Mathematical Programming 57 (1992) 121-143
North-Holland

Polynomial algorithms for LP over a subring of the algebraic integers with applications to LP with circulant matrices

Ilan Adler and Peter A. Beling

Department of Industrial Engineering and Operations Research, University of California, Berkeley, USA

Received 15 May 1991
Revised manuscript received 20 September 1991

We show that a modified variant of the interior point method can solve linear programs (LPs) whose coefficients are real numbers from a subring of the algebraic integers. By defining the encoding size of such numbers to be the bit size of the integers that represent them in the subring, we prove the modified algorithm runs in time polynomial in the encoding size of the input coefficients, the dimension of the problem, and the order of the subring. We then extend the Tardos scheme to our case, obtaining a running time which is independent of the objective and right-hand side data. As a consequence of these results, we are able to show that LPs with real circulant coefficient matrices can be solved in strongly polynomial time. Finally, we show how the algorithm can be applied to LPs whose coefficients belong to the extension of the integers by a fixed set of square roots.

Key words: Linear programming, polynomial-time algorithms, strongly polynomial-time algorithms, circulant matrices, algebraic numbers.

1. Introduction

The question of whether a linear programming problem can be solved in polynomial time was answered in a landmark paper by Khachiyan in 1979. In fact, both Khachiyan's ellipsoid method [10] and Karmarkar's interior point method [9] solve linear programs (LPs) with rational coefficients in time that is polynomial in the number of input coefficients and the total number of bits in a binary encoding of the problem data. Nevertheless, several interesting questions concerning the complexity of linear programming remain open. One of the main open questions is usually stated as: Is there a strongly polynomial algorithm for linear programming? Following standard usage (cf. [19]), we say an algorithm for linear programming is *strongly polynomial* if:

(S1) it consists of the elementary operations addition, subtraction, multiplication, division, and comparison;

(S2) the total number of elementary operations performed by the algorithm is polynomial in the number of input items (i.e., the number of coefficients in the

Correspondence to: Dr. Ilan Adler, Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA.

matrices and vectors that define the linear program); and

(S3) *when applied to a rational instance*, the (binary encoding) size of the numbers that occur during the course of the algorithm is bounded by a polynomial in the size and total number of input items.

The conditional nature of (S3) stems from differences between rational-number and real-number models of computation. The standard model of computation for rational numbers is derived from the Turing machine. Consequently, the time required to perform an elementary arithmetic operation (S1) depends on the bit size of the operands. Requirement (S3) ensures that the time the algorithm spends performing any intermediate calculation is polynomial in the input size. In the case of real numbers, linear programming is usually modeled in terms of a machine that can perform any of the elementary operations (S1) in constant time, regardless of the nature of the operands. (See [4] for a rigorous treatment of general computation with real numbers.)

Taking advantage of the dichotomy implied by (S3), we can split the question of the existence of a strongly polynomial algorithm into two easier questions:

(A) If the data is rational, does there exist an algorithm satisfying (S1), (S2), and (S3)?

(B) If the data is real, does there exist an algorithm satisfying (S1) and (S2)?

To date, efforts to find a strongly polynomial algorithm for linear programming follow one of two main approaches, distinguished by whether they are directed at question (A) or at question (B). Those efforts directed at question (A) involve modifying existing polynomial-time algorithms, such as the ellipsoid method or variants of the interior point algorithm, so that their running times become independent of the size of at least part of the input data. In a key result along these lines, Tardos [19] showed that a LP can be solved in time that is independent of the size of the data in the objective and right-hand side vectors. As a consequence, LPs in which the coefficient matrix has "small" rational entries, such as those that arise frequently in combinatorial optimization, can be solved in strongly polynomial time. Recently, Norton, Plotkin, and Tardos [16] extended Tardos' results by giving an algorithm whose running time is independent of the size of the data in a fixed number of rows or columns of the coefficient matrix.

Before discussing work directed at question (B), we mention some issues concerning the existing polynomial algorithms for rational LPs. Both the ellipsoid and interior point methods depend in a fundamental way on upper and lower bounds on the magnitude of certain numbers related to basic solutions of the LP. These bounds allow the algorithms to be properly initiated and terminated, and are themselves part of the theoretical complexity of the algorithms. If the problem is rational, the bounds are a function of the bit size of the problem data and can be computed in polynomial time. If the problem is not rational, it is still possible to compute the necessary upper bounds in polynomial time, but no polynomial method for computing the lower bounds is known [13].

The second approach toward finding a strongly polynomial algorithm for linear programming focuses on answering question (B) for special classes of LPs. Considering the discussion above, it is not surprising that these efforts generally involve the construction of algorithms that are greatly different from the existing polynomial algorithms for rational data. The work of Megiddo provides important examples of this second approach; in [11] a strongly polynomial algorithm is given for feasibility problems in which at most two variables appear in each inequality, and in [12] a strongly polynomial algorithm is given for problems in which the number of variables is fixed. Interestingly, the combination of the ideas in this latter algorithm with those in [19] led to the algorithm in [16].

In this paper, we show that linear programs whose coefficient matrices are circulant can be solved in strongly polynomial time. (LPs of this kind are related to discrete convolution and arise frequently in image processing.) In proving this result we are led to an analysis of polynomial-time algorithms for linear programming in a subring of the algebraic integers. Specifically, we show that a variant of the interior point method can solve LPs whose coefficients are real members of the set

$$W_p = \left\{ \alpha \colon \alpha = \sum_{j=0}^{p-1} a_j \omega^j;\ a_j \text{ integer } \forall j \right\},$$

where $\omega = e^{2\pi i/p}$ is the first primitive pth root of unity. (We lose no generality by working with the subring W_p instead of the subfield $\Omega_p = \{\gamma \colon \gamma = \sum_{j=0}^{p-1} q_j\omega^j;\ q_j$ rational $\forall j\}$.) The restriction to W_p allows us to define the "encoding size" of the input number α to be the bit size of the integers $a_0, \ldots, a_{p-1}$ in the representation $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$.

The key to our construction is our ability to obtain "reasonable" upper and lower bounds on certain quantities involving the basic solutions of the LP. These bounds are a function of p (the order of the subring in which we work) and the encoding size of the data. We use these bounds to show that the modified algorithm's running time is polynomial in the dimension of the LP, the order p of the subring, and the encoding size of the data. We then proceed to modify the scheme given by Tardos [19] for rational data so that it works with data from W_p. The modified Tardos scheme gives us an algorithm whose running time is independent of the encoding size of the objective and right-hand side data (in fact, the objective and right-hand side vectors are allowed to be arbitrary real numbers).

Finally, we show that the numbers belonging to the extension of the integers by a set of positive integer square roots are embedded in W_p, for some known p. Using our earlier results, we then obtain an algorithm to solve LPs with such coefficients in time that is polynomial in the problem dimension, the encoding size of the input data, and the product of all the square roots.

The paper is organized in the following manner: In Section 2, we discuss circulant matrices and show that LPs with a circulant coefficient matrix can be polynomially transformed into an equivalent LP in which the entries of the coefficient matrix are small in magnitude and belong to W_n, where n is the dimension of the matrix. We also show that if the original data is rational then these entries are integers. In this case, the Tardos scheme [19] gives us a strongly polynomial algorithm directly. In Section 3, we analyze general LPs whose coefficients belong to W_p. In particular, we modify a variant of the interior point algorithm to solve these problems in polynomial time. In Section 4, we use the results of Section 3 and the scheme in [19] to obtain an algorithm that runs in time independent of the encoding size of the objective and right-hand side data. In Section 5, we show how our results can be applied to LPs whose coefficients belong to an extension of the integers by a set of square roots. Finally, in Section 6, we conclude with some remarks.

2. Linear programming with circulant matrices

In this section we bound the complexity of linear programs whose coefficient matrices are circulant.

Definition. The $n \times n$ matrix A (possibly with complex entries) is said to be *circulant* if and only if it has the following form:

$$A = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \vdots & & \ddots & & \vdots \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{bmatrix}.$$

Given a vector $a^{\mathrm{T}} = (a_0, \ldots, a_{n-1})$, we shall use circ (a) to denote the circulant matrix defined above. (If M is a matrix or vector, we use M^{T} to denote the transpose of M.)

Circulant matrices are closely related to the vector operation of discrete convolution; the discrete convolution of the n-vectors a and b is defined to be the n-vector $\mathrm{circ}(a)^{\mathrm{T}}b$. Discrete convolution appears in a variety of physical models, particularly in the area of signal processing. Our interest in convolution and circulant matrices developed as part of computational work in tomography of the Earth, an area which shares with signal processing the inverse imaging problem of reconstructing a true signal from an observed but distorted one. Below, we give a simplistic account of a problem in inverse imaging and show how linear programs with circulant coefficient matrices arise in this context.

A discrete approximation to the spectrum of a light source can be obtained by passing light from the source through a spectrometer and measuring its intensity at each of a discrete set of wavelengths. Because spectrometers have finite resolution, the intensity recorded at any particular wavelength is usually contaminated by the intensity of the source at neighboring wavelengths; thus, the observed spectrum is a somewhat "smeared" version of the true one. This distortion can be modeled using convolution and the characteristics of the spectrometer and the source. Specifically, if the n-vector x contains the true intensity of the source at n equally-spaced wavelengths and the vector b represents the intensity actually observed at those wavelengths, a reasonable model of the distortion phenomenon is given by $b = \mathrm{circ}(a)^{\mathrm{T}}x$. The vector a is derived from known characteristics of the spectrometer.

A common problem in this context is the reconstruction of the true spectrum from a set of observed intensities. Under the above model, this task is equivalent to solving the equation $\mathrm{circ}(a)^{\mathrm{T}}x = b$ for the unknown x. It is often the case, however, that any one of a number of true spectra could lead to the observed data, implying $\mathrm{circ}(a)$ is singular and the distortion equation has no unique solution. In this case, instead of a unique solution, one usually looks for a physically valid solution that minimizes a linear functional with some interesting interpretation. As a simple example, one may desire the nonnegative spectrum with the smallest intensity in the jth wavelength that is consistent with the observed spectrum. The desired spectrum, of course, is the solution to the linear program $\{\min e_j^{\mathrm{T}}x \mid \mathrm{circ}(a)^{\mathrm{T}}x = b, x \geqslant 0\}$, where e_j is the jth unit vector.

This last problem belongs to the class of LPs we analyze in this section. Specifically, we consider the following standard-form problem:

$$\text{(P)} \qquad \begin{aligned} \min \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & Ax = b, \\ & x \geqslant 0, \end{aligned}$$

where $a, b, c, x \in \mathbb{R}^n$ and $A = \mathrm{circ}(a)$.

To analyze (P), we need to develop a number of the properties of circulant matrices. We begin along these lines by defining a matrix with the interesting property that it diagonalizes every circulant matrix.

Definition. Let ω be the first primitive nth root of unity; that is,

$$\omega = \mathrm{e}^{2\pi \mathrm{i}/n},$$

where $\mathrm{i} = \sqrt{-1}$. Then we define the $n \times n$ matrix F_n to be

$$F_n = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega^1 & \cdots & \omega^{n-2} & \omega^{n-1} \\ \vdots & & \ddots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^2 & \omega^1 \end{bmatrix}.$$

In general, the jkth component of F_n is $\omega^{(j-1)(k-1)}$.

The matrix F_n is usually called the nth order *Fourier matrix*, the label Fourier coming from the fact that the product $F_n a$ can be used to define a discrete Fourier transform of the n-vector a. By using the identity $\omega^j = \omega^{j \bmod n}$ and by manipulating geometric series, it is easy to show that $\sqrt{1/n}\, F_n$ is unitary; that is, $F_n^{-1} = (1/n)F_n^*$, where F_n^* denotes the Hermitian transpose of F_n (the jkth component of F_n^* is $\omega^{(1-j)(k-1)}$). In an effort to keep notation simple, we omit the index n on F_n when the size of the Fourier matrix is clear from the context of the discussion.

We now state several well-known results on circulant matrices. Detailed proofs of these results can be found in [5].

Proposition 2.1. (*i*) *Let $A = \mathrm{circ}(a)$ be an $n \times n$ circulant matrix. Then the columns of the nth order Fourier matrix F are eigenvectors of A and the entries of the n-vector Fa are the corresponding eigenvalues.*

(ii) *A is an $n \times n$ circulant matrix if, and only if, $A = (1/n)FGF^*$ for some $n \times n$ diagonal matrix G.* □

As an easy corollary, we have the following useful fact:

Corollary 2.1. *Let A and B be circulant matrices of the same order. Then $AB = BA$.* □

Proposition 2.1 can also be used to show that the class of circulant matrices is closed under the operations of addition, multiplication, (Hermitian) transposition, inversion, and pseudoinversion. This last operation, which is formally defined below, will be of fundamental importance when we turn to linear programming.

Definition. Given a matrix M, we use the term *pseudoinverse* of M to denote the unique matrix M^+ that satisfies

(i) $MM^+M = M$;
(ii) $M^+MM^+ = M^+$;
(iii) $MM^+ = (MM^+)^*$;
(iv) $M^+M = (M^+M)^*$.

(Some authors refer to M^+ as the generalized inverse or the Penrose–Moore inverse of M).

It is particularly easy to express the pseudoinverse of a circulant matrix in terms of the Fourier diagonalization given in Proposition 2.1.

Proposition 2.2. *Let G be an $n \times n$ diagonal matrix and let $A = (1/n)FGF^*$. Then, the pseudoinverse of A is given by $A^+ = (1/n)FG^+F^*$, where G^+ is an $n \times n$ diagonal matrix with entries*

$$G^+_{jj} = \begin{cases} 1/G_{jj} & \text{if } G_{jj} \neq 0, \\ 0 & \text{otherwise.} \end{cases} \qquad \square$$

It follows from Proposition 2.2 and part (ii) of Proposition 2.1 that the matrix A is circulant if and only if its pseudoinverse is circulant.

As the next result shows, it is easy to compute the pseudoinverse inverse of a circulant matrix.

Proposition 2.3. *Let A be a given $n \times n$ circulant matrix. Then it is possible to compute A^+ from A using only Gaussian elimination and a constant number of matrix multiplications, each of order n.*

Proof. Consider the system $A^3X = A$, where X is an $n \times n$ matrix of unknowns. This system has at least one solution, namely $(A^+)^2$. To see this, note that since A and A^+ are both circulant they commute by Corollary 2.1, and so

$$A^3(A^+)^2 = AA^+AA^+A = AA^+A = A,$$

by property (i) of the pseudoinverse. Thus we can find a particular solution $\tilde{X}$ to $A^3X = A$ by Gaussian elimination. Now,

$$(A^+)^2A^3\tilde{X} = (A^+)^2A.$$

Hence,

$$A\tilde{X} = A^+,$$

which completes the proof. □

We turn now to an analysis of the linear program (P). Our strategy for establishing the strongly polynomial-time solvability of (P) is centered around a problem transformation. Briefly, we transform the given problem into an equivalent problem in which most of the information in the coefficient matrix has been "pushed" to the right-hand side. We then show that the transformed problem can be solved in time independent of its right-hand side.

We begin with a simple observation that will be of use in transforming (P) into an equivalent but more transparent problem.

Proposition 2.4. *Let $M \in \mathbb{C}^{r\times s}$, $g \in \mathbb{C}^r$, $v \in \mathbb{C}^s$, and suppose that the system $Mv = g$ has a solution. Then, $\{v \mid Mv = g\} = \{v \mid M^+Mv = M^+g\}$.*

Proof. The proof follows easily from the observation that the null space of M is the same as the null space of M^+M by property (i) of the pseudoinverse. □

It follows from the last proposition that if $Ax = b$ is consistent then (P) is equivalent to the following problem:

$$(\mathrm{P}') \qquad \begin{aligned} \min \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & nA^+Ax = nA^+b, \\ & x \geq 0. \end{aligned}$$

An interesting aspect of this transformation is that the entries of nA^+A, the coefficient matrix of (P′), have a particularly simple and useful representation in terms of the nth roots of unity. This representation is the subject of the next proposition.

Proposition 2.5. *Let A be an $n \times n$ circulant matrix. Then each entry α of the matrix nA^+A can be written in the form*

$$\alpha = \sum_{j=0}^{n-1} d_j \omega^j,$$

where $\omega = \mathrm{e}^{2\pi \mathrm{i}/n}$ and where d_j is either 0 or 1.

Proof. By Propositions 2.1 and 2.2, we can write

$$A^+A = \{(1/n)FG^+F^*\}\{(1/n)FGF^*\}$$

for some $n \times n$ diagonal matrix G. Using the identity $(1/n)F^*F = I$, we then have

$$nA^+A = FDF^*,$$

where D is a diagonal matrix with ones and zeros in positions on the diagonal that correspond to the positions of the nonzero and zero elements of G, respectively. Proof of the statement follows easily from this expression by using the definitions of F and F^* and the identity $\omega^j = \omega^{j \bmod n}$. $\square$

We shall use the representation given in Proposition 2.5 to show that the entries in the coefficient matrix of problem (P′) are nice in some sense. Toward this end, it is useful to view these entries in terms of a standard notion from algebraic number theory, which we discuss next.

Definition. A complex number α is called an *algebraic integer* if there exists integers $d_0, \ldots, d_{k-1}$ such that $\alpha^k + d_{k-1}\alpha^{k-1} + \cdots + d_0 = 0$.

Note that because it satisfies the polynomial equation $\omega^n - 1 = 0$, the number $\omega = \mathrm{e}^{2\pi \mathrm{i}/n}$ is an algebraic integer by the above definition.

Proposition 2.6. *The set of algebraic integers is closed under addition, multiplication, and negation.* $\square$

Since ω is an algebraic integer, Proposition 2.6 implies that if α has the form $\alpha = \sum_{j=0}^{n-1} d_j \omega^j$, where d_j is integer, then α is an algebraic integer.

Next, we present a standard result from number theory that will be of key importance when we derive complexity bounds for rational instances of (P).

Proposition 2.7. *An algebraic integer is rational if and only if it is an integer.* $\square$

Proofs of Propositions 2.6 and 2.7 can be found in most texts on algebraic number theory (see, e.g., [8]).

If the coefficient matrix A in (P) is rational we can use Propositions 2.6 and 2.7 to make a strong statement about the form of the entries in $nA^{+}A$, the coefficient matrix of (P′):

Proposition 2.8. *If A is an $n \times n$ circulant matrix with rational entries, then the entries of $nA^{+}A$ are all integers.*

Proof. It follows immediately from Propositions 2.5 and 2.6 that each entry of $nA^{+}A$ is an algebraic integer. But since each entry of A is rational, we see by Proposition 2.3 that each entry of A^{+}, and hence each entry of $nA^{+}A$, is also rational. Proof of the proposition follows by noting that since each entry of $nA^{+}A$ is both a rational number and an algebraic integer, it must be an integer by Proposition 2.7. □

Armed with the above results, we can easily bound the complexity of rational instances of (P):

Theorem 2.1. *Suppose that A is an $n \times n$ circulant matrix and that the entries of A, b, and c are rational. Then* (P) *can be solved in strongly polynomial time.*

Proof. As a preprocessing step in the solution of (P), one can use Gaussian elimination to check the consistency of $Ax = b$. If $Ax = b$ is inconsistent, then obviously (P) is infeasible and no further work is required. If $Ax = b$ is consistent, then by Proposition 2.3 one can convert (P) into the equivalent problem (P′) using Gaussian elimination and matrix multiplication. In either case the dominant computational work is Gaussian elimination, which is a strongly polynomial operation (see [6]).

In Proposition 2.5 we established that every entry α in the coefficient matrix $nA^{+}A$ has the form $\alpha = \sum_{j=0}^{n-1} d_j \omega^j$, where d_j is either 0 or 1. Since $|\omega^j| = 1$ for all j, we see that $|\alpha| \leq n$. But because we assume that A is rational, α must be an integer by proposition 2.8. Thus, the coefficient matrix of (P′) contains integers of absolute value at most n. Moreover, the right-hand side and objective vectors of (P′) contain only rational numbers, since by Proposition 2.3 we can construct (P′) from (P) using only Gaussian elimination and matrix multiplication.

Tardos [19] showed that LPs of the form (P′) with rational coefficients can be solved in time polynomial in the problem dimension and the binary encoding size of the numbers in the coefficient matrix. By the above arguments, the binary encoding size of $nA^{+}A$ is bounded by a polynomial in the dimension of (P′). Hence, (P′), and therefore (P), can be solved in strongly polynomial time. □

For the purpose of analyzing LPs with circulant coefficient matrices and real coefficients (including those in the objective and right-hand side), we adopt a model of computation that allows constant time addition, subtraction, multiplication, division, and comparison of real numbers. Under this model we have the following result:

Theorem 2.2. *Suppose that A is an $n \times n$ circulant matrix and that the entries of A, b, and c are real numbers. Then* (P) *can be solved in strongly polynomial time.*

To prove Theorem 2.2 we need a number of new results concerning the numbers in nA^+A. These results, which we shall obtain in Sections 3 and 4, lead in a natural way to an analysis of linear programming in a subring of the algebraic integers. The proof of Theorem 2.2 follows directly from this analysis and is given at the end of Section 4.

3. LP over a subring of the algebraic integers

We consider the following primal–dual pair of LPs:

$$\text{(P)} \qquad \begin{aligned} \min \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

$$\text{(D)} \qquad \begin{aligned} \max \quad & b^{\mathrm{T}}y \\ \text{s.t.} \quad & A^Ty + z = c, \\ & z \geq 0, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c, x \in \mathbb{R}^n$. Throughout this section we assume that all entries of A, b, and c are real and belong to the following set:

$$W_p = \left\{ \alpha \colon \alpha = \sum_{j=0}^{p-1} a_j\omega^j;\ a_j \in \mathbb{Z}\ \forall j \right\},$$

where $\omega = \mathrm{e}^{2\pi \mathrm{i}/p}$ and where $\mathbb{Z}$ denotes the integers. We shall use V_p to denote the set of all real members of W_p; that is, $V_p = W_p \cap \mathbb{R}$.

Note that, because $w^j = w^{j \bmod p}$, the set W_p forms a subring of the complex numbers (i.e., W_p is closed under addition, multiplication, and negation).

It is also important to note that we lose no generality by working with the set W_p instead of the larger set $\Omega_p = \{\gamma \mid \gamma = \sum_{j=0}^{p-1} q_j\omega^j;\ q_j \text{ rational } \forall j\}$. It is well known from number theory that Ω_p is precisely the subfield $\Gamma_p = \{\gamma \mid \gamma = \alpha/\beta;\ \beta \neq 0;\ \alpha, \beta \in W_p\}$. Hence, by multiplying through by a common denominator we can transform any LP whose coefficients belong to Γ_p $(= \Omega_p)$ into an equivalent problem whose coefficients all belong to W_p.

We now develop some properties of W_p that will prove useful in the analysis of linear programming to follow. We begin by giving explicit expressions for the addition and multiplication of members of W_p.

Proposition 3.1. *Let $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$ and $\beta = \sum_{j=0}^{p-1} b_j\omega^j$, where a_j and b_j are integers for all j, and let $a^{\mathrm{T}} = (a_0, \ldots, a_{p-1})$ and $b^{\mathrm{T}} = (b_0, \ldots, b_{p-1})$. Then*:

(i) *If $\gamma = \alpha + \beta$, then $\gamma = \sum_{j=0}^{p-1} c_j\omega^j$, where $c = (c_0, \ldots, c_{p-1})^{\mathrm{T}} = a + b$.*

(ii) *If $\gamma = \alpha\beta$, then $\gamma = \sum_{j=0}^{p-1} c_j\omega^j$, where $c = (c_0, \ldots, c_{p-1})^{\mathrm{T}} = \mathrm{circ}(a)^{\mathrm{T}}b$.*

Proof. Statement (i) is obvious. Statement (ii) follows directly from the fact that $\omega^j = \omega^{j \bmod p}$. □

Our immediate goal is to establish upper and lower bounds on certain functions of $\alpha \in W_p$. To obtain these bounds, we need a measure of the magnitude of the coefficients $a_0, \ldots, a_{p-1}$ in the representation $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$. This representation, however, is not unique for general p, since $\sum_{j=0}^{p-1} \omega^j = 0$ for $p > 1$. The question of uniqueness of representation motivates us to measure the magnitude of the representation of α in the following manner:

Definition. Given $\alpha \in W_p$, we define the *representation height* of α with respect to W_p to be:

$$S_p(\alpha) = \min\left\{\sum_{j=0}^{p-1} |a_j|:\ \alpha = \sum_{j=0}^{p-1} a_j\omega^j;\ a_j \in \mathbb{Z}\ \forall j\right\},$$

where $\omega = \mathrm{e}^{2\pi\mathrm{i}/p}$. We include the index p on $S_p(\alpha)$ only when there is danger of confusion regarding the set W_p with respect to which the representation height of α is to be taken. (Representation height should not be confused with any of the various notions of height encountered in number theory.)

Next, we state the main algebraic and metric properties of representation height.

Proposition 3.2. *Let $\alpha, \beta \in W_p$. Then*:

(i) *$S(a\alpha + b\beta) \leq |a|S(\alpha) + |b|S(\beta)$ for any integers a and b.*

(ii) *$S(\alpha\beta) \leq S(\alpha)S(\beta)$.*

(iii) *$|\alpha| \leq S(\alpha)$.*

(iv) *If $\alpha \neq 0$, $|\alpha| \geq (S(\alpha))^{1-p}$.*

Proof. (i) This statement follows directly from the definition of the function S and part (ii) of Proposition 3.1.

(ii) Define vectors $a, b \in \mathbb{Z}^p$ such that $S(\alpha) = \|a\|_1$, $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$, $S(\beta) = \|b\|_1$, and $\beta = \sum_{j=0}^{p-1} b_j\omega^j$. (By the definition of S such vectors must exist.) Let $\gamma = \alpha\beta$. Then, by part (ii) of Proposition 3.1, we have $\gamma = \sum_{j=0}^{p-1} c_j\omega^j$, where $c = (c_0, \ldots, c_{p-1})^{\mathrm{T}} = \mathrm{circ}(a)^{\mathrm{T}}b$. Hence,

$$S(\gamma) \leq \|c\|_1 = \|\mathrm{circ}(a)^{\mathrm{T}}b\|_1 \leq \|\mathrm{circ}(a)\|_1\|b\|_1 = \|a\|_1\|b\|_1 = S(\alpha)S(\beta).$$

(iii) Let $a \in \mathbb{Z}^p$ such that $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$. Then we have $|\alpha| = |\sum_{j=0}^{p-1} a_j\omega^j| \leq \sum_{j=0}^{p-1} |a_j||\omega^j| = \sum_{j=0}^{p-1} |a_j| \leq S(\alpha)$, as desired.

(iv) Define $a \in \mathbb{Z}^p$ such that $S(\alpha) = \|a\|_1$ and $\alpha = \sum_{j=0}^{p-1} a_j\omega^j$. Now consider $A = \operatorname{circ}(a)$. It follows from part (i) of Proposition 2.1 that A is a normal matrix (i.e., $AA^{\mathrm{T}} = A^{\mathrm{T}}A$). This fact implies that the eigenvalues of AA^{T} are $|\hat{a}_0|^2, |\hat{a}_1|^2, \ldots, |\hat{a}_{p-1}|^2$, where $\hat{a}_j = (Fa)_j$ is (again by Proposition 2.1) the jth eigenvalue of A. Let $J = \{j \mid \hat{a}_j \neq 0\}$. Then by a standard result of linear algebra (see, e.g., [7]) we have

$$\prod_{j \in J} |\hat{a}_j|^2 = \text{sum of all principle minors of } AA^{\mathrm{T}} \text{ of order } |J|,$$

where $|J|$ denotes the cardinality of the set J. Because the entries of AA^{T} are all integers, this last equality implies $\prod_{j \in J} |\hat{a}_j|^2 \geq 1$. Note that the index 1 belongs to J, since we assume that $\alpha = \hat{a}_1 \neq 0$. Now, using the inequality $|\hat{a}_k| = |\sum_{j=0}^{p-1} a_j\omega^{jk}| \leq \sum_{j=0}^{p-1} |a_j| = S(\alpha)$, we have

$$|\alpha| \geq (S(\alpha))^{1-|J|} \geq (S(\alpha))^{1-p},$$

which completes the proof of the proposition. □

We shall use Proposition 3.2 to derive several useful results concerning matrices and systems of equations whose coefficients belong to W_p. These results are most easily stated in terms of the quantities introduced below.

Definition. Let M be an $r \times s$ matrix all of whose entries M_{jk} belong to W_p. We define the *representation height of the matrix M* to be

$$T(M) = \max_{j,k}\{S(M_{jk})\}.$$

Let l denote the rank of M, and let

$$\Delta(M) = (lT(M))^l.$$

Then we define the *representation size of the matrix M* to be

$$L(M) = \log(\Delta(M)).$$

We use similar notation when discussing linear programs. Let (Q) denote the problem $\{\min f^{\mathrm{T}}v \mid Mv = g,\ v \geq 0\}$, where the entries of M, g, and f all belong to W_p. Let Ψ denote the set of all entries in M, g, and f. We define the *representation height of the linear program* (Q) to be

$$T(M, g, f) = \max_{\alpha \in \Psi}\{S(\alpha)\}.$$

Let l denote the rank of M, and let

$$\Delta(M, g, f) = (lT(M, g, f))^l.$$

Then we define the *representation size of the linear program* (Q) to be

$$L(M, g, f) = \log(\Delta(M, g, f)).$$

We use analogous notation with respect to systems of linear inequalities. Specifically, we define the *representation height and size of the system* $\{Mv = g,\ v \geq 0\}$ to be the representation height and size of the linear program $\{\min 0v \mid Mv = g,\ v \geq 0\}$. We write these quantities as $T(M, g)$ and $L(M, g)$.

Having fixed notation, we now state some key properties of matrix determinants.

Proposition 3.3. *Let B be an $r \times r$ nonsingular matrix all of whose entries B_{jk} belong to W_p, and let $\Delta = \Delta(B)$. Then:*

(i) $\det(B) \in W_p$.

(ii) $S(\det(B)) \leq \Delta$.

(iii) $\Delta^{1-p} \leq |\det(B)| \leq \Delta$.

Proof. Let J be the set of all $r!$ permutations of $(1, 2, \ldots, r)$ and let $j = (j_1, \ldots, j_r)$ be a member of J. Then, by the definition of the determinant of a matrix, we can write

$$\det(B) = \sum_{j \in J} (\pm)(B_{1j_1} \cdots B_{rj_r}).$$

Statement (i) follows immediately from this expansion and the fact that W_p is closed under the operations of addition, multiplication, and negation. Now, taking the representation height of both sides of the expansion and using Proposition 3.2 gives

$$\begin{aligned} S(\det(B)) &= S\Big(\sum_{j \in J} (\pm)(B_{1j_1} \cdots B_{rj_r})\Big) \leq \sum_{j \in J} S(B_{1j_1}) \cdots S(B_{rj_r}) \\ &\leq \sum_{j \in J} (T(B))^r = r!(T(B))^r \leq (rT(B))^r = \Delta. \end{aligned}$$

This proves statement (ii). Statement (iii) follows directly from Proposition 3.2 and statements (i) and (ii). □

As a corollary we have the following result:

Corollary 3.1. *Let M and g be $r \times s$ and $r \times 1$ matrices, respectively, all of whose entries belong to W_p, and let $\Delta = \Delta(M, g)$. If $\bar{v}$ is a basic solution to the system $Mv = g$, then every nonzero component $\bar{v}_j$ of $\bar{v}$ satisfies $\Delta^{-p} \leq |\bar{v}_j| \leq \Delta^p$.*

Proof. The statement follows trivially from Cramer's rule and Proposition 3.3. □

Using the above results, we can modify almost any variant of the interior point method [9] or the ellipsoid method [10] to solve problem (P) in polynomial time. In the remaining part of this section, we shall show how to modify the primal-dual path following algorithm and its analysis as presented in [14] and [15]. Since most of the algorithm and analysis are not affected by the change from rationals to V_p, we present only the necessary modifications.

For the purposes of the complexity analysis, we assume that we have a machine that performs addition, subtraction, multiplication, division, and comparison of real numbers in constant time per operation. Implicit in our derivation of the basic complexity results is the assumption that, for any instance of (P), a bound on the representation size of the instance is part of the input data. (For our purposes, this assumption is essentially equivalent to the requirement that $S(\alpha)$ be part of the input for every coefficient α in the problem instance.) Later, we adopt a more natural input scheme and show that the basic complexity results extend easily to this case.

In the following two propositions, we use the properties of the set W_p and the function S to establish a theoretical stopping point for an interior point algorithm applied to (P) and (D). Because the discussion involves systems of inequalities and linear programs, we shall assume that the matrices and vectors we encounter are real.

Proposition 3.4. *Let M and g be $r \times s$ and $r \times 1$ matrices, respectively, all of whose components belong to V_p. Let $0 < \delta \leqslant \Delta^{-2p}$ be given, where $\Delta = \Delta(M, g)$. Let e_r denote the r-vector of all ones. Then the open system $Mv < g + \delta e_r$ is feasible if and only if the closed system $Mv \leqslant g$ is feasible. Moreover, given a solution to one system we can find a solution to the other system in time polynomial in r and s.*

Proof. The proof is an adaptation of that given by Papadimitriou and Steiglitz [17, Lemma 8.7, pp. 173–174] for an analogous result concerning systems with rational coefficients. The main difference lies in the use of the properties of representation height (Proposition 3.2) to derive lower bounds on integral polynomial forms involving members of V_p.

Obviously, any solution to the closed system is also a solution to the open system. Conversely, let $\bar{v}$ be a solution to $Mv < g + \delta e_r$. We shall show how to construct a solution to $Mv \leqslant g$.

Let $M_{j\cdot}$ denote the jth row of M, and consider the set of indices

$$J = \{j \mid g_j \leqslant M_{j\cdot}\bar{v} < g_j + \delta\}.$$

We may assume that, for all k, $M_{k\cdot} = \sum_{j \in J_1} \beta_{kj} M_{j\cdot}$ for some linearly independent subset J_1 of J, since otherwise we can increase the cardinality of J by solving a corresponding Gale system (cf. [17]).

By Cramer's rule and Proposition 3.3, the coefficient β_{kj} can be written in the form $\beta_{kj} = \gamma_{kj}/|\phi|$, where $\gamma, \phi \in W_p$ and $S(\gamma) \leqslant \Delta$ and $S(\phi) \leqslant \Delta$.

Now, let $\tilde{v}$ be a solution to the equations $M_{j\cdot} v = g_j$, for $j \in J_1$. Then, for $k \notin J_1$, we have

$$\begin{aligned} |\phi|(M_{k\cdot}\tilde{v} - g_k) &= \sum_{j \in J_1} \gamma_{kj} M_{j\cdot}\tilde{v} - |\phi| g_k \\ &= \sum_{j \in J_1} \gamma_{kj} g_j - |\phi| g_k \qquad\qquad (*) \\ &= -\sum_{j \in J_1} \gamma_{kj}(M_{j\cdot}\bar{v} - g_j) + |\phi|(M_{k\cdot}\bar{v} - g_k) \\ &< \delta\left(\sum_{j \in J_1} |\gamma_{kj}| + |\phi|\right) \\ &\leqslant \delta(s+1)\Delta \\ &\leqslant \Delta^{2(1-p)}. \end{aligned}$$

But by taking representation heights of equality (*) we have

$$S(|\phi|(M_{k\cdot}\tilde{v}-g_k)) \leq \sum_{j\in J_1} S(\gamma_{kj})S(g_j)+S(|\phi|g_k) \leq \Delta\left(\sum_{j\in J_1} S(g_j)+S(g_k)\right) \leq \Delta^2.$$

From this last inequality and Proposition 3.2, we see that $|\phi|(M_{k\cdot}\tilde{v}-g_k) \geq \Delta^{2(1-p)}$ if and only if $|\phi|(M_{k\cdot}\tilde{v}-g_k)>0$. But since we also have $|\phi|(M_{k\cdot}\tilde{v}-g_k)<\Delta^{2(1-p)}$, it must be that $|\phi|(M_{k\cdot}\tilde{v}-g_k)\leq 0$. Hence $(M_{k\cdot}\tilde{v}-g_k)\leq 0$, and we see that $\tilde{v}$ is the desired solution to $Mv\leq g$. Proof of the proposition follows by noting that we can construct $\tilde{v}$ in time polynomial in r and s using Gaussian elimination. □

Proposition 3.5. *Let the primal-dual pair* (P) *and* (D) *be given and let all the entries in A, b, and c belong to V_p. Assume we have a point $\bar{w}=(\bar{x},\bar{y},\bar{z})$ feasible to* (P) *and* (D) *that satisfies $\bar{x}^{\mathrm{T}}\bar{z}\leq(6n\Delta)^{-24p}$, where $\Delta=\Delta(A,b,c)$. Then from $\bar{w}$, we can find a point $w^*=(x^*,y^*,z^*)$ in no more than $\mathrm{O}(n^3)$ arithmetic operations, such that x^* is optimal to* (P) *and y^* and z^* are optimal to* (D).

Proof. The proof is essentially the same as that given for the rational case in [15, Proposition 4.2], provided Proposition 3.4 is used in place of the analogous result employed there. □

Next, we state the minimum improvement made in the duality gap during each iteration of the primal-dual path following algorithm.

Proposition 3.6. *Let $w^0=(x^0,y^0,z^0)$ be a valid initial point for the primal-dual path following algorithm given in* [14] *when applied to* (P) *and* (D). *Then the algorithm generates a sequence of feasible points $w^k=(x^k,y^k,z^k)$ satisfying*

$$(x^k)^{\mathrm{T}}z^k\leq(x^0)^{\mathrm{T}}z^0(1-q\sqrt{1/n})^k,$$

for some constant $0<q\leq 1$.

Proof. The statement follows directly from the fact that the convergence proofs given for the algorithm in [14] do not rely on rationality of the input data. □

We can now state the main results of the section.

Theorem 3.1. *Let $\delta\geq L(A,b,c)$ be given. Then, under the model of computation discussed above, problems* (P) *and* (D) *can be solved in time polynomial in p, n, and δ.*

Proof. For rational data, it is shown in [15] that the solutions to (P) and (D) can be obtained by solving a pair of artificial problems whose size is order of the size of (P) and (D). The artificial pair has the property that a valid starting point, with known duality gap, is readily available for the primal-dual algorithm. In our case, it is a straightforward exercise to show that we can construct a similar pair of artificial problems whose representation size is order of $L(A, b, c)$ and whose starting point has a duality gap that is $2^{O(p\delta)}$. Using Propositions 3.5 and 3.6, it is easy to show that, given this initial duality gap, the number of iterations performed by the primal-dual algorithm is polynomial in p, n, and δ. Proof of the theorem then follows by noting that the work performed in each iteration of the algorithm is polynomial in n. $\square$

Theorem 3.1 is derived under the assumption that a bound on the representation size of (P) is input along with the problem coefficients. As an alternative, we may consider a model of input based on an integer representation of the problem coefficients. In this model, we assume that every number $\alpha \in V_p$ in an instance of (P) is encoded for input as a set of integers $a_0, \ldots, a_{p-1}$, such that $\alpha = \sum_{j=0}^{p-1} a_j \omega^j$. To ensure α can be calculated from its integer representation, we also assume that the machine has available the real part of ω ($=\cos 2\pi/p$), or that it can calculate this number.

We work with a different measure of input size in the integer-based model. We define the *encoding size* of $\alpha \in W_p$ to be the sum of the binary encoding sizes of the integers $a_0, \ldots, a_{p-1}$ in the above expansion. Similarly, we define the encoding size of a matrix or LP to be the sum of encoding sizes of its coefficients.

We now restate our earlier complexity bounds in terms of the integer-based input model.

Theorem 3.2. *Under the model of computation discussed above, problems* (P) *and* (D) *can be solved in time polynomial in p, n, and the encoding size of* (P).

Proof. The result follows immediately from Theorem 3.1 by noting that $L(A, b, c)$ is bounded by a polynomial in the encoding size of problem (P). $\square$

4. Tardos scheme for LP over a subring of the algebraic integers

In this section, we use the results of Section 3 to modify the Tardos scheme for solving combinatorial linear programs [19]. The modifications permit us to solve problem (P) of Section 3 in time polynomial in n, p, and the size of the matrix A, independent of the objective and right-hand side data. (In fact, the objective and right-hand side coefficients may be arbitrary real numbers.)

We shall follow the presentation of Tardos' algorithm given by Schrijver in [18, Section 15.2], although we present only those key propositions needed for the switch from rationals to V_p.

We consider (P) and (D), the primal-dual pair of LPs defined at the beginning of Section 3. In order to show that these problems can be solved in polynomial time independent of the numbers in b and c, we need several sensitivity results and a guarantee that we can find a feasible solution in time independent of the size of the right-hand side. We begin with the sensitivity results (Propositions 4.1 and 4.2).

Let z' be defined as the solution of the following problem:

$$\begin{aligned} \min \quad & \|z\|_2 \\ \text{s.t.} \quad & A^{\mathrm{T}}y + z = c. \end{aligned}$$

As in [18], we may assume $z' \neq 0$ without loss of generality. Now define the vector c' as $c' = (mn\Delta^{p+1}/\|z'\|_\infty)z'$. Note that c' can replace c in (P) without changing the optimal solution. Hence, we lose no generality by assuming that the objective vector c in (P) already has the form of c'.

Now, let $\tilde{c} = (\lceil c_1 \rceil, \ldots, \lceil c_n \rceil)^{\mathrm{T}}$, where $\lceil c_k \rceil$ denotes the smallest integer not less than c_k, and consider the following primal-dual pair of LPs:

$$\begin{aligned} (\mathrm{P}') \quad \min \quad & \tilde{c}^{\mathrm{T}}x \\ \text{s.t.} \quad & Ax = b, \\ & x \geqslant 0, \\ (\mathrm{D}') \quad \max \quad & b^{\mathrm{T}}y \\ \text{s.t.} \quad & A^{\mathrm{T}} + z = \tilde{c}, \\ & z \geqslant 0. \end{aligned}$$

Using the properties of representation height and following the proofs given for rational problems by Schrijver [18], one can prove the following sensitivity results.

Proposition 4.1. *Let* $(\tilde{y}, \tilde{z})$ *be an optimal solution to* (D′) *and let* $\Delta = \Delta(A, b, \tilde{c})$. *Then,* $\|\tilde{z}\|_\infty \geqslant m\Delta^{p+1}$. □

The proof of Proposition 4.1, which is independent of the numerical properties of the coefficients of (D′), is essentially the same as that given in [18, Lemma A, expression (29), pp. 195–196].

Proposition 4.2. *Let* $(\tilde{y}, \tilde{z})$ *be an optimal solution to* (D′), *and suppose that* (D) *has an optimal solution. Let* $\Delta = \Delta(A, b, \tilde{c})$. *Then there exists an optimal solution* (y^*, z^*) *to* (D) *such that*:

(i) $\|y^* - \tilde{y}\|_\infty < m\Delta^p$.

(ii) $z_k^* > 0$, *where* $k = \arg\max_j \{\tilde{z}_j\}$.

Proof. (i) Let δ be an upper bound on the absolute value of each entry of B^{-1}, where B is any nonsingular submatrix of A. Then it can be shown (cf. [18, Theorem 10.5, p. 126]) that there exists an optimal solution y^* to (D) such that

$$\|y^* - \tilde{y}\|_\infty \leqslant m\delta \|c - \lceil c \rceil\|_\infty < m\delta.$$

By Cramer's rule, if B is a nonsingular submatrix of A then each entry of B^{-1} is a ratio of subdeterminants of A. But since the entries of A belong to W_p, Proposition 3.2 gives us the bound $\delta \leqslant \Delta\Delta^{p-1} = \Delta^p$. Hence we have $\|y^* - \tilde{y}\|_\infty < m\Delta^p$, as desired.

(ii) Let $k = \arg\max_j\{([c] - A^{\mathrm{T}}\tilde{y})_j\}$, and let y^* be the solution to (D) from (i). Then, letting $A^{\mathrm{T}}_{k\cdot}$ denote the kth row of A^{T} and using statement (i) and Proposition 4.1, we have

$$\begin{aligned} A^{\mathrm{T}}_{k\cdot} y^* &\leq A^{\mathrm{T}}_{k\cdot}\tilde{y} + (\|A^{\mathrm{T}}_{k\cdot}\|_1)(\|y^* - \tilde{y}\|_\infty) < A^{\mathrm{T}}_{k\cdot}\tilde{y} + \Delta(m\Delta^p) \\ &\leq c - m\Delta^{p+1} + \Delta(m\Delta^p) \leq c, \end{aligned}$$

which completes the proof. □

Next, we give a nondegeneracy result that is key in proving that a feasible solution to a linear program can be found in time independent of the size of the right-hand side (cf. [18, Lemma B]).

Proposition 4.3. *Let A be an $m \times n$ matrix of rank m, all of whose entries belong to V_p. Let $f = ((\Delta^p + 1), (\Delta^p + 1)^2, \ldots, (\Delta^p + 1)^n)^{\mathrm{T}}$, where $\Delta = \Delta(A)$. Then every basic solution of the system $Ax = Af$ is nondegenerate (i.e., for every $m \times m$ nonsingular submatrix B of A, the vector $B^{-1}Af$ has no zero components).*

Proof. By Cramer's rule, it suffices to prove that if M is any $m \times (m-1)$ submatrix of A with rank $m-1$, then the matrix formed by adjoining Af to M, namely $[Mc \vdots Af]$, is nonsingular. Proceeding along these lines, we use the definition of f and a well-known expansion for determinants to write

$$\det[M \vdots Af] = \sum_{k=1}^{n} (\Delta^p + 1)^k \det A_{(k)},$$

where $A_{(k)} = [M \vdots A_{\cdot k}]$ and $A_{\cdot k}$ denotes the kth column of A. Let l be largest index for which $\det A_{(k)} \neq 0$. (Such an index must exist since rank $(A) = m$.) By Proposition 3.3, if $\det A_{(k)} \neq 0$ then $\Delta^{1-p} \leq |\det(A_{(k)})| \leq \Delta$. In particular,

$$(\Delta^p + 1)^l |\det(A_{(k)})| \geq (\Delta^p + 1)^l (\Delta^{1-p}).$$

But,

$$\begin{aligned} &\sum_{k=1}^{l-1} (\Delta^p + 1)^k |\det(A_{(k)})| \\ &\quad\leq \sum_{k=1}^{l-1} (\Delta^p + 1)^k \Delta = \Delta \cdot \frac{(\Delta^p + 1)^l - (\Delta^p + 1)}{\Delta^p} < (\Delta^p + 1)^l \Delta^{1-p}. \end{aligned}$$

Hence, $\det[M \vdots Af] = \sum_{k=1}^{l} (\Delta^p + 1)^k \det A_{(k)} \neq 0$, which completes the proof. □

By following the Tardos scheme as presented in [18, Section 15.2], one can easily (though admittedly rather tediously) verify that Propositions 4.1 through 4.3 contain all the modifications to the proofs of Tardos' algorithm necessary for the switch from rationals to V_p.

Theorem 4.1. *Let $\delta \geqslant L(A)$ be given. Then problems* (P) *and* (D) *can be solved in time polynomial in p, n, and δ.*

Proof. Note that the effort involved in rounding off the objective and right-hand side vectors in Tardos' algorithm depends only on the size of m, n, and $\Delta(A)$ and not on the size of b or c, since these vectors are scaled before rounding. Therefore, the rounding procedure is polynomial even if b and c are real.

Because we have available the polynomial-time algorithm for LPs with coefficients in V_p developed in Section 3, the result follows by the arguments in [18, Section 15.2] together with Propositions 4.1 through 4.3. □

The next theorem follows immediately from Theorem 4.1 by noting that $L(A)$ is bounded by a polynomial in the encoding size of the matrix A.

Theorem 4.2. *Problems* (P) *and* (D) *can be solved in time polynomial in p, n, and the encoding size of the matrix A.* □

We are now able to prove the claim, made in Section 2, that LPs with real circulant coefficient matrices are strongly polynomial.

Proof of Theorem 2.2. By the discussion in Section 2, we have an a priori bound on the representation size of the coefficient matrix, namely $L(nA^{+}A) \leqslant 2n \log n$. Proof of the theorem follows directly from this bound and Theorem 4.1. □

5. LP in quadratic field extensions

In this section, we use our earlier results to obtain complexity bounds for linear programs in which the coefficients are integer linear combinations of integer square roots. In particular, we consider LPs whose coefficients belong to $\mathbb{Z}(d_1, \ldots, d_k)$, where we define $\mathbb{Z}(d_1, \ldots, d_k)$ to be the additive and multiplicative ring generated by $1, \sqrt{d_1}, \ldots, \sqrt{d_k}$; that is, $\mathbb{Z}(d_1, \ldots, d_k)$ consists of all numbers that have the form $\sum a_j(\sqrt{d_1})^{j_1}(\sqrt{d_2})^{j_2} \cdots (\sqrt{d_k})^{j_k}$, where a_j is an integer and the summation runs over all (distinct) k-tuples $J_j = (j_1, \ldots, j_k)$ with elements that are either 0 or 1. Because we are interested in linear programs with real coefficients, we shall assume that the d_j are positive.

Our strategy is to find an integer q such that the set $\mathbb{Z}(d_1, \ldots, d_k)$ is embedded in the set W_q. Using this result, we bound the representation height of $\alpha \in \mathbb{Z}(d_1, \ldots, d_k)$ with respect to W_q by a function of the coefficients in the representation of α in terms of the cross products of the $\sqrt{d_j}$. We then apply the results of Sections 3 and 4, which bound the complexity of a LP by a function of its representation size.

We begin by stating a key result due originally to Gauss. Proof can be found in many advanced texts on number theory (see, e.g., [8]).

Proposition 5.1. *Let p be an odd prime and let $\omega = \mathrm{e}^{2\pi\mathrm{i}/p}$. Then,*

$$\sum_{j=0}^{p-1} \omega^{j^2} = \begin{cases} \sqrt{p} & \text{if } p = 1 \bmod 4, \\ \mathrm{i}\sqrt{p} & \text{if } p = 3 \bmod 4. \end{cases} \qquad \Box$$

We use Proposition 5.1 to characterize the square root of any prime number in terms of the roots of unity.

Proposition 5.2. *Let p be a prime. Then W_{4p} contains $\sqrt{p}$. Moreover $S_{4p}(\sqrt{p}) \leq p$.*

Proof. There are three cases to consider.

Case 1: $p = 1 \bmod 4$. Using Proposition 5.1, we have $\sqrt{p} = \sum_{j=0}^{p-1} \mathrm{e}^{2\pi\mathrm{i}j^2/p} = \sum_{j=0}^{p-1} \mathrm{e}^{2\pi\mathrm{i}4j^2/4p} \in W_{4p}$.

Case 2: $p = 3 \bmod 4$. Again using Proposition 5.1, we have $(-\mathrm{i})\mathrm{i}\sqrt{p} = \mathrm{e}^{-\pi\mathrm{i}/2} \sum_{j=0}^{p-1} \mathrm{e}^{2\pi\mathrm{i}4j^2/4p} = \sum_{j=0}^{p-1} \mathrm{e}^{2\pi\mathrm{i}(4j^2-p)/4p} \in W_{4p}$.

Case 3: $p = 2$. It is easy to verify that $\mathrm{e}^{2\pi\mathrm{i}/8} + \mathrm{e}^{2\pi\mathrm{i}7/8} = \sqrt{2}$. It is clear from the above expansions that $S_{4p}(\sqrt{p}) \leq p$ in all three cases. $\Box$

Note that if $p = 1 \bmod 4$ then, by Proposition 5.1, $\sqrt{p}$ also belongs to the smaller set W_p. For simplicity we avoid distinguishing this case further.

To make full use of Proposition 5.2, we need the following simple observation.

Proposition 5.3. *Let n and k be positive integers. Then W_{nk} contains W_k.*

Proof. By definition W_{nk} contains $\omega^j = \mathrm{e}^{2\pi\mathrm{i}nj/nk} = \mathrm{e}^{2\pi\mathrm{i}j/k}$, for all $j = 0, \ldots, k-1$. Because W_{nk} is a ring it also contains all integer linear combinations of the ω^j. But this last set is just W_k. $\Box$

Next, we use Propositions 5.2 and 5.3 to characterize an embedding of $\mathbb{Z}(d) = \{\alpha \mid \alpha = a + b\sqrt{d}\,;\ a, b \in \mathbb{Z}\}$ for d not necessarily prime.

Proposition 5.4. *Let d be a positive integer. Then W_{4d} contains $\mathbb{Z}(d)$. Moreover $S_{4d}(\sqrt{d}) \leq d$.*

Proof. By the prime factorization theorem for integers (see, e.g., [8]) we know that $d = a^2 p_1 \cdots p_k$ for some positive integer a and primes $p_1, \ldots, p_k$. Because $4d$ is an integer multiple of $4p_j$, we know by Proposition 5.3 that W_{4d} contains W_{4p_j}. Hence by Proposition 5.2, W_{4d} also contains $\sqrt{p_j}$ for all $j = 1, \ldots, k$. The statement that W_{4d} contains $\mathbb{Z}(d)$ follows trivially from this last result and the fact that W_{4d} is a ring. Now, by Proposition 5.3, we have $S_{4p_j}(\sqrt{p_j}) \leq p_j$. But by the definition of representation height, the relation $W_{4p_j} \subset W_{4d}$ implies that $S_{4d}(\sqrt{p_j}) \leq S_{4p_j}(\sqrt{p_j})$. Hence, we can write

$$\begin{aligned} S_{4d}(\sqrt{d}) &= S_{4d}(\sqrt{a^2 p_1 \cdots p_k}) \\ &\leq S_{4d}(a) S_{4d}(\sqrt{p_1}) \cdots S_{4d}(\sqrt{p_k}) \leq a^2 p_1 \cdots p_k \leq d, \end{aligned}$$

which completes the proof of the proposition. $\Box$

Making an easy generalization of Proposition 5.4, we next characterize an embedding of the ring generated by a number of square roots.

Proposition 5.5. *Let $d_1, \ldots, d_k$ be positive integers and let $d = \prod_{j=1}^{k} d_j$. Then, W_{4d} contains $\mathbb{Z}(d_1, \ldots, d_k)$.*

Proof. Since $4d$ is an integer multiple of $4d_j$, it follows from Propositions 5.3 and 5.4 that W_{4d} contains $\mathbb{Z}(d_j)$, for all $j = 1, \ldots, k$. But W_{4d} is closed under addition and multiplication and so also must contain $\mathbb{Z}(d_1, \ldots, d_k)$. □

Proposition 5.6. *Let $d_1, \ldots, d_k$ be positive integers and let $d = \prod_{j=1}^{k} d_j$. Let $\alpha \in \mathbb{Z}(d_1, \ldots, d_k)$ have the representation $\alpha = \sum a_j(\sqrt{d_1})^{j_1}(\sqrt{d_2})^{j_2} \cdots (\sqrt{d_k})^{j_k}$, where a_j is an integer and the summation runs over all (distinct) k-tuples $J_j = (j_1, \ldots, j_k)$ with elements that are either 0 or 1. Then, $S_{4d}(\alpha) \leq d \sum |a_j|$.*

Proof. By the properties of representation height, we have

$$\begin{aligned} S_{4d}(\alpha) &= S_{4d}(\textstyle\sum a_j(\sqrt{d_1})^{j_1}(\sqrt{d_2})^{j_2} \cdots (\sqrt{d_k})^{j_k}) \\ &\leq \textstyle\sum |a_j| S_{4d}(\sqrt{d_1})^{j_1} S_{4d}(\sqrt{d_2})^{j_2} \cdots S_{4d}(\sqrt{d_k})^{j_k}. \end{aligned}$$

But the relation $W_{4d_j} \subset W_{4d}$ implies that $S_{4d}(\sqrt{d_j}) \leq S_{4d_j}(\sqrt{d_j}) \leq d_j$, where the last inequality follows from Proposition 5.4. Thus we have $S_{4d}(\alpha) \leq d \sum |a_j|$, as desired. □

We now use these results to bound the complexity of LPs whose coefficients belong to $\mathbb{Z}(d_1, \ldots, d_k)$. As in Sections 3 and 4, we assume that we have a machine that can perform arithmetic operations on real numbers in constant time per operation, and that the machine has available $\sqrt{d_j}$ for all j, or that it can calculate these numbers. We also assume that every number $\alpha \in \mathbb{Z}(d_1, \ldots, d_k)$ in a problem instance is encoded for input as a set of 2^k integers that are the coefficients in the representation of α in terms of the cross products of the $\sqrt{d_j}$. We define the *encoding size* of α to be the sum of the binary encoding sizes of these coefficients, and we define the encoding size of a matrix to be the sum of the encoding sizes of its entries.

Theorem 5.1. *Let (P) be the standard form LP defined at the beginning of section 3. Suppose that all the coefficients in (P) belong to $\mathbb{Z}(d_1, \ldots, d_k)$ for positive integers $d_1, \ldots, d_k$ with $d = \prod_{j=1}^{k} d_j$. Then (P) can be solved in time polynomial in n, d, and the encoding size of the matrix A.*

Proof. Using Proposition 5.6, it is easy to show that $L(A)$ is bounded by a polynomial in $\log d$ and the encoding size of the matrix A. Proof of the theorem then follows from Theorem 4.1. □

If d is fixed, Theorem 5.1 implies that (P) can be solved in time polynomial in the problem dimension and encoding size. We improve these results in a sequel paper [1], obtaining a bound which, although exponential in k, is polynomial in the bit size of d and the problem encoding size.

6. Remarks

In light of the results obtained here on the complexity of LPs with coefficients from W_p, it may be worth investigating what interesting classes of real numbers can be embedded in W_p. It is well-known that every finite Abelian extension of the rationals can be embedded in the extension of the rationals by the pth root of unity, for some p (see, e.g., [8]). Therefore, there may be other classes of LPs whose complexity can be bounded in the manner developed in Section 5 for LPs with coefficients from a square-root extension of the integers.

More ambitiously, one may ask if it is possible to obtain complexity results for other classes of algebraic numbers without first embedding the numbers in W_p. Recently, by using some additional material from number theory in conjunction with an approach similar to that of Sections 3 and 4, we have obtained such results for all algebraic numbers [1].

In some sense, our approach in this paper has been to encode a set of algebraic numbers as integers for input to a machine that performs real arithmetic. It is natural to ask whether the requirement for real arithmetic can be relaxed to the point where all computations are performed symbolically, using integer arithmetic only. In fact this can be done both here and in the more general context of an algorithm for all algebraic numbers. A report on this topic is under preparation [2].

In light of the results in [1], which extend the linear programming results given here to general algebraic numbers, it is worth investigating whether other classes of problems can be shown to be strongly polynomial by arguments similar to those used for circulant LPs. Equivalently, we can ask whether there are other simultaneously diagonalizable families of matrices whose diagonalizing matrix is composed of algebraic numbers (not necessarily from the subring W_p) that are small in the appropriate sense. Indeed, it appears that such families do exist and that considerable progress in identifying them can be made by using results from the theory of group representations. We plan to report on this topic in a subsequent paper.

Acknowledgement

This research was funded by the National Science Foundation under grant DMS8810192. We greatly appreciate this financial support.

References

[1] I. Adler and P.A. Beling, "Polynomial algorithms for linear programming over the algebraic numbers," *Proceedings of the 24th Annual ACM Symposium on the Theory of Computing* (1992) 483-494.

[2] I. Adler and P.A. Beling, "Rational arithmetic in linear programming over the algebraic numbers," manuscript (1991).

[3] P.A. Beling, "Linear programming over the algebraic numbers," Ph.D. dissertation, University of California (Berkeley, CA, 1991).

[4] L. Blum, M. Shub and S. Smale, "On a theory of computation and complexity over the real numbers; NP-completeness, recursive functions and universal machines," *Bulletin of the AMS* 21 (1989) 1-46.

[5] P.J. Davis, *Circulant Matrices* (Wiley, New York, 1979).

[6] J. Edmonds, "System of distinct representatives and linear algebra," *Journal of Research of the National Bureau of Standards Section B* 71 (1967) 241-245.

[7] R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).

[8] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory* (Springer, New York, 1972).

[9] N. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica* 4 (1984) 373-395.

[10] L. Khachiyan, "A polynomial algorithm in linear programming," *Soviet Mathematics Doklady* 20 (1979) 191-194.

[11] N. Megiddo, "Towards a genuinely polynomial algorithm for linear programming," *SIAM Journal on Computing* 12 (1983) 347-353.

[12] N. Megiddo, "Linear programming in linear time when the dimension is fixed," *Journal of the Association for Computing Machinery* 31 (1984) 114-127.

[13] N. Megiddo, "On solving the linear programming problem approximately," *Contemporary Mathematics* (1990).

[14] R.D.C. Monteiro and I. Adler, "Interior path following primal-dual algorithms. Part I: Linear programming," *Mathematical Programming* 44 (1989) 27-41.

[15] R.D.C. Monteiro and I. Adler, "Interior path following primal-dual algorithms. Part II: Convex quadratic programming," *Mathematical Programming* 44 (1989) 43-66.

[16] C. Norton, S. Plotkin and E. Tardos, "Using separation algorithms in fixed dimension," *Proceedings of the 1st ACM/SIAM Symposium on Discrete Algorithms* (1990) 377-387.

[17] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Prentice-Hall, Englewood Cliffs, NJ, 1982).

[18] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, New York, 1987).

[19] E. Tardos, "A strongly polynomial algorithm to solve combinatorial linear programs," *Operations Research* 34 (1986) 250-256.