Mathematical Programming 35 (1986) 225-235
North-Holland

ON THE EXPECTED NUMBER OF LINEAR COMPLEMENTARITY CONES INTERSECTED BY RANDOM AND SEMI-RANDOM RAYS

Nimrod MEGIDDO

The IBM Almaden Research Center, 650 Harry Road, San José, CA 95120, USA and Department of Statistics, Tel Aviv University, Tel Aviv, Israel

Received 30 December 1984
Revised manuscript received 21 August 1984

Lemke's algorithm for the linear complementarity problem follows a ray which leads from a certain fixed point (traditionally, the point $(1,\ldots,1)^{\mathrm{T}}$) to the point given in the problem. The problem also induces a set of 2^n cones, and a question which is relevant to the probabilistic analysis of Lemke's algorithm is to estimate the expected number of times a (semi-random) ray intersects the boundary between two adjacent cones. When the problem is sampled from a spherically symmetric distribution this number turns out to be exponential. For an n-dimensional problem the natural logarithm of this number is equal to $\ln(\tau)n+\mathrm{o}(n)$, where τ is approximately 1.151222. This number stands in sharp contrast with the expected number of cones intersected by a ray which is determined by two random points (call it *random*). The latter is only $(n/2)+1$. The discrepancy between linear behavior (under the 'random' assumption) and exponential behavior (under the 'semi-random' assumption) has implications with respect to recent analyses of the average complexity of the linear programming problem. Surprisingly, the semi-random case is very sensitive to the fixed point of the ray, even when that point is confined to the positive orthant. We show that for points of the form $(\varepsilon,\varepsilon^2,\ldots,\varepsilon^n)^{\mathrm{T}}$ the expected number of facets of cones cut by a semi-random ray tends to $\frac{1}{8}n^2+\frac{3}{8}n$ when ε tends to zero.

Key words: Linear complementarity, Lemke's algorithm, probabilistic analysis.

1. Introduction

The linear complementarity problem (LCP) is the following: Given $M\in\mathbb{R}^{n\times n}$ and $q\in\mathbb{R}^n$, find $z, w\in\mathbb{R}^n$ such that

$$w=Mz+q,\quad w^{\mathrm{T}}z=0,\quad z\geqslant 0\quad\text{and}\quad w\geqslant 0.$$

The LCP has received during the past fifteen years much attention within the mathematical programming community. The reader who is not familiar with it may consult [5] for background and references.

The linear programming problem can be formulated as an LCP and Lemke's algorithm, which we describe below, always solves it (whereas it may fail in general). Smale [10], [11] analyzed the average performance of Lemke's algorithm when

This work was done while the author was visiting Stanford University and XEROX-PARC and was supported in part by the National Science Foundation under grants MCS-8300984, ECS-8218181 and ECS-8121741.

applied to the linear programming problem. In the context of linear programming this algorithm is also known as the 'Self-Dual Simplex Algorithm' [4]. Smale's analysis yields an upper-bound which is quite different from bounds recently obtained by others [1, 6] for other related expected values. Murty [8] proved that the LCP may require an exponential number of steps in the worst-case. Some results on the average LCP are given in [7], [9] and [12]. Even though Lemke's algorithm does not always solve the linear complementarity problem, the results of the present paper suggest that its average number of steps is exponential. We show that the expected number of basic solutions relative to points on the line connecting the starting point e and the given vector q is exponential.

2. Background

We now describe briefly what is called 'Lemke's algorithm'. First, we note that the LCP may be interpreted as a problem of inverting a piecewise linear mapping $F:\mathbb{R}^n \to \mathbb{R}^n$ as follows. Given $x=(x_1,\ldots,x_n)^T$, let $x_i^+ = \max(x_i, 0)$, $x_i^- = \min(x_i, 0)$ $(i=1,\ldots,n)$ and $x^+=(x_1^+,\ldots,x_n^+)^T$ and $x^-=(x_1^-,\ldots,x_n^-)^T$. Given M, define

$$F(x) = Mx^+ + x^-.$$

It is easy to verify that by identifying x^+ with z and $-x^-$ with x we obtain an equivalent description of our LCP: Given M and q, find x such that $F(x) = -q$. If $q \geqslant 0$ then we may select $x=-q$ as a solution since F coincides with the identity on the negative orthant.

Let $e=(1,\ldots,1)^T \in \mathbb{R}^n$. The algorithm starts from $-e$ and inverts F along the line segment connecting $-e$ with $-q$. The algorithm succeeds only if $F^{-1}[(1-t)(-e)+t(-q)] \neq \emptyset$ for every $t \in [0,1]$. When the algorithm succeeds the effort is bounded by the number of orthants $R^S \subseteq \mathbb{R}^n$ ($R^s = \{x \in \mathbb{R}^n\colon x_i \geqslant 0$ if $i \in S$ and $x_i \leqslant 0$ otherwise$\}$) such that the line segment $[-e, -q]$ intersects the image of R^S under F. If the algorithm fails then the effort prior to the failure, namely, the number of pivot-steps, is bounded by the number of orthants crossed by the inverse image up to that point. Smale [10] argued that the number of these orthants is equal to the number of *facets* of orthants, whose images under F intersect $[-e, -q]$, plus one. Now, the restriction of F to R^S is linear with an underlying matrix M^S defined as follows. The ith column of M^S is equal to the ith column of M if $i \in S$; otherwise, the ith column of M^S is equal to $-e^i$ (i.e., a unit n-vector with a negative unity in the ith position). As observed by Smale, $F(R^S) \cap (-e, -q] \neq \emptyset$ if and only if $-q$ belongs to the cone spanned by the columns of M^S together with e.

An *orthant-facet* $R^{S,i}$ is defined by $R^{S,i} = R^S \cap \{x \in \mathbb{R}^n\colon x_i = 0\}$. Obviously, $F(R^{S,i}) \cap (-e, -q] \neq \emptyset$ if and only if $-q$ belongs to the cone spanned by the columns of M^S with e replacing the ith column of M^S. We denote by $\mathrm{con}(C)$ a cone spanned by a set $C \subseteq \mathbb{R}^n$ and by $\mathrm{lin}(C)$ the linear subspace spanned by C. Also, we identify any matrix with the set of its columns, so that our operators con and lin are well-defined for matrices.

Our goal will be to evaluate the average number of orthant-facets whose images meet the line segment $(-e, -q]$, under the probabilistic model used by Smale. In this model the vector q and the matrix M are sampled (independently) from spherically symmetric distributions on $\mathbb{R}^n$ and $\mathbb{R}^{n\times n}$, respectively. The radial parts of these distributions are immaterial for our purposes here. It is thus convenient to assume that the entries of both M and q are sampled independently from the standard normal distribution. We denote the expected number of orthant-facets discussed above by $\rho_{\mathrm{lcp}}(n)$. It turns out that $\rho_{\mathrm{lcp}}(n)$ is very sensitive to the starting point of the ray. It is exponential if the starting poing is $-e$, whereas it is only quadratic if the starting point is of the form $(-\varepsilon, -\varepsilon^2, \ldots, -\varepsilon^n)^{\mathrm{T}}$, for ε sufficiently small. A result equivalent to the latter was also obtained by Todd [12].

3. The exponential result

Definition. Let $C \subseteq \mathbb{R}^n$. (i) The *volume*, $V(C)$, of C is defined to be the probability that a vector x, drawn from a spherically symmetric distribution on $\mathbb{R}^n$, belongs to con C. (ii) The *relative volume*, rel $V(C)$, of C is defined to be the probability that a vector x, drawn from a spherically symmetry distribution on lin C, belongs to con C.

The following lemma is also proved in [7].

Lemma 1. *Let $C \subseteq \mathbb{R}^n$ be such that* lin C *has dimension less than n, and let $u \in \mathbb{R}^n$ be a random vector (i.e., u is drawn rom a spherically symmetric distribution on $\mathbb{R}^n$). Under these conditions,*

$$\mathscr{E}(\text{rel } V(C \cup \{u\})) = \tfrac{1}{2} \text{rel } V(C)$$

where $\mathscr{E}$ denotes expectation.

Proof. Let $C^+ = C \cup \{u\}$ and $C^- = C \cup \{-u\}$. Obviously,

$$\mathscr{E}(\text{rel } V(C^+)) = \varepsilon(\text{rel } V(C^-)).$$

Now, the probability that a random x in $\text{lin}(C \cup \{u\})$ belongs to C is zero. It is therefore sufficient to prove that

$$\mathscr{E}(\text{rel } V(C \cup \{u, -u\})) = \text{rel } V(C).$$

Let $\{b^1, \ldots, b^k\}$ be an orthonormal basis for lin C and let b^{k+1} be a unit vector in $\text{lin}(C \cap \{u\})$ which is orthogonal to lin C. Random vectors in $\text{lin}(C \cup \{u\})$ can be generated by selecting their coefficients relative to $\{b^1, \ldots, b^{k+1}\}$ independently from the standard normal distribution. In particular, let $(u_1, \ldots, u_{k+1})$ and $(x_1, \ldots, x_{k+1})$ represent u and another random vector x, respectively. Now, $c \in \text{con}(C \cup \{u, -u\})$ if and only if there is a $v \in \text{con } C$ and a real λ such that $x = \lambda u + v$.

Necessarily, $\lambda = x_{k+1}/u_{k+1}$ (or else $x \notin \text{lin}\, C$) so that $x \in \text{lin}(C \cup \{u\})$ belongs to $\text{con}(C \cup \{u\})$ if and only if $x - (x_{k+1}/u_{k+1})u \in \text{con}\, C$. However, the vector $x - (x_{k+1}/u_{k+1})u$ is spherically symmetrically distributed in lin C and the probability of that event is hence equal to rel $V(C)$. □

We denote by $\Phi(t)$ the standard normal cumulative distribution function and by $\phi(t)$ the standard normal density function.

Lemma 2

$$\text{rel } V(-e, e^1, \ldots, e^k) = \sqrt{\frac{n-k}{2\pi}} \int_0^\infty (\Phi(t))^k (\mathrm{e}^{-(1/2)t^2})^{n-k}\, \mathrm{d}t.$$

Proof. A random vector $x \in \mathbb{R}^n$ belongs to $\text{lin}(-e, e^1, \ldots, e^k)$ if and only if $x_{k+1} = \cdots = x_n$. Given that x satisfies these conditions, x belongs to $\text{con}(-e, e^1, \ldots, e^k)$ if and only if $x_{k+1} \leqslant 0$ and $x_i \geqslant x_{k+1}$ for $i = 1, \ldots, k$. Assuming that $x_1, \ldots, x_n$ are independent standard normal variates, we claim the following: (i) The conditional distribution of x_{k+1}, given that $x_{k+1} = \cdots = x_n$, is normal with mean zero and variance $1/(n-k)$; this follows from the fact that the norm of the vector $(x_{k+1}, \ldots, x_n)$ (when $x_{k+1} = \cdots = x_n$) is equal to $\sqrt{n-k}\,x_{k+1}$ and the (signed) distance between the origin and the point $(x_{k+1}, \ldots, x_n)$ has the standard normal distribution (given $x_{k+1} = \cdots = x_n$). (ii) The conditional distributions of $x_1, \ldots, x_k$ (given the same condition) are standard normal and $x_1, \ldots, x_k, x_{k+1}$ remain independent under the condition $x_{k+1} = \cdots = x_n$. Thus, given that $x_{k+1} = t$, the probability that $x_i \geqslant x_{k+1}$ for $i = 1, \ldots, k$ is equal to $(1 - \Phi(t))^k$ and integration over the negative domain of x_{k+1} yields

$$\text{rel } V(-e, e^1, \ldots, e^k) = \frac{1}{\sqrt{2\pi}(1/\sqrt{n-k})} \int_{-\infty}^0 (1 - \Phi(t))^k \mathrm{e}^{-t^2/2/(n-k)}\, \mathrm{d}t$$

from which our lemma follows easily. □

Corollary. *Let C be a cone spanned by $\{-e, e^1, \ldots, e^k\}$ together with $n-k-1$ more random vectors in $\mathbb{R}^n$. Then*

$$\mathscr{E}V(C) = \frac{1}{2^{n-k-1}} \sqrt{\frac{n-k}{2\pi}} \int_0^\infty (\Phi(t))^k (\mathrm{e}^{-(1/2)t^2})^{n-k}\, \mathrm{d}t,$$

where $\mathscr{E}$ denotes the expectation operator.

We can now describe the expected number of cones as follows.

Theorem 1

$$\rho_{\text{lcp}}(n) = \sum_{k=0}^{n-1} n \binom{n-1}{k} \frac{1}{2^{n-k-1}} \sqrt{\frac{n-k}{2\pi}} \int_0^\infty (\Phi(t))^k (\mathrm{e}^{-(1/2)t^2})^{n-k}\, \mathrm{d}t.$$

Proof. We evaluate the number of orthant-facets $R^{S,i}$ whose images intersect the line segment $(-e, -q]$. Without loss of generality assume $i \notin S$. The probability that the image of $R^{S,i}$ intersects $(-e, -q]$ is equal to the volume of a set like the one described in the Corollary (with $k = |S|$). Taking the sum of volumes of all possible sets like that establishes the proof. □

For estimating the asymptotic behavior of $\rho_{\mathrm{lcp}}(n)$ we define the following function:

$$G(n) = \int_0^\infty (\Phi(t) + \tfrac{1}{2}\mathrm{e}^{-(1/2)t^2})^{n-1}\phi(t)\,\mathrm{d}t.$$

Consider, first, the function

$$g(t) = \Phi(t) + \tfrac{1}{2}\mathrm{e}^{-(1/2)t^2}.$$

Lemma 3. (i) $g(0) = g(\infty) = 1$. (ii) $g(t) > 1$ *for every* $t > 0$. (ii) $g(t)$ *has a unique local maximum at* $t = \sqrt{2/\pi}$.

Proof. The proof is immediate by observing the derivative

$$g'(t) = \phi(t)(1 - \tfrac{1}{2}\sqrt{2\pi}\,t).$$

Denote

$$\tau = g\left(\sqrt{\frac{2}{\pi}}\right) = \Phi\left(\sqrt{\frac{2}{\pi}}\right) + \tfrac{1}{2}\mathrm{e}^{-1/\pi} \approx 1.151222.$$

Lemma 4. (i) *For every* ε $(0 < \varepsilon < \tau - 1)$ *there is a consant* $c = c(\varepsilon)$ *such that* $G(n) > c(\varepsilon)(\tau - \varepsilon)^{n-1}$. (ii) $G(n) < \frac{1}{2}\tau^{n-1}$.

Proof. (i) Given ε, let $a_1 < a_2$ be the positive numbers such that $g(a_1) = g(a_2) = \tau - \varepsilon$ whose existence follows from lemma 3. Let $c(\varepsilon) = \Phi(a_2) - \Phi(a_1)$. Thus

$$G(n) > \int_{a_1}^{a_2} (g(t))^{n-1}\phi(t)\,\mathrm{d}t > (g(a_1))^{n-1}\int_{a_1}^{a_1} \phi(t)\,\mathrm{d}t = c(\varepsilon)(\tau - \varepsilon)^{n-1}.$$

(ii) Since the maximum of $g(t)$ is attained at $\sqrt{2/\pi}$ it follows that

$$G(n) < \left(g\left(\sqrt{\frac{2}{\pi}}\right)\right)^{n-1}\int_0^\infty \phi(t)\,\mathrm{d}t = \tfrac{1}{2}\tau^{n-1}. \qquad \square$$

Theorem 2. $\ln(\rho_{\mathrm{lcp}}(n)) = \ln(\tau)n + \mathrm{o}(n)$.

Proof. We know that

$$\rho_{\mathrm{lcp}}(n)=\sum_{k=0}^{n-1} n\binom{n-1}{k}\frac{1}{2^{n-k-1}}\sqrt{\frac{n-k}{2\pi}}\int_0^\infty (\Phi(t))^k(\mathrm{e}^{-(1/2)t^2})^{n-k}\,\mathrm{d}t$$

$$=n\int_0^\infty \sum_{k=0}^{n-1}\sqrt{n-k}\binom{n-1}{k}(\Phi(t))^k(\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2})^{n-1-k}\phi(t)\,\mathrm{d}t.$$

Thus,

$$\rho_{\mathrm{lcp}}(n)<n^{3/2}\int_0^\infty (\Phi(t)+\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2})^{n-1}\phi(t)\,\mathrm{d}t=n^{3/2}G(n)$$

and, similarly,

$$\rho_{\mathrm{lcp}}(n)>nG(n).$$

The theorem now follows from Lemma 4. □

4. What is a good starting point?

We can easily generalize the results of the previous section to any starting point, but it should be noticed that only points in the negative orthant can serve as starting points for the actual algorithm. We will later see which of these points are best.

We first compute rel $V(-a, e^{i_1},\dots,e^{i_k})$ where a is any n-vector (but only positive ones are meaningful for the algorithm). Let $N=\{1,\dots,n\}$.

Lemma 5. *Let $S=\{i_1,\dots,i_s\}$ be a proper subset of N $(s<n)$. Under these conditions*

$$\operatorname{rel} V(-a, e^{i_1},\dots,e^{i_s})=\sqrt{\frac{\sum_{i\notin S} a_i^2}{2\pi}}\int_0^\infty \prod_{i\in S}\Phi(a_i t)\prod_{i\notin S}\mathrm{e}^{-(1/2)t^2a_i^2}\,\mathrm{d}t.$$

Proof. The proof is a straightforward generalization of that of Lemma 2. A random vector $x\in\mathbb{R}^n$ belongs to $\operatorname{lin}(-a, e^{i_1},\dots,e^{i_s})$ if and only if $(x_i)_{i\notin S}$ is proportional to $(a_i)_{i\notin S}$. Suppose x satisfies this condition and let λ denote the coefficient of proportionality. Now, x belongs to $\operatorname{con}(-a, e^{i_1},\dots,e^{i_s})$ if and only if $\lambda\le 0$ and $x_i\ge\lambda a_i$ for $i\in S$. The conditional density function of λ, given that $x_i=\lambda a_i$ $(i\notin S)$ is

$$f(\lambda)=\frac{\prod\phi(\lambda a_i)}{\int_{-\infty}^{\infty}\prod_{i\notin S}\phi(\lambda a_i)\,\mathrm{d}\lambda}=\sqrt{\frac{\sum_{i\notin S}a_i^2}{2\pi}}\prod_{i\notin S}\mathrm{e}^{-(1/2)t^2a_i^2}.$$

The conditional distributions of $x_{i_1},\dots,x_{i_s}$ (given λ) are standard normal. Given $\lambda<0$, probability that all the other coefficients are nonnegative is simply

$\prod_{i\in S}\Phi(-\lambda a_i)$. Integration of this quantity (multiplied by the conditional density of λ) over the negative values of λ proves our lemma. $\square$

Corollary. *Let C be a cone spanned by $\{-a, e^{i_1}, \ldots, e^{i_s}\}$ $(s<n)$ together with $n-s-1$ more random vectors in $\mathbb{R}^n$. Then*

$$\mathscr{E}V(C)=2\sqrt{\frac{\sum_{i\notin S} a_i^2}{2\pi}}\int_0^\infty \prod_{i\in S}\Phi(a_i t)\prod_{i\notin S}\tfrac{1}{2}\,\mathrm{e}^{-(1/2)t^2a_i^2}\,\mathrm{d}t.$$

We can now write a formula for the expected number of orthant-facets, met by a ray starting at $-a$ and leading to a random point $-q$. In fact all we have to do is sum up the previously computed relative volumes, over all *proper* subsets S of N, each multiplied by $n-|S|$; the latter follows from the fact that given the choice of the unit vectors (that is, the set S), we still have $n-|S|$ different ways to choose the position of $-a$ within the matrix. Thus,

Theorem 3

$$\rho_{\mathrm{lcp}}(n)=2\sum_{\substack{S\subseteq N\\ S\neq N}}(n-|S|)\sqrt{\frac{\sum_{i\notin S} a_i^2}{2\pi}}\int_0^\infty \prod_{i\in S}\Phi(a_i t)\prod_{i\notin S}\tfrac{1}{2}\,\mathrm{e}^{-(1/2)t^2a_i^2}\,\mathrm{d}t.$$

We note that $\rho_{\mathrm{lcp}}(n)$ is homogeneous as a function of the vector a. Moreover, it is symmetric in the components of a. Assume, without loss of generality that $a_1\geqslant a_2\geqslant\cdots\geqslant a_n$. The following identity is easy to verify:

$$\sum_{\substack{S\subseteq N\\ S\neq N}}\prod_{i\in S}A_i\prod_{i\notin S}B_i=\sum_{i=1}^n\left\{B_i\prod_{j=1}^{i-1}A_j\prod_{j=i+1}^{n}(A_j+B_j)\right\}.$$

We will apply this identity with $A_i=\Phi(a_i t)$ and $B_j=\frac{1}{2}\mathrm{e}^{-(1/2)t^2a_j^2}$. We get the following lower bound for $\rho_{\mathrm{lcp}}(n)$:

$$\begin{aligned}\rho_{\mathrm{lcp}}(n)&\geqslant 2\sum_{\substack{S\subseteq N\\ S\neq N}}(n-|S|)\frac{\max_{i\notin S}a_i}{\sqrt{2\pi}}\int_0^\infty\prod_{i\in S}\Phi(a_it)\prod_{i\notin S}\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_i^2}\,\mathrm{d}t\\ &\leqslant 2\sum_{i=1}^n\left\{\frac{a_i}{\sqrt{2\pi}}\int_0^\infty\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_i^2}\prod_{j=1}^{i-1}\Phi(a_jt)\prod_{j=i+1}^{n}(\Phi(a_jt)+\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_j^2})\,\mathrm{d}t\right\}.\end{aligned}$$

Note that on the other hand we have a close upper bound:

$$\begin{aligned}\rho_{\mathrm{lpc}}(n)&\leqslant 2\sum_{\substack{S\subseteq N\\ S\neq N}}(n-|S|)\frac{\sqrt{n-|S|}\,\max_{i\notin S}a_i}{\sqrt{2\pi}}\int_0^\infty\prod_{i\in S}\Phi(a_it)\prod_{i\notin S}\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_i^2}\,\mathrm{d}t\\ &\leqslant 2\sum_{i=1}^n\left\{\frac{(n-i)^{3/2}a_i}{\sqrt{2\pi}}\int_0^\infty\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_i^2}\prod_{j=1}^{i-1}\Phi(a_jt)\prod_{j=i+1}^{n}(\Phi(a_jt)+\tfrac{1}{2}\mathrm{e}^{-(1/2)t^2a_j^2})\,\mathrm{d}t\right\}.\end{aligned}$$

The upper and the lower bounds differ by a factor of $n^{3/2}$ so that from a point of view of polynomial versus exponential we have a sharp estimate. Consider the following functions $J_i(n)$ for $i=1,\dots,n$:

$$J_i(n)=\frac{2a_i}{\sqrt{2\pi}}\int_0^\infty \tfrac{1}{2}e^{-(1/2)t^2a_i^2}\prod_{j=1}^{i-1}\Phi(a_jt)\prod_{j=i+1}^{n}(\Phi(a_jt)+\tfrac{1}{2}e^{-(1/2)t^2a_j^2})\,dt$$

$$=\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-(1/2)t^2}\prod_{j=1}^{i-1}\Phi\left(\frac{a_j}{a_i}t\right)\prod_{j=i+1}^{n}\left\{\Phi\left(\frac{a_j}{a_i}t\right)+\tfrac{1}{2}e^{-(1/2)t^2(a_j/a_i)^2}\right\}dt.$$

We have now come closer to understanding the effect of the starting point a. The integral can be interpreted as a weighted average of a product of $n-1$ factors. We know from Section 3 that each factor of the form $\Phi(a_jt)+\frac{1}{2}e^{-(1/2)t^2a_j^2}$ is a function of t which is always grater than 1 and attains its maximum at the point where $a_jt=\sqrt{2/\pi}$. Thus, the worst case is when all the a_j's are equal, that is $a=e$. In that case the maxima of the factors coincide, yielding an exponential integrand. On the other hand, each of the factors tends to 1 when t tends either to ∞ or to 0. Attempting to achieve maximum separation among the peaks of the factors, it is now clear that we should choose the a_i's so that the ratio of any two of them is very large.

We are now led naturally to the choice of $a_j=\varepsilon^j$, $j=1,\dots,n$, subject to which we obtain

$$J_i(n)=\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-(1/2)t^2}\prod_{j=1}^{i-1}\Phi(\varepsilon^{j-i}t)\prod_{j=i+1}^{n}(\Phi(\varepsilon^{j-i}t)+\tfrac{1}{2}e^{-(1/2)t^2\varepsilon^{2(j-1)}})\,dt.$$

Consider the latter integral over an interval $[\varepsilon^{0.5},\varepsilon^{-0.5}]$, letting ε tend to zero. Suppose $\varepsilon^{0.5}\leqslant t\leqslant\varepsilon^{-0.5}$. For $j\leqslant i-1$ we have $\varepsilon^{j-i}t\geqslant\varepsilon^{-0.5}$, while for $j\geqslant i+1$ we have $\varepsilon^{j-i}t\leqslant\varepsilon^{0.5}$. It follows that $J_i(n)$ simply tends to $\frac{1}{2}$ as ε tends to zero. Moreover, this behavior is independent of the distribution being normal. Note that in the limit all that counts is that $\Phi(0)=\frac{1}{2}$ and $\phi(0)$ is finite. Returning to the expression in Theorem 3, we can now argue about the limit $\rho^0(n)$ of $\rho_{\text{lcp}}(n)$, when ε tends to zero, as follows. For each proper subset $S\subseteq N$, let $i(S)$ denote the smallest i such that $i\notin S$. The contribution of S to $\rho^0(n)$ depends only on $i(S)$ and the cardinality of S. Specifically, this contribution is asymptotically equal to

$$(n-|S|)\frac{\varepsilon^{i(S)}}{\sqrt{2\pi}}\int_{\varepsilon^{0.5}}^{\varepsilon^{-0.5}} e^{-(1/2)t^2\varepsilon^{2i(S)}}\prod_{j\in S}\Phi(\varepsilon^jt)\prod_{\substack{j\notin S\\ j\neq i(S)}}\tfrac{1}{2}e^{-(1/2)t^2\varepsilon^{2j}}\,dt.$$

Asymptotically, every $j\geqslant i(S)$ contributes a factor of $\frac{1}{2}$ whereas every $j<i(S)$ contributes a factor of 1. Thus, the contribution of S in the limit is $(n-|S|)2^{-(n-i(S)+1)}$. It follows that

$$\rho^0(n)=\sum_{\substack{S\subseteq N\\ S\neq N}}(n-|S|)2^{-(n-i(S)+1)}$$

$$=\sum_{i=1}^{n}2^{-(n-i+1)}\sum_{s=i-1}^{n-1}(n-s)\binom{n-i}{s-i+1}$$

$$= \sum_{j=0}^{n-1} 2^{-(j+1)} \sum_{k=0}^{j} (k+1)\binom{j}{k}$$

$$= \tfrac{1}{2} \sum_{j=0}^{n-1} \left(\frac{j}{2}+1\right) = \tfrac{1}{8}n^2 + \tfrac{3}{8}n.$$

We have thus established the following

Theorem 4. *The expected number of times, which a semi-random ray starting at the point* $(\varepsilon, \varepsilon^2, \ldots, \varepsilon^n)^{\mathrm{T}}$ *intersects facets of linear complementarity cones, tends to* $\frac{1}{8}n^2+\frac{3}{8}n$ *when* ε *tends to zero.*

We note that an equivalent result was also obtained by Todd [12]. It follows that the number of cones intersected by a semi-random ray in the latter case is not greater than $\frac{1}{4}n^2+\frac{3}{4}n$ since in the extreme case the inverse image of the ray may consist of paths crossing two adjacent orthants and the facet between them.

5. Discussion

Suppose that instead of a fixed vector a we take a random vector $u \in \mathbb{R}^n$. We may now ask what is the expected number of orthant-facets whose images under F intersect the line segment $(-u, -q]$. Is the latter a good approximation to $\rho_{\mathrm{lcp}}(n)$? This question is interesting since such an approximation argument has been suggested recently by several people with reference to the average numer of steps for solving linear programming problems. It turns out that replacing the fixed vector a by random vector u simplifies the analysis considerably both in the LCP context and in the linear programming context. However, the question of evaluating $\rho(m, n)$ (the average number of steps for solving a linear programming problem of dimension $m \times n$) is still open and there is still a debate whether replacing e by a random u yields a good approximation.

Now, for each orthant-facet $R^{S,i}$ the probability of its image intersecting $(-u, -q]$ is equal to the expected volume of a set of the form $\{-u, e^1, \ldots, e^{k-1}, v^1, \ldots, v^{n-k}\}$, where $v^1, \ldots, v^{n-k}$ are random vectors. Obviously,

$$\operatorname{rel} V(e^1, \ldots, e^{k-1}) = \frac{1}{2^{k-1}}$$

and by Lemma 1 we get

$$\mathscr{E}(V(-u, e^1, \ldots, e^{k-1}, v^1, \ldots, v^{n-k})) = \frac{1}{2^n}.$$

Since there are precisely $n2^{n-1}$ orthant-facets, it follows that the expected number of orthant-facets whose images intersect a random line segment (i.e., one whose endpoints are sampled independently from a spherically symmetric distribution) is equal to

$$(n2^{n-1})\frac{1}{2^n}=\frac{n}{2}.$$

The latter is also proved in [9]. Thus, 'symmetrizing' yields considerably different results. We may remark here that the same 'symmetizing' trick yields for a linear programming problem of order $m \times n$ the quantity

$$(m+n)\binom{m+n}{m}\frac{1}{2^{m+n}}$$

as the expected number of orthant-facets whose images intersect a random line segment. This follows easily by observing that under this assumption the probability associated with any orthant-facet is precisely $2^{-(m+n)}$. Similar results are in [1] and [6]. It is not inconceivable that the expected number of steps it takes the self-dual algorithm (starting at $(1, \ldots, 1)^T$) to solve the linear programming problem is also exponential. However, it is now known (recently observed by Adler, Megiddo and Todd [2, 3, 12]) that starting points of the form $(1, \varepsilon, \varepsilon^2, \ldots)$ yield quadratic performance for the linear programming problem.

References

[1] I. Adler, "The expected number of pivots needed to solve parametric linear programs and the efficiency of the Self-Dual Simplex method", Technical Report, Department of Industrial Engineering and Operations Research, University of California (Berkeley, CA, June 1983).

[2] I. Adler and N. Megiddo, "A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension", in: *Proceedings of the 16th Annual ACM Symposium on Theory of Computing* (ACM, New York, 1984). Also: *Journal of ACM,* to appear.

[3] I. Adler, N. Megiddo and M.J. Todd, "New results on the average behavior of simplex algorithms", *Bulletin of the AMS* 11 (1984) 378-381.

[4] G.B. Dantzig, *Linear programming and extensions* (Princeton University Press, Princeton, NJ, 1963).

[5] B.C. Eaves and H. Scarf, "The solution of systems of piecewise linear equations", *Mathematics of Operations Research* 1 (1976) 1-27.

[6] M. Haimovich, "The simplex algorithm is very good!—On the expected number of pivot steps and related properties of random linear programs", manuscript, Columbia University (New York, NY, April 1983).

[7] R. Howe, "Linear complementarity and the average volume of simplicial cones", Cowles Foundation Discussion Paper No. 670, Yale University, (New Haven, CT, June 1983).

[8] K.G. Murty, "Computational complexity of complementary pivot methods", *Mathematical Programming Studies* 7 (1978) 61-73.

[9] R. Saigal, "On some average results for random complementarity problems", Technical Report, Department of Industrial Engineering, Northwestern University (Evanston, IL, June 1983).

[10] S. Smale, "On the average number of steps of the simplex method of linear programming", *Mathematical Programming* 27 (1983) 241-262.

[11] S. Smale, "The problem of the average speed of the simplex method", in: A. Bachem, M. Grötschel and B. Korte, eds, *Mathematical programming: The state of the art* (Springer-Verlag, 1983) pp. 530-539.

[12] M.J. Todd, "Polynomial expected behavior of a pivoting algorithm for linear complementarity and linear programming problems", Techical Report No. 595, School of Operations Research and Industrial Engineering, Cornell University (Ithaca, NY, November 1983).