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

Franz Baader

SEKI Report SR-90-01

Unification in Commutative Theories, Hilbert's Basis Theorem and Gröbner Bases 1)

Franz Baader

DFKI

Postfach 2080, Erwin-Schrödingerstraße, D-6750 Kaiserslautern, F.R.G.

Unification in a commutative theory E may be reduced to solving linear equations in the corresponding semiring S(E) (Nutt (1988)). The unification type of E can thus be characterized by algebraic properties of S(E). The theory of abelian groups with n commuting homomorphisms corresponds to the semiring $\mathbb{Z}[X_1,...,X_n]$. 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 $\mathbb{Z}[X_1,...,X_n]$ 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 (Baader (1988)). This can also be proved by using the fact that the corresponding semiring, namely $\mathbb{N}[X]$, is not noetherian. An other example of a semiring (even ring), which is not noetherian, is the ring $\mathbb{Z}<X_1,...,X_n>$, where X_1, ..., X_n ($n > 1$) are non-commuting indeterminates. This semiring corresponds to the theory of abelian groups with n non-commuting homomorphisms. Surprisingly, by construction of a Gröbner Base algorithm for right ideals in $\mathbb{Z}<X_1,...,X_n>$, it can be shown that this theory is unitary unifying.

1. Introduction

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

In general we do not need the set of all E-unifiers. A *complete set of E-unifiers*, i.e. 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

$\sigma =_E \theta$ iff $x\sigma =_E x\theta$ for all variables x occurring in s_i or t_i for some i, i = 1, ..., n,

$\sigma \leq_E \theta$ iff there exists a substitution λ such that $\sigma =_E \theta \circ \lambda$.

In this case 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 $\theta \leq_E \sigma$.

1) This research was done while the author was still at the IMMD 1, University Erlangen.

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 w.r.t. 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. If the terms may contain free constants, we talk about unification with constants, else about unification without constants (see Baader (1988), Section 7). If nothing else is specified, "unification" means "unification without constants". For more information about unification theory and the unification hierarchy consult Siekmann (1988).

Unification in the empty theory (which is unitary) plays an important rôle in automated theorem proving, term rewriting and logic programming. Generalizations to E-unification usually require that E is finitary (see e.g. Stickel (1985), Jouannaud-Kirchner (1986) and Jaffar-Lassez-Maher (1984)). A finitary theory most used in this context is the theory of abelian semigroups (monoids), i.e. the theory of an associative, commutative binary operation (with a neutral element). Unification algorithms for this theory (see e.g. Livesey-Siekmann (1978), Stickel (1981), Fages (1984), Fortenbacher (1985), Büttner (1986), Herold (1987)) make use of the fact that unifiers correspond to solutions of systems of linear equations in the semiring $\mathbb{N}$ (see Eilenberg (1974) or Kuich-Salomaa (1986) for the definition and properties of semirings). The same phenomenon occurs for the theory of abelian groups where the semiring is $\mathbb{Z}$ (Lankford-Butler-Brady (1984)) and for the theory of idempotent abelian monoids where the 2-element boolean semiring $\mathcal{B}$ is used (Livesey-Siekmann (1978), Baader-Büttner (1988)).

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 (1988). 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 (resp. finitary w.r.t. unification with constants). The above mentioned results for abelian monoids etc. 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 (1989) these conditions were modified to algebraic characterizations of unification type unitary for constant-free unification 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 w.r.t. unification with constants), if the finitely generated free objects are finite (Baader (1988)).

Werner Nutt (Nutt (1988)) 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 (1985) and Section 6 of this paper), which thus provide the desired algorithm.

The theory of abelian monoids with a homomorphism is of type zero (Baader (1988)). This can also be demonstrated using the fact that the corresponding semiring, namely

$\mathbb{N}[X]$, is not noetherian (Section 4).
Another example of a semiring which is not noetherian is the ring $\mathbb{Z}\langle X,Y\rangle$, where X, Y are non-commuting indeterminates. This semiring corresponds to the theory of abelian groups with two (non-commuting) 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 non-commuting 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 (1988) and show how the corresponding semirings may be obtained in this framework.

An equational theory E defines a *variety V(E)*, i.e. the class of all algebras (of the given signature Ω) which 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 $F_E(X)$.

Let F(E) be the class of all free algebras $F_E(X)$ with finite sets X and let C(E) be the category which has the elements of F(E) as objects and the homomorphisms between these elements as morphisms. Note that the coproduct of $F_E(X)$ and $F_E(Y)$ in C(E) is given by $F_E(X \dot\cup Y)$ (where $\dot\cup$ means disjoint union). Thus $F_E(X)$ is the coproduct of the isomorphic objects $F_E(x)$ for $x \in X$.

Let $\Gamma = \langle s_i = t_i;\ 1 \le i \le n \rangle_E$ 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 s_i and t_i as elements of $F_E(X)$. Since we do not distinguish between $=_E$-equivalent unifiers, any E-unifier of Γ can be regarded as a homomorphism of $F_E(X)$ into $F_E(Y)$ for some finite set Y (of variables). Let $I = \{ x_1, ..., x_n \}$ be a set of cardinality n. We define homomorphisms

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

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

This categorical reformulation of E-unification (due to Rydeheard-Burstall (1985)) allows to characterize the class of *commutative theorie*s by properties of the category C(E) of finitely generated E-free objects: C(E) has to be a *semiadditive category* (see Herrlich-Strecker (1973) and Baader (1988)). In order to give a more algebraic definition of commutative theories we need some more notation.

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 $f(e,...,e) =_E e$, i.e. in any algebra $A \in V(E)$, $f(e,...,e) = e$ holds. Note that for nullary f this means $f =_E e$.

Let **K** be a class of algebras (of signature Ω). An n-ary *implicit operation* in **K** is a family $f = \{ f_A;\ A \in \mathbf{K} \}$ of mappings $f_A\colon A^n \to A$ which is compatible with all homomor-

phisms, i.e. for any homomorphism $h: A \to B$ with $A, B \in \mathbf{K}$ and all $a_1, ..., a_n \in A$, $f_A(a_1,...,a_n)h = f_B(a_1h,...,a_nh)$ holds. In the following we omit the index and just write f for any f_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(E) such that

(a) The constant e is a neutral element for $*$ in any algebra $A \in F(E)$.

(b) For any n-ary function symbol $f \in \Omega$, any algebra $A \in F(E)$ and any $s_1, ..., s_n, t_1, ..., t_n \in A$ we have $f(s_1 * t_1,...,s_n * t_n) = f(s_1,...,s_n) * f(t_1,...,t_n)$.

In Baader (1988) the following properties for commutative theories E are shown within a categorical framework, using well-known results for semiadditive categories:

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

(2.3) The implicit operation $*$ of Definition 2.1 is associative and commutative. It induces a binary operation + on any morphism set $\mathrm{hom}(F_E(X),F_E(Y))$ 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. 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 $\mathrm{hom}(F_E(X),F_E(Y))$ and it is a neutral element for + on $\mathrm{hom}(F_E(X),F_E(Y))$

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

(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 $M_{\sigma\delta} = M_\sigma \cdot M_\delta$.

Werner Nutt (Nutt (1988)) observed that commutative theories are (modulo a translation of the signature) what he calls monoidal theories 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 **1** be an arbitrary set of cardinality 1. Property (2.3) yields that $\mathrm{hom}(F_E(\mathbf{1}),F_E(\mathbf{1}))$ 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 n-th power and copower of $F_E(\mathbf{1})$. Thus, 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).

We now give some examples of commutative theories, whose unification properties will be considered in subsequent sections of this paper. In all these examples, the implicit operation is given by a function symbol, which is associative and commutative in the corresponding theory. Additional examples of commutative theories can be found in Baader (1988).

EXAMPLES 2.6. We consider the following signatures:
$\Sigma := \{ \cdot, 1, h \}$, where $\cdot$ is binary, 1 is nullary and h is unary.
For $n \geq 0$, $\Omega_n := \{ \cdot, 1, {}^{-1}, h_1, ..., h_n \}$, 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 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 (non-commuting) 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 $\mathrm{AGnHC} := \mathrm{AGnH} \cup \{ h_i(h_j(x)) = h_j(h_i(x)); 1 \leq i < j \leq n \}$.

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 $\cdot$) satisfies 2b 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 $A \in F(E)$ 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 (1989). As a consequence we derive, that unification in a commutative theory E means solving systems of linear equations in the semiring S(E). This yields algebraic characterizations of the unification types.

THEOREM 3.1. A commutative theory E is unitary iff it satisfies the following condition:
Let y be an arbitrary variable. For any E-unification problem $< \sigma = \tau >_E$ (where σ, τ: $F_E(I) \rightarrow F_E(X)$) there are finitely many E-unifiers $\alpha_1, ..., \alpha_r$: $F_E(X) \rightarrow F_E(y)$ such that any E-unifier δ: $F_E(X) \rightarrow F_E(y)$ is representable as

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

where λ_i: $F_E(y) \rightarrow F_E(y)$ are morphisms.

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

COROLLARY 3.2. A commutative theory E is unitary iff the corresponding semiring S(E) satisfies the following condition: For any n, m ≥ 1 and any pair M_σ, M_τ of m×n-matrices over S(E) the set

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

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

THEOREM 3.3. Let E be a unitary commutative theory. Then E is finitary w.r.t. unification with constants iff the following condition holds:

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

(1) The elements of M are morphisms $\mu: F_E(Y) \to F_E(X)$ satisfying $\delta\mu = 1$.

(2) The elements of $N = \{ \nu_1, \ldots, \nu_r \}$ are morphisms $\nu_i: F_E(Y) \to F_E(Z_i)$ 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 w.r.t. unification with constants.

COROLLARY 3.4. Let E be a unitary commutative theory. Then E is finitary w.r.t. unification with constants, if the following condition holds in S(E):

Let A be any m×n-matrices 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 $N := \{ \underline{x} \in S(E)^n;\ A\underline{x} = 0 \}$, i.e. there exist finitely many $\underline{m}_1, \ldots, \underline{m}_k \in S(E)^n$ such that $M = \{ \underline{m}_i + \underline{n};\ \underline{n} \in N \text{ and } 1 \leq i \leq k \}$.

Note that the semimodule N is finitely generated, since E is unitary and N = U(A,0), where 0 is the m×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 $\underline{b}$.

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 w.r.t. unification with constants.

4. A Commutative Theory of Unification Type Zero

In 1972 Plotkin conjectured, that there exists an equational theory E which has unification type zero. But only in 1983, Fages and Huet constructed the first example of an equational theory of this type. Schmidt-Schauß (1986) and the present author (1986) showed that the theory of idempotent semigroups is of unification type zero and in Baader (1987) I have 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, which are more natural than the original example of Fages and Huet.
In Baader (1988) it is shown that the theory AIMH of idempotent abelian monoids with a homomorphism has 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 has type zero. Since commutative theories are either unitary or of unification type zero (Baader (1988), Theorem 6.1), it is sufficient to show, that the semiring S(AMH) does not satisfy the condition of Corollary 3.2.
Let σ: $F_{AMH}(x) \to F_{AMH}(x)$ be a morphism of C(AMH). Then there are $k \geq 0$ and a_0, ..., $a_k \in \mathbb{N}$ such that $x\sigma = x^{a_0}h(x^{a_1})...h^k(x^{a_k})$. We associate with the morphism σ the polynomial $p_\sigma = a_0 + a_1X + ... + a_kX^k \in \mathbb{N}[X]$. 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 $S(AMH) \cong \mathbb{N}[X]$.
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 $\underline{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)}) = (1, X + X^2 + ... + X^{n+1}, X^n)$ 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.

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 an ideal 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.
An ideal $I \subseteq \mathbb{N}[X]$ 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 $g = 1$. But $1 \notin I$.

PROPOSITION 4.3. The theory AMH has unification type zero.
PROOF. Assume that AMH has not 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 $I_{1+3} = \{ p_1 + p_3$; 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 semi-

module is not specific for the semiring $\mathbb{N}[X]$. More general, let S be a semiring which 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-Nutt (1989)).

5. AGnHC-Unification and Hilbert's Basis Theorem

It is easy to see that S(AGnHC) is isomorphic to the ring $\mathbb{Z}[X_1,...,X_n]$, i.e. the polynomial ring over $\mathbb{Z}$ in the (commuting) indeterminates $X_1, ..., X_n$. To establish the condition of Corollary 3.2, we have to consider systems of homogeneous linear equations in $\mathbb{Z}[X_1, ...,X_n]$, i.e. systems $f_{1i}x_1 + ... + f_{ki}x_k = 0$ ($i = 1, ..., s$), where the coefficients f_{ij} and the desired solutions are elements of $\mathbb{Z}[X_1,...,X_n]$. The set of solutions $\underline{I} \subseteq (\mathbb{Z}[X_1,...,X_n])^k$ is a $\mathbb{Z}[X_1,...,X_n]$-module, which is finitely generated by Hilbert's Basis Theorem and the fact that $\mathbb{Z}$ is a noetherian ring (see e.g. Jacobson (1980)). Thus AGnHC is unitary w.r.t. unification without constants. Since $\mathbb{Z}[X_1,...,X_n]$ is a ring, Corollary 3.5 applies and we have proved

PROPOSITION 5.1. (Nutt (1988))
For any $n \geq 0$ the theory AGnHC is unitary and it is also unitary w.r.t. 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 $\mathbb{Z}[X_1,...,X_n]$ effectively. The next section describes one possible answer to this problem.

6. Solving Linear Equations in $\mathbb{Z}[X_1,...,X_n]$ using Gröbner Bases

Buchberger (1985) describes an effective method, which constructs finitely many generators of the solutions of a single equation $f_1x_1 + ... + f_kx_k = 0$, where the f_i and the desired solutions are elements of $K[X_1,...,X_n]$ for a field K. This method may also be used for $\mathbb{Z}[X_1,...,X_n]$ (see Buchberger (1985) for Gröbner Bases of polynomials over $\mathbb{Z}$ and Kandry-Rody-Kapur (1988) for Gröbner Bases of polynomials over a euclidean ring), but the proof of its 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-Sasaki-Kobayashi (1986), where Gröbner base theory is extended to modules over $K[X_1,...,X_n]$.

First we recall some facts and notations concerning Gröbner bases:

(6.1) *Admissible term orderings.*
Let $T_n := \{ X_1^{k_1}...X_n^{k_n};\ k_1, ..., k_n \in \mathbb{N} \}$ be the set of all terms (i.e. monomials with coefficient 1) in $\mathbb{Z}[X_1,...,X_n]$. With respect to multiplication of polynomials, T_n is a commuta-

tive monoid (with neutral element $1 = X_1^0 \ldots X_n^0$), which is isomorpic 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 (1953), Zaiceva (1953) and, more recently, by Robbiano (1985) and Martin (1988):

Any compatible linear ordering $<$ on T_n is completely determined by a n×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: $X_1^{k_1} \ldots X_n^{k_n} < X_1^{h_1} \ldots X_n^{h_n}$ iff the first non-zero 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 non-zero entry is greater than zero.

(6.2) *Rewriting with polynomials.*

For a polynomial f and a term t which 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 w.r.t. $<$ and coeff(t,f) = $a \in \mathbb{Z}$ 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 $\mathbb{Z}[X_1,\ldots,X_n]$:

$f \rightarrow_F g$ iff (1) f contains a term t with coefficient a.

(2) F contains a polynomial h such that $HT(h) = t \cdot s$ (for some $s \in T_n$) and $|HC(h)| \leq |a|$.

(3) $g = f - h \cdot b \cdot s$, where $a = b \cdot HC(h) + c$ for $0 \leq c < |HC(h)|$, $b, c \in \mathbb{Z}$.

Let $\xrightarrow{*}_F$ (resp. $\xrightarrow{+}_F$) denote the reflexive, transitive (resp. transitive) closure of $\rightarrow_F$. It can be shown (using a multiset extension of $<$) that $\xrightarrow{+}_F$ 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 $f - g \in \langle F \rangle$. This congruence is the reflexive, transitive and symmetric closure of $\rightarrow_F$ (Bachmair-Buchberger (1980)).

(6.3) *Gröbner bases and S-polynomials.*

Let I be an ideal in $\mathbb{Z}[X_1,\ldots,X_n]$ and B let be a finite set of polynomials. B is a *Gröbner base* for I iff $\langle B \rangle = I$ and $\rightarrow_B$ is confluent. Since $\xrightarrow{+}_B$ is noetherian, confluence is equivalent to local confluence and this property can be tested with the help of finitely many critical pairs, which are here called S-polynomials.

Let $g_1 = c_1 \cdot t_1 + R(g_1)$ and $g_2 = c_2 \cdot t_2 + R(g_2)$ 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* $S(g_1,g_2)$ 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

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

Now B is a Gröbner base iff for every pair of polynomials in B the S-polynomial reduces to 0 w.r.t. $\to_B$.

If B is a Gröbner base for the ideal I, then $f \in I$ iff $f \overset{*}{\to}_B 0$ and $f \equiv_I g$ iff f and g reduce to the same $\to_B$-irreducible element. Thus we can decide ideal membership for I, if we have a Gröbner base for I. But a Gröbner base can always be constructed, if a finite set of generators of I (which always exists by Hilbert's Basis Theorem) is given.

(6.4) *Buchberger's algorithm.*

Let I be an ideal in $\mathbb{Z}[X_1,...,X_n]$ and F be a finite set of polynomials such that $\langle F \rangle = I$. As described in (6.3), we can effectively test whether F is a Gröbner base for I. If F is not a Gröbner base, we can extend F by the $\to_F$-irreducibles of those S-polynomials, which do not reduce to 0, and test again. This completion procedure always terminates with a finite Gröbner base for I (see e.g. Kandry-Rody-Kapur (1988) for more details). This termination property is a consequence of Dicksons Lemma (Dickson (1913)), which holds for free commutative monoids, but not for free monoids (see e.g. Mora (1985)).

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

Let (*) $f_1x_1 + ... + f_rx_r = f_0$ be an (inhomogeneous) linear equation in $\mathbb{Z}[X_1,...,X_n]$. According to Section 3 we have to find one solution for (*) and finitely many generators of the solutions of the homogeneous equation (**) $f_1x_1 + ... + f_rx_r = 0$.

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

Obviously, (*) has a solution iff $f_0 \in I$. Hence, if (*) has a solution, then f_0 reduces to 0 w.r.t $\to_B$. This yields $p_1, ..., p_s \in \mathbb{Z}[X_1,...,X_n]$ 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 $|z^{(1)}, ..., |z^{(L)}$ of the solutions of the equation (++) $g_1x_1 + ... + g_sx_s = 0$. Then $P \cdot |z^{(1)}, ..., P \cdot |z^{(L)}$ are solutions of (**), but in general they do not generate all solutions. Let E_r be the $r \times r$ identity matrix and let $|t^{(1)}, ..., |t^{(r)}$ 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 6.5. The finitely many vectors $P \cdot |z^{(1)}$, ..., $P \cdot |z^{(L)}$, $|t^{(1)}$, ..., $|t^{(r)}$ are solutions of (**) and they generate all solutions of this equation.

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

We now show how to solve the equation (++) $g_1x_1 + ... + g_sx_s = 0$, if $B = \{ g_1, ..., g_s \}$ is a Gröbner base.

For a set $\{ q_1, ..., q_s \}$ of polynomials the *complexity measure* $BS(q_1,...,q_s)$ is defined as follows: Let $t := \max\{ HT(q_1), ..., HT(q_s)\}$ and for all i, $1 \le i \le s$, let $a_i := \mathrm{coeff}(t,q_i)$ (Note that $a_i = 0$ for $HT(q_i) < t$). Then $BS(q_1,...,q_s) := (|a_1| + ... + |a_s|) \cdot t$.

Now t is called the term and $|a_1| + ... + |a_s|$ the coefficient of $BS(q_1,...,q_s)$. We define

$$a \cdot t = BS(q_1,...,q_s) < BS(q_1',...,q_s') = a' \cdot t' \text{ iff } t < t' \text{ or } t = t' \text{ and } a < a'.$$

Let $S(g_i,g_j) = s_i \cdot g_i - a \cdot s_j \cdot g_j = b \cdot s_i \cdot t_i + s_1 \cdot R(g_i) - a \cdot s_j \cdot R(g_j)$ be the S-polynomial of g_i and g_j (see 6.3). Since B is a Gröbner base, we have $S(g_i,g_j) \xrightarrow{*}_B 0$. This derivation yields polynomials $w_1, ..., w_s$ such that

$$S(g_i,g_j) = \sum_{k=1}^{k=s} w_k \cdot g_k$$

and $BS(w_1 \cdot g_1,...,w_s \cdot g_s) = c \cdot t'$ for some $t' < s_i \cdot t_i$, if $b = 0$, or $BS(w_1 \cdot g_1,...,w_s \cdot g_s) = b \cdot s_i \cdot t_i$, if $b \neq 0$.

Now $s_i \cdot g_i - a \cdot s_j \cdot g_j = S(g_i,g_j) = w_1 \cdot g_1 + ... + w_s \cdot g_s$ implies $w_1 \cdot g_1 + ... + (w_i - s_i) \cdot g_i + ... + (w_j + a \cdot s_j) \cdot g_i + ... + w_s \cdot g_s = 0$. Thus $|w_{ij} := (w_1,...,w_i - s_i,...,w_j + a \cdot s_j,...,w_s)^T$ is a solution of the equation (++).

LEMMA 6.6. The finitely many vectors $|w_{ij}$ generate all solutions of (++).

PROOF. Let $|p = (p_1,...,p_s)^T$ be a solution of (++) and let $t = \max\{ HT(g_1p_1), ..., HT(g_sp_s) \}$. We prove the lemma by induction on $BS(g_1p_1,...,g_sp_s)$. Since $g \cdot |p = 0$, there exist i, j such that $HT(g_ip_i) = t = HT(g_jp_j)$ and $HC(g_ip_i)$ and $HC(g_jp_j)$ have different sign. Without loss of generality we assume that $c_i := HC(g_i) \ge HC(g_j) =: c_j > 0$. Obviously, $t_i := HT(g_i)$ and $t_j := HT(g_j)$ are divisors of t and thus $\mathrm{lcm}(t_i,t_j) = s_it_i = s_jt_j$ divides t, i.e. there exists r with $rs_it_i = rs_jt_j = t$.

We now consider the case $HC(g_ip_i) > 0$ and $HC(g_jp_j) < 0$ (the other case is similar). The vector $|q = (q_1,...,q_s)^T := |p + r \cdot |w_{ij}$ is a solution of (++) and we have $g_1q_1 = g_1p_1 + g_1rw_1$, ..., $g_iq_i = g_ip_i + g_irw_i - g_irs_i$, ..., $g_jq_j = g_jp_j + g_jrw_j + g_jars_j$, ..., $g_sq_s = g_sp_s + g_srw_s$.

If $\max\{ HT(g_1q_1), ..., HT(g_sq_s) \} < t$, the lemma is proved by induction, since the term of BS has decreased. Otherwise $\max\{ HT(g_1q_1), ..., HT(g_sq_s) \} = t$ and we have to calcu-

late the coefficient of $BS(g_1q_1,...,g_sq_s)$. The triangle inequality yields

$$BS(g_1q_1,...,g_sq_s) \le BS(g_1p_1,...,g_ip_i - g_irs_i,...,g_jp_j + g_jars_j,...,g_sp_s) + b\cdot t,$$

since $BS(g_1rw_1,...,g_srw_s) = r\cdot b\cdot s_i\cdot t_i = b\cdot t$ (for $b > 0$) or $BS(g_1rw_1,...,g_srw_s)$ has a term which is smaller than t (for $b = 0$).

We have $|coeff(t,g_ip_i - g_irs_i)| = |coeff(t,g_ip_i)| - c_i$ (since $coeff(t,g_ip_i) = HC(g_ip_i) \ge c_i \ge 0$) and $|coeff(t,g_jp_j + g_jars_j)| < |coeff(t,g_jp_j)| + ac_j$ (since $coeff(t,g_jp_j) = HC(g_jp_j) < 0$ and $coeff(t,g_jars_j) = ac_j > 0$).

Thus $BS(g_1p_1,...,g_ip_i - g_irs_i,...,g_jp_j + g_jars_j,...,g_sp_s) < BS(g_1p_1,...,g_sp_s) + (ac_j - c_i)\cdot t$ and, since $c_i = a\cdot c_j + b$, $BS(g_1q_1,...,g_sq_s) < BS(g_1p_1,...,g_sp_s)$. This completes the proof of Lemma 6.6 by induction on BS.

Now we have completely described a method to solve linear equations in $\mathbb{Z}[X_1,...,X_n]$.

EXAMPLE 6.7. 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 Gröbner base for the Ideal I, generated by f_1, f_2 and f_3. Let $<$ be the admissible ordering defined by the matrix

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

With respect to this ordering, the Buchberger algorithm yields the Gröbner base $B = \{ g_1, g_2, g_3, g_4, g_5 \}$, 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 w.r.t. $\rightarrow_B$, we get the matrix Q such that $\underline{g}\cdot Q = \underline{f}$. In our example

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

We now determine whether $f_0 \in I = \langle B \rangle$, i.e. whether f_0 reduces to 0 w.r.t. $\rightarrow_B$:

$f_0 \rightarrow_B f_0 - g_5\cdot XY = XYZ^3 - X^3Y^3Z^2 \rightarrow_B f_0 - g_5\cdot XY + g_3\cdot XY^2Z = XYZ^3 - XY^2Z^3 \rightarrow_B f_0 - g_5\cdot XY + g_3\cdot XY^2Z + g_4\cdot XYZ = XYZ^3 - XYZ^3 = 0.$

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 $f_1x_1 + f_2x_2 + f_3x_3 = f_0$:

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

The next step is to determine the solutions $|w_{ij}$ of the equation $g_1x_1 + ... + g_5x_5 = 0$.

$S(g_1,g_2) = g_1 \cdot X - g_2 \cdot Z = -X^2YZ + Z^2 = -g_3$ 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 $\mathrm{lw}_{1,2} = (-X, Z, -1, 0, 0)^T$.

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

Similar computations yield the other vectors lw_{ij}:

$\mathrm{lw}_{1,4} = (-Z, 0, 0, XY, 0)^T$, $\quad \mathrm{lw}_{1,5} = (-XY, 0, -Z, YZ+Z, Y^2)^T$,

$\mathrm{lw}_{2,3} = (0, -Z, Y, 1, 0)^T$, $\quad \mathrm{lw}_{2,4} = (0, -Z^2, Z, X^2Y, 0)^T$,

$\mathrm{lw}_{2,5} = (0, -Z^2, 0, YZ+Z, Y^2)^T$, $\quad \mathrm{lw}_{3,4} = (0, 0, -Z, X^2, 1)^T$,

$\mathrm{lw}_{3,5} = (0, 0, -Z, Z, Y)^T$, $\quad \mathrm{lw}_{4,5} = (0, 0, 0, -X^2+Z, Y-1)^T$.

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

$P \cdot \mathrm{lw}_{1,2} = (0, 0, 0)^T$, $\quad P \cdot \mathrm{lw}_{1,3} = (0, 0, 0)^T$,

$P \cdot \mathrm{lw}_{1,4} = (0, X^2Y^2 - Z, -XY^2Z + XYZ)^T$, $\quad P \cdot \mathrm{lw}_{1,5} = (-XY) \cdot P \cdot \mathrm{lw}_{1,4}$,

$P \cdot \mathrm{lw}_{2,3} = (0, 0, 0)^T$, $\quad P \cdot \mathrm{lw}_{2,4} = X \cdot P \cdot \mathrm{lw}_{1,4}$,

$P \cdot \mathrm{lw}_{2,5} = P \cdot \mathrm{lw}_{1,5} = (-XY) \cdot P \cdot \mathrm{lw}_{1,4}$, $\quad P \cdot \mathrm{lw}_{3,4} = (0, 0, 0)^T$,

$P \cdot \mathrm{lw}_{3,5} = -P \cdot \mathrm{lw}_{2,4} = (-X) \cdot P \cdot \mathrm{lw}_{1,4}$, $\quad P \cdot \mathrm{lw}_{4,5} = P \cdot \mathrm{lw}_{3,5} = (-X) \cdot P \cdot \mathrm{lw}_{1,4}$.

The solution $P \cdot \mathrm{lw}_{1,4} = (0, X^2Y^2 - Z, -XY^2Z + XYZ)^T$ 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

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

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

EXAMPLE 6.8. 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 $<\{ f_1, f_2, f_3 \}> = \mathbb{Z}[X]$ and that $B = \{ g_1 \}$ for $g_1 = 1$ is the corresponding Gröbner base. The transformation matrices are $P = (1, -1, 0)^T$ and $Q = (X, X-1, -X^2)$.

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

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

i.e. $(X-1, -X, 0)^T$ and $(-X^2, X^2, -1)^T$, 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 $\mathbb{Z}<X_1,...,X_n>$, i.e. the polynomial ring over $\mathbb{Z}$ in the non-commuting indeterminates $X_1, ..., X_n$. Unfortunately, for $n \geq 2$ this ring is not noetherian (see Mora (1985)) and the membership problem for finitely generated two-sided ideals is undecidable (Kandry-Rody-Weispfenning (1988)). Fortunately, we are not interested in two-sided ideals, but only in right ideal. The solutions of a homogeneous equation $f_1x_1 + ... + f_rx_r = 0$ are only closed under right multiplication and the inhomogeneous equation $f_1x_1 + ... + f_rx_r = f_0$ has a solution iff f_0 is a member of the right ideal generated by $f_1, ..., f_r$. Though, for $n \geq 2$, $\mathbb{Z}<X_1,...,X_n>$ is not even right noetherian (i.e. there are right ideals in $\mathbb{Z}<X_1,...,X_n>$, which are not finitely generated), the set of solutions of a homogeneous equation $f_1x_1 + ... + f_rx_r = 0$ is a finitely generated right $\mathbb{Z}<X_1,...,X_n>$-semimodule and the membership problem for finitely generated right ideals is decidable in $\mathbb{Z}<X_1,...,X_n>$ (see Section 8 and 9). This yields

PROPOSITION 7.1. For any $n \geq 0$ the theory AGnH is unitary and it is also unitary w.r.t. unification with constants.

8. "Gröbner bases" for finitely generated right ideals in $\mathbb{Z}<X_1,...,X_n>$

The construction of Gröbner bases for finitely generated right ideals in $K<X_1,...,X_n>$, where K is a field, is very easy (Mora (1985)). For $\mathbb{Z}<X_1,...,X_n>$ one has to be more careful.

The rôle of terms in the commutative case is now played by words over the alphabet $\Sigma_n := \{ X_1, ..., X_n \}$. 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.

A total ordering $<$ on W_n is called *admissible* iff the following two conditions hold:

(1) For all $s, t, r \in W_n$ $s < t$ implies $sr < tr$ (compatibility with right concatenation).

(2) For all $s \in W_n$ the set $\{ t \in W_n; t < s \}$ is finite.

LEMMA 8.1. Let $<$ be an admissible 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 admissible orderings are graded lexicographical orderings and, more general,

all suffle-compatible total orders (see Leeb-Pirillo (1988)). The complete characterization of all concatenation-compatible (resp. right concatenation-compatible) linear orderings is still an open problem.
We now extend admissible orderings to monomials and polynomials.

DEFINITION 8.2. Let < be an admissible ordering on W_n.

(1) Let a, b $\in \mathbb{Z}$ and s, t $\in W_n$. Then as < bt iff s < t or s = t and $|a| < |b|$ or s = t and $|a| = |b|$ and a < b. This defines a well-ordering on the monomials of $\mathbb{Z}\langle X_1,...,X_n\rangle$.

(2) Let $f = \Sigma a_i s_i$ and $g = \Sigma b_i t_i$ be two polynomials, i.e. elements of $\mathbb{Z}\langle X_1,...,X_n\rangle$. Then we define f < g iff $\{ ... a_i s_i, ... \} << \{ ... b_i t_i, ... \}$, where << denotes the multiset ordering (see Dershowitz-Manna (1979)) induced by the ordering < on monomials.
(3) Let f be a polynomial. We write f = at + R(f) if t is the maximal (w.r.t. <) word in f (t = HW(f)) and a is the coefficient of t in f (a = HC(f)).
(4) For a set F of polynomials in $\mathbb{Z}\langle X_1,...,X_n\rangle$, the reduction relation $\rightarrow_F$ is defined as in Section 6, 6.3.

For $K\langle X_1,...,X_n\rangle$, Mora (1985) has described a very easy algorithm, which transforms a finite set F of polynomials into a "Gröbner base" (see Mora (1985) 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 HW(f) is a prefix of 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, which generates the same right ideal as F and has the following property:
For two different elements f and g of G, HW(f) and HW(g) are not comparable w.r.t. the prefix-ordering (i.e. for any word r, $HW(f) \cdot r \neq HW(g)$ and $HW(g) \cdot r \neq HW(f)$).

For $\mathbb{Z}\langle X_1,...,X_n\rangle$, we encounter the following problem: For $f = a \cdot t + R(f)$ and $g = b \cdot t \cdot r + R(g)$ with t, r $\in W_n$, a, b $\in \mathbb{Z}$ and $|a| > |b|$, HW(f) is prefix of HW(g), but the head monomial of g can not be reduced by f. If, in addition, b devides 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 describtion of an algorithm, which transforms a finite set of polynomials $\{ f_1, ..., f_m \}$ into a "Gröbner base", which defines the same right ideal.
In the beginning, $F_0 := \{ f_1, ..., f_m \}$ 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. While there are $f := f_i$ and $g := f_j$ in F_k such that

(1) (i, j) is not marked and
(2) $f = a \cdot t + R(f)$ and $g = b \cdot tr + R(g)$ for some a, b $\in \mathbb{Z}$ and t, r $\in W_n$,

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 $f_1 := f - g \cdot c = d \cdot t + R(f) - R(g) \cdot c$ and $F_{k+1} := (F_k \setminus \{ f \}) \cup \{ f_1 \}$. We do not have to mark (i, j), since $f = f_i$ 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 $g_1 := g - f \cdot cr = d \cdot tr + R(g) - R(f) \cdot cr$ 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 devide $|a|$.

Let $a = bc + d$ for some c, d such that $0 < d < |b| < |a|$. We define $g_1 := f \cdot r - g \cdot c = d \cdot tr + R(f) \cdot r - R(g) \cdot c$. Since the words occurring in $R(f) \cdot r$ and $R(g) \cdot c$ are smaller than tr, we have $HW(g_1) = tr$, $HC(g_1) = d$ and $R(g_1) = R(f) \cdot r - R(g) \cdot c$. 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 $|b|$ devides $|a|$, i.e. there exists c such that $a = bc$.

Define $g_1 := f \cdot r - g \cdot c = R(f) \cdot r - R(g) \cdot c$. Now $g_1 < g$, but since $|c| \neq 1$, g can not be represented using g_1. We distinguish the following cases:

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

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

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 $|HC(g_1)| \geq |HC(h)|$.

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. $h \notin F_k$ and there is no polynomial in F_k which has HW(h) as head word.

Thus the monomial HC(h)HW(h) has been reduced in some previous step. It is easy to see, that then HC(h)HW(h) can also be reduced by $\rightarrow_{F_k}$. If we have $|HC(g_1)| \geq |HC(h)|$, g_1 can be reduced and we proceed as in Case 4.1.2.

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

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

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 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 $\mathbb{Z}$<a,b,c,d>. We take the graded lexicograpical ordering with $a > b > c > d$ as admissible ordering (i.e. $u < v$ iff $|u| < |v|$ or $|u| = |v|$ and $u <_{lex} v$) 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 $HW(f_{11}) = HW(f_4)$ 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 occured 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 $F_0 = \{ f_1, ..., f_m \}$ 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 $<<$ induced by the ordering $<$ on polynomials (see Definition 8.2). Since $<$ is well-founded, the multiset extension $<<$ is also well-founded.
For the Cases 1, 2, 3 and 4.1.1, $F_k >> F_{k+1}$. Case 4.1.2 and the according subcase of 4.1.3 can not occur infinitely often in successive steps, because then $g_1 > g_2 > g_3 > ...$ 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 $t \leq \max\{ HW(f_1), ..., HW(f_m) \}$. Since $<$ is admissible, 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 term, which has disappeared in some former step, appears again – can only occur finitely often for a certain term, because the absolut 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

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

should be interpreted as follows: the a_i are monomials in $\mathbb{Z}<X_1,...,X_n>$, 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 $h = \Sigma_{h_i \in F_k} h_i \cdot a_i$ 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 $g \in F_k$ and $HW(g) = HW(f) \geq HW(f_1)$. 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 similar.

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

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

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) is satisfied, since only in Case 4 both f and g remain in F_{k+1}.

(2) In Case 4 we have $g_1 := f\cdot r - g\cdot c = R(f)\cdot r - R(g)\cdot c$ and thus $HW(g_1) < HW(g) = HW(f\cdot r) = tr$. There is some g_i such that $g_1 \overset{*}{\to}_{Fk} g_i$ (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 $HW(g_i) \leq HW(g_1)$ and $g_1 = g_i + \Sigma_{h_i \in Fk}\, h_i\cdot a_i$ for monomials a_i with $HW(h_i\cdot a_i) \leq HW(g_1)$. Lemma 8.6 yields $g_1 = g_i + \Sigma_{h_i' \in Fk+1}\, h_i'\cdot b_i$ for monomials b_i with $HW(h_i'\cdot b_i) \leq HW(g_1)$ and since $g_i \in F_{k+1}$ or $g_i = 0$ we have $g_1 = \Sigma_{h_i'' \in Fk+1}\, h_i''\cdot c_i$ for monomials c_i with $HW(h_i''\cdot c_i) \leq HW(g_1)$. 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 is easy to see that, for any k, $\langle F_k\rangle = \langle F_{k+1}\rangle$.

This lemma and the next proposition shows, that it is reasonable to call the result of Algorithm 8.3 a 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 w.r.t. $\to_G$.

PROOF. The proof is similar to the proof of Lemma 2.4 in Mora (1985).

Obviously, $f \in \langle G\rangle$ means $f = \Sigma_{g_i \in G}\, g_i\cdot a_i$ for some monomials a_i. Let $t := \max\{ \ldots HW(g_i\cdot a_i) \ldots \}$ and $I := \{ i;\, HW(g_i\cdot a_i) = t \}$.

Case 1. $|I| = 1$.

Then $HW(f) = t$ and (for $I = \{ j \}$ and $a_j = c_j\cdot r_j$ ($c_j \in \mathbb{Z}$, $r_j \in W_n$)) $HW(f) = t = HW(g_j)\cdot r_j$ and $HC(f) = HC(g_j)\cdot c_j$. 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 $f_1 \overset{*}{\to}_G 0$ and thus $f \to_G f_1 \overset{*}{\to}_G 0$.

Case 2. $|I| > 1$.
Let i, j be two different elements of I and let $a_i = c_i \cdot r_i$, $a_j = c_j \cdot r_j$ ($c_i, c_j \in \mathbb{Z}$, $r_i, r_j \in W_n$). Since $HW(g_i) \cdot r_i = t = HW(g_j) \cdot r_j$, either $HW(g_i)$ is a prefix of $HW(g_j)$ or vice versa. Without loss of generality we assume $HW(g_i) = HW(g_j) \cdot r$ for some $r \in W_n$. By Lemma 8.7, $HC(g_j) = HC(g_i) \cdot c$ for some $c \in \mathbb{Z}$ and $g_j \cdot r - g_i \cdot c = \Sigma_{h_k \in G}\, h_k \cdot b_k$, where $HW(h_k \cdot b_k) < HW(g_i) = HW(g_j \cdot r)$. Hence $g_j \cdot r_j - g_i \cdot r_i c = (g_j \cdot r - g_i \cdot c) \cdot r_i = \Sigma_{h_k \in G}\, h_k \cdot (b_k r_i)$, where $HW(h_k \cdot (b_k r_i)) < HW(g_i) \cdot r_i = t$.
Now $f = (g_j \cdot r_j - g_i \cdot r_i c) \cdot c_j + g_i \cdot (c_i + cc_j) r_i + \Sigma_{m \neq i,j}\, g_m \cdot a_m$
$= \Sigma_{h_k \in G}\, h_k \cdot (b_k c_j r_i) + g_i \cdot (c_i + cc_j) r_i + \Sigma_{m \neq i,j}\, g_m \cdot a_m$ 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 $I = \langle\{ f_1, \ldots, f_m \}\rangle$ be a finitely generated right ideal in $\mathbb{Z}\langle X_1,\ldots,X_n\rangle$. We apply Algorithm 8.3 to $F_0 = \{ f_1, \ldots, f_m \}$ and get a set G of polynomials. Now $f \in I$ iff f can be reduced to 0 w.r.t. $\rightarrow_G$. If f is $\rightarrow_G$-irreducible, then $f \in I$ iff $f = 0$. Otherwise we can effectively find some g such that $f \rightarrow_G g$ 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 "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 that described in Section 6.
Let (*) $f_1x_1 + \ldots + f_rx_r = f_0$ 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 (**) $f_1x_1 + \ldots + f_rx_r = 0$.
Let $G = \{ g_1, \ldots, g_s \}$ be the output of Algorithm 8.3. when started with input $\{ f_1, \ldots, f_r \}$. There exist an $r \times s$-matrix P and an $s \times r$-matrix Q with entries in $\mathbb{Z}\langle X_1,\ldots,X_n\rangle$ such that $\underline{f} \cdot P = \underline{g}$ and $\underline{g} \cdot Q = \underline{f}$. This matrices can be obtained as by-products of Algorithm 8.3.
Obviously, (*) has a solution iff $f_0 \in \langle\{ f_1, \ldots, f_r \}\rangle = \langle G\rangle$. Hence, if (*) has a solution, Proposition 8.9 implies that f_0 reduces to 0 w.r.t $\rightarrow_G$. This yields $p_1, \ldots, p_s \in \mathbb{Z}[X_1,\ldots,X_n]$ 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 $|z^{(1)}, \ldots, |z^{(L)}$ of the set of solutions of the equation

$$(++)\ g_1x_1 + \ldots + g_sx_s = 0.$$

As in Section 6 one can show

LEMMA 9.1. The vectors $P\cdot \mathrm{l}z^{(1)}, \ldots, P\cdot \mathrm{l}z^{(L)}$ and the columns of the matrix $PQ - E_r$ 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 $HW(g_i) = HW(g_j)\cdot r$ for some $r \in W_n$, the equation (++) has no nontrivial solutions. Otherwise, let i, j ($i \neq j$) be indices, such that $HW(g_i) = HW(g_j)\cdot r$ for some $r \in W_n$.
By Lemma 8.7, $HC(g_j) = HC(g_i)\cdot c$ for some $c \in \mathbb{Z}$, $r \neq 1$ and

$$g_j\cdot r - g_i\cdot c = \sum_{k=1}^{k=r} g_k\cdot h_k$$

for polynomials $h_k \in \mathbb{Z}\langle X_1,\ldots,X_n\rangle$ with $HW(g_k\cdot h_k) < HW(g_i)$. 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 $\mathrm{l}q_{ij} := (q_1,\ldots, q_s)^T$ is a solution of (++).

LEMMA 9.2. The finitely many vectors $\mathrm{l}q_{ij}$ generate all solutions of (++).

PROOF. Let $\mathrm{l}p = (p_1,\ldots,p_s)^T$ be a nontrivial solution of (++). The complexity of such a solution is given by (t, α), where $t := \max\{ HW(g_ip_i);\ 1 \leq i \leq s \}$ and $\alpha := |\{ i;\ 1 \leq i \leq s$ and $HW(g_ip_i) = t \}|$.
Since $g\cdot \mathrm{l}p = 0$ and $\mathrm{l}p$ is not trivial, α has to be greater than 1. Hence there exist i, j ($i \neq j$) such that $HW(g_i)HW(p_i) = t = HW(g_j)HW(p_j)$. Without loss of generality we assume that $HW(g_j)$ is a prefix of $HW(g_i)$. Thus $HW(g_i) = HW(g_j)\cdot r$ and $HC(g_j) = HC(g_i)\cdot c$ for some $r \in W_n$ and $c \in \mathbb{Z}$ and $HW(p_j) = r\cdot HW(p_i)$. Let $c_i := HC(p_i)$ and $c_j := HC(p_j)$.
The vector $\mathrm{l}q_{ij}$, which was defined above, is a solution of (++). We define a new solution $(p_1',\ldots, p_s')^T = \mathrm{l}p' := \mathrm{l}p + \mathrm{l}q_{ij}\cdot c_jHW(p_i)$ and show that it has smaller complexity than $\mathrm{l}p$. To that purpose we have to consider the words $HW(g_kp_k')$ for all k, $1 \leq k \leq s$.

CASE 1. $k \neq i, j$.
We have $g_kp_k' = g_kp_k + g_kh_kc_jHW(p_i)$ and $HW(g_k\cdot h_k) < HW(g_i)$. This implies that $HW(g_kh_kc_jHW(p_i)) < HW(g_i)HW(p_i) = t$. Thus $HW(g_kp_k') = t$, if $HW(g_kp_k) = t$, and otherwise, $HW(g_kp_k') < t$.

Case 2. $k = i$.
We have $g_ip_i' = g_ip_i + g_icc_jHW(p_i)$. Hence $HW(g_ip_i') = t$ if $c_i + cc_j \neq 0$ and $HW(g_ip_i') < t$ if $c_i + cc_j = 0$.

Case 3. $k = j$.

$$\begin{aligned} g_jp_j' &= g_jp_j + g_jh_jc_jHW(p_i) - g_jrc_jHW(p_i) \\ &= HC(g_j)c_jt + R(g_jp_j) + g_jh_jc_jHW(p_i) - HC(g_j)c_jHW(g_j)rHW(p_i) - R(g_j)rc_jHW(p_i) \\ &= R(g_jp_j) + g_jh_jc_jHW(p_i) - R(g_j)rc_jHW(p_i), \text{ since } rHW(p_i) = HW(g_j). \end{aligned}$$

This shows that $HW(g_jp_j') < t$.

Thus we have seen that the complexity of the solution lp' is smaller than the complexity of lp and the lemma is proved by induction.

EXAMPLE 9.3. As an example we consider the homogeneous linear equation $f_1x_1 + \ldots + f_4x_4 = 0$ in $\mathbb{Z}$<a,b,c,d> 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

$$Q = \begin{pmatrix} 0 & 1 & d & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & -3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 1 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 0 & -1 & 3 & 0 & 0 \\ 1 & c & -2c & 2d & -5d \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

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

(1) $g_1\cdot c - g_2\cdot 3 = g_3$ and thus $lq_{1,2} = (-c, 3, 1, 0, 0)^T$.

(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 $lq_{1,4} = (-d, 0, -3, 3, -2)^T$.

We now apply P, to get the corresponding solutions of $f_1x_1 + \ldots + f_4x_4 = 0$:

$P\cdot lq_{1,2} = (0, 0, 0, 0)^T$ and $P\cdot lq_{1,4} = (-9, 6c + 15d, -9, -6)^T$.

$$\text{The matrix } PQ - E_4 \text{ is } \begin{pmatrix} 0 & 0 & -9 & 3 \\ 0 & 0 & 6c+15d & -2c-5d \\ 0 & 0 & -9 & 3 \\ 0 & 0 & -6 & 2 \end{pmatrix}.$$

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

10. Conclusion

The categorical reformulation of E-unification allows to characterize the class of commutative theories by properties of the category C(E) of finitely generated E-free objects: C(E) has to be a semiadditive category. The definition of semiadditive categories provides an algebraic structure on the morphism sets, which 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 syntactic 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, i.e. the theory of abelian groups with n commuting homomorphisms, can be derived with the help of well-known algebraic methods (e.g. Gröbner Base algorithms) to solve linear equations in $\mathbb{Z}[X_1,...,X_n]$. In order to obtain a unification algorithm for the theory AGnH of abelian groups with n non-commuting homomorphisms, we developed a Gröbner base algorithm for the ring $\mathbb{Z}<X_1,...,X_n>$ of polynomials over $\mathbb{Z}$ in n non-commuting indeterminates. Since Dicksons Lemma (Dickson (1913)), which is used for $\mathbb{Z}[X_1,...,X_n]$ to prove termination of the Gröbner Base algorithm, does not hold for $\mathbb{Z}<X_1,...,X_n>$, 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 admissible ordering. Hence it would be interesting to have a complete characterization of all admissible orderings for $\mathbb{Z}<X_1,...,X_n>$.

References

Baader, F. (1986). The Theory of Idempotent Semigroups is of Unification Type Zero. *J. Automated Reasoning* **2**.

Baader, F. (1987). Unification in Varieties of Idempotent Semigroups. *Semigroup Forum* **36**.

Baader, F. (1988). Unification in Commutative Theories. To appear in *J. Symbolic Computation.*

Baader, F. (1989). Unification Properties of Commutative Theories: A Categorical Treatment. Proceedings of the Summer Conference on Category Theory and Computer Science, Manchester (England).

Baader, F., Büttner, W. (1988). Unification in Commutative Idempotent Monoids. *TCS* **56**.

Baader, F., Nutt, W. (1989). In preparation.

Bachmair, L., Buchberger, B. (1980). A Simplified Proof of the Characterization Theorem of Gröbner-Bases. *ACM-SIGSAM Bulletin* **14**.

Buchberger, B. (1985). Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory. In Bose, N. K. (Ed.). *Recent Trends in Multidimensional System Theory.*

Büttner, W. (1986). Unification in the Data Structure Multiset. *J. Automated Reasoning* **2**.

Cohn, P.M. (1965). *Universal Algebra.* New York: Harper and Row.

Dershowitz, N., Manna, Z. (1979). Proving Termination with Multiset Orderings. *J. ACM* **22**.

Dickson, L.E. (1913). Finiteness of the Odd Perfect and Primitive Abundant Numbers with n Distinct Factors. *Amer. J. Math.* **35**.

Eilenberg, S. (1974). *Automata, Languages and Machines, Volume A.* New York: Academic Press.

Fages, F. (1984). Associative-Commutative Unification. Proceedings of the CADE '84 Napa (USA). *Springer Lec. Notes Comp. Sci.* **170**.

Fages, F., Huet, G. (1986). Complete Sets of Unifiers and Matchers in Equational Theories. *TCS* **43**.

Fortenbacher, A. (1985). An Algebraic Approach to Unification under Associativity and Commutativity. Proceedings of the RTA '85 Dijon (France). *Springer Lec. Notes Comp. Sci.* **202**.

Freyd, P. (1964). *Abelian Categories.* New York: Harper and Row.

Furukawa, A., Sasaki, T., Kobayashi, H. (1986). Gröbner Basis of a Module over $K[X_1,...,X_n]$ and Polynomial Solutions of Systems of Linear Equations. Proceedings of the Symsac '86, Waterloo, Ontario.

Grätzer, G. (1968). *Universal Algebra*. Princeton: Van Nostrand Company.

Herold, A. (1987). Combination of Unification Algorithms in Equational Theories. Ph.D. Dissertation, Universität Kaiserslautern.

Herrlich, H., Strecker, G.E. (1973). *Category Theory*. Boston: Allyn and Bacon Inc.

Huet, G. (1980). Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems. *J. ACM* **27**.

Jaffar, J., Lassez, J.L., Maher, M.J. (1984). A Theory of Complete Logic Programs with Equality. *J. Logic Programming* **1**.

Jacobson, N. (1980). *Basic Algebra II*. San Francisco: Freeman and Company.

Jouannaud, J.P., Kirchner, H. (1986). Completion of a Set of Rules Modulo a Set of Equations. *SIAM J. Comp.* **15**.

Kandri-Rody, A., Kapur, D. (1988). Computing a Gröbner Basis of a Polynomial Ideal over a Euclidean Domain. *J. Symbolic Computation* **6**.

Kandri-Rody, A., Weispfenning, V. (1988). Non-Commutative Gröbner Bases in Algebras of Solvable Type. Preprint.

Kuich, W., Salomaa, A. (1986). *Semirings, Automata, Languages*. Berlin: Springer Verlag.

Lankford, D., Butler, G., Brady, B. (1984). Abelian Group Unification Algorithms for Elementary Terms. *Contemporary Mathematics* **29**.

Leeb, K., Pirillo, G. (1988). Shuffle-Compatible Total Orders. To appear in *Annali Matematica*.

Livesey, M., Siekmann, J. (1978). Unification in Sets and Multisets. SEKI Technical Report, Universität Kaiserslautern.

Martin, U. (1988). A Geometrical Approach to Multiset Orderings. Technical Report, University of London, RHBNC.

Mora, F. (1986). Gröbner Bases for Non-Commutative Polynomial Rings. Proceedings of the AAECC3. *Springer Lec. Notes Comp. Sci.* **228**.

Nutt, W. (1988). Talk at the Second Workshop on Unification, Val d' Ajol (France).

Plotkin, G. (1972). Building in Equational Theories. *Machine Intelligence* **7**.

Robbiano, L. (1985). Term Orderings on the Polynomial Ring. Proceedings of the EUROCAL '85. *Springer Lec. Notes Comp. Sci.* **204**.

Rydeheard, D.E., Burstall, R.M. (1985). A Categorical Unification Algorithm. Proceedings of the Workshop on Category Theory and Computer Programming. *Springer Lec. Notes Comp. Sci.* **240**.

Schmidt-Schauß, M. (1986). Unification under Associativity and Idempotence is of Type Nullary. *J. Automated Reasoning* **2**.

Siekmann, J. (1988). Unification Theory. *J. Symbolic Computation* **7**, Special Issue on Unification.

Stickel, M. (1981). A Unification Algorithm for Associative-Commutative Functions. *J. ACM* **28**.

Stickel, M. (1985). Automated Deduction by Theory Resolution. *J. Automated Reasoning* **1**.

Trevisan, G. (1953). Classifica ione dei semplici ordinamenti di un gruppo libero commutativo con m generatori. *Rendiconti del Seminario Matematico della Universita di Padova* **22**.

Zaiceva, M.I. (1953). On the Set of Ordered Abelian Groups. *Uspehi Matem. Nauk (N.S)* **8**.