An Algorithm for the Construction of a Normal Basis

Kurt Girstmair

Institut für Mathematik, Universität Innsbruck, Technikerstrasse 25/7,
A-6020 Innsbruck, Austria
E-mail: Kurt.Girstmair@uibk.ac.at

Communicated by M. Pohst
Received April 20, 1998

We present an algorithm for the construction of a normal basis of a Galois extension of degree \( n \) in characteristic 0. The algorithm requires \( O(n^4) \) multiplications in the ground field. It is based on representation theory but does not require the knowledge of representation theoretical data (like characters). © 1999 Academic Press

1. INTRODUCTION

Let \( L \) be a Galois extension of the field \( K \) with Galois group \( G = \{ s_1, ..., s_n \} \) (so \( n = [L : K] \) and \( y \in L \). One says \( y \) generates a normal basis of \( L \) over \( K \) if \((s_1(y), ..., s_n(y)) \) is a basis of the \( K \)-vector space \( L \). We simply say, by abuse of notation, \( y \) is a normal basis of \( L/K \) if this holds. Moreover, we assume char(\( K \)) = 0 for the time being, since we are interested in this case mainly. The literature contains quite a number of proofs for the existence of a normal basis \( y \), some of them of a purely field theoretic nature (e.g., \([4, 11]\)) and others using representation theory (\([1, 3, 10]\)). But the question of how to actually find a normal basis seems to be less well investigated (except for the case of finite fields, which we have ruled out). Maybe this is the reason why the authors of \([8]\) claimed that “apart from some very special extensions ... up to now no algorithm has been found to construct a normal basis ...” In reality the situation is not that bad. For instance, the most common field theoretic proof is constructive by its nature, but possibly requires an enormous number of trials (of order of magnitude \( n^6 \) or so) in the worst case. We shall briefly discuss the algorithmic aspects of this proof in Section 2.

The main purpose of this note is the deterministic construction of a normal basis that succeeds by running through a certain loop at most \( n \) (in many cases considerably fewer) times. The loop itself requires the solution of standard problems in linear algebra with \( n \times n \) matrices over \( K \) (Section 3). Although this method relies on representation theory, it does not demand...
knowledge of representation theoretical data. It turns out, however, that this kind of data (e.g., the complex character table of $G$) is quite useful if it is known: The problem can be reduced to “smaller” problems (smaller matrices) which are independent of each other and, thus, can be treated simultaneously. A detailed analysis of our algorithm is given in Sections 4 and 5.

At this point some other papers dealing with the algorithmic construction of normal bases—in the special case of an abelian or cyclic group $G$—should be mentioned, namely [5, 7, 9].

2. THE STANDARD METHOD

In the above setting let $(x_1, ..., x_n)$ be a basis of the $K$-vector space $L$. Let $z = (z_1, ..., z_n)$ be an $n$-tuple of elements of $K$ and consider the element

$$y = \sum_{j=1}^{n} x_j z_j \in L.$$ 

Most field theoretic proofs of the existence of a normal basis (e.g. [6, p. 229; 4; 11]) come down to the following fact: $y$ is a normal basis if, and only if, $z$ is not a zero of the polynomial

$$D = \det \left( \sum_{j=1}^{n} s_k s_l(x_j) Z_{k,l} \right)$$

in $n$ indeterminates $Z_1, ..., Z_n$ over $L$. In principle, an $n$-tuple $z$ of the desired kind can be found by trial and error. But as far as we know all upper bounds for the number of trials are extremely high. For example, since $D$ has degree (at most) $n$, a suitable $z$ is certainly contained in the cube

$$\{ z \in \mathbb{Z}^n; 0 \leq z_1, ..., z_n \leq n \}$$

of $(n+1)^n$ elements, so $(n+1)^n$ is such a bound. This defect seems to be inevitable as long as no specific information about the polynomial $D$ is known. Indeed, a homogeneous polynomial of degree $n$ in $n$ indeterminates has $m = \binom{2n-1}{n}$ coefficients, so one may prescribe $m - 1$ of its zeros arbitrarily; the number $m$ is much larger than $2^n$ for $n \geq 5$, say. Therefore, one has to reckon on upper bounds of this order of magnitude, which is definitely beyond real possibilities for $n \geq 50$.

Nevertheless, finding a normal basis in this way by a reasonable number of attempts is not hopeless. Let us look at one attempt: The said huge number of coefficients prohibits the storage of the polynomial $D$. Consequently,
one will compute $D(z)$ as a determinant for each individual $z \in K^n$ in question. We disregard problems that may arise from the action of the group elements $s_k$: One has either to store the $n$ matrices $A_k = (A_{k,j})_j$ (of dimension $n \times n$ over $K$) defined by

$$s_k(x_i) = \sum_{j=1}^{n} A_{k,j} x_j, \quad k = 1, ..., n,$$

or to compute these matrices where necessary. We also disregard problems that may arise from the arithmetic of the ground field $K$ (such as extremely "long" numbers). Apart from these possible problems the computation of the determinant $D(z)$ requires $O(n^3)$ multiplications in $L$. This amounts to $O(n^4)$ multiplications in $K$, provided that one works with a primitive element of $L$ over $K$ and reduction modulo its minimal polynomial. We have learned from the referee's comment that this number can be reduced to $O(n^3)$:

Consider the matrix $B$ whose columns are $A_1 z^T, ..., A_n z^T$, where $z^T$ means the transpose of $z$. Then it is not hard to see that $y$ is a normal basis if, and only if, $\det(B) \neq 0$. But computing the matrix $B$ requires $O(n^3)$ $K$-multiplications; as its entries are in $K$, its determinant is not more expensive.

The method described below also needs, under the same premises, $O(n^3)$ multiplications in $K$ for the main loop. Since it suffices to run this loop at most $n$ times, we obtain a total of $O(n^4)$.

3. BASIC FEATURES OF THE ITERATIVE METHOD

First we describe this method in a rather general context, which will be specialized in the next section. Let $A$ be a finite-dimensional semisimple $K$-algebra and $V$ a (left) $A$-module that is isomorphic to (the $A$-module) $A$. Thus, there exists an element $v \in V$ such that the map

$$A \to V; a \mapsto av$$

is an $A$-linear isomorphism. We show how to detect such an element $v$. Our method relies, quite substantially, on the following assumption: Let $I$ be a (left) ideal of $A$, i.e., a (left) $A$-submodule of $A$, and let a $K$-basis of $I$ be given. Then we assume that it is possible to find a $K$-basis of an $A$-linear complement $I'$ of $A$, i.e., of a (left) ideal $I'$ such that

$$A = I \oplus I'.$$

The algorithm starts with an arbitrary element $u \in V$, $u \neq 0$. The main loop produces a new element $u' \in V$ such that the module $Au'$ is strictly larger than $U = Au$ as long as $U \neq V$. Therefore, the $K$-dimension of $V$ is
an obvious upper bound for the number of iterations of this loop. Given
the said element $u$, we consider the $\mathcal{A}$-linear map

$$r_u : \mathcal{A} \to V; \alpha \mapsto r_u(\alpha) = x u.$$ 

Let $\mathcal{I}$ be the kernel of $r_u$. If $\mathcal{I} = 0$, we are done. Otherwise we determine
bases of $\mathcal{I}$ and of an $\mathcal{A}$-linear complement $\mathcal{J}'$ of $\mathcal{I}$. On writing $1 \in \mathcal{A}$ in
terms of these bases, we obtain elements $e \in \mathcal{I}$ and $e' \in \mathcal{J}'$ such that

$$1 = e + e'.$$

These elements are idempotents of $\mathcal{A}$ satisfying $ee' = e'e = 0$ and $\mathcal{I} = e\mathcal{A}e,$
$\mathcal{J}' = e'\mathcal{A}e'$. We define the $\mathcal{K}$-linear map

$$l_e : V \to V; \omega \mapsto l_e(\omega) = \varepsilon \omega.$$ 

Now the main point is the observation that the image $l_e(V) = \varepsilon V$ is not
contained in $U = \mathcal{A}u$. We shall prove this below (Proposition 1). Consequently, we can find an element $w \in V$ such that $\varepsilon w \notin U$ (for example, each
$K$-basis of $V$ contains such an element). Put $\varepsilon = u + \varepsilon w$. Then $\varepsilon \varepsilon' = \varepsilon' \varepsilon = \varepsilon w$, since $\varepsilon$ is in the kernel of $r_u$. This means $\varepsilon w \in \mathcal{A}u'$ and so $\mathcal{A}u' \not\supset U$.
On the other hand, $\varepsilon \varepsilon' = u$, because of $\varepsilon u = (1 - \varepsilon) u = u$ and $\varepsilon' = 0$. But then $u \in \mathcal{A}u'$ and $U \subseteq \mathcal{A}u'$. So $\varepsilon'$ has the desired property.

**Proposition 1.** In the above setting let $V$ be isomorphic to $\mathcal{A}$ (as an
$\mathcal{A}$-module) and $e \in \mathcal{A}$ an idempotent element that generates the kernel $\mathcal{I}$ of
$r_u$. If $eV$ is contained in $U = \mathcal{A}u$, then $e = 0$ (and $r_u$ is an isomorphism).

**Proof.** Let $B \subseteq \mathcal{A}$ be an isotypical component of $\mathcal{A}$; so $B$ has the
shape $B = \mathcal{A} \eta$ for some central idempotent $\eta$. Further, there is, up to
isomorphy, exactly one simple (left) $B$-module $W$. We show $e \eta = 0$. Since $e$
$1 \in \mathcal{A}$ is the sum of all possible $\eta$'s, this gives $e = 0$.

To this end let $U'$ be an arbitrary $\mathcal{A}$-linear complement of $U$ in $V$, so
$V = U \oplus U'$. Since $U'$ is an $\mathcal{A}$-module, $eU' \subseteq U'$. On the other hand,
$eV \subseteq U$, so $eU' \subseteq U \cap U' = 0$.

**Case 1:** $\eta U' \neq 0$. Then $U'$ contains a $B$-module $W'$ isomorphic to the
said simple module $W$. Let $\text{End}(W')$ denote the algebra of $K$-linear
endomorphisms of $W'$. Consider the $K$-linear map

$$\mathcal{B} \to \text{End}(W'); \alpha \mapsto l_\alpha,$$

where $l_\alpha$ is defined as usual by $l_\alpha(w) = \alpha w$. Wedderburn's theory says that
this map is injective (more precisely, it defines an isomorphism between $\mathcal{B}$
and the algebra of $D$-linear endomorphisms of $W^\circ$, $D$ being the $K$-division algebra of $B$-linear endomorphisms of $W$). But we have

$$\varepsilon\eta W^\circ \subseteq \varepsilon\eta U^\circ = \eta\varepsilon U^\circ = 0,$$

since $\eta$ is central and $\varepsilon U^\circ = 0$. This means $l_{\eta\varepsilon} = 0$ and, by injectivity, $\varepsilon\eta = 0$.

**Case 2**: $\eta U^\circ = 0$. Then $\eta V \subseteq \eta U$ and, therefore, $\eta V = \eta U$. Hence $\eta U$ is the isotypical component of $W$ in $V$. Because of $V \cong \mathcal{A}$, $\eta U$ is isomorphic to $\mathcal{B}$ as an $\mathcal{A}$-module; in particular, the $K$-dimensions of $\mathcal{B}$ and $\eta U$ are the same. On the other hand, the map

$$\mathcal{B} \to \eta U; \alpha \mapsto \alpha\eta$$

is surjective (recall $\mathcal{B} = \eta\mathcal{A}$, $\eta$ is central, and $U = \eta\mathcal{A}$). But a $K$-linear surjection between spaces of equal (finite) dimensions is an isomorphism. This isomorphism maps $\varepsilon\eta = \eta\varepsilon \in \mathcal{B}$ onto $0$, since $\varepsilon\eta = 0$. We obtain $\varepsilon\eta = 0$.

---

**4. THE ITERATIVE METHOD IN DETAIL**

As above, let $G = \{s_1, \ldots, s_n\}$ be the Galois group of $L$ over $K$. Let $\mathcal{A}$ be the group ring

$$\mathcal{A} = K[G] = \bigoplus_{k=1}^{n} Ks_k$$

of $G$ over $K$. Then $L$ is a $K[G]$-module in the usual way. It is isomorphic to $K[G]$ by the normal basis theorem. Consequently, the method of the foregoing section can be applied to $V = L$, provided that $\text{char}(K)$ does not divide $n$: In this case $K[G]$ is semisimple, and a standard argument shows that the above assumption concerning the $\mathcal{A}$-linear complement of a (left) ideal $\mathcal{I}$ of $\mathcal{A} = K[G]$ applies, too. Indeed, take an arbitrary $K$-linear projection $\pi$ of $K[G]$ onto $\mathcal{I}$ and define $\hat{\pi}: K[G] \to \mathcal{I}$ by

$$\hat{\pi}(x) = n^{-1} \sum_{k=1}^{n} s_k \pi(s_k^{-1} x).$$

Then $\hat{\pi}$ is a $K[G]$-linear projection onto $\mathcal{I}$ and its kernel $\mathcal{J}$ has the desired property.

In the following, however, we confine ourselves to a case where the construction of a $K[G]$-linear complement is much simpler: Let $K$ be a subfield of the field of complex numbers that is closed under complex conjugation (so $c \in K$ implies $\bar{c} \in K$).
We need some additional notations. If \( \mathbf{w} = (w_1, \ldots, w_n) \) is an \( n \)-tuple of vectors in a \( K \)-vector space \( W \) and \( A = (A_{jk}) \) is an \( n \times r \) matrix with entries in \( K \), we write

\[
\mathbf{v} = \mathbf{w} A
\]

for the \( r \)-tuple \( \mathbf{v} = (v_1, \ldots, v_r) \) given by

\[
v_k = \sum_{j=1}^{n} w_j A_{jk} = \sum_{j=1}^{n} A_{jk} w_j \in W.
\]

In other words, the row \( \mathbf{v} \) arises from the row \( \mathbf{w} \) by formal matrix multiplication. Accordingly, each \( \alpha \in k[G] \) takes the shape

\[
\alpha = s \cdot a,
\]

where \( s = (s_1, \ldots, s_n) \) is the \( K \)-basis of \( k[G] \) formed by the group elements and \( a = (a_1, \ldots, a_n)^T \) is a column of elements of \( K \). For any two elements \( \alpha = s \cdot a \) and \( \beta = s \cdot b \) of \( k[G] \), we define the scalar product

\[
\langle \alpha, \beta \rangle = \langle a, b \rangle = \sum_{j=1}^{n} a_j b_j.
\]

This scalar product is \( G \)-invariant, i.e., \( \langle s_k \cdot \alpha, s_k \cdot \beta \rangle = \langle \alpha, \beta \rangle \) for each \( s_k \in G \). Hence the orthogonal complement

\[
\mathcal{I} = \{ \beta \in k[G]; \langle \alpha, \beta \rangle = 0 \text{ for each } \alpha \in \mathcal{I} \}
\]

of an ideal \( \mathcal{I} \) in \( k[G] \) is a \( k[G] \)-linear complement, too—a fact that simplifies the performance of our method considerably.

Let \( \mathcal{X} = (x_1, \ldots, x_n) \) be a \( K \)-basis of \( L \). For any \( \alpha \in k[G] \), let \( M(\alpha) \) denote the matrix of the \( K \)-linear map \( l_\alpha \colon L \to L; w \mapsto \alpha w \) relative to the basis \( \mathcal{X} \); in other words,

\[
\alpha \mathcal{X} = (\alpha x_1, \ldots, \alpha x_n) = \mathcal{X} M(\alpha).
\]

Similarly, for \( w \) in \( L \), \( M(w) \) denotes the matrix of the map \( r_w \colon k[G] \to L; \beta \mapsto \beta w \) with respect to the bases \( \mathcal{X} \) and \( \mathcal{X} \); so \( (s_1(w), \ldots, s_n(w)) = \mathcal{X} M(w) \).

As in Section 2 our measure of complexity is the required number of \( K \)-multiplications: Hence we say that a certain procedure is \( O(n^r) \) if it needs at most \( O(n^r) \) multiplications in \( K \). We also assume, as in the said section, that we have the matrices \( M(s_k), k = 1, \ldots, n \), to hand. If this is not the case—for reasons of limited memory capacity—the successive computation of all of these matrices may require \( n \) multiplications of \( n \times n \) matrices over \( K \). Since this problem occurs at two points in the main loop, the said shortage
of memory may increase the costs of our algorithm to $O(n^5)$ instead of $O(n^4)$. But in this case the standard method is also more expensive: Each attempt is $O(n^4)$ instead of $O(n^3)$.

Now suppose that $u = \chi b$ is a nonzero element of $L$ given by the column $b = (b_1, \ldots, b_n)^T$ of elements of $K$, for instance, $u = x_1$. We compute the matrix $B = M(u)$, whose columns are $M(s_1) b, \ldots, M(s_n) b$; this computation is $O(n^3)$. An element $\alpha = s a \in K[G]$ is in the kernel $\mathcal{F}$ of $r_a$ if, and only if, $Ba = 0$. Using elementary operations with rows and, if necessary, some interchanges of columns, one transforms $B$ into a matrix of the shape

$$
\begin{pmatrix}
I_q & C \\
0 & 0
\end{pmatrix},
$$

where $I_q$ is the $q \times q$ unit matrix and $C$ is a $q \times (n - q)$ matrix. This transformation is also $O(n^3)$. We neglect the possible column changes (for instance, on renumbering the group elements $s_1, \ldots, s_n$). Then the columns of the matrices

$$
\begin{pmatrix}
- C \\
I_{n-q}
\end{pmatrix}, \quad \begin{pmatrix}
I_q \\
C^T
\end{pmatrix},
$$

form a basis of the nullspace of $B$ and of its orthogonal complement, respectively. Therefore, the computation of $e \in \mathcal{F}$ and $e' \in \mathcal{F}' = \mathcal{F}^\perp$ comes down to the solution of the system of linear equations

$$
\begin{pmatrix}
- C \\
I_{n-q}
\end{pmatrix} \begin{pmatrix}
L \\
C^T
\end{pmatrix} z = d,
$$

in the unknowns $z = (z_1, \ldots, z_n)^T$, the column $d$ being defined by $1 = s_d$ ($\in K[G]$). This is $O(n^3)$ again. Next put

$$
e = \begin{pmatrix}
- C \\
I_{n-q}
\end{pmatrix} \begin{pmatrix}
L \\
C^T
\end{pmatrix} z = d,
$$

Because of $e = \chi e$, we are in a position to establish the matrix

$$
E = M(e) = \sum_{j=1}^{n} e_j M(s_j)
$$

now (which is $O(n^3)$). Let $E_k$ denote the $k$th column of $E$. Then $e x_k = \chi E_k$, $k = 1, \ldots, n$. Since $s L \not\subseteq U = K[G] u$, there is an index $k$ such that $e x_k \notin \langle s_1(u), \ldots, s_k(u) \rangle$, i.e., $E_k$ is not in the space spanned by the columns of $B$. 


Finding such an index $k$ is $O(n^3)$, too: The matrix $B$ has rank $q$, so it can be brought into the shape

$$B = \begin{pmatrix} \mathbb{I}_q & 0 \\ D & 0 \end{pmatrix}$$

by means of column operations and interchanges of rows (which is $O(n^3)$). We neglect the row changes. Then $E_k$ is in the column space of $B$ if, and only if,

$$0 = E_{1k} \bar{B}_1 + \cdots + E_{qk} \bar{B}_q - E_k,$$

where the $\bar{B}_i$'s are the respective columns of $\bar{B}$. But checking this relation is $O(n^2)$ for each $k$, which proves our assertion. An appropriate $k$ being known, we restart the loop with

$$u' = u + xk = y(b + E_k)$$

instead of $u$.

Since all steps in the loop are $O(n^3)$, this is true for the loop as a whole. Accordingly, the construction of a normal basis by this method is $O(n^4)$.

5. SOME ADDITIONAL REMARKS

The algorithm, as described in the foregoing section, can be improved in various ways. For example, it is advisable to choose, in the final part of the loop, the index $k$ in such a way that the module $K[G]\langle x_k \rangle$ is large. In other words, one will compute the rank of the matrix

$$(M(s_1) E_k | \cdots | M(s_n) E_k)$$

and eventually try another $k$ if this rank is small.

The main loop need not be iterated as many as $n$ times in most cases: Let

$$K[G] = \sum_{j=1}^{r} \mathcal{B}_j$$

be the decomposition of the group ring into isotypical components, so

$$\mathcal{B}_j \cong W_j^{n_j}$$

\hfill (\ast)
for some simple \( K[G] \)-module \( W_j, j = 1, \ldots, r \). Therefore, \( K[G] \) and \( L \) split into \( p = p_1 + \cdots + p_r \), simple \( K[G] \)-modules each, so we are certainly done after running the main loop \( p \) times. It is well-known that
\[
p \leq d_1 + \cdots + d_r,
\]
the \( d \)’s denoting the degrees of the irreducible complex characters of the group \( G \).

If these characters are known, it is easy to write down the central idempotents \( \eta_j \) that generate the \( B_j \)’s (as \( K[G] \)-modules): Let \( \chi \) be an irreducible complex character of \( G \) of degree \( d \), say, and \( \chi_1, \chi_2, \ldots, \chi_l \) its \( K \)-conjugates. Then \( \psi = \chi_1 + \cdots + \chi_l \) is a character with values in \( K \) (though not necessarily the character of a \( K[G] \)-module). But the element
\[
(d/n) \sum_{j=1}^{n} \psi(s_j) s_j
\]
is one of the idempotents \( \eta_j \), and \( B_j = \eta_j L \) has the \( K \)-dimension \( ld^2 \). All idempotents \( \eta_j, j = 1, \ldots, r \), arise in this way. One can apply the iterative method to \( A = B_j \) and \( V = \eta_j L \) instead of \( K[G] \) and \( L \). If we have found elements \( y_j \in \eta_j L \) such that \( \eta_j L = \eta_j y_j \), we get a normal basis \( y \) of \( L/K \) on setting \( y = y_1 + \cdots + y_r \). Of course, the transition to these smaller algebras \( B_j \) and modules \( \eta_j L \) also requires some work. But the main loop becomes much simpler thereby, and one may run it for all \( B_j \)’s simultaneously (in the sense of parallel processing). For instance, consider \( K = \mathbb{Q} \) and the group \( G = \text{PSL}(3, 2) \) of order \( n = 168 \). The algebras \( B_j \) have the \( \mathbb{Q} \)-dimensions 1, 18, 36, 49, and 64, and the respective exponents \( p_j \) in \((*)\) are 1, \( \leq 3 \), 6, 7, \( \leq 8 \) (in two cases we give upper bounds only, since we did not calculate the Schur indices in question). Thus, a normal basis is detected after at most 25 runs of the main loop, or after at most 8 runs of up to 5 parallel loops. The “Atlas of Finite Groups” [2] contains the relevant data for many examples of this kind.

ACKNOWLEDGMENT

The author thanks the referee for his enlightening comments on complexity questions.

REFERENCES