Unification in Commutative Theories

Franz Baader

Institut für Mathematische Maschinen und Datenverarbeitung (Informatik) 1,
Universität Erlangen-Nürnberg, Martensstraße 3, 8520 Erlangen, F.R.G.

(Received 22 December 1987)

A general framework for unification in "commutative" theories is investigated which is based on a categorical reformulation of theory unification. We thus obtain the well-known results for abelian groups, abelian monoids and idempotent abelian monoids as well as some new results as corollaries to a general theorem. In addition, it is shown that constant-free unification problems in "commutative" theories are either unitary or of unification type zero and we give an example of a "commutative" theory of type zero.

1. Introduction

Equational theories which are of unification type "finitary" play an important role in automated theorem provers with built-in equational theories (see e.g. Plotkin (1972), Nevins (1974), Slage (1974) or Stickel (1985)) and in generalizations of the Knuth-Bendix algorithm (see e.g. Huet (1980), Peterson-Stickel (1981), Jouannaud (1983) and Jouannaud-Kirchner (1986)). Examples of finitary theories are the theory of abelian groups (Lankford-Butler-Brady (1984)), the theory of abelian monoids (Livesey-Siekmann (1978), Stickel (1981), Fages (1984), Fortenbacher (1985), Büttner (1986), Herold (1987)) and the theory of idempotent abelian monoids (Livesey-Siekmann (1978), Baader-Büttner (1988)). The proofs of these finitary-results make use of the following property which the three theories have in common: The finitely generated free objects are direct products of the free objects in one generator.

This paper is concerned with equational theories which satisfy this and some additional properties. In Section 5 we give a characterization of these theories which justifies the name "commutative theories". A categorical reformulation of E-unification (Rydeheard-Burstall (1985)) shows that commutative theories correspond to semiadditive categories, i.e. categories which allow a binary operation on morphisms distributing with the composition of morphisms (Section 4). Using this fact we get sufficient conditions for commutative theories to be finitary. The above mentioned results for abelian groups etc. and some new results (for abelian monoids with an involution, idempotent abelian monoids with an involution, abelian groups with an involution and abelian groups of exponent m) can thus be obtained as corollaries to a general theorem. This shows which parts of the proofs are common for all these theories and which parts are specific for the theory in question. Furthermore we shall show that constant-free unification problems in commutative theories are either unitary or of unification type zero and we give an
example of a commutative theory of type zero.

Before starting with the details, I would like to point out two advantages of a categorical setting for the description of unification problems. First, unification theory is not only interested in specific unification algorithms, but also in general results for whole classes of theories. Therefore an appropriate level of abstraction has to be found which allows to exhibit common structures. This paper shows that — at least for "commutative" theories — categories yield such a level of abstraction. Second, well-known results about certain categories — here semiaadditive categories — can be exploited to obtain unification theoretic results.

In the following we assume that the reader is familiar with the basic notions of universal algebra (see e.g. Cohn (1965), Gratzer (1968)). For more information about unification theory see Siekmann (1986). The composition of mappings and morphisms will be written from left to right, i.e. \( f \circ g \) means first \( f \) and then \( g \). Consequently we use suffix notation for mappings.

2. E-unification

Let \( E \) be an equational theory and \( =_E \) be the equality of terms, induced by \( E \). We assume that terms are \( \Omega \)-terms (with variables) for a given signature \( \Omega \). For a function symbol \( f \) in \( \Omega \) we shall write \( f \) for its realization in any \( \Omega \)-algebra. An E-unification problem is a finite set of equations denoted by \( \Gamma = \langle s_i = t_i; 1 \leq i \leq n \rangle_E \) where \( s_i \) and \( t_i \) are terms. A substitution \( \theta \) is called an E-unifier of \( \Gamma \) iff \( s_i \theta = t_i \theta \) for each \( i, i = 1, \ldots, n \).

The set of all E-unifiers of \( \Gamma \) is denoted by \( U_E(\Gamma) \). We are mostly interested in complete sets of E-unifiers, i.e. sets of E-unifiers from which \( U_E(\Gamma) \) may be generated by instantiation. More formally, 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 \( i, i = 1, \ldots, n \),

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

In this case \( \sigma \) is called an E-instance of \( \theta \). As usual the quasi-ordering \( \leq_E \) induces an equivalence relation \( =_E \) on \( U_E(\Gamma) \), namely \( \sigma =_E \theta \) iff \( \sigma \leq_E \theta \) and \( \theta \leq_E \sigma \). A complete set \( cU_E(\Gamma) \) of E-unifiers of \( \Gamma \) 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 \( \theta \leq_E \sigma \).

A minimal complete set \( \mu U_E(\Gamma) \) is a complete set of E-unifiers of \( \Gamma \) satisfying the minimality condition

3. For all \( \sigma, \theta \in \mu U_E(\Gamma) \) \( \sigma \leq_E \theta \) implies \( \sigma = \theta \).

A set \( \mu U_E(\Gamma) \) may not always exist, but if it does it is unique up to \( =_E \)-equivalence (Fages-Huet (1986)). Consequently equational theories may be classified according to the cardinality or existence of \( \mu U_E(\Gamma) \) as follows:

1. \( \mu U_E(\Gamma) \) exists for all E-unification problems \( \Gamma \) and has at most one element then \( E \) is called unitary.
(2) If $\mu U_E(\Gamma)$ exists for all $E$-unification problems $\Gamma$ and has finite cardinality then $E$ is called \textit{finitary}.

(3) If $\mu U_E(\Gamma)$ exists for all $E$-unification problems $\Gamma$ and for some $E$-unification problem is denumerable then $E$ is called \textit{infinitary}.

(4) If for some $E$-unification problem $\Gamma$, $\mu U_E(\Gamma)$ does not exist then $E$ is said to be of \textit{unification type zero}.

3. A Categorical Reformulation of $E$-unification

An equational theory $E$ defines a \textit{variety} $V(E)$, i.e. the class of all algebras (of the given signature $\Omega$) which satisfy each identity of $E$. For any set $X$ of generators, $V(E)$ contains a \textit{free algebra over} $V(E)$ \textit{with generators} $X$, which will be denoted by $F_E(X)$. Thus any mapping of $X$ into an algebra $B \in V(E)$ can be uniquely extended to a homomorphism of $F_E(X)$ into $B$.

Let $\Gamma = \{ s_i \equiv t_i \mid 1 \leq i \leq n \}$ 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 $\Gamma$ 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

$$\sigma, \tau: F_E(I) \rightarrow F_E(X)$$

by $x_i\sigma := s_i$ and $x_i\tau := t_i$ ($i = 1, ..., n$).

Now $\delta: F_E(X) \rightarrow F_E(Y)$ is an $E$-unifier of $\Gamma$ 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 $\sigma, \tau: F_E(I) \rightarrow F_E(X)$ in the following category:

\textbf{Definition 3.1.} Let $E$ be an equational theory and $V$ be a denumerable set. Then the category $C(E)$ is defined as follows:

(1) The objects of $C(E)$ are the algebras $F_E(X)$ for finite subsets $X$ of $V$. We denote the class of these objects by $F(E)$.

(2) The morphisms of $C(E)$ are the homomorphisms between these objects.

(3) The composition of morphisms is the usual composition of mappings.

Note that in $C(E)$ epimorphisms need not be surjective. But the isomorphisms (in the categorical sense) are just the bijective homomorphisms. Two objects $F_E(X)$, $F_E(Y)$ of $C(E)$ are isomorphic iff $|X| = |Y|$. An $E$-unifiers of the unification problem $\langle \sigma = \tau \rangle_E$ is a morphism $\delta$ such that $\sigma\delta = \tau\delta$. For morphisms $\sigma: F_E(X) \rightarrow F_E(Y)$, $\gamma: F_E(X) \rightarrow F_E(Z)$ we have $\sigma \leq_E \gamma$ iff there is a morphism $\lambda: F_E(Z) \rightarrow F_E(Y)$ such that $\sigma = \gamma\lambda$.

Now the notions complete and minimal complete set of $E$-unifiers and unification type of a theory $E$ are defined as in Section 2.

In this paper we shall consider equational theories $E$ such that $C(E)$ is a semiadditive
category. Thus the well-known structure of these categories (see Freyd (1964) and Herrlich-Strecker (1973)) can be exploited to obtain results about unification properties of these theories.

4. Semiadditive Categories

Before defining semiadditive categories we recall some basic concepts of category theory. Let \( C \) be a category and \( A, B \) be objects of \( C \). We denote by \( \text{hom}(A,B) \) the set of morphisms with domain \( A \) and codomain \( B \). The identity morphism in \( \text{hom}(A,A) \) is denoted by \( 1_A \) or just \( 1 \). We say that the object \( P \) is a **product** of \( A, B \) iff there exist morphisms \( p_1: P \to A, p_2: P \to B \) such that for every pair of morphisms \( f: X \to A, g: X \to B \) there is a unique morphism \( h: X \to P \) such that the product diagram of Figure 4.1 commutes.

\[
\begin{array}{ccc}
X & \xrightarrow{h} & P \\
\downarrow{f} & & \downarrow{p_1} \\
A & \xrightarrow{p_2} & B \\
\end{array}
\]

**FIGURE 4.1**

A product of two objects may not exist, but if it exists it is unique up to isomorphism. We denote the product of \( A \) and \( B \) by \( A \times B \) and call the corresponding morphisms projections. The dual of the product is the coproduct. An object \( S \) is a **coproduct** of \( A, B \) iff there exist morphisms \( u_1: A \to S, u_2: B \to S \) such that for every pair of morphisms \( f: A \to X, g: B \to X \) there is a unique morphism \( h: S \to X \) such that the coproduct diagram of Figure 4.1 commutes. We denote the coproduct of \( A \) and \( B \) by \( A + B \) and call the corresponding morphisms injections. Products and coproducts of more than two objects are defined in an analogous way. Given a coproduct \( S \) of the objects \( A_1, \ldots, A_n \) and a product \( P \) of the objects \( B_1, \ldots, B_m \), every morphism \( f: S \to P \) is uniquely determined by the matrix \( M_f = (f_{i,j}) \) where \( f_{i,j} := u_i f p_j \in \text{hom}(A_i, B) \) for \( i = 1, \ldots, n \) and \( j = 1, \ldots, m \). For \( n = 1 \) (resp. \( m = 1 \)) we take \( u_1 = 1 \) (resp. \( p_1 = 1 \)) in this definition.

An object \( A \) is called **initial** (terminal) iff for every object \( B \), \( \text{hom}(A,B) \) (\( \text{hom}(B,A) \)) is a singleton. An object which is both initial and terminal is called **zero object**. If \( C \) has a zero object \( 0 \) we define the zero morphism \( 0_{A,B}: A \to B \) to be the composite of the unique morphism in \( \text{hom}(A,0) \) and the unique morphism in \( \text{hom}(0,B) \). It is easy to see that in this definition it does not matter which zero object of \( C \) is used. Let \( f: C \to A, g: B \to C \) be morphisms. Then we have \( f \circ 0_{A,B} = 0_{C,B} \) and \( 0_{A,B} \circ g = 0_{A,C} \). In the following we shall omit the index and write \( 0 \) for any zero morphism.

Now we can define semiadditive categories:
DEFINITION 4.2. A category C is *semiadditive* iff

1. C has a zero object.
2. For every pair of objects there is a coproduct.
3. For any pair of objects A, B there is a binary operation "+" on hom(A,B) such that
   3.1. 0_{A,B} is a neutral element for "+" on hom(A,B).
   3.2. For any objects A, B, C, D and any morphisms a, b ∈ hom(A,B), c ∈ hom(C,A) and d ∈ hom(B,D) we have c(a + b) = ca + cb and (a + b)d = ad + bd.

The following theorem yields an alternative characterization of semiadditive categories.

THEOREM 4.3. Let C be a category satisfying (1) and (2) of Definition 4.2 and let A+B with the morphisms u_1, u_2 be a coproduct of A and B. The morphisms p_1: A+B → A, p_2: A+B → B are defined by the commuting diagrams of Figure 4.4.

![Figure 4.4](image URL)

Then the following statements are equivalent:

1. For any pair of objects A, B the coproduct A+B is also the product of A, B relative to the morphisms p_1, p_2 defined above.
2. C is semiadditive.

PROOF. The following sketch of the proof is included to give an idea of how these two properties are linked. Complete proofs can be found in Freyd (1964) and Herrlich-Strecker (1973).

If C satisfies (1), "+" may be defined as follows:

For a, b ∈ hom(A,B)  a + b := (a,b) o (1).

Recall that (a,b) and (1) are the unique morphisms such that the diagrams of Figure 4.5 commute.

![Figure 4.5](image URL)
It can be shown that "+" satisfies (3.1) and (3.2) of Definition 4.2. Moreover "+" is associative and commutative (see e.g. Freyd (1964) pp. 47-49).

Conversely let \( C \) satisfy (2) and let \( a : X \to A \), \( b : X \to B \) be morphisms. Then \( x = au_1 + bu_2 \) is the unique morphism such that \( x \pi_1 = a \) and \( x \pi_2 = b \) (see Freyd (1964) p. 50).

**Corollary 4.6.** The operation "+" of Definition 4.2 is unique and thus associative and commutative.

**Proof.** See Freyd (1964) p. 60 or Herrlich-Strecker (1973), Theorem 40.13.

Let \( C \) be a semiadditive category and let \( A \) (resp. \( B, C \)) be coproduct and product of the objects \( A_1 \) (resp. \( B_j, C_k \)). The morphisms \( a : A \to B \) and \( b : B \to C \) are uniquely determined by the matrices \( M_a = (a_{ij}) \) and \( M_b = (b_{jk}) \). It can be shown (see Freyd (1964) p. 49) that the usual rules of matrix multiplication hold, i.e. we have

**Corollary 4.7.** Let \( c = ab \) and \( M_c = (c_{ik}) \). Then \( c_{ik} = \sum_j a_{ij} b_{jk} \).

5. Commutative Theories

We now characterize the equational theories for which \( C(E) \) is a semiadditive category. A constant symbol (i.e. a nullary function symbol) \( \varepsilon \in \Omega \) is called idempotent in \( E \), iff for any \( f \in \Omega \) we have \( f(\varepsilon, \ldots, \varepsilon) = \varepsilon \), i.e. in any algebra \( A \in V(E) \), \( f(\varepsilon, \ldots, \varepsilon) = \varepsilon \) holds. Note that for nullary \( f \) this means \( f = \varepsilon \).

Let \( K \) be a class of algebras (of signature \( \Omega \)). An \( n \)-ary implicit operation in \( K \) is a family \( \{ f_A : A \in K \} \) of mappings \( f_A : A^n \to A \) which is compatible with all homomorphisms, i.e. for any homomorphism \( h : A \to B \) with \( A, B \in K \) and all \( a_1, \ldots, a_n \in A \), \( f_A(a_1, \ldots, a_n)h = f_B(a_1h, \ldots, a_nh) \) holds. In the following we omit the index and just write \( f \) for any \( f_A \). Obviously an \( \Omega \)-term induces an implicit operation on any class of \( \Omega \)-algebras.

**Example 5.1.** Let \( \Omega \) consist of a nullary function symbol \( \varepsilon \) and a binary function symbol \( \ast \). \( E := \{ x \ast x = e, x \ast e = x, (x \ast (e \ast y)) \ast (e \ast z) = (z \ast (e \ast y)) \ast (e \ast x) \} \).

The \( \Omega \)-terms \( x \ast (e \ast y) \) and \( e \ast x \) define a binary implicit operation \( \ast \) and a unary implicit operation \( ^{-1} \) in \( V(E) \) and \( F(E) \) as follows:

Let \( A \in V(E) \) and \( a, b \in A \). Then \( a \ast b := a \ast (e \ast b) \) and \( a^{-1} := e \ast a \).

It is easy to see, that \( \ast, ^{-1} \) satisfy the abelian group axioms in any \( A \in V(E) \). The constant symbol \( \varepsilon \) is idempotent in \( E \).

Since \( V(E) \) contains (\( F(E) \) consists of) all \( E \)-free algebras with finite set of generators, any implicit operation in \( V(E) \) (resp. \( F(E) \)) is given by an \( \Omega \)-term (see Lawvere (1963)).

In the following we assume that \( E \) is not trivial, i.e. that \( V(E) \) contains algebras of car-
The following proposition characterizes the theories $E$ for which $C(E)$ has a zero object.

**Proposition 5.2.** The following three conditions are equivalent:

1. $C(E)$ contains a zero object.
2. $|F_E(\emptyset)| = 1$.
3. $\Omega$ contains a constant symbol $e$, which is idempotent in $E$.

**Proof.** The equivalence of (2) and (3) is obvious. The definition of $E$-free objects shows that $F_E(\emptyset)$ is initial in $C(E)$. If in addition $|F_E(\emptyset)| = 1$ it is also terminal. This proves "2$\rightarrow$1". Conversely, since $F_E(\emptyset)$ is initial, any zero object of $C(E)$ is isomorphic to $F_E(\emptyset)$. It is easy to see that for $|F_E(\emptyset)| > 1$ we have at least two morphisms of $F_E(x)$ into $F_E(\emptyset)$ and for $|F_E(\emptyset)| = 0$ we have no morphism of $F_E(x)$ into $F_E(\emptyset)$. Hence "1$\rightarrow$2" holds. □

Evidently the zero morphism of $F_E(x)$ into $F_E(Y)$ is defined by $x \mapsto e$ for $x \in X$.

Condition (2) of Definition 4.2 holds in $C(E)$ for any equational theory $E$. In fact, the coproduct of $F_E(X)$ and $F_E(Y)$ is given by $F_E(X \uplus Y)$ where $\uplus$ means disjoint union. Note that $F_E(X)$ is the coproduct of the objects $F_E(x)$ for $x \in X$ and the $F_E(x)$ are isomorphic to each other. We now consider Condition (3) of Definition 4.2.

**Proposition 5.3.** Let $C(E)$ contain a zero object and let $e$ be the constant symbol which is idempotent in $E$. Then the following assertions are equivalent:

1. $C(E)$ is semiadditive.
2. There is a binary implicit operation $\circ$ in $F(E)$ such that
   
   2.1. The constant $e$ is a neutral element for $\circ$ in any algebra $A \in F(E)$.
   2.2. For any $n$-ary function symbol $f \in \Omega$, any algebra $A \in F(E)$ and any $s_1, \ldots, s_n, t_1, \ldots, t_n \in A$ we have $f(s_1 \circ t_1, \ldots, s_n \circ t_n) = f(s_1, \ldots, s_n) \circ f(t_1, \ldots, t_n)$.

**Proof.** We first prove (1) implies (2). The operation $+$ on the morphisms of $C(E)$ induces an operation $\circ$ in the $F_E(X)$ as follows: Given $s, t \in F_E(X)$ we define morphisms $\sigma, \tau: F_E(x) \to F_E(X)$ by $\sigma: x \mapsto s, \tau: x \mapsto t$. Now define $s \circ t := x(\sigma + \tau)$.

The operation $\circ$ is implicit, since for $\lambda: F_E(X) \to F_E(Y)$ we have $(s \circ t)\lambda = x(\sigma + \tau)\lambda = x(\sigma\lambda + \tau\lambda) = (s\lambda) \circ (t\lambda)$. Assertion (2.1) holds, since the zero morphism is neutral for $\circ$. To show that (2.2) holds, we consider morphisms $\sigma, \tau: F_E(x_1, \ldots, x_n) \to F_E(X)$, defined by $x_i\sigma = s_i$ ( $i = 1, \ldots, n$ ) and $x_i\tau = t_i$ ( $i = 1, \ldots, n$ ). Then $x_i(\sigma + \tau) = s_i \circ t_i$ and since $\sigma + \tau$ is a homomorphism we have

$$f(x_1, \ldots, x_n)(\sigma + \tau) = f(x_1(\sigma + \tau), \ldots, x_n(\sigma + \tau)) = f(s_1 \circ t_1, \ldots, s_n \circ t_n).$$

On the other hand we may consider $\gamma: F_E(\emptyset) \to F_E(x_1, \ldots, x_n)$ defined by $\gamma = f(x_1, \ldots, x_n)$.
This yields \( f(x_1, \ldots, x_n)(\sigma + \tau) = (x_1\sigma + \tau) = \langle x_1\sigma + x_1\tau \rangle = (x_1\sigma) \ast (x_1\tau) = f(s_1, \ldots, s_n) \ast f(t_1, \ldots, t_n) \).

Conversely, let \( E \) satisfy (2). The implicit operation \( \ast \) induces an operation \( + \) on the morphisms of \( C(E) \). Let \( \sigma, \tau: F_E(X) \to F_E(Y) \) and \( s \in F_E(X) \). Then \( s(\sigma + \tau) := (s\sigma) \ast (s\tau) \).

Using (2.2) it is easy to show that \( \sigma + \tau \) is really a morphism of \( C(E) \), i.e. a homomorphism of \( F_E(X) \) into \( F_E(Y) \). Obviously (2.1) yields (3.1) of Definition 4.2. The fact that \( \ast \) is an implicit operation yields (3.2) of Definition 4.2. 

The implicit operation \( \ast \) of Example 5.1 satisfies (2.1) and (2.2) of the proposition. Thus for the theory \( E \) of the example, \( C(E) \) is semiadditive.

**COROLLARY 5.4.**
The implicit operation of Theorem 5.3 is associative and commutative.

**PROOF.** This is an immediate consequence of Corollary 4.6. 

Note that the explicit operation \( \ast \) of Example 5.1 is neither commutative nor associative. Corollary 5.4 justifies the following definition.

**DEFINITION 5.5.**
An equational theory \( E \) is called commutative iff \( C(E) \) is semiadditive.

We now consider examples of commutative theories. In all these examples the implicit operation is given by a function symbol which is associative and commutative in the corresponding theory.

**EXAMPLES 5.6.** We consider the following signatures:

1. The theory of abelian monoids. The signature is \( \Omega_1 \) and 
   \( AM := \{ \cdot, 1 \} \) where \( \cdot \) is binary and 1 is nullary.
2. The theory AIM of idempotent abelian monoids. The signature is \( \Omega_1 \) and 
   \( AIM := AM \cup \{ x \cdot x = x \} \).
3. The theory AIMH of idempotent abelian monoids with a homomorphism. The signature is \( \Omega_3 \) and 
   \( AIMH := AIM \cup \{ h(x \cdot y) = h(x) \cdot h(y), h(1) = 1 \} \).
4. The theory AMI of abelian monoids with an involution. The signature is \( \Omega_2 \) and 
   \( AMI := AM \cup \{ x \cdot x = x \} \).
5. The theory AIMI of idempotent abelian monoids with an involution. We have signature \( \Omega_3 \) and 
   \( AIMI := AMI \cup \{ x \cdot x = x \} \).
(6) The theory $\text{AG}_m$ of abelian groups of exponent $m$ ($m \in \mathbb{N}$) is given by
signature $\Omega_2$ and $\text{AG}_m := \text{AM} \cup \{ x \cdot x^{-1} = 1, x^m = 1 \}$. $\text{AG} = \text{AG}_0$ is the theory of abelian groups. Any variety of abelian groups is defined in this way.

(7) The theory $\text{AGI}$ of abelian groups with an involution.
We take signature $\Omega$ and define $\text{AGI} := \text{AG} \cup \text{AMI}$.

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 (2.2) of Proposition 5.3 for $f = \cdot$ iff $(a \cdot b) \cdot (c \cdot d) = (a \cdot c) \cdot (b \cdot d)$ holds in any algebra $A \in F(E)$.
We shall now consider unification in commutative theories and, in the end, determine the unification types of the theories defined above.

6. Commutative Theories and Unification

First we show the following theorem:

THEOREM 6.1. Commutative theories are either unitary or of unification type zero.

This is an easy consequence of the following two lemmata. Now let $E$ be a commutative theory.

LEMMA 6.2. Let $\Gamma = \langle \sigma = \tau > E$ be an $E$-unification problem and let $\{ \gamma_1, \ldots, \gamma_n \}$ be a finite complete set of $E$-unifiers of $\Gamma$. Then there exists an $E$-unifier $\gamma$ of $\Gamma$ such that the singleton $\{ \gamma \}$ is a complete set of $E$-unifiers of $\Gamma$.

PROOF. We have $\sigma, \tau: F_E(I) \to F_E(X)$ and $\gamma_i: F_E(X) \to F_E(Y_i)$.

With $Y = Y_1 \circ \ldots \circ Y_n$, $F_E(Y)$ is coproduct and product of the $F_E(Y_i)$. Let $p_1, \ldots, p_n$ be the corresponding projections. Then there exists a unique morphism $\gamma: F_E(X) \to F_E(Y)$ such that $\gamma p_i = \gamma_i$ for $i = 1, \ldots, n$. The morphism $\gamma$ is an $E$-unifier of $\Gamma$, since $\sigma \gamma = \tau \gamma$ iff $\sigma \gamma p_i = \tau \gamma p_i$ for $i = 1, \ldots, n$ (by the definition of product).

Let $\delta$ be an $E$-unifier of $\Gamma$. Since $\{ \gamma_1, \ldots, \gamma_n \}$ is complete there is an index $i$ and a morphism $\lambda$ such that $\delta = \gamma_i \lambda$. But now $\delta = (\gamma p_i) \lambda = \gamma (p_i \lambda)$. Hence $\{ \gamma \}$ is complete. \( \square \)

LEMMA 6.3. Let $\Gamma = \langle \sigma = \tau > E$ be an $E$-unification problem and let $U = \{ \gamma_1, \gamma_2, \gamma_3, \ldots \}$ be an infinite set of $E$-unifiers of $\Gamma$ such that the $\gamma_i$ do not lie (w.r.t. $\leq_E$) below a single $E$-unifier of $\Gamma$. Then there does not exist a minimal complete set $\mu U_E(\Gamma)$.

PROOF. We have $\sigma, \tau: F_E(I) \to F_E(X)$ and $\gamma_i: F_E(X) \to F_E(Y_i)$.

The morphisms $\delta_n$ are defined as follows: $\delta_1$ is just $\gamma_1$.

Let $\delta_n: F_E(X) \to F_E(Z_n)$ be already defined and let $F_E(Z_{n+1})$ with the projections $p_1, p_2$ be a product of $F_E(Y_{n+1})$ and $F_E(Z_n)$. Then $\delta_{n+1}: F_E(X) \to F_E(Z_{n+1})$ is defined to be the
unique morphism such that \( \delta_{n+1}p_1 = \gamma_{n+1} \) and \( \delta_{n+1}p_2 = \delta_n \) (see Figure 6.4).

\[
\begin{array}{ccc}
F_E(X) & \xrightarrow{\delta_n} & F_E(Z) \\
\gamma_{n+1} & \downarrow & \delta_{n+1} \\
F_E(Y_{n+1}) & \xrightarrow{p_1} & F_E(Z_{n+1}) \\
\delta_n & \downarrow & \delta_{n+1} \\
F_E(Z) & \xrightarrow{\theta} & F_E(Y)
\end{array}
\]

**Figure 6.4**

It is easy to see that the morphisms \( \delta_n \) are \( E \)-unifiers of \( F \) and that \( \delta_n \leq E \delta_{n+1} \) for all \( n \geq 1 \). The condition imposed on \( U \) implies, that the increasing chain \( \delta_1 \leq E \delta_2 \leq E \delta_3 \leq E \ldots \) has no upper bound in \( U_E(\Gamma) \).

Now we assume that \( \mu U_E(\Gamma) \) exists and derive a contradiction. Since \( \mu U_E(\Gamma) \) is complete there is \( \theta \in \mu U_E(\Gamma) \) such that \( \delta_1 \leq E \theta \). The fact that \( \delta_1 \leq E \delta_2 \leq E \delta_3 \leq E \ldots \) has no upper bound in \( U_E(\Gamma) \) yields an \( n \geq 1 \) satisfying \( \delta_n \leq E \theta \) but not \( \delta_{n+1} \leq E \theta \). Let \( \theta : F_E(X) \to F_E(Y) \) and let \( F_E(Z) \) with the the projections \( q_1, q_2 \) be the product of \( F_E(Z_{n+1}) \) and \( F_E(Y) \). The morphism \( \theta : F_E(X) \to F_E(Z) \) is defined to be the unique morphism such that \( \theta q_1 = \delta_{n+1} \) and \( \theta q_2 = \delta_n \) (see the right diagram of Figure 6.4). Obviously \( \theta \) is an \( E \)-unifier of \( \Gamma \), \( \delta_{n+1} \leq E \theta \) and \( \theta \leq E \theta \). Since \( \mu U_E(\Gamma) \) is complete there is \( \theta' \in \mu U_E(\Gamma) \) such that \( \theta \leq E \theta' \). Now \( \theta \leq E \theta' \) for \( \theta, \theta' \in \mu U_E(\Gamma) \) yields \( \theta = \theta' \) by minimality of \( \mu U_E(\Gamma) \). But then \( \delta_{n+1} \leq E \theta' \leq E \theta = \theta \) is a contradiction.

A similar argument was used in Baader (1987) to show that most varieties of idempotent semigroups are defined by type zero theories. In Section 8 it will be shown that the theory AIMH of Example 5.6 is of type zero. In the remaining part of this section we establish a sufficient condition for a commutative theory to be unitary.

Let \( \Gamma = < \sigma = \tau >_E \) (where \( \sigma, \tau : F_E(\Gamma) \to F_E(X) \)) be an \( E \)-unification problem.

**Lemma 6.5.** Let \( \delta : F_E(X) \to F_E(Z) \) be an \( E \)-unifier of \( \Gamma \) and let \( y \) be an arbitrary variable. Then \( \{ \delta \} \) is a complete set of \( E \)-unifiers of \( \Gamma \), if for any \( E \)-unifier \( \gamma : F_E(X) \to F_E(Y) \) of \( \Gamma \) there is a morphism \( \lambda : F_E(Z) \to F_E(Y) \) such that \( \gamma = \delta \lambda \).

**Proof.** Let \( \alpha : F_E(X) \to F_E(Y) \) be an \( E \)-unifier of \( \Gamma \) and let \( |Y| = n \). Hence \( F_E(Y) \) is the \( n \)-th power of \( F_E(Y) \). We call the corresponding projections \( p_1, \ldots, p_n \). Since \( \alpha p_i : F_E(X) \to F_E(Y) \) are \( E \)-unifier of \( \Gamma \) the assumption yields morphisms \( \lambda_i : F_E(Z) \to F_E(Y) \) such that \( \alpha p_i = \delta \lambda_i \).
Let $\lambda: F_\mathcal{E}(Z) \to F_\mathcal{E}(Y)$ be the unique morphism such that $\lambda p_i = \lambda_i$ for $i = 1, ..., n$. Now $\alpha p_i = \delta \lambda_i = (\delta \lambda)p_i$ for $i = 1, ..., n$ and thus $\alpha = \delta \lambda$. $\square$

Hence if we want to check completeness of a set of $\mathcal{E}$-unifiers, we only have to consider unifiers which introduce a single variable. Therefore the following condition only considers unifiers of $F_\mathcal{E}(X)$ into $F_\mathcal{E}(y)$.

**CONDITION 6.6.** Let $y$ be an arbitrary variable. For any $\mathcal{E}$-unification problem $< \sigma = \tau >_\mathcal{E}$ (where $\sigma, \tau: F_\mathcal{E}(I) \to F_\mathcal{E}(X)$) there are finitely many $\mathcal{E}$-unifiers $\alpha_1, ..., \alpha_r: F_\mathcal{E}(X) \to F_\mathcal{E}(y)$ such that any $\mathcal{E}$-unifier $\delta: F_\mathcal{E}(X) \to F_\mathcal{E}(y)$ is representable as

$$\delta = \sum_{i=1}^{r} \alpha_i n_i \quad (n_i \in \mathbb{N}).$$

Here $\emptyset$ is the zero morphism and for $n \in \mathbb{N} \setminus \{0\}$ the symbol $n$ stands for $1 + ... + n$-times ($1$ is the identity morphism). Note that $n\alpha = \alpha + ... + \alpha = n\alpha$.

**THEOREM 6.7.** A commutative theory $\mathcal{E}$ satisfying Condition 6.6 is unitary.

**PROOF.** Let $\Gamma = < \sigma = \tau >_\mathcal{E}$ be an $\mathcal{E}$-unification problem and let $\alpha_1, ..., \alpha_r: F_\mathcal{E}(X) \to F_\mathcal{E}(y)$ be as in the condition. For $Z = \{ z_1, ..., z_r \}, F_\mathcal{E}(Z)$ is the $r$-th power and the $r$-th co-power of $F_\mathcal{E}(y)$. Let $p_1, ..., p_r (u_1, ..., u_r)$ be the corresponding projections (injections) such that $u_i p_i = 1$ and $u_i p_j = 0$ (for $i \neq j$).

We define an $\mathcal{E}$-unifier $\alpha: F_\mathcal{E}(X) \to F_\mathcal{E}(Z)$ as follows: $\alpha$ is the unique morphism such that $\alpha p_i = \alpha_i$ for $i = 1, ..., n$. Thus $\alpha = (\alpha_1, ..., \alpha_r)$.

Obviously $\alpha$ is an $\mathcal{E}$-unifier of $\Gamma$. It remains to be shown that $\{ \alpha \}$ is complete. Because of Lemma 6.5 we only have to consider $\mathcal{E}$-unifiers $\beta: F_\mathcal{E}(X) \to F_\mathcal{E}(y)$. Condition 6.6 yields non-negative integers $n_1, ..., n_r$ such that

$$\beta = \sum_{i=1}^{r} \alpha_i n_i.$$

Let $\lambda$ be the unique morphism such that $u_i \lambda = n_i$ for $i = 1, ..., n$.

Thus $\lambda = \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix}$ in matrix notation. Now $\alpha \lambda = (\alpha_1, ..., \alpha_r) \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \sum_{i=1}^{r} \alpha_i n_i = \beta$,

which completes the proof of the theorem. $\square$

If $F_\mathcal{E}(X)$ and $F_\mathcal{E}(y)$ are finite, the set $\text{hom}(F_\mathcal{E}(X), F_\mathcal{E}(y))$ is also finite and Condition 6.6 trivially holds. This yields

**COROLLARY 6.8.** Let $\mathcal{E}$ be a commutative theory for which the finitely generated free objects are finite. Then $\mathcal{E}$ is unitary.
The theories AIM, AIMI and $AG_m$ for $m \geq 1$ are examples of commutative theories satisfying the assumption of Corollary 6.8. Thus we have

**PROPOSITION 6.9.** The theories AIM, AIMI and $AG_m$ for $m \geq 1$ are unitary.

Direct proofs for AIM can be found in Livesey-Siekmann (1978) and Baader-Böttner (1988).

### 7. Unification with Constants in Commutative Theories

In some applications of E-unifications it is advantageous to consider – in addition to the variables – syntactical constants. These constants have no meaning in the theory $E$, i.e. they behave like variables w.r.t. $=_E$. But they differ from variables in that they must not be replaced by substitutions. In the categorical context this can be formulated as follows: Let $V$ (vid. Definition 3.1) be the disjoint union of denumerable sets $U$ and $C$. A $c$-morphism is a morphism of $C(E)$ which is the identity on elements of $C$. The subcategory $C_c(E)$ of $C(E)$ is obtained from $C(E)$ by restricting the morphisms to $c$-morphisms. Now $E_c$-unification is defined as E-unification with $C_c(E)$ in place of $C(E)$.

In the following let $I, W, X, Y, Z$ (resp. $A, B$) denote finite subsets of $U$ (resp. $C$) and let $E$ be a commutative theory. A unitary theory $E$ need not be unitary w.r.t. $E_c$-unification. In this section a sufficient condition is established for a commutative theory to be finitary w.r.t. $E_c$-unification. Let $\Gamma = < \sigma = \tau >_E$, where $\sigma, \tau: F_E(I) \to F_E(X \cup A)$, be an $E_c$-unification problem. First note that we may confine ourselves to $E_c$-unifiers which do not introduce new constants, i.e. to $c$-morphisms $\gamma: F_E(X \cup A) \to F_E(Y \cup A)$. Otherwise the additional constants can be replaced by new variables which yields an $E_c$-unifier $\delta$ such that $\gamma \leq_E \delta$ (see Baader-Böttner (1988), Lemma 3.1).

**CONDITION 7.1.** For any morphism (of $C(E)$) $\delta: F_E(A) \to F_E(Z)$ there exist finite sets $M, N$ such that:

1. The elements of $M$ are morphisms $\mu: F_E(Z) \to F_E(A)$ satisfying $\delta \mu = 1$.
2. The elements of $N = \{ v_1, ..., v_r \}$ are morphisms $v_i: F_E(Z) \to F_E(A)$ with $\delta v_i = 0$.
3. For any $\lambda: F_E(Z) \to F_E(A)$ with $\delta \lambda = 1$ there are $\mu \in M$ and non-negative integers $n_1, ..., n_r$ satisfying:

$$\lambda = \mu + \sum_{i=1}^{r} n_i v_i$$

**THEOREM 7.2.**

A unitary commutative theory satisfying Condition 7.1 is finitary w.r.t. $E_c$-unification.

**PROOF.** We assume that $E$ is unitary and satisfies Condition 7.1. Let $\Gamma = < \sigma = \tau >_E$, where $\sigma, \tau: F_E(I) \to F_E(X \cup A)$, be an $E_c$-unification problem. First we consider $\Gamma$ as E-unification problem, i.e. the elements of $A$ are treated as variables. Let $\delta: F_E(X \cup A) \to$
be an $E$-unifier of $\Gamma$ such that $\{ \delta \}$ is a complete set of $E$-unifiers. Since any $E_c$-unifier $\gamma: F_E(X \cup A) \rightarrow F_E(Y \cup A)$ is also an $E$-unifier there is a morphism $\lambda$ (which need not be a $c$-morphism) satisfying $\gamma = \delta \lambda$. Since $E$ is commutative, $F_E(X \cup A)$ (resp. $F_E(Y \cup A)$) is coproduct and product of $F_E(X)$ and $F_E(A)$ (resp. $F_E(Y)$ and $F_E(A)$). Thus $\gamma$ is uniquely determined by a matrix

$$
\begin{pmatrix}
\gamma_1 & \gamma_2 \\
\gamma_3 & \gamma_4
\end{pmatrix}
$$

where $\gamma_1: F_E(X) \rightarrow F_E(Y)$, $\gamma_2: F_E(X) \rightarrow F_E(A)$,
$\gamma_3: F_E(A) \rightarrow F_E(Y)$, $\gamma_4: F_E(A) \rightarrow F_E(A)$.

Since $\gamma$ is an $E_c$-unifier, we have $\gamma_3 = 0$ and $\gamma_4 = 1$. Accordingly $\delta$ is determined by a matrix

$$
\begin{pmatrix}
\delta_1 \\
\delta_2
\end{pmatrix}
$$

where $\delta_1: F_E(X) \rightarrow F_E(Z)$,
$\delta_2: F_E(A) \rightarrow F_E(Z)$,

and $\lambda$ is determined by a matrix $(\lambda_1, \lambda_2)$ where $\lambda_1: F_E(Z) \rightarrow F_E(Y)$ and $\lambda_2: F_E(Z) \rightarrow F_E(A)$. Now $\gamma = \delta \lambda$ yields $\delta_1 \lambda_1 = \gamma_1$, $\delta_1 \lambda_2 = \gamma_2$, $\delta_2 \lambda_1 = 0$ and $\delta_2 \lambda_2 = 1$.

Applying Condition 7.1 for $\delta_2: F_E(A) \rightarrow F_E(Z)$ we get finite sets $M, N$ of morphisms satisfying (1), (2), (3) of 7.1. Thus there are $\mu \in M$ and $n_1, \ldots, n_r \in \mathbb{N}$ such that

$$
\lambda_2 = \mu + \sum_{i=1}^{r} n_i \eta_i
$$

Since $\delta_2 \lambda_1 = 0 = 0 \lambda_1$, $\lambda_1$ is an $E$-unifier of the $E$-unification problem $\Delta = \langle \delta_2 = 0 \rangle_E$ where $0: F_E(A) \rightarrow F_E(Z)$ is the zero morphism. Let $\kappa$ be a complete set of $E$-unifiers of $\Delta$, where $\kappa: F_E(Z) \rightarrow F_E(W)$. Then there is a morphism $\rho: F_E(W) \rightarrow F_E(Y)$ with the property $\lambda_1 = \kappa \rho$.

Let $X_1, \ldots, X_r$ be subsets of $U$ of cardinality $|A|$. Then $F_E(X_1) \times \cdots \times F_E(X_r) \times F_E(A)$ is an $(r+1)$th power of $F_E(A)$. Thus the matrix $(\kappa \nu_1 \ldots \nu_r \mu)$ defines a morphism $F_E(Z) \rightarrow F_E(W) \times F_E(X_1) \times \cdots \times F_E(X_r) \times F_E(A)$. Note that this morphism only depends on $\delta_2$ but not on $\lambda$. Furthermore we have

$$
(\kappa \nu_1 \ldots \nu_r \mu)
$$

$$
= (\kappa \rho + \sum_{i=1}^{r} n_i \eta_i)
$$

and the morphism

$$
\begin{pmatrix}
\rho & 0 \\
0 & n_1 \\
\vdots & \\
0 & n_r \\
0 & 1
\end{pmatrix}
$$

: $F_E(W \otimes X_1 \otimes \cdots \otimes X_r \otimes A) \rightarrow F_E(Y \otimes A)$ is obviously a $c$-morphism.
We define $K = \{ \delta \circ (\kappa \nu_1 ... \nu_r \mu) ; \mu \in M \}$, which is a finite set of morphisms of $F_E(X \cup A)$ into $F_E(W \cup X_1 \cup ... \cup X_r \cup A)$. It remains to be shown that $K$ is a complete set of $E_c$-unifiers of $\Gamma$. Since $\delta$ is an $E$-unifier of $\Gamma$, the elements of $K$ are also $E$-unifiers of $\Gamma$. They are $E_c$-unifiers, because,

$$
\begin{pmatrix}
\delta_1 \\
\delta_2
\end{pmatrix}
(\kappa \nu_1 ... \nu_r \mu) =
\begin{pmatrix}
\delta_1 \kappa & \delta_1 \nu_1 & ... & \delta_1 \nu_r & \delta_1 \mu \\
\delta_2 \kappa & \delta_2 \nu_1 & ... & \delta_2 \nu_r & \delta_2 \mu
\end{pmatrix}
$$

and Condition 7.1 asserts $\delta_2 \nu_1 = ... = \delta_2 \nu_r = 0$, $\delta_2 \mu = 1$, and we have $\delta_2 \kappa = 0$ by definition of $\kappa$.

In the preceding we have seen that an arbitrary $E_c$-unifier $\gamma$ is of the form $\gamma = \delta \lambda$ where

$$
\lambda = (\kappa \nu_1 ... \nu_r \mu) \begin{pmatrix}
\rho & 0 \\
0 & n_1 \\
... \\
0 & n_r \\
0 & 1
\end{pmatrix}
$$

for some $n_1, ..., n_r \in N$. Thus

$$
\gamma = (\delta \circ (\kappa \nu_1 ... \nu_r \mu)) \begin{pmatrix}
\rho & 0 \\
0 & n_1 \\
... \\
0 & n_r \\
0 & 1
\end{pmatrix}
$$

which shows completeness of $K$.

But then a minimal complete set of $E_c$-unifiers exists and is also finite. This completes the proof of Theorem 7.2. 

This proof shows that the cardinality of the minimal complete set is bounded by the cardinality of $M$.

**COROLLARY 7.3.** If the set $M$ of Condition 7.1 is a singleton for all $E_c$-unification problems $\Gamma$, then $E$ is unitary w.r.t $E_c$-unification.

As in the case of $E$-unification we have

**COROLLARY 7.4.** Let $E$ be a commutative theory for which the finitely generated free objects are finite. Then $E$ is finitary w.r.t. $E_c$-unification.

And thus

**PROPOSITION 7.5.** The theories $AIM$, $AIMI$ and $AG_m$ for $m \geq 1$ are finitary w.r.t. unification with constants.

Note that $AG_m$ for $m \geq 1$ is even unitary. This can be shown analogously to the proof of
Unification in Commutative Theories

8. Examples

In the following section we shall consider the theories of example 5.6 which do not satisfy the finiteness condition for the finitely generated free objects.

8. Examples

In the first part of this section it is shown that the theory $\text{AIMH}$ of idempotent abelian monoids with a homomorphism (see Example 5.6) is of unification type zero.

We consider the unification problem $F = \langle h(x_1)h(x_2) = x_2h^2(x_3) \rangle_{\text{AIMH}}$. For $n \geq 0$ the substitutions $\theta_n$ are defined by

\[
\begin{align*}
  x_1\theta_n &= y, \\
  x_2\theta_n &= h(y)h^2(y)...h^{n+1}(y), \\
  x_3\theta_n &= h^n(y).
\end{align*}
\]

We have $(h(x_1)h(x_2))\theta_n = h(y)h^2(y)...h^{n+1}(y) = (x_2h^2(x_3))\theta_n$ which shows that the $\theta_n$ are $\text{AIMH}$-unifiers of $F$.

\[\text{LEMMA 8.1. Let } \gamma \text{ be an } \text{AIMH}-\text{unifier of } F \text{ such that } \theta_n \leq \text{AIMH } \gamma. \text{ Then } x_3\gamma \text{ contains } h^n(z) \text{ as factor for some variable } z.\]

\[\text{PROOF. Let } \lambda \text{ be a substitution such that } \theta_n = \text{AIMH } \gamma \circ \lambda. \text{ Since } x_1\gamma\lambda = \text{AIMH } x_1\theta_n = y, \text{ } x_1\gamma \text{ has to contain a variable } z \text{ such that } z\lambda = y. \text{ But then } h(z) \text{ appears in } (h(x_1)h(x_2))\gamma = \text{AIMH } (x_2h^2(x_3))\gamma. \text{ This implies that } h(z) \text{ is a factor of } x_2\gamma, \text{ which yields } h^2(z) \text{ on the left side of the above equation. Let } h^i(z) (i \geq 2) \text{ be a factor of } (h(x_1)h(x_2))\gamma. \text{ If } h^i(z) \text{ is not in } (h^2(x_3))\gamma, \text{ we have } h^i(z) \text{ in } x_3\gamma \text{ and thus } h^{i+1}(z) \text{ in } (h(x_1)h(x_2))\gamma. \text{ Hence to stop this process, there must be an } i \geq 2 \text{ such that } h^{i-2}(z) \text{ is a factor of } x_3\gamma. \text{ But then } h^{i-2}(y) \text{ is a factor of } x_3\gamma\lambda, \text{ which yields } i-2 = n. \]

An immediate consequence of this lemma is, that there can be no $\text{AIMH}$-unifier of $F$ which lies (w.r.t. $\leq_{\text{AIMH}}$) above $\theta_n$ for all $n \geq 0$. Thus $\Gamma$ and $\U = \{ \theta_1, \theta_2, \ldots \}$ satisfy the assertion of Lemma 6.3 and we have proved

\[\text{PROPOSITION 8.2. The theory } \text{AIMH} \text{ is of unification type zero.}\]

This is the first example of a commutative theory of type zero. In the remaining part of this section it will be shown that the theories AM, AMI, AG and AGI (see Example 5.6) are unitary (resp. finitary w.r.t. unification with constants).

Let $X = \{ x_1, \ldots, x_n \}$, $Y = \{ y_1, \ldots, y_m \}$ and let $\delta: F_{\text{AG}}(X) \to F_{\text{AG}}(Y)$ be a morphism of $\text{C(AG)}$. Then $\delta$ is uniquely determined by an $n \times m$-matrix $M_{\delta} \in M_{n,m}(\mathbb{Z})$ where $(M_{\delta})_{ij} \in \mathbb{Z}$ is the exponent of $y_j$ in $x_i\delta$. Sum and composition of morphisms correspond
to sum and product of matrices, i.e. $M_{\sigma \delta} = M_{\sigma}M_{\delta}$ and $M_{\sigma + \delta} = M_{\sigma} + M_{\delta}$. The morphism $\delta$ is the zero morphism iff all entries of $M_{\delta}$ are zero and for $X = Y$, $\delta$ is the identity morphism iff $M_{\delta}$ is the unit matrix $E$.

Accordingly any morphism $\delta: F_{AM}(X) \rightarrow F_{AM}(Y)$ of $C(AM)$ corresponds to a matrix $M_{\delta} \in M_{n,m}(N)$.

For a morphism $\delta: F_{AGI}(X) \rightarrow F_{AGI}(Y)$ of $C(AGI)$ we define matrices $A_{\delta}, B_{\delta}$ as follows: $A_{\delta} = (a_{i,j})$ where $a_{i,j} \in Z$ is the exponent of $y_j$ in $x_i \delta$.

$B_{\delta} = (b_{i,j})$ where $b_{i,j} \in Z$ is the exponent of $h(y_j)$ in $x_i \delta$.

We associate with the morphism $\delta$ the matrix

$$M_\delta = \begin{pmatrix} A_\delta & B_\delta \\ B_\delta & A_\delta \end{pmatrix} \in M_{2n,2m}(Z)$$

It is easy to see that $M_{\sigma \delta} = M_{\sigma}M_{\delta}$ and $M_{\sigma + \delta} = M_{\sigma} + M_{\delta}$. The set

$$D_{n,m}(Z) = \left\{ \begin{pmatrix} A & B \\ B & A \end{pmatrix}; A, B \in M_{n,m}(Z) \right\}$$

with addition of matrices is a subgroup of the abelian group $M_{2n,2m}(Z)$.

Accordingly any morphism $\delta: F_{AMI}(X) \rightarrow F_{AMI}(Y)$ of $C(AMI)$ corresponds to a matrix $M_{\sigma} \in D_{n,m}(N)$. The set

$$D_{n,m}(N) = \left\{ \begin{pmatrix} A & B \\ B & A \end{pmatrix}; A, B \in M_{n,m}(N) \right\}$$

with addition of matrices is a stable subsemigroup of $M_{2n,2m}(N)$.

A subsemigroup $T$ of a commutative semigroup $S$ is called stable iff for all $a, b \in S$ $a, a+b \in T$ implies $b \in T$. Note that $S$ is a stable subsemigroup of $S$.

For morphisms $\sigma, \tau, \delta, \lambda$ of $C(AG)$ (resp. $C(AM), C(AGI), C(AMI)$) we have

$\sigma \delta = \tau \delta$ iff $M_{\sigma \delta} = M_{\tau \delta}$ iff $M_{\sigma}M_{\delta} = M_{\tau}M_{\delta}$ iff $(M_{\sigma} - M_{\tau})M_{\delta} = 0$ and

$\sigma \lambda = 1$ iff $M_{\sigma \lambda} = M_1 = E$ iff $M_{\sigma}M_{\lambda} = E$.

Thus Condition 6.6 and Condition 7.1 for AG, AM, AGI and AMI translate into statements about certain solutions of systems of linear diophantine equations.

**Lemma 8.3.** Let $G$ be a subgroup of $M_{n,m}(Z)$ and let $A$ be an element of $M_{k,n}(Z)$. The solutions of the equation $AX = 0$ in $G$ are $Z$-linear combinations of $r \leq n.m$ basic solutions in $G$.

**Proof.** Let $U = \{ Y \in M_{n,m}(Z); AY = 0 \}$ be the set of solutions in $M_{n,m}(Z)$. $U$ and hence $U \cap G$ are subgroups of the free abelian group of rank $n.m$ $M_{n,m}(Z)$.

Hence $U \cap G$ is a free abelian group of rank $r \leq n.m$ (see e.g. Kurosh (1960), p. 145).


If we take these basic solutions and their negatives we get any solution in $G$ as $\mathbb{N}$-linear combination of $s \leq 2n-m$ solutions in $G$. This yields Condition 6.6 for $AG$ and $AGI$. The solutions of the inhomogeneous equation are obtained in the usual way.

**Lemma 8.4.** Let $G$ be a subgroup of $M_{n,m}(\mathbb{Z})$ and let $A, B$ be elements of $M_{k,m}(\mathbb{Z}), M_{k,m}(\mathbb{Z})$. Let $Y_0 \in G$ be an arbitrary solution of $AX = B$. Any solution $Y \in G$ is of the form $Y = Y_0 + Z$ where $Z$ is a solution of $AX = 0$.

This yields Condition 7.1 for $AG$ and $AGI$. Note that we need only one special solution $Y_0$ of $AX = B$. Thus we have

**Proposition 8.5.** The theories $AG$ and $AGI$ are unitary. In addition they are unitary w.r.t. unification with constants.

See Lankford-Butler-Brady (1984) for a direct proof of the result for $AG$.

**Lemma 8.6.** Let $S$ be a stable subsemigroup of $M_{n,m}(\mathbb{N})$ and let $A$ be an element of $M_{k,n}(\mathbb{Z})$. The solutions of the equation $AX = 0$ in $S$ are $\mathbb{N}$-linear combinations of finitely many basic solutions.

**Proof.** The proof is similar to that of Corollary 9.19 of Clifford-Preston (1967).

On the elements of $M_{n,m}(\mathbb{N})$ we define the partial ordering $\leq$ component wise. Since $\leq$ on $\mathbb{N}$ is a well partial ordering (wpo), its n-m-fold cartesian product on $M_{n,m}(\mathbb{N})$ is also a wpo (see Nash-Williams (1963)). The set $H = \{Y \in S; AY = 0\}$ is a stable subsemigroup of $M_{n,m}(\mathbb{N})$. Let $H_0$ be the set of minimal elements of $H$. $H_0$ is finite (finite antichains) and any element of $H$ lies above a minimal element (finite chains). We now show that $H_0$ is the set of basic solutions we are searching for. For any solution $Y \in H \setminus H_0$ there is a solution $Y_0 \in H_0$ such that $Y < Y_0$, i.e. there exists $Z \in M_{n,m}(\mathbb{N})$ with $Y = Y_0 + Z$. Since $H$ is stable we have $Z \in H$ and obviously $Z < Y$. By noetherian induction we thus have proved the lemma.

**Lemma 8.7.** Let $S$ be a stable subsemigroup of $M_{n,m}(\mathbb{N})$ and let $A, B$ be elements of $M_{k,m}(\mathbb{Z}), M_{k,m}(\mathbb{Z})$. There exists a finite set $T_0$ of solutions of $AX = B$ in $S$ such that any solution $Y$ of $AX = B$ in $S$ is of the form $Y = Y_0 + Z$ where $Y_0 \in T_0$ and $Z \in S$ is a solution of $AX = 0$.

**Proof.** Let $T_0$ be the finite set of minimal elements in $T = \{Y \in S; AY = B\}$. For $Y \in T_0$ we have $Y = Y + 0$. Otherwise $Y_0 < Y$ for some $Y_0 \in T_0$, i.e. $Y = Y_0 + Z$ for some $Z \in M_{n,m}(\mathbb{N})$. Since $S$ is stable and $Y, Y_0 \in S$ we have $Z \in S$. Now $B = AY = AY_0 + AZ = B + AZ$ yields $AZ = 0$.\[\square\]
Lemma 8.6 and 8.7 establish Condition 6.6 and 7.1 for AM and AMI. Thus we have

**PROPOSITION 8.8.** The theories AM and AMI are unitary and they are finitary w.r.t. unification with constants.

For alternative proofs of this result for AM see the references in the Introduction. Effective methods to solve systems of linear diophantine equations in \( \mathbb{Z} \) can be found in Niven-Zuckerman (1972) and Knuth (1973). For solutions in \( \mathbb{N} \) see e.g. Makanin (1977) (Lemma 1.1), Huet (1978), Fortenbacher (1985), Lambert (1987) and Clausen-Fortenbacher (1988). Efficient unification algorithms for the theories AG, AM, AGI and AMI depend upon efficient implementations of these methods.

**9. Conclusion**

In this paper we were less interested in deriving efficient unification algorithms for a specific theory. Instead, we gave a general framework for unification in the whole class of commutative theories. An important result is the fact that commutative theories where the finitely generated objects are finite, are always unitary (finitary w.r.t. unification with constants). But even in this case the construction of an efficient unification algorithm which computes the most general unifier is yet another problem. This algorithm should produce unifiers which introduce a minimal number of variables (i.e. the number \( r \) of E-unifiers in Condition 6.6 should be as small as possible). In the case of unification with constants we want to obtain a minimal complete set (i.e. the set \( M \) of Condition 7.1 has to be as small as possible) rather than just a complete set (see e.g. Baader-Büttner (1988) where this problem is solved for AIM).

**References**


