ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 22, Number 4, Fall 1992

DO SUBSPACES HAVE DISTINGUISHED BASES?

DANIEL R. FARKAS AND EDWARD L. GREEN

While trying to develop a computer program to calculate resolutions for modules over path algebras, the second author conjectured the existence of an abstract version of the Gram-Schmidt process. Given a basis for a vector space, there seemed to be an "algorithmically preferred" basis for each subspace. Although this idea is quite simple-minded, it does not appear explicitly in any of the standard treatments of elementary linear algebra. On the other hand, mathematics teachers will recognize our observation as a concrete description of what we have all noticed and tried to explain when teaching Gaussian elimination. In clarifying the obvious we provide some insights into the construction of Gröbner bases, a fundamental tool in computational algebra.

We wish to take advantage of the ordering in an ordered basis for a vector space. Sometimes a concrete space comes equipped with a natural ordered basis and, sometimes, as we shall see in an application to diagonalizability, the ordering can be quite arbitrary.

Example 1. Let K be a field and let $V = K^n$ be the vector space of n-tuples with coordinates from K. The standard basis $e_1, \dots, e_n$ has a standard well-ordering, namely $e_1 < e_2 < \cdots < e_n$. In our discussion of row echelon form we refer to the *reverse* ordering on the standard basis: $e_n < e_{n-1} < \cdots < e_1$.

We introduce definitions and notations which will be used in the remainder of the paper. Let V be a vector space over a field K with a given basis B which is well-ordered by $<$. Each $v \in V$ can be written in a unique way as a linear combination of members of B; if $b \in B$ and its coefficient in this linear combination is nonzero, we will say that b *occurs in* v. The maximal $b \in B$ (by the ordering of B) which occurs in v is called the *tip of* v. If X is a nonempty subset of V, then $\mathrm{TIP}(X)$ will consist of all basis elements in B which occur as the tip of

Received by the editors on March 26, 1990, and in revised form on September 10, 1990.
Both authors were partially supported by grants from the National Science Foundation.

some nonzero vector in X. The complement of $\mathrm{TIP}(X)$ in B is denoted $\mathrm{NONTIP}(X)$.

The novelty of our presentation lies in the utilization of the next definition. Let X be a nonempty subset of V. A vector $x \in X$ is *sharp* for X provided its tip appears in x with coefficient 1 and no other basis element which occurs in x ever occurs as the tip of any other vector in X. Thus, if b occurs in x and $b < \mathrm{TIP}(x)$, then $b \in \mathrm{NONTIP}(X)$. The collection of vectors sharp for X is denoted by $\mathrm{SH}(X)$. For emphasis, we point out that $\mathrm{SH}(X)$ is a subset of X uniquely determined by the given basis B and its ordering. Our main goal is to prove that the set of sharp vectors for a subspace always constitutes a basis for the subspace. In what follows, W denotes a nonzero subspace of V.

Lemma 1. *If* $x, y \in \mathrm{SH}(W)$, *then* $\mathrm{TIP}(x) = \mathrm{TIP}(y)$ *if and only if* $x = y$.

Proof. If x and y have the same tip b, then $x-y$ is a linear combination of basis vectors smaller than b which occur in either x or y. But these basis elements are in $\mathrm{NONTIP}(W)$. Since $x - y$ has no tip, $x - y = 0$. $\square$

Lemma 2. $\mathrm{SH}(W)$ *is a linearly independent set.*

Proof. Let $x_1, \dots, x_n$ be distinct elements of $\mathrm{SH}(W)$. By Lemma 1, the tip of x_i cannot occur in x_j for $j \neq i$. Consequently, if $\alpha_i \in K$ and $\sum \alpha_i x_i = 0$, then $\sum \alpha_i \mathrm{TIP}(x_i) = 0$. It follows that each α_i is zero. $\square$

Theorem 3. *Let* V *be a vector space with a well-ordered basis* B *and let* W *be a subspace of* V. *Then* $\mathrm{SH}(W)$ *is a basis for* W.

Proof. To clarify the argument, we introduce some suggestive notation. If $x \in \mathrm{SH}(W)$ and $w \in W$, let $\langle w, x \rangle$ denote the coefficient of the basis element $\mathrm{TIP}(x)$ in the expansion of w as a linear combination of members of B.

The key step is to observe that each element of $\mathrm{TIP}(W)$ appears as the tip of some sharp vector for W. Indeed, suppose not. By well-ordering, there is a minimal basis vector b which lies in $\mathrm{TIP}(W)$ but not in $\mathrm{TIP}(\mathrm{SH}(W))$; choose $w \in W$ so that b is its tip and b has coefficient 1 in w. The minimal choice of b implies that all other tips of W which occur in w lie in $\mathrm{TIP}(\mathrm{SH}(W))$. Consider

$$w' = w - \sum_{x \in \mathrm{SH}(W)} \langle w, x \rangle \cdot x.$$

Then the unique tip which occurs in w' is b. That is, w' is sharp for W. We reach the contradiction that $b \in \mathrm{TIP}(\mathrm{SH}(W))$.

Now take an arbitrary $u \in W$. Since $u - \sum_{x \in \mathrm{SH}(W)} \langle u, x \rangle \cdot x$ has no tip, we must have

$$(*) \qquad u = \sum_{x \in \mathrm{SH}(W)} \langle u, x \rangle \cdot x.$$

□

The reader will notice that the formula $(*)$ is some sort of projection formula with $\mathrm{SH}(W)$ playing the role of an orthonormal basis. As an illustration of this analog, notice that $\mathrm{NONTIP}(W)$ is the basis of a *canonical* subspace complementary to W in V.

The notion of sharp basis allows us to give a particularly transparent proof that the restriction of a diagonalizable linear transformation to an invariant subspace is diagonalizable. Suppose V is a finite dimensional space with ordered basis $v_1, v_2, \ldots, v_n$ and T is a linear transformation on V such that $T(v_j) = \lambda_j v_j$. Assume that we are given a T-invariant subspace W of V. The proof consists of observing that a sharp vector for W is an eigenvector for T. For suppose that $v = v_k + \sum_{i \in N} \alpha_i v_i \in W$ has tip v_k and $v_i \in \mathrm{NONTIP}(W)$ for each $i \in N$. If the eigenvalue $\lambda_k = 0$, then $T(v)$ is a linear combination of nontips for W and, consequently, $T(v) = 0$. If $\lambda_k \neq 0$, then $(1/\lambda_k)T(v)$ is also a sharp vector for W with the same tip as v; apply Lemma 1.

As another illustration of these results, we show that Gaussian elimination provides a method for finding a basis of sharp vectors, given a subspace spanned by a set of vectors in n-space. We also obtain the uniqueness of the reduced row-echelon form without any further work. Suppose that we have vectors $x_1, \dots, x_r$ in Euclidean n-space K^n. Let M be the $r \times n$ matrix whose rows are $x_1, \dots, x_r$ and let M^* be the reduced row-echelon form of M with nonzero rows $x_1^*, \dots, x_t^*$. Giving the reverse ordering to the standard basis of K^n (see Example 1), we see that $x_1^*, \dots, x_t^*$ are all of the sharp vectors for the row space of M^*. Since Gaussian elimination does not change the row space, we see that Gaussian elimination provides an algorithm to find the basis of sharp vectors for the span of $x_1, \dots, x_n$. Moreover, the uniqueness of the set of sharp vectors yields the uniqueness of the reduced row-echelon form. Finally, the formula $(*)$ given at the end of Theorem 3 explicitly states why the basis of sharp vectors is "nice." That is, if x is in the span of $x_1, \dots, x_r$, then $x = \sum_{i=1}^{t} \langle x, x_i^* \rangle \cdot x_i^*$.

We now make a jump in sophistication.

Example 2. Consider the commutative polynomial ring $R = K[x_1, \dots, x_n]$ as a vector space over the field K. It has a basis B which consists of all monomials together with 1. Notice that B is a cancellative monoid; that is, if $ab = ac$, then $b = c$. If we order the variables $1 < x_1 < x_2 < \cdots < x_n$, then B can be totally ordered by using degree and lexicographic ordering. That is, if $m = x_1^{a_1} \cdots x_n^{a_n}$ and $m' = x_1^{b_1} \cdots x_n^{b_n}$, then $m < m'$ if either $\sum a_i < \sum b_i$ or if $\sum a_i = \sum b_i$ and there is a $1 \leq j \leq n$ so that $a_i = b_i$ for $i < j$ and $a_j < b_j$. Notice that $\leq$ is a well ordering and is compatible with multiplication in B.

In this example B comes equipped with an intrinsic partial order, divisibility. Explicitly, $x_1^{e_1} \cdots x_n^{e_n}$ divides $x_1^{f_1} \cdots x_n^{f_n}$ when $e_1 \leq f_1, \dots,$ and $e_n \leq f_n$. This can also be regarded as the point-wise partial order on the n-fold Cartesian product of the natural numbers with the usual ordering. Divisibility enjoys an often proved property that has been attributed to Dickson (cf. [**4**]): any infinite subset of B contains two monomials which are comparable by divisibility. Equivalently, $\mathbf{N}^n$ has no infinite antichains. (An *antichain* is a set of pairwise incomparable elements.) It is not difficult to verify this assertion by induction on n.

Fix a degree-lexicographic order $<$ on B and let $\mathcal{I}$ be a nonzero ideal of the polynomial ring R. A finite set of polynomials $G = \{z_1, \dots, z_m\}$ in $\mathcal{I}$ is called a *Gröbner basis* for $\mathcal{I}$ if the ideal generated by the tips of

G contains the tips of all polynomials in $\mathcal{I}$, i.e.,

$$\text{ideal generated by } \mathrm{TIP}(G) = \text{ ideal generated by } \mathrm{TIP}(\mathcal{I}).$$

It is straightforward to show that a Gröbner basis for $\mathcal{I}$ generates $\mathcal{I}$. A distinguished Gröbner basis is lurking behind all of the clutter which has accumulated to this point. Let $\min \mathrm{TIP}(\mathcal{I})$ denote the collection of those monomials which are minimal in $\mathrm{TIP}(\mathcal{I})$ with respect to divisibility. As we observed in the previous paragraph, $\min \mathrm{TIP}(\mathcal{I})$ is finite. A polynomial in $\mathrm{SH}(\mathcal{I})$ is *minimally sharp* provided its tip lies in $\min \mathrm{TIP}(\mathcal{I})$; obviously, there are finitely many of these.

Theorem 4. *Let $\mathcal{I}$ be a nonzero ideal of R. If $<$ is a degree-lexicographic ordering on the monomials, then the finite set of minimally sharp polynomials of $\mathcal{I}$ constitutes a Gröbner basis for $\mathcal{I}$.*

Proof. It suffices to prove that every member of $\mathrm{TIP}(\mathcal{I})$ is divisible (in the monoid B) by some element in $\min \mathrm{TIP}(\mathcal{I})$. If $b \in \min \mathrm{TIP}(\mathcal{I})$, choose $c \in \min \mathrm{TIP}(\mathcal{I})$ minimal with respect to c dividing b. We know from Theorem 3 that c is the tip of some sharp polynomial which is, necessarily, minimally sharp. $\square$

The astute reader will notice that Theorem 4 is a constructive proof of the Hilbert Basis Theorem. It is the first step of a theory initiated by Hermann and developed by Buchberger [**4**], providing a framework for machine computations which answer questions about commutative rings.

Our construction of the set of minimally sharp polynomials illustrates the concept of a *reduced* Gröbner basis as described in [**4**, Theorem 8.3]. It is worth noting that $\mathrm{NONTIP}(\mathcal{I})$ is determined by the minimally sharp polynomials and vice versa. This can be stated more precisely in the next result whose proof is left to the reader.

Proposition 5. *Let R, $\mathcal{I}$, and $<$ be as in Theorem* 4. *Then* $\mathrm{NONTIP}(\mathcal{I})$ *is the set of monic monomials which are not divisible by an element in* $\min \mathrm{TIP}(\mathcal{I})$. *Furthermore,* $\min \mathrm{TIP}(\mathcal{I})$ *is the set of monomials not in* $\mathrm{NONTIP}(\mathcal{I})$ *whose proper divisors all lie in* $\mathrm{NONTIP}(\mathcal{I})$.

Example 3. Let $R = K[x, y]$ where K is a field of characteristic zero, and let B be the ordered monoid described in Example 2, subject to $x < y$. Suppose $\mathcal{I}$ is the ideal generated by $xy^2 - x^2$ and $x^2y - y^2$. The minimally sharp polynomials for $\mathcal{I}$ turn out to be $xy^2 - x^2$, $x^2y - y^2$, $y^3 - x^3$, and $x^4 - y^2$. The point, of course, is that this list can be calculated algorithmically where, after comparing common factors of the tips of a generating set of polynomials, simple operations on polynomials are used to create a new generating set from the previous one [**3**, Section 3]. Having the minimal sharp polynomials we conclude $\min \mathrm{TIP}(\mathcal{I}) = \{xy^2, x^2y, y^3, x^4\}$. By Proposition 5, we see that $\mathrm{NONTIP}(\mathcal{I}) = \{1, x, x^2, x^3, y, y^2, xy\}$. As observed earlier, the K-linear span of $\mathrm{NONTIP}(\mathcal{I})$ is complementary to the subspace $\mathcal{I}$. (Each polynomial $f \in R$ decomposes uniquely as $f_1 + f_2$ where $f_1 \in \mathcal{I}$ and f_2 lies in this complement, an example of the so-called "rest of f" or *normal form* of a polynomial [**3**, Definition 2.3 or **2**, Corollary 8.2].) In particular, R/I is seven-dimensional.

We close this note by clarifying the noetherian argument which appeared implicitly in Theorem 4. Recall that if $\leq$ is a partial order on a set Y then a subset $X \subseteq Y$ is an *order ideal* provided that $x \in X$ and $y \geq x$ imply that $y \in X$. The next lemma is due to Higman; a parital order satisfying any of the equivalent properties is called a *well partial ordering*.

Lemma 6 ([**1**]). *The following conditions on a partially ordered set Y are equivalent*:

(i) *The ascending chain condition holds for the order ideals of Y.*

(ii) *Every infinite sequence of elements of Y has an infinite ascending subsequence.*

(iii) *Every infinite sequence of elements of Y has an ascending subsequence of length* 2.

(iv) *There exist in Y neither an infinite strictly descending sequence nor an infinite antichain.*

In Example 2, divisibility on B is a well partial order. The degree-lexicographic order is *monoidal* in the following sense: $1 \leq a$ for all

$a \in B$ and $b \leq c$ implies $bd \leq cd$ (and $db \leq dc$) for all $d \in B$. There is a lovely interconnection between these properties.

Lemma 7 ([**3**, Lemma 1.3]). *Assume that M is a monoid which is cancellative on each side and that $\leq$ is a monoidal total order on M.*

(1) *If a divides b, then $a \leq b$.*

(2) *If left divisibility on M is a well partial order, then $\leq$ is a well order.*

Proof. (1) We are supposing that $ac = b$ for some $c \in M$.

If $a > b$, then $ac > bc$. But $c \geq 1$ implies that $bc \geq b$. Hence, $b = ac > b$, a contradiction. Therefore, $a \leq b$.

(2) Apply (1) and, for example, (iii) of Lemma 6. □

As one consequence, the assumption that $\leq$ is a degree-lexicographic ordering in Theorem 4 can be replaced with the hypothesis that $\leq$ is a monoidal total order on the collection of monomials. These same ideas can be exploited to give a very short proof for a theorem of J. Lewin.

Example 4. Consider the free algebra $F = K\langle x_1, \dots, x_n \rangle$. As a vector space over K, it has a basis B consisting of all words in the alphabet $x_1, \dots, x_n$. Now suppose that L is a semigroup ideal of B, a nonempty subset closed under left and right multiplication by elements in B. Let $\overline{F}$ be the nonomial algebra obtained by factoring out the two-sided algebra ideal generated by L. It is not difficult to check that $\overline{F}$ has as basis $\overline{B} = B \backslash L$. Moreover, one can obtain the *multiplication table* for $\overline{B}$ by contracting L to zero; if the product of two words in $\overline{B}$ lies in L, their product in $\overline{F}$ is 0.

Theorem 8 ([**2**]). *If a monomial algebra $\overline{F}$ is right noetherian, then it is finitely presented. That is, the ideal generated by the collection of monomial relations, L, is finitely generated as a bimodule or two-sided ideal.*

Proof. Given $a, b \in B$ we say that b is a *subword of* a when there exist $u, v \in B$ such that $a = ubv$. The partial order of being a subword is the noncommutative analogue of divisibility. Thus the role of the minimal tips for the ideal generated by L is played by

$$\min(L) = \{a \in L | \text{ no proper subword of } a \text{ lies in } L\}.$$

It suffices to prove that $\min(L)$ is finite.

Let $\leq$ denote the partial order of left divisibility on $\overline{B}$. If X is an order ideal of $\overline{B}$, then the vector space span of X is a right ideal of $\overline{F}$. Since $\overline{F}$ is right noetherian, condition (i) of Lemma 6 tells us that $\leq$ is a well partial order. Consequently, $\overline{B}$ has no infinite antichains under $\leq$.

Suppose $\min(L)$ is infinite. If $a \in \min(L)$ write $a = i(a)r(a)$ where $i(a)$ is the initial letter of a. Then $r(a) \in \overline{B}$. Since the alphabet is finite, there exists a member x_j of the alphabet such that

$$\{r(a) | x_j r(a) \in \min(L)\}$$

is infinite. It follows that at least two elements in this set are comparable, say $r(a) < r(b)$. Then $x_j r(a) < x_j r(b)$ and, so, $a < b$. But now a is a proper subword of b while both lie in $\min(L)$. $\square$

REFERENCES

1. G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. **2** (1952), 326–336.

2. J. Lewin, *A matrix representation for associative algebras*. I, Trans. Amer. Math. Soc. **188** (1974), 293–308.

3. F. Pauer and M. Pfeifhofer, *The theory of Gröbner bases*, L'Enseignement Mathématique, 2nd Série **34** (1988), 215–232.

4. L. Robbiano, *Gröbner bases: a foundation for commutative algebra*, Notes distributed at Computers and Mathematics 1989, MIT, June, 1989.

Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061