4.1. Newton Series

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

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

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

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

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

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

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

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

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

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

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

Then

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

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

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

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

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

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

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

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

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

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

PROOF. The function f is bounded by Lemma 3.6.3. Denote by B the supremum of the set of values of f on $[a, b]$. If f does not take the maximum value, then $f(x) \neq B$ 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 3.6.3. But the difference $B - f(x)$ takes arbitrarily small values, because $B-\varepsilon$ does not bound $f(x)$. Therefore $\frac{1}{B-f(x)}$ is not bounded. This is in contradiction to Lemma 3.6.3, which states that a locally bounded function is bounded. The same arguments prove that $f(x)$ takes its minimal value on $[a,b]$. □

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Problems.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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