# LECTURE 14

# **COMPLEX NUMBERS.** POLYNOMIALS

#### **COMPLEX NUMBERS. BASIC DEFINITIONS**

A complex number is a number given by the expression

$$z = a + ib \tag{1}$$

where a and b are *real* numbers and i is the so-called *imaginary* unit, which is defined as

$$i = \sqrt{-1}$$
 or  $i^2 = -1$  (2)

a is called the *real part*, and b, the *imaginary part* of the complex number. They are designated, respectively, as follows:

$$a = \operatorname{Re} z, \ b = \operatorname{Im} z$$

If a=0, then the number 0+ib=ib is a pure imaginary; if b=0, then we have the real number a+i0=a. Two complex numbers z=a+ib and  $\overline{z}=a-ib$  that differ solely in the sign of the imaginary part are called *conjugate complex numbers*.

We agree upon the two following basic definitions.

(1) Two complex numbers  $z_1 = a_1 + ib_1$  and  $z_2 = a_2 + ib_2$  are equal,  $z_1 = z_2$ , if

$$a_1 = a_2, \quad b_1 = b_2$$

that is, if their real parts are equal and their imaginary parts are equal.

(2) A complex number z is equal to zero

$$z=a+ib=0$$

if and only if a = 0, b = 0.

1. Geometric representation of complex numbers. Any complex number z = a + ib may be represented in the xy-plane as a point A(a, b) with coordinates a and b. Conversely, every point M(x, y) of the plane is associated with a complex number z = x + iy. The plane on which complex numbers are represented is called the plane of the complex variable z, or the complex plane (Fig. 162, the encircled z symbol indicates that this is the complex plane).

Points of the plane of the complex variable z lying on the x-axis correspond to real numbers (b=0). Points lying on the y-axis represent pure imaginary numbers, since a=0. Therefore,

in the complex plane, the y-axis is called the *imaginary axis*, or axis of *imaginaries*, and the x-axis is the real axis, or axis of reals.

Joining the point A(a, b) to the origin, we get a vector  $\overline{OA}$ . In certain instances, it is convenient to consider the vector  $\overline{OA}$  as the geometric representation of the complex  $y \neq \mathbb{Z}$  A(a,b)

2. Trigonometric form of a complex number. Denote by  $\varphi$  and  $r \ (r \ge 0)$  the polar coordinates of the point A(a, b) and consider the origin as the pole and the positive direction of the x-axis, the polar axis. Then (Fig. 162) we have the familiar relationships



$$a = r \cos \varphi, \quad b = r \sin \varphi$$

and, hence, the complex number may be given in the form

$$a+ib=r\cos\varphi+ir\sin\varphi$$
 or  $z=r(\cos\varphi+i\sin\varphi)$  (3)

The expression on the right is called the *trigonometric form* (or *polar form*) of the complex number z = a + ib; r is termed the modulus of the complex number z,  $\varphi$  is the *argument* (*amplitude* or *phase*) of the complex number z. They are designated as

$$r = |z|, \quad \varphi = \arg z$$
 (4)

The quantities r and  $\varphi$  are expressed in terms of a and b as follows:

$$r = \sqrt{a^2 + b^2}, \quad \varphi = \operatorname{Arctan} \frac{b}{a}$$

To summarize, then,

$$|z| = |a+ib| = \sqrt{a^2 + b^2}$$
  
arg z = arg (a+ib) = Arctan  $\frac{b}{a}$  (5)

The amplitude of a complex number is considered positive if it is reckoned from the positive x-axis counterclockwise, and negative, in the opposite sense. The amplitude  $\varphi$  is obviously not determined uniquely but up to term  $2\pi k$ , where k is any integer.

Note. The conjugate complex numbers z = a + ib and  $\overline{z} = a - ib$  have equal moduli  $|z| = |\overline{z}|$  and their arguments (amplitudes) are equal in absolute value but differ in sign:  $\arg z = -\arg \overline{z}$ .

It will be noted that the real number A can also be written in the form (3), namely:

$$A = |A| (\cos 0 + i \sin 0) \text{ for } A > 0$$
  
$$A = |A| (\cos \pi + i \sin \pi) \text{ for } A < 0$$

The modulus of the complex number 0 is zero: |0| = 0. Any angle  $\varphi$  may be taken for amplitude zero. Indeed, for any angle  $\varphi$ we have

$$0 = 0 \left( \cos \varphi + i \sin \varphi \right)$$

#### **BASIC OPERATIONS ON COMPLEX NUMBERS**

1. Addition of complex numbers. The sum of two complex numbers  $z_1 = a_1 + ib_1$  and  $z_2 = a_2 + ib_2$  is a complex number defined by the equation

$$z_1 + z_2 = (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2)$$
(1)

From (1) it follows that the addition of complex numbers depicted as vectors is performed by the rule of the addition of vectors (Fig. 163a).



2. Subtraction of complex numbers. The difference of two complex numbers  $z_1 = a_1 + ib_1$  and  $z_2 = a_2 + ib_2$  is a complex number such that when it is added to  $z_2$  it yields  $z_1$ . It is easy to see that

$$z_1 - z_2 = (a_1 + ib_1) - (a_2 + ib_2) = (a_1 - a_2) + i(b_1 - b_2)$$
(2)

It will be noted that the modulus of the difference of two complex numbers is equal to the distance between the points representing these numbers in the plane of the complex variable (Fig. 163b)

$$|z_1-z_2| = \sqrt{(a_1-a_2)^2 + (b_1-b_2)^2}$$

3. Multiplication of complex numbers. The product of two complex numbers  $z_1 = a_1 + ib_1$  and  $z_2 = a_3 + ib_2$  is a complex number obtained when these two numbers are multiplied as binomials by the rules of algebra, provided that

 $i^2 = -1$ ,  $i^3 = -i$ ,  $i^4 = (-i) \cdot i = -i^2 = 1$ ,  $i^5 = i$ , etc. and, generally, for any integral k,

 $i^{4k} = 1$ ,  $i^{4k+1} = i$ ,  $i^{4k+2} = -1$ ,  $i^{4k+3} = -i$ 

From this rule we get

$$z_1 z_2 = (a_1 + ib_1) (a_2 + ib_2) = a_1 a_2 + ib_1 a_3 + ia_1 b_2 + i^2 b_1 b_2$$

or

$$z_1 z_2 = (a_1 a_2 - b_1 b_2) + i (b_1 a_2 + a_1 b_2)$$
(3)

Let the complex numbers be written in trigonometric form

$$z_1 = r_1 (\cos \varphi_1 + i \sin \varphi_1), \quad z_2 = r_2 (\cos \varphi_2 + i \sin \varphi_2)$$

then

$$z_1 z_2 = r_1 (\cos \varphi_1 + i \sin \varphi_1) r_2 (\cos \varphi_2 + i \sin \varphi_2)$$
  
=  $r_1 r_2 [\cos \varphi_1 \cos \varphi_2 + i \sin \varphi_1 \cos \varphi_2 + i \cos \varphi_1 \sin \varphi_2]$   
+  $i^2 \sin \varphi_1 \sin \varphi_2] = r_1 r_2 [(\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2)]$   
+  $i (\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2)]$   
=  $r_1 r_2 [\cos (\varphi_1 + \varphi_2) + i \sin (\varphi_1 + \varphi_2)]$ 

Thus,

$$z_{1}z_{2} = r_{1}r_{2} \left[ \cos{(\varphi_{1} + \varphi_{2})} + i\sin{(\varphi_{1} + \varphi_{2})} \right]$$
(3')

i.e., the product of two complex numbers is a complex number, the modulus of which is equal to the product of the moduli of the factors, and the amplitude is equal to the sum of the amplitudes of the factors.

Note 1. The product of two conjugate complex numbers z=a+iband  $\overline{z}=a-ib$  is, by virtue of (3), expressed as follows:

 $\overline{zz} = a^2 + b^2$ 

or

$$\overline{zz} = |z|^2 = |\overline{z}|^2$$

The product of two conjugate complex numbers is equal to the square of the modulus of each number.

4. Division of complex numbers. The *division* of complex numbers is defined as the inverse operation of multiplication.

Suppose we have  $z_1 = a_1 + ib_1$ ,  $z_2 = a_3 + ib_2$ ,  $|z_2| = \sqrt{a_2^2 + b_3^2} \neq 0$ . Then  $\frac{z_1}{z_2} = z$  is a complex number such that  $z_1 = z_3 z$ . If

$$\frac{a_1+ib_1}{a_2+ib_2} = x+iy$$

15\*

then

$$a_1 + ib_1 = (a_2 + ib_2)(x + iy)$$

or

$$a_1 + ib_1 = (a_2x - b_2y) + i(a_2y + b_2x)$$

x and y are found from the system of equations

$$a_1 = a_2 x - b_2 y, \ b_1 = b_3 x + a_2 y$$

Solving this system we get

$$x = \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2}, \qquad y = \frac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2}$$

and finally we have

$$z = \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + i \frac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2}$$
(4)

Actually, complex numbers are divided as follows: to divide  $z_1 = a_1 + ib_1$  by  $z_2 = a_2 + ib_2$ , multiply the dividend and divisor by a complex number conjugate to the divisor (that is, by  $a_2 - ib_2$ ). Then the divisor will be a real number; dividing the real and imaginary parts of the dividend by it, we get the quotient

$$\frac{a_1+ib_1}{a_2+ib_2} = \frac{(a_1+ib_1)(a_2-ib_3)}{(a_2+ib_2)(a_2-ib_3)} = \frac{(a_1a_2+b_1b_2)+i(a_2b_1-a_1b_2)}{a_2^2+b_2^2} = \frac{a_1a_2+b_1b_2}{a_2^2+b_3^2} + i\frac{a_2b_1-a_1b_2}{a_2^2+b_3^2}$$

If the complex numbers are given in the trigonometric form

$$z_1 = r_1 (\cos \varphi_1 + i \sin \varphi_1), \quad z_2 = r_2 (\cos \varphi_2 + i \sin \varphi_2)$$

then

$$\frac{z_1}{z_2} = \frac{r_1 (\cos \varphi_1 + i \sin \varphi_1)}{r_2 (\cos \varphi_2 + i \sin \varphi_2)} = \frac{r_1}{r_2} \left[ \cos (\varphi_1 - \varphi_2) + i \sin (\varphi_1 - \varphi_2) \right]$$
(5)

To verify this equation, multiply the divisor by the quotient:

$$r_{2} (\cos \varphi_{2} + i \sin \varphi_{2}) \frac{r_{1}}{r_{2}} [\cos (\varphi_{1} - \varphi_{2}) + i \sin (\varphi_{1} - \varphi_{2})]$$
  
=  $r_{2} \frac{r_{1}}{r_{2}} [\cos (\varphi_{2} + \varphi_{1} - \varphi_{2}) + i \sin (\varphi_{2} + \varphi_{1} - \varphi_{2})] = r_{1} (\cos \varphi_{1} + i \sin \varphi_{1})$ 

Thus, the modulus of the quotient of two complex numbers is equal to the quotient of the moduli of the dividend and the divisor; the amplitude of the quotient is equal to the difference between the amplitudes of the dividend and divisor.

Note 2. From the rules of operations involving complex numbers it follows that the operations of addition, subtraction, multiplication and division of complex numbers yield a complex number. If the rules of operations on complex numbers are applied to real numbers, these being regarded as a special case of complex numbers, they will coincide with the ordinary rules of arithmetic.

Note 3. Returning to the definitions of a sum, difference, product and quotient of complex numbers, it is easy to show that if each complex number in these expressions is replaced by its conjugate, then the results of the aforementioned operations will yield conjugate numbers, whence, as a particular case, we have the following theorem.

**Theorem.** If in a polynomial with real coefficients

 $A_0x^n + A_1x^{n-1} + \ldots + A_n$ 

we put, in place of x, the number a+ib and then the conjugate number a-ib the results of these substitutions will be mutually conjugate.

# **POWERS AND ROOTS OF COMPLEX NUMBERS**

1. Powers. From formula (3') of the preceding section it follows that if n is a positive integer, then

$$[r(\cos\varphi + i\sin\varphi)]^{n} = r^{n}(\cos n\varphi + i\sin n\varphi)$$
(1)

This formula is called De Moivre's formula. It shows that when a complex number is raised to a positive integral power the modulus is raised to this power, and the amplitude is multiplied by the exponent.

Now consider another application of De Moivre's formula.

Setting r = 1 in this formula, we get

 $(\cos \varphi + i \sin \varphi)^n = \cos n\varphi + i \sin n\varphi$ 

Expanding the left-hand side by the binomial theorem and equating the real and imaginary parts, we can express  $\sin n\varphi$  and  $\cos n\varphi$  in terms of powers of  $\sin \varphi$  and  $\cos \varphi$ . For instance, if n = 3 we have

 $\cos^3 \varphi + i 3 \cos^2 \varphi \sin \varphi - 3 \cos \varphi \sin^2 \varphi - i \sin^3 \varphi = \cos 3\varphi + i \sin 3\varphi$ Making use of the condition of equality of two complex numbers, we get

$$\cos 3\varphi = \cos^3 \varphi - 3 \cos \varphi \sin^2 \varphi$$
$$\sin 3\varphi = -\sin^3 \varphi + 3 \cos^2 \varphi \sin \varphi$$

2. Roots. The *n*th root of a complex number is another complex number whose *n*th power is equal to the radicand, or

$$\sqrt[n]{r(\cos\varphi + i\sin\varphi)} = \rho(\cos\psi + i\sin\psi)$$

if

$$\rho^n (\cos n\psi + i \sin n\psi) = r (\cos \varphi + i \sin \varphi)$$

Since the moduli of equal complex numbers must be equal, while their amplitudes may differ by a multiple of  $2\pi$ , we have

$$\rho^n = r$$
,  $n\psi = \varphi + 2k\pi$ 

Whence we find

$$\rho = \sqrt[n]{r}, \quad \psi = \frac{\varphi + 2k\pi}{n}$$

where k is any integer,  $\sqrt[n]{r}$  is the principal (positive real) root of the positive number r. Therefore,

$$\sqrt[n]{r(\cos\varphi + i\sin\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right)$$
(2)

Giving k the values 0, 1, 2, ..., n-1, we get n different values of the root. For the other values of k, the amplitudes will differ from those obtained by a multiple of  $2\pi$ , and, for this reason, root values will be obtained that coincide with those considered.

Thus, the *n*th root of a complex number has *n* different values. The *n*th root of a real nonzero number A also has *n* values, since a real number is a special case of a complex number and

may be represented in trigonometric form:



if A > 0, then  $A = |A| (\cos 0 + i \sin 0)$ if A < 0, then  $A = |A| (\cos \pi + i \sin \pi)$ 

**Example 1.** Find all the values of the cube root of unity. Solution. We represent unity in trigonometric form:  $1 = \cos 0 + i \sin 0$ 

By formula (2) we have

$$\sqrt[3]{1} = \sqrt[3]{\cos 0 + i \sin 0} = \cos \frac{0 + 2k\pi}{3} + i \sin \frac{0 + 2k\pi}{3}$$

Setting k equal to 0, 1, 2, we find three values of the root:

$$x_1 = \cos 0 + i \sin 0 = 1, \ x_2 = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3},$$
$$x_3 = \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}$$

Noting that

$$\cos\frac{2\pi}{3} = -\frac{1}{2}$$
,  $\sin\frac{2\pi}{3} = \frac{\sqrt{3}}{2}$ ,  $\cos\frac{4\pi}{3} = -\frac{1}{2}$ ,  $\sin\frac{4\pi}{3} = -\frac{\sqrt{3}}{2}$ 

we get

$$x_1 = 1, \ x_2 = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \ x_3 = -\frac{1}{2} - i \frac{\sqrt{3}}{2}$$

In Fig. 164, the points A, B, C are geometric representations of the roots obtained.

# 3. Solution of a binomial equation. An equation of the form $x^n = A$

is called a binomial equation. Let us find its roots.

If A is a real positive number, then

$$x = \sqrt[n]{A} \left( \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \right) \ (k = 0, \ 1, \ 2, \ \dots, \ n-1)$$

The expression in the brackets gives all the values of the nth root of 1.

If A is a real negative number, then

$$x = \sqrt[n]{|A|} \left( \cos \frac{\pi + 2k\pi}{n} + i \sin \frac{\pi + 2k\pi}{n} \right)$$

The expression in the brackets gives all the values of the nth root of -1.

If A is a complex number, then the values of x are found from formula (2).

Example 2. Solve the equation

 $x^4 = 1$ 

Solution.

$$x = \sqrt[4]{\cos 2k\pi + i \sin 2k\pi} = \cos \frac{2k\pi}{4} + i \sin \frac{2k\pi}{4}$$

Setting k equal to 0, 1, 2, 3, we get

$$x_{1} = \cos 0 + i \sin 0 = 1$$
  

$$x_{2} = \cos \frac{2\pi}{4} + i \sin \frac{2\pi}{4} = i$$
  

$$x_{3} = \cos \frac{4\pi}{4} + i \sin \frac{4\pi}{4} = -1$$
  

$$x_{4} = \cos \frac{6\pi}{4} + i \sin \frac{6\pi}{4} = -i$$

#### EXPONENTIAL FUNCTION WITH COMPLEX EXPONENT AND ITS PROPERTIES

Let z = x + iy. If x and y are real variables, then z is called a complex variable. To each value of the complex variable z in the xy-plane (the complex plane) there corresponds a definite point (see Fig. 162).

**Definition.** If to every value of the complex variable z of a certain range of complex values there corresponds a definite value of another complex quantity w, then w is a function of the complex variable z. Functions of a complex variable are denoted by w = f(z) or w = w(z).

Here, we consider the exponential function of a complex variable:

or

 $w = e^{x + iy}$ 

 $w = e^z$ 

The complex values of the function w are defined as follows \*:

$$e^{x+iy} = e^x \left(\cos y + i \sin y\right) \tag{1}$$

that is

$$w(z) = e^x (\cos y + i \sin y) \tag{2}$$

Examples:

1. 
$$z = 1 + \frac{\pi}{4}i$$
,  $e^{1 + \frac{\pi}{4}i} = e\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = e\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right)$ .  
2.  $z = 0 + \frac{\pi}{2}i$ ,  $e^{0 + \frac{\pi}{2}i} = e^{0}\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) = i$ .

3. z=1+i,  $e^{1+i}=e^1(\cos 1+i\sin 1)\approx 0.54+i\cdot 0.83$ . 4. z=x is a real number,  $e^{x+0i}=e^x(\cos 0+i\sin 0)=e^x$  is an ordinary exponential function.

# Properties of an exponential function.

1. If  $z_1$  and  $z_2$  are two complex numbers, then

$$e^{z_1+z_3} = e^{z_1}e^{z_3} \tag{3}$$

Proof. Let

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2$$

then

$$e^{z_1+z_2} = e^{(x_1+iy_1)+(x_2+iy_2)} = e^{(x_1+x_2)+i(y_1+y_2)}$$
  
=  $e^{x_1}e^{x_2}[\cos(y_1+y_2)+i\sin(y_1+y_2)]$  (4)

On the other hand, by the theorem of the product of two complex numbers in trigonometric form we will have

$$e^{z_1}e^{z_2} = e^{x_1 + iy_1}e^{x_2 + iy_2} = e^{x_1} \left(\cos y_1 + i\sin y_1\right)e^{x_2} \left(\cos y_2 + i\sin y_2\right)$$
$$= e^{x_1}e^{x_2} \left[\cos \left(y_1 + y_2\right) + i\sin \left(y_1 + y_2\right)\right]$$
(5)

In (4) and (5) the right sides are equal, hence the left sides are equal too:

$$e^{z_1+z_2}=e^{z_1}e^{z_2}$$

2. The following formula is similarly proved:

$$e^{z_1 - z_2} = \frac{e^{z_1}}{e^{z_2}} \tag{6}$$

<sup>\*</sup> The advisability of this definition of the exponential function of a complex variable will also be shown later on (see Sec. 13.21 and Sec. 16.18 of Vol. II).

3. If m is an integer, then

$$(e^z)^m = e^{mz} \tag{7}$$

For m > 0, this formula is readily obtained from (3); if m < 0, then it is obtained from formulas (3) and (6).

4. The identity

$$e^{z+2\pi i}=e^z \tag{8}$$

holds.

Indeed, from (3) and (1) we get

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z (\cos 2\pi + i \sin 2\pi) = e^z$$

From identity (8) it follows that the exponential function  $e^z$  is a periodic function with a *period* of  $2\pi i$ .

5. Let us now consider the complex quantity

$$w = u(x) + iv(x)$$

where u(x) and v(x) are real functions of a real variable x. This is a complex function of a real variable.

(a) Let there exist the limits

$$\lim_{x \to x_0} u(x) = u(x_0), \quad \lim_{x \to x_0} v(x) = v(x_0)$$

Then  $u(x_0) + iv(x_0) = w_0$  is called the *limit* of the complex variable w.

(b) If the derivatives u'(x) and v'(x) exist, then we shall call the expression

$$w'_{x} = u'(x) + iv'(x)$$
 (9)

the *derivative* of a complex function of a real variable with respect to a real argument.

Let us now consider the following exponential function:

$$w = e^{\alpha x + i\beta x} = e^{(\alpha + i\beta)x}$$

where  $\alpha$  and  $\beta$  are real constants and x is a real variable. This is a complex function of a real variable, which function may be rewritten, according to (1), as follows:

$$w = e^{ex} \left[ \cos \beta x + i \sin \beta x \right]$$

or

$$w = e^{ax} \cos \beta x + i e^{ax} \sin \beta x$$

Let us find the derivative  $w'_x$ . From (9) we have

$$w'_{x} = (e^{\alpha x} \cos \beta x)' + i (e^{\alpha x} \sin \beta x)'$$
  
=  $e^{\alpha x} (\alpha \cos \beta x - \beta \sin \beta x) + ie^{\alpha x} (\alpha \sin \beta x + \beta \cos \beta x)$   
=  $\alpha [e^{\alpha x} (\cos \beta x + i \sin \beta x)] + i\beta [e^{\alpha x} (\cos \beta x + i \sin \beta x)]$   
=  $(\alpha + i\beta) [e^{\alpha x} (\cos \beta x + i \sin \beta x)] = (\alpha + i\beta) e^{(\alpha + i\beta)x}$ 

To summarize then, if  $w = e^{(\alpha + i\beta)x}$ , then  $w' = (\alpha + i\beta)e^{(\alpha + i\beta)x}$  or  $[e^{(\alpha+i\beta)x}]' = (\alpha+i\beta)e^{(\alpha+i\beta)x}$ (10)

Thus, if k is a complex number (or, in the special case, a real number) and x is a real number, then

$$(e^{kx})' = ke^{kx} \tag{9'}$$

We have thus obtained the ordinary formula for differentiating an exponential function. Further,

$$(e^{kx})'' = [(e^{kx})']' = k (e^{kx})' = k^2 e^{kx}$$

and for arbitrary n

$$(e^{kx})^{(n)} = k^n e^{kx}$$

We shall need these formulas later on.

# EULER'S FORMULA. THE EXPONENTIAL FORM OF A COMPLEX NUMBER

Putting x=0 in formula (1) of the preceding section, we get  $e^{iy} = \cos y + i \sin y$ (1)

This is Euler's formula, which expresses an exponential function with an imaginary exponent in terms of trigonometric functions. Replacing y by -y in (1) we get

$$e^{-iy} = \cos y - i \sin y \tag{2}$$

From (1) and (2) we find  $\cos y$  and  $\sin y$ :

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}$$
  

$$\sin y = \frac{e^{iy} - e^{-iy}}{2i}$$
(3)

These formulas are used in particular to express powers of  $\cos \varphi$ and  $\sin \varphi$  and their products in terms of the sine and cosine of multiple arcs.

Example 1. 
$$\cos^2 y = \left(\frac{e^{iy} + e^{-iy}}{2}\right)^2 = \frac{1}{4} \left(e^{i2y} + 2 + e^{-i2y}\right)$$
  

$$= \frac{1}{4} \left[ (\cos 2y + i \sin 2y) + 2 + (\cos 2y - i \sin 2y) \right]$$

$$= \frac{1}{4} \left( 2 \cos 2y + 2 \right) = \frac{1}{2} \left( 1 + \cos 2y \right)$$
Example 2.  $\cos^2 \varphi \sin^2 \varphi = \left(\frac{e^{i\varphi} + e^{-i\varphi}}{2}\right)^2 \left(\frac{e^{i\varphi} - e^{-i\varphi}}{2i}\right)^2$   

$$= \frac{\left(e^{i2\varphi} - e^{-i2\varphi}\right)^2}{4 \cdot 4i^2} = -\frac{1}{8} \cos 4\varphi + \frac{1}{8}$$

The exponential form of a complex number. Let us represent a complex number in trigonometric form:

$$z = r (\cos \varphi + i \sin \varphi)$$

where r is the modulus of the complex number and  $\varphi$  is the argument (amplitude) of the complex number. By Euler's formula,

$$\cos\varphi + i\sin\varphi = e^{i\varphi} \tag{4}$$

Thus, any complex number may be represented in the so-called exponential form:

$$z = r e^{l\varphi}$$

**Example 3.** Represent the numbers 1, *i*, -2, -*i* in exponential form. Solution.  $1 = \cos 2k\pi + i \sin 2k\pi = e^{2k\pi i}$ 

$$i = \cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = e^{\frac{\pi}{2}i}$$
$$-2 = 2(\cos\pi + i \sin\pi) = 2e^{\pi i}$$
$$-i = \cos\left(-\frac{\pi}{2}\right) + i \sin\left(-\frac{\pi}{2}\right) = e^{-\frac{\pi}{2}i}$$

By Properties (3), (6), (7), Sec. 7.4, of an exponential function, it is easy to operate on complex numbers in exponential form.

Suppose we have

$$z_1 = r_1 e^{i\varphi_1}, \quad z_2 = r_2 e^{i\varphi_2}$$

then

$$z_1 \cdot z_2 = r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$$
(5)

$$\frac{z_1}{z_2} = \frac{r_1 e^{i\varphi_1}}{r_2 e^{i\varphi_2}} = \frac{r_1}{r_2} e^{i(\varphi_1 - \varphi_2)}$$
(6)

$$z^n = (re^{l\varphi})^n = r^n e^{in\varphi} \tag{7}$$

$$\sqrt[n]{re^{i\varphi}} = \sqrt[n]{r}e^{i\frac{\varphi + 3k\pi}{n}} \quad (k = 0, 1, 2, ..., n-1)$$
(8)

Formula (5) coincides with (3') of Sec. 7.2; (6), with (5) of Sec. 7.2; (7), with (1) of Sec. 7.3; (8) with (2) of Sec. 7.3.

#### FACTORING A POLYNOMIAL

The function

$$f(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_n$$

where *n* is an integer, is known as a *polynomial* or a *rational integral function* of *x*; the number *n* is called the *degree of the polynomial*. Here, the coefficients  $A_0, A_1, \ldots, A_n$  are real or complex numbers; the independent variable *x* can also take on both real and complex values. The *root* of a polynomial is that value of the variable *x* at which the polynomial becomes zero. **Theorem 1 (Remainder Theorem).** Division of a polynomial f(x) by x - a yields a remainder equal to f(a).

**Proof.** The quotient obtained by the division of f(x) by x-a is a polynomial  $f_1(x)$  of degree one less than that of f(x), and the remainder is a constant R. We can thus write

$$f(x) = (x - a) f_1(x) + R$$
 (1)

This equation holds for all values of x different from a (division by x - a when x = a is meaningless).

Now let x approach a. Then the limit of the left side of (1) will equal f(a), and the limit of the right side will equal R. Since the functions f(x) and  $(x-a)f_1(x) + R$  are equal for all  $x \neq a$ , their limits are likewise equal as  $x \rightarrow a$ , that is, f(a) = R.

**Corollary.** If a is a root of the polynomial, that is, if f(a) = 0, then x - a divides f(x) without remainder and, hence, f(x) is represented in the form of a product

$$f(x) = (x - a) f_1(x)$$

where  $f_1(x)$  is a polynomial.

**Example 1.** The polynomial  $f(x) = x^3 - 6x^2 + 11x - 6$  becomes zero for x = 1; thus, f(1) = 0, and so x - 1 divides this polynomial without remainder:

 $x^3-6x^2+11x-6=(x-1)(x^2-5x+6)$ 

Let us now consider equations in one unknown, x.

Any number (real or complex) which, when substituted into the equation in place of x, converts the equation into an identity is called a *root* of the equation.

**Example 2.** The numbers  $x_1 = \frac{\pi}{4}$ ,  $x_2 = \frac{5\pi}{4}$ ,  $x_3 = \frac{9\pi}{4}$ , ... are the roots of the equation  $\cos x = \sin x$ .

If the equation is of the form P(x) = 0, where P(x) is a polynomial of degree *n*, it is called an *algebraic* equation of degree *n*. From the definition it follows that the roots of an algebraic equation P(x) = 0 are the same as are the roots of the polynomial P(x).

Quite naturally the question arises: Does every equation have roots?

In the case of nonalgebraic equations, the answer is no: there are nonalgebraic equations which do not have a single root, either real or complex; for example, the equation  $e^x = 0$ .\*

<sup>\*</sup> Indeed, if the number  $x_1 = a + ib$  were the root of this equation, we would have the identity  $e^{a+ib} = 0$  or (by Euler's formula)  $e^a (\cos b + i \sin b) = 0$ . But  $e^a$  cannot equal zero for any real value of a; neither is  $\cos b + i \sin b$  equal to zero (because the modulus of this number is  $\sqrt{\cos^2 b + \sin^2 b} = 1$  for any b). Hence, the product  $e^a (\cos b + i \sin b) \neq 0$ , i.e.,  $e^{a+ib} \neq 0$ ; but this means that the equation  $e^x = 0$  has no roots.

But in the case of an algebraic equation the answer is yes. This is given by the fundamental theorem of algebra.

**Theorem 2 (Fundamental Theorem of Algebra).** Every rational integral function f(x) has at least one root, real or complex.

The proof of this theorem is given in higher algebra. Here we accept it without proof.

With the aid of the fundamental theorem of algebra it is easy to prove the following theorem.

**Theorem 3.** Every polynomial of degree n may be factored into n linear factors of the form x - a and a factor equal to the coefficient of  $x^n$ .

**Proof.** Let f(x) be a polynomial of degree n:

$$f(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_n$$

By virtue of the fundamental theorem, this polynomial has at least one root; we denote it by  $a_1$ . Then, by a corollary of the remainder theorem, we can write

$$f(x) = (x - a_1) f_1(x)$$

where  $f_1(x)$  is a polynomial of degree n-1;  $f_1(x)$  also has a root. We designate it by  $a_2$ . Then

$$f_1(x) = (x - a_2) f_2(x)$$

where  $f_2(x)$  is a polynomial of degree n-2. Similarly,

$$f_{2}(x) = (x - a_{3}) f_{3}(x)$$

Continuing this process of factoring out linear factors, we arrive at the relation

$$f_{n-1}(x) = (x - a_n) f_n$$

where  $f_n$  is a polynomial of degree zero, i.e., some specified number. This number is obviously equal to the coefficient of  $x^n$ ; that is,  $f_n = A_0$ .

On the basis of the equations obtained we can write

$$f(x) = A_0 (x - a_1) (x - a_2) \dots (x - a_n)$$
(2)

From the expansion (2) it follows that the numbers  $a_1, a_2, \ldots, a_n$  are roots of the polynomial f(x), since upon the substitution  $x = a_1$ ,  $x = a_2, \ldots, x = a_n$  the right side, and hence, the left, becomes zero.

**Example 3.** The polynomial  $f(x) = x^3 - 6x^2 + 11x - 6$  becomes zero when

$$x = 1, x = 2, x = 3$$

Therefore,

$$x^3 - 6x^2 + 11x - 6 = (x - 1)(x - 2)(x - 3)$$

No value x = a that is different from  $a_1, a_2, \ldots, a_n$  can be a root of the polynomial f(x), since no factor on the right side of (2) vanishes when x = a. Whence the following proposition.

A polynomial of degree n cannot have more than n distinct roots. But then the following theorem obtains.

**Theorem 4.** If the values of two polynomials of degree n,  $\varphi_1(x)$  and  $\varphi_2(x)$ , coincide for n+1 distinct values  $a_0, a_1, a_2, \ldots, a_n$  of the argument x, then these polynomials are identical.

**Proof.** Denote the difference of the polynomials by f(x):

$$f(x) = \varphi_1(x) - \varphi_2(x)$$

It is given that f(x) is a polynomial of degree not higher than *n* that becomes zero at the points  $a_1, \ldots, a_n$ . It can therefore be represented in the form

$$f(x) = A_0 (x - a_1) (x - a_2) \dots (x - a_n)$$

But it is given that f(x) also vanishes at the point  $a_0$ . Then  $f(a_0) = 0$ and not a single one of the linear factors equals zero. For this reason,  $A_0 = 0$  and then from (2) it follows that the polynomial f(x) is identically equal to zero. Consequently,  $\varphi_1(x) - \varphi_2(x) \equiv 0$ or  $\varphi_1(x) \equiv \varphi_2(x)$ .

Theorem 5. If a polynomial

$$P(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_{n-1} x + A_n$$

is identically equal to zero, all its coefficients equal zero. **Proof.** Let us write its factorization using formula (2):

$$P(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_{n-1} x + A_n = A_0 (x - a_1) \ldots (x - a_n) \quad (1')$$

If this polynomial is identically equal to zero, it is also equal to zero for some value of x different from  $a_1, \ldots, a_n$ . But then none of the bracketed values  $x - a_1, \ldots, x - a_n$  is equal to zero, and, hence,  $A_0 = 0$ .

Similarly it is proved that  $A_1 = 0$ ,  $A_2 = 0$ , and so forth.

**Theorem 6.** If two polynomials are identically equal, the coefficients of one polynomial are equal to the corresponding coefficients of the other.

This follows from the fact that the difference between the polynomials is a polynomial identically equal to zero. Therefore, from the preceding theorem all its coefficients are zeros.

**Example 4.** If the polynomial  $ax^3 + bx^2 + cx + d$  is identically equal to the polynomial  $x^3 - 5x$ , then a = 0, b = 1, c = -5, and d = 0.

#### THE MULTIPLE ROOTS OF A POLYNOMIAL

If, in the factorization of a polynomial of degree n into linear factors

$$f(x) = A_0 (x - a_1) (x - a_2) \dots (x - a_n)$$
(1)

certain linear factors turn out the same, they may be combined, and then factorization of the polynomial will yield

$$f(x) = A_0 (x - a_1)^{k_1} (x - a_2)^{k_2} \dots (x - a_m)^{k_m}$$
(1')

Here

 $k_1+k_2+\ldots+k_m=n$ 

In this case, the root  $a_1$  is called a root of multiplicity  $k_1$ , or a  $k_1$ -tuple root,  $a_2$ , a root of multiplicity  $k_2$ , etc.

**Example.** The polynomial  $f(x) = x^3 - 5x^2 + 8x - 4$  may be factored into the following linear factors:

$$f(x) = (x-2)(x-2)(x-1)$$

This factorization may be written as follows:

$$f(x) = (x-2)^2(x-1)$$

The root  $a_1 = 2$  is a double root,  $a_2 = 1$  is a simple root.

If a polynomial has a root a of multiplicity k, then we will consider that the polynomial has k coincident roots. Then from the theorem of factorization of a polynomial into linear factors we get the following theorem.

Every polynomial of degree n has exactly n roots (real or complex). Note. All that has been said of the roots of the polynomial

$$f(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_n$$

may obviously be formulated in terms of the roots of the algebraic equation

$$A_0x^n + A_1x^{n-1} + \ldots + A_n = 0$$

Let us now prove the following theorem.

**Theorem.** If, for the polynomial f(x),  $a_1$  is a root of multiplicity  $k_1 > 1$ , then for the derivative f'(x) this number is a root of multiplicity  $k_1 - 1$ .

**Proof.** If  $a_1$  is a root of multiplicity  $k_1 > 1$ , then it follows from formula (1') that

$$f(x) = (x - a_1)^{k_1} \varphi(x)$$

where  $\varphi(x) = (x - a_2)^{k_2} \dots (x - a_m)^{k_m}$  does not become zero at  $x = a_1$ ; that is,  $\varphi(a_1) \neq 0$ . Differentiating, we get

$$f'(x) = k_1 (x - a_1)^{k_1 - 1} \varphi(x) + (x - a_1)^{k_1} \varphi'(x)$$
  
=  $(x - a_1)^{k_1 - 1} [k_1 \varphi(x) + (x - a_1) \varphi'(x)]$ 

Put

$$\psi(x) = k_1 \varphi(x) + (x - a_1) \varphi'(x)$$

Then

$$f'(x) = (x - a_1)^{k_1 - 1} \psi(x)$$

and here

$$\psi(a_1) = k_1 \varphi(a_1) + (a_1 - a_1) \varphi'(a_1) = k_1 \varphi(a_1) \neq 0$$

In other words,  $x = a_1$  is a root of multiplicity  $k_1 - 1$  of the polynomial f'(x). From the foregoing proof it follows that if  $k_1 = 1$ , then  $a_1$  is not a root of the derivative f'(x). From the proved theorem it follows that  $a_1$  is a root of multiplicity that  $a_1$  is a root of multiplicity the second second

From the proved theorem it follows that  $a_1$  is a root of multiplicity  $k_1-2$  for the derivative f''(x), a root of multiplicity  $k_1-3$  for the derivative  $f'''(x) \ldots$ , and a root of multiplicity one (simple root) for the derivative  $f^{(k_1-1)}(x)$  and is not a root for the derivative  $f^{(k_1)}(x)$ , or

$$f(a_1) = 0, f'(a_1) = 0, f''(a_1) = 0, \dots, f^{(k_1-1)}(a_1) = 0$$

but

$$f^{(k)}(a_1) \neq 0$$

### FACTORING A POLYNOMIAL IN THE CASE OF COMPLEX ROOTS

In formula (1), Sec. 7.7, the roots  $a_1, a_2, \ldots, a_n$  may be either real or complex. We have the following theorem.

**Theorem.** If a polynomial f(x) with real coefficients has a complex root a + ib, it also has a conjugate root a - ib.

**Proof.** Substitute, in the polynomial f(x), a+ib in place of x, raise to a power and collect separately terms containing i and those not containing i; we then get

$$f(a+ib)=M+iN,$$

where M and N are expressions that do not contain i.

Since a+ib is a root of the polynomial, we have

$$f(a+ib) = M+iN = 0$$

whence

$$M = 0, N = 0$$

Now substitute the expression a-ib for x in the polynomial. Then (on the basis of Note 3 at the end of Sec. 7.2) we get the conjugate of the number M + iN, or

$$f(a-ib) = M - iN$$

Since M = 0 and N = 0, we have f(a-ib) = 0; a-ib is a root of the polynomial.

Thus, in the factorization

$$f(x) = A_0(x - a_1)(x - a_2) \dots (x - a_n)$$

the complex roots enter as conjugate pairs.

Multiplying together the linear factors that correspond to a pair of complex conjugate roots, we get a trinomial of degree two with real coefficients:

$$[x-(a+ib)] [x-(a-ib)] = [(x-a)-ib] [(x-a)+ib] = (x-a)^2+b^2 = x^2-2ax+a^2+b^2 = x^2+px+q$$

where p = -2a,  $q = a^2 + b^2$  are real numbers.

If the number a+ib is a root of multiplicity k, the conjugate number a-ib must be a root of the same multiplicity k, so that factorization of the polynomial will yield the same number of linear factors x-(a+ib) as those of the form x-(a-ib).

Thus, a polynomial with real coefficients may be factored into factors with real coefficients of the first and second degree of corresponding multiplicity; that is,

$$f(x) = A_0 (x - a_1)^{k_1} (x - a_2)^{k_2} \dots (x - a_r)^{k_r} (x^2 + p_1 x + q_1)^{l_1} \dots (x^2 + p_s x + q_s)^{l_s}$$

where

$$k_1+k_2+\ldots+k_r+2l_1+\ldots+2l_s=n$$

# INTERPOLATION. LAGRANGE'S INTERPOLATION FORMULA

Let it be established, in the study of some phenomenon, that there is a functional relationship between the quantities y and xwhich describes the quantitative aspect of the phenomenon; the function  $y = \varphi(x)$  is unknown, but

experiment has established the values of this function  $y_0, y_1, y_2, \ldots, y_n$  for certain values of the argument  $x_0, x_1, x_2, \ldots, x_n$ , in the interval [a, b].

The problem is to find a function (as simple as possible from the computational standpoint; for example, a polynomial) which will represent the unknown function  $y = \varphi(x)$  on the interval [a, b] either exactly or approximately. In more



Fig. 165

abstract fashion the problem may be formulated as follows: given on the interval [a, b] the values of an unknown function  $y = \varphi(x)$  at n+1 distinct points  $x_0, x_1, \ldots, x_n$ :

$$y_0 = \varphi(x_0), \ y_1 = \varphi(x_1), \ \ldots, \ y_n = \varphi(x_n)$$

It is required to find a **polynomial** P(x) of degree  $\leq n$  that approximately expresses the function  $\varphi(x)$ .

It is natural to take a polynomial whose values at the points  $x_0, x_1, x_2, \ldots, x_n$  coincide with the corresponding values  $y_0, y_1$ ,

 $y_2, \ldots, y_n$  of the function  $\varphi(x)$  (Fig. 165). Then the problem, which is called the "problem of *interpolating* a function", is formulated thus: for a given function  $\varphi(x)$  find a polynomial P(x) of degree  $\leq n$ , which, for the given values of  $x_0, x_1, \ldots, x_n$ , will take on the values

$$y_0 = \varphi(x_0), y_1 = \varphi(x_1), \ldots, y_n = \varphi(x_n)$$

For the desired polynomial, take a polynomial of degree n of the form

$$P(x) = C_0 (x - x_1) (x - x_2) \dots (x - x_n) + C_1 (x - x_0) (x - x_2) \dots (x - x_n) + C_2 (x - x_0) (x - x_1) (x - x_3) \dots (x - x_n) + \dots + C_n (x - x_0) (x - x_1) \dots (x - x_{n-1})$$
(1)

and define the coefficients  $C_0, C_1, \ldots, C_n$  so that the following conditions are fulfilled:

$$P(x_0) = y_0, P(x_1) = y_1, \ldots, P(x_n) = y_n$$
 (2)

In (1) put  $x = x_0$ ; then, taking into account (2), we get

$$y_0 = C_0 (x_0 - x_1) (x_0 - x_2) \dots (x_0 - x_n)$$

whence

$$C_{0} = \frac{y_{0}}{(x_{0} - x_{1}) (x_{0} - x_{2}) \dots (x_{0} - x_{n})}$$

Then, setting  $x = x_1$ , we get

$$y_1 = C_1 (x_1 - x_0) (x_1 - x_2) \dots (x_1 - x_n)$$

whence

$$C_1 = \frac{y_1}{(x_1 - x_0) (x_1 - x_2) \dots (x_1 - x_n)}$$

In the same way we find

$$C_{2} = \frac{y_{2}}{(x_{2} - x_{0}) (x_{2} - x_{1}) (x_{2} - x_{3}) \dots (x_{2} - x_{n})}$$
  

$$\cdots$$
  

$$C_{n} = \frac{y_{n}}{(x_{n} - x_{0}) (x_{n} - x_{1}) (x_{n} - x_{2}) \dots (x_{n} - x_{n-1})}$$

Substituting these values of the coefficients into (1), we get

$$P(x) = \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)}y_0 + \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)}y_1 + \frac{(x-x_0)(x-x_1)(x-x_3)\dots(x-x_n)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)\dots(x_2-x_n)}y_2 + \dots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})}y_n$$
(3)

This formula is called Lagrange's interpolation formula.

Let it be noted, without proof, that if  $\varphi(x)$  has a derivative of the (n+1)th order on the interval [a, b], the error resulting from replacing the function  $\varphi(x)$  by the polynomial P(x), i. e., the quantity  $R(x) = \varphi(x) - P(x)$ , satisfies the inequality

$$|R(x)| < |(x-x_0)(x-x_1)...(x-x_n)| \frac{1}{(n+1)!} \max |\varphi^{(n+1)}(x)|$$

Note. From Theorem 4, Sec. 7.6, it follows that the polynomial P(x) which we found is the only one that satisfies the given conditions.

There are other interpolation formulas, one of which (Newton's) is considered in Sec. 7.10.

**Example.** From experiment we get the values of the function  $y = \varphi(x)$ ;  $y_0 = 3$  for  $x_0 = 1$ ,  $y_1 = -5$  for  $x_1 = 2$ ,  $y_2 = 4$  for  $x_2 = -4$ . It is required to represent the function  $y = \varphi(x)$  approximately by a polynomial of degree two. Solution. From (3) we have (for n = 2):

$$P(x) = \frac{(x-2)(x+4)}{(1-2)(1+4)} 3 + \frac{(x-1)(x+4)}{(2-1)(2+4)} (-5) + \frac{(x-1)(x-2)}{(-4-1)(-4-2)} 4$$

or

$$P(x) = -\frac{39}{30}x^2 - \frac{123}{30}x + \frac{252}{30}$$

#### **NEWTON'S INTERPOLATION FORMULA**

Suppose we know (n+1) values of a function  $\varphi(x)$ , namely  $y_0, y_1, \ldots, y_n$  for (n+1) values of the argument  $x_0, x_1, \ldots, x_n$ . The values of the argument are equally spaced. We denote the constant difference of the arguments by h. This yields a table of values of the unknown function  $y = \varphi(x)$  for respective values of the argument.

x	<b>x</b> 0	$x_1 = x_0 + h$	$x_2 = x_0 + 2h$		$x_n = x_0 + nh$
y	y <sub>o</sub>	<b>y</b> 1	y 2	••••	<b>y</b> n

Let us set up a polynomial of degree not greater than n that takes on appropriate values for the corresponding values of x. This polynomial will represent the function  $\varphi(x)$  in approximate fashion.

We introduce the following notation:

$$\Delta y_0 = y_1 - y_0, \quad \Delta y_1 = y_2 - y_1, \quad \Delta y_2 = y_3 - y_2, \dots$$
  

$$\Delta^2 y_0 = y_2 - 2y_1 + y_0 = \Delta y_1 - \Delta y_0, \quad \Delta^2 y_1 = \Delta y_2 - \Delta y_1, \dots$$
  

$$\Delta^3 y_0 = y_3 - 3y_2 + 3y_1 - y_0 = \Delta^2 y_1 - \Delta^2 y_0, \dots$$
  

$$\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0$$

These are the so-called first, second,  $\ldots$ , *nth* differences.

We write down a polynomial that takes on the values  $y_0$ ,  $y_1$  for  $x_0$  and  $x_1$ , respectively. This is a polynomial of the first degree,

$$P_{1}(x) = y_{0} + \Delta y_{0} \frac{x - x_{0}}{h}$$
(1)

Indeed

$$P_1(x)|_{x=x_0} = y_0, \quad P_1|_{x=x_1} = y_0 + \Delta y_0 \frac{h}{h} = y_0 + (y_1 - y_0) = y_1$$

Now write down the polynomial that takes on the values  $y_0$ ,  $y_1$ ,  $y_2$  for  $x_0$ ,  $x_1$ ,  $x_2$ , respectively. This is a polynomial of degree 2:

$$P_{2}(x) = y_{0} + \Delta y_{0} \frac{x - x_{0}}{h} + \frac{\Delta^{2} y_{0}}{2!} \frac{x - x_{0}}{h} \left(\frac{x - x_{0}}{h} - 1\right)$$
(2)

Indeed,

$$P_{2}|_{x=x_{0}} = y_{0}, \quad P_{2}|_{x=x_{1}} = y_{1}, \\P_{2}|_{x=x_{0}} = y_{0} + \Delta y_{0} \cdot 2 + \frac{\Delta^{2} y_{0}}{2!} \frac{2h}{h} \left(\frac{2h}{1} - 1\right) = y_{2}$$

A polynomial of degree three will look like this:

$$P_{s}(x) = y_{0} + \Delta y_{0} \frac{x - x_{0}}{h} + \frac{\Delta^{2} y_{0}}{2!} \frac{x - x_{0}}{h} \left(\frac{x - x_{0}}{h} - 1\right) + \frac{\Delta^{3} y_{0}}{3!} \frac{x - x_{0}}{h} \left(\frac{x - x_{0}}{h} - 1\right) \left(\frac{x - x_{0}}{h} - 2\right)$$
(3)

Finally, a polynomial of degree *n* taking on the values  $y_0$ ,  $y_1$ ,  $y_2$ , ...,  $y_n$  for the respective values  $x_0$ ,  $x_1$ ,  $x_2$ , ...,  $x_n$  will be of the form

$$P_{n}(x) = y_{0} + \Delta y_{0} \frac{x - x_{0}}{h} + \frac{\Delta^{2} y_{0}}{2!} \frac{x - x_{0}}{h} \left(\frac{x - x_{0}}{h} - 1\right) + \dots + \frac{\Delta^{n} y_{0}}{n!} \frac{x - x_{0}}{h} \left(\frac{x - x_{0}}{h} - 1\right) \dots \left[\frac{x - x_{0}}{h} - (n - 1)\right]$$
(4)

This can be seen at once by direct substitution. This is the Newton interpolation formula (or the Newton interpolation polynomial).

Actually, the Lagrange polynomial and the Newton polynomial are identical for the given table of values but are written differently, since a polynomial of degree not exceeding n and assuming (n+1) values for (n+1) given values of x is found in unique fashion.

In many cases, Newton's interpolation polynomial is more convenient than Lagrange's interpolation polynomial. The peculiarity of this polynomial lies in the fact that when passing from a polynomial of degree k to one of degree k+1 the first (k+1)terms remain unchanged, and we add one new term, which for all preceding values of the argument is zero.

Note. The Lagrange interpolation formula [see formula (3), Sec. 7.9] and the Newton interpolation formula [see formula (4), Sec. 7.10] are used to determine values of a function on the interval  $x_0 < x < x_n$ . If these formulas are used to find values of the function for  $x < x_0$  (this can be done for small  $|x - x_0|$ ), then we say that the *table is extrapolated backward*. If the value of the function is sought for  $x > x_n$ , then we say that the *table is extrapolated forward*.

#### NUMERICAL DIFFERENTIATION

Suppose the values of some unknown function  $\varphi(x)$  are given in tabular form, say, by the table of Sec. 7.10. It is required to approximate the derivative of the function. The problem is solved by constructing the Lagrange (or Newton) interpolation polynomial and then taking the derivative of that polynomial.

Since equally spaced tables of the argument are ordinarily employed, we will make use of the Newton interpolation formula. Suppose we have three values of the function,  $y_0$ ,  $y_1$ ,  $y_2$ , for the values  $x_0$ ,  $x_1$ ,  $x_2$  of the argument. Then write down polynomial (2) of Sec. 7.10 and differentiate it to get the approximate value of the derivative function on the interval  $x_0 \leq x \leq x_2$ ,

$$\varphi'(x) \approx P'_{2}(x) = \frac{\Delta y_{0}}{h} + \frac{\Delta^{2} y_{0}}{2h} \left( 2 \frac{x - x_{0}}{h} - 1 \right)$$
 (1)

For  $x = x_0$  we have

$$\varphi'(x_0) \approx P'_2(x_0) = \frac{\Delta y_0}{h} - \frac{\Delta^2 y_0}{2h}$$
(2)

If we consider a third-degree polynomial [see (3), Sec. 7.10], then differentiation yields the following expression for the derivative:

$$\varphi'(x) \approx P'_{3}(x) = \frac{\Delta y_{0}}{h} + \frac{\Delta^{3} y_{0}}{2h} \left( 2 \frac{x - x_{0}}{h} - 1 \right) + \frac{\Delta^{3} y_{0}}{2 \cdot 3h} \left[ 3 \left( \frac{x - x_{0}}{h} \right)^{2} - 6 \left( \frac{x - x_{0}}{h} \right) + 2 \right]$$
(3)

In particular, for  $x = x_0$ , we get

$$\varphi'(x_0) \approx P'_3(x) = \frac{\Delta y_0}{h} - \frac{\Delta^2 y_0}{2h} + \frac{\Delta^3 y_0}{3h}$$
 (4)

Using formula (4), Sec. 7.10, we approximate the derivative for  $x = x_0$  as

$$\varphi'(x_0) \approx P'_n(x) = \frac{\Delta y_0}{h} - \frac{\Delta^2 y_0}{2h} + \frac{\Delta^3 y_0}{3h} - \frac{\Delta^4 y_0}{4h} + \dots$$
 (5)

Note that for a function having derivatives, the difference  $\Delta y_0$  is an infinitesimal of the first order,  $\Delta^2 y_0$  is an infinitesimal of the second order,  $\Delta^3 y_0$  is an infinitesimal of the third order, etc., relative to h.

# ON THE BEST APPROXIMATION OF FUNCTIONS BY POLYNOMIALS. CHEBYSHEV'S THEORY

A natural question arises from what was discussed in Secs. 7.9 and 7.10. If a continuous function  $\varphi(x)$  is given on a closed interval [a, b], can this function be represented approximately in the form of a polynomial P(x) to any preassigned degree of accuracy? In other words, is it possible to choose a polynomial P(x)such that the absolute difference between  $\varphi(x)$  and P(x) at all points of the interval [a, b] is less than any preassigned positive number  $\varepsilon$ ? The following theorem, which we give without proof, answers this question in the affirmative.\*

Weierstrass' Approximation Theorem. If a function  $\varphi(x)$  is continuous on a closed interval [a, b], then for every  $\varepsilon > 0$  there exists a polynomial P(x) such that  $|\varphi(x) - P(x)| < \varepsilon$  at all points of the interval.

The Soviet mathematician Academician S. N. Bernstein gave the following method for the direct construction of such polynomials that are approximately equal to the continuous function  $\varphi(x)$  on the given interval.

Let  $\varphi(x)$  be continuous on the interval [0, 1]. We write the expression

$$B_n(x) = \sum_{m=0}^n \varphi\left(\frac{m}{n}\right) C_n^m x^m (1-x)^{n-m}$$

Here,  $C_n^m$  are binomial coefficients,  $\varphi\left(\frac{m}{n}\right)$  is the value of the given function at the point  $x = \frac{m}{n}$ . The expression  $B_n(x)$  is an *n*th degree polynomial called the *Bernstein polynomial*.

<sup>\*</sup> It will be noted that the Lagrange interpolation formula [see (3) Sec. 7.9] cannot yet answer this question. Its values are equal to those of the function at the points  $x_0, x_1, x_2, \ldots, x_n$ , but they may be very far from the values of the function at other points of the interval [a, b].

If an arbitrary  $\varepsilon > 0$  is given, one can choose a Bernstein polynomial (that is, select its degree *n*) such that for all values of *x* on the interval [0, 1], the following inequality will hold:

$$|B_n(x)-\varphi(x)| < \varepsilon$$

It should be noted that consideration of the interval [0, 1], and not an arbitrary interval [a, b], is not an essential restriction of generality, since by changing the variable x = a + t (b - a) it is possible to convert any interval [a, b] into [0, 1]. In this case, an *n*th degree polynomial will be transformed into a polynomial of the same degree.

The creator of the theory of best approximation of functions by polynomials is the Russian mathematician P. L. Chebyshev (1821-1894). In this field, he obtained the most profound results, which exerted a great influence on the work of later mathematicians. Studies involving the theory of articulated mechanisms, which are widely used in machines, served as the starting point of Chebyshev's theory. While studying these mechanisms he arrived at the problem of finding, among all polynomials of a given degree with leading coefficient unity, a polynomial of least deviation from zero on the given interval. He found these polynomials, which subsequently became known as *Chebyshev polynomials*. They possess many remarkable properties and at present are a powerful tool of investigation in many problems of mathematics and engineering.