A NOTE ON GROWTH FUNCTIONS OF ALGEBRAS AND SEMIGROUPS

George M. Bergman

Let R be an associative algebra over a field k, generated by a finite set X. Let V be the k-subspace of R spanned by $X \cup \{1\}$. One is interested in the rate of growth of the function $g_{X,R}(n) = \dim_k(V^n)$, where V^n is the subspace of R spanned by all products of n elements of V (equivalently, of X.) We show in section 1 below that the growth function of any k-algebra is also the growth function of a semigroup with 0, and vice versa. In section 2 we show that any such function with less than quadratic rate of growth has at most linear growth, recovering a result of ?
In section 3 we give examples of various growth-behaviors.

1. *Irreducible monomials.* All semigroups will have 1, and all rings and algebras will be associative with 1.

Let k be a field and R a finitely generated associative k-algebra. To be a little more formal than above, let X be a finite generating *family* for R, i.e. a finite set given with a map into R, $x \mapsto x_R$, whose image generates R as a k-algebra. Let us choose a total ordering for X, and totally order the free semigroup $\langle X \rangle$ on X by setting $a < b$ if a is shorter than b, or has the same length as b but precedes it lexicographically. We now define the subset $W_{X,R} \subseteq \langle X \rangle$ of "irreducible monomials" by taking $a \in W_{X,R}$ if and only if the image a_R of a in R is *not* equal to a k-linear combination of the elements b_R where $b < a$ in $\langle X \rangle$. It is not hard to see that the elements a_R $(a \in W_{X,R})$ form a k-basis for R, and that $g_{X,R}(n) = \text{card}\ \{a \in W_{X,R} \mid \text{length}(a) \leq n\}$.

A semigroup with zero will mean a pair $(S,0)$, where S is a semigroup, and 0 an element satisfying $a0=0a=0$ for all $a \in S$. When we speak of generating such an object, it will be understood that 0, like 1, is given to us free.

Work done while the author was partially supported by NSF grant .

$X_S \cup \{1\}$

If $(S,0)$ is generated by the finite family X, we define $g_{X,(S,0)}(n)$ to be the number of nonzero elements of X_S^n; and if X is ordered we define $W_{X,(S,0)}$ to be the set of $a \in \langle X \rangle$ such that no relation $a_S = 0$ or $a_S = b_S$ $(b < a$ in $\langle X \rangle)$ holds in S. For ordinary semigroups S we likewise have have obvious definitions of $g_{X,S}$ and $W_{X,S}$. If $(S,0)$ is a semigroup with zero and k a field, the semigroup-with-zero-algebra $k(S,0)$ will mean the factor-algebra of the ordinary semigroup algebra kS by the 1-dimensional ideal spanned by $0 \in S$.

Lemma 1. Let X be an ordered set, W a subset of $\langle X \rangle$, and k a field. Then the following conditions are equivalent:

(i) $W = W_{X,(S,0)}$ for some semigroup with zero $(S,0)$ generated by an image of X.

(ii) $W = W_{X,R}$ for some k-algebra R generated by an image of X.

(iii) Every subword of an element of W lies in W.

Further, if S is a semigroup generated by an image of X, $W_{X,S}$ satisfies the above conditions, and conversely, if $W \subseteq X$ satisfies the above conditions, then either W or $W \cup \{a\}$ for some $a \in \langle X \rangle$ has the form $W_{X,S}$ for some semigroup S.

Proof. We get (i) $\Rightarrow$ (ii) by forming the semigroup-with-zero-algebra $k(S,0)$. (ii) $\Rightarrow$ (iii) follows from the observation that if an element $a \in \langle X \rangle$ satisfies a relation $a_R = \sum_{b<a} \alpha_b\, b_R$ in R, then any element $a'aa''$ of which it is a subword also satisfies such a relation, $(a'aa'')_R = \sum \alpha_b\,(a'ba'')_R$. For (iii) $\Rightarrow$ (i) we construct $(S,0)$ by identifying together all elements of $\langle X \rangle$ not in W into a single element 0.

Given a semigroup S, we may adjoin a zero-element, whence we see that $W_{X,S}$ satisfies (i). Conversely, given a semigroup with zero $(S,0)$, we may regard S as an ordinary semigroup, and we find that $W_{X,S}$ will have at most one element more than $W_{X,(S,0)}$. ∎

2. Growth rates. Let X be a finite set, and W a subset of the free semigroup $\langle X\rangle$, such that every subword of an element of W lies in W. We shall find it most convenient to look, not at the function "number of elements of W of length $\leq n$", but at "number of elements of W of length exactly n". We shall show that unless this function exceeds n for all n, it is bounded. Namely, if for some $d > 0$ W contains $\leq d$ words of length d, we shall show that every sufficiently long word in W can be broken into a short initial segment, a long periodic segment of period $\leq d$, and a short final segment, and that the number of possibilities for each segment is bounded.

(Originally I planned on a compactness proof, in which I would consider the set $\breve{W}$ of infinite words $\breve{a}\colon \mathbb{Z} \rightarrow X$, which were "limits" of words in W, show that these were periodic and finite in number, and then work back to W. The idea of looking at infinite words was helpful — I would probably not have found the proof without it. But the lemmas needed to make this proof work turned out to yield the desired result more easily without reference to infinite words.)

We must begin with some Lemmas. For any $m > 0$, we shall say that a word $a = a_1 \ldots a_n$ of length $n \geq m$ has period m if $a_i = a_{i+m}$ for all i such that $1 \leq i,\ i+m \leq n$. (Remark: under this definition, a word a may have periods which are not multiples of its minimal period. E.g., $(xxy)^r xx$, in addition to the periods $3, 6, \ldots, 3r$, also has periods $3r+1$, $3r+2$. This phenomenon would vanish if we restricted the definition of period by requiring $m \leq n/2$, as may be deduced from the next Lemma.)

Lemma 2. Suppose $a = a_1 \ldots a_n \in X^n$ is periodic, of minimal period $m \leq n$. Then if a has two equal subwords

(1) $$a_{i+1} \ldots a_{i+r} = a_{j+1} \ldots a_{j+r}$$

of length $r \geq m-1$, then their locations in a differ by an exact multiple of m, i.e. $m \mid j-i$.

Proof. By the periodicity assumption we may define a map $\bar{a}\colon \mathbb{Z}/m \to X$ by $\bar{a}_{(i)} = a_i$ $(1 \leq i \leq n)$; we may think of $\bar{a}$ as a "cyclic word". If $r \geq m$, (1) tells us that this cycle is invariant under translation by $j-i$, from which it follows that $\bar{a}$ has period g.c.d.$(m, j-i)$, hence so does a. As m is the minimal period of a, we must have $m \mid j-i$, as claimed. If $r = m-1$, then by comparing the number of occurrences of each element of X on each side of (1) and in $\bar{a}$, we deduce that the unique terms of $\bar{a}$ not represented on each side of (1), which we may write $\bar{a}_{(i)}$ and $\bar{a}_{(j)}$ respectively, must be equal, so again (1) implies that $\bar{a}$ is invariant under translation by $j-i$, whence again $m \mid j-i$. ∎

Lemma 3. Suppose $a = a_1 \ldots a_h \in X^h$, and d is a positive integer such that a has $\leq d$ distinct subwords of length d.

(i) If $h \geq 2d$, then a has a subword $a_{j+1} \ldots a_{j+r}$ periodic of some period $n \leq d$, and having length $r \geq d+n$.

(ii) If

(2) $$a_{i+1} \ldots a_{i+r}$$

is a subword of a whose minimal period m satisfies

(3) $m \leq d$ (small enough period),

(4) $r \geq d+m$ (long enough word),

(5) $i \geq d-m+1$ (enough room on left of (2) in a),

then the subword $a_i \ldots a_{i+r}$ formed by adding one more term from a to the left end of (2) is also periodic of minimal period m.

(iii) Likewise, if (2) is a subword of a whose minimal period m satisfies (3), (4) and

(6) $i+r \leq h-(d-m+1)$ (enough room on the right of (2) in a)

then the word $a_{i+1} \ldots a_{i+r+1}$ is also periodic of period m.

Hence

(iv) If $h \geq 2d$ then for some $m \leq d$, a may be written as the product of a word of length $d-m$, a word periodic of period m, and another word of length $d-m$.

<u>Proof.</u> (i) The subword $a_1 \dots a_{2d}$ of a has $d+1$ subwords of length d, hence at least two must be equal. Say

(7) $$a_{j+1} \dots a_{j+d} = a_{j+n+1} \dots a_{j+n+d} \quad (j \geq 0, \ n > 0, \ j+n \leq d).$$

It is easy to see that the word $a_{j+1} \dots a_{j+n+d}$ is periodic of period n, as required.

(ii) By possibly dropping an initial segment of a, we may assume for notational convenience that

(8) $$i = d-m+1 \quad (\text{cf. } (5)).$$

Now, as in (i), let us find j and n such that (7) holds. From (4), (7) and (8) it follows that the two sides of (7), as subwords of a, each overlap (2) in at least their last $m-1$ terms. But by (7) the last $m-1$ terms of these two words are the same. Applying Lemma 1, we conclude that $m \mid n$. Now from (8) and the last inequality of (7) we can deduce that $j < i$, hence the left-hand side of (7) contains the term a_i which we wish to attach to (2). The corresponding term on the right-hand-side of (7) is a_{i+n}, which by (4) and (8) is a term of (2). Hence as $m \mid n$, the value of a_i is indeed the value needed to extend the periodicity of (2).

(iii) holds by the symmetric argument, and (iv) follows. ▮

We can now deduce

<u>Proposition 4.</u> Let X be a finite set, and $W \subseteq \langle X \rangle$ a set of words such that $ab \in W \Longrightarrow a \in W, b \in W$. If for some integer $d > 0$ W has $\leq d$ words of length exactly d, then for all $h \geq d$, W has $\leq d^3$ words of length exactly h.

<u>Proof.</u> Let $a \in W$ be a word of length $h \geq d$. If $h \leq 2d$, we see that a is determined by its initial subword of length d and its final subword of length d, and there are at most d possibilities for each, hence at most d^2 possibilities for a. If $h \geq 2d$, then by our above result a can be written $a_1 a_2 a_3$ where a_2 is a periodic word, of some period $m \leq d$, and a_1, a_3 each have length $d-m$. It is not hard to show that $a_1 a_2$ is determined by the first $2d$ terms of a (by (i) we can find a periodicity beginning in this segment, and by (iii) this will propagate to the right), i.e. by two blocks of length d, and a_3 is determined by the final

block of length $\leq d$, giving $\leq d^3$ possibilities in all.∎

3. Examples and remarks.

3.1. Example. Let d be a positive integer, and take any family of infinite periodic words, $\hat{a},\ldots,\hat{c} \in X^{\mathbb{Z}}$, whose minimal periods sum to d, and none of which is a translate of another. Let W be the set of all finite subwords of these infinite words. It is easy to see that W has at most d elements of any length h. If $h \geqslant d$, then one can show using Lemma 2 that no subword of length h of one of our infinite periodic words can coincide with a subword of length h in another (use the fact that the sum of the two periods is $\leq d$), and deduce that there are precisely d distinct words of length h in W.

3.2. Example. Let p, q be positive integers and $d = p+q+1$. Let W consist of all words in x, y, z whose subwords of length d all belong to the list of d words

$$y^p x^{q+1},\ \ldots,\ y x^{p+q},\ x^{p+q+1},\ x^{p+q} z,\ \ldots,\ x^{p+1} z^q.$$

Then for $h \geqslant 2d$, W contains all words of the form

$$y^e\, x^{h-e-f}\, z^f \quad (0 \leq e \leq p,\ \ 0 \leq f \leq q),$$

or $(e+1)(f+1)$ elements in all. Taking $e \approx f$ we get $(e+1)(f+1) \approx d^2/4$. This is the largest number of words (as a function of d) that I have been able to realize

3.3. Remark. The above examples illustrate two ways of getting sets W satisfying the hypothesis of Proposition 4 and having fairly large numbers of words in each degree. In the first example we had a large number of possibilities for the periodic part of a word; in the second, a large number of possibilities for the initial and final subwords. I suspect that these two approaches are incompatible, in a sense which I will sketch. It seems likely that if the periodicity of a word a breaks off e terms before the left-hand side, then the e subwords of length d, $a_1 \ldots a_d$, $a_2 \ldots a_{d+1}$, $\ldots$, $a_e \ldots a_{d+e}$, first of all cannot occur in the periodic part of <u>any</u> word in W, and secondly cannot occur at the left-hand side of any word with periodic part different (up to translation) from that of a. Hence if words of W involve both several different periodic parts and several different beginning- and and end-segments connecting with each, our system is low on interchangeability of parts, and we cannot get a large total number of words. Note that if the above suggestion that a given initial segment is not compatible with more than one periodic part is true, it follows that a word in W of length h is determined by its initial length-d segment and its final length-d segment, reducing the estimate in Proposition 4 of the number of such words to d^2.

3.4. Example. Let d be any positive integer and let W consist of all words in x and y whose length-d subwords all belong to the set of $d+1$ words

$$x^{d-1}y,\ x^{d-2}yx,\ \ldots,\ yx^{d-1},\ x^d.$$

Then W contains the free semigroup generated by yx^{d-1} and x, and so has exponential rate of growth. So in Proposition 4, the hypothesis that W has no more than d elements of length d cannot be weakened to "no more than $d+1$ ...".

3.5. Example. Let S be an arbitrary subset of the positive integers, and let W consist of all words in x, y of the forms

$$x^i,\quad x^iy,\quad yx^i,\ yx^iy$$

where i is unrestricted in the first three cases, but in the last case, the length, $i+2$, is required to lie in S. Then any subword of a word in W belongs to W. The number of words of length $h \geqslant 2$ in W is 4 if $h \in S$, 3 otherwise. This shows that the count functions associated with such sets W can be quite disorderly even if bounded.

3.6. Example. Let f, g be nondecreasing positive integer valued functions on the nonnegative integers, such that

$$f(n),\quad n/f(n),\quad g(n),\quad n/g(n)$$

are all unbounded functions of n. Let W denote the set of all words in x, y, z of the form

$$x^i y^j z^k \text{ such that } i \leq f(j) \text{ if } k \neq 0, \text{ and } k \leq f(j) \text{ if } i \neq 0.$$

Again, any subword of a member of W lies in W. The rate of growth of W is essentially that of $n\, f(n)\, g(n)$. We omit the details, but this example shows that one can achieve any growth rate (modulo the appropriate equivalence relation) between linear and cubic (or if we look at the growth of the set of words of less than or equal to a given length, between quadratic and fourth degree) and by easy modifications of this trick we can get the rate of growth of x^α for any $\alpha \geq 1$.