exhaust all primitive pth roots of unity. Therefore, for a polynomial $f_p(x)$ we can consider the Lagrange resolvent

$$r_1 = \alpha + \beta\alpha^g + \beta^2\alpha^{g^2} + \cdots + \beta^{p-1}\alpha^{g^{p-1}}.$$

Set $r_i = r_1(\alpha, \beta^i)$. It is easy to verify that $r_1(\alpha^g, \beta^i) = \beta^{-i} r_1(\alpha, \beta^i)$. Hence, the quantities r_1^{p-1} and $r_i r_1^{p-1-i}$ do not vary when α is replaced with α^g. As in the case $p = 11$, we can prove with the help of this property and Theorem 6.4.1 that r_1^{p-1} and $r_i r_1^{p-1-i}$ are polynomials in β with integer coefficients. For β an expression in radicals can be obtained by the induction because β is a root of unity of degree $p - 1 < p$. The formula for α is as follows:

$$\alpha = \frac{1}{p-1}\left(r_1 + \frac{r_2 r_1^{p-3}}{r_1^{p-3}} + \frac{r_3 r_1^{p-4}}{r_1^{p-4}} + \cdots\right).$$

This formula gives an unambiguous expression for α after one of the values of $r_1 = \sqrt[p-1]{r_1^{p-1}}$ is chosen.

§6.5. The Abel theorem on the unsolvability in radicals of the general quintic equation

Lagrange's works spurred many geometers (this was the common name for all mathematicians of that time) to begin searching for a proof of the impossibility to solve in radicals the general quintic equation and higher degree equations. In 1788–1813 there appeared several papers of the Italian mathematician **Paulo Ruffini** (1765–1822). Following Lagrange, he considered substitutions of roots of equations and it was he who coined the term the *group of substitutions*. His series of papers culminated in the proof of the theorem on impossibility to solve in radicals the general equations of the fifth and higher degrees.

Regrettably, this proof had an essential gap. Without justification Ruffini assumed that the radicals can be rationally expressed in terms of the roots of the initial equation (cf. Theorem 6.5.4 below).

The Norwegian mathematical genius **Niels Henrik Abel** (1802–1829) was the first to give a complete proof of the theorem on unsolvability of the general quintic equation. He exposed his proof in the memoir *Proof on Impossibility of an Algebraic Solution of General Fifth Degree Equations* published in the first issue of Crelle's journal in 1826.

We say that the equation

$$F(x) = x^n + c_1 x^{n-1} + \cdots + c_n = 0 \tag{5.1}$$

is the *general nth degree equation* if its coefficients $c_1, \ldots, c_n$ are independent variables over the ground field L. In what follows we will assume that $L = \mathbb{Q}$.

Adjoining $c_1, \ldots, c_n$ to $\mathbb{Q}$, we get the field $\Delta = \mathbb{Q}(c_1, \ldots, c_n)$. This field is called the *rationality field of equation* (5.1).

Having attached to Δ the roots $\alpha_1, \ldots, \alpha_n$ of equation (5.1) we get the field $\Delta(F) = \Delta(\alpha_1, \ldots, \alpha_n)$, called the *normal field* of equation (5.1) or the *Galois field* of this equation.

We will say that equation (5.1) is *solvable in radicals* if $\Delta(F)$ is contained in the extension R of Δ obtained after attaching to Δ certain radicals

$$\rho_1 = \sqrt[p]{a_1},\ \rho_2 = \sqrt[q]{a_2},\ \ldots,\ \rho_m = \sqrt[s]{a_m},$$

where

$$a_1 \in \Delta,\ a_2 \in \Delta(\rho_1),\ a_3 \in \Delta(\rho_1, \rho_2),\ \ldots,\ a_m \in \Delta(\rho_1, \ldots, \rho_{m-1}).$$

EXAMPLE. Let $F(x) = x^2 + c_1 x + c_2$. Then $\Delta = \mathbb{Q}(c_1, c_2)$ and $\Delta(F) = \Delta(\sqrt{a_1})$, where $a_1 = c_1^2 - 4c_2 \in \Delta$.

Observe that the exponents $p, q, \ldots, s$ of the radicals $\rho_1, \rho_2, \ldots, \rho_m$ can be assumed to be primes. Indeed, if $p = lm$, then instead of adjoining the radical $\rho_1 = \sqrt[p]{a_1}$ we may consecutively adjoin the radicals $\rho = \sqrt[l]{a_1}$ and $\rho_1 = \sqrt[m]{\rho}$. Therefore, in what follows we will only consider adjoining the radicals with prime exponents.

Suppose that equation (5.1) is solvable in radicals. Adjoin to Δ the primitive roots of unity $\varepsilon_1, \ldots, \varepsilon_m$ whose degrees are equal to the degrees of the radicals $\rho_1, \ldots, \rho_m$, respectively. Denote the obtained field by K.

Since $\Delta \subset K$, it follows that

$$\Delta(F) \subset \Delta(\rho_1, \ldots, \rho_m) \subset K(\rho_1, \ldots, \rho_m).$$

To prove Abel's theorem, we will need four auxiliary statements, Theorems 6.5.1–6.5.4.

6.5.1. THEOREM. *Let p be a prime and k a field of zero characteristic. The polynomial $x^p - a$ is reducible over k if and only if $a = b^p$ for some $b \in k$.*

PROOF. Suppose $x^p - a = f(x)g(x)$, where $f(x)$ and $g(x)$ are polynomials over k. Let ε be a primitive pth root of unity and $\beta = \sqrt[p]{a}$. Then

$$f(x) = x^r + c_1 x^{r-1} + \cdots + c_r = (x - \varepsilon^{n_1}\beta) \cdots (x - \varepsilon^{n_r}\beta).$$

Hence, $\pm\varepsilon^l \beta^r = c_r \in k$, where $l = n_1 + \cdots + n_r$. Since $(\varepsilon^l)^p = 1$, it follows that $(\pm\beta^r)^p = (c_r)^p$, i.e., $a^r = (\pm c_r)^p$. The number p is prime and $1 \leq r = \deg f < p$; hence, $rs + pt = 1$ for certain integers s and t. Therefore, $a = a^{rs}a^{pt} = (\pm c_r a^t)^p = b^p$, where $b = \pm c_r a^t \in k$.

It is also clear that if $a = b^p$, then $x^p - a$ is reducible because it is divisible by $x - b$. □

6.5.2. THEOREM. *Let s be a prime and $a_i \in k = K(\rho_1, \ldots, \rho_{i-1})$. If $\rho_i = \sqrt[s]{a_i} \notin k$, then $\rho_i^l \in k$ if and only if l is divisible by s.*

PROOF. If $l = ns$, then $\rho_i^l = a_i^n \in k$ since $a_i \in k$. Now suppose that $\rho_i^l = a \in k$ and $l = sq + r$, where $0 < r < s$. Then $a = \rho_i^l = (a_i)^q \rho_i^r$ and, therefore, $\rho_i^r = b$, where $b = a\,(a_i)^{-q} \in k$.

Over k, the polynomials $x^s - a_i$ and $x^r - b$ have a common root ρ_i; hence, they have a common divisor whose degree does not exceed $r < s$. In particular, the polynomial $x^s - a_i$ is reducible over k. Theorem 6.5.1 implies that $a_i = b^s$, where $b \in k$. Clearly, $b = \varepsilon\rho_i$, where ε is a primitive root of unity of degree s. Since $\varepsilon \in K \subset k$, it follows that $\rho_i \in k$. Contradiction. □

We may assume that $\rho_1, \ldots, \rho_m$ is a *minimal* sequence of radicals (of prime degrees) required to compute a root α of equation (5.1), i.e., any other such sequence contains at least m radicals. In what follows we will only consider minimal sequences of radicals. Under this assumption the following statement holds.

6.5.3. THEOREM. *Let $\rho_1, \dots, \rho_m$ be a minimal sequence of radicals needed to compute a root α of equation* (5.1). *Then α can be represented in the form*

$$\alpha = u_0 + \rho + u_2\rho^2 + \cdots + u_{s-1}\rho^{s-1},$$

where s is the degree of ρ_m, $\rho = \sqrt[s]{a}$, $a \in k = K(\rho_1, \dots, \rho_{m-1})$ and $u_i \in k$.

PROOF. Since $\alpha \in K(\rho_1, \dots, \rho_m) = k(\rho_m)$ and $\rho_m^s \in k$, we have

$$\alpha = b_0 + b_1\rho_m + b_2\rho_m^2 + \cdots + b_{s-1}\rho_m^{s-1}, \tag{5.2}$$

where $b_i \in k$. The only difficulty is to ensure that $b_1 = 1$. By the assumption, $\alpha \notin k$ so that at least one of the numbers $b_1, \dots, b_{s-1}$ is nonzero. Let $b_l \neq 0$ for some l such that $1 \leq l < s$. Set $\rho = b_l\rho_m^l$. Since s is a prime, $ul + vs = 1$ for certain integers u and v. Moreover, we have

$$\rho^u = b_l^u \rho_m^{ul} = b_l^u \rho_m^{1-vs} = b_l^u a^{-v} \rho_m,$$

i.e., $\rho_m = c\rho^u$, where $c = b_l^{-u}a^v \in k$. Since $\rho_m \notin k$, it follows that $\rho \notin k$. It is also clear that $\rho^s = b_l^s \rho_m^{ls} = b_l^s a^l \in k$.

In (5.2) replace ρ_m with $c\rho^u$ taking into account that $b_l\rho_m^l = \rho$. As a result, we get

$$\alpha = b_0 + b_1 c\rho^u + b_2 c^2 \rho^{2u} + \cdots + \rho + \cdots + b_{s-1}c^{s-1}\rho^{(s-1)u}. \tag{5.3}$$

Theorem 6.5.2 implies that $\rho^t \in k$ if and only if t is divisible by s. Since u and s are relatively prime, the elements $1, \rho^u, \rho^{2u}, \dots, \rho^{(s-1)u}$ are linearly independent over k and the set of these elements coincides with the set $1, \rho, \rho^2, \dots, \rho^{s-1}$ (perhaps, ordered differently). Thus, formula (5.3) gives the required expression for α:

$$\alpha = b_0 + \rho + b_2'\rho^2 + \cdots + b_{s-1}'\rho^{s-1}. \qquad \square$$

6.5.4. THEOREM. *The minimal sequence of radicals $\rho_1, \dots, \rho_m$ necessary to calculate a root α of polynomial* (5.1) *can be selected so that $\rho_1, \dots, \rho_m$ are polynomials over K of the roots $\alpha_1, \dots, \alpha_n$ of polynomial* (5.1).

PROOF. Start with an arbitrary minimal sequence $\rho_1, \dots, \rho_m$. By Theorem 6.5.3 we can replace ρ_m with a radical ρ of the same degree s so that

$$\alpha = u_0 + \rho + u_1\rho^2 + \cdots + u_{s-1}\rho^{s-1},$$

where $u_i \in k = K(\rho_1, \dots, \rho_{m-1})$ and $\rho^s = a \in k$. Let us show that for any root ξ of the polynomial $x^s - a$

$$\alpha(\xi) = u_0 + \xi + u_1\xi^2 + \cdots + u_{s-1}\xi^{s-1}$$

is a root of polynomial (5.1). Substitute $x = \alpha(\xi)$ in the polynomial

$$F(x) = x^n + c_1x^{n-1} + \cdots + c_n.$$

Taking into account that $\xi^s = a \in k$ we get an expression of the form

$$b_0 + b_1\xi + \cdots + b_{s-1}\xi^{s-1},$$

where $b_i \in k$. The polynomials $x^s - a$ and $b_0 + b_1x + \cdots + b_{s-1}x^{s-1}$ have a common root ρ; hence, they have a common divisor over k. By Theorem 6.5.1 the polynomial $x^s - a$ is irreducible over k; hence, $b_0 = b_1 = \cdots = b_{s-1} = 0$. This means that if ξ

is a root of the polynomial $x^s - a$, then $\alpha(\xi)$ is a root of polynomial (5.1). Let ε be a primitive root of unity of degree s. Then $\xi = \varepsilon^r \rho$; hence,

$$\alpha_{r+1} = u_0 + \varepsilon^r \rho + u_2 \varepsilon^{2r} \rho^2 + \cdots + u_{s-1} \varepsilon^{(s-1)r} \rho^{s-1}$$

for $r = 0, 1, \ldots, s-1$ are roots of polynomial (5.1).

For example, for $s = 3$ we get

$$\begin{aligned} \alpha_1 &= u_0 + \rho + u_2 \rho^2, \\ \alpha_2 &= u_0 + \varepsilon\rho + u_2 \varepsilon^2 \rho^2, \\ \alpha_3 &= u_0 + \varepsilon^2 \rho + u_2 \varepsilon \rho^2. \end{aligned}$$

Since $1 + \varepsilon + \varepsilon^2 = 0$, we have

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 3u_0, \\ \alpha_1 + \varepsilon^{-1}\alpha_2 + \varepsilon^{-2}\alpha_3 &= 3\rho, \\ \alpha_1 + \varepsilon^{-2}\alpha_2 + \varepsilon^{-1}\alpha_3 &= 3u_2\rho^2. \end{aligned}$$

Therefore, $\rho = \frac{1}{3}\left(\alpha_1 + \varepsilon^2\alpha_2 + \varepsilon\alpha_3\right)$. For $s > 3$ we get more cumbersome formulas but the arguments remain the same. The proof of the theorem for the last radical ρ_m is completed.

Let us now turn to ρ_{m-1}. We have shown above (for $s = 3$) that the expressions $u_0, \rho, u_2\rho^2, \ldots, u_{s-1}\rho^{s-1}$ can be polynomially expressed in terms of roots $\alpha_1, \ldots, \alpha_n$ of polynomial (5.1). Moreover, they lie in the field $K(\rho_1, \ldots, \rho_{m-1})$, so that each of the values indicated can be represented in the form

$$v_0 + v_1\rho_{m-1} + v_2\rho_{m-1}^2 + \cdots + v_{t-1}\rho_{m-1}^{t-1},$$

where $v_i \in K(\rho_1, \ldots, \rho_{m-2})$. The sequence of radicals $\rho_1, \ldots, \rho_m$ is minimal, so that the equations $v_1 = v_2 = \cdots = v_{t-1} = 0$ cannot be simultaneously satisfied for all the quantities because otherwise we could have excluded ρ_{m-1}. Therefore, there exists a relation of the form

$$v_0 + v_1\rho_{m-1} + v_2\rho_{m-1}^2 + \cdots + v_{t-1}\rho_{m-1}^{t-1} = r(\alpha_1, \ldots, \alpha_n),$$

where $v_i \in K(\rho_1, \ldots, \rho_{m-2})$, not all elements $v_1, \ldots, v_{t-1}$ vanish and $r(\alpha_1, \ldots, \alpha_n)$ is a polynomial over K. Consider the polynomial

$$G(x) = \prod \left(x - r(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(n)})\right),$$

where the product runs over all the permutations $\sigma \in S_n$. The coefficients of G are the symmetric polynomials of the roots of polynomial (5.1), so that they can be polynomially expressed in terms of the coefficients of polynomial (5.1). Thus, G is a polynomial over K and

$$\beta = v_0 + v_1\rho_{m-1} + \cdots + v_{t-1}\rho_{m-1}^{t-1}$$

is a root of this polynomial. It is also clear that the root β can be expressed by means of the radicals (with the help of the sequence of radicals $\rho_1, \ldots, \rho_{m-1}$). By Theorem 6.5.3 replacing ρ_{m-1} with the radical ρ' of the same degree we may assume that $v_1 = 1$. We can now apply to ρ' the same arguments as we applied to ρ. Iterating the arguments for ρ_{m-2}, and so on, down to ρ_1 completes the proof. □

Now we can pass to the proof of Abel's theorem proper.

6.5.5. THEOREM (Abel). *For $n \geq 5$ it is impossible to express the roots of the general nth degree polynomial in radicals.*

PROOF. Suppose that a certain root α_1 of the general nth degree polynomial

$$x^n + c_1 x^{n-1} + c_2 x^{n-2} + \cdots + c_n$$

can be expressed in radicals. Then by Theorems 6.5.1–6.5.4 there exists an expression of α_1 in radicals of the following particular form. The root α_1 is obtained by consecutively adjoining the radicals $\rho_1, \ldots, \rho_m$ of prime degrees to the ground field, and these radicals, in their turn, are polynomials in the roots $\alpha_1, \ldots, \alpha_n$ of the initial polynomial. More precisely, let $\varepsilon_1, \ldots, \varepsilon_m$ be primitive roots of unity whose degrees are equal to the degrees of the radicals $\rho_1, \ldots, \rho_m$, respectively, $\Delta = \mathbb{Q}(c_1, \ldots, c_n)$, and

$$K = \Delta(\varepsilon_1, \ldots, \varepsilon_m) = \mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m, c_1, \ldots, c_n).$$

Then α_1 can be polynomially expressed over K in terms of $\rho_1, \ldots, \rho_m$, i.e.,

$$\alpha_1 = r(\rho_1, \ldots, \rho_m, c_1, \ldots, c_n),$$

where r is a polynomial over $\mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m)$. In their turn, $\rho_1, \ldots, \rho_m$ can be polynomially expressed over K in terms of $\alpha_1, \ldots, \alpha_n$, i.e.,

$$\rho_i = r_i(\alpha_1, \ldots, \alpha_n, c_1, \ldots, c_n),$$

where r_i is a polynomial over $\mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m)$. Since we deal with the general polynomial of degree n, we may assume that $\alpha_1, \ldots, \alpha_n$ are independent variables and $c_1, \ldots, c_n$ are (up to a sign) the elementary symmetric polynomials of $\alpha_1, \ldots, \alpha_n$.

Let us show that for $n \geq 5$ the assumption on solvability in radicals of the general algebraic equation of degree n leads to a contradiction. To this end consider the permutation

$$T = \begin{pmatrix} 123456 \ldots n \\ 234516 \ldots n \end{pmatrix}$$

that cyclically permutes the first 5 elements, the others being fixed. Let us prove that under the action of T on the roots $\alpha_1, \ldots, \alpha_n$ the first radical ρ_1 does not change. Since

$$\rho_1 = r_1(\alpha_1, \ldots, \alpha_n, c_1, \ldots, c_n) = \sqrt[p]{a_1},$$

where a_1 is a polynomial of $c_1, \ldots, c_n$ over the field $\mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m)$, the equation $\rho_1^p = a_1$ can be considered as a relation of the form

$$\varphi(\alpha_1, \ldots, \alpha_n, c_1, \ldots, c_n) = 0,$$

where φ is a polynomial over $\mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m)$.

Let us show that any relation of this form is preserved under any permutation of the roots $\alpha_1, \ldots, \alpha_n$. Let $\beta_1 = \alpha_{i_1}, \ldots, \beta_n = \alpha_{i_n}$, where $i_1, \ldots, i_n$ is a permutation of the numbers $1, 2, \ldots, n$. Then

$$\varphi(\beta_1, \ldots, \beta_n, d_1, \ldots, d_n) = 0,$$

where $d_i = c_i(\beta_1, \ldots, \beta_n)$. Clearly, $d_i = c_i(\alpha_1, \ldots, \alpha_n) = c_i$ because the functions c_i are symmetric. Hence,

$$\varphi(\alpha_{i_1}, \ldots, \alpha_{i_n}, c_1, \ldots, c_n) = 0.$$

Thus, the relation $\rho_1^p = a_1$ is preserved under the action of T on the roots $\alpha_1, \ldots, \alpha_n$, i.e., $T(\rho_1^p) = T(a_1)$. Clearly, $T(\rho_1^p) = T(\rho_1)^p$. Since a_1 only depends on the symmetric functions of roots, $T(a_1) = a_1$. Therefore, $T(\rho_1) = \varepsilon_1^\lambda \rho_1$ and $T^m(\rho_1) = \varepsilon_1^{m\lambda}\rho_1$. But $T^5 = I$ is the identity substitution, hence, $\varepsilon_1^{5\lambda}\rho_1 = T^5(\rho_1) = \rho_1$, i.e., $\varepsilon_1^{5\lambda} = 1$.

Let us now turn to the substitutions

$$U = \begin{pmatrix} 123456\ldots n \\ 124536\ldots n \end{pmatrix}, \quad V = \begin{pmatrix} 123456\ldots n \\ 231456\ldots n \end{pmatrix}.$$

It is easy to verify that $U^3 = V^3 = 1$; hence, $U(\rho_1) = \varepsilon_1^\mu \rho_1$ and $V(\rho_1) = \varepsilon_1^\nu \rho_1$, and $\varepsilon_1^{3\mu} = \varepsilon_1^{3\nu} = 1$. Moreover, $UV = T$; hence,

$$T(\rho_1) = VU(\rho_1) = \varepsilon_1^{\mu+\nu}\rho_1.$$

Hence, $\varepsilon_1^\lambda = \varepsilon_1^{\mu+\nu}$ so that $\varepsilon_1^\lambda = \varepsilon_1^{6\lambda}\varepsilon_1^{-5\lambda} = \varepsilon_1^{6(\mu+\nu)} = 1$ because $\varepsilon_1^{5\lambda} = \varepsilon_1^{6\mu} = \varepsilon_1^{6\nu} = 1$. As a result we get $T(\rho_1) = \rho_1$.

Passing consecutively to the radicals $\rho_2, \ldots, \rho_m$ we similarly get $T(\rho_i) = \rho_i$ for $i = 2, \ldots, m$.

Since $\rho_i = r_i(\alpha_1, \ldots, \alpha_n, c_1, \ldots, c_n)$, it follows that the equation

$$\alpha_1 = r(\rho_1, \ldots, \rho_m, c_1, \ldots, c_n)$$

can be considered as a relation between $\alpha_1, \ldots, \alpha_n, c_1, \ldots, c_n$ over $\mathbb{Q}(\varepsilon_1, \ldots, \varepsilon_m)$. This relation is preserved under the action of T, i.e.,

$$T(\alpha_1) = r(T(\rho_1), \ldots, T(c_n)) = r(\rho_1, \ldots, c_n)$$

since $T(c_i) = c_i$ and $T(\rho_i) = \rho_i$. Therefore, $T(\alpha_1) = \alpha_1$. On the other hand, by the definition of T we get $T(\alpha_1) = \alpha_2$; hence, $\alpha_1 = \alpha_2$. The relation $\alpha_1 = \alpha_2$ contradicts the independence of the roots of the general equation. □

§6.6. The Tschirnhaus transformations. Quintic equations in Bring's form

In 1683 in the journal *Acta Eruditorum* **E.W. von Tschirnhaus**[1] (1651–1708) published a method for transformation of algebraic equations which, Tschirnhaus believed, enabled one to solve in radicals the equation of any degree. **Leibniz** immediately announced that Tschirnhaus' claim on the universality of this transformation was not valid. The catch is that in order to solve a quintic equation with the help of the Tschirnhaus transformations one has to solve an equation of degree 24.

Still, the Tschirnhaus transformation has important applications. For example, with its help any quintic equation without multiple roots can be reduced to the form $y^5 + 5y = a$ and in the process we only have to solve equations of degrees 2 and 3. In Chapter 7 we will show that equations of such a form can then be solved using theta functions.

[1]The mathematicians often write Tschirnhausen, but as is clear from the works of historians of mathematics, the correct spelling is *Tschirnhaus*. (Regrettably, his original works were inaccessible for us.) *The authors.*