Lagrange's factorization for multiple polynomial roots

Roland Ven, Nicolas Boulenguez

january 2020

Abstract

A well-known result by Galois leaves no hope for an explicit algorithm based on elementary computations and splitting a general polynom into factors with degree 1. This article describes a method by Lagrange at least producing factors *with only simple roots*.

Most contents were originally found in a book by Georges Valiron.

Keywords: polynomial factorization, simple root, multiple root, square-free polynomial.

Given a polynomial $P(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$, $n \geqslant 3$, a_i a complex number, the Lagrange method determines explicit polynoms $P_1(X), P_2(X), \cdots, P_q(X)$ such that

$$P(X) = P_1(X) \times (P_2(X))^2 \times \cdots \times (P_q(X))^q$$

and each $P_j(X)$ only has *simple* roots, or is the constant 1. It does *not require to compute the roots* of $P(X)$.

The separation of the roots when the coefficients of $P(X)$ are *real* will be considered after the description of the method.

Notation LCD$(U(X), V(X))$ designates one of the Lowest Common Divisors of two (non-zero) polynoms $U(X)$ et $V(X)$. All possible values only differ by a non-zero constant factor.

Prerequisites

- Elementary polynomial computations, described for example by `wikipedia`, are sufficient in order to apply the method, namely derivation (term after term), division (by decreasing powers), and LCD (with Euclid's algorithm). Most software tools implement them, but they are easily applied manually to polynome with limited degrees.
- In order to read the proof, you need to understand classical properties of $\mathbb{C}[X]$, especially the link between multiple roots and derivation, but also some abstract caracterisations of the LCD of two non-zero polynoms.

1 Presentation of the method

1.1 Definition of the $P_j(X)$ polynoms

Let $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$, $n \geqslant 3$, a_i some complex numbers.
Let $x_1, \cdots, x_r$ be the (distinct) roots of $P(X)$.

Without any need to compute these roots, we know that

$$P(X) = (X - x_1)^{h(1)} \times \cdots \times (X - x_r)^{h(r)} \qquad \text{where root } x_k \text{ has multiplicity } h(k).$$

Let q be the greatest of all $h(k)$ in this equation, in other words the highest multiplicity order among all roots of $P(X)$.

For each integer j from 1 to q, let us define $P_j(X)$ in all possible cases.

- If $P(X)$ has no root of order j, $P_j(X) = 1$.
- If x_i is the *only* root of $P(X)$ of order j, $P_j(X) = (X - x_i)$.
- If $P(X)$ has at least two roots of order j, $P_j(X) = \prod_i (X - x_i)$, this product running over all roots x_i of $P(X)$ with a multiplicity equal to j.

In the relation $P(X) = (X - x_1)^{h(1)} \times \cdots \times (X - x_r)^{h(r)}$, if at least two x_i have the same order j, let us consider $\prod_i (X - x_i)^j$, this product covering *all* roots x_i of $P(x)$ with multiplicity j. Then

$$\prod_i (X - x_i)^j = \left(\prod_i (X - x_i)\right)^j = \left(P_j(X)\right)^j.$$

so $P(X) = P_1(X) \times (P_2(X))^2 \times \cdots \times \left(P_q(X)\right)^q$, where q is the highest multiplicity of a root of $P(X)$.

Remark. When $P(X)$ only has simple roots, this reads: $P(X) = P_1(X)$.

Remark. The definition of $P_j(X)$ makes it explicit that $P_j(X)$ can only have simple roots, and that for $j \neq k$, $P_j(X)$ and $P_k(X)$ share no common root.

1.2 Main property

From now on, we will use shorter notations: $P = P(X)$, $P_1 = P_1(X)$, *etc.* Also P' will designate the polynomial derivative of P. With these notations, $P = P_1 P_2^2 \cdots P_q^q$.

Property 1 ((proof at this end of this article)). *With the notations above for the P_j polynoms, $P = P_1 P_2^2 \cdots P_q^q = \prod_{j=1}^{q} P_j^j$ implies*

$$\mathrm{LCD}(P, P') = P_2 P_3^2 \cdots P_q^{q-1} = \prod_{j=2}^{q} P_j^{j-1}.$$

Remark. When P only has simple roots, this reads: $\mathrm{LCD}(P, P') = 1$.

1.3 Algorithm

Let

$$\begin{aligned} Q_0 &= P & &= P_1 P_2^2 \cdots P_q^q & &\text{and} \\ Q_1 &= \mathrm{LCD}(Q_0, Q_0') & &= P_2 \cdots P_q^{q-1} & &\text{according to the property above.} \end{aligned}$$

It has allready been stated that all roots of P_j must be simple, and that P_j and P_k have no root in common when $j \neq k$. So the roots of P_2 are the roots of Q_1 with multiplicity 1, and so on… the roots of P_q are the roots of Q_1 with multiplicity $q-1$. Property 1 then applies to $Q_1 = P_2 \cdots P_q^{q-1}$ (with a shift in indices).

$$
\begin{aligned}
Q_2 &= \mathrm{LCD}(Q_1, Q_1') &&= P_3 \cdots P_q^{q-2} && \text{and so on until}\\
Q_{q-2} &= \mathrm{LCD}(Q_{q-3}, Q_{q-3}') &&= P_{q-1} P_q^2 &&\\
Q_{q-1} &= \mathrm{LCD}(Q_{q-2}, Q_{q-2}') &&= P_q && \text{and finally}\\
Q_q &= \mathrm{LCD}(Q_{q-1}, Q_{q-1}') &&= 1 && \text{because all roots of } P_q \text{ are simple}
\end{aligned}
$$

The Q_i polynomials can be computed by polynomial derivation, then determination of a LCD of two non-zero polynomials (for example with Euclid’s algorithm), *without computing* the roots of $P(X)$, or prior knowledge of the P_j polynomials. The algorithm stops when the LCD is a constant. The number of iterations then *determines the value of* q, and the last non-constant LCD *gives the value of* P_q.

By construction, $Q_t = P_{t+1} \cdots P_q Q_{t+1}$ for t from 0 to $q-2$. Let R_t be the quotient of Q_t by its successor.

$$
\begin{aligned}
R_1 &= \frac{Q_0}{Q_1} &&= P_1 \cdots P_q &&\\
R_2 &= \frac{Q_1}{Q_2} &&= P_2 \cdots P_q && \text{and so on until}\\
R_{q-1} &= \frac{Q_{q-2}}{Q_{q-1}} &&= P_{q-1} P_q && \text{and finally}\\
R_q &= \frac{Q_{q-1}}{Q_q} &&= P_q && \text{already known, and consistent since } Q_q = 1
\end{aligned}
$$

Each P_j is then the quotient of R_j by R_{j+1} for j from 1 to $q-1$.

$$
\begin{aligned}
P_1 &= \frac{R_1}{R_2} &&\\
P_2 &= \frac{R_2}{R_3} && \text{and so on until}\\
P_{q-1} &= \frac{R_{q-1}}{R_q} &&\\
P_q &= R_q && \text{is already known}
\end{aligned}
$$

This effective process leads to an *explicit* factorization $P = P_1 P_2^2 \cdots P_q^q$,. The search of root of P is then split in smaller searches for the roots of each P_j, which are all simple roots (of course, the final product can omit the factors where $P_j = 1$). Moreover, when P has at least a multiple root, the degree of each P_j is strictly inferior to the degree of P, leading to new computing opportunities.

The algorithm is summarized formally by

$$
Q_0 = P \qquad Q_{i+1} = \mathrm{LCD}(Q_i, Q_i') \qquad R_{i+1} = \frac{Q_i}{Q_{i+1}} \qquad P_{i+1} = \frac{R_{i+1}}{R_{i+2}}
$$

and the following pseudo-code, parametrized by the polynom `p` that we want to factorize and a procedure `process` describing what we intend to with the resulting factors.

```
q1 := lcd (p, derivative (p));
r1 := exact_division (p, q1);
loop
   q2 := lcd (q1, derivative (q1));
   r2 := exact_division (q1, q2);
   process (exact_division (r1, r2));
   exit when degree (q1) = 0;
   q1 := q2;
   r1 := r2;
end loop;
```

For practical purposes, it is worth mentioning that most software libraries provide an alternative division wich performs much faster when the remainder is known in advance to be zero.

1.4 Example

This case study is an addition to the original book by Georges Valiron. The initial degree is 4 so the roots could, tediously, be computed with direct methods.

$P = X^4 + X^3 - 3X^2 - 5X - 2 = Q_0$, so $Q_0' = 4X^3 + 3X^2 - 6X - 5$.

$Q_1 = \mathrm{LCD}(Q_0, Q_0') = X^2 + 2X + 1$ (either with a symbolic computing software or with Euclid's algorithm), $Q_1' = 2X + 2$

$Q_2 = \mathrm{LCD}(Q_1, Q_1') = X + 1$, $Q_2' = 1$.

$Q_3 = \mathrm{LCD}(Q_2, Q_2') = 1$, leading to $q = 3$, and $P_3 = Q_2 = X + 1$.

At this step, $P = Q_0 = X^4 + X^3 - 3X^2 - 5X - 2$, $Q_1 = X^2 + 2X + 1$, $Q_2 = X + 1 = P_3$ and $q = 3$. We can now proceed:

$$\begin{aligned}
R_1 &= \frac{Q_0}{Q_1} = X^2 - X - 2 \\
R_2 &= \frac{Q_1}{Q_2} = X + 1 && \text{and finally} \\
R_3 &= P_3 = X + 1
\end{aligned}$$

This allows to compute:

$$\begin{aligned}
P_1 &= \frac{R_1}{R_2} = \frac{X^2 - X - 2}{X + 1} = X - 2 \\
P_2 &= \frac{R_2}{R_3} = \frac{X + 1}{X + 1} = 1 && \text{and} \\
P_3 &= X + 1 && \text{which was already known}
\end{aligned}$$

Finally: $P = (X - 2).(X + 1)^3$ (omitting $P_2 = 1$).

As far as the authors can tell with only partial copies of the original book by Georges Valiron, the following material was not explicitly covered there.

2 Case of real polynomials

2.1 Validity of the method

The process only requires derivations, divisions and computations of a lowest common divisor. Because such operations produce real results from real arguments, when P has real coefficients, so are all intermediate results, especially the final P_j polynomials.

2.2 Separation of the roots

The P_j polynomials may simplify the classical problem of separating the roots of $P(X) \in \mathbb{R}[X]$.

For example, it is common when applying the Sturm theorem, to replace P with

$$R_1 = \frac{P}{\mathrm{LCD}(P, P')}$$

in order to avoid multiple roots. The new polynom R_1 has the same roots than P, but for R_1 these are all *simple* roots

Lagrange's method extends this idea, and, when possible, splits R_1 into real polynoms of smaller degree, to which the Sturm theorem applies separately. Of course, a smaller degree allows new opportunities, like the determination of an exact formula for some roots.

2.3 Approximation of the roots

Many methods for the approximation of roots of numerical equations, like Newton'method, are subject to numeric instability when encountering an horizontal tangent. For polynomial equations, Lagrange's method transforms this problematic case into an optimization, because the equation can then be replaced with separate equations with a strictly smaller degree and only simple roots.

3 Proof of property 1

We will only need the following classical properties in $\mathbb{C}[X]$.

1. When P only has simple roots, P and P' have no common root. Now recall that in $\mathbb{C}[X]$, two polynoms without common root have a LCD equal to 1, in this case $\mathrm{LCD}(P, P') = 1$.
2. When P has a multiple root, let $D(X) = \prod_i (X - x_i)^{h(i)-1}$, the product covering all roots x_i of P with multiplicity $h(i) \geqslant 2$. When x_i is the only root of P with multiplicity $h(i) \geqslant 2$, $D(X)$ equals $(X - x_i)^{h(i)-1}$
   (a) The roots shared by P and P' are exactly the roots of P with multiplicity at least 2.
   (b) Each root x_i with multiplicity $h(i) \geqslant 2$ of polynom P is also root with multiplicity $h(i) - 1$ of its derivative P'.
   (c) These two properties imply that $P = D{\times}H$ and $P' = D{\times}N$, where H and N are non-zero polynomials sharing neither roots nor (non constant) divisors.
   (d) Since $P = D \times H$, $P' = D \times N$, H and N have no common divisor, it follows that D is a LCD of P and P'.
   (e) Finally, let us gather the $(X - x_i)^j$ factors in $D(X)$ as done in $P(X)$ in section 1. The definition of the P_j now reads

$$D = P_2 P_3^2 \cdots P_q^{q-1} = \prod_{j=2}^{q} P_j^{j-1}.$$