AAECC manuscript No.
(will be inserted by the editor)

Gröbner Basis Cryptosystems

Peter Ackermann and Martin Kreuzer

Fachbereich Mathematik, Universität Dortmund, D-44221 Dortmund, Germany
e-mail: peter.ackermann, martin.kreuzer@mathematik.uni-dortmund.de

January 14, 2005

Abstract In the first sections we extend and generalize Gröbner basis theory to submodules of free right modules over monoid rings. Over free monoids, we adapt the known theory for right ideals and prove versions of Macaulay's basis theorem, the Buchberger criterion, and the Buchberger algorithm. Over monoids presented by a finitely generated convergent string rewriting system we generalize Madlener's Gröbner basis theory based on prefix reduction from right ideals to right modules. After showing how these Gröbner basis theories relate to classical group-theoretic problems, we use them as a basis for a new class of cryptosystems that are generalizations of the cryptosystems described in [2] and [8]. Well known cryptosystems such as RSA, El-Gamal, Polly Cracker, Polly Two and a braid group cryptosystem are shown to be special cases. We also discuss issues related to the security of these Gröbner basis cryptosystems.

Key words Gröbner basis, cryptosystem, monoid ring

Mathematics Subject Classification (1991): Primary 94A60; Secondary 11T71, 13P10, 16-08

1 Introduction

In recent years, algebraic cryptosystems experienced a considerable increase in active interest, mainly due to the braid group based cryptosystem suggested in [1]. In this paper we want to unite recent attempts at group based cryptosystems with earlier suggestions by Neal Koblitz and others in a commutative setting (see [8] and [14]). Unfortunately, these earlier suggestions met a very polemic response by the Gröbner basis community (see the paper by Boo Barkee *et al.* [2]; note that Boo was the name of Moss Sweedler's

former dog). In the following years, there were only scattered attempts to rescue these Polly Cracker type cryptosystems, for instance in [29] and [18]. We find it particularly ironic that one of the authors of that polemic paper [2], Teo Mora aka. Theo Moriarty, had already helped to develop the very Gröbner basis theory for non-commutative rings which we feel is destined to overcome those initial objections (see [23], [24], and [25]).

Let us explain the basic setup. For non-commutative rings, there are several types of Gröbner basis theories. The natural way of defining Gröbner bases for free associative algebras (i.e. non-commutative polynomial rings) has been generalized to so-called basic algebras by Edward Green and his co-workers (see [6], [10], and [11]). These algebras are characterized by the property that they have a multiplicative vector space basis on which there exists a term ordering. Alas, for the monoid and group rings we are interested in such bases are usually not available. Thus we resort to another generalization of Mora's approach: Klaus Madlener, Birgit Reinert, and their co-workers (see [19], [20], [21], [22], and [27]) have successfully described a Gröbner basis theory for monoid and group rings based on a presentation of the monoid or group by a convergent term rewriting system.

In Sections 2, 3, and 4 we extend this theory to the case of submodules of finitely generated right or two-sided free modules over monoid rings. Furthermore, we prove a general version of Macaulay's basis theorem (see 2.2) which plays a key role in our cryptosystems, and we formulate generalizations of Buchberger's algorithm for enumerating Gröbner bases (see 2.7 and 4.12). In Section 5 we show how one can solve well-known problems (such as the word problem, the subgroup problem, the conjugacy problem, and the conjugator search problem) if one succeeds in computing a Gröbner basis of the appropriate module.

Then we introduce a new class of Gröbner basis cryptosystems in Section 6. We show that this class contains the original Polly Cracker cryptosystems (see [8] and [14]), Ly's Polly Two system (see [29] and [30]), the RSA cryptosystem (see [28]), the ElGamal cryptosystem (see [5]), and certain group based cryptosystems derived from [15]. Other special cases are the non-commutative polynomial cryptosystems of [26]. The underlying idea of our cryptosystems is straightforward: the plain text units are normal forms of elements w.r.t. a finitely generated module, encryption is achieved by adding a random element of the module, and decryption uses a Gröbner basis of the module and a reduction process to compute the normal form again. The security of such cryptosystems is based on a number of facts:

- Gröbner bases are usually difficult to compute
- the attacker knows only part of the module for which he wants to compute a Gröbner basis; the Gröbner basis of this part may be infinite
- the action of the monoid ring on the module can encode hard combinatorial and number theoretic problems
- the structure of the base ring may encode hard algebraic and combinatorial problems

In Section 7 we consider the efficiency of the computations that are involved when using the cryptosystem. Moreover we discuss how one can meet various attacks on the system.

A more general discussion about these and related issues follows in the last section, together with some suggestions for generating further secure instances of Gröbner basis cryptosystems.

2 Gröbner Bases for Right Modules over Free Monoid Rings

In the following we let $\Sigma = \{x_1, \ldots, x_n\}$ be a finite alphabet. The monoid of words (or terms) generated by Σ will be denoted by Σ^*. Its elements are of the form $w = x_{i_1} \cdots x_{i_s}$ with $i_1, \ldots, i_s \in \{1, \ldots, n\}$. Its neutral element is the empty word λ, and its multiplication is given by concatenation.

Let K be a field. The *free monoid ring* (or the *free associative algebra* or the *non-commutative polynomial ring*) of Σ^* over K is the set consisting of all formal sums $\sum_{i=1}^{s} c_i w_i$ with $c_i \in K$ and $w_i \in \Sigma^*$ together with the obvious addition and the multiplication defined by extending the multiplication in Σ^* linearly.

Our goal in this section is to develop some parts of a Gröbner basis theory for submodules of finitely generated free right modules over the free monoid ring $K[\Sigma^*]$. We shall content ourselves with introducing those results that are necessary to define and study the cryptosystems in Section 6. We begin by recalling the basics of the Gröbner basis theory for two-sided and right ideals of $K[\Sigma^*]$ described in [20], [21], [22], and [27].

Let σ be a *term ordering* on Σ^*, i.e. a total ordering such that:

1. The inequality $w_1 \leq_\sigma w_2$ implies $w_3 w_1 w_4 \leq_\sigma w_3 w_2 w_4$ for all elements $w_1, w_2, w_3, w_4 \in \Sigma^*$.
2. Every descending chain of terms $w_1 \geq_\sigma w_2 \geq_\sigma \cdots$ in Σ^* is eventually stationary, i.e. σ is a well-ordering.

Then every $f \in K[\Sigma^*] \setminus \{0\}$ has a unique representation $f = \sum_{i=1}^{s} c_i w_i$ with $c_i \in K \setminus \{0\}$ and $w_i \in \Sigma^*$ such that $w_1 >_\sigma w_2 >_\sigma \cdots >_\sigma w_s$. The term $\mathrm{LT}_\sigma(f) = w_1$ is called the *leading term* of f. Moreover, we let $\mathrm{LC}_\sigma(f) = c_1$ be the *leading coefficient* of f and $\mathrm{LM}_\sigma(f) = c_1 w_1$.

Given a two-sided ideal $I \subseteq K[\Sigma^*]$, the *leading term ideal* of I is defined to be the two-sided ideal generated by the leading terms of the non-zero elements of I and is denoted by $\mathrm{LT}^*_\sigma(I) = \langle \mathrm{LT}_\sigma(f) \mid f \in I \setminus \{0\} \rangle$. A set of elements $\{f_i \mid i \in \Lambda\} \subseteq I$ is called a *σ-Gröbner basis* of I if $\mathrm{LT}^*_\sigma(I) = \langle \mathrm{LT}_\sigma(f_i) \mid i \in \Lambda \rangle$.

Similarly, given a right ideal $I \subseteq K[\Sigma^*]$, the *right leading term ideal* of I is defined to be the right ideal generated by the leading terms of the non-zero elements of I and is denoted by $\mathrm{LT}_\sigma(I) = \langle \mathrm{LT}_\sigma(f) \mid f \in I \setminus \{0\} \rangle_\varrho$. (The subscript ϱ will be used to denote right modules, right generation, etc.) A set of elements $\{f_i \mid i \in \Lambda\} \subseteq I$ is called a *right σ-Gröbner basis* of I if $\mathrm{LT}_\sigma(I) = \langle \mathrm{LT}_\sigma(f_i) \mid i \in \Lambda \rangle_\varrho$.

Now we extend this theory to submodules of free right modules over the ring $K[\Sigma^*]$. Let Φ be a finite or countably infinite set, and let F_ϱ be the free right $K[\Sigma^*]$-module with basis $\{e_i \mid i \in \Phi\}$. The elements of F_ϱ can be represented as formal sums $\sum_{i\in\Phi} e_i f_i$ where all but finitely many of the elements $f_i \in K[\Sigma^*]$ are zero. Furthermore, let $U \subseteq F_\varrho$ be a finitely generated right submodule. In this situation we introduce Gröbner bases as follows.

Definition 2.1 *A* **term** *in F_ϱ is an element of the form $e_i m$ with $i \in \Phi$ and $m \in \Sigma^*$. The set of all terms in F will be denoted by $\mathbb{T}(F_\varrho)$. If $t = e_i m \in \mathbb{T}(F_\varrho)$ is a term and $c \in K$ by ct we will mean $e_i cm$.*

A **module term ordering** *on $\mathbb{T}(F_\varrho)$ is a total ordering τ such that:*

1. *The inequality $t_1 \leq_\tau t_2$ implies $t_1 w \leq_\tau t_2 w$ for all $t_1, t_2 \in \mathbb{T}(F_\varrho)$ and $w \in \Sigma^*$.*
2. *Every descending chain of terms $t_1 \geq_\tau t_2 \geq_\tau \cdots$ in $\mathbb{T}(F_\varrho)$ is eventually stationary, i.e. τ is a well-ordering.*

Given a module term ordering τ and an element $v = \sum_{i=1}^{s} c_i t_i \in F_\varrho \setminus \{0\}$ with $c_i \in K \setminus \{0\}$ and $t_i \in \mathbb{T}(F_\varrho)$ satisfying $t_1 >_\tau \cdots >_\tau t_s$, we say that $\mathrm{LT}_\tau(v) = t_1$ is the **leading term** *of v and $\mathrm{LC}_\tau(v) = c_1$ its* **leading coefficient**. *Moreover, we let $\mathrm{LM}_\tau(v) = c_1 t_1$.*

The right submodule $\mathrm{LT}_\tau(U) = \langle \mathrm{LT}_\tau(v) \mid v \in U \setminus \{0\} \rangle_\varrho$ of F_ϱ is called the **(right) leading term module** *of U, the set $\mathrm{LT}_\tau\{U\} = \{\mathrm{LT}_\tau(v) \mid v \in U \setminus \{0\}\}$ the* **leading term set** *of U.*

Finally, a set of non-zero vectors $\{v_i \mid i \in \Lambda\}$ in U is called a **(right) τ-Gröbner basis** *of U if we have $\mathrm{LT}_\tau\{U\} = \{\mathrm{LT}_\tau(v_i) w \mid i \in \Lambda, w \in \Sigma^*\}$.*

Note that $G = \{v_i \mid i \in \Lambda\}$ is a right τ-Gröbner basis of U if and only if $\mathrm{LT}_\tau(U) = \langle \mathrm{LT}_\tau(v_i) \mid i \in \Lambda \rangle_\varrho$.

The following result will become essential for our cryptosystems.

Proposition 2.2 (Macaulay's Basis Theorem)
Let τ be a module term ordering on $\mathbb{T}(F_\varrho)$ and $U \subseteq F_\varrho$ a right submodule. Then the residue classes of the terms in $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$ form a K-basis of F_ϱ/U.

Proof For $b \in F_\varrho$ let $\bar{b} \in F_\varrho/U$ denote the corresponding residue class. First suppose the residue classes of $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$ do not generate F_ϱ/U. Let $m \in F_\varrho$ such that $\bar{m} \notin \langle \bar{b} \mid b \in \mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\} \rangle_K =: A$. Since τ is a well ordering we may assume that m has minimal leading term with respect to τ among all elements of F whose residue classes are not contained in A. If $\mathrm{LT}_\tau(m) \in \mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$ then $m - \mathrm{LM}_\tau(m)$ also has residue class not in A, but has a smaller leading term, a contradiction. If $\mathrm{LT}_\tau(m) \in \mathrm{LT}_\tau\{U\}$ then there exists $u \in U$ with $\mathrm{LT}_\tau(u) = \mathrm{LT}_\tau(m)$. But then $m - \frac{\mathrm{LC}_\tau(m)}{\mathrm{LC}_\tau(u)} u$ has residue class not in A but has a smaller leading term than m, giving again a contradiction. Therefore the residue classes of $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$ generate F_ϱ/U.

Now suppose that $\sum_{i=1}^{k} c_i \bar{b}_i = \bar{0}$ with $k \geq 1$, $c_i \in K \setminus \{0\}$, and terms $b_i \in \mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$. This means that $\sum_{i=1}^{k} c_i b_i \in U$. But then $\mathrm{LT}_\tau(\sum_{i=1}^{k} c_i b_i) \in \mathrm{LT}_\tau\{U\} \cap (\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\})$, a contradiction. □

A Gröbner basis of U allows us to compute for each residue class in F_ϱ/U a representative that is a K-linear combination of elements of $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$. Gröbner bases can be characterized via the corresponding *reduction systems*.

Definition 2.3 *Let $v, w \in F_\varrho$. If there exist a term $e_i m_1 \in \mathrm{Supp}(w)$ and $m_2 \in \Sigma^*$ such that $\mathrm{LT}_\tau(v) m_2 = e_i m_1$ then we say that v* **reduces** *w to $w' = w - c\,\mathrm{LC}_\tau(v)^{-1} v m_2$* **in one step**, *denoted by $w \xrightarrow{v} w'$. Here $c \in K$ is the coefficient of $e_i m_1$ in w. If $G \subseteq F_\varrho$ is a set of vectors we let $\xrightarrow{G}$ denote the reflexive and transitive closure of $\bigcup_{g \in G} \xrightarrow{g}$. This means that we write $v \xrightarrow{G} w$ if there exists a sequence $v \xrightarrow{g_1} v_1 \xrightarrow{g_2} \cdots \xrightarrow{g_k} w$ of reduction steps where $k \geq 0, g_i \in G$. By $\xleftrightarrow{G}$ we denote the reflexive, symmetric, and transitive closure of $\bigcup_{g \in G} \xrightarrow{g}$.*

A reduction system $\xrightarrow{G}$ is called **Noetherian** *if there are no infinite rewriting sequences. It is called* **confluent** *if for all v, w_1, w_2 such that $v \xrightarrow{G} w_1$ and $v \xrightarrow{G} w_2$ there exists w_3 such that $w_1 \xrightarrow{G} w_3$ and $w_2 \xrightarrow{G} w_3$. It is called* **locally confluent** *if for all $g_1, g_2 \in G$ and all v, w_1, w_2 such that $v \xrightarrow{g_1} w_1$ and $v \xrightarrow{g_2} w_2$ there exists w_3 such that $w_1 \xrightarrow{G} w_3$ and $w_2 \xrightarrow{G} w_3$. A reduction system that is Noetherian and confluent is called* **convergent**.

If $\xrightarrow{G}$ is a Noetherian reduction system then $\xrightarrow{G}$ is confluent if it is locally confluent. In general this implication does not hold.

Proposition 2.4 *Let τ be a module term ordering on $\mathbb{T}(F_\varrho)$, let $U \subseteq F_\varrho$ be a right submodule, and let $G = \{g_1, \ldots, g_s\} \subseteq U \setminus \{0\}$ be a generating set for U. Furthermore we assume that the field K has effective arithmetic.*

1. *The term rewriting system $\xrightarrow{G}$ is convergent if and only if G is a τ-Gröbner basis of U.*
2. *Let G be a τ-Gröbner basis of U, and let $v \in F_\varrho$. There exists a unique element $\mathrm{NF}_U(v) \in F_\varrho$ with $v \xrightarrow{G} \mathrm{NF}_U(v)$ and such that $\mathrm{NF}_U(v)$ is irreducible with respect to $\xrightarrow{G}$. This element is effectively computable. It is called the* **normal form** *of v w.r.t. U. It does not depend on the choice of the Gröbner basis G.*
3. *For every $v \in F_\varrho$, the normal form $\mathrm{NF}_U(v)$ is a K-linear combination of the elements in $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$.*

Proof ¿From the fact that τ is a module term ordering it immediately follows that the term rewriting system $\xrightarrow{G}$ is Noetherian.

Suppose that G is a τ-Gröbner basis for U. Let $m \in F_\varrho$ and suppose that $m', m'' \in F_\varrho$ are irreducible w.r.t $\xrightarrow{G}$ and $m \xrightarrow{G} m'$, $m \xrightarrow{G} m''$. Then

$m'-m''$ is irreducible as well and $m'-m'' \in U$. But since G is a τ-Gröbner basis we must have $m'-m''=0$ showing that $\xrightarrow{G}$ is confluent.

Now suppose that $\xrightarrow{G}$ is convergent. For all $u \in U$, we have $u \xleftrightarrow{G} 0$. But if two elements are equivalent with respect to some convergent rewriting system they have the same irreducible normal form, so $u \xrightarrow{G} 0$, and hence $\mathrm{LT}_\tau(u) \in \{\mathrm{LT}_\tau(v_i)w \mid i=1,\ldots,s,\ w \in \Sigma^*\}$. Altogether, we have shown the first claim.

Macaulay's Basis Theorem yields the equality of vector spaces $F_\varrho = U \oplus \langle \mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}\rangle_K$. An element $v = u + w \in F$ with $u \in U$ and $w \in \langle \mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}\rangle_K$ is reduced by $\xrightarrow{G}$ to w in finitely many steps, yielding the second and third claim. □

Definition 2.5 *Let $v, v' \in F_\varrho$. If there exists a term $r \in \Sigma^*$ such that $\mathrm{LT}_\tau(v)r = \mathrm{LT}_\tau(v')$ or $\mathrm{LT}_\tau(v) = \mathrm{LT}_\tau(v')r$ then the vector $S(v,v') = \mathrm{LC}_\tau(v)^{-1}vr - \mathrm{LC}_\tau(v')^{-1}v'$ or $S(v,v') = \mathrm{LC}_\tau(v)^{-1}v - \mathrm{LC}_\tau(v')^{-1}v'r$, respectively, is called the* **S-vector** *of v and v'.*

In this setting we have the following criterion for Gröbner bases.

Proposition 2.6 (The Buchberger Criterion)
A set of generators G of a right submodule $U \subseteq F_\varrho$ is a τ-Gröbner basis of U if and only if we have $S(g_1,g_2) \xrightarrow{G} 0$ for every S-vector $S(g_1,g_2)$ of elements $g_1, g_2 \in G$.

Proof If G is a τ-Gröbner basis then by Proposition 2.4 $\xrightarrow{G}$ is convergent. Since every S-vector $S(g_1,g_2)$ of elements g_1, g_2 is in U or, equivalently, $S(g_1,g_2) \xleftrightarrow{G} 0$ this implies that $S(g_1,g_2) \xrightarrow{G} 0$.

Now suppose that $S(g_1,g_2) \xrightarrow{G} 0$ for every S-vector of elements $g_1, g_2 \in G$. We want to show that $\xrightarrow{G}$ is confluent. Since $\xrightarrow{G}$ is Noetherian it suffices to show that $\xrightarrow{G}$ is locally confluent, i.e. if a vector $m \in F_\varrho$ can be reduced to both $m_1 = m - c_j g_j w_j$ and $m_2 = m - c_k g_k w_k$ in one step where $c_j, c_k \in K \setminus \{0\}$, $g_j, g_k \in G$ and $w_j, w_k \in \Sigma^*$ then there exists $m_3 \in F_\varrho$ such that $m_1 \xrightarrow{G} m_3$ and $m_2 \xrightarrow{G} m_3$. First suppose that in this situation $\mathrm{LT}_\tau(g_j)w_j \neq \mathrm{LT}_\tau(g_k)w_k$. Without loss of generality let $\mathrm{LT}_\tau(g_j)w_j >_\tau \mathrm{LT}_\tau(g_k)w_k$. If $\mathrm{LT}_\tau(g_k)w_k$ is not a term in $g_j w_j$ then $m_1 \xrightarrow{g_k} m - c_j g_j w_j - c_k g_k w_k$ and $m_2 \xrightarrow{g_j} m - c_j g_j w_j - c_k g_k w_k$. If $\mathrm{LT}_\tau(g_k)w_k$ is a term in $g_j w_j$ with coefficient $c \in K \setminus \{0\}$ then $m_1 \xrightarrow{g_k} m - c_j g_j w_j - (c_k - \frac{c_j c}{\mathrm{LC}_\tau(g_k)})g_k w_k$ and $m_2 \xrightarrow{g_j} m - c_j g_j w_j - c_k g_k w_k \xrightarrow{g_k} m - c_j g_j w_j - (c_k - \frac{c_j c}{\mathrm{LC}_\tau(g_k)})g_k w_k$. So, if $\mathrm{LT}_\tau(g_j)w_j \neq \mathrm{LT}_\tau(g_k)w_k$ then m_3 exists independently on any conditions on $\xrightarrow{G}$.

Now suppose $\mathrm{LT}_\tau(g_j)w_j = \mathrm{LT}_\tau(g_k)w_k$. Then $m_2 - m_1 = c_j g_j w_j - c_k g_k w_k$ is a multiple of $S(g_j,g_k)$ so by our assumption we have $m_2 - m_1 \xrightarrow{G} 0$. Let $m_2 - m_1 \xrightarrow{g_{i_1}} f_1 = (m_2 - m_1) - c_{i_1} g_{i_1} w_{i_1} \xrightarrow{g_{i_2}} \cdots \xrightarrow{g_{i_l}} 0$ with

$g_{i_j} \in G$ be a sequence of reduction steps. Then $\mathrm{LT}_\tau(g_{i_1})w_{i_1} \in \mathrm{Supp}(m_1)$ or $\mathrm{LT}_\tau(g_{i_1})w_{i_1} \in \mathrm{Supp}(m_2)$. Let c'_1 and c'_2 be the coefficient of $\mathrm{LT}_\tau(g_{i_1}w_{i_1})$ in m_1 and m_2 respectively where at most one of c'_1 and c'_2 is zero. With a slight abuse of notation in the case that one of the coefficients is zero we have $c_{i_1} = c'_2 - c'_1$, $m_1 \xrightarrow{g_{i_1}} h'_1 = m_1 - c'_1 g_{i_1} w_{i_1}$ and $m_2 \xrightarrow{g_{i_1}} h'_2 = m_2 - c'_2 g_{i_1} w_{i_1}$, so $f_1 = h'_2 - h'_1$. By induction on l there exist $h_1, h_2 \in F_\varrho$ such that $m_1 \xrightarrow{G} h_1$, $m_2 \xrightarrow{G} h_2$ and $h_2 - h_1 = 0$ yielding the confluence of $\xrightarrow{G}$. □

Easy examples show that not every right submodule of F_ϱ has a finite Gröbner basis, even if both F_ϱ and the submodule are finitely generated. However, the following generalization of Buchberger's algorithm (and of the Knuth-Bendix completion procedure) provides us at least with a procedure for enumerating a Gröbner basis. We say that a selection strategy for elements from a set B is *fair* if no element stays in B forever, i.e. if it is not possible that the addition of new elements to B delays the selection forever.

Theorem 2.7 (The Buchberger Algorithm)
Let U be a right submodule of F_ϱ, and let $G = \{g_1, \ldots, g_s\}$ be a system of generators of U. Consider the following sequence of instructions:

1. *Let $s' = s$ and $B = \{(i,j) \mid 1 \le i < j \le s'$ and g_i, g_j have a nontrivial S-vector$\}$.*
2. If $B = \emptyset$, return the result G. Otherwise choose a pair $(i,j) \in B$ using a fair strategy and delete it from B.
3. Using $\xrightarrow{G}$, reduce $S(g_i, g_j)$ as much as possible, i.e. until we reach an element $S'(g_i, g_j)$ that is irreducible with respect to $\xrightarrow{G}$. If the result is zero, continue with step 2.
4. Increase s' by one. Append $g_{s'} = S'(g_i, g_j)$ to G, and append the set of pairs $\{(i, s') \mid g_i, g_{s'}$ have a nontrivial S-vector$\}$ to B. Then continue with step 2.

This is a procedure that enumerates a τ-Gröbner basis G of U. If U has a finite τ-Gröbner basis, it stops after finitely many steps and the resulting tuple G is a finite τ-Gröbner basis of U.

Proof For every pair g, g' in the output G we have $S(g, g') \xrightarrow{G} 0$ by construction of G if the S-vector of g, g' exists. By Proposition 2.6 the set G is a Gröbner basis.

If there exists a finite Gröbner basis $G' = \{g'_1, \ldots, g'_k\}$ then, since the procedure enumerates a Gröbner basis G, for each $j = 1, \ldots, k$ there is $g_{i_j} \in G$ such that $\mathrm{LT}_\tau(g_{i_j})$ divides $\mathrm{LT}_\tau(g'_j)$. But then $\mathrm{LT}_\tau\{U\} = \{\mathrm{LT}_\tau(g'_j)w \mid j = 1, \ldots, k, w \in \Sigma^*\} \subseteq \{\mathrm{LT}_\tau(g_{i_j})w \mid j = 1, \ldots, k, w \in \Sigma^*\} \subseteq \{\mathrm{LT}_\tau(g_i)w \mid i = 1, \ldots, \max\{i_1, \ldots, i_k\}, w \in \Sigma^*\} \subseteq \mathrm{LT}_\tau\{U\}$ shows that $\{g_1, \ldots, g_{\max\{i_1,\ldots,i_k\}}\}$ is a Gröbner basis of U. Therefore, after the procedure has appended $g_{\max\{i_1,\ldots,i_k\}}$ to G, we have $S(g_i, g_j) \xrightarrow{G} 0$ for all $(i,j) \in B$. Thus no element is appended to G anymore and the procedure halts after treating all $(i,j) \in B$. □

3 Gröbner Bases for Two-Sided Modules over Free Monoid Rings

In this section we give a brief account of a Gröbner basis theory for submodules of two-sided free modules over free monoid rings, since this theory will be applied in Section 5 to the conjugator search problem.

Let F be a two-sided free module over $K[\Sigma^*]$ with basis $\{e_i \mid i \in \Phi\}$ where Φ is finite or countably infinite. By this we mean the $K[\Sigma^*]$-bimodule consisting of elements of the form $\sum_{i\in\Phi}\sum_{j\in\mathbb{N}} f_{ij}e_i g_{ij}$ where all but finitely many of the elements $f_{ij}, g_{ij} \in K[\Sigma^*]$ are zero and where we identify cfe_ig and fe_icg for $c \in K$ and $f, g \in K[\Sigma^*]$. Multiplication is defined by $(h_1, h_2)fe_ig = h_1fe_igh_2$ for $(h_1, h_2) \in K[\Sigma^*] \times K[\Sigma^*]$. Furthermore, let $U \subseteq F$ be a two-sided submodule. In this setting, Definition 2.1 can be adjusted as follows.

Definition 3.1 *A* **term** *in F is an element of the form me_im' with $i \in \Phi$ and $m, m' \in \Sigma^*$. The set of all terms in F will be denoted by $\mathbb{T}(F)$.*

A **module term ordering** *on $\mathbb{T}(F)$ is a total ordering τ such that:*

1. *The inequality $t_1 \leq_\tau t_2$ implies $w_3t_1w_4 \leq_\tau w_3t_2w_4$ for all $t_1, t_2 \in \mathbb{T}(F)$ and $w_3, w_4 \in \Sigma^*$.*
2. *Every descending chain of terms $t_1 \geq_\tau t_2 \geq_\tau \cdots$ in $\mathbb{T}(F)$ is eventually stationary, i.e. τ is a well-ordering.*

Given a module term ordering τ and a non-zero element $v = \sum_{i=1}^{s} c_it_i \in F$ with $c_i \in K \setminus \{0\}$ and $t_i \in \mathbb{T}(F)$ satisfying $t_1 >_\tau \cdots >_\tau t_s$, we say that $\mathrm{LT}^*_\tau(v) = t_1$ is the **leading term** *of v and $\mathrm{LC}^*_\tau(v) = c_1$ its* **leading coefficient**. *Moreover, we let $\mathrm{LM}^*_\tau(v) = c_1t_1$.*

The two-sided submodule $\mathrm{LT}^*_\tau(U) = \langle \mathrm{LT}_\tau(v) \mid v \in U \setminus \{0\}\rangle$ of F is called the **(two-sided) leading term module** *of U, the set $\mathrm{LT}^*_\tau\{U\} = \{\mathrm{LT}^*_\tau(v) \mid v \in U \setminus \{0\}\}$ the* **leading term set** *of U.*

Finally, a set of vectors $\{v_i \mid i \in \Lambda\}$ in $U \setminus \{0\}$ is called a **(two-sided) τ-Gröbner basis** *of U if we have $\mathrm{LT}^*_\tau\{U\} = \{w_1\,\mathrm{LT}^*_\tau(v_i)w_2 \mid i \in \Lambda, w_1, w_2 \in \Sigma^*\}$.*

Since we are not going to use them, we leave it to the reader to write down the two-sided versions of Propositions 2.2 and 2.4. However, we want to explain the computation of two-sided Gröbner bases. The definition of S-vectors and critical pairs now reads as follows.

Definition 3.2 *Let $f, g \in F$. If there exist $\ell, r \in \Sigma^*$ such that we have $\mathrm{LT}^*_\tau(f) = \ell\,\mathrm{LT}^*_\tau(g)r$ or $\mathrm{LT}^*_\tau(f)r = \ell\,\mathrm{LT}^*_\tau(g)$ or $\ell\,\mathrm{LT}^*_\tau(f) = \mathrm{LT}^*_\tau(g)r$ or $\ell\,\mathrm{LT}^*_\tau(f)r = \mathrm{LT}^*_\tau(g)$ then the vector $S(f,g) = \mathrm{LC}^*_\tau(f)^{-1}f - \mathrm{LC}^*_\tau(g)^{-1}\ell gr$ or $S(f,g) = \mathrm{LC}^*_\tau(f)^{-1}fr - \mathrm{LC}^*_\tau(g)^{-1}\ell g$ or $S(f,g) = \mathrm{LC}^*_\tau(f)^{-1}\ell f - \mathrm{LC}^*_\tau(g)^{-1}gr$ or $S(f,g) = \mathrm{LC}^*_\tau(f)^{-1}\ell fr - \mathrm{LC}^*_\tau(g)^{-1}g$, respectively, is called the* **S-vector** *of f and g.*

Using this definition, both the Buchberger Criterion 2.6 and the Buchberger Algorithm 2.7 hold true without modifications and with the same proofs. Thus there is also a procedure for enumerating a two-sided Gröbner basis of a two-sided submodule of the two-sided free module F.

4 Gröbner Bases for Modules over Monoid Rings

In this section we consider a finitely presented monoid $M = \Sigma^*/\sim_W$ where $\sim_W$ is the equivalence relation on Σ^* generated by a finite number of relations $w_1 \sim w_1', \ldots, w_r \sim w_r'$. We shall use a multiplicative notation for M. Let K be a field. The *monoid ring* of M over K is the K-algebra

$$K[M] = \{\sum_{i=1}^{s} a_i t_i \mid a_1, \ldots, a_s \in K \setminus \{0\},\ t_1, \ldots, t_s \in M\}$$

with K-basis M and multiplication induced by extending the multiplication of M linearly.

Remark 4.1 For a finitely presented monoid M as above, the monoid ring of M has a presentation $K[M] \cong K[\Sigma^*]/I_M$ where $K[\Sigma^*]$ is the free monoid ring and I_M is the two-sided ideal $I_M = \langle w_1 - w_1', \ldots, w_r - w_r' \rangle$.

Assumptions. From now on we shall always assume that there exists a term ordering σ on Σ^* such that $w_i >_\sigma w_i'$ for $i = 1, \ldots, r$, i.e. for the relations defining M, implying that the string rewriting system $\xrightarrow{W}$ generated by $w_i \xrightarrow{W} w_{i'}$ for $i = 1, \ldots, r$ is Noetherian. Moreover, we require that this string rewriting system is convergent. Unless explicitly stated otherwise, the elements of M shall be presented by the corresponding irreducible words in Σ^* with respect to $\xrightarrow{W}$. For $a, b \in M$, let ab denote the product of a and b in M, $a \odot b$ denote the concatenation of the corresponding words in Σ^*, and $\equiv$ denote the identity of words. In particular it follows that ab is the normal form of $a \odot b$ w.r.t. $\xrightarrow{W}$.

Remark 4.2 The assumption that $\xrightarrow{W}$ is a convergent string rewriting system implies that the ideal I_M has a finite two-sided σ-Gröbner basis $G = \{g_1, \ldots, g_r\}$ where $g_i = w_i - w_i'$ for $i = 1, \ldots, r$, and every element $f \in K[\Sigma^*]$ can be effectively reduced via $\xrightarrow{G}$ to a unique normal form $\mathrm{NF}_{I_M}(f)$.

The remainder of this section is devoted to introducing and studying Gröbner bases for submodules of right free modules over $K[M]$. In the following we collect the necessary results generalizing the theory of prefix rewriting and prefix Gröbner bases for one-sided ideals of $K[M]$ developed in [20], [21], [22], and [27]. A Gröbner basis theory for the two-sided case can be introduced in a similar way. However, we do not need it here so we omit it.

Assumption. Let Φ be a finite or countably infinite set, and let $\overline{F}_\varrho$ be the free right $K[M]$-module with basis $\{\bar{e}_i \mid i \in \Phi\}$. The elements of $\overline{F}_\varrho$ are of the form $\sum_{i \in \Phi} \bar{e}_i f_i$ where only finitely many of the elements $f_i \in K[M]$ are non-zero. Furthermore, let $\overline{U} \subseteq \overline{F}_\varrho$ be a finitely generated right submodule, and let τ be a module term ordering on $\mathbb{T}(F_\varrho)$ that is compatible with the

term ordering σ on Σ^*, i.e. for all $i \in \Phi$ and $w_1, w_2 \in \Sigma^*$ with $w_1 <_\sigma w_2$ we have $e_i w_1 <_\tau e_i w_2$. Using the above assumption, we can view τ as an ordering on the set of terms

$$\mathbb{T}(\overline{F}_\varrho) = \{\bar{e}_i\, m \mid i \in \Phi,\ m \in M\}$$

of $\overline{F}_\varrho$. Although τ is a module term ordering, we cannot expect an inequality $\bar{e}_i m_1 \leq_\tau \bar{e}_i m_2$ to imply $\bar{e}_i m_1 m_3 \leq_\tau \bar{e}_i m_2 m_3$ where $i \in \Phi$ and $m_1, m_2, m_3 \in M$ because the reductions via $\xrightarrow{W}$ may destroy the inequality. So, for $v \in \overline{F}_\varrho$ and $m \in M$, we may have $\mathrm{LT}_\tau(vm) \neq \mathrm{LT}_\tau(v)m$. To get a Noetherian rewriting system, the appropriate definition of reduction is the following.

Definition 4.3 *Let $v, w \in \overline{F}_\varrho \setminus \{0\}$. If there exist a term $\bar{e}_i m_1 \in \mathrm{Supp}(w)$ and $m_2 \in M$ such that $\mathrm{LT}_\tau(vm_2) = \bar{e}_i m_1$ then we say that v* **strongly reduces** *w to $w - c\,\mathrm{LC}_\sigma(vm_2)^{-1} vm_2$ in one step. Here $c \in K$ is the coefficient of $\bar{e}_i m_1$ in w.*

Unfortunately, in general one cannot decide whether an equation $sx = t$ is solvable for x in the monoid M. Therefore one cannot decide whether a vector $w \in \overline{F}_\varrho$ can be strongly reduced by another one. To make this decision feasible, we introduce a weaker kind of reduction.

Definition 4.4 *Let $v, w \in \overline{F}_\varrho \setminus \{0\}$. If there exist a term $\bar{e}_i\, m_1 \in \mathrm{Supp}(w)$ and $m_2 \in M$ such that $\mathrm{LT}_\tau(v) \odot m_2 \equiv \bar{e}_i m_1$, we say that v* **prefix reduces** *w to $w' = w - c\,\mathrm{LC}_\tau(v)^{-1} vm_2$ in one step and we write $w \xrightarrow{v}_\pi w'$. Here $c \in K$ is the coefficient of $\bar{e}_i m_1$ in w.*

In the situation of this definition we have $\mathrm{LT}_\tau(vm_2) = \mathrm{LT}_\tau(v) \odot m_2$. Therefore the term rewriting system generated by prefix reduction steps is Noetherian. By using prefix reduction instead of strong reduction we gain computability, but another problem arises. We have to pay the price that given a set of generators G of U and a vector $v \in U$ we do not necessarily have $v \xleftrightarrow{G}_\pi 0$. This additional property can be achieved by *prefix saturation.*

Definition 4.5 *A subset $S \subseteq \overline{F}_\varrho$ is called* **prefix saturated** *if we have $vm \xrightarrow{S}_\pi 0$ in one step for all $v \in S$ and all $m \in M$.*

Proposition 4.6 *Let $S \subseteq \overline{F}_\varrho$ be prefix saturated. Then we have $v \xleftrightarrow{S}_\pi 0$ for all $v \in \langle S \rangle_\varrho$.*

Proof Let $v \in \langle S \rangle_\varrho$. We write $v = \sum_{i=1}^k c_i w_i m_i$ with $c_i \in K$, $w_i \in S$, and $m_i \in M$. We proceed by induction on k. For $k = 0$, the claim holds obviously true. Since S is prefix saturated, there exists an element $\tilde{w} \in S$ such that $w_k m_k \xrightarrow{\tilde{w}}_\pi 0$. Hence we have $c_k w_k m_k = \tilde{c}\,\tilde{w}\tilde{m}$ for some element $\tilde{m} \in M$. Now we distinguish two cases. If $\mathrm{LT}_\tau(\tilde{w}) \odot \tilde{m}$ is not a term in the support of v then we have $\sum_{i=1}^{k-1} c_i w_i m_i \xrightarrow{\tilde{w}}_\pi v$. Otherwise, we have $v \xrightarrow{\tilde{w}}_\pi v - \tilde{c}\,\tilde{w}\tilde{m} \xleftarrow{\tilde{w}}_\pi \sum_{i=1}^{k-1} c_i w_i m_i$. The assertion now follows from the inductive hypothesis. $\square$

There are examples for vectors $v \in \overline{F}_\varrho$ for which there does not exist a finite prefix saturated set S such that $\langle v \rangle_\varrho = \langle S \rangle_\varrho$. However, the following procedure enumerates such a set.

Proposition 4.7 *Let $M = \Sigma^* / \sim_W$ be a finitely presented monoid, where the eqivalence relation $\sim_W$ corresponds to the convergent term rewriting system $\xrightarrow{W}$ generated by $w_i \longrightarrow w_i'$ for $i = 1, \ldots, r$. Given a vector $v \in \overline{F}_\varrho$, consider the following sequence of instructions.*

1. *Let $\mathrm{Sat}(v) = \{v\}$ and $A = \{v\}$.*
2. *Using a fair strategy, choose a vector $w \in A$ and delete it from A. Write $t = \mathrm{LT}_\tau(w) = \bar{e}_i m$ with $i \in \Phi$ and $m \in M$. Let $C(t) = \{m' \in \Sigma^* \mid m \equiv m'' \odot r$ and $r \odot m' \equiv w_j$ for some $j \in \{1, \ldots, r\}, r \neq \lambda\}$.*
3. *If $C(t) = \emptyset$, continue with step 5). Otherwise, choose $m' \in C(t)$ and delete it from $C(t)$.*
4. *Compute the normal form $\tilde{w}$ of wm' with respect to $\xrightarrow{W}$. If $\tilde{w} \neq 0$ and $\tilde{w}$ does not reduce to zero in one step using $\xrightarrow{\mathrm{Sat}(v)}$ then append $\tilde{w}$ to $\mathrm{Sat}(v)$ and to A and continue with step 3.*
5. *If $A = \emptyset$ then return $\mathrm{Sat}(v)$ and stop. Otherwise, continue with step 2.*

This is a procedure that enumerates a prefix saturated set $\mathrm{Sat}(v)$ that generates the right module $\langle v \rangle_\varrho$.

Proof For a contradiction, suppose that the resulting set $\mathrm{Sat}(v)$ is not prefix saturated. Then there exist $w \in \mathrm{Sat}(v)$ and $m \in M$ such that the term $\mathrm{LT}_\tau(w) \odot m \in F_\varrho$ is minimal among all elements for which wm does not reduce to zero via $\xrightarrow{\mathrm{Sat}(v)}$. Then $\mathrm{LT}_\tau(w) \odot m \in F_\varrho$ is necessarily reducible with respect to $\xrightarrow{W}$. Let $\mathrm{LT}_\tau(w) = e_i m'$ with $m' \in M$. Since $m'm$ is reducible with respect to $\xrightarrow{W}$, there exist decompositions $m' \equiv m_1' \odot m_2'$ and $m \equiv m_1 \odot m_2$ such that $m_1, m_2, m_1', m_2' \in M$ and $m_2' \neq \lambda$ and $m_2' \odot m_1$ is the left side of $\xrightarrow{W}$.

Consequently, we have $m_1 \in C(\mathrm{LT}_\tau(w))$. Let $\tilde{w} = \mathrm{NF}_W(w \odot m_1)$. In step 4 of the procedure there are two possibilities. Either $\tilde{w}$ reduces to zero via $\xrightarrow{\mathrm{Sat}(v)}$ or $\tilde{w}$ is appended to $\mathrm{Sat}(v)$ and A. In the former case there exists an element $v' \in \mathrm{Sat}(v)$ such that $\mathrm{LT}_\tau(wm_1) \equiv \mathrm{LT}_\tau(v') \odot m_3$ and $wm_1 = cv'm_3$ for some $c \in K$ and $m_3 \in M$. Then the relations $\mathrm{LT}_\tau(w) \odot m >_\tau \mathrm{LT}_\tau(wm_1) \odot m_2 \equiv \mathrm{LT}_\tau(v') \odot m_3 \odot m_2 \geq_\tau \mathrm{LT}_\tau(v') \odot (m_3 m_2)$ and the minimality of $\mathrm{LT}_\tau(w) \odot m$ imply that wm reduces to zero via $\xrightarrow{\mathrm{Sat}(v)}_\pi$, a contradiction. It remains to consider the case that $\tilde{w}$ does not reduce to zero via $\xrightarrow{\mathrm{Sat}(v)}$ and hence is appended to $\mathrm{Sat}(v)$. The relations $\mathrm{LT}_\tau(w) \odot m \equiv \mathrm{LT}_\tau(w) \odot m_1 \odot m_2 >_\tau \mathrm{LT}_\tau(wm_1) \odot m_2 \equiv \mathrm{LT}_\tau(\tilde{w}) \odot m_2$ and the minimality of $\mathrm{LT}_\tau(w) \odot m$ imply that $wm = \tilde{w}m_2$ reduces to zero via $\xrightarrow{\mathrm{Sat}(v)}_\pi$, a contradiction again. □

In analogy to Proposition 2.4 we now introduce Gröbner bases for prefix rewriting.

Definition 4.8 *Let $\overline{U}$ be a right submodule of $\overline{F}_\varrho$. A set $G \subseteq \overline{U}$ is called a* **prefix Gröbner basis** *of $\overline{U}$ if we have $u \xleftrightarrow{G}_\pi 0$ for all $u \in \overline{U}$ and if $\xrightarrow{G}$ is confluent.*

Remark 4.9 In analogy to Definition 2.1 the following holds: A set $G \subseteq \overline{U}$ is a Gröbner basis of $\overline{U}$ if and only if $\mathrm{LT}_\tau\{\overline{U}\} = \{\mathrm{LT}_\tau(g) \odot m \mid g \in G, m \in M\}$.

Macaulay's Basis Theorem also holds: If $\overline{U}$ is a right submodule of $\overline{F}_\varrho$ then the residue classes of the terms in $\mathbb{T}(\overline{F}_\varrho) \setminus \mathrm{LT}_\tau(\overline{U})$ form a K-basis of $\overline{F}_\varrho/\overline{U}$.

In order to characterize prefix Gröbner bases by a Buchberger criterion, we need to define the proper notion of S-vectors.

Definition 4.10 *Let $v, w \in \overline{F}_\varrho$. If there exists an element $m \in M$ such that $\mathrm{LT}_\tau(v) \odot m \equiv \mathrm{LT}_\tau(w)$ or $\mathrm{LT}_\tau(v) \equiv \mathrm{LT}_\tau(w) \odot m$ then the element $S(v,w) = \mathrm{LC}_\tau(v)^{-1} vm - \mathrm{LC}_\tau(w)^{-1} w$ or $S(v,w) = \mathrm{LC}_\tau(v)^{-1} v - \mathrm{LC}_\tau(w)^{-1} wm$, respectively, is called the* **S-vector** *of v and w.*

Proposition 4.11 (Buchberger Criterion for Prefix Gröbner Bases) *Let $G \subseteq \overline{F}_\varrho$ be a prefix saturated set. Then the set G is a prefix τ-Gröbner basis for $\langle G \rangle_\varrho$ if and only if we have $S(v,w) \xrightarrow{G}_\pi 0$ for all S-vectors of elements $v, w \in G$.*

Proof Let G be a prefix Gröbner basis, and let $S(v,w)$ be an S-vector of elements $v, w \in G$. Since $S(v,w) \in \langle G \rangle_\varrho$, by Proposition 4.6 we have $S(v,w) \xleftrightarrow{G}_\pi 0$. Now the fact that 0 is irreducible with respect to $\xrightarrow{G}_\pi$ easily implies that we actually have $S(v,w) \xrightarrow{G}_\pi 0$.

Conversely, assume that $S(v,w) \xrightarrow{G}_\pi 0$ for all S-vectors of $v, w \in G$. If we want to show that $\xrightarrow{G}_\pi$ is confluent it suffices to consider the critical situations as in Proposition 2.6. These critical situations correspond to S-vectors of elements of G and resolve if for all S-vectors $S(v,w)$ we have $S(v,w) \xrightarrow{G}_\pi 0$. The remaining claim follows from Proposition 4.6. □

Finally, we provide a procedure for computing prefix Gröbner bases.

Theorem 4.12 (Buchberger Algorithm for Prefix Gröbner Bases) *Let $\overline{U}$ be a right submodule of $\overline{F}_\varrho$, let $G = \{g_1, \ldots, g_s\}$ by a system of generators of $\overline{U}$, and let τ be a module term ordering on $\mathbb{T}(F)$. Consider the following sequence of instructions.*

1. *Let $H = G$, $A = G$, $S_h = \{h\}$ for all $h \in G$, $s' = s$ and $B = \{(i,j) \mid 1 \leq i < j \leq s',\ g_i, g_j$ have a non-trivial S-vector$\}$.*
2. *If A is empty, return H and stop. Otherwise start a computation of $\mathrm{Sat}(h)$ using Proposition 4.7 for every $h \in A$. If this procedure yields $\mathrm{Sat}(h) = S_h$, remove h from A. Otherwise, stop the procedure when it has computed a finite subset S'_h of $\mathrm{Sat}(h)$ consisting of S_h and at*

least one more element. Append the elements of $\bigcup_{h\in A}(S'_h \setminus S_h)$ *to* H*, increase* s' *by the number* s'' *of these new elements, and write* $H = \{h_1, \ldots, h_{s'}\}$*. Set* $S_h = S'_h$ *and append* $\{(i,j) \mid 1 \leq i < j \leq s',\ j > s' - s'',\ h_i, h_j$ have a non-trivial S-vector$\}$ *to the set* B*.*

3. *If* $B = \emptyset$*, continue with step 2. Otherwise, use a fair strategy to choose a pair* $(i,j) \in B$ *and delete it from* B*.*
4. *Using* $\xrightarrow{H}_\pi$*, reduce* $S(h_i, h_j)$ *as much as possible and call the result* $S'(h_i, h_j)$*. If* $S'(h_i, h_j) = 0$*, continue with step 2. Otherwise, increase* s' *by one, set* $h_{s'} = S'(h_i, h_j)$ *and append* $h_{s'}$ *to* H *and* A*. Set* $S_{h_{s'}} = \{h_{s'}\}$*, append* $\{(i,s') \mid 1 \leq i < s',\ h_i, h_{s'}$ have a non-trivial S-vector$\}$ *to the set* B *and continue with step 2.*

This is a procedure that enumerates a prefix τ*-Gröbner basis* H *of* $\overline{U}$*.*

Proof The set H enumerated by the procedure is prefix saturated as it is the union of prefix saturated sets $\mathrm{Sat}(h)$. By construction we have $S(h_i, h_j) \xrightarrow{H}_\pi 0$ for all $h_i, h_j \in H$ for that $S(h_i, h_j)$ exists. The assertion now follows from Propositions 4.6 and 4.11. □

5 Some Classical Problems

As above, we let $M = \Sigma^*/\sim_W$ be a finitely presented monoid, we let Φ be a finite or countably infinite set, and we let $\overline{F}_\varrho$ be the free right $K[M]$-module with basis $\{\bar{e}_i \mid i \in \Phi\}$. We continue to operate under the general assumptions introduced above. Moreover, we suppose that a right submodule $\overline{U} \subseteq \overline{F}_\varrho$ is given by a finite tuple of generators $\mathcal{U} = (\bar{u}_1, \ldots, \bar{u}_s)$ and that we know a prefix Gröbner basis of $\overline{U}$.

In this setting a number of classical problems for groups and monoids can be solved using Gröbner basis techniques. Therefore the original computational problem becomes the problem to find the correct Gröbner basis. Let us illustrate the method with some examples.

Proposition 5.1 (The Word Problem for Free Right Modules) *Given two vectors* $\bar{v}, \bar{w} \in \overline{F}_\varrho$*, we write* $\bar{v} - \bar{w} = \sum_{i\in\Phi} \bar{e}_i f_i$ *with* $f_i \in K[M]$*. Then the following conditions are equivalent:*

1. $\bar{v} = \bar{w}$
2. *For all* $i \in \Phi$ *we have* $\mathrm{NF}_{I_M}(f_i) = 0$*.*

Notice that if we write the terms in the support of f_i in their normal form with respect to $\xrightarrow{W}$, we only have to check whether these normal forms are zero.

In free right $K[M]$-modules, we can solve the submodule membership problem as follows.

Proposition 5.2 (The Submodule Membership Problem) *For a vector* $\bar{v} \in \overline{F}_\varrho$ *the following conditions are equivalent:*

1. $\bar{v} \in \overline{U}$
2. $\bar{v} + \overline{U} = 0$ *in the module* $\overline{F}/\overline{U}$
3. $\bar{v} \xrightarrow{G}_{\pi} 0$ *for some prefix Gröbner basis* G *of* $\overline{U}$.

Proof This follows immediately from Proposition 4.6.

The generalized word problem (also called the submonoid membership problem) was discussed in [20]. It was shown that it leads to a subalgebra membership problem in $K[M]$. For groups the situation is somewhat more accessible, since the subgroup membership problem is equivalent to a right ideal membership problem in $K[M]$.

The next interesting monoid and group theoretic problems are the conjugacy problem and the conjugator search problem. Let us indicate some methods for solving them using Gröbner bases.

Definition 5.3 *Let* $f_1, \ldots, f_s \in K[M]$. *The two-sided submodule of the two-sided free module* $F = \oplus_{i=1}^{s} K[M]e_iK[M]$ *generated by all elements* $\sum_{i=1}^{s} g_i e_i h_i$ *such that* $g_1 f_1 h_1 + \cdots + g_s f_s h_s = 0$ *is called the* **syzygy module** *of the tuple* $(f_1, \ldots, f_s)$. *We shall denote it by* $\mathrm{Syz}_{K[M]}(f_1, \ldots, f_s)$.

In the case $M = \Sigma^*$, there exist explicit descriptions of algorithms to compute syzygy modules (see [7] and [10]). For the general case, we can either lift the computation to $K[\Sigma^*]$ or construct a similar algorithm (see [3]). The computation of syzygy modules and the following easy proposition help us achieve our goal.

Proposition 5.4 *Let* M *be a group. For* $\bar{w}_1, \bar{w}_2 \in M$, *the following conditions are equivalent:*

1. $\bar{w}_1 = \bar{w}_3 \bar{w}_2 \bar{w}_3^{-1}$ *for some* $\bar{w}_3 \in M$
2. $\mathrm{Syz}_{K[M]}(w_1, w_2) \cap \{e_1\bar{w} - \bar{w}e_2 \mid \bar{w} \in M\} \neq \emptyset$

Proof It suffices to note that we have $\bar{w}_1 = \bar{w}_3 \bar{w}_2 \bar{w}_3^{-1}$ if and only if $\bar{w}_1 \bar{w}_3 = \bar{w}_3 \bar{w}_2$, i.e. if and only if $e_1\bar{w}_3 - \bar{w}_3 e_2$ is a syzygy of $(\bar{w}_1, \bar{w}_2)$. □

Hence the conjugacy and the conjugator search problems have been reduced to finding certain very simple elements in a syzygy module. The latter task can be achieved by a straightforward generalization of the method in the commutative case (see [16] and [17]).

6 Gröbner Basis Cryptosystems

In this section we will propose a public key cryptographic primitive based on the Gröbner basis theory in the setting described so far. Note that to actually use this Gröbner basis cryptosystem one has to find instances that guarantee efficient computations where needed and security under certain

assumptions. We show how one can realize well known public key cryptosystems as special cases of this cryptographic framework. Besides these examples we do not give concrete instances, however we will discuss how one can meet various attacks on the system in the next section.

In the following we continue to work in the described setting and use the assumptions of the earlier sections. In particular, we let $M = \Sigma^*/\sim_W$ be a finitely presented monoid and assume that elements of M can be represented by normal forms that can be efficiently computed, e.g. by a convergent term rewriting system $\xrightarrow{W}$. Let σ be a $\xrightarrow{W}$-admissible term ordering on Σ^*, F_ϱ be the free right $K[\Sigma^*]$-module with basis $\{e_i \mid i \in \Phi\}$ and let τ be a module term ordering on $\mathbb{T}(F_\varrho)$ that is compatible with σ. Let $\overline{F}_\varrho$ be the free right $K[M]$-module with basis $\{\bar{e}_i \mid i \in \Phi\}$ and let $\overline{U}$ be a right submodule of $\overline{F}_\varrho$. Finally, we assume that we know a τ-Gröbner basis G of $\overline{U}$ w.r.t. some reduction system that allows efficient computations of normal forms, and therefore the set $O_\tau(\overline{U}) = \mathbb{T}(\overline{F}_\varrho) \setminus \mathrm{LT}_\tau\{\overline{U}\}$.

Definition 6.1 *A* **Gröbner basis cryptosystem** *consists of the following data.*

1. Public information: *the free module* $\overline{F}_\varrho$, *the set* $O_\tau(\overline{U})$, *and finitely many vectors* $\bar{u}_1, \ldots, \bar{u}_s \in \overline{U}$
2. Secret key: *a prefix Gröbner basis* G *of* $\overline{U}$.
3. Encryption procedure: *A plaintext is a vector* $m \in \langle O_\tau(\overline{U})\rangle_K$, *i.e. a linear combination* $m = \bar{e}_{\lambda_1} c_1 w_1 + \cdots + \bar{e}_{\lambda_r} c_r w_r$ *such that* $c_i \in K$, $\lambda_i \in \Phi$, $w_i \in M$, *and* $\bar{e}_i w_i \in O_\tau(\overline{U})$. *Then the corresponding ciphertext is the vector* $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s$ *with suitably (e.g. randomly) chosen* $f_1, \ldots, f_s \in K[M]$.
4. Decryption procedure: *Using* $\xrightarrow{G}$, *compute* $m = \mathrm{NF}_{\tau,\bar{U}}(w)$.

Note that the right choice of $f_1, \ldots, f_s \in K[M]$ in the encryption procedure can be crucial for the security of the cryptosystem and will depend on the concrete setting in which the primitive is used. For some settings this issue will be discussed in the next section.

Remark 6.2 We shall also consider the following *variant*: for the ciphertext we construct a pair $w = (f_0, mf_0 + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s)$ where $f_0 \in K[M]$ is a further randomly chosen element. Then the decoding procedure consists of computing $\mathrm{NF}_{\tau,\bar{U}}(mf_0 + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s) = \mathrm{NF}_{\tau,\bar{U}}(mf_0)$ and "dividing" by f_0 to obtain m. In this way we achieve some additional data hiding: the summand mf_0 on the right hand-side has the same shape as the other summands. However there is no general method for performing the "division" $\mathrm{NF}_{\tau,\bar{U}}(mf_0) \mapsto m$. We have to provide an explicit procedure in every individual example.

Let us collect some easy remarks about the merits of such a cryptosystem.

Remark 6.3 Let a Gröbner basis cryptosystem be given as above.

1. If an attacker can compute G, he can break the cryptosystem. In general, the computation of Gröbner bases is EXPSPACE-hard.
2. The attacker knows $\bar{u}_1, \ldots, \bar{u}_s$ and $O_\tau(\overline{U})$, but not a system of generators of $\overline{U}$. We can make his task difficult by choosing $\bar{u}_1, \ldots, \bar{u}_s$ such that a Gröbner basis of $\langle \bar{u}_1, \ldots, \bar{u}_s \rangle_\varrho$ is hard to compute.
3. The advantage of using modules (rather than ideals in $K[M]$) is that one can encode hard combinatorial or number theoretic problems in the action of the terms on the canoncial basis vectors (see the examples below).
4. The free module F_ϱ is not required to be finitely generated. Any concrete calculation will involve only finitely many components.

Now we give some examples of Gröbner basis cryptosystems. In particular, we show that many classical cryptosystems can be realized as Gröbner basis cryptosystems.

Example 6.4 Let $K = \mathbb{F}_q$ be a finite field, where $q = p^e$ with a prime number p and $e > 0$. Let M be the monoid $M = \mathbb{N}^n = \Sigma^* / \sim_W$ where $\Sigma = \{x_1, \ldots, x_n\}$ and we require the relations $W = \{x_j x_i \sim x_i x_j \mid 1 \leq i < j \leq n\}$. We use the free right module of rank one, i.e. $\overline{F}_\varrho = K[M] = K[x_1, \ldots, x_n]$ is the commutative polynomial ring. Choose a point $(a_1, \ldots, a_n) \in \mathbb{F}_p^n$. Let $\overline{U} = (x_1 - a_1, \ldots, x_n - a_n)$ and choose elements $\bar{u}_1, \ldots, \bar{u}_s \in \overline{U}$, i.e. $\bar{u}_i(a_1, \ldots, a_n) = 0$. Consider the following Gröbner basis cryptosystem.

1. *Public information:* The one-dimensional free right module $\overline{F}_\varrho$, the set $O_\tau(\overline{U}) = \{1\}$, and the commutative polynomials $\bar{u}_1, \ldots, \bar{u}_s$.
2. *Secret key:* The point $(a_1, \ldots, a_n) \in \mathbb{F}_q^n$ corresponding to the Gröbner basis $G = \{x_1 - a_1, \ldots, x_n - a_n\}$ of the ideal $\overline{U}$.
3. *Encryption procedure:* A plaintext $m \in \mathbb{F}_q$ is encrypted as the polynomial $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s$ with randomly chosen polynomials $f_1, \ldots, f_s \in K[M]$.
4. *Decryption procedure:* Compute $m = w(a_1, \ldots, a_n) = \mathrm{NF}_{\tau, \bar{U}}(w)$.

This is Neal Koblitz' **polly cracker** cryptosystem (cf. [8] and [14]). Its disadvantage is that the attacker knows that there is an element in the set $w + \bar{u}_1 \cdot K[M] + \cdots + \bar{u}_s \cdot K[M]$ that has support $\{1\}$. Hence many coefficients have to vanish. This allows a linear algebra attack (see [2], [8] and [14]).

A number of improvements of Koblitz' original approach have been proposed (see for instance [18] and [29]). Many of them fit our scheme.

Example 6.5 In the setting of the preceding example, choose a second commutative polynomial ring $Q = K[y_1, \ldots, y_m]$ and polynomials $g_1, \ldots, g_m$ in $K[M]$. In this way there is a K-algebra homomorphism $\phi : Q \longrightarrow K[M]$ given by $\phi(y_i) = g_i$ for $i = 1, \ldots, m$. Choose a point $(\xi_1, \ldots, \xi_n) \in \mathbb{F}_p^n$ and elements $f_1, \ldots, f_s \in Q$ such that $\phi(f_1), \ldots, \phi(f_s) \in \overline{U} = (x_1 - \xi_1, \ldots, x_n - \xi_n)$. Now construct the following Gröbner basis cryptosystem.

1. *Public information:* The rings $K[M]$ and Q, the homomorphism ϕ, the term $O_\tau(\overline{U}) = \{1\}$, and the polynomials $f_1, \ldots, f_s \in Q$.
2. *Secret key:* The point $(\xi_1, \ldots, \xi_n) \in K^n$, or equivalently, the Gröbner basis $\{x_1 - \xi_1, \ldots, x_n - \xi_n\}$ of the ideal $\bar{U} = (x_1 - \xi_1, \ldots, x_n - \xi_n)$ in $K[M]$.
3. *Encryption procedure:* We proceed in a similar way to the variant above. A plaintext is an element $m \in K$. We choose random polynomials $h \in (f_1, \ldots, f_s)$ and $h' \in \ker(\phi)$ and a random exponent $\kappa \in \mathbb{N}^n$. Then we send $(y^\kappa, my^\kappa + h + h')$ where $y = (y_1, \ldots, y_m)$. In other words, an attacker knows the pair $(\phi(y)^\kappa, m\phi(y)^\kappa + \phi(h))$.
4. *Decryption procedure:* Compute $\bar{v} = [m\phi(y)^\kappa + \phi(h)](\xi_1, \ldots, \xi_n) = m\phi(y)^\kappa(\xi_1, \ldots, \xi_n)$ and obtain $m = \bar{v}/[\phi(y)^\kappa(\xi_1, \ldots, \xi_n)]$.

This is Le Van Ly's **Polly Two** cryptosystem (cf. [29]). Compared to Polly Cracker, it has the advantage that the usual linear algebra attacks do not work. It appears that an attacker has no choice but to compute a (possibly hard) Gröbner basis. Supposedly hard concrete instances of this cryptosystem have been suggested (see [30]).

Example 6.6 Let $K = \mathbb{F}_2$, let $\Sigma = \{x, y\}$, and let $M = \mathbb{N}^2 = \Sigma^*/\sim_W$ with $W = \{yx \sim xy\}$. Then $K[M] = K[\Sigma^*]/\langle yx - xy\rangle = K[x, y]$ is a commutative polynomial ring in two indeterminates. Moreover, let $p, q \gg 0$ be two distinct prime numbers, let $n = pq$, and let $\Pi = (\mathbb{Z}/n\mathbb{Z})^\times$ be the set of residue classes prime to n. We use the free module $\overline{F}_\varrho = \bigoplus_{i=0}^{n-1} e_i\, K[x, y]$ and the term ordering $\tau = \texttt{DegLexPos}$. Choose a number $\epsilon \in (\mathbb{Z}/(p-1)(q-1)\mathbb{Z})^\times$ and compute the inverse d of ϵ in $(\mathbb{Z}/(p-1)(q-1)\mathbb{Z})^\times$.

1. *Public information:* The module $\overline{F}_\varrho$ (and thus the number n), the set $O_\tau(\overline{U}) = \{e_0, \ldots, e_{n-1}\}$, the number ϵ and the vectors $\{u_1, \ldots, u_s\} = \{e_i x - e_{i^\epsilon \bmod n} \mid i = 0, \ldots, n-1\} \cup \{e_i xy - e_i \mid i = 0, \ldots, n-1\}$.
2. *Secret key:* The secret key consists of the primes p and q and the number d. Equivalently, the secret key is the τ-Gröbner basis $G = \{u_1, \ldots, u_s\} \cup \{e_i y - e_{i^d \bmod n} \mid i = 0, \ldots, n-1\}$ of $\overline{U} = \langle G\rangle$.
3. *Encryption procedure:* A plaintext is a vector $e_m \in O_\tau(\overline{U})$. To encrypt it, we form $e_m + (e_m xy - e_m) - (e_m x - e_{m^\epsilon \bmod n})y \in e_m + \overline{U}$ to obtain the ciphertext $w = e_{m^\epsilon \bmod n} y$.
4. *Decryption procedure:* Compute $\mathrm{NF}_{\tau,\bar{U}}(e_{m^\epsilon \bmod n} y) = e_{m^{\epsilon d} \bmod n} = e_m$.

It is easy to see that this is the Gröbner basis version of the **RSA** cryptographic primitive (see [28]) which is used to derive concrete instances of practical cryptosystems that are widely used in practice. If an attacker is able to factor n, he can break the code. This is equivalent to being able to find d. In the Gröbner basis version, the problem the attacker faces is that he does not know the Gröbner basis elements $e_i y - e_{i^d \bmod n}$.

Also discrete logarithms can be used in Gröbner basis cryptosystems.

Example 6.7 Let $K = \mathbb{F}_2$, let $\Sigma = \{x\}$, and let $M = \Sigma^* = \mathbb{N}$. Then $K[M] = K[x]$ is a polynomial ring. Moreover, let $p \gg 0$ be a prime number.

We use the $K[x]$-module $F_\varrho = \bigoplus_{i=1}^{p-1} \epsilon_i K[x] \oplus \bigoplus_{j=1}^{p-1} e_j K[x]$ where ϵ_i, e_j are the standard basis vectors. Let g be a generator of the multiplicative group $\mathbb{F}_p^\times$, and let $\tau = \texttt{DegPos}$ with $\epsilon_i >_\tau e_j$ for all $i, j = 1, \ldots, p-1$. Choose a number $a \in \{1, \ldots, p-1\}$ and compute $b = g^a \bmod p$. Now we introduce the following Gröbner basis cryptosystem.

1. *Public information:* The module F_ϱ, the set $O_\tau(U) = \{e_1, e_2, \ldots, e_{p-1}\}$, the number b, and the vectors $\{u_1, \ldots, u_s\} = \{\epsilon_1 - e_1\} \cup \{\epsilon_i x - \epsilon_{gi} \mid i = 1, \ldots, p-1\} \cup \{e_j x - e_{bj} \mid j = 1, \ldots, p-1\}$ where all indices are computed modulo p.
2. *Secret key:* The number $a \in \{1, \ldots, p-1\}$, or equivalently, the τ-Gröbner basis $G = \{u_1, \ldots, u_s\} \cup \{\epsilon_i - e_{i^a} \mid i = 1, \ldots, p-1\}$ of $\overline{U} = \langle G \rangle$.
3. *Encryption procedure:* A plaintext is of the form $e_1 + e_m$ with a number $m \in \{0, \ldots, p-1\}$. Using the variant, we randomly choose a number $k \in \{0, \ldots, p-1\}$, form $(e_1 + e_m)x^k$ and send the ciphertext $w = \epsilon_{g^k} + e_{mb^k} \in (e_1 + e_m)x^k + \langle u_1, \ldots, u_s \rangle_\varrho$.
4. *Decryption procedure:* We compute $\mathrm{NF}_{\tau,\bar{U}}(w) = e_{b^k} + e_{mb^k}$. Since $e_{b^k} + e_{mb^k} \overset{G}{\longleftrightarrow} (e_1 + e_m)x^k$, we have to "divide" this vector by x^k. To this end, it suffices to compute $m = (mb^k)/(b^k)$ in $\mathbb{F}_p$ and to form $e_1 + e_m$.

Clearly, this is the Gröbner basis version of the **ElGamal** cryptosystem (see [5]). It can be broken if the attacker is able to compute the discrete logarithm a of $b = g^a$ or k of g^k. In the Gröbner basis version, an attacker can only reduce using $\epsilon_{g^k} \xrightarrow{u_i} \cdots \xrightarrow{u_j} x^k \epsilon_1 \xrightarrow{u_1} x^k e_1$ which takes $k \gg 0$ reduction steps. If one knows a, one can get rid of the vector ϵ_{g^k} by using just one reduction step $\epsilon_{g^k} \longrightarrow e_{g^{ka}} = e_{b^k}$.

The next example uses non-commutative polynomials. In order to prevent linear algebra attacks (see next section) T. Rai suggested in his recent doctoral thesis [26] to construct Gröbner basis cryptosystems based on two-sided ideals. The corresponding Gröbner basis theory was sketched in [20], [21], [22], [27] and Section 3.

Example 6.8 Let K be a (finite) field, let $\Sigma = \{x_1, \ldots, x_n\}$, and let $M = \Sigma^*$. Then $K[M]$ is a non-commutative polynomial ring. We choose a two-sided ideal $I \subseteq K[M]$ for which we know a finite (two-sided) Gröbner basis $G = \{g_1, \ldots, g_t\}$ with respect to some term ordering τ.

1. *Public information:* The ring $K[M]$, the set $O_\tau(U)$, and a finite subset $\{u_1, \ldots, u_s\} \in I$ such that computing a Gröbner basis of $\langle u_1, \ldots, u_s \rangle$ is infeasible.
2. *Secret key:* The τ-Gröbner basis G of I.
3. *Encryption procedure:* A plaintext m is an element in $\langle O_\tau(U) \rangle_K$. The corresponding ciphertext is $w = m + f_1 u_1 g_1 + \cdots + f_s u_s g_s$ where the non-commutative polynomials f_i, g_i are suitably chosen so that in the computation of w leading term cancellation occurs (see [26], Section 4.1).

4. *Decryption procedure:* Compute $m = \mathrm{NF}_{\tau,\bar{U}}(w)$ using the Gröbner basis G.

In [26] several concrete instances of these cryptosystems are proposed. They offer good resistence to linear algebra attacks because using indeterminate coefficients for the polynomials f_i and g_j leads to systems of quadratic equations in these coefficients which cannot be solved using linear algebra. However, one has to take great care to make these cryptosystems secure against attackers who are able to compute partial Gröbner bases (see [26], Chapter 4).

Our approach is flexible enough to include recent attempts at group based cryptosystems. For instance, the following Gröbner basis cryptosystem relies on the difficulty of solving the conjugator search problem in certain groups.

Example 6.9 Let K be a field, and let $M = \Sigma^*/\sim_W$ be a finitely presented group. We use the free right $K[M]$-module $\overline{F} = \bigoplus_{\bar{w}\in M} \epsilon_{\bar{w}} K[M] \oplus \bigoplus_{\bar{w}\in M} e_{\bar{w}} K[M]$ (possibly of infinite rank). Moreover, let $\tau = \texttt{llexPos}$ be such that $\epsilon_{\bar{w}} >_\tau e_{\bar{u}}$ for all $w, u \in M$. Choose $a, g \in M$ and compute $g' = a^{-1}ga$. Now consider the following Gröbner basis cryptosystem.

1. *Public information:* The module $\overline{F}_\varrho$, the elements $g, g' \in M$, a set $B \subseteq \{c \in M \mid ca = ac\}$, the set $O_\tau(U) = \{e_{\bar{w}} \mid \bar{w} \in M\}$, and the vectors $\{u_\lambda \mid \lambda \in \Lambda\} = \{\epsilon_i h - \epsilon_{h^{-1}ih} \mid i, h \in M\} \cup \{\epsilon_g - e_{g'}\} \cup \{e_j k - e_{k^{-1}jk} \mid j, k \in M\}$.
2. *Secret key:* The element $a \in M$, or equivalently, the τ-Gröbner basis $G = \{u_\lambda \mid \lambda \in \Lambda\} \cup \{\epsilon_i - e_{a^{-1}ia} \mid i \in M\}$ of the submodule $\overline{U} = \langle G\rangle_\varrho$ of $\overline{F}_\varrho$.
3. *Encryption procedure:* Randomly choose an element $b \in B$. A plaintext $m \in M$ is written in the form $\epsilon_g + e_{g'\tilde{m}}$, where $\tilde{m} = bmb^{-1}$. Then we multiply by b and use the elements u_λ to obtain the ciphertext $w = \epsilon_{b^{-1}gb} + e_{b^{-1}g'\tilde{m}b}$.
4. *Decryption procedure:* Find $\mathrm{NF}_{\tau,\bar{U}}(w) = e_{a^{-1}g''a} + e_{b^{-1}g'bm} = e_{b^{-1}g'b} + e_{b^{-1}g'b\,m}$ first, where $g'' = b^{-1}gb$. Then determine m from $m = (b^{-1}g'b)^{-1}(b^{-1}g'bm)$.

As one can readily check, this is a Gröbner basis version of an ElGamal like cryptosystem based on a group with a "hard" Diffie-Hellman conjugacy problem, i.e. the problem to find $a^{-1}b^{-1}gba$ given g, $a^{-1}ga$ and $b^{-1}gb$ where a and b commute. One can solve this problem if given g and $g' = a^{-1}ga$ one can find a_1, a_2 such that $a_1 g a_2 = g'$ and a_1, a_2 commute with the elements from B. The advantage of knowing the Gröbner basis is that one can pass from $\epsilon_{g''}$ to the corresponding e_i without going through $\epsilon_g \longrightarrow e_{g'}$. The computation of that Gröbner basis is equivalent to finding a.

To perform the encryption step explicitly one has to perform the following simple computations in the group: Conjugate g' with b to obtain $b^{-1}g'b$ and multiply by the plaintext m. Conjugate g with b to obtain $b^{-1}gb$.

If we want to decrypt the ciphertext $\epsilon_{b^{-1}gb}+e_{b^{-1}g'\tilde{m}b}$ knowing the secret a an explicit decryption amounts to performing the following: Conjugate $b^{-1}gb$ with a to obtain the Gröbner basis element $\epsilon_{b^{-1}gb}-e_{a^{-1}b^{-1}gba}$, reduce w via this element in one step to $\mathrm{NF}_{\tau,\bar{U}}(w) = e_{a^{-1}b^{-1}gba}+e_{b^{-1}a^{-1}gabm}$ and obtain m by multiplying the inverse of the first index by the second index.

So all computations performed to encrypt and decrypt are actually computations in the group M.

In [15] braid groups have been suggested for this kind of cryptosystems. However in [4] it is shown that there is a polynomial time algorithm solving the Diffie-Hellman conjugacy problem in braid groups. If one chooses reasonable parameters this algorithm is not feasible today but it seems that a braid group based version of this cryptosystem is not secure in the future.

7 Efficiency and Security Considerations

Although we are not going to propose concrete examples of Gröbner basis cryptosystems, we are now going to discuss some issues one has to confront if one tries to construct hard instances.

A. Efficiency. Both for encryption and decryption, the users of Gröbner basis cryptosystems have to be able to compute efficiently in the ring $K[M]$ where K is a computable field and M a finitely presented monoid. The complexity of the multiplication in M is controlled by the convergent term rewriting system $\xrightarrow{W}$. However, for efficient computations in F_ϱ we also have to make sure that the supports of the elements we use do not get too large.

In particular, this constraint has to be taken into account when one has to compute the normal form $\mathrm{NF}_{\tau,G}$ in the decryption procedure. To make the necessary reduction steps feasible, we have to choose the Gröbner basis G suitably. Some possibilities are apparent from the examples above:

a) If G and w are binomials, all reduction steps yield binomials, i.e. the support of all elements consists of at most two terms.

b) If G and w are homogeneous and have bounded degrees with respect to some grading, there may exist bounds on the number of terms in the support of the elements computed during $w \xrightarrow{W} \mathrm{NF}_{\tau,\bar{U}}(w)$ and on the number of reduction steps.

c) If the set $\mathbb{T}(F_\varrho) \setminus \mathrm{LT}_\tau\{U\}$ is small, the coefficients of $\mathrm{NF}_{\tau,\bar{U}}(w)$ can be found by applying suitable K-linear maps to the terms in the support of w (see Example 6.4).

B. Linear Algebra Attacks. Several types of linear algebra attacks have been proposed that apply to special Gröbner basis cryptosystems.

1) The basic type is the attack proposed in the original paper [8]. In the equation $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s$, the attacker regards the coefficients of $f_1, \ldots, f_s$ as unknowns and tries to solve the resulting linear system of

equations. In our setup, it is possible to make this attack infeasible: By choosing a large set $O_\tau(U)$, we can make the plaintext m "similar" to the ciphertext w. By using a module of large rank, we can make the solution of this linear system infeasible. Moreover, since we are working over a monoid or group ring, many products $(e_i t)t'$ with $e_i t \in \mathrm{Supp}(\bar{u}_j)$ and $t' \in \mathrm{Supp}(f_j)$ can be made to yield the same term, so that the corresponding coefficients cannot be recovered.

2) The "intelligent" linear algebra attack suggested by H.W. Lenstra and described in [14] is based on the idea that in the equation $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s$ one can guess the terms t occuring in the support of $\bar{u}_1, \ldots, \bar{u}_s$ if $t \cdot \mathrm{Supp}(f_i)$ intersects $\mathrm{Supp}(w)$, and that the list of all such terms is not too large. As before, in our approach this attack can be repelled in several ways, namely by working over group rings or by using a free module of large rank. In each case sufficient cancellation happens during the computation of the cyphertext.

C. The Differential Attack. In [12] and [13] D. Hofheinz and R. Steinwandt described a "differential" attack on the Polly Cracker cryptosystem. This attack uses the observation that in an expansion $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_s f_s$ the quotients of terms in the support of w sometimes allow conclusions about the shape of the supports of the elements $\bar{u}_1, \ldots, \bar{u}_s$. In our setting, this attack can be repelled in the very same way as the "intelligent" linear algebra attack described above.

D. The Attack Using Characteristic Terms. If a representation $w = m + \bar{u}_1 f_1 + \cdots + \bar{u}_2 f_2$ is such that there are terms in w that do not belong to $O_\tau(U)$ and therefore not to $\mathrm{Supp}(m)$ then it is sometimes possible to reveal individual messages by performing suitable linear algebra on the coefficients of w and $f_1, \ldots, f_s$, in particular when there exist "characteristic terms", i.e. terms that occur in just one of the elements f_i. By recognizing multiples of these terms in the ciphertext one can reconstruct a constant message unit. As before this attack rests on the fact that plaintext units are small, i.e. that $O_\tau(U)$ is small. Furthermore, if several products $t \cdot t'$ with $t \in \mathrm{Supp}(\bar{u}_i)$ and $t' \in \mathrm{Supp}(f_i)$ contribute to one coefficient of w this attack becomes infeasible. Thus the defensive measures described above apply.

E. Chosen Ciphertext Attacks. In the proposed cryptosystems the receiver has no method for detecting invalid ciphertexts. In addition, since decryption is K-linear, the chosen ciphertext attacks described in [9] and [14] are possible. However by using suitable hash functions the system can be made secure in the way described in [30]: The sender appends a suitable random value to his message, computes the hash value of the result, and transmits the ciphertext of the message, the ciphertext of the random value, and the hash value.

8 Discussion and Further Suggestions

Let us point out some reasons for the choices we made in presenting Gröbner basis cryptosystems and some possibilities for further generalization and improvement.

• In the original Polly Cracker cryptosystem, an attacker has several advantages that allow him to use linear algebra methods. For instance, he knows (or guesses) that the normal form of the ciphertext with respect to an unknown Gröbner basis is very simple. Since no cancellation occurs when one multiplies terms in a polynomial ring, this means that many coefficients have to vanish. Using a similar system over the monoid ring of a large monoid with a sufficient amount of cancellation foils this attack. This is the reason why we think it is advisable to use monoid rings as base rings.

• The fact that Gröbner basis theory works for modules and not just rings gives us another degree of freedom: we can encode the action of a monoid on a set (namely the set of basis vectors of a free module). Hard instances of such actions are known.

• By leaving the world of commutative rings, we gain another advantage. In most cases, submodules of free modules over non-commutative rings do not have a finite Gröbner basis, and even the computation of partial ("truncated") Gröbner bases may not be practical. Therefore it is not difficult to create sets of vectors $\{u_1, u_2, \ldots\}$ such that the module $\langle u_1, u_2, \ldots\rangle$ has no "reasonable" Gröbner basis whereas a larger module (that is kept secret) does.

• An alternative, large class of non-commutative algebras for which there exists a well-developed Gröbner basis theory is the class of path algebras (see e.g. [10], [11]). Using these algebras as base rings, we can introduce a new kind of complexity: the oriented graph underlying a path algebra can incorporate hard combinatorial problems. We believe that this aspect deserves to be examined further in the future.

• Altogether, we can conclude that Gröbner basis cryptosystems allow us to combine several difficult computational problems coming from separate areas of mathematics. We think they form a suitable framework for searching for new hard instances.

Acknowledgements The authors are grateful to the Graduiertenkolleg "Mathematische und ingenieurwissenschaftliche Methoden für sichere Datenübertragung und Informationsvermittlung" and to H. Dobbertin for supporting the Dortmund workshop. Moreover, we thank D. Grigoriev, G. Rosenberger, and Springer Verlag for enabling us to disseminate these results to a wide audience.

References

1. I. Anshel, M. Anshel, and D. Goldfeld, An algebraic method for public-key cryptography, Math. Res. Lett. **6** (1999), 287–291

2. Boo Barkee *et al.*, Why you cannot even hope to use Gröbner bases in public key cryptography: an open letter to a scientist who failed and a challenge to those who have not yet failed, J. Symb. Comput. **18** (1994), 497–501
3. H. Bluhm, Syzygienberechnung in nichtkommutativen Polynomringen, Diplomarbeit, Universität Dortmund 2005
4. J.H. Cheon and B. Jun, A polynomial time algorithm for the braid Diffie-Hellman conjugacy problem, in: Advances in Cryptology – CRYPTO 2003, Lect. Notes Comp. Sci. **2729**, Springer, 2003, pp. 212–225
5. T. ElGamal, A public key cryptosystem and a signature scheme based on discrete logarithms, IEEE Trans. Inform. Theory **31** (1985), 469–472
6. D. Farkas, C. Feustel, and E. Green, Synergy in the theories of Gröbner bases and path algebras, Canad. J. Math. **45** (1993), 727–739
7. D. Farkas, E. Green, E. Kirkman, and J. Kuzmanovich, Constructing projective resolutions, Comm. in Alg. **21** (1993), 1869–1887
8. M. Fellows and N. Koblitz, Combinatorial cryptosystems galore!, Contemp. Math. **168** (1994), 51–61
9. W. Geiselmann and R. Steinwandt, Cryptanalysis of Polly Cracker, IEEE Transactions on Inf. Theory **48** (2002), 2990–2991
10. E. Green, Noncommutative Gröbner bases and projective resolutions, in: P. Dräxler (ed.) et al., *Computational methods for representations of groups and algebras*, Proc. Euroconference Essen, Germany, April 1–5, 1997, Progress in Math. **173**, Birkhäuser, Basel 1999, 29–60
11. E. Green, Multiplicative bases, Gröbner bases, and right Gröbner bases, J. Symb. Comput. **29** (2000), 601–623
12. D. Hofheinz, Angriffe auf das Public-Key-System Polly Cracker, Studienarbeit, Universität Karlsruhe 2003
13. D. Hofheinz and R. Steinwandt, A "differential" attack on Polly Cracker, Extended abstract in Proceedings of the 2002 IEEE International Symposium on Information Theory, p.211
14. N. Koblitz, *Algebraic aspects of cryptography*, Alg. in Comp. and Math. **3**, Springer, Heidelberg 1998
15. K.H. Ko, S.J. Lee, J.H. Cheon, J.W. Han, J.S. Kang, and C. Park, New public-key cryptosystems using braid groups, in: Advances in cryptology – CRYPTO 2000, Lect. Notes Comp. Sci. **1880**, Springer, 2000, pp. 166–183
16. M. Kreuzer and L. Robbiano, *Computational commutative algebra 1*, Springer, Heidelberg 2000
17. M. Kreuzer and L. Robbiano, *Computational commutative algebra 2*, Springer, Heidelberg 2004 (to appear)
18. F. Levy-dit-Vehel and L. Perret, A Polly Cracker system based on satisfiability, Progress Comp. Sci. Applied Logic **23**, Birkhäuser, 2004, pp. 177–192
19. K. Madlener and F. Otto, Some applications of prefix-rewriting in monoids, groups, and rings, Reports on Computer Algebra **22**, Universität Kaiserlautern 1998
20. K. Madlener and B. Reinert, Computing Gröbner bases in monoid and group rings, in: M. Bronstein (ed.), Proc. Conf. ISSAC 1993, ACM Press, New York 1993, 254–263
21. K. Madlener and B. Reinert, Relating rewriting techniques on monoids and rings: congruences on monoids and ideals in monoid rings, Theor. Comp. Sci. **208** (1998), 3–31
22. K. Madlener and B. Reinert, String rewriting and Gröbner bases – a general approach to monoid and group rings, in: M. Bronstein, J. Grabmeier, and V.

Weispfenning (eds.), *Symbolic Rewriting Techniques*, Proc. Workshop Monte Verita 1995, Progress in Comp. Sci. and Appl. Logic **15**, Birkhäuser, Basel 1998, pp. 127–180
23. T. Mora, Gröbner bases for non-commutative polynomial rings, in: Proc. Conf. AAECC 4, Lect. Notes Comp. Sci. **229**, Springer, Berlin 1986, pp. 353–362
24. T. Mora, Gröbner bases in non-commutative algebras, in: Proc. Conf. ISSAC 1988, Lect. Notes Comp. Sci. **358**, Springer, Berlin 1989, pp. 150–161
25. T. Mora, An introduction to commutative and noncommutative Gröbner bases, Theor. Comp. Sci. **134** (1994), 131–173
26. T. Rai, Infinite Gröbner bases and noncommutative Polly Cracker cryptosystems, dissertation, Virginia Polytechnic Institute, Blacksburg 2004
27. B. Reinert, On Gröbner bases in monoid and group rings, dissertation, Universität Kaiserslautern 1995
28. R. Rivest, A. Shamir, and L.N. Adleman, A method for obtaining digital signatures and public key cryptosystems, Comm. of the ACM **21** (1978), 120–126
29. L. Ly, Gröbner Basen und das Kryptoverfahren Polly Two, dissertation, Universität Bochum 2003
30. L. Ly, Polly two – a new algebraic polynomial-based public-key scheme, this volume