Galois Theory

Arnau Mas

2019

These are notes gathered during the subject *Teoria de Galois* as taught by Francesc Perera between September 2019 and January 2020.

Chapter 1

Preliminaries

1.1 The solution of low degree polynomial equations

It is surely well-known to any aspiring mathematician that there exist no general formulas for the solutions of polynomial equations of degree five and higher. This implies, of course, that such formulas exist for equations of degree fourth and lower. Indeed, the solution of linear equations is trivial and the quadratic formula should be more than well-known by this point. In this section we present a derivation of the solutions of both the quadratic and cubic equations.

1.1.1 The quadratic equation

First, note that we can, without loss of generality, assume that we are working with a monic equation since we may always divide through by the leading coefficient to obtain an equation with the same solutions and with leading coefficient 1. Thus, we are trying to solve $x^2 + bx + c = 0$. The standard method is completing the square, that is to write $x^2 + bx + c$ as a square, and one achieves so by adding and substracting $\frac{b^2}{4}$:

$$x^2 + bx + c = x^2 + bx + \frac{b^2}{4} - \frac{b^2}{4} + c = \left(x + \frac{b}{2}\right)^2 - \frac{b^2}{4} + c.$$

Then the solutions to the original equation must satisfy

$$\left(x + \frac{b}{2}\right)^2 = \frac{b^2}{4} - c.$$

If the term on the right is not a square in the field we are working over then there are no solutions in that field. On the other hand, if it is a square then it has two square roots and the solutions to the original equation are

$$x = -\frac{b}{2} \pm \frac{1}{2}\sqrt{b^2 - 4c},$$

which is the well known quadratic formula.

1.1.2 The cubic equation

Less well-known is the formula for the solutions of the cubic equation. Whereas the quadratic formula had been known to the greeks and babylonians, the cubic formula was discovered later during the fifteenth century. There were several italian mathematicians involved in its discovery: Cardano, Ferrari and del Ferro among others. The question of the original discoverer is a contemptious matter.

The first step in the solution is a change of variables to eliminate the quadratic term. If $x = y - \frac{1}{3}b$ then the original (monic) polynomial becomes

$$\begin{aligned} x^3 + bx^2 + cx + d &= y^3 - by^2 + \frac{1}{3}b^2y - \frac{1}{27}b^3 + by^2 - \frac{2}{3}b^2y + \frac{1}{9}b^3 + cy - \frac{1}{3}bc + d \\ &= y^3 + \left(c - \frac{1}{3}b^2\right)y + \frac{2}{27}b^3 - \frac{1}{3}bc + d. \end{aligned}$$

Therefore we only need to be able to solve cubics of the form $x^3 + px + q = 0$.

The basic trick is similar to completing the square. We have the identity

$$(u+v)^3 = u^3 + 3u^2v + 3uv^2 + v^3 = u^3 + 3uv(u+v) + v^3,$$

and rearranging we obtain $(u+v)^3 - 3uv(u+v) - u^3 - v^3 = 0$. One then notices that there are cubic and linear terms in $u+v$ but no quadratic terms. Then one tries to solve for u and v to then obtain x as $u+v$. u and v must satisfy $-3uv = p$ and $-u^3 - v^3 = q$. Multiplying this second condition by u^3 we get

$$u^6 + qu^3 + u^3v^3 = 0,$$

and using the fact that $uv = -\frac{1}{3}p$ we arrive at

$$u^6 + qu^3 - \frac{p^3}{27} = 0,$$

which is quadratic in u^3. If we instead had multiplied through by v^3 we would have arrived to the same equation for v^3 instead.

Up to now nothing we have done relied on any additional assumption on the field have been working over. From this point, however, the nature of the solutions will depend on the behaviour of radicals in the field in question. We will assume we are working in $\mathbb{C}$. We can then solve for u^3 and v^3 to find

$$\begin{aligned} u^3 &= -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \\ v^3 &= -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}. \end{aligned}$$

The ambiguity with the signs is eliminated due to the fact that $u^3 + v^3 = -q$ so we find the only good options are those in which the signs of the square root terms are opposite, so that they will cancel when added. Since we only care about the sum of u and v we might as well choose

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

and

$$v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}.$$

There are three possibilities for u and three for v. Indeed, every complex number has three roots and if a is one of them then so are ωa and $\omega^2 a$ where $\omega = e^{\frac{2\pi}{3}i}$. Not every combination of them leads to a solution of the cubic —if it were so we would have more than three roots and a cubic polynomial can only have three roots— since they are constrained by the relation $3uv = -p$. So, once we find u and v that satisfy this then so will ωu and $\omega^2 v$, as well as $\omega^2 u$ and ωv since $\omega^3 = 1$.

All together, one of the solutions to the cubic $x^3 + px + q = 0$ is given by

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

which is known as Cardano's formula. If we undo the change of variable to eliminate the quadratic term and use $p = c - \frac{1}{3}b^2$ and $q = \frac{2}{27}b^3 - \frac{1}{3}bc + d$ then we obtain the cubic formula in all of its glory:

$$x = -\frac{b}{3} + \sqrt[3]{\left(-\frac{b^3}{27} + \frac{bc}{6} - \frac{d}{2}\right) + \sqrt{\left(-\frac{b^3}{27} + \frac{bc}{6} - \frac{d}{2}\right)^2 + \left(\frac{c}{3} - \frac{b^2}{9}\right)^3}} + \sqrt[3]{\left(-\frac{b^3}{27} + \frac{bc}{6} - \frac{d}{2}\right) - \sqrt{\left(-\frac{b^3}{27} + \frac{bc}{6} - \frac{d}{2}\right)^2 + \left(\frac{c}{3} - \frac{b^2}{9}\right)^3}}.$$

1.2 Polynomial rings

The study of polynomial rings is particularly important in Galois theory since they play an important role in many of the constructions and definitions of the theory. Of special relevance is the study of their quotient rings.

1.2.1 The universal property of polynomial rings

Given a ring[1] R its polynomial ring $R[x]$ can be defined in various ways. Most commonly, the elements of $R[x]$ are said to be "formal sums" of the form

$$\sum_{k=1}^{n} a_k x^k$$

where the a_k are elements of R and x is refered to as an "indeterminate" or some other similarly ambiguous term. This definition may feel imprecise to the more technically inclined reader. A more exact definition of $R[x]$ is as the set of sequences of elements of R with finite suport. The sum is defined pointwise and the product is defined in a convoluted manner which correspond to the way one would multiply

[1] We will always assume that we are dealing with commutative rings with identity unless otherwise stated.

polynomials with repeated application of the distribuitive law. Then there is a canonical inclusion $R \hookrightarrow R[x]$ by way of $a \mapsto (a, 0, \cdots)$. And if we define x to be the sequence $(0, 1, \cdots)$ then we recover the more standard presentation of $R[x]$.

This discussion, however, is about what a programmer would call the implementation details and it misses the bigger picture. How $R[x]$ is constructed is not really what is relevant here. It is much more illuminating to think about what we want out of $R[x]$ instead. For one, $R[x]$ should contain R. We could require $R \subseteq R[x]$, but let's be more general and allow for an injective morphism $\iota\colon R \hookrightarrow R[x]$ that picks out a copy of R inside $R[x]$. The other important aspect of $R[x]$ is the indeterminate. The way to formalize it is with what is known as a universal property: for any morphism $\phi\colon R \to S$ and distinguished element $s \in S$ there is a unique morphism $\tilde{\phi}\colon R[x] \to S$ such that $\tilde{\phi} \circ \iota = \phi$ and $\tilde{\phi}(x) = s$. That is, $\tilde{\phi}$ must agree with ϕ on R and it must send x to s. This does indeed uniquely determine $\tilde{\phi}$. It can be shown that this determines $R[x]$ up to unique isomorphism, meaning there is a unique isomorphism between any two rings that satisfy the universal property. So you can construct $R[x]$ in whatever way you like so long as the result satisfies the universal property.

One last remark about polynomial rings in many variables: once we have defined the polynomial ring of a ring $\mathbb{R}$, we can then proceed inductively to define the polynomial ring on n variables as $R[x_1, \cdots, x_n] = R[x_1, \cdots, x_{n-1}][x_n]$. Polynomial rings in more than one variables also satisfy a universal property, which is essentially the same as before except now we have to specify where $\tilde{\phi}$ sends all of the x_i.

This universal property is not just of theoretical importance, it is also extermely practical. Indeed, it provides a very quick way of specifying morphisms on a polynomial ring. All you need is to specify how it acts on the coeficients and where it sends the indeterminate and you're done.

Example 1.1. These are various examples of morphisms defined on a polynomial ring making use of the universal property.

(*i*) For any element $\alpha \in R$ we can define the evaluation morphism $\mathrm{ev}_\alpha\colon R[x] \to R$ such that $\mathrm{ev}_\alpha|_R = \mathrm{id}_R$ and $\mathrm{ev}_\alpha(x) = \alpha$. That is, simply evaluate a polynomial on the element $\alpha \in R$. The element we evaluate at need not be an element of R, in fact it can be an element of any ring which contains R.

(*ii*) A trick that is often used when working with polynomials is a change of variable. This idea can be made formal in terms of an automorphism of $R[x]$. Say we wanted to make the change $y = x + 1$, or $x = y - 1$. This amounts to defining a morphism $\phi\colon R[x] \to R[x]$ such that $\phi|_R = \mathrm{id}_R$ and $\phi(x) = y - 1$. ϕ does not move the coeficients and changes x to $y - 1$. More generally, we could perform a change of the form $\phi(x) = ay + b$. In order for ϕ to be an isomorphism, a must be invertible. Indeed, ϕ^{-1} sends x to $a^{-1}(x - b)$. Then, when showing that a certain polynomial is irreducible, for instance, we can do any change of variable we please and rest assured that the resulting polynomial will be irreducible if and only if the original one was irreducible, for irreducibility is preserved under isomorphism.

(*iii*) Any permutation $\sigma \in \mathfrak{S}_n$ induces an isomorphism on $R[x, \ldots, x_n]$ by permuting the variables according to σ. Indeed, let ϕ_σ be the unique morphism that is

the identity on R and such that $\phi_\sigma(x_i) = x_{\sigma(i)}$. You can check that $\phi_\sigma \circ \phi_\tau = \phi_{\sigma\circ\tau}$. And as a corollary $\phi_\sigma^{-1} = \phi_{\sigma^{-1}}$. This is in fact an action of the symmetric group $\mathfrak{S}_n$ on $R[x_1, \cdots, x_n]$. The polynomials invariant under this action, $R[x_1, \cdots, x_n]^{\mathfrak{S}_n}$ are known as the *symmetric polynomials*.

$\bigtriangledown$