On a Theorem by R. Dedekind

Erich Kaltofen
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, New York 12181

To Arjen K. Lenstra

On the occasion of his successful doctoral thesis defense

1. Introduction

A lemma by C. F. Gauss [3, article 42] states that if a polynomial over a unique factorization domain is irreducible it remains irreducible over the field of quotients of this domain. This theorem does not hold, in general, if the coefficient domain is the ring of algebraic integers $\mathcal{O}_K$ of a number field K. Using an old generalization of Gauss' lemma by R. Dedekind [1] it is shown in [8] that the additional denominator needed to express the factorization of a univariate polynomial over K can be choosen equal to the leading coefficient of the polynomial to be factored. Several papers on factoring multivariate polynomials over algebraic number fields [4, 5, 6, 7, 8] have made reference to this fact without showing that the univariate lemma generalizes. In section 2 we will prove a multivariate version of first Dedekind's theorem and then a slightly more general multivariate version of the theorem in [8] designating possible denominators. Since the leading coefficient of a multivariate polynomial is not uniquely determined, we can further optimize the choice of a sufficient denominator needed to express the multivariate factors, a phenomenon which seems not to have been noticed before.

2. Main Results

We first state and prove Dedekind's generalization of Gauss' lemma. But first we need to clarify our notation. We say that an algebraic integer $a \in \mathcal{O}_K$ divides $b \in \mathcal{O}_K$, $a|b$, iff there exists a number $c \in \mathcal{O}_K$ such that $ac = b$.

Theorem 1 [1]: Let K be a number field, $f(x) = a_\ell x^\ell + \cdots + a_0$, $g(x) = b_m x^m + \cdots + b_0 \in \mathcal{O}_K[x]$, $h(x) = f(x)g(x) = c_n x^n + \cdots + c_0$, $n = \ell + m$, and let $d \in \mathcal{O}_K$ such that $d \mid c_k$ for all $0 \leq k \leq n$. Then $d \mid a_i b_j$ for all $0 \leq i \leq \ell$ and $0 \leq j \leq m$.

Proof [2, Sec. 4]: We first prove that $d \mid a_\ell b_j$ for all $0 \leq j \leq m$. For $j = m$, $d \mid c_n = a_\ell b_m$ by assumption. Let $\zeta_1, \ldots, \zeta_m$ be the roots of $g(x)$ and $\zeta_{m+1}, \ldots, \zeta_{m+\ell}$ the roots of $f(x)$. Multiplying $g(x)$ by a_ℓ,

$$a_\ell g(x) = c_n x^m + a_\ell b_{m-1} x^{m-1} + \cdots + a_\ell b_0$$

we see that

$$(-1)^{m-j} \frac{a_\ell b_j}{c_n} = s_{m-j}(\zeta_1, \ldots, \zeta_m)$$

where s_i denotes the i-th basic symmetric function

$$s_i(\zeta_1, \ldots, \zeta_m) = \sum_{1 \leq m_1 < \cdots < m_i \leq m} \zeta_{m_1} \cdots \zeta_{m_i} .$$

Now, with $t = n!$, the coefficients of

$$z^t + e_{t-1} z^{t-1} + \cdots + e_0 = \prod_{\sigma \in S_n} (z - s_{m-j}(\zeta_{\sigma(1)}, \ldots, \zeta_{\sigma(m)})) ,$$

S_n the set of permutations on $\{1, \ldots, n\}$, are symmetric functions in $\zeta_1, \ldots, \zeta_n$. Moreover, each individual ζ_k occurs to power $t - i$ in e_i. By the fundamental theorem on symmetric functions, e_i can be written as an integral polynomial in

$$s_1(\zeta_1, \ldots, \zeta_n) = -\frac{c_{n-1}}{c_n} , \ldots, s_n(\zeta_1, \ldots, \zeta_n) = (-1)^n \frac{c_0}{c_n}$$

of total degree $t - i$. Therefore, $d^{t-i} \mid c_n^{t-i} e_i$ and

$$\left(\frac{c_n z}{d}\right)^t + \frac{c_n e_{n-1}}{d}\left(\frac{c_n z}{d}\right)^{t-1} + \cdots + \frac{c_n^t e_0}{d^t} = 0$$

for $z = s_{m-j}(\zeta_1,\ldots,\zeta_m)$ and has coefficients in $\mathcal{O}_K$. Thus $c_n\, s_{m-j}(\zeta_1,\ldots,\zeta_m)/d = (-1)^{m-j}\, a_\ell b_j/d \in \mathcal{O}_K$. For arbitrary $a_i b_j$ the theorem now follows by induction on ℓ. For $\ell = 0$, $d \mid c_j = a_\ell b_j$, $0 \le j \le m$, by assumption. By our previous conclusions, the coefficients of

$$(f(x) - a_\ell x^\ell)\, g(x) = (c_{n-1} - a_\ell b_{m-1})\, x^{n-1} + \cdots + (c_\ell - a_\ell b_0) x^\ell$$

$$+ c_{\ell-1} x^{\ell-1} + \cdots + c_0$$

are divisible by d, thus by induction hypothesis, are the products $a_i b_j$, $0 \le i \le \ell-1$, $0 \le j \le m$. □

We next show that theorem 1 generalizes to multivariate polynomials.

Theorem 2: Let K be a number field,

$$f(x_1,\ldots,x_v) = \sum_{0 \le i_j \le \ell} a_{i_1,\ldots,i_v}\, x_1^{i_1} \cdots x_v^{i_v} \in \mathcal{O}_K[x_1,\ldots,x_v],$$

$$g(x_1,\ldots,x_v) = \sum_{0 \le i_j \le m} b_{i_1,\ldots,i_v}\, x_1^{i_1} \cdots x_v^{i_v} \in \mathcal{O}_K[x_1,\ldots,x_v],$$

$$h = f\,g = \sum_{0 \le i_j \le n} c_{i_1,\ldots,i_v}\, x_1^{i_1} \cdots x_v^{i_v}, \quad n = \ell + m,$$

and let $d \in \mathcal{O}_K$ be such that $d \mid c_{i_1,\ldots,i_v}$ for all $0 \le i_j \le n$.
Then $d \mid a_{i_1,\ldots,i_v}\, b_{k_1,\ldots,k_v}$ for all $0 \le i_j \le \ell$, $0 \le k_j \le m$.

Proof: We use Kronecker's homomorphism on f, g and h, that is we substitute

$$y^{(n+1)^{j-1}} \quad \text{for } x_j,\ 1 \le j \le v.$$

Thus,

$$f(y) = f\left(y, y^{(n+1)}, \ldots, y^{(n+1)^{(v-1)}}\right) = \sum a_{i_1,\ldots,i_v}\, y^{i_1+(n+1)i_2+\cdots+(n+1)^{v-1} i_v}$$

has as its coefficients 0 or the individual $a_{i_1,\ldots,i_v}$. Similarly, we map g and h to $\overline{g}(y)$ and $\overline{h}(y) \in \mathcal{O}_K[y]$. Theorem 1 now applies to $\overline{f}\,\overline{g} = \overline{h}$ and proves our statement. $\square$

We finally prove a multivariate version of Lemma 7.1 in [8]. However, we shall use a weaker notion for monicity which yields a slight improvement to the denominator prediction methods discussed later. We first define:

The polynomial $h(x_1,\ldots,x_v) \in K[x_1,\ldots,x_v]$, K a number field, is called <u>weakly normalized</u> if one of its coefficients ($\in K$) is a unit in $\mathcal{O}_K$.

<u>Theorem 3:</u> Let K be a number field and let $h(x_1,\ldots,x_v) \in (1/r)\mathcal{O}_K[x_1,\ldots,x_v]$ with $r \in \mathcal{O}_K$, $r \neq 0$. Assume that $f\,g = h$ with $f, g \in K[x_1,\ldots,x_v]$ weakly normalized. Then $f, g \in (1/r)\mathcal{O}_K[x_1,\ldots,x_v]$.

<u>Proof:</u> Let $f, g \in (1/s)\ \mathcal{O}_K[x_1,\ldots,x_v]$, $s \in \mathcal{O}_K$, such that $r \mid s$ (i.e. s may not be optimal). Then $s^2h = (sf)\ (sg)$ with $sf, sg \in \mathcal{O}_K[x_1,\ldots,x_v]$ and s^2/r divides all coefficients of s^2h. By theorem 2, s^2/r must divide all products $(sa_{i_1,\ldots,i_v})\ (sb_{k_1,\ldots,k_v})$ of coefficients $a_{i_1,\ldots,i_v}$ of f and $b_{k_1,\ldots,k_v}$ of g. This, in turn, is equivalent to

$$ra_{i_1,\ldots,i_v} \cdot b_{k_1,\ldots,k_v} \in \mathcal{O}_K$$

which, if we choose $a_{i_1,\ldots,i_v}$ the unit coefficient in f or $b_{k_1,\ldots,k_v}$ the unit coefficient in g, shows that

$$ra_{i_1,\ldots,i_v},\ rb_{k_1,\ldots,k_v} \in \mathcal{O}_K \quad . \quad \square$$

In theunivariate case $(v = 1)$ it is sufficient to choose f, g and h monic in order to enforce the weak normalization assumption. In the multivariate case one sufficient condition is that the non-zero monomials in f, g and h of maximum exponent vector with respect to a lexicographical ordering are normalized to 1. In fact, besides lexicographical orderings, any linear ordering on the exponent vectors of the monomials which satisfies

(+) $(i_1,\dots,i_v) \prec (j_1,\dots,j_v) \Longrightarrow (i_1+k_1,\dots,i_v+k_v) \prec (j_1+k_1,\dots,j_v+k_v)$

could be selected. One such ordering is

$$(i_1,\dots,i_v) \prec (j_1,\dots,j_v) \text{ iff } \begin{array}{l} i_1+\dots+i_v < j_1+\dots+j_v \text{ or} \\ i_1+\dots+i_v = j_1+\dots+j_v \text{ and} \\ (i_1,\dots,i_v) \prec_{lexico} (j_1,\dots,j_v). \end{array}$$

3. Application to Factorization Algorithms

Factorization algorithms for polynomial over an algebraic number field K have been devised by several authors. Theorem 3 enters when estimates for occurring rational numerators and denomiators are sought. The following representation for the coefficient domain K is usually adopted. First we choose $\alpha \in \mathcal{O}_K$ such that $Q(\alpha) = K$ by virtue of its minimal polynomial $\mu(\alpha) \in Z[\alpha]$. The polynomial f to be factored then can be transformed by multiplication with a rational integer to an element

$$f(x_1,\dots,x_v) \in (Z[\alpha]/(\mu))[x_1,\dots,x_v] \quad .$$

However, unlike in the integral case, the factorization of $f = g_1 \cdots g_t$, $g_i \in \mathcal{O}_K[x_1,\dots,x_v]$, may not have an associate factorization in $(Z[\alpha]/(\mu))[x_1,\dots,x_v]$. The reasons are twofold.

1) $Z[\alpha]$ can be a proper subset of $\mathcal{O}_K$. One can prove, however, that

$$Z[\alpha] \subseteq \mathcal{O}_K \subseteq \left(\tfrac{1}{d} Z\right)[\alpha] \subseteq \left(\tfrac{1}{D} Z\right)[\alpha]$$

where $D = \text{discriminant}(\mu) = \mp \text{resultant}(\mu, \mu')$ and $d^2 \mid D$. In fact, $D = d^2\Delta$ with Δ being the discriminant of K.

2) $f \in \mathcal{O}_K[x_1,\dots,x_v]$ may factor in $K[x_1,\dots,x_v]$ but not in $\mathcal{O}_K[x_1,\dots,x_v]$. One example, taken from [8], is $K = Q(\sqrt{-5})$, $\mathcal{O}_K = Z[\sqrt{-5}]$ $(-5 \equiv 3 \bmod 4)$, $f(x) = 2x^2 + 2x + 3 = \frac{1}{2}\left(2x + 1+\sqrt{-5}\right)(2x + 1-\sqrt{-5})$. Applying theorem 3 to $x^2 + x + \frac{3}{2}$ with $r = 2$ both monic linear factors must be elements in $\left(\frac{1}{2} Z[\sqrt{-5}]\right)[x]$ which

they are: $x + \frac{1+\sqrt{-5}}{2}$, $x + \frac{1-\sqrt{-5}}{2}$.

This example shows that, in the univariate case ($v=1$), monicity is one way of enforcing the weakly normalization assumption. But it is not the only possibility: E.g., we could have applied theorem 3 to $\frac{2}{3}x^2 + \frac{2}{3}x + 1 = \left(\frac{1+\sqrt{-5}}{3}x+1\right)\left(\frac{1-\sqrt{-5}}{3}x+1\right)$ choosing the constant coefficients 1 ($r = 3$).

In the multivariate case, one has even more choices. Let $a_{k_1,\ldots,k_v}(\alpha) \in Z[\alpha]$ be the coefficient of $x_1^{k_1} \ldots x_v^{k_v}$ in $f(x_1,\ldots,x_v)$ such that $(k_1,\ldots,k_v)$ is maximal with respect to some ordering $\prec$ satisfying($\dagger$). We now multiply f with $a_{k_1,\ldots,k_v}(\alpha)^{-1}$ mod $\mu(\alpha)$ and get $a_{k_1,\ldots,k_v}(\alpha)^{-1} f(x_1,\ldots,x_v) \in (1/r\ Z[\alpha])[x_1,\ldots,x_v]$ with $r \in Z$. Therefore, applying theorem 3 to a factorization of $f = g_1 \ldots g_t$ such that all non-zero mononials of maximum order w.r.t $\prec$ have coefficient 1, we get $g_i \in 1/r\ \mathcal{O}_K[x_1,\ldots,x_v]$.

It might not be apparent, at this point, why the minimization of r is of computational advantage. It mainly depends at which moment in the multivariate Hensel algorithm one switches from the mod p^k representation of numeric coeffi- back to rational ones. [4] suggests doing this before lifting the minor variables, whereas [6, 7, 8] at the very end for the recovery of the true factors. In the first case, comparing the denominators of the univariate factorization of $f(w_1,\ldots, w_{v-1},x_v)$, $w_i \in \mathcal{O}_K$, to rD (or rd, if d is known) might help discover some extraneous factors, but we believe this is not too helpful. Keeping the mod p^k representation of rationals to the very end, just before the trial diversions, seems a much better idea. One thus can keep p^k small by firstly working with the minimal r and secondly, one can recover the true denominators dr with $d^2 | D$ by computing a continued fraction approximation of the residues and p^k.

REFERENCES

[1] Dedekind, R.: Über einen arithmetischen Satz von Gauss. Mitteilungen der mathematischen Gesellschaft zu Prag, 1892.

[2] Fricke, R.: Lehrbuch der Algebra, Bd. 3. Braunschweig: Friedr. Vieweg & Sohn, 1928.

[3] Gauss, C. F.: Disquisitiones Arithmeticae. Leipzig, 1801.

[4] Kaltofen, E.: Polynomial-Time Reductions from Multivariate to Bi- and Univariate Integral Polynomial Factorization. SIAM J. on Comp., to appear.

[5] Landau, S.: Factoring Polynomials over Algebraic Number Fields is in Polynomial Time. Siam J. on Comp., to appear.

[6] Lenstra, A. K.: Factoring Multivariate Polynomials over Algebraic Number Fields. Manuscript, 1983.

[7] Wang, P.: Factoring Multivariate Polynomials over Algebraic Number Fields. Math. Comp. 30, 324-336 (1976).

[8] Weinberger, P. and Rothschild, L.: Factoring Polynomials over Algebraic Number Fields. ACM Trans Math. Software, 2, 335-350 (1976).