \bigsqcup_{k\in K} I_k$

Proof of Sum Partition Theorem. At first we prove the following Sum Transposition Formula for finite J

$$\sum_{i\in I}\sum_{j\in J} a_{ij} = \sum_{j\in J}\sum_{i\in I} a_{ij} \tag{1.3.2}$$

Indeed, if J contains just two elements, this formula turns into Term-Wise Addition Formula. The proof of this formula is the same as for series. Suppose the formula is proved for any set, which contains less elements than J does. Decompose J into union of two nonempty subsets $J_1 \sqcup J_2$. Then applying only Term-Wise Addition and Lemmas **2**,**3** **4**, we get

$$\sum_{i\in I}\sum_{j\in J} a_{ij} = \sum_{i\in I}\sum_{j\in J} a_{ij}[j \in J] = \sum_{i\in I}\sum_{j\in J}(a_{ij}[j \in J_1] + a_{ij}[j \in J_2]) =$$
$$= \sum_{i\in I}\sum_{j\in J} a_{ij}[j \in J_1] + \sum_{i\in I}\sum_{j\in J} a_{ij}[j \in J_2] = \sum_{i\in I}\sum_{j\in J_1} a_{ij} + \sum_{i\in I}\sum_{j\in J_2} a_{ij}.$$

But the last two sums can be transposed by induction hypothesis. After such transposition one gets

$$\sum_{j\in J_1}\sum_{i\in I} a_{ij} + \sum_{j\in J_2}\sum_{i\in I} a_{ij} = \sum_{j\in J}[j \in J_1]\sum_{i\in I} a_{ij} + \sum_{j\in J}[j \in J_2]\sum_{i\in I} a_{ij} =$$
$$= \sum_{j\in J}([j \in J_1] + [j \in J_2])\sum_{i\in I} a_{ij} = \sum_{j\in J}[j \in J]\sum_{i\in I} a_{ij} = \sum_{j\in J}\sum_{i\in I} a_{ij}$$

and the Sum Transposition Formula for finite J is proved. Consider the general case. To prove $\leq$ in (1.3.2), consider a finite $K \subset I$. By the Finite Transposition Formula the subsum $\sum\limits_{i\in K}\sum\limits_{j\in J} a_{ij}$ is equal to $\sum\limits_{j\in J}\sum\limits_{i\in K} a_{ij}$. But this sum is term-wise majorized by the right-hand side sum in (1.3.2). Therefore the left-hand side does not exceed the right-hand side by All-for-One principle.

To derive the Sum Partition Theorem from the Transposition formula, pose $a_{ij} = a_i[i \in I_j]$. Then $a_i = \sum\limits_{j\in J} a_{ij}$ and (1.3.1) turns into (1.3.2). This completes the proof of the Sum Partition Theorem.

Blocking For a given a series $\sum\limits_{k=0}^{\infty} a_k$ and an increasing sequence of natural numbers $\{n_k\}_{k=0}^{\infty}$ starting with $n_0 = 0$ one defines a new series $\sum\limits_{k=0}^{\infty} A_k$ by the rule $A_k = \sum\limits_{i=n_k}^{n_{k+1}-1} a_i$. The series $\sum\limits_{k=0}^{\infty} A_k$ is called *blocking of* $\sum\limits_{k=0}^{\infty} a_k$ *by* $\{n_k\}$.

The Sum Partition Theorem implies that the sums of blocked and unblocked series coincide. Blocking formalizes putting of brackets. Therefore the Sum Partition Theorem implies *Sequential Associativity Law*: *Placing brackets does not change the sum of series.*

Estimation of the Euler series Let us compare the Euler series with $\sum_{k=0}^{\infty} \frac{1}{2^k}$, blocked by $\{2^n\}$ to $\sum_{k=1}^{\infty} a_k$. The sum $\sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k^2}$ consists of 2^n summands, all of which are less then the first one, which is $\frac{1}{2^{2n}}$. As $2^n \frac{1}{2^{2n}} = \frac{1}{2^n}$, it follows that $a_n \leq \frac{1}{2^n}$ for each n. Summing these inequalities, one gets $\sum_{k=1}^{\infty} a_k \leq 2$.

Now let us estimate how many terms of Euler's series one needs to take into account to find its sum up to the eighteenth digit. To do this, we need to estimate its tail. The arguments above give $\sum_{k=2^n}^{\infty} \frac{1}{k^2} \leq \frac{1}{2^{n-1}}$. To obtain a lower estimation, let us remark that all terms of sum $\sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k^2}$ exceed $\frac{1}{2^{2(n+1)}}$. As the number of summands is 2^n, one gets $a_n \geq \frac{1}{4 \cdot 2^n}$. Hence $\sum_{k=2^n}^{\infty} \frac{1}{k^2} \geq \frac{1}{2^{n+1}}$. Since $2^{10} = 1024 \simeq 10^3$, one gets $2^{60} \simeq 10^{18}$. So, to provide accuracy 10^{-18} one needs to sum approximately 10^{18} terms. This task is inaccessible even for a modern computer. How did Euler manage to do this? He invented a summation formula (Euler-MacLaurin formula) and transformed this slowly convergent series into non-positive divergent (!) one, whose partial sum containing as little as ten terms gave 18-digit accuracy. The whole calculation took him an evening. To introduce this formula, one needs to know integrals and derivatives. We will do this later.

Problems.

1. Find $\sum_{k=1}^{\infty} \frac{1}{(2k)^2}$ and $\sum_{k=1}^{\infty} \frac{1}{(2k-1)^2}$, assuming $\sum_{k=1}^{\infty} \frac{1}{k^2} = \pi^2/6$.

2. Prove convergence of $\sum_{k=1}^{\infty} \frac{1}{k\sqrt{k}}$.

3. Estimate how many terms of series $\sum_{k=1}^{\infty} \frac{1}{n^3}$ is necessary to sum for calculation of its sum with precision 10^{-3}.

4. Estimate the value of $\sum_{k=1}^{\infty} \frac{1}{2^n} \frac{1}{k}$.

5. Prove equality $\sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k = \sum_{j,k \in \mathbb{N}} a_j b_k$.

6. Estimate how many terms of Harmonic series gives the sum surpassing 1000

7. Prove the Dirichlet formula

$$\sum_{k=1}^{n-1} \sum_{i=1}^{k+1} a_{ki} = \sum_{i=1}^{n} \sum_{k=i}^{n-1} a_{ki}.$$

8. Evaluate $\sum\limits_{i,j \in N} \dfrac{1}{2^i 3^j}$.
9. Evaluate $\sum\limits_{i,j \in N} \dfrac{i+j}{2^i 3^j}$.
10. Represent an unordered sum $\sum\limits_{i+j<n} a_{ij}$ as a double sum.
11. Evaluate $\sum\limits_{i,j \in N} \dfrac{ij}{2^i 3^j}$.
12. Change the summation order in $\sum\limits_{i=0}^{\infty} \sum\limits_{j=0}^{2i} a_{ij}$.
13. Define by Iverson notation the following function:
    - (integral part) $[x]$
    - $|x|$ (module)
    - $\operatorname{sgn} x$ (signum)
    - $n!$ (factorial)
14. Define only by formulas the expression $[p \text{ is prime}]$

1.4 Infinite Products

On the contents of the lecture In this lecture we become acquainted with infinite products. The famous Euler Identity will be proved. We will find out that $\zeta(2)$ is another name for the Euler series. And we will see how the Euler's decomposition of the sine into product works to sum up the Euler Series.

Definition. *Product of an infinite sequence of numbers $\{a_k\}$, such that $a_k \geq 1$ for all k is defined as the least number majorizing all partial products $\prod_{k=1}^{n} a_k = a_1 a_2 \dots a_n$*

A sequence of natural numbers is called *essentially finite* if all but finitely many of its elements are equal to zero. Denote by $\mathbb{N}^\infty$ the set of all essentially finite sequences of natural numbers.

Theorem 1. *For any given sequence of positive series $\sum\limits_{k=0}^{\infty} a_k^j$, $j = 1, 2, \dots$, such that $a_0^j = 1$ for all j one has*

$$\prod_{j=1}^{\infty} \sum_{k=0}^{\infty} a_k^j = \sum_{\{k_j\} \in \mathbb{N}^\infty} \prod_{j=1}^{\infty} a_{k_j}^j \tag{1.4.1}$$

Summands on the right-hand side of (1.4.1) usually contain factors which are less than one. But each of the summands contains only finitely many factors different from 1. So the summands are in fact finite products.

Proof. For a sequence $\{k_j\} \in \mathbb{N}^\infty$ define its *length* as maximal j for which $k_j \neq 0$ and its *maximum* as the value of its maximal term. The length of zero sequence is defined as 0.

Consider a finite subset $S \subset \mathbb{N}^\infty$. Consider the partial sum $\sum\limits_{\{k_j\} \in S} \prod_{k=1}^{\infty} a_{k_j}^j$. To estimate it, denote by L the maximal length of elements of S and denote by M the greatest of maxima of $\{k_j\} \in S$. In this case

$$\sum_{\{k_j\} \in S} \prod_{j=1}^{\infty} a_{k_j}^j = \sum_{\{k_j\} \in S} \prod_{j=1}^{L} a_{k_j}^j \leq \sum_{\{k_j\} \in \mathbb{N}_M^L} \prod_{j=1}^{L} a_{k_j}^j = \prod_{j=1}^{L} \sum_{k=0}^{M} a_k^j \leq \prod_{j=1}^{\infty} \sum_{k=0}^{\infty} a_k^j,$$

where $\mathbb{N}_M^L$ denotes the set of all finite sequences $\{k_1, k_2, \dots k_L\}$ of natural numbers such that $k_i \leq M$. By All-for-One this implies one of the required inequalities, namely, $\geq$.

To prove the opposite inequality, we prove that for any natural L one has

$$\prod_{j=1}^{L} \sum_{k=0}^{\infty} a_k^j = \sum_{\{k_j\} \in \mathbb{N}^L} \prod_{j=1}^{L} a_{k_j}^j, \tag{1.4.2}$$

where $\mathbb{N}^L$ denotes the set of all finite sequences $\{k_1, \dots k_L\}$ of natural numbers. The proof is by induction on L.

Lemma 2. *For any families $\{a_i\}_{i \in I}$ $\{b_j\}_{j \in J}$ of nonnegative numbers, one has*

$$\sum_{i \in I} a_i \sum_{j \in J} b_j = \sum_{(i,j) \in I \times J} a_i b_j \tag{1.4.3}$$

Proof. Since $I \times J = \bigsqcup_{i \in I} \{i\} \times J$ by the Sum Partition Theorem one gets:

(1.4.4)

$$\sum_{(i,j) \in I \times J} a_i b_j = \sum_{i \in I} \sum_{(i,j) \in \{i\} \times J} a_i b_j = \sum_{i \in I} \sum_{j \in J} a_i b_j = \sum_{i \in I} a_i \sum_{j \in J} b_j = \sum_{j \in J} b_j \sum_{i \in I} a_i$$

□

Case $L = 2$ follows from Lemma **2**, because $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$. The induction step is done as follows

$$\prod_{j=1}^{L+1} \sum_{k=0}^{\infty} a_k^j = \sum_{k=0}^{\infty} a_k^{L+1} \prod_{j=1}^{L} \sum_{k=0}^{\infty} a_k^j = \sum_{k \in \mathbb{N}} a_k^{L+1} \sum_{\{k_j\} \in \mathbb{N}^L} \prod_{j=1}^{L} a_{k_j}^j = \sum_{\{k_j\} \in \mathbb{N}^{L+1}} \prod_{j=1}^{L+1} a_{k_j}^j$$

The left-hand side of (1.4.2) is a partial product for the left-hand side of (**1**) and the right-hand side of (1.4.2) is a subsum of the right-hand side of (**1**). Consequently, all partial products of the right-hand side in (**1**) do not exceed its left-hand side. This proves the inequality $\leq$.

□

Euler's Identity Our next goal is to prove the *Euler Identity*

(1.4.5)
$$\boxed{\sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \prod_{p=1}^{\infty} \left(1 - \frac{1}{p^\alpha}\right)^{-[p \text{ is prime}]}}$$

Here α is any rational (or even irrational) positive number.

The product on the right-hand side is called the *Euler Product.* The series on the left-hand side is called *Dirichlet series.* Each factor of the Euler Product expands into geometric series $\sum_{k=0}^{\infty} \frac{1}{p^{k\alpha}}$. By Theorem 1.4.1 the product of these geometric series is equal to the sum of products of the type $p_1^{-k_1\alpha} p_2^{-k_2\alpha} \dots p_n^{-k_n\alpha} = N^{-\alpha}$. Here $\{p_i\}$ are different prime numbers, $\{k_i\}$ are positive natural numbers and $p_1^{k_1} p_2^{k_2} \dots p_n^{k_n} = N$. But each product $p_1^{k_1} p_2^{k_2} \dots p_n^{k_n} = N$ is a natural number, different products represent different numbers and any natural number has a unique representation of this sort. This is exactly what is called Principal Theorem of Arithmetics. That is the above decomposition of the Euler product expands in Dirichlet series.

Convergence of Dirichlet series

Theorem 3. *The Dirichlet series* $\sum_{k=1}^{\infty} \frac{1}{n^s}$ *converges iff* $s > 1$

Proof. Consider $\{2_k\}$ packing of the series. Then the n-th term of the packed series one estimates from the up as $\sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k^s} \leq \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{(2^n)^s} = 2^n \frac{1}{2^{ns}} = 2^{n-ns} = (2^{1-s})^n$. If $s > 1$ then $2^{1-s} < 1$ and the packed series is term-wise majorized by convergent geometric progression. Hence it converges. In case the of the Harmonic series ($s = 1$) the n-th term of its packing one estimates from below as $\sum\limits_{k=2^n}^{2^{n+1}-1} \frac{1}{k} \geq \sum\limits_{k=2^n}^{2^{n+1}-1} \frac{1}{2^{n+1}} = 2^n \frac{1}{2^{n+1}} = \frac{1}{2}$. That is why the harmonic series diverges. A Dirichlet series for $s < 1$ term-wise majorizes the Harmonic series and so diverges. $\square$

Riemann ζ-function Function $\zeta(s) = \sum\limits_{k=1}^{\infty} \frac{1}{n^s}$ is called *Riemann ζ-function.* It is of great importance in the number theory.

The simplest application of the Euler's Identity represents Euler's proof of infinity of the set of primes. The divergence of the *harmonic series* $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ implies the Euler Product has to contain infinitely many factors to diverge.

Euler proved essentially more exact result: the series of reciprocal primes diverges $\sum \frac{1}{p} = \infty$.

Summing via multiplication Multiplication of series gives rise to a new approach to evaluating their sums. Consider the geometric series $\sum\limits_{k=0}^{\infty} x^k$. Then

$$\left(\sum_{k=0}^{\infty} x^k\right)^2 = \sum_{j,k\in\mathbb{N}^2} x^j x^k = \sum_{m=0}^{\infty} \sum_{j+k=m} x^j x^k = \sum_{m=0}^{\infty} (k+1)x^m$$

As $\sum\limits_{k=0}^{\infty} x^k = \frac{1}{1-x}$ one gets $\sum\limits_{k=0}^{\infty} (k+1)x^k = \frac{1}{(1-x)^2}$.

Sine-product Now we are ready to understand how two formulas

$$\text{(1.4.6)} \qquad \frac{\sin x}{x} = \prod_{k=1}^{\infty}\left(1 - \frac{x^2}{k^2\pi^2}\right) \qquad \sin x = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(}2k+1)!$$

which appeared in Legends, yield an evaluation of the Euler Series. Since at the moment we do not know how to multiply infinite sequences of numbers which are less than one, we invert the product in the first formula. We get

$$\text{(1.4.7)} \qquad \frac{x}{\sin x} = \prod_{k=1}^{\infty}\left(1 - \frac{x^2}{k^2\pi^2}\right)^{-1} = \prod_{k=1}^{\infty}\sum_{j=0}^{\infty} \frac{x^{2j}}{k^{2j}\pi^{2j}}$$

To avoid negative numbers, we interpret the series $\sum\limits_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ in the second of formulas (1.4.6) as the difference $\sum\limits_{k=0}^{\infty} \frac{x^{4k+1}}{(4k+1)!} - \sum\limits_{k=0}^{\infty} \frac{x^{4k+3}}{(4k+3)!}$. Substituting this expression for $\sin x$ in $\frac{x}{\sin x}$ and cancelling out x, we get

$$\text{(1.4.8)} \qquad \frac{x}{\sin x} = \frac{1}{1 - \sum\limits_{k=1}^{\infty} (-1)^{k+1} \frac{x^{2k}}{(2k+1)!}} = \sum_{j=0}^{\infty}\left(\sum_{k=1}^{\infty}(-1)^{k+1} \frac{x^{2k}}{(2k+1)!}\right)^j$$

All terms on the right-hand side starting with $j = 2$ are divisible by x^4. Consequently the only summands with x^2 on the right-hand side is $\frac{x^2}{6}$. On the other hand in (1.4.7) after expansion into sum by Theorem 1.4.1 terms with x^2 give the series $\sum\limits_{k=1}^{\infty} \frac{x^2}{k^2\pi^2}$. Comparing these results, one gets $\sum\limits_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Problems.

1. Prove $\prod_{n=1}^{\infty} 1.1 = \infty$
2. Prove identity $\prod_{n=1}^{\infty} a_n^2 = (\prod_{n=1}^{\infty} a_n)^2$ $(a_n \geq 1)$
3. Does $\prod_{n=1}^{\infty}(1 + \frac{1}{n})$ converge?
4. Evaluate $\prod_{n=2}^{\infty} \frac{n^2}{n^2-1}$
5. Prove divergence $\prod\limits_{1}^{\infty}(1 + \frac{1}{k})^{[k \text{ is prime}]}$
6. Evaluate $\prod_{n=3}^{\infty} \frac{n(n+1)}{(n-2)(n+3)}$
7. Evaluate $\prod_{n=2}^{\infty} \frac{n^2-1}{n^2-4}$
8. Evaluate $\prod_{n=1}^{\infty}(1 + \frac{1}{n(n+2)})$
9. Evaluate $\prod_{n=1}^{\infty} \frac{(2n+1)(2n+7)}{(2n+3)(2n+5)})$
10. Evaluate $\prod_{n=2}^{\infty} \frac{n^3+1}{n^3-1}$
11. Prove inequality $\prod_{k=2}^{\infty}(1 + \frac{1}{k^2}) \geq \sum\limits_{k=2}^{\infty} \frac{1}{k^2}$
12. Prove convergence of the Wallis product $\prod \frac{4k^2}{4k^2-1}$
13. Evaluate $\sum\limits_{k=1}^{\infty} \frac{1}{k^4}$ applying (1.4.6).
14. Prove $\prod_{n=2}^{\infty} \frac{n^2+1}{n^2} < \infty$
15. Multiply geometric series by itself and get a power series expansion for $(1-x)^{-2}$
16. Define $\tau(n)$ as the number of divisors of n, then $\zeta^2(x) = \sum\limits_{k=1}^{\infty} \dfrac{\tau(n)}{n^x}$
17. Define $\phi(n)$ as the number of numbers which are less than n are relatively prime to n, then $\dfrac{\zeta(x-1)}{\zeta(x)} = \sum\limits_{k=1}^{\infty} \dfrac{\phi(n)}{n^x}$
18. Define $\mu(n)$ (Möbius function) as follows: $\mu(1) = 1$, $\mu(n) = 0$, if n is divisible by the square of a prime number, $\mu(n) = (-1)^k$, if n is the product of k different prime numbers. Then $\dfrac{1}{\zeta(x)} = \sum\limits_{k=1}^{\infty} \dfrac{\mu(n)}{n^x}$

19. * Prove $\sum_{k=1}^{\infty} \frac{[k \text{ is prime}]}{k} = \infty$

20. * Prove identity $\prod_{n=0}^{\infty}(1 + x^{2^n}) = \frac{1}{1-x}$

1.5 Telescopic sums

On the content of this lecture. In this lecture we learn the main secret of the elementary summation theory. We will evaluate series via its partial sums. We introduce *factorial powers*, which are easy to sum. Following Stirling we expand $\frac{1}{1+x^2}$ into a series of negative factorial powers and apply this expansion to evaluate the Euler series with Stirling's accuracy 10^{-8}.

The series $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)}$. In the first lecture we have calculated infinite sums directly without invoking partial sums. Now we present a dual approach to summing series. According this approach at first one finds a formula for the n-th partial sum and then substitutes in this formula infinity instead of n. The series $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)}$ gives a simple example for this method. A key to sum it up is the following identity

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

Because of this identity $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)}$ turns into the sum of differences

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{n}-\frac{1}{n+1}\right)+\dots \tag{1.5.1}$$

Its nth partial sum is equal to $1-\frac{1}{n+1}$. Substituting in this formula $n=+\infty$, one gets 1 as its ultimate sum.

Telescopic sums. The sum (1.5.1) represents a *telescopic sum*. This name is used for sums of the form $\sum\limits_{k=0}^{n}(a_k - a_{k+1})$. The value of such a telescopic sum is determined by the values of the first and the last of a_k, similarly to a telescope, whose thickness is determined by the radii of the external and internal rings. Indeed,

$$\sum_{k=0}^{n}(a_k - a_{k+1}) = \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_{k+1} = a_0 + \sum_{k=1}^{n} a_k - \sum_{k=0}^{n-1} a_{k+1} - a_{n+1} = a_0 - a_{n+1}.$$

The same arguments for infinite telescopic sum give

$$\sum_{k=0}^{\infty}(a_k - a_{k+1}) = a_0. \tag{1.5.2}$$

But this proof works only if $\sum\limits_{k=0}^{\infty} a_k < \infty$. This is untrue for $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)}$, owing to divergence of the Harmonic series. But the equality 1.5.2 holds also if a_k tends to 0 as k tends to infinity. Indeed, in this case a_0 is the least number majorizing all $a_0 - a_n$, the nth partial sums of $\sum\limits_{k=0}^{\infty} a_k$.

Differences. For a given sequence $\{a_k\}$ one denotes by $\{\Delta a_k\}$ the sequence of differences $\Delta a_k = a_{k+1} - a_k$ and calls latter sequence as *difference* of $\{a_k\}$. The main formula of elementary summation theory is

$$(1.5.3) \qquad \boxed{\sum_{k=0}^{n-1} \Delta a_k = a_n - a_0}$$

To telescope a series $\sum_{k=0}^{\infty} a_k$ it is sufficient to find a sequence $\{A_k\}$ such that $\Delta A_k = a_k$. On the other hand the sequence of sums $A_n = \sum_{k=0}^{n-1} a_k$ has difference $\Delta A_n = a_n$. Therefore, we see that to telescope a sum is equivalent to find a formula for partial sums. This lead to concept of *telescopic function.* For a function $f(x)$ we introduce its difference $\Delta f(x)$ as $f(x+1) - f(x)$. A function $f(x)$ telescopes $\sum a_k$, if $\Delta f(k) = a_k$ for all k.

Often the sequence $\{a_k\}$ that we would like to telescope has the form $a_k = f(k)$ for some function. Then we are searching for a *telescopic function* $F(x)$ for $f(x)$, i.e. a function such that $\Delta F(x) = f(x)$.

To evaluate the difference of a function is usually much easier than to telescope it. For this reason one has evaluated the differences of all basic functions and organized a *table of differences.* In order to telescope a given function, look in this table to find a table function whose difference coincides with or is close to given function.

For example, the differences of x^n for $n \leq 3$ are $\Delta x = 1$, $\Delta x^2 = 2x + 1$, $\Delta x^3 = 3x^2 + 3x + 1$. To telescope $\sum_{k=1}^{\infty} k^2$ we choose in this table x^3. Then $\frac{\Delta x^3}{3} - x^2 = x + \frac{1}{3} = \frac{\Delta x^2}{2} - \Delta \frac{x}{6}$. Therefore, $x^2 = \Delta\left(\frac{x^3}{3} - \frac{x^2}{2} + \frac{x}{6}\right)$. This immediately implies the following formula for sums of squares:

$$(1.5.4) \qquad \sum_{k=1}^{n-1} k^2 = \frac{2n^3 - 3n^2 + n}{6}$$

Factorial powers The usual powers x^n have a complicated differences. So-called *factorial powers* $x^{\underline{k}}$ have simpler differences. For any number x and any natural number k, let $x^{\underline{k}}$ denote $x(x-1)(x-2)\dots(x-k+1)$, and by $x^{\underline{-k}}$ we denote $\frac{1}{(x+1)(x+2)\dots(x+k)}$. At last we define $x^{\underline{0}} = 1$. The factorial power satisfies the following *addition law*

$$(1.5.5) \qquad \boxed{x^{\underline{k+m}} = x^{\underline{k}}(x-k)^{\underline{m}}}$$

We leave to the reader to check this rule for all integers m, k. The power $n^{\underline{n}}$ for a natural n coincides with the factorial $n! = 1 \cdot 2 \cdot 3 \cdots n$. The main property of factorial powers is given by

$$(1.5.6) \qquad \boxed{\Delta x^{\underline{n}} = n x^{\underline{n-1}}}$$

The proof is straightforward:

$$(x+1)^{\underline{k}} - x^{\underline{k}} = (x+1)^{\underline{1+(k-1)}} - x^{\underline{(k-1)+1}} = (x+1)x^{\underline{k-1}} - x^{\underline{k-1}}(x-k+1) = kx^{\underline{k-1}}$$

Applying this formula one can easily telescope any *factorial polynomial,* i.e. an expression of the form $a_0 + a_1 x^{\underline{1}} + a_2 x^{\underline{2}} + a_3 x^{\underline{3}} + \dots a_n x^{\underline{n}}$. Indeed, the explicit formula for the telescoping function is $a_0 x^{\underline{1}} + \frac{a_1}{2} x^{\underline{2}} + \frac{a_2}{3} x^{\underline{3}} + \frac{a_3}{4} x^{\underline{4}} + \dots \frac{a_n}{n+1} x^{\underline{n+1}}$. Therefore, another strategy to telescope x^k is to represent it as a factorial polynomial.

For example, to represent x^2 as factorial polynomial, consider $a + bx + cx^{\underline{2}}$, a general factorial polynomial of degree 2. We are looking for $x^2 = a + bx + cx^{\underline{2}}$. Substituting $x = 0$ in this equality one gets $a = 0$. Substituting $x = 1$, one gets $1 = b$, and finally for $x = 2$ one has $4 = 2 + 2c$. Hence $c = 1$. As result $x^2 = x + x^{\underline{2}}$. And the telescoping function is given by $\frac{x^{\underline{2}}}{2} + \frac{x^{\underline{3}}}{3} = \frac{1}{2}(x^2 - x) + \frac{1}{3}(x(x^2 - 3x + 2)) = \frac{1}{6}(2x^3 - 3x^2 + x)$. And we have once again proved the formula (1.5.4).

Stirling Estimation of the Euler series. We will expand $\frac{1}{(1+x)^2}$ into a series of negative factorial powers in order to telescope it. A natural first approximation to $\frac{1}{(1+x^2)}$ is $x^{\underline{-2}} = \frac{1}{(x+1)(x+2)}$. We represent $\frac{1}{(1+x)^2}$ as $x^{\underline{-2}} + R_1(x)$, where $R_1(x) = \frac{1}{(1+x)^2} - x^{\underline{-2}} = \frac{1}{(x+1)^2(x+2)}$. The remainder $R_1(x)$ is in a natural way approximated by $x^{\underline{-3}}$. If $R_1(x) = x^{\underline{-3}} + R_2(x)$ then $R_2(x) = \frac{2}{(x+1)^2(x+2)(x+3)}$. Further, $R_2(x) = 2x^{\underline{-4}} + R_3(x)$, where $R_3(x) = \frac{2\cdot 3}{(x+1)^2(x+2)(x+3)(x+4)} = \frac{3!x^{\underline{-4}}}{x+1}$. The above calculations led to conjecture

$$\frac{1}{(1+x)^2} = \sum_{k=0}^{n-1} k! x^{\underline{-k-2}} + \frac{n! x^{\underline{-n-1}}}{x+1}. \tag{1.5.7}$$

This conjecture is easily proved by induction. The remainder $R_n(x) = \frac{n!x^{\underline{-n-1}}}{x+1}$ represents the difference $\frac{1}{(1+x)^2} - \sum_{k=0}^{n-1} k! x^{\underline{-2-k}}$. Owing to the inequality $x^{\underline{-1-n}} \leq \frac{1}{(n+1)!}$, which is valid for all $x \geq 0$, the remainder decreases to 0 as n increases to infinity. This implies

Theorem 1. *For all $x \geq 0$ one has*

$$\frac{1}{(1+x)^2} = \sum_{k=0}^{\infty} k! x^{\underline{-2-k}} \tag{1.5.8}$$

To calculate $\sum_{k=p}^{\infty} \frac{1}{(1+k)^2}$, replace all summands by the expressions (1.5.7). We will get $\sum_{k=p}^{\infty} \left(\sum_{m=0}^{n-1} m! k^{\underline{-2-m}} + \frac{n! k^{\underline{-1-n}}}{k+1} \right)$. Changing the order of summation we have $\sum_{m=0}^{n-1} m! \sum_{k=p}^{\infty} k^{\underline{-2-m}} + \sum_{k=p}^{\infty} \frac{n! k^{\underline{-1-n}}}{k+1}$. Since $\frac{x^{\underline{-1-m}}}{1+m}$ telescopes the sequence $\{k^{\underline{-2-m}}\}$, we have $\sum_{k=p}^{\infty} k^{\underline{-2-m}} = \frac{p^{\underline{-1-m}}}{1+m}$. Denote the sum of remainders $\sum_{k=p}^{\infty} \frac{n! k^{\underline{-1-n}}}{k+1}$ by $R(n,p)$. Then for all natural p and n one has

$$\sum_{k=p}^{\infty} \frac{1}{(1+k)^2} = \sum_{m=0}^{n-1} m! \frac{p^{\underline{-1-m}}}{1+m} + R(n,p) \tag{1.5.9}$$

For $p = 0$ and $n = +\infty$, the right-hand side turns into the Euler series, and one could get a false impression that we get nothing new. But $k^{\underline{-2-n}} \leq \frac{k^{\underline{-1-n}}}{k+1} \leq (k-1)^{\underline{-2-n}}$, hence

$$\frac{n!p^{\underline{-1-n}}}{1+n} = \sum_{k=p}^{\infty} n!k^{\underline{-2-n}} \leq R(n,p) \leq \sum_{k=p}^{\infty} n!(k-1)^{\underline{-2-n}} = \frac{n!(p-1)^{\underline{-1-n}}}{1+n}.$$

Since $(p-1)^{\underline{-1-n}} - p^{\underline{-1-n}} = (1+n)(p-1)^{\underline{-2-n}}$, there is a $\theta \in (0,1)$ such that

$$R(n,p) = \frac{n!p^{\underline{-1-n}}}{1+n} + \theta n!(p-1)^{\underline{-2-n}}.$$

Finally we get:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \sum_{k=0}^{p-1} \frac{1}{(1+k)^2} + \sum_{k=0}^{n-1} k!\frac{p^{\underline{-1-k}}}{1+k} + \theta n!(p-1)^{\underline{-2-n}} \tag{1.5.10}$$

For $p = n = 3$ this formula turns into

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{4} + \frac{1}{40} + \frac{1}{180} + \frac{\theta}{420}$$

For $p = n = 10$ one gets $R(10,10) \leq 10!9^{\underline{-12}}$. After cancellations one has $\frac{1}{2\cdot 11\cdot 12\cdot 13\cdot 14\cdot 15\cdot 17\cdot 19}$. This is approximately $2\cdot 10^{-8}$. Therefore $\sum\limits_{k=0}^{10-1} \frac{1}{(k+1)^2} + \sum\limits_{k=0}^{10-1} \frac{k!10^{\underline{-1-k}}}{1+k}$ is inferior than the sum of the Euler series by only $2\cdot 10^{-8}$. In such a way one can in one hour calculate eight digits of $\sum\limits_{k=1}^{\infty} \frac{1}{k^2}$ after the decimal point. It is not a bad result, but it is still far from Euler's 18 digits. For $p = 10$ to provide 18 digits one has to sum essentially more than 100 terms of the series. This is a bit too much for a person, but is possible for a computer.

Problems.

1. Telescope $\sum k^3$
2. Represent x^4 as a factorial polynomial.
3. Evaluate $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+2)}$
4. Evaluate $\sum\limits_{k=1}^{\infty} \frac{1}{k(k+1)(k+2)(k+3)}$
5. If $\Delta a_k \geq \Delta b_k$ for all k and $a_1 \geq b_1$ then $a_k \geq b_k$ for all k
6. $\Delta(x+a)^{\underline{n}} = n(x+a)^{\underline{n-1}}$
7. Prove Archimedes's inequality $\frac{n^3}{3} \leq \sum\limits_{k=1}^{n-1} k^2 \leq \frac{(n+1)^3}{3}$

8. Telescope $\sum\limits_{k=1}^{\infty} \frac{k}{2^k}$

9. Prove the inequalities $\frac{1}{n} \geq \sum\limits_{k=n+1}^{\infty} \frac{1}{k^2} \geq \frac{1}{n+1}$

10. Prove that the degree of $\Delta P(x)$ is less than the degree of $P(x)$ for any polynomial $P(x)$.

11. Relying on $\Delta 2^n = 2^n$, prove that $P(n) < 2^n$ eventually for any polynomial $P(x)$.

12. Prove $\sum\limits_{k=0}^{\infty} k!(x-1)^{\underline{-1-k}} = \frac{1}{x}$

1.6 Complex series

On the contents of the lecture. The complex numbers hide the key to the Euler Series. The summation theory developed for positive series now extends to complex series. We will see that complex series can help to sum real series.

Cubic equation Complex numbers arose in connection with solution of cubic equation. The substitution $x = y - \frac{a}{3}$ reduces the general cubic equation $x^3 + ax^2 + bx + c = 0$ to

$$y^3 + py + q = 0 \tag{1.6.1}$$

The reduced equation one solves by the following trick. One looks for a root in the form $y = \alpha + \beta$. Then $(\alpha + \beta)^3 + p(\alpha + \beta) + q = 0$ or $\alpha^3 + \beta^3 + 3\alpha\beta(\alpha + \beta) + p(\alpha + \beta) + q = 0$ The latter equality one reduces to the system

$$\begin{aligned} \alpha^3 + \beta^3 &= -q \\ 3\alpha\beta &= -p \end{aligned} \tag{1.6.2}$$

Raising the second equation into cube one gets

$$\begin{aligned} \alpha^3 + \beta^3 &= -q \\ 27\alpha^3\beta^3 &= -p^3 \end{aligned} \tag{1.6.3}$$

Now α^3, β^3 are roots of the quadratic equation

$$x^2 + qx - \frac{p^3}{27} \tag{1.6.4}$$

called a *resolution* of the original cubic equation. Sometimes the resolution has no roots, while the cubic equation always has a root. Nevertheless one can evaluate a root of the cubic equation with the help of its resolution. To do this one simply ignores that numbers under square roots are negative.

For example consider the following cubic equation

$$x^3 - \frac{3}{2}x - \frac{1}{2} = 0 \tag{1.6.5}$$

Then (1.6.2) turns into

$$\begin{aligned} \alpha^3 + \beta^3 &= \frac{1}{2} \\ \alpha^3\beta^3 &= \frac{1}{8} \end{aligned} \tag{1.6.6}$$

Corresponding resolution is $t^2 - \frac{t}{2} + \frac{1}{8} = 0$ and its roots are $t_{1,2} = \frac{1}{4} \pm \sqrt{\frac{1}{16} - \frac{1}{8}} = \frac{1}{4} \pm \frac{1}{4}\sqrt{-1}$. Then the desired root of cubic equation is given by

$$\sqrt[3]{\frac{1}{4}(1 + \sqrt{-1})} + \sqrt[3]{\frac{1}{4}(1 - \sqrt{-1})} = \frac{1}{\sqrt[3]{4}}\left(\sqrt[3]{1 + \sqrt{-1}} + \sqrt[3]{1 - \sqrt{-1}}\right) \tag{1.6.7}$$

It turns out that the latter expression one uniquely interprets as a real number which is a root of the equation 1.6.5. To evaluate it consider the following expression

$$\sqrt[3]{(1+\sqrt{-1})^2}-\sqrt[3]{(1+\sqrt{-1})}\sqrt[3]{(1-\sqrt{-1})}+\sqrt[3]{(1-\sqrt{-1})^2} \tag{1.6.8}$$

Since

$$(1+\sqrt{-1})^2 = 1^2+2\sqrt{-1}+\sqrt{-1}^2 = 1+2\sqrt{-1}-1 = 2\sqrt{-1},$$

the left summand of (1.6.8) is equal to $\sqrt[3]{2\sqrt{-1}} = \sqrt[3]{2}\sqrt[3]{\sqrt{-1}} = \sqrt[3]{2}\sqrt{\sqrt[3]{-1}} = \sqrt[3]{2}\sqrt{-1}$. Similarly $(1-\sqrt{-1})^2 = -2\sqrt{-1}$, and the right summand of (1.6.8) turns into $-\sqrt[3]{2}\sqrt{-1}$. Finally $(1+\sqrt{-1})(1-\sqrt{-1}) = 1^2-\sqrt{-1}^2 = 2$ and the central one is $-\sqrt[3]{2}$. As result the whole expression (1.6.8) is evaluated as $-\sqrt[3]{2}$.

On the other hand one evaluates the product of (1.6.7) and (1.6.8) by usual formula as the sum of cubes

$$\frac{1}{\sqrt[3]{4}}((1+\sqrt{-1})+(1-\sqrt{-1})) = \frac{1}{\sqrt[3]{4}}((1+1)+(\sqrt{-1})-\sqrt{-1})) = \frac{1}{\sqrt[3]{4}}(2+0) = \sqrt[3]{2}.$$

Consequently (1.6.7) is equal to $\frac{\sqrt[3]{2}}{-\sqrt[3]{2}} = -1$. And -1 is a true root of (1.6.5).

Arithmetic of complex numbers. In the sequel we use i instead of $\sqrt{-1}$. There are two basic ways to represent a complex number. Representation $z = a+ib$, where a and b are real numbers we call *Cartesian form* of z. The numbers a and b are called respectively *real* and *imaginary* parts of z and denoted by $\operatorname{Re} z$ and by $\operatorname{Im} z$ respectively. Addition and multiplication of complex numbers are defined via their real and imaginary parts as follows

$$\operatorname{Re}(z_1+z_2) = \operatorname{Re} z_1 + \operatorname{Re} z_2 \tag{1.6.9}$$

$$\operatorname{Im}(z_1+z_2) = \operatorname{Im} z_1 + \operatorname{Im} z_2 \tag{1.6.10}$$

$$\operatorname{Re}(z_1z_2) = \operatorname{Re} z_1 \operatorname{Re} z_2 - \operatorname{Im} z_1 \operatorname{Im} z_2 \tag{1.6.11}$$

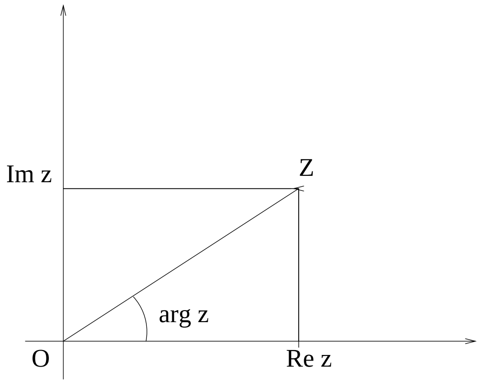
$$\operatorname{Im}(z_1z_2) = \operatorname{Re} z_1 \operatorname{Im} z_2 + \operatorname{Im} z_1 \operatorname{Re} z_2 \tag{1.6.12}$$

Trigonometric form of a complex number is $z = \rho(\cos\phi + i\sin\phi)$, where $\rho \geq 0$ is called *module* or *absolute value* of a complex number z and is denoted $|z|$, and ϕ is called its *argument*. The argument of a complex number is defined modulo 2π. We denote by $\operatorname{Arg}(z)$ the set of all arguments of z, and by $\arg z$ elements of $\operatorname{Arg} z$ which satisfy the inequalities $-\pi < \arg z \leq \pi$. So $\arg z$ is uniquely defined for all complex numbers. The $\arg z$ is called *principal argument* of z.

The number $a-bi$ is called *conjugated* to $z = a+bi$ and denoted $\overline{z}$. One has $z\overline{z} = |z|^2$. This allows us to express z^{-1} as $\frac{\overline{z}}{|z|^2}$.

If $z = a+ib$ then $|z| = \sqrt{a^2+b^2}$ and $\arg z = \operatorname{arctg}\frac{b}{a}$. One represents a complex numbers $z = a+bi$ as a points Z of the plane with coordinates (a,b). Then $|z|$ is equal to the distance from Z to the origin O. And $\arg z$ represents the angle between the axis of abscises and the ray $\overrightarrow{OZ}$. Addition of the complex numbers corresponds to usual vector addition. And the usual triangle inequality turns into the *module inequality*:

$$|z+\zeta| \leq |z|+|\zeta|.$$



The multiplication formula for complex numbers in the trigonometric form is especially simple:

$$r(\cos\phi + i\sin\phi)r'(\cos\psi + i\sin\psi) = rr'(\cos(\phi+\psi) + i\sin(\phi+\psi)) \tag{1.6.13}$$

Indeed, the left-hand side and the right-hand side of (1.6.13) transform to

$$rr'(\cos\phi\cos\psi - \sin\phi\sin\psi) + irr'(\sin\phi\cos\psi + \sin\psi\cos\phi).$$

That is the module of the product is equal to the product of modules and the argument of product is equal to the sum of arguments

$$\operatorname{Arg} z_1 z_2 = \operatorname{Arg} z_1 \oplus \operatorname{Arg} z_2$$

Any complex number is uniquely defined by its module and argument.

The multiplication formula allows us to prove by induction the following:

(Moivre Formula) $$(\cos\phi + i\sin\phi)^n = (\cos n\phi + i\sin n\phi)$$

Sum of complex series. Now is the time to extend our summation theory to series made of complex numbers. We extend the whole theory without any losses to so called absolutely convergent series. Series $\sum_{k=1}^{\infty} z_k$ with arbitrary complex terms is called *absolutely convergent*, if the series $\sum_{k=1}^{\infty} |z_k|$ of absolute values converges.

For any real number x one defines two nonnegative number its *positive x^+ and negative x^- parts* as $x^+ = x[x \geq 0]$ and $x^- = -x[x < 0]$. The following identities characterize the positive and negative partes of x

$$x^+ + x^- = |x| \qquad x^+ - x^- = x \tag{1.6.14}$$

Now the sum of an absolute convergent series of real numbers is defined as follows:

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_k^+ - \sum_{k=1}^{\infty} a_k^- \tag{1.6.15}$$

That is from the sum all positive summands one subtracts the sum of modules of all negative summands. The series on the right-hand side converge, because $a_k^+ \leq |a_k|$, $a_k^- \leq |a_k|$ and $\sum\limits_{k=1}^{\infty} |a_k| < \infty$.

For an absolutely convergent complex series $\sum\limits_{k=1}^{\infty} z_k$ we define the real and imaginary parts of its sum separately by formulas

$$\operatorname{Re}\sum_{k=1}^{\infty} z_k = \sum_{k=1}^{\infty} \operatorname{Re} z_k \qquad \operatorname{Im}\sum_{k=1}^{\infty} z_k = \sum_{k=1}^{\infty} \operatorname{Im} z_k \tag{1.6.16}$$

The series in the right-hand sides of these formulas are absolutely convergent, since $|\operatorname{Re} z_k| \leq |z_k|$ and $|\operatorname{Im} z_k| \leq |z_k|$.

Theorem 1. *For any pair of absolutely convergent series* $\sum\limits_{k=1}^{\infty} a_k$ *and* $\sum\limits_{k=1}^{\infty} b_k$ *its term-wise sum* $\sum\limits_{k=1}^{\infty} (a_k + b_k)$ *absolutely converges and*

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k \tag{1.6.17}$$

Proof. First, remark that absolute convergence of the series on the left-hand side follows from the Module Inequality $|a_k + b_k| \leq |a_k| + |b_k|$ and absolute convergence of the series on the right-hand side.

Now consider the case of real numbers. Representing all sums in (**1**) as differences of its positive and negative parts and separating positive and negative terms in different sides one transforms (**1**) into

$$\sum_{k=1}^{\infty} a_k^+ + \sum_{k=1}^{\infty} b_k^+ + \sum_{k=1}^{\infty}(a_k + b_k)^- = \sum_{k=1}^{\infty} a_k^- + \sum_{k=1}^{\infty} b_k^- + \sum_{k=1}^{\infty}(a_k + b_k)^+$$

But this equality is true due to term-wise addition for positive series and the following identity,

$$x^- + y^- + (x + y)^+ = x^+ + y^+ + (x + y)^-$$

Moving terms around turns this identity into

$$(x + y)^+ - (x + y)^- = (x^+ - x^-) + (y^+ - y^-),$$

which is true due to the identity $x^+ - +x^- = x$.

In the complex case the equality (1.6.17) splits into two equalities, one for real and another for imaginary parts. As for real series the term-wise addition is already proved, we can write the following chain of equalities

$$\operatorname{Re}\left(\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k\right) = \operatorname{Re}\sum_{k=1}^{\infty} a_k + \operatorname{Re}\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} \operatorname{Re} a_k + \sum_{k=1}^{\infty} \operatorname{Re} b_k =$$
$$\sum_{k=1}^{\infty}(\operatorname{Re} a_k + \operatorname{Re} b_k) = \sum_{k=1}^{\infty} \operatorname{Re}(a_k + b_k) = \operatorname{Re}\sum_{k=1}^{\infty}(a_k + b_k),$$

which proves the equality of real parts in (1.6.17). The same proof works for imaginary parts. □

Sum Partition Theorem. An unordered sum of a family of complex numbers is defined by the same formulas (1.6.15) and (1.6.16). Since for positive series non-ordered sum coincides with the ordered one, we get the same coincidence for all absolutely convergent series. Hence the commutativity law holds for all absolutely convergence series.

Theorem 2. *If* $I = \sqcup_{j\in J} I_j$ *and* $\sum\limits_{k=1}^{\infty} |a_k| < \infty$ *then* $\sum_{j\in J}\left|\sum_{i\in I_j} a_i\right| < \infty$ *and*

$$\sum_{j\in J}\sum_{i\in I_j} a_i = \sum_{i\in I} a_i \tag{1.6.18}$$

Proof. At first consider the case of real summands. By definition $\sum\limits_{i\in I} a_i = \sum\limits_{i\in I} a_i^+ - \sum\limits_{i\in I} a_i^-$ By Sum Partition Theorem positive series one transforms the original sum into $\sum\limits_{j\in J}\sum\limits_{i\in I_j} a_i^+ - \sum\limits_{j\in J}\sum\limits_{i\in I_j} a_i^-$ Now by the Term-Wise Addition applied at first to external and after to internal sums one gets

$$\sum_{j\in J}\left(\sum_{i\in I_j} a_i^+ - \sum_{i\in I_j} a_i^-\right) = \sum_{j\in J}\sum_{i\in I_j}(a_i^+ - a_i^-) = \sum_{j\in J}\sum_{i\in I_j} a_i.$$

So the Sum Partition Theorem is proved for all absolutely convergent real series. And it immediately extends to absolutely convergent complex series by its splitting into real and imaginary parts. □

Theorem 3 (Term-Wise Multiplication). *If* $\sum\limits_{k=1}^{\infty} |z_k| < \infty$ *then* $\sum\limits_{k=1}^{\infty} |cz_k| < \infty$ *for any (complex)* c *and* $\sum\limits_{k=1}^{\infty} cz_k = c\sum\limits_{k=1}^{\infty} z_k$

Proof. Term-wise Multiplication for positive numbers gives the first statement of the theorem $\sum\limits_{k=1}^{\infty} |cz_k| = \sum\limits_{k=1}^{\infty} |c||z_k| = |c|\sum\limits_{k=1}^{\infty} |z_k|$. The further proof is divided into five cases.

At first suppose c is positive and z_k real. Then $cz_k^+ = cz_k^+$ and by virtue of Term-Wise Multiplication for positive series we get

$$\sum_{k=1}^{\infty} cz_k = \sum_{k=1}^{\infty} cz_k^+ - \sum_{k=1}^{\infty} cz_k^- = c\sum_{k=1}^{\infty} z_k^+ - c\sum_{k=1}^{\infty} z_k^- = c\left(\sum_{k=1}^{\infty} z_k^+ - \sum_{k=1}^{\infty} z_k^-\right) = c\sum_{k=1}^{\infty} z_k$$

The second case. Let $c = -1$ and z_k be real. In this case

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

The third case. Let c be real and z_k complex. In this case $\operatorname{Re} cz_k = c\operatorname{Re} z_k$ and the two cases above imply the Term-Wise Multiplication for any real c.

Hence

$$\operatorname{Re}\sum_{k=1}^{\infty} cz_k = \sum_{k=1}^{\infty}\operatorname{Re} cz_k = \sum_{k=1}^{\infty} c\operatorname{Re} z_k = c\sum_{k=1}^{\infty}\operatorname{Re} z_k = c\operatorname{Re}\sum_{k=1}^{\infty} z_k = \operatorname{Re} c\sum_{k=1}^{\infty} z_k$$

The same is true for imaginary parts.

The fourth case. Let $c = i$ and z_k be complex. Then $\operatorname{Re} iz_k = -\operatorname{Im} z_k$ and $\operatorname{Im} iz_k = \operatorname{Re} z_k$. So one gets for real parts

$$\operatorname{Re}\sum_{k=1}^{\infty} iz_k = \sum_{k=1}^{\infty}\operatorname{Re}(iz_k) = \sum_{k=1}^{\infty} -\operatorname{Im} z_k =$$
$$-\sum_{k=1}^{\infty}\operatorname{Im} z_k = -\operatorname{Im}\sum_{k=1}^{\infty} z_k = \operatorname{Re} i\sum_{k=1}^{\infty} z_k$$

The general case. Let $c = a + bi$ with real a, b. Then

$$c\sum_{k=1}^{\infty} z_k = a\sum_{k=1}^{\infty} z_k + ib\sum_{k=1}^{\infty} z_k = \sum_{k=1}^{\infty} az_k + \sum_{k=1}^{\infty} ibz_k = \sum_{k=1}^{\infty}(az_k + ibz_k) = \sum_{k=1}^{\infty} cz_k$$

□

Multiplication of Series For given two series $\sum\limits_{k=0}^{\infty} a_k$ and $\sum\limits_{k=0}^{\infty} b_k$, one defines their *convolution* as a series $\sum\limits_{n=0}^{\infty} c_n$, where $c_k = \sum\limits_{k=0}^{n} a_k b_{n-k}$.

Theorem 4 (Cauchy). *For any absolutely convergent series* $\sum\limits_{k=0}^{\infty} a_k$ *and* $\sum\limits_{k=0}^{\infty} b_k$ *the convolution* $\sum\limits_{k=0}^{\infty} c_k$ *absolutely converges and*

$$\sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k \tag{1.6.19}$$

Proof. Consider the double series $\sum\limits_{i,j} a_i b_j$. Then by Sum Partition Theorem its sum is equal to

$$\sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} a_i b_j\right) = \sum_{j=0}^{\infty} b_j\left(\sum_{i=0}^{\infty} a_i\right) = \left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right).$$

On the other hand, $\sum\limits_{i,j} a_i b_j = \sum\limits_{n=0}^{\infty}\sum\limits_{k=0}^{n+1-1} a_k b_{n-k}$. But the last sum is just the convolution.

This proof goes through for positive series. In general case we have to prove absolute convergence of the double series. But this follows from

$$\left(\sum_{k=0}^{\infty}|a_k|\right)\left(\sum_{k=0}^{\infty}|b_k|\right) = \sum_{k=0}^{\infty}|c_k|$$

□

Module Inequality.

$$\left|\sum_{k=1}^{\infty} z_k\right| \leq \sum_{k=1}^{\infty} |z_k| \tag{1.6.20}$$

Let $z_k = x_k + iy_k$. The summation of inequalities $-|x_k| \leq x_k \leq |x_k|$ gives $-\sum_{k=1}^{\infty}|x_k| \leq \sum_{k=1}^{\infty} x_k \leq \sum_{k=1}^{\infty}|x_k|$, which means $\left|\sum_{k=1}^{\infty} x_k\right| \leq \sum_{k=1}^{\infty}|x_k|$. The same inequality is true for y_k. Consider $z'_k = |x_k| + i|y_k|$ Then $|z_k| = |z'_k|$ and $\left|\sum_{k=1}^{\infty} z_k\right| \leq \left|\sum_{k=1}^{\infty} z'_k\right|$. Therefore it is sufficient to prove inequality (1.6.20) for z'_k, that is for numbers with non-negative real and imaginary parts. Now supposing x_k, y_k to be nonnegative one gets the following chain of equivalent transformations of (1.6.20):

$$\left(\sum_{k=1}^{\infty} x_k\right)^2 + \left(\sum_{k=1}^{\infty} y_k\right)^2 \leq \left(\sum_{k=1}^{\infty} |z_k|\right)^2$$

$$\sum_{k=1}^{\infty} x_k \leq \sqrt{\left(\sum_{k=1}^{\infty} |z_k|\right)^2 - \left(\sum_{k=1}^{\infty} y_k\right)^2}$$

$$\sum_{k=1}^{n} x_k \leq \sqrt{\left(\sum_{k=1}^{\infty} |z_k|\right)^2 - \left(\sum_{k=1}^{\infty} y_k\right)^2} \qquad \forall n = 1, 2, \dots$$

$$\sum_{k=1}^{\infty} y_k \leq \sqrt{\left(\sum_{k=1}^{\infty} |z_k|\right)^2 - \left(\operatorname{Re}\sum_{k=1}^{n} x_k\right)^2} \qquad \forall n = 1, 2, \dots$$

$$\sum_{k=1}^{m} y_k \leq \sqrt{\left(\sum_{k=1}^{\infty} |z_k|\right)^2 - \left(\sum_{k=1}^{n} x_k\right)^2} \qquad \forall n, m = 1, 2, \dots$$

$$\left(\sum_{k=1}^{n} x_k\right)^2 + \left(\sum_{k=1}^{m} y_k\right)^2 \leq \left(\sum_{k=1}^{\infty} |z_k|\right)^2 \qquad \forall m, n = 1, 2, \dots$$

$$\left|\sum_{k=1}^{N} z_k\right| = \sqrt{\left(\sum_{k=1}^{N} x_k\right)^2 + \left(\sum_{k=1}^{N} y_k\right)^2} \leq \sum_{k=1}^{\infty} |z_k| \qquad \forall N = 1, 2, \dots$$

The inequalities of the last system hold because $\left|\sum_{k=1}^{N} z_k\right| \leq \sum_{k=1}^{N} |z_k| \leq \sum_{k=1}^{\infty} |z_k|$.

Complex geometric progression The sum of a geometric progression with a complex ratio is given by the same formula

$$\sum_{k=0}^{n-1} z^k = \frac{1 - z^n}{1 - z} \tag{1.6.21}$$

And the proof is the same as in the case of real numbers. But the meaning of this formula is different. Any complex formula is in fact a pair of formulas. Any complex equation is in fact a pair of equations.

In particular, for $z = q(\sin\phi + i\cos\phi)$ the real part of the left-hand side of (1.6.21) owing to Moivre Formula turns into $\sum_{k=0}^{n-1} q^k \sin k\phi$ and the right-hand side turns into $\sum_{k=0}^{n-1} q^k \cos k\phi$. So the formula for geometric progression splits into two formulas which allow us to telescope some trigonometric series.

Especially interesting is the case with the ratio $\varepsilon_n = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$. In this case the geometric progression cyclically takes the same values, because $\varepsilon_n^n = 1$. The terms of this sequence are called *the roots of unity*, because they satisfy the equation $z^n - 1 = 0$.

Lemma 5.

$$(z^n - 1) = \prod_{k=1}^{n}(z - \varepsilon_n^k) \tag{1.6.22}$$

Proof. Denote by $P(z)$ the right-hand side product. This polynomial has degree n has the major coefficient 1 and has all ε_n^k as its roots. Then the difference $(z^n - 1) - P(z)$ is a polynomial of degree $< n$ which has n different roots. Such polynomial has to be 0 by virtue of the following general theorem. □

Theorem 6. *Number of roots of any nonzero complex polynomial does not exceed its degree.*

Proof. Proof is by induction on degree of $P(z)$. Polynomial of degree 1 has form $az + b$ and the only root $\frac{b}{a}$. Suppose our theorem is proved for any polynomial of degree $< n$. Consider a polynomial $P(z) = a_0 + a_1 z + \cdots + a_n z$ of degree n, where coefficient are complex numbers. Suppose it has at least n roots $z_1 \ldots z_n$. Consider the polynomial $P^*(z) = a_n \prod_{k=1}^{n}(z - z_k)$. The difference $P(z) - P^*(z)$ has degree $< n$ and has at least n roots (all z_k). By the induction hypothesis this difference is zero. Hence, $P(z) = P^*(z)$. But $P^*(z)$ has only n roots. Indeed, for any z different from all z_k one has $|z - z_k| > 0$. Therefore $|P^*(z)| = |a_n| \prod_{k=1}^{n} |z - z_k| > 0$. □

Blocking conjugated roots one gets a pure real formula.

$$z^n - 1 = (z - 1) \prod_{k=1}^{(n-1)/2} \left(z^2 - 2z\cos\frac{2k\pi}{n} + 1 \right) \tag{1.6.23}$$

Complexification of series Complex numbers are effectively applied to sum up so called *trigonometric series,* i.e. series of the type $\sum_{k=0}^{\infty} a_k \cos kx$ and $\sum_{k=0}^{\infty} a_k \sin kx$. For example, to sum series $\sum_{k=1}^{\infty} q^k \sin k\phi$ one couple it with its dual $\sum_{k=0}^{\infty} q^k \cos k\phi$ and form a complex series $\sum_{k=0}^{\infty} q^k(\cos k\phi + i\sin k\phi)$. The last is a complex geometric series. Its sum is $\frac{1}{1-z}$, where $z = \cos\phi + i\sin\phi$. Now the sum of the sine series $\sum_{k=1}^{\infty} q^k \sin k\phi$ is equals to $\operatorname{Im}\frac{1}{1-z}$, the imaginary part of the complex series, and the real part of the complex series coincides with the cosine series. In particular, for $q = 1$, one has $\frac{1}{1-z} = \frac{1}{1+\cos\phi + i\sin\phi}$. To evaluate the real and imaginary parts one multiplies both numerator and denominator by $1 + \cos\phi - i\sin\phi$. Then one gets $(1-\cos\phi)^2 + \sin^2\phi = 1 - 2\cos^2\phi + \cos^2\phi + \sin^2\phi = 2 - 2\cos\phi$ as the denominator. Hence $\frac{1}{1-z} = \frac{1-\cos\phi + i\sin\phi}{2-2\cos\phi} = \frac{1}{2} + \frac{1}{2}\cot\frac{\phi}{2}$. And we get two remarkable formulas for the sum of the divergent series

$$\sum_{k=0}^{\infty} \cos k\phi = \frac{1}{2} \qquad \sum_{k=1}^{\infty} \sin k\phi = \frac{1}{2}\cot\frac{\phi}{2} \tag{1.6.24}$$

For $\phi = 0$ the left series turns into $\sum_{k=0}^{\infty} (-1)^k$. The evaluation of the Euler series via this cosine series is remarkably short, it takes one line. But one has to know integrals and a something else to justify this evaluation.

Problems.

1. Find real and imaginary parts for $\dfrac{1}{1-i}$, $\left(\dfrac{1-i}{1+i}\right)^3$, $\dfrac{i^5+2}{i^{19}+1}$, $\dfrac{(1+i)^5}{(1-i)^3}$.
2. Find trigonometric form for -1, $1+i$, $\sqrt{3}+i$.
3. Prove that $z_1 z_2 = 0$ implies either $z_1 = 0$ or $z_2 = 0$.
4. Prove the distributivity law for complex numbers.
5. Analytically prove the inequality $|z_1 + z_2| \leq |z_1| + |z_2|$.
6. Evaluate $\sum_{k=1}^{n-1} \frac{1}{z_k(z_k+1)}$, where $z_k = 1 + kz$.
7. Evaluate $\sum_{k=1}^{n-1} z_k^2$, where $z_k = 1 + kz$.
8. Evaluate $\sum_{k=1}^{n-1} \dfrac{\sin k}{2^k}$.
9. Solve $z^2 = i$.
10. Solve $z^2 = 3 - 4i$.
11. Telescope $\sum_{k=1}^{\infty} \frac{\sin 2n}{3^n}$.
12. Prove that the conjugated to a root of polynomial with real coefficient is the root of the polynomial.
13. Prove that $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$.
14. Prove that $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.
15. * Solve $8x^3 - 6x - 1 = 0$
16. Evaluate $\sum_{k=1}^{\infty} \dfrac{\sin n}{2^n}$

17. Evaluate $\sum_{k=1}^{\infty} \frac{\sin 2n}{3^n}$

18. Prove absolute convergence of $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ for any z

19. For which z the series $\sum_{k=1}^{\infty} \frac{z^k}{k}$ absolutely converges?

20. Multiply a geometric series onto itself several times applying Cauchy formula

21. Find series for $\sqrt{1+x}$ by method of indefinite coefficients

22. Does series $\sum_{k=1}^{\infty} \frac{\sin k}{k}$ absolutely converge?

23. Does series $\sum_{k=1}^{\infty} \frac{\sin k}{k^2}$ absolutely converge?

Chapter 2

Integrals

2.1 Natural Logarithm.

On the contents of the lecture.

> In the beginning of Calculus was the Word, and the Word was with Arithmetic, and the Word was *Logarithm*[1]

Logarithmic tables Multiplication is much more difficult than addition. The Logarithm reduces multiplication to addition. Inventions of logarithms were one of the great achievements of our civilization.

In early times, when logarithms was unknown instead of them one used trigonometric functions. The following identity

$$2\cos x\cos y = \cos(x+y) + \cos(x-y)$$

can be applied to calculate products via tables of cosines. To multiply numbers x and y, one represents them as cosines $x = \cos a$, $y = \cos b$ using the cosine table. Then one evaluates $(a+b)$ and $(a-b)$ and find their cosines in the table. Finally, the results are summed and divided by 2. That is all. A single multiplication requires 4 searches in the table of cosines, two additions, one subtraction and one division by 2.

A logarithmic function $l(x)$ is a function such that $l(xy) = l(x)+l(y)$ for any x and y. If one has a logarithmic table, to evaluate the product xy one has to find in logarithmic table $l(x)$ and $l(y)$ then sum them and find the antilogarithm of the sum. This is much easier.

The idea of logarithms arose in 1544, when M. Stiefel compared a geometric and arithmetic progressions. The addition of exponents corresponds to the multiplication of powers. Hence consider a number close to 1, say, 1.000001. Calculate the sequence of its powers and place them to the left column. Place to the right column the corresponding values of exponents, which are just the line numbers. The logarithmic table is ready.

Now to multiply two numbers x and y, one finds them (or their approximations) in the left column of the logarithmic table, reads their logarithms from the right column. Sum the logarithms and find the value of the sum in the right column. Next to this sum in the left column the product xy stands. The first tables of such logarithms were composed by John Napier in 1614.

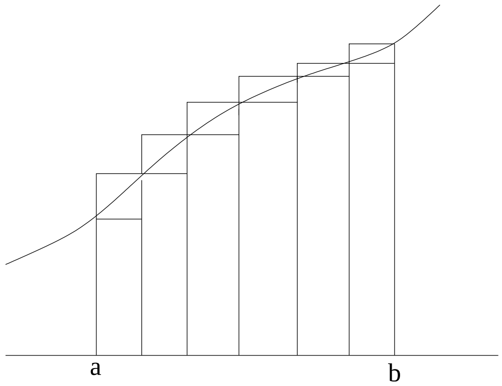
Area of curvilinear trapezium. Recall that a sequence is said to be monotone, if it is either increasing or decreasing. The minimal interval, which contains all elements of a given sequence of points will be called *supporting interval* of the sequence.

And a sequence is called *exhausting* for an interval I if I is the supporting interval of the sequence.

Let f be a non-negative function defined on $[a,b]$. The set $\{(x,y) \mid x \in [a,b] \text{ and } 0 \le y \le f(x)\}$ is called a *curvilinear trapezium* under the graph of f over the interval $[a,b]$.

To estimate the area of a curvilinear trapezium under the graph of f over $[a,b]$, one chooses an exhausting sequence $\{x_i\}_{i=0}^n$ for $[a,b]$ and consider the

[1] $\lambda o\gamma o\varsigma$ is Greek "word", $\alpha\varrho\iota\theta\mu o\varsigma$ means "number"



following sums:

$$(2.1.1) \qquad \sum_{k=0}^{n-1} f(x_k)|\Delta x_k| \qquad \sum_{k=0}^{n-1} f(x_{k+1})|\Delta x_k| \quad (\text{where } \Delta x_k = x_{k+1} - x_k)$$

We will call the first of them the *receding sum* and the second, the *advancing sum* of the sequence $\{x_k\}$ for the function f. If the function f is monotone the area of the curvilinear trapezium is contained between these two sums. To see this, consider the following step-figures: $\cup_{k=0}^{n-1}[x_k, x_{k+1}] \times [0, f(x_k)]$ and $\cup_{k=0}^{n-1}[x_k, x_{k+1}] \times [0, f(x_{k+1})]$. If f and $\{x_k\}$ both increase or both decrease the first step-figure is contained in the curvilinear trapezium and the second step-figure contains the trapezium with possible exception of a vertical segments $[a \times [0, f(a)]$ or $[b \times [0, f(b)]$. If one of f and $\{x_k\}$ increases and the other decreases, then the step-figures switch the roles. The back sum equals the area of the first step-figure, and the advancing sum equals the area of the second one. Thus we have proved the following lemma.

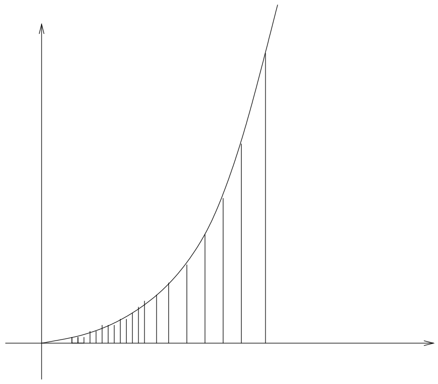
Lemma 1. *Let f be a monotone function, S the area of the curvilinear trapezium under the graph of f over $[a, b]$. Then for any sequence $\{x_k\}_{k=0}^n$ exhausting $[a, b]$ the area S is contained between $\sum_{k=0}^{n-1} f(x_k)|\Delta x_k|$ and $\sum_{k=0}^{n-1} f(x_{k+1})|\Delta x_k|$.*

Fermat's quadratures of parabolas. In 1636 Pierre Fermat proposed an ingenious trick to determine the area below the curve $y = x^a$.

If $a > -1$ then consider any interval of the form $[0, B]$. Choose a positive $q < 1$. Then the infinite geometric progression $B, Bq, Bq^2, Bq^3 \dots$ exhausts $[0, B]$ and the values of function for this sequence also form a geometric progression $B^a, q^aB^a, q^{2a}B^a, q^{3a}B^a, \dots$. Then both the receding and advancing sums turn into geometric progressions:

$$\sum_{k=0}^{\infty} B^a q^{ka}(q^kB - q^{k+1}B) = B^{a+1}(1-q)\sum_{k=0}^{\infty} q^{k(a+1)} = \frac{B^{a+1}(1-q)}{1-q^{a+1}}$$

$$\sum_{k=0}^{\infty} B^a q^{(k+1)a}(q^kB - q^{k+1}B) = B^{a+1}(1-q)\sum_{k=0}^{\infty} q^{(k+1)(a+1)} = \frac{B^{a+1}(1-q)q^a}{1-q^{a+1}}.$$



For a natural a, one has $\frac{1-q}{1-q^{a+1}} = \frac{1}{1+q+q^2+\dots q^a}$. As q tends to 1 both sums converge to $\frac{B^{a+1}}{a+1}$. This is the area of the curvilinear trapezium. Let us remark that for $a < 0$ this trapezium is unbounded, nevertheless it has finite area if $a > -1$.

If $a < -1$, then consider an interval in the form $[B, \infty]$. Choose a positive $q > 1$. Then the infinite geometric progression $B, Bq, Bq^2, Bq^3 \dots$ exhausts $[B, \infty]$ and the values of function for this sequence also form a geometric progression $B^a, q^a B^a, q^{2a} B^a, q^{3a} B^a, \dots$. The receding and advancing sums are

$$\sum_{k=0}^{\infty} B^a q^{ka}(q^{k+1}B - q^k B) = B^{a+1}(q-1)\sum_{k=0}^{\infty} q^{k(a+1)} = \frac{B^{a+1}(q-1)}{1-q^{a+1}}$$

$$\sum_{k=0}^{\infty} B^a q^{(k+1)a}(q^{k+1}B - q^k B) = B^{a+1}(1-q)\sum_{k=0}^{\infty} q^{(k+1)(a+1)} = \frac{B^{a+1}(q-1)q^a}{1-q^{a+1}}.$$

If a is an integer set $p = q^{-1}$. Then $\frac{q-1}{1-q^{a+1}} = q\frac{1-p}{1-p^{|a|-1}} = q\frac{1}{1+p+p^2+\dots p^{n-2}}$. As q tends to 1 both sums converge to $\frac{B^{a+1}}{|a|-1}$. This is the area of the curvilinear trapezium.

For $a > -1$ the area of the curvilinear trapezium under the graph of x^a over $[A, B]$ is equal to the difference between the areas of trapezia over $[0, B]$ and $[0, A]$. Hence this area is $\frac{B^{a+1}-A^{a+1}}{a+1}$.

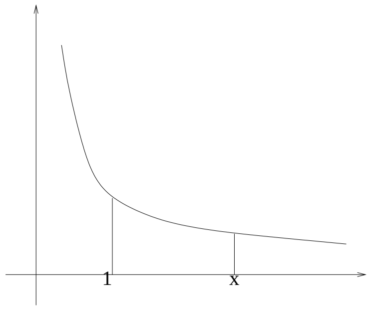
For $a < -1$ one can evaluates the area of the curvilinear trapezium under the graph of x^a over $[A, B]$ as the difference between the areas of trapezia over $[A, \infty]$ and $[B, \infty]$. The result is expressed by the same formula $\frac{B^{a+1}-A^{a+1}}{a+1}$.

Theorem 2 (Fermat). *The area below the curve $y = x^a$ over interval $[A, B]$ is equal to $\frac{B^{a+1}-A^{a+1}}{a+1}$ for $a \neq 1$.*

We have proved this theorem for integer a, but Fermat proved it for all real $a \neq -1$.

Natural Logarithm. In the case $a = -1$ the geometric progression for areas of step-figures turns into an arithmetic progression. This means that the area below hyperbola is a logarithm! This discovery was made by Gregory in 1647.

The figure bounded from above by the graph of hyperbola $y = 1/x$, from below by segment $[a, b]$ of the axis of abscissas and on each side by vertical lines



passing through the end points of the interval, is called a *hyperbolic trapezium over* $[a, b]$.

The area of hyperbolic trapezium over $[1, x]$ with $x > 1$ is called the *natural logarithm* of a x, it is denoted by $\ln x$. For a positive number $x < 1$ its logarithm is defined as the negative number whose absolute value coincides with the area of hyperbolic trapezium over $[x, 1]$. At last, $\ln 1$ is defined as 0.

Theorem 3 (on logarithm). *Natural logarithm is an increasing function defined for all positive numbers. For each pair of positive numbers x, y*

$$\ln xy = \ln x + \ln y. \tag{2.1.2}$$

Proof. Consider the case $x, y > 1$. The difference $\ln xy - \ln y$ is the area of the hyperbolic trapezium over $[y, xy]$. And we have to prove that it is equal to $\ln x$, the area of trapezium over $[1, x]$. Choose a large number n. Let $q = x^{1/n}$. Then $q^n = x$. The finite geometric progression $\{q^k\}_{k=0}^n$ exhausts $[1, x]$. Then the receding and advancing sums are

$$\sum_{k=0}^{n-1} q^{-k}(q^{k+1} - q^k) = n(q-1) \qquad \sum_{k=0}^{n-1} q^{-k-1}(q^{k+1} - q^k) = \frac{n(q-1)}{q}. \tag{2.1.3}$$

Now consider the sequence $\{xq^k\}_{k=0}^n$ exhausting $[x, xy]$. Its receding sum

$$\sum_{k=0}^{n-1} x^{-1}q^{-k}(xq^{k+1} - xq^k) = n(q-1)$$

just coincides with the receding sum (2.1.3) for $\ln x$. The same is true for the advancing sum. As a result we obtain for any natural n one has the following inequalities:

$$n(q-1) \geq \ln x \geq \frac{n(q-1)}{q} \qquad n(q-1) \geq \ln xy - \ln y \geq \frac{n(q-1)}{q} \tag{2.1.4}$$

This implies that $|\ln xy - \ln x - \ln y|$ does not exceed the difference between the the receding and advancing sums. The statement of Theorem **3** in the case $x, y > 1$ will be proved when we will prove that this difference can be made arbitrarily small by a choice of n. This will be deduced from the following general lemma.

Lemma 4. *Let f be a monotone function over interval $[a,b]$ and a sequence $\{x_k\}_{k=0}^n$ that exhausts $[a,b]$. Then*

$$\left|\sum_{k=0}^{n-1} f(x_k)\Delta x_k - \sum_{k=0}^{n-1} f(x_{k+1})\Delta x_k\right| \le |f(b)-f(a)| \max_{k<n} |\Delta x_k| \tag{2.1.5}$$

The proof of the lemma is a straightforward calculation. To shorten the notation, set $\Delta f(x_k) = f(x_{k+1}) - f(x_k)$.

$$\left|\sum_{k=0}^{n-1} f(x_k)\Delta x_k - \sum_{k=0}^{n-1} f(x_{k+1})\Delta x_k\right| = \left|\sum_{k=0}^{n-1} \Delta f(x_k)\Delta x_k\right| \le$$
$$\sum_{k=0}^{n-1} |\Delta f(x_k)| \max|\Delta x_k| = \max|\Delta x_k| \sum_{k=0}^{n-1} |\Delta f(x_k)| =$$
$$\max|\Delta x_k| \left|\sum_{k=0}^{n-1} \Delta f(x_k)\right| = \max|\Delta x_k||f(b)-f(a)|$$

The equality $\left|\sum_{k=0}^{n-1} \Delta f(x_k)\right| = \sum_{k=0}^{n-1} |\Delta f(x_k)|$ holds, as $\Delta f(x_k)$ have the same signs due to the monotonicity of f.

The value $\max|\Delta x_k|$ is called *maximal step* of the sequence $\{x_k\}$. For the sequence $\{q^k\}$ of $[1,x]$ its maximal step is equal to $q^n - q^{n-1} = q^n(1-q^{-1}) = x(1-q)/q$. It tends to 0 as q tends to 1. In our case $|f(b)-f(a)| = 1 - \frac{1}{x} < 1$. By Lemma **4** the difference between the receding and advancing sums could be made arbitrarily small. This completes the proof in the case $x, y > 1$.

Consider the case $xy = 1$, $x > 1$. We need to prove the following

$$\ln 1/x = -\ln x \tag{inversion rule}$$

As above, put $q^n = x > 1$. The sequence $\{q^{-k}\}_{k=0}^n$ exhausts $[1/x, 1]$. The corresponding receding sum $\sum_{k=0}^{n-1} q^{k+1}(q^{-k} - q^{-k-1}) = \sum_{k=0}^{n-1}(q-1) = n(q-1)$ coincides with its counterpart for $\ln x$. The same is true for the advancing one. The same arguments as above prove $|\ln 1/x| = \ln x$. The sign of $\ln 1/x$ is defined as minus because $1/x < 1$. This proves the inversion rule.

Now consider case $x < 1$, $y < 1$. Then $1/x > 1$ and $1/y > 1$ and by the first case $\ln 1/xy = (\ln 1/x + \ln 1/y)$. Replacing all terms of this equation according to the inversion rule, one gets $-\ln xy = -\ln x - \ln y$ and finally $\ln xy = \ln x + \ln y$.

The next case is $x > 1$, $y < 1$, $xy < 1$. Since both $1/x$ and xy are less then 1, then by the previous case $\ln xy + \ln 1/x = log\frac{xy}{x} = \ln y$. Replacing $\ln 1/x$ by $-\ln x$ one gets $\ln xy - \ln x = \ln y$ and finally $\ln xy = \ln x + \ln y$.

The last case, $x > 1$, $y < 1$, $xy > 1$ is proved by $\ln xy + \ln 1/y = \ln x$ and replacing $\ln 1/y$ by $-\ln y$. □

Base of logarithm Natural or hyperbolic logarithms are not the only logarithmic function. Other popular logarithms are decimal ones. In computer science one prefers binary logarithms. Different logarithmic functions are distinguished by its *bases*. The base of a logarithmic function $l(x)$ is defined as a number b for which $l(b) = 1$. Logarithms with the base b are denoted by $\log_b x$.

What is the base of natural logarithms? This is the second most important constant in mathematics (after π). It is an irrational number denoted by e which is equal to $2.71828182845905\ldots$. It was Euler who introduced these number and notation.

Well, e is the number such that the area of hyperbolic trapezium over $[1, e]$ is 1. Consider geometric progression q^n for $q = 1 + \frac{1}{n}$. All summands in the corresponding hyperbolic receding sum for this progression are equal to $\frac{q^{k+1}-q^k}{q^k} = q - 1 = \frac{1}{n}$. Hence the receding sum for the interval $[1, q^n]$ is equal to 1 and it is greater than $\log q^n$. Consequently $e > q^n$. The summands of the advancing sum in this case are equal to $\frac{q^{k+1}-q^k}{q^{k+1}} = 1 - \frac{1}{q} = \frac{1}{n+1}$. Hence the advancing sum for the interval $[1, q^{n+1}]$ is equal to 1. It is less than the corresponding logarithm. Consequently, $e < q^{n+1}$. Thus we have proved the following estimates for e:

$$\left(1 + \frac{1}{n}\right)^n < e < \left(1 + \frac{1}{n}\right)^{n+1} \tag{2.1.6}$$

We see that $\left(1 + \frac{1}{n}\right)^n$ rapidly tends to e as n tends to infinity.

Problems.

1. Prove that $\ln x/y = \ln x - \ln y$
2. Prove that $\ln 2 < 1$
3. Prove that $\ln 3 > 1$
4. Prove that $x > y$ implies $\ln x > \ln y$.
5. Is $\ln x$ bounded?
6. Prove that $\dfrac{1}{n+1} < \ln(1 + 1/n) < \dfrac{1}{n}$
7. Prove that $\dfrac{x}{1+x} < \ln(1 + x) < x$
8. Prove Fermat theorem for $a = 1/2, 1/3, 2/3$
9. Prove unboundedness of $\frac{n}{\ln n}$.
10. Compare $\left(1 + \frac{1}{n}\right)^n$ and $\left(1 + \frac{1}{n+1}\right)^{n+1}$
11. Prove monotonicity of $\frac{n}{\ln n}$.
12. Prove that $\sum_{k=2}^{n-1} \frac{1}{k} < \ln n < \sum_{k=1}^{n-1} \frac{1}{k}$
13. Prove that $\ln(1 + x) > x - \frac{x^2}{2}$
14. Estimate integral part of $\ln 1000000$
15. Prove that $\ln \dfrac{x+y}{2} \geq \dfrac{\ln x + \ln y}{2}$

16. Prove convergence $\sum_{k=1}^{\infty}(\frac{1}{k}-\ln(1+\frac{1}{k}))$

17. Prove that

$$\left(n+\frac{1}{2}\right)^{-1} \leq \ln\left(1+\frac{1}{n}\right) < \frac{1}{2}\left(\frac{1}{n}+\frac{1}{n+1}\right)$$

18. * Prove that

$$\frac{1}{1\cdot 2}+\frac{1}{3\cdot 4}+\frac{1}{5\cdot 6}+\cdots=\ln 2$$

2.2 Definite Integral.

On the contents of the lecture. Areas of curvilinear trapezia play an extraordinary important role in mathematics. They generate a key concept of calculus — the concept of *integral*.

Three Basic Rules For a nonnegative function f its integral $\int_a^b f(x)\,dx$ along the interval $[a,b]$ is defined just as the area of the curvilinear trapezium below the graph of f over $[a,b]$. We allow a function to take infinite values. Let us remark that changing of the value of function in one point does not affect the integral, because the area of the line is zero. That is why we allow the functions under consideration to be undefined in a finite number of points of the interval.

Immediately from the definition one gets the following three *basic rules of integration*:

Rule of constant	$\int_a^b f(x)\,dx = c(b-a)$,	if $f(x)=c$	for $x \in (a,b)$
Rule of inequality	$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$,	if $f(x) \leq g(x)$	for $x \in (a,b)$
Rule of partition	$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$		for $b \in (a,c)$

Partition. Let $|J|$ denote the length of an interval J. Let us say that a sequence $\{J_k\}_{k=1}^n$ of disjoint open subintervals of an interval I is a *partition* of I, if $\sum_{k=1}^n |I_k| = |I|$ The boundary of a partition $P = \{J_k\}_{k=1}^n$ is defined as the difference $I \setminus \cup_{k=1}^n J_k$ and is denoted ∂P.

For any finite subset S of an interval I, which contains the ends of I there is a unique partition of I, which has this set as the boundary. Such partition is called *generated* by S. For a monotone sequence $\{x_k\}_{k=0}^n$ the generated partition is $\{(x_{k-1}, x_k)\}_{k=1}^n$.

Piecewise constant functions. A function $f(x)$ is called *partially constant* on a partition $\{J_k\}_{k=1}^n$ of $[a,b]$ if it is constant on any J_k. The Rules of Constant and Partition immediately imply:

$$\int_a^b f(x)\,dx = \sum_{k=1}^n f(J_k)|J_k|. \tag{2.2.1}$$

Proof. Indeed, the integral splits into the sum of integrals over $J_k = [x_{k-1}, x_k]$, and the function takes the value $f(J_k)$ in (x_{k-1}, x_k). □

A function is called piecewise constant over an interval, if it is partially constant with respect to some finite partition of the interval.

Lemma 1. *Let f and g be piecewise constant functions over $[a,b]$. Then* $\int_a^b (f(x) \pm g(x))\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx$.

Proof. First, suppose $f(x) = c$ is constant on the interval (a, b). Let g takes the value g_k over the interval (x_k, x_{k+1}) for an exhausting $\{x_k\}_{k=0}^n$. Then $f(x)+g(x)$ takes values $(c+g_k)$ over (x_k, x_{k+1}). Hence $\int_a^b (f(x)+g(x))\,dx = \sum_{k=0}^{n-1}(c+g_k)|\Delta x_k|$ due to (2.2.1). Splitting this sum and applying (2.2.1) to the both summands, one gets $\sum_{k=0}^{n-1} c|\Delta x_k| + \sum_{k=0}^{n-1} g_k|\Delta x_k| = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$. This proves the case of constant f.

Now let f be partially constant on the partition generated by $\{x_k\}_{k=0}^n$. Then, by the partition rule, $\int_a^b (f(x) + g(x))\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} (f(x) + g(x))\,dx$. As f is constant on any (x_{k-1}, x_k), for any k one gets $\int_{x_{k-1}}^{x_k} (f(x) + g(x))\,dx = \int_{x_{k-1}}^{x_k} f(x)\,dx + \int_{x_{k-1}}^{x_k} g(x)\,dx$. Summing up these equalities one completes the proof of Lemma **1** for the sum.

The statement about differences follows from the addition formula applied to $g(x)$ and $f(x) - g(x)$. □

Lemma 2. *For any monotone nonnegative function f on the interval $[a, b]$ and for any $\varepsilon > 0$ there is such piecewise constant function f_ε such that $f_\varepsilon \leq f(x) \leq f_\varepsilon(x) + \varepsilon$.*

Proof. $f_\varepsilon(x) = \sum_{k=0}^{\infty} k\varepsilon[k\varepsilon \leq f(x) < (k+1)\varepsilon]$.

□

Theorem 3 (Addition Theorem). *Let f and g be nonnegative monotone functions defined on $[a, b]$. Then*

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \tag{2.2.2}$$

.

Proof. Let f_ε and g_ε be ε-approximations of f and g respectively provided by Lemma **2**. Set $f^\varepsilon(x) = f_\varepsilon(x) + \varepsilon$ and $g^\varepsilon(x) = g_\varepsilon(x) + \varepsilon$. Then $f_\varepsilon(x) \leq f(x) \leq f^\varepsilon(x)$ and $g_\varepsilon(x) \leq g(x) \leq g^\varepsilon(x)$ for $x \in (a, b)$. Summing and integrating these inequalities in different order gives

$$\int_a^b (f_\varepsilon(x) + g_\varepsilon(x))\,dx \leq \int_a^b (f(x) + g(x))\,dx \leq \int_a^b (f^\varepsilon(x) + g^\varepsilon(x))\,dx$$

$$\int_a^b f_\varepsilon(x)\,dx + \int_a^b g_\varepsilon(x)\,dx \leq \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \leq \int_a^b f^\varepsilon(x)\,dx + \int_a^b g^\varepsilon(x)\,dx.$$

Due to Lemma **1**, the left-hand sides of these inequalities coincide, as well as the right-hand sides. Hence the difference between the central parts does not exceed

$$\int_a^b (f^\varepsilon(x) - f_\varepsilon(x))\,dx + \int_a^b (g^\varepsilon(x) - g_\varepsilon(x))\,dx \leq 2\varepsilon(b-a). \tag{2.2.3}$$

Hence, for any positive ε

$$\left| \int_a^b (f(x) + g(x))\,dx - \int_a^b f(x)\,dx - \int_a^b g(x)\,dx \right| < 2\varepsilon(b-a). \tag{2.2.4}$$

This implies that the left-hand side vanishes.

□

Term by term integration of functional series.

Lemma 4. *Let $\{f_n\}_{n=1}^\infty$ be a sequence of nonnegative nondecreasing functions and p be a piecewise constant function. If $\sum\limits_{k=1}^\infty f_k(x) \geq p(x)$ for all $x \in [a,b]$ then $\sum\limits_{k=1}^\infty \int\limits_a^b f_k(x)\,dx \geq \int\limits_a^b p(x)\,dx$*

Proof. Let p be a piecewise constant function with respect to $\{x_i\}_{i=0}^n$. Choose any positive ε. Since $\sum\limits_{k=1}^\infty f_k(x_i) \geq p(c)$, eventually one has $\sum\limits_{k=1}^m f_k(x_i) > p(x_i) - \varepsilon$. Fix m such that this inequality holds simultaneously for all $\{x_i\}_{i=0}^n$. Let $[x_i, x_{i+1}]$ be an interval where $p(x)$ is constant. Then for any $x \in [x_i, x_{i+1}]$ one has these inequalities: $\sum\limits_{k=1}^m f_k(x) \geq \sum\limits_{k=1}^m f_k(x_k) > p(x_k) - \varepsilon = p(x) - \varepsilon$. Consequently for all $x \in [a,b]$ one has inequality $\sum\limits_{k=1}^m f_k(x) > p(x) - \varepsilon$. Taking integrals gives $\int\limits_a^b \sum\limits_{k=1}^m f_k(x)\,dx \geq \int\limits_a^b (p(x) - \varepsilon)\,dx = \int\limits_a^b p(x)\,dx - \varepsilon(b-a)$. By the Addition Theorem $\int\limits_a^b \sum\limits_{k=1}^m f_k(x)\,dx = \sum\limits_{k=1}^m \int\limits_a^b f_k(x)\,dx \leq \sum\limits_{k=1}^\infty \int\limits_a^b f_k(x)\,dx$. Therefore $\sum\limits_{k=1}^\infty \int\limits_a^b f_k(x)\,dx \geq \int\limits_a^b p(x)\,dx - \varepsilon(b-a)$ for any positive ε. This implies the inequality $\sum\limits_{k=1}^\infty \int\limits_a^b f_k(x)\,dx \geq \int\limits_a^b p(x)\,dx$. □

Theorem 5. *For any sequence $\{f_n\}_{n=1}^\infty$ of nonnegative nondecreasing functions on an interval $[a,b]$*

$$\int_a^b \sum_{k=1}^\infty f_k(x)\,dx = \sum_{k=1}^\infty \int_a^b f_k(x)\,dx \tag{2.2.5}$$

Proof. Since $\sum\limits_{k=1}^n f_k(x) \leq \sum\limits_{k=1}^\infty f_k(x)$ for all x, by integrating one gets

$$\int_a^b \sum_{k=1}^n f_k(x)\,dx \leq \int_a^b \sum_{k=1}^\infty f_k(x)\,dx.$$

By the the Addition Theorem the left-hand side is equal to $\sum\limits_{k=1}^{n}\int\limits_a^b f_k(x)\,dx$, which is a partial sum of $\sum\limits_{k=1}^{\infty}\int\limits_a^b f_k(x)\,dx$. Then by All-for-One one gets the inequality $\sum\limits_{k=1}^{\infty}\int\limits_a^b f_k(x)\,dx \leq \int\limits_a^b \sum\limits_{k=1}^{\infty} f_k(x)\,dx$.

To prove the opposite inequality for any positive ε, we apply Lemma **2** to find a piecewise constant function F_ε, such that $F_\varepsilon(x) \leq \sum\limits_{k=1}^{\infty} f_k(x)dx$ and $\int\limits_a^b \sum\limits_{k=1}^{\infty}(f_k(x) - F_\varepsilon(x))\,dx < \varepsilon$. On the other hand, by Lemma **4** one gets

$$\sum_{k=1}^{\infty}\int_a^b f_k(x)\,dx \geq \int_a^b F_\varepsilon(x)\,dx.$$

Together these inequalities imply $\sum\limits_{k=1}^{\infty}\int\limits_a^b f_k(x)\,dx + \varepsilon \geq \int\limits_a^b \sum\limits_{k=1}^{\infty} f_k(x)\,dx$. As the last inequality holds for all $\varepsilon > 0$, it holds also for $\varepsilon = 0$ □

Theorem 6 (Mercator,1668). *For any* $x \in (-1, 1]$ *one has*

$$\boxed{\log(1+x) = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}x^k}{k}} \tag{2.2.6}$$

Proof. Consider $x \in [0, 1)$. Since $\int\limits_0^x t^k\,dt = \frac{t^{k+1}}{k+1}$ due to the Fermat Theorem **2**, term-wise integration of the geometric series $\sum\limits_{k=0}^{\infty} t^k$ over interval $[0, x]$ for $x < 1$ gives $\int\limits_0^x \frac{1}{1-t}\,dt = \sum\limits_{k=0}^{\infty}\int\limits_0^x t^k\,dt = \sum\limits_{k=0}^{\infty}\frac{x^{k+1}}{k+1}$.

Lemma 7. $\int\limits_0^x \frac{1}{1-t}\,dt = \ln(1-x)$

Proof of Lemma. Construct a translation of the plane which transforms the curvilinear trapezium below $\frac{1}{1-t}$ over $[0, x]$ into the trapezium for $\ln(1-x)$. Indeed, the reflection of the plane $((x, y) \to (2-x, y))$ along the line $x = 1$ transforms this trapezium to the curvilinear trapezium under $\frac{1}{x-1}$ over $[2-x, 2]$. The parallel translation by 1 to the left of the latter trapezium $(x, y) \to (x-1, y)$ transform it just in logarithmic trapezium for $\ln(1-x)$. □

The Lemma proves the Mercator Theorem for negative x. To prove it for positive x, set $f_k(x) = x^{2k-1} - x^{2k}$. All functions f_k are nonnegative on $[0, 1]$ and $\sum\limits_{k=1}^{\infty} f_k(x) = \frac{1}{1+x}$. Term-wise integration of this equality over $[0, x]$ gives (2.2.6), modulo equality $\int\limits_0^x \frac{1}{1+t}\,dt = \int\limits_1^x \frac{1}{t}\,dt$. The latter is proved by parallel translation of the plane. Let us remark, that in the case $x = 1$ the series $\sum\limits_{k=1}^{\infty}\frac{(-1)^{k+1}x^k}{k}$ is not absolutely convergent, and under its sum we mean $\sum\limits_{k=1}^{\infty} \frac{1}{2k(2k-1)} = \sum\limits_{k=1}^{\infty} (\frac{1}{2k-1} - \frac{1}{2k})$. And the above proof proves just this fact. □

The arithmetic mean of Mercator's series evaluated at x and $-x$ gives *Gregory's Series*

$$\boxed{\frac{1}{2}\ln\frac{1+x}{1-x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \dots} \tag{2.2.7}$$

Gregory's series converges much faster than Mercator's one. For example, putting $x = \frac{1}{3}$ in (2.2.7) one gets

$$\ln 2 = \frac{2}{3} + \frac{2}{3\cdot 3^3} + \frac{2}{5\cdot 3^5} + \frac{2}{7\cdot 3^7} + \dots.$$

Problems.

1. Prove that $\left|\int\limits_a^b f(x)\,dx\right| \le \int\limits_a^b |f(x)|\,dx$

2. Prove the following formulas via piece-wise constant approximations:

(multiplication formula) $$\int\limits_a^b \lambda f(x)\,dx = \lambda \int\limits_a^b f(x)\,dx$$

(shift formula) $$\int\limits_a^b f(x)\,dx = \int\limits_{a+c}^{b+c} f(x-c)\,dx$$

(reflection formula) $$\int\limits_0^a f(x)\,dx = \int\limits_{-a}^{0} f(-x)\,dx$$

(compression formula) $$\int\limits_0^a f(x)\,dx = \frac{1}{k}\int\limits_0^{ka} f\left(\frac{x}{k}\right)\,dx$$

3. Evaluate $\int\limits_0^{2\pi} (\sin x + 1)\,dx$.

4. Prove the inequality $\int\limits_{-2}^{2} (2 + x^3 2^x)\,dx > 8$.

5. Prove $\int\limits_0^{2\pi} x(\sin x + 1)\,dx < 2\pi$.

6. $\int\limits_{100\pi}^{200\pi} \frac{x+\sin(x)}{x}\,dx \le 100\pi + \frac{1}{50}\pi$.

7. Denote by s_n the area of

$$\{(x,y) \mid 0 \le x \le 1, (1-x)\log n + x\ln(n+1) \le y \le \ln(1+x)\}$$

. Prove that $\sum\limits_{k=1}^{\infty} s_k < \infty$.

8. Prove that $\sum\limits_{k=1}^{2n} (-1)^{k+1}\frac{x^k}{k} < \ln(1+x) < \sum\limits_{k=1}^{2n+1} (-1)^{k+1}\frac{x^k}{k}$ for $x > 0$.

9. Compute the logarithms of the primes 2,3,5,7 with accuracy 0.01.

10. Evaluate $\int\limits_0^1 \sqrt{x}\, dx$.

11. * Evaluate $\int\limits_0^{\pi} \sin x\, dx$.

2.3 Stieltjes Integral

On the contents of the lecture. The Stieltjes relativization of integral makes the integral flexible. We learn the main transformations of integrals. They allow us to evaluate a lot of integrals.

Basic rules. *Parametric curve* is a mapping of an interval into the plane. In cartesian coordinates a parametric curve can be presented as a pair of functions $x(t), y(t)$. The first function $x(t)$ represents the value of abscises at the moment t, and the second $y(t)$ is ordinate at the same moment. We define the integral $\int_a^b f(t)\,dg(t)$ for a nonnegative function f called *integrand* and with respect to a nondecreasing *continuous* function g, called *differand* as the area below the curve $f(t), g(t) \mid t \in [a,b]$.

A monotone function f is called *continuous* over interval $[a,b]$ if it takes all intermediate values, that is the image $f[a,b]$ of $[a,b]$ coincides with $[f(a), f(b)]$. If it is not continuous for some $y \in [f(a), f(b)] \setminus f[a,b]$, there is a point $x(y) \in [a,b]$ with the following property: $f(x) < y$ if $x < x(y)$ and $f(x) > y$ if $x > x(y)$. Let us define a *generalized preimage* $f^{[-1]}(y)$ of a point $y \in [f(a), f(b)]$ either as its usual preimage $f^{-1}(y)$ if it is not empty, or as $x(y)$ in the opposite case.

Now the curvilinear trapezium below the curve $f(t), g(t)$ over $[a,b]$ is defined as $\{(x,y) \mid 0 \le y \le g(f^{[-1]}(x))\}$.

The basic rules for relative integrals transforms into

Rule of constant	$\int_a^b f(t)\,dg(t) = c(g(b) - g(a)),$	if $f(t) = c$	for $t \in (a,b)$
Rule of inequality	$\int_a^b f_1(t)\,dg(t) \le \int_a^b f_2(t)\,dg(t),$	if $f_1(t) \le f_2(t)$	for $t \in (a,b)$
Rule of partition	$\int_a^c f(t)\,dg(t) = \int_a^b f(t)\,dg(t) + \int_b^c f(t)\,dg(t)$		for $b \in (a,c)$

Addition theorem. The proof of others properties of integral is based on the piece-wise constant functions. For any number x, let us define its ε-integral part as $\varepsilon[x/\varepsilon]$. Immediately from the definition one gets.

Lemma 1. *For any monotone nonnegative function f on the interval $[a,b]$ and for any $\varepsilon > 0$, the function $[f]_\varepsilon$ is piecewise constant such that $[f(x)]_\varepsilon \le f(x) \le [f(x)]_\varepsilon + \varepsilon$ for all x.*

Theorem 2 (on multiplication). *For any nonnegative monotone f, and continuous nondecreasing g and any positive constant c one has*

$$\int_a^b cf(x)\,dg(x) = c\int_a^b f(x)\,dg(x) = \int_a^b f(x)\,dcg(x) \tag{2.3.1}$$

Proof. For the piecewise constant $f_\varepsilon = [f]_\varepsilon$, the proof is by a direct calculation. Hence

$$\int_a^b cf_\varepsilon(x)\,dg(x) = c\int_a^b f_\varepsilon(x)\,dg(x) = \int_a^b f_\varepsilon(x)\,dcg(x) = I_\varepsilon. \tag{2.3.2}$$

Now let us estimate the differences between integrals from (2.3.1) and their approximations from (2.3.2). For example, for the right-hand side integrals one has:

$$\int_a^b f\,dcg - \int_a^b f_\varepsilon\,dcg = \int_a^b (f - f_\varepsilon)\,dcg \leq \int_a^b \varepsilon\,dcg = \varepsilon(cg(b) - cg(a)). \tag{2.3.3}$$

Hence $\int_a^b f\,dcg = I_\varepsilon + \varepsilon_1$, where $\varepsilon_1 \leq c\varepsilon(g(b) - g(a))$. The same argument proves $c\int_a^b f\,dg = I_\varepsilon + \varepsilon_2$ and $\int_a^b cf\,dg = I_\varepsilon + \varepsilon_3$, where $\varepsilon_2, \varepsilon_3 \leq c\varepsilon(g(b) - g(a))$. Then the pairwise differences between integrals of (2.3.1) do not exceed $2c\varepsilon(g(b) - g(a))$. Consequently they are less than any positive number, that is, they are zeros.

□

Theorem 3 (Addition Theorem). *Let f_1, f_2 be nonnegative monotone functions and g_1, g_2 be a nondecreasing continuous functions over $[a, b]$, then*

$$\int_a^b (f_1(t) + f_2(t))\,dg_1(t) = \int_a^b f_1(t)\,dg_1(t) + \int_a^b f_2(t)\,dg_1(t) \tag{2.3.4}$$

$$\int_a^b f_1(t)\,d(g_1(t) + g_2(t)) = \int_a^b f_1(t)\,dg_1(t) + \int_a^b f_1(t)\,dg_2(t) \tag{2.3.5}$$

Proof. For piecewise constant integrands the both equalities follows from the Rule of Constant and the Rule of Partition. To prove (2.3.4) replace f_1 and f_2 in both parts by $[f_1]_\varepsilon$ and $[f_2]_\varepsilon$. We get equality and denote by I_ε the common value of both sides of this equality. Then by (2.3.3) both integral on the right-hand side differs from they approximation at most by $\varepsilon(g_1(b) - g_1(a))$, therefore the right-hand side of (2.3.4) differs from I_ε at most by $2\varepsilon(g_1(b) - g_1(a))$. The same is true for the left-hand side of (2.3.4). This follows immediately from (2.3.3) in case $f = f_1 + f_2$, $f_\varepsilon = [f_1]_\varepsilon + [f_2]_\varepsilon$ and $g = g_1$. Consequently, the difference between left-hand and right-hand sides of the (2.3.4) does not exceed $4\varepsilon(g_1(b) - g_1(a))$. As ε can be chosen arbitrarily small this difference has to be zero.

The proof of (2.3.5) is even simpler. Denote by I_ε the common value of both parts of (2.3.5) where f_1 is changed by $[f_1]_\varepsilon$. By (2.3.3) one can estimate the differences between integrals of (2.3.5) and their approximations as being $\leq \varepsilon(g_1(b) + g_2(b) - g_1(a) - g_2(a))$ for the left-hand side, and as $\leq \varepsilon(g_1(b) - g_1(a))$ and $\leq \varepsilon(g_2(b) - g_2(a))$ for corresponding integrals of the right-hand side of (2.3.5). So both sides differs from I_ε at most on $\leq \varepsilon(g_1(b) - g_1(a) + g_2(b) - g_2(a))$. Hence they difference vanishes. □

Differential forms An expression of the type $f_1dg_1 + f_2dg_2 + \cdots + f_ndg_n$ is called a *differential form*. One can add differential forms and multiply it by functions. The integral of a differential form $\int_a^b (f_1\,dg_1 + f_2dg_2 + \cdots + f_ndg_n)$

is defined as sum of integrals $\sum\limits_{k=1}^{n}\int\limits_a^b f_k\,dg_k$. Two differential forms are called equivalent on the interval $[a,b]$ if their integrals are equal for all subintervals of $[a,b]$. For the sake of brevity we denote the differential form $f_1dg_1 + f_2dg_2 + \cdots + f_ndg_n$ by FdG, where $F = \{f_1, \dots, f_n\}$ is a collection of integrands and $G = \{g_1, \dots, g_n\}$ is a collection of differands.

Theorem 4 (on multiplication). *Let FdG and $F'dG'$ be two differential forms with positive increasing integrand and continuous increasing differands, which are equivalent on $[a,b]$. Then their products by any increasing function f on $[a,b]$ are equivalent on $[a,b]$ too.*

Proof. If f is constant then the statement follows from the multiplication formula. If f is piece-wise constant, then divide $[a,b]$ into intervals where it is constant and prove the equality for parts and after collect the results by the Partition Rule. In the general case,

$$0 \le \int\limits_a^b fF\,dG - \int\limits_a^b [f]_\varepsilon F\,dG \le \int\limits_a^b \varepsilon F\,dG = \varepsilon\int\limits_a^b F\,dG.$$

Since $\int\limits_a^b [f]_\varepsilon F'\,dG' = \int\limits_a^b [f]_\varepsilon F\,dG$, one concludes that

$$\left|\int\limits_a^b fF'\,dG' - \int\limits_a^b fF\,dG\right| \le \varepsilon\int\limits_a^b F\,dG + \varepsilon\int\limits_a^b F'\,dG'$$

The right-hand side of this inequality can be made arbitrarily small. Hence the left-hand side is 0. □

Integration by parts.

Theorem 5. *If f and g are continuous nondecreasing nonnegative functions on $[a,b]$ then $d(fg)$ is equivalent to $fdg + gdf$.*

Proof. Consider $[c,d] \subset [a,b]$. The integral $\int\limits_c^d f\,dg$ represents the area below the curve $(f(t), g(t))_{t\in[c,d]}$. And the integral $\int\limits_c^d g\,df$ represents the area on the left of the same curve. Its union is equal to $[0, f(d)] \times [0, g(d)] \setminus [0, f(c)] \times [0, g(c)]$. The area of this union is equal to $(f(d)g(d) - f(c)g(c) = \int\limits_c^d dfg$. On the other hand the area of this union is the sum of the areas of curvilinear trapezia representing the integrals $\int\limits_c^d f\,dg$ and $\int\limits_c^d g\,df$. □

Change of variable Consider a Stieltjes integral $\int\limits_a^b f(\tau)\,dg(\tau)$ and suppose there is a continuous nondecreasing mapping $\tau\colon [t_0, t_1] \to [a,b]$, such that

$\tau(t_0) = a$ and $\tau(t_1) = b$. The composition $g(\tau(t))$ is a continuous nondecreasing function and the curve $\{(f(\tau(t), g(\tau(t))) \mid t \in [t_0, t_1]\}$ just coincides with the curve $\{(f(\tau), g(\tau)) \mid \tau \in [a, b]$. Hence, the following equality holds; it is known as the Formula of Change of Variable:

$$\int_{t_0}^{t_1} f(\tau(t))\, dg(\tau(t)) = \int_{\tau(t_0)}^{\tau(t_1)} f(\tau)\, dg(\tau) \tag{2.3.6}$$

For differentials this means that the equality $F(x)dG(x) = F'(x)dG'(x)$ conserves if one substitutes instead of an independent variable x a function.

Differential Transformations

dx^n. Integration by parts for $f(t) = g(t) = t$ gives $dt^2 = tdt + tdt$. Hence $tdt = \frac{t^2}{2}$. If we already know that $dx^a = adx^{a-1}$, then $dx^{a+1} = d(xx^a) = xdx^a + x^a dx = axx^{a-1}dx + x^a dx = (a+1)x^a dx$. This proves the Fermat Theorem for natural a.

$d\sqrt[n]{x}$. To evaluate $d\sqrt[n]{x}$ substitute $x = y^n$ into equality $dy^n = ny^{n-1}dy$. One gets $dx = \frac{nx\,d\sqrt[n]{x}}{\sqrt[n]{x}}$, hence $d\sqrt[n]{x} = \frac{1}{n}\frac{\sqrt[n]{x}dx}{x}$.

$\ln x dx$. We know $d\ln x = \frac{dx}{x}$. Integration by parts gives $\ln x dx = d(x\ln x) - xd\ln x = d(x\ln x) - dx = d(x\ln x - x)$.

Problems.

1. $dx^{2/3}$
2. dx^{-1}
3. $x \ln x\, dx$
4. $d\ln^2 x$
5. $\ln^2 x\, dx$
6. de^x
7. Investigate the convergence of $\sum\limits_{k=1}^{\infty} \frac{1}{k\ln k}$.

2.4 Asymptotics of sums

On the contents of the lecture. We become at last acquainted with the fundamental concept of *limit*. We extend the notion of sum of series and discover that change of the order of summands can affect the ultimate sum. Finally we derive the famous Stirling formula for $n!$.

Asymptotic formulas. Mercator series shows how useful series can be for evaluating of integrals. In this lecture we will use integrals to evaluate both partial and ultimate sums of series. Rarely one has an explicit formula for partial sums of a series. There are lots of important cases where such a formula does not exist. For example, it is known that partial sums of the Euler series cannot be expressed as a finite combination of elementary functions. When an explicit formula is not available, one tries to find a so-called *asymptotic formulas*. An asymptotic formula for a partial sum S_n of a series is a formula of the type $S_n = f(n) + R(n)$ where f is a known function called *principal part* and $R(n)$ is a *reminder*, which is small, in some sense, with respect to the principal part. Today we will get an asymptotic formula for partial sums of the harmonic series.

Infinitesimally small sequences. The simplest asymptotic formula has a constant as its principal part and an infinitesimally small remainder. One says that a sequence $\{z_k\}$ is *infinitesimally small* and writes $\lim z_k = 0$, if z_k tends to 0 as n tends to infinity. That is for any positive ε eventually (i.e. beginning with some n) holds $|z_k| < \varepsilon$. With Iverson notation, this definition can be expressed in the following clear form:

$$[\{z_k\}_{k=1}^{\infty} \text{ is infinitesimally small }] = \prod_{m=1}^{\infty} 2\left|\sum_{n=1}^{\infty}(-1)^n \prod_{k=1}^{\infty}[m[k>n]|z_k|<1]\right|$$

Three basic properties of infinitesimally small sequences immediately follow from the definition:

- if $\lim a_k = \lim b_k = 0$ then $\lim(a_k + b_k) = 0$
- if $\lim a_k = 0$ then $\lim a_k b_k = 0$ for any bounded sequence $\{b_k\}$
- if $a_k \leq b_k \leq c_k$ for all k and $\lim a_k = \lim c_k = 0$, then $\lim b_k = 0$

The third property is called *squeeze rule*.

Today we needs just one properties of infinitesimally small, that is

Theorem 1 (Addition theorem). *If sequences $\{a_k\}$ and $\{b_k\}$ are infinitesimally small, than its sum and its difference are infinitesimally small too.*

Proof. Let ε be a positive number. Then $\varepsilon/2$ also is positive number. And by definition of infinitesimally small, the inequalities $|a_k| < \varepsilon/2$ and $|b_k| < \varepsilon/2$ hold eventually beginning with some n. Then for $k > n$ one has $|a_k \pm b_k| \leq |a_k| + |b_k| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Limit of sequence

Definition. *A sequence $\{z_k\}$ of (complex) numbers converges to a number z if $\lim z - z_k = 0$. The number z is called limit of the sequence $\{z_k\}$ and denoted by $\lim z_k$.*

The infinite sum represents a particular case of a limit as demonstrate the following.

Theorem 2. *The partial sums of an absolute convergent series $\sum_{k=1}^{\infty} z_k$ converges to its sum.*

Proof. $|\sum_{k=1}^{n-1} z_k - \sum_{k=1}^{\infty} z_k| = |\sum_{k=n}^{\infty} z_k| \leq \sum_{k=n}^{\infty} |z_k|$. Since $\sum_{k=1}^{\infty} |z_k| > \sum_{k=1}^{\infty} |z_k| - \varepsilon$, there is a partial sum such that $\sum_{k=1}^{n-1} |z_k| > \sum_{k=1}^{\infty} |z_k| - \varepsilon$. Then for all $m \geq n$ one has $\sum_{k=m}^{\infty} |z_k| \leq \sum_{k=n}^{\infty} |z_k| < \varepsilon$. □

Conditional convergence. The concept of limit of sequence leads to the notion of convergence generalizing the absolute convergence.

Series $\sum_{k=1}^{\infty} a_k$ is called (conditionally) *convergent* if $\lim \sum_{k=1}^{n} a_k = A + \alpha_n$, where $\lim \alpha_n = 0$. The number A is called its ultimate sum.

The following theorem gives a lot of examples of conditionally convergent series, which are not absolutely convergent. By $[[n]]$ we denote the even part of the number n i.e. $[[n]] = 2[n/2]$

Theorem 3 (Leibniz). *For any of positive decreasing infinitesimally small sequence $\{a_n\}$, the series $\sum_{k=1}^{\infty} (-1)^{k+1} a_k$ converges.*

Proof. Denote the difference $a_k - a_{k+1}$ by Δa_k. The series $\sum_{k=1}^{\infty} \Delta a_{2k-1}$ and $\sum_{k=1}^{\infty} \Delta a_{2k}$ are positive and convergent, because their term-wise sum is $\sum_{k=1}^{\infty} \Delta a_k = a_1$. Hence $S = \sum_{k=1}^{\infty} \Delta a_{2k-1} \leq a_1$. Denote by S_n the partial sum $\sum_{k=1}^{n-1} (-1)^{k+1} a_k$. Then $S_{2n} = \sum_{k=1}^{n-1} \Delta a_{2n-1} = S + \alpha_n$, where $\lim \alpha_n = 0$. Then $S_n = S_{[[n]]} + a_n[n \text{ is odd}] + \alpha_{[[n]]}$. As $a_n[n \text{ is odd}] + \alpha_{[[n]]}$ is infinitesimally small, this implies the theorem. □

Lemma 4. *Let f be a non-increasing nonnegative function. Then the series $\sum_{k=1}^{\infty} (f(k) - \int_k^{k+1} f(x)\,dx)$ is positive and convergent and has sum $c_f \leq f(1)$.*

Proof. The integration of inequalities $f(k) \geq f(x) \geq f(k+1)$ over $[k, k+1]$ gives $f(k) \geq \int_k^{k+1} f(x)\,dx \geq f(n+1)$. This proves the positivity of the series and allows majorize it by the telescopic series $\sum_{k=1}^{\infty} (f(k) - f(k+1)) = f(1)$. □

Theorem 5 (integral test on convergence). *If a nonnegative function $f(x)$ decreases monotonically on $[1,+\infty)$, then $\sum\limits_{k=1}^{\infty} f(k)$ converges iff $\int\limits_1^{\infty} f(x)\,dx < \infty$.*

Proof. Since $\int\limits_1^{\infty} f(x)\,dx = \sum\limits_{k=1}^{\infty} \int\limits_k^{k+1} f(x)\,dx$, one has $\sum\limits_{k=1}^{\infty} f(k) = c_f + \int\limits_1^{\infty} f(x)\,dx$. $\square$

Euler constant. The sum $\sum\limits_{k=1}^{\infty} \left(\frac{1}{k} - \ln(1+\frac{1}{k})\right)$, which is c_f for $f(x) = \frac{1}{x}$, is called *Euler's constant* and denoted by γ. Its first ten digits are 0.5772156649....

Harmonic numbers The sum $\sum\limits_{k=1}^{n} \frac{1}{k}$ is denoted H_n and is called the n-th *harmonic number*.

Theorem 6. $H_n = \ln n + \gamma + o_n$ *where* $\lim o_n = 0$

Proof. Since $\ln n = \sum\limits_{k=1}^{n-1}(\ln(k+1) - \ln k) = \sum\limits_{k=1}^{n-1} \ln(1+\frac{1}{k})$, one has $\ln n + \sum\limits_{k=1}^{n-1} \left(\frac{1}{k} - \ln(1+\frac{1}{k})\right) = H_{n-1}$. But $\sum\limits_{k=1}^{n-1} \left(\frac{1}{k} - \ln(1+\frac{1}{k})\right) = \gamma + \alpha_n$, where $\lim \alpha_n = 0$. Therefore $H_n = \ln n + \gamma + (\frac{1}{n} + \alpha_n)$. $\square$

Alternating harmonic series The *alternating harmonic series* $\sum\limits_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ is a conditionally convergent series due to Leibniz's theorem, and it is not absolutely convergent. To find its sum we apply our Theorem **6** on the asymptotics of harmonic numbers.

Denote by $S_n = \sum\limits_{k=1}^{n} \frac{(-1)^{k+1}}{k}$ the partial sum. Then $S_n = H'_n - H''_n$, where $H'_n = \sum\limits_{k=1}^{n} \frac{1}{k}[k \text{ is odd}]$ and $H''_n = \sum\limits_{k=1}^{n} \frac{1}{k}[k \text{ is even}]$. Since $H''_{2n} = \frac{1}{2}H_n$ and $H'_{2n} = H_{2n} - H''_{2n} = H_{2n} - \frac{1}{2}H_n$ one gets

$$(2.4.1)\quad S_{2n} = H_{2n} - \frac{1}{2}H_n - \frac{1}{2}H_n = H_{2n} - H_n = \\ \ln 2n + \gamma + o_{2n} - \ln n - \gamma - o_n = \ln 2 + (o_{2n} - o_n).$$

Consequently $S_n = \ln 2 + (o_{[[n]]} - o_{[n/2]} + \frac{(-1)^{n+1}}{n}[n \text{ is odd}])$. As the sum in brackets is infinitesimally small, one gets

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

The same arguments for a permutated alternating harmonic series give

$$(2.4.2)\qquad 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} \dots = \frac{3}{2}\log 2.$$

Indeed, in this case its $3n$-th partial sum is $S_{3n} = H'_{4n} - H''_{2n} = H_{4n} - \frac{1}{2}H_{2n} - \frac{1}{2}H_n = \log 4n + \gamma + o_{4n} - \frac{1}{2}(\log 2n + \gamma + o_{2n} + \log n + \gamma + o_n) = \log 4 - \frac{1}{2}\log 2 + o'_n = \frac{3}{2}\log 2 + o'_n$, where $\lim o'_n = 0$. Since the difference between S_n and S_{3m} where $m = [n/3]$ is infinitesimally small, this proves (2.4.2).

Stirling's Formula. We will try to estimate $\ln n!$. The integration of inequalities $\ln[x] \leq \ln x \leq \ln[x+1]$ over $[1,n]$ gives $\ln(n-1)! \leq \int_1^n \ln x\,dx \leq \ln n!$. Let us estimate the difference D between $\int_1^n \ln x\,dx$ and $\frac{1}{2}(\ln n! + \ln(n-1)!)$.

(2.4.3)

$$D = \int_1^n (\ln x - \frac{1}{2}(\ln[x] + \ln[x+1]))\,dx = \sum_{k=1}^{n-1} \int_0^1 \left(\ln(k+x) - \ln\sqrt{k(k+1)}\right)\,dx.$$

To prove that all summands on the left-hand side are nonnegative, we apply the following general lemma.

Lemma 7. $\int_0^1 f(x)\,dx = \int_0^1 f(1-x)\,dx$ *for any function*

Proof. The reflection of the plane across the line $y = \frac{1}{2}$ transforms the curvilinear trapezium of $f(x)$ over $[0,1]$ into curvilinear trapezium of $f(1-x)$ over $[0,1]$. □

Lemma 8. $\int_0^1 \ln(k+x)\,dx \geq \ln\sqrt{k(k+1)}$

Proof. Due to Lemma **7** one has

$$\int_0^1 \ln(k+x)\,dx = \int_0^1 \ln(k+1-x)\,dx = \int_0^1 \frac{1}{2}(\ln(k+x) + \ln(k+1-x))\,dx =$$

$$= \int_0^1 \ln\sqrt{(k+x)(k+1-x)}\,dx = \int_0^1 \ln\sqrt{k(k+1)+x-x^2}\,dx \geq$$

$$\int_0^1 \ln\sqrt{k(k+1)}\,dx = \ln\sqrt{k(k+1)}$$

□

Integration of the inequality $\ln(1+x/k) \leq x/k$ over $[0,1]$ gives

$$\int_0^1 \ln(1+x/k)\,dx \leq \int_0^1 \frac{x}{k}\,dx = \frac{1}{2k}.$$

This estimate together with the inequality $\ln(1+1/k) \geq 1/(k+1)$ allow us to estimate the summands from the right-hand side of (2.4.3) in the following way:

(2.4.4)

$$\int_0^1 \ln(k+x) - \ln\sqrt{k(k+1)}\,dx = \int_0^1 \ln(k+x) - \ln k - \frac{\ln(k+1) - \ln k}{2}\,dx =$$

$$= \int_0^1 \ln\left(1+\frac{x}{k}\right) - \frac{1}{2}\ln\left(1+\frac{1}{k}\right)\,dx \leq \frac{1}{2k} - \frac{1}{2(k+1)}$$

We see that $D_n \leq \sum_{k=1}^{\infty} \frac{1}{2k} - \frac{1}{2(k+1)} = \frac{1}{2}$ for all n. Denote by D_∞ the sum (2.4.3) for infinite n. Then $R_n = D_\infty - D_n = \frac{\theta}{2n}$ for some nonnegative $\theta < 1$, and we get

$$(2.4.5)\quad D_\infty - \frac{\theta}{2n} = \int_1^n \ln x\,dx - \frac{1}{2}(\ln n! + \ln(n-1)!) = \int_1^n \ln x\,dx - \ln n! + \frac{\ln n}{2}$$

Substituting in (2.4.5) the value of the integral $\int_1^n \ln x\,dx = \int_1^n d(x\ln x - x) = (n\ln n - n) - (1\ln 1 - 1) = n\ln n - n + 1$, one gets

$$(2.4.6)\qquad \ln n! = n\ln n - n + \frac{\ln n}{2} + (1 - D_\infty) + \frac{\theta}{2n}$$

Now we know that $1 \geq (1 - D_\infty) \geq \frac{1}{2}$, but it is possible to evaluate the value of D_∞ with more accuracy. Later we will prove that $1 - D_\infty = \sqrt{2\pi}$.

Problems.

1. Does $\sum_{k=1}^{\infty} \sin k$ converge?
2. Does $\sum_{k=1}^{\infty} \sin k^2$ converge?
3. Evaluate $1 + \frac{1}{2} - \frac{2}{3} + \frac{1}{4} + \frac{1}{5} - \frac{2}{6} + \cdots - \frac{2}{3n} + \frac{1}{3n+1} + \frac{1}{3n+2} \cdots$
4. If $\lim \frac{a_{n+1}}{a_n} < 1$, then $\sum_{k=1}^{\infty} a_k$ converge.
5. If $\sum_{k=1}^{\infty} |a_k - a_{k-1}| < \infty$, then $\{a_k\}$ converges.
6. Prove convergence $\sum_{k=1}^{\infty} \dfrac{(-1)^{[\sqrt{k}]}}{k}$
7. Convergence $\sum_{k=1}^{\infty} \frac{1}{\ln^3 k}$
8. Convergence $\sum_{k=1}^{\infty} \frac{1}{k \ln k \sqrt{\ln \ln k}}$
9. Convergence $\sum_{k=1}^{\infty} \frac{1}{k \ln k (\ln \ln k)^2}$
10. Convergence $\sum_{k=1}^{\infty} \frac{1}{k \ln k}$ and asymptotic formula?
11. Convergence $\sum_{k=1}^{\infty} \frac{1}{k \ln^2 k}$

12. Which partial sum of above series is 0.01 close to its ultimate sum?

13. Evaluate $\sum_{k=1}^{\infty} \frac{1}{k \ln^2 k}$ with precision 0.01

14. Evaluate $\int_1^3 \ln x \, d[x]$

15. Express the Stirling constant via the Wallis product $\frac{\pi}{2} = \prod_{n=1}^{\infty} \frac{2n}{2n-1} \frac{2n}{2n+1}$

2.5 Quadrature of circle

On the contents of the lecture. We extend the concept of the integral to complex function. We evaluate a very important integral $\oint \frac{dz}{z}$, by applying Archimedes theorem on the area of circular sector. As a consequence, we evaluate the Wallis product and Stirling Constant.

Definition of complex integral. To specify an integral of a complex function one has to indicate not only its limits, but also the *path of integration*. A path of integration is a mapping $p\colon [a,b] \to \mathbb{C}$, of an interval $[a,b]$ of the real line into complex plane. The integral of a complex differential form $f dg$ (here f and g are complex functions of complex variable) along the path p is defined via separate integration of different combinations of real and imaginary parts in the following way:

$$\int_a^b \operatorname{Re} f(p(t))\, d \operatorname{Re} g(p(t)) - \int_a^b \operatorname{Im} f(p(t))\, d \operatorname{Im} g(p(t)) \\ + i \int_a^b \operatorname{Re} f(p(t))\, d \operatorname{Im} g(p(t)) + i \int_a^b \operatorname{Im} f(p(t))\, d \operatorname{Re} g(p(t)) \tag{2.5.1}$$

Two complex differential form are called equal if their integrals coincide for all paths. So, the definition above can be written shortly as $f dg = \operatorname{Re} f d \operatorname{Re} g - \operatorname{Im} f d \operatorname{Im} g + i \operatorname{Re} f d \operatorname{Im} g + i \operatorname{Im} f d \operatorname{Re} g$.

The integral $\int \frac{dz}{z}$. The Integral is the principal concept of Calculus and $\int \frac{dz}{z}$ is the principal integral. Let us evaluate it along the path $p(t) = \cos t + i \sin t$, $t \in [0, \phi]$, which goes along the arc of the circle of the length $\phi \leq \pi/2$. Since $\frac{1}{\cos t + i \sin t} = \cos t - i \sin t$, one has

$$\int_p \frac{1}{z}\, dz = \int_0^\phi \cos t\, d \cos t + \int_0^\phi \sin t\, d \sin t - i \int_0^\phi \sin t\, d \cos t + i \int_0^\phi \cos t\, d \sin t \tag{2.5.2}$$

Its real part transforms into $\int_0^\phi \frac{1}{2}\, d \cos^2 t + \int_0^\phi \frac{1}{2}\, d \sin^2 t = \int_0^\phi \frac{1}{2}\, d(\cos^2 t + \sin^2 t) = \int_0^\phi \frac{1}{2}\, d1 = 0$. An attentive reader has to object: the integrals were defined only for differential forms with non-decreasing differand, while $\cos t$ decreases.

Sign rule. Let us define the integral for any differential form $f dg$ with any continuous monotone differand g and any integrand f of a constant sign (i.e, non-positive or non-negative). The definition relies on the following *Sign Rule*

$$\int_a^b -f\, dg = -\int_a^b f\, dg = \int_a^b f\, d(-g) \tag{2.5.3}$$

If f is of constant sign, and g is monotone, then among the forms fdg, $-fdg$, $fd(-g)$ and $-fd(-g)$ there is just one with non-negative integrand and non-decreasing differand. For this form, the integral was defined earlier, for the other cases it is defined by the Sign Rule.

Thus the integral of a negative function against an increasing differand and the integral of a positive function against a decreasing differand are negative. And the integral of a negative function against a decreasing differand is positive.

The Sign Rule agrees with the Constant Rule: the formula $\int_a^b c\,dg = c(g(b) - g(a))$ remains true either for negative c or decreasing g.

The Partition Rule also is not affected by this extension of Integral.

The Inequality Rule takes the following form: if $f_1(x) \leq f_2(x)$ for all $x \in [a, b]$ then $\int_a^b f_1(x)\,dg(x) \leq \int_a^b f_2(x)\,dg(x)$ for non-decreasing differand and $\int_a^b f_1(x)\,dg(x) \geq \int_a^b f_2(x)\,dg(x)$ for non-increasing g.

Change of variable. Now all integrals in (2.5.2) are defined. The next objection concerns transformation $\cos t d\cos t = \frac{1}{2} d\cos^2 t$. This transformation based on a decreasing change of variable $x = \cos t$ in $dx^2/2 = xdx$. But what happens with an integral when one applies a decreasing change of variable. The curvilinear trapezium, which represents the integral, does not change at all under any change of variable, even for a non-monotone one. Hence the only thing that may happen is a change of sign. And the sign changes by the Sign Rule, simultaneously on the both sides of equality $dx^2/2 = xdx$. If the integrals of xdx and dx^2 were positive, both integrals of $\cos t d\cos t$ and $\cos^2 t$ are negative and have the same absolute value. These arguments work in the general case:

The decreasing change of variable reverses the sign of the integral.

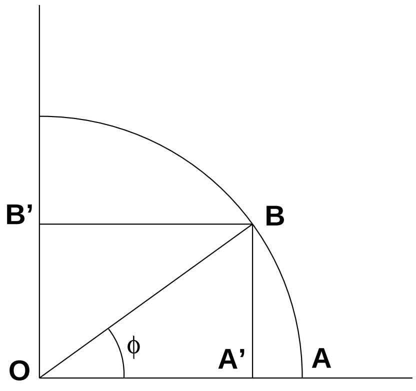
Addition Formula. The next question concerns legitimacy of addition of differentials, which appeared in the calculation $d\cos^2 t + d\sin^2 t = d(\cos^2 t + \sin^2 t) = 0$, where differands are not *comonotone:* $\cos t$ decreases, while $\sin t$ increases. The addition formula in its full generality will be proved in the next lecture, but this special case is not difficult to prove. Our equality is equivalent to $d\sin^2 t = -d\cos^2 t$. By the Sign Rule $-d\cos^2 t = d(-\cos^2 t)$, but $-\cos^2 t$ is increasing. And by the Addition Theorem $d(-\cos^2 t + 1) = d(-\cos^2 t) + d1 = d(-\cos^2 t)$. But $-\cos^2 t + 1 = \sin^2 t$. Hence our evaluation of the real part of (2.5.2) is justified.

Trigonometric integrals. We proceed to the evaluation of the imaginary part of (2.5.2), which is $\cos t\,d\sin t - \sin t\,d\cos t$. This is a simple geometric problem.

The integral of $\sin t\,d\cos t$ is negative as $\cos t$ is decreasing on $[0, \frac{\pi}{2}]$, and its absolute value is equal to the area of the curvilinear triangle $A'BA$, which is obtained from the circular sector OBA with area $\phi/2$ by deletion of the triangle $OA'B$, which has area $\frac{1}{2}\cos\phi\sin\phi$. Thus $\int_0^\phi \sin t\,d\cos t$ is $\phi/2 - \frac{1}{2}\cos\phi\sin\phi$

The integral of $\cos t\, d \sin t$ is equal to the area of curvilinear trapezium $OB'BA$. The latter consists of a circular sector OBA with area $\phi/2$ and a triangle $OB'B$ with area $\frac{1}{2}\cos\phi\sin\phi$. Thus $\int\limits_0^\phi \cos t\; d\sin t = \phi/2 + \frac{1}{2}\cos\phi\sin\phi$

As result we get $\int\limits_p \frac{1}{z}\, dz = i\phi$. This result has a lot of consequences. But today we restrict our attention to integrals of $\sin t$ and $\cos t$.



Multiplication of differentials. We have proved

$$\cos t\, d \sin t - \sin t\, d \cos t = dt. \tag{2.5.4}$$

Multiplying this equality by $\cos t$, one gets

$$\cos^2 t\, d \sin t - \sin t \cos t\, d \cos t = \cos t\, dt.$$

Replacing $\cos^2 t$ by $(1 - \sin^2 t)$ and moving $\cos t$ into the differential, one transforms the left-hand side as

$$d \sin t - \sin^2 t\, d \sin t - \frac{1}{2} \sin t\, d \cos^2 t = d \sin t - \frac{1}{2} \sin t\, d \sin^2 t - \frac{1}{2} \sin t\, d \cos^2 t.$$

We already know that $d \sin^2 t + d \cos^2 t$ is zero. Now we have to prove the same for the product of this form by $\frac{1}{2}\sin t$. The arguments are the same: we multiply by $\frac{1}{2}\sin t$ the equivalent equality $d\sin^2 t = d(-\cos^2 t)$ whose differands are increasing. This is a general way to extend the theorem on multiplication of differentials to the case of any monotone functions. We will do it later. Now we get just $d \sin t = \cos t dt$.

Further, the multiplication of the left-hand side of (2.5.4) by $\sin t$ gives

$$\sin t \cos t d \sin t - \sin^2 t d \cos t = \frac{1}{2} \cos t d \sin^2 t - d \cos t + \frac{1}{2} \cos t d \cos^2 t = -d \cos t.$$

So we get $d \cos t = -\sin t dt$.

Theorem 1. *$d \sin t = \cos t dt$ and $d \cos t = -\sin t dt$*

We have proved this equality only for $[0, \pi/2]$. But due to the well-known symmetries it suffices.

Application of trigonometric integrals.

Lemma 2. *For any convergent infinite product of factors ≥ 1 one has*

$$\lim \prod_{k=1}^{n} p_k = \prod_{k=1}^{\infty} p_k \tag{2.5.5}$$

Proof. Let ε be a positive number. Then $\prod\limits_{k=1}^{\infty} p_k > \prod\limits_{k=1}^{\infty} p_k - \varepsilon$, and by All-for-One there is n such that $\prod\limits_{k=1}^{n} p_k > \prod\limits_{k=1}^{\infty} p_k - \varepsilon$. Then for any $m > n$ one has the inequalities $\prod\limits_{k=1}^{\infty} p_k \geq \prod\limits_{k=1}^{m} p_k > \prod\limits_{k=1}^{\infty} p_k - \varepsilon$. Therefore $|\prod\limits_{k=1}^{m} p_k - \prod\limits_{k=1}^{\infty} p_k| < \varepsilon$. $\square$

Wallis product. Set $I_n = \int\limits_0^\pi \sin^n x\,dx$. Then $I_0 = \int\limits_0^\pi 1\,dx = \pi$ and $I_1 = \int\limits_0^\pi \sin x\,dx = -\cos\pi + \cos 0 = 2$. For $n \geq 2$, let us replace the integrand $\sin^n x$ by $\sin^{n-2} x(1-\cos^2 x)$ and obtain

$$\begin{aligned}
I_n &= \int_0^\pi \sin^{n-2} x(1-\cos^2 x)\,dx = \int_0^\pi \sin^{n-2} x - \int_0^\pi \sin^{n-2} x \cos x\; d\sin x = \\
&= I_{n-2} - \frac{1}{n-1}\int_0^\pi \cos x\, d\sin^{n-1}(x) = I_{n-2} - \int_0^\pi d(\cos x \sin^{n-1} x) + \\
&\qquad + \int_0^\pi \sin^{n-1} x\, d\cos x = I_{n-2} - \frac{1}{n-1} I_n
\end{aligned}$$

We get the recurrence relation $I_n = \frac{n-1}{n} I_{n-2}$, which gives the formula

$$I_{2n} = \pi\frac{(2n-1)!!}{2n!!} \qquad I_{2n-1} = 2\frac{(2n-2)!!}{(2n-1)!!} \tag{2.5.6}$$

where $n!!$ denotes the product $n(n-2)(n-4)\dots(n \mod 2+1)$. Since $\sin^n x \leq \sin^{n-1} x$ for all $x \in [0,\pi]$, the sequence $\{I_n\}$ decreases. Since $I_n \leq I_{n-1} \leq I_{n-2}$, one gets $\frac{n-1}{n} = \frac{I_n}{I_{n-2}} \leq \frac{I_{n-1}}{I_{n-2}} \leq 1$. Hence $\frac{I_{n-1}}{I_{n-2}}$ differs from 1 less than $\frac{1}{n}$. Consequently, $\lim \frac{I_{n-1}}{I_{n-2}} = 1$. In particular, $\lim \frac{I_{2n+1}}{I_{2n}} = 1$. Substituting in the last formula the expressions of I_n from (2.5.6) one gets

$$\lim \frac{\pi}{2}\frac{(2n+1)!!(2n-1)!!}{2n!!2n!!} = 1 \tag{2.5.7}$$

Therefore this is the famous Wallis Product

$$\frac{\pi}{2} = \lim \frac{2n!!2n!!}{(2n-1)!!(2n+1)!!} = \prod_{n=1}^{\infty} \frac{4n^2}{4n^2-1} \tag{2.5.8}$$

Stirling constant. In the Lecture 2.4 we have proved that

$$\ln n! = n\ln n - n + \frac{1}{2}\ln n + \sigma + o_n, \tag{2.5.9}$$

where o_n is infinitesimally small and σ is a constant. Now we are ready to determine this constant. Consider the difference $\ln 2n! - 2\ln n!$ by (2.5.9) it expands into $(2n\ln 2n - 2n + \frac{1}{2}\ln 2n + \sigma + o_{2n}) - 2(n\ln n - n + \frac{1}{2}\ln n + \sigma + o_n) = 2n\ln 2 + \frac{1}{2}\ln 2n - \ln n - \sigma + o'_n$, where $o'_n = o_{2n} - 2o_n$ is infinitesimally small. Then σ can be presented as $\sigma = 2\ln n! - \ln 2n! + 2n\ln 2 + \frac{1}{2}\ln n + \frac{1}{2}\ln 2 - \ln n + o'_n$. Multiplying by 2 one gets $2\sigma = 4\ln n! - 2\ln 2n! + 2\ln 2^{2n} - \ln n + \ln 2 + 2o'_n$ Hence $2\sigma = (\lim 4\ln n! - 2\ln 2n! + 2\ln 2^{2n} - \ln n + \ln 2)$ Switching to the products and keeping in mind identities $n! = n!!(n-1)!!$ and $n!2^n = 2n!!$ one gets
(2.5.10)

$$\sigma^2 = \lim \frac{n!^4 2^{4n+1}}{(2n!)^2 n} = \lim \frac{2\cdot(2n!!)^4}{(2n!!)^2(2n-1)!!^2 n} \lim \frac{2\cdot(2n!!)^2(2n+1)}{(2n-1)!!(2n+1)!!n} = 2\pi$$

Problems.

1. Evaluate $\int\sqrt{1-x^2}\,dx$
2. Evaluate $\int\frac{1}{\sqrt{1-x^2}}\,dx$
3. Evaluate $\int\sqrt{5-x^2}\,dx$
4. Evaluate $\int\cos^2 x\,dx$
5. Evaluate $\int\tan x\,dx$
6. Evaluate $\int\sin^4 x\,dx$
7. Evaluate $\int\sin x^2\,dx$
8. Evaluate $\int\tan x\,dx$
9. Evaluate $\int x^2\sin x\,dx$
10. Evaluate $\frac{d\arcsin x}{dx}$
11. Evaluate $\int\arcsin x\,dx$
12. Evaluate $\int e^x\cos x\,dx$

2.6 Virtually monotone functions

Monotonization of integrand. Let us say that a pair of functions f_1, f_2 *monotonize* a function f, if f_1 is a non-negative and non-decreasing, f_2 is a non-positive and non-increasing and $f = f_1 + f_2$.

Lemma 1. *Let $f = f_1 + f_2$ and $f = f_1' + f_2'$ be two monotonizations of f. Then for any monotone h one has $f_1\, dh + f_2\, dh = f_1' dh + f_2' dh$.*

Proof. Our equality is equivalent to $f_1\, dh - f_2'\, dh = f_1'\, dh - f_2\, dh$. By the sign rule this turns into $f_1\, dh + (-f_2')\, dh = f_1'\, dh + (-f_2)\, dh$. Now all integrands are nonnegative and for non-decreasing h we can apply the Addition Theorem and transform the inequality into $(f_1 - f_2')\, dh = (f_1' - f_2)\, dh$. This is true because $(f_1 - f_2') = (f_1' - f_2)$.

The case of non-increasing differand is reduced to the case of non-decreasing one by the transformation $f_1\, d(-h) + f_2\, d(-h) = f_1' d(-h) + f_2' d(-h)$, which is based on the Sign Rule.

□

A function which has a monotonization is called *virtually monotone*.

We define the integral $\int\limits_a^b f\, dg$ for any virtually monotone integrand f and any continuous monotone differand g via a monotonization $f = f_1 + f_2$ by

$$\int\limits_a^b f\, dg = \int\limits_a^b f_1\, dg + \int\limits_a^b f_2\, dg. \tag{2.6.1}$$

Lemma **1** demonstrates that this definition does not depend on the choice of a monotonization.

Lemma 2. *Let f and g be virtually monotone functions; then $f + g$ is virtually monotone and $fdh + gdh = (f + g)dh$ for any continuous monotone h.*

Proof. Let h be nondecreasing. Consider monotonizations $f = f_1 + f_2$ and $g = g_1 + g_2$. Then $fdh + gdh = f_1dh + f_2dh + g_1dh + g_2dh$ by definition via monotonization of the integrand. By virtue of the Addition Theorem **3** this turns into $(f_1 + g_1)dh + (f_2 + g_2)dh$. But the pair of brackets monotonize $f + g$. Hence $f + g$ is proved to be virtually monotone and the later expression is $(f + g)dh$ by definition, via monotonization of integrand. The case of non-increasing h is reduced to the previous case via $-fd(-h) - gd(-h) = -(f + g)d(-h)$. □

Lemma on locally constant functions. Let us say that a function $f(x)$ is *locally constant* at a point x if $f(y) = f(x)$ for all y sufficiently close to x, i.e for all y from an interval $(x - \varepsilon, x + \varepsilon)$.

Lemma 3. *A function f which is locally constant at any point of an interval is constant.*

Proof. Suppose $f(x)$ is not constant on $[a, b]$. We will construct by induction a sequence of intervals $I_k = [a_k, b_k]$, such that $I_0 = [a, b]$, $I_{k+1} \subset I_k$, $|b_k - a_k| \geq 2|b_{k+1} - a_{k+1}|$ and the function f is not constant on each I_k. First step. Let $c = (a + b)/2$, as f is not constant $f(x) \neq f(c)$ for some x. Then choose $[x, c]$ or

$[c, x]$ as for $[a_1, b_1]$. On this interval f is not constant. The same are all further steps. The intersection of the sequence is a point such that any its neighborhood contains some interval of the sequence. Hence f is not locally constant at this point. ☐

Lemma 4. *If $f(x)$ is a continuous monotone function and $a < f(x) < b$ then $a < f(y) < b$ for all y sufficiently close to x.*

Proof. If f takes values greater than b, than it takes value b and if $f(x)$ takes values less than a then it takes value a due to continuity. Then $[f^{-1}(a), f^{-1}(b)]$ is the interval where inequalities hold. ☐

Lemma 5. *Let g_1, g_2 be continuous* comonotone *functions. Then $g_1 + g_2$ is continuous and monotone, and for any virtually monotone f one has*

$$fdg_1 + fdg_2 = fd(g_1 + g_2) \tag{2.6.2}$$

Proof. Suppose $g_1(x) + g_2(x) < p$, let $\varepsilon = p - g_1(x) - g_2(x)$. Then $g_1(y) < g_1(y) + \varepsilon/2$ and $g_2(y) < g_2(y) + \varepsilon/2$ for all y sufficiently close to x. Hence $g_(y)+g_2(y) < p$ for all y sufficiently close to x. The same is true for the opposite inequality. Hence $\operatorname{sgn}(g_1(x) + g_2(x) - p)$ is locally constant at all points where it is not 0. But it is not constant if p is an intermediate value, hence it is not locally constant, hence it takes value 0. At this point $g_1(x) + g_2(x) = p$ and the continuity of $g_1 + g_2$ is proved.

Consider a monotonization $f = f_1 + f_2$. Let g_i be nondecreasing. By definition via monotonization of the integrand, the left-hand side of (2.6.2) turns into $(f_1dg_1 + f_2dg_1) + (f_1dg_2 + f_2dg_2) = (f_1dg_1 + f_1dg_2) + (f_2dg_1 + f_2dg_2)$. By the Addition Theorem **3** $f_1dg_1 + f_1dg_2 = f_1d(g_1 + g_2)$. And the equality $f_2dg_1 + f_2dg_2 = f_2d(g_1 + g_2)$ follows from $(-f_2)dg_1 + (-f_2)dg_2 = (-f_2)d(g_1 + g_2)$ by the Sign Rule. Hence the left-hand side is equal to $f_1d(g_1 + g_2) + f_2d(g_1 + g_2)$, which coincides with the right-hand side of (2.6.2) by definition via monotonization of integrand. The case of non-increasing differands is taken care of via transformation of (2.6.2) by the Sign Rule into $fd(-g_1) + fd(-g_2) = fd(-g_1 - g_2)$. ☐

Lemma 6. *Let $g_1+g_2 = g_3+g_4$ where all $(-1)^k g_k$ are non-increasing continuous functions. Then $fdg_1 + fdg_2 = fdg_3 + fdg_4$ for any virtually monotone f*

Proof. Our equality is equivalent to $fdg_1 - fdg_4 = fdg_3 - fdg_2$. By the Sign Rule it turns into $fdg_1 + fd(-g_4) = fdg_3 + fd(-g_2)$. Now all differands are nondecreasing and by Lemma **5** it transforms into $fd(g_1 - g_4) = fd(g_3 - g_2)$. This is true because $g_1 - g_4 = g_3 - g_2$. ☐

Monotonization of differand. A monotonization by continuous functions is called continuous. Virtually monotone function which has a continuous monotonization is called *continuous.* The integral for any virtually monotone integrand f against a virtually monotone continuous differand g is defined via a continuous virtualization $g = g_1 + g_2$ of the differand

$$\int_a^b f\,dg = \int_a^b f\,dg_1 + \int_a^b f\,dg_2 \tag{2.6.3}$$

The integral is well-defined because of Lemma **6**.

Theorem 7 (Addition Theorem). *For any virtually monotone functions f, f' and any virtually monotone continuous g, g' hold $fdg + f'dg = (f + f')dg$ and $fdg + fdg' = fd(g + g')$*

Proof. To prove $fdg + f'dg = (f + f')dg$, consider a continuous monotonization $g = g_1 + g_2$. Then by definition of integral for virtually monotone differand this equality turns into $(fdg_1 + fdg_2) + (f'dg_1 + f'dg_2) = (f + f')dg_1 + (f + f')dg_2$. After rearranging it turns into $(fdg_1 + f'dg_1) + (fdg_2 + f'dg_2) = (f + f')dg_1 + (f + f')dg_2$ But this is true due to Lemma **2**.

To prove $fdg + fdg' = fd(g + g')$, consider monotonizations $g = g_1 + g_2$, $g' = g'_1 + g'_2$. Then $(g_1 + g'_1) + (g_2 + g'_2)$ is monotonization for $g + g'$. And by the definition of integral for virtually monotone differand our equality turns into $fdg_1 + fdg_2 + fdg'_1 + fdg'_2$ □

Change of variable.

Lemma 8. *If f is virtually monotone and g is monotone, then $f(g(x))$ is virtually monotone.*

Proof. Let $f_1 + f_2$ be a monotonization of f. If h is non-decreasing then $f_1(h(x)) + f_2(h(x))$ gives a monotonization of $f(g(x))$. If h is decreasing the monotonization is given by $(f_2(h(x))+c)+(f_1(h(x))-c)$ where c is a sufficiently large constant, to provide positivity of the first brackets and negativity of the second one. □

The following natural convection is applied to define the integral with reversed limits.

$$\int_a^b f(x)\,dg(x) = -\int_b^a f(x)\,dg(x)$$

Theorem 9 (on change of variable). *If $h\colon [a,b] \to [h(a), h(b)]$ is monotone, $f(x)$ is virtually monotone, and $g(x)$ is virtually monotone continuous, then* $\int_a^b f(h(t))\,dg(h(t)) = \int_{h(a)}^{h(b)} f(x)\,dg(x)$

Proof. Let $f = f_1 + f_2$ and $g = g_1 + g_2$ be a monotonization and a continuous monotonization of f and g respectively. The $\int_a^b f(h(t))\,dg(h(t))$ splits into sum of four integrals. $\int_a^b f_i(h(t))\,dg_j(h(t))$ where f_i are of constant sign and g_j are monotone continuous. This integrals coincide with corresponding integrals $\int_{h(a)}^{h(b)} f_i(x)\,dg_i(x)$. Indeed its absolute values are the area of the same curvilinear trapezium. And their signs determined by the Sign Rule are the same. □

Integration by parts. We have established the Integration by parts formula for non-negative and non-decreasing differential forms. Now we extend it to the case of continuous monotone forms. In the first case f and g are non-decreasing. In this case choose a positive constant c sufficiently large to provide positivity of $f+c$ and $g+c$ on the interval of integration. Then $d(f+c)(g+c) =$

$(f+c)d(g+c)+(g+c)d(f+c)$. On the other hand $d(f+c)(g+c) = dfg+cdf+cdg$ and $(f+c)d(g+c)+(g+c)d(f+c) = fdg+cdg+cdf$. Compare these results to get $dfg = fdg + gdf$. Now if f is increasing and g decrease the $-g$ increase and we get $-dfg = df(-g) = fd(-g) + (-g)df = -fdg - gdf$, which lead to $dfg = fdg+gdf$. The other cases: f decreasing, g increasing and both decreasing are proved by the same arguments. The extension of the Integration by Parts to piece-wise monotone forms immediately follows by the Partition Rule.

Variation. Define *variation of a sequence* of numbers $\{x_k\}_{k=1}^n$ as the sum $\sum\limits_{k=1}^{\infty} |x_{k+1} - x_k|$. Define variation of a function f along a sequence $\{x_k\}_{k=0}^n$ as the variation of sequence $\{f(x_k)\}_{k=0}^n$. Define a *chain* on an interval $[a,b]$ as a nondecreasing sequence $\{x_k\}_{k=0}^n$ such that $x_0 = a$ and $x_n = b$. Define a *partial variation* of f on an interval $[a,b]$ as its variation along a chain on the interval.

The least number surpassing all partial variations function f over $[a,b]$ is called *the (ultimate) variation of a function* $f(x)$ on an interval $[a,b]$ and denote by $\operatorname{var}_f[a,b]$.

Lemma 10. *For any function f one has inequality* $\operatorname{var}_f[a,b] \geq |f(b) - f(a)|$. *If f is a monotone function on $[a,b]$, then* $\operatorname{var}_f[a,b] = |f(b) - f(a)|$

Proof. Inequality $\operatorname{var}_f[a,b] \geq |f(b) - f(a)|$ follows immediately from the definition because $\{a,b\}$ is a chain. For monotone f this all partial variations are telescopic sums equal to $|f(b) - f(a)|$ □

Theorem 11 (additivity of variation). $\operatorname{var}_f[a,b] + \operatorname{var}_f[b,c] = \operatorname{var}_f[a,c]$

Proof. Consider a chain $\{x_k\}_{k=0}^n$ of $[a,c]$, which contains b. In this case the variation of f along $\{x_k\}_{k=0}^n$ splits into sums of partial variations of f along $[a,b]$ and along $[b,c]$. As a partial variations does not exceed an ultimate. We get that in this case the variation of f along $\{x_k\}_{k=0}^n$ does not exceed $\operatorname{var}_f[a,b]+\operatorname{var}_f[b,c]$.

If $\{x_k\}_{k=0}^n$ does not contain b, let us add b to the chain. Then in the sum expressing the partial variation of f, summand $|f(x_{i+1} - f(x_i)|$ changes by sum $|f(b) - f(x_i)| + |f(x_{i+1} - f(b)|$ which is greater or equal. Hence the variation does not decrease after such modification. But the variation along the modified chain does not exceed $\operatorname{var}_f[a,b] + \operatorname{var}_f[b,c]$ as was proved above. As all partial variations of f over $[a,c]$ do not exceed $\operatorname{var}_f[a,b] + \operatorname{var}_f[b,c]$, the same is true for the ultimate variation.

To prove the opposite inequality we consider a *relaxed* inequality $\operatorname{var}_f[a,b] + \operatorname{var}_f[b,c] \leq \operatorname{var}_f[a,c] + \varepsilon$ where ε is an positive number. Choose chains $\{x_k\}_{k=0}^n$ on $[a,b]$ and $\{y_k\}_{k=0}^m$ on $[b,c]$ such that corresponding partial variation of f be $\geq \operatorname{var}_f[a,b]+\varepsilon/2$ and $\geq \operatorname{var}_f[b,c]+\varepsilon/2$ respectively. As the union of these chains is a chain on $[a,c]$ the sum of these partial variations is a partial variation of f on $[a,c]$. Consequently this sum is less or equal to $\operatorname{var}_f[a,c]$. On the other hand it is greater or equal to $\operatorname{var}_f[a,b] + \varepsilon/2 + \operatorname{var}_f[b,c] + \varepsilon/2$. Comparing this results gives just the relaxed inequality. As relaxed inequality is proved for all $\varepsilon > 0$ it also holds for $\varepsilon = 0$. □

Lemma 12. *For any functions f, g one has inequality* $\operatorname{var}_{f+g}[a,b] \leq \operatorname{var}_f[a,b] + \operatorname{var}_g[a,b]$

Proof. Since $|f(x_{k+1})+g(x_{k+1})-f(x_k)-g(x_k)| \leq |f(x_{k+1})-f(x_k)|+|g(x_{k+1})-g(x_k)|$, the variation of $f+g$ along any sequence does not exceed the sum of the variations of f and g along the sequence. Hence all partial variations of $f+g$ do not exceed $\operatorname{var}_f[a,b] + \operatorname{var}_g[a,b]$, and so the same is true for the ultimate variation. $\square$

Lemma 13. *For any function of finite variation on $[a,b]$, functions $\operatorname{var}_f[a,x]$ and $\operatorname{var}_f[a,x] - f(x)$ are both nondecreasing functions of x.*

Proof. The nondecreasing of $\operatorname{var}_f[a,x]$ follows from nonnegativity and additivity of variation. If $x > y$ then the inequality $\operatorname{var}_f[a,x] - f(x) \geq \operatorname{var}_f[a,y] - f(y)$ is equivalent to $\operatorname{var}_f[a,x] - \operatorname{var}_f[a,y] \geq f(x) - f(y)$. This is true because $\operatorname{var}_f[a,x] - \operatorname{var}_f[a,y] = \operatorname{var}_f[x,y] \geq |f(x) - f(y)|$. $\square$

Lemma 14. $\operatorname{var}_{f^2}[a,b] \leq 2(|f(a)| + \operatorname{var}_f[a,b]) \operatorname{var}_f[a,b]$

Proof. For all $x, y \in [a,b]$ one has $|f(x) + f(y)| = |2f(a) + f(x) - f(a) + f(y) - f(a)| \leq 2|f(a)| + \operatorname{var}_f[a,x] + \operatorname{var}_f[a,y] \leq 2|f(a)| + 2\operatorname{var}_f[a,b]$. Hence $\sum\limits_{k=0}^{n-1} |f^2(x_{k+1}) - f^2(x_k)| = \sum\limits_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)||f(x_{k+1} + f(x_k)| \leq 2(|f(a)| + \operatorname{var}_f[a,b]) \sum\limits_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \leq 2(|f(a)| + \operatorname{var}_f[a,b]) \operatorname{var}_f[a,b]$ $\square$

Lemma 15. *If $\operatorname{var}_f[a,b] < \infty$ and $\operatorname{var}_g[a,b] < \infty$, then $\operatorname{var}_{fg}[a,b] < \infty$*

Proof. $4fg = (f+g)^2 - (f-g)^2$ $\square$

Theorem 16. *The function f is virtually monotone on $[a,b]$ if and only if it has a finite variation.*

Proof. Since monotone functions have finite variation on finite intervals, and the variation of a sum does not exceed the sum of variations one gets that all virtually monotone functions have finite variation. On the other hand if f has a finite variation then $f = (\operatorname{var}_f[a,x] + c) + (f(x) - \operatorname{var}_f[a,x] - c)$, the functions in the brackets are monotone due to Lemma **13**, and choosing constant c sufficiently large one obtains that the second bracket is negative. $\square$

Problems.

1. Evaluate $\int\limits_1^i z^2\,dz$
2. Prove that $1/f(x)$ has finite variation if it is bounded.
3. Prove $\int\limits_a^b f(x)\,dg(x) \leq \max_{[a,b]} f \operatorname{var}_g[a,b]$

Chapter 3

Derivatives

3.1 Newton-Leibniz Formula

Motivation. Consider the following problem: for a given function F find a function f such that $dF(x) = f(x)\,dx$, over $[a,b]$, that is, $\int\limits_c^d f(t)\,dt = F(d)-F(c)$ for any subinterval $[c,d]$ of $[a,b]$.

Suppose that such an f exists. Since the value of f at a single point does not affects the integrals, we cannot say anything about the value of f at any given point. But if f is continuous at a point x_0, its value is uniquely defined by F.

To be precise, the difference quotient $\frac{F(x)-F(x_0)}{x-x_0}$ tends to $f(x_0)$ as x tends to x_0. Indeed, $F(x) = F(x_0) + \int\limits_{x_0}^x f(t)\,dt$. Furthermore, $\int\limits_{x_0}^x f(t)\,dt = f(x_0)(x-x_0) + \int\limits_{x_0}^x (f(t)-f(x_0))\,dt$. Also, $|\int\limits_{x_0}^x (f(t)-f(x_0)\,dt| \leq \operatorname{var}_f[x_0,x]|x-x_0|$. Consequently

$$\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| \leq \operatorname{var}_f[x,x_0] \tag{3.1.1}$$

However, $\operatorname{var}_f[x,x_0]$ can be made arbitrarily small by choosing x sufficiently close to x_0, since $\operatorname{var}_f x_0 = 0$.

Infinitesimally small functions. A set is called a *neighborhood* of a point x if it contains all points *sufficiently close* to x, that is all points y such that $|y-x|$ is less then a positive number ε.

We will say that a function f is *locally bounded* (above) by a constant C at a point x, if $f(x) \leq C$ for all y sufficiently close to x.

A function $o(x)$ is called *infinitesimally small* at x_0, if $|o(x)|$ is locally bounded at x_0 by any $\varepsilon > 0$.

Lemma 1. *If the functions o and ω are infinitesimally small at x_0 then $o \pm \omega$ are infinitesimally small at x_0.*

Proof. Given a $\varepsilon > 0$. Let O_1 be a neighborhood of x_0 where $|o(x)| < \varepsilon/2$, and O_2 be a neighborhood of x_0 where $|\omega(x)| < \varepsilon/2$, then $O_1 \cap O_2$ is a neighborhood where both inequalities hold. Hence for all $x \in O_1 \cap O_2$ one has $|o(x) \pm \omega(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Lemma 2. *If $o(x)$ is infinitesimally small at x_0 and $f(x)$ is locally bounded at x_0, then $f(x)o(x)$ is infinitesimally small.*

Proof. The neighborhood where $|f(x)o(x)|$ is bounded by a given $\varepsilon > 0$ can be constructed as the intersection of a neighborhood U, where $|f(x)|$ is bounded by a constant C, and a neighborhood V, where $|o(x)|$ is bounded by ε/C. □

Definition. *One says that a function $f(x)$ tends to A as x tends to x_0 and writes $\lim\limits_{x\to x_0} f(x) = A$, if $f(x) = A + o(x)$ on the complement of x_0, where $o(x)$ is infinitesimally small at x_0.*

Corrolary 3. *If limits $\lim\limits_{x\to x_0} f(x)$ and $\lim\limits_{x\to x_0} g(x)$ exist, then $\lim\limits_{x\to x_0} (f(x)+g(x))$ also exists and $\lim\limits_{x\to x_0} (f(x)+g(x)) = \lim\limits_{x\to x_0} f(x) + \lim\limits_{x\to x_0} g(x)$.*

Proof. This follows immediately from Lemma **1**. □

Lemma 4. *If the limits* $\lim_{x\to x_0} f(x)$ *and* $\lim_{x\to x_0} g(x)$ *exist, then also* $\lim_{x\to x_0} f(x)g(x)$ *and* $\lim_{x\to x_0} f(x)g(x) = \lim_{x\to x_0} f(x) \lim_{x\to x_0} g(x)$ *exist.*

Proof. If $f(x) = A + o(x)$ and $g(x) = B + \omega(x)$, then $f(x)g(x) = AB + A\omega(x) + Bo(x) + \omega(x)o(x)$, where $A\omega(x)$, $Bo(x)$ and $\omega(x)o(x)$ all are infinitesimally small at x_0 by Lemma **2**, and their sum is infinitesimally small by Lemma **1**. □

Definition. *A function f is called* continuous *at x_0, if* $\lim_{x\to x_0} f(x) = f(x_0)$.

A function is said to be continuous (without mentioning a point), if it is continuous at all points under consideration.

The following lemma gives a lot of examples of continuous functions.

Lemma 5. *If f is a monotone function on $[a,b]$ such that $f[a,b] = [f(a), f(b)]$ then f is continuous.*

Proof. Given a positive ε. Suppose f is nondecreasing. For a given point x denote by $x^\varepsilon = f^{-1}(f(x)+\varepsilon)$ and $x_\varepsilon = f^{-1}(f(x)-\varepsilon)$. Then $[x_\varepsilon, x^\varepsilon]$ contains a neighborhood of x, and for any $y \in [x_\varepsilon, x^\varepsilon]$ one has $f(x)+\varepsilon = f(x_\varepsilon) \leq f(y) \leq f(x^\varepsilon) = f(x)+\varepsilon$. Hence the inequality $|f(y) - f(x)| < \varepsilon$ holds locally at x for any ε. □

The following theorem immediately follows from the Corollary **3** and Lemma **4**.

Theorem 6. *If the functions f and g are continuous at x_0, then $f+g$ and fg are continuous at x_0.*

The following property of continuous function is very important.

Theorem 7. *If f is continuous at x_0 and g is continuous at $f(x_0)$, then $g(f(x))$ is continuous at x_0.*

Proof. Given $\varepsilon > 0$, we have to find a neighborhood U of x_0, such that $|g(f(x)) - g(f(x_0))| < \varepsilon$ for $x \in U$. As $\lim_{y\to f(x_0)} g(y) = g(f(x_0))$, there exists a neighborhood V of $f(x_0)$ such that $|g(y) - g(y_0)| < \varepsilon$ for $y \in V$. Thus it is sufficient to find a U such that $f(U) \subset V$. And we can do this. Indeed, by definition of neighborhood there is $\delta > 0$, such that V contains $V_\delta = \{y \mid |y - f(x_0)| < \delta\}$. Since $\lim_{x\to x_0} f(x) = f(x_0)$, there is a neighborhood U of x_0 such that $|f(x) - f(x_0)| < \delta$ for all $x \in U$. Then $f(U) \subset V_\delta \subset V$. □

Definition. *A function f is called* differentiable *at a point x_0 if the* difference quotient $\frac{f(x)-(f_0)}{x-x_0}$ *has a limit as x tends to x_0. This limit is called the* derivative *of the function F at the point x_0, and denoted* $f'(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$

Immediately from the definition one evaluates the derivative of linear function.

$$(ax+b)' = a \tag{3.1.2}$$

The following lemma is a direct consequence of Lemma **3**.

Lemma 8. *If f and g are differentiable at x_0, then $f+g$ is differentiable at x_0 and $(f+g)'(x_0) = f'(x_0) + g'(x_0)$*

Linearization. Let f be differentiable at x_0. Denote by $o(x)$ the difference $\frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)$. Then

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x)(x - x_0), \tag{3.1.3}$$

where $o(x)$ is infinitesimally small at x_0. We will call such a representation a *linearization* of $f(x)$.

Lemma 9. *If f is differentiable at x_0, then it is continuous at x_0.*

Proof. All summands but $f(x_0)$ on the right-hand side of (3.1.3) are infinitesimally small at x_0; hence $\lim\limits_{x\to x_0} f(x) = f(x_0)$. □

Lemma 10 (on uniqueness of linearization). *If $f(x) = a + b(x - x_0) + o(x)(x - x_0)$, where $\lim\limits_{x\to x_0} o(x) = 0$, then f is differentiable and at x_0, $a = f(x_0)$ and $b = f'(x_0)$.*

Proof. The difference $f(x) - f(x_0)$ is infinitesimally small at x_0, because f is continuous at x_0, and the difference $f(x) - a = b(x - x_0) + o(x)(x - x_0)$ is infinitesimally small by the definition of linearization. Hence $f(x_0) - a$ is infinitesimally small. But it is constant, hence $f(x_0) - a = 0$.

Thus we established $a = f(x_0)$. The difference $\frac{f(x)-a}{x-x_0} - b = o(x)$ is infinitesimally small as well as $\frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)$. But $\frac{f(x)-f(x_0)}{x-x_0} = \frac{f(x)-a}{x-x_0}$. Therefore $b - f'(x_0)$ is infinitesimally small. That is $b = f'(x_0)$. □

Lemma 11. *If f and g are differentiable at x_0, then fg is differentiable at x_0 and $(fg)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0)$*

Proof. Consider lineariations $f(x_0) + f'(x_0)(x - x_0) + o(x)(x - x_0)$ and $g(x_0) + g'(x_0)(x - x_0) + \omega(x)(x - x_0)$. Their product is $f(x_0)g(x_0) + (f'(x_0)g(x_0) + f(x_0)g'(x_0))(x - x_0) + (f(x)\omega(x) + f(x_0)o(x))(x - x_0)$. This is the linearization of $f(x)g(x)$ at x_0, because $f\omega$ and go are infinitesimally small at x_0. □

Theorem 12. *If f is differentiable at x_0, and g is differentiable at $f(x_0)$ then $g(f(x))$ is differentiable at x_0 and $(g(f(x_0)))' = g'(f(x_0))f'(x_0)$*

Proof. Denote $f(x_0)$ by y_0 and substitute into the linearization $g(y) = g(y_0) + g'(y_0)(y - y_0) + o(y)(y - y_0)$ another linearization $y = f(x_0) + f'(x_0)(x - x_0) + \omega(x)(x - x_0)$. Since $y - y_0 = f'(x_0)(x - x_0) + \omega(x)(x - x_0)$, we get $g(y) = g(y_0) + g'(y_0)f'(x_0)(x - x_0) + g'(y_0)(x - x_0)\omega(x) + o(f(x))(x - x_0)$. Due to the Lemma on linearization, it is sufficient to prove that $g'(y_0)\omega(x) + o(f(x))$ is infinitesimally small at x_0. The first summand is obviously infinitesimally small. To prove that the second one also is infinitesimally small, we remark that $o(f(x_0)) = 0$ and $o(y)$ is continuous at $f(x_0)$ and that $f(x)$ is continuous at x_0 due to Lemma **9**. Hence by Theorem 3.1 the composition is continuous at x_0 and infinitesimally small. □

Theorem 13. *Let f be a virtually monotone function on $[a, b]$. Then $F(x) = \int_a^x f(t)\,dt$ is virtually monotone and continuous on $[a, b]$. It is differentiable at any point x_0 where f is continuous, and $F'(x_0) = f(x_0)$.*

Proof. If f has a constant sign, then F is monotone. So, if $f = f_1 + f_2$ is a monotonization of f, then $\int\limits_a^x f_1(x)\,dx + \int\limits_a^x f_1(x)\,dx$ is a monotonization of $F(x)$. This proves that $F(x)$ is virtually monotone.

To prove continuity of $F(x)$ at x_0, fix a constant C which bounds f in some neighborhood U of x_0. Then for $x \in U$ one proves that $|F(x) - F(x_0)|$ is infinitesimally small via inequalities $|F(x)-F(x_0)| = |\int\limits_{x_0}^x f(x)\,dx| \leq |\int\limits_{x_0}^x C\,dx| = C|x-x_0|$.

Now suppose f is continuous at x_0. Then $o(x) = f(x_0) - f(x)$ is infinitesimally small at x_0. Therefore $\lim\limits_{x\to x_0} \frac{1}{x-x_0} \int\limits_{x_0}^x o(x)\,dx = 0$. Indeed for any $\varepsilon > 0$ the inequality $|o(x)| \leq \varepsilon$ holds over $[x_\varepsilon, x_0]$ for some x_ε. Hence $|\int\limits_{x_0}^x o(x)\,dx| \leq |\int\limits_{x_0}^x \varepsilon\,dx| = \varepsilon|x-x_0|$ for any $x \in [x_0, x_\varepsilon]$.

Then $F(x) = F(x_0)+f(x_0)(x-x_0)+(\frac{1}{x-x_0}\int\limits_{x_0}^x o(t)\,dt)(x-x_0)$ is a linearization of $F(x)$ at x_0. □

Corrolary 14. *The functions* ln, sin, cos *are differentiable and* $\ln'(x) = \frac{1}{x}$, $\sin' = \cos$, $\cos' = -\sin$.

Proof. Since $d\sin x = \cos x\,dx$, $d\cos x = -\sin x\,dx$, due to Theorem **13** both $\sin x$ and $\cos x$ are continuous, and, as they are continuous, the result follows from Theorem **13**. And $\ln' x = \frac{1}{x}$, by Theorem **13**, follows from continuity of $\frac{1}{x}$. The continuity follows from Lemma **5**. □

Since $\sin'(0) = \cos 0 = 1$ and $\sin 0 = 0$, the linearization of $\sin x$ at 0 is $x + xo(x)$. This implies the following very important equality

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 \tag{3.1.4}$$

Lemma 15. *If $f'(x) > 0$ for all $x \in [a,b]$, then $f(b) > f(a)|$*

Proof. Suppose $f(a) \geq f(b)$. We construct a sequence of intervals $[a,b] \supset [a_1,b_1] \supset [a_2,b_2] \supset \dots$. such that their lengths tend to 0 and $f(a_k) \geq f(b_k)$. All steps of construction are the same. The general step is: let m be the middle point of $[a_k,b_k]$. If $f(m) \leq f(a_k)$ we set $[a_{k+1},b_{k+1}] = [a_k,m]$, otherwise $f(m) > f(a_k) \geq f(b_k)$ and we set $[a_{k+1},b_{k+1}] = [m,b_k]$.

Now consider a point x belonging to all $[a_k,b_k]$. Let $f(y) = f(x) + (f'(x) + o(x))(y-x)$ be the linearization of f at x. Let U be neighborhood where $|o(x)| < f'(x)$. Then $\operatorname{sgn}(f(y)-f(x)) = \operatorname{sgn}(y-x)$ for all $y \in U$. However for some n we get $[a_n,b_n] \subset U$. If $a_n \leq x < b_n$ we get $f(a_n) \leq f(x) < f(b_n)$ else $a_n < x$ and $f(a_n) < f(x) \leq f(b_n)$. In the both cases we get $f(a_n) < f(b_n)$. This is contradiction with our construction of the sequence of intervals. □

Theorem 16. *If $f'(x) = 0$ for all $x \in [a,b]$, then $f(x)$ is constant.*

Proof. Set $k = \frac{f(b)-f(a)}{b-a}$. If $k < 0$ then $g(x) = f(x) - kx/2$ has derivative $g'(x) = f'(x) - k/2 > 0$ for all x. Hence by Lemma **15** $g(b) > g(a)$ further $f(b) - f(a) > k(b-a)/2$. This contradict to definition of k. If $k > 0$ then the same contradiction one gets considering $g(x) = -f(x) + kx/2$. □

Theorem 17 (Newton-Leibniz). *If $f'(x)$ is a continuous virtually monotone function on an interval $[a, b]$, then $\int\limits_a^b f'(x)\,dx = f(b) - f(a)$*

Proof. Due to Theorem **13**, the derivative of the difference $\int\limits_a^x f'(t)\,dt - f(x)$ is zero. Hence the difference is constant by Theorem **16**. Substituting $x = a$ we find the constant which is $f(a)$. Consequently, $\int\limits_a^x f'(t)\,dt - f(x) = f(a)$ for all x. In particular, for $x = b$ we get the Newton-Leibniz formula.

□

Problems.

1. $(1/x)'$, $\sqrt{x}'$, $(\sqrt{\sin x^2})'$
2. $\exp' x$
3. $\operatorname{arctg}' x$, $\tan' x$
4. $|x|'$, $\operatorname{Re} z'$
5. $f'(x) \equiv 1 \iff f(x) = x + const$
6. $\left(\int\limits_x^{x^2} \frac{\sin t}{t}\,dt\right)'_x$
7. $\sqrt{1-x^2}'$
8. $\left(\int\limits_0^1 \frac{\sin kt}{t}\,dt\right)'_k$
9. If f is continuous at a and $\lim\limits_{n\to\infty} x_n = a$ then $\lim\limits_{n\to\infty} f(x_n) = f(a)$.
10. $\left(\int\limits_0^y [x]\,dx\right)'_y$.
11. $arcsin' x$
12. $\int \frac{dx}{2+3x^2}$
13. If $f'(x) < 0$ for all $x < m$ and $f'(x) > 0$ for all $x > m$ then $f'(m) = 0$
14. If $f'(x)$ is bounded on $[a, b]$ then f is virtually monotone

3.2 Exponential Function.

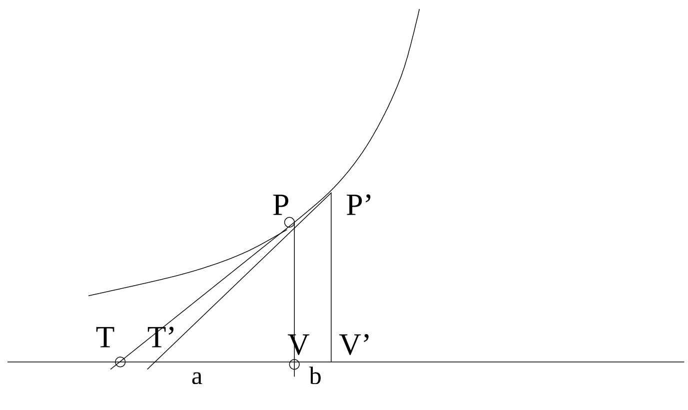
On the contents of the lecture. We solve the principal differential equation $y' = y$. Its solution, the exponential function, is expanded into a power series. We become acquainted with hyperbolic functions. And, finally, we prove irrationality of e.

Debeaune's problem. In 1638 F. Debeaune posed Descartes the following geometrical problem: find a curve $y(x)$ such that for each point P the distances between V and T, the points where the vertical and the tangent line cut the x-axis, are always equal to a given constant a. Despite the efforts of Descartes and Fermat, this problem remained unsolved for nearly 50 years. In 1684 Leibniz solved the problem via infinitesimal analysis of this curve: let x, y be a given point P (see the picture). Then increase x by a small increment of b, so that y increases almost by yb/a. Indeed, in small the curve is considered as the line. Hence the point P' of the curve with vertical projection V', one consider lying on the line TP. Hence the triangle $TP'V'$ is similar to TPV. As $TV = a$, $TV' = b + a$ this similarity gives equality $\frac{a+b}{y+\Delta y} = \frac{a}{y}$ which gives $\Delta y = yb/a$.

Repeating we obtain a sequence of values

$$y,\ y\left(1+\frac{b}{a}\right),\ y\left(1+\frac{b}{a}\right)^2,\ y\left(1+\frac{b}{a}\right)^3,\dots \tag{3.2.1}$$

We see that "in small" $y(x)$ transforms an arithmetic progression into a geometric one. This is inverse to what the logarithm does. And the solution is a function which is inverse to a logarithmic function. Such functions are called *exponential.*



Tangent line and derivative. A tangent line to a smooth convex curve at a point x is defined as a straight line such that the line intersects the curve just at x and the whole curve lies on one side of the line.

We state that the equation of the tangent line to the graph of function f at a point x_0 is just the principal part of linearization of $f(x)$ at x_0. In other words, the equation is $y = f(x_0) + (x - x_0)f'(x_0)$.

First, consider the case of a horizontal tangent line. In this case $f(x_0)$ is either maximal or minimal value of $f(x)$.

Lemma 1. *If a function $f(x)$ is differentiable at an extremal point x_0, then $f'(x_0) = 0$.*

Proof. Consider the linearization $f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x))(x - x_0)$. Denote $x - x_0$ by Δx, and $f(x) - f(x_0)$ by $\Delta f(x)$. If we suppose that $f'(x_0) \neq 0$, then, for sufficiently small Δx, we get $|o(x \pm \Delta x)| < |f'(x)|$, hence $\operatorname{sgn}(f'(x_0) + o(x_0 + \Delta x)) = \operatorname{sgn}(f'(x_0) + o(x_0 - \Delta x))$, and $\operatorname{sgn} \Delta f(x) = \operatorname{sgn} \Delta x$. Therefore the sign of $\Delta f(x)$ changes whenever the sign of Δx changes. The sign of $\Delta f(x)$ cannot be positive, if $f(x_0)$ is the maximal value of $f(x)$, and it cannot be negative, if $f(x_0)$ is the minimal value. This is the contradiction. $\square$

Theorem 2. *If a function $f(x)$ is differentiable at x_0 and its graph is convex, then a tangent line to the graph of $f(x)$ at x_0 is $y = f(x_0) + f'(x_0)(x - x_0)$.*

Proof. Let $y = ax + b$ be the equation of a tangent line to the graph $y = f(x)$ at the point x_0. Since $ax + b$ passes through x_0, one has $ax_0 + b = f(x_0)$, therefore $b = f(x_0) - ax_0$, and it remains to prove that $a = f'(x_0)$. If the tangent line $ax + b$ is not horizontal, consider the function $g(x) = f(x) - ax$. At x_0 it takes either a maximal or a minimal value and $g'(x_0) = 0$ by Lemma **1**. On the other hand, $g'(x_0) = f'(x_0) - a$. $\square$

Differential equation. The Debeaune problem leads to a so called differential equation on $y(x)$. To be precise, the equation of the tangent line to $y(x)$ at x_0 is $y = y(x_0) + y'(x_0)(x - x_0)$. So the x-coordinate of the point T can be found from equation $0 = y(x_0) + y'(x_0)(x - x_0)$. The solution is $x = x_0 - \frac{y(x_0)}{y'(x_0)}$. The x-coordinate of V is just x_0. Hence TV is equal to $\frac{y(x_0)}{y'(x_0)}$. And Debeaune's requirement is $\frac{y(x_0)}{y'(x_0)} = a$. Or $ay' = y$. Equations that include derivatives of functions are called *differential equations*. The equation above is the simplest differential equation. Its solution takes one line. Indeed passing to differentials one get $ay'\,dx = y\,dx$, further $ady = y\,dx$, then $a\frac{dy}{y} = dx$ and $a\,d\ln y = dx$. Hence $a \ln y = x + c$ and finally $y(x) = \exp(c + \frac{x}{a})$, where $\exp x$ denotes the function inverse to the natural logarithm and c is an arbitrary constant.

Exponenta. The function inverse to the natural logarithm is called the *exponential function.* We shall call it *exponenta* to distinguish it from other exponential functions.

Theorem 3. *The exponenta is the unique solution of the differential equation $y' = y$ such that $y'(0) = 1$.*

Proof. Differentiation of equality $\ln \exp x = x$ gives $\frac{\exp' x}{\exp x} = 1$. Hence $\exp x$ satisfies the differential equation $y' = y$. For $x = 0$ this equation gives $\exp'(0) = \exp 0$. But $\exp 0 = 1$ as $\ln 1 = 0$.

For the converse, let $y(x)$ be a solution of $y' = y$. The derivative of $\ln y$ is $\frac{y'}{y} = 1$. Hence the derivative of $\ln y(x) - x$ is zero. By Theorem **16** from the previous lecture, this implies $\ln y(x) - x = c$ for some constant c. If $y'(0) = 1$, then $y(0) = 1$ and $c = \ln 1 - 0 = 0$. Therefore $\ln y(x) = x$ and $y(x) = \exp \ln y(x) = \exp x$. $\square$

Exponential series. Our next goal is to prove that

$$\exp x = 1 + x + \frac{x^2}{2} + \frac{x^3}{2 \cdot 3} + \dots + \frac{x^k}{k!} + \dots = \sum_{k=0}^{\infty} \frac{x^n}{n!}, \tag{3.2.2}$$

where $0! = 1$. This series is absolutely convergent for any x. Indeed, the ratio of its subsequent terms is $\frac{x}{n}$ tends to 0, hence it is eventually majorized by any geometric series.

Hyperbolic functions. To prove that the function presented by series (3.2.2) is virtually monotone, consider its odd and even parts. These partes represent so called *hyperbolic functions*: hyperbolic sine $\operatorname{sh} x$, and cosine $\operatorname{ch} x$.

$$\operatorname{sh}(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \qquad \operatorname{ch}(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \tag{3.2.3}$$

The hyperbolic sine is an increasing function, as all odd powers are increasing over the whole line. The hyperbolic cosine is increasing for positive x and decreasing for negative. Hence both are virtually monotone; and so is their sum.

Consider the integral $\int_0^x \operatorname{sh} t\, dt$. As all terms of the series representing sh are increasing, we can integrate the series termwise. This integration gives $\operatorname{ch} x$. As $\operatorname{sh} x$ is locally bounded, $\operatorname{ch} x$ is continuous by Theorem **13**. Consider the integral $\int_0^x \operatorname{ch} t\, dt$, here we also can integrate the series representing ch termwise, because for positive x all the terms are increasing, and for negative x, decreasing. The integration gives $\operatorname{sh} x$. Since the continuity of $\operatorname{ch} x$ was already proved. Further, by Theorem **13** we get that $\operatorname{sh} x$ is differentiable and $\operatorname{sh}' x = \operatorname{ch} x$. Now returning to equality $\operatorname{ch} x = \int_0^x \operatorname{sh} t\, dt$ we get $\operatorname{ch}' x = \operatorname{sh} x$, as $\operatorname{sh} x$ is continuous.

Therefore $(\operatorname{sh} x + \operatorname{ch} x)' = \operatorname{ch} x + \operatorname{sh} x$. And $\operatorname{sh} 0 + \operatorname{ch} 0 = 0 + 1 = 1$. Now by the above theorem **3** one gets $\exp x = \operatorname{ch} x + \operatorname{sh} x$.

Other exponential functions. The exponenta as function inverse to logarithm transforms sums into products. That is for all x and y one has

$$\exp(x + y) = \exp x \exp y \tag{3.2.4}$$

A function which has this property (i.e., transform sums into products) is called *exponential*.

Theorem 4. *For any positive a there is a unique differentiable function denoted by a^x called exponential function to base a, such that $a^1 = a$ and $a^{x+y} = a^x a^y$ for any x, y. This function is defined by the formula* $\exp a \ln x$.

Proof. Consider $l(x) = \ln a^x$. This function has the property $l(x + y) = l(x) + l(y)$. Therefore its derivative at any point is the same: it is equal to $k = \lim_{x \to 0} \frac{l(x)}{x}$. Hence the function $l(x) - kx$ is constant, because its derivative is 0. This constant is equal to $l(0)$, which is 0. Indeed $l(0) = l(0 + 0) = l(0) + l(0)$. Thus

$\ln a^x = kx$. Substituting $x = 1$ one gets $k = \ln a$. Hence $a^x = \exp(x \ln a)$. So if a differentiable exponential function with base a exists, it coincides with $\exp(x \ln a)$. On the other hand it is easy to see that $\exp(x \ln a)$ satisfies all requirements for an exponential function to base a, that is $\exp(1 \ln a) = a$, $\exp((x + y) \ln a) = \exp(x \ln a) \exp(y \ln a)$; and it is differentiable as composition of differentiable functions. □

Powers. Hence for any positive a and any real b one defines the number a^b as

$$a^b = \exp(b \ln a)$$

a is called base, and b is called exponent. For rational b this definition agrees with the old definition. Indeed if $b = \frac{p}{q}$ then properties of exponent and logarithm imply $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.

Earlier, we have defined logarithms to base b as the number c, called *logarithm of b to base a*, if $a^b = c$ and denoted $c = \log_a b$.

The basic properties of powers are collected here.

Theorem 5.

$$\left(a^b\right)^c = a^{(bc)} \qquad a^{b+c} = a^b a^c \qquad (ab)^c = a^c b^c \qquad \log_a b = \frac{\log a}{\log b}$$

Power functions The power operation allows us to define power function x^α for any real degree α. Now we can prove the equality $(x^\alpha)' = \alpha x^{\alpha-1}$ in its full value. Indeed, $(x^\alpha)' = (\exp(\alpha \ln x))' = \exp'(\alpha \ln x)(\alpha \ln x)' = \exp(\alpha \ln x)\frac{\alpha}{x} = \alpha x^{\alpha-1}$

Infinite products via the Logarithm.

Lemma 6. *Let $f(x)$ be a function continuous at x_0. Then for any sequence $\{x_n\}$ such that $\lim_{n\to\infty} x_n = x_0$ one has $\lim_{n\to\infty} f(x_n) = f(x_0)$*

Proof. For any given $\varepsilon > 0$ there is a neighborhood U of x_0 such that $|f(x) - f(x_0)| \leq \varepsilon$ for $x \in U$. As $\lim_{n\to\infty} x_n = x_0$ eventually $x_n \in U$. Hence eventually $|f(x_n) - f(x_0)| < \varepsilon$. □

As we already have remarked the infinite sums and infinite products are limits of partial products.

$$\ln \prod_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} \ln p_k \tag{3.2.5}$$

Proof.

$$\exp(\sum_{k=1}^{\infty} \ln p_k) = \exp(\lim_{n\to\infty} \sum_{k=1}^{n} \log p_k) = \lim_{n\to\infty} \exp(\sum_{k=1}^{n} \log p_k) =$$
$$\lim_{n\to\infty} \prod_{k=1}^{n} p_k = \prod_{k=1}^{\infty} p_k.$$

Hence $\exp(\sum_{k=1}^{\infty} \ln p_k) = \prod_{k=1}^{\infty} p_k$. Taking logarithm of both sides one gets (3.2.5)

□

Symmetric arguments prove the following

$$\exp \sum_{k=1}^{\infty} a_k = \prod_{k=1}^{\infty} \exp a_k \tag{3.2.6}$$

Irrationality of e. The expansion of exponent into power series gives an expansion in series for e which is $\exp 1$.

Lemma 7. *For any natural n one has $\frac{1}{n+1} < en! - [en!] < \frac{1}{n}$*

Proof. $en! = \sum_{k=0}^{\infty} \frac{n!}{k!}$. The partial sum $\sum_{k=0}^{n} \frac{n!}{k!}$ is an integer. The tail $\sum_{k=n+1}^{\infty} \frac{n!}{k!}$ is termwise majorized by geometric series $\sum_{k=1}^{\infty} \frac{1}{(n+1)^k} = \frac{1}{n}$. On the other hand the first summand of the tail is $\frac{1}{n+1}$. Consequently the tail has its sum between $\frac{1}{n+1}$ and $\frac{1}{n}$. □

Theorem 8. *The number e is irrational*

Proof. Suppose $e = \frac{p}{q}$ where p and q are natural. Then $eq!$ is a natural number. But it is not an integer by Lemma **7**. □

Problems.

1. Prove the inequalities $1 + x \leq \exp x \leq \frac{1}{1-x}$
2. Prove the inequalities $\frac{x}{1+x} \leq \log(1 + x) \leq x$
3. Evaluate $\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n$
4. Evaluate $\lim_{n\to\infty} \left(1 + \frac{2}{n}\right)^n$
5. Evaluate $\lim_{n\to\infty} \left(1 + \frac{1}{n^2}\right)^n$
6. Find the derivative of x^x.
7. $x > y$ implies $\exp x > \exp y$
8. Express via e: $\exp 2$, $\exp(1/2)$, $\exp(2/3)$, $\exp(-1)$
9. Prove that $\exp(m/n) = e^{\frac{m}{n}}$
10. Prove that $\exp x > 0$ for any x.
11. Prove the addition formulas

$$\operatorname{ch}(x+y) = \operatorname{ch}(x)\operatorname{ch}(y)+\operatorname{sh}(x)\operatorname{sh}(y) \quad \operatorname{sh}(x+y) = \operatorname{sh}(x)\operatorname{ch}(y)+\operatorname{sh}(y)\operatorname{ch}(x)$$

12. Prove that $\Delta \operatorname{sh}(x-0.5) = \operatorname{sh} 0.5 \operatorname{ch}(x)$ $\Delta \operatorname{ch}(x-0.5) = \operatorname{sh} 0.5 \operatorname{sh}(x)$.

13. Prove $\operatorname{sh} 2x = 2 \operatorname{sh} x \operatorname{ch} x$

14. Prove $\operatorname{ch}^2(x) - \operatorname{sh}^2(x) = 1$

15. Solve the equation $\operatorname{sh} x = 4/5$

16. Express via e the sum $\sum\limits_{k=1}^{\infty} k/k!$

17. Express via e the sum $\sum\limits_{k=1}^{\infty} k^2/k!$

18. Prove that $\{\frac{\exp k}{k^n}\}$ is unbounded

19. The product $\prod(1+p_n)$ converges iff the sum $\sum p_n$ $(p_n \geq 0)$ converges.

20. Determine convergence $\prod \dfrac{e^{1/n}}{1+\frac{1}{n}}$

21. Does $\prod n(e^{1/n}-1)$ converges?

22. Prove divergence of $\sum\limits_{k=1}^{\infty} \frac{[k-prime]}{k}$.

23. Expand a^x into a power series

24. Geometrical sense of $\operatorname{sh} x$ and $\operatorname{ch} x$

25. Evaluate $\lim\limits_{n\to\infty} \sin \pi e n!$

26. Does the series $\sum\limits_{k=1}^{\infty} \sin \pi e n!$ converge?

27. * Prove irrationality of e^2

3.3 Euler Formula

Complex Newton-Leibniz. For a function of a complex variable $f(z)$ the derivative is defined by the same formula $f'(z_0) = \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$. We will denote it also by $\frac{df(z)}{dz}$, to distinguish from derivative of pathes: complex valued functions of real variable. For a path $p(t)$ its derivative will be denoted either $p'(t)$ or $\frac{dp(t)}{dt}$. The Newton-Leibniz formula for real functions can be expressed The equality $\frac{df(t)}{dt}dt = df(t)$. Now we extend this formula to complex functions.

The linearization of a complex function $f(z)$ at z_0 has the same form $f(z_0)+f'(z_0)(z-z_0)+o(z)(z-z_0)$, where $o(z)$ is an infinitesimally small function of complex variable. The same arguments as for real numbers prove the basic rules of differentiation: the derivative of sum, product and composition.

Theorem 1. $\frac{dz^n}{dz} = nz^{n-1}$

Proof. $\frac{dz}{dz} = 1$ one gets immediately from the definition of the derivative. Suppose the equality $\frac{dz^n}{dz} = nz^{n-1}$ is proved for n. Then $\frac{dz^{n+1}}{dz} = \frac{dzz^n}{dz} = z\frac{dz^n}{dz} + z^n\frac{dz}{dz} = znz^{n-1}+z^n = (n+1)z^n dz$. And the theorem is proved by induction. □

A smooth path is a differentiable mapping $p\colon [a,b] \to \mathbb{C}$ with continuous bounded derivative. A function $f(z)$ of complex variable is called *virtually monotone* if for any smooth path $p(t)$ functions $\operatorname{Re} f(p(t))$ and $\operatorname{Im} f(p(t))$ are virtually monotone.

Lemma 2. *If $f'(z)$ is bounded, then $f(z)$ is virtually monotone.*

Proof. Consider a smooth path p. Then $\frac{df(p(t))}{dt} = f'(p(t))p'(t)$ is bounded by some K. Due to Lemma **15** one has $|f(p(t)) - f(p(t_0))| \leq K|t-t_0|$. Hence any partial variation of $f(p(t))$ does not exceed $K(b-a)$. Therefore $\operatorname{var}_{f(p(t))}[a,b] \leq K$. □

Theorem 3. *If a complex function $f(z)$ has a bounded virtually monotone continuous complex derivative over the image of a smooth path $p\colon [a,b] \to \mathbb{C}$, then $\int_p f'(z)\,dz = f(p(b)) - f(p(a))$.*

Proof. $\frac{df(p(t))}{dt} = f'(p(t))p'(t) = \frac{d\operatorname{Re} f(p(t))}{dt} + i\frac{d\operatorname{Im} f(p(t))}{dt}$. All functions here are continuous and virtually monotone by hypothesis. Passing to differential forms one gets $\frac{df(p(t))}{dt}\,dt = \frac{d\operatorname{Re} f(p(t))}{dt}\,dt + i\,dt\operatorname{Im} f(p(t))\,dt = d(\operatorname{Re} f(p(t))) + i\,d(\operatorname{Im} f(p(t))) = d(\operatorname{Re} f(p(t)) + i\operatorname{Im} f(p(t))) = d(f(p(t))$. Hence $\int_p f'(z)\,dz = \int_p df(z)$.

□

Corrolary 4. *If $f'(z) = 0$ then $f(z)$ is constant*

Proof. Consider $p(t) = z_0 + (z-z_0)t$, then $f(z) - f(z_0) = \int_p f'(\zeta)\,d\zeta = 0$. □

Differentiation of series. Let us say that a series of complex numbers $\sum_{k=1}^{\infty} a_k$ majorizes (eventually) another such series $\sum_{k=1}^{\infty} b_k$ if $|b_k| \leq |a_k|$ for all k (resp. for k onward some n).

The series $\sum_{k=1}^{\infty} kc_k(z-z_0)^{k-1}$ is called a *formal derivative* of $\sum_{k=0}^{\infty} c_k(z-z_0)^k$.

Lemma 5. *Any power series* $\sum_{k=0}^{\infty} c_k(z-z_0)^k$ *eventually majorizes its formal derivative* $\sum_{k=0}^{\infty} kc_k(z_1-z_0)^{k-1}$ *if* $|z_1-z_0| < |z-z_0|$.

Proof. The ratio of n-th term of the derivative to the n-th term of the series tends to 0 as n tends to infinity. Indeed, this ratio is $\frac{k(z_1-z_0)^k}{(z-z_0)^k} = kq^k$, where $|q| < 1$ since $|z_1 - z_0| < |z - z_0$. The fact that $\lim_{n\to\infty} nq^n = 0$ follows from convergence of $\sum_{k=1}^{\infty} kq^k$ which we already have proved before. This series is eventually majorized by any geometric series $\sum_{k=0}^{\infty} AQ^k$ with $Q > q$. □

A path $p(t)$ is called *monotone* if both $\operatorname{Re} p(t)$ and $\operatorname{Im} p(t)$ are monotone.

Lemma 6. *Let* $p\colon [a,b] \to \mathbb{C}$ *be a smooth monotone path, and let* $f(z)$ *be virtually monotone. If* $|f(p(t))| \leq c$ *for* $t \in [a,b]$ *then* $\left|\int_p f(z)\,dz\right| \leq 4c|p(b) - p(a)|$

Proof. Integration of the inequalities $-c \leq \operatorname{Re} f(p(t)) \leq c$ against $d\operatorname{Re} z$ along the path gives $\left|\int_p \operatorname{Re} f(z)\,d\operatorname{Re} z\right| \leq c|\operatorname{Re} p(b) - \operatorname{Re} p(a)| \leq c|p(b)-p(a)|$. The same arguments prove $\left|\int_p \operatorname{Im} f(z)\,d\operatorname{Im} z\right| \leq c|\operatorname{Im} p(b) - \operatorname{Im} p(a)| \leq c|p(b) - p(a)|$. The sum of these inequalities gives $\left|\operatorname{Re}\int_p f(z)\,dz\right| \leq 2c|\operatorname{Re} p(b) - \operatorname{Re} p(a)|$. The same arguments yields $\left|\operatorname{Im}\int_p f(z)\,dz\right| \leq 2c|\operatorname{Re} p(b) - \operatorname{Re} p(a)|$. And the addition of the two last inequalities allow us to accomplish the proof of the Lemma because $\left|\int_p f(z)\,dz\right| \leq \left|\operatorname{Re}\int_p f(z)\,dz\right| + \left|\int_p f(z)\,dz\right|$. □

Lemma 7. $|z^n - \zeta^n| \leq n|z-\zeta| \max\{|z^{n-1}|, |\zeta^{n-1}|\}$

Proof. $(z^n - \zeta^n) = (z-\zeta)\sum_{k=0}^{n-1} z^k\zeta^{n-k-1}$ and $|z^k\zeta^{n-k-1}| \leq \max\{|z^{n-1}|, |\zeta^{n-1}|\}$. □

A *linear path* from z_0 to z_1 is defined as a linear mapping $p\colon [a,b] \to \mathbb{C}$, such that $p(a) = z_0$ and $p(b) = z_1$, that is $p(t) = z_0(t-a) + (z_1-z_0)(t-a)/(b-a)$.

We denote by $\int\limits_a^b f(z)\,dz$ the integral along the linear path from a to b.

Lemma 8. *For any complex z, ζ and natural $n > 0$ one has*

$$|z^n - z_0^n - nz_0^{n-1}(z - z_0)| \leq 2n(n-1)|z - z_0|^2 \max\{|z|^{n-2}, |z_0|^{n-2}\} \tag{3.3.1}$$

Proof. By the Newton-Leibniz formula, $z^n - z_0^n = \int\limits_{z_0}^{z} n\zeta^{n-1}\,d\zeta$. Further,

$$\int\limits_{z_0}^{z} n\zeta^{n-1}\,d\zeta = \int\limits_{z_0}^{z} nz_0^{n-1}\,d\zeta + \int\limits_{z_0}^{z} n(\zeta^{n-1} - z_0^{n-1})\,d\zeta =$$

$$nz_0^{n-1} + \int\limits_{z_0}^{z} n(\zeta^{n-1} - z_0^{n-1})\,d\zeta.$$

Consequently, the left-hand side of (3.3.1) is equal to $\left|\int\limits_{z_0}^{z} n(\zeta^{n-1} - z_0^{n-1})\,d\zeta\right|$. Due to Lemma **7** the absolute value of the integrand along the linear path does not exceed $(n-1)|z-z_0|\max\{|z^{n-2}|, |z_0^{n-2}|\}$. Now the estimation of the integral by Lemma **6** gives just the inequality (3.3.1). □

Theorem 9. *If $\sum\limits_{k=0}^{\infty} c_k(z_1 - z_0)^k$ converges absolutely, then $\sum\limits_{k=0}^{\infty} c_k(z - z_0)^k$ and $\sum\limits_{k=1}^{\infty} kc_k(z - z_0)^{k-1}$ absolutely converge provided by $|z - z_0| < |z_1 - z_0|$, and the function $\sum\limits_{k=1}^{\infty} kc_k(z - z_0)^{k-1}$ is the complex derivative of $\sum\limits_{k=0}^{\infty} c_k(z - z_0)^k$.*

Proof. Series $\sum\limits_{k=0}^{\infty} c_k(z - z_0)^k$ and its formal derivative are eventually majorized by $\sum\limits_{k=0}^{\infty} c_k(z_1 - z_0)^k$ if $|z - z_0| \leq |z_1 - z_0|$ by the Lemma **5**. Hence they absolutely converge in the circle $|z - z_0| \leq |z_1 - z_0|$. Consider

$$R(z) = \sum_{k=0}^{\infty} c_k(z - z_0)^k - \sum_{k=0}^{\infty} c_k(\zeta - z_0)^k - (z - \zeta)\sum_{k=1}^{\infty} kc_k(\zeta - z_0)^{k-1}. \tag{3.3.2}$$

To prove that the formal derivative is the derivative of $\sum\limits_{k=0}^{\infty} c_k(z - z_0)^k$ at ζ it is sufficient to prove that $R(z) = o(z)(z - \zeta)$, where $o(z)$ is infinitesimally small at ζ. One has $R(z) = \sum\limits_{k=1}^{\infty} c_k\left((z - z_0)^k - (\zeta - z_0)^k - k(\zeta - z_0)^{k-1}\right)$. By Lemma **8** one gets the following estimate: $|R(z)| \leq \sum\limits_{k=1}^{\infty} 2|c_k|k(k-1)|z - \zeta|^2|z_2 - z_0|^{n-2}$, where $|z_2 - z_0| = \max\{|z - z_0|, |\zeta - z_0|\}$. Hence all we need now is to prove that $\sum\limits_{k=1}^{\infty} 2k(k-1)|c_k||z_2 - z_0|^{k-2}|z - \zeta|$ is infinitesimally small at ζ. And this in its turn follows from the convergence of $\sum\limits_{k=1}^{\infty} 2k(k-1)|c_k||z_2 - z_0|^{k-2}$. The later may be deduced from the Lemma **5**. Indeed, consider z_3, such that $|z_2 - z_0| < |z_3 - z_0| < |z_1 - z_0|$. The convergence of $\sum\limits_{k=1}^{\infty} k|c_k||z_3 - z_0|^{k-1}$ follows from the convergence of $\sum\limits_{k=0}^{\infty} |c_k||z_1 - z_0|^k$ by the Lemma **5**. And convergence of $\sum\limits_{k=2}^{\infty} k(k-1)|c_k||z_2 - z_0|^{k-2}$ follows from convergence of $\sum\limits_{k=1}^{\infty} k|c_k||z_3 - z_0|^{k-1}$ by the same Lemma □

Corrolary 10. *Let* $f(z) = \sum\limits_{k=0}^{\infty} c_k z^k$ *absolutely converges for* $|z| < r$*, and* a, b *have absolute values less then* r*. Then* $\int\limits_a^b f(z)\,dz = \sum\limits_{k=0}^{\infty} \frac{c_k}{k+1}(b^{k+1} - a^{k+1})$

Proof. Consider $F(z) = \sum\limits_{k=0}^{\infty} \frac{c_k z^{k+1}}{k+1}$. This series is term-wise majorized by the series of $f(z)$, hence it converges absolutely for $|z| < r$. By the Theorem **9** $f(z)$ is its derivative for $|z| < r$. In our case $f(z)$ is differentiable and its derivative is bounded by $\sum\limits_{k=0}^{\infty} k|c_k|r_0^k$, where $r_0 = \max\{|a|, |b|\}$. Hence $f(z)$ is continuous and virtually monotone and our result now follows from the Theorem **3**. □

Exponenta in $\mathbb{C}$. The exponenta for any complex number z is defined as $\exp z = \sum\limits_{k=0}^{\infty} \frac{z^k}{k!}$. The definition works because the series $\sum\limits_{k=0}^{\infty} \frac{z^k}{k!}$ absolutely converges for any $z \in \mathbb{C}$.

Theorem 11. . *The exponenta is a differentiable function of complex variable with derivative* $\exp' z = \exp z$*, such that for all complex* z*,* ζ *holds the following* addition formula $\exp(z + \zeta) = \exp z \exp \zeta$.

Proof. The derivative of the exponenta can be evaluated termwise by the Theorem **9**. And this evaluation gives $\exp' z = \exp z$. To prove the addition formula consider the following function $r(z) = \frac{\exp(z+\zeta)}{\exp z}$. Differentiation of the equality $r(z)\exp z = \exp(z + \zeta)$ gives $r'(z)\exp z + r(z)\exp z = \exp(z + \zeta)$. Division by $\exp z$ gives $r'(z) + r(z) = r(z)$. Hence $r(z)$ is constant. This constant is determined by substitution $z = 0$ as $r(z) = \exp \zeta$. This proves the addition formula. □

Lemma 12. *Let* $p\colon [a, b] \to \mathbb{C}$ *be a smooth path contained in the complement of a neighborhood of* 0*. Then* $\exp \int\limits_p \frac{1}{\zeta}\,d\zeta = \frac{p(b)}{p(a)}$.

Proof. First consider the case when p is contained in a circle $|z - z_0| < |z_0|$ with center $z_0 \neq 0$. In this circle, $\frac{1}{z}$ expands in a power series:

$$\frac{1}{\zeta} = \frac{1}{z_0 - (z_0 - \zeta)} = \frac{1}{z_0}\frac{1}{1 - \frac{z_0 - \zeta}{z_0}} = \sum_{k=0}^{\infty} \frac{(z_0 - \zeta)^k}{z_0^{k+1}} \tag{3.3.3}$$

The integration of this series is possible to do termwise due to Corollary **10**. Hence the result of the integration does not depends on path. And the Theorem **9** provides differentiability of the termwise integral and possibility of its termwise differentiation. Such differentiation simply gives the original series, which represents $\frac{1}{z}$ in this circle.

Consider the function $l(z) = \int\limits_{z_0}^{z} \frac{1}{\zeta}\, d\zeta$. Then $l'(z) = \frac{1}{z}$. The derivative of the composition $\exp l(z)$ is $\frac{\exp l(z)}{z}$. Hence the composition satisfies the differential equation $y'z = y$. We search for a solution of this equation in the form $y = zw$. Then $y' = w + w'z$ and our equation turns into $wz + w'z^2 = wz$. Therefore $w' = 0$ and w is constant. To find this constant substitute $z = z_0$ and get $1 = \exp 0 = \exp l(z_0) = wz_0$. Hence $w = \frac{1}{z_0}$ and $\exp l(z) = \frac{z}{z_0}$.

To prove the general case consider a partition $\{x_k\}_{k=0}^{n}$ of $[a, b]$. Denote by p_k the restriction of p over $[x_k, x_{k+1}]$. Choose the partition so small that $|p(x) - p(x_k)| < |p(x_k)|$ for all $x \in [x_k, x_{k+1}]$. Then any p_k satisfies the requirement of the above considered case. Hence $\exp \int\limits_{p_k} \frac{1}{\zeta}\, d\zeta = \frac{p(x_{k+1})}{p(x_k)}$. Further $\exp \int\limits_{p} \frac{1}{\zeta}\, d\zeta = \exp \sum\limits_{k=0}^{n-1} \int\limits_{p_k} \frac{1}{\zeta}\, d\zeta = \prod_{k=0}^{n-1} \frac{p(x_{k+1})}{p(x_k)} = p(b)/p(a)$. ☐

Theorem 13 (Euler Formula). *For any real ϕ one has*

$$\boxed{\exp i\phi = \cos\phi + i\sin\phi} \tag{3.3.4}$$

Proof. In Lecture 2.5 we have evaluated $\int\limits_{p} \frac{1}{z}\, dz = i\phi$ for $p(t) = \cos t + i \sin t$, $t \in [0, \phi]$. Hence Lemma **12** applied to $p(t)$ immediately gives the Euler formula. ☐

Trigonometric functions in $\mathbb{C}$. The Euler formula gives power series expansions for $\sin x$ and $\cos x$

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{3.3.5}$$

These expansions are used to define the trigonometric functions for complex variable. On the other hand the Euler formula allows us to express trigonometric functions via the exponenta:

$$\sin z = \frac{\exp(iz) - \exp(-iz)}{2i} \qquad \cos z = \frac{\exp(iz) + \exp(-iz)}{2} \tag{3.3.6}$$

The other trigonometric functions tan, cot, sec, cosec are defined for complex variable by usual formulas via sin and cos.

Problems.

1. Evaluate $\sum\limits_{k=1}^{\infty} \frac{\sin k}{k!}$

2. Prove the formula of Joh. Bernoulli $\int\limits_{0}^{1} x^x\, dx = \sum\limits_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^k}$

3. Find $\ln(-1)$.

4. Solve the equation $\exp z = i$.
5. Evaluate i^i.
6. Prove $\sin z = \dfrac{e^{iz} - e^{-iz}}{2i}$ $\cos z = \dfrac{e^{iz} + e^{-iz}}{2}$
7. Prove the identity $\sin^2 z + \cos^2 z = 1$
8. Solve the equation $\sin z = 5/3$
9. Evaluate $\sum\limits_{k=1}^{\infty} \frac{\sin n}{n!}$
10. Solve the equation $\cos z = 2$
11. Evaluate $\sum\limits_{k=0}^{\infty} \frac{\cos n}{n!}$
12. Evaluate $\oint\limits_{|z|=1} \dfrac{dz}{z^2}$
13. Evaluate $\sum\limits_{k=1}^{\infty} q^k \dfrac{\sin kx}{k}$
14. Expand into a power series $e^x \cos x$

3.4 Abel's Theorem

On the contents of the lecture. Expansion of logarithm into power series will be extended to the complex case. We learn very important Abel's transformation of sum. This transformation is a discrete analogue of integrations by parts. Abel's theorem on the limit of power series will be applied to evaluation of trigonometric series related to logarithm. The concept of Abel's sum of a divergent series will be introduced.

Principal branch of the Logarithm. Since $\exp(x+iy) = e^x(\cos y + i\sin y)$ one gets the following formula for the logarithm: $\operatorname{Log} z = \ln|z| + i\operatorname{Arg} z$, where $\operatorname{Arg} z = \arg z + 2\pi k$. We see that the logarithm is a multi-valued function, that is why one usually choose a *branch* of logarithm to work. For our purposes it is sufficient to consider the *principal* branch of logarithm:

$$\ln z = \ln|z| + i\arg z, \qquad -\pi < \arg z \leq \pi. \tag{3.4.1}$$

The principal branch of the logarithm is a differentiable function of complex variable with derivative $\frac{1}{z}$, inverse to $\exp z$. This branch is not continuous at negative numbers. However its restriction on upper half-plane is continuous and even differentiable at negative numbers.

Lemma 1. *For any nonnegative z one has $\int_1^z \frac{1}{\zeta}\,d\zeta = \ln z$.*

Proof. If $\operatorname{Im} z \neq 0$, the segment $[0,z]$ is contained in the circle $|\zeta - z_0| < |z_0|$ for $z_0 = \frac{|z|^2}{\operatorname{Im} z}$. In this circle $\frac{1}{\zeta}$ expands into a power series, which one can integrate termwise. Since for z^k the result of integration depends only on the ends of integration path, the same is true for power series. Hence, we can change path of integration without changing the result. Consider the following path: $p(t) = \cos t + i\sin t$, $t \in [0, \arg z]$. We know the integral $\int_p \frac{1}{\zeta}\,d\zeta = i\arg z$. This path terminates at $\frac{z}{|z|}$. Continue this path by the linear path to z. The integral $\int_{z/|z|}^{z} \frac{1}{\zeta}\,d\zeta = \int_1^{|z|} \frac{1}{z/|z|t}\,dtz/|z| = \int_1^{|z|} \frac{1}{t}\,dt = \ln|z|$. Therefore $\int_1^z \frac{1}{\zeta}\,d\zeta = \int_p \frac{1}{\zeta}\,d\zeta + \int_{z/|z|}^{z} \frac{1}{\zeta}\,d\zeta = i\arg z + \ln|z|$. □

Logarithmic series. In particular for $|1-z| < 1$ the termwise integration of the series $\frac{1}{\zeta} = \sum_{k=0}^{\infty} (1-\zeta)^k$ gives the complex Mercator series:

$$\ln(1+z) = \sum_{k=1}^{\infty} (-1)^{k+1}\frac{z^k}{k}. \tag{3.4.2}$$

Substitute in this series $-z$ for z and subtract the obtained series from (3.4.2)to get the complex Gregory series:

$$\ln\frac{1+z}{1-z} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{2k+1} \tag{3.4.3}$$

In particular for $z = ix$, one has $\left|\frac{1+ix}{1-ix}\right| = 1$ and $\arg\frac{1+ix}{1-ix} = 2\operatorname{arctg} x$. Therefore one gets

$$\operatorname{arctg} x = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1} \tag{3.4.4}$$

Since $\arg(1+e^{i\phi}) = \operatorname{arctg}\frac{\sin\phi}{1+\cos\phi} = \operatorname{arctg}\tan(\phi/2) = \frac{\phi}{2}$. The substitution of $\exp(i\phi)$ for z in the Mercator series $\ln(1+e^{i\phi}) = \sum\limits_{k=1}^{\infty}(-1)^{k+1}\frac{e^{ik\phi}}{k}$ gives for the imaginary parts:

$$\sum_{k=0}^{\infty}(-1)^{k+1}\frac{\sin k\phi}{k} = \frac{\phi}{2} \tag{3.4.5}$$

However the last substitution is not correct, because $|e^{i\phi}| = 1$ and (3.4.2) is proved only for $|z| < 1$. To justify it we will prove a general Theorem, due to Abel.

Summation by parts. Consider two sequences $\{a_k\}_{k=1}^{n}$, $\{b_k\}_{k=1}^{n}$. The difference of their product $\Delta a_k b_k = a_{k+1}b_{k+1} - a_k b_k$ can be presented as

$$\Delta(a_k b_k) = a_{k+1}\Delta b_k + b_k\Delta a_k \tag{3.4.6}$$

Summation of these equalities gives

$$a_n b_n - a_1 b_1 = \sum_{k=1}^{n-1} a_{k+1}\Delta b_k + \sum_{k=1}^{n-1} b_k\Delta a_k \tag{3.4.7}$$

A permutation of the latter equality gives so called *Abel's transformation* of sums

$$\sum_{k=1}^{n-1} b_k\Delta a_k = a_n b_n - a_1 b_1 - \sum_{k=1}^{n-1} a_{k+1}\Delta b_k \tag{3.4.8}$$

Abel's theorem. One writes $x \to a - 0$ instead of $x \to a$ and $x < a$, and $x \to a + 0$ means $x > a$ and $x \to a$.

Theorem 2 (Abel). *If* $\sum\limits_{k=0}^{\infty} a_k$ *converges, then* $\lim\limits_{x\to 1-0}\sum\limits_{k=0}^{\infty} a_k x^k$

Proof. $\sum\limits_{k=0}^{\infty} a_k x^k$ converges absolutely for $|x| < 1$, because of boundedness of $\{a_k\}$.

Given an $\varepsilon > 0$. Set $A(n,m) = \sum\limits_{k=n}^{m} a_k$, $A(n,m)(x) = \sum\limits_{k=n}^{m} a_k x^k$. Choose N so large that

$$|A(0,n) - A(0,\infty)| < \frac{\varepsilon}{9} \quad (\quad \forall n > N) \tag{3.4.9}$$

Applying the Abel transformation for any $m > n$ one gets

(3.4.10)
$$A(n,m) - A(n,m)(x) = \sum_{k=n}^{m} a_k(1-x^k) = (1-x)\sum_{k=n}^{m} \Delta A(n-1,k-1)\sum_{j=0}^{k-1} x^j =$$
$$(1-x)\left(A(n-1,m)\sum_{j=0}^{m} x^j - A(n-1,n)\sum_{j=0}^{n} x^j - \sum_{k=n}^{m} A(n-1,k)x^k\right)$$

By (3.4.9) for $n > N$, one gets $|A(n-1,m)| = |(A(0,m)-A)+(A-A(0,n))| \leq \varepsilon/9 + \varepsilon/9 = 2\varepsilon/9$. Hence, we can estimate the absolute value of the expression in the brackets (3.4.10) from above by $\frac{2\varepsilon/3}{1-x}$. As result we get

$$|A(n,m) - A(n,m)(x)| \leq \frac{2\varepsilon}{3} \qquad \forall m \geq n > N, \quad \forall x \tag{3.4.11}$$

Since $\lim\limits_{x\to 1-0} A(0,N)(x) = A(0,N)$ one chooses δ so small, that for $x > 1-\delta$ the following inequality holds:

$$|A(0,N) - A(0,N)(x)| < \frac{\varepsilon}{3} \tag{3.4.12}$$

Summing up this inequality with (3.4.11) for $n = N+1$ one gets:

$$|A(0,m) - A(0,m)(x)| < \varepsilon \quad \forall m > N, |1-x| < \delta \tag{3.4.13}$$

Passing to limits as m tends to infinity the latter inequality gives

$$|A(0,\infty) - A(0,\infty)(x)| \leq \varepsilon, \quad \text{ for } |1-x| < \delta \tag{3.4.14}$$

□

Leibniz series. As the first application of the Abel Theorem evaluate the Leibniz series $\sum\limits_{k=0}^{\infty} \frac{(-1)^k}{2n+1}$. This series converges by the Leibniz Theorem. By the Abel Theorem its sum is $\lim\limits_{x\to 1-0} \sum\limits_{k=0}^{\infty} \frac{(-1)^k x^k}{2n+1} = \lim\limits_{x\to 1-0} \operatorname{arctg} x = \operatorname{arctg} 1 = \frac{\pi}{4}$. We get the following remarkable equality:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots \tag{3.4.15}$$

Abel sum of series. One defines the *Abel sum* of a series $\sum\limits_{k=0}^{\infty} a_k$ as the limit $\lim\limits_{x\to 1-0} \sum\limits_{k=0}^{\infty} a_k x^k$. The series which has Abel sum are called *Abel summable.* The Abel Theorem shows that all convergent series has Abel sum coinciding with its usual sum. However there are a lot of series, which has Abel sum, but do not converge.

Abel's inequality. Consider a series $\sum_{k=1}^{\infty} a_k b_k$, where partial sums $A_n = \sum_{k=1}^{n-1} a_k$ are bounded by some constant A and the sequence $\{b_k\}$ is monotone. Then $\sum_{k=1}^{n-1} a_k b_k = \sum_{k=1}^{n-1} b_k \Delta A_k = A_n b_n - A_1 b_1 + \sum_{k=1}^{n-1} A_{k+1} \Delta b_k$. Since $\sum_{k=1}^{n-1} |\Delta b_k| = |b_n - b_1|$, one gets the following inequality:

$$\left| \sum_{k=1}^{n-1} a_k b_k \right| \leq 3A \max\{|b_k|\} \tag{3.4.16}$$

Convergence test.

Theorem 3. *Let the sequence of partial sums* $\sum_{k=1}^{n-1} a_k$ *be bounded, and* $\{b_k\}$ *be non-increasing and infinitesimally small. Then* $\sum_{k=1}^{\infty} a_k b_k$ *converges to its Abel sum, if the latter exists.*

Proof. The difference between a partial sum $\sum_{k=1}^{n-1} a_k b_k$ and the Abel sum is equal to

$$\lim_{x \to 1-0} \sum_{k=1}^{n-1} a_k b_k (1 - x^k) + \lim_{x \to 1-0} \sum_{k=n}^{\infty} a_k b_k x^k \tag{3.4.17}$$

The first limit is zero, the second limit can be estimated by Abel's inequality from above by $3Ab_n$. It tends to 0 as n tends to infinity. □

Application. Now we are ready to prove the equality (3.4.5). The series $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin kx}{k}$ has an Abel sum. Indeed,

$$\lim_{q \to 1-0} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{q^k \sin kx}{k} = \operatorname{Im} \lim_{q \to 1-0} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(qe^{ix})^k}{k} =$$
$$\operatorname{Im} \lim_{q \to 1-0} \ln(1 + qe^{ix}) = \operatorname{Im} \ln(1 + e^{ix}).$$

The sums $\sum_{k=1}^{n-1} \sin kx = \operatorname{Im} \sum_{k=1}^{n-1} e^{ikx} = \operatorname{Im} \frac{1 - e^{inx}}{1 - e^{ix}}$ are bounded. And $\frac{1}{k}$ is decreasing and infinitesimally small. Hence we can apply the Theorem **3**.

Problems.

1. Evaluate $1 + \frac{1}{2} - \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \frac{1}{6} - \frac{1}{7} - \frac{1}{8} \dots$
2. Evaluate $\sum_{k=1}^{\infty} \frac{\sin 2k}{k}$

3. $\sum_{k=1}^{\infty} \frac{\cos n\phi}{n} = -\ln|2\sin\frac{\phi}{2}|$, $(0 < |\phi| \le \pi)$

4. $\sum_{k=1}^{\infty} \frac{\sin n\phi}{n} = \frac{\pi-\phi}{2}$, $(0 < \phi < 2\pi)$

5. $\sum_{k=0}^{\infty} \frac{\cos(2n+1)\phi}{2n+1} = \frac{1}{2}\ln|2\cot\frac{\phi}{2}|$, $(0 < |\phi| < \pi)$

6. $\sum_{k=0}^{\infty} \frac{\sin(2n+1)\phi}{2n+1} = \frac{\pi}{4}$, $(0 < \phi < \pi)$

7. $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{\cos n\phi}{n} = \ln\left(2\cos\frac{\phi}{2}\right)$, $(-\pi < \phi < \pi)$

8. Find the Abel sum of $1 - 1 + 1 - 1 \ldots$

9. Find the Abel sum of $1 - 1 + 0 + 1 - 1 + 0 \ldots$

10. A periodic series, such that the sum of the period is zero, has an Abel sum.

11. Telescope $\sum_{k=1}^{\infty} \dfrac{k^2}{2^k}$

12. Evaluate $\sum_{k=0}^{n-1} k\cos kx$

13. Estimate from above $\sum_{k=n}^{\infty} \frac{\sin kx}{k^2}$

14. * If $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ and their convolution $\sum_{k=0}^{\infty} c_k$ converge then $\sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k$

3.5 Residues Theory

On the contents of the lecture. At last, the reader learns something, which Euler did not know, and what he would highly appreciate. The residues theory allows one to evaluate a lot of integrals which were not accessible by Newton-Leibniz formula.

Monotone curve. A *monotone curve* Γ is defined as a subset of the complex plane, which is the image of a monotone path. Nonempty intersections of vertical and horizontal lines with a monotone curve are either points or closed intervals.

The points of the monotone curve which have extremal sum of real and imaginary parts are called its *endpoints*, other points of the curve are called its *interior* points.

A continuous injective monotone path p whose image coincides with Γ is called a *parameterization* of Γ

Lemma 1. *Let $p_1\colon [a,b] \to \mathbb{C}$ and $p_2\colon [c,d] \to \mathbb{C}$ be two parameterizations of the same monotone curve Γ. Then $p_1^{-1}p_2\colon [c,d] \to [a,b]$ is a continuous monotone bijection.*

Proof. Set $P_i(t) = \operatorname{Re} p_i(t) + \operatorname{Im} p_i(t)$. Then P_1 and P_2 are continuous and strictly monotone. And $p_1(t) = p_2(\tau)$ iff $P_1(t) = P_2(\tau)$. Hence $p_1^{-1}p_2 = P_1^{-1}P_2$. Since P_1^{-1} and P_2 are monotone continuous, the composition $P_1^{-1}P_2)$ is monotone continuous. □

Orientation. One says that two parameterizations p_1 and p_2 of a monotone curve Γ have the *same orientation*, if $p_1^{-1}p_2$ is increasing, and one says that they have the *opposite orientations*, if $p_1^{-1}p_2$ is decreasing.

Orientation divides all parameterizations of a curve into two classes. All elements of one orientation class has the same orientation and all elements of different classes have the opposite orientation.

An oriented curve is a curve with fixed *circulation direction*. A choice of orientation means distinguishing of one of the orientation classes as positive, corresponding to the oriented curve. For a monotone curve, to specify its orientation, it is sufficient to indicate which of its endpoints is its beginning and which is the end. Then all positively oriented parameterizations start with its beginning and finish at its end, and negatively oriented parameterizations do the opposite.

If an oriented curve is denoted by Γ, then its *body*, the curve without orientation, is denoted $|\Gamma|$ and the curve with the same body but with opposite orientation is denoted $-\Gamma$.

If Γ' is a monotone curve which is contained in an oriented curve Γ, then one defines the *induced orientation* on Γ' by Γ as the orientation of a parameterization of Γ' which extends to a positive parameterization of Γ.

Line integral. One defines the integral $\int\limits_\Gamma f(z)\,dg(z)$ along a oriented monotone curve Γ as the integral $\int\limits_p f(z)\,dg(z)$, where p is a positively oriented parameterization of Γ. This definition does not depend on the choice of p, because different parameterizations are obtained from each other by an increasing change of the variable (Lemma **1**).

One defines a *partition of a curve* Γ by a point x as a pair of monotone curves Γ_1, Γ_2, such that $\Gamma = \Gamma_1 \cup \Gamma_2$ and $\Gamma_1 \cap \Gamma_2 = x$. And we write in this case $\Gamma = \Gamma_1 + \Gamma_2$.

The Partition Rule for the line integral is

$$\int_{\Gamma_1+\Gamma_2} f(z)\, dg(z) = \int_{\Gamma_1} f(z)\, dg(z) + \int_{\Gamma_2} f(z)\, dg(z), \tag{3.5.1}$$

where the orientations on Γ_i are induced by an orientation of Γ. To prove the Partition Rule consider a positive parameterization $p\colon [a,b] \to \Gamma$. Then the restrictions of p over $[a, p^{-1}(x)]$ and $[p^{-1}(x), b]$ give a positive parameterizations of Γ_1 and Γ_2. Hence the equality (3.5.1) follows from $\int_a^{p^{-1}(x)} f(z)\, dg(z) + \int_{p^{-1}(x)}^b f(z)\, dg(z) = \int_a^b f(z)\, dg(z)$.

A sequence of oriented monotone curves $\{\Gamma_k\}_{k=1}^n$ with disjoint interiors is called a *chain* of monotone curves and denoted by $\sum_{k=1}^n \Gamma_k$. The body of a chain $C = \sum_{k=1}^n \Gamma_k$ is defined as $\cup_{k=1}^n |\Gamma_k|$ and denoted by $|C|$. The interior of the chain is defined as the union of interiors of its elements.

The integral of a form $f\,dg$ along the chain is defined as $\int_{\sum_{k=1}^n \Gamma_k} f\,dg = \sum_{k=1}^n \int_{\Gamma_k} f\,dg$.

One says that two chains $\sum_{k=1}^n \Gamma_k$ and $\sum_{k=1}^m \Gamma'_k$ have the same orientation, if the orientations induced by Γ_k and Γ'_j on $\Gamma_k \cap \Gamma'_j$ coincide in the case when $\Gamma_k \cap \Gamma'_j$ has a nonempty interior. Two chains with the same body and orientation are called *equivalent*.

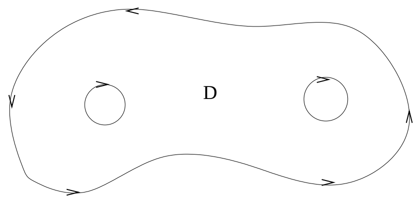
Lemma 2. *If two chains $C = \sum_{k=1}^n \Gamma_k$ and $C' = \sum_{k=1}^m \Gamma'_k$ are equivalent then the integrals along these chains coincide for any form $f dg$.*

Proof. For any interior point x of the chain C, one defines the subdivision of C by x as $\Gamma_j^+ + \Gamma_j^- + \sum_{k=1}^n \Gamma_k[k \neq j]$, where Γ_j is the curve containing x and $\Gamma_j^+ + \Gamma_j^-$ is the partition of Γ by x. The subdivision does not change the integral along the chain due to the Partition Rule.

Hence we can subdivide C step by step by endpoints of C' to construct a chain Q which endpoints include all endpoints of P'. And the integral along Q is the same as along P. Another possibility to construct Q is to subdivide C' by endpoints of C. This construction shows that the integral along Q coincides with the integral along C'. Hence the integrals along C and C' coincide. $\square$

Due to this lemma, one can introduce the integral of a differential form along any oriented piecewise monotone curve Γ. To do this one consider a *monotone* *partition* of Γ into a sequence $\{\Gamma_k\}_{k=1}^n$ of monotone curves with disjoint interiors and equip all Γ_k with induced orientation. One gets a chain and the integral along this chain does not depends on partition.

Contour integral. A *domain* D is defined as a connected bounded part of the plane with piecewise monotone boundary. The boundary of D denoted ∂D is the union of finitely many monotone curves. And we suppose that $\partial D \subset D$, that is we consider a closed domain.



For a monotone curve Γ, which is contained in the boundary of a domain D, one defines the *induced orientation* of Γ by D as the orientation of a parameterization which leaves D on the left during the movement along Γ around D.

One introduces the integral $\oint\limits_{\partial D} f(z)dg(z)$ as the integral along any chain whose body coincides with ∂D and orientations of curves are induced by D.

Due to the Lemma **2** the choice of chain does not affect the integral.

Lemma 3. *Let D be a domain and l be either a vertical or a horizontal line, which bisects D into two parts: D' and D'' lying on the different sides of l. Then* $\oint\limits_{\partial D} f(z)dz = \oint\limits_{\partial D'} f(z)dz + \oint\limits_{\partial D''} f(z)dz$

Proof. The line l intersects the boundary of D in a finite sequence of points and intervals $\{J_k\}_{k=1}^m$.

Set $\partial' D = \partial D \cap \partial D'$ and $\partial'' D = \partial D \cap \partial D''$. The intersection $\partial' D \cap \partial'' D$ consists of finitely many points. Indeed, the interior points of J_k do not belong to this intersection, because their small neighborhoods have points of D only from the one side of l. Hence

$$\int\limits_{\partial' D} f(z)\,dz + \int\limits_{\partial'' D} f(z)\,dz = \oint\limits_{\partial D} f(z)dz \tag{3.5.2}$$

The boundary of D' consists of $\partial' D$ and some number of intervals. And the boundary of D'' consists of $\partial'' D$ and the same intervals, but with opposite orientation. Therefore

$$L = \int\limits_{l\cap\partial D'} f(z)\,dz = -\int\limits_{l\cap\partial D''} f(z)\,dz. \tag{3.5.3}$$

On the other hand

$$\oint_{\partial D'} f(z)dz = \int_{\partial' D} f(z)\,dz + L \text{ and } \oint_{\partial D''} f(z)dz = \int_{\partial'' D} f(z)\,dz - L,$$

hence

$$\oint_{\partial D'} f(z)dz + \oint_{\partial D''} f(z)dz = \int_{\partial' D} f(z)\,dz + \int_{\partial'' D} f(z)\,dz = \oint_{\partial D} f(z)dz$$

$\square$

Lemma 4 (on estimate). *If $|f(z)| \leq B$ for any z from a body of a chain $C = \sum_{k=1}^{n} \Gamma_k$, then* $\left|\int_C f(z)\,dz\right| \leq 4Bn \operatorname{diam}|C|$

Proof. By the Lemma **6** for any k one has $\left|\int_{\Gamma_k} f(z)\,dz\right| \leq 4B|A_k - B_k| \leq 4B \operatorname{diam}|C|$ where A_k and B_k are endpoints of Γ_k. The summation of these inequalities proves the Lemma. $\square$

Theorem 5 (Cauchy). *If a function f is complex differentiable in a domain D then* $\oint_{\partial D} f(z)dz = 0$.

Proof. Fix a rectangle R with sides parallel to coordinate axis which contains D and denote by $|R|$ its area and by P its perimeter.

The proof is by contradiction. Suppose $\left|\oint_{\partial D} f(z)\,dz\right| \neq 0$. Denote by c the ratio of $\left|\oint_{\partial D} f(z)\,dz\right| / |R|$. We will construct a nested sequence of rectangles $\{R_k\}_{k=0}^{\infty}$ such that

- $R_0 = R$, $R_{k+1} \subset R_k$;
- R_{2k} is similar to R;
- $\left|\oint_{\partial(R_k \cap D)} f(z)\,dz\right| \geq c|R_k|$, where $|R_k|$ is the area of R_k.

The induction step. Suppose R_k is already constructed. Divide R_k in two equal rectangular R'_k and R''_k by drawing either vertical, if k is even, or horizontal, if k is odd, interval joining middles of the opposite sides of R_k. Set $D_k = D \cap R_k$, $D' = D \cap R'_k$, $D'' = D \cap R''_k$. We state that at least one of the following inequalities holds:

$$\left|\oint_{\partial D'} f(z)dz\right| \geq c|R'_k| \qquad \left|\oint_{\partial D''} f(z)dz\right| \geq c|R''_k|. \tag{3.5.4}$$

Indeed, in the opposite case one gets

$$\left|\oint_{\partial D'} f(z)dz + \oint_{\partial D''} f(z)dz\right| < c|R'_k| + c|R'_k| = c|R_k| \tag{3.5.5}$$

Since $\oint\limits_{\partial D'} f(z)dz + \oint\limits_{\partial D''} f(z)dz = \oint\limits_{\partial D_k} f(z)dz$ by the Lemma **3** we get the contradiction with the hypothesis $\left|\int\limits_{p_k} f(z)\,dz\right| \geq c|R_k|$. Hence, one of the inequalities (3.5.4) holds. If the first inequality holds we set $R_{k+1} = R_k'$ else we set $R_{k+1} = R_k''$.

After constructing the sequence $\{R_k\}$, consider the point z_0 belonging to $\cap_{k=1}^{\infty} R_k$. This point belongs to D, because all its neighborhood contains points of D. Consider the linearization $f(z) = f(z_0) + f'(z_0)(z - z_0) + o(z)(z - z_0)$. Since $\oint\limits_{\partial D_k} (f(z_0) + f'(z_0)(z - z_0))dz = 0$ one gets

$$\left|\oint\limits_{\partial D_k} o(z)(z - z_0)dz\right| = \left|\oint\limits_{\partial D_k} f(z)dz\right| \geq c|R_k| \tag{3.5.6}$$

The boundary of D_k is contained in the union $\partial R_k \cup R_k \cap \partial D$. Consider a monotone partition $\partial D = \sum\limits_{k=1}^{n} \Gamma_k$. Since the intersection of R_k with a monotone curve is a monotone curve, one concludes that $\partial D \cap R_k$ is a union of at most n monotone curves. As ∂R_k consists of 4 monotone curves we get that ∂D_k is as a body of a chain with at most $4 + n$ monotone curves.

Denote by P_k the perimeter of R_k. And suppose that $o(x)$ is bounded in R_k by a constant o_k. Then $|o(x)(z - z_0)| \leq P_k o_k$ for all $z \in R_k$.

Since $\operatorname{diam} \partial D_k \leq \frac{P_k}{2}$ by the Estimation Lemma **4**, we get the following inequality:

$$\left|\oint\limits_{\partial D_k} o(z)(z - z_0)dz\right| \leq 4(4 + n)P_k o_k \frac{P_k}{2} = 2(4 + n)o_k P_k^2 \tag{3.5.7}$$

The ratio $P_k^2/|R_k|$ is constant for even k. Therefore the inequalities (3.5.6) and (3.5.7) contradict each other for $o_k < \frac{c|R_k|}{2(4+n)P_k^2} = \frac{c|R|}{2(4+n)P^2}$. However the inequality $|o(x)| < \frac{c|R|}{2(4+n)P^2}$ holds for some neighborhood V of z_0 as $o(x)$ is infinitesimally small at z_0. This is a contradiction, because V contains some R_{2k}. □

Residues. By $\oint\limits_{z_0}^{r} f(z)\,dz$ we denote the integral along the boundary of the disk $\{|z - z_0| \leq r\}$.

Lemma 6. *If a function $f(z)$ is complex differentiable in the domain D with the exception of a finite set of points $\{z_k\}_{k=1}^n$. Then $\oint\limits_{\partial D} f(z)dz = \sum\limits_{k=1}^{n} \oint\limits_{z_k}^{r} f(z)\,dz$, where r is so small that all disks $|z - z_k| < r$ are contained in D and disjoint.*

Proof. Denote by D' the complement of the union of the disks in D. Then $\partial D'$ is the union of ∂D and the boundary circles of the disks. By the Cauchy Theorem $\oint\limits_{\partial D'} f(z)dz = 0$. On the other hand this integral is equal to the sum $\oint\limits_{\partial D} f(z)dz$ and the sum of integrals along boundaries of the circles. The orientation induced by D' onto the boundaries of these circles is opposite to the orientation induced from the circles. Hence $0 = \oint\limits_{\partial D'} f(z)dz = \oint\limits_{\partial D} f(z)dz - \sum\limits_{k=1}^{n} \oint\limits_{z_k}^{r} f(z)\,dz$. □

A *singular point* of a complex function is defined as a point where either the function or its derivative are not defined. A singular point is called isolated, if it has a neighborhood, where it is the only singular point.

One defines the *residue* of f at a point z_0 and denotes it as $\operatorname{res}_{z_0} f$ as the limit $\lim\limits_{r\to 0} \frac{1}{2\pi i} \oint\limits_{z_0}^{r} f(z)dz$. The above lemma shows that this limit exists for any isolated singular point and moreover, that all integrals along sufficiently small circumferences in this case are the same.

Since in all regular points the residues are 0 the conclusion of the Lemma **6** for a function with finitely many singular points can be presented in the form:

$$\boxed{\oint\limits_{\partial D} f(z)dz = 2\pi i \sum_{z\in D} \operatorname{res}_z f} \tag{3.5.8}$$

An isolated singular point z_0 is called a *simple pole* of a function $f(z)$, if there exists a nonzero limit $\lim\limits_{z\to z_0} f(z)(z-z_0)$.

Lemma 7. *If z_0 is a simple pole of $f(z)$ then* $\operatorname{res}_{z_0} f = \lim\limits_{z\to z_0}(z-z_0)f(z)$

Proof. Set $L = \lim\limits_{z\to z_0}(z-z_0)f(z)$. Then $f(z) = \frac{L}{z} + \frac{o(z)}{(}z-z_0)$, where $o(z)$ is infinitesimally small at z_0. Hence

$$\oint\limits_{z_0}^{r} \frac{o(z)\,dz}{z-z_0} = \oint\limits_{z_0}^{r} f(z)\,dz - \oint\limits_{z_0}^{r} \frac{L}{z-z_0}\,dz. \tag{3.5.9}$$

Since the second integral from the right-hand side of (3.5.9) is equal to $2L\pi i$ and another one is equal to $2\pi i \operatorname{res}_{z_0} f$ for sufficiently small r, we conclude that the integral from the left-hand side also is constant for sufficiently small r. To prove that $L = \operatorname{res}_{z_0} f$ we have to prove that this constant $c = \lim\limits_{r\to 0} \oint\limits_{z_0}^{r} \frac{o(z)}{z-z_0}\,dz$ is 0. Indeed, assume that $|c| > 0$. Then there is a neighborhood U of z_0 such that $|o(z)| \leq \frac{|c|}{32}$ for all $z \in U$. Then one gets a contradiction by estimation of $\left|\oint\limits_{z_0}^{r} \frac{o(z)\,dz}{z-z_0}\right|$ (which is equal to $|c|$ for sufficiently small r) from above by $\frac{|c|}{\sqrt{2}}$ for r less than the radius of U. Indeed, the integrand is bounded by $\frac{|c|}{32r}$ and the path of integration (the circle) can be divided into four monotone curves of diameter $r\sqrt{2}$: quarters of the circle. Hence by Lemma on Estimate one gets $\left|\oint\limits_{z_0}^{r} \frac{o(z)\,dz}{z-z_0}\right| \leq 16\sqrt{2}\frac{|c|}{32} = \frac{|c|}{\sqrt{2}}$. □

Remark 1. *Denote by $\Gamma(r,\phi,z_0)$ an arc of the circle $|z-z_0| = r$, which angle measure is ϕ. Under hypothesis of Lemma* **7** *the same arguments prove the following* $\lim\limits_{r\to 0} \int\limits_{\Gamma(\phi,r,0z)} f(z)\,dz = i\phi \lim\limits_{f}(z)(z-z_0)$.

Problems.

1. Evaluate $\oint\limits_{1}^{1} \dfrac{1}{1+z^4}$

2. Evaluate $\oint\limits_{0}^{1} \frac{dz}{\sin z}$

3. Evaluate $\oint\limits_{0}^{1} \frac{dz}{e^z-1}$

4. Evaluate $\oint\limits_{0}^{1} \frac{dz}{z^2}$

5. Evaluate $\oint\limits_{0}^{1} \sin\frac{1}{z}\,dz$

6. Evaluate $\oint\limits_{0}^{1} ze^{\frac{1}{z}}\,dz$

7. Evaluate $\oint\limits_{0}^{5/2} z^2 \cot \pi z\,dz$

8. Evaluate $\oint\limits_{2}^{\frac{1}{2}} \frac{z\,dz}{(z-1)(z-2)^2}$

9. $\int\limits_{-\pi}^{+\pi} \frac{d\phi}{5+3\cos\phi}$

10. $\int\limits_{-\pi}^{+\pi} \frac{d\phi}{(1+\cos^2\phi)^2}$

11. $\int\limits_{0}^{2\pi} \frac{d\phi}{(1+\cos\phi)^2}$

12. $\int\limits_{-\infty}^{+\infty} \frac{dx}{1+x^4}$

13. $\int\limits_{0}^{+\infty} \frac{dx}{(1+x^2)(4+x^2)}$

14. $\int\limits_{-\infty}^{+\infty} \frac{1+x^2}{1+x^4}$

15. $\int\limits_{-\infty}^{+\infty} \frac{x^3}{1+x^6}\,dx$

3.6 Analytic Functions

On the contents of the lecture. This lecture introduces the reader into phantastically beautiful world of analytic functions. Integral Cauchy formula, Taylor series, Fundamental Theorem of Algebra. The reader will see all of these treasures in a single lecture.

Theorem 1 (Integral Cauchy Formula). *If function f is complex differentiable in the domain D, then for any interior point $z \in D$ one has*

$$\boxed{f(z) = \frac{1}{2\pi i} \oint\limits_{\partial D} \frac{f(\zeta)\,dz}{\zeta - z}} \tag{3.6.1}$$

Proof. The function $\frac{f(z)}{z-z_0}$ has its only singular point inside the circle. This is z_0, which is a simple pole. The residue of $\frac{f(z)}{z-z_0}$ by Lemma **7** is equal to $\lim\limits_{z\to z_0}(z-z_0)\frac{f(z)}{z-z_0} = \lim\limits_{z\to z_0} f(z) = f(z_0)$. And by the formula (3.5.8) the integral is equal to $2\pi i f(z_0)$. □

Lemma 2. *Let $\sum\limits_{k=1}^{\infty} f_k$ be a series of virtually monotone complex functions, which is termwise majorized by a convergent positive series $\sum\limits_{k=1}^{\infty} c_k$ on a monotone curve Γ (that is $|f_k(z)| \leq c_k$ for natural k and $z \in \Gamma$) and such that $F(z) = \sum\limits_{k=1}^{\infty} f_k(z)$ is virtually monotone. Then*

$$\sum_{k=1}^{\infty} \int\limits_{\Gamma} f_k(z)\,dz = \int\limits_{\Gamma} \sum_{k=1}^{\infty} f_k(z)\,dz \tag{3.6.2}$$

Proof. By the Lemma on Estimate one has the following inequalities:

$$\left|\int\limits_{\Gamma} f_k(z)\,dz\right| \leq 4c_k \operatorname{diam}\Gamma \qquad \left|\int\limits_{\Gamma} \sum_{k=n}^{\infty} f_k(z)\,dz\right| \leq 4 \operatorname{diam}\Gamma \sum_{k=n}^{\infty} c_k \tag{3.6.3}$$

Set $F_n(z) = \sum\limits_{k=1}^{n-1} f_k(z)$. By the left inequality of (3.6.3), the module of difference between $\int\limits_{\Gamma} F_n(z)\,dz = \sum\limits_{k=1}^{n-1} \int\limits_{\Gamma} f_k(z)\,dz$ and the left-hand side of (3.6.2) does not exceed $4 \operatorname{diam}\Gamma \sum\limits_{k=n}^{\infty} c_k$. Hence this module is infinitesimally small as n tends to infinity. On the other hand, by the right inequality of (3.6.3) one gets $\left|\int\limits_{\Gamma} F_n(z)\,dz - \int\limits_{\Gamma} F(z)\,dz\right| \leq 4 \operatorname{diam}\Gamma \sum\limits_{k=n}^{\infty} c_k$. This implies that the difference between the left-hand and right-hand sides of (3.6.2) is infinitesimally small as n tends to infinity. But this difference does not depend on n. Hence it is zero. □

Lemma 3. *If a real function f defined over interval $[a, b]$ is locally bounded, then it is bounded.*

Proof. The proof is by contradiction. Suppose that f is unbounded. Divide the interval $[a, b]$ in half. Then the function has to be unbounded at least on one of the halves. Consider this half and divide it in half. Choose the half where the function is unbounded. So we construct a nested infinite sequence of intervals converging to a point, which coincides with the intersection of all the intervals. And f is obviously not locally bounded at this point. $\square$

Corrolary 4. *A complex function $f(z)$ continuous on the boundary of a domain D is bounded on ∂D.*

Proof. Consider a path $p\colon [a, b] \to \partial D$. Then $|f(p(t))|$ is continuous on $[a, b]$, hence it is locally bounded, hence it is bounded. Since ∂D can be covered by images of finitely many pathes this implies boundedness of f over ∂D. $\square$

Theorem 5. *If a function $f(z)$ is complex differentiable in the disk $|z-z_0| \leq R$, then for $|z - z_0| < R$*

$$f(z) = \sum_{k=0}^{\infty}(z - z_0)^k \oint_{z_0}^{R} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}}\, d\zeta, \tag{3.6.4}$$

where the series on the right-hand side absolutely converges for $|z - z_0| < R$.

Proof. Fix a point z such that $|z - z_0| < R$ and consider ζ as a variable. For $|\zeta - z_0| > |z - z_0|$ one has

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{\zeta - z_0}\frac{1}{1 - \frac{z-z_0}{\zeta-z_0}} = \sum_{k=0}^{\infty} \frac{(z - z_0)^k}{(\zeta - z_0)^{k+1}} \tag{3.6.5}$$

On the circle $|\zeta - z_0| = R$ the series on the right-hand side is majorized by a convergent series $\sum\limits_{k=0}^{\infty} \frac{|z-z_0|^k}{R^{k+1}}$ for $r > |z - z_0|$. The function $f(\zeta)$ is bounded on $|\zeta - z_0| = R$ by the Corollary **4**. Therefore after multiplication of (3.6.5) by $f(\zeta)$ all conditions of the Lemma **2** are satisfied. The termwise integration gives:

$$f(z) = \oint_{z_0}^{R} \frac{f(\zeta)}{\zeta - z} d\zeta = \sum_{k=0}^{\infty}(z - z_0)^k \oint_{z_0}^{R} \frac{f(\zeta)\, d\zeta}{(\zeta - z_0)^{k+1}} \tag{3.6.6}$$

$\square$

Analytic functions. A function $f(z)$ of complex variable is called an *analytic function* in a point z_0 if there is a positive ε such that $f(z) = \sum\limits_{k=0}^{\infty} a_k(z - z_0)^k$ for all z from a disk $|z - z_0| \leq \varepsilon$ and the series absolutely converges. Since one can differentiate power series termwise (Theorem **9**), any function which is analytic at z is also complex differentiable at z. The theorem **5** gives a converse. Thus, we get the following:

Corrolary 6. *A function $f(z)$ is analytic at z, iff it is complex differentiable in some neighborhood of z.*

Theorem 7. *If f is analytic at z then f' is analytic at z. If f and g are analytic at z then $f+g$, $f-g$, fg are analytic at z. If f is analytic at z and g is analytic at $f(z)$ then $g(f(z))$ is analytic at z.*

Proof. The termwise differentiation of a power series representing f in a neighborhood of z gives the power series for its derivative. Hence f' is analytic. The differentiability of $f \pm g$, fg and $g(f(z))$ follows from corresponding differentiation rules. □

Lemma 8 (on isolated zeros). *If $f(z)$ is analytic and is not identically equal to 0 in some neighborhood of z_0, then $f(z) \neq 0$ for all $z \neq z_0$ sufficiently close to z_0.*

Proof. Let $f(z) = \sum_{k=0}^{\infty} c_k(z-z_0)^k$ in a neighborhood U of z_0. Let c_m be the first nonzero coefficient. Then $\sum_{k=m}^{\infty} c_k(z-z_0)^{k-m}$ converges in U to a differentiable function $g(z)$ by Theorem **9**. Since $g(z_0) = c_m \neq 0$ and $g(z)$ is continuous at z_0, the inequality $g(z) \neq 0$ holds for all z sufficiently close to z_0. As $f(z) = g(z)(z-z_0)^m$, the same is true for $f(z)$. □

Theorem 9 (Uniqueness Theorem). *If two power series $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ and $\sum_{k=0}^{\infty} b_k(z-z_0)^k$ converge in a neighborhood of z_0 and their sums coincide for some infinite sequence $\{z_k\}_{k=1}^{\infty}$ such that $z_k \neq z_0$ for all k and $\lim_{k\to\infty} z_k = z_0$, then $a_k = b_k$ for all k.*

Proof. Set $c_k = a_k - b_k$. Then $f(z) = \sum_{k=0}^{\infty} c_k(z-z_0)^k$ has a non-isolated zero at z_0. Hence $f(z) = 0$ in a neighborhood of z_0. We get a contradiction by considering function $g(z) = \sum_{k=m}^{\infty} c_k(z-z_0)^{k-m}$, which is nonzero for all z sufficiently close to z_0 (cf. the proof of Lemma on isolated zero), and satisfies equation $f(z) = g(z)(z-z_0)^m$. □

Taylor series. Set $f^{(0)} = f$ and by induction define the $(k+1)$-th derivative $f^{(k+1)}$ of f as the derivative of its k-th derivative $f^{(k)}$. For the first and the second derivatives one prefers notation f' and f''. For example, the k-th derivative of z^n is $n^{\underline{k}}z^{n-k}$. (Recall that $n^{\underline{k}} = n(n-1)\dots(n-k+1)$)

The following series is called the *Taylor series* of a function f at point z_0:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k \tag{3.6.7}$$

The Taylor series is defined for any analytic function, because an analytic function has derivative of any order due to Theorem **7**.

Theorem 10. *If a function f is analytic in the disk $|z-z_0| < r$ then $f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k$ for any z from the disk.*

Proof. By the Theorem **5**, $f(z)$ is presented in the disk by a convergent power series $\sum_{k=0}^{\infty} a_k(z-z_0)^k$. To prove our theorem we prove that

$$a_k = \oint_{z_0}^{R} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}}\, d\zeta = \frac{f^{(k)}(z_0)}{k!}. \tag{3.6.8}$$

Indeed, $a_0 = f(z_0)$ and termwise differentiating of $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ applied n times gives $f^{(n)}(z) = \sum_{k=n}^{\infty} k^{\underline{n}} a_k(z-z_0)^k$. Putting $z = z_0$, one gets $f^{(n)}(z_0) = n^{\underline{n}} a_n = a_n n!$. □

Theorem 11 (Liouville). *If a function f is analytic and bounded on the whole complex plane, then f is constant.*

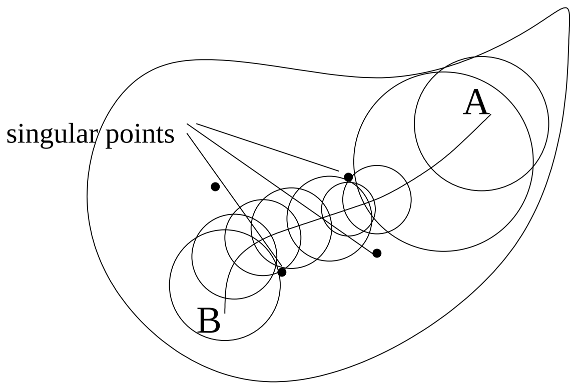
Proof. If f is analytic on the whole plane then $f(z) = \sum_{k=0}^{\infty} a_k z^k$, where a_k is defined by (3.6.8). If $|f(z)| \leq B$ by Lemma on Estimate one gets

$$|a_k| = \left| \oint_{0}^{R} \frac{f(\zeta)}{z^{k+1}}\, d\zeta \right| \leq 4 \cdot 4 \frac{B}{R^{k+1}} \frac{R}{\sqrt{2}} = \frac{C}{R^k} \tag{3.6.9}$$

Consequently a_k for $k > 0$ is infinitesimally small as R tends to infinity. But a_k does not depend on R, hence it is 0. Therefore $f(z) = a_0$. □

Theorem 12 (Fundamental Theorem of Algebra). *Any nonconstant polynomial $P(z)$ has a complex root.*

Proof. If $P(z)$ has no roots the function $f(z) = \frac{1}{P(z)}$ is analytic on the whole plane. Since $\lim_{z\to\infty} f(z) = 0$ the inequality $|f(z)| < 1$ holds for $|z| = R$ if R is sufficiently large. Therefore the estimation (3.6.9) for k-th coefficient of f holds with $B = 1$ for sufficiently large R. Hence the same arguments as in proof of Liouville Theorem show that $f(z)$ is constant. This is a contradiction. □



Analytic continuation.

Lemma 13. *If an analytic function $f(z)$ has finitely many singular points in a domain D and a non isolated zero at a point $z_0 \in D$ then $f(z) = 0$ for all regular $z \in D$*

Proof. For any nonsingular point $z \in D$, we construct a sequence of sufficiently small disks $D_0, D_1, D_2, \dots D_n$ without singular points with the following properties: 1) $z_0 \in D_0 \subset U$, 2) $z \in D_n$, 3) z_k, the center of D_k, belongs to D_{k-1} for all $k > 0$. Then by induction we prove that $f(D_k) = 0$. First step: if z_0 is a non-isolated zero of f, then the Taylor series of f vanishes at z_0 by Uniqueness theorem. But this series represents $f(z)$ on D_0 due to Theorem **10**, since D_0 does not contain singular points. Hence, $f(D_0) = 0$. Suppose we have proved already that $f(D_k) = 0$. Then z_{k+1} is a non-isolated zero of f by the third property of the sequence $\{D_k\}_{k=0}^n$. Consequently, the same arguments as above for $k = 0$ prove that $f(D_{k+1}) = 0$. And finally we get $f(z) = 0$. □

Consider any formula which you know from school about trigonometric functions. For example, $\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$. The above Lemma implies that this formula remains true for complex x and y. Indeed, consider the function $T(x,y) = \tan(x+y) - \frac{\tan x + \tan y}{1 - \tan x \tan y}$. For a fixed x function $T(x,y)$ is analytic and has finitely many singular points in any disk. This function has non-isolated zeros in all real points, hence this function is zero in any disk intersecting the real line. This implies that $T(x,y)$ is zero for all y. The same arguments applied to $T(x,y)$ with fixed y and variable x prove that $T(x,y)$ is zero for all complex x, y.

The same arguments prove the following theorem

Theorem 14. *If some analytic relation between analytic functions holds on a curve Γ, it holds for any $z \in \mathbb{C}$, which can be connected with Γ by a paths avoiding singular points of the functions.*

Lemma 15. $\sin t \geq \frac{2t}{\pi}$ *for* $t \in [0, \pi/2]$

Proof. Let $f(t) = \sin t - \frac{2t}{\pi}$. Then $f'(x) = \cos t - \frac{2}{\pi}$. Set $y = \arccos \frac{2}{\pi}$. Then $f'(x) \geq 0$ for $x \in [0, y]$. Therefore f is nondecreasing on $[0, y]$, and nonnegative, because $f(0) = 0$. On the interval $[y, \pi/2]$ the derivative of f is negative. Hence $f(x)$ is non-increasing and nonnegative, because its value on the end of the interval is 0. □

Lemma 16 (Jordan). *Let $f(z)$ be an analytic function in the upper half-plane such that $\lim\limits_{z \to \infty} f(z) = 0$. Denote by Γ_R the upper half of the circle $|z| = R$. Then for any natural m*

$$\lim_{R \to \infty} \int\limits_{\Gamma_R} f(z) \exp(miz)\, dz = 0 \tag{3.6.10}$$

Proof. Consider the parameterization $z(t) = R\cos t + Ri \sin t$, $t \in [0, \pi]$ of Γ_R. Then the integral (3.6.10) turns into

$$\begin{aligned} \int\limits_0^\pi f(z) \exp(iRm \cos t - Rm \sin t)\, d(R\cos t + Ri \sin t) = \\ = \int\limits_0^\pi Rf(z) \exp(iRm \cos t) \exp(-Rm \sin t)(-\sin t + i \cos t)\, dt. \end{aligned} \tag{3.6.11}$$

If $|f(z)| \leq B$ on Γ_R, then $|f(z)\exp(iRm\cos t)(-\sin t + i\cos t)| \leq B$ on Γ_R. And the module of the integral (3.6.11) can be estimated from above by

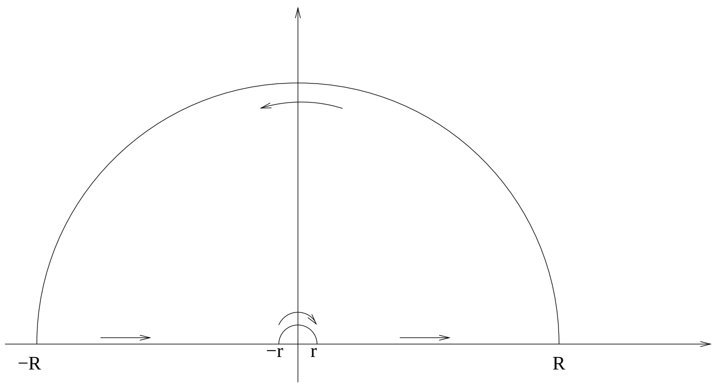
$$BR\int_0^\pi \exp(-Rm\sin t)\,dt.$$

Since $\sin(\pi - t) = \sin t$, the latter integral is equal to $2BR\int_0^{\pi/2} \exp(-Rm\sin t)\,dt$. Since $\sin t \geq \frac{2t}{\pi}$, the latter integral does not exceed

$$2BR\int_0^{\pi/2} \exp(-2Rmt/\pi)\,dt = 2BR\frac{1-\exp(-Rm)}{2Rm} \leq \frac{B}{m} \tag{3.6.12}$$

Since B can be chosen arbitrarily small for sufficiently large R, this proves Lemma. □

Evaluation of $\int_{-\infty}^{+\infty} \frac{\sin x}{x}\,dx = \lim_{N\to\infty}\int_{-N}^{N} \frac{\sin x}{x}\,dx$. Since $\sin x = \operatorname{Im} e^{ix}$ our integral is equal to $\operatorname{Im}\int_{-\infty}^{+\infty} \frac{e^{iz}}{z}\,dz$. Set $\Gamma(r) = \{z \mid |z| = r, \operatorname{Im} z \geq 0\}$. This is a semicircle. Let us orient it counter-clockwise, so that its initial point is r.



Consider the domain $D(R)$ bounded by semicircles $-\Gamma(r)$, $\Gamma(R)$ and intervals $[-R,-r]$ and $[r,R]$, where $r = \frac{1}{R}$ and $R > 1$. The function $\frac{e^{iz}}{z}$ has not singular points inside $D(R)$. Hence $\oint_{\partial D(R)} \frac{e^{iz}}{z} dz = 0$. Hence for any R

$$\int_{-r}^{-R} \frac{e^{iz}}{z}\,dz + \int_r^R \frac{e^{iz}}{z}\,dz = \int_{\Gamma(r)} \frac{e^{iz}}{z}\,dz - \int_{\Gamma(R)} \frac{e^{iz}}{z}\,dz \tag{3.6.13}$$

The second integral on the right-hand side tends to 0 as R tends to infinity due to Jordan's Lemma. The function $\frac{e^{iz}}{z}$ has a simple pole at 0, hence the first integral on the right-hand side of (3.6.13) tends to $\pi i \operatorname{res} \frac{e^{iz}}{z} = \pi i$ due to Remark

1. As a result, the right-hand side of (3.6.13) tends to πi as R tends to infinity. Consequently the left-hand side of (3.6.13) also tends to πi as $R \to \infty$. The imaginary part of left-hand side of (3.6.13) is equal to $\int\limits_{-R}^{R} \frac{\sin x}{x}\,dx - \int\limits_{-r}^{r} \frac{\sin x}{x}\,dx$. The last integral tends to 0 as $r \to 0$, because $|\frac{\sin x}{x}| \leq 1$. Hence $\int\limits_{-R}^{R} \frac{\sin x}{x}\,dx$ tends to π as $R \to \infty$. Finally $\int\limits_{-\infty}^{+\infty} \frac{\sin x}{x}\,dx = \pi$.

Problems.

1. Prove that an even analytic function f, i.eà function that $f(z) = f(-z)$ has Taylor series at 0 consisting only of even powers
2. Prove that analytic function which has Taylor series only with even powers is an even function.
3. If an analytic function $f(z)$ takes real values on $[0, 1]$, then $f(x)$ is real for any real x.
4. $\int\limits_{-\infty}^{+\infty} \dfrac{1}{1+x^4}$
5. $\int\limits_{-\pi}^{+\pi} \dfrac{d\phi}{5 + 3\cos\phi}$
6. Evaluate $\int\limits_{0}^{\infty} \frac{x^2}{(x^2+a^2)^2}\,dx \ (a > 0)$
7. $\int\limits_{-\infty}^{+\infty} \frac{x \sin x}{x^2+4x+20}\,dx$
8. $\int\limits_{0}^{\infty} \frac{\cos ax}{x^2+b^2}\,dx \ (a, b > 0)$

Chapter 4

Differences

4.1 Newton Series

On the contents of the lecture. The formula with 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 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 has 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 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 $P(x) = \sum\limits_{k=0}^{n+1} a_k x^k$ 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 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 **1**, hence $P(x)$ is constant. □

Lemma 3 (Newton 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 $\Delta P(x) = \sum\limits_{k=1}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}}$ by the induction hypothesis. Denote by $Q(x)$ the Newton series $\sum\limits_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!} x^{\underline{k}}$ for

$P(x)$. Then

$$(4.1.2)\quad \Delta Q(x) = \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!}(x+1)^{\underline{k}} - \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!}x^{\underline{k}} = \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!}\Delta x^{\underline{k}} =$$

$$\sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{k!}kx^{\underline{k-1}} = \sum_{k=0}^{\infty} \frac{\Delta^k P(0)}{(k-1)!}x^{\underline{k-1}} = \sum_{k=0}^{\infty} \frac{\Delta^k (\Delta P(0))}{k!}x^{\underline{k}} = \Delta P(x)$$

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

Lemma 4 (Lagrange Formula). *For any sequence $\{y_k\}_{k=0}^n$, the polynomial $L_n(x) = \sum\limits_{k=0}^{n}(-1)^{n-k}\frac{y_k}{k!(n-k)!}\frac{x^{\underline{n+1}}}{x-k}$ has the property $L_n(k) = y_k$ for $0 \le k \le n$.*

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

Lemma 5. *For any function f and for any natural number $m \le n$ one has $f(m) = \sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!} m^{\underline{k}}$.*

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, $L_n(x) = \sum\limits_{k=0}^{\infty} \frac{\Delta^k f(0)}{k!}x^{\underline{k}} = \sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!}x^{\underline{k}}$ by Lemma **3**. Putting $x = m$ in the latter equality, one gets $f(m) = L_n(m) = \sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!}m^{\underline{k}}$. □

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 *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 *infinum* and denoted by $\inf A$.

Theorem 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**. Denote by B the supremum of the set of values of f on $[a, b]$. If f does not take the maximum value, then $f(x) \neq B$ 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**. 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**, which states that locally bounded function is bounded. The same arguments prove that $f(x)$ takes its minimal value on $[a, b]$. □

Theorem 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 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 **1**. $\square$

Lemma 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 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 $(n-1)$ derivative has a root. But the $(n-1)$-th derivative of f' is the n-th derivative of f. $\square$

Theorem 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 $x \in \{0, 1, \ldots n\}$ and any ξ, due to Lemma **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 $f(x) = \sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!} x^{\underline{k}} + C x^{\underline{k+1}}$. The function $F(y) = f(y) - \sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!} x^{\underline{k}} - C y^{\underline{k+1}}$ has roots $0, 1, \ldots, n, x$. Hence its $n+1$-th derivative has a root $\xi \in I$. Since $\sum\limits_{k=0}^{n} \frac{\Delta^k f(0)}{k!} x^{\underline{k}}$ is 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 $\sum\limits_{k=0}^{\infty} \frac{\Delta^k f(0)}{k!} x^{\underline{k}}$ 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

$$(1+x)^y = \sum_{k=0}^{\infty} \frac{x^k y^{\underline{k}}}{k!} \tag{4.1.3}$$

For fixed y and variable x, the formula (4.1.3) represents the famous Newton binomial series. Our proof is not correct. We apply Newton interpolation formula, proved only for polynomials, to an exponential function. But original Newton's 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 binomial series in the 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 of 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

$$(4.1.4) \qquad (1-x^2)^{-\frac{1}{2}} = \sum_{k=0}^{\infty} \frac{(-x^2)^k(-\frac{1}{2})^{\underline{k}}}{k!} = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{2k!!} x^{2k},$$

One gets the series for arcsin by termwise integration of (4.1.4). The result is

$$(4.1.5) \qquad \arcsin x = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{2k!!} \frac{x^{2k+1}}{2k+1}$$

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

Theorem 10. *For any complex z and ζ such that $|z| < 1$, the series $\sum\limits_{k=0}^{\infty} \frac{z^k \zeta^{\underline{k}}}{k!}$ absolutely converges to $(1+z)^\zeta = \exp(\zeta \log(1+z))$.*

Proof. The analytic function $\exp\zeta\log(1+z)$ 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 Taylor series of the function is $\sum\limits_{k=0}^{\infty} \frac{\zeta^{\underline{k}} z^k}{k!}$. □

On the boundary of convergence. Since $(1+z)^\zeta$ has the 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

$$(4.1.6) \qquad \sum_{k=0}^{\infty} \frac{x^{\underline{k}}}{k!} = 2^x$$

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 Leibniz Theorem, one needs only to prove that $\lim\limits_{n\to\infty} \frac{x^{\underline{n}}}{n!} = 0$. Transform this limit into $\lim\limits_{n\to\infty} \frac{x}{n} \prod\limits_{k=1}^{n-1} (\frac{x}{k} - 1)$. The product $\prod\limits_{k=1}^{n-1} (\frac{x}{k} - 1)$ 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 majorated 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 $(1+x)^n = \sum\limits_{k=0}^{n} \frac{n^{\underline{k}} x^k}{k!}$. Because $(a+b)^n = a^n(1+\frac{b}{a})^n$, one gets the following famous formula

$$(a+b)^n = \sum_{k=0}^{n+1} \frac{n^{\underline{k}}}{k!} a^k b^{n-k}. \tag{4.1.7}$$

This is the formula that is usually called Newton 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* $\mathbf{S}^a$ by formula $\mathbf{S}^a f(x) = f(x+a)$ for any function. 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 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 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 to all usual numeric laws, and all identities which holds for numbers extends 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 the for any polynomial P the degree of $O(P)$ is less than the degree of P. For example, operation of difference $\Delta = \mathbf{S} - 1$ and operation $\mathbf{D}$ of differentiation $\mathbf{D}f(x) = f'(x)$ are decreasing. For a decreasing operation O, any power series $\sum\limits_{k=0}^{\infty} a_k O^k$ 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 $(1+\Delta)^y = \sum\limits_{k=0}^{\infty} \frac{\Delta^k y^{\underline{k}}}{k!}$ represents $\mathbf{S}^y$. And the equality $\mathbf{S}^y = \sum\limits_{k=0}^{\infty} \frac{\Delta^k y^{\underline{k}}}{k!}$, which is in fact the Newton Interpolation Formula for polynomials, 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, $\mathbf{S}^x = \sum\limits_{k=0}^{n} \frac{n^{\underline{k}}\Delta_n^k}{k!} = \sum\limits_{k=0}^{\infty} \frac{n^{\underline{k}}}{n^k}\frac{(n\Delta_n)^k}{k!}$. Now we follow Euler "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 $\mathbf{S}^x = \sum\limits_{k=0}^{\infty} \frac{x^k \mathbf{D}^k}{k!}$. Our proof is copied from the Euler proof in his Introductio of $\lim\limits_{n\to\infty}(1+\frac{x}{n})^n = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$. This substitution of infinity means passing to 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 expansion of functions into power series in Euler's times. Euler also applied it to get power expansions for trigonometric functions.

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. The 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. If p is prime, then $\frac{p^{\underline{k}}}{k!}$ is divisible by p.
4. $\frac{n^{\underline{k}}}{k!} = \frac{n^{\underline{n-k}}}{(n-k)!}$
5. Deduce plain binomial theorem from multiplication of series for exponenta.
6. Deduce formula for Catalan numbers.
7. Prove that $\Delta^k x^{\underline{n}} x^{\underline{m}} = 0$ for $k < n$.
8. Prove that $\sum_{k=0}^{n} (-1)^k \frac{n^{\underline{k}}}{k!} = 0$
9. Get a differential equation for 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. A sequence $\{a_k\}$ such that $\Delta^2 a_k \geq 0$ satisfies inequality $\max\{a_1, \dots a_n\} \geq a_k$ for any k between 1 and n
12. $\sum_{k=0}^{\infty} (-1)^k \frac{x^{\underline{2k}}}{2k!} = 2^{x/2} \cos \frac{x\pi}{4}$
13. $\sum_{k=0}^{\infty} (-1)^k \frac{x^{\underline{2k+1}}}{(2k+1)!} = 2^{x/2} \sin \frac{x\pi}{4}$
14. $\Delta^n 0^p$ 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 the telescoping of powers. These formulas involve a remarkable 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:

$$\psi_m(x) = \sum_{k=0}^{m} B_{m-k} \frac{(m-1)^{\underline{m-k-1}}}{(m-k)!} x^k = \sum_{k=0}^{m} B_{m-k} \frac{(m-1)!}{k!(m-k)!} x^k = \sum_{k=0}^{m} B_{m-k} \frac{(m-1)^{\underline{k-1}}}{k!} x^k \tag{4.2.2}$$

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

$$\phi_{m+1}(x) = \sum_{k=0}^{\infty} \frac{\Delta^k 0^m}{(k+1)!} x^{\underline{k+1}} \tag{4.2.3}$$

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

Polynomials $\phi_m(x)$, as follows from Lemma **2**, are characterized by two conditions:

$$\Delta\phi_m(x) = x^{m-1}, \qquad \phi_m(1) = 0 \tag{4.2.4}$$

Lemma 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 the Lemma **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 $B_m(x) = m\phi_m(x) + B_m$. 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:

$$\Delta B_m(x) = mx^{m-1} \quad (m > 0) \tag{4.2.5}$$

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

$$B_m(0) = B_m(1) = B_m \quad \text{for } (m > 1) \text{ and } \quad B_1(0) = -B_1(1) = B_1 \tag{4.2.6}$$

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

$$B'_m(x) = mB_{m-1}(x) \tag{4.2.7}$$

Indeed, $B'_m(x) = m\phi'_m(x) = m((m-1)\phi_{m-1}(x) + \phi'_m(0)) = mB_{m-1}(x)$, by Lemma **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 Taylor formula gives the following representation of the Bernoulli polynomial:

$$B_m(x) = \sum_{k=0}^{m} \frac{m^{\underline{k}}B_{m-k}}{k!}x^k \tag{4.2.8}$$

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

$$\int_0^1 B_m(x)\,dx = 0 \quad (m > 0) \tag{4.2.9}$$

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 allows us to evaluate Bernoulli polynomials recursively. Thus, $B_1(x)$ has 1 as leading coefficient and zero integral on $[0,1]$, this allows us 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 (4.2.9) 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 (4.2.9) and so on. Thus we obtain the following theorem:

Theorem 2 (characterization). *If a sequence of polynomials $\{P_n(x)\}$ satisfies the following conditions:*

- $P_0(x) = 1$,
- $\int\limits_0^1 P_n(x)\,dx = 0$ *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 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 conditions of the Theorem **2**. Indeed, $T_0 = B_0 = 1$, $\int\limits_0^1 T_n(x)\,dx = (-1)^n \int\limits_1^0 B_n(x)\,dx = 0$ and

$$T_n(x)' = (-1)^n B_n'(1-x) = (-1)^n n B_{n-1}(1-x)(1-x)' = (-1)^{n+1} n B_{n-1}(x) = nT_{n-1}(x)$$

□

Lemma 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 **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 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 $B_n'(x)$ 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 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 the Lemma **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 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** on roots and 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 0 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. $\square$

Lemma 6 (Lagrange Formula). *If f is a differentiable function on $[a,b]$, then there is a $\xi \in (a,b)$, such that*

$$f(b) = f(a) + f'(\xi)\frac{f(b)-f(a)}{b-a} \tag{4.2.10}$$

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 (4.2.10) gives the equality. $\square$

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

$$B(x,t) = \sum_{k=0}^{\infty} B_k(x)\frac{t^k}{k!} \tag{4.2.11}$$

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 $\sum\limits_{k=0}^{\infty} kB_{k-1}(x)\frac{t^k}{k!} = tB(x,t)$. Consequently $(\log B(x,t))'_x = \frac{B'_x(x,t)}{B(x,t)} = t$ and $\log B(x,t) = xt + c(t)$, 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 $B(0,t) = k(t) = \sum\limits_{k=0}^{\infty} B_k\frac{t^k}{k!}$. 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 $\sum\limits_{k=0}^{\infty} \Delta B_k(x)\frac{t^k}{k!} = \sum\limits_{k=0}^{\infty} kB_{k-1}(x)\frac{t^k}{k!} = tB(x,t)$. Comparing these expressions we get explicit formulas for the generating functions of Bernoulli numbers:

$$k(t) = \frac{t}{\exp t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!}t^k\,, \tag{4.2.12}$$

and Bernoulli polynomials:

$$B(x,t) = \sum_{k=+}^{0-1} B_k(x)\frac{t^k}{k!} = \frac{t\exp(tx)}{\exp t - 1} \tag{4.2.13}$$

From (4.2.11) one gets $t = (\exp t - 1)\sum\limits_{k=0}^{\infty} B_k\frac{t^k}{k!}$. Substituting $\exp t - 1 = \sum\limits_{k=1}^{\infty}\frac{t^k}{k!}$ in this equality, by the Uniqueness Theorem, one gets the equalities for the coefficients of the power series

$$\sum_{k=1}^{n} \frac{B_{n-k}}{(n-k)!k!} = 0 \quad \text{for } n > 1 \tag{4.2.14}$$

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

$$B_n = \sum_{k=0}^{n} \frac{B_k n^{\underline{k}}}{k!} \quad \text{for } n > 1 \tag{4.2.15}$$

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

$$\phi_m(1+x) = \frac{B_m(x+1)}{m} - \frac{B_m}{m} = \sum_{k=0}^{m} B_{m-k} \frac{(m-1)^{\underline{k-1}}}{k!} x^k = \psi_m(x) \tag{4.2.16}$$

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 (4.2.16) for positive degrees.

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

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

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

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 (4.2.15)

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

Problems.

1. Evaluate $\int_0^1 B_n(x) \sin 2\pi x$.
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 $|B_n(x)| \leq |B_n|$ for even n.
5. Prove the inequality $|B_n(x)| \leq \frac{n}{4}|B_{n-1}|$ 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 $B_m(nx) = n^{m-1} \sum\limits_{k=0}^{n-1} B_m\left(x + \frac{k}{n}\right)$

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 $B_n = \sum\limits_{k=0}^{\infty} (-1)^k \frac{\Delta^k 0^n}{k+1}$

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 18 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

$$f(x+1) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!}. \tag{4.3.1}$$

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

$$\Delta = \exp \mathbf{D} - \mathbf{1}. \tag{4.3.2}$$

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 (4.3.2), 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

$$\Delta^{-1} = \mathbf{D}^{-1} \frac{\mathbf{D}}{\exp \mathbf{D} - \mathbf{1}} = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathbf{D}^{k-1} = \mathbf{D}^{-1} - \frac{1}{2}\mathbf{1} + \sum_{k=1}^{\infty} \frac{B_{2k}}{2k!} \mathbf{D}^{2k-1} \tag{4.3.3}$$

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:

$$\sum_{k=1}^{n-1} f(k) = \int_1^n f(x)\,dx - \frac{f(n) - f(1)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} (f^{(2k-1)}(n) - f^{(2k-1)}(1)) \tag{4.3.4}$$

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 the absolute value $\frac{k!}{10^{k+2}}$. Hence the module of k-th term of the summation formula does not exceed $\frac{B_k}{k10^{k+2}}$. For an accuracy of 18 digit places it is sufficient to sum up the first 14 terms of the series, only 8 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 $\frac{B_{16}}{16 \cdot 10^{16}}$. 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}
$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{30}$	$\frac{1}{42}$	$-\frac{1}{30}$	$\frac{5}{66}$	$-\frac{691}{2730}$	$\frac{7}{6}$	$-\frac{3617}{510}$	$\frac{43867}{798}$	$-\frac{174611}{330}$

We see from the table 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 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 $\Sigma_m^n a_k$ the sum $\frac{1}{2}a_m + \sum_{k=m-1}^{n-1} a_k + \frac{1}{2}a_n$.

Lemma 2. *For any natural p, q and a function $f(x)$ one has*

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

Proof. Applying Lemma **1** to $f(x+k)$ one gets $\frac{1}{2}(f(k+1)+f(k)) = \int_0^1 f(x+k)\,dx + \int_0^1 f'(x+k)B_1(x)\,dx = \int_k^{k+1} f(x)\,dx + \int_k^{k+1} f'(x)B_1\{x\}\,dx$. Summing up these equalities for k from p to q, one proves the Lemma. □

Lemma 3. *For $m > 0$ and a function f one has*

$$\int_p^q f(x)B_m\{x\}\,dx = \frac{B_{m+1}}{m+1}(f(q)-f(p)) - \int_p^q f'(x)B_{m+1}\{x\}\,dx \tag{4.3.5}$$

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 (**3**) is obtained by a simple integration by parts. □

Theorem 4. *For any function f and natural numbers n and m one has:*

$$\Sigma_1^n f(k) = \int_1^n f(x)\,dx + \sum_{k=1}^{m-1} \frac{B_{k+1}}{(k+1)!}\left(f^{(k)}(n) - f^{(k)}(1)\right) + \frac{(-1)^{m+1}}{m!}\int_1^n f^{(m)}(x)B_m\{x\}\,dx \tag{4.3.6}$$

Proof. The proof is by induction on m. For $m = 1$, formula (4.3.6) is just given by Lemma **2**. Suppose (4.3.6) 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 **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 $m > 0$. $\square$

Estimation of remainder. For $m = \infty$, (4.3.6) turns into (4.3.4). Denote

$$R_m = \frac{(-1)^{m+1}}{m!}\int_1^n f^{(m)}(x)B_m\{x\}\,dx \tag{4.3.7}$$

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

Lemma 5. *$R_{2m} = R_{2m+1}$ for any $m > 1$*

Proof. Because $B_{2m+1} = 0$, the only thing which changes in (4.3.6) when one passes from $2m$ to $2m+1$ is the remainder. Hence its value does not change either. $\square$

Lemma 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 $-\int_0^{0.5} f(1-x)B_{2m+1}(x)\,dx$:

$$\int_0^1 f(x)B_{2m+1}(x)\,dx = \int_0^{0.5} f(x)B_{2m+1}(x)\,dx + \int_{0.5}^1 f(x)B_{2m+1}(x)\,dx = \int_0^{0.5}(f(x) - f(1-x))B_{2m+1}(x)\,dx \tag{4.3.8}$$

$B_{2m+1}(x)$ 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 in (4.3.8) has constant sign. Consequently the integral itself has the same sign as the integrand has. $\square$

Lemma 7. *If $f^{(2m+1)}(x)$ and $f^{(2m+3)}(x)$ are comonotone for $x \geq 1$ then*

$$R_{2m} = \theta_m \frac{B_{2m+2}}{(2m+2)!}(f^{(2m+1)}(n) - f^{(2m+1)}(1)) \quad 0 \leq \theta_m \leq 1 \tag{4.3.9}$$

Proof. The signs of R_{2m+1} and R_{2m+3} are opposite. Indeed, by Lemma **5** $\operatorname{sgn} B_{2m} = -\operatorname{sgn} B_{2m+2}$ and $\operatorname{sgn}(f^{(2m+1)}(n) - f^{(2m+1)}(1)) = \operatorname{sgn}(f^{(2m+3)}(n) - f^{(2m+3)}(1)$, owing to comonotonity condition. Hence $\operatorname{sgn} R_{2m+1} = -\operatorname{sgn} R_{2m+3}$ by Lemma **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 **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 8. *If $f^{(k)}$ and $f^{(k+2)}$ are comonotone for any $k > 1$, then*

$$\left| \Sigma_1^n f(k) - \int_1^n f(x)\,dx - \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!}\left(f^{(2k-1)}(n) - f^{(2k-1)}(1)\right) \right| \leq \left| \frac{B_{2m+2}}{(2m+2)!}\left(f^{(2m+1)}(n) - f^{(2m+1)}\right) \right| \tag{4.3.10}$$

Hence the value of the error which gives the summation formula (4.3.4) 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 9. *Suppose that $\int_1^\infty |f^{(k)}(x)|\,dx < \infty$, $\lim_{x\to\infty} f^{(k)}(x) = 0$ 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 $\theta_m \in [0, 1]$ holds*

(4.3.11)

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

Lemma 10. *Under the condition of the theorem, for any $m \geq K$,*

$$\frac{(-1)^m}{m!} \int_p^\infty f^{(m)}(x) B_m\{x\}\,dx = -\theta_m \frac{B_{2m+2}}{(2m+2)!} f^{(2m+1)}(p) \tag{4.3.12}$$

Proof. By Lemma **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 (4.3.12), pass to the limit as q tends to infinity. □

Proof. To get (4.3.11) we change the form of the remainder R_K for (4.3.6). 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 equality 4.3.5 to the interval $[n,\infty)$, one gets

$$-\frac{(-1)^{k+1}B_k}{k!}\int\limits_n^\infty B_k\{x\}f^{(k)}(x)\,dx = \frac{(-1)^{k+1}B_{k+1}}{(k+1)!}f^{(k)}(n) - \frac{(-1)^{k+2}B_{k+1}}{(k+1)!}\int\limits_n^\infty B_{k+1}\{x\}f^{(k+1)}(x)\,dx \tag{4.3.13}$$

Iterating this formula one gets
(4.3.14)

$$R_K = \int\limits_1^\infty B_K\{x\}f^{(K)}\,dx + \sum_{k=K}^{m}\frac{B_{k+1}}{(k+1)!}f^{(k)}(n) + \frac{(-1)^m}{m!}\int\limits_n^\infty B_m\{x\}f^{(m)}(x)\,dx$$

Here we take into account the equalities $(-1)^k B_k = B_k$ and $(-1)^{m+2} = (-1)^m$. Now we substitute this expression of R_K into (4.3.6) and set

$$C = (-1)^{K+1}\int\limits_1^\infty B_K\{x\}f^{(K)}(x)\,dx - \frac{f(1)}{2} - \sum_{k=1}^{K-1}\frac{B_{k+1}}{(k+1)!}f^{(k)}(1). \tag{4.3.15}$$

$\square$

Stirling formula. The logarithm satisfies all condition of Theorem **9** with $K = 2$. Its k-th derivative at n is equal to $\frac{(-1)^{k+1}(k-1)!}{n^k}$. Thus (4.3.11) for the logarithm turns into
(4.3.16)

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

By (4.3.15), the constant is $\sigma = \int\limits_1^\infty \frac{B_2\{x\}}{x^2}\,dx - \frac{B_2}{2}$. 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:

$$n! = \sqrt{2\pi n}\; n^n e^{-n+\frac{\Theta}{12n}} \tag{4.3.17}$$

Problems.

1. Write the Euler-Maclaurin series telescoping $\frac{1}{x}$
2. Prove the uniqueness of the constant in Euler-Maclaurin formula.
3. Calculate 10 digit places of $\sum\limits_{k=1}^{\infty}\frac{1}{n^3}$
4. Calculate 8 digit places of $\sum\limits_{k=1}^{1000000}\frac{1}{k}$
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 are expressed in terms of Gamma-function. In particular it will help us to 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 other 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 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

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

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

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

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

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

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

(4.4.3)

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

The function

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

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

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

Theorem 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, $\operatorname{var}_f[1,n] = \sum\limits_{k=1}^{n} \frac{1}{k} < \ln n + \gamma$. 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 $\lim\limits_{n\to\infty} \frac{\ln n}{n} = 0$, we get $v = 0$. Hence $f - \digamma$ is constant, and it is zero if $f(1) = \digamma(1)$. □

Lemma 3. $\digamma' = \mathbb{F}$.

Proof. To prove that $\digamma'(x) = \mathbb{F}(x)$, consider $F(x) = \int\limits_1^x \mathbb{F}(t)\,dt$. This function is monotone, because $F'(x) = \mathbb{F}(x) \geq 0$. Further $(\Delta F)' = \Delta F' = \Delta\mathbb{F}(x) = -\frac{1}{(1+x)^2}$. It follows $\Delta F = \frac{1}{1+x} + c$, where c is a constant. By the Theorem **2** it follows $F(x+1) - cx - \gamma = \digamma(x)$. Hence $\digamma(x)' = F'(x+1) + c = \mathbb{F}(x)$. This proves that $\digamma'$ is differentiable and has finite variation. As $\Delta\digamma(x) = \frac{1}{1+x}$ it follows $\Delta\digamma'(x) = -\frac{1}{(1+x)^2}$. We get that $\digamma'(x) = \mathbb{F}(x)$ by theorem **1**. □

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

$$\left(\ln\left(1 + \frac{x+1}{k}\right) - \frac{x+1}{k}\right) - \left(\ln\left(1+\frac{x}{k}\right) - \frac{x}{k}\right) =$$
$$\left(\ln(x+k+1) - \ln k - \frac{x+1}{k}\right) - \left(\ln(x+k) - \ln k - \frac{x}{k}\right) = \Delta\ln(x+k) - \frac{1}{k}$$

Hence

$$\begin{aligned}(4.4.5)\quad \Delta F(x) = -\Delta \ln x - \sum_{k=1}^{\infty}\left(\Delta \ln(x+k) - \frac{1}{k}\right) = \\ \lim_{n\to\infty}\left(-\Delta \ln x - \sum_{k=1}^{n-1}\left(\Delta \ln(x+k) - \frac{1}{k}\right)\right) = \\ \lim_{n\to\infty}\left(\ln x - \ln(n+x) + \sum_{k=1}^{n-1}\frac{1}{k}\right) = \\ \ln x + \lim_{n\to\infty}(\ln(n) - \ln(n+x)) + \lim_{n\to\infty}\left(\sum_{k=1}^{n-1}\frac{1}{k} - \ln n\right) = \ln x + \gamma\end{aligned}$$

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

$$(4.4.6)\qquad \Theta(x) = -\gamma x - \ln x - \sum_{k=1}^{\infty}\left(\ln\left(1+\frac{x}{k}\right) - \frac{x}{k}\right)$$

Theorem 4. *The series* (4.4.6) *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 $\frac{x}{1+x} \leq \ln(1+x) \leq x$ implies

$$(4.4.7)\qquad |\ln(1+x) - x| \leq \left|\frac{x}{1+x} - x\right| = \left|\frac{x^2}{1+x}\right|.$$

Denote by ε the distance from x to the closest negative integer. Then due to (4.4.7), the series $\sum\limits_{k=1}^{\infty} \ln\left(\left(1+\frac{y}{k}\right) - \frac{y}{k}\right)$ is termwise majorized by the convergent series $\sum\limits_{k=1}^{\infty} \frac{x^2}{\varepsilon k^2}$. This proves the absolute convergence of (4.4.6).

Since $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n-1}(\ln(1+\frac{1}{k}) - \frac{1}{k}) = \lim\limits_{n\to\infty}(\ln n - \sum\limits_{k=1}^{n-1}\frac{1}{k}) = -\gamma$, 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:

$$(4.4.8)\qquad f(\theta x + \overline{\theta} y) \leq \theta f(x) + \overline{\theta} f(y) \qquad \forall\, \theta \in [0,1]$$

Immediately from the definition it follows that

Lemma 5. *Any linear function $ax + b$ is convex.*

Lemma 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 7. *If $f(p) = f(q) = 0$ and f is convex, then $f(x) \geq 0$ 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 $f(q) \leq f(x)\theta + f(p)\overline{\theta} = f(x)$, it follows $f(x) \geq f(q) = 0$. For $x < p$ one has $p = x\theta + q\overline{\theta}$ for $\theta = \frac{q-p}{q-x}$. Hence $0 = f(p) \leq f(x)\theta + f(q)\overline{\theta} = f(x)$. □

Lemma 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 $t^{\underline{2}} = t(1 - t) < 0$ we get inequality $F(t) \leq F(0)\overline{\theta} + tF(1)$. Since $F(\theta) = f(x\overline{\theta} + y\theta)$ this is just the inequality of convexity. □

Lemma 9. *If f is convex, then $0 \leq f(a) + \theta\Delta f(a) - f(a+\theta) \leq \Delta^2 f(a-1)$ for any a and any $\theta \in [0, 1]$*

Proof. Since $a + \theta = \overline{\theta}a + \theta(a + 1)$ we get $f(a + \theta) \leq f(a)\overline{\theta} + f(a + 1)\theta = f(a) + \theta\Delta f(a)$. On the other hand, the convex function $f(a + x) - f(a) - x\Delta f(a-1)$ has roots -1 and 0. By Lemma **7** it is nonnegative for $x > 0$. Hence $f(a + \theta) \geq f(a) + \theta\Delta f(a - 1)$. It follows that $f(a) + \theta\Delta f(a) - f(a + \theta) \geq f(a) + \theta\Delta f(a) - f(a) - \theta\Delta f(a - 1) = \theta\Delta^2 f(a - 1)$ □

Theorem 10. *$\Theta(x)$ is the unique convex function that telescopes $\ln x$ and $\Theta(1) = 1$.*

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

$$\begin{aligned} (4.4.9) \quad \phi(c\theta + d\overline{\theta}) - \theta\phi(c) - \overline{\theta}\phi(d) &= (f(c\theta + d\overline{\theta}) - \theta f(c) - \overline{\theta} f(d)) - \\ &(\Theta(c\theta + d\overline{\theta}) - \theta\Theta(c) - \overline{\theta}\Theta(d)) \leq \theta\Theta(c) + \overline{\theta}\Theta(d) - \Theta(c\theta + d\overline{\theta}). \end{aligned}$$

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

$$\phi(c\theta + d\overline{\theta}) \leq \theta\phi(c) + \overline{\theta}\phi(d) + \varepsilon. \tag{4.4.10}$$

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 **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 $\theta L(c) + \overline{\theta}L(d) - L(c\theta + d\overline{\theta}) = 0$ less than by ε. The inequality (4.4.10) 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 11. $\Theta''(1 + x) = ⅎ(x)$

Proof. The function $F(x) = \int\limits_1^x \digamma(t)\,dt$ is convex because its second derivative is $\underline{\digamma}$. The difference of $F' = \digamma$ is $\frac{1}{1+x}$. Hence $\Delta F(x) = \ln(x+1) + c$, where c is some constant. It follows $F(x-1) - cx + c = \Theta(x)$. Hence Θ is twice differentiable and its second derivative is $\underline{\digamma}$. $\square$

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 (4.4.6) gives a representation of Gamma function in so-called *canonical Weierstrass form*:

$$\Gamma(x) = \frac{e^{-\gamma x}}{x} \prod_{k=1}^{\infty} \left(1 + \frac{x}{k}\right)^{-1} e^{\frac{x}{k}} \tag{4.4.11}$$

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

$$\Gamma(x+1) = x\Gamma(x) \tag{4.4.12}$$

Since $\Theta(1) = 0$, according to (**4**), one proves by induction that $\Gamma(n) = (n-1)!$ using (4.4.12).

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

Theorem 12 (characterization). *$\Gamma(x)$ is the unique logarithmically convex function defined for $x > 0$, which satisfies equation* (4.4.12) *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 conditions of Theorem **4**. Hence, $\ln f(x) = \Theta(x)$ and $f(x) = \Gamma(x)$. $\square$

Theorem 13 (Euler). *For any $x \geq 0$ one has $\Gamma(x) = \int\limits_0^\infty t^{x-1}e^{-t}\,dt$*

Let us check that the integral satisfies all conditions of Theorem **12**. For $x = 1$ the integral gives $\int\limits_0^\infty e^{-t}\,dt = -e^{-t}\,|_0^\infty = 1$. The integration by parts $\int\limits_0^\infty t^x e^{-t}\,dt = -\int\limits_0^\infty t^x\,de^{-t} = -t^x e^{-t}\,|_0^\infty + \int\limits_0^\infty e^{-t} x t^{x-1}\,dx$ proves that it satisfies the gamma-equation (4.4.12). It remains to prove logarithmic convexity of the integral.

Lemma 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 $F(t) \le L(t)$ is already proved for fractions with denominators of degree $\le n$. 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 $F(r^\pm) \le L(r^\pm)$. Since $r = \frac{r^+ + r^-}{2}$, by mean inequality one has $F(r) \le \frac{f(r^+)+f(r^-)}{2} \le \frac{L(r^+)+L(r^-)}{2} = L(\frac{r^+ + r^-}{2}) = L(r)$.

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 $|L(p)-L(t)| \le |p-t||f(y)-f(x)| < |F(t)-L(t)|$. Therefore $F(p) \le L(p) < F(t)$. 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 15 (Cauchy-Bunyakovski-Schwarz).

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \le \int_a^b f^2(x)\,dx \int_a^b g^2(x)\,dx \tag{4.4.13}$$

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

$$t^2 \int_a^b g^2(x)\,dx + 2t \int_a^b f(x)g(x)\,dx + \int_a^b f^2(x)\,dx = 0 \tag{4.4.14}$$

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 Euler integral. The integral is obviously an increasing function, hence by the mean criterion it is sufficient to prove the following inequality:

$$\left(\int_0^\infty t^{\frac{x+y}{2}-1} e^{-t}\,dt\right)^2 \le \int_0^\infty t^{x-1}e^{-t}\,dt \int_0^\infty t^{y-1}e^{-t}\,dt \tag{4.4.15}$$

This inequality turns into the Cauchy-Bunyakovski-Schwarz inequality (4.4.13) 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

$$\prod_{k=1}^\infty \{(1 - x/n)\exp(x/n)\} = \frac{-e^{\gamma x}}{x\Gamma(-x)} \qquad \prod_{k=1}^\infty \{(1 + x/n)\exp(-x/n)\} = \frac{e^{-\gamma x}}{x\Gamma(x)} \tag{4.4.16}$$

One can evaluate a lot of products by splitting them into parts which have this canonical form (4.4.16). For example, consider the product $\prod_{k=1}^\infty \frac{(}{1} - \frac{x^2}{k^2})$. Division by n^2 transforms it into $\prod\limits_{k=1}^{\infty}(1-\frac{1}{2n})^{-1}(1+\frac{1}{2n})^{-1}$. Introducing multipliers $e^{\frac{1}{2n}}$ and $e^{-\frac{1}{2n}}$, one gets a canonical form

$$\prod_{k=1}^{\infty}\left\{\left(1-\frac{1}{2n}\right)e^{\frac{1}{2n}}\right\}^{-1}\prod_{k=1}^{\infty}\left\{\left(1+\frac{1}{2n}\right)e^{-\frac{1}{2n}}\right\}^{-1} \tag{4.4.17}$$

Now we can apply (4.4.16) for $x=\frac{1}{2}$. The first product of (4.4.17) 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 $\prod\limits_{k=1}^{\infty}\left(1+\frac{x}{n}\right)\left(1+\frac{2x}{n}\right)\left(1-\frac{3x}{n}\right)$
2. Evaluate the product $\prod\limits_{k=1}^{\infty}\frac{k(5+k)}{(3+k)(2+k)}$
3. The sum of logarithmically convex functions is logarithmically convex.
4. Prove $\Gamma(x)=\lim\limits_{n\to\infty}\frac{n!n^x x^{\underline{-n}}}{x}$
5. Prove $\prod\limits_{k=1}^{\infty}\frac{k}{x+k}\left(\frac{k+1}{k}\right)^x=\Gamma(x+1)$
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 the product $\prod(x - x_k)$ where x_k are its roots. By analogy, Euler conjectured that $\sin x$ can be expanded into the product

$$\sin x = \prod_{k=-\infty}^{\infty} (x - k\pi). \tag{4.5.1}$$

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

$$\prod_{k=-\infty}^{\infty} \left(1 - \frac{x}{k\pi}\right) = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right). \tag{4.5.2}$$

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 Wallis product in (4.5.2) multiplied by $x = \frac{\pi}{2}$. Hence the value of (4.5.2) is 1, which coincides with $\sin \frac{\pi}{2}$. Thus it is natural to expect that $\sin x$ coincides with the product (4.5.2).

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

$$\operatorname{ctg}(x) = \sum_{k=-\infty}^{\infty} \frac{1}{x - k\pi} = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2x}{x^2 - k^2\pi^2} \tag{4.5.3}$$

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

$$\cot z = \frac{1}{z} + \sum_{k=1}^{\infty}(-1)^k \frac{4^k B_{2k}}{(2k)!} z^{2k-1} \tag{4.5.4}$$

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

$$\frac{1}{z^2 - n^2\pi^2} = -\frac{1}{n^2\pi^2}\frac{1}{1 - \frac{z^2}{n^2\pi^2}} = -\sum_{k=0}^{\infty}\frac{z^{2k}}{(n\pi)^{2k+2}} \tag{4.5.5}$$

into (4.5.3) and changing the order of summation, one gets:

$$\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\frac{z^{2k}}{(n\pi)^{2k+2}} = \sum_{k=0}^{\infty}\frac{z^{2k}}{\pi^{2k+2}}\sum_{n=1}^{\infty}\frac{1}{n^{2k+2}} \tag{4.5.6}$$

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

Lemma 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* $\frac{1}{\pi^{n+1}}\sum\limits_{k=1}^{\infty}\frac{1}{k^{n+1}}$ *for any odd* n.

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

$$\boxed{(-1)^n\frac{4^n B_{2n}}{2n!} = -\frac{1}{\pi^{2n}}\sum_{k=1}^{\infty}\frac{1}{k^{2n}}} \tag{4.5.7}$$

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

Exploring the cotangent.

Lemma 2. $|\cot z| \leq 4$ *provided* $|\operatorname{Im} z| \geq 1$

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

$$\cot z = i\frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} = i\frac{e^{2iz} + 1}{e^{2iz} - 1} \tag{4.5.8}$$

is less than 4. For $y \geq 1$ the same arguments work for the representation of $\cot z$ as $i\frac{1+e^{-2iz}}{1-e^{-2iz}}$. □

Lemma 3. $|\cot(\pi/2 + iy)| \leq 4$ *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 numerator of this fraction does not exceed $e - e^{-1}$ for $t \in [-1, 1]$ and denominator is greater than 1. This proves the inequality for $y \in [-1, 1]$. For others 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. *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 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 form $\{k\pi\}$. Since $\lim\limits_{z\to 0} z\cot z = \lim\limits_{z\to 0} \frac{z\cos z}{\sin z} = \lim\limits_{z\to 0} \frac{z}{\sin z} = \frac{1}{\sin' 0} = 1$. We get that 0 is a simple pole of $\cot z$ with residue 1 and others poles have the same residue because of periodicity of $\cot z$. □

Lemma 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*

(4.5.9)

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

.

Proof. In our situation every singular points of $f(z)\cot z$ in D are either π-integer or 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 the Residue Theory. □

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

Lemma 6. $\operatorname{ctg}(z+\pi) = \operatorname{ctg}(z)$ *for any* z

Proof. $\operatorname{ctg}(z+\pi) = \lim\limits_{n\to\infty} \sum\limits_{k=-n}^{n} \frac{1}{z+\pi-k\pi} = \lim\limits_{n\to\infty} \sum\limits_{k=-n-1}^{n-1} \frac{1}{z+k\pi} = \lim\limits_{n\to\infty} \frac{1}{z-(n+1)\pi} + \lim\limits_{n\to\infty} \frac{1}{z-n\pi} + \lim\limits_{n\to\infty} \sum\limits_{k=-(n-1)}^{(n-1)} \frac{1}{z+\pi-k\pi} = 0 + 0 + \operatorname{ctg}(z)$ □

Lemma 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 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 $|z^2 - k^2\pi^2| \ge k^2\pi^2$. Hence $|\operatorname{ctg}(z)|$ is term-wise majorized by $\frac{1}{\pi} + \sum\limits_{k=1}^{\infty} \frac{1}{k^2\pi^2} < 2$. □

Lemma 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 $|C(z)| \le \frac{2}{\pi} + \sum\limits_{k=1}^{\infty} \frac{1}{k^2} \le 1 + 2 = 3$. □

Lemma 9. *For any $z \neq k\pi$ and domain D which contains z and whose boundary does not contain π-integers, one has*

$$\oint_{\partial D} \frac{\operatorname{ctg}(\zeta)}{\zeta - z} d\zeta = 2\pi i \operatorname{ctg}(z) + 2\pi i \sum_{k=-\infty}^{\infty} \frac{1}{k\pi - z}[k\pi \in D] \tag{4.5.10}$$

Proof. As was proved in Lecture 3.6, series $\sum\limits_{k=-\infty}^{\infty} \frac{1}{(\zeta - z)(\zeta - k\pi)}$ admits term-wise 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

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

□

Lemma 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 (4.5.10) 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 (4.5.10) allows us to evaluate the residues. □

Theorem 11. $\cot z = \frac{1}{z} + \sum\limits_{k=1}^{\infty} \frac{2z}{z^2 - k^2\pi^2}$

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 $R(z) = \frac{1}{2\pi i} \oint\limits_{\partial D} \frac{R(\zeta)}{\zeta - z} d\zeta$ 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

(4.5.11)
$$R(z) - R(z_0) = \frac{1}{2\pi i} \oint_{\partial D} R(\zeta) \left(\frac{1}{\zeta - z} - \frac{1}{\zeta - z_0} \right) d\zeta = \frac{1}{2\pi i} \oint_{\partial D} \frac{R(z)(z - z_0)}{(\zeta - z)(\zeta - z_0)}$$

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 **2**, **3**, **7**, and **8** the integrand of (4.5.11) eventually is bounded by $\frac{7|z - z_0|}{n^2}$. The contour of integration consists of four monotone curve of diameter $< 2n\pi$. By the Lemma on estimate, 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 $\operatorname{ctg}(\pi/2) = \lim\limits_{n\to\infty} \sum\limits_{k=-n}^{n} \frac{1}{\pi/2-k\pi} = \frac{2}{\pi} \lim\limits_{n\to\infty} \sum\limits_{k=-n}^{n} \frac{1}{1-2k}$.
This limit is zero because $\sum\limits_{k=-n}^{n} \frac{1}{1-2k} = \sum\limits_{k=-n}^{0} \frac{1}{1-2k} + \sum\limits_{k=1}^{n} \frac{1}{1-2k} = \sum\limits_{k=0}^{n} \frac{1}{2k+1} + \sum\limits_{k=1}^{n} -\frac{1}{2k-1} = \frac{1}{2n+1}$. □

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

Theorem 12. *For any rational function $R(z)$, which is not singular in integers and has degree ≤ -2, one has* $\sum\limits_{k=-\infty}^{\infty} R(n) = -\sum_z \operatorname{res} \pi \cot(\pi z) R(z)$

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

Factorization of $\sin x$**.** Theorem **11** with πz substituted for z gives the following:

$$\pi \cot \pi z = \sum_{k=-\infty}^{\infty} \frac{1}{z-k} \tag{4.5.12}$$

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 $\sum\limits_{k=-\infty}^{\infty} \frac{1}{k}[k \neq 0]$ to provide convergence:

$$\sum_{k=-\infty}^{-1} \left(\frac{1}{z-k} + \frac{1}{k}\right) + \sum_{k=0}^{\infty} \left(\frac{1}{z-k} + \frac{1}{k+1}\right) = \sum_{k=1}^{\infty} \left(-\frac{1}{k} + \frac{1}{z+k}\right) + \sum_{k=1}^{\infty} \left(\frac{1}{z+1-k} + \frac{1}{k}\right) \tag{4.5.13}$$

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 digamma function:

$$-\digamma(z) + \digamma(1-z) = \pi \cot \pi z \tag{4.5.14}$$

Since $\Theta''(z+1) = \digamma'(z) = \bar{\digamma}(z)$ (Lemma **11**) it follows $\Theta'(1+z) = \digamma(z) + c$ 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 Gamma-function*:

$$\Gamma(1-z)\Gamma(z) = \frac{\pi}{\sin \pi z} \tag{4.5.15}$$

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

$$\prod_{k=1}^{\infty}\left\{\left(1-\frac{x}{n}\right)e^{\frac{x}{n}}\right\}^{-1}\prod_{k=1}^{\infty}\left\{\left(1+\frac{x}{n}\right)e^{-\frac{x}{n}}\right\}^{-1} \tag{4.5.16}$$

The first product of (4.5.16) 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

$$\frac{1}{\Gamma(x)\Gamma(1-x)}=x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2}\right). \tag{4.5.17}$$

Comparing this to (4.5.15) and substituting πx for x we get the Euler formula:

$$\sin x = x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{\pi^2k^2}\right) \tag{4.5.18}$$

$\square$

Problems.

1. Expand $\tan z$ into power series
2. Evaluate $\sum_{k=1}^{\infty}\frac{1}{1+k^2}$
3. Evaluate $\sum_{k=1}^{\infty}\frac{1}{1+k^4}$

Bibliography

[1] E. Hairer, G. Wanner, *Analysis by Its Hystory*, Undergraduate Texts in Mathematics, Springer, 1997

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

[3] R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994

[4] L. Euler, *Introductio in analysin infinitorum*, t.1, Lausannae, 1748