

SPARSE SQUARES OF POLYNOMIALS

JOHN ABBOTT

ABSTRACT. We answer a question left open in an article of Coppersmith and Davenport which proved the existence of polynomials whose powers are sparse, and in particular polynomials whose squares are sparse (i.e., the square has fewer terms than the original polynomial). They exhibit some polynomials of degree 12 having sparse squares, and ask whether there are any lower degree complete polynomials with this property. We answer their question negatively by reporting that no polynomial of degree less than 12 has a sparse square, and explain how the substantial computation was effected using the system CoCoA.

1. INTRODUCTION

A polynomial has a sparse square if the number of terms in the square is strictly smaller than the number of terms in the polynomial itself. The problem we tackle in this article is that of finding the lowest degree in which there are polynomials with sparse squares.

The general question of existence of polynomials (in characteristic 0) with sparse powers was addressed and resolved in [2] where some remarkable theorems are proved covering both the case which interests us as well as several generalizations. They report that the problem had already attracted prior attention from several authors, including Erdős, Freud, Rényi, Schinzel and Verdenius. Furthermore, they gave explicitly some new polynomials with sparse squares of lower degree than any previously known example. Their polynomials are

$$(1 + 2x - 2x^2 + 4x^3 - 10x^4 + 50x^5 + 125x^6)(1 + \alpha x^6),$$

where α can take any one of eight distinct values, of which six are rational; the rational values are -110 , -253 , $-55/2$ and 15625 divided by each of these.

We answer one of the questions they left open, namely whether there is a still lower degree polynomial having a sparse square. They correctly observed that the answer can be found merely by performing a finite computation; the difficulty lies in the scale of this computation which, at the time they wrote their article, was intractably large. Since then, significant advances have been made both in computer hardware and in the algorithms/programs used for calculation. These advances, coupled with careful exploitation of the symmetries of the problem, are sufficient to bring the task within reach.

We describe how we undertook the computation using the system CoCoA [3] which offered a convenient environment in which to manage and effect the calculation of numerous Gröbner bases, an essential step of our solution. Note that the

Received by the editor February 1, 2000.

2000 *Mathematics Subject Classification*. Primary 11C04; Secondary 12Y05.

results in [2] were for “complete” polynomials, i.e., of the form $\sum_{i=0}^d a_i x^i$ where all the a_i are nonzero. In Section 4 we also consider “noncomplete” polynomials (i.e., where some of the a_i may be zero). We conclude from our extensive calculations that there is no polynomial of degree less than 12 (in characteristic 0) having a sparse square.

The author would like to thank Professor Robbiano and the other members of the CoCoA Team for their valuable help and suggestions.

2. ASSUMPTIONS AND NOTATION

We shall say that a polynomial f has a *sparse square* if the number of nonzero coefficients (“the number of terms”) in f^2 is strictly smaller than the number of terms in f . Our search will be for $f \in \overline{\mathbb{Q}}[x]$ having sparse squares, where $\overline{\mathbb{Q}}$ is the algebraic closure of the rationals.

The two following lemmas will be useful. Lemma 1 is the key to making the search finite: it allows us to ascertain the existence of polynomials with sparse squares by determining the solvability of one or more systems of polynomial equations. This latter we can do by means of Gröbner bases [1]. Lemma 2 helps shorten the search when we look for noncomplete polynomials in Section 4.

We shall say that a polynomial $f = \sum_{i=0}^d a_i x^i$ has *zero pattern* Z to mean that $Z \subseteq \{1, 2, \dots, d-1\}$ and $a_i = 0 \Leftrightarrow i \in Z$. In particular, a complete polynomial has zero pattern $Z = \emptyset$. Note that we exclude the indices 0 and d from Z , so f is not divisible by x and its degree is d .

Lemma 1. *Let d be a positive integer, and Z a zero pattern for a polynomial of degree d . Let $S = \{0, 1, \dots, d\} \setminus Z$, so the generic polynomial with zero pattern Z is $f = \sum_{s \in S} a_s x^s$, where the a_s are nonzero variables, and its square is $f^2 = \sum_{j=0}^{2d} b_j x^j$, where the b_j are polynomials of degree at most 2 in $\mathbb{Z}[a_0, a_1, \dots, a_d]$. Then there is a polynomial with zero pattern Z having a sparse square if and only if there is a subset of indices $J \subset \{0, 1, 2, \dots, 2d\}$ of cardinality at least $|Z| + d + 1$ such that the polynomial system $\{b_j : j \in J\}$ admits a solution with every a_s being nonzero.*

Proof. Both directions are simple. □

Lemma 2. *Any polynomial having a sparse square contains at least 5 terms.*

Proof. A polynomial containing a single term has a square containing a single term. A polynomial containing two terms has a square containing three. A polynomial containing three or more terms has a square containing at least four terms. Let $f = a_1 x^{d_1} + a_2 x^{d_2} + \dots + a_{t-1} x^{d_{t-1}} + a_t x^{d_t}$, where the a_i are nonzero, the d_i are in descending order, and $t \geq 3$. Then

$$f^2 = a_1^2 x^{2d_1} + 2a_1 a_2 x^{d_1+d_2} + \dots + 2a_{t-1} a_t x^{d_{t-1}+d_t} + a_t^2 x^{2d_t}$$

and these four terms do not vanish or cancel with other terms. As the square contains at least four terms, a polynomial with a sparse square must have strictly more than four. □

3. THE CASE OF COMPLETE POLYNOMIALS

We begin by studying the case of complete polynomials, as this is simpler to describe. The noncomplete case will be addressed in the next section where the ideas developed here will be extended.

3.1. A large but finite search. We shall perform the search degree by degree. Let d be the degree in which we shall seek a polynomial with a sparse square. The generic complete polynomial of degree d is just $f = \sum_{i=0}^d a_i x^i$ with the a_i being nonzero unknowns. By Lemma 1 we know that a complete polynomial of degree d exists if and only if there are nonzero values for the a_i which annihilate $\{b_j : j \in J\}$ for some index set J of cardinality at least $d + 1$. We choose to determine this by iterating over all the subsets of $\{b_0, b_1, \dots, b_{2d}\}$ having cardinality $d + 1$, and for each subset determining whether such nonzero values of the a_i exist.

The Weak Nullstellensatz tells us that a system of polynomial equations is insoluble if and only if the associated ideal contains the element 1: here “insoluble” means that there is no solution in the algebraic closure of the field generated by the coefficients of the polynomials. All systems which we shall consider contain only polynomials with integer coefficients, hence the Weak Nullstellensatz tells us about solvability in $\overline{\mathbb{Q}}$, the algebraic closure of the rationals. We can determine effectively if an ideal contains 1 by computing a Gröbner basis for it: the ideal contains 1 if and only if the Gröbner basis contains a constant (regardless of the term-ordering chosen).

We now explain what has to be done in each iteration. Let the set of indices to be considered this iteration be $J = \{j_1, j_2, \dots, j_{d+1}\}$. It is not sufficient to determine the solvability of the polynomial system $\{b_j : j \in J\}$ since we must also include a side condition to restrict the a_i to being nonzero. This side condition can be expressed in various ways, one of which is by including the equation $z \prod_{i=0}^d a_i - 1 = 0$ in the system, where z is a new indeterminate. The ideal associated to the system is

$$I = \text{Ideal}\left(b_{j_1}, b_{j_2}, \dots, b_{j_{d+1}}, z \prod_{i=0}^d a_i - 1\right).$$

All that remains is to compute a Gröbner basis of this ideal, and check whether it contains a constant. Unfortunately such a direct approach is too costly.

3.2. Refinements. The search strategy outlined above needs considerable refinement before the whole computation becomes tractable, most particularly the search in degree 11. The calculation can be accelerated in two ways: by reducing the number of subsets to consider or by speeding up the processing for each subset. Our ideas were guided by the suspicion that the ideal will almost always contain 1, thus our efforts are concentrated on this particular case.

Symmetries. To shorten our search we exploit three symmetries which preserve the number of terms in a polynomial and also in its square:

- A. $f \mapsto \alpha f$: we can multiply the polynomial f by a nonzero number α ;
- B. $f(x) \mapsto f(\alpha x)$: we can rescale the variable in f by a nonzero number α ;
- C. $f(x) \mapsto x^d f(1/x)$: we can reverse the order of the coefficients in f .

Reducing the number of subsets. A moment’s consideration suffices to see that subsets containing 0, 1, $2d - 1$ or $2d$ can be discarded immediately as none of the coefficients $b_0 = a_0^2$, $b_1 = 2a_0 a_1$, $b_{2d-1} = 2a_{d-1} a_d$ and $b_{2d} = a_d^2$ can ever be zero for nonzero a_i . Furthermore, we can eliminate roughly half the remaining combinations by using Symmetry C: we need to compute a Gröbner basis for only one of the subsets S and $S' = \{2d - s : s \in S\}$; though when S and S' coincide we save nothing.

Example 1. When $d = 11$, if we did not use Symmetry C, then we would have to consider $50388 = \binom{23-4}{12}$ subsets: f^2 has 23 coefficients but we ignore the four indices 0, 1, $2d - 1$ and $2d$. To see how many we must consider if we exploit Symmetry C, we need to know how often S and S' coincide. They coincide only when S is “palindromic”, so S must have six indices in between 2 and 10, and the other six arranged as a “mirror image about 11” between 12 and 20; thus there are $84 = \binom{9}{6}$ such subsets. Hence Symmetry C allows us to reduce the number of subsets to try to $25236 = \frac{1}{2} \left(\binom{23-4}{12} + \binom{9}{6} \right)$, i.e., just slightly more than half.

Speeding up each subset. The cost of computing a Gröbner basis generally increases rapidly as the number of indeterminates involved increases, thus any way to get rid of some indeterminates is potentially interesting. In fact, by exploiting Symmetries A and B we can reduce the number of indeterminates by two. We may assume that a_d and a_{d-1} are both equal to 1: firstly use Symmetry B to make the coefficients of x^d and x^{d-1} equal in f , and then use Symmetry A to reduce them both to 1. For this problem there does not seem to be any other obvious way of getting rid of further indeterminates.

One disadvantage of eliminating a_d and a_{d-1} as above is that the system of equations is no longer homogeneous, yet homogeneity is a valuable asset for Gröbner basis computations. Various ways of preserving homogeneity were tried experimentally, but none proved to be as fast as replacing a_d and a_{d-1} by 1.

Observe that the side condition (that the a_i be nonzero) led us to include a single polynomial of high degree in the system of equations to be solved, viz. $z \prod_{i=0}^d a_i - 1$. An alternative could be to include many degree 2 polynomials: $z_i a_i - 1$ for i running from 0 to $d-2$, but this involves using lots of extra indeterminates (namely z_0, \dots, z_{d-2}), and turned out to be much slower in practice.

Instead, a useful variation was to weaken the side condition so that just a_0 is restricted to being nonzero; that is replacing $z \prod_{i=0}^d a_i - 1 = 0$ by simply $z a_0 - 1 = 0$. This reduced computation time significantly, and in most cases this weaker condition was enough to prove insolubility. In those cases where the weaker condition was insufficient to rule out solutions, we then returned to the full side condition.

Our final refinement arose from the observation that each time we compute a Gröbner basis we are seeking a solution to a system of $d + 1$ polynomials in $d - 1$ indeterminates, and typically we find that the system is insoluble. Since we have only $d - 1$ indeterminates, we would not generally expect to be able to solve a system with more than $d - 1$ polynomials. Moreover, if we can prove that some d -tuple of the polynomials b_j is insoluble, then obviously any system containing that tuple must also be insoluble. Thus, with a single Gröbner basis computation we could dispose of the $d - 3$ subsets (of cardinality $d + 1$) which contain that particular d -tuple.

There remained the question of which d -tuples to select, and the answer hinged on two matters: the first was how to cover “efficiently” (i.e., without too much overlap) all the cases, and the second was that some d -tuples of polynomials led to very costly Gröbner basis computations. The solution adopted was to use a partial cover of cheap d -tuples of polynomials, and then to handle the uncovered subsets individually. Our measure of “cheapness” was derived empirically by studying the behaviour in degrees 9 and 10. The earlier reduction to the case $a_d = a_{d-1} = 1$ means that the top few coefficients, $b_{2d-2}, b_{2d-3}, b_{2d-4}$ and b_{2d-5} , are simpler than the others. The measure of cheapness is simply that subsets containing more of

these four polynomials are cheaper, and, among those containing the same number of these simple polynomials, the cheaper ones are those containing more polynomials with higher indices.

The partial cover was determined using a simple greedy algorithm. At the outset none of the subsets of cardinality $d + 1$ is marked as covered. First, we took those d -tuples which covered $d - 3$ unmarked subsets, starting with the cheapest d -tuples. As each d -tuple was selected, we marked as covered all the subsets containing it. When there was no further tuple covering $d - 3$ unmarked subsets, we then repeated the process but for tuples covering only $d - 4$ unmarked subsets, and so on until there was no d -tuple left which covered more than 2 subsets. At every stage, however, the most expensive d -tuples were excluded from consideration. Those subsets left unmarked at the end were tested individually. For example, in degree 11 we obtained a partial cover of about 3550 tuples which left about 3400 subsets uncovered; these figures should be compared with the total number of subsets, viz. about 25000 after exploiting Symmetry C.

Experimentally this approach worked well: only a small proportion of subsets were left uncovered by the partial cover, and almost all d -tuples in the cover chosen were quickly shown to be insoluble. A summary of times and number of Gröbner bases computed is given in Table 1 in Section 5. It should be noted that these times are quite sensitive to seemingly small changes in the measure of “cheapness”.

A single Gröbner basis. The vanishing of some subset of $\{b_2, b_3, \dots, b_{2d-2}\}$ of cardinality $d + 1$ is equivalent to the simultaneous vanishing of all products of $d - 3$ of these b_j . Based on this fact, a referee observed that the question could be resolved by computing a Gröbner basis of just a single ideal generated by $\binom{2d-3}{d-3}$ monomials in new symbols B_2, \dots, B_{2d-2} together with the polynomials $B_j - b_j$ for $j = 2, 3, \dots, 2d - 2$, and the side condition. Computationally this approach encounters difficulties even in degree 7: we stopped the computation after more than 30 hours’ CPU when the memory space in use exceeded 350 megabytes.

4. THE CASE OF NONCOMPLETE POLYNOMIALS

In the previous section we considered only “complete” polynomials with sparse squares. Similar techniques can be used to look for noncomplete polynomials, f , with sparse squares. Previously we searched through various distributions of zeroes amongst the coefficients of the square. Now we must search through distributions of zeroes amongst the coefficients of both f itself and f^2 . As previously, we conducted the search degree by degree. Most of the refinements we used before were adapted to this case—a notable exception being the use of (partial) covers. The reversal symmetry was used in two ways: to reduce the number zero patterns in f to be checked, and when the zero pattern in f was invariant under reversal then, as before, roughly half the zero patterns in f^2 could be eliminated.

In the case of complete polynomials we observed that we could exclude all index sets containing at least one of $0, 1, 2d - 1$, and $2d$. This technique can be generalized for noncomplete polynomials. We assume that the polynomial has a prescribed zero pattern, i.e., the polynomial is of the form $f = \sum_{s \in S} a_s x^s$ where the a_s are nonzero, and the support S is a fixed subset of $\{0, 1, \dots, d\}$. Viewing the a_s as transcendentals and f^2 as a polynomial in x , we classify the indices $t \in \{0, 1, \dots, 2d\}$ into three categories:

- (i) the coefficient of x^t in f^2 is identically zero;

- (ii) the coefficient of x^t in f^2 is of the form $a_{t/2}^2$ or $2a_i a_{t-i}$ for some i ;
- (iii) the rest.

Note that category (ii) represents terms which can never be zero whenever the a_s are nonzero, hence we can exclude from the search all subsets containing indices of this category. In practice, so that f is genuinely a polynomial of degree d , we insisted that the support contain both 0 and d ; furthermore, by Lemma 2 we need test only polynomials with support of cardinality 4 or more.

The classification of indices may easily be read off the square of $\hat{f} = \sum_{s \in S} x^s$. Let t be an index in the range 0 to $2d$, and let c_t be the coefficient of x^t in \hat{f}^2 . Then t is in category (i) whenever $c_t = 0$, t is in category (ii) whenever $c_t = 1$ or $c_t = 2$, otherwise t is in category (iii).

Example 2. Consider polynomials of degree 11 having zero coefficients for precisely x^3, x^4, x^7 and x^8 . Such polynomials have exactly 8 terms, so we want to know if the square could have 7 or fewer terms. Put $\hat{f} = 1 + x + x^2 + x^5 + x^6 + x^9 + x^{10} + x^{11}$. Now consider \hat{f}^2 : it has no zero coefficients, fourteen coefficients equal to 1 or 2, and nine others. Thus there are fourteen indices of category (ii), and hence the square of any polynomial having the given pattern of zero coefficients will have at least fourteen terms, and so the polynomial cannot possibly have a sparse square. Therefore no searching through subsets is needed in this case.

We ran the computer search as described here from degree 5 to degree 11. By Lemma 2 we know that there cannot be any noncomplete polynomial of degree less than 5 with a sparse square. In every case, the corresponding ideal contained 1, proving nonexistence of a polynomial with that support having a sparse square. In Section 5 we give a summary of the total number of Gröbner bases actually computed and the total time taken in each degree. The computations in degrees 5, 6 and 7 were very fast, and are excluded from the table.

5. TABLE OF RESULTS

Table 1 summarizes the statistics produced by our experiments. As already hinted above, the programs used for the “complete” and “noncomplete” cases were distinct; the latter being less refined. Therefore we distinguish in the table the data

TABLE 1

Degree	G-bases	Time
8 (complete)	118	6
8 (noncomplete)	290	10
9 (complete)	460	45
9 (noncomplete)	2300	69
10 (complete)	1800	2200
10 (noncomplete)	17000	1200
11 (complete)	7000	370000
11 (noncomplete)	119000	128000

given by the two different programs—this is indicated together with the degree. The column headed “Time” gives the CPU time in seconds on a 433MHz Digital Alpha running DEC Unix; the version of CoCoA used was 3.7. The column headed “G-bases” indicates the approximate number (to two significant digits) of Gröbner bases computed during the whole calculation.

6. CONCLUSION

Coppersmith and Davenport had already done the computations to show that no complete polynomial of degree less than 8 had a sparse square. We have now done them for degrees 8, 9, 10 and 11, and in every case have found that there was no solution, i.e., there is no complete polynomial in any of these degrees having a sparse square. Similarly, our search for noncomplete polynomials having a sparse square showed that none exists of degree less than 12.

Even though the underlying mathematics is rigorous, as is common with results dependent on lengthy computations, our conclusions must be tempered by the fact that our implementations have not been *proved correct* in any rigorous sense (i.e., correctness of processor, compiler, the CoCoA system, and our code). So, with the foregoing provisos, we have proved that there are no polynomials of degree less than 12 having sparse squares.

From our table of results it is quite evident that the computations in degree 11 account for almost all the total time, and it seems reasonable to conclude that a complete search in degree 12 using our techniques would still be prohibitively long. So, we leave open the matter of confirming that the only polynomials of degree 12 having sparse squares are the eight given in [2], up to the symmetries A, B and C. Indeed, their solutions are really four pairs related by the symmetry $f(x) \mapsto \alpha x^{12} f(\frac{1}{5x})$ where the scalar α depends on the choice of f .

We did conduct a search for complete polynomials of degree 12 lying in $\overline{\mathbb{F}}_p[x]$ having sparse squares for the prime $p = 29641$. The search found the images in $\mathbb{F}_p[x]$ of the polynomials given in [2], and no others. Unfortunately, we cannot deduce from this anything about the existence or nonexistence of other solutions of degree 12 in $\overline{\mathbb{Q}}[x]$.

REFERENCES

- [1] W Adams, P Lounstaunau, *An Introduction to Gröbner Bases*, Graduate Studies in Mathematics **3**, Amer. Math. Soc., Providence, 1994. MR **95g**:13025
- [2] D Coppersmith, J Davenport, “Polynomials whose powers are sparse” *Acta Arithmetica* **58** (1991), 79–87. MR **92h**:12001
- [3] A Capani, G Niesi, L Robbiano, *CoCoA: Computations in Commutative Algebra* <http://cocoa.dima.unige.it/>

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI GENOVA, ITALY
E-mail address: abbott@dim.unige.it