Unification in Commutative Theories, Hilbert's Basis Theorem, and Gröbner Bases

FRANZ BAADER

German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany

Abstract. Unification in a communitative theory E may be reduced to solving linear equations in the corresponding semiring $S(\mathrm{E})$ [37]. The unification type of E can thus be characterized by algebraic properties of $S(\mathrm{E})$. The theory of Abelian groups with n commuting homomorphisms corresponds to the semiring $\mathbb{Z}[X_1, \ldots, X_n]$. Thus, Hilbert's Basis Theorem can be used to show that this theory is unitary. But this argument does not yield a unification algorithm. Linear equations in $\mathbb{Z}[X_1, \ldots, X_n]$ can be solved with the help of Gröbner Base methods, which thus provide the desired algorithm. The theory of Abelian monoids with a homomorphism is of type zero [4]. This can also be proved by using the fact that the corresponding semiring, namely $\mathbb{N}[X]$, is not Noetherian. Another example of a semiring (even ring) that is not Noetherian is the ring $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$, where $X_1, \ldots, X_n$ $(n > 1)$ are noncommuting indeterminates. This semiring corresponds to the theory of Abelian groups with n noncommuting homomorphisms. Surprisingly, by construction of a Gröbner Base algorithm for right ideals in $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$, it can be shown that this theory is unitary unifying.

Categories and Subject Descriptors: F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems—*computations on polynomials*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*computations on discrete structures*; *pattern matching*; F.4.1 [**Mathematical Logic and Formal Languages**]: Mathematical Logic—*mechanical theorem proving*; I.1.2 [**Algebraic Manipulation**]: Algorithms—*algebraic algorithms*; I.2.3 [**Artificial Intelligence**]: Deduction and Theorem Proving—*resolution*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Equational reasoning, Gröbner bases, unification

1. *Introduction*

E-unification is concerned with solving term equations modulo an equational theory E. More formally, let E be an equational theory and $=_{\mathrm{E}}$ be the equality of terms, induced by E. An *E-unification problem* Γ is a finite set of equations $\langle s_i = t_i;\ 1 \le i \le n\rangle_{\mathrm{E}}$ where s_i and t_i are terms. A substitution θ is called an *E-unifier* of Γ iff $s_i\theta =_{\mathrm{E}} t_i\theta$ for each $i, i = 1, \ldots, n$. The set of all E-unifiers of Γ is denoted by $U_{\mathrm{E}}(\Gamma)$.

This research was carried out while the author was a member of IMMD1, University of Erlangen.
Author's address: German Research Center for Artificial Intelligence (DFKI), Stuhlsatzenhausweg 3, D-6600 Saarbrücken 11, Germany.

In general, we do not need the set of all E-unifiers. A *complete set of E-unifiers*, that is, a set of E-unifiers from which all E-unifiers may be generated by E-instantiation, is sufficient. More precisely, we extend $=_E$ to $U_E(\Gamma)$, and define a quasi-ordering $\leq_E$ on $U_E(\Gamma)$ by

$$\sigma =_E \theta \quad \text{iff} \quad x\sigma =_E x\theta \text{ for all variables } x \text{ occurring in } s_i \text{ or } t_i \text{ for some } i, i = 1, \ldots, n,$$

$$\sigma \leq_E \theta \quad \text{iff} \quad \text{there exists a substitution } \lambda \text{ such that } \sigma =_E \theta \circ \lambda.$$

If $\sigma \leq_E \theta$, then σ is called an E-instance of θ.

A *complete set* $cU_E(\Gamma)$ *of E-unifiers* of Γ is defined as

(1) $cU_E(\Gamma) \subseteq U_E(\Gamma)$,
(2) For all $\theta \in U_E(\Gamma)$, there exists $\sigma \in cU_E(\Gamma)$ such that $\theta \leq_E \sigma$.

For reasons of efficiency, this set should be as small as possible. Thus, we are interested in *minimal complete sets of E-unifiers*, that means complete sets where two different elements are not comparable with respect to E-instantiation. The *unification type* of a theory E is defined with reference to the cardinality and existence of minimal complete sets. The theory E is *unitary* (*finitary*, *infinitary*) iff minimal complete sets of E-unifiers always exist and their cardinality is at most one (always finite, at least once infinite). E has unification *type zero* iff there is an E-unification problem without minimal complete set of E-unifiers. Please note that the signature over which the terms of the unification problems may be built is important for the definition of the unification type of a theory. If the terms of the problems may only contain symbols that occur in an identity of E, then one talks about *elementary E-unification*. If the terms of the problems may contain additional "free" constants, one talks about *E-unification with constants*, and if they may contain additional "free" function symbols of arbitrary arity, one talks about *general E-unification*. These additional symbols may, for example, arise as Skolem constants or Skolem functions in the context of automated theorem proving. In the present paper, we shall restrict our attention to elementary unification and unification with constants. If nothing else is specified, "unification" will mean "elementary unification." Sometimes, we shall also use the notion "unification without constants" to distinguish elementary unification from unification with constants. For more information about unification theory and the unification hierarchy, consult Siekmann [43].

Unification in the empty theory (which is unitary with respect to general unification) plays an important role in automated theorem proving, term rewriting and logic programming. Generalizations to E-unification usually require that E is finitary (see e.g., Stickel [45], Jouannaud and Kirchner [26], Huet [23], and Jaffar et al. [25]). A finitary theory most used in this context is the theory of Abelian semigroups (monoids), that is, the theory of an associative, commutative binary operation (with a neutral element). Unification algorithms for this theory (see, e.g., Livesey and Siekmann [34], Stickel [44], Fages [15], Fortenbacher [17], Büttner [10], and Herold [21]) make use of the fact that unifiers correspond to solutions of systems of linear equations in the semiring $\mathbb{N}$ of nonnegative integers (see Eilenberg [14] or Kuich and Salomaa [31] for the definition and properties of semirings). The same phenomenon occurs for the theory of Abelian groups where the semiring is the ring $\mathbb{Z}$ of integers (Lankford et al. [32]) and for the theory of idempotent Abelian monoids where the 2-element Boolean semiring $\mathscr{B}$ is used (Livesey and Siekmann [34], Baader and Büttner [6]).

These three theories belong to the class of commutative theories (roughly speaking, theories where the finitely generated free objects are direct products of the free objects in one generator), which were defined in Baader [4]. In that paper, it is shown that constant-free unification in commutative theories is either unitary or of type zero, and there are given sufficient conditions for a commutative theory to be unitary (respectively, finitary with respect to unification with constants). The above-mentioned results for Abelian monoids, Abelian groups, and idempotent Abelian monoids and some new results (for Abelian monoids with an involution, idempotent Abelian monoids with an involution, Abelian groups with an involution, Abelian groups of exponent m) could thus be obtained as corollaries to a general theorem. In Baader [5], these conditions were modified to algebraic characterizations of unification type unitary for unification without constants, and type finitary for unification with constants in commutative theories. An interesting consequence of these characterizations is the fact that commutative theories are always unitary (finitary with respect to unification with constants), if the finitely generated free objects are finite [4].

Werner Nutt [37, 38] observed that commutative theories are (modulo a translation of the signature) what he calls *monoidal theories*, and that unification in these theories may always be reduced to solving linear equations in certain semirings. He pointed out that the theory of Abelian groups with a homomorphism corresponds to the semiring $\mathbb{Z}[X]$. Thus, Hilbert's Basis Theorem can be used to prove that the theory of Abelian groups with a homomorphism is unitary. But this argument does not yield a unification algorithm. Linear equations in $\mathbb{Z}[X]$ can be solved with the help of Gröbner Base methods (see Buchberger [9] and Section 6 of this paper), which thus provide the desired algorithm.

The theory of Abelian monoids with a homomorphism is of type zero. This was shown in Baader [4] using purely combinatorial arguments. In Section 4 of the present paper, we shall see that this can also be proved algebraically, by using the fact that the corresponding semiring, namely $\mathbb{N}[X]$, is not Noetherian.

Another example of a semiring that is not Noetherian is the ring $\mathbb{Z}\langle X,Y\rangle$, where X, Y are noncommuting indeterminates. This semiring corresponds to the theory of Abelian groups with two (noncommuting) homomorphisms. Surprisingly, by construction of a Gröbner Base algorithm for right ideals in $\mathbb{Z}\langle X,Y\rangle$, I was able to show that this theory is unitary unifying. Of course, this result can be extended to an arbitrary, finite number of noncommuting indeterminates (Section 8 and 9).

2. *Commutative Theories*

In this section, we give a definition of commutative theories, recall some of the properties derived in Baader [5], and show how the corresponding semirings may be obtained within this framework.

An equational theory E defines a *variety* $V(E)$, that is, the class of all algebras (of the given signature Ω) that satisfy each identity of E. For any set X of generators, V(E) contains a *free algebra over* V(E) *with generators* X, which will be denoted by $F_{\mathrm{E}}(X)$.

Let $F(\mathrm{E})$ be the class of all free algebras $F_{\mathrm{E}}(X)$ with finite sets X, and let $C(\mathrm{E})$ be the category which has the elements of $F(\mathrm{E})$ as objects and the homomorphisms between these elements as morphisms. Note that the coproduct of $F_{\mathrm{E}}(X)$ and $F_{\mathrm{E}}(Y)$ in $C(\mathrm{E})$ is given by $F_{\mathrm{E}}(X \cup Y)$ (where $\cup$ means disjoint union). If $|X| = |Y|$, the algebras $F_{\mathrm{E}}(X)$ and $F_{\mathrm{E}}(Y)$ are isomorphic. Thus, $F_{\mathrm{E}}(X)$ is the coproduct of the isomorphic objects $F_{\mathrm{E}}(x)$ for $x \in X$, where x is used as abbreviation for the singleton set $\{x\}$.

Let $\Gamma = \langle s_i = t_i;\ 1 \le i \le n\rangle_{\mathrm{E}}$ be an E-unification problem and X be the (finite) set of variables x occurring in some s_i or t_i. Evidently, we can consider the terms s_i and t_i as elements of $F_{\mathrm{E}}(X)$. Since we do not distinguish between $=_{\mathrm{E}}$-equivalent unifiers, any E-unifier of Γ can be regarded as a homomorphism of $F_{\mathrm{E}}(X)$ into $F_{\mathrm{E}}(Y)$ for some finite set Y (of variables). Let $I = \{x_1, \dots, x_n\}$ be a set of cardinality n. We define homomorphisms

$$\sigma, \tau \colon F_{\mathrm{E}}(I) \to F_{\mathrm{E}}(X) \text{ by } x_i\sigma := s_i \quad \text{and} \quad x_i\tau := t_i \ (i = 1, \dots, n).$$

Now $\delta\colon F_{\mathrm{E}}(X) \to F_{\mathrm{E}}(Y)$ is an E-unifier of Γ iff $x_i\sigma\delta = s_i\delta = t_i\delta = x_i\tau\delta$ for $i = 1, \dots, n$, that is, iff $\sigma\delta = \tau\delta$. Thus, an E-unification problem can be written as a pair $\langle \sigma = \tau\rangle_{\mathrm{E}}$ of morphisms $\sigma, \tau\colon F_{\mathrm{E}}(I) \to F_{\mathrm{E}}(X)$ in the category $C(\mathrm{E})$. An E-unifiers of the unification problem $\langle \sigma = \tau\rangle_{\mathrm{E}}$ is a morphism δ such that $\sigma\delta = \tau\delta$.

Motivated by this categorical reformulation of E-unification (due to Rydeheard and Burstall [41]), the class of *commutative theories* is defined by properties of the category $C(\mathrm{E})$ of finitely generated E-free objects as follows: An equational theory E is commutative iff the corresponding category $C(\mathrm{E})$ is a *semiadditive category* (see Herrlich and Strecker [22], Freyd [18], and Baader [4] for the definition of semiadditive categories). In order to give a more algebraic definition of commutative theories, we need some additional notation from universal algebra [11, 20].

A constant symbol (i.e., a nullary function symbol) $e \in \Omega$ is called *idempotent in E* iff for any $f \in \Omega$ we have $f(e, \dots, e) =_{\mathrm{E}} e$, that is, in any algebra $A \in V(\mathrm{E})$, $f(e, \dots, e) = e$ holds. Note that for nullary f this means $f =_{\mathrm{E}} e$.

Let $\mathbf{K}$ be a class of algebras (of signature Ω). An n-ary *implicit operation* in $\mathbf{K}$ is a family $o = \{o_A;\ A \in \mathbf{K}\}$ of mappings $o_A\colon A^n \to A$ that is compatible with all homomorphisms, that is, for any homomorphism $\phi\colon A \to B$ with $A, B \in \mathbf{K}$ and all $a_1, \dots, a_n \in A$, $o_A(a_1, \dots, a_n)\phi = o_B(a_1\phi, \dots, a_n\phi)$ holds. In the following, we shall omit the index and just write o for any o_A. Obviously, an Ω-term induces an implicit operation on any class of Ω-algebras.

Definition 2.1. An equational theory E is called *commutative* iff the following holds:

(1) Ω contains a constant symbol e which is idempotent in E.
(2) There is a binary implicit operation $*$ in $F(\mathrm{E})$ such that
  (a) The constant e is a neutral element for $*$ in any algebra $A \in F(\mathrm{E})$.
  (b) For any n-ary function symbol $f \in \Omega$, any algebra $A \in F(\mathrm{E})$, and any $s_1, \dots, s_n, t_1, \dots, t_n \in A$, we have $f(s_1 * t_1, \dots, s_n * t_n) = f(s_1, \dots, s_n) * f(t_1, \dots, t_n)$.

In Baader [4], the following properties of commutative theories E are shown within a categorical framework, using well-known results for semiadditive categories.

Property 2.2. $|F_E(\emptyset)| = 1$ and $F_E(\emptyset)$ is the zero object of $C(E)$.

Property 2.3. The implicit operation $*$ of Definition 2.1 is associative and commutative. It induces a binary operation $+$ on any morphism set $\hom(F_E(X), F_E(Y))$ as follows: Let $\sigma, \tau: F_E(X) \to F_E(Y)$ and $s \in F_E(X)$. Then, $s(\sigma + \tau) := (s\sigma)*(s\tau)$.

This operation is also associative and commutative, and it distributes with the composition of morphisms. Let e be the idempotent constant required in the definition of commutative theories. Then, the morphism $0: F_E(X) \to F_E(Y)$ defined by $x \mapsto e$ for all $x \in X$ is the zero morphism in $\hom(F_E(X), F_E(Y))$, and it is a neutral element for $+$ on $\hom(F_E(X), F_E(Y))$.

Property 2.4. The coproduct $F_E(X \cup Y)$ of $F_E(X)$ and $F_E(Y)$ is also the product of these objects, that is, $F_E(X \dot\cup Y) \cong F_E(X) \times F_E(Y)$.

Property 2.5. Consider $\sigma: F_E(X) \to F_E(Y)$. Let u_x for $x \in X$ (p_y for $y \in Y$) be the injections of the coproduct $F_E(X)$ (projections of the product $F_E(Y)$). Then, σ is uniquely determined by the matrix $M_\sigma = (u_x \sigma p_y)_{x \in X, y \in Y}$. For $\sigma, \tau: F_E(X) \to F_E(Y)$ and $\delta: F_E(Y) \to F_E(Z)$, we have $M_{\sigma+\tau} = M_\sigma + M_\tau$ and $M_{\sigma\delta} = M_\sigma \cdot M_\delta$.

Nutt [37, 38] observed that commutative theories are, modulo a translation of the signature, what he calls *monoidal theories* (see Baader and Nutt [7] for a proof), and that unification in a monoidal theory E may be reduced to solving linear equations in a certain semiring $S(E)$.

In our framework, this semiring can be obtained as follows: Let $\mathbf{1}$ be an arbitrary set of cardinality 1. Property (2.3) yields that $\hom(F_E(\mathbf{1}), F_E(\mathbf{1}))$ with addition "+" and composition as multiplication is a *semiring*, which shall be *denoted by* $S(E)$. Any $F_E(x)$ is isomorphic to $F_E(\mathbf{1})$ and for $|X| = n$, $F_E(X)$ is thus nth power and copower of $F_E(\mathbf{1})$. That means that, for $\sigma: F_E(X) \to F_E(Y)$, the entries $u_x \sigma p_y$ of the $|X| \times |Y|$-matrix M_σ may all be considered as elements of $S(E)$. Hence, all morphisms of $C(E)$ can be written as matrices over the semiring $S(E)$. Addition and composition of morphisms correspond to addition and multiplication of matrices over $S(E)$ as stated in (2.5).

Now we shall give some examples of commutative theories in which the unification properties will be considered in subsequent sections of this paper. In all these examples, the implicit operation is given by a function symbol of the signature that is associative and commutative in the corresponding theory. Additional examples of commutative theories can be found in Baader [4].

Example 2.6. We consider the following signatures: $\Sigma := \{\cdot, 1, h\}$, where $\cdot$ is binary, 1 is nullary, and h is unary, and for $n \geq 0$, $\Omega_n := \{\cdot, 1, ^{-1}, h_1, \ldots, h_n\}$, where $\cdot$ is binary, 1 is nullary, and $^{-1}$ and the h_i are unary.

(1) The theory AMH of *Abelian monoids with a homomorphism*. The signature is Σ and $\mathrm{AMH} := \{x \cdot 1 = x, x \cdot (y \cdot z) = (x \cdot y) \cdot z, x \cdot y = y \cdot x, h(x \cdot y) = h(x) \cdot h(y), h(1) = 1\}$.
(2) The theory of AIMH of *idempotent Abelian monoids with a homomorphism*. The signature is Σ and $\mathrm{AIMH} := \mathrm{AMH} \cup \{x \cdot x = x\}$.
(3) The theory AGnH of *Abelian groups with n (noncommuting) homomorphisms*. We take signature Ω_n and define $\mathrm{AGnH} := \{x \cdot 1 = x, x \cdot (y \cdot z) = (x \cdot y) \cdot z, x \cdot y = y \cdot x, x \cdot x^{-1} = 1\} \cup \{h_i(x \cdot y) = h_i(x) \cdot h_i(y); 1 \leq i \leq n\}$.

(4) The theory AGnHC of *Abelian groups with n commuting homomorphisms*. The signature is Ω_n and AGnHC := AGnH $\cup \{h_i(h_j(x)) = h_j(h_i(x));\ 1 \leq i < j \leq n\}$.

It is easy to see that these theories are commutative. Note that the implicit operation induced by the term $x \cdot y$ (for a binary function symbol "·") satisfies (2)(b) of Definition 2.1 for $f = \cdot$ iff $(a \cdot b) \cdot (c \cdot d) = (a \cdot c) \cdot (b \cdot d)$ holds in any algebra $A \in F(\mathrm{E})$, and for $f = h$ (for a unary function symbol h) iff $h(x \cdot y) = h(x) \cdot h(y)$ holds.

3. *Unification in Commutative Theories*

In this section, we recall the characterizations of unification type unitary (finitary for unification with constants) for commutative theories given in Baader [5] within the categorical framework. As a consequence, we derive that unification in a commutative theory E means solving systems of linear equations in the semiring $S(\mathrm{E})$. This yields algebraic characterizations of the unification types that are similar to those given in Nutt [38] and Baader and Nutt [7].

THEOREM 3.1. *A commutative theory E is unitary with respect to unification without constants iff it satisfies the following condition*:

Let y be an arbitrary variable. For any E-unification problem $\langle \sigma = \tau \rangle_E$ *(where* σ, τ: $F_E(I) \to F_E(X)$*) there are finitely many E-unifiers* $\alpha_1, \ldots, \alpha_r$: $F_E(X) \to F_E(y)$ *such that any E-unifier* δ: $F_E(X) \to F_E(y)$ *can be represented as*

$$\delta = \sum_{i=1}^{i=r} \alpha_i \lambda_i,$$

where λ_i: $F_E(y) \to F_E(y)$ *are morphisms of* $C(E)$.

If we translate morphisms into matrices over $S(\mathrm{E})$, we obtain the following reformulation of Theorem 3.1:

COROLLARY 3.2. *A commutative theory E is unitary with respect to unification without constants iff the corresponding semiring* $S(E)$ *satisfies the following condition*: *For any* $n, m \geq 1$ *and any pair* M_σ, M_τ *of* $m \times n$*-matrices over* $S(E)$ *the set*

$$U(M_\sigma, M_\tau) := \left\{ \underline{x} \in S(E)^n;\ M_\sigma \underline{x} = M_\tau \underline{x} \right\}$$

is a finitely generated right $S(E)$*-semimodule, that is, there are finitely many* $\underline{x}_1, \ldots, \underline{x}_r \in S(E)^n$ *such that* $U(M_\sigma, M_\tau) = \{\underline{x}_1 s_1 + \cdots + \underline{x}_r s_r;\ s_1, \ldots, s_r \in S(E)\}$.

THEOREM 3.3. *Let E be a commutative theory that is unitary with respect to unification without constants. Then E is finitary with respect to unification with constants iff the following condition holds*:

For any morphism (of $C(E)$*)* δ: $F_E(X) \to F_E(Y)$, *there exist finite sets* M, N *such that*:

(1) *The elements of* M *are morphisms* μ: $F_E(Y) \to F_E(X)$ *satisfying* $\delta\mu = 1$.
(2) *The elements of* $N = \{\nu_1, \ldots, \nu_r\}$ *are morphisms* ν_i: $F_E(Y) \to F_E(Z_i)$ *with* $\delta\nu_i = 0$.

(3) *For any* $\lambda: F_E(Y) \to F_E(X)$ *with* $\delta\lambda = 1$, *there are* $\mu \in M$ *and morphisms* $\lambda_1, \ldots, \lambda_r$ (*where* $\lambda_i: F_E(Z_i) \to F_E(X)$) *satisfying*

$$\lambda = \mu + \sum_{i=1}^{i=r} \nu_i \lambda_i .$$

The translation of morphisms into matrices over S(E) yields a sufficient condition for E to be finitary with respect to unification with constants.

COROLLARY 3.4. *Let E be a unitary commutative theory. Then E is finitary with respect to unification with constants, if the following condition holds in* $S(E)$:

Let A *be any* $m \times n$*-matrix over* $S(E)$ *and let* $\underline{b}$ *be any element of* $S(E)^m$. *Then the set* $M := \{\underline{x} \in S(E)^n;\ A\underline{x} = \underline{b}\}$ *is a finite union of cosets of the (finitely generated) right* $S(E)$*-semimodule* $N := \{\underline{x} \in S(E)^n;\ A\underline{x} = 0\}$, *that is, there exist finitely many* $\underline{m}_1, \ldots, \underline{m}_k \in S(E)^n$ *such that* $M = \{\underline{m}_i + \underline{n};\ \underline{n} \in N \text{ and } 1 \le i \le k\}$.

Note that the semimodule N is finitely generated since E is unitary and $N = U(A, 0)$, where 0 is the $m \times n$ zero matrix. From Theorem 3.3, we can only deduce that the condition of the corollary is sufficient since in Theorem 3.3 we talk about specific inhomogeneous equations AX = E, while in Corollary 3.4 the right-hand side of the equation is an arbitrary vector $\underline{b}$. Nutt [37] and Baader and Nutt [7] consider a different condition for unification with constants, which turns out to be a characterization of type finitary. The difference between the two conditions stems from the different treatment of unification with constants. Baader [4, 5] generalizes Stickel's approach to AC-unification with constants (Stickel [44]), whereas Nutt [37] builds up on the approach as, for example, by Herold [21].

Assume that S(E) is a ring and let $\underline{x}_0$ be an arbitrary solution of the inhomogeneous equation $A\underline{x} = \underline{b}$. Then any solution $\underline{y}$ of $A\underline{x} = \underline{b}$ is of the form $\underline{y} = \underline{x}_0 + \underline{z}$, where $\underline{z} := \underline{y} - \underline{x}_0$ is a solution of the homogeneous equation $A\underline{x} = 0$. This proves

COROLLARY 3.5. *Let E be a unitary commutative theory such that* $S(E)$ *is a ring. Then E is unitary with respect to unification with constants.*

4. *A Commutative Theory of Unification Type Zero*

In 1972, Plotkin [33] conjectured that there exists an equational theory E that is of unification type zero. But it wasn't until 1983 that Fages and Huet [16] constructed the first example of an equational theory of this type. Schmidt-Schauß [42] and the present author [2] showed that the theory of idempotent semigroups is of unification type zero, and in Baader [3], it is proved that almost all varieties of idempotent semigroups are defined by type zero theories. This provides us with countably many examples of type zero theories that are more natural than the original example of Fages and Huet.

In Baader [4], it is shown with the help of purely combinatorial arguments that the theory AIMH of idempotent Abelian monoids with a homomorphism is of type zero. The same proof can be used for AMH, the theory of Abelian monoids with a homomorphism, in place of AIMH. This section contains a more algebraic proof of the fact that AMH is of type zero. This algebraic proof is easier and better comprehensible than the original one. Since commutative theories are either unitary or of unification type zero (Baader [4, Theorem 6.1]), it is sufficient to show that the semiring S(AMH) does not satisfy the condition of Corollary 3.2.

Let σ: $F_{\mathrm{AMH}}(x) \to F_{\mathrm{AMH}}(x)$ be a morphism of C(AMH). Then there are $k \geq 0$ and $a_0, \ldots, a_k \in \mathbb{N}$ such that $x\sigma =_{\mathrm{AMH}} x^{a_0}h(x^{a_1}) \cdots h^k(x^{a_k})$. We associate with the morphism σ the polynomial $p_\sigma = a_0 + a_1X + \cdots + a_kX^k \in \mathbb{N}[X]$. It is easy to see that $p_{\sigma\delta} = p_\sigma \cdot p_\delta$ and $p_{\sigma+\delta} = p_\sigma + p_\delta$, which shows that $S(\mathrm{AMH}) \cong \mathbb{N}[X]$.

We consider the linear equation $(*)$ $Xx_1 + Xx_2 = x_2 + X^2x_3$, which has to be solved by a vector $\underline{p} = (p_1, p_2, p_3)$ in $(\mathbb{N}[X])^3$. Obviously, for any $n \geq 0$, the vector $\underline{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)}) = (1, X + X^2 + \cdots + X^{n+1}, X^n)$ is a solution of $(*)$.

LEMMA 4.1. *There does not exist a solution* $\underline{p}$ *of* $(*)$ *in* $(\mathbb{N}[X])^3$ *such that* $p_1 + p_3 = 1$.

PROOF. For $p_1 = 0$ and $p_3 = 1$ we get $Xp_2 = p_2 + X^2$, which yields $(X - 1)p_2 = X^2$ in $\mathbb{Z}[X]$. But $X - 1$ is not a divisor of X^2. The case $p_1 = 1$ and $p_3 = 0$ leads to a similar contradiction. □

Similarly to ideals in rings one can define semiideals in semirings. A subset I of $\mathbb{N}[X]$ is a *semiideal* iff it is closed under addition (i.e., $f, g \in I$ implies $f + g \in I$) and multiplication with elements of $\mathbb{N}[X]$ (i.e., $f \in I$ and $g \in \mathbb{N}[X]$ imply $fg \in I$). The semiideal I is *finitely generated* iff there exist $f_1, \ldots, f_n \in I$ such that $I = \{f_1g_1 + \cdots + f_ng_n;\ g_1, \ldots, g_n \in \mathbb{N}[X]\}$.

It is easy to see that $I_{1+3} := \{p_1 + p_3;$ there exists p_2 such that (p_1, p_2, p_3) solves $(*)\}$ is a semiideal in $\mathbb{N}[X]$. We know that $1 + X^n \in I_{1+3}$ for any $n \geq 0$ and $1 \notin I_{1+3}$.

LEMMA 4.2. *A semiideal* $I \subseteq \mathbb{N}[X]$ *such that* $1 + X^n \in I$ *for any* $n \geq 0$ *and* $1 \notin I$ *is not finitely generated.*

PROOF. Evidently $1 + X^n = f \cdot g$ for $f, g \in \mathbb{N}[X]$ or $1 + X^n = f + g$ for $f, g \in \mathbb{N}[X] \setminus \{0\}$ implies $f = 1$ or $g = 1$. Since $1 \notin I$, this means that a sum $1 + X^n = f + g$ with $f, g \in I \setminus \{0\}$ is impossible, and that a factorization $1 + X^n = f \cdot g$ with $f \in \mathbb{N}[X]$, $g \in I$ cannot be a real factorization of $1 + X^n$, that is, g has to be $1 + X^n$ itself. This shows that $1 + X^n$ cannot be generated by other elements of I. □

PROPOSITION 4.3. *The theory AMH is of unification type zero.*

PROOF. Assume that AMH is not of type zero. Then AMH is unitary and, by Corollary 3.2, $\underline{I} := \{\underline{p} \in (\mathbb{N}[X])^3;\ \underline{p}$ is a solution of $(*)\}$ is a finitely generated right $\mathbb{N}[X]$-semimodule. But then the semiideal $I_{1+3} = \{p_1 + p_3;$ there exists p_2 such that $(p_1, p_2, p_3) \in \underline{I}\}$ would also be finitely generated, which contradicts Lemma 4.2. □

The fact that the set of solutions of the equation $(*)$ is not a finitely generated right semimodule is not specific for the semiring $\mathbb{N}[X]$. More general, let S be a semiring that is not a ring (that means that there exists $s \in S$ such that for all $t \in S$, $s + t \neq 0$). Then the right $S[X]$-semimodule $I := \{\underline{p} \in (S[X])^3;\ \underline{p}$ is a solution of $(*)\}$ is not finitely generated (Baader and Nutt [7]).

5. *AGnHC-Unification and Hilbert's Basis Theorem*

It is easy to see that S(AGnHC) is isomorphic to the ring $\mathbb{Z}[X_1,\ldots,X_n]$, that is, the polynomial ring over $\mathbb{Z}$ in the (commuting) indeterminates $X_1,\ldots,X_n$. To establish the condition of Corollary 3.2, we have to consider systems of homogeneous linear equations in $\mathbb{Z}[X_1,\ldots,X_n]$, that is, systems $f_{1i}x_1 + \cdots + f_{ki}x_k = 0$ $(i = 1,\ldots,s)$, where the coefficients f_{ij} and the desired solutions are elements of $\mathbb{Z}[X_1,\ldots,X_n]$. The set of solutions $\underline{I} \subseteq (\mathbb{Z}[X_1,\ldots,X_n])^k$ is a $\mathbb{Z}[X_1,\ldots,X_n]$-module, which is finitely generated by Hilbert's Basis Theorem and the fact that $\mathbb{Z}$ is a Noetherian ring (see, e.g., Jacobson [24]). Thus, AGnHC is unitary with respect to unification without constants. Since $\mathbb{Z}[X_1,\ldots,X_n]$ is a ring, Corollary 3.5 applies and we have proved the following:

PROPOSITION 5.1 [38]. *For any $n \geq 0$, the theory AGnHC is unitary with respect to unification without constants, and it is also unitary with respect to unification with constants.*

This proof of Proposition 5.1 does not yield an AGnHC-unification algorithm because we still do not know how to solve linear equations in $\mathbb{Z}[X_1,\ldots,X_n]$ effectively. The next section describes one possible answer to this problem.

6. *Solving Linear Equations in $\mathbb{Z}[X_1,\ldots,X_n]$ Using Weak Gröbner Bases*

Buchberger [9] describes an effective method which constructs finitely many generators of the solutions of a single equation $f_1x_1 + \cdots + f_kx_k = 0$ where the f_i and the desired solutions are elements of $K[X_1,\ldots,X_n]$ for a field K. This method may also be used for $\mathbb{Z}[X_1,\ldots,X_n]$, but one has to be very careful in the details, and thus the proof of correctness becomes more involved. Systems of equations can then be solved by successive substitution. A more efficient approach to solving systems of equations is described in Furukawa et al. [19] where Gröbner base theory is extended to modules over $K[X_1,\ldots,X_n]$. Furukawa et al. also mention that their approach can be extended to $\mathbb{Z}[X_1,\ldots,X_n]$, but they do not give any details or proofs.

Gröbner bases for polynomials over $\mathbb{Z}$ have been considered in for example, Buchberger [9], and more general for polynomials over Euclidean rings in Kandri-Rody and Kapur [27–29]. However, in the present paper, we shall consider a rewrite relation on polynomials (see 6.2), that is different from those used by Buchberger and Kandri-Rody and Kapur. As a consequence, we shall not get Gröbner bases (in the sense of Buchberger and Kandri-Rody and Kapur), but only "weak" Gröbner bases (see 6.3). But it turns out that weak Gröbner bases are sufficient for the purpose of equation solving. An advantage of our rewrite relation, as compared to the relation used by Kandri-Rody and Kapur, is that the proof of Lemma 6.4—which is crucial for the proof of correctness of this method of equation solving—becomes more obvious. In addition, we thus get a presentation that is very similar to the one used in Sections 8 and 9 for the noncommutative case. Finally, though we cannot just refer to known results on Gröbner bases [8] (e.g., to get Proposition 6.5), we do not have to invest more work. Lemma 6.4 and the arguments used in the proof of Proposition 6.5 are needed anyway for the proof of Proposition 6.8.

First we recall some facts and notations concerning Gröbner bases:

Fact 6.1. *Admissible term orderings.* Let $T_n := \{X_1^{k_1} \cdots X_n^{k_n};\ k_1,\ldots,k_n \in \mathbb{N}\}$ be the set of all terms (i.e., monomials with coefficient 1) in $\mathbb{Z}[X_1,\ldots,X_n]$.

With respect to multiplication of polynomials, T_n is a commutative monoid (with neutral element $1 = X_1^0 \cdots X_n^0$), which is isomorphic to the additive monoid $\mathbb{N}^n$.

A linear ordering $<$ on T_n is called *compatible* iff for all $r, s, t \in T_n$, $r < s$ implies $rt < st$, and it is called *admissible* iff it is compatible and satisfies $1 < s$ for all $s \in T_n$. It is easy to see that a compatible linear ordering on T_n is admissible iff it is Noetherian.

Complete descriptions of all compatible linear orderings have been given by Trevisan [46], Zaiceva [47] and more recently by Robbiano [40] and Martin [35]: Any compatible linear ordering $<$ on T_n is completely determined by a $n \times s$ matrix $U_<$ of $s \leq n$ orthogonal vectors $u_1, \ldots, u_s \in \mathbb{R}^n$ of $\mathbb{Q}$-dimension n as follows: $X_1^{k_1} \cdots X_n^{k_n} < X_1^{h_1} \cdots X_n^{h_n}$ iff the first nonzero element of $(h_1 - k_1, \ldots, h_n - k_n) \cdot U_<$ is greater than zero.

It is easy to see that the compatible linear ordering $<$ is admissible iff in any row of $U_<$, the first nonzero entry is greater than zero.

An admissible ordering $<$ on terms can be extended to monomials and polynomials as follows: Let $a, b \in \mathbb{Z}$ and $s, t \in T_n$. Then, $as < bt$ iff (i) $s < t$ or (ii) $s = t$ and $|a| < |b|$ or (iii) $s = t$ and $|a| = |b|$ and $a < b$. This defines a well-ordering on the monomials of $\mathbb{Z}[X_1, \ldots, X_n]$.

Let $f = \Sigma a_i s_i$ and $g = \Sigma b_i t_i$ be two polynomials, that is, elements of $\mathbb{Z}[X_1, \ldots, X_n]$. Then, we define $f < g$ iff $\{\cdots a_i s_i, \cdots\} \ll \{\cdots b_i t_i, \cdots\}$, where $\ll$ denotes the multiset ordering (see Dershowitz and Manna [12]) induced by the ordering $<$ on monomials. This ordering on polynomials is also Noetherian.

Fact 6.2. *Rewriting with Polynomials*. For a polynomial f and a term t that occurs in f, coeff(t, f) denotes the coefficient of t in f. If t does not occur in f, we define coeff$(t, f) := 0$. Let $<$ be an admissible ordering and let $f = a \cdot t + g$ be a polynomial in $\mathbb{Z}[X_1, \ldots, X_n]$ such that $t \in T_n$ is the greatest term in f with respect to $<$ and coeff$(t, f) = a \in \mathbb{Z} \setminus \{0\}$ is the coefficient of t in f. Then, t is called *head-term* of f (HT(f)), a is called *head-coefficient* of f (HC(f)), $a \cdot t$ is called *head-monomial* of f (HM(f)) and $g = f -$ HM(f) is called *rest* of f (R(f)).

A set F of polynomials induces the following *rewrite relation* on $\mathbb{Z}[X_1, \ldots, X_n]$:

$f \to_F g$ iff (1) f contains a term t with coefficient a,
(2) F contains a polynomial h such that $t = \mathrm{HT}(h) \cdot s$ (for some $s \in T_n$) and $|\mathrm{HC}(h)| \leq |a|$,
(3) $g = f - h \cdot b \cdot s$, where $a = b \cdot \mathrm{HC}(h) + c$ for $0 \leq c < |\mathrm{HC}(h)|$, $b, c \in \mathbb{Z}$.

Let $\xrightarrow{*}_F$ (respectively, $\xrightarrow{+}_F$) denote the reflexive, transitive (respectively, transitive) closure of $\to_F$. It is easy to see that $f \to_F g$ implies $f > g$, and thus $\xrightarrow{+}_F$ is Noetherian. The set F generates an ideal $\langle F \rangle$ in $\mathbb{Z}[X_1, \ldots, X_n]$, and this ideal induces a congruence $\equiv_{\langle F \rangle}$, namely $f \equiv_{\langle F \rangle} g$ iff $f - g \in \langle F \rangle$. Obviously, the reflexive, transitive, and symmetric closure of $\to_F$ is contained in this congruence. However, $\equiv_{\langle F \rangle}$ can be larger than this reflexive, transitive, and symmetric closure since the rewrite relation defined above does not satisfy the "unique remainder property" required in Kandri-Rody and Kapur [28].

Fact 6.3. *Weak Gröbner Bases and S-Polynomials.* Let I be an ideal in $\mathbb{Z}[X_1, \ldots, X_n]$, and let B be a finite set of polynomials. B is a *weak Gröbner base* for I iff $\langle B \rangle = I$ and any element of I can be reduced to 0 with respect to $\to_{\mathrm{B}}$. This is weaker than the definition of Gröbner base where it is required that each $\equiv_I$-class has a unique $\to_{\mathrm{B}}$-irreducible element. For weak Gröbner bases, we only require that the class of 0, namely I, has the unique irreducible element 0. But as for Gröbner bases, the property of being a weak Gröbner base can be localized with the help of so-called S-polynomials.

Let $g_1 = c_1 \cdot t_1 + \mathrm{R}(g_1)$ and $g_2 = c_2 \cdot t_2 + \mathrm{R}(g_2)$ be elements of B such that $c_1 \geq c_2 \geq 0$ (without loss of generality, we assume that the head coefficients of the polynomials in B are positive). The *S-polynomial* $\mathrm{S}(g_1, g_2)$ of g_1 and g_2 is defined as follows: Let $s_1 \cdot t_1 = s_2 \cdot t_2 = \mathrm{lcm}(t_1, t_2)$ and $c_1 = a \cdot c_2 + b, 0 \leq b < c_2 \leq c_1, a \geq 1$. Then

$$\mathrm{S}(g_1, g_2) := s_1 \cdot g_1 - a \cdot s_2 \cdot g_2 = b \cdot s_1 \cdot t_1 + s_1 \cdot \mathrm{R}(g_1) - a \cdot s_2 \cdot \mathrm{R}(g_2).$$

Now B is a weak Gröbner base iff for every pair of polynomials in B the S-polynomial reduces to 0 with respect to $\to_{\mathrm{B}}$. The proof of this fact requires the following technical lemma (which has an easy, but somewhat tedious proof).

LEMMA 6.4. *Let $0 \neq f = at + g$ be a polynomial such that $a \geq 0$ and g contains only terms smaller than t. Let $G = \{g_1, \ldots, g_s\}$ be a set of polynomials such that $f \xrightarrow{*}_G 0$. Then there exist polynomials $w_1, \ldots, w_s$ such that*

$$f = \sum_{k=1}^{k=s} w_k \cdot g_k,$$

and

(1) *for $a = 0$: $max\{HT(w_1 g_1), \ldots, HT(w_s g_s)\} < t$,*
(2) *for $a > 0$: $max\{HT(w_1 g_1), \ldots, HT(w_s g_s)\} = t$, $coeff(t, w_1 g_1) \geq 0, \ldots, coeff(t, w_s g_s) \geq 0$, and $a = coeff(t, w_1 g_1) + \cdots + coeff(t, w_s g_s)$.*

PROOF. By Noetherian induction with respect to $\to_{\mathrm{G}}$, applied to f.

Case 1. $a = 0$.

Then, $g \neq 0$ and there exists g' with $f = g \to_{\mathrm{G}} g' \xrightarrow{*}_{\mathrm{G}} 0$. Assume that the first reduction is done by the polynomial $g_i \in G$. Then $g' = g - qs_i g_i$ where $\mathrm{HT}(g) = s_i \mathrm{HT}(g_i)$ and $\mathrm{HC}(g) = q\mathrm{HC}(g_i) + b, 0 \leq b < |\mathrm{HC}(g)|$. Obviously, $\mathrm{HT}(qs_i g_i) = \mathrm{HT}(g) < t$.

Case 1.1. If $g' = 0$, then $g = qs_i g_i$, and thus we can take $w_\nu := 0$ for $\nu \neq i$ and $w_i := qs_i$.

Case 1.2. If $g' > 0$, then induction yields polynomials $u_1, \ldots, u_s$ such that $at + g' = g' = u_1 g_1 + \cdots + u_s g_s$, and $\mathrm{HT}(u_\nu g_\nu) < t$ for $\nu = 1, \ldots, s$. Thus, we can take $w_\nu := u_\nu$ for $\nu \neq i$ and $w_i := u_i + qs_i$.

Case 2. $a > 0$.

Case 2.1. Assume that the first reduction step of $f \xrightarrow{*}_{\mathrm{G}} 0$ is applied inside of g, that is, $f = at + g \to_{\mathrm{G}} at + g' \xrightarrow{*}_{\mathrm{G}} 0$ and for some g_i in G, $g' = g - qs_i g_i$ where $\mathrm{HT}(g) = s_i \mathrm{HT}(g_i)$ and $\mathrm{HC}(g) = q\mathrm{HC}(g_i) + b$, $0 \leq b < |\mathrm{HC}(g)|$. We can now apply the induction hypothesis to $at + g'$, and thus we get $u_1, \ldots, u_s$ such that $at + g' = u_1 g_1 + \cdots + u_s g_s$, $\text{coeff}(t, u_1 g_1) \geq 0, \ldots, \text{coeff}(t, u_s g_s) \geq 0$, and $a = \text{coeff}(t, u_1 g_1) + \cdots + \text{coeff}(t, u_s g_s)$.

Hence, we can take $w_\nu := u_\nu$ for $\nu \neq i$ and $w_i := u_i + qs_i$. Please note that, since $\text{HT}(qs_i g_i) < t$, $\text{coeff}(t, w_\nu g_\nu) = \text{coeff}(t, u_\nu g_\nu)$ for all ν.

Case 2.2. Assume that the first reduction step of $f \xrightarrow{*}_G 0$ is applied to at. Thus, there exists a polynomial $g_i = c_i t_i + R(g_i)$ in G, a term s_i, and integers b, c such that $t = s_i t_i$, $a = |a| \geq |c_i|$, $a = c_i b + c$, $0 \leq c < |c_i|$.

Case 2.2.1. $c = 0$, that is, $a = c_i b$.

Then, $f \rightarrow_G g' = g - bs_i R(g_i) \xrightarrow{*}_G 0$, and $\text{HT}(g') < t$. By induction we get polynomials $u_1, \ldots, u_s$ such that $0t + g' = g' = u_1 g_1 + \cdots + u_s g_s$, and $\text{HT}(u_\nu g_\nu) < t$ for $\nu = 1, \ldots, s$. Thus, we can take $w_\nu := u_\nu$ for $\nu \neq i$ and $w_i := u_i + bs_i$. We have $\text{coeff}(t, w_\nu g_\nu) = 0$ for $\nu \neq i$, and $\text{coeff}(t, w_i g_i) = \text{coeff}(t, bs_i g_i) = c_i b = a > 0$.

Case 2.2.2. $c > 0$ (this is the most interesting case because here the exact definition of our rewrite relation on polynomials becomes important).

Then, $f \rightarrow_G ct + g' = ct + g - bs_i R(g_i) \xrightarrow{*}_G 0$, and $\text{HT}(g') < t$. By induction, we get polynomials $u_1, \ldots, u_s$ such that $ct + g' = u_1 g_1 + \cdots + u_s g_s$, $\max\{\text{HT}(u_1 g_1), \ldots, \text{HT}(u_s g_s)\} = t$, $\text{coeff}(t, u_1 g_1) \geq 0, \ldots, \text{coeff}(t, u_s g_s) \geq 0$, and $c = \text{coeff}(t, u_1 g_1) + \cdots + \text{coeff}(t, u_s g_s)$. We define $w_\nu := u_\nu$ for $\nu \neq i$ and $w_i := u_i + bs_i$. Then we have $\text{coeff}(t, w_\nu g_\nu) = \text{coeff}(t, u_\nu g_\nu) \geq 0$ for $\nu \neq i$, and $\text{coeff}(t, w_i g_i) = \text{coeff}(t, u_i g_i) + \text{coeff}(t, bs_i g_i) = \text{coeff}(t, u_i g_i) + c_i b$. Consequently, $\text{coeff}(t, w_1 g_1) + \cdots + \text{coeff}(t, w_s g_s) = c + c_i b = a$.

It remains to be shown that $\text{coeff}(t, w_i g_i) = \text{coeff}(t, u_i g_i) + c_i b \geq 0$. Assume that $c_i b < 0$. Because of $0 \leq c < |c_i| \leq |c_i b|$ and $a = c_i b + c$, this would imply $a < 0$, which contradicts the assumptions of the lemma. This completes the proof of Lemma 6.4. □

In the sequel, the following notations will be convenient:

(1) Let $h_1, \ldots, h_m$ be elements of $\mathbb{Z}[X_1, \ldots, X_n]$. We denote the $1 \times m$-matrix $(h_1, \ldots, h_m)$ by $\underline{h}$, and the $m \times 1$ matrix $(h_1, \ldots, h_m)^{\mathrm{T}}$ (here $^{\mathrm{T}}$ denotes the transpose of matrices) by $|h$.

(2) For a sequence $q_1, \ldots, q_s$ of polynomials, the *complexity measure* $\text{BS}(q_1, \ldots, q_s)$ is defined as follows: If all the q_i's are zero, then $\text{BS}(q_1, \ldots, q_s) := 0 \cdot 1 = 0 \cdot X_1^0 \cdots X_n^0$.

Otherwise, let $t := \max\{\text{HT}(q_1), \ldots, \text{HT}(q_s)\}$, and for all $i, 1 \leq i \leq s$, let $a_i := \text{coeff}(t, q_i)$ (Note that $a_i = 0$ for $\text{HT}(q_i) < t$). Then $\text{BS}(q_1, \ldots, q_s) := (|a_1| + \cdots + |a_s|) \cdot t$.

Now t is called the term and $|a_1| + \cdots + |a_s|$ the coefficient of $\text{BS}(q_1, \ldots, q_s)$. We define $a \cdot t = \text{BS}(q_1, \ldots, q_s) < \text{BS}(q_1', \ldots, q_s') = a' \cdot t'$ iff $t < t'$ or $t = t'$ and $a < a'$.

(3) Let $\mathrm{B} = \{g_1, \ldots, g_s\}$ be a set of polynomials, and let $\mathrm{S}(g_i, g_j) = s_i \cdot g_i - a \cdot s_j \cdot g_j = b \cdot s_i \cdot t_i + s_i \cdot \mathrm{R}(g_i) - a \cdot s_j \cdot \mathrm{R}(g_j)$ be the S-polynomial of g_i and g_j (see 6.3). If we assume that $\mathrm{S}(g_i, g_j) \xrightarrow{*}_{\mathrm{B}} 0$, then the assumptions of Lemma 6.4 are satisfied. Thus, we get polynomials $w_1, \ldots, w_s$ such that $\mathrm{S}(g_i, g_j) = \underline{g} \cdot |w$, and $\text{BS}(w_1 \cdot g_1, \ldots, w_s \cdot g_s) = c \cdot t'$ for some $t' < s_i \cdot t_i$, if $b = 0$, or $\text{BS}(w_1 \cdot g_1, \ldots, w_s \cdot g_s) = b \cdot s_i \cdot t_i$, if $b \neq 0$.

Now $s_i \cdot g_i - a \cdot s_j \cdot g_j = \mathrm{S}(g_i, g_j) = w_1 \cdot g_1 + \cdots + w_s \cdot g_s$ implies $w_1 \cdot g_1 + \cdots + (w_i - s_i) \cdot g_i + \cdots + (w_j + a \cdot s_j) \cdot g_j + \cdots + w_s \cdot g_s = 0$, and thus $|w_{ij} := (w_1, \ldots, w_i - s_i, \ldots, w_j + a \cdot s_j, \ldots, w_s)^{\mathrm{T}}$ satisfies $\underline{g} \cdot |w_{ij} = 0$.

PROPOSITION 6.5. *B is a weak Gröbner base iff for every pair of polynomials in B the S-polynomial reduces to* 0 *with respect to* $\to_B$.

PROOF

(1) Let B be a weak Gröbner base. It is easy to see that, for polynomials $g_i, g_j \in \mathrm{B}$, the S-polynomial $\mathrm{S}(g_i, g_j)$ is in $\langle \mathrm{B} \rangle$, and thus reduces to 0 by the definition of weak Gröbner base.
(2) Let f_0 be an element of $\langle \mathrm{B} \rangle$. That means that there exist polynomials $p_1, \ldots, p_s$ such that $\underline{g} \cdot |p = f_0$. The **if**-part of the proposition is now proved by nested induction on $\to_{\mathrm{B}}$ and $\mathrm{BS}(g_1 p_1, \ldots, g_s p_s)$. If $f_0 = 0$, there is nothing to prove. Otherwise, $\mathrm{BS}(g_1 p_1, \ldots, g_s p_s) = a \cdot t$ for a positive integer a and a term t.

Case 1. Assume that for $\nu = 1, \ldots, s$, the coefficients $\mathrm{coeff}(t, g_\nu p_\nu)$ are of the same sign. Consequently, $\mathrm{HT}(f_0) = t$ and $|\mathrm{HC}(f_0)| = a$. Let i be an index such that $\mathrm{HT}(g_i p_i) = t$. Then we have $|\mathrm{HC}(g_i)| \leq |\mathrm{HC}(g_i p_i)| \leq a$, and $\mathrm{HT}(g_i)\mathrm{HT}(p_i) = t = \mathrm{HT}(f_0)$. This shows that f_0 can re reduced by g_i, that is, there exists a polynomial f_1 with $f_0 \to_{\mathrm{B}} f_1$. Since f_1 is also an element of $\langle \mathrm{B} \rangle$, we get $f_1 \overset{*}{\to}_{\mathrm{B}} 0$ by induction.

Case 2. Assume that there exist i, j such that $\mathrm{HT}(g_i p_i) = t = \mathrm{HT}(g_j p_j)$, and $\mathrm{HC}(g_i p_i)$ and $\mathrm{HC}(g_j p_j)$ have different sign.

Without loss of generality, we assume that $c_i := \mathrm{HC}(g_i) \geq \mathrm{HC}(g_j) =: c_j > 0$. Obviously, $t_i := \mathrm{HT}(g_i)$ and $t_j := \mathrm{HT}(g_j)$ are divisors of t, and thus lcm $(t_i, t_j) = s_i t_i = s_j t_j$ divides t, that is, there exists $r \in \mathrm{T}_n$ with $r s_i t_i = r s_j t_j = t$.

We consider the case $\mathrm{HC}(g_i p_i) > 0$ and $\mathrm{HC}(g_j p_j) < 0$ (the other case is similar, we just add $(-r) \cdot |w_{ij}$ instead of $r \cdot |w_{ij}$ in the definition of $|q$ below).

The vector $|q = (q_1, \ldots, q_s)^{\mathrm{T}} := |p + r \cdot |w_{ij}$ satisfies $\underline{g} \cdot |q = \underline{g} \cdot |p + r \cdot (\underline{g} \cdot |w_{ij}) = \underline{g} \cdot |q = f_0$, and we have $g_1 q_1 = g_1 p_1 + g_1 r w_1, \ldots, g_i q_i = g_i p_i + g_i r w_i - g_i r s_i, \ldots, g_j q_j = g_j p_j + g_j r w_j + g_j a r s_j, \ldots, g_s q_s = g_s p_s + g_s r w_s$.

If $\max\{\mathrm{HT}(g_1 q_1), \ldots, \mathrm{HT}(g_s q_s)\} < t$, the lemma is proved by induction since the term of BS has decreased. Otherwise, $\max\{\mathrm{HT}(g_1 q_1), \ldots, \mathrm{HT}(g_s q_s)\} = t$, and we have to calculate the coefficient of $\mathrm{BS}(g_1 q_1, \ldots, g_s q_s)$. The triangle inequality yields

$$\begin{aligned} \mathrm{BS}(g_1 q_1, \ldots, g_s q_s) \leq\; & \mathrm{BS}(g_1 p_1, \ldots, g_i p_i - g_i r s_i, \ldots, g_j p_j + g_j a r s_j, \ldots, g_s p_s) \\ & + b \cdot t, \end{aligned}$$

since $\mathrm{BS}(g_1 r w_1, \ldots, g_s r w_s) = r \cdot b \cdot s_i \cdot t_i = b \cdot t$ (for $b > 0$) or $\mathrm{BS}(g_1 r w_1, \ldots, g_s r w_s)$ has a term that is smaller than t (for $b = 0$).

We have $|\mathrm{coeff}(t, g_i p_i - g_i r s_i)| = |\mathrm{coeff}(t, g_i p_i)| - c_i$ (since $\mathrm{coeff}(t, g_i p_i) = \mathrm{HC}(g_i p_i) \geq c_i \geq 0$) and $|\mathrm{coeff}(t, g_j p_j + g_j a r s_j)| < |\mathrm{coeff}(t, g_j p_j)| + a c_j$ (since $\mathrm{coeff}(t, g_j p_j) = \mathrm{HC}(g_j p_j) < 0$ and $\mathrm{coeff}(t, g_j a r s_j) = a c_j > 0$).

Thus $\mathrm{BS}(g_1 p_1, \ldots, g_i p_i - g_i r s_i, \ldots, g_j p_j + g_j a r s_j, \ldots, g_s p_s) < \mathrm{BS}(g_1 p_1, \ldots, g_s p_s) + (a c_j - c_i) \cdot t$ and, since $c_i = a \cdot c_j + b$, $\mathrm{BS}(g_1 q_1, \ldots, g_s q_s) < \mathrm{BS}(g_1 p_1, \ldots, g_s p_s)$. This completes the proof of Proposition 6.5 by induction on BS. $\square$

The proposition shows how to decide whether a given set of polynomials is a weak Gröbner base: Just calculate the finitely many S-polynomials and try to reduce them to 0. Once we have a weak Gröbner base for I, we can decide ideal membership for I: For a given polynomial f, we apply reductions until we reach an irreducible element g (this happens because the rewrite relation is Noetherian). If f is in I, g is also in I, and thus has to reduce to 0 by the definition of weak Gröbner bases. Since g is irreducible, this means that g has to be 0. Thus, $f \in I$ iff $g = 0$ (where g is an arbitrary irreducible element obtained by reducing f).

But a weak Gröbner base can always be constructed, if a finite set of generators of I (which always exists by Hilbert's Basis Theorem) is given.

Fact 6.6. *Buchberger's algorithm.* Let I be an ideal in $\mathbb{Z}[X_1, \dots, X_n]$, and let F be a finite set of polynomials such that $\langle F \rangle = I$. As described above, we can effectively test whether F is a weak Gröbner base for I. If F is not a weak Gröbner base, we can extend F by the $\to_{\mathrm{F}}$-irreducibles of those S-polynomials that do not reduce to 0, and test again. This completion procedure always terminates with a finite weak Gröbner base for I. The termination proof is identical to the one given for example, by Kandri-Rody and Kapur [28] for the termination of their Gröbner base construction. Please note that this termination property is a consequence of Dickson's Lemma [13], which holds for free commutative monoids, but not for free monoids (see for example, [36]).

Now we are ready to describe the method for solving linear equations over $\mathbb{Z}[X_1, \dots, X_n]$. Let $(*)$ $f_1 x_1 + \cdots + f_r x_r = f_0$ be an (inhomogeneous) linear equation in $\mathbb{Z}[X_1, \dots, X_n]$. According to Section 3 we have to find one solution for $(*)$ and finitely many generators of the solutions of the homogeneous equation $(**)$ $f_1 x_1 + \cdots + f_r x_r = 0$.

First, we construct a weak Gröbner base $\mathrm{B} = \{g_1, \dots, g_s\}$ for $I := \langle \{f_1, \dots, f_r\} \rangle$. Since $\langle \mathrm{B} \rangle = I$, there exist an $r \times s$-matrix P and an $s \times r$ matrix Q with entries in $\mathbb{Z}[X_1, \dots, X_n]$ such that $\underline{f} \cdot P = \underline{g}$ and $\underline{g} \cdot Q = \underline{f}$. These matrices can easily be obtained as by-products of the weak Gröbner base construction.

Obviously, $(*)$ has a solution iff $f_0 \in I$. Hence, if $(*)$ has a solution, then f_0 reduces to 0 with respect to $\to_{\mathrm{B}}$. By keeping track of how the polynomials of B are used in this reduction process, we get polynomials $p_1, \dots, p_s \in \mathbb{Z}[X_1, \dots, X_n]$ such that $\underline{g} \cdot |p = f_0$. But then $P \cdot |p$ is a solution of $(*)$.

Now we assume that we already have finitely many generators $|z^{(1)}, \dots, |z^{(L)}$ of the solutions of the equation $(++)$ $g_1 x_1 + \cdots + g_s x_s = 0$. Then $P \cdot |z^{(1)}, \dots, P \cdot |z^{(L)}$ are solutions of $(**)$, but in general they do not generate all solutions. Let E_r be the $r \times r$ identity matrix and let $|t^{(1)}, \dots, |t^{(r)}$ be the columns of the matrix $PQ - E_r$. Since $\underline{f} \cdot (PQ - E_r) = \underline{f} \cdot PQ - \underline{f} \cdot E_r = \underline{g} \cdot Q - \underline{f} = \underline{0}$, these columns are solutions of $(**)$.

LEMMA 6.7. *The finitely many vectors* $P \cdot |z^{(1)}, \dots, P \cdot |z^{(L)}, |t^{(1)}, \dots, |t^{(r)}$ *are solutions of* $(**)$, *and they generate all solutions of this equation.*

PROOF. Let $|q = (q_1, \dots, q_r)^{\mathrm{T}}$ be an arbitrary solution of $(**)$. Then $Q \cdot |q$ is a solution of $(++)$ and thus there are $a_1, \dots, a_L \in \mathbb{Z}[X_1, \dots, X_n]$ such that $Q \cdot |q = a_1 \cdot |z^{(1)} + \cdots + a_L |z^{(L)}$. Now $|q = PQ \cdot |q - (PQ - E_r) \cdot |q = a_1 \cdot (P \cdot |z^{(1)}) + \cdots + a_L \cdot (P \cdot |z^{(L)}) + q_1 \cdot |t^{(1)} + \cdots + q_r \cdot |t^{(r)}$. $\square$

Now we show how to solve the equation $(++)$ $g_1x_1 + \cdots + g_sx_s = 0$, if $\mathrm{B} = \{g_1, \ldots, g_s\}$ is a weak Gröbner base. In fact, we already have defined the finitely many generators of all solutions of $(++)$. In the paragraph preceding Proposition 6.5, we have seen that an S-polynomial $\mathrm{S}(g_i, g_j)$ which reduces to 0 yields a solution $|w_{i,j}$ of $(++)$. Since B is a weak Gröbner base, all S-polynomials reduce to zero, and thus yield such a solution.

PROPOSITION 6.8. *The finitely many vectors $|w_{ij}$ generate all solutions of* $(++)$.

PROOF. Let $|p = (p_1, \ldots, p_s)^{\mathrm{T}}$ be a nontrivial solution of $(++)$, and let $t = \max\{\mathrm{HT}(g_1p_1), \ldots, \mathrm{HT}(g_sp_s)\}$. We prove the lemma by induction on $\mathrm{BS}(g_1p_1, \ldots, g_sp_s)$. Since $g \cdot |p = 0$, there exist i, j such that $\mathrm{HT}(g_ip_i) = t = \mathrm{HT}(g_jp_j)$, and $\mathrm{HC}(g_ip_i)$ and $\mathrm{HC}(g_jp_j)$ have different sign. Thus, the assumptions of Case 2 in the proof of Proposition 6.5 are satisfied (where $f_0 = 0$). In that proof, we have shown that one gets a new solution $|q$ from $|p$ by adding or subtracting $r \cdot |w_{ij}$, and that this new solution is smaller with respect to the complexity measure BS. Thus, the proposition is proved by induction. □

Now we have completely described a method to solve linear equations in $\mathbb{Z}[X_1, \ldots, X_n]$. In the remainder of this section, the method will be demonstrated by two examples.

Example 6.9. As an example, consider the equation $f_1x_1 + f_2x_2 + f_3x_3 = f_0$ for $f_0 = X^3YZ^2 - X^3Y^3Z^2$, $f_1 = X^3YZ - XZ^2$, $f_2 = XY^2Z - XYZ$ and $f_3 = X^2Y^2 - Z$.

First, we have to calculate a weak Gröbner base for the Ideal I, generated by f_1, f_2, and f_3. Let $<$ be the admissible ordering defined by the matrix

$$M_< = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \begin{array}{l} \text{(that means: first order by total degree and, within} \\ \qquad\qquad\quad \text{a given degree, order lexicographically} \\ \qquad\qquad\quad \text{with } X < Y < Z). \end{array}$$

With respect to this ordering, the Buchberger algorithm yields the weak Gröbner base $\mathrm{B} = \{g_1, g_2, g_3, g_4, g_5\}$, where $g_1 = f_2$, $g_2 = f_3$, $g_3 = X^2YZ - Z^2$, $g_4 = YZ^2 - Z^2$ and $g_5 = X^2Z^2 - Z^3$. By keeping track of how the g_i are generated in this process, we obtain the transformation matrix P such that $\underline{f} \cdot P = \underline{g}$ and, by reduction of the f_j with respect to $\to_{\mathrm{B}}$, we get the matrix Q such that $\underline{g} \cdot Q = \underline{f}$. In our example

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -X & XY & -ZX - X^3Y \\ 0 & 1 & Z & -YZ + Z & Z^2 + X^2YZ - X^2Z \end{pmatrix}$$

and

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We now determine whether $f_0 \in I = \langle B \rangle$, that is, whether f_0 reduces to 0 with respect to $\to_B$: $f_0 \to_B f_0 - g_5 \cdot XY = XYZ^3 - X^3Y^3Z^2 \to_B f_0 - g_5 \cdot XY + g_3 \cdot XY^2Z = XYZ^3 - XY^2Z^3 \to_B f_0 - g_5 \cdot XY + g_3 \cdot XY^2Z + g_4 \cdot XYZ = XYZ^3 - XYZ^3 = 0$.

Thus, $f_0 = g_1 \cdot 0 + g_2 \cdot 0 + g_3 \cdot (-XY^2Z) + g_4 \cdot (-XYZ) + g_5 \cdot XY \in \langle B \rangle = I$, and we can use the transformation matrix P to obtain a solution of the equation $f_1x_1 + f_2x_2 + f_3x_3 = f_0$:

$$P \cdot (0, 0, -XY^2Z, -XYZ, XY)^{\mathrm{T}} = (0, -X^2YZ - X^4Y^2, X^3Y^2Z - X^3YZ)^{\mathrm{T}}.$$

The next step is to determine the solutions $|w_{ij}$ of the equation $g_1x_1 + \cdots + g_5x_5 = 0$. $S(g_1, g_2) = g_1 \cdot X - g_2 \cdot Z = -X^2YZ + Z^2 = -g_3$, and thus $g_1 \cdot (-X) + g_2 \cdot Z + g_3 \cdot (-1) + g_4 \cdot 0 + g_5 \cdot 0 = 0$. That means $|w_{1,2} = (-X, Z, -1, 0, 0)^{\mathrm{T}}$.

We have $S(g_1, g_3) = g_1 \cdot X - g_3 \cdot Y = -X^2YZ + YZ^2 = -g_3 - Z^2 + YZ^2 = -g_3 + g_4$, and thus we get $|w_{1,3} = (-X, 0, Y - 1, 1, 0)^{\mathrm{T}}$.

Similar computations yield the other vectors $|w_{ij}$:

$$|w_{1,4} = (-Z, 0, 0, XY, 0)^{\mathrm{T}}, \qquad |w_{1,5} = (-XY, 0, -Z, YZ + Z, Y^2)^{\mathrm{T}},$$

$$|w_{2,3} = (0, -Z, Y, 1, 0)^{\mathrm{T}}, \qquad |w_{2,4} = (0, -Z^2, Z, X^2Y, 0)^{\mathrm{T}},$$

$$|w_{2,5} = (0, -Z^2, 0, YZ + Z, Y^2)^{\mathrm{T}}, \qquad |w_{3,4} = (0, 0, -Z, X^2, 1)^{\mathrm{T}},$$

$$|w_{3,5} = (0, 0, -Z, Z, Y)^{\mathrm{T}}, \qquad |w_{4,5} = (0, 0, 0, -X^2 + Z, Y - 1)^{\mathrm{T}}.$$

Now we use the transformation matrix P to obtain solutions of the homogeneous equation $f_1x_1 + f_2x_2 + f_3x_3 = 0$:

$$\begin{aligned}
P \cdot |w_{1,2} &= (0, 0, 0)^{\mathrm{T}}, & P \cdot |w_{1,3} &= (0, 0, 0)^{\mathrm{T}}, \\
P \cdot |w_{1,4} &= (0, X^2Y^2 - Z, & & \\
&\qquad -XY^2Z + XYZ)^{\mathrm{T}}, & P \cdot |w_{1,5} &= (-XY) \cdot P \cdot |w_{1,4}, \\
P \cdot |w_{2,3} &= (0, 0, 0)^{\mathrm{T}}, & P \cdot |w_{2,4} &= X \cdot P \cdot |w_{1,4}, \\
P \cdot |w_{2,5} &= P \cdot |w_{1,5} & & \\
&= (-XY) \cdot P \cdot |w_{1,4}, & P \cdot |w_{3,4} &= (0, 0, 0)^{\mathrm{T}}, \\
P \cdot |w_{3,5} &= -P \cdot |w_{2,4} & P \cdot |w_{4,5} &= P \cdot |w_{3,5} \\
&= (-X) \cdot P \cdot |w_{1,4}, & &= (-X) \cdot P \cdot |w_{1,4}.
\end{aligned}$$

The solution $P \cdot |w_{1,4} = (0, X^2Y^2 - Z, -XY^2Z + XYZ)^{\mathrm{T}}$ thus obtained does not generate all solutions of $f_1x_1 + f_2x_2 + f_3x_3 = 0$. In addition, we need the columns of the matrix

$$P \cdot Q - E_3 = \begin{pmatrix} -1 & 0 & 0 \\ -X^2 & 0 & 0 \\ XZ & 0 & 0 \end{pmatrix}.$$

Thus, all solutions of the homogeneous equation $f_1x_1 + f_2x_2 + f_3x_3 = 0$ are generated by the two solutions $(0, X^2Y^2 - Z, -XY^2Z + XYZ)^{\mathrm{T}}$ and $(-1, -X^2, XZ)^{\mathrm{T}}$.

Example 6.10. As a second example, we consider the equation $Xx_1 + Xx_2 = x_2 + X^2x_3$ of Section 4, but now we want to solve it in $\mathbb{Z}[X]$. Hence, we have to solve the homogeneous equation $f_1x_1 + f_2x_2 + f_3x_3 = 0$ for $f_1 = X$, $f_2 = X - 1$ and $f_3 = -X^2$. It is easy to see that $\langle\{f_1, f_2, f_3\}\rangle = \mathbb{Z}[X]$, and that $\mathrm{B} = \{g_1\}$ for $g_1 = 1$ is the corresponding weak Gröbner base. The transformation matrices are $P = (1, -1, 0)^{\mathrm{T}}$ and $Q = (X, X-1, -X^2)$.

Obviously, the equation $g_1x_1 = 0$ has only the trivial solution $x_1 = 0$. Thus, the columns of

$$P \cdot Q - E_3 = \begin{pmatrix} X-1 & X-1 & -X^2 \\ -X & -X & X^2 \\ 0 & 0 & -1 \end{pmatrix},$$

that is, $(X-1, -X, 0)^{\mathrm{T}}$ and $(-X^2, X^2, -1)^{\mathrm{T}}$, generate all solutions of $Xx_1 + Xx_2 = x_2 + X^2x_3$ in $(\mathbb{Z}[X])^3$.

7. *AGnH-Unification*

It is easy to see that S(AGnH) is isomorphic to the ring $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$, that is, the polynomial ring over $\mathbb{Z}$ in the noncommuting indeterminates $X_1, \ldots, X_n$. Unfortunately, for $n \geq 2$, this ring is not Noetherian (see Mora [36]), and the membership problem for finitely generated two-sided ideals is undecidable (Kandri-Rody and Weispfenning [30]). Fortunately, we are not interested in two-sided ideals, but only in right ideals. The solutions of a homogeneous equation $f_1x_1 + \cdots + f_rx_r = 0$ are only closed under right multiplication, and the inhomogeneous equation $f_1x_1 + \cdots + f_rx_r = f_0$ has a solution iff f_0 is a member of the right ideal generated by $f_1, \ldots, f_r$. Though, for $n \geq 2$, $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$ is not even right Noetherian (i.e., there are right ideals in $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$ that are not finitely generated), the set of solutions of a homogeneous equation $f_1x_1 + \cdots + f_rx_r = 0$ is a finitely generated right $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$-semimodule, and the membership problem for finitely generated right ideals is decidable in $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$ (see Section 8 and 9). This yields;

PROPOSITION 7.1. *For any $n \geq 0$, the theory AGnH is unitary with respect to unification without constants, and it is also unitary with respect to unification with constants.*

8. *Weak Gröbner Bases for Finitely Generated Right Ideals in $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$*

The construction of Gröbner bases for finitely generated right ideals in $K\langle X_1, \ldots, X_n\rangle$, where K is a field, is very easy (Mora [36], see also Apel and Lassner [1]). For $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$, one has to be more careful.

The role of terms in the commutative case is now played by words over the alphabet $\Sigma_n := \{X_1, \ldots, X_n\}$. Let W_n be the set of these words, i.e., the free monoid generated by Σ_n, and let 1 denote the empty word. For W_n, the definition of admissible term orderings as given in 6.1 is not sufficient to ensure termination of the algorithm (see 8.3). A total ordering $<$ on W_n is called (1) *right compatible* iff for all $s, t, r \in \mathrm{W}_n$, $s < t$ implies $sr < tr$, and it is called (2) *bounded* iff for all $s \in \mathrm{W}_n$ the set $\{t \in \mathrm{W}_n;\ t < s\}$ is finite. The role of the admissible orderings in the commutative case is now played by bounded, right compatible orderings.

LEMMA 8.1. *Let $<$ be a bounded, right compatible ordering on W_n.*

(1) *$<$ is order-isomorphic to ω, and thus Noetherian.*
(2) *$1 < t$ for all $t \in W_n \setminus \{1\}$.*
(3) *$s = tr$ for $r \neq 1$ implies $s > t$.*

Examples of bounded, right compatible orderings are graded lexicographical orderings, and more general, all shuffle-compatible total orders (see Leeb and Pirillo [33]). The complete characterization of all concatenation-compatible (respectively, right concatenation-compatible) linear orderings is still an open problem.

Definition 8.2. Let $<$ be a bounded, right compatible ordering on W_n.

(1) As described in 6.1 for admissible orderings on T_n, one can also extend bounded, right compatible orderings on W_n to monomials and polynomials in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$.
(2) Let f be a polynomial. We write $f = at + \mathrm{R}(f)$ if t is the maximal (with respect to $<$) word in f ($t = \mathrm{HW}(f)$) and a is the coefficient of t in f ($a = \mathrm{HC}(f)$).
(3) For a set F of polynomials in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$, the reduction relation $\to_{\mathrm{F}}$ is defined as in Section 6.2.

For $K\langle X_1, \ldots, X_n \rangle$, Mora [36] has described a very easy algorithm that transforms a finite set F of polynomials into a "Gröbner base" (see Mora [36] for the definition of Gröbner bases in this case).

Start with $F_0 := F$. As long as there are polynomials f, g in F_k, such that $\mathrm{HW}(f)$ is a prefix of $\mathrm{HW}(g)$, g can be reduced by f to a smaller polynomial g'. Define $F_{k+1} := (F_k \setminus \{g\} \cup \{g'\}$ and continue with F_{k+1} in place of F_k.

This process terminates after finitely many steps, and yields a finite set G of polynomials that generates the same right ideal as F and has the following property:

For two different elements f and g of G, $\mathrm{HW}(f)$, and $\mathrm{HW}(g)$ are not comparable with respect to the prefix-ordering (i.e., for any word r, $\mathrm{HW}(f) \cdot r \neq \mathrm{HW}(g)$ and $\mathrm{HW}(g) \cdot r \neq \mathrm{HW}(f)$).

For $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$, we encounter the following problem: For $f = a \cdot t + \mathrm{R}(f)$ and $g = b \cdot t \cdot r + \mathrm{R}(g)$ with $t, r \in \mathrm{W}_n$, $a, b \in \mathbb{Z}$ and $|a| > |b|$, $\mathrm{HW}(f)$ is prefix of $\mathrm{HW}(g)$, but the head monomial of g cannot be reduced by f. If, in addition, b divides a, it may become necessary to increase the actual set of polynomials (see Case 4 below). Since Dickson's Lemma does not hold for free monoids, we have to be very careful, if we want to obtain a terminating algorithm.

Algorithm 8.3. This is the informal description of an algorithm which transforms a finite set of polynomials $\{p_1, \ldots, p_m\} \subseteq \mathbb{Z}\langle X_1, \ldots, X_n \rangle$ into a weak Gröbner base that defines the same right ideal.

In the beginning, $F_0 := \{p_1, \ldots, p_m\}$ and all pairs of indices are unmarked.

Assume that F_k ($k \geq 0$) is already defined. If there is the zero polynomial 0 in F_k, we erase it. As long as there are $f := p_i$ and $g := p_j$ in F_k such that

(1) (i, j) is not marked and
(2) $f = a \cdot t + \mathrm{R}(f)$ and $g = b \cdot tr + \mathrm{R}(g)$ for some $a, b \in \mathbb{Z}$ and $t, r \in \mathrm{W}_n$, we do the following:

Case 1. $r = 1$.

Without loss of generality, we may assume that $|a| \geq |b|$. Let $a = bc + d$ for some c, d such that $0 \leq d < |b| \leq |a|$.

Define $f_1 := f - g \cdot c = d \cdot t + \mathrm{R}(f) - \mathrm{R}(g) \cdot c$ and $F_{k+1} := (F_k \setminus \{f\}) \cup \{f_1\}$. We do not have to mark (i, j) since $f = p_i$ is removed.

Obviously, $f_1 < f$ and $f = f_1 + g \cdot c$. Hence, F_{k+1} generates the same right ideal as F_k, but f is replaced by the smaller polynomial f_1.

Case 2. $r \neq 1$ and $|a| \leq |b|$.

Let $b = ac + d$ for some c, d such that $0 \leq d < |a| \leq |b|$. Define $g_1 := g - f \cdot cr = d \cdot tr + \mathrm{R}(g) - \mathrm{R}(f) \cdot cr$ and $F_{k+1} := (F_k \setminus \{g\}) \cup \{g_1\}$.

Obviously, $g_1 < g$ and $g = g_1 + f \cdot cr$. Hence, F_{k+1} generates the same right ideal as F_k, but g is replaced by the smaller polynomial g_1.

Case 3. $r \neq 1$, $|a| > |b|$ and $|b|$ does not divide $|a|$.

Let $a = bc + d$ for some c, d such that $0 < d < |b| < |a|$. We define $g_1 := f \cdot r - g \cdot c = d \cdot tr + \mathrm{R}(f) \cdot r - \mathrm{R}(g) \cdot c$. Since the words occurring in $\mathrm{R}(f) \cdot r$ and $\mathrm{R}(g) \cdot c$ are smaller than tr, we have $\mathrm{HW}(g_1) = tr$, $\mathrm{HC}(g_1) = d$ and $\mathrm{R}(g_1) = \mathrm{R}(f) \cdot r - \mathrm{R}(g) \cdot c$. Obviously, $g_1 < g$, $g_1 \in \langle F_k \rangle$ and the pair g_1, g satisfies Case 1. Hence, we define $g_2 := g - g_1 \cdot c_1$ (where $b = dc_1 + d_1$, $0 \leq d_1 < d$) and $F_{k+1} := (F_k \setminus \{g\}) \cup \{g_1, g_2\}$. Since $g_1, g_2 < g$ and $g = g_2 + g_1 \cdot c$, F_{k+1} generates the same right ideal as F_k, but g is replaced by the two smaller polynomials g_1 and g_2.

Case 4. $r \neq 1$, $|a| > |b|$ and $|b|$ divides $|a|$, that is, there exists c such that $a = bc$.

Define $g_1 := f \cdot r - g \cdot c = \mathrm{R}(f) \cdot r - \mathrm{R}(g) \cdot c$. Now $g_1 < g$, but since $|c| \neq 1$, g cannot be represented using g_1 and f. Thus, the problem is that we should like to add g_1, but we are not allowed to remove the larger polynomial g since this would possibly change the generated right ideal. We distinguish the following cases:

Case 4.1. There is $h \in \bigcup_{i \leq k} F_i$ with the property $\mathrm{HW}(g_1) = \mathrm{HW}(h)$.

We choose h such that $|\mathrm{HC}(h)|$ is minimal.

Case 4.1.1. $h \in F_k$ and $|\mathrm{HC}(g_1)| < |\mathrm{HC}(h)|$.

We have $g_1 < h$ and h may be reduced by g_1 to some $h_1 < h$ (see Case 1). Define $F_{k+1} := (F_k \setminus \{h\}) \cup \{g_1, h_1\}$ and mark (i, j). F_{k+1} generates the same right ideal as F_k, but h is replaced by the two smaller polynomials g_1 and h_1.

Case 4.1.2. $h \in F_k$ and $|\mathrm{HC}(g_1)| \geq |\mathrm{HC}(h)|$.

Then g_1 may be reduced by h to a smaller polynomial g_2 (see Case 1). If $g_2 = 0$, $F_{k+1} := F_k$ and we mark (i, j). Otherwise we continue with g_2 in place of g_1.

Case 4.1.3. $h \notin F_k$.

That means that $h \in F_i$ for some $i < k$, but h has been removed in some iteration of the algorithm between step i and step k.

First, assume that $|\mathrm{HC}(g_1)| \geq |\mathrm{HC}(h)|$. Then, the head monomial HC $(g_1)\mathrm{HW}(g_1)$ of g_1 can be reduced by h, and thus is $\rightarrow_{F_i}$-reducible. We want to show that $\mathrm{HC}(g_1)\mathrm{HW}(g_1)$ can also be reduced by $\rightarrow_{F_k}$.

To that purpose, assume that the monomial $a \cdot r$ is reducible by some polynomial $p = b \cdot s + \mathrm{R}(p) \in F_j$, that is, $r = ss'$ for some words s' and $|a| \geq |b|$. If p is in F_{j+1}, then $a \cdot r$ is also reducible with respect to $\rightarrow_{F_{j+1}}$. Assume that $p \notin F_{j+1}$. By considering the cases where a polynomial is removed, one finds that F_{j+1} contains a polynomial $q = c \cdot t + \mathrm{R}(q)$ that reduces the head monomial of p, that is, $s = tt'$ for some word t' and $|b| \geq |c|$. But then $r = ss' = t(t's')$ and $|a| \geq |b| \geq |c|$ yield that q reduces $a \cdot r$.

Thus, if $\mathrm{HC}(g_1)\mathrm{HW}(g_1)$ can be reduced with respect to $\rightarrow_{F_i}$, it can also be reduced with respect to $\rightarrow_{F_{i+1}}, \ldots, \rightarrow_{F_k}$.

To sum up, we know that for $|\mathrm{HC}(g_1)| \geq |\mathrm{HC}(h)|$, g_1 can be reduced by $\rightarrow_{F_k}$. Thus, we can proceed as in Case 4.1.2.

Otherwise, that is, if $|\mathrm{HC}(g_1)| < |\mathrm{HC}(h)|$, we define $F_{k+1} := F_k \cup \{g_1\}$ and mark (i, j).

Case 4.2. There is no $h \in \bigcup_{i \leq k} F_i$ with the property $\mathrm{HW}(g_1) = \mathrm{HW}(h)$.

In this case, we also define $F_{k+1} := F_k \cup \{g_1\}$ and mark (i, j).

This completes the description of Algorithm 8.3. We shall soon show that this algorithm always terminates with a finite set of polynomials G whose properties justify the name weak Gröbner base. But first, we consider an example.

Example 8.4. Let $f_1 = 2abc - bc$, $f_2 = 3ab - 2b$, $f_3 = 5abd - bc$ and $f_4 = bc - 5bd$ be polynomials in $\mathbb{Z}\langle a, b, c, d\rangle$. We take the graded lexicographical ordering with $a > b > c > d$ as bounded, right compatible ordering (i.e., $u < v$ iff $|u| < |v|$ or $|u| = |v|$ and $u <_{\mathrm{lex}} v$), and *run Algorithm* 8.3 *with input* $F_0 := \{f_1, f_2, f_3, f_4\}$.

(1) For f_1 and f_2, we have Case 3.
Define $f_5 := f_2 \cdot c - f_1 = abc - bc$ and $f_6 := f_1 - f_5 \cdot 2 = bc$. Now f_1 is replaced by f_5, f_6, which yields $F_1 = \{f_2, f_3, f_4, f_5, f_6\}$. We have $f_1 = f_5 \cdot 2 + f_6$.
(2) For f_2 and f_3, we have Case 2.
Define $f_7 := f_3 - f_2 \cdot d = 2abd - bc + 2bd$ and replace f_3 by f_7, which yields $F_2 = \{f_2, f_4, f_5, f_6, f_7\}$. We have $f_3 = f_7 + f_2 \cdot d$.
(3) For f_2 and f_5, we have Case 4.
Define $f_8 = f_2 \cdot c - f_5 \cdot 3 = bc = f_6$. Hence, we have Case 4.1.2, and since f_6 reduces f_8 to 0, $F_3 = F_2 = \{f_2, f_4, f_5, f_6, f_7\}$, and the index pair (2,5) is marked.
(4) For f_2 and f_7, we have Case 3.
Define $f_9 := f_2 \cdot d - f_7 = abd - 4bd + bc$ and $f_{10} = f_7 - f_9 \cdot 2 = -3bc + 10bd$. Now f_7 is replaced by f_9 and f_{10}, which yields $F_4 = \{f_2, f_4, f_5, f_6, f_9, f_{10}\}$. We have $f_7 = f_{10} + f_9 \cdot 2$.

(5) For f_2 and f_9, we have Case 4.
Define $f_{11} := f_2 \cdot d - f_9 \cdot 3 = -3bc + 10bd$. Now $\mathrm{HW}(f_{11}) = \mathrm{HW}(f_4)$ and f_4 reduces f_{11} to the polynomial $f_{12} := f_{11} + f_4 \cdot 3 = -5bd$ (Case 4.1.2). We continue with f_{12} in place of f_{11}, and have Case 4.2 since bd has not yet occurred as head word. Hence, $F_5 := F_4 \cup \{f_{12}\}$ and $(2,5)$ and $(2,9)$ are already marked.

(6) For f_4 and f_6, we have Case 1.
Define $f_{13} := f_4 - f_6 = f_{12}$ and $F_6 := F_5 \setminus \{f_4\} = \{f_2, f_5, f_6, f_9, f_{10}, f_{12}\}$.

(7) For f_6 and f_{10}, we have Case 1.
Define $f_{14} := f_{10} + f_6 \cdot 3 = 10bd$ and $F_7 := \{f_2, f_5, f_6, f_9, f_{12}, f_{14}\}$.

(8) For f_{12} and f_{14}, we have Case 1.
Since $f_{14} = f_{12} \cdot (-2)$, f_{14} can be eliminated and we get $F_8 = \{f_2, f_5, f_6, f_9, f_{12}\}$, where $(2,5)$ and $(2,9)$ are marked.

Hence, Algorithm 8.3 terminates with $G := F_8 = \{f_2, f_5, f_6, f_9, f_{12}\}$. The elements of G are $g_1 := f_2 = 3ab - 2b$, $g_2 := f_5 = abc - bc$, $g_3 := f_6 = bc$, $g_4 := f_9 = abd - 4bd + bc$, and $g_5 := f_{12} = -5bd$.

LEMMA 8.5. *For any finite input set $F_0 = \{f_1, \ldots, f_m\}$ of polynomials, Algorithm* 8.3 *always terminates.*

PROOF. We consider the F_k's as multisets of polynomials which are ordered by the multiset ordering $\ll$ induced by the ordering $<$ on polynomials (see Definition 8.2). Since $<$ is well-founded, the multiset extension $\ll$ is also well-founded.

For the Cases 1, 2, 3, and 4.1.1, $F_k \gg F_{k+1}$. Case 4.1.2 and the corresponding subcase of 4.1.3 cannot occur infinitely often in successive steps because then $g_1 > g_2 > g_3 > \cdots$ would be an infinite descending $<$-chain. That means that after finitely many steps $g_i = 0$ or Case 4.1.1, the other subcase of 4.1.3 or Case 4.2 occur.

For the Cases 4.1.3 and 4.2, F_{k+1} is larger than F_k. But these cases can only occur finitely often during the whole run of the algorithm. First note that all words t occurring in some polynomial of some F_k satisfy $t \leq \max\{\mathrm{HW}(f_1), \ldots, \mathrm{HW}(f_m)\}$. Since $<$ is bounded, there are only finitely many words with this property. Hence, Case 4.2 can only occur finitely often. Case 4.1.3—where a head word which has disappeared in some former step appears again—can only occur finitely often for a certain word because the absolute value of the head coefficient gets smaller each time. □

Before we can state the next lemma, we have to introduce a new notation (or rather an abuse of the usual notation). Let F be a finite set of polynomials. The expression

$$f = \sum_{h_i \in F} h_i \cdot a_i,$$

should be interpreted as follows: the a_i are monomials in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$, f is a finite sum of the polynomials $h_i \cdot a_i$, but *an element of F may occur more than once* in this sum, and each occurrence may have a different coefficient a_i.

LEMMA 8.6. *Let $t \in W_n$ be a word, and F_k be the set of polynomials obtained after some iterations of Algorithm* 8.3. *Assume that h is a polynomial, and that $h = \sum_{h_i \in F_k} h_i \cdot a_i$ for monomials a_i with $HW(h_i \cdot a_i) < t$. Then $h =$*

$\Sigma_{h'_i \in F_{k+1}} h'_i \cdot b_i$ *for monomials* b_i *with* $HW(h'_i \cdot b_i) < t$.

PROOF. For the Cases 4.1.3 and 4.2, we have $F_k \subseteq F_{k+1}$, and thus we can use the given sum. In Case 1, $F_{k+1} := (F_k \setminus \{f\}) \cup \{f_1\}$ and $f = f_1 + g \cdot c$. In addition, we have $g \in F_{k+1}$ and $\mathrm{HW}(g) = \mathrm{HW}(f) \geq \mathrm{HW}(f_1)$. Thus a term $f \cdot a_j$ in the sum $h = \Sigma_{h_i \in F_k} h_i \cdot a_i$ can be replaced by $f_1 \cdot a_j + g \cdot ca_j$. The other cases can be treated similarly. □

The next lemma will play a role that is similar to the one played by Lemma 6.4 in the commutative case.

LEMMA 8.7. *Let* G *be the output of Algorithm* 8.3 (*i.e., the actual set* F_k *when the algorithm terminates*) *and let* $f = a \cdot t + R(f)$ *and* $g = b \cdot tr + R(g)$ *be elements of* G. *Then the following holds*:

(1) $a = bc$ *for some* $c \in \mathbb{Z}$, $|c| \neq 1$ *and* $r \neq 1$.
(2) *The S-polynomial* $g_1 := f \cdot r - g \cdot c = R(f) \cdot r - R(g) \cdot c$ *can be obtained as a finite sum*

$$g_1 = \sum_{h_i \in G} h_i \cdot a_i,$$

where the a_i *are monomials in* $\mathbb{Z}\langle X_1, \ldots, X_n\rangle$ *and* $HW(h_i \cdot a_i) \leq HW(g_1) < HW(g) = HW(f \cdot r)$.

PROOF. Since Algorithm 8.3 has terminated, the index pair corresponding to f and g is marked. Thus, for some k, f and g are in F_k and they are selected by the algorithm.

(1) Property (1) of the lemma is satisfied, since, only in Case 4, both f and g remain in F_{k+1}.
(2) In Case 4 we have $g_1 := f \cdot r - g \cdot c = \mathrm{R}(f) \cdot r - \mathrm{R}(g) \cdot c$, and thus $\mathrm{HW}(g_1) < \mathrm{HW}(g) = \mathrm{HW}(f \cdot r) = tr$. There is some g_i such that $g_1 \overset{*}{\rightarrow}_{F_k} g_i$ (see Case 4.1.2 and the first subcase of 4.1.3) and $g_i \in F_{k+1}$ or $g_i = 0$. Hence, $\mathrm{HW}(g_i) \leq \mathrm{HW}(g_1)$ and $g_1 = g_i + \Sigma_{h_i \in F_k} h_i \cdot a_i$ for monomials a_i with $\mathrm{HW}(h_i \cdot a_i) \leq \mathrm{HW}(g_1)$. Lemma 8.6 yields $g_1 = g_i + \Sigma_{h'_i \in F_{k+1}} h'_i \cdot b_i$ for monomials b_i with $\mathrm{HW}(h'_i \cdot b_i) \leq \mathrm{HW}(g_1)$, and since $g_i \in F_{k+1}$ or $g_i = 0$ we have $g_1 = \Sigma_{h''_i \in F_{k+1}} h''_i \cdot c_i$ for monomials c_i with $\mathrm{HW}(h''_i \cdot c_i) \leq \mathrm{HW}(g_1)$. By Lemma 8.6, g_1 can be represented by such a sum for all F_m with $m \geq k + 1$. Thus, we have proved the lemma. □

Let $F \subseteq \mathbb{Z}\langle X_1, \ldots, X_n\rangle$ be a set of polynomials. In the following, $\langle F\rangle$ denotes the right ideal generated by F.

LEMMA 8.8. *Let* G *be the output of Algorithm* 8.3 *if started with input* F_0. *Then* $\langle G\rangle = \langle F_0\rangle$.

PROOF. It has already been pointed out during the description of the algorithm that in any case $\langle F_k\rangle = \langle F_{k+1}\rangle$. □

This lemma and the next proposition shows that it is reasonable to call the result of Algorithm 8.3 a weak Gröbner base.

PROPOSITION 8.9. *Let* G *be the output of Algorithm* 8.3. *Then any* $f \in \langle G\rangle$ *can be reduced to* 0 *with respect to* $\rightarrow_G$.

PROOF. The proof is similar to the proof of Lemma 2.4 in Mora [36], and the proof of Proposition 6.5 above. Obviously, $f \in \langle G\rangle$ means $f = \Sigma_{g_i \in G} g_i \cdot a_i$

for some monomials a_i. If $f = 0$, then there is nothing to prove. Otherwise, let $t := \max\{\cdots \mathrm{HW}(g_i \cdot a_i) \cdots\}$ and $I := \{i;\ \mathrm{HW}(g_i \cdot a_i) = t\}$.

Case 1. $|I| = 1$. Then $\mathrm{HW}(f) = t$ and (for $I = \{j\}$ and $a_j = c_j \cdot r_j$ ($c_j \in \mathbb{Z}$, $r_j \in \mathrm{W}_n$)) $\mathrm{HW}(f) = t = \mathrm{HW}(g_j) \cdot r_j$ and $\mathrm{HC}(f) = \mathrm{HC}(g_j) \cdot c_j$. Hence, f can be reduced by g_j to the smaller polynomial $f_1 := f - g_j \cdot a_j \in \langle G \rangle$. By induction we get $f_1 \xrightarrow{*}_{\mathrm{G}} 0$ and thus $f \rightarrow_{\mathrm{G}} f_1 \xrightarrow{*}_{\mathrm{G}} 0$.

Case 2. $|I| > 1$. Let i, j be two different elements of I, and let $a_i = c_i \cdot r_i$, $a_j = c_j \cdot r_j$ ($c_i, c_j \in \mathbb{Z}$, $r_i, r_j \in \mathrm{W}_n$). Since $\mathrm{HW}(g_i) \cdot r_i = t = \mathrm{HW}(g_j) \cdot r_j$, either $\mathrm{HW}(g_i)$ is a prefix of $\mathrm{HW}(g_j)$ or vice versa. Without loss of generality we assume $\mathrm{HW}(g_i) = \mathrm{HW}(g_j) \cdot r$ for some $r \in W_n$. By Lemma 8.7, $\mathrm{HC}(g_j) = \mathrm{HC}(g_i) \cdot c$ for some $c \in \mathbb{Z}$, and $g_j \cdot r - g_i \cdot c = \sum_{h_k \in G} h_k \cdot b_k$ where $\mathrm{HW}(h_k \cdot b_k) < \mathrm{HW}(g_i) = \mathrm{HW}(g_j \cdot r)$. Hence, $g_j \cdot r_j - g_i \cdot r_i c = (g_j \cdot r - g_i \cdot c) \cdot r_i = \sum_{h_k \in G} h_k \cdot (b_k r_i)$, where $\mathrm{HW}(h_k \cdot (b_k r_i)) < \mathrm{HW}(g_i) \cdot r_i = t$.

Now,

$$
\begin{aligned}
f &= (g_j \cdot r_j - g_i \cdot r_i c) \cdot c_j + g_i \cdot (c_i + c c_j) r_i + \sum_{\nu \neq i, j} g_\nu \cdot a_\nu \\
&= \sum_{h_k \in G} h_k \cdot (b_k c_j r_i) + g_i \cdot (c_i + c c_j) r_i + \sum_{\nu \neq i, j} g_\nu \cdot a_\nu
\end{aligned}
$$

yields a representation of f as a sum where $|I|$ is smaller. □

COROLLARY 8.10. *The membership problem for finitely generated right ideals in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$ is decidable.*

PROOF. Let $I = \langle \{p_1, \ldots, p_m\} \rangle$ be a finitely generated right ideal in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$. We apply Algorithm 8.3 to $F_0 = \{p_1, \ldots, p_m\}$, and get a set G of polynomials. Now $f \in I$ iff f can be reduced to 0 with respect to $\rightarrow_{\mathrm{G}}$. If f is $\rightarrow_{\mathrm{G}}$-irreducible, then $f \in I$ iff $f = 0$. Otherwise, we can effectively find some g such that $f \rightarrow_{\mathrm{G}} g$ and $f \in I$ iff $g \in I$. Thus, Corollary 8.10 is proved by induction. □

9. *Solving Linear Equations in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$*

In the previous section, we have shown how to compute weak Gröbner bases for finitely generated right ideals in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$. In this section, these bases are used to solve linear equations in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$. The method is very similar to the one described in Section 6.

Let $(*)$ $f_1 x_1 + \cdots + f_m x_m = f_0$ be an (inhomogeneous) linear equation in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$. We have to find one solution for $(*)$ and finitely many generators of the solutions of the homogeneous equation $(**)$ $f_1 x_1 + \cdots + f_m x_m = 0$.

Let $G = \{g_1, \ldots, g_s\}$ be the output of Algorithm 8.3 when started with input $\{f_1, \ldots, f_m\}$. There exist an $m \times s$-matrix P and an $s \times m$-matrix Q with entries in $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$ such that $\underline{f} \cdot P = \underline{g}$ and $\underline{g} \cdot Q = \underline{f}$. These matrices can be obtained as by-products of Algorithm 8.3.

Obviously, $(*)$ has a solution iff $f_0 \in \langle \{f_1, \ldots, f_m\} \rangle = \langle G \rangle$. Hence, if $(*)$ has a solution, Proposition 8.9 implies that f_0 reduces to 0 with respect to $\rightarrow_{\mathrm{G}}$. By keeping track of how the polynomials of B are used in this reduction process, we get $p_1, \dots, p_s \in \mathbb{Z}[X_1, \dots, X_n]$ such that $\underline{g} \cdot |p = f_0$. But then $P \cdot |p$ is a solution of $(*)$.

We now assume that we already have finitely many generators $|z^{(1)}, \dots, |z^{(L)}$ of the set of solutions of the equation

$$(++) g_1 x_1 + \cdots + g_s x_s = 0.$$

As in Section 6, one can show

LEMMA 9.1. *The vectors $P \cdot |z^{(1)}, \dots, P \cdot |z^{(L)}$ and the columns of the matrix $PQ - E_m$ are solutions of $(**)$, and they generate all solutions of this equation.*

We now show how to compute the finitely many generators of the solutions of $(++)$. If there do not exist i, j $(i \neq j)$ such that $\mathrm{HW}(g_i) = \mathrm{HW}(g_j) \cdot r$ for some $r \in \mathrm{W}_n$, the equation $(++)$ has no nontrivial solutions. Otherwise, let i, j $(i \neq j)$ be indices such that $\mathrm{HW}(g_i) = \mathrm{HW}(g_j) \cdot r$ for some $r \in \mathrm{W}_n$.

By Lemma 8.7, $\mathrm{HC}(g_j) = \mathrm{HC}(g_i) \cdot c$ for some $c \in \mathbb{Z}$, $r \neq 1$, and $g_j \cdot r - g_i \cdot c = \sum_{k=1}^{k=s} g_k \cdot h_k$ for *polynomials* $h_k \in \mathbb{Z}\langle X_1, \dots, X_n \rangle$ with $\mathrm{HW}(g_k \cdot h_k) < \mathrm{HW}(g_i)$. Obviously, h_i has to be 0. If we define $q_k := h_k$ for $k \neq i, j, q_i := h_i + c = c$, and $q_j := h_j - r$, then $|q_{ij} := (q_1, \dots, q_s)^{\mathrm{T}}$ is a solution of $(++)$.

LEMMA 9.2. *The finitely many vectors $|q_{ij}$ generate all solutions of* $(++)$.

PROOF. Let $|p = (p_1, \dots, p_s)^{\mathrm{T}}$ be a nontrivial solution of $(++)$. The complexity of such a solution is given by (t, α) where $t := \max\{\mathrm{HW}(g_i p_i); 1 \leq i \leq s\}$ and $\alpha := |\{i; 1 \leq i \leq s \text{ and } \mathrm{HW}(g_i p_i) = t\}|$.

Since $\underline{g} \cdot |p = 0$ and $|p$ is not trivial, α has to be greater than 1. Hence there exist i, j $(i \neq j)$ such that $\mathrm{HW}(g_i)\mathrm{HW}(p_i) = t = \mathrm{HW}(g_j)\mathrm{HW}(p_j)$. Without loss of generality, we assume that $\mathrm{HW}(g_j)$ is a prefix of $\mathrm{HW}(g_i)$. Thus, $\mathrm{HW}(g_i) = \mathrm{HW}(g_j) \cdot r$ and $\mathrm{HC}(g_j) = \mathrm{HC}(g_i) \cdot c$ for some $r \in \mathrm{W}_n$ and $c \in \mathbb{Z}$, and $\mathrm{HW}(p_j) = r \cdot \mathrm{HW}(p_i)$. Let $c_i := \mathrm{HC}(p_i)$ and $c_j := \mathrm{HC}(p_j)$.

The vector $|q_{ij}$ that was defined above is a solution of $(++)$. We define a new solution $(p'_1, \dots, p'_s)^{\mathrm{T}} = |p' := |p + |q_{ij} \cdot c_j \, \mathrm{HW}(p_i)$, and show that it has smaller complexity than $|p$. To that purpose, we have to consider the words $\mathrm{HW}(g_k p'_k)$ for all k, $1 \leq k \leq s$.

Case 1. $k \neq i, j$. We have $g_k p'_k = g_k p_k + g_k h_k c_j \, \mathrm{HW}(p_i)$ and $\mathrm{HW}(g_k \cdot h_k) < \mathrm{HW}(g_i)$. This implies that $\mathrm{HW}(g_k h_k c_j \, \mathrm{HW}(p_i)) < \mathrm{HW}(g_i)\mathrm{HW}(p_i) = t$. Thus, $\mathrm{HW}(g_k p'_k) = t$ iff $\mathrm{HW}(g_k p_k) = t$.

Case 2. $k = i$. We have $g_i p'_i = g_i p_i + g_i c c_j \, \mathrm{HW}(p_i)$. Hence, $\mathrm{HW}(g_i p'_i) = t$ if $c_i + c c_j \neq 0$, and $\mathrm{HW}(g_i p'_i) < t$ if $c_i + c c_j = 0$.

Case 3. $k = j$.

$$\begin{aligned} g_j p'_j &= g_j p_j + g_j h_j c_j \, \mathrm{HW}(p_i) - g_j r c_j \, \mathrm{HW}(p_i) \\ &= \mathrm{HC}(g_j) c_j t + \mathrm{R}(g_j p_j) + g_j h_j c_j \, \mathrm{HW}(p_i) \\ &\quad - \mathrm{HC}(g_j) c_j \, \mathrm{HW}(g_j) r \, \mathrm{HW}(p_i) - \mathrm{R}(g_j) r c_j \, \mathrm{HW}(p_i) \\ &= \mathrm{R}(g_j p_j) + g_j h_j c_j \, \mathrm{HW}(p_i) - \mathrm{R}(g_j) r c_j \, \mathrm{HW}(p_i) \end{aligned}$$

since $r\mathrm{HW}(p_i) = \mathrm{HW}(g_j)$.

This shows that $\mathrm{HW}(g_j p'_j) < t$.

Thus, we have seen that the complexity of the solution $|p'$ is smaller than the complexity of $|p$, and the lemma is proved by induction. □

Example 9.3. As an example, we consider the homogeneous linear equation $f_1x_1 + \cdots + f_4x_4 = 0$ in $\mathbb{Z}\langle a,b,c,d\rangle$ for the polynomials $f_1 = 2abc - bc$, $f_2 = 3ab - 2b$, $f_3 = 5abd - bc$ and $f_4 = bc - 5bd$ of Example 8.4.

We have seen that Algorithm 8.3 terminates with $G = \{g_1, g_2, g_3, g_4, g_5\}$ where $g_1 = 3ab - 2b$, $g_2 = abc - bc$, $g_3 = bc$, $g_4 = abd - 4bd + bc$, and $g_5 = -5bd$. The transformation matrices P, Q such that $\underline{f} \cdot P = \underline{g}$ and $\underline{g} \cdot Q = \underline{f}$ are

$$Q = \begin{pmatrix} 0 & 1 & d & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & -3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 1 \end{pmatrix}$$

and

$$P = \begin{pmatrix} 0 & -1 & 3 & 0 & 0 \\ 1 & c & -2c & 2d & -5d \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

All solutions of the equation $g_1x_1 + \cdots + g_5x_5 = 0$ are generated by $|q_{1,2}$ and $|q_{1,4}$:

(1) $g_1 \cdot c - g_2 \cdot 3 = g_3$, and thus $|q_{1,2} = (-c, 3, 1, 0, 0)^{\mathrm{T}}$.
(2) $g_1 \cdot d - g_4 \cdot 3 = f_{11} = f_{12} - f_4 \cdot 3 = f_{12} - (f_6 + f_{12}) \cdot 3 = f_{12} \cdot (-2) + f_6(-3) = g_5 \cdot (-2) + g_3(-3)$, and thus $|q_{1,4} = (-d, 0, -3, 3, -2)^{\mathrm{T}}$.

The matrix $PQ - E_4$ is

$$\begin{pmatrix} 0 & 0 & -9 & 3 \\ 0 & 0 & 6c + 15d & -2c - 5d \\ 0 & 0 & -9 & 3 \\ 0 & 0 & -6 & 2 \end{pmatrix}.$$

This yields the new solution $(3, -2c - 5d, 3, 2)^{\mathrm{T}}$ and since $|q_{1,4} = (3, -2c - 5d, 3, 2)^{\mathrm{T}} \cdot (-3)$, the solution $(3, -2c - 5d, 3, 2)^{\mathrm{T}}$ generates all solutions of $f_1x_1 + \cdots + f_4x_4 = 0$ in $\mathbb{Z}\langle a,b,c,d\rangle$.

10. *Conclusion*

The categorical reformulation of E-unification allows to characterize the class of commutative theories by properties of the category $C(\mathrm{E})$ of finitely generated E-free objects: $C(\mathrm{E})$ has to be a semiadditive category. The definition of semiadditive categories provides an algebraic structure on the morphism sets that can be used to obtain algebraic characterizations of the unification types. This shows the connection between unification in commutative theories and equation solving in linear algebra. The very common syntactical approach to equational unification, which only uses the defining axioms, is thus replaced by a more semantic approach, which works with algebraic properties of the defined algebras.

Hence, unification algorithms for the commutative theory AGnHC, that is, the theory of Abelian groups with n commuting homomorphisms, can be derived by applying well-known algebraic methods (e.g., Gröbner Base algorithms) to solve linear equations in $\mathbb{Z}[X_1, \ldots, X_n]$. In order to obtain a unification algorithm for the theory AGnH of Abelian groups with n noncommuting homomorphisms, we have developed a Gröbner base algorithm for the ring $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$ of polynomials over $\mathbb{Z}$ in n noncommuting indeterminates. Since Dickson's Lemma (Dickson [13]), which is used for $\mathbb{Z}[X_1, \ldots, X_n]$ to prove termination of the Gröbner Base algorithm, does not hold for $\mathbb{Z}\langle X_1, \ldots, X_n \rangle$, we had to be very careful to obtain a terminating algorithm. As in the commutative case, the performance of the algorithm depends on the choice of the ordering. Hence, it would be very interesting to have a complete characterization of all bounded, right compatible orderings for W_n.

ACKNOWLEDGMENTS. I should like to thank the anonymous referees for their comments and suggestions; in particular for calling my attention to the difference between weak Gröbner bases and Gröbner bases.

REFERENCES

1. APEL, J., AND LASSNER, W. An extension of Buchberger's algorithm and calculations in enveloping fields of lie algebras. *J. Symb. Computation 6* (1988), 361–370.
2. BAADER, F. The theory of idempotent semigroups is of unification type zero. *J. Automat. Reas. 2* (1986), 283–286.
3. BAADER, F. Unification in varieties of idempotent semigroups. *Semigroup Forum 36* (1987), 127–145.
4. BAADER, F. Unification in commutative theories. *J. Symb. Computation 8* (1989), 479–497.
5. BAADER, F. Unification properties of commutative theories: A categorical treatment. In *Proceedings of the Summer Conference on Category Theory and Computer Science* (Manchester, England). 1989, pp. 273–299.
6. BAADER, F., AND BÜTTNER, W. Unification in commutative idempotent monoids. *Theoret. Comput. Sci. 56* (1988), 345–353.
7. BAADER, F., AND NUTT, W. Adding homomorphisms to commutative/monoidal theories, or how algebra can help in equational unification. In *Proceedings of the RTA'91* (Como, Italy). Lecture Notes in Computer Science, vol. 488. Springer-Verlag, New York, 1991, pp. 124–135.
8. BACHMAIR, L., AND BUCHBERGER, B. A simplified proof of the characterization theorem of Gröbner-bases. *ACM-SIGSAM Bulletin 14* (1980), 29–34.
9. BUCHBERGER, B. Gröbner bases: An algorithmic method in polynomial ideal theory. In *Recent Trends in Multidimensional System Theory*, N. K. Bose, ed., Reidel, Dordrecht, Germany, 1985, pp. 184–232.
10. BÜTTNER, W. Unification in the data structure multiset. *J. Automat. Reas. 2* (1986), 75–88.
11. COHN, P. M. *Universal Algebra*. Harper and Row, New York, 1965.
12. DERSHOWITZ, N., AND MANN, Z. Proving termination with multiset orderings. *Commun. ACM 22*, 4 (Apr. 1979), 465–475.
13. DICKSON, L. E. Finiteness of the odd perfect and primitive abundant numbers with n distinct factors. *Amer. J. Math. 35* (1913), 413–422.
14. EILENBERG, S. *Automata, Languages and Machines, Volume A*. Academic Press, Orlando, Fla., 1974.
15. FAGES, F. Associative-commutative unification. In *Proceedings of the CADE '84* (Napa, Fla.). Lecture Notes in Computer Science, vol. 170. Springer-Verlag, New York, 1984, pp. 205–220.
16. FAGES, F., AND HUET, G. Complete sets of unifiers and matchers in equational theories. *Theoret Comput. Sci. 43* (1986), 189–200.
17. FORTENBACHER, A. An algebraic approach to unification under associativity and commutativity. In *Proceedings of the RTA '85* (Dijon, France). Lecture Notes in Computer Science, vol. 202. Springer-Verlag, New York, 1985, pp. 381–397.
18. FREYD, P. *Abelian Categories*. Harper and Row, New York, 1964.
19. FURUKAWA, A., SASAKI, T., AND KOBAYASHI, H. Gröbner basis of a module over $K[X_1, \ldots, X_n]$ and polynomial solutions of systems of linear equations. In *Proceedings of the*

Symsac '86 (Waterloo, Ont., Canada). ACM, New York, 1986, pp. 222–224. (An extended version is available as internal report.)

20. GRÄTZER, G. *Universal Algebra*. Van Nostrand Company, Princeton, N.J., 1968.
21. HEROLD, A. Combination of unification algorithms in equational theories, Ph.D. Dissertation, Universität Kaiserslautern, 1987.
22. HERRLICH, H., AND STRECKER, G. E. *Category Theory*. Allyn and Bacon, Boston, Mass., 1973.
23. HUET, G. Confluent reductions: Abstract properties and applications to term rewriting systems. *J. ACM 27*, 4 (Oct. 1980), 797–821.
24. JACOBSON, N. *Basic Algebra II*. Freeman and Company, San Francisco, Calif., 1980.
25. JAFFAR, J., LASSEZ, J. L., AND MAHER, M. J. A theory of complete logic programs with equality. *J. Logic Prog. 1* (1984), 175–184.
26. JOUANNAUD, J. P., AND KIRCHNER, H. Completion of a set of rules modulo a set of equations. *SIAM J. Comput. 15* (1986), 1155–1194.
27. KANDRI-RODY, A., AND KAPUR, D. Algorithms for computing the Gröbner basis of polynomial ideals over various Euclidean rings. In *Proceedings of EUROSAM 84*. Lecture Notes in Computer Science, vol. 174. Springer-Verlag, New York, 1984, pp. 197–206.
28. KANDRI-RODY, A., AND KAPUR, D. An algorithms for computing the Gröbner basis of a polynomial ideal over a Euclidean ring. General Electric Research and Development Report No. 84CRD045. General Electric, Schnectady, N.Y.
29. KANDRI-RODY, A., AND KAPUR, D. Computing a Gröbner basis of a polynomial ideal over a Euclidean domain. *J. Symb. Comput. 6*, (1988), 37–57.
30. KANDRI-RODY, A., AND WEISPFENNING, V. Non-commutative Gröbner bases in algebras of solvable type. *J. Symb. Comput. 9* (1990), 1–26.
31. KUICH, W., AND SALOMAA, A. *Semirings, Automata, Languages*. Springer-Verlag, New York, 1986.
32. LANKFORD, D., BUTLER, G., AND BRADY, B. Abelian group unification algorithms for elementary terms. *Cont. Math. 29* (1984), 193–199.
33. LEEB, K., AND PIRILLO, G. Shuffle-compatible total orders. *Ann. Mat. Pura Appl. 153* (1988), 1–26.
34. LIVESEY, M., AND SIEKMANN, J. Unification in sets and multisets. SEKI Tech. Rep. Universität Karlsruhe.
35. MARTIN, U. A geometrical approach to multiset orderings. Tech. Rep. Univ. London, London, England, 1988.
36. MORA, F. Gröbner bases for non-commutative polynomial rings. In *Proceedings of the Algebraic Algorithms and Error-Correcting Codes*. Lecture Notes in Computer Science, vol. 229. Springer-Verlag, New York, 1986, pp. 353–362.
37. NUTT, W. Talk at the second workshop on unification (Val d' Ajol, France), 1988.
38. NUTT, W. Unification in monoidal theories. In *Proceedings of the CADE '90* (Kaiserslautern, Germany). Lecture Notes in Computer Science, vol. 449. Springer-Verlag, New York, 1990, pp. 618–632.
39. PLOTKIN, G. Building in equational theories. *Mach. Int. 7* (1972), 73–90.
40. ROBBIANO, L. Term orderings on the polynomial ring. In *Proceedings of the EUROCAL '85*. Lecture Notes in Computer Science, vol. 204. Springer-Verlag, New York, 1985, pp. 513–517.
41. RYDEHEARD, D. E., AND BURSTALL, R. M. A categorical unification algorithm. In *Proceedings of the Workshop on Category Theory and Computer Programming*. Lecture Notes in Computer Science, vol. 240, Springer-Verlag, New York, 1985, pp. 493–505.
42. SCHMIDT-SCHAUß, M. Unification under associativity and idempotence is of type nullary. *J. Autom. Reas. 2* (1986), 277–282.
43. SIEKMANN, J. Unification theory. *J. Symb. Computat. 7* (Special issue on unification) (1988), 315–337.
44. STICKEL, M. A unification algorithm for associative-commutative functions. *J. ACM 28* (1981), 423–434.
45. STICKEL, M. Automated deduction by theory resolution. *J. Automat. Reas. 1* (1985), 333–355.
46. TREVISAN, G. Classificazione dei semplici ordinamenti di un gruppo libero commutativo con m generatori. *Rendiconti del Seminario Matematico della Universita di Padova 22* (1953), 143.
47. ZAICEVA, M. I. On the set of ordered Abelian groups. *Uspehi Matem. Nauk (N.S.) 8* (1953), 135–137.

RECEIVED JUNE 1989; REVISED NOVEMBER 1990; ACCEPTED JUNE 1991