6.966 Algebra and Computation		October 8, 1998
	Lecture 9	
Lecturer: Madhu Sudan		*Scribe: Amos Beimel*

In this lecture we will discuss Hilbert's irreducibility and show how to use it for black-box factorization of multivariate polynomials.

1 Hilbert Irreducibility Theorem

1.1 Motivating Discussion

Let f be a multivariate polynomial over the integers, that is $f(x_1, x_2, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$. We want to check if f is irreducible. Clearly, this can be done using an exhaustive search. Hilbert asked if this can be done more efficiently. His suggestion was to substitute the variables, that is, let $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$, where $a_i \in \mathbb{Z}$, and check if $f(a_1, a_2, \ldots, a_n)$ is a prime. If f is reducible, say

$$f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n) \cdot h(x_1, x_2, \ldots, x_n),$$

then "usually" $g(a_1, a_2, \ldots, a_n) \neq 1$ and $h(a_1, a_2, \ldots, a_n) \neq 1$ and $f(a_1, a_2, \ldots, a_n)$ is composite. Hilbert showed that some converse of this statement is true; if $f(a_1, a_2, \ldots, a_n)$ is often a composite then f is reducible.

The previous example is, in some sense, a reduction in the wrong direction since over the integers it is easier to check irreducibility then to check primality. However, we want to use this idea over arbitrary fields. The first obstacle is that $f(a_1, a_2, \ldots, a_n)$ is always a product of two elements of the field. To overcome this problem we use *partial substitutions*. That is, we substitute the values of the variables such that the result is a polynomial in few variables. A simple partial substitution is to substitute all the variables except for x_1: given $f(x_1, x_2, \ldots, x_n)$ we choose $a_2, \ldots, a_n$ and consider the polynomial $f(x_1, a_2, \ldots, a_n)$ (this is a univariate polynomial in x_1). This substitution and every other substitution which results in univariate polynomial are not good enough since over algebraically closed fields (e.g., $\mathbb{C}$) every univariate polynomial always factors. We will use substitutions which result in bivariate polynomials. This is a better idea since over every field there are irreducible bivariate polynomials, e.g., $x - y^2$.

1.2 The Theorem

We use the following substitution rule: for $f(x, y_1, y_2, \ldots, y_n) \in F[x, y_1, y_2, \ldots, y_n]$ and $\hat{a}, \hat{b} \in F^n$ we consider the bivariate polynomial $f(x, a_1t + b_1, a_2t + b_2, \ldots, a_nt + b_n)$. If f is reducible, say $f = p \cdot q$, then with high probability $p(x, a_1t + b_1, a_2t + b_2, \ldots, a_nt + b_n)$ and $q(x, a_1t + b_1, a_2t + b_2, \ldots, a_nt + b_n)$ are not constant, and $f(x, a_1t + b_1, a_2t + b_2, \ldots, a_nt + b_n)$ is reducible. The next theorem states that the converse is true if f is monic in x. In Section 1.7 we show how to check irreducibility of non-monic polynomials.

Theorem 1 *Let $S \subseteq F$ be a finite set, and $f(x, y_1, y_2, \ldots, y_n)$ be a monic polynomial in x whose total degree is at most d. If*

$$\frac{\partial f}{\partial x} \neq 0$$

and

$$\Pr_{\hat{a}, \hat{b}}[f(x, a_1t + b_1, a_2t + b_2, \ldots, a_nt + b_n) \text{ is reducible }] \geq \frac{O(d^5)}{|S|},$$

where $\hat{a}$ and $\hat{b}$ are chosen uniformly and independently from S^n, then f is reducible.

The rest of this section is devoted to the proof of this theorem.

1.3 A Warm-up Lemma

In this section we prove that if with "high" probability $f(x, b_1, \ldots, b_n)$ is not square free then f is reducible. This lemma will be used in the sequel to assume that $f(x, 0, \ldots, 0)$ is square free. Furthermore, the proof of this lemma is a warm-up for the proof of Theorem 1. This lemma requires that $\frac{\partial f}{\partial x} \neq 0$, which is not really a restriction when the characteristic of F is zero, but is more problematic over finite fields. We show how to check irreducibility over finite fields in Section 1.8. We start with a simple observation.

Observation 2 *If the polynomial $f(x, y_1, \ldots, y_n)$ is monic in x and $g(x, y_1, \ldots, y_n)$ divides f then g is monic in x.*

Lemma 3 *If $\frac{\partial f}{\partial x} \neq 0$ and*

$$\Pr_{\hat{b} \in S^n} [\ f(x, b_1, \ldots, b_n) \text{ is not square free }] > \frac{2d^2}{|S|},$$

then f is reducible.

Proof Recall that $f(x, b_1, \ldots, b_n)$ is a univariate polynomial in $F[x]$. We denote the discriminant of $f(x, b_1, \ldots, b_n)$ by $\zeta_{\hat{b}}$. If $f(x, b_1, \ldots, b_n)$ is not square free for some $\hat{b}$ then

$$\gcd\left(f(x, b_1, \ldots, b_n), \frac{df(x, b_1, \ldots, b_n)}{dx}\right) \neq 1.$$

That is, the discriminant $\zeta_{\hat{b}}$ is 0.

Now, view $f(x, y_1, \ldots, y_n)$ as a polynomial in $F(y_1, \ldots, y_n)[x]$, that is, as a polynomial in x whose coefficients are polynomials in $y_1, \ldots, y_n$. The discriminant of f is a polynomial in $y_1, \ldots, y_n$ which we denote $\zeta(y_1, \ldots, y_n)$. Clearly, $\zeta(b_1, \ldots, b_n) = \zeta_{\hat{b}}$. By the conditions of the lemma, with probability greater than $(2d^2)/|S|$ over the $\hat{b}$'s, the discriminant $\zeta(b_1, \ldots, b_n)$ is zero. Thus, $\zeta(y_1, \ldots, y_n)$ is a polynomial of the n variables $y_1, \ldots, y_n$ which has degree at most $2d^2$, and is zero with "high" probability. By the Schwartz-Zippel lemma, the discriminant of f, namely $\zeta(y_1, \ldots, y_n)$, is identically zero, i.e., $\gcd(f, \frac{\partial f}{\partial x}) \neq 1$. Since $\frac{\partial f}{\partial x} \neq 0$, this means that f is reducible.[1] ■

We conclude that either with high probability $f(x, b_1, \ldots, b_n)$ is not square free and by Lemma 3 the polynomial f is reducible, or there exists some $\hat{b}$ such that $f(x, b_1, \ldots, b_n)$ is square free. We fix such $\hat{b}$ and consider the polynomial $f(x, y_1 + b_1, \ldots, y_n + b_n)$ which is reducible if and only if $f(x, y_1 + b_1, \ldots, y_n + b_n)$ is reducible. Thus,

$$\Pr_{\hat{a}}[f(x, a_1 t + b_1, a_2 t + b_2, \ldots, a_n t + b_n) \text{ is reducible }] \geq \frac{O(d^5)}{|S|},$$

where the probability is only taken over the vectors $\hat{a}$ chosen uniformly from S^n. To simplify the notations, we "forget" the fixed $\hat{b}$ and assume that $f(x, 0, \ldots, 0)$ is square free.

1.4 Ideas of the proof

We next summarize the ideas of the proof of Lemma 3 and then we show how we use this ideas for proving the the theorem. We have shown that if $f(x, b_1, \ldots, b_n)$ is not square free then its discriminant is zero. We observe that the discriminant of $f(x, b_1, \ldots, b_n)$ is obtained from the discriminant of $f(x, y_1, \ldots, y_n)$ by substituting $y_1 = b_1, \ldots, y_n = b_n$. Thus, the discriminant of $f(x, y_1, \ldots, y_n)$ is a polynomial that is zero for "many" substitutions and therefore must be identically zero. However, if the discriminant of f is zero then f is reducible.

[1] Over fields of characteristic zero we can deduce that f is not square free. This is not necessarily true over finite fields.

A similar idea is used to prove Theorem 1. We use Hensel's lifting lemma to find a factor of the univariate polynomial $f(x, a_1t, \ldots, a_nt)$. If $f(x, a_1t, \ldots, a_nt)$ is reducible (and it is with "high" probability) then we succeed in finding a factor (since we can assume that $f(x, 0, \ldots, 0)$ is square free). We use the same procedure to try and find a factor of f. However, in the latter case we do not know if the conditions that guaranteed the success for finding a factor hold. Nevertheless, we observe that we succeed if some determinant is zero. Similar to the proof of Lemma 3, we substitute $y_1 = a_1, \ldots, y_n = a_n$ in this determinant and observe that since we managed to factor $f(x, a_1t, \ldots, a_nt)$ then under this substitution this determinant is zero. This implies that the determinant is zero for many substitutions and therefore must be identically zero and we succeed in factoring f.

1.5 A Technical Claim

In this section we define some auxiliary polynomial and investigate some of their properties. First, we define

$$p(x, t, y_1, y_2, \ldots, y_n) \stackrel{\text{def}}{=} f(x, y_1t, y_2t, \ldots, y_nt).$$

Next, we define for every $\hat{a} \in F^n$ a bivariate polynomial

$$f_{\hat{a}}(x, t) \stackrel{\text{def}}{=} f(x, a_1t, a_2t, \ldots, a_nt).$$

Notice that $p(x, t, a_1, a_2, \ldots, a_n) = f_{\hat{a}}(x, t)$, and $p(x, 1, y_1, \ldots, y_n) = f(x, y_1, \ldots, y_n)$. We prove that it is enough to prove that p is reducible.

Claim 4 *If p is reducible then f is reducible.*

Proof Assume $p = \alpha \cdot \beta$ for some polynomials α and β which are not constants. Recall that f is monic in x.[2] This implies that p, α and β are monic in x.

Thus,

$$f(x, y_1, \ldots, y_n) = p(x, 1, y_1, \ldots, y_n) = \alpha(x, 1, y_1, \ldots, y_n) \cdot \beta(x, 1, y_1, \ldots, y_n).$$

Since f is irreducible then either $\alpha(x, 1, y_1, \ldots, y_n)$ is constant or $\beta(x, 1, y_1, \ldots, y_n)$ is constant. W.l.o.g., assume that $\alpha(x, 1, y_1, \ldots, y_n) \equiv 1$. Since α is non-constant, there is a non-zero polynomial $\alpha_1(x, t, y_1, \ldots, y_n)$ such that

$$\alpha(x, t, y_1, \ldots, y_n) = 1 + (1 - t)\alpha_1(x, t, y_1, \ldots, y_n).$$

This contradicts the fact that α is monic in x. ■

1.6 Using Hensel's Lifting Lemma

This section is the main technical part of the proof of Theorem 1. We show how to use the factoring algorithm provided by Hensel's lifting lemma to prove that p is reducible. As explained in Section 1.3, we can assume that $f(x, 0, \ldots, 0)$ is square free.

Since $f(x, 0, \ldots, 0)$ is monic in x and $f(x, a_1, \ldots, a_n)$ factors for some $\hat{a}$, there are non-constant polynomials $g_0(x)$ and $h_0(x)$ such that

$$f(x, 0, \ldots, 0) = g_0(x) \cdot h_0(x). \tag{1}$$

It must hold that $\gcd(g_0(x) \cdot h_0(x)) = 1$ (since $f(x, 0, \ldots, 0)$ is square free), and we can choose g_0 and h_0 such that g_0 is monic and irreducible. Since $f(x, 0, \ldots, 0) = f_{\hat{a}}(x, t) \pmod{t}$, we get

$$f_{\hat{a}}(x, t) = g_0(x) \cdot h_0(x) \pmod{t}. \tag{2}$$

[2] If f is not monic in x then p might be reducible. For example, if $f(x, y_1, y_2) = xy_1 + y_2$ then $p(x, t, y_1) = t(xy_1 + y_2)$ is reducible.

For every $\hat{a} \in F^n$ we can apply Hensel's lifting lemma to $f_{\hat{a}}$ and find polynomials $g_{k,\hat{a}}(x,t)$ and $h_{k,\hat{a}}(x,t)$ such that

$$f_{\hat{a}}(x,t) = g_{k,\hat{a}}(x,t) \cdot h_{k,\hat{a}}(x,t) \pmod{t^k}. \tag{3}$$

Furthermore, $g_{k,\hat{a}}(x,t) = g_0(x) \pmod{t}$ and $h_{k,\hat{a}}(x,t) = h_0(x) \pmod{t}$.

Claim 5 *There exists a polynomial $g_k(x,t,y_1,\ldots,y_n)$ such that for every $\hat{a} \in F^n$*

$$g_{k,\hat{a}}(x,t) = g_k(x,t,a_1,\ldots,a_n).$$

Proof Equation (1) implies that

$$f(x,y_1,\ldots,y_n) = g_0(x) \cdot h_0(x) \pmod{y_1,\ldots,y_n} \tag{4}$$

(since $f(x,0,\ldots,0) = f(x,y_1,\ldots,y_n) \pmod{y_1,\ldots,y_n}$). Thus, we can apply Hensel's lifting lemma to f and find polynomials $g'_k(x,y_1,\ldots,y_n)$ and $h'_k(x,y_1,\ldots,y_n)$ such that

$$f(x,y_1,\ldots,y_n) = g'_k(x,y_1,\ldots,y_n) \cdot h'_k(x,y_1,\ldots,y_n) \pmod{(y_1,\ldots,y_n)^k}. \tag{5}$$

Now, define

$$g_k(x,t,y_1,\ldots,y_n) \stackrel{\text{def}}{=} g'_k(x,y_1 t,\ldots,y_n t),$$

and

$$h_k(x,t,y_1,\ldots,y_n) \stackrel{\text{def}}{=} h'_k(x,y_1 t,\ldots,y_n t).$$

It holds that

$$\begin{aligned} p(x,t,y_1,\ldots,y_n) &= f(x,y_1 t,\ldots,y_n t) \\ &= g'_k(x,y_1 t,\ldots,y_n t) \cdot h'_k(x,y_1 t,\ldots,y_n t) \pmod{(y_1,\ldots,y_n)^k} \\ &= g_k(x,t,y_1,\ldots,y_n) \cdot h_k(x,t,y_1,\ldots,y_n) \pmod{t^k} \end{aligned}$$

(the last equality holds since in any monomial in g_k and h_k the degree of t is the sum of the degrees of the y_i's in the monomial). In particular,

$$f_{\hat{a}}(x,t) = p(x,t,a_1,\ldots,a_n) = g_k(x,t,a_1,\ldots,a_n) \cdot h_k(x,t,a_1,\ldots,a_n) \pmod{t^k}.$$

By the uniqueness guaranteed by Hensel's lemma we deduce $g_k(x,t,a_1,\ldots,a_n) = g_{k,\hat{a}}(x,t)$. ■

As we proved in Lectures 7, if $f_{\hat{a}}$ is reducible, g_0 and h_0 have no common factors, and g_0 is irreducible then for $k = O(d^2)$ there are polynomials $g_{\hat{a}}(x,t), \ell_{k,\hat{a}}(x,t) \neq 0$ such that

$$\begin{aligned} g_{\hat{a}}(x,t) &= g_{k,\hat{a}}(x,t) \cdot \ell_{k,\hat{a}}(x,t) \pmod{t^k}, && (6) \\ \deg_x(g_{\hat{a}}(x,t)) &< \deg_x(f_{\hat{a}}(x,t)), \text{ and } \deg_t(g_{\hat{a}}(x,t)) \leq \deg_t(f_{\hat{a}}(x,t)). && (7) \end{aligned}$$

We next prove that there are polynomials g, ℓ_k satisfying similar properties with respect to p.

Claim 6 *There are polynomials $g(x,t,y_1,\ldots,y_n), \ell_k(x,t,y_1,\ldots,y_n) \neq 0$ such that*

$$\begin{aligned} g(x,t,y_1,\ldots,y_n) &= g_k(x,t,y_1,\ldots,y_n) \cdot \ell_k(x,t,y_1,\ldots,y_n) \pmod{t^k}, && (8) \\ \deg_x(g(x,t,y_1,\ldots,y_n)) &< \deg_x(p(x,t,y_1,\ldots,y_n)), && (9) \\ \deg_t(g(x,t,y_1,\ldots,y_n)) &\leq \deg_x(p(x,t,y_1,\ldots,y_n)), && (10) \\ \sum_{i=1}^{n} \deg_{y_i}(g(x,t,y_1,\ldots,y_n)) &= O(k^5). && (11) \end{aligned}$$

Proof Using Claim 5 and Equation (6) can get

$$g_{\hat{a}}(x,t) = g_k(x,t,a_1,\ldots,a_n) \cdot \ell_{k,\hat{a}}(x,t) \pmod{t^k}. \tag{12}$$

This is a homogeneous linear system, where the unknowns are the coefficients of $g_{\hat{a}}$ and $\ell_{\hat{a},k}$. The degree of x in all these polynomials is at most d (where d is the total degree of f) and the degree of t in all these polynomials is less than $k = O(d^2)$. Thus, the number of unknowns in the system, denoted m is $O(d^3)$. If $f_{\hat{a}}$ is reducible, then this system has a solution in which the degrees of x and t in $g_{\hat{a}}$ are as specified in (7). These degree restrictions are expressed as additional linear equations which state that some coefficients are equal to zero.

Now, consider the following homogeneous linear system

$$g(x,t,y_1,\ldots,y_n) = g_k(x,t,y_1,\ldots,y_n) \cdot \ell_k(x,t,y_1,\ldots,y_n) \pmod{t^k}.$$

We view g, g_k and ℓ_k as polynomials in $F(y_1,\ldots,y_n)[x,t]$, and this system also has m unknowns. If this system has a no solution (in the field $F(y_1,\ldots,y_n)$) where the degrees of x and t in g are "small", the linear system defines an $m \times m$ square submatrix that has a non-zero determinant. This determinant is a polynomial $D(y_1,\ldots,y_n)$ whose total degree is at most $mk = O(d^5)$ (since every entry in the matrix is a polynomial of total degree less than k). However, for every $\hat{a}$ such that $f_{\hat{a}}$ is reducible the determinant $D(a_1,\ldots,a_n)$ is zero (since (12) has a solution). By the Schwartz-Zippel lemma $D(y_1,\ldots,y_n) \equiv 0$, thus we can solve the linear system and find g and ℓ_k. Since this is a homogeneous system, we can always find a solution in $F[y_1,\ldots,y_n]$ (and not just in $F(y_1,\ldots,y_n)$). Furthermore, using Gaussian elimination, we obtain a solution in which the total degree of each coefficient is $O(k^5)$. ■

As proved in previous lectures, every polynomial satisfying (8) – (10) can be used to find a non-trivial factor of p. The problem is that this factor is in the field $F(y_1,\ldots,y_n)$. However, we know that for every $\hat{a} \in F^n$ the polynomials $g(x,t,a_1,\ldots,a_n)$ and $\ell_k(x,t,a_1,\ldots,a_n)$ are a solution of (6) and (7). Thus, as proved in Lectures 7 and 8,

$$\gcd_x(p(x,t,a_1,\ldots,a_n), g(x,t,a_1,\ldots,a_n)) \quad \neq \quad 1, \tag{13}$$

which implies that the resultant of $p(x,t,a_1,\ldots,a_n)$ and $g(x,t,a_1,\ldots,a_n)$ is zero. That is, the resultant of $p(x,t,y_1,\ldots,y_n)$ and $g(x,t,y_1,\ldots,y_n)$ is zero for every substitution of $y_1,\ldots,y_n$, i.e., the resultant is identically zero. Hence, p has a non-trivial factor, and the polynomials p and f are reducible. This completes the proof of Theorem 1.

1.7 Testing Irreducibility of Non-monic Polynomials

In Theorem 1 we assumed that f is monic in x. We show how to remove this assumption. Given a polynomial $r(x,y_1,\ldots,y_n)$ which has degree at least 1 in x, we construct a polynomial $f(x,y_1,\ldots,y_n)$ that is monic in x such that f is reducible then r is reducible, and if r has a "good" factorization (as described in Claim 8) then f is reducible. Furthermore, from the factorization of f, one can easily construct the factorization of r.

We next present the construction. Let $r(x,y_1,\ldots,y_n) = \sum_{i=0}^{e} r_i(y_1,\ldots,y_n)x^i$ where e is the degree of x in r, that is, $r_e \neq 0$. We define

$$f(x,y_1,\ldots,y_n) \stackrel{\text{def}}{=} (r_e(y_1,\ldots,y_n))^{e-1} \cdot r\left(\frac{x}{r_e(y_1,\ldots,y_n)}, y_1,\ldots,y_n\right). \tag{14}$$

Claim 7 *The function f is a polynomial in $F[x,y_1,\ldots,y_n]$ which is monic in x. Furthermore, if f is reducible then r is reducible.*

Proof By definition of f it holds that

$$\begin{aligned} f(x, y_1, \ldots, y_n) &= \sum_{i=0}^{e} (r_e(y_1, \ldots, y_n))^{e-1-i} r_i(y_1, \ldots, y_n) x^i \\ &= x^e + \sum_{i=0}^{e-1} (r_e(y_1, \ldots, y_n))^{e-1-i} r_i(y_1, \ldots, y_n) x^i. \end{aligned}$$

Thus, f is a polynomial in $F[x, y_1, \ldots, y_n]$ that is monic in x.

Furthermore,

$$r(x, y_1, \ldots, y_n) \cdot (r_e(y_1, \ldots, y_n))^{e-1} = f(x \cdot r_e(y_1, \ldots, y_n), y_1, \ldots, y_n).$$

Assume that f is reducible, say $f = g \cdot h$. Then,

$$r(x, y_1, \ldots, y_n) \cdot (r_e(y_1, \ldots, y_n))^{e-1} = g(x \cdot r_e(y_1, \ldots, y_n), y_1, \ldots, y_n) \cdot h(x \cdot r_e(y_1, \ldots, y_n), y_1, \ldots, y_n).$$

Thus, there is some integer $e_1 \geq 0$ and polynomials g' and h' such that

$$g(x \cdot r_e(y_1, \ldots, y_n), y_1, \ldots, y_n) = r_e(y_1, \ldots, y_n)^{e_1} g'(x, y_1, \ldots, y_n)$$

and

$$h(x \cdot r_e(y_1, \ldots, y_n), y_1, \ldots, y_n) = r_e(y_1, \ldots, y_n)^{e-e_1-1} h'(x, y_1, \ldots, y_n).$$

Therefore, $r(x, y_1, \ldots, y_n) = g'(x, y_1, \ldots, y_n) \cdot h'(x, y_1, \ldots, y_n)$. Since f is monic in x, the polynomials g and h are monic in x as well, and it is not possible that g' or h' are constant. Thus, we get a factorization of r. ■

We next claim that if r is reducible and has a "good" factorization then f is reducible. The proof is similar to the proof of Claim 7 and therefore omitted.

Claim 8 *If $r = g(x, y_1, \ldots, y_n) \cdot h(x, y_1, \ldots, y_n)$ for some polynomials g and h such that $\deg_x(g) > 0$ and $\deg_x(h) > 0$ then f is reducible.*

Therefore, we get an efficient procedure to check if a polynomial $r(x, y_1, \ldots, y_n)$ is irreducible.

- If $\gcd(r_0(y_1, \ldots, y_n), r_1(y_1, \ldots, y_n), \ldots, r_e(y_1, \ldots, y_n)) \neq 1$ then r is reducible.
- Otherwise, construct f as described in (14).
- Test, using Hilbert's irreducibility theorem, if f is reducible. That is, choose $\hat{a}$ and $\hat{b}$ at random from S^n. If $f(x, ta_1 + b_1, \ldots, ta_n + b_n)$ is reducible, output "r is reducible".

Notice that if the degree of r is d then the degree of f is at most d^2, thus to use Hilbert's irreducibility theorem we need a field F with at least $\Omega(d^{10})$ elements.

1.8 Checking Irreducibility over Finite Fields

Theorem 1 requires that $\frac{\partial f}{\partial x} \neq 0$. Over fields with characteristic zero this only means that we have to choose a variable that appears in the polynomial. However, over finite fields there are irreducible polynomials r whose degree in x_1 is at least 1 but $\frac{\partial r(x_1, \ldots, x_n)}{\partial x_1} = 0$ (e.g., consider the irreducible polynomial $x_1^6 + x_2$ over $GF(2)$). Nevertheless, if $\frac{\partial r(x_1, \ldots, x_n)}{\partial x_i} = 0$ for every i then the degree of each variable in each monomial is divisible by p – the characteristic of the field. Thus, if r is a non-constant polynomial and F is a finite field then r is a p-th power of some polynomial.

We are left with the case that $\frac{\partial r(x_1, \ldots, x_{n+1})}{\partial x_i} \neq 0$ for some i. In this case if we covert r to a function f that is monic in x_i as described in Section 1.7, then $\frac{\partial f(y_1, \ldots, x_n)}{\partial x_i} \neq 0$. Thus, we can use the test of Section 1.7 to check if r is reducible.

2 Black-box Factorization of Multivariate Polynomials

The task of black-box factorization of multivariate polynomials is defined as follows. Given a black-box access to a polynomial f, construct black-boxes for its factors where each black-box has access to the black-box for f. We emphasize that we do not know the number of factors in advance and we have to compute it. This is done in a preprocessing stage, in which we also "identify" each black-box with a factor of f. Later on, we are given an assignment and an index i and we need to return the value of the i-th factor on this assignment. Both the preprocessing stage and the computing stage use the black-box for f.

We use Hilbert's irreducibility theorem as well as an algorithm for factoring polynomials with 3 variables for the black-box factorization. The preprocessing stage we describe is randomized and has a small probability of failing. However, if the preprocessing stage succeeded then we always compute the correct value of the factor (provided that we manage to factor some polynomial with 3 variables). In the rest of this section we assume that we are working over a field that is big enough, f is monic in x, and $\frac{\partial f}{\partial x} \neq 0$.[3]

2.1 Preprocessing Stage

Our first task is to determine the number of factors of f. Assume that f has $n+1$ variables $x, y_1, \ldots, y_n$, it is monic in x, the derivative $\frac{\partial f}{\partial x}$ is non-zero, and the total degree of f is at most d. Furthermore, let $S \subseteq F$ be a set of size $\Omega(d^6/\epsilon)$ (for some $\epsilon < 1$). To compute the number of factors we use the following simple algorithm:

- Chose $\hat{a}, \hat{b}$ at random from S^n.
- Factor $f_{\hat{a},\hat{b}}(x,t) \stackrel{\text{def}}{=} f(x, a_1t_1+b_1, \ldots, a_nt_n+b_n)$. Let ℓ be the number of factors and $\tilde{f}^1(x,t), \ldots, \tilde{f}^\ell(x,t)$ be the factors.

Assume that f has ℓ' irreducible factors $f^1(x, y_1, \ldots, y_n), \ldots, f^{\ell'}(x, y_1, \ldots, y_n)$. Since f is monic in x, all the factors are monic in x. For a fixed i, the probability that the polynomial $f^i_{\hat{a},\hat{b}}(x,t) \stackrel{\text{def}}{=} f^i(x, a_1t_1 + b_1, \ldots, a_nt_n + b_n)$ is irreducible is at most $O(d^5/|S|)$. Since $\ell' \leq d$, the probability that there exists some f^i that is reducible is at most $O(d^6/|S|) \leq \epsilon$. Each $f^i_{\hat{a},\hat{b}}(x,t)$ is not constant since f^i is monic in x. These two facts imply that $\ell' = \ell$ with probability at least $1-\epsilon$. By renaming $f^1, \ldots, f^\ell$ we can assume that $\tilde{f}^i(x,t) = f^i_{\hat{a},\hat{b}}(x,t)$. We define the i-th black-box as computing the factor f^i of f.

We should remark that to factor $f_{\hat{a},\hat{b}}(x,t)$ we need to have its representation as a bivariate polynomial while we only have a black-box to it. However, we can interpolate and find the $\binom{d+1}{2}$ coefficients of $f_{\hat{a},\hat{b}}$.

2.2 Computing the Value of Box B_i

Given $\alpha, \beta_1, \ldots, \beta_n$ we want to compute $f^i(\alpha, \beta_1, \ldots, \beta_n)$. For this we define the function

$$g(x, t_1, t_2) \stackrel{\text{def}}{=} f(x, a_1t_1 + b_1 + (\beta_1 - b_1)t_2, \ldots, a_nt_n + b_n + (\beta_n - b_n)t_2).$$

The function g is a restriction of f to a plane that contains both the point $\alpha, \beta_1, \ldots, \beta_n$ and the line $a_1t+b_1, \ldots, a_nt+b_n$. That is, $g(\alpha, 0, 1) = f(\alpha, \beta_1, \ldots, \beta_n)$ and $g(x, t, 0) = f_{\hat{a},\hat{b}}(x,t)$. The following algorithm computes $f^i(\alpha, \beta_1, \ldots, \beta_n)$.

1. Factor $g(x, t_1, t_2)$ into factors $g^1, \ldots, g^{\ell''}$.
2. Find an index j such that $g^j(x, t, 0) = \tilde{f}^i$.
3. Output $g^j(\alpha, 0, 1)$.

[3] In the black box scenario, we cannot use the construction of Section 1.7 since we do not know what the coefficient of the largest power of x is.

Lemma 9 *If the preprocessing stage succeeded and the factorization of g is correct, then the above algorithm is correct.*

Proof We will prove that we can always find an index j in Step 2 of the algorithm, and

$$g^j(x,t_1,t_2) = f^i(x, a_1t_1 + b_1 + (\beta_1 - b_1)t_2, \ldots, a_nt_n + b_n + (\beta_n - b_n)t_2).$$

Thus, $g^j(\alpha, 0, 1) = f^i(\alpha, \beta, \ldots, \beta_n)$ as required.

First,

$$g(x,t_1,t_2) = \prod_{i=1}^{\ell} f^i(x, a_1t_1 + b_1 + (\beta_1 - b_1)t_2, \ldots, a_nt_n + b_n + (\beta_n - b_n)t_2). \tag{15}$$

Since f is monic in x, every factor f^i is monic in x and $f^i(x, a_1t_1+b_1+(\beta_1-b_1)t_2, \ldots, a_nt_n+b_n+(\beta_n-b_n)t_2)$ is not constant. Thus, g has at least ℓ irreducible factors. Now,

$$f_{\hat{a},\hat{b}}(x,t) = g(x,t,0) = \prod_{j=1}^{\ell''} g^j(x,t,0).$$

However, g is monic in x and therefore $g^j(x,t_1,t_2)$ is monic in x and $g^j(x,t,0)$ is non-constant. Since $f_{\hat{a},\hat{b}}$ has ℓ irreducible factors, g has at most ℓ factors. Thus, g has exactly ℓ factors which are described in Equation (15). Since $\tilde{f}^i(x,t) = f^i_{\hat{a},\hat{b}}(x,t)$, the lemma follows. ■