

Computational Algebra

Willem de Graaf

Contents

1	Gröbner Bases	1
1.1	Gröbner bases	1
1.1.1	Polynomial rings	1
1.1.2	Polynomials in one variable	3
1.1.3	Monomial orders	3
1.1.4	Polynomial division for multivariate polynomials	5
1.1.5	Monomial ideals	6
1.1.6	Gröbner bases	9
1.1.7	Computing Gröbner bases	10
1.2	Applications of Gröbner bases	15
1.2.1	Solving polynomial equations	15
1.2.2	Applications of Gröbner bases in geometry	18
1.2.3	An application in combinatorics: Alon's non-vanishing theorem	23
1.2.4	An application in cryptography: Polly-cracker	24
1.3	Exercises	25
2	Integer Factorisation	29
2.1	The Miller-Rabin primality test	29
2.2	Factorisation	32
2.2.1	The method of Fermat	32
2.2.2	The continued fraction method (CFRAC)	33
2.2.3	The elliptic curve method (ECM)	43
2.2.4	Complexity	50
2.3	Primality proving with elliptic curves	50
2.4	Exercises	51
3	Polynomial Factorisation	53
3.1	Some generalities on polynomials	53
3.2	Berlekamp's algorithm	55
3.3	The algorithm of Cantor-Zassenhaus	59
3.4	Factorization of polynomials over \mathbb{Q}	60
3.4.1	Hensel lifting	61
3.4.2	Hensel lifting for more factors	67
3.5	Exercises	68

Chapter 1

Gröbner Bases

The term “Gröbner bases” was introduced by Bruno Buchberger



in his 1965 Ph.D. thesis (*An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal*, University of Innsbruck) under the supervision of Wolfgang Gröbner. However, similar notions in different settings were around before that date, and have been developed after. In 1900 Gordan appears to have used a very similar concept. Furthermore, a “Gröbner basis theory” for free Lie algebras was developed by Shirshov in 1962 (Some algorithmic problems for Lie algebras, *Sib. Mat. Zh.*). Also, the so-called Knuth-Bendix algorithm (Knuth and Bendix, Simple word problems in universal algebra, 1970) is a similar development in the theory of finitely-presented groups.

In this chapter we describe the theory of Gröbner bases, and some of its applications, for ideals in a polynomial ring.

Throughout we use the notation $\mathbb{N} = \mathbb{Z}_{\geq 0}$.

1.1 Gröbner bases

1.1.1 Polynomial rings

Throughout k will denote a field, and $k[x_1, \dots, x_n]$ will be the polynomial ring in n variables, with coefficients in k . That is

$$k[x_1, \dots, x_n] = \left\{ \sum_{i_1, \dots, i_n} c_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n} \mid c_{i_1 \dots i_n} \in k \right\}.$$

Example 1.1.1 In the cases where we have one, two or three variables we also write, respectively, $k[x]$, $k[x, y]$ and $k[x, y, z]$.

A more compact notation for the exponents is useful on many occasions. For $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_i \in \mathbb{N}$, we define $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Such an element is called a *monomial*.

Example 1.1.2 x^2yz^{10} is a monomial, and $2x^2y + 3yz$ is a polynomial.

So in general we can write a polynomial as $\sum_{\alpha} c_{\alpha} x^{\alpha}$.

Recall that a ring R is said to be *commutative* if $fg = gf$ for all $f, g \in R$. So polynomial rings are commutative.

Definition 1.1.3 Let R be a commutative ring. A subset $I \subset R$ is called an ideal if

- $0 \in I$;
- if $f, g \in I$ then also $f + g \in I$;
- if $f \in I$ and $g \in R$ then $gf \in I$.

In this chapter we deal exclusively with polynomial rings $k[x_1, \dots, x_n]$. In such rings we have a particular class of examples of ideals, that are described as follows.

Let $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ and set

$$I = \left\{ \sum_{i=1}^s g_i f_i \mid g_i \in k[x_1, \dots, x_n] \right\}.$$

Then obviously I is an ideal of $k[x_1, \dots, x_n]$. Later we will see that every ideal of $k[x_1, \dots, x_n]$ is of this form.

Notation. We will denote the ideal I above with $I = \langle f_1, \dots, f_s \rangle$, and say that I is generated by f_1, \dots, f_s .

Fundamental problem: "Given $I = \langle f_1, \dots, f_s \rangle$ and $g \in k[x_1, \dots, x_n]$ decide whether $g \in I$."

Example 1.1.4 Let $R = k[x, y, z]$, $f_1 = xyz - xy$, $f_2 = x^2y - yz$ and $g = yz^2 - yz$. The question is whether $g \in \langle f_1, f_2 \rangle$.

Observe that

$$xf_1 - zf_2 = -x^2y + yz^2$$

so we add f_2 and find

$$xf_1 - zf_2 + f_2 = yz^2 - yz = g$$

and hence $g = xf_1 + (1 - z)f_2 \in \langle f_1, f_2 \rangle$.

Example 1.1.5 Let $I \subset k[x, y, z]$ be the ideal generated by $f_1 = x^2yz - yz - x$, $f_2 = xy^2z - xy - y$, $f_3 = xyz^2 - xy - z$. Then $x^2 - z^2$ and $y - z$ lie in I . But it certainly is not obvious.

We see that checking whether $g \in \langle f_1, \dots, f_s \rangle$ can be a rather complicated thing. But at least we know that we have to find h_i with $f = \sum_i h_i f_i$. To show that $g \notin \langle f_1, \dots, f_s \rangle$ seems even more difficult.

1.1.2 Polynomials in one variable

If we are dealing with polynomials in a single variable, then the fundamental problem has an easy solution.

Lemma 1.1.6 *Let $I \subset k[x]$ be an ideal. Then there exists $h \in k[x]$ with $I = \langle h \rangle$.*

PROOF. Let h be a nonzero element of I , of minimal degree. Let $f \in I$. Then $f = qh + r$ with $q, r \in k[x]$ and $\deg(r) < \deg(h)$.

Since $r = f - qh \in I$ we get that $r = 0$ as h is of minimal degree in I . It follows that $f = qh$. \square

Now let $I = \langle h \rangle \subset k[x]$ and $f \in k[x]$. From the proof of the preceding lemma we get the following procedure to check whether $f \in I$:

- 1) Write $f = qh + r$ with $\deg(r) < \deg(h)$. (Polynomial division.)
- 2) If $r = 0$ then $f \in I$ otherwise $f \notin I$.

(Note that if I is generated by h , then h is of minimal degree in I .)

We want to find some analogous procedure in the multivariate case. The key remark is that in the case of one variable we use the degree of a polynomial. In other words, we use an ordering on the monomials: $x^m < x^n$ if $m < n$. We start by generalising that to the multivariate situation.

1.1.3 Monomial orders

Recall the notation $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where $\alpha = (\alpha_1, \dots, \alpha_n)$.

Definition 1.1.7 *An order \leq on the set $\{x^\alpha\}$ of the monomials of $k[x_1, \dots, x_n]$ is called a monomial order if*

- 1) \leq is total, that is for all α, β we have $x^\alpha \leq x^\beta$ or $x^\beta \leq x^\alpha$.
- 2) \leq is multiplicative, that is, if $x^\alpha \leq x^\beta$ then $x^\alpha x^\gamma \leq x^\beta x^\gamma$ for all γ .
- 3) There are no infinite descending chains $x^{\alpha(1)} > x^{\alpha(2)} > \dots$ (in other words, \leq is a well-ordering).

We identify $\{x^\alpha \mid \alpha \in \mathbb{N}^n\}$ and \mathbb{N}^n by $x^\alpha \leftrightarrow \alpha$.

So from an order \leq on the set of monomials we get an order on \mathbb{N}^n by $\alpha < \beta$ if $x^\alpha < x^\beta$. Then the conditions for \leq to be a monomial order translate to:

- 1) \leq is total.
- 2) If $\alpha \leq \beta$ then $\alpha + \gamma \leq \beta + \gamma$ for all γ .
- 3) There are no infinite descending chains.

In the sequel, when we speak about monomial orders we freely interchange between an order on the set of monomials and on \mathbb{N}^n , whichever is the most convenient.

The following observation is useful.

Lemma 1.1.8 Set $\gamma_0 = (0, \dots, 0) \in \mathbb{N}^n$ and let \leq be a monomial order. Then $\alpha > \gamma_0$ for all $\alpha \in \mathbb{N}^n$, $\alpha \neq \gamma_0$.

PROOF. If $\alpha < \gamma_0$ then $\alpha + \alpha < \alpha + \gamma_0 = \alpha$. In the same way we find that $(k+1)\alpha < k\alpha$. So we have constructed an infinite descending set. \square

Example 1.1.9 (Lexicographical order.) Is denoted $<_{\text{lex}}$. By definition, $\alpha <_{\text{lex}} \beta$ if $\alpha_i < \beta_i$ where i is the minimal index with $\alpha_i \neq \beta_i$.

For example in $k[x, y, z]$ we have

$$xy^2z^{100} <_{\text{lex}} x^2yz$$

and

$$xyz <_{\text{lex}} xy^3.$$

The claim is that $<_{\text{lex}}$ is a monomial order.

PROOF. 1) It is obvious that it is a total order: if $\alpha \neq \beta$ then there is an index i with $\alpha_i \neq \beta_i$.

2) Suppose that $\alpha <_{\text{lex}} \beta$ and let i be minimal with $\alpha_i \neq \beta_i$, and hence $\alpha_i < \beta_i$. Then i is also the minimal index with $(\alpha + \gamma)_i \neq (\beta + \gamma)_i$ and

$$(\alpha + \gamma)_i = \alpha_i + \gamma_i < \beta_i + \gamma_i = (\beta + \gamma)_i.$$

Therefore $(\alpha + \gamma) <_{\text{lex}} (\beta + \gamma)$.

3) The most difficult thing to show is that there are no infinite descending chains. So suppose there is one: $\alpha(1) >_{\text{lex}} \alpha(2) >_{\text{lex}} \dots$. We use induction on n (i.e., the number of variables) to obtain a contradiction.

For $n = 1$ it is obvious that there are no infinite descending chains (here the order reduces to the normal degree order).

Assume that there are no infinite descending chains, with respect to $<_{\text{lex}}$, when there are $n - 1$ variables.

Observe that $0 \leq \alpha(i)_1 \leq \alpha(1)_1$.

Hence

$$\{\alpha(i) \mid i \geq 1\} = \bigcup_{k=0}^{\alpha(1)_1} C_k$$

where $C_k = \{\alpha(i) \mid i \geq 1, \alpha(i)_1 = k\}$.

So there has to be a k_0 with C_{k_0} infinite. Then we set $\tilde{C}_{k_0} = \{(\alpha_2, \dots, \alpha_n) \mid \alpha \in C_{k_0}\}$.

But this is an infinite descending chain with respect to the $<_{\text{lex}}$ order on $n - 1$ variables.

So we have our contradiction. \square

Example 1.1.10 (Graded lexicographical order.) Notation $<_{\text{glex}}$.

Define $|\alpha| = \alpha_1 + \dots + \alpha_n$; this is the degree of the monomial x^α . By definition $\alpha <_{\text{glex}} \beta$ if

- $|\alpha| < |\beta|$

- or $|\alpha| = |\beta|$ and $\alpha <_{\text{lex}} \beta$.

It is obvious that $<_{\text{glex}}$ is a monomial order.

1.1.4 Polynomial division for multivariate polynomials

As before we write $R = k[x_1, \dots, x_n]$. In this section we fix a monomial order \leq on \mathbb{N}^n .

Let $f = \sum_{\alpha} c_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$. α be maximal with respect to \leq with $c_{\alpha} \neq 0$. Then x^{α} is called the *leading monomial* of f , and denoted $\text{LM}(f)$, whereas c_{α} is called the *leading coefficient* and denoted by $\text{LC}(f)$.

With this notation we have the following division algorithm, which is a generalisation of the division algorithm for polynomials in one variable.

Algorithm 1.1.11 (Polynomial division with remainder)

Given: $f_1, \dots, f_s \in R, g \in R$.

We compute: $r \in R$ such that $g = h_1 f_1 + \dots + h_s f_s + r$ with $h_i \in R$ and no monomial of r is divisible by an $\text{LM}(f_i)$.

- 1) Set $\bar{g} := g, r := 0$.
- 2) If $\bar{g} = 0$ we stop.
- 3) Write $x^{\gamma} = \text{LM}(\bar{g}), c = \text{LC}(\bar{g})$. If there is an f_i such that $\text{LM}(f_i)$ divides x^{γ} then we set

$$\bar{g} := \bar{g} - \frac{c}{\text{LC}(f_i)} x^{\beta} f_i$$

where $x^{\beta} \text{LM}(f_i) = x^{\gamma}$. Otherwise we set

$$\bar{g} := \bar{g} - c x^{\gamma} \quad \text{e} \quad r := r + c x^{\gamma}.$$

Return to 2).

Proposition 1.1.12 *The algorithm 1.1.11 terminates correctly.*

PROOF. In round of the algorithm $\text{LM}(\bar{g})$ decreases. Hence the algorithm terminates as there are no infinite descending chains.

Not that throughout the execution of the algorithm we have $g - (\bar{g} + r) \in I$, where $I = \langle f_1, \dots, f_s \rangle$. Indeed, this is surely true initially (when $\bar{g} = g, r = 0$). Assume it holds for \bar{g} and r . If they get replaced by \bar{g}_1, r_1 , with

$$\bar{g}_1 = \bar{g} - \frac{c}{\text{LC}(f_i)} x^{\beta} f_i \quad \text{e} \quad r_1 = r,$$

then

$$g - (\bar{g}_1 + r_1) = g - (\bar{g} + r) + \frac{c}{\text{LC}(f_i)} x^{\beta} f_i$$

so also $g - (\bar{g}_1 + r_1) \in I$. If on the other hand $\bar{g}_1 = \bar{g} - c x^{\gamma}$ and $r_1 = r + c x^{\gamma}$ then

$$g - (\bar{g}_1 + r_1) = g - (\bar{g} + r) \in I.$$

So the property continues to hold in the next round of the algorithm. (The property $g - (\bar{g} + r) \in I$ is said to be a *loop-invariant* of the algorithm.)

So upon termination we have $\bar{g} = 0$ and $g - r \in I$ and therefore

$$g - r = h_1 f_1 + \dots + h_s f_s.$$

By construction we see that no $\text{LM}(f_i)$ divides a monomial in r . □

Remark 1.1.13 We note that the algorithm, and hence its output, depends strongly on the chosen monomial order. In the next example we see that there is another element of choice in the algorithm that can also affect the output.

Example 1.1.14 Let $f_1 = xy + 1$, $f_2 = y + 1$ e $g = xy - y$. We use the monomial order \leq_{glex} .

We can perform the algorithm in two ways:

- First alternative:

- 1) $\bar{g} = xy - y$, $r = 0$;
- 2) $\bar{g} = \bar{g} - f_1 = -y - 1$;
- 3) $\bar{g} = \bar{g} + f_2 = 0$;
- 4) $r = 0$.

- Second alternative:

- 1) $\bar{g} = xy - y$, $r = 0$;
- 2) $\bar{g} = \bar{g} - xf_2 = -x - y$;
- 3) $\bar{g} = \bar{g} + x = -y$, $r = -x$;
- 4) $\bar{g} = \bar{g} + f_2 = 1$;
- 5) $\bar{g} = \bar{g} - 1 = 0$, $r = -x + 1$;
- 6) $r = -x + 1$.

Definition 1.1.15 Let r be the output of Algorithm 1.1.11 with input $f_1, \dots, f_s \in R$, and $g \in R$. Then r is called a remainder of g modulo f_1, \dots, f_s .

So from this example we see that there can be more than one remainder of g modulo f_1, \dots, f_s . So the main problem is not solved as easily as in the univariate case. The main idea is to find a different generating set $G = \{g_1, \dots, g_t\}$ of the ideal $I = \langle f_1, \dots, f_s \rangle$, with the property that the remainder of a $g \in R$ modulo G is unique, and moreover it is zero if and only if $g \in I$. Such a G will be called a Gröbner basis.

First we study what happens in a relatively easy case.

1.1.5 Monomial ideals

Definition 1.1.16 Let $A \subseteq \mathbb{N}^n$ (possibly infinite). Then the ideal

$$I = \langle x^\alpha \mid \alpha \in A \rangle \subset R = k[x_1, \dots, x_n]$$

is called a monomial ideal.

Lemma 1.1.17 Let $I = \langle x^\alpha \mid \alpha \in A \rangle$ be a monomial ideal. Then $h \in I$ if and only if all of its monomials lie in I .

PROOF. “ \Leftarrow ”. Obvious.

“ \Rightarrow ”. Let $h \in I$ and write $h = \sum_{\alpha \in A} c_{\alpha} h_{\alpha} x^{\alpha}$ with certain $h_{\alpha} \in R$ and a finite number of the coefficients $c_{\alpha} \in k$ are nonzero.

Let x^{γ} be a monomial of h . So x^{γ} is a monomial of at least one $h_{\alpha} x^{\alpha}$. But

$$h_{\alpha} = \sum_{i=1}^m \mu_i x^{\beta(i)} \implies h_{\alpha} x^{\alpha} = \sum_{i=1}^m \mu_i x^{\beta(i)+\alpha}.$$

So there is an i with

$$x^{\gamma} = x^{\beta(i)} x^{\alpha} \in I.$$

□

Lemma 1.1.18 *Let the notation be as in the preceding lemma. We have $x^{\gamma} \in I$ if and only if $x^{\gamma} = x^{\beta} x^{\alpha}$ for certain $\beta \in \mathbb{N}^n$ and $\alpha \in A$.*

PROOF. Analogous to the proof of Lemma 1.1.17. □

Now we define a partial order \preceq on \mathbb{N}^n by $\alpha \preceq \beta$ if $\alpha_i \leq \beta_i$ for all i .

Example 1.1.19 We have $(1, 2, 3) \preceq (1, 3, 5)$. Furthermore, $(1, 2, 3) \not\preceq (1, 3, 2)$ and $(1, 3, 2) \not\preceq (1, 2, 3)$.

Hence not all elements are comparable with the order \preceq . We say that $\alpha, \beta \in \mathbb{N}^n$ with $\alpha \not\preceq \beta$ and $\beta \not\preceq \alpha$ are *uncomparable* with respect to \preceq .

A set $A \subseteq \mathbb{N}^n$ of which any pair of elements is uncomparable is called an *antichain*.

Lemma 1.1.20 *Every antichain in \mathbb{N}^n is finite.*

PROOF. For $n = 1$ this is obvious as in that case \preceq is equal to the normal order \leq on the inegers. (So an antichain has at most one element.)

Let $n > 1$ minimal such that there is an infinite antichain $A \subseteq \mathbb{N}^n$. Let $\alpha \in A$. Then for all $\beta \in A$ with $\beta \neq \alpha$ there exists i with $\beta_i < \alpha_i$ (and j with $\beta_j > \alpha_j$), because α and β are not comparable. Then

$$A \setminus \{\alpha\} = \bigcup_{i=1}^n \{\beta \in A \mid \beta_i < \alpha_i\}.$$

So since A is infinite, at least one set $B_{i_0} = \{\beta \in A \mid \beta_{i_0} < \alpha_{i_0}\}$ is infinite. Write $\delta = \alpha_{i_0}$. Then $B_{i_0} = \bigcup_{j=0}^{\delta-1} \{\beta \in A \mid \beta_{i_0} = j\}$.

So at least one of these sets is infinite; let it be $C = \{\beta \in A \mid \beta_{i_0} = j_0\}$. Set

$$C' = \{(\beta_1, \dots, \beta_{j_0-1}, \beta_{j_0+1}, \dots, \beta_n) \mid \beta = (\beta_1, \dots, \beta_n) \in C\}.$$

But then also C' is an antichain and $|C'| = |C| = \infty$.

As $C' \subseteq \mathbb{N}^{n-1}$ we have reached a contradiction. □

Lemma 1.1.21 (Dickson) *Let $A \subseteq \mathbb{N}^n$ e $I = \langle x^{\alpha} \mid \alpha \in A \rangle \subset R = k[x_1, \dots, x_n]$. Then there exists an $A' \subset A$, with A' finite and $I = \langle x^{\alpha} \mid \alpha \in A' \rangle$.*

In other words, a monomial ideal is always generated by a finite number of monomials.

PROOF. Let $B \subseteq A$ a maximal antichain (that means that there is no antichain $B' \subseteq A$ with $B' \supseteq B$ and $B' \neq B$). By Lemma 1.1.20 B is finite.

Observe that for all $\alpha \in \mathbb{N}^n$ the set $\{\beta \in \mathbb{N}^n \mid \beta \preceq \alpha\}$ is finite.

Set $A' = B \cup \bigcup_{\alpha \in B} \{\beta \in A \mid \beta \preceq \alpha\}$. We claim that $I = \langle x^\alpha \mid \alpha \in A' \rangle$. The inclusion “ \supseteq ” is obvious because $A' \subseteq A$. In order to show “ \subseteq ” let $\beta \in A$. We have to show that there is a $\gamma \in \mathbb{N}^n$ and $\alpha \in A'$ with $x^\beta = x^\gamma x^\alpha$.

If $\beta \in B \subseteq A'$ then there is nothing to show. Otherwise $\beta \notin B$ and hence there is an $\alpha \in B$ such that α and β are comparable (otherwise $B \cup \{\beta\}$ is a bigger antichain). So there are two possibilities: $\alpha \preceq \beta$ or $\beta \preceq \alpha$.

If $\alpha \preceq \beta$ then we take $\gamma = \beta - \alpha$ and hence $x^\beta = x^\gamma x^\alpha$. If $\beta \preceq \alpha$ then $\beta \in A'$ so we can take $\gamma = (0, \dots, 0)$ and $\alpha = \beta$. \square

Theorem 1.1.22 *Let \leq be an order on \mathbb{N}^n with*

- 1) \leq is total,
- 2) if $\alpha \leq \beta$ then $\alpha + \gamma \leq \beta + \gamma$ for all $\gamma \in \mathbb{N}^n$,
- 3) $\alpha \geq 0 = (0, \dots, 0)$ for all $\alpha \in \mathbb{N}^n$.

Then \leq is a monomial order.

PROOF. We must show that there are no infinite descending chains. So suppose that we have $\alpha(1) \geq \alpha(2) \geq \dots$

Let $A = \{\alpha(i) \mid i \geq 1\}$ and $I = \langle x^\alpha \mid \alpha \in A \rangle$. By Dickson's lemma we have that there exists a finite $A' \subset A$, with $I = \langle x^\alpha \mid \alpha \in A' \rangle$.

Write $A' = \{\beta(1), \dots, \beta(r)\}$ with $\beta(1) > \beta(2) > \dots > \beta(r)$. We claim that $\alpha(i) \geq \beta(r)$ for all i .

Indeed, $x^{\alpha(i)} \in I$ so by Lemma 1.1.18 there exists a $\beta(j)$ e $\gamma \in \mathbb{N}^n$ with $x^{\alpha(i)} = x^\gamma x^{\beta(j)}$, that is, $\alpha(i) = \gamma + \beta(j)$.

By 3) we know that $\gamma \geq 0 = (0, \dots, 0)$ hence

$$\alpha(i) = \gamma + \beta(j) \geq 0 + \beta(j) = \beta(j) \geq \beta(r).$$

Let m be such that $\alpha(m) = \beta(r)$. Since $\alpha(m) \geq \alpha(m+1) \geq \dots \geq \beta(r)$ and $\alpha(m) = \beta(r)$ it follows $\alpha(m) = \alpha(m+1) = \dots$. So the chain cannot be infinite. \square

Theorem 1.1.23 (Hilbert's basis theorem) *Every ideal of $R = k[x_1, \dots, x_n]$ is generated by a finite number of elements.*

PROOF. Let $I \subset R$ be an ideal and $J = \langle \text{LM}(f) \mid f \in I \rangle$ the ideal generated by the leading monomials of the elements of I (with respect to a fixed monomial order).

Dickson's lemma implies that there exist $g_1, \dots, g_s \in I$ with

$$J = \langle \text{LM}(g_i) \mid 1 \leq i \leq s \rangle.$$

These generate I . Indeed, let $f \in I$, then using polynomial division with remainder we find $r \in R$ such that there are $h_1, \dots, h_s \in R$ with $f = h_1g_1 + \dots + h_sg_s + r$ and no monomial of r is divisible by a $\text{LM}(g_i)$.

But $r = f - \sum h_i g_i \in I$ hence $\text{LM}(r)$ lies in J . So there exists i such that $\text{LM}(g_i)$ divides $\text{LM}(r)$ (Lemma 1.1.18) but this is impossible except when $r = 0$. It follows that $f = \sum h_i g_i$. \square

Corollary 1.1.24 *If I_k , for $k \geq 1$, are ideals of R with $I_1 \subseteq I_2 \subseteq I_3 \subseteq \dots$ then there exists m with $I_m = I_{m+1} = I_{m+2} = \dots$*

PROOF. Let $J = \cup_{k \geq 1} I_k$ which is an ideal of R . Theorem 1.1.23 says that there are $g_1, \dots, g_s \in J$ with $J = \langle g_1, \dots, g_s \rangle$. Let k_i be such that $g_i \in I_{k_i}$ and set $m = \max(k_1, \dots, k_s)$. Then $g_1, \dots, g_s \in I_m$ and hence $I_{m+i} = I_m$ because $I_m \subseteq I_{m+i}$ is given and $I_{m+i} \subseteq J = I_m$. \square

1.1.6 Gröbner bases

Definition 1.1.25 *Let $I \subset R = k[x_1, \dots, x_n]$ be an ideal. Then the ideal $\langle \text{LM}(f) \mid f \in I \rangle$ is called the initial ideal of I . We write $\langle \text{LM}(I) \rangle = \langle \text{LM}(f) \mid f \in I \rangle$.*

Definition 1.1.26 *Let $I \subset R = k[x_1, \dots, x_n]$ be an ideal, and $G \subset I$ with $\langle \text{LM}(I) \rangle = \langle \text{LM}(g) \mid g \in G \rangle$. Then G is called a Gröbner basis of I .*

Remark 1.1.27

- $G = I$ is a Gröbner basis of I .
- From the proof of Hilbert's basis theorem we immediately get that every ideal has a finite Gröbner basis.

Lemma 1.1.28 *Let $I \subseteq k[x_1, \dots, x_n]$ be an ideal. Then $G \subset I$ is a Gröbner basis of I if and only if for all $f \in I$ there is a $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(f)$.*

PROOF. We have that G is a Gröbner basis of I if and only if

$$\langle \text{LM}(g) \mid g \in G \rangle = \langle \text{LM}(I) \rangle$$

so if and only if for all $f \in I$ there exists $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(f)$ (Lemma 1.1.18). \square

Proposition 1.1.29 *Let $I \subset R = k[x_1, \dots, x_n]$ be an ideal and $G \subset I$ a Gröbner basis. Let $f \in R$ and $r \in R$ a remainder of f modulo G . Then r is uniquely determined, that is, it does not depend on the choices made during the division algorithm. Moreover, $r = 0$ if and only if $f \in I$.*

PROOF. Suppose that

$$f = h_1g_{i_1} + \dots + h_sg_{i_s} + r_1 = \bar{h}_1g_{j_1} + \dots + \bar{h}_tg_{j_t} + r_2$$

with g_{i_k} and g_{j_k} in G and $r_i \in R$ with the property that none of their monomials are divisible by a $\text{LM}(g)$ with $g \in G$.

Then

$$r_1 - r_2 = h_1g_{i_1} + \dots + h_sg_{i_s} + \bar{h}_1g_{j_1} - \dots - \bar{h}_tg_{j_t} \in I$$

so by Lemma 1.1.28 there exists $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(r_1 - r_2)$. But $\text{LM}(r_1 - r_2)$ is a monomial of r_1 or of r_2 (or of both). Hence $r_1 - r_2 = 0$ so that $r_1 = r_2$.

If $f \in I$ then $r_1 \in I$ (because $r_1 = f - \sum_{u=1}^s h_u g_{i_u}$). So there exists $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(r_1)$ (Lemma 1.1.28) whence $r_1 = 0$. The other direction is trivial. \square

1.1.7 Computing Gröbner bases

As before we write $R = k[x_1, \dots, x_n]$ and we fix a monomial order \leq . Let $G \subset R$.

Let $r \in R$ be a remainder of $f \in R$ modulo G . Then we write $r = \bar{f}^G$.

If $\bar{f}^G = 0$ then we say that f *reduces* to 0 modulo G .

Let $f_1, f_2 \in R \setminus \{0\}$ and write $x^{\alpha(i)} = \text{LM}(f_i)$. Let γ be defined by $\gamma_j = \max(\alpha(1)_j, \alpha(2)_j)$. Then x^γ is called the *least common multiple* of $x^{\alpha(1)}$ and $x^{\alpha(2)}$.

Furthermore,

$$S(f_1, f_2) = \frac{x^\gamma}{\text{LC}(f_1)\text{LM}(f_1)} f_1 - \frac{x^\gamma}{\text{LC}(f_2)\text{LM}(f_2)} f_2$$

is called the *S-polynomial* of f_1 and f_2 .

Example 1.1.30 Let $f_1 = 2xyz^2 - xy$ and $f_2 = 3x^2y^2z - 5xyz$. We use the order $<_{\text{glex}}$. The least common multiple of the leading monomials of f_1 and f_2 is $x^\gamma = x^2y^2z^2$ and

$$S(f_1, f_2) = \frac{xy}{2}(2xyz^2 - xy) - \frac{z}{3}(3x^2y^2z - 5xyz) = -\frac{1}{2}x^2y^2 + \frac{5}{3}xyz^2,$$

hence $\text{LM}(S(f_1, f_2)) = x^2y^2$.

We see that a remainder of $S(f_1, f_2)$ modulo $\{f_1, f_2\}$ is not zero because x^2y^2 is not divisible by $\text{LM}(f_1)$ or by $\text{LM}(f_2)$.

In particular, we see that $\{f_1, f_2\}$ is not a Gröbner basis of the ideal generated by it.

An element of the form cx^α with $c \in k$ is called a *term*.

Lemma 1.1.31 Let $G \subset R = k[x_1, \dots, x_n]$ and $f \in R$. Then if f reduces to zero modulo G we have $f = h_1g_1 + \dots + h_sg_s$ with $g_i \in G$ (not necessarily distinct) and where the $h_i \in R$ are terms with

$$\text{LM}(f) = \text{LM}(h_1g_1) > \text{LM}(h_2g_2) > \dots > \text{LM}(h_sg_s).$$

PROOF. This follows from the division algorithm. \square

The next theorem gives us a criterion with which we can decide whether a given set $G \subset R$ is a Gröbner basis (of the ideal it generates) or not.

Theorem 1.1.32 Let $I \subset R$ be an ideal generated by a set $G \subset I$. Then G is a Gröbner basis of I if and only if $\overline{S(g_1, g_2)}^G = 0$ for all $g_1, g_2 \in G$.

PROOF.

“ \Rightarrow ” If G is a Gröbner basis, then since $S(g_1, g_2) \in I$, this reduces to zero modulo G (Proposition 1.1.29).

“ \Leftarrow ” Let $f \in I$; we show that there exists a $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(f)$. Then with Lemma 1.1.28 we conclude that G is a Gröbner basis of I .

We can write $f = h_1g_1 + \dots + h_sg_s$ where the $h_i \in R$ are terms and $g_i \in G$. Write $m_i = \text{LM}(h_i g_i)$ and $m = m_1$. After maybe reordering, we may assume that

$$m = m_1 = m_2 = \dots = m_v > m_{v+1} \geq \dots \geq m_s.$$

Let us also assume that $\text{LC}(g_i) = 1$ (if $\text{LC}(g_i)$ is not 1 then the same proof will go through, but the notation becomes a bit heavier).

Choose an expression of the form $f = h_1g_1 + \dots + h_sg_s$ with m minimal, and among the set of those expressions with v minimal.

Suppose that $v > 1$ and write $h_i = c_i x^{\alpha(i)}$. Then

$$\begin{aligned} f &= c_1 x^{\alpha(1)} g_1 + \dots + c_s x^{\alpha(s)} g_s \\ &= c_1 (x^{\alpha(1)} g_1 - x^{\alpha(2)} g_2) + (c_1 + c_2) x^{\alpha(2)} g_2 + \dots + c_s x^{\alpha(s)} g_s. \end{aligned}$$

Now let x^γ be the least common multiple of $\text{LM}(g_1)$ and $\text{LM}(g_2)$ and write $x^\delta = m = \text{LM}(x^{\alpha(1)} g_1) = \text{LM}(x^{\alpha(2)} g_2)$. Then

$$x^{\alpha(1)} g_1 - x^{\alpha(2)} g_2 = \frac{x^\delta}{\text{LM}(g_1)} g_1 - \frac{x^\delta}{\text{LM}(g_2)} g_2$$

because $x^\delta = x^{\alpha(i)} \text{LM}(g_i)$.

Hence

$$x^{\alpha(1)} g_1 - x^{\alpha(2)} g_2 = x^{\delta-\gamma} \left(\frac{x^\gamma}{\text{LM}(g_1)} g_1 - \frac{x^\gamma}{\text{LM}(g_2)} g_2 \right) = x^{\delta-\gamma} S(g_1, g_2).$$

But $S(g_1, g_2)$ reduces to zero modulo G . Hence by Lemma 1.1.31

$$S(g_1, g_2) = u_1 g_{i_1} + \dots + u_t g_{i_t}$$

where u_i is a term and

$$x^\gamma > \text{LM}(S(g_1, g_2)) = \text{LM}(u_1 g_{i_1}) > \text{LM}(u_2 g_{i_2}) > \dots > \text{LM}(u_t g_{i_t}).$$

Therefore

$$x^{\alpha(1)} g_1 - x^{\alpha(2)} g_2 = \sum_{j=1}^t x^{\delta-\gamma} u_j g_{i_j}$$

e $\text{LM}(x^{\delta-\gamma} u_j g_{i_j}) < x^\delta = m$. So since

$$f = c_1 (x^{\alpha(1)} g_1 - x^{\alpha(2)} g_2) + (c_1 + c_2) x^{\alpha(2)} g_2 + \dots + c_s x^{\alpha(s)} g_s$$

after substituting we obtain an expression of the form $f = h_1g_1 + \dots + h_sg_s$ with v decreased, or in case $v = 2$ and $c_1 + c_2 = 0$, with m decreased. But this is impossible, hence $v = 1$.

It follows that $m = \text{LM}(x^{\alpha(1)} g_1) = \text{LM}(f)$ and hence $\text{LM}(g_1)$ divides $\text{LM}(f)$. \square

Algorithm 1.1.33

Given: $\{g_1, \dots, g_s\}$ generating the ideal $I \subset k[x_1, \dots, x_n]$ and a fixed monomial order \leq .

We compute: a Gröbner basis of I with respect to \leq .

1. Set $G_0 := \{g_1, \dots, g_s\}$, and $i := 0$.
2. If $\overline{S(f, g)}^G$ for all $f, g \in G_i$ then G_i is a Gröbner basis and we stop.
3. If there are $f, g \in G_i$ with $r := \overline{S(f, g)}^G \neq 0$ then we set $G_{i+1} := G_i \cup \{r\}$, $i := i + 1$, and return to 2.

Proposition 1.1.34 Then algorithm 1.1.33 terminates correctly.

PROOF.

When the algorithm terminates G_i is a Gröbner basis of I because

- it generates I as it contains g_1, \dots, g_s and $G_i \subset I$,
- by Theorem 1.1.32 G_i is a Gröbner of the ideal it generates.

Now to termination. Consider the ideals $J_i = \langle \text{LM}(g) \mid g \in G_i \rangle$. We claim that $G_{i+1} \supsetneq G_i$ implies that $J_{i+1} \supsetneq J_i$. Indeed, $G_{i+1} = G_i \cup \{r\}$ and $\text{LM}(r)$ is not divisible by any $\text{LM}(g)$ for $g \in G_i$. Hence $\text{LM}(r) \notin J_i$ (Lemma 1.1.18). But $\text{LM}(r) \in J_{i+1}$ so $J_{i+1} \supsetneq J_i$.

By Corollary 1.1.24 there are no infinite increasing chains of ideals in R . Hence the algorithm must terminate. \square

Example 1.1.35 Let $g_1 = xyz - xy$ and $g_2 = x^2y - yz$. We use the order $<_{\text{glex}}$ with $x >_{\text{glex}} y >_{\text{glex}} z$.

We have $G_0 = \{g_1, g_2\}$ and

$$\begin{aligned} S(g_1, g_2) &= xg_1 - zg_2 \\ &= x^2yz - x^2y - (x^2yz - yz^2) \\ &= -x^2y + yz^2 \\ &\xrightarrow{g_2} yz^2 - yz \\ &= g_3. \end{aligned}$$

So $G_1 = \{g_1, g_2, g_3\}$. Next

$$\begin{aligned} S(g_1, g_3) &= zg_1 - xg_3 \\ &= xyz^2 - xyz - (xyz^2 - xyz) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} S(g_2, g_3) &= z^2g_2 - x^2g_3 \\ &= x^2yz^2 - yz^3 - (x^2yz^2 - x^2yz) \\ &= x^2yz - yz^3 \\ &\xrightarrow{g_1} -yz^3 + x^2y \\ &\xrightarrow{g_3} x^2y - yz^2 \\ &\xrightarrow{g_2} -yz^2 + yz \\ &\xrightarrow{g_3} 0. \end{aligned}$$

Hence $G_1 = \{g_1, g_2, g_3\}$ is a Gröbner basis of I .

When computing a Gröbner basis one can use a number of tricks that often make life easier. First of all, one can divide every element by a constant in order to make the leading coefficient equal to 1. A second trick is based on the following results.

Lemma 1.1.36 *Let $f, g \in k[x_1, \dots, x_n]$ and assume that the least common multiple of $\text{LM}(f)$ and $\text{LM}(g)$ is the product $\text{LM}(f)\text{LM}(g)$ (so the greatest common divisor of $\text{LM}(f)$ and $\text{LM}(g)$ is 1). Let $u, v \in k[x_1, \dots, x_n]$ be such that $\text{LM}(u) < \text{LM}(f)$, $\text{LM}(v) < \text{LM}(g)$. Then $ug + vf$ reduces to 0 modulo $\{f, g\}$.*

PROOF. The leading monomial of $ug + vf$ occurs in ug or in vf but not in both. Indeed, if $\text{LM}(ug) = \text{LM}(vf)$ then $\text{LM}(u)\text{LM}(g) = \text{LM}(v)\text{LM}(f)$. So since $\text{LM}(g)$ and $\text{LM}(f)$ have no common factors it follows that $\text{LM}(g)$ divides $\text{LM}(v)$. But $\text{LM}(v) < \text{LM}(g)$ so this is not possible.

It follows that $\text{LM}(ug) \neq \text{LM}(vf)$ so the biggest of the two is $\text{LM}(ug + vf)$.

Suppose that $\text{LM}(ug + vf) = \text{LM}(ug)$. Hence $\text{LM}(ug + vf)$ is divisible by $\text{LM}(g)$ with factor $\text{LM}(u)$. So in the division algorithm $ug + vf$ is replaced by

$$ug + vf - c\text{LM}(u)g = (u - c\text{LM}(u))g + vf$$

where $c \in k$.

We see that we have obtained an expression of the same form. So the division algorithm proceeds until it obtains zero. \square

Lemma 1.1.37 *Let $f, g \in k[x_1, \dots, x_n]$ such that $\text{LM}(f)$ and $\text{LM}(g)$ have no common factors. Then $S(f, g)$ reduces to 0 modulo $\{f, g\}$.*

PROOF. Write $f = c_1x^\alpha + r_1$ and $g = c_2x^\beta + r_2$, with $x^\alpha = \text{LM}(f)$ and $x^\beta = \text{LM}(g)$. Then the least common multiple of x^α and x^β is $x^{\alpha+\beta}$. It follows that

$$\begin{aligned} S(f, g) &= \frac{x^{\alpha+\beta}}{c_1x^\alpha}(c_1x^\alpha + r_1) - \frac{x^{\alpha+\beta}}{c_2x^\beta}(c_2x^\beta + r_2) \\ &= \frac{1}{c_1}r_1x^\beta - \frac{1}{c_2}r_2x^\alpha \\ &= \frac{1}{c_1c_2}r_1(g - r_2) - \frac{1}{c_1c_2}r_2(f - r_1) \\ &= \frac{1}{c_1c_2}(r_1g - r_2f) \end{aligned}$$

and $\text{LM}(r_1) < \text{LM}(f)$, $\text{LM}(r_2) < \text{LM}(g)$. By Lemma 1.1.36, $r_1g - r_2f$ reduces to zero modulo $\{f, g\}$. \square

Conclusion. When calculating a Gröbner basis we don't need to check $S(f, g)$ if $\text{LM}(f)$ and $\text{LM}(g)$ don't have common factors.

Example 1.1.38 One problem with Gröbner bases is that on many occasions they are difficult to compute because they can be very big. As example consider

$$f_1 = x^3 + y + z^2 - 1, \quad f_2 = x^2 + y^3 + z - 1, \quad f_3 = x + y^2 + z^3 - 1.$$

A Gröbner basis of the ideal generated by these polynomials with respect to $<_{\text{lex}}$ is:

$$\begin{aligned}
& x + y + 7467678205870943/65719095621171465*z^{25} + \\
& \quad 1906548593713883/4381273041411431*z^{24} + \\
& \quad 8191728039103700/13143819124234293*z^{23} - \\
& \quad 18118488190106417/65719095621171465*z^{22} - \\
& \quad 40197573291173158/13143819124234293*z^{21} - \\
& \quad 60670103256071060/13143819124234293*z^{20} - \\
& \quad 152406006738162748/65719095621171465*z^{19} + \\
& \quad 80624458839890252/13143819124234293*z^{18} + \\
& \quad 649226374687018607/65719095621171465*z^{17} + \\
& \quad 618150565337585002/65719095621171465*z^{16} + \\
& \quad 353034549878420633/65719095621171465*z^{15} + \\
& \quad 238802559571273606/65719095621171465*z^{14} - \\
& \quad 11655150926515255/4381273041411431*z^{13} - \\
& \quad 1972928184800696599/65719095621171465*z^{12} - \\
& \quad 775940019593808291/21906365207057155*z^{11} - \\
& \quad 1345970642704934506/65719095621171465*z^{10} + \\
& \quad 2087861421527788238/65719095621171465*z^9 + \\
& \quad 3117371817891097726/65719095621171465*z^8 + \\
& \quad 374132655781594758/21906365207057155*z^7 - \\
& \quad 437890362723893518/65719095621171465*z^6 - \\
& \quad 60670429063906802/1991487746096105*z^5 - \\
& \quad 42413702095534382/65719095621171465*z^4 - \\
& \quad 28132496530919371/21906365207057155*z^3 + \\
& \quad 702615426148730872/65719095621171465*z^2 - \\
& \quad 80217895633896968/21906365207057155*z - 1, \\
& y^2 - y - 7467678205870943/65719095621171465*z^{25} - \\
& \quad 1906548593713883/4381273041411431*z^{24} - \\
& \quad 8191728039103700/13143819124234293*z^{23} + \\
& \quad 18118488190106417/65719095621171465*z^{22} + \\
& \quad 40197573291173158/13143819124234293*z^{21} + \\
& \quad 60670103256071060/13143819124234293*z^{20} + \\
& \quad 152406006738162748/65719095621171465*z^{19} - \\
& \quad 80624458839890252/13143819124234293*z^{18} - \\
& \quad 649226374687018607/65719095621171465*z^{17} - \\
& \quad 618150565337585002/65719095621171465*z^{16} - \\
& \quad 353034549878420633/65719095621171465*z^{15} - \\
& \quad 238802559571273606/65719095621171465*z^{14} + \\
& \quad 11655150926515255/4381273041411431*z^{13} + \\
& \quad 1972928184800696599/65719095621171465*z^{12} + \\
& \quad 775940019593808291/21906365207057155*z^{11} + \\
& \quad 1345970642704934506/65719095621171465*z^{10} - \\
& \quad 2087861421527788238/65719095621171465*z^9 - \\
& \quad 3117371817891097726/65719095621171465*z^8 - \\
& \quad 374132655781594758/21906365207057155*z^7 +
\end{aligned}$$

$$\begin{aligned}
& 437890362723893518/65719095621171465*z^6 + \\
& 60670429063906802/1991487746096105*z^5 + \\
& 42413702095534382/65719095621171465*z^4 + \\
& 50038861737976526/21906365207057155*z^3 - \\
& 702615426148730872/65719095621171465*z^2 + \\
& 80217895633896968/21906365207057155*z, \\
y*z & + 114351761873236539/148963283407988654*z^{25} + \\
& 209291487801391117/446889850223965962*z^{24} + \\
& 52236360062652115/148963283407988654*z^{23} - \\
& 1468425502479577274/223444925111982981*z^{22} - \\
& 858875628508250933/223444925111982981*z^{21} - \\
& 616034236519743323/223444925111982981*z^{20} + \\
& 8794492882346418295/446889850223965962*z^{19} + \\
& 4815691490104347169/446889850223965962*z^{18} + \\
& 1149199885582915487/446889850223965962*z^{17} - \\
& 1088785538668067356/223444925111982981*z^{16} - \\
& 2515633540839994712/223444925111982981*z^{15} + \\
& 6870511665918010699/446889850223965962*z^{14} - \\
& 33317003410103522629/446889850223965962*z^{13} - \\
& 1138745048309366852/74481641703994327*z^{12} - \\
& 2405659933320304387/223444925111982981*z^{11} + \\
& 16532107971394002975/148963283407988654*z^{10} + \\
& 38492638078382481727/446889850223965962*z^9 - \\
& 17305956248039834578/223444925111982981*z^8 - \\
& 3042243532386622694/223444925111982981*z^7 - \\
& 8470564338716612201/74481641703994327*z^6 + \\
& 109793303496803963/1194892647657663*z^5 - \\
& 1834169915554595837/74481641703994327*z^4 + \\
& 23672654350351476275/446889850223965962*z^3 - \\
& 2965819909779524184/74481641703994327*z^2 + \\
& 3212763377507509769/446889850223965962*z, \\
z^{26} & - 9*z^{23} + 29*z^{20} - 6*z^{18} - 11*z^{17} - 14*z^{16} + 27*z^{15} - 110*z^{14} + \\
& 37*z^{13} + 4*z^{12} + 163*z^{11} + 36*z^{10} - 173*z^9 + 28*z^8 - 146*z^7 + \\
& 208*z^6 - 94*z^5 + 89*z^4 - 91*z^3 + 37*z^2 - 5*z
\end{aligned}$$

1.2 Applications of Gröbner bases

In this second section we see a few applications of Gröbner bases. Of course, they solve the main problem: given an ideal $I \subset k[x_1, \dots, x_n]$ and $f \in k[x_1, \dots, x_n]$ decide if $f \in I$. In order to solve this problem we compute a Gröbner basis G of I . Then $f \in I$ if and only if $\bar{f}^G = 0$.

1.2.1 Solving polynomial equations

Definition 1.2.1 A field k is called algebraically closed if every polynomial in $k[x]$ has a zero in k .

Example 1.2.2 For example $k = \mathbb{C}$.

The problem of this section is given f_1, \dots, f_s in $k[x_1, \dots, x_n]$, to decide if there is a vector $(a_1, \dots, a_n) \in k^n$ with

$$f_1(a_1, \dots, a_n) = \dots = f_s(a_1, \dots, a_n) = 0$$

and in the affirmative case to find such (a_1, \dots, a_n) .

If the field is algebraically closed the first part of the problem can be solved with Hilbert's Nullstellensatz.

Theorem 1.2.3 (Hilbert's Nullstellensatz) *Let k be an algebraically closed field and $R = k[x_1, \dots, x_n]$. Let $f_1, \dots, f_s \in R$. Then there exists $\bar{a} = (a_1, \dots, a_n) \in k^n$ with $f_1(\bar{a}) = \dots = f_s(\bar{a}) = 0$ if and only if the ideal $I = \langle f_1, \dots, f_s \rangle$ does not contain 1.*

In the proof we use a lemma that we do not prove.

Lemma 1.2.4 *Let $k \subset K$ be fields. Suppose that there are $\zeta_1, \dots, \zeta_s \in K$ with the property that every element of K can be written as polynomial in ζ_1, \dots, ζ_s with coefficients in k . Then K is algebraic over k .*

The intuitive idea of the proof of this lemma is as follows. If there is a transcendental $z \in K$, then $k[z]$ is isomorphic to the polynomial ring in one variable, and $k(z)$ is isomorphic to the rational function field in one variable. But that field is not finitely generated over k .

PROOF.(of the Nullstellensatz). “ \Rightarrow ” Obvious. If \bar{a} exists, then $f(\bar{a}) = 0$ for all $f \in I$; hence $1 \notin I$.

“ \Leftarrow ” Every infinite chain of ideals $I = I_0 \subseteq I_1 \subseteq I_2 \subseteq \dots$ has the property that there exists r with $I_r = I_{r+1} = I_{r+2} = \dots$ (Corollary 1.1.24). So there exists a maximal ideal of R , $J \supset I$, $J \neq R$ (note that $1 \in I$ if and only if $I = R$). So $K = k[x_1, \dots, x_n]/J$ is a field.

Write $\zeta_i = x_i \bmod J$; then every element of K is a polynomial in the ζ_i with coefficients in k . Hence K is algebraic over k (Lemma 1.2.4). Since k is algebraically closed we have $k = K$.

Hence $x_i \bmod J = a_i$ for certain $a_i \in k$ and therefore $x_i - a_i \in J$. But the ideal $\langle x_1 - a_1, \dots, x_n - a_n \rangle$ is maximal and hence $J = \langle x_1 - a_1, \dots, x_n - a_n \rangle$.

We conclude that for all $f \in J$ we have $f(a_1, \dots, a_n) = 0$ and in particular we get that for all $f \in I$. \square

The work of David Hilbert (1862 - 1943)



is of great importance for the development and use of Gröbner bases. However, interestingly, his methods were not algorithmical at all. His proof of what is now known as Hilbert's basis theorem shows that every ideal has a finite number of generators, without giving a clue as to what these generators are, or how many are needed. This prompted Gordan, famously,

to exclaim “Das ist nicht Mathematik. Das ist Theologie.” (This is not mathematics. This is theology.)

Using the Nullstellensatz we can decide whether or not a given set of polynomial equations has a solution in an algebraically closed field. Indeed, we compute a Gröbner basis G of the ideal generated by the polynomials. Then 1 lies in the ideal if and only if $\bar{1}^G = 0$.

Now we turn to the problem of actually determining a solution, if one exists. The following lemma is trivial, but nonetheless important.

Lemma 1.2.5 *Let I be the ideal generated by $f_1, \dots, f_s \in k[x_1, \dots, x_n]$, and let $\bar{a} = (a_1, \dots, a_n) \in k^n$. Then $f_1(\bar{a}) = \dots = f_s(\bar{a}) = 0$ if and only if $f(\bar{a}) = 0$ for all $f \in I$.*

PROOF. Note that $f \in I$ if and only if there exist $h_1, \dots, h_s \in R$ with $f = h_1 f_1 + \dots + h_s f_s$. \square

From this lemma we see that solving the polynomial equations $f_1 = \dots = f_s = 0$ is equivalent to solving the equations $g_1 = \dots = g_t = 0$, where $\{g_1, \dots, g_t\}$ is any generating set of the ideal generated by the f_i . Next we see that Gröbner bases with respect to the lexicographical order are particularly useful in this respect.

Definition 1.2.6 *Let $I \subset k[x_1, \dots, x_n]$ be an ideal and $0 \leq l \leq n-1$. Then $I \cap k[x_{l+1}, \dots, x_n]$, which is an ideal of $k[x_{l+1}, \dots, x_n]$, is called the l -th elimination ideal of I .*

The l -th elimination ideal eliminates the first l variables.

Theorem 1.2.7 (Elimination theorem) *Set $R = k[x_1, \dots, x_n]$, and use the order $<_{\text{lex}}$ with $x_1 >_{\text{lex}} x_2 >_{\text{lex}} \dots >_{\text{lex}} x_n$. Let $I \subset R$ be an ideal and G a Gröbner basis of I with respect to $<_{\text{lex}}$. Then $G \cap k[x_{l+1}, \dots, x_n]$ is a Gröbner basis of $I \cap k[x_{l+1}, \dots, x_n]$.*

PROOF. Set $G_l = G \cap k[x_{l+1}, \dots, x_n]$ and $I_l = I \cap k[x_{l+1}, \dots, x_n]$.

We need to show that $\langle \text{LM}(I_l) \rangle = \langle \text{LM}(g) \mid g \in G_l \rangle$.

“ \supset ” Obvious as $G_l \subset I_l$.

“ \subset ” Let $f \in I_l$. Then $f \in I$ and since G is a Gröbner basis of I , there exists $g \in G$ such that $\text{LM}(g)$ divides $\text{LM}(f)$. But $f \in k[x_{l+1}, \dots, x_n]$ so also $\text{LM}(g) \in k[x_{l+1}, \dots, x_n]$.

Now let x^α be a monomial of g , $x^\alpha \neq \text{LM}(g)$. If $\alpha_i > 0$ for some $i \leq l$ then $x^\alpha >_{\text{lex}} \text{LM}(g)$. But that is impossible.

Hence $x^\alpha \in k[x_{l+1}, \dots, x_n]$ and therefore $g \in k[x_{l+1}, \dots, x_n]$. It follows that $g \in G_l$ whence $\text{LM}(f) \in \langle \text{LM}(g) \mid g \in G_l \rangle$. \square

Conclusion. With a Gröbner basis with respect to $<_{\text{lex}}$ we immediately find Gröbner bases for

$$\begin{aligned} I \cap k[x_n] \\ I \cap k[x_{n-1}, x_n] \\ \vdots \\ I \cap k[x_2, \dots, x_n] \\ I \cap k[x_1, \dots, x_n] = I. \end{aligned}$$

So the Gröbner basis has a triangular structure that may help to solve the corresponding polynomial equations.

Note: it is possible that $I \cap k[x_{l+1}, \dots, x_n] = 0$.

Example 1.2.8 Consider the polynomials

$$f_1 = x^2 + y + z - 1, \quad f_2 = x + y^2 + z - 1, \quad f_3 = x + y + z^2 - 1.$$

We want to solve the system $f_1 = f_2 = f_3 = 0$.

A Gröbner basis with respect to $>_{\text{lex}}$ consists of the polynomials

$$\begin{aligned} g_1 &= x + y + z^2 - 1 \\ g_2 &= y^2 - y - z^2 + z \\ g_3 &= 2yz^2 + z^4 - z^2 \\ g_4 &= z^6 - 4z^4 + 4z^3 - z^2 = z^2(z-1)^2(z^2 + 2z - 1). \end{aligned}$$

With $I = \langle f_1, f_2, f_3 \rangle$ we see that $I \cap k[z]$ is generated by g_4 , $I \cap k[y, z]$ is generated by g_2, g_3, g_4 .

Also we see that $g_4 = 0$ gives a finite number of values of z , namely 0, 1 and $-1 \pm \sqrt{2}$. If $z = 0$ then also $g_3 = 0$, but $g_2 = 0$ implies that y can be 0 or 1. If $y = z = 0$, then from $g_1 = 0$ it follows that $x = 1$. We find the solution $(1, 0, 0)$. Going on like this we find all solutions:

$$\begin{aligned} &(1, 0, 0) \\ &(0, 1, 0) \\ &(0, 0, 1) \\ &(-1 - \sqrt{2}, -1 - \sqrt{2}, -1 - \sqrt{2}) \\ &(-1 + \sqrt{2}, -1 + \sqrt{2}, -1 + \sqrt{2}). \end{aligned}$$

1.2.2 Applications of Gröbner bases in geometry

Definition 1.2.9 Let $I \subset k[x_1, \dots, x_n]$ be an ideal. Then

$$\sqrt{I} = \{f \in k[x_1, \dots, x_n] \mid \text{esiste } m > 0 \text{ con } f^m \in I\}$$

is called the radical of I .

Remark 1.2.10 \sqrt{I} is an ideal of $k[x_1, \dots, x_n]$.

PROOF. Obviously $0 \in \sqrt{I}$. Let $f \in \sqrt{I}$ and $g \in k[x_1, \dots, x_n]$, so $f^m \in I$ for a certain $m > 0$. Hence $g^m f^m = (gf)^m \in I$ so that $gf \in \sqrt{I}$.

If $f, g \in \sqrt{I}$ then we let $m > 0$ be such that $f^m, g^m \in I$. Then

$$(f + g)^{2m} = \sum_{i=0}^{2m} \binom{2m}{i} f^i g^{2m-i} \in I$$

so $f + g \in \sqrt{I}$. □

Lemma 1.2.11 Let $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$ be an ideal and $f \in k[x_1, \dots, x_n]$. Then $f \in \sqrt{I}$ if and only if 1 is contained in the ideal $J = \langle f_1, \dots, f_s, 1 - yf \rangle \subset k[x_1, \dots, x_n, y]$.

PROOF. “ \Rightarrow ” Suppose that $f^m \in I$. Then $y^m f^m \in J$ (because $I \subset J$). But also

$$(1 - y^m f^m) = (1 - yf)(1 + yf + (yf)^2 + \dots + (yf)^{m-1}) \in J$$

so $1 \in J$.

“ \Leftarrow ” If $1 \in J$ then there exist $p_i, q \in k[x_1, \dots, x_n, y]$ with

$$1 = \sum_{i=1}^s p_i(x_1, \dots, x_n, y) f_i + q(x_1, \dots, x_n, y)(1 - yf).$$

Now we formally substitute y with $\frac{1}{f}$ and get

$$1 = \sum_{i=1}^s p_i(x_1, \dots, x_n, \frac{1}{f}) f_i.$$

Note that the $p_i(x_1, \dots, x_n, \frac{1}{f})$ are rational expressions with denominators f^r for certain $r \geq 0$.

Hence there is an m with $f^m p_i(x_1, \dots, x_n, \frac{1}{f}) \in k[x_1, \dots, x_n]$ and therefore

$$f^m = \sum_{i=1}^s f^m p_i(x_1, \dots, x_n, \frac{1}{f}) f_i$$

so $f^m \in I$, in other words $f \in \sqrt{I}$. □

So with Gröbner bases one can check whether $f \in \sqrt{I}$ or not. This leads to the following algorithm.

Algorithm 1.2.12

Given: f_1, \dots, f_s generating the ideal $I \subset k[x_1, \dots, x_n]$, and a $f \in k[x_1, \dots, x_n]$.

We decide $f \in \sqrt{I}$ or not.

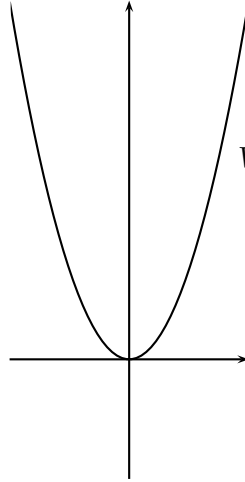
1. Compute a Gröbner basis G of $J = \langle f_1, \dots, f_s, 1 - yf \rangle$;
2. If $\bar{1}^G = 0$ then $f \in \sqrt{I}$, otherwise $f \notin \sqrt{I}$.

Definition 1.2.13 Let $W = k^n$ be a vector space over k of dimension n . Let $I = \langle f_1, \dots, f_s \rangle$ be an ideal of $k[x_1, \dots, x_n]$. Then the set

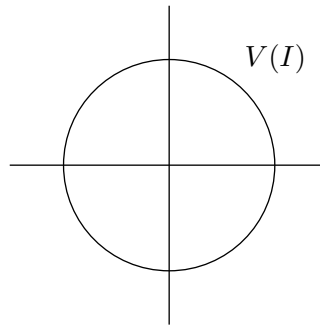
$$V(I) = \{(w_1, \dots, w_n) \in W \mid f(w_1, \dots, w_n) = 0 \text{ per ogni } f \in I\}$$

is called the closed set corresponding to I .

Example 1.2.14 Let $k = \mathbb{R}$. Let $I \subset k[x, y]$ be generated by $y - x^2$. Then $V(I)$ consists of the points of the graph



Similarly, if $I = \langle x^2 + y^2 - 1 \rangle$, then $V(I)$ consists of the points of



Theorem 1.2.15 (Hilbert's Nullstellensatz, strong form) *Let k be an algebraically closed field and $I \subset k[x_1, \dots, x_n]$ an ideal. Let $J \subset k[x_1, \dots, x_n]$ be the ideal defined by*

$$J = \{f \in k[x_1, \dots, x_n] \mid f(w) = 0, \forall w \in V(I)\}.$$

Then $J = \sqrt{I}$.

PROOF. If $f \in \sqrt{I}$ then there exists $m > 0$ with $f^m \in I$ and hence $f^m(w) = 0$ for all $w \in V(I)$. But $f^m(w) = (f(w))^m$ so $f(w) = 0$ for all $w \in V(I)$. Therefore $f \in J$.

Now take $f \in J$, that is, $f(w) = 0$ for all $w \in V(I)$. Let I be generated by $f_1, \dots, f_s \in k[x_1, \dots, x_n]$. Set $\tilde{I} = \langle f_1, \dots, f_s, 1 - yf \rangle \subset k[x_1, \dots, x_n, y]$. Let $w = (w_1, \dots, w_n, w_{n+1}) \in k^{n+1}$. We consider two cases:

- $(w_1, \dots, w_n) \in V(I)$. Then $f_i(w) = 0$ for $i = 1, \dots, s$. But $f(w) = 0$ so $(1 - yf)(w) \neq 0$.
- $(w_1, \dots, w_n) \notin V(I)$. Then there is f_i with $f_i(w) \neq 0$.

The conclusion is that $\{w = (w_1, \dots, w_{n+1}) \mid h(w) = 0, \forall h \in \tilde{I}\} = \emptyset$. By the weak form of Hilbert's Nullstellensatz (Teorema 1.2.3), $1 \in \tilde{I}$ (k is algebraically closed). Hence by Lemma 1.2.11, $f \in \sqrt{I}$. \square

Remark 1.2.16 One has

$$V(I) \cap V(J) = V(I + J)$$

where $I + J$ is the ideal

$$I + J = \{f + g \mid f \in I, g \in J\}.$$

PROOF. Let $v \in V(I) \cap V(J)$ and $h = f + g \in I + J$ with $f \in I$ and $g \in J$. Then

$$h(v) = f(v) + g(v) = 0 + 0 = 0$$

so $v \in V(I + J)$.

Conversely, let $v \in V(I + J)$. Since $I \subset I + J$ we get $h(v) = 0$ for all $h \in I$, whence $v \in V(I)$. Analogously we get $v \in V(J)$. Therefore $v \in V(I) \cap V(J)$. \square

From this it follows that for $X \subset k^n$ there exists a unique minimal closed set containing X . This is called the *closure* of X . Now let

$$\text{Id}(X) = \{f \in k[x_1, \dots, x_n] \mid f(v) = 0, \forall v \in X\}$$

which is an ideal of $k[x_1, \dots, x_n]$. Then the closure of X is $V(\text{Id}(X))$.

Theorem 1.2.17 Let k be an algebraically closed field, and $R = k[x_1, \dots, x_n]$, and $I = \langle f_1, \dots, f_s \rangle \subset R$. Let $\pi_l: k^n \rightarrow k^{n-l}$ be defined by $\pi_l(v_1, \dots, v_n) = (v_{l+1}, \dots, v_n)$.

Let $I_l = k[x_{l+1}, \dots, x_n] \cap I$, the l -th elimination ideal of I . Then $V(I_l)$ is the closure of $\pi_l(V(I))$.

Example 1.2.18 Let $I = \langle xy - 1 \rangle \subset k[x, y]$. Consider $\pi_1: k^2 \rightarrow k$ with $\pi_1(v_1, v_2) = v_2$. Then

$$V(I) = \{(v_1, v_2) \mid v_1 v_2 = 1\} \implies \pi_1(V(I)) = k \setminus \{0\}.$$

Now $I_1 = 0$ which implies $V(I_1) = k$.

PROOF. (of Theorem 1.2.17) “ \supset ”. Let $v = (v_1, \dots, v_n) \in V(I)$. Then $f(v) = 0$ for all $f \in I$. In particular for $f \in I_l \subset k[x_{l+1}, \dots, x_n]$ we have $f(\pi_l(v)) = f(v_{l+1}, \dots, v_n) = 0$.

So $V(I_l)$ is a closed set containing $\pi_l(V(I))$. hence $V(I_l)$ contains the closure of $\pi_l(V(I))$.

“ \subset ” Let $J = \text{Id}(\pi_l(V(I))) \subset k[x_{l+1}, \dots, x_n]$. We claim that $J \subset \sqrt{I_l}$. Indeed, $f \in J$ implies $f(v_{l+1}, \dots, v_n) = 0$ for all $(v_1, \dots, v_n) \in V(I)$. It follows that $f(v) = 0$ for all $v \in V(I)$ and hence, by Hilbert’s Nullstellensatz (Theorem 1.2.15), $f \in \sqrt{I}$. In other words, $f^m \in I$.

But $f \in k[x_{l+1}, \dots, x_n]$ so also f^m lives there. Therefore $f^m \in I \cap k[x_{l+1}, \dots, x_n] = I_l$. It follows that $f \in \sqrt{I_l}$.

We conclude that $V(J) \supset V(\sqrt{I_l}) = V(I_l)$. \square

Conclusion. In view of the elimination theorem, using Gröbner bases we can find the closure of $\pi_l(U)$, where $U \subset k^n$ is a closed set.

As an application we show how to find the closure of the image of a regular function.

A regular function $h: k^n \rightarrow k^m$ is given by m polynomials $h_1, \dots, h_m \in k[x_1, \dots, x_n]$, such that $h(v) = (h_1(v), \dots, h_m(v))$.

Let $X = V(I) \subset k^n$ where $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$. We want to find the closure of $h(X)$.

Let

$$\Gamma = \{(v, h(v)) \mid v \in X\} \subset k^{n+m}.$$

This is called the *graph* of h .

Observe that $\Gamma \subset k^{n+m}$ is closed. Indeed, let J be the ideal of $k[x_1, \dots, x_n, y_1, \dots, y_m]$ generated by

$$\{f_1, \dots, f_s, y_1 - h_1, \dots, y_m - h_m\}.$$

Then $(v, w) \in \Gamma$ if and only if $p(v, w) = 0$ for all $p \in J$ ($w = h(v)$ if and only if $(y_i - h_i)(v, w) = w_i - h_i(v) = 0$).

By computing a Gröbner basis we can compute the closure of $\pi_n(\Gamma)$ where $\pi_n: k^{n+m} \rightarrow k^m$ is defined by

$$\pi_n(v_1, \dots, v_n, w_1, \dots, w_m) = (w_1, \dots, w_m).$$

Now $\pi_n(\Gamma) = h(X)$ so we see that we can compute the closure of $h(X)$ (here we always assume the field k to be algebraically closed).

Example 1.2.19 Let $S = \{(t^2, t^3) \mid t \in \mathbb{C}\}$. We compute the closure of S . Define $h: \mathbb{C} \rightarrow \mathbb{C}^2$ by $h(t) = (t^2, t^3)$. Then $S = \{h(t) \mid t \in \mathbb{C}\}$. Consider the graph of h :

$$\Gamma = \{(u, h(u)) \mid u \in \mathbb{C}\}.$$

We have that $\Gamma = V(I)$ where $I = \langle y_1 - t^2, y_2 - t^3 \rangle \subset k[t, y_1, y_2]$.

We compute a Gröbner basis of I with respect to $\langle_{\text{lex}} t \rangle_{\text{lex}} y_1 \rangle_{\text{lex}} y_2$, and obtain

$$G = \{t^2 - y_1, ty_1 - y_2, ty_2 - y_1^2, y_1^3 - y_2^2\}.$$

So by the elimination theorem, a Gröbner basis of $I \cap k[y_1, y_2]$ is $G \cap k[y_1, y_2] = \{y_1^3 - y_2^2\}$. The conclusion is that the closure of S is $V(J)$ where J is generated by $y_1^3 - y_2^2$.

Here we have $V(I) = S$, but this is not true in general.

The set S is the parametric form of a curve in \mathbb{C}^2 , whereas the representation as $V(J)$ is called the implicit form of the curve. It is not true in general that the parametric form and the implicit form define exactly the same set of points.

Divertimento: the real Nullstellensatz

In general the Nullstellensatz does not hold when the field is not algebraically closed. However, when the field is \mathbb{R} , then an interesting variant of the theorem is true. It is called the *real Nullstellensatz*, (see G. Stengle. A Nullstellensatz and a Positivstellensatz in semi-algebraic geometry, Math. Ann. (1974), pp. 87-97).

For this let I be an ideal of $R = \mathbb{R}[x_1, \dots, x_n]$. Then its real radical is defined as

$$\sqrt[\mathbb{R}]{I} = \{f \in R \mid f^{2m} + \sum_i r_i^2 \in I \text{ for some } m > 0, r_i \in R\}.$$

With this definition the real Nullstellensatz holds: let J be as in the strong form of the Nullstellensatz, then

$$J = \sqrt[\mathbb{R}]{I}.$$

Example: $I = \langle x_1^2 + x_2^2 \rangle$; then it is not difficult to see that

$$\sqrt[\mathbb{R}]{I} = \langle x_1, x_2 \rangle.$$

1.2.3 An application in combinatorics: Alon's non-vanishing theorem

Here we let k be a field. Let T_1, \dots, T_n be finite subsets of k , and write $t_i = |T_i|$. Then we form the set of points $T = T_1 \times T_2 \times \dots \times T_n \subset k^n$. We set

$$I(T) = \{f \in k[x_1, \dots, x_n] \mid f(p) = 0 \text{ for all } p \in T\}.$$

It is straightforward to see that $I(T)$ is an ideal of $k[x_1, \dots, x_n]$. For $1 \leq i \leq n$ set

$$f_i = \prod_{s \in T_i} (x_i - s).$$

These are elements of $k[x_i]$, and obviously lie in $I(T)$.

Lemma 1.2.20 $\{f_1, \dots, f_n\}$ is a Gröbner basis of $I(T)$ (with respect to any monomial ordering).

PROOF. The leading monomial of f_i is $x_i^{t_i}$. Pairwise they have no common factors; therefore the f_i form a Gröbner basis of the ideal they generate.

We need to show that the f_i generate $I(T)$. For that we use induction on n . For $n = 1$ it is obvious, so suppose $n > 1$.

Let $f \in I(T)$ and write $f = g_0 + g_1 x_n + \dots + g_r x_n^r$, with $g_i \in k[x_1, \dots, x_{n-1}]$. Suppose that $f \notin \langle f_1, \dots, f_n \rangle$, and choose f with this property such that r is minimal.

First suppose that $g_i(u_1, \dots, u_{n-1}) = 0$ for all i and $(u_1, \dots, u_{n-1}) \in T_1 \times \dots \times T_{n-1}$. Then by induction the g_i lie in the ideal $\langle f_1, \dots, f_{n-1} \rangle \subset k[x_1, \dots, x_{n-1}]$. Hence $f \in \langle f_1, \dots, f_{n-1} \rangle \subset k[x_1, \dots, x_n]$. So we obtain a contradiction.

Now assume that there is a $(u_1, \dots, u_{n-1}) \in T_1 \times \dots \times T_{n-1}$ with $g_j(u_1, \dots, u_{n-1}) \neq 0$ for a certain j . Set $\hat{f} = f(u_1, \dots, u_{n-1}, x_n)$. Then \hat{f} is not the zero polynomial. But it vanishes on T_n . Hence it is a multiple of f_n . In particular, $r \geq t_n$. Now set $\tilde{f} = f - g_r x_n^{r-t_n} f_n$. Then $\tilde{f} \in I(T)$, but $\tilde{f} = h_0 + \dots + h_{r-1} x_n^{r-1}$, with $h_i \in k[x_1, \dots, x_n]$. So again we obtain a contradiction.

The conclusion is that $I(T) = \langle f_1, \dots, f_n \rangle$. □

Theorem 1.2.21 (Alon's non-vanishing theorem) Let $p \in k[x_1, \dots, x_n]$ be of degree

$$\sum_{i=1}^n (t_i - 1).$$

Suppose that the coefficient of $x_1^{t_1-1} \dots x_n^{t_n-1}$ in p is nonzero. Then there is $u = (u_1, \dots, u_n) \in T$ with $p(u) \neq 0$.

PROOF. We use a degree compatible term ordering, i.e., if $|\alpha| < |\beta|$ then $x^\alpha < x^\beta$.

Suppose that p vanishes on all elements of T . Then $p \in I(T)$. So by the previous lemma we can write

$$p = u_{i_1} f_{i_1} + \dots + u_{i_m} f_{i_m}$$

where the u_{i_k} are terms, and

$$\text{LM}(p) = \text{LM}(u_{i_1} f_{i_1}) > \text{LM}(u_{i_2} f_{i_2}) > \dots > \text{LM}(u_{i_m} f_{i_m}).$$

In particular this means that the degree of every $\text{LM}(u_{i_1}f_{i_1})$ is at most $\sum_{i=1}^n(t_i - 1)$.

Note that $\deg(f_i) = t_i$. Suppose that the monomial $x_1^{t_1-1} \cdots x_n^{t_n-1}$ occurs in a $u_{i_k}f_{i_k}$. Then the degree of $u_{i_k}f_{i_k}$ has to exceed $\sum_{i=1}^n(t_i - 1)$, which is excluded. It follows that $x_1^{t_1-1} \cdots x_n^{t_n-1}$ occurs in no $u_{i_k}f_{i_k}$, which is a contradiction. \square

Next we give an application in combinatorics of Alon's non-vanishing theorem.

Let p be a prime and $A, B \subset \mathbb{F}_p$. Then we define

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

Theorem 1.2.22 (Cauchy-Davenport) *Let A, B be non-empty subsets of \mathbb{F}_p . Then*

$$|A + B| \geq \min(p, |A| + |B| - 1).$$

PROOF. First assume that $|A| + |B| \leq p + 1$. Suppose that there exists a set $C \subset \mathbb{F}_p$ with $|C| = |A| + |B| - 2$, and $A + B \subset C$. (In other words, suppose that $|A + B| < |A| + |B| - 1$.) Put

$$f = \prod_{c \in C} (x + y - c) \in \mathbb{F}_p[x, y].$$

Set $T_1 = A$ and $T_2 = B$. Then f vanishes on $T = T_1 \times T_2$. Also $\deg(f) = |C| = t_1 - 1 + t_2 - 1$. Furthermore, the coefficient of $x^{t_1-1}y^{t_2-1}$ is

$$\binom{t_1 - 1 + t_2 + 1}{t_1 - 1}.$$

But this is nonzero in \mathbb{F}_p as $t_1 - 1 + t_2 - 1 = |A| + |B| - 2 \leq p - 1 < p$. So we have obtained a contradiction with Alon's theorem.

If $|A| + |B| > p + 1$; then we take subsets $A' \subset A$, $B' \subset B$ such that $|A'| + |B'| = p + 1$. Then

$$|A + B| \geq |A' + B'| \geq p.$$

(So $|A + B| = p$.) \square

1.2.4 An application in cryptography: Polly-cracker

Here we briefly describe a cryptosystem based on the fact that it is difficult in general to compute Gröbner bases.

Alice wants to receive secret messages from Bob. To this end, Alice chooses a finite field \mathbb{F}_q (for example \mathbb{F}_2) and works in $R = \mathbb{F}_q[x_1, \dots, x_n]$

Alice chooses a $y \in \mathbb{F}_q^n$ and some polynomials $f_1, \dots, f_s \in R$ with $f_i(y) = 0$. The public key is f_1, \dots, f_s and the secret key is y .

In order to send a message Bob first encodes it as an $m \in \mathbb{F}_q$, and then chooses polynomials $h_1, \dots, h_s \in R$. Finally he sends the polynomials $f = m + \sum h_i f_i$ to Alice.

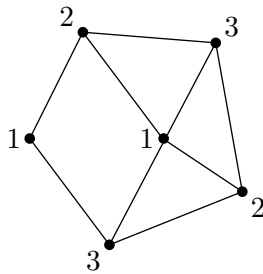
Alice can read the message, because $f(y) = m + \sum h_i f_i(y) = m$.

With Gröbner bases one can break the system. If G is a Gröbner basis of $\langle f_1, \dots, f_s \rangle$ then $f^G = m$.

So Alice needs to choose the polynomials f_i in such a way that a Gröbner basis of the ideal $\langle f_1, \dots, f_s \rangle$ is very hard to compute.

A strategy to choose the f_i is the following: One takes a problem that is known to be very hard to solve (a so-called NP-complete problem, for example), and one reformulates a particularly difficult instance of that problem in terms of polynomial equations. Then one can realistically hope that a Gröbner basis is hard to compute.

For example, we can consider the 3-colouring problem from graph theory. A graph is given as $\Gamma = (V, E)$ (V : vertices, E : edges). A 3-colouring is a function $\varphi: V \rightarrow \{1, 2, 3\}$ with the property that if $(v, w) \in E$ then $\varphi(v) \neq \varphi(w)$. For example:



Now finding a 3-colouring is in general a very difficult problem (it is NP-complete).

Now we take \mathbb{F}_2 as base field, and the variables $t_{v,i}$ for $v \in V$ and $i \in \{1, 2, 3\}$. Let $B = B_1 \cup B_2 \cup B_3$ where

$$\begin{aligned} B_1 &= \{t_{v,1} + t_{v,2} + t_{v,3} - 1 \mid v \in V\} \\ B_2 &= \{t_{v,i}t_{v,j} \mid v \in V, 1 \leq i < j \leq 3\} \\ B_3 &= \{t_{u,i}t_{v,i} \mid (u, v) \in E, 1 \leq i \leq 3\}. \end{aligned}$$

Set $t_{v,i} = 1$ if the vertex v has colour i , and $t_{v,i} = 0$ otherwise. This defines a zero of all polynomials in B . Conversely, if we have a zero of all elements of B , then we can find a 3-colouring. Indeed, from the vanishing of the elements of $B_1 \cup B_2$ we find the colour of every vertex; and the vanishing of the elements of B_3 expresses that two adjacent vertices cannot have the same colour.

Experiments show that it is difficult to break the Polly Cracker system with Gröbner bases. However, there are other methods to try and do that, that make the system insecure. For example, observe that if $f = m + \sum_i h_i f_i$ is the message, then

$$m = f(0) - \sum_i h_i(0)f_i(0).$$

So it is enough to know the constant terms of the h_i in order to reconstruct the message. Sometimes it is possible to get the constant terms by studying how it is possible to obtain the polynomial f from the f_i .

1.3 Exercises

1. Consider the order $<_{\text{rlex}}$ defined by $x^\alpha <_{\text{rlex}} x^\beta$ if $\alpha_k < \beta_k$ where k is maximal with $\alpha_k \neq \beta_k$. Prove that $<_{\text{rlex}}$ is a monomial order.
2. We use $<_{\text{glex}}$ (graded lexicographical). Let $f_1 = y^2 - z$, $f_2 = z^3 - y$, $f_3 = z^2 - 1$, and $g = xy^2z^2 + xy - yz$. Find all possible remainders of g modulo f_1, f_2, f_3 .

3. Compute $S(f, g)$ for $f = 4x^2z - 7y^2$, $g = xyz^2 + 3xz^4$, and $f = x^4y - z^2$, $g = 3xz^2 - y$, using the order $<_{\text{lex}}$.
4. Let $g_1 = x - z^2, g_2 = y - z^3 \in k[x, y, z]$. Prove that $G = \{g_1, g_2\}$ is a Gröbner basis with respect to $<_{\text{lex}}$. Prove that it is not a Gröbner basis with respect to $<_{\text{glex}}$. Compute a Gröbner basis of the ideal generated by G with respect to $<_{\text{glex}}$.
5. Let $g_1 = x^2y - 1, g_2 = xy^2 - x$. Compute a Gröbner basis of the ideal $I \subset k[x, y]$ generated by g_1, g_2 , with respect to $<_{\text{glex}}$.
6. Let $I \subset k[x, y, z]$ be the ideal generated by $g_1 = x^2yz - yz - x$, $g_2 = xy^2z - xy - y$, $g_3 = xyz^2 - xy - z$. We use the order $<_{\text{glex}}$.
 - (a) Compute $g_4 = S(g_1, g_2)$ (reduced modulo g_1, g_2, g_3), and $g_5 = S(g_1, g_3)$ (reduced modulo g_1, \dots, g_4).
 - (b) Compute $S(g_2, g_5)$, reduced modulo g_1, \dots, g_5 .
 - (c) Prove that I is also generated by $h_1 = x^2z^2 - z^2 - x$, $h_2 = xz^3 - xz - z$, $h_3 = y - z$.
 - (d) Compute $h_4 = S(h_1, h_2)$, $h_5 = S(h_1, h_4)$, $h_6 = S(h_2, h_5)$.
 - (e) Prove that I is also generated by h_2, h_3, h_5, h_6 . Show that they form a Gröbner basis of I .
 - (f) Find $u_1, u_2, u_3 \in k[x, y, z]$ with $y - z = u_1g_1 + u_2g_2 + u_3g_3$
7. Let $g_1 = x^2 + 2y^2 - 3$, $g_2 = xy - y^2 + 3$ in $\mathbb{C}[x, y]$. Let I be the ideal generated by g_1, g_2 .
 - (a) Find a Gröbner basis of I with respect to $<_{\text{lex}}$.
 - (b) Find a Gröbner basis of $I \cap \mathbb{C}[y]$.
 - (c) Find all solutions of the equations $g_1 = g_2 = 0$.
8. Let $g_1 = x^2 + y^2 + z^2 - 4$, $g_2 = x^2 + 2y^2 - 5$, $g_3 = xz - 1$ in $\mathbb{C}[x, y, z]$. Let $I = \langle g_1, g_2, g_3 \rangle$.
 - (a) We use the monomial order $<_{\text{lex}}$. Compute $g_4 = S(g_1, g_2)$ (reduced modulo g_1, g_2, g_3), and $g_5 = S(g_1, g_3)$ (reduced modulo g_1, \dots, g_4), and $g_6 = S(g_3, g_5)$ (reduced modulo g_1, \dots, g_5).
 - (b) Prove that g_4, g_5, g_6 generate I , and that they form a Gröbner basis of I .
 - (c) Compute a Gröbner basis of $I \cap \mathbb{C}[y, z]$ and of $I \cap \mathbb{C}[z]$.
 - (d) Find all solutions of $g_1 = g_2 = g_3 = 0$.
9. Let $I \subset \mathbb{C}[x, y, z]$ be the ideal generated by $g_1 = x^2yz - yz - x$, $g_2 = xy^2z - xy - y$, $g_3 = xyz^2 - xy - z$. Let $f_1 = x - z^4 + z^2$, $f_2 = y - z$, $f_3 = z^7 - 2z^5 + z^3 - z$.
 - (a) Prove that $I = \langle f_1, f_2, f_3 \rangle$ (hint: use the result of exercise 6).
 - (b) Show that $\{f_1, f_2, f_3\}$ is a Gröbner basis of I with respect to the order $<_{\text{lex}}$ (with $x >_{\text{lex}} y >_{\text{lex}} z$).
 - (c) Prove that f_3 is square free (one can compute $\text{gcd}(f_3, f_3')$).
 - (d) How many points $a = (a_1, a_2, a_3) \in \mathbb{C}^3$ are there with $f_1(a) = f_2(a) = f_3(a) = 0$?
10. Let $I = \langle x^2, y^2 \rangle \subset k[x, y]$. Prove that $x + y \in \sqrt{I}$, by computing a Gröbner basis.

11. (Rational parametrisation.) Let

$$C = \left\{ \left(\frac{2t}{t^2 + 1}, \frac{1 - t^2}{t^2 + 1} \right) \mid t \in \mathbb{C}, t^2 \neq -1 \right\}.$$

Let $f_1 = (t^2 + 1)y_1 - 2t$, $f_2 = (t^2 + 1)y_2 - 1 + t^2 \in \mathbb{C}[t, y_1, y_2]$, and set

$$\Gamma = \{(s, u_1, u_2) \in \mathbb{C}^3 \mid f_1(s, u_1, u_2) = f_2(s, u_1, u_2) = 0\}.$$

- (a) Let $\pi_1 : \mathbb{C}^3 \rightarrow \mathbb{C}^2$ be defined by $\pi_1(s, u_1, u_2) = (u_1, u_2)$. Prove that $\pi_1(\Gamma) = C$.
- (b) Use the order $<_{\text{lex}}$, with $t >_{\text{lex}} y_1 >_{\text{lex}} y_2$. Compute $S(f_1, f_2)$ (reduced modulo f_1, f_2) and obtain f_3 . In the same way compute $S(f_2, f_3)$ to obtain f_4 and $S(f_3, f_4)$ to obtain f_5 .
- (c) Prove that $\{f_3, f_4, f_5\}$ is a Gröbner basis of the ideal generated by f_1, f_2 , with respect to $<_{\text{lex}}$.
- (d) Find an ideal $J \subset \mathbb{C}[y_1, y_2]$ such that $V(J)$ is the closure of C .

12. Consider

$$S = \{(uv, uv^2, u^2) \in \mathbb{C}^3 \mid u, v \in \mathbb{C}\}.$$

A Gröbner basis of $\langle uv - x, uv^2 - y, u^2 - z \rangle$, with respect to the order $<_{\text{lex}}$ with $u >_{\text{lex}} v >_{\text{lex}} x >_{\text{lex}} y >_{\text{lex}} z$ is

$$\{u^2 - z, uv - x, ux - vz, uy - x^2, v^2z - x^2, vx - y, vyz - x^3, x^4 - y^2z\}.$$

Find an ideal $I \subset \mathbb{C}[x, y, z]$ such that the closure of S is $V(I)$. Find points on $V(I)$ that do not lie in S . (So $V(I)$ is strictly bigger than S .)

Chapter 2

Integer Factorisation

There are three types of algorithms that have to do with the problem to factorise an integer in its prime factors:

- 1) primality testing (probabilistic): such a test proves that n is not prime, or that n is prime with high probability (such tests are typically fast),
- 2) algorithms for proving that a given number n , that tests have shown to be probably prime, in fact is prime,
- 3) algorithms for factorising n , knowing that n is not prime.

The first algorithm that comes to mind, to solve all problems, is to divide n by all integers upto \sqrt{n} . For big n this becomes impossibly slow. However, it can be used to eliminate small factors ($\leq 10^5$ for example).

In this chapter we treat some other methods for these three problems. In the first section we describe the Miller-Rabin test, which is one of the most efficient primality tests. The second section has several methods for the last problem, i.e., we have an integer n and we know that it is not prime. The problem is to factorise it. The main algorithms that we describe are CFRAC (based on continued fractions) and ECM (the elliptic curve method). Finally in the last section we briefly touch upon a method for proving that a given number, which is prime with high probability, in fact is prime. This method uses elliptic curves, and is often denoted by the acronym ECPP (elliptic curve primality proving).

2.1 The Miller-Rabin primality test

The Miller-Rabin primality test was invented by Gary Miller



who gave a deterministic version of the test in 1976 (whose correctness depends on the Generalised Riemann Hypothesis), and Michael Rabin



who in the 1980 developed the probabilistic version of the test, that is still in use today.

Here we describe the test, following the article by René Schoof: *Four primality testing algorithms*, Algorithmic number theory; MSRI Publications 44, Cambridge University Press, Cambridge 2008, 101–126.

Lemma 2.1.1 *Let p be an odd prime and $e \geq 1$, x integers. Then $x^2 = 1 \pmod{p^e}$ implies $x = \pm 1 \pmod{p}$.*

PROOF. From $x^2 = 1 \pmod{p^e}$ we get that p^e divides $(x+1)(x-1)$. But p cannot divide both factors as p is odd. Hence p^e divides one of them; whence the statement. \square

Lemma 2.1.2 *Let C_n be the cyclic group of order n , with generator g (so the order of g is n). Let $e \geq 1$ and $d = \gcd(e, n)$. There are exactly d elements $h \in C_n$ with $h^e = 1$.*

PROOF. Write $h = g^j$, where $1 \leq j \leq n$. Then $h^e = 1$ if and only if $g^{ej} = 1$, which is equivalent to $n|ej$. In turn this is the same as saying $\frac{n}{d}|\frac{e}{d}j$. But the latter happens if and only if $\frac{n}{d}|j$. So for j we have the possibilities $j = \frac{n}{d}, 2\frac{n}{d}, \dots, d\frac{n}{d}$. \square

Theorem 2.1.3 *Let $n > 9$ be an integer that is composite and odd. Write $n - 1 = 2^k m$, where $k \geq 1$ and m is odd. Set*

$$B = \{x \in (\mathbb{Z}/n\mathbb{Z})^* \mid x^m = 1 \text{ or there is } i \text{ with } 0 \leq i < k \text{ and } x^{m2^i} = -1\}.$$

Then

$$\frac{|B|}{\varphi(n)} \leq \frac{1}{4}$$

(where $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^*|$).

PROOF. Let $l \in \mathbb{Z}$ be maximal such that 2^l divides $p - 1$ for all prime factors p of n . Set $B' = \{x \in (\mathbb{Z}/n\mathbb{Z})^* \mid x^{m2^{l-1}} = \pm 1\}$. We claim that $B \subset B'$. To see this, let $x \in B$. If $x^m = 1$ then $x^{m2^{l-1}} = 1$. Secondly, suppose that $x^{m2^i} = -1$ for an i with $0 \leq i < k$. By x we also denote the integer congruent to $x \pmod{n}$, and $1 < x < n$. Then $x^{m2^i} = 1 \pmod{p}$ for all primes p that divide n . So for all such p , the exact power of 2 that divides the order of $x \pmod{p}$ is 2^{i+1} . So $2^{i+1}|p-1$ (as $(\mathbb{Z}/p\mathbb{Z})^*$ has $p-1$ elements). It follows that $i < l$. Furthermore, if $i = l-1$ then $x^{m2^{l-1}} = -1$ and if $i < l-1$ then $x^{m2^{l-1}} = 1$. The claim is proved.

Now write $n = \prod_{i=1}^s p_i^{e_i}$, where the p_i are distinct (odd) primes. Then by the chinese remainder theorem we get

$$(\mathbb{Z}/n\mathbb{Z})^* = (\mathbb{Z}/p_1^{e_1}\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/p_s^{e_s}\mathbb{Z})^*.$$

Now the group $(\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*$ is cyclic of order $(p_i - 1)p_i^{e_i-1}$. So by Lemma 2.1.2 the number of $y \in (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*$ with $y^{m2^{l-1}} = 1$ is $\gcd((p_i - 1)p_i^{e_i-1}, m2^{l-1}) = \gcd(p_i - 1, m)2^{l-1}$ (by the definition of l we see that $2^{l-1} | p_i - 1$; furthermore, $p_i \nmid m$). For this reason, the number of $x \in (\mathbb{Z}/n\mathbb{Z})^*$ with $x^{m2^{l-1}} = 1$ is $\prod_{i=1}^s \gcd(p_i - 1, m)2^{l-1}$. Analogously, the number of solutions in $(\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*$ of the equation $X^{m2^l} = 1$ is $\gcd(p_i - 1, m)2^l$. This is twice the number of solutions to the equation $X^{m2^{l-1}} = 1$. So by Lemma 2.1.1, it follows that the number of solutions to the equation $X^{m2^{l-1}} = -1$ is the same as the number of solutions to $X^{m2^{l-1}} = 1$. We conclude that

$$|B'| = 2 \prod_{i=1}^s \gcd(p_i - 1, m)2^{l-1}.$$

Now suppose that $|B|/\varphi(n) > \frac{1}{4}$. Since $B \subset B'$ and $\varphi(n) = \prod_{i=1}^s (p_i - 1)p_i^{e_i-1}$ we get

$$\frac{1}{4} < 2 \prod_{i=1}^s \frac{\gcd(p_i - 1, m)2^{l-1}}{(p_i - 1)p_i^{e_i-1}}. \quad (2.1)$$

Observe that $\gcd(p_i - 1, m)2^{l-1}$ divides $\frac{p_i-1}{2}$. From this it follows that the right hand side of (2.1) is smaller than $2^{1-s} \prod_{i=1}^s \frac{1}{p_i^{e_i-1}}$. So $s \leq 2$. Suppose first that $s = 2$. Then also $e_1 = e_2 = 1$, so that $n = pq$, where p, q are distinct primes. Then (2.1) is equivalent to

$$2 > \frac{p-1}{\gcd(p-1, m)2^l} \frac{q-1}{\gcd(q-1, m)2^l}.$$

Both factors on the right are positive integers; so both have to be 1. Hence $p-1 = \gcd(p-1, m)2^l$ and $q-1 = \gcd(q-1, m)2^l$. Write $p-1 = a_p 2^l$, $q-1 = a_q 2^l$ with a_p, a_q odd. Then a_p, a_q divide m . Consider the equation $pq = 1 + 2^k m$ modulo a_p . Since $p \equiv 1 \pmod{a_p}$ we get $a_p | q-1$. Analogously, $a_q | p-1$. This implies that $a_p = a_q$, which is a contradiction.

It follows that $s = 1$ and $n = p^e$, $e \geq 2$. Then (2.1) is equivalent to

$$4 > \frac{1}{2} \frac{(p-1)p^{e-1}}{\gcd(p-1, m)2^{l-1}} = \frac{p-1}{\gcd(p-1, m)2^l} p^{e-1} \geq p^{e-1}$$

(as $\gcd(p-1, m)2^l$ divides $p-1$). So the only possibility is $p = 3$, $e = 2$, but that contradicts $n > 9$. \square

Algorithm 2.1.4 *Given:* $n > 9$ odd.

We return: FALSE (meaning that n certainly is not prime) or FAIL (meaning that n is likely to be prime).

1. Take a random integer x with $1 < x < n$ (uniform distribution).
2. If $\gcd(x, n) > 1$ return FALSE.

3. If $x^{n-1} \neq 1 \pmod n$ then return FALSE.
4. Write $n - 1 = 2^k m$, with m odd. Compute the minimal $i \geq 0$ with $x^{m2^i} = 1 \pmod n$. If $i > 0$ and $x^{m2^{i-1}} \neq -1 \pmod n$ then return FALSE. Otherwise return FAIL.

Note that the i in Step 4 exists as $x^{m2^k} = 1 \pmod n$ by Step 3.

Theorem 2.1.5 *If the previous algorithm returns FALSE then n is composite. Furthermore, the probability of n being composite and the algorithm returning FAIL is $\leq \frac{1}{4}$.*

PROOF. Suppose that $i > 0$ and $x^{m2^i} = 1 \pmod n$. If n is prime, then by Lemma 2.1.1 we see that $x^{m2^{i-1}} = \pm 1 \pmod n$. But it cannot be 1 as i is minimal. If it is not -1 either, then we conclude that n cannot be prime. If FAIL is returned then either $x^m = 1 \pmod n$ or $x^{m2^{i-1}} = -1 \pmod n$, implying that $x \in B$ (the set from Theorem 2.1.3). (NB: after Step 2 we know that $x \pmod n$ lies in $(\mathbb{Z}/n\mathbb{Z})^*$.) But the probability of a random element lying in B is $\leq \frac{1}{4}$. \square

Example 2.1.6 Let $n = 1729$. This is a Carmichael number, i.e., it is not prime, but $x^{n-1} = 1 \pmod n$ for all positive integers $x < n$. We have $n - 1 = 2^6 \cdot 27$. So we take $m = 27$. Let $x = 3$. Then $x^{27} = 664 \pmod n$ and $x^{27 \cdot 2} = 1 \pmod n$. Hence the Miller-Rabin test shows that n is not prime.

2.2 Factorisation

2.2.1 The method of Fermat

Note that we may assume that n is odd. Set

$$A = \{(a, b) \mid 0 < a \leq b \text{ e } n = ab\}$$

and

$$B = \{(s, t) \mid 0 \leq s < t \text{ e } n = t^2 - s^2\}.$$

There is a bijection $\sigma: A \rightarrow B$ with $\sigma(a, b) = \left(\frac{b-a}{2}, \frac{b+a}{2}\right)$.

Indeed:

$$\left(\frac{b+a}{2}\right)^2 - \left(\frac{b-a}{2}\right)^2 = ab = n$$

and

$$\sigma^{-1}(s, t) = (t - s, t + s).$$

The conclusion is that finding a factorisation $n = ab$ is equivalent to finding s and t with $n = t^2 - s^2$. So we have the following algorithm.

Algorithm 2.2.1 *Given: n odd and non-prime.*

We compute: $a, b > 1$ with $n = ab$.

1. Set $t_0 = \lceil \sqrt{n} \rceil = \min\{k \in \mathbb{Z} \mid k \geq \sqrt{n}\}$.

2. If $t_i^2 - n = s_i^2$ is a square then set $a = (t_i + s_i)$, and $b = (t_i - s_i)$. Otherwise set $t_{i+1} = t_i + 1$, and continue.

This algorithm terminates correctly because if $n = t^2 - s^2$ with $s < t$ then $t^2 = n + s^2$ implies $t > \sqrt{n}$.

Example 2.2.2 Take $n = 39$. then

$$\begin{aligned} t_0 = 7 &\Rightarrow 7^2 - 39 = 10 \\ t_1 = 8 &\Rightarrow 8^2 - 39 = 25 = 5^2 \end{aligned}$$

hence $(8 + 5)(8 - 5) = 39$.

This algorithm works well if s is small, because in that case t is close to \sqrt{n} and hence only a few steps are needed, even if the factors themselves are big. So if one wants to construct a number that is difficult to factor (as in the cryptosystem RSA, for example) then the factors cannot be chosen close to each other.

Pierre de Fermat (1601 - 1665)



was a french lawyer, in mathematics most famous for his work in number theory, although he did not publish most of his results, preferring to mention the problems he had solved to other mathematicians and challenging them to also come up with a solution.

2.2.2 The continued fraction method (CFRAC)

Factor bases

Some more advanced methods use a similar idea to fermat's algorithm. But instead of looking for s and t with $n = t^2 - s^2$, they try to find x and y with

- $x^2 - y^2 = 0 \pmod n$;
- $0 < y < x < n$;
- $x + y \neq n$.

If we have such x, y then we can factorise n because $(x + y)(x - y)$ is divisible by n , but $x + y$ and $x - y$ are not. Hence $a = \gcd(x + y, n)$ and $b = \gcd(x - y, n)$ are nontrivial factors of n (that means $a, b \neq 1$ and $a, b \neq n$).

Example 2.2.3 Take $n = 377$, then

$$335^2 = 256 \pmod{377} = 16^2 \pmod{377}$$

so with $x = 335$ and $y = 16$ we get

$$\text{mcd}(335 + 16, n) = 13 \quad \text{e} \quad \text{mcd}(335 - 16, n) = 29$$

and $n = 13 \cdot 29$.

So the idea is to find x with $0 < x < n$ such that $x^2 \bmod n$ is a square y^2 . Not all x can be chosen for this: for example if $n = 15$ and $x = 2$, then $x^2 \bmod n = 4$ but $y = 2$ doesn't work as $y = x$. But also $x = 13$ is not good as $x^2 \bmod n = 4$ but with $y = 2$ we get $x + y = n$. (Note that in this case $x = -2 \bmod n$.) A better try is $x = 7$ with which we find $7^2 \bmod n = 4$ and

$$\text{mcd}(7 + 2, 15) = 3 \quad \text{e} \quad \text{mcd}(7 - 2, 15) = 5.$$

Notation. By $x \bmod n$ we denote the unique integer with $-\frac{n}{2} < s \leq \frac{n}{2}$, congruent to $x \bmod n$.

Definition 2.2.4 A set $B = \{p_0 = -1, p_1, \dots, p_r\}$ with p_i prime, is called a factor base.

Definition 2.2.5 An integer k is called a B -number (or B -smooth) if $k = p_0^{e_0} p_1^{e_1} \cdots p_r^{e_r}$.

In order to find x such that $x^2 \bmod n$ is a square we try the following scheme:

- we choose a factor base;
- we generate a lot of integers b_i such that $b_i^2 \bmod n$ is B -smooth.

The aim is to find a product $b_{i_1} \cdots b_{i_m}$ such that

$$(b_{i_1} \cdots b_{i_m})^2 \bmod n = p_0^{e_0} p_1^{e_1} \cdots p_r^{e_r}$$

with e_i even for all i , and hence $p_1^{e_1} \cdots p_r^{e_r} = y^2$.

Example 2.2.6 Like in Example 2.2.3 we take $n = 377$ and $B = \{-1, 2\}$. Then

$$\begin{aligned} 19^2 &= -16 \bmod n \text{ and } -16 = p_0 p_1^2 \\ 97^2 &= -16 \bmod n \text{ and } -16 = p_0 p_1^2. \end{aligned}$$

Therefore $(19 \cdot 97)^2 = p_0^2 p_1^4 \bmod n$. So we set $x = 19 \cdot 97 = 335 \bmod n$, and $y = 16$. Indeed, $\text{mcd}(335 + 16, 377) = 13$ and $\text{mcd}(335 - 16, 377) = 29$ are factors of n .

So the idea is to combine the factorisations of several $b_i \bmod n$ such that a square arises. A famous method for generating b_i with good properties for this purpose is based on the classical theory of continued fractions.

Continued fractions

A continued fraction is an expression of the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4}}}}$$

with $a_i \in \mathbb{Z}$, $a_0 > 0$ and $a_i \geq 1$ for $i \geq 1$.

Notation. The expression above is denoted by $[a_0; a_1, a_2, a_3, a_4]$.

More in general we say that a continued fraction is a sequence of the form $[a_0; a_1, \dots, a_n]$, with $a_i \in \mathbb{Z}$, $a_0 \geq 0$, $a_i \geq 1$ for $i \geq 1$. This corresponds to a rational number γ , which is defined as follows. If $n = 0$ then $\gamma = a_0$. If $n > 0$, then let γ' be the rational number corresponding to $[a_1; a_2, \dots, a_n]$, and set

$$\gamma = a_0 + \frac{1}{\gamma'}.$$

Let $x > 0$ in \mathbb{R} . We want to find a_0, a_1, \dots such that the continued fractions $[a_0; a_1, \dots, a_k]$ are good approximations of x ; we want that they converge to x when $k \rightarrow \infty$ (it is not clear yet that it is always possible to find such a_i , we will show that in this section).

The a_i , if they exist, must be defined in the following way: Write $x = a_0 + \frac{1}{x_1}$ with $a_0 = \lfloor x \rfloor = \max\{n \in \mathbb{Z} \mid n \leq x\}$.

Set $x_1 = \frac{1}{x - a_0}$. Next, write $x_1 = a_1 + \frac{1}{x_2}$ with $a_1 = \lfloor x_1 \rfloor$ and hence $x_2 = \frac{1}{x_1 - a_1}$ and so on.

More formally, set $x_0 = x$ and for $i \geq 0$:

$$\begin{aligned} a_i &= \lfloor x_i \rfloor, \\ x_{i+1} &= \frac{1}{x_i - a_i}. \end{aligned}$$

Example 2.2.7 Let $x = \sqrt{2}$. Then $x_0 = x$ and

$$\begin{aligned} a_0 = 1 &\Rightarrow x_1 = \frac{1}{\sqrt{2} - 1} = 1 + \sqrt{2} \\ a_1 = 2 &\Rightarrow x_2 = \frac{1}{\sqrt{2} - 1} = 1 + \sqrt{2} \\ a_2 = 2 &\Rightarrow x_3 = \frac{1}{\sqrt{2} - 1} = 1 + \sqrt{2} \end{aligned}$$

continuing like this we find $a_i = 2$ for all $i \geq 1$. Now, for example,

$$[1; 2, 2, 2, 2] = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}} = \frac{41}{29}.$$

We have that $2 - (\frac{41}{29})^2 = \frac{1}{841}$; so it is a reasonable approximation of $\sqrt{2}$.

Lemma 2.2.8 For $x \in \mathbb{Q}$ the algorithm for computing the continued fraction corresponding to x terminates (that is, at a certain point we find $x_i - a_i = 0$). The resulting continued fraction is equal to x .

PROOF. Write $x = \frac{a}{b}$ and $a = q_1b + r_1$ with $0 \leq r_1 < b$. Then

$$a_0 = q_1 \quad \text{e} \quad x_1 = \frac{1}{\frac{a}{b} - q_1} = \frac{b}{r_1}.$$

Now write $b = q_2r_1 + r_2$ with $0 \leq r_2 < r_1$ and get

$$a_1 = q_2 \quad \text{e} \quad x_2 = \frac{r_1}{r_2}.$$

So we find a sequence of r_i with $0 \leq r_{i+1} < r_i$; therefore the algorithm must terminate.

By induction on the number of steps one shows that the continued fraction that comes out is equal to x . \square

In general, when the continued fraction for x produces the integers a_0, a_1, \dots , we write $\frac{b_i}{c_i} = [a_0; a_1, \dots, a_i]$ where $\text{mcd}(b_i, c_i) = 1$. These rational numbers are called the *convergents* of the continued fraction for x .

Example 2.2.9 For $x = \sqrt{2}$ we get

$$1 + \frac{1}{2 + \frac{1}{2}} = \frac{7}{5} = \frac{b_2}{c_2}.$$

Theorem 2.2.10 Let a_0, a_1, \dots , be the integers that are produced by the algorithm for computing the continued fraction for $x \in \mathbb{R}$, $x > 0$. Set

$$\begin{aligned} \beta_0 &= a_0, & \beta_1 &= a_0a_1 + 1, & \beta_i &= a_i\beta_{i-1} + \beta_{i-2}, \\ \gamma_0 &= 1, & \gamma_1 &= a_1, & \gamma_i &= a_i\gamma_{i-1} + \gamma_{i-2}, \end{aligned}$$

for $i \geq 2$. Then

- 1) $\beta_i\gamma_{i-1} - \beta_{i-1}\gamma_i = (-1)^{i-1}$, $i \geq 1$;
- 2) $\text{mcd}(\beta_i, \gamma_i) = 1$;
- 3) $b_i = \beta_i$ e $c_i = \gamma_i$ dove $b_i, c_i \in \mathbb{Z}_{\geq 0}$ sono tali che $\text{mcd}(b_i, c_i) = 1$ e $\frac{b_i}{c_i} = [a_0; a_1, \dots, a_i]$.

PROOF. 1) We use induction. For $i = 1$ we have

$$\beta_1\gamma_0 - \beta_0\gamma_1 = (a_0a_1 + 1)1 - a_0a_1 = 1 = (-1)^0.$$

For $i \geq 1$ we get, using the induction hypothesis

$$\begin{aligned} \beta_{i+1}\gamma_i - \beta_i\gamma_{i+1} &= (a_{i+1}\beta_i + \beta_{i-1})\gamma_i - \beta_i(a_{i+1}\gamma_i + \gamma_{i-1}) \\ &= \beta_{i-1}\gamma_i - \beta_i\gamma_{i-1} \\ &= -(-1)^{i-1} \\ &= (-1)^i. \end{aligned}$$

2) If p divides γ_i and β_i then by 1) it also divides $(-1)^{i-1}$, which is impossible for a prime p .

3) Also here we use induction on i . For $i = 0$ we get

$$\frac{b_0}{c_0} = a_0 = \frac{\beta_0}{\gamma_0}$$

and for $i = 1$

$$\frac{b_1}{c_1} = a_0 + \frac{1}{a_1} = \frac{\beta_1}{\gamma_1}.$$

Note that $[a_0; a_1, \dots, a_i, a_{i+1}]$ is obtained from $[a_0; a_1, \dots, a_i]$ by substituting $a_i + \frac{1}{a_{i+1}}$ for a_i . By induction we now have

$$[a_0; a_1, \dots, a_i] = \frac{\beta_i}{\gamma_i} = \frac{a_i \beta_{i-1} + \beta_{i-2}}{a_i \gamma_{i-1} + \gamma_{i-2}}$$

hence

$$\begin{aligned} [a_0; a_1, \dots, a_{i+1}] &= \frac{\left(a_i + \frac{1}{a_{i+1}}\right) \beta_{i-1} + \beta_{i-2}}{\left(a_i + \frac{1}{a_{i+1}}\right) \gamma_{i-1} + \gamma_{i-2}} \\ &= \frac{a_{i+1}(a_i \beta_{i-1} + \beta_{i-2}) + \beta_{i-1}}{a_{i+1}(a_i \gamma_{i-1} + \gamma_{i-2}) + \gamma_{i-1}} \\ &= \frac{a_{i+1} \beta_i + \beta_{i-1}}{a_{i+1} \gamma_i + \gamma_{i-1}} \\ &= \frac{\beta_{i+1}}{\gamma_{i+1}}. \end{aligned}$$

It follows that $\beta_{i+1} = b_{i+1}$ and $\gamma_{i+1} = c_{i+1}$ (since $\text{mcd}(\beta_{i+1}, \gamma_{i+1}) = 1$). □

Remark 2.2.11 It follows that $b_i c_{i-1} - c_i b_{i-1} = (-1)^{i-1}$. Divide by $c_i c_{i-1}$ to get

$$\frac{b_i}{c_i} - \frac{b_{i-1}}{c_{i-1}} = \frac{(-1)^{i-1}}{c_i c_{i-1}}$$

and hence

$$\left| \frac{b_i}{c_i} - \frac{b_{i-1}}{c_{i-1}} \right| = \frac{1}{c_i c_{i-1}}.$$

But $c_i > c_{i-1} > \dots \geq 1$ so

$$\left| \frac{b_i}{c_i} - \frac{b_{i-1}}{c_{i-1}} \right| \longrightarrow 0.$$

Theorem 2.2.12 *The notation is as in Theorem 2.2.10. Then $\frac{b_i}{c_i}$ converges to x when $i \rightarrow \infty$.*

PROOF. For $i \geq 0$ we set $\tilde{x}_i = x_i - a_i$. Then we claim that

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_i + \tilde{x}_i}}}.$$

Indeed, for $i = 0$ we have $a_0 + \tilde{x}_0 = x_0 = x$. For $i > 0$ we use induction; by the induction hypothesis we get

$$a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots + \frac{1}{a_i + \tilde{x}_i}}} = \tilde{x}_1,$$

which establishes the claim.

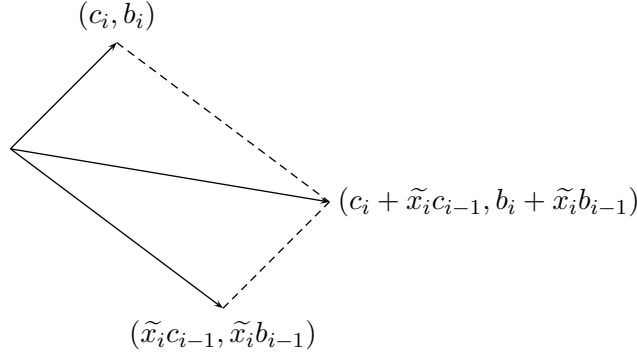
Conclusion: x is obtained from $[a_0; a_1, \dots, a_{i+1}]$ by substituting $\frac{1}{\tilde{x}_i}$ for a_{i+1} . But

$$[a_0; a_1, \dots, a_{i+1}] = \frac{b_{i+1}}{c_{i+1}} = \frac{a_{i+1}b_i + b_{i-1}}{a_{i+1}c_i + c_{i-1}}$$

hence

$$x = \frac{\frac{1}{\tilde{x}_i}b_i + b_{i-1}}{\frac{1}{\tilde{x}_i}c_i + c_{i-1}} = \frac{b_i + \tilde{x}_i b_{i-1}}{c_i + \tilde{x}_i c_{i-1}}.$$

Now we observe the following. If we take a vector from the origin to a point (s, t) , then $\frac{t}{s}$ is its slope. Now we take the vectors that go from the origin to the points (c_i, b_i) and $(\tilde{x}_i c_{i-1}, \tilde{x}_i b_{i-1})$. Summing them we get the vector that goes to the point $(c_i + \tilde{x}_i c_{i-1}, b_i + \tilde{x}_i b_{i-1})$.



Hence $\frac{b_i + \tilde{x}_i b_{i-1}}{c_i + \tilde{x}_i c_{i-1}}$ lies between $\frac{b_i}{c_i}$ and $\frac{b_{i-1}}{c_{i-1}}$.

From this observation we have that the distance between $\frac{b_i}{c_i}$ and $\frac{b_{i+1}}{c_{i+1}}$ goes to 0. Hence $\frac{b_i}{c_i}$ converges to a limit that has to be x (as x is always in the middle). \square

Factorisation with continued fractions

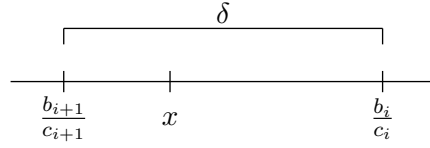
Lemma 2.2.13 *Let $x \in \mathbb{R}$, $x > 1$. Let $\frac{b_i}{c_i}$ be the convergents of the continued fraction for x . Then*

$$|b_i^2 - c_i^2 x^2| < 2x.$$

PROOF. By the observation in the proof of Theorem 2.2.12 we have

$$\left| \frac{b_{i+1}}{c_{i+1}} - \frac{b_i}{c_i} \right| = \frac{1}{c_{i+1}c_i}.$$

Set $\delta = \frac{1}{c_{i+1}c_i}$.



Now

$$|b_i^2 - c_i^2 x^2| = |c_i^2 x^2 - b_i^2| = c_i^2 \left| x - \frac{b_i}{c_i} \right| \left| x + \frac{b_i}{c_i} \right|.$$

And $\left| x - \frac{b_i}{c_i} \right| < \delta$ while

$$\left| x + \frac{b_i}{c_i} \right| = x + \frac{b_i}{c_i} = 2x + \frac{b_i}{c_i} - x < 2x + \delta$$

hence

$$|b_i^2 - c_i^2 x^2| < c_i^2 \delta (2x + \delta).$$

Therefore

$$|b_i^2 - c_i^2 x^2| - 2x < 2x \left(-1 + \delta c_i^2 + \frac{\delta^2 c_i^2}{2x} \right) = 2x \left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{2x c_{i+1}^2} \right).$$

But $\frac{1}{2x c_{i+1}^2} < \frac{1}{c_{i+1}}$ because $x > 1$ and $c_{i+1} \geq 1$. Moreover, $c_i + 1 \leq c_{i+1}$ because $c_i < c_{i+1}$. Hence

$$|b_i^2 - c_i^2 x^2| - 2x < 2x \left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{c_{i+1}} \right) < 2x \left(-1 + \frac{c_{i+1}}{c_{i+1}} \right) = 0$$

and therefore

$$|b_i^2 - c_i^2 x^2| < 2x.$$

□

Corollary 2.2.14 *Let $n > 16$ be an integer and $\frac{b_i}{c_i}$ the convergents of the continued fraction of \sqrt{n} . Then $|b_i^2 \pmod n| < 2\sqrt{n}$.*

PROOF. By Lemma 2.2.13 we have $|b_i^2 - c_i^2 n| < 2\sqrt{n}$. But $b_i^2 - c_i^2 n$ is an integer congruent to $b_i^2 \pmod n$ and contained in the interval $(-2\sqrt{n}, 2\sqrt{n})$. Hence it is contained in the interval $(-\frac{n}{2}, \frac{n}{2})$ (as $n > 16$). It follows that $b_i^2 - c_i^2 n$ is equal to $b_i^2 \pmod n$. □

Idea of the factorisation algorithm: Let $B = \{p_0 = -1, p_1, \dots, p_r\}$ be a factor base. Let $\frac{b_i}{c_i}$ be the convergents of the continued fraction of \sqrt{n} .

Then $b_i^2 \pmod n$ is always “small”. So among the $b_i^2 \pmod n$ there are often B -smooth numbers. So one can reasonably hope to quickly find b_{i_1}, \dots, b_{i_t} such that $(b_{i_1} \cdots b_{i_t})^2 \pmod n$ is a square.

Algorithm 2.2.15 *Given: n which is not prime.*

Calculate: two factors of n .

1. Choose a factor base $B = \{p_0 = -1, p_1, \dots, p_r\}$.
2. Compute the continued fraction for \sqrt{n} and find a_0, a_1, a_2, \dots .
3. Compute $b_0 = a_0$, $b_1 = a_0a_1 + 1$, $b_{i+1} = a_{i+1}b_i + b_{i-1}$.
4. For the b_i such that $b_i^2 \pmod n$ is B -smooth we write $b_i^2 \pmod n = p_0^{e_0} \cdots p_r^{e_r}$ until we find a product $(b_{i_1} \cdots b_{i_t})^2 \pmod n = p_0^{e_0} \cdots p_r^{e_r}$ with e_i even for all i . Set $x = b_{i_1} \cdots b_{i_t}$ and $y = p_0^{\frac{e_0}{2}} \cdots p_r^{\frac{e_r}{2}}$.
5. If we are not unlucky, then $\text{mcd}(n, x + y)$ and $\text{mcd}(n, x - y)$ are factors of n .

Remark 2.2.16 The algorithm is based on heuristic ideas; it is not guaranteed that it manages to factor n . However, we know that n is composite. Hence if we do not find the factorisation in reasonable time we can try again; for example with a larger factor base. Implementations have shown that the algorithm works well in practice.

Example 2.2.17 (Factorisation of Fermat numbers)

The Fermat numbers are $F_n = 2^{2^n} + 1$.

Fermat thought that these were all prime numbers. In fact, F_n is prime for $0 \leq n \leq 4$. However, Euler has shown that

$$F_5 = 641 \cdot 6700417.$$

In 1880 Landry proved that

$$F_6 = 274177 \cdot 67280421310721.$$

And in 1970 Morison and Brillhart, using the continued fraction method on an early computer, got

$$F_7 = 59649589127497217 \cdot P_{22}$$

where P_{22} is a prime of 22 digits. (On my laptop this factorisation now takes less than a millisecond.)

Remark 2.2.18 In order to compute the continued fraction for \sqrt{n} the following observation can help. For $\alpha \in \mathbb{R}$, $\alpha > 0$ and $m \in \mathbb{Z}$, $m > 0$, we have $\lfloor \frac{\alpha}{m} \rfloor = \lfloor \frac{\lfloor \alpha \rfloor}{m} \rfloor$.

PROOF. Write $\alpha = u + \beta$ with u an integer and $\beta \in \mathbb{R}$, $0 \leq \beta < 1$. Write $u = qm + r$ with $0 \leq r < m$. Then

$$\left\lfloor \frac{\alpha}{m} \right\rfloor = \left\lfloor \frac{u}{m} + \frac{\beta}{m} \right\rfloor = \left\lfloor q + \frac{r + \beta}{m} \right\rfloor = q$$

because $r + \beta < m$ and hence $\frac{r + \beta}{m} < 1$. On the other hand,

$$\left\lfloor \frac{\lfloor \alpha \rfloor}{m} \right\rfloor = \left\lfloor \frac{u}{m} \right\rfloor = \left\lfloor q + \frac{r}{m} \right\rfloor = q.$$

□

Example 2.2.19 We want to factorise $n = 33$.

The continued fraction for $\sqrt{33}$:

$$\begin{array}{ll}
 a_0 = 5 & \tilde{x}_0 = x - a_0 = \sqrt{33} - 5 \\
 & x_1 = \frac{1}{\sqrt{33}-5} = \frac{5+\sqrt{33}}{8} \\
 a_1 = 1 & \tilde{x}_1 = x_0 - a_1 = \frac{-3+\sqrt{33}}{8} \\
 & x_2 = \frac{1}{\tilde{x}_1} = \frac{8}{-3+\sqrt{33}} = \frac{3+\sqrt{33}}{3} \\
 a_2 = 2 & \tilde{x}_2 = x_1 - a_2 = \frac{-3+\sqrt{33}}{3} \\
 & x_3 = \frac{3+\sqrt{33}}{8} \\
 a_3 = 1 & \tilde{x}_3 = x_2 - a_3 = \frac{-5+\sqrt{33}}{8} \\
 & x_4 = 5 + \sqrt{33} \\
 a_4 = 10 & \dots
 \end{array}$$

The b_i :

$$\begin{array}{ll}
 b_0 = a_0 = 5 & b_0^2 \pmod{33} = -8 \\
 b_1 = a_0 a_1 + 1 = 6 & b_1^2 \pmod{33} = 3 \\
 b_2 = a_2 b_1 + b_0 = 17 & b_2^2 \pmod{33} = -8.
 \end{array}$$

Hence

$$x = b_0 b_2 = 5 \cdot 17 = 19 \pmod{33} \quad \text{e} \quad y = 8$$

so that

$$\text{mcd}(x + y, n) = \text{mcd}(27, 33) = 3 \quad \text{e} \quad \text{mcd}(x - y, n) = \text{mcd}(11, 33) = 11.$$

Example 2.2.20 Let us factorise $n = 197209$. Taking the b_i modulo n we get

$$\begin{array}{lll}
 a_0 = 444 & b_0 = 444 & b_0^2 \pmod{n} = -73 = -1 \cdot 73 \\
 a_1 = 12 & b_1 = 5329 & b_1^2 \pmod{n} = 145 = 5 \cdot 29 \\
 a_2 = 6 & b_2 = 31428 & b_2^2 \pmod{n} = -37 = -1 \cdot 37 \\
 a_3 = 23 & b_3 = 159316 & b_3^2 \pmod{n} = 720 = 2^4 \cdot 3^2 \cdot 5 \\
 a_4 = 1 & b_4 = 191734 & b_4^2 \pmod{n} = -143 = -11 \cdot 13 \\
 a_5 = 5 & b_5 = 131941 & b_5^2 \pmod{n} = 215 = 5 \cdot 43 \\
 a_6 = 3 & b_6 = 193139 & b_6^2 \pmod{n} = -656 = -2^4 \cdot 41 \\
 a_7 = 1 & b_7 = 127871 & b_7^2 \pmod{n} = 33 = 3 \cdot 11 \\
 a_8 = 26 & b_8 = 165232 & b_8^2 \pmod{n} = -136 = -2^3 \cdot 17 \\
 a_9 = 6 & b_9 = 133218 & b_9^2 \pmod{n} = 405 = 3^4 \cdot 5
 \end{array}$$

Hence $x = b_3 \cdot b_9 = 126308 \pmod{n}$ and $y = 2^2 \cdot 3^3 \cdot 5 = 540$. Now $\text{mcd}(x + y, n) = 991$ and $\text{mcd}(x - y, n) = 199$.

In 1999, Brent, using ECM (Elliptic Curve Method) has found the factorisation of the Fermat number F_{10}

$$45592577 \cdot 6487031809 \cdot 4659775785220018543264560743076778192897 \cdot P_{252}$$

where P_{252} is a prime of 252 digits. (This factorisation now takes about 102 seconds on my laptop.)

Divertimento: solving the Pell equation

Let d be a positive square-free integer. The problem is to find integral solutions $x, y \in \mathbb{Z}$ to the equation

$$x^2 - dy^2 = 1$$

which is called the *Pell equation*. Here we show how this can be done using continued fractions.

First of all we consider the continued fraction of \sqrt{d} . We get real numbers x_0, x_1, \dots and integers a_0, a_1, \dots with

$$x_0 = \sqrt{d}, \quad a_i = \lfloor x_i \rfloor, \quad x_{i+1} = \frac{1}{x_i - a_i}.$$

Moreover we get the convergents $\frac{b_i}{c_i}$.

It is a fact, which we do not prove here, that the continued fraction of \sqrt{d} is periodic. This means that there is an $s > 1$ such that $x_{s+1} = x_1$, which also implies $a_{s+1} = a_1$. We take $s > 1$ to be the smallest such number. This also means that $x_{ks+1} = x_1$ for $k = 1, 2, \dots$. So since $x_i = a_1 + \frac{1}{x_{i+1}}$ we get

$$x_{ks} = a_{ks} + \frac{1}{x_1} = a_1 + \frac{1}{x_1}.$$

Setting $\tilde{x}_i = x_i - a_i$ we have, as seen in the proof of Theorem 2.2.12

$$\sqrt{d} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_i + \tilde{x}_i}}}.$$

Now $a_{ks} + \tilde{x}_{ks} = a_{ks} + \frac{1}{x_1} = a_{ks} + x_0 - a_0 = a_{ks} - a_0 + \sqrt{d}$. Set $u = a_{ks} - a_0$, which lies in \mathbb{Z} . Then we see that we get \sqrt{d} from the ks -th convergent $\frac{b_{ks}}{c_{ks}}$ by substituting $a_{ks} \mapsto u + \sqrt{d}$. For brevity we write $j_0 = ks$. Then $\frac{b_{ks}}{c_{ks}} = (a_{j_0} b_{j_0-1} + b_{j_0-2}) / (a_{j_0} c_{j_0-1} + c_{j_0-2})$. Hence

$$\sqrt{d} = \frac{(u + \sqrt{d})b_{j_0-1} + b_{j_0-2}}{(u + \sqrt{d})c_{j_0-1} + c_{j_0-2}}.$$

Multiplying by the denominator on the right-hand side we get

$$(uc_{j_0-1} + c_{j_0-2})\sqrt{d} + dc_{j_0-1} = b_{j_0-1}\sqrt{d} + ub_{j_0-1} + b_{j_0-2},$$

yielding $uc_{j_0-1} + c_{j_0-2} = b_{j_0-1}$ and $ub_{j_0-1} + b_{j_0-2} = dc_{j_0-1}$. We multiply the first equation by b_{j_0-1} , the second by c_{j_0-1} and subtract. This gives

$$b_{j_0-1}c_{j_0-2} - b_{j_0-2}c_{j_0-1} = b_{j_0-1}^2 - dc_{j_0-1}^2.$$

By Theorem 2.2.10, the left-hand side of this equation is $(-1)^{j_0}$. We conclude:

- if s is even then the $(s - 1)$ -st convergent gives a solution to the Pell equation,
- if s is odd, then the $(2s - 1)$ -st convergent gives a solution to the Pell equation.

Example 2.2.21 Consider $d = 14$. Then for the continued fraction of \sqrt{d} we get

i	a_i	x_{i+1}	b_i	c_i
0	3	$\frac{\sqrt{14+3}}{5}$	3	1
1	1	$\frac{\sqrt{14+2}}{2}$	4	1
2	2	$\frac{\sqrt{14+2}}{5}$	11	3
3	1	$\sqrt{14+3}$	15	4
4	6	$\frac{\sqrt{14+3}}{5}$	101	27

We see that $s = 4$, and therefore $x = 15$, $y = 4$ give a solution to the Pell equation. Indeed, $15^2 - 14 \cdot 4^2 = 1$.

The Pell equation has a long and interesting history. Apparently it was considered by Archimedes who posed the “cattle problem”, whose solution goes via a Pell equation. It was also studied by the indian mathematician Brahmagupta (598 - 670), who was the first to give a method for solving it. Much later Fermat challenged other mathematicians to find a method to solve the equation. He accompanied his challenge with some particularly difficult cases; therefore it is very plausible that Fermat knew how to solve the equation. The challenge was mainly taken up by english mathematicians as Brouncker and Wallis. Wallis published a book, also describing a method to solve the Pell equation, which was due to Brouncker. However, when Euler worked on the equation, he confused Brouncker with Pell, and called it the Pell equation. For this reason it is still called like this, although Pell never worked on it. Euler came up with a rudimentary form of the continued fraction method to solve the equation. Later Lagrange proved rigorously that this method is correct.

There are many interesting questions related to the Pell equation. One is how the solution (x, y) grows with d . For example, when $d = 139$ the smallest solution is $x = 77563250$, $y = 6578829$. For more information on this and other problems we refer to the book by Michael Jacobson and Hugh Williams, *Solving the Pell equation*, Springer Verlag 2009.

2.2.3 The elliptic curve method (ECM)

In this section we describe a method for factoring integers that uses elliptic curves. It is due to H. W. Lenstra, Jr.



Elliptic curves

Let k be a field and $x^3 + ax + b \in k[x]$ a polynomial with three distinct roots. Then

$$E(k) = \{(x, y) \in k^2 \mid y^2 = x^3 + ax + b\} \cup \{O\}$$

where O is a symbol, is called an *elliptic curve*.

Remark 2.2.22 Let $f(x) = x^3 + ax + b$. Then f has three distinct roots if and only if there does not exist $\alpha \in k$ with $f(\alpha) = f'(\alpha) = 0$.

Let $P = (\alpha, \beta) \in E(k)$; we want to construct the line tangent to $E(k)$ in P . To compute the slope we consider the derivative of $y^2 = x^3 + ax + b$, that is

$$2y \frac{dy}{dx} = 3x^2 + a = f'(x).$$

If f has three distinct roots then it is impossible that $\beta^2 = f(\alpha) = 0$ and $f'(\alpha) = 0$ at the same time. So the tangent line has a well defined slope, namely $\frac{3\alpha^2+a}{2\beta}$ ($\beta = 0$ means that the line is vertical).

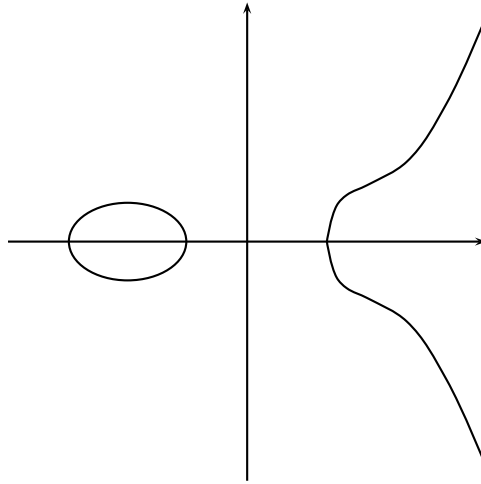
Write $x^3 + ax + b = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ and define

$$\Delta = - \prod_{i < j} (\alpha_i - \alpha_j)^2$$

which is called the *discriminant* of the polynomial $x^3 + ax + b$. The latter has three distinct roots if and only if $\Delta \neq 0$. Furthermore, it is possible to show that $\Delta = 4a^3 + 27b^2$. (We will not go into that here.)

Conclusion: $y^2 = x^3 + ax + b$ defines an elliptic curve if and only if $\Delta = 4a^3 + 27b^2 \neq 0$.

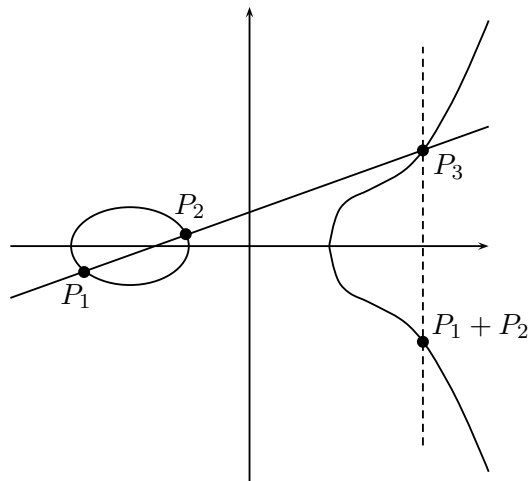
If $k = \mathbb{R}$ we can make a graph of the curve. It will look like this:



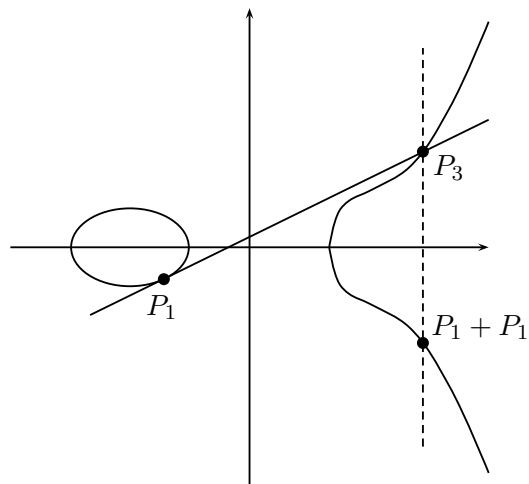
Note that the graph has to be symmetric with respect to the x -axis, as the equation is $y^2 = f(x)$. Hence $(x_0, y_0) \in E(k)$ implies $(x_0, -y_0) \in E(k)$.

One of the main points of elliptic curves is that we can define an addition $+$ on it that makes it into an abelian group. This is done as follows:

- The neutral element is O , i.e., $P + O = O + P = P$ for all $P \in E(k)$.
- If $P_1, P_2 \in E(k)$, $P_1 \neq P_2$ and $P_1, P_2 \neq O$ then we construct the line ℓ through P_1 and P_2 . The line ℓ intersects $E(k)$ in a third point $P_3 = (x_3, y_3)$, and we set $P_1 + P_2 = (x_3, -y_3)$.



- If $P_1 = P_2 \in E(k)$, $P_1 \neq O$, then ℓ will be the tangent line at $E(k)$ through P_1 ; after which we do the same thing.



- If the line ℓ is vertical then the sum will be O .

Theorem 2.2.23 *The set $E(k)$ with the operation $+$ is an abelian group.*

Here we do not prove this theorem; the difficult point is to show the associative law:

$$(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3).$$

Addition formulas

Let $P_1, P_2 \in E(k)$, $P_1 \neq P_2$, $P_1, P_2 \neq O$ where $P_i = (x_i, y_i)$.

The line ℓ through P_1 and P_2 is

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1).$$

Set $m = \frac{y_2 - y_1}{x_2 - x_1}$; we compute the points of intersection of ℓ and the curve $y^2 = x^3 + ax + b$.

We get

$$(m(x - x_1) + y_1)^2 = x^3 + ax + b \iff x^3 - m^2x^2 + ux + v = 0$$

for certain u and v .

Two solutions of this equation are x_1 and x_2 because P_1 and P_2 are on the curve and on the line. Hence

$$x^3 - m^2x^2 + ux + v = (x - x_1)(x - x_2)(x - x_3) = x^3 - (x_1 + x_2 + x_3)x^2 + \dots$$

so $x_1 + x_2 + x_3 = m^2$. Therefore

$$\begin{aligned} x_3 &= m^2 - x_1 - x_2 \\ y_3 &= m(x_3 - x_1) + y_1. \end{aligned}$$

It follows that $P_1 + P_2 = (m^2 - x_1 - x_2, -m(x_3 - x_1) - y_1)$.

We note that the sum does not depend on a and b . However, the points P_1 and P_2 of course do.

If $P_1 = P_2$ then we consider the tangent line ℓ . Take the derivative with respect to x of $y^2 = x^3 + ax + b$, i.e., $2y \frac{dy}{dx} = 3x^2 + a$. We get

- If $y_1 \neq 0$ then

$$m = \frac{3x_1^2 + a}{2y_1}$$

so the line ℓ has equation $y = m(x - x_1) + y_1$. It follows that $P_1 + P_1 = (x_3, -y_3)$ with

$$\begin{aligned} x_3 &= m^2 - 2x_1 \\ y_3 &= m(x_3 - x_1) + y_1. \end{aligned}$$

- If $y_1 = 0$ then $P_1 + P_1 = O$. (Note that it is impossible that $y_1 = 3x_1^2 + a = 0$.)

Example 2.2.24 Take $y^2 = x^3 + 17$ and $P_1 = (-1, 4)$. Then

$$\begin{aligned} m &= \frac{3(-1)^2 + 0}{8} = \frac{3}{8} \\ x_3 &= \frac{9}{64} + 2 = \frac{137}{64} \\ y_3 &= \frac{3}{8} \left(\frac{137}{64} + 1 \right) + 4 = \frac{2651}{512}. \end{aligned}$$

So $P_1 + P_1 = \left(\frac{137}{64}, -\frac{2651}{512} \right)$.

Factorisation with elliptic curves

The main idea is to use the group law on an elliptic curve together with the reduction modulo p of its points.

Let p be a prime, and define

$$A_p = \left\{ \frac{\alpha}{\beta} \in \mathbb{Q} \mid p \text{ does not divide } \beta \right\}.$$

N.B. If p does not divide β then there exists $\gamma \in \mathbb{Z}$ with $\gamma\beta = 1 \pmod{p}$. Write $\gamma = \beta^{-1} \pmod{p}$.

Define $\psi_p: A_p \rightarrow \mathbb{F}_p$, with $\psi_p\left(\frac{\alpha}{\beta}\right) = \alpha\beta^{-1} \pmod{p}$.

Let $a, b \in \mathbb{Z}$ such that $4a^3 + 27b^2 \not\equiv 0 \pmod{p}$. Then

$$E(\mathbb{F}_p) = \{(x, y) \in \mathbb{F}_p^2 \mid y^2 = x^3 + ax + b\} \cup \{O\}$$

is an elliptic curve over \mathbb{F}_p .

Moreover,

$$E(\mathbb{Q}) = \{(x, y) \in \mathbb{Q} \mid y^2 = x^3 + ax + b\} \cup \{O\}$$

is an elliptic curve over \mathbb{Q} .

Define $\varphi_p: E(\mathbb{Q}) \rightarrow E(\mathbb{F}_p)$ in the following way. Let $P = \left(\frac{\alpha}{\beta}, \frac{\gamma}{\delta}\right) \in E(\mathbb{Q})$. If p does not divide β, δ then

$$\varphi_p(P) = \left(\psi_p\left(\frac{\alpha}{\beta}\right), \psi_p\left(\frac{\gamma}{\delta}\right)\right)$$

otherwise $\varphi_p(P) = O$.

Theorem 2.2.25 *The map φ_p is a group homomorphism. In other words, $\varphi_p(P_1 + P_2) = \varphi_p(P_1) + \varphi_p(P_2)$.*

Here we only prove a special case of this theorem.

Definition 2.2.26 *Let p be a prime and $P = \left(\frac{u_1}{v_1}, \frac{u_2}{v_2}\right) \in \mathbb{Q}^2$. Then we say that P is good for p if p does not divide v_1 or v_2 .*

Lemma 2.2.27 *Let $E(k) = \{(x, y) \mid y^2 = x^3 + ax + b\} \cup \{O\}$ be an elliptic curve with $a, b \in \mathbb{Z}$ and $p > 2$ a prime not dividing $\Delta = 4a^3 + 27b^2$. Let $\varphi_p: E(\mathbb{Q}) \rightarrow E(\mathbb{F}_p)$ be the reduction modulo p . Let $P_1, P_2 \in E(\mathbb{Q})$ be two points that are good for p . Then $\varphi_p(P_1 + P_2) = \varphi_p(P_1) + \varphi_p(P_2)$.*

PROOF. Let $P_i = (x_i, y_i)$. Here we also write $x \pmod{p}$ for $\psi_p(x)$.

- Suppose $P_1 \neq P_2$, $\varphi_p(P_1) \neq \varphi_p(P_2)$ and $x_1 \pmod{p} \neq x_2 \pmod{p}$. In this case one uses the same addition formulas over \mathbb{Q} and over \mathbb{F}_p . Since $\text{mod } p$ respects addition and multiplication, we get that $\varphi_p(P_1 + P_2) = \varphi_p(P_1) + \varphi_p(P_2)$. The same reasoning holds when $P_1 = P_2$. (Then of course $\varphi_p(P_1) = \varphi_p(P_2)$).

- Suppose that $P_1 \neq P_2$ and $x_1 \pmod{p} = x_2 \pmod{p}$. Write $x_2 = x_1 + p^r x$ with $x \in \mathbb{Q}$, $x \pmod{p} \neq 0$. Then $y_1^2 \equiv y_2^2 \pmod{p}$ because $x_1^3 + ax_1 + b = x_2^3 + ax_2 + b \pmod{p}$ and hence $y_1 = \pm y_2 \pmod{p}$. We calculate

$$\begin{aligned} y_2^2 &= x_2^3 + ax_2 + b \\ &= (x_1 + p^r x)^3 + a(x_1 + p^r x) + b \\ &= x_1^3 + ax_1 + b + (3x_1^2 + a)p^r x + *p^{r+1} \\ &= y_1^2 + (3x_1^2 + a)p^r x + *p^{r+1} \end{aligned}$$

where $*$ is a rational number we don't want to write down.

Hence

$$(y_2 + y_1)(y_2 - y_1) = (3x_1^2 + a)p^r x + *p^{r+1}.$$

Now we distinguish three cases. In the first we have $y_1 = y_2 \pmod{p}$, and $y_2 + y_1 \not\equiv 0 \pmod{p}$. In this case

$$y_2 - y_1 = (3x_1^2 + a)p^r x (y_2 + y_1)^{-1} + *p^{r+1}$$

for some $*$.

To compute $P_1 + P_2$ in $E(\mathbb{Q})$ we use

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{3x_1^2 + a}{y_2 + y_1} + *p$$

where $x_2 - x_1 = p^r x$.

Hence $m \bmod p = \frac{3x_1^2 + a}{2y_1} \bmod p$ which is exactly the m that is used for computing $\varphi_p(P_1) + \varphi_p(P_2)$. Hence

$$\varphi(P_1 + P_2) = \varphi_p(P_1) + \varphi_p(P_2).$$

In the second case we have $y_1 = -y_2 \bmod p$ and $y_2 - y_1 \neq 0 \bmod p$. Then

$$m = \frac{y_2 - y_1}{p^r x}.$$

Since $r > 0$ we get $\varphi_p(P_1 + P_2) = O$, which is equal to $\varphi_p(P_1) + \varphi_p(P_2)$.

In the third case we have $y_2 + y_1 = 0 \bmod p$ and $y_2 - y_1 = 0 \bmod p$. Then, since $p > 2$, we get $y_1 \bmod p = y_2 \bmod p = 0$. Hence $\varphi_p(P_1) = \varphi_p(P_2) = (\psi_p(x_1), 0)$. Hence $\varphi_p(P_1) + \varphi_p(P_2) = O$. Also $3x_1^2 + a \neq 0 \bmod p$, otherwise $p \mid \Delta$. So p^r is the highest power of p that divides $(y_2 + y_1)(y_2 - y_1)$. Since p divides both factors we get that $y_2 - y_1$ is only divisible by p^s for some $s < r$. Therefore, when computing the sum in $E(\mathbb{Q})$ we use an m of the form $\frac{a}{b}$ with $p \mid b$. It follows that $\varphi_p(P_1 + P_2) = O$. \square

The idea lying behind the elliptic curve method for factorising integers is the following. The group $E(\mathbb{F}_p)$ is finite, so for $P \in E(\mathbb{F}_p)$ there exists k with

$$\underbrace{P + \dots + P}_k = O.$$

Let $Q \in E(\mathbb{Q})$ with $\varphi_p(Q) = P$. Then by Theorem 2.2.25

$$\varphi(\underbrace{Q + \dots + Q}_k) = \underbrace{P + \dots + P}_k = O.$$

Hence $kQ = Q + \dots + Q$ is equal to 0 or it has coordinates with a denominator divisible by p .

Algorithm 2.2.28 *Given:* a composite number n .

We compute: a factor of n

1. We choose a curve $y^2 = x^3 + ax + b$ with a point P on $E(\mathbb{Q})$.
2. Let $d = \gcd(n, 4a^3 + 27b^2)$. If $d = n$ we choose a different curve. If $d > 1$ and $d < n$ we have found a factor of n .
3. Assume $d = 1$. Set $k = \text{lcm}(2, 3, \dots, K)$ for a certain K .
4. Compute $kP = \left(\frac{u_1}{v_1}, \frac{u_2}{v_2}\right)$. Let $d_i = \gcd(v_i, n)$. If d_i is bigger than 1, and smaller than n , then we have found a factor. Otherwise we retry with a bigger K , or a different curve.

N.B. We know that n is not prime. The method works if

- (let p be a prime dividing n),
- if the order of $\varphi_p(Q)$ divides k then $\varphi_p(kQ) = k\varphi_p(Q) = O$; so if $kQ \neq O$ then p divides the denominator of u or of v .

Theorem 2.2.29 (Hasse) *We have*

$$p + 1 - 2\sqrt{p} \leq |E(\mathbb{F}_p)| \leq p + 1 + 2\sqrt{p}.$$

Remark 2.2.30 If $|E(\mathbb{F}_p)|$ divides k then we find p . For example when $K \geq p + 1 + 2\sqrt{p}$.

Example 2.2.31 Suppose that 61 divides n . Let E_a be the curve $y^2 = x^3 + ax + 3 - a$ and $P = (1, 2)$. We have

a	$ E_a(\mathbb{F}_{61}) $	order of $(P \bmod 61)$
-4	68	68
-3	54	9
-2	76	38
-1	69	69
0	48	12
1	48	4
2	66	33
3	72	6
4	67	67

Now $4P = O \in E_1(\mathbb{Q})$ and $6P = \left(\frac{51825601}{16208676}, *\right) \in E_3(\mathbb{Q})$ and 61 divides 16208676.

Remark 2.2.32 It is enough to compute kP modulo n . Observe that if $rP = \left(\frac{u_1}{v_1}, \frac{u_2}{v_2}\right)$ and $\text{mcd}(v_i, n) \neq 1$, then we have found a factor. Otherwise if $\text{mcd}(v_i, n) = 1$ then there exist c_i, d_i with $c_i v_i + d_i n = 1$ and hence $\frac{u_i}{v_i} \bmod n = u_i c_i \bmod n$. This allows us to work with integer coordinates modulo n .

Example 2.2.33 Lets factorise $n = 39$. Let E be the curve given by $y^2 = x^3 + x - 1$ and $P = (1, 1)$ a point on it. We have $\Delta = 31$ and $\text{mcd}(31, 39) = 1$. Take $k = 6$.

We have $2P = (2, -3)$. Let's compute $4P$:

$$m = \frac{3x_1^2 + a}{y_1} = \frac{12 + 1}{-6} = -\frac{13}{6}$$

and since $\text{mcd}(6, 39) = 3$ we have found a factor.

Example 2.2.34 Now for something more difficult. Take $n = 185$ with the curve of Example 2.2.33 and $k = 8$. Again $\Delta = 31$ and $\text{mcd}(31, 185) = 1$ so $y^2 = x^3 + x - 1$ defines an elliptic curve modulo every prime that divides n .

Also, $2P = (2, -3)$.

Now we compute $4P$. Again $m = -\frac{13}{6}$. Now

$$\begin{array}{rcl} 185 & = & 1 \cdot 185 + 0 \cdot 6 \\ 6 & = & 0 \cdot 185 + 1 \cdot 6 \\ \hline 5 & = & 1 \cdot 185 - 30 \cdot 6 \\ 1 & = & -1 \cdot 185 + 31 \cdot 6 \end{array}$$

hence $6^{-1} \bmod 185 = 31$. It follows that

$$m \bmod 185 = -13 \cdot 31 \bmod 185 = 152 \bmod 185.$$

So

$$\begin{aligned} x_3 &= m^2 - 2x_1 = 152^2 - 2 \cdot 2 \\ &= 160 \bmod 185 \\ y_3 &= m(x_3 - x_1) - y_1 = 152(160 - 2) - 3 \\ &= -37 \bmod 185. \end{aligned}$$

Hence $4P = (160, 37) \bmod 185$.

Now we compute $8P$. We have

$$m = \frac{3x_1^2 + a}{2y_1} = \frac{3 \cdot 160^2 + 1}{74}.$$

Furthermore,

$$\begin{array}{r} 185 = 1 \cdot 185 + 0 \cdot 74 \\ 74 = 0 \cdot 185 + 1 \cdot 74 \\ \hline 37 = 1 \cdot 185 - 2 \cdot 74 \\ 0 = -2 \cdot 185 + 5 \cdot 74 \end{array}$$

so $\text{mcd}(185, 74) = 37$ and $185 = 5 \cdot 37$.

Here we note that the order of P in $E(\mathbb{Q})$ is ∞ , while the order of P in $E(\mathbb{F}_{37})$ is 8.

2.2.4 Complexity

Here we list some known facts about the complexity of the methods that we have seen (and some that we haven't seen), without bothering about the proofs.

- The algorithm trying division by $k = 2, 3, \dots, \sqrt{n}$ has complexity $O(p)$ (order p), where p is the smallest prime factor of n . But, in the worst case, $O(p) = O(\sqrt{n})$ and $\sqrt{n} = n^{\frac{1}{2}} = \exp(\frac{1}{2} \log n)$ and $\log n \cong$ "number of digits of n " = size of n . So $\exp(\frac{1}{2} \log n)$ is exponential in the input size.

- The elliptic curve method has estimated complexity $O(\exp \sqrt{\log p \log(\log p)})$ where p is the smallest factor of n . This is called "subexponential".

- The continued fraction method has estimated complexity $O(\exp \sqrt{2 \log p \log(\log p)})$.

- Two other famous methods are: the quadratic sieve with complexity $O(\exp \sqrt{\log n \log(\log n)})$, and the number field sieve, with complexity $O(\exp(\log)^{1/3}(\log(\log p))^{2/3})$.

2.3 Primality proving with elliptic curves

In 1986 Goldwasser and Kilian published a method for proving that a given n , of which it has been established that it is prime with high probability, in fact is prime. It uses elliptic curves, and is based on the following theorem.

Theorem 2.3.1 *Let E be an elliptic curve with equation $y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Z}$. Suppose that for $1 \leq i \leq r$ there are distinct prime numbers p_i and points $P_i \in E(\mathbb{Q})$, such that*

- P_i is good for n , $1 \leq i \leq r$,
- $p_i P_i$ is not good for n , $1 \leq i \leq r$,
- $p_1 \cdots p_r > (n^{\frac{1}{4}} + 1)^2$.

Then n is prime.

PROOF. Let p be a prime factor of n . Since the P_i are good for n , they are also good for p . So we get points $\varphi_p(P_i) \in E(\mathbb{F}_p)$. Moreover, $p_i \varphi_p(P_i) = O$ for all i . So $\varphi_p(P_i)$ has order p_i . Hence p_i divides $|E(\mathbb{F}_p)|$ and therefore so does $p_1 \cdots p_r$. So we get

$$(n^{\frac{1}{4}} + 1)^2 < p_1 \cdots p_r \leq |E(\mathbb{F}_p)| < p + 1 + 2\sqrt{p} = (\sqrt{p} + 1)^2$$

implying that $p > \sqrt{n}$. So the prime factors of n are greater than \sqrt{n} . We conclude that n is prime. \square

Example 2.3.2 Let $n = 1873$ and let E be given by $y^2 = x^3 - 3x + 6$; and $P = (1, 2)$. Then $101P$ is not good for n . (Note that this can be established by doing some additions modulo n .) But $101 > (n^{\frac{1}{4}} + 1)^2 = 57.4 \dots$. It follows that n is prime.

The basic problem with this test is to find a suitable elliptic curve E and points P_i on it that are good for n , but for which there exist relatively small primes p_i such that $p_i P_i$ are not good for n . (Note that the p_i need to be small with respect to n , otherwise the proof that the p_i are prime is as difficult as proving this for n .) In 1993 Atkin and Morain published a method for that, using the theory of complex multiplication. In practice this works surprisingly well, and huge numbers have been shown to be prime with this method.

2.4 Exercises

1. The smallest Carmichael number is $n = 561$. Show that it is not prime, using the Miller-Rabin test.
2. Factorise 203 and 899 with the method of Fermat.
3. In this exercise we find the factorisation of $n = 8616460799$ with Fermat's method. So we need to find $t > s$ such that $t^2 - s^2 = n$.
 - (a) We have $n \bmod 3 = 2$. Show that $k^2 \bmod 3$ is 0 or 1 for $k \in \mathbb{Z}$. Prove that if $t^2 \bmod 3 = 1$, then $t^2 - n$ cannot be a square. Conclude that t has to be divisible by 3.
 - (b) We have $n \bmod 8 = 7$. Show that $k^2 \bmod 8$ is 0,1,4. Prove that if $t^2 - n$ is a square, then $t^2 = 0 \bmod 8$. Conclude that t is divisible by 4.
 - (c) Prove that t is divisible by 12.
 - (d) Now try for t all multiples of 12, bigger than $\lceil \sqrt{n} \rceil$. Factorise n in a product of two smaller integers.

- (e) The English economist W. S. Jevons (1835-1882) wrote: "Given any two numbers, we may by a simple and infallible process obtain their product, but it is quite another matter when a large number is given, to determine its factors. Can the reader say what two numbers multiplied together produce the number 8616460799? I think it unlikely that anyone but myself will ever know."

4. Find the first 7 convergents of the continued fraction of $\sqrt{5}$.
5. Factorise 299, 341, 1537 and 1139 with the continued fraction method.
6. Let $a > 1$ be an integer and

$$\theta = \frac{a + \sqrt{a^2 + 4}}{2}.$$

Show that for the continued fraction of θ we have $a = a_0 = a_1 = a_2, \dots$

7. Let

$$x = \frac{1 + \sqrt{5}}{2}.$$

Let a_0, a_1, \dots be the integers corresponding to the continued fraction of x .

- (a) Prove that $a_i = 1$ for $i \geq 0$.
- (b) Let F_n be the n -th Fibonacci number (that is $F_0 = 1, F_1 = 1, F_{n+1} = F_n + F_{n-1}$). Prove that the convergents of the continued fraction of x are $\frac{F_{n+1}}{F_n}$ for $n \geq 0$.

8. Consider the elliptic curve with equation $y^2 = x^3 + x - 1$. Let $P = (1, 1) \in E$. Compute $3P = P + P + P$.
9. In this exercise we factorise $n = 65 = 5 \cdot 13$ with the elliptic curve method. Let E be the elliptic curve with equation $y^2 = x^3 + 2x + 1$ and $P = (1, 2)$, a point.
- (a) Check that the equation defines an elliptic curve modulo every prime that divides n .
- (b) Compute all points of $E(\mathbb{F}_5)$ (here it can be handy to make a table of all squares modulo 5, and all $x^3 + 2x + 1$ for $x \in \mathbb{F}_5$). Show that $k\varphi_5(P) = O$ implies that k is divisible by 7. (Here $\varphi_5(P)$ is the reduction modulo 5 of P .)
- (c) Compute all points of $E(\mathbb{F}_{13})$. Prove that $4\varphi_{13}(P) = O$.
- (d) Compute $4P$ modulo n , and factorise n .
10. Again we consider $n = 65$, but now we take the elliptic curve E with equation $y^2 = x^3 + 3x + 2$ and point $P = (2, 4)$.
- (a) Check that the equation defines an elliptic curve modulo every prime that divides n .
- (b) Compute all points of $E(\mathbb{F}_5)$. Show that $k\varphi_5(P) = O$ if and only if k is divisible by 5.
- (c) Compute $5P$ modulo n and factorise n .
11. Factorise $n = 115$ with the elliptic curve method. (For example with the curve $y^2 = x^3 + 2x - 3$ and point $P = (2, 3)$.)

Chapter 3

Polynomial Factorisation

The problem considered in this chapter is to find the factorisation of a polynomial in $k[x]$, where k is a field.

The algorithms depend heavily on k . Here we treat algorithms for finite fields $k = \mathbb{F}_q$, with $q = p^n$, and $k = \mathbb{Q}$.

3.1 Some generalities on polynomials

Lemma 3.1.1 *For $f, g \in k[x]$ there exist unique $q, r \in k[x]$ with $\deg(r) < \deg(g)$ and $f = qg + r$ (division with remainder).*

Lemma 3.1.2 *Let $I \subset k[x]$ be an ideal. Then I is generated by a single element, that is, there is a $g \in k[x]$ with $I = \{fg \mid f \in k[x]\}$.*

PROOF. Let $g \in I$ be a nonzero element of minimal degree. For $h \in I$ write $h = qg + r$ with $\deg(r) < \deg(g)$. This implies $r \in I$ whence $r = 0$ and $h = qg$. \square

Lemma 3.1.3 *Let $f_1, f_2 \in k[x]$. Then there exists a unique monic $g \in k[x]$ with*

- 1) g divides f_1, f_2 ;
- 2) if h divides f_1, f_2 then h divides g .

PROOF. Let $I \subset k[x]$ be the ideal generated by f_1, f_2 , that is $I = \{h_1f_1 + h_2f_2 \mid h_i \in k[x]\}$. Then by the previous lemma I is generated by a monic g (of minimal degree). Now $f_1, f_2 \in I$ hence g divides f_1 and f_2 .

If h divides f_1, f_2 then $g = h_1f_1 + h_2f_2$ (since $g \in I = \langle f_1, f_2 \rangle$); we conclude that h divides g .

In order to show uniqueness, let $g' \in k[x]$ with the same properties as g . Then by 1) and 2), g' divides g . Analogously g divides g' . Since g and g' are monic we get $g = g'$. \square

Definition 3.1.4 *The polynomial g from the previous lemma is called the greatest common divisor of f_1 and f_2 .*

From the proof of Lemma 3.1.3 it follows that there exist h_1, h_2 such that $h_1 f_1 + h_2 f_2 = g$. In the same way as for the integers we have a euclidean algorithm for computing the greatest common divisor, and the h_1, h_2 . It is based on the following lemma.

Lemma 3.1.5 Write $f_1 = qf_2 + r$ with $\deg(r) < \deg(f_2)$. Then $\gcd(f_1, f_2) = \gcd(f_2, r)$.

PROOF. Set

$$D_1 = \{h \in k[x] \mid h \text{ divides } f_1, f_2\} \quad \text{and} \quad D_2 = \{h \in k[x] \mid h \text{ divides } f_2, r\}.$$

Then $D_1 = D_2$. So the elements of maximal degree of D_1 and D_2 coincide. \square

To compute $\gcd(f_1, f_2)$ we replace (f_1, f_2) by (f_2, r) and so on. At a certain point we find the pair $(g, 0)$ and $\gcd(g, 0) = g$.

Example 3.1.6 Let $f_1 = x^7 + 1, f_2 = x^4 + x^2 + x \in \mathbb{F}_2[x]$. Then

$$f_1 = (x^3 + x + 1)f_2 + x^3 + x + 1.$$

Set $f_3 = x^3 + x + 1$, then

$$f_2 = xf_3 + 0$$

hence $\gcd(f_1, f_2) = \gcd(f_2, f_3) = \gcd(f_3, 0) = f_3$.

N.B. A polynomial $h \in k[x]$ is called *irreducible* if $h \notin k$ and for every factorisation $h = ab$ with $a, b \in k[x]$ we have that a or b lies in k .

Lemma 3.1.7 Let $a \in k[x]$ be irreducible and suppose that a divides bc where $b, c \in k[x]$. Then a divides b or a divides c .

PROOF. If a does not divide b then $\gcd(a, b) = 1$ so there exist h_1, h_2 with $h_1 a + h_2 b = 1$ whence $c = h_1 ac + h_2 bc$. Therefore a divides c . \square

Theorem 3.1.8 Let $f \in k[x]$, then there exist $c \in k$, irreducible $f_1, \dots, f_r \in k[x]$ and $e_1, \dots, e_r \in \mathbb{Z}_{>0}$ with $f = cf_1^{e_1} \cdots f_r^{e_r}$. Moreover, this factorisation is “essentially unique”; meaning that if we have another one $f = dg_1^{d_1} \cdots g_s^{d_s}$, then $c = d, r = s$, and after a permutation of the indices, $f_i = g_i, e_i = d_i$.

PROOF. Note that c is the coefficient of the highest degree monomial in f , hence it is uniquely determined. So without loss of generality we may assume that f is monic. If f is irreducible then there is nothing to prove. So suppose that $f = ab$ where $a, b \in k[x]$ are monic with $\deg(a), \deg(b) < \deg(f)$. By induction on the degree a and b are products of irreducible polynomials, hence so is f .

To show uniqueness suppose that $f = f_1^{e_1} \cdots f_r^{e_r} = g_1^{d_1} \cdots g_s^{d_s}$, where also the g_i are irreducible and monic.

By Lemma 3.1.7 f_1 divides a g_i . But g_i is irreducible and monic, hence $f_1 = g_i$. Cancelling both from the products, and by induction on the degree we conclude that the factorisation is essentially unique. \square

Corollary 3.1.9 *Let $p, q, r \in k[x]$ with $\gcd(q, r) = 1$. Then $\gcd(p, qr) = \gcd(p, q) \gcd(p, r)$.*

PROOF. Let $f, g \in k[x]$. Then by the previous theorem we can write $f = c_1 f_1^{e_1} \cdots f_m^{e_m}$, $g = c_2 f_1^{d_1} \cdots f_m^{d_m}$, where the f_i are monic and irreducible and $e_i, d_i \in \mathbb{Z}_{\geq 0}$. From this it follows that $\gcd(f, g) = f_1^{a_1} \cdots f_r^{a_r}$, where $a_i = \min\{d_i, e_i\}$.

So since $\gcd(q, r) = 1$, the factorisation of qr is obtained by appending the factorisation of r to that of q . This implies the statement of the corollary. \square

Problem. Given a monic $f \in k[x]$ find monic irreducible f_1, \dots, f_r such that $f = f_1^{e_1} \cdots f_r^{e_r}$.

3.2 Berlekamp's algorithm

In this section we describe an algorithm due to Elwyn Berlekamp (1940-)



for factorising a polynomial over a finite field. Berlekamp is known, among other things, for his work in coding theory and game theory.

Definition 3.2.1 *Let $f \in k[x]$ be monic and $f = f_1^{e_1} \cdots f_r^{e_r}$ with f_i irreducible and monic. Then the $f_i^{e_i}$ are called the primary factors of f .*

Lemma 3.2.2 *Let $f \in \mathbb{F}_q[x]$ be monic and $v \in \mathbb{F}_q[x]$ with $v^q = v \pmod{f}$. Then*

$$f = \prod_{a \in \mathbb{F}_q} \gcd(f, v - a).$$

PROOF. In $\mathbb{F}_q[Y]$ we have

$$Y^q - Y = \prod_{a \in \mathbb{F}_q} (Y - a)$$

(because $a^q = a$ for all $a \in \mathbb{F}_q$). Substituting v for Y we obtain

$$v^q - v = \prod_{a \in \mathbb{F}_q} (v - a).$$

But $v^q - v$ is divisible by f , so $f = \gcd(f, v^q - v)$.

Observe that $\gcd(v - a, v - b) = 1$ if $a \neq b$, $a, b \in \mathbb{F}_q$. Indeed, if g divides $v - a$ and $v - b$ then g divides $b - a \neq 0$.

Using Corollary 3.1.9 we see that

$$\begin{aligned} f &= \gcd(f, v^q - v) \\ &= \gcd(f, \prod_{a \in \mathbb{F}_q} (v - a)) \\ &= \prod_{a \in \mathbb{F}_q} \gcd(f, v - a). \end{aligned}$$

□

Notation. With $\mathbb{F}_q[x]/(f)$ we denote the quotient ring $\mathbb{F}_q[x]$ modulo the ideal generated by f .

If $f = x^n + \dots + a_0$ then every element of $\mathbb{F}_q[x]/(f)$ has a unique representative of the form $b_0 + b_1x + \dots + b_{n-1}x^{n-1}$.

Lemma 3.2.3 *Let $f \in \mathbb{F}_q[x]$ be monic and*

$$V = \{v \in \mathbb{F}_q[x]/(f) \mid v^q - v \bmod f\}.$$

Write $f = f_1^{e_1} \dots f_r^{e_r}$ where the f_i are irreducible and distinct. Then V is a vector space over \mathbb{F}_q of dimension r .

PROOF. Let $v, w \in V$, then

$$(v + w)^q = v^q + w^q = v + w \bmod f$$

and hence $v + w \in V$.

Furthermore, let $\alpha \in \mathbb{F}_q$, then

$$(\alpha v)^q = \alpha^q v^q = \alpha v^q = \alpha v \bmod f$$

so $\alpha v \in V$.

It follows that V is a vector space over \mathbb{F}_q .

Let R_1, \dots, R_s be rings. We recall the construction of the direct sum: $R = \bigoplus_{i=1}^s R_i$ is the set consisting of (r_1, \dots, r_s) with $r_i \in R_i$; the ring operations are defined as follows

$$(r_1, \dots, r_s) + (r'_1, \dots, r'_s) = (r_1 + r'_1, \dots, r_s + r'_s)$$

and

$$(r_1, \dots, r_s)(r'_1, \dots, r'_s) = (r_1 r'_1, \dots, r_s r'_s).$$

Then R is again a ring.

By the chinese remainder theorem for polynomial rings there exists a ring isomorphism $\varphi: \mathbb{F}_q[x]/(f) \rightarrow \bigoplus_{i=1}^r \mathbb{F}_q[x]/(f_i^{e_i})$. Here $\varphi(h \bmod f) = (h \bmod f_1^{e_1}, \dots, h \bmod f_r^{e_r})$.

For $v \in V$ we write $\varphi(v) = (v_1, \dots, v_r)$. Then

$$v \in V \iff v^q = v \bmod f \iff v_i^q = v_i \bmod f_i^{e_i}.$$

By Lemma 3.2.2 we have

$$f_i^{e_i} = \prod_{a \in \mathbb{F}_q} \gcd(f_i^{e_i}, v_i - a). \quad (3.1)$$

But $v_i - a$ and $v_i - b$ are coprime when $a, b \in \mathbb{F}_q$ and $a \neq b$. Hence in (3.1) we only have factors $v_i - a$ for a unique $a \in \mathbb{F}_q$. In other words $f_i^{e_i} = \gcd(f_i^{e_i}, v_i - a)$, implying that $f_i^{e_i}$ divides $v_i - a$. It follows that $v_i = a \pmod{f_i^{e_i}}$ and therefore $\varphi(V) \subset \mathbb{F}_q \oplus \dots \oplus \mathbb{F}_q$ where $\mathbb{F}_q \subset \mathbb{F}_q[x]/(f_i^{e_i})$ are the constant polynomials (more precisely: the cosets that have a representative in \mathbb{F}_q).

But $a^q = a$ for all $a \in \mathbb{F}_q$, whence $\mathbb{F}_q \oplus \dots \oplus \mathbb{F}_q \subset \varphi(V)$.

It follows that $\varphi(V) = \mathbb{F}_q \oplus \dots \oplus \mathbb{F}_q$ (r summands), implying that V is a vector space of dimension r . \square

Now we give Berlekamp's algorithm for finding the primary factors of a monic $f \in \mathbb{F}_q[x]$.

Algorithm 3.2.4 *Given: $f \in \mathbb{F}_q[x]$, monic.*

Compute: the primary factors of f .

1. Let $\{v_1 = 1, v_2, \dots, v_r\}$ be a basis of $V = \{v \in \mathbb{F}_q[x]/(f) \mid v^q = v \pmod{f}\}$.
2. Set $P_1 = \{f\}$ and for $j = 2, \dots, r$ let
 - 2a) P_j be the set of consisting of the nontrivial elements of the sets $\{\gcd(h, v_j - a) \mid a \in \mathbb{F}_q\}$, where $h \in P_{j-1}$.
3. Return P_r .

Lemma 3.2.5 *The algorithm of Berlekamp finds the primary factors of f .*

PROOF. Write $P = P_r$. We have to show that

- f is the product of the elements in P ;
- the elements of P are coprime;
- the elements of P are powers of irreducible polynomials.

First we claim that the product of the elements of P_j is f . This is certainly true for $j = 1$. Suppose the claim holds for P_{j-1} . Let $h \in P_{j-1}$, then h divides f , whence $v_j^q = v_j \pmod{h}$. So by Lemma 3.2.2,

$$h = \prod_{a \in \mathbb{F}_q} \gcd(h, v_j - a).$$

Also the elements of P_j are coprime. Indeed, the elements of P_j are of the form $\gcd(h, v_j - a)$, for $a \in \mathbb{F}_q$, $h \in P_{j-1}$. Now for different $a \in \mathbb{F}_q$ such factors are coprime. But also for different h (and the same a), because by induction the elements of P_{j-1} are coprime.

Now let $h \in P$. Suppose that h is not a power of an irreducible. Then from what is said above it follows that h is divisible by $f_i^{e_i}$ and by $f_j^{e_j}$, $i \neq j$. To ease notation we assume that $i = 1$ and $j = 2$.

We fix a $j \geq 1$. Then for $j \leq l \leq r$ and $g \in P_l$ there exists $a_j \in \mathbb{F}_q$ with $v_j = a_j \pmod{g}$. Again we use induction to show this. Note that the elements of P_j are factors of $v_j - a$, for various $a \in \mathbb{F}_q$. Hence for $l = j$ the statement holds. Now let $g \in P_l$, $l > j$. Then g is a factor of a $h \in P_{l-1}$. By induction $v_j = a_j \pmod{h}$, hence also $v_j = a_j \pmod{g}$.

We apply this to $h \in P = P_r$ to conclude that there is $a_j \in \mathbb{F}_q$ with $v_j = a_j \pmod{h}$ for $1 \leq j \leq r$.

Now let $v \in V$ and write $v = \sum_{j=1}^r \beta_j v_j$, $\beta_j \in \mathbb{F}_q$. Then

$$v = \sum_{j=1}^r \beta_j a_j \pmod{h}.$$

Set $a_v = \sum_{j=1}^r \beta_j a_j$; then $v = a_v \pmod{h}$, and $a_v \in \mathbb{F}_q$.

Since $f_1^{e_1}, f_2^{e_2}$ divide h we get $v = a_v \pmod{f_1^{e_1}}$ and $v = a_v \pmod{f_2^{e_2}}$. But $\varphi: V \rightarrow \mathbb{F}_q \oplus \dots \oplus \mathbb{F}_q$ with $\varphi(v) = (v \pmod{f_1^{e_1}}, \dots, v \pmod{f_r^{e_r}})$ is an isomorphism.

But $\varphi(v) = (a_v, a_v, *, \dots, *)$, in other words, the first two coordinates are always equal. This implies that φ is not surjective, and hence we have obtained a contradiction. \square

Remaining problem. Given $f \in \mathbb{F}_q[x]$, $f = g^e$ with g irreducible, find g and e .

For this we compute $f' = \frac{d}{dx} f = eg'g^{e-1}$. There are two cases:

1) $f' = 0$. Then we claim that f is a polynomial in x^p where $q = p^m$, p a prime. Indeed, the exponents of x that occur in f are divisible by p . Hence $f = h(x^p) = h_1(x)^p$. We compute h_1 , which again is a power of g , and continue with h_1 .

2) $f' \neq 0$. Then $g = \frac{f}{\gcd(f, f')}$ since $\gcd(f, f') = g^{e-1}$.

Example 3.2.6 Let $f = x^4 + x^3 + x + 1 \in \mathbb{F}_2[x]$. Write $v \in \mathbb{F}_2[x]/(f)$ as $v = a_0 + a_1x + a_2x^2 + a_3x^3$, $a_i \in \mathbb{F}_2$ (more precisely: the last expression is the unique representative modulo f of degree $< \deg(f)$ of a class in $\mathbb{F}_2[x]/(f)$).

Since $a_i^2 = a_i$ we get $v^2 = a_0 + a_1x^2 + a_2x^4 + a_3x^6$. Modulo f we have:

$$\begin{aligned} x^4 &= x^3 + x + 1 \\ x^5 &= x^4 + x^2 + x = x^3 + x + 1 + x^2 + x \\ &= x^3 + x^2 + 1 \\ x^6 &= x^4 + x^3 + x = x^3 + x + 1 + x^3 + x \\ &= 1 \end{aligned}$$

hence

$$\begin{aligned} v^2 &= a_0 + a_1x^2 + a_2(x^3 + x + 1) + a_3 \pmod{f} \\ &= a_0 + a_2 + a_3 + a_2x + a_1x^2 + a_2x^3 \pmod{f} \end{aligned}$$

and therefore

$$\begin{cases} a_0 + a_2 + a_3 = a_0 \\ a_2 = a_1 \\ a_1 = a_2 \\ a_2 = a_3 \end{cases} \iff a_1 = a_2 = a_3$$

so a basis of V is $\{v_1 = 1, x + x^2 + x^3\}$. So, in particular, $r = 2$.

In step 2a) we replace f with

$$\begin{aligned} h_1 &= \gcd(f, x + x^2 + x^3) = x^2 + x + 1 \\ h_2 &= \gcd(f, 1 + x + x^2 + x^3) = x^2 + 1. \end{aligned}$$

So $P = \{h_1, h_2\}$. We have $h_1' = 1$ so that $\gcd(h_1, h_1') = 1$, thus h_1 is irreducible. Furthermore, $h_2' = 0$ implies $h_2 = (x + 1)^2$ and $(x + 1)' = 1$ so it is irreducible.

We conclude that $f = (x + 1)^2(x^2 + x + 1)$ is the factorisation of f .

3.3 The algorithm of Cantor-Zassenhaus

The algorithm of Berlekamp has the disadvantage to have to compute $\gcd(f, v-a)$ for $a \in \mathbb{F}_q$. If q gets large this is very laborious. Here we sketch a different approach, due to Cantor and Zassenhaus, based on the possibility of making random choices. For this we assume that q is odd.

Lemma 3.3.1 *Let $\gamma \in \mathbb{F}_q^*$, then $\gamma^{\frac{q-1}{2}} = \pm 1$. Moreover, $\gamma^{\frac{q-1}{2}} = 1$ if and only if γ is a square in \mathbb{F}_q and $\gamma^{\frac{q-1}{2}} = -1$ if and only if γ is not a square.*

PROOF. Let $\alpha \in \mathbb{F}_q^*$ be a primitive element: $\alpha^{q-1} = 1$, $\alpha^k \neq 1$ for $k < q-1$. Then $\gamma = \alpha^i$, and

$$\gamma^{\frac{q-1}{2}} = \left(\alpha^{\frac{q-1}{2}}\right)^i.$$

Remark 3.3.2 We have that $\left(\alpha^{\frac{q-1}{2}}\right)^2 = 1$ implies $\alpha^{\frac{q-1}{2}} = \pm 1$ (there are only two roots of 1 in \mathbb{F}_q) but $\alpha^{\frac{q-1}{2}} \neq 1$ so $\alpha^{\frac{q-1}{2}} = -1$.

Hence if γ is a square then $\gamma = \alpha^{2k}$, and hence $\gamma^{\frac{q-1}{2}} = (\alpha^{q-1})^k = 1$. If γ is not a square then $\gamma = \alpha^{2k+1}$ so that $\gamma^{\frac{q-1}{2}} = \alpha^{\frac{q-1}{2}}(\alpha^{q-1})^k = -1$. \square

Consider the space

$$V = \{v \in \mathbb{F}_q[x]/(f) \mid v^q = \text{mod } f\}.$$

Let $v \in V$ and view V inside $\mathbb{F}_q[x]$. Then

$$v^q - v = v \left(v^{\frac{q-1}{2}} - 1\right) \left(v^{\frac{q-1}{2}} + 1\right).$$

Note that these factors are coprime.

We have

$$f = \gcd(f, v^q - v) = \gcd(f, v) \gcd(f, v^{\frac{q-1}{2}} - 1) \gcd(f, v^{\frac{q-1}{2}} + 1).$$

This factorisation of f is trivial (that is $f = f \cdot 1 \cdot 1$) if and only if f divides one of v , $v^{\frac{q-1}{2}} - 1$, $v^{\frac{q-1}{2}} + 1$.

We may assume that $\deg(v) < \deg(f)$ so f does not divide v .

Let $\varphi: \mathbb{F}_q[x]/(f) \rightarrow \bigoplus_{i=1}^r \mathbb{F}_q[x]/(f_i^{e_i})$ be the isomorphism, where the $f_i^{e_i}$ are the primary factors of f . Then $\varphi(V) = \bigoplus_{i=1}^r \mathbb{F}_q$ implies that $\varphi(v) = (a_1, \dots, a_r)$, $a_i \in \mathbb{F}_q$.

Now f divides $v^{\frac{q-1}{2}} - 1$ if and only if $v^{\frac{q-1}{2}} = 1 \pmod f$ and hence if and only if $\varphi(v)^{\frac{q-1}{2}} = (1, \dots, 1)$; or $a_i^{\frac{q-1}{2}} = 1$ for all i . By Lemma 3.3.1 we have that $a_i \in \mathbb{F}_q^*$ is a square for all i .

Analogously f divides $v^{\frac{q-1}{2}} + 1$ if and only if $a_i \in \mathbb{F}_q^*$ is not a square for all i .

Set

$$\begin{aligned} C &= \{(a_1, \dots, a_r) \in \mathbb{F}_q^{*r} \mid a_i \text{ is a square } \forall i\} \\ &\cup \{(a_1, \dots, a_r) \in \mathbb{F}_q^{*r} \mid a_i \text{ is not a square } \forall i\}. \end{aligned}$$

We have

$$|C| = \left(\frac{q-1}{2}\right)^r + \left(\frac{q-1}{2}\right)^r = 2 \left(\frac{q-1}{2}\right)^r.$$

If we take $v \in V$ by random choice (where we use a uniform distribution) then the probability that we do not find a factorisation is

$$\mathcal{P}(\varphi(v) \in C) = \frac{|C|}{|V|} = \frac{2 \left(\frac{q-1}{2}\right)^r}{q^r} = 2 \left(\frac{1 - \frac{1}{q}}{2}\right)^r < \frac{1}{2}.$$

So if k times we randomly choose a v the probability of not finding a factorisation is less than $\left(\frac{1}{2}\right)^k$.

3.4 Factorization of polynomials over \mathbb{Q}

First we show that the factorization of polynomials over \mathbb{Q} is the same as the factorization of polynomials over \mathbb{Z} .

Definition 3.4.1 Let $f \in \mathbb{Z}[x]$, $f = a_0 + a_1x + \dots + a_nx^n$, $a_i \in \mathbb{Z}$. Then

$$c(f) = \gcd(a_0, a_1, \dots, a_n)$$

is called the content of f .

Lemma 3.4.2 Let $f, g \in \mathbb{Z}[x]$ with $c(f) = c(g) = 1$, then $c(fg) = 1$.

PROOF. Suppose that $c(fg) > 1$. Then there exists a prime p dividing all coefficients of fg .

For $h \in \mathbb{Z}[x]$ we write \bar{h} for the polynomial $h \bmod p$ which lies in $\mathbb{F}_p[x]$. Then $0 = \overline{fg} = \bar{f}\bar{g} \neq 0$ because $\bar{f} \neq 0 \neq \bar{g}$ as they have content 1. \square

Lemma 3.4.3 (Gauss) Let $f, g \in \mathbb{Z}[x]$. Then $c(fg) = c(f)c(g)$.

PROOF. Write $a = c(f)$, $b = c(g)$. Then $\frac{f}{a}, \frac{g}{b} \in \mathbb{Z}[x]$. But

$$fg = ab \frac{f}{a} \frac{g}{b} \Rightarrow c(fg) = c\left(ab \frac{f}{a} \frac{g}{b}\right) = ab \cdot c\left(\frac{f}{a} \frac{g}{b}\right).$$

Now $c\left(\frac{f}{a}\right) = c\left(\frac{g}{b}\right) = 1$ hence, by Lemma 3.4.2, $c\left(\frac{f}{a} \frac{g}{b}\right) = 1$ and therefore $c(fg) = ab$. \square

Theorem 3.4.4 Let $f \in \mathbb{Z}[x]$ and suppose that there exist $g, h \in \mathbb{Q}[x]$ with $f = gh$. Then there exist $a, b \in \mathbb{Q}$ such that $ag, bh \in \mathbb{Z}[x]$ and $f = (ag)(bh)$.

In other words, every factorisation of f in $\mathbb{Q}[x]$ comes from a factorisation in $\mathbb{Z}[x]$.

PROOF. We can assume that $c(f) = 1$ (otherwise we divide by $c(f)$). Let $a, b \in \mathbb{Q}$ be such that $ag, bh \in \mathbb{Z}[x]$ and $c(ag) = c(bh) = 1$. We have

$$abf = (ag)(bh) \in \mathbb{Z}[x] \quad \text{e} \quad c(abf) = c(ag)c(bh) = 1.$$

Write $ab = \frac{s}{t}$ with $t \geq 1$; then all coefficients of f are divisible by t because $\frac{s}{t}f \in \mathbb{Z}[x]$. But $c(f) = 1$ so $t = 1$.

Now $c(abf) = c(sf) = 1$ hence $s = \pm 1$. Therefore $(ag)(bh) = \pm f$. If this is $-f$ then we replace a with $-a$ and obtain $f = (ag)(bh)$. \square

So in order to factorise $f \in \mathbb{Q}[x]$ we can make some reductions

- assume $f \in \mathbb{Z}[x]$ (otherwise we multiply by an $s \in \mathbb{Z}$);
- the problem does not change if we look for factors in $\mathbb{Z}[x]$ (by Theorem 3.4.4);
- assume that f is square-free; otherwise we write $f = g^2h$ and then $f' = 2gg'h + g^2h'$ so g divides $\gcd(f, f')$; so we can factorise $\gcd(f, f')$ and $\frac{f}{\gcd(f, f')}$.

Idea. We factorise $f \bmod p$, p a prime, and we “lift” the factors to elements of $\mathbb{Z}[x]$. This process is called *Hensel lifting*.

3.4.1 Hensel lifting

Kurt Hensel (1861 - 1941)



invented p -adic numbers. The next theorem shows that a factorisation of a polynomial modulo p leads to a p -adic factorisation.

Theorem 3.4.5 (Hensel) *Let $f \in \mathbb{Z}[x]$ and p a prime, not dividing the leading coefficient of f . Let $g, h \in \mathbb{Z}[x]$ be such that*

- $\deg(g) + \deg(h) = \deg(f)$;
- $\gcd(g, h) = 1 \bmod p$;
- $f = gh \bmod p^k$ for a certain integer $k > 0$.

Then there exist $u, v \in \mathbb{Z}[x]$ with $\deg u < \deg g$, $\deg v \leq \deg h$; and after setting $\tilde{g} = g + p^k u$, $\tilde{h} = h + p^k v$, we have $\deg(\tilde{h}) = \deg h$, and

$$f = \tilde{g}\tilde{h} \bmod p^{k+1}.$$

Moreover, the u and v with these properties are uniquely determined modulo p .

PROOF. Set $\tilde{g} = g + p^k u$, $\tilde{h} = h + p^k v$ for certain $u, v \in \mathbb{Z}[x]$. Then

$$f - \tilde{g}\tilde{h} = f - gh - p^k vg - p^k uh - p^{2k} uv$$

and this has to be 0 modulo p^{k+1} .

Observe that $f - gh$ is divisible by p^k .

So we want u, v such that

$$\frac{f - gh}{p^k} - vg - uh = 0 \bmod p.$$

Set $e = \frac{f - gh}{p^k}$.

We know that $\gcd(g, h) = 1 \pmod p$. So there exist $a, b \in \mathbb{Z}[x]$ with $ag + bh = 1 \pmod p$. Hence $ea + ebh = e \pmod p$.

So we could take $v = ea$ and $u = eb$ but this can increase the degree. The trick is to first fix eb . Write $eb = qg + r$ with $\deg(r) < \deg(g)$. Then

$$eag + (qg + r)h = e \pmod p \quad \Leftrightarrow \quad (ea + qh)g + rh = e \pmod p.$$

Now we try $v = ea + qh \pmod p$ and $u = r \pmod p$. Of course, we are only interested in u and v modulo p .

With this choice we have $f = \tilde{g}\tilde{h} \pmod{p^{k+1}}$ and $\deg(u) < \deg(g)$. We need to check the degree of v .

Since $e = \frac{f-gh}{p^k}$ we get $\deg(e) \leq \deg(f)$. Moreover, $uh + vg = e \pmod p$ and $\deg(uh) < \deg(f)$ because $\deg(u) < \deg(g)$. Hence $\deg(vg \pmod p) \leq \deg(f)$, and therefore

$$\deg(v \pmod p) \leq \deg(h).$$

It also follows that $\deg(\tilde{h}) \leq \deg(h)$. However, if this inequality is strict then $f = \tilde{g}\tilde{h} \pmod{p^{k+1}}$ implies that p divides the leading coefficient of f , contrary to the assumption on p . Hence $\deg(\tilde{h}) = \deg(h)$.

In order to show uniqueness, let $u_0, v_0 \in \mathbb{Z}[x]$ have the same properties. Then $v_0g + u_0h = vg + uh \pmod p$, or

$$(u_0 - u)h = (v_0 - v)g \pmod p.$$

Since g, h are coprime modulo p this implies that g divides $u_0 - u$. But that is impossible because of the degrees; hence $u_0 - u = 0$. From that we also get $v - v_0 = 0$. \square

Remark 3.4.6 Note that the \tilde{g}, \tilde{h} in the conclusion of the preceding theorem satisfy the same properties as g and h , except that k has changed to $k + 1$. So we can continue and obtain a factorisation $\pmod{p^{k+2}, p^{k+3}, \dots}$. In the limit this then yields a p -adic factorisation.

The proof of the theorem also gives a method for constructing \tilde{g}, \tilde{h} . One computes

- $e = \frac{f-gh}{p^k}$;
- a, b with $ag + bh = 1 \pmod p$ (extended euclidean algorithm);
- q, r with $eb = qg + r$ and $\deg(r) < \deg(g)$ (division with remainder);
- $u = r \pmod p$ and $v = ea + qh \pmod p$.

Then $\tilde{g} = g + p^k u$ e $\tilde{h} = h + p^k v$.

Example 3.4.7 Let $f = x^4 + 2x^3 - 3x^2 - 4x - 1$, $g = x^2 + 1$ and $h = x^2 + 2x + 2$. We have

$$gh = x^4 + 2x^3 + 3x^2 + 2x + 2 = f \pmod 3$$

hence

$$e = \frac{f-gh}{3} = -2x^2 - 2x - 1 = x^2 + x + 2 \pmod 3.$$

We do all computations modulo p because we are interested in u and v only modulo p . In particular e, a, b, q, v can be computed modulo p . We compute (everything modulo 3):

$$\begin{array}{rcl} h & = & x^2 + 2x + 2 = 1 \cdot h + 0 \cdot g \\ g & = & x^2 + 1 = 0 \cdot h + 1 \cdot g \\ \hline & & 2x + 1 = h - g \\ & & 2x^2 + x = xh - xg \\ & & x + 1 = xh + (1 - x)g \\ & & 2 = (x + 1)h - xg \\ & & 1 = (2x + 2)h + xg. \end{array}$$

Hence $a = x, b = 2x + 2$ modulo 3. Therefore, modulo 3,

$$eb = (x^2 + x + 2)(2x + 2) = 2x^3 + x^2 + 1 = (2x + 1)g + x = qg + r.$$

So $u = r = x$ and

$$\begin{aligned} v &= ea + qh = (x^2 + x + 2)x + (2x + 1)(x^2 + 2x + 2) = 3x^3 + 6x^2 + 8x + 2 \\ &= 2x + 2 \pmod{3}. \end{aligned}$$

We see that $\deg(v) < \deg(h) \pmod{p}$ but it is not guaranteed that v has degree $\leq \deg(h)$ without the modulo p operation.

Now

$$\begin{aligned} g_1 &= g + 3u = x^2 + 3x + 1 \\ h_1 &= h + 3(2x + 2) = x^2 + 8x + 8. \end{aligned}$$

Indeed we have

$$f - g_1h_1 = -9x^3 - 36x^2 - 36x - 9 = 0 \pmod{9}.$$

For $k = 2$ we get

$$e = \frac{f - g_1h_1}{9} = -x^3 - 4x^2 - 4x - 1 = 2x^3 + 2x^2 + 2x + 2 \pmod{3}.$$

Since $g_1 = g \pmod{p}$ and $h_1 = h \pmod{p}$ the same a and b of the previous step can also be used here. So

$$eb = x^4 + 2x^3 + 2x^2 + 2x + 1 = (x^2 + 2x + 1)g_1 + 0 = qg + r.$$

We get $u = r = 0$ and

$$\begin{aligned} v &= ea + qh_1 = (2x^3 + 2x^2 + 2x + 2)x + (x^2 + 2x + 1)(x^2 + 8x + 8) \\ &= 2x + 2 \pmod{3}. \end{aligned}$$

Now

$$\begin{aligned} g_2 &= g_1 + 9u = g_1 = x^2 + 3x + 1 \\ h_2 &= h_1 + 9(2x + 2) = x^2 + 26x + 26. \end{aligned}$$

And $f = g_2h_2 \pmod{27}$. Continuing we find

$$\begin{aligned} g_3 &= x^2 + 3x + 1 \\ h_3 &= x^2 + 80x + 80. \end{aligned}$$

And $f = g_3h_3 \pmod{81}$, and so on.

So, starting with a factorisation modulo p , we find factorisations modulo p^k , for any k we want. Now we describe how we can get the factorisation in $\mathbb{Z}[x]$ from the factorisation modulo p^k , when k is big enough.

Definition 3.4.8 Let $f = \sum_{i=0}^n f_i x^i \in \mathbb{C}[x]$. The norm of f is

$$\|f\| = \sqrt{\sum_{i=0}^n |f_i|^2}.$$

Lemma 3.4.9 Let $h \in \mathbb{C}[x]$, $a \in \mathbb{C}$. Then

$$\|(x-a)h\| = |a| \|(x-\bar{a}^{-1})h\|.$$

PROOF. Write $h = \sum_{i=0}^m h_i x^i$. We have

$$\begin{aligned} (x-a)h &= h_m x^{m+1} + (h_{m-1} - ah_m)x^m + \dots + (h_0 - ah_1)x - ah_0 \\ &= \sum_{i=0}^{m+1} (h_{i-1} - ah_i)x^i \end{aligned}$$

where we set $h_{-1} = h_{m+1} = 0$.

In general we have

$$\begin{aligned} |u-v|^2 &= (u-v)(\bar{u}-\bar{v}) = u\bar{u} - u\bar{v} - \bar{u}v + v\bar{v} \\ &= |u|^2 - u\bar{v} - \bar{u}v + |v|^2. \end{aligned}$$

Hence

$$\begin{aligned} \|(x-a)h\|^2 &= \sum_{i=0}^{m+1} |h_{i-1} - ah_i|^2 \\ &= \sum (|h_{i-1}|^2 - \bar{a}h_{i-1}\bar{h}_i - a\bar{h}_{i-1}h_i + |ah_i|^2) \\ &= \sum (|h_i|^2 - \bar{a}h_{i-1}\bar{h}_i - a\bar{h}_{i-1}h_i + |\bar{a}h_{i-1}|^2) = \sum |\bar{a}h_{i-1} - h_i|^2. \end{aligned}$$

Because $\sum |h_{i-1}|^2 = \sum |h_i|^2$ and

$$|ah_i|^2 = |a|^2|h_i|^2 = |\bar{a}h_i|^2.$$

Therefore

$$\begin{aligned} \|(x-a)h\|^2 &= \|(\bar{a}x-1)h\|^2 \\ &= |\bar{a}| \|(x-\bar{a}^{-1})h\|^2 \\ &= |a| \|(x-\bar{a}^{-1})h\|^2. \end{aligned}$$

□

Theorem 3.4.10 (Landau-Mignotte) Let $f \in \mathbb{Z}[x]$ and $g \in \mathbb{Z}[x]$ be a factor of f , $\deg(g) = m$. Write $g = \sum_{i=0}^m g_i x^i$. Then

$$|g_i| \leq \binom{m}{i} \|f\|.$$

PROOF. Let b_1, \dots, b_s and a_1, \dots, a_t in \mathbb{C} be the zeros of f “with multiplicity” (so a zero of multiplicity k appears k times), where $|b_i| > 1$ and $|a_i| \leq 1$.

Write $f = \sum_{i=0}^n f_i x^i$. Then

$$f = f_n \prod_{i=1}^s (x - b_i) \prod_{j=1}^t (x - a_j)$$

so by Lemma 3.4.9

$$\|f\| = \|f_n \prod_{i=1}^s (x - b_i) \prod_{j=1}^t (x - a_j)\| = |a_1 \cdots a_t| \|f_n \prod_{j=1}^t (x - \bar{a}_j^{-1}) \prod_{i=1}^s (x - b_i)\|.$$

Write $h = \sum_i h_i x^i$, then $\|h\| \geq |h_0|$ since $\|h\|^2 = \sum |h_i|^2$. Hence

$$\|f\| \geq |a_1 \cdots a_t| \|f_n\| |\bar{a}_1^{-1} \cdots \bar{a}_t^{-1}| |b_1 \cdots b_s|.$$

For $z \in \mathbb{C}$, $|z\bar{z}^{-1}| = 1$, so we get

$$\|f\| \geq |f_n b_1 \cdots b_s|.$$

Now let $\gamma_1, \dots, \gamma_m$ be the zeros of g . Then $g = g_m(x - \gamma_1) \cdots (x - \gamma_m)$, whence

$$g_i = (-1)^i g_m \sigma_{m-i}(\gamma_1, \dots, \gamma_m)$$

where $\sigma_i(x_1, \dots, x_m)$ is the i -th elementary symmetric polynomial in x_1, \dots, x_m ; that is

$$\sigma_i(x_1, \dots, x_m) = \sum_{1 \leq j_1 < j_2 < \dots < j_i \leq m} x_{j_1} \cdots x_{j_i}.$$

We have that $\sigma_i(x_1, \dots, x_m)$ is a sum of $\binom{m}{i}$ monomials of degree i .

We assume that the γ_i are ordered in such a way that $|\gamma_1| \geq |\gamma_2| \geq \dots \geq |\gamma_m|$. Then

$$\begin{aligned} |\sigma_i(\gamma_1, \dots, \gamma_m)| &\leq \binom{m}{i} |\gamma_1 \cdots \gamma_i| \\ &\leq \binom{m}{i} |b_1 \cdots b_s| \\ &\leq \binom{m}{i} \frac{\|f\|}{|f_n|} \end{aligned}$$

whence

$$|g_i| \leq |g_m| \binom{m}{m-i} \frac{\|f\|}{|f_n|} = |g_m| \binom{m}{i} \frac{\|f\|}{|f_n|}.$$

We have $|g_m| \leq |f_n|$ because g_m divides f_n . Therefore

$$|g_i| \leq \binom{m}{i} \|f\|.$$

□

Lemma 3.4.11 Let $g \in \mathbb{Z}[x]$, $g = \sum_{i=0}^m g_i x^i$, and $B > 0$ with $|g_i| < B$. Let M be an integer with $M > 2B$ and write $g \bmod M = \sum_{i=0}^m \tilde{g}_i x^i$ where $\tilde{g}_i \in (-\frac{M}{2}, \frac{M}{2}]$. Then

$$g = \sum_{i=0}^m \tilde{g}_i x^i$$

in other words: $g = g \bmod M + kM$ and $g_i = \tilde{g}_i$.

PROOF. Suppose that $\tilde{g}_i \neq g_i$. Then $g_i - \tilde{g}_i = kM$ with $|k| \geq 1$. But $|\tilde{g}_i| \leq \frac{M}{2}$, so

$$|g_i| = |kM + \tilde{g}_i| \geq \frac{M}{2} > B$$

but this is a contradiction. \square

Based on these results we have the following procedure for factorising a polynomial f in $\mathbb{Z}[x]$. We assume that f is square-free. We choose a prime p such that p does not divide the leading coefficient of f , and such that $f \bmod p$ is also square-free. For simplicity first we assume that $f \bmod p$ is the product of at most two irreducible factors.

- If $f \bmod p$ is irreducible then $f \in \mathbb{Z}[x]$ is irreducible, and we stop.
- Otherwise $f \bmod p = (g \bmod p)(h \bmod p)$, where $g \bmod p$ and $h \bmod p$ in $\mathbb{F}_p[x]$ are irreducible and distinct. (These factors can be found using Berlekamp's algorithm.)
- By Hensel lifting we find g_k and h_k such that $f = g_k h_k \bmod p^k$.
- Using the theorem of Landau-Mignotte we find a B so that the coefficients of a factor of f have absolute value $< B$.

Let $M = p^{k_0}$ with k_0 such that $M > 2B$. Then we find \tilde{g}_i, \tilde{h}_i such that

$$\begin{aligned} g_{k_0} \bmod M &= \sum \tilde{g}_i x^i, & \tilde{g}_i &\in \left(-\frac{M}{2}, \frac{M}{2}\right] \\ h_{k_0} \bmod M &= \sum \tilde{h}_j x^j, & \tilde{h}_j &\in \left(-\frac{M}{2}, \frac{M}{2}\right]. \end{aligned}$$

Then $f = g_{k_0} h_{k_0}$ is the factorisation of f or f is irreducible. Indeed, if f is not irreducible then $f = \bar{g}\bar{h}$, for certain $\bar{g}, \bar{h} \in \mathbb{Z}[x]$. (We know that f has at most two factors over \mathbb{Z} , as this is the case modulo p .) Now $f \bmod p^{k_0} = g_{k_0} h_{k_0} \bmod p^{k_0} = (\bar{g} \bmod p^{k_0})(\bar{h} \bmod p^{k_0})$. However, by Theorem 3.4.5, the factorisation $f \bmod p^{k_0} = g_{k_0} h_{k_0} \bmod p^{k_0}$ is unique. Hence $g_{k_0} \bmod p^{k_0} = \bar{g} \bmod p^{k_0}$ and $h_{k_0} \bmod p^{k_0} = \bar{h} \bmod p^{k_0}$ (or the other way round, of course).

Example 3.4.12 As in Example 3.4.7, let $f = x^4 + 2x^3 - 3x^2 - 4x + 1$. Then $\|f\|^2 = 1 + 4 + 9 + 16 + 1 = 31$. Let $\alpha_2 x^2 + \alpha_1 x + \alpha_0$ be a factor of f . Then

$$|\alpha_2| \leq \binom{2}{2} \sqrt{31}, \quad |\alpha_1| \leq \binom{2}{1} \sqrt{31}, \quad |\alpha_0| \leq \binom{2}{0} \sqrt{31}$$

so $|\alpha_i| \leq 2\sqrt{31}$. Hence $|\alpha_i| \leq 11$, since $\alpha_i \in \mathbb{Z}$, so we can take $B = 11$.

We know that $f = g_3 h_3 \bmod 27$ where $g_3 = x^2 + 3x + 1$ and $h_3 = x^2 + 26x + 26$.

We choose $M = 27 > 2B = 22$. Now

$$\begin{aligned} g_3 \bmod M &= x^2 + 3x + 1 \\ h_3 \bmod M &= x^2 - x - 1 \end{aligned}$$

and $f = (x^2 + 3x + 1)(x^2 - x - 1)$; so we have factorised f .

3.4.2 Hensel lifting for more factors

Given $f \in \mathbb{Z}[x]$ and $f_1, \dots, f_r \in \mathbb{Z}[x]$, and p a prime such that f_i, f_j are coprime modulo p . We assume that $\deg(f_1) + \dots + \deg(f_r) = \deg(f)$, $f = f_1 \cdots f_r \pmod{p}$.

We compute $\tilde{f}_1, \dots, \tilde{f}_r \in \mathbb{Z}[x]$ with $\tilde{f}_i = f_i \pmod{p}$ and $f = \tilde{f}_1 \cdots \tilde{f}_r \pmod{p^k}$ for a certain k .

- 1) Let $m = \lceil \frac{r}{2} \rceil$ and $g = f_1 \cdots f_m$, $h = f_{m+1} \cdots f_r$.
- 2) By Hensel lifting we compute $\tilde{g}, \tilde{h} \in \mathbb{Z}[x]$ with $\tilde{g} = g \pmod{p}$, $\tilde{h} = h \pmod{p}$ and $f = \tilde{g}\tilde{h} \pmod{p^k}$.
- 3) Recursively we compute $\tilde{f}_1, \dots, \tilde{f}_r$ with $\tilde{f}_i = f_i \pmod{p}$, $\tilde{g} = \tilde{f}_1 \cdots \tilde{f}_m \pmod{p^k}$, $\tilde{h} = \tilde{f}_{m+1} \cdots \tilde{f}_r \pmod{p^k}$.
- 4) $f = \tilde{f}_1 \cdots \tilde{f}_r \pmod{p^k}$.

The algorithm

Now we describe a procedure for factorising a square-free $f \in \mathbb{Z}[x]$.

- 1) We choose a prime p such that $f \pmod{p}$ is square-free, and p does not divide the leading coefficient of f .
- 2) With Landau-Mignotte we compute a B such that

$$|\text{coefficient of a factor of } f| \leq B$$

and we choose k such that $M = p^k > 2B$.

- 3) Let f_1, \dots, f_r the irreducible factors of $f \pmod{p}$ (obtained with Berlekamp's algorithm).
- 4) By Hensel lifting we get $\tilde{f}_1, \dots, \tilde{f}_r \in \mathbb{Z}[x]$ such that $f = \tilde{f}_1 \cdots \tilde{f}_r \pmod{p^k}$ and $\tilde{f}_i = f_i \pmod{p}$.
- 5) For all $S \subset \{1, \dots, r\}$ we compute

$$g_S = \prod_{i \in S} \tilde{f}_i \quad \text{and} \quad h_S = \prod_{i \notin S} \tilde{f}_i.$$

We write the coefficients of g_S and h_S modulo M in $(-\frac{M}{2}, \frac{M}{2}]$.

If $f = g_S h_S$ then we found a factorisation. If for all S , $f_S \neq g_S h_S$ then f is irreducible.

Example 3.4.13 (Swinnerton-Dyer polynomials) Let $p_1 = 2, p_2 = 3, \dots, p_n$ be the first n primes, and consider the field

$$F_n = \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_n}).$$

Let f_n be the minimal polynomial of $\alpha_n = \sqrt{p_1} + \dots + \sqrt{p_n}$. Using Galois theory one shows that f_n has degree 2^n and the roots of f_n are

$$\epsilon_1 \sqrt{p_1} + \dots + \epsilon_n \sqrt{p_n}$$

where $\epsilon_i \in \{-1, 1\}$. Using a bit of algebraic number theory one sees that $f_n \in \mathbb{Z}[x]$ (the $\sqrt{p_i}$ lie in the ring of integers of F_n , hence so does α_n ; therefore $f_n \in \mathbb{Z}[x]$). Now let p be a prime, then all polynomials $x^2 - p_i$ have their roots in the quadratic extension \mathbb{F}_{p^2} of \mathbb{F}_p . It follows that $f_n \bmod p$ factorises into linear factors over \mathbb{F}_{p^2} . Hence over \mathbb{F}_p , it factorises into linear or quadratic factors. So $f_n \bmod p$ has at least 2^{n-1} factors over \mathbb{F}_p . Now f_n , being the minimal polynomial of α_n , is irreducible in $\mathbb{Q}[x]$. But in order to show that using the algorithm outlined in this section one needs to inspect at least $2^{2^{n-1}}$ subsets S . So we see that at least for some polynomials this algorithm is not without its drawbacks.

3.5 Exercises

1. Prove that $x^4 + x + 1 \in \mathbb{F}_2[x]$ is irreducible.
2. Factorise $x^7 + 1 \in \mathbb{F}_2[x]$ and $x^4 + x + 1 \in \mathbb{F}_3[x]$. (In coding theory, the factorisation of $x^7 + 1 \in \mathbb{F}_2[x]$ implies that there are $8 = 2^3$ cyclic codes of length 7 over \mathbb{F}_2 . The factorisation is used to find their generator polynomials; see J. H. van Lint, *Introduction to Coding Theory*, Chapter 6.)
3. Let $f = x^4 + 5x^3 + 6x^2 - x - 1$ and $g = x^2 + 1$, $h = x^2 + 2x + 2$. Then $g \bmod 3$ and $h \bmod 3$ are irreducible in $\mathbb{F}_3[x]$ and $f = gh \bmod 3$. Let g_k, h_k be the polynomials obtained from g, h by Hensel lifting. (So $g_1 = g$, $h_1 = h$, and $f = g_k h_k \bmod 3^k$.)
 - (a) Compute g_k, h_k for $k = 2, 3$.
 - (b) Prove that $g_k = x^2 + 3x + 1$ and $h_k = x^2 + 2x + 3^k - 1$ for $k \geq 2$.
 - (c) Show that the coefficients of a factor of f are ≤ 24 in absolute value. Factorise f .
4. Let $f = x^5 + 2x^3 + 2x^2 + x + 1$. Factorise $f \bmod 3$, and prove that $f \in \mathbb{Q}[x]$ is irreducible.
5. In this exercise we show that $f = x^4 + 1 \in \mathbb{Q}[x]$ is irreducible, but $f \bmod p$ factorises modulo every prime p .
 - (a) Factorise $f \bmod 2$ (hint: it is easy to find a root of f).
 - (b) Let p be an odd prime. Let V be the space of $v \in \mathbb{F}_p[x]/\langle f \rangle$ such that $v^p = v \bmod f$. Compute a basis of V , treating the cases $p = 1 + 4k$, k even and odd, and $p = 3 + 4k$, k even and odd, separately. Conclude that $f \bmod p$ factorises for all primes p .
 - (c) Factorise $f \bmod 5$.
 - (d) By Hensel lifting find $g, h \in \mathbb{Z}[x]$ with $f = gh \bmod 125$.
 - (e) Find a bound on the coefficients of a possible factor of f in $\mathbb{Z}[x]$ (Landau-Mignotte). Prove that f is irreducible.