Polynomial GCD

Application of discrete models

Previous documents (Divisors, prime numbers; Greatest common divisor) discussed divisors, divisibility, greatest common divisor and other related concepts for integers. In this document, we will introduce the equivalents of these concepts for polynomials.

In fact, we will define these concepts more generally, in *rings*. This abstraction allows us to easily apply them to both integers and polynomials, and it opens the future possibility to extend them to other structures. More specifically, the discussion will be based on a particular type of rings: the *integral domains* (see: Algebraic structures). Reminder: the main integral domains are the set of integers ($\mathbb{Z}$), the polynomial rings ($K[x]$, where K is any integral domain), and all fields (e.g. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and any $\mathbb{Z}_p$ for prime p).

1 Divisibility

First, we generalize the concept of divisibility ($b \mid a$), which is already well-known for integers:

Definition 1.1. In an integral domain R, $a \in R$ is **divisible** by $b \in R$ (or b is a **divisor** of a) if $\exists c \in R : a = bc$. Notation: $b \mid a$.

Example. In $\mathbb{Z}$, $2 \mid 6$, because $6 = 2 \cdot 3$.

Example. In $\mathbb{R}[x]$, $2x^2 - x - 1$ is divisible by $x - 1$, because $2x^2 - x - 1 = (x - 1)(2x + 1)$.

Example. In fields, divisibility is trivial, because all elements are divisible by all nonzero elements, since $c := b/a$ always exists. For example, in $\mathbb{Q}$, 5 is divisible by 2, because $5 = 2 \cdot (5/2)$.

The point of divisibility for polynomials is that it has a strong connection with the roots: if $b \mid a$, then all roots of b are also roots of a (and at least as many times as for b). For example, 1 is a root of $x - 1$, and $x - 1 \mid 2x^2 - x - 1$, so we know that 1 must be a root of $2x^2 - x - 1$ too.

Just like for integers, divisibility of polynomials can be checked using division with remainder: if $b \neq 0$, then $b \mid a \iff a \bmod b = 0$.

To discuss divisibility in general, we need to define some new notions:

Definition 1.2. In an integral domain R, $u \in R$ is a **unit** if it divides the identity element.

Example. $\mathbb{Z}$ has two units: 1 and -1 (the divisors of 1).

Example. In $\mathbb{R}[x]$, all nonzero constant polynomials are units.

Example. In fields, all nonzero elements are units.

Example. In general, the units of $K[x]$ are the units of K (as constant polynomials), e.g. in $\mathbb{Z}[x]$, the units are 1 and -1, and in $\mathbb{Q}[x]$, all nonzero constant polynomials are units.

Note: do not confuse *units* with the *identity element*, as the latter is unique, but there can be more than one units in an integral domain (as we have seen above).

(Notice how the notion of "unit" is exactly the same as the "invertible element", see: Algebraic structures.)

We also need another important notion related to divisibility:

Definition 1.3. In an integral domain R, $a \in R$ and $b \in R$ are **associates** if they divide each other: $a \mid b$ and $b \mid a$.

Example. In $\mathbb{Z}$, the associates are integers with the same magnitude, e.g. 5 and -5.

Example. In $\mathbb{R}[x]$, the associates are the constant multiples of each other, e.g. $x-1$, $2x-2$, $3x-3$, $-x+1$, $-x/2+1/2$ etc. are all associates.

Example. In fields, any two nonzero elements are associates.

Notice that the associates are always unit multiples of each other, e.g. in $\mathbb{Z}$, $-5 = -1 \cdot 5$, in $\mathbb{R}[x]$, $2x-2 = 2(x-1)$ etc. The point of associates is that they behave in exactly the same way regarding divisibility, e.g. $2 \mid 6 \implies 2 \mid -6,\ -2 \mid 6,\ -2 \mid -6$. Therefore, we can say that the associates are "essentially" the same.

2 Irreducible elements

The concept of *prime numbers* can be generalized as follows:

Definition 2.1. In an integral domain R, $a \in R$ is **irreducible** if it is nonzero, non-unit, and its only divisors are units and associates of a.

The "nonzero, non-unit" corresponds to excluding 0 and 1 from the definition of prime numbers, and "its only divisors are units and associates of a" corresponds to "its only divisors are 1 and n".

Similarly, we can generalize *composite numbers*:

Definition 2.2. In an integral domain R, $a \in R$ is **reducible** if it is nonzero, non-unit, and not irreducible.

Example. The irreducible elements of $\mathbb{Z}$ are the prime numbers and their negatives: $2, 3, 5, 7, \ldots$, $-2, -3, -5, -7, \ldots$. (For simplicity, we defined prime numbers to be positive, i.e. we chose one representative from each pair of associates. But in general, we don't have such a convenient choice, so we need to include all associates into the notion of irreducible elements.)

Example. $x^2 - 2$ is irreducible in $\mathbb{Z}[x]$, but reducible in $\mathbb{R}[x]$, because $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$.

Example. Fields have no irreducible or reducible elements, because all nonzero elements are units.

Note: the associates of irreducible elements are also irreducible.

In polynomial rings ($K[x]$), the irreducible polynomials have the following properties:

1. Over fields, all linear polynomials are irreducible, because the divisors of $p = ax+b$ are constant and linear polynomials, the former of which are units, and the latters are p's associates.
2. Irreducible polynomials of degree ≥ 2 cannot have roots (in the coefficient ring K), otherwise we could extract a root factor, which would mean that the polynomial is not irreducible.
3. In $\mathbb{C}[x]$, *only* linear polynomials are irreducible, because the fundamental theorem of algebra (Basics of polynomials, Theorem 4.8) states that all nonconstant polynomials have roots.
4. In $\mathbb{R}[x]$, only linear and quadratic polynomials can be irreducible, e.g. $2x + 3$ and $x^2 + 1$ are irreducible over $\mathbb{R}$. (But the quadratics are not always irreducible, e.g. $x^2 - 1 = (x-1)(x+1)$.)
5. But in $\mathbb{Z}[x]$ and $\mathbb{Q}[x]$, irreducible polynomials can have arbitrary degree, e.g. all polynomials of the form $x^n - 2$ are irreducible.

For integers, the significance of prime numbers was shown by the fundamental theorem of arithmetic (Divisors, prime numbers, Theorem 3.4), which guaranteed that the prime factorization is unique. If we try to generalize this, we'll find that it's not true for all integral domains. But we can distinguish those ones for which it works:

Definition 2.3. An integral domain R is called a **unique factorization domain (UFD)** if all nonzero non-unit elements can be written uniquely as a product of irreducible elements, up to associatedness and the order of the factors.

Example. $\mathbb{Z}$ is a UFD by the fundamental theorem of arithmetic. For example, $-10 \in \mathbb{Z}$ can be factored in multiple ways: $-10 = (-2)\cdot 5 = 2\cdot(-5) = 5\cdot(-2) = (-5)\cdot 2$, but these factorizations differ only in order and associatedness (i.e. sign), i.e. they are "essentially" the same. (For the fundamental theorem of arithmetic, we could get around associatedness by working with only positive integers.)

Example. All fields are trivially UFDs (since they have no nonzero non-unit elements).

Example. $\mathbb{R}[x]$ is a UFD, e.g. $x^2-1 = (x+1)(x-1) = (2x+2)(x/2-1/2) = (-3x+3)(-x/3-1/3)$ etc., but these factorizations differ only in order and associatedness (i.e. constant multiplier).

Example. If K is a UFD, then $K[x]$ is also a UFD, i.e. $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{Z}_5[x]$ etc. are all UFDs.

In Sage, the `factor()` function (see: Divisors, prime numbers) works for arbitrary UFDs, with the following caveat: it selects the irreducible elements (from its associates) so that they are as simple as possible (e.g. in $\mathbb{Z}$, it selects the positive ones), but if it doesn't work that way (e.g. if the factored number is negative), then it adds a unit (e.g. -1) to the beginning of the product. The result of `factor()` is a `Factorization` object, which behaves like a sequence (e.g. it can be iterated through by a `for` loop), but the unit, if any, is not part of this sequence, instead, it can be queried separately by `unit()`. For example:

```
sage: F = factor(-60)
sage: F
-1 * 2^2 * 3 * 5
sage: for (p, k) in F:
....:     print(p, k)
2 2
3 1
5 1
sage: F.unit()
-1
sage: P.<x> = QQ[]
sage: F = factor(2*x^2 - 1/2)
sage: F
(2) * (x - 1/2) * (x + 1/2)
sage: [p for (p, k) in F]
[x - 1/2, x + 1/2]
sage: F.unit()
2
```

3 Greatest common divisor

Now we extend the greatest common divisor and least common multiple (see the document called Greatest common divisor) from integers to arbitrary integral domains:

Definition 3.1. In an integral domain R, a **greatest common divisor** of $a \in R$ and $b \in R$ is $c \in R$ if $c \mid a \ \wedge \ c \mid b \ \wedge \ \forall d \in R : d \mid a \ \wedge \ d \mid b \implies d \mid c$. Notation: $\gcd(a, b)$.
Furthermore, a and b are **coprime** if their greatest common divisor is a unit.

Definition 3.2. In an integral domain R, a **least common multiple** of $a \in R$ and $b \in R$ is $c \in R$ if $a \mid c \ \wedge \ b \mid c \ \wedge \ \forall d \in R : a \mid d \ \wedge \ b \mid d \implies c \mid d$. Notation: $\mathrm{lcm}(a, b)$.

Example. In $\mathbb{Z}$, a greatest common divisor of 12 and 20 is 4, and a least common multiple is 60.

Example. In $\mathbb{R}[x]$, a greatest common divisor of $p = x^2 - 2x + 1 = (x-1)^2$ and $q = 2x^2 - x - 1 = (x-1)(2x+1)$ is $x - 1$, and a least common multiple is $(x-1)^2(2x+1) = 2x^3 - 3x^2 + 1$.

But in such a general context, there are several problems with these concepts. First, they do not necessarily exist. But fortunately, in the important special cases that we care about, they do exist:

Theorem 3.3. *In a unique factorization domain (UFD), any two elements have a greatest common divisor and a least common multiple.*

So in $\mathbb{Z}$, $\mathbb{Z}[x]$, $\mathbb{R}[x]$ etc., gcd and lcm do always exist. (The theorem is not surprising if we consider how gcd and lcm can be calculated from the prime factorization in $\mathbb{Z}$.)

But there is another problem: in general, gcd and lcm are not unique!

Example. In $\mathbb{Z}$, the above definition for $\gcd(6, 4)$ is satisfied by both 2 and -2. In the original $\mathbb{Z}$-only definition, the only reason gcd was unique was that we restricted the result to $\mathbb{N}$.

Example. For the above $p = x^2 - 2x + 1$ and $q = 2x^2 - x - 1$, in $\mathbb{R}[x]$, the following polynomials all qualify as their greatest common divisors: $r_1 = x - 1$, $r_2 = -x + 1$, $r_3 = 2x - 2$, $r_4 = -x/2 + 1/2$ etc. (But in $\mathbb{Z}[x]$, only r_1 and r_2 do, because there $r_3 \nmid p$ and $r_4 \notin \mathbb{Z}[x]$.)

But even if we can't have complete uniqueness, there is uniqueness "in some sense":

Theorem 3.4. *The greatest common divisor and the least common multiple, if they exist, are unique up to associatedness.*

Example. In $\mathbb{Z}$, gcd is unique *up to sign*: $\gcd(6, 4)$ can be both 2 and -2, but nothing else. And if we restrict the gcd to $\mathbb{N}$, then it is completely unique.

Example. For polynomials in $\mathbb{R}[x]$, gcd is unique *up to a constant multiplier*: in the above example, $r_1 = -r_2 = 1/2 \cdot r_3 = -2r_4$. For complete uniqueness, we can require e.g. that the leading coefficient of the gcd must be 1 (e.g. then $r_1 = x - 1$ is the "true" gcd). (But in $\mathbb{Z}[x]$, this is not always possible.)

The following theorem shows a reason why the polynomial gcd is interesting:

Theorem 3.5. *If $p, q, r \in K[x]$ and $r = \gcd(p, q)$, then the roots of r are exactly the common roots of p and q, and their multiplicity in r is the smaller one of their multiplicities in p and q.*

Example. $p = x^2 - 2x + 1$ has only one root: 1 (a double root); $q = 2x^2 - x - 1$ has two roots: 1 and $-1/2$ (simple roots); so $r = \gcd(p, q)$ (or any other r_i) has only one root: 1 (as a simple root).

So if we need to find the common roots of two polynomials, we only need to calculate the roots of their gcd, which is usually easier, because it has a smaller degree; and if it is a constant, i.e. the two polynomials are coprime, then we know immediately that they cannot have common roots.

4 Euclidean algorithm

In the document 'Greatest common divisor', we have seen the Euclidean algorithm that calculates the gcd of two integers. The same algorithm can be applied to polynomials too, because they also have a 'division with remainder' operation (see the document: Polynomial operations).

Example. The Euclidean algorithm for $p = 2x^3 + x^2 - 2x - 1$ and $q = x^2 - 2x + 1$ works as follows:

$$\begin{aligned} r_0 &:= p &&= 2x^3 + x^2 - 2x - 1, \\ r_1 &:= q &&= x^2 - 2x + 1, \\ r_2 &:= r_0 \bmod r_1 &&= 6x - 6, \\ r_3 &:= r_1 \bmod r_2 &&= 0, \end{aligned}$$

therefore $\gcd(p, q) = r_2 = 6x - 6$. But this is not the only correct answer, because we have seen that the greatest common divisor is not unique. That is, we can multiply the above result by any nonzero constant, e.g. $\gcd(p, q) = -2r_2 = -12x + 12$ and $\gcd(p, q) = \frac{1}{6}r_2 = x - 1$ are also correct.

Example. The Euclidean algorithm for $p = x^4 - 4x^3 + 1$ and $q = x^3 - 3x^2 + 1$ works as follows:

$$\begin{aligned} r_0 &:= p = x^4 - 4x^3 + 1, \\ r_1 &:= q = x^3 - 3x^2 + 1, \\ r_2 &:= r_0 \bmod r_1 = -3x^2 - x + 2, \\ r_3 &:= r_1 \bmod r_2 = 16/9 \cdot x - 11/9, \\ r_4 &:= r_2 \bmod r_3 = -27/256, \\ r_5 &:= r_3 \bmod r_4 = 0, \end{aligned}$$

therefore $\gcd(p, q) = r_4 = -27/256$, or indeed any constant multiple of this is also correct, e.g. $\gcd(p, q) = 256r_4 = -27$ or $\gcd(p, q) = -256/27 \cdot r_4 = 1$. In general, if the result is a constant, then any other nonzero constant is also correct, and then p and q are coprime (i.e. they cannot have common roots). Note: since constant multipliers don't change the correctness of the result, we can also multiply any earlier polynomial in the algorithm by arbitrary constants to simplify the calculations, e.g. we can use the fraction-free $\tilde{r}_3 := 9r_3 = 16x - 11$ instead of r_3, and so on.

Important: in general, the Euclidean algorithm works only over *fields*, e.g. in $\mathbb{R}[x]$ or $\mathbb{Q}[x]$, but not in $\mathbb{Z}[x]$, because the algorithm can introduce fractions (as in the last example). This is true even if a gcd exists in $\mathbb{Z}[x]$, e.g. $\gcd(p, q) = 1 \in \mathbb{Z}[x]$ was true above – but we only learnt this by first calculating another gcd in $\mathbb{Q}[x]$: $-27/256$.

Not only the regular Euclidean algorithm, but its extended version can also be applied to polynomials, just like for integers (see the document called Greatest common divisor):

Example. For the above $p = 2x^3 + x^2 - 2x - 1$ and $q = x^2 - 2x + 1$, the extended Euclidean algorithm (EEA) works as follows:

$2x^3 + x^2 - 2x - 1$	1	0	
$x^2 - 2x + 1$	0	1	$(-(2x + 5))$
$6x - 6$	1	$-2x - 5$	$(-(\frac{1}{6}x - \frac{1}{6}))$
0	$-\frac{1}{6}x + \frac{1}{6}$	$\frac{1}{3}x^2 + \frac{1}{2}x + \frac{1}{6}$	

That is, $\gcd(p, q) = 6x - 6 = p \cdot 1 + q \cdot (-2x - 5)$. Or we can multiply this by any constant, e.g. $\gcd(p, q) = x - 1 = p \cdot \frac{1}{6} + q \cdot (-\frac{1}{3}x - \frac{5}{6})$ is also true.

The extended Euclidean algorithm is based on the following theorem:

Theorem 4.1 (Bézout's identity for polynomials)**.** *If K is a field, and $r \in K[x]$ is a greatest common divisor of $p \in K[x]$ and $q \in K[x]$, then there exist $u, v \in K[x]$ for which $pu + qv = r$.*

Note: it is still important that the polynomials are over a field. For example, if we consider the last example, but in $\mathbb{Z}[x]$, then we have $\gcd(p, q) = x - 1 \in \mathbb{Z}[x]$, but the equation $pu + qv = x - 1$ has no solution for u and v in $\mathbb{Z}[x]$, only in $\mathbb{Q}[x]$.

In Sage, `gcd(p, q)` and `xgcd(p, q)` work for polynomials in the same way as for integers (except that for polynomials, `gcd(p, q)` is not the only possible solution, since the gcd is not unique):

```
sage: P.<x> = QQ[]
sage: p = 2*x^3 + x^2 - 2*x - 1
sage: q = x^2 - 2*x + 1
sage: gcd(p, q)
x - 1
sage: xgcd(p, q)
(x - 1, 1/6, -1/3*x - 5/6)
sage: p * _[1] + q * _[2]   # check the result
x - 1
```

We can define the Euclidean algorithm not only for integers and polynomials, but also more generally. For this, we need to generalize division with remainder, which is the key operation of the algorithm.

The problem is that this operation doesn't work in all integral domains. For example, $\mathbb{Z}[x]$ is an integral domain (even a UFD, see above, so the gcd always exists), but the division with remainder only works for polynomials over fields, e.g. in $\mathbb{Q}[x]$. Therefore, we need to narrow down the class of rings where division with remainder is possible, and so the Euclidean algorithm can be performed. We will call such rings *Euclidean domains* below.

If we compare the division theorem for integers (Divisors and prime numbers, Theorem 1.1) and for polynomials (Polynomial operations, Theorem 3.1), we can see that the main difference is how we defined that the remainder is "smaller" than the divisor: for polynomials, we used the degree, and for integers, we used the magnitude. But either way, some form of "smaller" is essential, because this is what guarantees that the Euclidean algorithm eventually terminates: each remainder calculation gives a "smaller" element, therefore it is assured that we reach 0 eventually. In general, we define this "smallness" using a function f, which assigns natural numbers to the elements of the ring.

Definition 4.2. An integral domain R is called a **Euclidean domain** if there is a function $f : R \setminus \{0\} \to \mathbb{N}$ with the following properties:

1. $\forall a, b \in R \setminus \{0\} : f(a) \leq f(ab)$,
2. $\forall a, b \in R, b \neq 0 : \exists q, r \in R : a = qb + r \ \wedge \ (r = 0 \ \vee \ f(r) < f(b))$.

Example. $\mathbb{Z}$ is a Euclidean domain, by choosing $f(a) := |a|$.

Example. $K[x]$ is a Euclidean domain for any field K, by choosing $f(a) := \deg a$. (Note: in the definition, we excluded 0 from the domain of f, so the zero polynomial doesn't pose a problem.)

Example. Every field is trivially a Euclidean domain, because every element is divisible by any nonzero element, so the remainder is always $r = 0$, and the function f has no role (e.g. it can be always $f(a) := 1$).

Theorem 4.3. *If R is a Euclidean domain, then it is also a unique factorization domain (UFD).*

(But this is not true in the other way, e.g. the above-mentioned $\mathbb{Z}[x]$ is a UFD, but not a Euclidean domain.)

(Note: the definition of the Euclidean domain doesn't require that r and q are unique, only that they exist. For example, for $b = 10 \in \mathbb{Z}$, it only requires that $|r| < 10$, i.e. $-9 \leq r \leq 9$, which allows multiple possibilities, e.g. $a = 36 = 3 \cdot 10 + 6 = 4 \cdot 10 + (-4)$. To make division with remainder unambiguous, we may need to give further restrictions depending on the ring, e.g. for integers, the division theorem required that $r \geq 0$. But for polynomials, we don't need further restrictions, because $\deg r < \deg b$ is sufficient for the uniqueness.)

5 Square-free factorization

Definition 5.1. In an integral domain R, $a \in R$ is **square-free** if there is no non-unit $b \in R$ for which $b^2 \mid a$.

Example. In $\mathbb{Z}$, 10 is square-free, because no square number divides it other than 1. But 24 is not square-free, because $2^2 \mid 24$. In general, an integer is square-free if it doesn't have repeated prime factors (e.g. $10 = 2 \cdot 5$, but $24 = 2^3 \cdot 3$).

Example. In $\mathbb{R}[x]$, $x^2+x-2 = (x-1)(x+2)$ is square-free, because no square of polynomial divides it other than constants; but $x^3 - 4x^2 - 3x + 18 = (x+2)(x-3)^2$ is not, because $(x-3)^2 \mid p$.

The point of square-free polynomials is that because they don't have multiple *factors*, they also cannot have multiple *roots* (see the root factors in Basics of polynomials).

(But the reverse is not true: if a polynomial is not square-free, it doesn't necessarily have multiple roots, e.g. $(x^2+1)^2(x-1) \in \mathbb{R}[x]$ has multiple *factors*, but not multiple *roots*, because (x^2+1) has no real roots.)

In this section, we will discuss how to get rid of the multiple factors – and therefore multiple roots – of a polynomial, or more precisely, how to turn them into simple roots (factors), i.e. make the polynomial square-free. By doing this, we can reduce the degree of the polynomial without changing its "essence", e.g. its roots. This is often the first step of complex polynomial algorithms such as root finding and factorization. (Furthermore, many root finding algorithms behave poorly for multiple roots, so it is even more important to "flatten" them, i.e. to turn them into simple roots.)

For this, a crucial tool will be the *formal derivative* of a polynomial (see the previous document). The following theorem shows why it is useful:

Theorem 5.2. *If $p \in K[x]$ is divisible by $r \in K[x]$ exactly k times ($k \geq 1$), i.e. $p = r^k q$, but $r \nmid q$, then p' is divisible by the same r exactly $k-1$ times. In particular, if $r = x - c$, then if c is a root of p with multiplicity k, then c is a root of p' with multiplicity $k-1$.*

Example. $p = x^3 - 4x^2 - 3x + 18 = (x+2)(x-3)^2$ has the double root $c = 3$, so it is a simple root of $p' = 3x^2 - 8x - 3 = (x-3)(3x+1)$. (Also, -2 is a simple root of p, and not a root of p' at all.)

(Proof: calculate the derivative of $p = r^k q$, using the properties of the derivative from the previous document:

$$p' = (r^k q)' = (r^k)'q + r^k q' = kr^{k-1}q + r^k q' = r^{k-1}(kq + rq'),$$

so r divides p' *at least* $k-1$ times. The '*exactly* $k-1$ times' follows from the fact that $r \nmid kq + rq'$, because $r \nmid q$.)

That is, p' preserves the *multiple* roots (factors) of p (each one with one lower multiplicity). However, the problem is that p' can have additional roots (factors) (e.g. $c = -1/3$ above). But we can easily filter these out by taking $\gcd(p, p')$, keeping *only* the multiple roots (factors) of p.

Example. Continuing the above example, $\gcd(p, p') = x - 3$, whose only root is the simple root $c = 3$. This shows clearly that the only multiple root of p is $c = 3$.

Note: if $\gcd(p, p')$ is constant, then we know immediately that p is square-free.

If we divide the polynomial by $\gcd(p, p')$, i.e. calculate $p \,/\, \gcd(p, p')$ (which is also a polynomial, because $\gcd(p, p') \mid p$), then it has exactly the same roots (factors) as p, but each one only once (because each r^k was divided by r^{k-1}). Thus, we got the square-free version of the polynomial p.

Example. For the above $p = (x+2)(x-3)^2$, its square-free version is $p \,/\, \gcd(p, p') = (x+2)(x-3)$.

Note: for number polynomials (e.g. in $\mathbb{R}[x]$), the above method works perfectly, but in other structures, e.g. in $\mathbb{Z}_m[x]$, we may encounter problems (even if $\mathbb{Z}_m$ is a field). For example, the derivative of $p = x^6 + x^3 + \overline{1} \in \mathbb{Z}_3[x]$ vanishes: $p' = \overline{0}$. This is a problem, because then $p \,/\, \gcd(p, p') = p/p = 1$, so we can't filter out the multiple roots – even though p has plenty of them: $p = x^6 + x^3 + \overline{1} = (x - \overline{1})^6$, i.e. it has a sextuple root, $c = \overline{1}$.

Continuing the above steps, let $u := \gcd(p, p')$ and $v := p/u = p \,/\, \gcd(p, p')$, and consider the polynomial $w := \gcd(u, v)$: just like u, it has exactly the multiple roots (factors) of p, but just like v, each one only once. We can then divide these multiple roots (factors) out of the polynomial, i.e. calculate v/w, which contains *exactly* the simple roots (factors) of p.

Example. If $p = (x-4)(x-5)(x-6)^2(x-7)^4$, then we get $u = \gcd(p, p') = (x-6)(x-7)^3$, $v = p/u = (x-4)(x-5)(x-6)(x-7)$, $w = \gcd(u, v) = (x-6)(x-7)$ and $v/w = (x-4)(x-5)$.

We can improve this calculation to separate not only the simple roots (factors), but also the double, triple, quadruple etc. ones. For example, we can get the double factors of p by calculating the simple factors of the above u, because u contains each factor of p with one lower multiplicity. Continuing this, we can eventually reach the following decomposition:

Definition 5.3. The **square-free factorization** of a polynomial $p \in K[x]$ is

$$p = a_1 a_2^2 a_3^3 \cdots a_m^m,$$

where each $a_i \in K[x]$ is square-free, and pairwise coprime (i.e. $\gcd(a_i, a_j)$ are constant if $i \neq j$).

Square-free factorization:

Input: $p \in \mathbb{R}[x]$
Output: $a_1, a_2, \ldots, a_i \in \mathbb{R}[x]$
$u := \gcd(p, p')$
$v := p/u$
$i := 1$
while u *is not constant* **do**
  $w := \gcd(u, v)$
  $a_i := v/w$
  $i := i + 1$
  $u := u/w$
  $v := w$
$a_i := v$
$a_1 := a_1 \cdot u$

Example. If $p = (x-4)(x-5)(x-6)^2(x-7)^4$, then the algorithm calculates the following polynomials:

Pre-loop: $u = (x-6)(x-7)^3$, $v = (x-4)(x-5)(x-6)(x-7)$.
Loop 1: $w = (x-6)(x-7)$, $a_1 = (x-4)(x-5)$, $u = (x-7)^2$, $v = (x-6)(x-7)$.
Loop 2: $w = (x-7)$, $a_2 = (x-6)$, $u = (x-7)$, $v = (x-7)$.
Loop 3: $w = (x-7)$, $a_3 = 1$, $u = 1$, $v = (x-7)$.
Post-loop: $a_4 = (x-7)$.

$$a_1 = (x-4)(x-5),\ a_2 = (x-6),\ a_3 = 1,\ a_4 = (x-7),$$
$$p = a_1 a_2^2 a_3^3 a_4^4 = \big((x-4)(x-5)\big)(x-6)^2 1^3 (x-7)^4.$$

Note: because the gcd of polynomials is not unique, the above algorithm doesn't return a unique square-free factorization either, only up to associatedness (separately for each a_i). But in any case, the input polynomial p can still be reassembled from the result exactly (this is why we multiplied a_1 by the remaining constant u in the last step).

The square-free factorization can be calculated much more efficiently than the full factorization of the polynomial. This comes useful for example for integrating rational functions (i.e. fractions of polynomials) (see: Calculus), where we need exactly the square-free factorization of the denominator.

6 Author

This document was written by Uray M. János, partly using the following sources:

[1] https://compalg.elte.gitlab-pages.hu/dimoa-web/tematika/dimoa (Hungarian)
[2] https://en.wikipedia.org/wiki/Polynomial_greatest_common_divisor
[3] H. Cohen (1996): A Course in Computational Algebraic Number Theory (chapter 3).
[4] K.O. Geddes, S.R. Czapor, G. Labahn (1992): Algorithms for Computer Algebra (chapters 2, 7 and 8).

Errors and suggestions can be reported here: uray.janos@inf.elte.hu.