HANDOUT – HOW TO SOLVE POLYNOMIALS OF DEGREE > 4

GREGORIO MALAJOVICH

1. Introduction

A *polynomial* is an expression of the form

$$f(x) = f_d x^d + f_{d-1} x^{d-1} + \cdots + f_1 x + f_0$$

with $f_d \neq 0$. The integer $d \geq 1$ is called the *degree* of the polynomial. If the coefficients f_j are all real numbers, we say that f is a *real* polynomial. If the coefficients are all complex numbers, we say that f is a *complex* polynomial. In this talk, all the polynomials are real or complex.

Remark 1. Every real polynomial is also a complex polynomial. An example of complex polynomial which is not a real polynomial is $x - \sqrt{-1}$.

A *root* of a polynomial is a real or complex number ζ such that $f(\zeta) = 0$. Real polynomials do not need to have real roots. For instance $f(x) = x^2 + 1$ has values ≥ 1 for all $x \in \mathbb{R}$. However, it admits a pair of complex *conjugate* roots, $+i$ and $-i$ where $i = \sqrt{-1}$.

Problem 1. If $f(x)$ is a real polynomial and $\zeta = a + bi$ is a root of f, show that its conjugate $\bar{\zeta} = a - bi$ is also a root of f.

Problem 2. Show that a complex polynomial of degree d has at most d distinct roots.

We will admit the following fact:

Theorem 1 (Foundamental Theorem of Algebra)**.** *Every complex polynomial admits at least one complex root.*

Solving a polynomial means finding all its roots. This is a very traditional mathematical problem, while the meaning of 'finding' could change over time. The ancient Sumerians knew around 2000 BC how to find the real roots of a quadratic equation through the quadratic formula. Formulas for the cubic and the quartic equation appeared in the XVI-th century, thanks to the works of Scipione del Ferro, Niccoló Tartaglia, Lodovico Ferrari, Gerolamo Cardano, Vieta and others. This was at the time an extremely competitive endeavor. To give an idea of the difficulty, those mathematicians did not know about negative numbers and had to work with a misterious 'minus minus' (later known as the imaginary number $i = \sqrt{-1}$). There was no symbolic notation before Vieta.

At that time it was common for mathematicians to challenge each other with problems. Vieta is said to have answered to such a challenge on the spot (solving a certain polynomial of degree 45).

At that time, 'solving' a polynomial meant to give a formula for its roots, using the 4 elementary operations and *radicals* (square root, cubic root, etc...). In 1824, Niels Abel proved that for polynomials of degree > 4, no such general formula could possibly exist.

Date: Berkeley Mathematical Circle, Oct 7, 2014.

Departamento de Matemática Aplicada, Instituto de Matemática, Universidade Federal do Rio de Janeiro. Caixa Postal 68530 21941-909 Rio de Janeiro - RJ - Brasil. url: www.labma.ufrj.br/~gregorio.

In the 1830', the Berlin Academy of Science offered a prize for the best method to solve polynomials. At that time, solving meant finding the roots within a prescribed number of decimal places. Computations could be performed by large groups of people (the 'computers') working with pen, paper, mathematical tables and eventually with logarithmic paper.

The prize was awarded to Karl Heinrich Gräffe (1799-1873) for a paper published in 1837. The priority for the discovery is doubtful, as similar methods were proposed by Germinal Dandelin (1794-1847) in 1826 and Nicolai Ivanovich Lobachevskii (1792-1856) in 1834. I will refer to this method as Gräffe iteration for short, but the three authors deserve credit.

It is mainly about their work that I will speak tonight.

2. Basic facts about polynomials

Problem 3. If f is a complex polynomial of degree d and g is a complex polynomial of degree $e < d$, then there exist polynomials q and r so that $f(x) \equiv g(x)q(x) + r(x)$. Furthermore, if f and g are real polynomials, then q and r are real polynomials also.

Problem 4. Every complex polynomial can be written as

$$f(x) = f_d(x - \zeta_1)(x - \zeta_2) \cdots (x - \zeta_d)$$

Problem 5. Let f has degree 3 with $f_3 = 1$. Write f_2, f_1, f_0 in terms of $\zeta_1, \dots, \zeta_d$.

We can generalize this formula by defining the elementary symmetric functions in d variables:

$$\begin{array}{rcl}
\sigma_0(z_1, \dots, z_d) & = & 1 \\
\sigma_1(z_1, \dots, z_d) & = & z_1 + z_2 + \cdots + z_d \\
\sigma_2(z_1, \dots, z_d) & = & z_1 z_2 + z_1 z_3 + \cdots + z_{d-1} z_d \\
 & \vdots & \\
\sigma_d(z_1, \dots, z_d) & = & z_1 z_2 \cdots z_d
\end{array}$$

In general, the *k-th elementary symmetric function on d variables* is the sum of the products of every subset of k variables:

$$\sigma_k(z_1, \dots, z_d) = \sum_{\substack{S \subseteq \{1, \dots, d\} \\ \#S = k}} \prod_{j \in S} z_j .$$

Problem 6. Verify that

$$f_j = (-1)^{d-j} \sigma_{d-j}(\zeta_1, \dots, \zeta_d)$$

Problem 7 (Hard. For those that are bored). A function in d variables $x_1, \dots x_d$ is said to be symmetric if it does not change when permuting the variables. Show that any symmetric polynomial variables $x_1, \dots, x_d$ can be written as a polynomial in $z_1 = \sigma_1(x_1, \dots, x_d)$, ..., $z_d = \sigma_d(x_1, \dots, x_d)$. For instance, $x_1^2 + x_2^2 = z_1^2 - 2z^2$.

3. Dandelin-Gräffe-Lobachevskii iteration

Gräffe iteration takes a degree d polynomial $f(x)$ into another degree d polynomial $(Gf)(x)$ given by

$$(Gf)(x) = (-1)^d f(\sqrt{x}) f(-\sqrt{x})$$

Problem 8. Show that $(Gf)(x)$ is really a polynomial in x. If the roots of f are $\zeta_1, \dots, \zeta_d$, then the roots of Gf are $\zeta_1^2, \dots, \zeta_d^2$.

This is why the Gräffe iteration is also known as the root-squaring iteration. But the interest of this arises because there is an easy formula to compute iterates.

Problem 9. Find a formula for the j-th coefficient of Gf in terms of the coefficients of f.

People use to compute a few iterates of f: $G(f)$, then $G(Gf) = (G \circ G)(f) f$, then $G(G(G(f))) = (G \circ G \circ G)(f)$ and so on. We will write

$$G^N f(x) = \left((\underbrace{G \circ G \circ \cdots \circ G}_{N \text{ times}})(f) \right)(x)$$

So if the roots of f are $\zeta_1, \ldots, \zeta_d$, the roots of g are $\zeta_1^{2^N}, \ldots, \zeta_d^{2^N}$.

To understand how to recover the roots of f from the *coefficients* of g, we will assume for the time being that

$$|\zeta_1| < |\zeta_2| < \cdots < |\zeta_d|.$$

Let $\rho = \max_{1 \le k < d} \frac{|\zeta_k|}{|\zeta_{k+1}|} < 1$. We write down the coefficients of g:

$$\begin{aligned}
g_d &= 1 \\
g_{d-1} &= -\left(\zeta_1^{2^N} + \zeta_2^{2^N} + \cdots + \zeta_d^{2^N} \right) \\
g_{d-2} &= \zeta_1^{2^N} \zeta_2^{2^N} + \zeta_1^{2^N} \zeta_3^{2^N} + \cdots + \zeta_{d-1}^{2^N} \zeta_d^{2^N} \\
&\vdots \\
g_0 &= (-1)^d \zeta_1^{2^N} \zeta_2^{2^N} \cdots \zeta_d^{2^N}
\end{aligned}$$

In general, $g_j = (-1)^{d-j} \sigma_{d-j}(\zeta_1^{2^N}, \cdots, \zeta_d^{2^N})$. We will approximate

$$\zeta_{d-j}^{2^N} \simeq -\frac{g_{j-1}}{g_j}.$$

Problem 10.

$$g_j = \zeta_{d-j+1}^{2^N} \cdots \zeta_d^{2^N} (1 + \delta_j)$$

where

$$|\delta_j| < \rho^{2^N} \binom{d}{j} < \rho^{2^N} 2^d$$

Suppose we want to compute g_j up to 3 decimal places. Then we need to guarantee $|\delta_j| < 10^{-4}$.

Problem 11. How many steps N of Gräffe iteration are sufficient to guarantee that $|\delta_j| < 10^{-4}$?

Problem 12. If we know $|\zeta|^{2^8}$ up to precision 10^{-4}, and if we can compute logarithms and exponentials with precision 10^{-6}, how precisely do we know $|\zeta|$?

I did not explain how to find ζ, just $|\zeta|$. One possibility is to solve the polynomials $f(x)$ and $f(x - \epsilon)$. Then one recovers $|\zeta|$ and $|\zeta + \epsilon|$.

Problem 13. How can one find ζ?

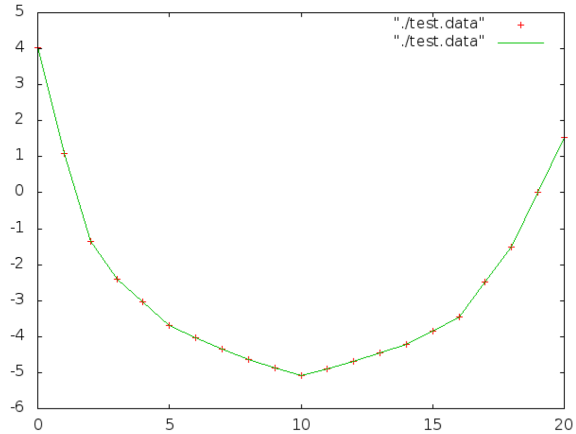

FIGURE 1. Newton diagram of a random real polynomial of degree 20

4. Real polynomials

From now on, we will always order the roots of a polynomial by non-decreasing modulus. That is, we assume that

$$|\zeta_1| \le |\zeta_2| \le \cdots \le |\zeta_d|$$

Also, we assume that if $|\zeta_j| = |\zeta_{j+1}|$ then $\zeta_{j+1} \ne \zeta_j = \bar{\zeta}_{j+1}$ and $|\zeta_{j-1}|\rho < |\zeta_j| = |\zeta_{j+1}| < \frac{1}{\rho}|\zeta_{j+2}|$ for some $0 \le \rho < 1$.

To find ζ_j, we look again at the coefficients of $g = G^N f$:

$$\begin{aligned} g_{d-j-1} &= (-1)^{j+1}\zeta_1^{2^N} \cdots \zeta_{j-1}^{2^N}\zeta_j^{2^N}\zeta_{j+1}^{2^N} + \cdots \\ g_{d-j} &= (-1)^{j}\zeta_1^{2^N} \cdots \zeta_{d-j+2}^{2^N}(\zeta_j^{2^N} + \bar{\zeta}_j^{2^N}) + \cdots \\ g_{d-j+1} &= (-1)^{j-1}\zeta_1^{2^N} \cdots \zeta_{d-j+2}^{2^N} + \cdots \end{aligned}$$

Problem 14. How to find $\zeta_j^{2^N}$ from the coefficients of g?

Another possibility was to plot the points (j, g_j) in logarithmic paper, at the correct scale. That is, to plot $(j, 2^{-N}\log|g_j|)$. Then fill the dots. This plot is called the *Newton diagram* (See Fig.1) Nowadays, we can use computers to do the same.

Problem 15. In the example above, the coefficient of the line passing by $(j-1, 2^{-N}\log|g_{d-j-1}|)$ and $(j+1, 2^{-N}\log|g_{d-j+1}|)$ is approximately $-\log|\zeta_j|$. How does one recognize conjugate roots?

The method we explained does not work if f has a double ot multiple root. The next exercises show how to guarantee that this is the case, assuming exact arithmetic.

Problem 16. Show that there is an algorithm to compute the greatest common divisor of two given polynomials. Hint: use the division with remainer. Hint2: it is the same algorithm used for integer gcd.

Problem 17 (Requires calculus). Show that if ζ is a double root of f, then ζ is a root of its derivative f'. What happens if ζ is a multiplicity m root of f?

Problem 18. Verify that the greatest common divisor of f and its derivative cannot have a double root.