JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
S 0894-0347(04)00451-5
Article electronically published on January 15, 2004

IDEAL MEMBERSHIP IN POLYNOMIAL RINGS OVER THE INTEGERS

MATTHIAS ASCHENBRENNER

Introduction

The following well-known theorem, due to Grete Hermann [20], 1926, gives an upper bound on the complexity of the ideal membership problem for polynomial rings over fields:

Theorem. *Consider polynomials $f_0, \dots, f_n \in F[X] = F[X_1, \dots, X_N]$ of (total) degree $\leq d$ over a field F. If $f_0 \in (f_1, \dots, f_n)$, then*

$$f_0 = g_1 f_1 + \cdots + g_n f_n$$

for certain $g_1, \dots, g_n \in F[X]$ whose degrees are bounded by β, where $\beta = \beta(N, d)$ depends only on N and d (and not on the field F or the particular polynomials $f_0, \dots, f_n$).

This theorem was a first step in Hermann's project, initiated by work of Hentzelt and Noether [19], to construct bounds for some of the central operations of commutative algebra in polynomial rings over fields. A simplified and corrected proof was published by Seidenberg [35] in the 1970s, with an explicit but incorrect bound $\beta(N, d)$. In [31, p. 92], it was shown that one may take

$$\beta(N, d) = (2d)^{2^N}.$$

We will reproduce a proof, using Hermann's classical method, in Section 3 below. Note that the computable character of this bound reduces the question of whether $f_0 \in (f_1, \dots, f_n)$ for given $f_j \in F[X]$ to solving an (enormous) system of linear equations over F. Hence, in this way one obtains a (naive) algorithm for solving the ideal membership problem for $F[X]$ (provided F is given in some explicitly computable manner). Later, Buchberger in his Ph.D. thesis (1965) introduced the important concept of a *Gröbner basis* and gave an algorithm for deciding ideal membership for $F[X]$ which is widely used today (see, e.g., [6]).

The doubly exponential nature of β above is essentially unavoidable, as a family of examples due to Mayr and Meyer [27] shows. In fact, they prove that ideal membership for $\mathbb{Q}[X]$ is exponential-space hard: the amount of space needed by *any* algorithm to decide ideal membership for $\mathbb{Q}[X]$ (or $\mathbb{Z}[X]$) grows exponentially in the size of the input. If we restrict to $f_0, \dots, f_n$ of a special form, often dramatic

Received by the editors May 2, 2003.
2000 *Mathematics Subject Classification.* Primary 13P10; Secondary 11C08.
Key words and phrases. Ideal membership over the integers, bounds, restricted power series.
Partially supported by the Mathematical Sciences Research Institute.

improvements are possible: for example, if $f_0 = 1$ (the situation of Hilbert's Nullstellensatz), then in the theorem we may replace the doubly exponential $(2d)^{2^N}$ by the single exponential bound d^N if $d > 2$ (due to Kollár [22]) and by 2^{N+1} if $d = 2$ (due to Sombra [38]). A number of results show the existence of single-exponential bounds in the (general) ideal membership problem for $F[X]$, under suitable geometric assumptions on the ideal $I = (f_1, \dots, f_n)$: for example if I is zero-dimensional or a complete intersection [7]. Membership in an unmixed ideal I can be decided in single-exponential time [10].

In this paper, we study the ideal membership problem over coefficient rings of an arithmetic nature, like the ring of integers $\mathbb{Z}$ (instead of over a field F). The following example shows that contrary to what happens over fields, if a bound d on the degree of $f_0, f_1, \dots, f_n \in \mathbb{Z}[X]$ is given and $f_0 \in (f_1, \dots, f_n)\mathbb{Z}[X]$, then *there is no uniform bound on the degrees of g_j's such that $f_0 = g_1 f_1 + \cdots + g_n f_n$, which depends only on N and d.* So any bound on the degree of the $g_1, \dots, g_n$ as a function of $f_0, f_1, \dots, f_n$ will necessarily also have to depend on the coefficients of the polynomials f_j.

Example. Let $p > 1$ and $d \geq 1$ be integers. We have $1 \in (1 - pX, p^d X)\mathbb{Z}[X]$, since

$$1 = \big(1 + pX + \cdots + p^{d-1}X^{d-1}\big)(1 - pX) + X^{d-1}p^d X,$$

with the degrees of $1+pX+\cdots+p^{d-1}X^{d-1}$ and X^{d-1} tending to infinity, as $d \to \infty$. Considering everything mod p^d, we see that $1 - pX$ is a unit in $(\mathbb{Z}/p^d\mathbb{Z})[X]$; indeed

$$1 \equiv \big(1 + pX + \cdots + p^{d-1}X^{d-1}\big)(1 - pX) \mod p^d.$$

Hence if $1 \equiv g(X)(1 - pX) \bmod p^d$ for some $g(X) \in \mathbb{Z}[X]$, then necessarily $g \equiv 1 + pX + \cdots + p^{d-1}X^{d-1} \bmod p^d$. It follows that if

$$1 = g(X)(1 - pX) + h(X)p^d X \qquad \text{with } g, h \in \mathbb{Z}[X],$$

then $\deg g, \deg h \geq d-1$. Taking for p a prime number and replacing $\mathbb{Z}$ by its localization $\mathbb{Z}_{(p)}$, the same example works if we consider polynomials with coefficients in $\mathbb{Z}_{(p)}$.

A decision procedure for the ideal membership problem for polynomial rings over $\mathbb{Z}$ has been known at least since the early 1970s; see, e.g., [4], [5], [12], [21], [32], [36], [37]. However, these results did not yield the existence of a *primitive* recursive algorithm, for any fixed $N \geq 3$, let alone the existence of bounds similar to the ones in Hermann's theorem for polynomial rings over fields. Indeed, it was suspected by some that this was one of the rare cases where a natural decision problem allows an algorithmic solution, but not a primitive recursive one. (See [2] for a survey of the history and the various proposals for computing in $\mathbb{Z}[X]$.)

Finding a decision procedure for ideal membership in $\mathbb{Z}[X]$ was central to Kronecker's ideology of constructive mathematics [11]. In fact, one may argue that he was primarily interested in what we would call today a *primitive recursive* algorithm. Thus, the task of finding a primitive recursive decision method for ideal membership in $\mathbb{Z}[X]$ has aptly been called "Kronecker's problem" in [13]. In this paper, Gallo and Mishra adapted Buchberger's algorithm for the construction of Gröbner bases and deduced a primitive recursive procedure to decide the ideal membership problem for $\mathbb{Z}[X]$, when the number of variables N is fixed. Analyzing their algorithm, they obtained the following bounds:

Theorem. *Let $f_0, \dots, f_n \in \mathbb{Z}[X] = \mathbb{Z}[X_1, \dots, X_N]$. If $f_0 \in (f_1, \dots, f_n)$, then*

$$f_0 = g_1 f_1 + \cdots + g_n f_n$$

for certain polynomials $g_1, \dots, g_n \in \mathbb{Z}[X]$ whose size $|g_j|$ is bounded by

$$W_{4N+8}\big(|f_0| + \cdots + |f_n| + N\big).$$

Here the size $|f|$ of a polynomial $f \in \mathbb{Z}[X]$ is a crude measure of its complexity and equals the maximum of the absolute values of the coefficients and the degrees of f with respect to each indeterminate; see [13, p. 346]. The function W_k is the kth function in the so-called Wainer hierarchy of primitive recursive functions; see [39]. Even for small k, these functions are already very rapidly growing: We have $W_0(n) = n + 1$, $W_1(n) = 2n + 1$, but W_2 grows asymptotically like the exponential $n \mapsto 2^n$, W_3 like the n-times iterated exponential function, and so on. These bounds are only primitive recursive for each fixed N; the growth rate of this bound as a function of N is similar to the notorious Ackermann function.

Gallo and Mishra's analysis of the complexity of their algorithm ultimately rests on an effective version of Hilbert's Basis Theorem for increasing chains of monomial ideals in $\mathbb{Z}[X]$. This approach is doomed to fail in providing bounds which are also primitive recursive for varying N: In general, even the length of an increasing chain of ideals in $\mathbb{Z}[X]$ with the kth ideal in the chain generated by monomials of degree at most kd can have a growth behavior similar to Ackermann's function, as a function of N and d. (See [28].)

In the present paper, we will give a proof of the following theorem. Given a polynomial $f \in \mathbb{Z}[X]$, we let $h(f)$ be the height of f, that is, the maximum of $\log|a|$ where a ranges over the non-zero coefficients of f, with $h(0) := 0$.

Theorem A. *If $f_0, f_1, \dots, f_n \in \mathbb{Z}[X] = \mathbb{Z}[X_1, \dots, X_N]$ are polynomials with $f_0 \in (f_1, \dots, f_n)$, whose degrees are at most d and whose heights are at most h, then*

$$f_0 = g_1 f_1 + \cdots + g_n f_n$$

for certain polynomials $g_1, \dots, g_n \in \mathbb{Z}[X]$ of degrees at most

$$\gamma(N, d, h) = (2d)^{2^{O(N \log(N+1))}}(h + 1).$$

In principle, the (universal) constant hidden in the O-notation can be made explicit; see Section 6 below. The bound γ on the degrees of the g_j's implies the existence of a similar (doubly exponential) bound on the heights of the g_j. As a consequence, we obtain a naive elementary recursive decision procedure for ideal membership in $\mathbb{Z}[X]$. In this paper we prove in fact a generalization of Theorem A with $\mathbb{Z}$ replaced by the ring of integers of a number field F, using an appropriate notion of height for elements of F.

The starting point for our proof of Theorem A is the simple observation that one can *localize* the question of whether $f_0 \in (f_1, \dots, f_n)$ and reduce it to finitely many subproblems in the following way: Using the classical method of Hermann (for $F = \mathbb{Q}$), one can test whether $f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X]$, and assuming this is so, we obtain, by clearing denominators, a representation

$$\delta f_0 = g_1 f_1 + \cdots + g_n f_n \qquad \text{with } \delta \in \mathbb{Z},\ \delta \neq 0,\ g_1, \dots, g_n \in \mathbb{Z}[X]. \tag{1}$$

Let $p_1, \dots, p_K$ be the different prime factors of δ. Then another necessary condition for $f_0 \in (f_1, \dots, f_n)$, besides $f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X]$, is that $f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p_k)}[X]$

for $k = 1, \dots, K$. Together with (1), these *necessary* conditions are also *sufficient* for $f_0 \in (f_1, \dots, f_n)$: If $f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p_k)}[X]$, then

$$\delta_k f_0 = g_{1k} f_1 + \cdots + g_{nk} f_n \qquad \text{for some } \delta_k \in \mathbb{Z} \setminus p_k\mathbb{Z} \text{ and } g_{jk} \in \mathbb{Z}[X]. \tag{2}$$

Since $\delta, \delta_1, \dots, \delta_K$ have no common prime factor, we find, by the Euclidean Algorithm, a linear combination of them that equals 1:

$$a\delta + a_1\delta_1 + \cdots + a_K\delta_K = 1 \qquad (a, a_1, \dots, a_K \in \mathbb{Z}). \tag{3}$$

Combining (1), (2) and (3), we get

$$f_0 = \big(a\delta + a_1\delta_1 + \cdots + a_K\delta_K\big) f_0 = \sum_{j=1}^{n} \big(ag_j + a_1 g_{j1} + \cdots + a_K g_{jK}\big) f_j,$$

which exhibits f_0 as an element of $(f_1, \dots, f_n)$.

Now note that given a prime p, we have $f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p)}[X]$ if and only if the homogeneous linear equation

$$f_1 y_1 + \cdots + f_n y_n - f_0 y_{n+1} = 0 \tag{4}$$

in the unknowns $y_1, \dots, y_{n+1}$ has a solution $(y_1, \dots, y_{n+1}) \in \big(\mathbb{Z}_{(p)}[X]\big)^{n+1}$ with $y_{n+1} = 1$. This reduces the problem of deciding whether $f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p)}[X]$ to the following two subproblems:

(a) constructing a collection of generators $z^{(1)}, \dots, z^{(L)} \in \big(\mathbb{Z}_{(p)}[X]\big)^{n+1}$ for the module of solutions (in $\mathbb{Z}_{(p)}[X]$) to equation (4), and
(b) deciding whether the ideal in $\mathbb{Z}_{(p)}[X]$ generated by the last components of the vectors $z^{(1)}, \dots, z^{(L)}$ contains 1.

Problem (b) can be easily treated by applying the effective Nullstellensatz for $\mathbb{Q}[X]$ and $\mathbb{F}_p[X]$ (or Hermann's Theorem). By a faithful flatness argument, it is possible to further reduce problem (a) to the construction of a set of generators for the $\mathbb{Q}[X]$-module of solutions to (4) in $\mathbb{Q}[X]$ and a set of generators $S \subseteq \big(\mathbb{Z}_{(p)}[X]\big)^{n+1}$ for the $\mathbb{Z}_p\langle X\rangle$-module of solutions to (4) in $\mathbb{Z}_p\langle X\rangle$. Here, $\mathbb{Z}_p\langle X\rangle$ denotes the ring of restricted power series with p-adic integer coefficients (see [8] or Section 2). The great advantage of the power series rings $\mathbb{Z}_p\langle X\rangle$ over polynomial rings over $\mathbb{Z}$ (or over $\mathbb{Z}_{(p)}$) is that they satisfy a Weierstraß Division and Preparation Theorem. Hermann's method for deciding ideal membership, that is, deciding solvability of a single inhomogeneous linear equation, has a variant which allows for the construction of a finite set of generators for the $\mathbb{Q}[X]$-module of solutions to the linear homogeneous equation (4) in $\mathbb{Q}[X]$. The key step in our argument is to adapt this method to explicitly construct the set S from above, that is, to show the *effective flatness* of $\mathbb{Z}_p\langle X\rangle$ as a $\mathbb{Z}_{(p)}[X]$-module. All computations take place in $\mathbb{Z}_{(p)}[X]$, and bounds for the heights of the polynomials occurring in each step can be found. This enables us to calculate the bound γ.

Theorem A naturally generalizes to *systems of linear equations* over polynomial rings, and as the sketch above already indicates, one also obtains information on homogeneous systems of linear equations. For example, the methods developed here lead to the following theorem on degree bounds for generators of syzygies:

Theorem B. *The $\mathbb{Z}[X]$-module of solutions $(y_1, \dots, y_n) \in \big(\mathbb{Z}[X]\big)^n$ of the equation*

$$f_1 y_1 + \cdots + f_n y_n = 0,$$

where $f_1, \dots, f_n \in \mathbb{Z}[X] = \mathbb{Z}[X_1, \dots, X_N]$ are of degree $\leq d$, is generated by solutions

$$y^{(1)}, \dots, y^{(K)} \in \big(\mathbb{Z}[X]\big)^n$$

whose entries are of degree at most $(2d)^{2^{O(N \log(N+1))}}$.

Note that this bound does not depend on the coefficients of the f_j's. (The number K of required generators also depends only on N, n and d.) Theorem B holds in a rather more general context, for any almost hereditary ring in place of $\mathbb{Z}$. See Section 4 below for the definition of almost hereditary rings and [1] for uniform degree bounds on syzygies for an even larger class of rings. The size of the coefficients of the entries of $y^{(k)}$ can be similarly estimated, by a bound that also depends on the heights of $f_1, \dots, f_n$.

Organization of the paper. We begin (in Section 1) by recalling basic definitions about absolute values on number fields, defining a height function on the algebraic closure of $\mathbb{Q}$, and establishing some auxiliary facts about it used later. In Section 2 we state some fundamental facts about the ring of restricted power series over a complete discrete valuation ring. Section 3 contains an exposition of Hermann's method for solving systems of linear equations in polynomial rings over fields. This is the basis for Section 4, where we give a proof of Theorem B modeled on this method. We also indicate two applications concerning bounds for some operations on finitely generated modules and a criterion for primeness of ideals in $\mathbb{Z}[X]$ in the style of [34]. In Section 5 we complement Theorem B for rings of integers in number fields by establishing bounds on the height of generators for syzygy modules. In Section 6 we use these results to prove Theorem A.

Notation and conventions. Throughout this paper $\mathbb{N} = \{0, 1, 2, \dots\}$ denotes the set of natural numbers.

Let R be a ring (here and below: always commutative with a unit element). The localization $S^{-1}R$, where S denotes the set of non-zero-divisors of R, is called the ring of fractions of R, denoted by $\mathrm{Frac}(R)$. If A is an $m \times n$-matrix with entries in R, the set of solutions in R^n to the homogeneous system of linear equations $Ay = 0$ is an R-submodule of R^n, which we denote by $\mathrm{Sol}_R(A)$. It is sometimes called the (first) **module of syzygies of** A**.** If R is coherent (e.g., if R is Noetherian), then $\mathrm{Sol}_R(A)$ is finitely generated. For submodules M, M' of an R-module L we write

$$(M' : M) := \big\{a \in R : am \in M' \text{ for all } m \in M\big\},$$

which is an ideal of R (containing the annihilator of M). If $\delta \in R$, then

$$(M : \delta) := \big\{m \in L : \delta m \in M\big\},$$

which is a submodule of L. If $L = R^m$ is a finitely generated free module and $R \to R'$ a ring homomorphism, then MR' denotes the R'-submodule of $(R')^m$ generated by the image of M.

By $X = (X_1, \dots, X_N)$ we always denote a tuple of N distinct indeterminates, where $N \in \mathbb{N}$. The (total) degree of a polynomial $0 \neq f \in R[X] = R[X_1, \dots, X_N]$ is denoted by $\deg(f)$, and the degree of f in X_i (where $i \in \{1, \dots, N\}$) is denoted by $\deg_{X_i}(f)$. By convention $\deg(0) := -\infty$ and $\deg_{X_i}(0) := -\infty$, where $-\infty < \mathbb{N}$. We extend this notation to finite tuples $f = (f_1, \dots, f_n)$ of polynomials in $R[X]$ by setting $\deg(f) := \max_j \deg(f_j)$ (the degree of f). Similarly we define $\deg_{X_i}(f)$.

The notions of *computable field* and *computable ring* are used in an informal way. We will say that a computable ring R is **syzygy-solvable** if there is an algorithm which, given $a_1, \dots, a_n \in R$, constructs a finite set of generators for the solutions to the homogeneous linear equation $a_1 y_1 + \cdots + a_n y_n = 0$. (This is called "finitely related" in [32].) For example, the prime fields $\mathbb{Q}$ and $\mathbb{F}_p$ are clearly syzygy-solvable, as is $\mathbb{Z}$, or more generally the ring of integers of any number field (see [9]).

1. Absolute values and height functions

We assume that the reader is familiar with the basic theory of absolute values on number fields as expounded in, say, [25, Chapter II], and the (absolute, logarithmic) height function on the algebraic closure of $\mathbb{Q}$ as used in diophantine geometry (see [24, Chapter 3]). We recall some definitions and a few basic facts used later.

Absolute values. We let $|\cdot|$ denote the usual (Euclidean) absolute value $|\cdot|$ on $\mathbb{Q}$, and for a prime number p we let $|\cdot|_p$ denote the p-adic absolute value on $\mathbb{Q}$: $|a|_v = p^{-v_p(a)}$ for $a \in \mathbb{Q}^\times$, where $v_p\colon \mathbb{Q}^\times \to \mathbb{Z}$ denotes the p-adic valuation on $\mathbb{Q}$. Let F be an algebraic number field of degree $d = [F : \mathbb{Q}]$, and let $R = \mathcal{O}_F$ be the ring of integers of F. If v is a finite place of F which lies over the prime number p, we write $v|p$. If v is an infinite place of F, we write $v|\infty$. To every place v of F we associate an absolute value $|\cdot|_v$ on F, normalized so that

(1) if $v|p$ for a prime p, then $|\cdot|_v$ extends the p-adic absolute value $|\cdot|_p$ on $\mathbb{Q}$,
(2) if $v|\infty$, then $|\cdot|_v$ extends the usual absolute value $|\cdot|$ on $\mathbb{Q}$.

For every place v of F we let F_v denote the completion of F with respect to the topology induced by $|\cdot|_v$ and $\mathbb{Q}_v \subseteq F_v$ the completion of $\mathbb{Q}$ with respect to the topology induced by the restriction of $|\cdot|_v$ to $\mathbb{Q}$. We put $d_v = [F_v : \mathbb{Q}_v]$. If $v|p$ is finite, then $\mathbb{Q}_p$ is the field of p-adic numbers. If $v|\infty$, then $\mathbb{Q}_v = \mathbb{R}$, and either $F_v = \mathbb{R}$ and $d_v = 1$ (in which case v is called real) or $F_v = \mathbb{C}$ and $d_v = 2$ (v is complex). Given $w = \infty$ or $w = p$ for a prime p, we have

$$d = \sum_{v|w} d_v,$$

where the sum ranges over all places v of F with $v|w$. We let M_F denote the set of all places of F, $M_F^\infty := \{v \in M_F : v|\infty\}$ the set of infinite places, and $M_F^0 := M_F \setminus M_F^\infty$ the set of finite places of F. We put $||a||_v := |a|_v^{d_v}$ for $v \in M_F$. With this normalization, the number field F satisfies the following **product formula**:

$$\prod_{v \in M_F} ||a||_v = 1 \qquad (a \in F^\times). \tag{1.1}$$

The assignment $v \mapsto \mathfrak{p}_v := \big\{r \in R : |r|_v < 1\big\}$ establishes a one-to-one correspondence between M_F^0 and the set of non-zero prime ideals of R. If $v|p$ is a finite place of F and $\mathfrak{p} = \mathfrak{p}_v$, then the absolute value $|\cdot|_v$ on F associated to v and the $\mathfrak{p}$-adic valuation on F are connected as follows:

$$|a|_v = p^{-v_{\mathfrak{p}}(a)/e_v} \qquad \text{for all } a \in F^\times.$$

Here e_v denotes the ramification index of v, that is, the unique integer such that $p = \pi^{e_v} u$ for some unit u of $R_{\mathfrak{p}}$ and some $\pi \in R_{\mathfrak{p}}$ with $v_{\mathfrak{p}}(\pi) = 1$. We have $e_v | d_v$; in fact, $\#(R/\mathfrak{p}) = p^{d_v/e_v}$.

Divisors. An M_F**-divisor** is a function $\mathfrak{c}\colon M_F \to \mathbb{R}$ such that

(1) $\mathfrak{c}(v) > 0$ for all $v \in M_F$;
(2) $\mathfrak{c}(v) = 1$ for all but finitely many $v \in M_F$;
(3) for each $v \in M_F^0$ there exists an element $a \in F$ with $\mathfrak{c}(v) = |a|_v$.

We shall sometimes write $|\mathfrak{c}|_v$ instead of $\mathfrak{c}(v)$, and we put $||\mathfrak{c}||_v := |\mathfrak{c}|_v^{d_v}$. We define the **size** of an M_F-divisor $\mathfrak{c}$ to be

$$||\mathfrak{c}||_F := \prod_v ||\mathfrak{c}||_v.$$

The product $\mathfrak{c} \cdot \mathfrak{d}$ of two M_F-divisors $\mathfrak{c}$ and $\mathfrak{d}$ is an M_F-divisor, and $||\mathfrak{c} \cdot \mathfrak{d}||_F = ||\mathfrak{c}||_F \cdot ||\mathfrak{d}||_F$. Given an M_F-divisor $\mathfrak{c}$, we let

$$L(\mathfrak{c}) := \big\{a \in F : |a|_v \le \mathfrak{c}(v) \text{ for all } v \in M_F\big\},$$

a finite set. Each non-zero fractional ideal I of F (i.e., a finitely generated R-submodule of F) determines a unique M_F-divisor $\mathfrak{c}_I$ such that

$$L(\mathfrak{c}_I) = \big\{a \in I : |a|_v \le 1 \text{ for all } v \in M_F^\infty\big\}.$$

We have $\mathfrak{c}(v) = 1$ for all $v \in M_F^\infty$ and

$$\mathfrak{c}_I(v_\mathfrak{p}) = p^{-v_\mathfrak{p}(I)/e_{v_\mathfrak{p}}}$$

for all primes $\mathfrak{p} \neq 0$ of R. Here $I = \prod_\mathfrak{p} \mathfrak{p}^{v_\mathfrak{p}(I)}$ is the unique representation of I as a product of non-zero prime ideals of R, with

$$v_\mathfrak{p}(I) = \min\big\{v_\mathfrak{p}(a) : a \in I\big\} \in \mathbb{Z}$$

and $v_\mathfrak{p}(I) = 0$ for almost all $\mathfrak{p}$. We have $||\mathfrak{c}_I||_F = 1/N(I)$, where

$$N(I) = \prod_\mathfrak{p} \#(R/\mathfrak{p})^{v_\mathfrak{p}(I)}$$

denotes the norm of I. If $I \subseteq R$, then $N(I) = \#(R/I)$.

Local heights. Given a place $v \in M_F$ and a non-empty finite set $S \subseteq F$, we put

$$|S|_v := \max\big\{|a|_v : a \in S\big\}, \qquad ||S||_v := |S|_v^{d_v}.$$

We define the (logarithmic) *local height* $h_v(S)$ of S at v by

$$h_v(S) := \log^+ ||S||_v.$$

Here $\log^+ r := \max\big\{0, \log r\big\}$ for $r \in \mathbb{R}^{>0}$ and $\log^+ 0 := 0$. We declare $h_v(\varnothing) := 0$. For a polynomial $f \in F[X]$ we put $||f||_v := ||S||_v$, where S is the set of coefficients of f. The *local height* of f at $v \in M_F$ is defined by

$$h_v(f) := \log^+ ||f||_v.$$

More generally, for $f_1, \dots, f_n \in F[X]$ we put

$$h_v(f_1, \dots, f_n) := \log^+ ||S||_v,$$

where S is the set of coefficients of $f_1, \dots, f_n$. Note $h_v(f_1, \dots, f_n) \ge 0$. Here are some other basic properties of h_v, immediate from the definition:

Lemma 1.1. *Let $v \in M_F$ and $a, a_1, \dots, a_n \in F$. Then*

(1) $h_v(a) = h_v(-a)$;
(2) $h_v(a^k) = k \cdot h_v(a)$ *for* $k \in \mathbb{N}$;
(3) $h_v(a_1 + \cdots + a_n) \le h_v(a_1, \dots, a_n) + \log n$ *if* $v \in M_F^\infty$;
(4) $h_v(a_1 + \cdots + a_n) \le h_v(a_1, \dots, a_n)$ *if* $v \in M_F^0$;

(5) $h_v(a_1 \cdots a_n) \leq h_v(a_1) + \cdots + h_v(a_n)$.

From (1) and (3)–(5) in the lemma we obtain:

Corollary 1.2. *If $A^{(1)}, \dots, A^{(m)}$ are $n \times n$-matrices with entries in F, then*

(1) $h_v\big(\det A^{(1)}, \dots, \det A^{(m)}\big) \leq n \cdot \big(h_v\big(A^{(1)}, \dots, A^{(m)}\big) + \log n\big)$ *if $v \in M_F^\infty$;*
(2) $h_v\big(\det A^{(1)}, \dots, \det A^{(m)}\big) \leq n \cdot h_v\big(A^{(1)}, \dots, A^{(m)}\big)$ *if $v \in M_F^0$.*

(Using Hadamard's inequality, it is possible to improve the term $\log n$ in (1) slightly, to $\frac{1}{2}\log n$.)

Global height. The (global) *height* of a finite set $S \subseteq F$ is defined in terms of the local heights:

$$h(S) := \frac{1}{d} \sum_{v \in M_F} h_v(S).$$

The (global) *height* of $f_1, \dots, f_n \in F[X]$ is the global height of its set of coefficients. The quantity $h(S)$ does not change if the field F is replaced by another algebraic number field containing the set S. Hence h gives rise to a height function, also denoted by h, on (finite subsets of) the algebraic closure of $\mathbb{Q}$. The product formula (1.1) implies that $h(a) = h(1/a)$ for all $a \in F^\times$.

We have $h(S') \leq h(S)$ for all subsets $S' \subseteq S$; hence $h(a) \leq h(S)$ for $a \in S$. Suppose that $S \neq \{0\}$, and let I denote the fractional ideal generated by S. Then

$$h(S) = \frac{1}{d}\left(\log N(\mathfrak{d}) + \sum_{v \in M_F^\infty} h_v(S)\right),$$

where $I = \mathfrak{b}/\mathfrak{d}$ is a factorization of I with $\mathfrak{b}, \mathfrak{d}$ relatively prime ideals of R. In particular, for $0 \neq a \in R$ we get

$$h(a) = h(1/a) = \frac{1}{d}\left(\log N(a) + \sum_{v \in M_F^\infty} h_v(1/a)\right). \tag{1.2}$$

Moreover, if $S \subseteq R$, then

$$h(S) = \frac{1}{d} \sum_{v \in M_F^\infty} h_v(S).$$

It follows that in this case $h(S) \leq d \cdot \max\big\{h(a) : a \in S\big\}$.

Example. For non-zero and relatively prime integers $r, s \in \mathbb{Z}$ we have $h(r/s) = \max\big\{\log |r|, \log |s|\big\}$. In particular $h(r) = \log |r|$ for $0 \neq r \in \mathbb{Z}$.

From Lemma 1.1 we get the following rules for estimating the behavior of h with respect to the elementary operations of F:

$$h(a) = h(-a), \tag{1.3}$$

$$h(a^k) = |k|h(a) \quad \text{for all } k \in \mathbb{Z}, \tag{1.4}$$

$$h(a_1 + \cdots + a_n) \leq h(a_1, \dots, a_n) + \log n, \tag{1.5}$$

$$h(a_1 \cdots a_n) \leq h(a_1) + \cdots + h(a_n). \tag{1.6}$$

From Corollary 1.2 we obtain the following bound on the height of determinants of $n \times n$-matrices $A^{(1)}, \dots, A^{(m)} \in F^{n \times n}$:

$$h\big(\det A^{(1)}, \dots, \det A^{(m)}\big) \leq n \cdot \big(h\big(A^{(1)}, \dots, A^{(m)}\big) + \log n\big). \tag{1.7}$$

The following facts will also be used later on:

Lemma 1.3. *For all $a \in F^\times$,*

$$\sum_{v \in M_F^0, v_\mathfrak{p}(a)>0} \log p \cdot v_\mathfrak{p}(a) \le d \cdot h(a).$$

Here the sum runs over all $v \in M_F^0$ such that $v_\mathfrak{p}(a) > 0$, with $\mathfrak{p} = \mathfrak{p}_v$ denoting the prime ideal of R corresponding to v and p the unique prime number such that $v|p$.

Proof. We have (using (1.4))

$$d \cdot h(a) = d \cdot h(1/a) \ge \sum_{v \in M_F^0} d_v/e_v \cdot \log p \cdot \max\{0, v_\mathfrak{p}(a)\} \ge \sum_{v \in M_F^0, v_\mathfrak{p}(a)>0} \log p \cdot v_\mathfrak{p}(a)$$

as claimed. □

It follows that given a non-zero element a of R there are at most $d \cdot h(a)/\log 2$ many absolute values $v \in M_F^0$ such that $v_\mathfrak{p}(a) > 0$, where $\mathfrak{p} = \mathfrak{p}_v$. Moreover $|v_\mathfrak{p}(a)| \le d \cdot h(a)/\log p$ for all $v \in M_F^0$, where $v|p$.

Lemma 1.4. *There exists a constant C_0, depending only on F, with the following property: Given ideals I and J of R with I properly contained in J, there exists $a \in J \setminus I$ of height at most $C_0 + \left(1 + \frac{1}{d}\right) \log N(I)$.*

Proof. By [25, Chapter V, §2, Theorem 1], we have, for every M_F-divisor $\mathfrak{c}$:

$$\#L(\mathfrak{c}) = B_F ||\mathfrak{c}||_F + O\big(||\mathfrak{c}||_F^{1-1/d}\big) \qquad \text{as } ||\mathfrak{c}||_F \to \infty,$$

where $B_F = \frac{2^{d_1}(2\pi)^{d_2}}{|D(F)|^{1/2}}$. Here d_1 and d_2 denote the number of real and complex places of F, respectively, and $D(F)$ denotes the discriminant of F. That is, there exists a positive real number C (only depending on F) such that for every M_F-divisor $\mathfrak{c}$ with $||\mathfrak{c}||_F \ge C$

$$\big|\#L(\mathfrak{c}) - B_F ||\mathfrak{c}||_F\big| \le C \cdot ||\mathfrak{c}||_F^{1-1/d}. \tag{1.8}$$

We may assume that $C \ge \left(\frac{B_F}{2}\right)^{\frac{d}{d-1}}$; hence $(C')^d \ge C$ for $C' := \frac{2C}{B_F}$. Let $t := C' \cdot N(I)^{1+\frac{1}{d}}$ and let $\mathfrak{d}$ be the M_F-divisor given by $\mathfrak{d}(v) = 1$ if $v \in M_F^0$ and $\mathfrak{d}(v) = t$ if $v \in M_F^\infty$. We consider the M_F-divisors $\mathfrak{c} = \mathfrak{c}_I \cdot \mathfrak{d}$ and $\mathfrak{c}' = \mathfrak{c}_J \cdot \mathfrak{d}$ of size $||\mathfrak{c}||_F = t^d/N(I)$ and $||\mathfrak{c}'||_F = t^d/N(J)$, respectively. We have $||\mathfrak{c}||_F \ge C$; hence (1.8) implies

$$\big|\#L(\mathfrak{c}) - B_F ||\mathfrak{c}||_F\big| \le C \cdot t^{d-1} N(I)^{1/d-1},$$

and similarly with $\mathfrak{c}$ replaced by $\mathfrak{c}'$. Since I is properly contained in J, we have $N(I) > N(J)$ and $L(\mathfrak{c}) \subseteq L(\mathfrak{c}')$. If $L(\mathfrak{c}) = L(\mathfrak{c}')$, then

$$B_F t^d \left(\frac{1}{N(J)} - \frac{1}{N(I)}\right) = \big|B_F ||\mathfrak{c}||_F - B_F ||\mathfrak{c}'||_F\big| \le 2C \cdot t^{d-1} N(I)^{1/d-1}$$

and hence

$$t \le C' N(I)^{1/d-1} \left(\frac{N(I)N(J)}{N(I) - N(J)}\right) < C' \cdot N(I)^{1+\frac{1}{d}},$$

a contradiction. Therefore $L(\mathfrak{c}') \setminus L(\mathfrak{c}) \ne \varnothing$, that is, there exists $a \in J \setminus I$ with $|a|_v \le t$ for all $v \in M_F^\infty$. It follows that

$$h(a) = \frac{1}{d} \sum_{v \in M_F^\infty} d_v \log^+ |a|_v \le \log^+ t \le C_0 + \left(1 + \frac{1}{d}\right) \log N(I)$$

where $C_0 = \log C'$ is a constant only depending on the number field F. □

Given any fractional ideal $I \neq \{0\}$ of F and a non-zero prime ideal $\mathfrak{p}$ of R, there exists an element b of

$$I^{-1} = \{a \in F : aI \subseteq R\}$$

such that $v_{\mathfrak{p}}(b) = -v_{\mathfrak{p}}(I)$. In Section 4 we will need the existence of such b having small height, for integral I:

Corollary 1.5. *There exists a constant C_1, depending only on F, with the following property: Given an ideal $I = (a_1, \dots, a_n)$ of R and a non-zero prime ideal $\mathfrak{p}$ of R with $v_{\mathfrak{p}}(I) > 0$, there exists an element b of I^{-1} such that $v_{\mathfrak{p}}(b) = -v_{\mathfrak{p}}(I)$ and*

$$h(b) \leq C_1\big(h(a_1, \dots, a_n) + 1\big).$$

Proof. Let $C_0 > 0$ be the constant from Lemma 1.4, and put $C_1 := C_0 + d + 2$. Let $i \in \{1, \dots, n\}$ be such that $v_{\mathfrak{p}}(a_i) = v_{\mathfrak{p}}(I)$, and put $J := (a_i) \cdot I^{-1}$, an ideal of R. By Lemma 1.4 there exists $a \in J \setminus J \cdot \mathfrak{p}$ with $h(a) \leq C_0 + \big(\frac{d+1}{d}\big) \log N(J \cdot \mathfrak{p})$, and by (1.2)

$$\frac{1}{d} \log N(J \cdot \mathfrak{p}) \leq \frac{1}{d} \log N(a_i) \leq h(a_i).$$

Hence the element $b = a/a_i \in I^{-1}$ satisfies

$$h(b) \leq h(a) + h(a_i) \leq C_0 + \left(1 + \frac{1}{d}\right) \log N(J \cdot \mathfrak{p}) + h(a_i) \leq C_0 + (d+2)h(a_i)$$

and has the required properties. □

Remarks.

(1) For $F = \mathbb{Q}$ the constant $C_1 = 1$ has the property claimed in the corollary: Given integers $a_1, \dots, a_n \in \mathbb{Z}$ and a prime number p, let $b = p^{-\mu}$, where $\mu = \min_i v_p(a_i)$. Then $ba_1, \dots, ba_n \in \mathbb{Z}$ and $h(b) = \log p^{\mu} \leq h(a_1, \dots, a_n)$.

(2) If the number field F is explicitly given, say in terms of its multiplication table for a $\mathbb{Z}$-basis $\omega_1, \dots, \omega_d$ of R, and the generators $a_1, \dots, a_n$ are also explicitly given (in terms of their coefficients in the basis $\omega_1, \dots, \omega_d$), then $b \in I^{-1}$ with $v_{\mathfrak{p}}(b) = -v_{\mathfrak{p}}(I)$ can be found effectively: By [9, pp. 202–205] we can compute a basis $b_1, \dots, b_m$ for the R-module I^{-1}; then $b = b_i$, where $v_{\mathfrak{p}}(b_i)$ is minimal, has the required property. (Perhaps, using [33], one could also obtain a more explicit constant C_1 in this way.)

2. Rings of restricted power series

Let $\mathcal{O}$ be a discrete valuation ring (DVR) with maximal ideal $\mathfrak{m} = t\mathcal{O}$. We write $v_{\mathfrak{m}} \colon \mathcal{O} \setminus \{0\} \to \mathbb{N}$ for the $\mathfrak{m}$-adic valuation associated to $\mathcal{O}$ (normalized so that $v_{\mathfrak{m}}(t) = 1$). We always consider $v_{\mathfrak{m}}$ extended to a map $v_{\mathfrak{m}} \colon \mathcal{O} \to \mathbb{N}_\infty$ by $v_{\mathfrak{m}}(0) := \infty$, where $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ with the usual conventions $\mathbb{N} < \infty$ and $r + \infty = \infty + r = \infty$ for all $r \in \mathbb{N}_\infty$. The residue field of $\mathcal{O}$ is denoted by $\overline{\mathcal{O}} = \mathcal{O}/\mathfrak{m}$, with residue homomorphism $a \mapsto \overline{a} \colon \mathcal{O} \to \overline{\mathcal{O}}$.

From now until further notice we assume that $\mathcal{O}$ is complete in the $\mathfrak{m}$-adic topology on $\mathcal{O}$. The completion of the polynomial ring $\mathcal{O}[X] = \mathcal{O}[X_1, \dots, X_N]$ with respect to the $\mathfrak{m}\mathcal{O}[X]$-adic topology on $\mathcal{O}[X]$ will be denoted by $\mathcal{O}\langle X\rangle = \mathcal{O}\langle X_1, \dots, X_N\rangle$. It may be regarded as a subring of the ring $\mathcal{O}[[X]]$ of formal power

series over $\mathcal{O}$, and it is called the ring of **restricted power series** with coefficients in $\mathcal{O}$. Its elements are the power series

$$f = \sum_{\nu} a_\nu X^\nu \in \mathcal{O}[[X]] \qquad (a_\nu \in \mathcal{O} \text{ for all } \nu)$$

such that $a_\nu \to 0$ (in the $\mathfrak{m}$-adic topology on $\mathcal{O}$) as $|\nu| \to \infty$. Here $\nu = (\nu_1, \dots, \nu_N)$ ranges over all multi-indices in $\mathbb{N}^N$, and $|\nu| = \nu_1 + \cdots + \nu_N$.

The $\mathfrak{m}$-adic valuation $v_\mathfrak{m} \colon \mathcal{O} \to \mathbb{N}_\infty$ extends to $\mathcal{O}\langle X\rangle$ by setting

$$v_\mathfrak{m}(f) = \min_{\nu} v_\mathfrak{m}(a_\nu) \qquad \text{for } f = \sum_{\nu} a_\nu X^\nu \in \mathcal{O}\langle X\rangle.$$

The map $v_\mathfrak{m} \colon \mathcal{O}\langle X\rangle \to \mathbb{N}_\infty$ is a valuation on the domain $\mathcal{O}\langle X\rangle$, that is, for all $f, g \in \mathcal{O}\langle X\rangle$ we have $v_\mathfrak{m}(fg) = v_\mathfrak{m}(f) + v_\mathfrak{m}(g)$ and $v_\mathfrak{m}(f+g) \geq \min\{v_\mathfrak{m}(f), v_\mathfrak{m}(g)\}$. (See [8, p. 44, Corollary 2].) We denote the image of $f \in \mathcal{O}\langle X\rangle$ under the canonical surjection $\mathcal{O}\langle X\rangle \to \mathcal{O}\langle X\rangle/t\mathcal{O}\langle X\rangle \cong \overline{\mathcal{O}}[X]$ by $\overline{f}$.

Suppose from now on that $N \geq 1$, and let $X' := (X_1, \dots, X_{N-1})$. Canonically $\mathcal{O}\langle X'\rangle \subseteq \mathcal{O}\langle X\rangle$, and every element $f \in \mathcal{O}\langle X\rangle$ can be written uniquely as

$$f = \sum_{i=0}^{\infty} f_i X_N^i \quad \text{with } f_i(X') \in \mathcal{O}\langle X'\rangle \text{ for all } i \in \mathbb{N}, \tag{2.1}$$

where the infinite sum converges with respect to the $\mathfrak{m}\mathcal{O}\langle X\rangle$-adic topology on $\mathcal{O}\langle X\rangle$. An element f of $\mathcal{O}\langle X\rangle$, expressed as in (2.1), is called **regular in X_N of degree** $s \in \mathbb{N}$ if its reduction $\overline{f} \in \overline{\mathcal{O}}[X]$ is unit-monic of degree s in X_N, that is,

(1) $\overline{f_s} \neq 0$, and
(2) $v_\mathfrak{m}(f_i) > 0$ for all $i > s$.

If $f \in \mathcal{O}\langle X'\rangle[X_N]$ is monic of X_N-degree s (so that in particular f is regular in X_N of degree s, as an element of $\mathcal{O}\langle X\rangle$), then f is called a **Weierstraß polynomial in X_N of degree s.** For a proof of the following standard facts see, e.g., [8].

Lemma 2.1. (Noether normalization) *Let R be a domain, $e > 1$, and let $f \in R[X_1, \dots, X_N] = R[X]$, $f \neq 0$, be of total degree $< e$. Then the R-automorphism $T_e \colon R[X] \to R[X]$ given by*

$$\begin{aligned} X_i &\mapsto X_i + X_N^{e^{N-i}} \qquad (\textit{for } 1 \leq i < N), \\ X_N &\mapsto X_N \end{aligned}$$

has the property that for some $s < e^N$ and non-zero $u \in R$

$$T_e(f) = uX_N^s + \textit{terms of lower degree.}$$

Applying this to $R = \overline{\mathcal{O}}$, one concludes:

Lemma 2.2. *Let $e > 1$ and suppose that the image of $f \in \mathcal{O}\langle X\rangle$ in $\overline{\mathcal{O}}[X]$ is non-zero of degree $< e$. Let $T_e \colon \mathcal{O}\langle X\rangle \to \mathcal{O}\langle X\rangle$ be the $\mathcal{O}$-automorphism defined by*

$$\begin{aligned} X_i &\mapsto X_i + X_N^{e^{N-i}} \qquad (\textit{for } 1 \leq i < N), \\ X_N &\mapsto X_N. \end{aligned}$$

Then $T_e(f)$ is regular in X_N of degree $< e^N$.

The ring of restricted power series has the following fundamental property:

Theorem 2.3. (Weierstraß Division Theorem for $\mathcal{O}\langle X\rangle$) *Let $g \in \mathcal{O}\langle X\rangle$ be regular in X_N of degree s. Then for each $f \in \mathcal{O}\langle X\rangle$ there are uniquely determined elements $q \in \mathcal{O}\langle X\rangle$ and $r \in \mathcal{O}\langle X'\rangle[X_N]$ with $\deg_{X_N} r < s$ such that $f = qg + r$.*

In particular, we get

$$\mathcal{O}\langle X\rangle/(g) \cong \mathcal{O}\langle X'\rangle \oplus \mathcal{O}\langle X'\rangle\overline{X_N} \oplus \cdots \oplus \mathcal{O}\langle X'\rangle\overline{X_N}^{s-1}$$

as $\mathcal{O}\langle X'\rangle$-algebras. (Here $\overline{X_N} = X_N \bmod g$.) Applying Weierstraß Division with $f = X_N^s$, we obtain the important corollary:

Corollary 2.4. (Weierstraß Preparation Theorem for $\mathcal{O}\langle X\rangle$) *Let $g \in \mathcal{O}\langle X\rangle$ be regular in X_N of degree s. There are a unique Weierstraß polynomial $w \in \mathcal{O}\langle X'\rangle[X_N]$ of degree s and a unique unit $u \in \mathcal{O}\langle X\rangle$ such that $g = u \cdot w$.*

From Weierstraß Preparation it follows that the ring $\mathcal{O}\langle X\rangle$ is Noetherian. Here is another useful consequence:

Corollary 2.5. *Let $w \in \mathcal{O}\langle X'\rangle[X_N]$ be a Weierstraß polynomial. Then the inclusion map $\mathcal{O}\langle X'\rangle[X_N] \subseteq \mathcal{O}\langle X\rangle$ induces an isomorphism*

$$\mathcal{O}\langle X'\rangle[X_N]/w\mathcal{O}\langle X'\rangle[X_N] \xrightarrow{\cong} \mathcal{O}\langle X\rangle/w\mathcal{O}\langle X\rangle.$$

Proof. The surjectivity of the map follows from the existence part of Weierstraß Division. For injectivity, we have to show: if $fw = g \in \mathcal{O}\langle X'\rangle[X_N]$ for some $f \in \mathcal{O}\langle X\rangle$, then $f \in \mathcal{O}\langle X'\rangle[X_N]$. This follows by Euclidean Division of g by the monic polynomial w in $\mathcal{O}\langle X'\rangle[X_N]$ and by the uniqueness statement in the Weierstraß Division Theorem. □

Now let $\mathcal{O}$ be an arbitrary DVR, not necessarily complete, with maximal ideal generated by t, and let $\widehat{\mathcal{O}}$ be the completion of $\mathcal{O}$ in the $\mathfrak{m}$-adic topology. We let $F = \operatorname{Frac}(\mathcal{O})$ be the fraction field of $\mathcal{O}$. The following lemma and its corollary below will become important in later sections.

Lemma 2.6. *If a (finite) system of linear equations over $\mathcal{O}[X]$ has a solution in $F[X]$ and in $\widehat{\mathcal{O}}\langle X\rangle$, then it has a solution in $\mathcal{O}[X]$.*

Proof. For simplicity, we just treat the case of a single linear equation

$$f_0 = f_1y_1 + \cdots + f_ny_n \qquad (f_0, f_1, \ldots, f_n \in \mathcal{O}[X]). \tag{2.2}$$

The general case is similar. From a solution in $F[X]$ we obtain, after clearing denominators, an integer $e \geq 1$ and polynomials $g_1, \ldots, g_n \in \mathcal{O}[X]$ such that

$$t^e f_0 = f_1g_1 + \cdots + f_ng_n. \tag{2.3}$$

Now $\widehat{\mathcal{O}}\langle X\rangle$ is faithfully flat over its subring $(S_e)^{-1}\mathcal{O}[X]$, where S_e is the multiplicative set $1 + t^e\mathcal{O}[X]$. (See [17, Theorems 4.9, 5.1].) So if (2.2) is solvable in $\widehat{\mathcal{O}}\langle X\rangle$, then there exist $h, h_1, \ldots, h_n \in \mathcal{O}[X]$ with

$$(1 + t^e h)f_0 = f_1h_1 + \cdots + f_nh_n. \tag{2.4}$$

Multiplying (2.3) on both sides by h and subtracting from (2.4), we obtain

$$f_0 = f_1(h_1 - hg_1) + \cdots + f_n(h_n - hg_n)$$

with $h_1 - hg_1, \ldots, h_n - hg_n \in \mathcal{O}[X]$ as desired. □

Corollary 2.7. *Let A be an $m \times n$-matrix over $\mathcal{O}[X]$. If*

$$y^{(1)}, \ldots, y^{(L)} \in \big(\mathcal{O}[X]\big)^n$$

generate the $F[X]$-module $\mathrm{Sol}_{F[X]}(A)$ of solutions of the homogeneous system of linear equations $Ay = 0$ in $F[X]$ and

$$z^{(1)}, \ldots, z^{(M)} \in \big(\mathcal{O}[X]\big)^n$$

generate the $\widehat{\mathcal{O}}\langle X\rangle$-module $\mathrm{Sol}_{\widehat{\mathcal{O}}\langle X\rangle}(A)$ of solutions of $Ay = 0$ in $\widehat{\mathcal{O}}\langle X\rangle$, then

$$y^{(1)}, \ldots, y^{(L)}, z^{(1)}, \ldots, z^{(M)}$$

generate the $\mathcal{O}[X]$-module $\mathrm{Sol}_{\mathcal{O}[X]}(A)$ of solutions of $Ay = 0$ in $\mathcal{O}[X]$. □

3. Hermann's method

In this section, we first give a presentation of Hermann's method for constructing generators for the solutions of systems of homogeneous linear equations over polynomial rings. We begin by adapting this approach so that it applies to systems of linear equations over any integral domain D. In the next section we will use a variant of Hermann's method in the case where $D = \mathcal{O}\langle X\rangle$ for a complete DVR $\mathcal{O}$. Here we present the case (treated by Hermann) where D is a polynomial ring over a field and deduce bounds on the degrees of generators for syzygy modules. Finally we show how this method can be modified to solve inhomogeneous systems.

Hermann's method in a general setting. Let D be an integral domain with fraction field K. (Typically, D is a ring of polynomials over an integral domain.) We consider a homogeneous system of linear equations

$$\text{(I)} \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

with coefficient matrix $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ having entries $a_{ij} \in D$. We are interested in effectively finding a set of generators for the module of syzygies $\mathrm{Sol}(A) = \mathrm{Sol}_D(A)$ of A. Of course, for this we may assume $A \neq 0$. We shall indicate here a reduction of this problem to a similar problem over a coefficient ring (a quotient of D) that is in many cases simpler than the domain D.

Let $r = \mathrm{rank}_K(A) \ge 1$ be the rank of A (considered as a matrix over K) and let Δ be an $r \times r$-submatrix of A with $\delta = \det \Delta \neq 0$. After rearranging the order of the equations and permuting the unknowns $y_1, \ldots, y_n$ in (I), we may assume that Δ is the upper left corner of A, i.e., $\Delta = (a_{ij})_{1 \le i,j \le r}$. Each row $a_i = (a_{i1}, \ldots, a_{in})$ with $r < i \le m$ is a K-linear combination of the first r rows $a_1, \ldots, a_r$, so (I) has the same solutions in D^n as the system

$$\text{(II)} \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Changing the notation, we let $r = m$ and $A = (a_{ij})_{\substack{1 \le i \le r \\ 1 \le j \le n}}$. So (II) can now be written as $Ay = 0$. Multiplying both sides of $Ay = 0$ on the left by the adjoint of

Δ, (II) turns into the system

$$\text{(S)} \qquad \begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

with $c_{ij}, d_i \in D$ for $1 \leq i \leq r < j \leq n$, which has the same solutions in D^n as (II) and as (I). We note the following $n - r$ linearly independent solutions of (S):

$$v^{(1)} = \begin{bmatrix} -c_{1,r+1} \\ \vdots \\ -c_{r,r+1} \\ \delta \\ 0 \\ \vdots \\ 0 \end{bmatrix}, v^{(2)} = \begin{bmatrix} -c_{1,r+2} \\ \vdots \\ -c_{r,r+2} \\ 0 \\ \delta \\ \vdots \\ 0 \end{bmatrix}, \ldots, v^{(n-r)} = \begin{bmatrix} -c_{1,n} \\ \vdots \\ -c_{r,n} \\ 0 \\ \vdots \\ 0 \\ \delta \end{bmatrix} \tag{3.1}$$

If δ is a unit, these vectors form in fact a basis for $\mathrm{Sol}(A)$. Suppose δ is not a unit, so $\overline{D} = D/\delta D \neq 0$. Then, reducing the coefficients in (S) modulo δ, the system (S) turns into the system

$$\text{($\overline{\text{S}}$)} \qquad \begin{bmatrix} \overline{c_{1,r+1}} & \cdots & \overline{c_{1n}} \\ \vdots & \ddots & \vdots \\ \overline{c_{r,r+1}} & \cdots & \overline{c_{rn}} \end{bmatrix} \begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

over $\overline{D}$. (Here $\overline{a}$ denotes the image of $a \in D$ in $\overline{D}$.)

Lemma 3.1. *Let $z^{(1)}, \ldots, z^{(M)} \in D^{n-r}$ be such that $\overline{z^{(1)}}, \ldots, \overline{z^{(M)}} \in \overline{D}^{n-r}$ generate the $\overline{D}$-module of solutions to ($\overline{\text{S}}$). The vectors $z^{(1)}, \ldots, z^{(M)}$ may be extended uniquely to vectors $y^{(1)}, \ldots, y^{(M)}$ in D^n which, together with the solutions of* (I) *in* (3.1)*, generate* $\mathrm{Sol}(A)$.

This fact is rather obvious, but what makes it useful is that under favorable circumstances $\overline{D}$ is "simpler" than D. (Note however that it may happen that $\overline{D}$ is not a domain anymore.) Let us consider an example where this can be exploited.

Hermann's method for $F[X]$. Assume that D is a polynomial ring over a field F, that is, $D = F[X] = F[X_1, \ldots, X_N]$. Let $N > 0$. Suppose first that F is *infinite.* In this case, after a linear change of variables, we may assume that

$$\delta = uX_N^e + \text{terms of lower } X_N\text{-degree}, \quad \text{with } e = \deg \delta > 0,\ u \in F^\times. \tag{3.2}$$

Then by Euclidean Division each element $\overline{a} \in \overline{D}$ can be uniquely written as

$$\overline{a} = a_0 + a_1\overline{X_N} + a_2\overline{X_N}^2 + \cdots + a_{e-1}\overline{X_N}^{e-1}$$

with $a_0, \ldots, a_{e-1} \in F[X'] = F[X_1, \ldots, X_{N-1}]$. In particular, each coefficient $\overline{c_{ij}}$ in ($\overline{\text{S}}$) can be written in this way. Note that $\deg_{X'} a_i \leq \deg_X a$ for all $0 \leq i < e$. Let us also write each unknown y_j in ($\overline{\text{S}}$), for $r < j \leq n$, as

$$y_j = y_{j0} + y_{j1}\overline{X_N} + \cdots + y_{j,e-1}\overline{X_N}^{e-1}$$

with new unknowns y_{jk} ($r < j \le n$, $0 \le k < e$) ranging over $D' = F[X']$. Each product $\overline{c_{ij}}y_j$ in $(\overline{\mathrm{S}})$ can then be written as

$$\beta_0(y_{j0},\dots,y_{j,e-1}) + \beta_1(y_{j0},\dots,y_{j,e-1})\overline{X_N} + \cdots + \beta_{e-1}(y_{j0},\dots,y_{j,e-1})\overline{X_N}^{e-1},$$

where each β_k is a linear form in $y_{j0},\dots,y_{j,e-1}$ with coefficients in D'. From this, it is routine to construct a homogeneous system of $r(e-1)$ linear equations in the $e(n-r)$ unknowns y_{jk} over D' whose solutions in D' are in one-to-one correspondence with the solutions of $(\overline{\mathrm{S}})$ in $\overline{D}$.

Computing degree bounds. For the sake of obtaining "good" bounds on the degrees of solutions, we modify the general construction sketched above, exploiting some more special features of $F[X]$. Put $d = \deg_{X_N} A$. Write each a_{ij} as

$$a_{ij} = a_{ij0} + a_{ij1}X_N + \cdots + a_{ijd}X_N^d \tag{3.3}$$

with $a_{ijk} \in F[X']$ and also each unknown y_j as

$$y_j = y_{j0} + y_{j1}X_N + \cdots + y_{j,rd-1}X_N^{rd-1} \tag{3.4}$$

with new unknowns y_{jk} ranging over $F[X']$. Then the ith equation in (II) yields $(r+1)d$ equations

$$\sum_{l=0}^{k}\sum_{j=1}^{n} a_{ijl}y_{j,k-l} = 0, \qquad 0 \le k < (r+1)d,$$

where we put $a_{ijl} := 0$ for $l > d$ and $y_{i,l} := 0$ for $l \ge rd$. In this way, we obtain a new system

$$A'y' = 0, \tag{I$'$}$$

where A' is an $\big(rd(r+1)\big) \times (nrd)$-matrix with entries in D' and

$$y' = \big[y_{1,0},\dots,y_{1,rd-1},\dots,y_{n,0},\dots,y_{n,rd-1}\big]^{\mathrm{tr}}, \tag{3.5}$$

whose solutions in D' are in one-to-one correspondence with the solutions of (II) in D of X_N-degree $< rd$. Note that the entries of A' are still of degree (in X') at most $\deg_X A$. If $N > 1$, then we can repeat the same procedure with (I$'$) instead of (I), etc., until we obtain a (huge) homogeneous system of linear equations over F. We can (effectively) find a finite set of generators for the F-vector space of solutions to this system, and reversing the process above, we obtain a finite set of generators for the original system (I): Suppose we have already found a finite set of generators for the D'-submodule $\mathrm{Sol}_{D'}(A')$ of $(D')^{nrd}$, where A' is the matrix constructed from A as above. That is, we have finitely many solutions $y^{(1)},\dots,y^{(M')}$ of (I) such that each solution to (I) of X_N-degree $< rd$ is a linear combination of $y^{(1)},\dots,y^{(M')}$. The solutions in (3.1) together with $y^{(1)},\dots,y^{(M')}$ form a set of generators for $\mathrm{Sol}(A) = \mathrm{Sol}_D(A)$: Given any solution $y = [y_1,\dots,y_n]^{\mathrm{tr}} \in \mathrm{Sol}(A)$, we can divide each y_j, $j = n-r+1,\dots,n$, by δ:

$$y_j = Q_{j-r}\delta + R_{j-r} \qquad (j = n-r+1,\dots,n)$$

with $Q_1,\dots,Q_{n-r} \in F[X]$ and $R_1,\dots,R_{n-r} \in F[X]$ of X_N-degree $< e$. Then

$$z = y - Q_1v^{(1)} - \cdots - Q_{n-r}v^{(n-r)} = \big[h_1,\dots,h_r,R_1,\dots,R_{n-r}\big]^{\mathrm{tr}}$$

is also a solution to (S), with $h_1,\dots,h_r \in F[X]$. Now

$$\delta h_i = -(c_{i,r+1}R_1 + \cdots + c_{in}R_{n-r}) \qquad \text{for } i = 1,\dots,r,$$

where the right-hand sides have X_N-degree $< rd + e$. Hence $\deg_{X_N} h < rd$ and therefore $\deg_{X_N} z < rd$. It follows that z is a D-linear combination of $y^{(1)}, \dots, y^{(M')}$, so y is a D-linear combination of $y^{(1)}, \dots, y^{(M')}, v^{(1)}, \dots, v^{(n-r)}$ as claimed.

Let $\alpha = \alpha(N, d, m)$ be the smallest natural number such that for all infinite fields F, a system of m homogeneous linear equations (I) over $D = F[X] = F[X_1, \dots, X_N]$ with all $\deg a_{ij}$ bounded from above by d is generated by the solutions of degree $\leq \alpha$. (By the considerations above, $\alpha(N, d, m)$ exists.) The derived system (I′) consists of at most $dm(m+1)$ equations in at most dn^2 unknowns, and $\deg_{X'}(A') \leq d$. From a set of generators of the solutions to (I′) of degree $\leq d'$ we can produce a set of generators of the solutions to (I) of degree $\leq d' + md$. We get the relation

$$\alpha(N, d, m) \leq \alpha\big(N - 1, d, dm(m+1)\big) + md$$

for $N > 0$. Noting that $\alpha(0, d, m) = 0$ for all d, m, we find that

$$\alpha(N, d, m) \leq (m+1)d + \big((m+1)d\big)^2 + \cdots + \big((m+1)d\big)^{2^{N-1}} \leq (2md)^{2^N}.$$

If F is any field, possibly finite, we work over $F' = F(T)$, an infinite field. Here, T is an indeterminate distinct from $X_1, \dots, X_N$. Given $y \in \big(F[T, X]\big)^n$, write $y = y(0) + y(1)T + y(2)T^2 + \cdots$ (a finite sum) with $y(k) \in \big(F[X]\big)^n$ for all k. If $\mathcal{G}$ is a generating set for $\mathrm{Sol}_{F'[X]}(A)$ consisting of elements of $\big(F[T, X]\big)^n$, then the collection of $y(k)$, where $y \in \mathcal{G}$ and $k \in \mathbb{N}$, generates $\mathrm{Sol}_{F[X]}(A)$. To sum up, we have shown the classical result:

Theorem 3.2. (Hermann [20], Seidenberg [35]) *For every polynomial ring $D = F[X_1, \dots, X_N]$ over a field F and $A \in D^{m \times n}$ of degree at most d, the solution module $\mathrm{Sol}_D(A)$ of the homogeneous system $Ay = 0$ is generated by the solutions of degree at most $\beta(N, d, m) = (2md)^{2^N}$.* □

Hermann's method for inhomogeneous systems. Again let D be a domain with fraction field K. Given an $m \times n$-matrix $A = (a_{ij})$ with entries $a_{ij} \in D$, we are now interested in determining for each column vector $b = [b_1, \dots, b_m]^{\mathrm{tr}} \in D^m$ whether the system

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{I$_b$}$$

(or $Ay = b$) is solvable for some $y = [y_1, \dots, y_n]^{\mathrm{tr}} \in D^n$, and if it is, effectively finding such a solution. Similarly to the case of homogeneous equations, this problem can be reduced to an analogous problem over a quotient of D: As above let Δ be an $r \times r$-submatrix of A with $\delta = \det \Delta \neq 0$, where $r = \mathrm{rank}_K(A) \geq 1$. Again we may assume that $\Delta = (a_{ij})_{1 \leq i,j \leq r}$. Each row $a_i = (a_{i1}, \dots, a_{in})$ with $r < i \leq m$ is a K-linear combination

$$a_i = \sum_{\varrho=1}^{r} \lambda_{i\varrho} a_\varrho \qquad (\lambda_{i\varrho} \in K)$$

of the first r rows $a_1, \dots, a_r$. So a *necessary condition* for (I$_b$) to have a solution in D^n is that

$$b_i = \sum_{\varrho=1}^{r} \lambda_{i\varrho} b_\varrho \qquad \text{for } r < i \leq m. \tag{NC}$$

(That is, $\operatorname{rank}_K(A) = \operatorname{rank}_K(A, b)$.) Assume (NC) holds. Then (I_b) has the same solutions in D^n as the system

$$(\mathrm{II}_b) \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_r \end{bmatrix}$$

Changing the notation, we let $r = m$, so (II_b) can now be written as $Ay = b$. Multiplying both sides of $Ay = b$ on the left by the adjoint Δ^{ad} of Δ, (II_b) turns into the system

$$(\mathrm{S}_b) \qquad \begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{bmatrix}$$

(with $c_{ij}, d_i \in D$ for $1 \le i \le r < j \le n$) which has the same solutions in D^n as (II_b) and as (I_b). Clearly, a sufficient condition for (S_b) to have a solution $y = [y_1, \ldots, y_n]^{\mathrm{tr}} \in D^n$ is that $d_1, \ldots, d_r$ are each divisible by δ. This will be the case if δ is a unit. A solution to (S_b) (and hence to (I_b)) is then given by

$$y_j = \begin{cases} d_j/\delta & \text{for } 1 \le j \le r, \\ 0 & \text{for } d < j \le n. \end{cases}$$

Suppose δ is not a unit, so $\overline{D} = D/\delta D \neq 0$. Then, reducing the coefficients in (S_b) modulo δ, the system (S_b) turns into

$$(\overline{\mathrm{S}_b}) \qquad \begin{bmatrix} \overline{c_{1,r+1}} & \cdots & \overline{c_{1n}} \\ \vdots & \ddots & \vdots \\ \overline{c_{r,r+1}} & \cdots & \overline{c_{rn}} \end{bmatrix} \begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \overline{d_1} \\ \vdots \\ \overline{d_r} \end{bmatrix}$$

over $\overline{D}$. The key fact here is the following (similar to Lemma 3.1):

Lemma 3.3. *Any $[y_{r+1}, \ldots, y_n]^{\mathrm{tr}} \in D^{n-r}$ with the property that $\left[\overline{y_{r+1}}, \ldots, \overline{y_n}\right]^{\mathrm{tr}}$ is a solution of the reduced system ($\overline{\mathrm{S}_b}$) can be augmented uniquely to a solution*

$$y = \left[y_1, \ldots, y_r, y_{r+1}, \ldots, y_n\right]^{\mathrm{tr}} \in D^n$$

of (S_b) and hence of (I_b). (In particular, (I_b) is solvable in D if and only if ($\overline{\mathrm{S}_b}$) is solvable in $\overline{D}$.)

In the case where $D = F[X]$ is a polynomial ring over a field F, we can again modify this reduction somewhat to facilitate the computation of bounds. Suppose that $N > 0$ and F is infinite. Then, after applying a linear change of variables, we may assume that δ has the form (3.2). By Euclidean Division we write each b_i as

$$b_i = \delta f_i + g_i \qquad \text{with } f_i, g_i \in D, \deg_{X_N} g_i < e.$$

The solutions of (II_b) in D^n are in one-to-one correspondence with the solutions in D^n of the system

$$(\mathrm{III}_b) \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} g_1 \\ \vdots \\ g_r \end{bmatrix}$$

with the same coefficient matrix A as (II_b). To see this, let

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_r \end{bmatrix}, \quad g = \begin{bmatrix} g_1 \\ \vdots \\ g_r \end{bmatrix}, \quad \text{and} \quad h = \begin{bmatrix} \Delta^{\mathrm{ad}} f \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in D^n.$$

Note that $b = \delta f + g$ and $Ah = \delta f$; so $y \in D^n$ is a solution to (II_b) if and only if $y - h \in D^n$ is a solution to (III_b). Moreover, if all a_{ij} and b_i have degree $\leq d$, and if (II_b) is solvable in D^n, then (III_b) even has a solution in D^n of X_N-degree $< rd$. In order to prove this, suppose $y = [y_1, \dots, y_n]^{\mathrm{tr}} \in D^n$ is a solution to (II_b). The polynomial δ, each c_{ij} and each d_i have degree at most rd. Subtracting from y appropriate multiples of the solutions $v^{(1)}, \dots, v^{(n-r)}$ (see (3.1)) to the homogeneous system $Ay = 0$ associated with (II_b), if necessary, we may assume that $\deg_{X_N} y_j < e \leq rd$ for $j = r+1, \dots, n$. Multiplying the equation $A(y-h) = g$ on both sides from the left by the adjoint Δ^{ad} of Δ, we get, for $j = 1, \dots, r$:

$$\delta(y_j - h_j) = \sum_{k=r+1}^{n} y_k c_{jk} + (\text{terms of } X_N\text{-degree} < e + rd).$$

It follows that $\deg_{X_N}\big(\delta(y_j - h_j)\big) < e + rd$ and thus $\deg_{X_N}(y_j - h_j) < rd$. So $y - h$ is a solution to (III_b) of X_N-degree $< rd$ as required.

Write each g_i as

$$g_i = g_{i0} + g_{i1} X_N + \cdots + g_{i,e-1} X_N^{e-1}$$

with $g_{i0}, \dots, g_{i,e-1} \in F[X']$, each a_{ij} in the form (3.3), and also each unknown y_j as in (3.4). Comparing the coefficients of equal powers of X_N on both sides, the ith equation in (III_b) yields $(r+1)d$ equations

$$\sum_{l=0}^{k} \sum_{j=1}^{n} a_{ijl} y_{j,k-l} = g_{ik}, \qquad 0 \leq k < (r+1)d,$$

with $a_{ijl} := 0$ for $l > d$, $y_{i,l} := 0$ for $l \geq rd$, $g_{il} := 0$ for $l \geq e$. We get a new system

$$A' y' = b', \tag{I'_b}$$

where A' is an $\big(rd(r+1)\big) \times (nrd)$-matrix with entries in D', b' is an $\big(rd(r+1)\big)$-column vector with components from D', and y' is as in (3.5), whose solutions in D' are in one-to-one correspondence with the solutions of (III_b) in D of X_N-degree $< rd$. So starting with a system (I) over $D = F[X_1, \dots, X_N]$, we have constructed a system (I'_b) over $D' = F[X_1, \dots, X_{N-1}]$ which is, assuming (NC), in some sense equivalent to it. Note that $\deg_{X'}(A', b') \leq d$.

Associated to (I'_b) we have the necessary condition

$$\mathrm{rank}_{K'}(A') = \mathrm{rank}_{K'}(A', b') \qquad (\text{where } K' = \mathrm{Frac}(D')) \tag{NC$'$}$$

for its solvability in D'. So if $N > 1$ and (NC$'$) holds, then we can repeat the procedure with (I'_b), until we obtain a system of linear equations over K. We can (effectively) decide whether this system has a solution over K, and if it does, find one, e.g., by Gaussian Elimination. Eventually we obtain a solution $y \in D^n$ of the original system (I_b) with $\deg y \leq \beta(N, d, m) = (2md)^{2^N}$, where $d = \deg(A, b)$.

If F is a finite field, we again work over the infinite field $F' = F(T)$. The algorithm described above allows us to test whether the system (I_b) has a solution $y' = [y'_1, \dots, y'_n]^{\mathrm{tr}} \in \big(F'[X]\big)^n$, and if it does, to effectively obtain such a solution with $\deg_X y' \leq \beta(N, d, m)$. Since the coefficients of the y'_j solve a certain system of

linear equations involving the coefficients of the a_{ij} and the b_i, we can also find a solution y *in* $F[X]$ with the $\deg y$ majorized by the same bound. This shows:

Theorem 3.4. (Hermann [20], Seidenberg [35]) *For every polynomial ring $D = F[X_1, \ldots, X_N]$ over a field F and $A \in D^{m\times n}$, $b \in D^m$ of degree at most d, if the system of linear equations $Ay = b$ has a solution in D^n, then it has such a solution of degree at most $(2md)^{2^N}$.* □

Remark. Theorems 3.2 and 3.4 above remain true for a polynomial ring $D = F[X_1, \ldots, X_N]$ over a von Neumann regular ring F. This follows easily from the fact that any von Neumann regular ring admits a faithfully flat embedding into a direct product of fields.

Given a Noetherian domain R with fraction field F and an ideal I of $R[X]$, there exists $\delta \in R$ such that $IF[X] \cap R[X] = (I : \delta)$. More generally, we have:

Corollary 3.5. *Let R be a domain with fraction field $F = \mathrm{Frac}(R)$, let $R[X] = R[X_1, \ldots, X_N]$, and let M be a finitely generated submodule of the free $R[X]$-module $R[X]^m$. Then there exists $\delta \in R$ with the property that for every domain R' extending R with fraction field F':*

$$MF'[X] \cap R'[X]^m = (MR'[X] : \delta). \tag{3.6}$$

If R is computable, then δ can be computed elementary recursively (in the ring operations of R) from a given finite collection of generators for M.

Note that M has the form

$$M = \left\{b \in D^m : Ay = b \text{ has a solution in } D^n\right\}$$

for some matrix A of size $m \times n$ (for some n) with coefficients in D, and

$$MS[X] = \left\{b \in S[X]^m : Ay = b \text{ has a solution in } S[X]^n\right\}$$

for every ring S extending R. The corollary follows by a slight modification of Hermann's method as described in the proof of Theorem 3.4 above. The difference is that we now apply Lemma 2.1 instead of a linear change of variables to bring the chosen minor of A in the form (3.2). (This works regardless of F being infinite or not.) Note that then (using the notation introduced earlier) the only arithmetic operations performed in constructing (A', b') from (A, b) are additions, subtractions, multiplications, and divisions by the leading coefficient u of δ. (The latter occur in the Euclidean Division by δ.) But u itself arises as a polynomial in the coefficients *of the entries of A only.* We leave the details to the reader. From Corollary 3.5 and the proof Theorem 3.2 we obtain:

Corollary 3.6. *Let R be a domain, $F = \mathrm{Frac}(R)$, and let A be an $m \times n$-matrix with entries in $D = R[X] = R[X_1, \ldots, X_N]$. There exist generators $u^{(1)}, \ldots, u^{(K)} \in \mathrm{Sol}_D(A)$ of $\mathrm{Sol}_{F[X]}(A)$ and $\delta \in R$ such that*

$$\mathrm{Sol}_D(A) = (M : \delta), \qquad \text{where } M = Du^{(1)} + \cdots + Du^{(K)}.$$

If R is computable, then the $u^{(k)}$ and δ can be computed elementary recursively (in the ring operations of R) from A. □

4. Effective flatness

The purpose of this section is to prove Theorem B from the introduction, in a more general setting. A ring R is called **hereditary** if every ideal of R is projective (as an R-module). A domain R is hereditary if and only if R is a Dedekind domain ([16, p. 27]). A domain R is called **almost Dedekind** if every localization $R_{\mathfrak{m}}$ of R at a maximal ideal $\mathfrak{m}$ of R is a DVR. (See [15, p. 434].) Somewhat more generally, we shall call a ring R **almost hereditary** if the ring of fractions $\mathrm{Frac}(R)$ of R is von Neumann regular and $R_{\mathfrak{m}}$ is a DVR for every maximal ideal $\mathfrak{m}$ of R. Every hereditary ring is almost hereditary ([16, pp. 27–28]). There exist examples of domains which are almost Dedekind but not Dedekind; see [15, pp. 516–518]. With this terminology, we have:

Theorem 4.1. *Let R be an almost hereditary ring and let $A = (a_{ij}) \in D^{m\times n}$, $A \neq 0$, where $D = R[X_1, \dots, X_N]$. The module of solutions to $Ay = 0$ in D is generated by elements of degree at most $(2m \deg A)^{2((N+1)^N-1)}$.*

Since an almost hereditary ring is semihereditary and hence coherent (see [16, p. 128]), finitely many such generators will suffice. Theorem 4.1 specializes to Theorem B when applied to $R = \mathbb{Z}$ and $m = 1$.

As a first step in the proof of this theorem, we show an easy local-global result:

Lemma 4.2. *Let R be a ring and let M be an $R[X]$-submodule of $R[X]^n$. For each maximal ideal $\mathfrak{m}$ of R let $v_{\mathfrak{m}}^{(1)}, \dots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})} \in M$ generate the $R_{\mathfrak{m}}[X]$-submodule $MR_{\mathfrak{m}}[X]$ of $R_{\mathfrak{m}}[X]^n$. Then $v_{\mathfrak{m}}^{(1)}, \dots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})}$, where $\mathfrak{m}$ ranges over all maximal ideals of R, generate the $R[X]$-module M.*

Proof. Let $y \in M$. Then for each maximal ideal $\mathfrak{m}$ of R there exist $\delta_{\mathfrak{m}} \in R \setminus \mathfrak{m}$ and $b_{1,\mathfrak{m}}, \dots, b_{K_{\mathfrak{m}},\mathfrak{m}} \in R[X]$ such that

$$\delta_{\mathfrak{m}} y = b_{\mathfrak{m},1} v_{\mathfrak{m}}^{(1)} + \cdots + b_{\mathfrak{m},K_{\mathfrak{m}}} v_{\mathfrak{m}}^{(K_{\mathfrak{m}})}. \tag{4.1}$$

The various $\delta_{\mathfrak{m}}$, where $\mathfrak{m}$ ranges over all maximal ideals of R, generate the unit ideal of R. Hence there exist maximal ideals $\mathfrak{m}_1, \dots, \mathfrak{m}_k$ of R (for some $k \in \mathbb{N}$) and $c_1, \dots, c_k \in R$ such that

$$1 = c_1 \delta_{\mathfrak{m}_1} + \cdots + c_k \delta_{\mathfrak{m}_k}.$$

Therefore

$$y = c_1(\delta_{\mathfrak{m}_1} y) + \cdots + c_k(\delta_{\mathfrak{m}_k} y).$$

Together with (4.1) this shows that y is an $R[X]$-linear combination of the $v_{\mathfrak{m}}^{(j)}$. □

Remark. Let M' be a submodule of M generated by $u^{(1)}, \dots, u^{(K)}$, and let δ be an element of the ideal $(M' : M)$ of R. Similarly to the proof of the lemma one shows that $u^{(1)}, \dots, u^{(K)}$ together with $v_{\mathfrak{m}}^{(1)}, \dots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})}$, where $\mathfrak{m}$ ranges over all maximal ideals of R containing δ, suffice to generate M. Below we will apply this in the situation where $u^{(1)}, \dots, u^{(K)} \in M$ are generators of the $F[X]$-module $MF[X]$ and δ satisfies $MF[X] \cap R[X]^n = (M' : \delta)$. (Here $F = \mathrm{Frac}(R)$ denotes the ring of fractions of R.)

Now let R be an almost hereditary ring and $0 \neq A = (a_{ij}) \in D^{m\times n}$, where $D = R[X]$, $X = (X_1, \dots, X_N)$. Then $F = \mathrm{Frac}(R)$ is von Neumann regular, and

$R_{\mathfrak{m}}$ is a DVR, for every maximal ideal $\mathfrak{m}$ of R. By virtue of the lemma applied to $M = \mathrm{Sol}_D(A)$, it suffices to find

$$v_{\mathfrak{m}}^{(1)}, \ldots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})} \in \mathrm{Sol}_D(A)$$

generating $MR_{\mathfrak{m}}[X] = \mathrm{Sol}_{R_{\mathfrak{m}}[X]}(A)$, for each maximal ideal $\mathfrak{m}$ of R, with $v_{\mathfrak{m}}^{(j)}$ of "small" degree. For the construction of the $v_{\mathfrak{m}}^{(j)}$ we may use Corollary 2.7, since $R_{\mathfrak{m}}$ is a DVR. Hence, given a maximal ideal $\mathfrak{m}$ of R, we need to find

(1) $y_{\mathfrak{m}}^{(1)}, \ldots, y_{\mathfrak{m}}^{(L_{\mathfrak{m}})} \in \mathrm{Sol}_{R_{\mathfrak{m}}[X]}(A)$ generating $\mathrm{Sol}_{\mathrm{Frac}(R_{\mathfrak{m}})[X]}(A)$ and
(2) $z_{\mathfrak{m}}^{(1)}, \ldots, z_{\mathfrak{m}}^{(M_{\mathfrak{m}})} \in \mathrm{Sol}_{R_{\mathfrak{m}}[X]}(A)$ generating $\mathrm{Sol}_{\widehat{R_{\mathfrak{m}}\langle X\rangle}}(A)$,

with $y_{\mathfrak{m}}^{(i)}$ and $z_{\mathfrak{m}}^{(j)}$ of degree at most $(2m \deg A)^{2((N+1)^N-1)}$. By Hermann's Theorem 3.2 from the last section we obtain $y_{\mathfrak{m}}^{(i)}$ satisfying (1), of degree bounded by

$$(2m \deg A)^{2^N} \leq (2m \deg A)^{2((N+1)^N-1)} \qquad (\text{for } N > 0).$$

The existence of the $z_{\mathfrak{m}}^{(j)}$ is a consequence of the following *effective flatness* result applied to the DVR $\mathcal{O} = R_{\mathfrak{m}}$:

Proposition 4.3. *Let $\mathcal{O}$ be a DVR with maximal ideal $\mathfrak{m}$ and $\mathfrak{m}$-adic completion $\widehat{\mathcal{O}}$, and let $A = (a_{ij}) \in \big(\mathcal{O}[X]\big)^{m\times n}$, $A \neq 0$. There exist solutions $z^{(1)}, \ldots, z^{(M)} \in \mathrm{Sol}_{\mathcal{O}[X]}(A)$ of degree at most $(2m \deg A)^{2((N+1)^N-1)}$ which generate $\mathrm{Sol}_{\widehat{\mathcal{O}}\langle X\rangle}(A)$.*

In proving this proposition, we proceed by induction on N, following Hermann's method as in the proof of Theorem 3.2, with $F[X]$ replaced by $\widehat{\mathcal{O}}\langle X\rangle$ and Weierstraß Division for $\widehat{\mathcal{O}}\langle X\rangle$ in place of Euclidean Division for $F[X]$. However, this procedure breaks down if $\delta \bmod t = 0$ for all $r\times r$-minors δ of A, since then Weierstraß Division by δ is *inapplicable*. To overcome this obstacle, we shall first transform our system

$$\text{(I)} \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

into an equivalent system for which $\delta \bmod t \neq 0$ for a suitable $r \times r$-minor δ of the new coefficient matrix. For this, by removing superfluous rows from A, we may of course assume that the rows of A are linearly independent over the fraction field $F(X)$ of $\mathcal{O}[X]$, i.e., $m = r = \mathrm{rank}_{F(X)}(A) \geq 1$. Let Δ be an $r \times r$-submatrix of A such that $v_{\mathfrak{m}}(\det \Delta)$ is *minimal* among all $r \times r$-submatrices of A. Without loss of generality, $\Delta = (a_{ij})_{1\leq i,j\leq r}$. As in Section 3, consider now the system

$$\text{(S)} \qquad \begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which is obtained by multiplying both sides of (I) from the left with the adjoint of Δ. It has the same solutions as (I) in any domain extending $\mathcal{O}[X]$. Here, $\delta = \det \Delta$, and the c_{ij} are certain signed $r \times r$-minors of A. In particular, $v_{\mathfrak{m}}(c_{ij}) \geq v_{\mathfrak{m}}(\delta)$ for

all i, j, by choice of Δ. We have the $n-r$ linearly independent solutions

$$v^{(1)} = \begin{bmatrix} -c_{1,r+1} \\ \vdots \\ -c_{r,r+1} \\ \delta \\ 0 \\ \vdots \\ 0 \end{bmatrix}, v^{(2)} = \begin{bmatrix} -c_{1,r+2} \\ \vdots \\ -c_{r,r+2} \\ 0 \\ \delta \\ \vdots \\ 0 \end{bmatrix}, \dots, v^{(n-r)} = \begin{bmatrix} -c_{1,n} \\ \vdots \\ -c_{r,n} \\ 0 \\ \vdots \\ 0 \\ \delta \end{bmatrix} \tag{4.2}$$

to the homogeneous system (S). Put $\mu = v_{\mathfrak{m}}(\delta)$ and $u^{(k)} = t^{-\mu}v^{(k)} \in \big(\mathcal{O}[X]\big)^n$ for $k = 1, \dots, n-r$. If $N = 0$, then $t^{-\mu}\delta$ is a unit in $\mathcal{O}$, so the solutions $u^{(1)}, \dots, u^{(n-r)}$ form a basis of $\mathrm{Sol}_{\mathcal{O}}(A)$ and hence of $\mathrm{Sol}_{\widehat{\mathcal{O}}}(A)$ (since $\widehat{\mathcal{O}}$ is flat over $\mathcal{O}$). Suppose now that $N > 0$. We let $e = r \deg A + 1$ and put $b_{ij} = T_e(a_{ij})$, where T_e is the $\widehat{\mathcal{O}}$-automorphism of $\widehat{\mathcal{O}}\langle X\rangle$ defined in Lemma 2.2. Then the system $By = 0$, where $B = (b_{ij}) \in \big(\mathcal{O}[X]\big)^{m\times n}$, has the same rank r as (I), and $y \in \widehat{\mathcal{O}}\langle X\rangle^n$ is a solution to (I) if and only if $T_e(y)$ is a solution to $By = 0$. Dividing all coefficients δ and c_{ij} in (S) by t^{μ} and applying T_e to the resulting system, we obtain a system

$$\begin{bmatrix} \varepsilon & & & d_{1,r+1} & \cdots & d_{1,n} \\ & \varepsilon & & d_{2,r+1} & \cdots & d_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & & \varepsilon\ d_{r,r+1} & \cdots & d_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{S_e}$$

which has the same solutions, in any domain extending $\mathcal{O}[X]$, as $By = 0$, where $d_{ij} \in \mathcal{O}[X]$ for all i, j and $\varepsilon \in \mathcal{O}[X]$ is regular in X_N of some degree $s < e^N$. This system has the $n-r$ linearly independent solutions $w^{(1)}, \dots, w^{(n-r)}$, where $w^{(k)} = T_e\big(u^{(k)}\big)$ for $k = 1, \dots, n-r$. Let $d = \deg_{X_B} B$, so $d < e^N$. (Note that $\deg_{X'}(b_{ij}) \le \deg_{X'}(a_{ij})$ for all i, j.) Write

$$b_{ij} = b_{ij0} + b_{ij1}X_N + \cdots + b_{ijd}X_N^d$$

with $b_{ij0}, \dots, b_{ijd} \in \mathcal{O}[X']$ and each unknown y_j as

$$y_j = y_{j0} + y_{j1}X_N + \cdots + y_{j,rd-1}X_N^{rd-1}$$

with new unknowns y_{jk} ($1 \le j \le n$, $0 \le k < rd$) ranging over $\widehat{\mathcal{O}}\langle X'\rangle$. The ith equation in $By = 0$ may then be written as

$$\sum_{l=0}^{k}\sum_{j=1}^{n} b_{ijl}y_{j,k-l} = 0, \qquad\qquad 0 \le k < (r+1)d,$$

where we put $b_{ijl} := 0$ for $l > d$ and $y_{i,l} := 0$ for $l \ge rd$. This gives rise to a system over $\mathcal{O}[X']$:

$$A'y' = 0, \tag{4.3}$$

consisting of $rd(r+1)$ homogeneous equations in the nrd unknowns $y' = (y_{jk})$, whose solutions in $\widehat{\mathcal{O}}\langle X'\rangle$ are in one-to-one correspondence with the solutions $y \in \big(\widehat{\mathcal{O}}\langle X'\rangle[X_N]\big)^n$ to $By = 0$ with $\deg_{X_N} y < rd$. From a finite set of generators of $\mathrm{Sol}_{\mathcal{O}[X']}(A')$ we thus obtain finitely many column vectors

$$y^{(1)}, \dots, y^{(M')} \in \big(\mathcal{O}[X]\big)^n$$

with the following property: each $y^{(i)}$ is a solution to $By = 0$ of X_N-degree $< rd$, and each solution $y \in \big(\widehat{\mathcal{O}}\langle X'\rangle[X_N]\big)^n$ to this system of linear equations with $\deg_{X_N} y < rd$ is an $\widehat{\mathcal{O}}\langle X'\rangle$-linear combination of $y^{(1)}, \ldots, y^{(M')}$. Consider now the solutions

$$u^{(1)}, \ldots, u^{(n-r)}, T_e^{-1}\big(y^{(1)}\big), \ldots, T_e^{-1}\big(y^{(M')}\big) \in \big(\mathcal{O}[X]\big)^n \tag{4.4}$$

to (I). We show:

Lemma 4.4. *The vectors in* (4.4) *generate the* $\widehat{\mathcal{O}}\langle X\rangle$*-module* $\mathrm{Sol}_{\widehat{\mathcal{O}}\langle X\rangle}(A)$.

Proof. Suppose that $x \in \big(\widehat{\mathcal{O}}\langle X\rangle\big)^n$ is any solution to $Ay = 0$, and let $y = T_e(x)$, a solution to $By = 0$. Since ε is regular in X_N of degree s, we can write, by Weierstraß Division in $\widehat{\mathcal{O}}\langle X\rangle$:

$$y_j = Q_{j-r}\varepsilon + R_{j-r} \qquad (j = n-r+1, \ldots, n)$$

with $Q_1, \ldots, Q_{n-r} \in \widehat{\mathcal{O}}\langle X\rangle$ and $R_1, \ldots, R_{n-r} \in \widehat{\mathcal{O}}\langle X'\rangle[X_N]$ of X_N-degree $< s$. Then

$$z = y - Q_1 w^{(1)} - \cdots - Q_{n-r} w^{(n-r)} = \big[h_1, \ldots, h_r, R_1, \ldots, R_{n-r}\big]^{\mathrm{tr}}$$

is also a solution to (S_e), with $h_1, \ldots, h_r \in \widehat{\mathcal{O}}\langle X\rangle$. Let $U \in \widehat{\mathcal{O}}\langle X\rangle$ be a unit and $W \in \widehat{\mathcal{O}}\langle X'\rangle[X_N]$ be a Weierstraß polynomial such that $\varepsilon = UW$. Since ε is polynomial in X_N, by Corollary 2.5 we also have $U \in \widehat{\mathcal{O}}\langle X'\rangle[X_N]$. The degree of ε in X_N is $\leq rd$, and the degree of W in X_N is s; hence U is of degree $\leq rd - s$ in X_N. Moreover,

$$W(Uh_i) = \varepsilon h_i = -(d_{i,r+1}R_1 + \cdots + d_{in}R_{n-r}) \in \widehat{\mathcal{O}}\langle X'\rangle[X_N] \tag{4.5}$$

for $i = 1, \ldots, r$. Since W is monic in X_N, it follows that $Uh_i \in \widehat{\mathcal{O}}\langle X'\rangle[X_N]$. Put $z = Uz' \in \big(\widehat{\mathcal{O}}\langle X'\rangle[X_N]\big)^n$, a solution to ($\mathrm{S}_e$). We claim that all entries of z have X_N-degree $< rd$: To see this, note that $\deg_{X_N} R_i < s$ for $i = 1, \ldots, n-r$ and $\deg_{X_N} U \leq d - s$; hence the last $n - r$ entries $UR_1, \ldots, UR_{n-r}$ of z are of X_N-degree $< rd$. For the first r entries $Uh_1, \ldots, Uh_r$ use that the right-hand side of (4.5) has X_N-degree $< rd + s$; since $\deg_{X_N} W = s$, we get $\deg_{X_N} Uh_i < rd$ for all $i = 1, \ldots, r$. It follows that z', and hence z, is an $\widehat{\mathcal{O}}\langle X\rangle$-linear combination of $y^{(1)}, \ldots, y^{(K')}$. Since U is a unit in $\widehat{\mathcal{O}}\langle X\rangle$, the solution y can be expressed as an $\widehat{\mathcal{O}}\langle X\rangle$-linear combination of the column vectors

$$w^{(1)}, \ldots, w^{(n-r)}, y^{(1)}, \ldots, y^{(M')} \in \big(\mathcal{O}[X]\big)^n.$$

Hence the solution $x = T_e^{-1}(y)$ to our original equation (I) is an $\widehat{\mathcal{O}}\langle X\rangle$-linear combination of the vectors in (4.4) as claimed. □

Remark. We can bound the degrees of the solutions in (4.4): We have $\deg u^{(k)} \leq r \deg A$ for $k = 1, \ldots, n-r$ and $\deg_{X'} T_e^{-1}\big(y^{(i)}\big) \leq \deg_{X'} y^{(i)}$ for $i = 1, \ldots, M'$. Moreover $\deg_{X_N} y^{(l)} < rd < re^N$ and thus

$$\deg_{X_N} T_e^{-1}\big(y^{(i)}\big) \leq e^{N-1} \deg y^{(i)} \leq e^{N-1}\big(\deg_{X'} y^{(i)} + re^N\big)$$

for $i = 1, \ldots, M'$.

Starting with (I) we successively obtain equivalent homogeneous matrix equations

$$A^{(N)}y^{(N)} = 0 \tag{H_N}$$

$$\vdots$$

$$A^{(\nu)}y^{(\nu)} = 0 \tag{H_ν}$$

$$\vdots$$

$$A^{(0)}y^{(0)} = 0, \tag{H_0}$$

where $0 \leq \nu \leq N$, $A^{(\nu)}$ is an $m(\nu) \times n(\nu)$-matrix with entries in the polynomial ring $\mathcal{O}[X_1, \dots, X_\nu]$ and

$$y^{(\nu)} = \Big[y_1^{(\nu)}, \dots, y_{n(\nu)}^{(\nu)}\Big]^{\mathrm{tr}}$$

is a vector of unknowns ranging over $\widehat{\mathcal{O}}\langle X_1, \dots, X_\nu\rangle$. So the *initial* equation (H_N) is just $Ay = 0$, and if $\nu > 0$, then the system $(\mathrm{H}_{\nu-1})$ is obtained from (H_ν) by the procedure described above (passage from A to A'). We have

$$m(\nu) \leq m(\nu+1)\big(m(\nu+1)+1\big)e(\nu+1)^\nu$$

for all $\nu = 0, \dots, N-1$, where $e(\nu) = m(\nu)\deg A^{(\nu)} + 1$. It follows that

$$e(\nu) \leq m(\nu+1)\big(m(\nu+1)+1\big)e(\nu+1)^\nu \deg A^{(\nu)} + 1.$$

Using that $\deg A^{(\nu)} \leq \deg A$, we get the estimate

$$e(\nu) \leq (m \deg A + 1)^{(N+1)^{N-\nu}} \tag{4.6}$$

for all $\nu = 0, \dots, N$. Let $\mathcal{B}(0) \subseteq \mathcal{O}^{n(0)}$ be a finite system of generators of $\mathrm{Sol}_{\mathcal{O}}(A^{(0)})$, and for every $\nu = 1, \dots, N$ let $\mathcal{B}(\nu) \subseteq \mathcal{O}[X_1, \dots, X_\nu]^{n(\nu)}$ be a system of generators for the module of solutions to (H_ν) in $\widehat{\mathcal{O}}\langle X_1, \dots, X_\nu\rangle$, with $\mathcal{B}(\nu)$ constructed from $\mathcal{B}(\nu-1)$ according to the process described above. For $\nu = 0, \dots, N$ let $\gamma(\nu)$ be the maximal degree of an element of $\mathcal{B}(\nu)$. Clearly $\gamma(0) = 0$, and by the remark following Lemma 4.4 we have

$$\gamma(\nu) \leq e(\nu)^{\nu-1}\big(\gamma(\nu-1) + m(\nu)e(\nu)^\nu\big) + \gamma(\nu-1).$$

The right-hand side can be further estimated from above by

$$e(\nu)^{\nu-1}\big(2\gamma(\nu-1) + m(\nu)e(\nu)^\nu\big) \leq e(\nu)^{2\nu-1}\big(\gamma(\nu-1) + m(\nu)\big).$$

Hence we get

$$\gamma(\nu) + 1 \leq e(\nu)^{2\nu}\big(\gamma(\nu-1) + 1\big)$$

for all $\nu = 1, \dots, N$. It follows that

$$\gamma(N) + 1 \leq e(N)^{2N}e(N-1)^{2(N-1)} \cdots e(1)^2,$$

and hence, using (4.6):

$$\gamma(N) \leq \big(m \deg A + 1\big)^{\varrho}$$

where $\varrho = 2\sum_{i=0}^{N-1}(N+1)^i(N-i)$. It is easy to see that $\varrho \leq 2\big((N+1)^N - 1\big)$. Hence every element of $\mathcal{B}(N)$ has degree $\leq (2m \deg A)^{2((N+1)^N-1)}$, finishing the proof of Proposition 4.3 and thus of Theorem 4.1. $\square$

Remark 4.5. As a consequence of Theorem 4.1, if R is an almost Dedekind domain that is syzygy-solvable, then there exists an (impractical) algorithm which, given an $m \times n$-matrix A with entries in $D = R[X]$, constructs a finite collection of generators for $\mathrm{Sol}_D(A)$. If $R = \mathbb{Z}$, or more generally, a computable principal ideal domain, we can also turn the proof of the theorem into such an algorithm: We first find generators $u^{(1)}, \ldots, u^{(K)} \in M := \mathrm{Sol}_D(A)$ for $\mathrm{Sol}_{F[X]}(A) = MF[X]$, where $F = \mathrm{Frac}(R)$, and $\delta \in R$ such that $\mathrm{Sol}_D(A) = (M' : \delta)$ where $M' = Du^{(1)} + \cdots + Du^{(K)}$. (See Corollary 3.6.) We may assume $A \neq 0$; hence $\delta \neq 0$. For every prime factor π of δ we now follow the inductive procedure outlined in the proof of Proposition 4.3 to construct generators $v_\pi^{(1)}, \ldots, v_\pi^{(K_\pi)} \in \mathrm{Sol}_D(A)$ for $\mathrm{Sol}_{\widehat{R_{(\pi)}\langle X\rangle}}(A)$. By the remark following Lemma 4.2, the solutions $u^{(1)}, \ldots, u^{(K)}$ together with the $v_\pi^{(1)}, \ldots, v_\pi^{(K_\pi)}$ (with π ranging over the prime factors of δ) generate $\mathrm{Sol}_D(A) = MF[X] \cap D^n$.

Sometimes Theorem 4.1 still holds for rings which are not almost hereditary:

Corollary 4.6. *Let R be an integrally closed almost Dedekind domain, and let S be the integral closure of R inside an algebraic closure of the fraction field F of R. Let A be an $m \times n$-matrix with entries in $S[X] = S[X_1, \ldots, X_N]$. Then* $\mathrm{Sol}_{S[X]}(A)$ *is generated by elements of degree at most* $(2m \deg A)^{2((N+1)^N - 1)}$.

Proof. Let F' be a finite field extension of F containing all the coefficients of the entries of A, and let R' be the integral closure of R in F'. Then R' is almost Dedekind. (See [15, (36.1)].) Since R' is a Prüfer domain and S a torsion-free R'-module, S is flat over R'. The claim now follows from Theorem 4.1. □

The corollary applies to $R = \mathbb{Z}$ (so $S =$ the ring of all algebraic integers).

Application 1: Bounds for module-theoretic operations. Let R be an almost Dedekind domain with fraction field $F = \mathrm{Frac}(R)$. We can exploit Theorem 4.1 to establish bounds for some basic operations on finitely generated submodules of free modules over $D = R[X] = R[X_1, \ldots, X_N]$. We say that a finitely generated D-submodule of D^m is **of type** d (where $d \in \mathbb{N}$) if it is generated by vectors of degree at most d.

Proposition 4.7. *Let M and M' be finitely generated submodules of the free D-module D^m of type d. Then the D-modules $(M : \delta)$ (where $\delta \in R$), $MF[X] \cap D^m$ and $M \cap M'$ as well as the ideal $(M' : M)$ are of type $\tau(N, d, m) = (2md)^{(N+1)^{O(N)}}$.*

Proof. Let $M = Dv^{(1)} + \cdots + Dv^{(n)}$ and $M' = Dw^{(1)} + \cdots + Dw^{(p)}$ with $v^{(i)}, w^{(j)} \in D^m$ of degree at most d. To find generators for the D-module $(M : \delta)$, we first find a finite set of generators $z^{(1)}, \ldots, z^{(K)} \in D^{n+m}$ for the D-module of solutions to the system of homogeneous equations

$$v^{(1)}y_1 + \cdots + v^{(n)}y_n + (-\delta e^{(1)})y_{n+1} + \cdots + (-\delta e^{(m)})y_{n+m} = 0.$$

Here $e^{(1)}, \ldots, e^{(m)}$ denote the unit vectors in D^m. Then clearly the K vectors consisting of the last m entries of $z^{(1)}, \ldots, z^{(K)}$ generate $(M : \delta)$. By Theorem 4.1 $(M : \delta)$ is of type $\tau(N, d, m)$. Using Corollary 3.5, this implies that $MF[X] \cap D^m$ is also of type $\tau(N, d, m)$.

In order to find generators for $M \cap M'$, it suffices to find generators for the D-module of solutions to the system of homogeneous equations

$$v^{(1)}y_1 + \cdots + v^{(n)}y_n = w^{(1)}y_{n+1} + \cdots + w^{(p)}y_{n+p}.$$

Moreover we have

$$(M' : M) = (M' : Dv^{(1)}) \cap \cdots \cap (M' : Dv^{(n)}),$$

and if $u^{(1)}, \dots, u^{(q)} \in D^m$ generate $M' \cap Dv$, where $v = [v_1, \dots, v_m]^{\mathrm{tr}} \in D^m$, then

$$(M' : Dv) = \bigcap_{j=1}^{m} \big(u_j^{(1)}/v_j, \dots, u_j^{(q)}/v_j\big).$$

Here $a/0 := 1$ for all $a \in R$. From this, one easily obtains the bounds on the type of $M \cap M'$ and $(M' : M)$ as claimed. □

Remarks.

(1) If R is syzygy-solvable, then generators for the D-modules $MF[X] \cap D^m$ and $M \cap M'$ and for the ideal $(M' : M)$ can be computed elementary recursively (in the basic operations of R) from given generators for M and M'. This follows from the proof of the proposition and Remark 4.5 above.
(2) By Corollary 4.6, the proposition remains true if R is replaced by the ring of algebraic integers.

Application 2: A criterion for primeness. The following lemma is well known; we leave the proof to the reader.

Lemma 4.8. *Let R be a ring and let I be an ideal of $R[X]$. Then I is prime if and only if the image of I in $(R/I \cap R)[X]$ is prime. If R is an integral domain with fraction field F and $I \cap R = (0)$, then I is prime if and only if $IF[X]$ is prime and $IF[X] \cap R[X] = I$.*

As a consequence, we obtain a test for primeness of an ideal in $\mathbb{Z}[X]$. Let $I = (f_1, \dots, f_n)$ with $f_1, \dots, f_n \in \mathbb{Z}[X]$ and $\delta \in \mathbb{Z}$ such that $I\mathbb{Q}[X] \cap \mathbb{Z}[X] = (I : \delta)$.

Corollary 4.9. *The ideal $I = (f_1, \dots, f_n)$ is prime if and only if one of the following holds:*

(1) *$I\mathbb{Q}[X]$ is prime and $(I : \delta) = I$, or*
(2) *there exists a prime factor p of δ such that $p \in I$ and the image of I in $\mathbb{F}_p[X]$ is a prime ideal.*

Combining Corollary 4.9 with Proposition 4.7 and a result from [34], we get a criterion for the primeness of an ideal of $\mathbb{Z}[X]$ which is polynomial in the degrees of the generators:

Proposition 4.10. *There exists $\varrho = \varrho(N) \in \mathbb{N}$ such that for each ideal I of $\mathbb{Z}[X]$ of type d, the following is true: I is prime if and only if $1 \notin I$, and for all $f, g \in \mathbb{Z}[X]$ of degree $\leq d^{\varrho}$, if $fg \in I$, then $f \in I$ or $g \in I$.*

Proof. By Proposition 4.7, for all ideals I of $\mathbb{Z}[X]$ of type d, the ideal $I\mathbb{Q}[X] \cap \mathbb{Z}[X]$ of $\mathbb{Z}[X]$ is of type $\tau(N, d) = (2d)^{(N+1)^{O(N)}}$. Moreover by [34] there exists $\varrho' = \varrho'(N) \in \mathbb{N}$ such that for each field F and each ideal J of $F[X]$ of type d, we have: J is prime if and only if $1 \notin J$, and for all $f, g \in F[X]$ of degree $\leq d^{\varrho'}$, if $fg \in J$, then $f \in J$ or $g \in J$. We claim that ϱ with $d^{\varrho} \geq \max\{\tau, d^{\varrho'}\}$ has the required properties. For this, let $I = (f_1, \dots, f_n)$ be an ideal of $\mathbb{Z}[X]$ of type d, and let $\delta \in \mathbb{Z}$ with $I\mathbb{Q}[X] \cap \mathbb{Z}[X] = (I : \delta)$. Suppose $1 \notin I$ and $fg \in I \Rightarrow f \in I$ or $g \in I$, for all $f, g \in \mathbb{Z}[X]$ of degree $\leq d^{\varrho}$. Then either $1 \in I\mathbb{Q}[X]$ or $I\mathbb{Q}[X]$ is prime, since $d^{\varrho} \geq d^{\varrho'}$. Similarly, if we have $p \in I$ for some prime divisor p of δ, then the image

of I in $\mathbb{F}_p[X]$ is a prime ideal. In this case, it follows by Corollary 4.9 (2) that I is prime. Otherwise, $f \in I$ for all $f \in \mathbb{Z}[X]$ of degree at most τ with $\delta f \in I$. Hence $1 \notin I\mathbb{Q}[X]$ (so $I\mathbb{Q}[X]$ is prime) and $(I : \delta) = I$. By Corollary 4.9 (1) it follows that also in this case I is prime as desired. □

It is clear that Corollary 4.9 and Proposition 4.10 hold, mutatis mutandis, for any PID R with fraction field F in place of $\mathbb{Z}$ and $\mathbb{Q}$, respectively.

5. Height bounds

Throughout this section we let F be a number field and $R = \mathcal{O}_F$ the ring of integers of F. Let $A = (a_{ij})$ be a non-zero $m \times n$-matrix with entries a_{ij} in $D = R[X] = R[X_1, \ldots, X_N]$. Let $d = \deg A$ and $h = h(A)$. As was shown in the previous section, we can explicitly bound the degrees of generators for the D-module $\mathrm{Sol}_D(A)$ in terms of d, m and N. We now want to bound the heights of those generators in a similar fashion (in terms of N, d, h, and m).

The local case. Let $\mathfrak{p} \neq 0$ be a prime ideal of R, and let $\mathcal{O} := R_{\mathfrak{p}}$ (a DVR). We first investigate the height of generators for $\mathrm{Sol}_{\mathcal{O}[X]}(A)$ and begin with the case $N = 0$:

Lemma 5.1. *Suppose that $a_{ij} \in R$ for all i, j, and let $r = \mathrm{rank}_F(A)$. The $\mathcal{O}$-module $\mathrm{Sol}_{\mathcal{O}}(A)$ of solutions in $\mathcal{O}^n$ to the system of homogeneous linear equations $Ay = 0$ is generated by $n - r$ many vectors whose height is bounded by*

$$C_2 \cdot r(h + \log r + 1).$$

Here C_2 is a constant only depending on F.

Proof. Let $v \in M_F$ denote the place of F associated with $\mathfrak{p}$, so $\mathfrak{p} = \mathfrak{p}_v$. We may assume that $\det \Delta \neq 0$, where $\Delta = (a_{ij})_{1 \leq i,j \leq r}$ (after permuting the unknowns in our system $Ay = 0$ if necessary). In fact, we may assume that the $\mathfrak{p}$-adic valuation $\mu := v_{\mathfrak{p}}(\det \Delta)$ of $\det \Delta$ is minimal among all $r \times r$-submatrices of A; cf. the proof of Proposition 4.3. Now $Ay = 0$ has the same solutions in any domain extending $\mathcal{O}$ as the system (S) obtained from $Ay = 0$ by multiplying both sides from the left with the adjoint of Δ (see Section 3). The entries $\delta = \det \Delta$ and c_{ij} $(1 \leq i \leq r < j \leq n)$ of the coefficient matrix of (S) are certain signed $r \times r$-minors of A. Let $v^{(1)}, \ldots, v^{(n-r)}$ be the $n - r$ linearly independent solutions to $Ay = 0$ listed in (3.1). By (1.7) we have $h\big(v^{(k)}\big) \leq r(h + \log r)$ for $k = 1, \ldots, n - r$. Corollary 1.5 implies that there exists an element b of F such that $v_{\mathfrak{p}}(b) = -\mu$, $bv^{(k)} \in R^n$ for all $k = 1, \ldots, n - r$, and $h(b) \leq C_1 r(h + \log r + 1)$. Here $C_1 > 0$ is a constant which only depends on F. The vectors $bv^{(1)}, \ldots, bv^{(n-r)} \in R^n$ generate $\mathrm{Sol}_{\mathcal{O}}(A)$ and are bounded in height by $C_2 r(h + \log r + 1)$, with $C_2 = 2C_1$. □

Remark. Note that the vectors $v^{(1)}, \ldots, v^{(n-r)} \in \mathrm{Sol}_R(A)$ as in the proof of the lemma generate $\mathrm{Sol}_F(A)$ and satisfy $h(v^{(k)}) \leq r(h + \log r)$ for $k = 1, \ldots, n - r$. Moreover, the element $0 \neq \delta \in R$ of height $h(\delta) \leq r(h + \log r)$ has the property that $\delta y \in Rv^{(1)} + \cdots + Rv^{(n-r)}$ for every $y \in \mathrm{Sol}_R(A)$.

We now consider the general case $N \geq 0$. By Proposition 4.3, the solution module $\mathrm{Sol}_{\mathcal{O}[X]}(A)$ is generated by solutions $y = [y_1, \ldots, y_n]^{\mathrm{tr}}$ with $y_1, \ldots, y_n \in \mathcal{O}[X]$ of

degree at most $\gamma = \gamma(N, d, m) := (2md)^{2((N+1)^N - 1)}$. Write

$$y_j = \sum_{|\nu| \leq \gamma} y_{j,\nu} X^\nu$$

with variables $y_{j,\nu}$ ranging over $\mathcal{O}$ and

$$a_{ij} = \sum_{|\mu| \leq d} a_{ij,\mu} X^\mu$$

with $a_{ij,\mu} \in \mathcal{O}$, where $1 \leq i \leq m$, $1 \leq j \leq n$ and $\nu, \mu \in \mathbb{N}^N$, $|\nu| \leq \gamma$, $|\mu| \leq d$. A polynomial in $X_1, \ldots, X_N$ of degree at most d has at most $M(N, d) = \binom{N+d}{N}$ monomials. Hence the solutions (in $\mathcal{O}[X]$) of every equation

$$a_{i1} y_1 + \cdots + a_{in} y_n = 0 \qquad (1 \leq i \leq m)$$

are in one-to-one correspondence with the solutions (in $\mathcal{O}$) of the system consisting of the $M(N, \gamma + d)$ homogeneous equations

$$\sum_j \sum_{\mu+\nu=\lambda} a_{ij,\mu} y_{j,\nu} = 0 \qquad (|\lambda| \leq \gamma + d)$$

in the $n \cdot M(N, \gamma)$ many variables $y_{j,\nu}$, with coefficients in $\mathcal{O}$. So the entire system $Ay = 0$, with coefficients in $\mathcal{O}[X]$, may be replaced by a certain homogeneous system of $m \cdot M(N, \gamma + d)$ equations in the variables $y_{j,\nu}$, having coefficients in $\mathcal{O}$. Applying the lemma above to the new system and using the estimate

$$m \cdot M(N, \gamma + d) \leq m \cdot (\gamma + d + 1)^N = \big(2m(d+1)\big)^{(N+1)^{O(N)}},$$

we get the following result, with C_2 as above.

Proposition 5.2. *For any $N \geq 0$, the $\mathcal{O}[X]$-module $\mathrm{Sol}_{\mathcal{O}[X]}(A)$ is generated by solutions of degree at most $(2md)^{(N+1)^{O(N)}}$ and height at most*

$$C_2 \cdot \big(2m(d+1)\big)^{(N+1)^{O(N)}} (h+1). \tag{5.1}$$

Here C_2 is a constant only depending on F. □

With $\beta = \beta(N, m, d) = (2md)^{2^N}$ we have

$$m \cdot M(N, \beta + d) \leq m \cdot (\beta + d + 1)^N = \big(2m(d+1)\big)^{2^{O(N+1)}}.$$

Using this estimate as well as Theorem 3.2 (in place of Proposition 4.3) and the remark following Lemma 5.1, one obtains a result similar to Proposition 5.2:

Lemma 5.3. *The $F[X]$-module $\mathrm{Sol}_{F[X]}(A)$ is generated by vectors $u^{(1)}, \ldots, u^{(K)} \in \mathrm{Sol}_{\mathcal{O}[X]}(A)$ of degree at most $(2md)^{2^N}$ and height at most*

$$\big(2m(d+1)\big)^{2^{O(N+1)}} (h+1). \tag{5.2}$$

Moreover, there exists a non-zero $\delta \in R$ of height bounded by (5.2) such that $\delta y \in Du^{(1)} + \cdots + Du^{(K)}$ for every $y \in \mathrm{Sol}_D(A)$. □

The global case. Proposition 5.2 and Lemma 5.3 now imply the existence of generators of $\mathrm{Sol}_D(A)$ of small height:

Corollary 5.4. *The D-module $\mathrm{Sol}_D(A)$ can be generated by solutions of degree at most $(2md)^{(N+1)^{O(N)}}$ and of height at most*

$$C_2 \cdot \big(2m(d+1)\big)^{(N+1)^{O(N)}}(h+1).$$

Here C_2 is a constant only depending on F.

Proof. By Lemma 5.3 we find $u^{(1)}, \dots, u^{(K)} \in M := \mathrm{Sol}_D(A)$ with the following properties: $u^{(1)}, \dots, u^{(K)}$ generate the $F[X]$-module $\mathrm{Sol}_{F[X]}(A) = MF[X]$, $\deg u^{(k)} \le (2md)^{2^N}$ for all k, and $h(u^{(k)})$ is bounded by (5.2), for each k. Moreover we find an element $0 \neq \delta \in R$ of height bounded by (5.2) such that $\mathrm{Sol}_D(A) = (M' : \delta)$ with $M' = Du^{(1)} + \cdots + Du^{(K)}$. For every maximal ideal $\mathfrak{m}$ of R we find generators $v_{\mathfrak{m}}^{(1)}, \dots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})} \in \mathrm{Sol}_D(A)$ of $MR_{\mathfrak{m}}[X] = \mathrm{Sol}_{R_{\mathfrak{m}}[X]}(A)$ having degree at most $(2md)^{(N+1)^{O(N)}}$ and height bounded by (5.1). By the remark following Lemma 4.2, the vectors $u^{(1)}, \dots, u^{(K)}$ and $v_{\mathfrak{m}}^{(1)}, \dots, v_{\mathfrak{m}}^{(K_{\mathfrak{m}})}$, where $\mathfrak{m}$ ranges over all maximal ideals of R containing δ, generate $\mathrm{Sol}_D(A) = MF[X] \cap R[X]^n$. □

Remark. The number of generators of $\mathrm{Sol}_D(A)$ can be bounded in a similar way: If δ is a unit in R, then $u^{(1)}, \dots, u^{(K)}$ generate $\mathrm{Sol}_D(A)$, and $K \le n \cdot M(N, \beta + d) = n\big(2m(d+1)\big)^{2^{O(N+1)}}$. In general, by the remark after Lemma 1.3, there are at most $[F : \mathbb{Q}] \cdot h(\delta)/\log 2$ many maximal ideals of R containing δ. So we have at most

$$n \cdot M(N, \gamma + d) \cdot \big(1 + [F : \mathbb{Q}] \cdot h(\delta)/\log 2\big) = n \cdot [F : \mathbb{Q}] \cdot \big(2m(d+1)\big)^{(N+1)^{O(N)}}(h+1)$$

generators in total.

6. Ideal membership

In this section we use the results obtained so far to give a proof of Theorem A from the introduction. We begin by studying ideal membership problems of a special form.

Bézout identities. Let R be a ring, $f_1, \dots, f_n \in R[X]$, and $d = \max_i \deg f_i$. We call a representation of 1 as a linear combination

$$1 = f_1 g_1 + \cdots + f_n g_n \tag{6.1}$$

of $f_1, \dots, f_n$ with coefficients $g_1, \dots, g_n \in R[X]$ a **Bézout identity** for $f_1, \dots, f_n$ in $R[X]$. If $R = F$ is a field, then from Hermann's Theorem 3.4 it follows that $1 \in (f_1, \dots, f_n)F[X]$ if and only if there exist $g_1, \dots, g_n \in F[X]$ of degree $\le (2d)^{2^N}$ satisfying the Bézout identity (6.1). By the effective version of Hilbert's Nullstellensatz due to Kollár [22], this bound may be improved substantially: if $1 \in (f_1, \dots, f_n)F[X]$, then there are $g_1, \dots, g_n \in F[X]$ of degrees $\le (3d)^N$ satisfying (6.1). For $F = \mathbb{Q}$ this means: if $1 \in (f_1, \dots, f_n)\mathbb{Q}[X]$, then there are $\delta \in \mathbb{Z} \setminus \{0\}$ and $g_1, \dots, g_n \in \mathbb{Z}[X]$ of degree $\le (3d)^N$ with

$$\delta = f_1 g_1 + \cdots + f_n g_n.$$

We have the following bound for the size of δ, obtained along the lines of Lemma 5.3 (i.e., Cramer's rule). From now on, F denotes a number field.

Lemma 6.1. *Suppose that $R = \mathcal{O}_F$ is the ring of integers of F. If we have $1 \in (f_1, \dots, f_n)F[X]$, then*

$$\delta = f_1 g_1 + \cdots + f_n g_n$$

for some $g_1, \dots, g_n \in R[X]$ of degree $\leq (3d)^N$ and some $\delta \in R$, $\delta \neq 0$, of height at most

$$\big(2(d+1)\big)^{O(N^2+1)}\big(h(f_1, \dots, f_n) + 1\big).$$

We now want to show that Kollár's degree bound over fields entails a similar bound for Bézout identities over rings of integers.

Proposition 6.2. *Suppose that $R = \mathcal{O}_F$. If $1 \in (f_1, \dots, f_n)$, then there exist $g_1, \dots, g_n \in R[X]$ with*

$$1 = f_1 h_1 + \cdots + f_n h_n$$

and

$$\deg h_i \leq [F : \mathbb{Q}] \cdot (3d)^{O(N^2)}\big(h_1(f_1, \dots, f_n) + 1\big)$$

for all $i = 1, \dots, n$.

Before we begin with the proof, we state an elementary lemma whose proof is left to the reader:

Lemma 6.3. *Let $U = (U_1, \dots, U_n)$, $V = (V_1, \dots, V_n)$ be tuples of pairwise distinct indeterminates over $\mathbb{Z}$, and let $e \geq 1$ be an integer. There exist polynomials*

$$g_1^{(e)}(U, V), \dots, g_n^{(e)}(U, V)$$

with non-negative integer coefficients such that

$$\big(1 + U_1 V_1 + \cdots + U_n V_n\big)^e = 1 + g_1^{(e)}(U, V)U_1 + \cdots + g_n^{(e)}(U, V)U_n \tag{6.2}$$

and $\deg_U g_j^{(e)} = e - 1$, $\deg_V g_j^{(e)} = e$.

We first show a local analogue of Proposition 6.2:

Lemma 6.4. *Suppose that $R = (\mathcal{O}_F)_{\mathfrak{p}}$, where $\mathfrak{p} \neq 0$ is a prime ideal of $\mathcal{O}_F$. If $1 \in (f_1, \dots, f_n)$, then*

$$1 = f_1 h_1 + \cdots + f_n h_n$$

for some $h_1, \dots, h_n \in R[X]$ of degree at most

$$[F : \mathbb{Q}] \cdot (3d)^{O(N^2)}\big(h(f_1, \dots, f_n) + 1\big)/\log p. \tag{6.3}$$

Here p is the unique prime number such that $\mathbb{Z} \cap \mathfrak{p} = p\mathbb{Z}$.

Proof. Suppose $1 \in (f_1, \dots, f_n)$. Then $1 \in (f_1, \dots, f_n)F[X]$; hence by Lemma 6.1 there exist $g_1, \dots, g_n \in R[X]$ of degree at most $(3d)^N$ and a non-zero $\delta \in R$ of height at most $\big(2(d+1)\big)^{O(N^2)}(h+1)$ such that

$$\delta = f_1 g_1 + \cdots + f_n g_n. \tag{6.4}$$

Here and below $h = h(f_1, \dots, f_n)$. If δ is a unit in R, then

$$1 = f_1(g_1/\delta) + \cdots + f_n(g_n/\delta)$$

is a Bézout identity for $f_1, \dots, f_n$ in $R[X]$, and $h_i := g_i/\delta$, $i = 1, \dots, n$, have the required properties. Suppose that δ is not a unit, so $e = v_{\mathfrak{p}}(\delta) \geq 1$. We have

$1 \in (\overline{f_1}, \dots, \overline{f_n})$, where $\overline{f}$ is the canonical image of $f \in R[X]$ in $(R/\mathfrak{p}R)[X]$. By Kollár's theorem [22] applied to the field $R/\mathfrak{p}R$, there exist $r_1, \dots, r_n \in R[X]$ with

$$1 - (r_1 f_1 + \cdots + r_n f_n) \in \mathfrak{p}R[X]$$

and $\deg r_j \leq (3d)^N$ for all $j = 1, \dots, n$. Specializing the U_i's to $(f_1, \dots, f_n)$ and the V_i's to $(-r_1, \dots, -r_n)$ in (6.2) gives $s_1, \dots, s_n \in R[X]$ and $s \in \mathfrak{p}^e R[X]$ such that

$$1 - (f_1 s_1 + \cdots + f_n s_n) = s. \tag{6.5}$$

We have $\deg s_j \leq e\big(d + (3d)^N\big) - d$ for all j; hence $\deg s \leq e\big(d + (3d)^N\big)$. From (6.4) and (6.5) we get

$$1 = f_1 s_1 + \cdots + f_n s_n + s = f_1 h_1 + \cdots + f_n h_n$$

with $h_j = s_j + (s/\delta) g_j \in R[X]$. We have

$$\deg(s g_j) \leq e\big(d + (3d)^N\big) + (3d)^N \leq e(3d)^{N+1},$$

and since $\big(2(d+1)\big)^{O(N^2+1)} = (3d)^{O(N^2)}$, we get

$$e \cdot \log p \leq [F : \mathbb{Q}] \cdot h(\delta) = [F : \mathbb{Q}] \cdot (3d)^{O(N^2)}(h+1)$$

by the remarks following Lemma 1.3, for $N > 0$, $d > 0$. It follows that $\deg h_j$ is bounded from above by (6.3), for $j = 1, \dots, n$. □

Now suppose that $R = \mathcal{O}_F$, and assume that $1 \in (f_1, \dots, f_n)$. Hence by Lemma 6.1 there are $g_1, \dots, g_n \in R[X]$ of degree at most $(3d)^N$ and a non-zero $\delta \in R$ of height at most $\big(2(d+1)\big)^{O(N^2+1)}(h+1)$ such that

$$\delta = f_1 g_1 + \cdots + f_n g_n.$$

For every prime ideal $\mathfrak{p}$ of R containing δ, and p the prime number generating the ideal $\mathbb{Z} \cap \mathfrak{p}$, we find $h_1^{(\mathfrak{p})}, \dots, h_n^{(\mathfrak{p})} \in R[X]$ of degree bounded by (6.3) as well as $\delta^{(\mathfrak{p})} \in R \setminus \mathfrak{p}$ such that

$$\delta^{(\mathfrak{p})} = f_1 h_1^{(\mathfrak{p})} + \cdots + f_n h_n^{(\mathfrak{p})}.$$

Let $\mathfrak{p}_1, \dots, \mathfrak{p}_K$ be the pairwise distinct prime ideals of R containing δ. Then there exist $a, a_1, \dots, a_K \in R$ such that

$$1 = a\delta + a_1 \delta^{(\mathfrak{p}_1)} + \cdots + a_K \delta^{(\mathfrak{p}_K)}.$$

Hence, letting $h_j = a g_j + a_1 h_j^{(\mathfrak{p}_1)} + \cdots + a_K h_j^{(\mathfrak{p}_K)} \in R[X]$ for $j = 1, \dots, n$, we get

$$\begin{aligned} f_1 h_1 + \cdots + f_n h_n &= a(f_1 g_1 + \cdots + f_n g_n) + \sum_{k=1}^{K} a_k \big(f_1 h_1^{(\mathfrak{p}_k)} + \cdots + f_n h_n^{(\mathfrak{p}_k)}\big) \\ &= a\delta + \sum_{k=1}^{K} a_k \delta^{(\mathfrak{p}_k)} = 1. \end{aligned}$$

From this Proposition 6.2 follows. □

Remark. By [23, Theorem 3.6], the height of the denominator δ in Lemma 6.1 can be bounded in terms of N, d, n and the height of $f_1, \dots, f_n$ by a bound which is

single-exponential in d and linear in $h(f_1, \dots, f_n)$, while at the same time retaining a single-exponential bound on the degrees of the g_j:

$$\deg g_i \leq 4Nd^N,$$
$$h(\delta, g_1, \dots, g_n) \leq 4N(N+1)d^N\big(h(f_1, \dots, f_n) + \log n + (N+7)\log(N+1)d\big).$$

This leads to the following improved degree bound in Proposition 6.2:

$$\deg h_i \leq [F:\mathbb{Q}] \cdot (N+1)^{O(1)} d^{2N}\big(h(f_1, \dots, f_n) + \log n + d\big) \qquad \text{for all } i.$$

In most cases this is much more precise. We decided to use the cruder bound on $h(\delta)$ in Lemma 6.1 and the ensuing degree bounds in Proposition 6.2, since they are independent of n. In the case $\mathcal{O}_F = \mathbb{Z}$ one could have also used Philippon's estimate [30] (without dependence on n)

$$\deg g_i \leq (N+2)d^N, \qquad h(\delta) \leq c(N) \cdot d^N\big(h(f_1, \dots, f_n) + 1\big),$$

where $c(N)$ depends exponentially on N.

Ideal membership. In the following we let $R = \mathcal{O}_F$ for a number field F. Let A be an $m \times n$-matrix with entries in $R[X]$ and let $b \in \big(R[X]\big)^m$ be a column vector.

Theorem 6.5. *If the system $Ay = b$ has a solution in $D = R[X]$, then it has such a solution of degree at most*

$$[F:\mathbb{Q}] \cdot C_2 \cdot \big(2m \deg(A,b)\big)^{(N+1)^{O(N)}} \cdot \big(h(A,b) + 1\big). \tag{6.6}$$

Here the constant C_2 depends only on F.

Proof. Put $d = \deg(A,b)$ and $h = h(A,b)$. By Corollary 5.4 there exist generators $z^{(1)}, \dots, z^{(K)}$ for the D-module of solutions to the system of homogeneous linear equations $(A, -b)z = 0$, where z is a vector of $n+1$ unknowns $z_1, \dots, z_{n+1}$, with

$$\deg\big(z^{(k)}\big) = (2md)^{(N+1)^{O(N)}}, \tag{6.7}$$

$$h\big(z^{(k)}\big) = C_2 \cdot \big(2m(d+1)\big)^{(N+1)^{O(N)}}(h+1) \tag{6.8}$$

for all $k = 1, \dots, K$. The constant C_2 only depends on the number field F. For each k let $z^{(k)}_{n+1} \in R[X]$ be the last component of $z^{(k)}$. Clearly, $Ay = b$ is solvable in $R[X]$ if and only if $1 \in \big(z^{(1)}_{n+1}, \dots, z^{(K)}_{n+1}\big)$. Moreover, if $h_1, \dots, h_K$ are elements of $R[X]$ such that

$$1 = h_1 z^{(1)}_{n+1} + \cdots + h_K z^{(K)}_{n+1},$$

then $y \in \big(R[X]\big)^n$ with

$$\begin{bmatrix} y \\ 1 \end{bmatrix} = h_1 z^{(1)} + \cdots + h_K z^{(K)}$$

is a solution to $Ay = b$. By Proposition 6.2 we find such $h_1, \dots, h_K$ with

$$\deg(h_k) \leq [F:\mathbb{Q}] \cdot \big(3 \max_l \deg\big(z^{(l)}\big)\big)^{O(N^2)}\big(\max_l h\big(z^{(l)}\big) + 1\big)$$

for all k. From (6.7) and (6.8) it follows that then the vector y has degree at most (6.6) as required. □

The doubly exponential degree bound on the solutions y in Theorem 6.5 implies a doubly exponential bound on $h(y)$. See [29] for good bounds on the height of solutions to linear equations over R.

For $m = 1$ the previous theorem yields:

Corollary 6.6. *Let $f_0, f_1, \dots, f_n \in R[X]$, and put $d = \deg(f_1, \dots, f_n)$, $h = h(f_1, \dots, f_n)$. If $f_0 \in (f_1, \dots, f_n)$, then there exist $g_1, \dots, g_n \in R[X]$ with*

$$f_0 = g_1 f_1 + \cdots + g_n f_n$$

and

$$\deg(g_1, \dots, g_n) \leq [F : \mathbb{Q}] \cdot C_2 \cdot (2d)^{(N+1)^{O(N)}}.$$

The constant C_2 depends only on F. □

Using the criterion for primeness of ideals in $\mathbb{Z}[X]$ given in Corollary 4.9, we get:

Corollary 6.7. *One can test elementary recursively whether finitely many given polynomials from $\mathbb{Z}[X]$ generate a prime ideal I.*

Proof. It is well known that the conditions "$I\mathbb{Q}[X]$ prime" and "$I \bmod p \subseteq \mathbb{F}_p[X]$ prime" (for a prime number p) can be tested elementary recursively [20], [35]. The condition "$I\mathbb{Q}[X] \cap \mathbb{Z}[X] = I$" may be tested elementary recursively using Proposition 4.7 and Corollary 6.6. □

See also [14] for an algorithm to test primeness of ideals in $\mathbb{Z}[X]$, which is however not even obviously primitive recursive.

Similarly to Theorem 6.5, using Proposition 5.2 and Lemma 6.4 in place of Corollary 5.4 and Proposition 6.2, respectively, one shows:

Theorem 6.8. *Let $\mathfrak{p}$ be a non-zero prime ideal of R and p the unique prime number with $\mathbb{Z} \cap \mathfrak{p} = p\mathbb{Z}$. If the system $Ay = b$ has a solution in $R_{\mathfrak{p}}[X]$, then it has such a solution of degree at most*

$$[F : \mathbb{Q}] \cdot C_2 \cdot \big(2m \deg(A, b)\big)^{(N+1)^{O(N)}} \cdot \big(h(A, b) + 1\big) / \log p.$$

Here C_2 is a constant which only depends on F. □

Final remarks. From Theorem 6.5 we obtain an effective reduction of the ideal membership problem for $R[X]$, where $R = \mathcal{O}_F$ for a number field F, to the solvability of a (huge) system of linear equations over R. As in the case of fields, this leads to a simple algorithm for deciding membership in ideals of $R[X]$. Certainly algorithms using Gröbner bases in $R[X]$ are much more effective in practice; it remains to establish doubly exponential degree and height bounds for the elements of Gröbner bases in $R[X]$. We plan to address this issue at a later point.

In [26], Mayr shows that ideal membership problems $f_0 \in (f_1, \dots, f_n)$ with $f_0, f_1, \dots, f_n \in \mathbb{Q}[X]$ can be decided by an algorithm which uses space that grows exponentially in the size of the input $f_0, \dots, f_n$. Together with [27] this establishes that ideal membership in $\mathbb{Q}[X]$ is exponential-space complete. The proof rests on an efficient parallel algorithm for computing the rank of $m \times m$-matrices over $\mathbb{Q}$ in time $O(\log^2 m)$ and the parallel computation thesis ("parallel time = sequential space"). By [27], ideal membership in $\mathbb{Z}[X]$ is exponential-space hard. Theorem 6.5 (and the reduction given in [26]) unfortunately only shows that ideal membership in $\mathbb{Z}[X]$ is exponential-space complete *provided* that solvability of systems of linear equations over $\mathbb{Z}$ can be decided using logarithmic space. However, this is even unknown for systems consisting of a single equation of the form $1 = ax + by$ $(a, b \in \mathbb{Z})$; see [18].

Acknowledgments

This paper is based on a part of the author's Ph.D. thesis [3] written under the direction of Lou van den Dries, whom he would like to thank for his guidance and advice. He is also grateful to Hendrik Lenstra and Bjorn Poonen for useful hints concerning the proof of Lemma 1.4 and to the referees for valuable suggestions.

References

1. M. Aschenbrenner, *Bounds and definability in polynomial rings*, submitted.
2. ———, *Kronecker's problem*, in preparation.
3. ———, *Ideal Membership in Polynomial Rings over the Integers*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2001.
4. C. Ayoub, *On constructing bases for ideals in polynomial rings over the integers*, J. Number Theory **17** (1983), no. 2, 204–225. MR 85m:13017
5. W. Baur, *Rekursive Algebren mit Kettenbedingungen*, Z. Math. Logik Grundlagen Math. **20** (1974), 37–46. MR 50:4269
6. T. Becker and V. Weispfenning, *Gröbner Bases*, Graduate Texts in Mathematics, vol. 141, Springer-Verlag, New York, 1993. MR 95e:13018
7. C. Berenstein and A. Yger, *Bounds for the degrees in the division problem*, Michigan Math. J. **37** (1990), 25–43. MR 91c:32004
8. S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean Analysis. A Systematic Approach to Rigid Analytic Geometry*, Grundlehren der Mathematischen Wissenschaften, vol. 261, Springer-Verlag, Berlin, 1984. MR 86b:32031
9. H. Cohen, *A Course in Computational Algebraic Number Theory*, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, Berlin, 1993. MR 94i:11105
10. A. Dickenstein, N. Fitchas, M. Giusti, and C. Sessa, *The membership problem for unmixed polynomial ideals is solvable in single exponential time*, Discrete Appl. Math. **33** (1991), 73–94. MR 92m:13025
11. H. M. Edwards, *Kronecker's views on the foundations of mathematics*, The History of Modern Mathematics, Vol. I (Poughkeepsie, NY, 1989), Academic Press, Boston, MA, 1989, pp. 67–77. MR 91b:01041
12. T. Evans, *Some connections between residual finiteness, finite embeddability and the word problem*, J. London Math. Soc. (2) **1** (1969), 399–403. MR 40:2589
13. G. Gallo and B. Mishra, *A solution to Kronecker's problem*, Appl. Algebra in Engrg. Comm. Comput. **5** (1994), no. 6, 343–370. MR 95i:13026
14. P. Gianni, B. Trager, and G. Zacharias, *Gröbner bases and primary decomposition of polynomial ideals*, J. Symbolic Comput. **6** (1988), no. 2-3, 149–167. MR 90f:68091
15. R. Gilmer, *Multiplicative Ideal Theory*, Queen's Papers in Pure and Applied Mathematics, vol. 12, Queen's University, Kingston, Ont., 1968. MR 37:5198
16. S. Glaz, *Commutative Coherent Rings*, Lecture Notes in Math., vol. 1371, Springer-Verlag, Berlin-Heidelberg-New York, 1989. MR 90f:13001
17. S. Greco and P. Salmon, *Topics in* $\mathfrak{m}$*-adic Topologies*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 58, Springer-Verlag, New York-Berlin, 1971. MR 44:190
18. R. Greenlaw, H. J. Hoover, and W. L. Ruzzo, *Limits to Parallel Computation: P-Completeness Theory*, Oxford University Press, Oxford, 1995. MR 96e:68033
19. K. Hentzelt and E. Noether, *Zur Theorie der Polynomideale und Resultanten*, Math. Ann. **88** (1923), 53–79.
20. G. Hermann, *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann. **95** (1926), 736–788.
21. A. Kandri-Rody and D. Kapur, *Computing a Gröbner basis of a polynomial ideal over a Euclidean domain*, J. Symbolic Comput. **6** (1988), no. 1, 37–57. MR 89h:13002
22. J. Kollár, *Sharp effective Nullstellensatz*, J. Amer. Math. Soc. **1** (1988), 963–975. MR 89h:12008
23. T. Krick, L. M. Pardo, and M. Sombra, *Sharp estimates for the arithmetic Nullstellensatz*, Duke Math. J. **109** (2001), no. 3, 521–598. MR 2002h:11060
24. S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, New York, 1983. MR 85j:11005

25. ———, *Algebraic Number Theory*, 2nd ed., Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994. MR 95f:11085
26. E. Mayr, *Membership in polynomial ideals over $\mathcal{Q}$ is exponential space complete*, STACS 89 (Paderborn, 1989), Lecture Notes in Comput. Sci., vol. 349, Springer, Berlin, 1989, pp. 400–406. MR 90i:68014
27. E. Mayr and A. Meyer, *The complexity of the word problems for commutative semigroups and polynomial ideals*, Adv. Math. **46** (1982), no. 3, 305–329. MR 84g:20099
28. G. Moreno Socías, *Length of polynomial ascending chains and primitive recursiveness*, Math. Scand. **71** (1992), no. 2, 181–205. MR 94d:13019
29. R. O'Leary and J. Vaaler, *Small solutions to inhomogeneous linear equations over number fields*, Trans. Amer. Math. Soc. **336** (1993), no. 2, 915–931. MR 93f:11032
30. P. Philippon, *Dénominateurs dans le théorème des zéros de Hilbert*, Acta Arith. **58** (1991), no. 1, 1–25. MR 92i:13008
31. B. Renschuch, *Beiträge zur konstruktiven Theorie der Polynomideale. XVII/1. Zur Hentzelt/Noether/Hermannschen Theorie der endlich vielen Schritte*, Wiss. Z. Pädagog. Hochsch. "Karl Liebknecht" Potsdam **24** (1980), no. 1, 87–99. MR 83d:13003
32. F. Richman, *Constructive aspects of Noetherian rings*, Proc. Amer. Math. Soc. **44** (1974), 436–441. MR 54:4937
33. D. Roy and J. L. Thunder, *Bases of number fields with small height*, Rocky Mountain J. Math. **26** (1996), no. 3, 1089–1098. MR 98d:11126
34. K. Schmidt-Göttsch, *Polynomial bounds in polynomial rings over fields*, J. Algebra **125** (1989), 164–180. MR 91c:12001
35. A. Seidenberg, *Constructions in algebra*, Trans. Amer. Math. Soc. **197** (1974), 273–313. MR 50:2141
36. ———, *What is Noetherian?*, Rend. Sem. Mat. Fis. Milano **44** (1974), 55–61. MR 54:4938
37. H. Simmons, *The solution of a decision problem for several classes of rings*, Pacific J. Math. **34** (1970), 547–557. MR 42:1658
38. M. Sombra, *A sparse effective Nullstellensatz*, Adv. in Appl. Math. **22** (1999), 271–295. MR 2000c:13041
39. S. S. Wainer, *A classification of the ordinal recursive functions*, Arch. Math. Logik Grundlagenforsch. **13** (1970), 136–153. MR 45:3207

Abstract. We present a new approach to the ideal membership problem for polynomial rings over the integers: given polynomials $f_0, f_1, \ldots, f_n \in \mathbb{Z}[X]$, where $X = (X_1, \ldots, X_N)$ is an N-tuple of indeterminates, are there $g_1, \ldots, g_n \in \mathbb{Z}[X]$ such that $f_0 = g_1 f_1 + \cdots + g_n f_n$? We show that the degree of the polynomials $g_1, \ldots, g_n$ can be bounded by $(2d)^{2^{O(N \log(N+1))}}(h+1)$ where d is the maximum total degree and h the maximum height of the coefficients of $f_0, \ldots, f_n$. Some related questions, primarily concerning linear equations in $R[X]$, where R is the ring of integers of a number field, are also treated.

Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, California 94720; Department of Mathematics, University of California at Berkeley, Evans Hall, Berkeley, California 94720

Current address: Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 851 S. Morgan St. (M/C 249), Chicago, Illinois 60607

E-mail address: maschenb@math.uic.edu