Polynomial-Time Algorithm for the Orbit Problem

R. KANNAN AND R. J. LIPTON

University of California, Berkeley, California

Abstract. The accessibility problem for linear sequential machines [12] is the problem of deciding whether there is an input x such that on x the machine starting in a given state q_1 goes to a given state q_2. Harrison shows that this problem is reducible to the following simply stated linear algebra problem, which we call the "orbit problem":

> Given (n, A, x, y), where n is a natural number and A, x, and y are $n \times n$, $n \times 1$, and $n \times 1$ matrices of rationals, respectively, decide whether there is a natural number i such that $A^i x = y$.

He conjectured that the orbit problem is decidable. No progress was made on the conjecture for ten years until Shank [22] showed that if n is fixed at 2, then the problem is decidable. This paper shows that the orbit problem for general n is decidable and indeed decidable in polynomial time. The orbit problem arises in several contexts; two of these, linear recurrences and the discrete logarithm problem for polynomials, are discussed, and we apply our algorithm for the orbit problem in these contexts.

Categories and Subject Descriptors: F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems—*computations on matrices*; *number-theoretic computations*; F.4.3 [**Mathematical Logic and Formal Languages**]: Formal Languages—*decision problems*; G.2.0 [**Discrete Mathematics**]: General

General Terms: Algorithms, Languages, Theory, Verification

Additional Key Words and Phrases: Linear sequential machines, linear transformations, matrix orbits

1. *Introduction*

A linear sequential machine M (LSM) is a transducer that functions as follows: At each time instant, M is in a state q among a set Q of its possible states. M receives an input over an alphabet Σ which is read one letter at a time. Depending on the current state q and the current letter read a, M outputs $\lambda(q, a)$, which is some letter of its output alphabet Δ and goes into state $\delta(q, a)$. Formally, a linear sequential machine M is a 5-tuple $(Q, \Sigma, \Delta, \delta, \lambda)$ such that there is a (finite or infinite) field F and integers n, k, and l such that $Q = F^n$, $\Sigma = F^k$, $\Delta = F^l$, and δ: $Q \times \Sigma \rightarrow Q$

A preliminary version of this paper appeared in Kannan, R., and Lipton, R. J. The orbit problem is decidable. In *Proceedings of the 12th Annual ACM Symposium on the Theory of Computing.* ACM, New York, 1980, pp. 252–261.

The work of R. Kannan was supported by the National Science Foundation under grants MCS 83-04770, MCS 81-05557, and MCS 77-09906. The work of R. J. Lipton was supported by the National Science Foundation under grant MCS 79-20409.

Authors' present addresses: R. Kannan, Department of Computer Science, Carnegie-Mellon University, Scheneley Park, Pittsburgh, Pa. 15213; R. J. Lipton, Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ.

and λ: $Q \times \Sigma \rightarrow \Delta$ are defined by suitably dimensioned matrices A, B, C, and D over F as follows:

$$\delta(q, a) = Aq + Ba; \qquad \lambda(q, a) = Cq + Da. \tag{1.1}$$

Harrison [12] wrote a comprehensive book on linear sequential machines. In that book, the *accessibility* problem for LSMs is the following: Given $M = (Q, \Sigma, \Delta, \delta, \lambda)$, A, B, C, and D defining δ and λ as in (1.1), and two states q_1 and q_2 in Q, determine whether there is an input that takes M from state q_1 to state q_2. Harrison shows that this problem is reducible to the *Orbit Problem*:

> Given A in $F^{n \times n}$, x, y in F^n, does there exist a nonnegative integer i such that $A^i x = y$?

We only consider the case in which $F = \mathcal{Q}$, the field of rational numbers. Shank's result for two dimensions also deals only with the field of rational numbers [22]. Although it would be nice to extend this to any computable field [18], the rational case provides us with so much structure that the extension does not seem automatic. We use the term *orbit* because the set $\{x, Ax, A^2x, A^3x, \ldots\}$ may be termed the orbit of x under A (with a slight abuse of notation). The main result of this paper is a polynomial-time algorithm to solve the orbit problem. As a consequence of this we are able to derive polynomial-time algorithms for the following two problems:

Problem 1. Given rationals $a_0, a_1, \ldots, a_{n-1}, r_0, \ldots, r_{n-1}$, and $y_1, \ldots, y_n$, consider the nth degree linear recurrence $r_{k+n} = \sum_{i=0}^{n-1} a_i \cdot r_{k+i}$ (for $k \geq 0$). Does there exist a k such that $r_{k+i} = y_i$ for $i = 0, 1, \ldots, n-1$? And if so, find such a k.

Problem 2. Given three polynomials $a(x)$, $b(x)$, and $d(x)$ over the rationals, does there exist a natural number i such that $(a(x))^i \equiv b(x) \pmod{d(x)}$? If so, find such an i.

The second problem is the polynomial analog of the well-known discrete logarithm problem for fields of prime characteristic [1]. Problem 1 is related to the long-standing open problem of determining whether a z rational sequence has a zero [19, 20]. The zero of a z rational sequence problem is the following: Given $a_0, a_1, \ldots, a_{n-1}, r_0, \ldots r_{n-1}$, determine whether there is a k for which r_k of the linear recurrence defined in Problem 1 is zero. It is not known whether this problem is decidable or not. We hope that some of the techniques of this paper can be extended to solve that problem. Finally, our solution of the orbit problem gives us as a special case a polynomial-time algorithm to solve the D0L ultimate growth equivalence [9]. In the course of solving the orbit problem we devise polynomial-time algorithms for finding the minimal polynomial of a matrix and for determining whether all the eigenvalues of a matrix are roots of unity. These simple by-products may be of independent value.

Our approach uses some number theory. To make the paper self-contained, we give below the definitions needed from number theory and algebra. For more details and proofs of statements, the reader is referred to Birkoff and MacLane [3] or Marcus [16].

An algebraic number is a root of a polynomial in $\mathcal{Q}[x]$, the ring of polynomials in the variable x over the rationals. An algebraic number is said to be an *algebraic integer* if it is the root of a monic polynomial with integer coefficients. For a square matrix A (algebraic number α), the minimal polynomial of A (of α) denoted $f_A(x)$

$(f_\alpha(x))$ is the least degree monic (i.e., with leading coefficient 1) polynomial in $\mathscr{Q}[x]$ such that $f_A(A) = O(f_\alpha(\alpha) = 0)$. For algebraic numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$, $\mathscr{Q}(\alpha_1, \ldots, \alpha_n)$ denotes the extension of the rationals by $\alpha_1, \ldots, \alpha_n$. $\mathscr{Q}(\alpha_1, \ldots, \alpha_n)$ is the field of all complex numbers that are equal to some polynomial in $\alpha_1, \ldots, \alpha_n$ with rational coefficients; it is called a number field.

Let I be the set of all algebraic integers. Then it is known that I is a ring. Thus for any number field F, $F \cap I$ is a ring. An *ideal* S of $F \cap I$ is a set satisfying the following conditions: S is a subgroup under addition, and $\alpha \in S$ and $\beta \in F \cap I$ together imply that $\alpha \cdot \beta \in S$. For any $\alpha \in F \cap I$ we define (α), *the ideal generated by* α, to be the smallest ideal of $F \cap I$ that contains α. Whereas the unique factorization theorem does not hold for number rings, it holds for ideals of number rings. To be more precise, an ideal S of $F \cap I$ is said to be a *prime ideal* if, whenever the product of two elements of $F \cap I$ belongs to S, one of them does too. (This is the analog of the definition of a prime number over the natural numbers—p is prime if, whenever it divides the product of two integers, it divides one of them.) For any two ideals S_1 and S_2, we define the product of S_1 and S_2, denoted $S_1 \cdot S_2$, to be the smallest ideal containing all products of the form $\alpha \cdot \beta$, where $\alpha \in S_1$ and $\beta \in S_2$. (In the case of ordinary integers, every ideal is *principal*, that is, it is the set of all multiples of some integer; thus, the product of two ideals (a) and (b) is precisely the ideal (ab); thus, the definition of product of ideals corresponds to ordinary product over the integers.)

The *fundamental theorem of ideal theory* (unique factorization theorem for ideals of a number ring) states that, in the domain of algebraic integers of any number field, every ideal can be expressed uniquely, except for order as the product of prime ideals. This will be useful for us later.

Let us now examine a plausible approach that tries to prove the decidability of the orbit problem by showing that some quantity associated with $A^i \cdot x$ grows with i and hence we can derive an upper bound on any i for which $A^i \cdot x = y$. Suppose for the moment that the roots of $f_A(x)$ are all distinct. Then it is known that A is diagonalizable, that is, that there exists a matrix S with complex entries such that SAS^{-1} is a diagonal matrix [3]. Call the diagonal matrix D. Then

$$A^i \cdot x = y \quad \text{implies} \quad S^{-1}(SAS^{-1})^i S \cdot x = y, \qquad \text{that is,} \quad D^i(Sx) = Sy.$$

Let x' be Sx and y' be Sy. Then $A^i x = y$ iff $(D_{jj})^i \cdot x'_j = y'_j$ for $j = 1, 2, \ldots, n$ since D is diagonal. Hence the problem is reduced to several problems of the form $\alpha^i = \beta$, where α and β are algebraic numbers. If α happens to be greater than 1 in absolute value, then clearly $|\alpha|^i$ monotonically increases with i and we can bound i. Similar reasoning holds when α is less than 1 in absolute value. If $|\alpha|$ equals 1 and α is a root of unity, then $\alpha^j = 1$ for some j and hence the only values of i to be checked are $1, 2, \ldots, j$. The real problem cases are when α has absolute value equal to 1 and is not a root of unity and when A is not diagonalizable. To handle all these cases without the use of cumbersome similarity transformations, etc., we use a natural relation between matrices and algebraic numbers.

We first outline here our method of attack. In the next section, the orbit problem is reduced to the *matrix power* problem:

> Given an $n \times n$ matrix A of rationals and a polynomial $q(x)$ with rational coefficients, does there exist a natural number i such that $A^i = q(A)$?

If now α is a root of the minimal polynomial of A, then $A^i = q(A) \Rightarrow \alpha^i = q(\alpha)$ (Theorem 2). We use this fact to solve our problem. A very brief outline of the solution is given below. We consider three cases:

If the minimal polynomial of A has a root α that is not an algebraic integer, then there is a prime ideal such that its maximum powers dividing the numerator and the denominator of α are different. This and the fact that the norm of any prime ideal is at least 2 lead to a bound on i.

If the minimal polynomial of A has a root that is an algebraic integer but not a root of unity, we use a theorem of Blanksby and Montgomery [5] that asserts that, if an algebraic integer α of degree n is not a root of unity, then it has a conjugate α' of magnitude at least $1 + [1/(30n^2 \log n)]$. $\alpha^i = q(\alpha)$ implies that $(\alpha')^i = q(\alpha')$. This enables us to get a bound on i.

If all the roots of the minimal polynomial A are roots of unity, we can determine exactly what the roots are. If there are no repeated roots, then i can be found by solving a system of linear congruences. If there are repeated roots, the situation is trickier: We argue that $A^i = q(A)$ iff $B^i = q(B)$ whenever B has the same minimal polynomial as A. We replace A by a matrix B which is the direct sum of Jordan blocks with elements from the splitting field of the minimal polynomial of A. The structure of B then enables us to solve the problem. Of course, this entails doing computations on algebraic numbers, which is easily accomplished by treating them as formal polynomials. It is known that the following can be done in polynomial time: finding the rank of a matrix of rationals, solving a system of linear equations over the rationals (e.g., [10] and [14]), and finding the greatest common divisor of polynomials [8]. These results are assumed in the paper.

1.1 NOTATION. $\mathscr{N}$, $\mathscr{Z}$, $\mathscr{Q}$ denote the set of natural numbers, integers, and rationals, respectively. For any field F, $F[x]$ denotes the ring of polynomials in the variable x with coefficients in F. If α is in a number field, the norm of α denoted $N(\alpha)$ is as usual the product of all its conjugates. For any polynomial $f(x) = f_0 + f_1x + f_2x^2 + \cdots + f_dx^d$ over the rationals, the length of the polynomial $f(x)$, denoted $|f|$, is the Euclidean length of the vector $(f_0, f_1, \ldots, f_d)$. For any field F, $F^{n \times n}$ denotes the set of all $n \times n$ matrices with entries in F. e_i denotes the ith unit vector in Euclidean n-space where n's value is clear from the context. (The ith unit vector has a 1 in the ith component and 0 everywhere else.) For any $n \times n$ matrix A, A^0 is the $n \times n$ identity matrix.

2. *Reduction to the Matrix Power Problem*

In this section we prove that the orbit problem, restated formally as Problem 3, is polynomial-time reducible to the matrix power problem (Problem 4), which is later stated in a slightly different form as Problem 5.

Problem 3. Given A in $\mathscr{Q}^{n \times n}$ and x, y in $\mathscr{Q}^n$, does there exist a nonnegative integer i such that $A^i x = y$?

Problem 4. Given A, D in $\mathscr{Q}^{n \times n}$, does there exist a nonnegative integer i such that $A^i = D$?

THEOREM 1. *The orbit problem* (*Problem* 3) *is polynomial-time reducible to the matrix power problem* (*Problem* 4).

PROOF. Let

$$\nu = \max l{:}x,\ A^1 \cdot x,\ A^2 \cdot x,\ \ldots,\ A^l \cdot x \quad \text{are linearly independent} \qquad (2.1)$$

and

$$C \text{ be the } n \times (\nu + 1) \text{ matrix } [x \,|\, Ax \,|\, \cdots \,|\, A^\nu x]. \qquad (2.2)$$

Note that since matrix multiplication and rank finding can be done in polynomial time, ν and C can be computed in polynomial time, given A and x. There are now two cases to consider.

Case 1: $\nu = n - 1$. Then C is an $n \times n$ invertible matrix and therefore

$$\begin{aligned} A^i x = y \quad & \text{iff} \quad A^i C = [y \mid Ay \mid A^2 y \mid \cdots \mid A^\nu y] \\ & \text{iff} \quad A^i = [y \mid Ay \mid \cdots \mid A^\nu y] C^{-1} = D \quad \text{(say).} \end{aligned} \tag{2.3}$$

Since D can be computed in polynomial time, we have completed the reduction in this case.

Case 2: $\nu \leq n - 2$. In this case we first reduce Problem 3 to a problem of the same form, but in $\nu + 1$ dimensions. Intuitively speaking, the column space of C—called the cyclic space generated by x under A—is where all of the $A^i x$ lie, and thus one is able to shift the entire scenario to this space, which is of dimension $\nu + 1$.

Let C' by an $n \times n$ matrix defined as follows: The first $\nu + 1$ columns of C' are the columns of C. The other $n - \nu - 1$ columns are chosen so that C' has rank n (this is clearly possible). Define

$$U = (C')^{-1}.$$

Then

$$U(A^i)x = e_{i+1} \qquad \text{for} \quad i = 0, 1, \ldots, \nu, \tag{2.4}$$

where the e_i denote the unit vectors (see notation).

Now we note that $A^i x = y$ is equivalent to $UA^iU^{-1}(Ux) = Uy$, which is equivalent to $(UAU^{-1})^i(Ux) = Uy$. Thus we may replace A by UAU^{-1}, x by Ux, and y by Uy; we assume henceforth that this replacement has been done. Thus we now know that for all i, $i = 0, 1, 2, \ldots,$

$$A^i x \text{ is an } n\text{-vector with 0 in the last } n - \nu - 1 \text{ coordinates} \ldots. \tag{2.5}$$

Thus the answer to the problem is *no*, if y had a nonzero entry among the last $n - \nu - 1$ coordinates, and in this case we are done. So assume this is not the case. We now argue that the problem can be reduced to a $\nu + 1$-dimensional problem by deleting all but the first $\nu + 1$ rows and columns of A and all but the first $\nu + 1$ entries in x and y. First, let B be the matrix obtained from A by setting all entries of its last $n - \nu - 1$ columns to zero. We show by induction on i that $A^i x = B^i x$ for all i. This is obvious for $i = 0, 1$. Assume it is true for i.

$$A^{i+1}x = A(A^i x) = A(B^i x) \quad \text{by the inductive assumption.}$$

But the last $n - \nu - 1$ entries of $B^i x$ are all zero; thus the product $A(B^i x)$ does not change if we alter the last $n - \nu - 1$ columns of A. Hence we have $A(B^i x) = B(B^i x)$, and the inductive proof is complete.

Now let P be the matrix obtained from B by setting the last $n - \nu - 1$ rows of B to zero. We prove by induction on i that $B^i x$ and $P^i x$ agree on the first $\nu + 1$ entries. This is clearly true for $i = 0, 1$. Suppose it is true for i. Let $B^i x = P^i x + v$, where v is a vector with zeros in the first $\nu + 1$ components. Then

$$B^{i+1}x = B(B^i x) = B(P^i x + v) = BP^i x$$

since Bv equals zero.

Now, since B and P differ in the last $n - \nu - 1$ rows only BP^ix equals $P^{i+1}x + v'$, where v' is a vector with zeros in the first $\nu + 1$ coordinates. This completes the inductive proof that B^ix and P^ix agree on the first $\nu + 1$ coordinates for all i.

Now we may proceed with the orbit problem as follows: Since we already know that the last $n - \nu - 1$ coordinates of A^ix and y are all zero for all i, we can replace A by P', the matrix consisting of the first $\nu + 1$ rows and columns of P, and x and y by vectors x' and y', consisting of their first $\nu + 1$ coordinates, respectively, and work on the $\nu + 1$-dimensional problem.

Thus we have shown a reduction of the orbit problem to the matrix power problem—either we have Case 1, by which the reduction is as shown above, or we have Case 2, by which the dimension of the problem is cut down by at least 1. So after at most $n - 1$ occurrences of Case 2, we must land in a one-dimensional problem solvable by inspection. We are tempted to conclude at this point that the reduction to the matrix power problem is polynomial time bounded. Unfortunately, we have only shown so far that the number of arithmetic operations—additions, subtractions, and multiplications—is bounded by a polynomial in the length of the input. We also have to argue that the size of all numbers involved (i.e., the number of bits needed to represent them) is bounded by a polynomial in the length of the input. This is far from obvious, since even if the magnitude of the numbers at most squares (i.e., the size doubles) every time we encounter Case 2, the size becomes exponential after n iterations. (For a general discussion of this topic see [14].) Fortunately, there is a simple reason why the numbers stay polynomially bounded in size in this reduction—we assert that Case 2 is encountered at most once. To see this, we show that for the P', x', and y' defined earlier, the cyclic space generated by x' under P is of dimension $\nu + 1$. By definition of ν the cyclic space generated by x under A has dimension $\nu + 1$, and indeed this space is spanned by the unit vectors $e_1, e_2, \dots, e_{\nu+1}$. This is of course the same as the cyclic space generated by x under B. Furthermore, since P^ix and B^ix agree on the first $\nu + 1$ coordinates for all i, the cyclic space generated by x' under P' has dimension $\nu + 1$, as claimed.

This implies that, when we apply the reduction recursively to the problem defined by P', x', y', we have Case 1.

We summarize the algorithm below.

procedure Reduction to matrix power problem

Comment. This procedure reduces the orbit problem (Problem 3) to the matrix power problem (Problem 4). The input is A, x, y of Problem 3, which are $n \times n$, $n \times 1$, and $n \times 1$, respectively.

```
do forever
    if $n = 1$, solve (1.1) and return the answer.
    if $x = 0$ then do the obvious.
    compute $\nu$ as defined in (2.1)
    if $\nu = n - 1$ then return $(A, D)$ where $D$ is as defined in (2.3)
    find $U$ such that $UA^ix = e_{i+1}$ for $i = 0, 1, 2, \dots, \nu$ (cf. 2.4)
    $A \Leftarrow UAU^{-1}$; $x \Leftarrow Ux$; $y \Leftarrow Uy$
    if the last $n - \nu - 1$ coordinates of $y$ are not all zero then return NO.
    $A \Leftarrow$ the $(\nu + 1) \times (\nu + 1)$ matrix containing the first $\nu + 1$ rows and columns of $A$. $x \Leftarrow$
    the vector of the first $\nu + 1$ components of $x$. Same for $y$
    set $n = \nu + 1$
end
```

This completes the proof of Theorem 1. □

Finally, we wish to reduce the matrix power problem (4) further to a modified version of it (Problem 5 below):

Problem 5. Given A in $\mathcal{Q}^{n \times n}$ and $q(x)$ in $\mathcal{Q}[x]$, determine whether there exists $i \in \mathcal{N}$ such that $A^i = q(A)$.

To see this reduction, note that if an instance of Problem 4 is to have a *yes* answer, D must be equal to $q(A)$ for some polynomial $q(x)$ in $\mathcal{Q}[x]$. We can assume without loss of generality that the degree of q is at most the degree of the minimal polynomial $m(x)$ of A, which is at most n. (This is because, for any polynomials $p(x)$, $r(x)$, if $p \equiv r \pmod{m(x)}$, then $p(A) = r(A)$). Thus we can solve the n^2 simultaneous equations $\sum_{j=0}^{n} q_j A^j = D$ in the variables $q_0, q_1, \ldots, q_n$. If there is no solution, then the answer to Problem 2 is *no*; otherwise, the problem is reduced to one of the form of Problem 5.

3. *Correspondence between Matrices and Algebraic Numbers*

THEOREM 2. *Let F by any field. Suppose A in $F^{n \times n}$ has minimal polynomial $p(x)$ in $F[x]$ and let $r(x)$, $q(x)$ be elements of $F[x]$. Then*

$$r(A) = q(A), \tag{3.1}$$

$$\Leftrightarrow r(x) \equiv q(x)\ (mod\, p(x)). \tag{3.2}$$

Further, if $F = \mathcal{Q}$ and $p(x)$ is irreducible over Q and has α as a root, then (3.1) *and* (3.2) *are equivalent to*

$$r(\alpha) = q(\alpha). \tag{3.3}$$

PROOF. Clearly to any $s(x)$ in $F[x]$ there corresponds a unique polynomial $s'(x)$ in $F[x]$ satisfying $s' \equiv s \pmod{p(x)}$ and degree s' less than degree p. By definition of p, $r(A) = r'(A)$ and $q(A) = q'(A)$. Thus

$$r(x) \equiv q(x) \pmod{p(x)} \Leftrightarrow r' = q' \Rightarrow r(A) = r'(A) = q'(A) = q(A).$$

Conversely, if $r(A) = q(A)$, then $p(x)$ divides $r(x) - q(x)$. □

For any algebraic number α, we denote by $f_\alpha(x)$ the monic irreducible polynomial in $\mathcal{Q}[x]$ satisfied by α and by n_α, the degree of this polynomial.

THEOREM 3. *There exists a polynomial $P(., ., .)$ such that for any algebraic number α that is not a root of unity and for any $q(x)$ in $\mathcal{Q}[x]$, if $\alpha^i = q(\alpha)$ for a natural number i, then i is at most $P(n_\alpha, \log|q|, \log|f_\alpha|)$. Further, if α is an sth root of unity, then either $I_s = \{i : \alpha^i = q(\alpha)\}$ is empty or $I_s = \{i_0 + zs : z \in \mathcal{Z}\}$, where i_0 is a fixed integer satisfying* $0 \le i_0 \le s - 1$.

PROOF. The case when α is a root of unity is obvious. Suppose that α is not a root of unity. If α is an algebraic integer, there is a conjugate θ of α such that $|\theta| \ge 1 + (1/(30n_\alpha^2 \log_e 6n_\alpha))$ [5]. Since θ is a conjugate of α, we have

$$\alpha^i = q(\alpha) \Leftrightarrow \theta^i = q(\theta) \Rightarrow i \le \frac{\log|q(\theta)|}{\log|\theta|}, \qquad \log|q(\theta)| \le \log((n+1)|\theta|^n|q|),$$

because $|\theta| \ge 1$.

Thus

$$\alpha^i = q(\alpha) \Rightarrow i \le n + \left\lceil \frac{\log(n+1) + \log|q|}{\log|\theta|} \right\rceil$$
$$\le n + [(60n^2 \log_e(6n))(\log(n+1) + \log|q|)]$$
$$\le p(n, \log|q|, log|f_\alpha|) \quad \text{(say)},$$

where p can be chosen to be a polynomial.

We now consider the case in which α is not an algebraic integer. (In particular, it is not a root of unity.) Let $\alpha_1, \alpha_2, \beta_1, \beta_2$, be algebraic integers such that $\alpha = \alpha_1/\alpha_2$ and $q(\alpha) = \beta_1/\beta_2$. Then $\alpha^i = q(\alpha) \Leftrightarrow \alpha_1^i\beta_2 = \alpha_2^i\beta_1$. Since α is not an algebraic integer, the ideals (α_1) and (α_2) are not equal. (The ideals (α_1) and (α_2) will be equal if and only if there is unit γ such that $\alpha_1 \cdot \gamma = \alpha_2$, but this cannot happen here since α_1/α_2 is not an algebraic integer.) Thus there is a prime ideal $\mathscr{P}$ such that $\mathscr{P}^{l_1} \| (\alpha_1)$, $\mathscr{P}^{l_2} \| (\alpha_2)$ and $l_1 \ne l_2$. (Here $\mathscr{P}^l \| (\alpha)$ means that $\mathscr{P}^l$ divides (α) and $\mathscr{P}^{l+1}$ does not.) Suppose further that $\mathscr{P}^{l_3} \| (\beta_2)$, $\mathscr{P}^{l_4} \| (\beta_1)$. Then $\alpha_1^i\beta_2 = \alpha_2^i\beta_1 \Rightarrow il_1 + l_3 = il_2 + l_4$ (by the unique factorization theorem for ideals of a number ring). Assume without loss of generality that l_2 is greater than l_1. Then

$$i = \frac{l_3 - l_4}{l_2 - l_1} \le \frac{l_3}{l_2 - l_1} \le l_3.$$

Thus we need to show only that l_3 cannot be too big. Since $\mathscr{P}^{l_3}$ divides (β_2), it follows that the norm of the ideal $(\mathscr{P})^{l_3}$ divides the norm of the ideal (β_2). But the norm of $\mathscr{P}$ is at least 2; thus $l_3 \le \log_2 N((\beta_2))$. $B\alpha$ is an algebraic integer for some rational integer B with $B \le |f_\alpha|$ [16]. Thus we can choose $\alpha_2 = B$; $\beta_2 = B^n$ and apply the above argument. We then have $\alpha^i = q(\alpha) \Rightarrow i \le n \cdot \log_2|f_\alpha|$. Taking $P(.,\ .,\ .)$ to be the maximum of the polynomials in the two cases, we have Theorem 3. □

Remark. The proof of Theorem 3 is as short as it is only because the remarkable result of Blanksby and Montgomery [5] is available to us. The result is a substantial strengthening of a result of Kronecker's [15], which showed that, if an algebraic integer is not a root of unity, then it has a conjugate of absolute value greater than 1. It was improved by Ore [17]. Schinzel and Zassenhaus [21] showed that, if α is an algebraic integer of degree n over $\mathscr{Q}$ and is not a root of unity, then it has a conjugate θ of absolute value at least $1 + (c/2^n)$, c a constant. Subsequent strengthening by Blanskby [4] increased the right-hand side to

$$1 + \left(\frac{c}{(\sqrt{2} + \epsilon)^n}\right).$$

We note that none of these earlier bounds would have sufficed to give a polynomial running time bound on our algorithm for the orbit problem since they do not yield Theorem 3.

4. *Polynomial Algorithm for Problem* 5

Given an instance A, $q(x)$ of Problem 5, we first find the minimal polynomial $f_A(x)$. The following obviously polynomial-time algorithm does the job, though not in the most efficient manner.

```
procedure MIN-POLY (A, n)
find $A^2, A^3, \ldots, A^n$
  for $i = 1$ step 1 until $n$ do:
    if the linear system of $n^2$ equations $\sum_{j=0}^{i} y_j A^j = 0$ has a solution $y = (y_0, y_1, \ldots, y_i)$ in
    the rationals with $y_i \neq 0$ then return $f_A(x) = \sum_{j=0}^{i} y_j x^j$
  end
end procedure
```

Thus the procedure returns the minimum-degree polynomial, which must obviously be a scalar multiple of the minimum polynomial. The procedure runs in polynomial time because the solution of simultaneous linear equations can be found in polynomial time.

The next procedure decides whether $f_A(x)$ has only roots of unity as its eigenvalues. This will be necessary for our algorithm.

```
procedure ROOTS OF UNITY
initialize $h_j(x) \leftarrow 1$ for $j = 1, 2, \ldots, (\text{degree}(f_A))^2$
  $f'_A(x) \leftarrow f_A(x)$
  Comment. At the end, $h_j(x)$ will have as its roots all the $j$th primitive roots of unity that
  $f_A(x)$ has. $f'_A(x)$ has all the roots of $f_A(x)$ that are not roots of unity as its roots.
for $j = 1$ step 1 until $(\text{degree}(f_A))^2$ do:
  $h_j(x) \leftarrow \text{GCD}(f'_A(x), (x^j - 1)^{\text{degree}(f_A)})$
  $f'_A - f'_A/h_j$
end
end procedure
```

LEMMA 1. *$f_A(x)$ the minimal polynomial of A has only roots of unity as its eigenvalues if and only if $f'_A(x)$ has degree zero at the end of the above procedure. Further, at the end of the procedure, $h_j(x)$ equals $(C_j(x))^{k_j}$ for some nonnegative integers k_j, where C_j is the jth cyclotomic polynomial (irreducible monic polynomial in $\mathcal{Q}[x]$ with $\exp(2\pi\sqrt{-1}/j)$ as a root.)*

PROOF. Let d be the degree of f_A. If the jth primitive root of unity is also a root of f_A, then clearly $C_j(x)$ divides f_A. Thus d is at least $\phi(j)$, the Euler ϕ function. From elementary number theory we get a crude bound $\phi(j) \geq \sqrt{j}$ (see e.g., [2]). Hence, if a jth primitive root of unity is a root of f_A, then $d \geq \sqrt{j}$. Further, the multiplicity of any root of f_A is at most d. Thus at the end of the procedure f'_A contains no roots of unity, but contains all other roots of f_A. This establishes the first statement of Lemma 1. The second statement follows from the fact that, when the algorithm finds $h_j(x)$, the only complex numbers that are roots of both f'_A and $(x^j - 1)$ are the primitive jth roots of unity.

At the conclusion of the last procedure we know which of the following three cases we are in and we handle the problem accordingly.

Case 1. There is an α, a nonroot of unity such that $f_A(\alpha) = 0$. Then by Theorem 2, $A^i = q(A) \Rightarrow \alpha^i = q(\alpha)$. By Theorem 3, i is at most $P(n_\alpha, \log|q|, \log|f_\alpha|)$. We try the polynomially many possible values of i to check whether any of them satisfy the matrix equation $A^i = q(A)$.

Case 2. $f_A(x) = \prod_{j=0}^{n} (C_j(x))^{k_j}$, where each k_j is 0 or 1 (no repeated roots). In this case

$$\begin{aligned} A^i = q(A) &\Leftrightarrow x^i \equiv q(x) \pmod{f_A(x)} && \text{(by Theorem 2)} \\ &\Leftrightarrow x^i \equiv q(x) \pmod{C_j(x)} && \text{for all } j \text{ such that } k_j = 1 \\ & && \text{(by Chinese remainder theorem).} \end{aligned}$$

Since the $C_j(x)$ for each j such that $k_j = 1$ is known from the procedures, we may find for each such j the unique i_j between zero and j (if one exists) such that $x^{i_j} \equiv q(x) \pmod{C_j(x)}$. If none exists, of course, we can stop after returning a negative answer to Problem 5. Clearly now, $A^i = q(A)$ iff $i \equiv i_j \pmod{j}$ for all j with k_j equal to 1. The answer to problem (5) can now be obtained by solving a system of simultaneous congruences, which can be done in polynomial time [23].

Case 3. $f_A(x) = \prod_{j=1}^{n} (C_j(x))^{k_j}$, where at least one k_j is 2 or greater. The proof that this last case can be handled in polynomial time is the subject of the rest of this section.

THEOREM 4. *Suppose A is an $n \times n$ matrix with only roots of unity as its eigenvalues, and suppose $q(x)$ is any polynomial with rational coefficients. Suppose also that the minimal polynomial $f_A(x)$ of A equals $\prod_{j=1}^{n} (C_j(x))^{k_j}$, where the C_j are cyclotomic polynomials. For each j such that $k_j \geq 1$, define a $k_j \times k_j$ matrix B_j as follows (with $\omega_j = exp(2\pi\sqrt{(-1)}/j)$):*[1]

$$B_j = \begin{bmatrix} \omega_j & 1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & \omega_j & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & \cdot & \cdot & 0 & \cdot & 0 \\ \cdot & \cdot & & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & & & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & \omega_j \end{bmatrix}. \tag{4.1}$$

Then,[2]

$$A^i = q(A) \Leftrightarrow \binom{i}{s-r} \omega_j^{i-(s-r)} = (q(B_j))_{r,s}, \tag{4.2}$$

for all j with $k_j \geq 1$ and for all s, r, with $k_j \geq s \geq r \geq 1$.

PROOF. $A^i = q(A) \Leftrightarrow x^i \equiv q(x) \pmod{C_j(x)^{k_j}} \forall j$ by the Chinese remainder theorem. Considering x^i and $q(x)$ as polynomials in $(\mathcal{Q}(\omega_j))[x]$, we have $x^i = q(x) (\bmod\, C_j(x)^{k_j}) \Rightarrow x^i \equiv q(x) \pmod{(x - \omega_j)^{k_j}}$. Conversely, if $(x - \omega_j)^{k_j}$ divides $x^i - q(x)$ in $(\mathcal{Q}(\omega_j))[x]$, then $p(x)$ would divide $(x^i - q(x))$ in $\mathcal{Q}[x]$ where p is the least degree polynomial in $\mathcal{Q}[x]$ that $(x - \omega_j)^{k_j}$ divides. Since $p(x) = C_j(x)^{k_j}$, we have shown

$$x^i \equiv q(x) \pmod{C_j^{k_j}} \Leftrightarrow x^i \equiv q(x) \pmod{(x - \omega_j)^{k_j}}. \tag{4.3}$$

It is a standard fact that the matrix B_j has minimal polynomial $(x - \omega_j)$. (To see this, let B_j equal $\omega_j I + M$ where M is strictly upper triangular. It is easily checked that $M^{k_j} = 0$ and that no lower power of M is zero. Thus $(B_j - \omega_j I)^{k_j}$ and no lower power of $B_j - \omega_j I$ is zero. This proves the assertion.) Applying Theorem 2 with $F = \mathcal{Q}(\omega_j)$, we see that (4.3) is equivalent to $B_j^i = q(B_j)$. To sum up, we have so far proved

$$A^i = q(A) \Leftrightarrow B_j^i = q(B_j) \qquad \text{for} \quad j : k_j \neq 0. \tag{4.4}$$

The lemma below explicitly calculates the entries of B_j^i and completes the proof of Theorem 4.

[1] Matrix (4.1) is called a Jordan block.

[2] (v^u) here stands for $u!/(v!(u - v)!)$ if $u \geq v$ and zero otherwise.

LEMMA 2. *For $k_j \geq s \geq r \geq 1$,*

$$(B_j^i)_{r,s} = \binom{i}{s-r}(\omega_j)^{i-(s-r)}.$$

The entries below the diagonal of B_j^i are zero.

PROOF. B_j can be considered to be the incidence matrix of a weighted directed graph G on k_j vertices: $v_1, v_2, \ldots, v_{k_j}$, where $(B_j)_{r,s}$ is the weight on the edge from vertex v_r to vertex v_s of G.

Then from basic graph theory (e.g. [6]), we get

$(B_j^i)_{r,s} = \sum_{P: P \text{ is a path of length } i \text{ from } v_r \text{ to } v_s}$, product of the weights of the edges of P.

For $s \geq r$ any path from v_r and v_s of length i must consist of $s - r$ edges of weight 1. The position of these $s - r$ edges uniquely determines the path. Thus there are

$$\binom{i}{s-r}$$

such paths. Each path has product of weights equal to $\omega_j^{i-(s-r)}$, proving the lemma and hence the theorem. □

We now handle case 3 as follows: Suppose that k_j is at least 2 (note that this must be true of some j in this case). With $s = 2$ and $r = 1$ in the theorem, we get $A^i = q(A) \Rightarrow i \cdot \omega_j^{i-1} = q(B_j)_{1,2}$. We first compute $(q(B_j))_{1,2}$. More precisely, we compute a polynomial $t(x)$ with rational coefficients such that $t(\omega_j) = (q(B_j))_{1,2}$. Now by Theorem 2,

$$i \cdot \omega_j^{i-1} = t(\omega_j) \Leftrightarrow i \cdot x^{i-1} = t(x) \pmod{C_j(x)}. \tag{4.5}$$

We now compute polynomials $P_1(x), P_2(x), \ldots, P_j(x)$ such that the degree of each of these polynomials is less than the degree of $C_j(x)$ and

$$x^l \equiv P_l(x) \pmod{C_j x}. \tag{4.6}$$

(Of course, $P_l(x) = x^l$ for $l < \phi(j) = \text{degree}(C_j)$.)

Next we find

$$S = \left\{l : \frac{t(x)}{P_l(x)} \text{ is an integer}\right\}. \tag{4.7a}$$

Further define

$$S' = \left\{l' : \exists l \in S \text{ such that } l' \equiv l \,(\text{mod}\, j)) \text{ and } \frac{t(x)}{P_l(x)} = l'\right\}. \tag{4.7b}$$

Note that $|S'|$ is at most $|S|$, which is at most j.

LEMMA 3. *Let S' be as defined in the last paragraph. Then $A^i = q(A)$ implies that i belongs to S'.*

PROOF. Modulo $C_j(x)$ there are exactly j distinct powers of x and these are x^1, $x^2, \ldots, x^j$, which equal to $P_1, P_2, \ldots, P_j$, respectively. Now suppose for some i, $A^i = q(A)$, and suppose that $i \equiv l \pmod{j}$ with $0 < l < j + 1$. Then $x^i \equiv P_l(x)$ $(\text{mod}\, C_j(x))$ and by (4.5) l belongs to S and i belongs to S'.

At this point, we are tempted to conclude that we are done: There are only polynomially many elements in S'. For each i in S' we could find A^i by repeated multiplication and check whether any one of them equals $q(A)$. Unfortunately, this is not polynomial time bounded since only the number of bits of i (not necessarily its magnitude) can be bounded by a polynomial in the length of the input. This suggests that perhaps we can find A^i by successive squaring instead of successive multiplications by A. Unfortunately, here again we encounter a problem—the number of bits needed to represent the square of a matrix may be double that needed to represent the matrix. Thus, the number of bits needed to represent $A^{i/2}$, for example, may be $i/2$ times the same number for A. This would again render the algorithm nonpolynomial time bounded. We solve this problem by appealing to (4.2). Instead of checking whether A^i equals $q(A)$, we check whether

$$\binom{i}{s-r}\omega_j^{i-(s-r)} = (q(B_j))_{r,s}. \tag{4.8}$$

We find S', which is clearly polynomial time computable. Next, for each candidate i in S', we check whether $A^i = q(A)$ by making use of (4.2). To this end we compute polynomials $t_{r,s}(x)$ for $1 \le r,\ s \le j$, each of degree less than that of C_j such that $(q(B_j))_{r,s} = t_{r,s}(\omega_j)$ for all r, s.

Then compute integers $i_{r,s}$ such that $1 \le i_{r,s} \le j$ and $i_{r,s} \equiv i - (s - r) \pmod{j}$. Clearly,

$$T_{r,s}(\omega_j) = \binom{i}{s-r} \cdot \omega_j^{i_{r,s}} \tag{4.9}$$

is the left-hand side of (4.8). The right-hand side of (4.9) is

$$\binom{i}{s-r} \cdot P_{i_{r,s}}(\omega_j).$$

Thus it suffices to compute the polynomial

$$T_{r,s} = \binom{i}{s-r} P_{i_{r,s}}(x)$$

and check whether $T_{r,s}$ equals the polynomial $t_{r,s}$ for each r, s. Obviously, this can be done in polynomial time.

We summarize the entire discussion by an algorithm for solving Problem 5.

procedure Problem 5

Comment. The procedure takes as input an $n \times n$ matrix A and a polynomial $q(x)$ and checks whether there is a nonnegative integer i such that $A^i = q(A)$.

Call procedures MIN-POLY and ROOTS OF UNITY.

If A has an eigenvalue other than a root of unity, then check for all i's up to the polynomial bound given by Theorem 3 whether $A^i = q(A)$; if there is a positive answer return YES, else return NO.

If A has no repeated roots, then solve the congruences described in Case 2 of Section 4 and return the answer.

```
choose any $j$ such that $k_j$ is at least two (cf., Lemma 1)
find $P_l(x)$ for $l = 1, 2, \ldots, j$ (cf., (4.6))
find the sets $S$, $S'$ (cf., (4.7))
if $S'$ is empty then return NO
For each element $i$ of $S'$ do:
  Call CHECK($i$)
  if CHECK($i$) returns YES then return YES
  end
  Return NO
  end procedure
```

```
procedure CHECK($i$)
For each $j$ such that $k_j$ is nonzero (cf., Lemma 1) do:
    Compute the polynomials $P_l(x)$ (4.6) $t_{r,s}(x)$ and $T_{r,s}(x)$ (cf., (4.9) and Lemma 3)
if for some $r$, $s$, $T_{r,s} \neq t_{r,s}$ then return NO
end
Return YES
end procedure
```

5. *Related Problems*

Consider Problem 1, the question of whether a linear recurrence of degree n ever reaches a fixed set of n values. Using the notation introduced in the definition of Problem 1, we define $r^{(k)}$ to be the n-column vector $(r_k, r_{k+1}, \ldots, r_{k+n-1})^t$. Then clearly $r^{(k+1)} = A \cdot r^{(k)}$, where A is a suitable $n \times n$ matrix. Now it is clear that the linear recurrence problem (Problem 1) has been rewritten as an orbit problem and thus can be solved in polynomial time.

To solve Problem 2, the polynomial discrete algorithm problem, first suppose that, given a polynomial $d(x)$, we can find a matrix M_d such that its minimal polynomial equals $d(x)$. Then by Theorem 2,

$$(a(x))^i \equiv b(x) \pmod{d(x)} \Leftrightarrow (a(M_d))^i = b(M_d).$$

We may of course find $a(M_d)$, $b(M_d)$ given M_d and solve the matrix power problem (Problem 4) in polynomial time. It remains to find M_d given d. The standard notion of a companion matrix of a polynomial comes to our rescue. (see, e.g., [13, sect. 6.7]). The companion matrix M_d of a monic polynomial $d(x) = d_0 + d_1x + d_2x^2 + \cdots + x^n$ is the $n \times n$ matrix

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -d_0 & -d_1 & -d_2 & \cdots & -d_{n-1} \end{bmatrix}.$$

This matrix has as its minimal polynomial $d(x)$ and is definitely easy to construct from d. Note further that the assumption that $d(x)$ is monic is not restrictive, since multiplying it through by a rational does not alter the discrete logarithm problem. Thus we have a polynomial-time algorithm for this problem.

It is clear that the zero of a $\mathscr{Z}$-rational function problem mentioned in the introduction can be rephrased as follows:

> Given an $n \times n$ matrix A of rationals, an n-vector x, and linearly independent vectors v_1, v_2, v_{n-1}, does there exist a natural number i such that A^ix belongs to the vector space V spanned by v_1, v_2, v_{n-1}?

The orbit problem replaces the $n - 1$ dimensional space V with a point. With some effort one should be able to extend the algorithm, hopefully with a polynomial time bound, to the case in which V is a one-dimensional subspace. The cases in which V has dimensions 2 or 3 seem harder, but an approach to solving the long-standing zero of a $\mathscr{Z}$-rational function problem is to increase the dimension of the vector space V progressively, of course aiming for decidability (not a low complexity).

ACKNOWLEDGMENTS. The authors thank Professor R. Karp for his generous support and Professor M. Harrison for bringing this problem to our attention. We also thank the referee whose ideas have substantially improved the presentation.

REFERENCES

(Note: References [7] and [11] are not cited in text.)

1. ADELMAN, L. A subexponential algorithm for the discrete logarithm problem. In *Proceedings of the 20th IEEE Symposium on the Foundations of Computer Science.* IEEE, New York, 1979, pp. 55–60.
2. APOSTOL, T. M. *Introduction to Analytic Number Theory.* Springer-Verlag, Berlin, 1976.
3. BIRKOFF, G., AND MACLANE, S. *A Brief Survey of Modern Algebra* 2nd ed. MacMillan, New York, 1965.
4. BLANKSBY, P. E. A note on algebraic integers. *J. Number Theory 1* (1969), 155–160.
5. BLANKSBY, P. E., AND MONTGOMERY, H. L. Algebraic integers near the unit circle. *Acta Arith.* (1971), 355–369.
6. BONDY, J. A., AND MURTY, U. S. R. *Graph Theory with Applications.* North Holland, New York, 1976.
7. COHN, H. *A Classical Invitation to Algebraic Numbers and Class Fields.* Springer-Verlag, New York, 1978.
8. COLLINS, G. E. Subresultants and reduced polynomial remainder sequences. *J. ACM 14*, 1 (1967), 128–142.
9. CULIK, K. Homomorphisms, decidability, equality and test sets. In *Formal Language Theory: Perspectives and Open Problems*, R. Book, Ed. Academic Press, Orlando, Fla., 1980.
10. EDMONDS, J. Systems of distinct representatives and linear algebra. *J. Res. NBS 71 B* (1971), 241–245.
11. FRUMKIN, M. A. Polynomial time algorithms in the theory of linear diophantine equations. In *Fundamentals of Computation Theory* (*New York*), M. Karpinsky, Ed. Lecture Notes in Computer Science, vol. 56. Springer-Verlag, New York, 1977, pp. 386–392.
12. HARRISON, M. *Lectures on Sequential Machines.* Academic Press, Orlando, Fla., 1969.
13. HERSTEIN, I. N. *Topics in Algebra*, 2nd ed. Xerox College Publishing, Lexington, Mass., 1975.
14. KANNAN, R. The size of numbers in the analysis of certain algorithms. Ph.D. dissertation, Operations Research Dept., Cornell Univ., Ithaca, N.Y., 1980.
15. KRONECKER, L. Zwei sätze über gleichungen mit ganzzahligen Koeffizienten. *J. Reine Angew. Math. 53* (1875), 173–175.
16. MARCUS, D. A. *Number Fields.* Springer-Verlag, Berlin, 1977.
17. ORE, O. *Les corps algebriques et la theorie des ideaux.* Hermann and Co., Paris, 1934.
18. RABIN, M. O. Computable algebra, general theory and theory of computable fields. *Trans. AMS 87* (1960), 341–360.
19. ROZENBERG, G., AND SALOMAA, A. The mathematical theory of L-systems. In *Advances in Information Systems Science*, vol. 6, J. Tou, Ed. Plenum, New York, 1976, pp. 186–187.
20. RUOHONEN, K. Zeros of $\mathscr{Z}$-rational functions and D0L-equivalence. *Theor. Comput. Sci. 3* (1976), 282–292.
21. SCHINZEL, A., AND ZASSENHAUS, H. A refinement of two theorems of Kronecker. *Mich. Math. J. 12* (1965), 81–84.
22. SHANK, H. S. The rational case of a matrix problem of Harrison. *Discrete Math. 28* (1979), 207–212.
23. VON ZUR GATHEN, J., AND SIEVEKING, M. Komplexität von Entscheidungsproblemen—ein Seminar. In E. Specker and V. Strassen, Eds. Lecture Notes in Computer Science, vol. 43. Springer-Verlag, New York, 1976, pp. 49–71.

RECEIVED DECEMBER 1983; REVISED JANUARY 1986; ACCEPTED FEBRUARY 1986