3

THE SHAPE OF THE SYMMETRIC SOLUTION SET

Götz Alefeld*, Vladik Kreinovich, and Günter Mayer*****

**Institut für Angewandte Mathematik, Universität Karlsruhe, D-76128Karlsruhe, Germany, email goetz.alefeld@mathematik.uni-karlsruhe.de*

***Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, email vladik@cs.utep.edu*

****Fachbereich Mathematik, Universität Rostock, D–18051 Rostock, Germany, email guenter.mayer@mathematik.uni-rostock.de*

ABSTRACT

We give a new deduction of the set of inequalities which characterize the solution set S of real linear systems $Ax = b$ with the $n \times n$ coefficient matrix A varying between a lower bound $\underline{A}$ and an upper bound $\overline{A}$, and with b similarly varying between $\underline{b}$ and $\overline{b}$. The idea of this deduction can also be used to construct a set of inequalities which describe the so–called symmetric solution set S_{sym}, i.e., the solution set of $Ax = b$ with $A = A^T$ varying between the bounds $\underline{A} = \underline{A}^T$ and $\overline{A} = \overline{A}^T$. This is the main result of our paper. We show that in each orthant S_{sym} is the intersection of S with sets of which the boundaries are quadrics.

1 INTRODUCTION

In [2], [3], [4], [5], [6], [7], and others, the intersection $S \cap O$ of the *solution set*

$$S := \{x \in \mathbf{R}^n \mid Ax = b,\ A \in [A],\ b \in [b]\} \tag{3.1}$$

with an orthant O is characterized by a set of inequalities which describe half–spaces. Here, $[A]$ denotes an $n \times n$ matrix with real compact intervals as entries, $[b]$ is a given vector with n real compact intervals as components, and $[A]$ is assumed to be *regular*, i.e., it contains only regular matrices as elements. In [2] it is also shown that for regular *symmetric* interval matrices $[A] = [A]^T$ the

R. B. Kearfott and V. Kreinovich (eds.), Applications of Interval Computations, 61–79.

symmetric solution set

$$S_{\text{sym}} := \{x \in \mathbf{R}^n \mid Ax = b,\ A = A^T \in [A],\ b \in [b]\} \subseteq S \tag{3.2}$$

is compact and connected. For 2×2 interval matrices, the intersection $S_{\text{sym}} \cap O$ was characterized as an intersection of S with sets of which the boundaries are quadrics. The proof could, however, not be generalized to the $n \times n$ case, $n > 2$. In the present paper, we handle this general problem by another technique based on results of Tarski [9] and Seidenberg [8]. In addition, we dispense with the regularity of $[A]$ in the definition of S and S_{sym}. It is obvious that these sets then no longer need to be bounded. First, we apply the new technique to obtain the description of $S \cap O$ by the above-mentioned set of inequalities. Then we show that the same technique can be used to characterize $S_{\text{sym}} \cap O$ as intersection of S with sets of which the boundaries are quadrics. This is the generalization of the results from [2] to the case $n > 2$. The proof is constructive, i.e., all the inequalities can be explicitly derived. However, their number $n_{\#}$ increases tremendously with the number n of rows and columns of $[A]$. Therefore, we restrict ourselves to describing the way in which to derive them. An optimal bound for $n_{\#}$ is given in the particular case of symmetric tridiagonal $n \times n$ matrices $[A]$, which can also be viewed as generalizations of 2×2 matrices.

We conclude our paper with several examples which illustrate the theory.

2 NOTATION

By $\mathbf{R}^n$, $\mathbf{R}^{n\times n}$, $\mathbf{IR}$, $\mathbf{IR}^n$, $\mathbf{IR}^{n\times n}$ we denote the set of real vectors with n components, the set of real $n \times n$ matrices, the set of intervals, the set of interval vectors with n components and the set of $n \times n$ interval matrices, respectively. By 'interval' we always mean a real compact interval. Interval vectors and interval matrices are vectors and matrices, respectively, with interval entries. We write intervals in brackets with the exception of degenerate intervals (so-called *point intervals*) which we identify with the element being contained, and we proceed similarly with interval vectors and interval matrices. As usual, we denote the lower and upper bound of an interval $[a]$ by $\underline{a}$ and $\overline{a}$, respectively. Similarly, we use $\underline{A}$ and $\overline{A}$ for the bounds of interval matrices $[A] \in \mathbf{IR}^{n\times n}$. We write $[A] = [\underline{A}, \overline{A}] = ([a_{ij}]) = ([\underline{a}_{ij}, \overline{a}_{ij}]) \in \mathbf{IR}^{n\times n}$ simultaneously, without further reference. By $[A]^T$ we mean the transposed matrix of $[A]$. We mention that $[A] = [A]^T$ is equivalent to $\underline{A} = \underline{A}^T$ and $\overline{A} = \overline{A}^T$.

For computations with interval quantities we refer to [1].

We denote any orthant of $\mathbf{R}^n$ by O, and the first orthant by O_1.

3 RESULTS

3.1 A Fundamental Lemma

We start this section with the following lemma, which goes back to results of Tarski [9] and Seidenberg [8]. This lemma is basic for all our subsequent considerations. In particular, it contains the constructive procedure which leads to the inequalities mentioned in Section 1.

Lemma 1. *Let $L_i = L_i(z_1, \ldots, z_n)$, $i = 1, \ldots, m$, be m given linear functions of n real variables $z_1, \ldots, z_n$:*

$$L_i(z_1, \ldots, z_n) = c_i + \sum_{j=1}^{n} l_{ij} z_j \ , \quad i \in \{1, \ldots, m\}.$$

Without loss of generality, let

$$\begin{aligned} l_{i1} &= 0 \quad \text{if} \quad i \in M^0 := \{1, \ldots, m_1\}, \\ l_{i1} &< 0 \quad \text{if} \quad i \in M^- := \{m_1 + 1, \ldots, m_2\}, \\ l_{i1} &> 0 \quad \text{if} \quad i \in M^+ := \{m_2 + 1, \ldots, m\}, \end{aligned}$$

where at least one of the sets M^0, M^-, M^+ is not empty. If $M^0 = \emptyset$ we set $m_1 := 0$, if $M^- = \emptyset$ we set $m_2 := m_1$, if $M^+ = \emptyset$ we set $m_2 := m$.

Let $\hat{m} := m_1 + (m_2 - m_1)(m - m_2)$. If $\hat{m} > 0$ then construct linear functions $\hat{L}_i := \hat{L}_i(z_2, \ldots, z_n)$, $i = 1, \ldots, \hat{m}$, in the following way:

i) If $M^0 \neq \emptyset$ then let

$$\hat{L}_i := L_i, \ i = 1, \ldots, m_1.$$

ii) If $M^- \neq \emptyset$ and $M^+ \neq \emptyset$ then for $p \in M^-$, $q \in M^+$ and $i = p + (q - m_2 - 1)(m_2 - m_1)$ (i.e., $p = m_1 + 1, \ldots, m_2$, $q = m_2 + 1, \ldots, m$, and therefore $i = m_1 + 1, \ldots, m_2 + (m - m_2 - 1)(m_2 - m_1) = \hat{m}$) let

$$\hat{L}_i := \frac{c_q}{l_{q1}} - \frac{c_p}{l_{p1}} + \sum_{j=2}^{n} \left(\frac{l_{qj}}{l_{q1}} - \frac{l_{pj}}{l_{p1}} \right) z_j.$$

No $\hat{L}_i$ is constructed if $\hat{m} = 0$.

Let the real numbers $\tilde{z}_2, \dots \tilde{z}_n$ be given. Then the following two assertions are equivalent.

a) There exists a real number $\tilde{z}_1$ such that

$$L_i(\tilde{z}_1, \dots, \tilde{z}_n) \leq 0, \quad i = 1, \dots, m.$$

b) The inequalities

$$\hat{L}_i(\tilde{z}_2, \dots, \tilde{z}_n) \leq 0, \quad i = 1, \dots, \hat{m}, \tag{3.3}$$

hold, where we define (3.3) to be true in the case $\hat{m} = 0$, i.e., if no condition $\hat{L}_i(\tilde{z}_2, \dots, \tilde{z}_n) \leq 0$ exists.

Remark. The procedure of constructing the functions $\hat{L}_i$ starting from the functions L_i is called an *elimination of* z_1.

Proof of Lemma 1.
a) $\Rightarrow$ b) : Let $L_i(\tilde{z}_1, \dots, \tilde{z}_n) \leq 0$ hold for $i = 1, \dots, m$. Then $\hat{L}_i(\tilde{z}_2, \dots, \tilde{z}_n) = L_i(\tilde{z}_1, \dots, \tilde{z}_n) \leq 0$ for $i = 1, \dots, m_1$. If $M^- = \emptyset$ and $M^+ = \emptyset$, respectively, then $m_2 = m_1$ and $m = m_2$, respectively; hence $\hat{m} = m_1$, and b) follows.

Assume now $M^- \neq \emptyset$ and $M^+ \neq \emptyset$. Let $p \in M^-$. Then $l_{p1} < 0$; hence $L_p(\tilde{z}_1, \dots, \tilde{z}_n) \leq 0$ is equivalent to

$$-\frac{c_p}{l_{p1}} - \sum_{j=2}^{n} \frac{l_{pj}}{l_{p1}} \tilde{z}_j \leq \tilde{z}_1 . \tag{3.4}$$

Analogously,

$$\tilde{z}_1 \leq -\frac{c_q}{l_{q1}} - \sum_{j=2}^{n} \frac{l_{qj}}{l_{q1}} \tilde{z}_j \quad \text{for } q \in M^+. \tag{3.5}$$

Combining (3.4) with (3.5) results in

$$-\frac{c_p}{l_{p1}} - \sum_{j=2}^{n} \frac{l_{pj}}{l_{p1}} \tilde{z}_j \leq -\frac{c_q}{l_{q1}} - \sum_{j=2}^{n} \frac{l_{qj}}{l_{q1}} \tilde{z}_j \quad \text{for } p \in M^-,\ q \in M^+, \tag{3.6}$$

which is equivalent to

$$\hat{L}_i(\tilde{z}_2, \dots, \tilde{z}_n) \leq 0 \quad \text{for } i = p + (q - m_2 - 1)(m_2 - m_1).$$

This shows b).

b) $\Rightarrow$ a) : Let $\hat{L}_i(\tilde{z}_2, \ldots, \tilde{z}_n) \le 0$ hold for $i = 1, \ldots, \hat{m}$. Then, by definition, $L_i(\tilde{z}_1, \ldots, \tilde{z}_n) \le 0$ for $i = 1, \ldots m_1$ and for any $\tilde{z}_1 \in \mathbf{R}$.
If $M^- = \emptyset$ choose any $\tilde{z}_1$ such that

$$\tilde{z}_1 \le \min_{q \in M^+} \left\{ -\frac{c_q}{l_{q1}} - \sum_{j=2}^{n} \frac{l_{qj}}{l_{q1}} \tilde{z}_j \right\} .$$

Then $L_q(\tilde{z}_1, \ldots, \tilde{z}_n) \le 0$ for each $q \in M^+$.
If $M^+ = \emptyset$ choose any $\tilde{z}_1$ such that

$$\max_{p \in M^-} \left\{ -\frac{c_p}{l_{p1}} - \sum_{j=2}^{n} \frac{l_{pj}}{l_{p1}} \tilde{z}_j \right\} \le \tilde{z}_1 .$$

which is equivalent to $L_p(\tilde{z}_1, \ldots, \tilde{z}_n) \le 0$ for each $p \in M^-$.
If $M^- \ne \emptyset$ and $M^+ \ne \emptyset$ then $\hat{L}_i(\tilde{z}_2, \ldots, \tilde{z}_n) \le 0$ for $i = m_1 + 1, \ldots, \hat{m}$. By the construction of $\hat{L}_i$ this means

$$-\frac{c_p}{l_{p1}} - \sum_{j=2}^{n} \frac{l_{pj}}{l_{p1}} \tilde{z}_j \le -\frac{c_q}{l_{q1}} - \sum_{j=2}^{n} \frac{l_{qj}}{l_{q1}} \tilde{z}_j \quad \text{for each } \; p \in M^-, \; q \in M^+ .$$

This implies

$$c_{\max} := \max_{p \in M^-} \left\{ -\frac{c_p}{l_{p1}} - \sum_{j=2}^{n} \frac{l_{pj}}{l_{p1}} \tilde{z}_j \right\} \le c_{\min} := \min_{q \in M^+} \left\{ -\frac{c_q}{l_{q1}} - \sum_{j=2}^{n} \frac{l_{qj}}{l_{q1}} \tilde{z}_j \right\} .$$

Choose $\tilde{z}_1 \in [c_{\max}, c_{\min}]$; then $L_i(\tilde{z}_1, \ldots \tilde{z}_n) \le 0$ for $i = m_1 + 1, \ldots, m$. This proves b). $\square$

3.2 Characterization of the Solution Set S by Inequalities

Using Lemma 1, we will characterize the solution set S by inequalities. To this end let $[A] \in \mathbf{IR}^{n \times n}$, $[b] \in \mathbf{IR}^n$ and $Ax = b$ with $A \in [A]$ and $b \in [b]$. Without loss of generality, let $x \in O_1$. Then

$$\begin{aligned} \underline{a}_{ij} x_j \le a_{ij} x_j \le \overline{a}_{ij} x_j \; , \qquad & i, j = 1, \ldots n, \\ \underline{b}_i \le \sum_{j=1}^{n} a_{ij} x_j \le \overline{b}_i \; , \qquad & i = 1, \ldots n. \end{aligned} \tag{3.7}$$

With

$$\tilde{z}_{ij} := a_{ij}x_j \ , \quad i,j = 1,\ldots n,$$

the n^2+n double inequalities (3.7) are equivalent to the $2(n^2+n)$ inequalities

$$\left.\begin{array}{l} \left.\begin{array}{lcl} \underline{a}_{ij}x_j - \tilde{z}_{ij} & \leq & 0 \\ -\overline{a}_{ij}x_j + \tilde{z}_{ij} & \leq & 0 \end{array}\right\} \quad j = 1,\ldots,n \\ \begin{array}{lcl} \underline{b}_i - \sum_{j=1}^{n} \tilde{z}_{ij} & \leq & 0 \\ -\overline{b}_i + \sum_{j=1}^{n} \tilde{z}_{ij} & \leq & 0 \end{array} \end{array}\right\} \quad i = 1,\ldots,n\ . \tag{3.8}$$

It is obvious that (3.8) is in the form prescribed by Lemma 1a) (not yet ordered in the way described there) with $\tilde{z}_{ij}$ instead of $\tilde{z}_k$. Therefore, if $\tilde{z}_{11}, \tilde{z}_{12}, \tilde{z}_{13}, \ldots, \tilde{z}_{nn}$ satisfy (3.8) then $\tilde{z}_{12}, \tilde{z}_{13}, \ldots, \tilde{z}_{nn}$ satisfy a system $\hat{L}_i(\tilde{z}_{12}, \tilde{z}_{13}, \ldots, \tilde{z}_{nn}) \leq 0$ with $\hat{L}_i$ from Lemma 1. Note that for fixed i and j the term $\tilde{z}_{ij}$ occurs exactly once in each line of (3.8), i.e., four times for each i. In the notation of Lemma 1, the values of the coefficients l_{ij} are only -1, 0, or 1, and $M^- \neq \emptyset$, $M^+ \neq \emptyset$. Elimination of $\tilde{z}_{11}$ means replacing the inequalities in (3.8)

$$\begin{array}{rcl} \underline{a}_{11}x_1 - \tilde{z}_{11} & \leq & 0 \\ -\overline{a}_{11}x_1 + \tilde{z}_{11} & \leq & 0 \\ \underline{b}_1 - \sum_{j=1}^{n} \tilde{z}_{1j} & \leq & 0 \\ -\overline{b}_1 + \sum_{j=1}^{n} \tilde{z}_{1j} & \leq & 0 \end{array}$$

by

$$\left.\begin{array}{rcl} \underline{a}_{11}x_1 & \leq & \overline{a}_{11}x_1 \\ \underline{a}_{11}x_1 & \leq & \overline{b}_1 - \sum_{j=2}^{n} \tilde{z}_{1j} \\ \underline{b}_1 - \sum_{j=2}^{n} \tilde{z}_{1j} & \leq & \overline{a}_{11}x_1 \\ \underline{b}_1 - \sum_{j=2}^{n} \tilde{z}_{1j} & \leq & \overline{b}_1 - \sum_{j=2}^{n} \tilde{z}_{1j} \Longleftrightarrow \underline{b}_1 \leq \overline{b}_1 \end{array}\right\} \tag{3.9}$$

while keeping the remaining ones. (Cf. the process of getting (3.6) from (3.4) and (3.5) in the proof of (3.3) in Lemma 1.) The first and the fourth inequality of (3.9) always hold; therefore we delete them. Then (3.9) is equivalent to

$$\begin{aligned} \underline{b}_1 - \overline{a}_{11}x_1 - \sum_{j=2}^{n} \tilde{z}_{1j} &\leq 0 , \\ -\overline{b}_1 + \underline{a}_{11}x_1 + \sum_{j=2}^{n} \tilde{z}_{1j} &\leq 0 . \end{aligned}$$

Together with the inequalities of (3.8) which do not contain $\tilde{z}_{11}$ yields a system of inequalities for which the arguments above can be repeated for $\tilde{z}_{12}, \tilde{z}_{13}, \ldots \tilde{z}_{nn}$. After a total of n^2 elimination steps we end at the following $n_{\#} = 2n$ relations.

$$\left.\begin{array}{l} \underline{b}_i \leq \sum_{j=1}^{n} \overline{a}_{ij}x_j \\ \sum_{j=1}^{n} \underline{a}_{ij}x_j \leq \overline{b}_i \end{array}\right\} \quad i = 1, \ldots, n. \tag{3.10}$$

Repeated application of Lemma 1 shows that the statements '$x \in S$' and 'x satisfies (3.10)' are equivalent for $x \in O_1$. Note that (3.10) is just the well-known characterization of S which can be found, e.g., in [2].

We assumed $x \in O_1$. If $x \notin O_1$ then the bounds in the first group of inequalities in (3.7) are changed for those j for which $x_j < 0$. For $x_j < 0$, they are $\overline{a}_{ij}x_j \leq a_{ij}x_j \leq \underline{a}_{ij}x_j$. The elimination process for the $\tilde{z}_{ij}$ proceeds now in an analogous manner. The final result is summarized in the following theorem (cf. [2]). Equivalent formulations can be found in [3], [4], [5], [6] and [7].

Theorem 1. *Let $[A] \in \mathbb{IR}^{n\times n}$ and let $[b] \in \mathbb{IR}^n$. Then $x \in S$ if and only if x satisfies*

$$\begin{aligned} &\sum_{j=1}^{n} \check{a}_{ij}x_j \leq \overline{b}_i \quad \text{with } \check{a}_{ij} := \begin{cases} \overline{a}_{ij} & \text{if } x_j < 0 \\ \underline{a}_{ij} & \text{if } x_j \geq 0 \end{cases} \\ &\underline{b}_i \leq \sum_{j=1}^{n} \hat{a}_{ij}x_j \quad \text{with } \hat{a}_{ij} := \begin{cases} \underline{a}_{ij} & \text{if } x_j < 0 \\ \overline{a}_{ij} & \text{if } x_j \geq 0 \end{cases} \end{aligned} \tag{3.11}$$

for $i = 1, \ldots, n$.

3.3 Characterization of the Symmetric Solution Set S_{sym} by Inequalities

Let the symmetric interval matrix $[A] = [A]^T$ be given. Of course, the solution set S of $[A]$ is characterized by Theorem 1 also in this particular case. We are now going to describe the symmetric solution set S_{sym}.

We want to apply Lemma 1 to show that the symmetric solution set S_{sym} can be characterized as an intersection of closed half–spaces and closed sets of which the boundary is formed by quadrics. This result generalizes that in [2] which was only valid for 2×2 matrices.

Due to the equality of a_{ij} and a_{ji} the numbers $\tilde{z}_{ij} := a_{ij}x_j$ and $\tilde{z}_{ji} := a_{ji}x_i$ are no longer independent if x_i and x_j are given. Therefore, the procedure for characterizing S must be modified in order to get a set of inequalities which describe S_{sym}. Instead of considering $a_{ij}x_j$ we will now use the $\dfrac{n(n+1)}{2}$ numbers

$$\tilde{z}_{ij} := a_{ij}x_ix_j, \quad i,j = 1,\dots n, \quad i \le j.$$

They can be introduced by multiplying (3.7) by x_i taking into account the sign of x_i. For simplicity we restrict ourselves (without loss of generality) to the first orthant. We then get the inequalities

$$\left.\begin{array}{ll} \underline{a}_{ij}x_ix_j \le a_{ij}x_ix_j = \tilde{z}_{ij} \le \overline{a}_{ij}x_ix_j, & i,j = 1,\dots n,\ i \le j \\ \underline{b}_ix_i \le \sum_{j=1}^{n} a_{ij}x_ix_j = \sum_{j=1}^{n} \tilde{z}_{ij} \le \overline{b}_ix_i, & i = 1,\dots n, \\ & \text{with } \tilde{z}_{ij} := \tilde{z}_{ji} \text{ in the case } j < i \end{array}\right\} \tag{3.12}$$

which are equivalent to

$$\left.\begin{array}{l} \left.\begin{array}{lcl} \underline{a}_{ij}x_ix_j - \tilde{z}_{ij} & \le & 0 \\ -\overline{a}_{ij}x_ix_j + \tilde{z}_{ij} & \le & 0 \end{array}\right\} \ j = i,\dots,n \\ \begin{array}{lcl} \underline{b}_ix_i - \sum_{j=1}^{n} \tilde{z}_{ij} & \le & 0 \\ -\overline{b}_ix_i + \sum_{j=1}^{n} \tilde{z}_{ij} & \le & 0 \end{array} \end{array}\right\} \quad i = 1,\dots,n. \tag{3.13}$$

with $\tilde{z}_{ij} = \tilde{z}_{ji}$. Lemma 1 can now be applied as it was in the proof of Theorem 1. Thus, if $\tilde{z}_{11}, \tilde{z}_{22}, \dots, \tilde{z}_{nn}$ are eliminated first, the number of non–trivial inequalities in (3.13) reduces from $2\left(\frac{n(n+1)}{2}+n\right) = n(n+1)+2n$ to $n(n+1)$. The inequalities now read

$$\left.\begin{array}{l} \left.\begin{array}{lcl} \underline{a}_{ij}x_ix_j - \tilde{z}_{ij} & \leq & 0 \\ -\overline{a}_{ij}x_ix_j + \tilde{z}_{ij} & \leq & 0 \end{array}\right\} \quad j = i+1,\ldots,n \\ \underline{b}_i x_i - \overline{a}_{ii}x_i^2 - \sum_{\substack{j=1\\ j\neq i}}^{n} \tilde{z}_{ij} \quad \leq \quad 0 \\ -\overline{b}_i x_i + \underline{a}_{ii}x_i^2 + \sum_{\substack{j=1\\ j\neq i}}^{n} \tilde{z}_{ij} \quad \leq \quad 0 \end{array}\right\} \quad i = 1,\ldots,n. \tag{3.14}$$

In order to show what happens when eliminating $\tilde{z}_{ij} = \tilde{z}_{ji}$ for $i \neq j$ we will choose $\tilde{z}_{12}$ as the next candidate. Here, we have to replace the six inequalities

$$\left.\begin{array}{rcl} \underline{a}_{12}x_1x_2 - \tilde{z}_{12} & \leq & 0 \\ -\overline{a}_{12}x_1x_2 + \tilde{z}_{12} & \leq & 0 \\ \underline{b}_1 x_1 - \overline{a}_{11}x_1^2 - \sum_{j=2}^{n} \tilde{z}_{1j} & \leq & 0 \\ -\overline{b}_1 x_1 + \underline{a}_{11}x_1^2 + \sum_{j=2}^{n} \tilde{z}_{1j} & \leq & 0 \\ \underline{b}_2 x_2 - \overline{a}_{22}x_2^2 - \sum_{\substack{j=1\\ j\neq 2}}^{n} \tilde{z}_{2j} & \leq & 0 \\ -\overline{b}_2 x_2 + \underline{a}_{22}x_2^2 + \sum_{\substack{j=1\\ j\neq 2}}^{n} \tilde{z}_{2j} & \leq & 0 \end{array}\right\} \tag{3.15}$$

by the six non–trivial inequalities

$$\begin{array}{rcl} \underline{a}_{12}x_1x_2 & \leq & \overline{b}_1 x_1 - \underline{a}_{11}x_1^2 - \sum_{j=3}^{n} \tilde{z}_{1j} \\ \underline{a}_{12}x_1x_2 & \leq & \overline{b}_2 x_2 - \underline{a}_{22}x_2^2 - \sum_{j=3}^{n} \tilde{z}_{2j} \\ \underline{b}_1 x_1 - \overline{a}_{11}x_1^2 - \sum_{j=3}^{n} \tilde{z}_{1j} & \leq & \overline{a}_{12}x_1x_2 \\ \underline{b}_2 x_2 - \overline{a}_{22}x_2^2 - \sum_{j=3}^{n} \tilde{z}_{2j} & \leq & \overline{a}_{12}x_1x_2 \end{array}$$

$$
\begin{aligned}
\underline{b}_1 x_1 - \overline{a}_{11} x_1^2 - \sum_{j=3}^{n} \tilde{z}_{1j} &\leq \overline{b}_2 x_2 - \underline{a}_{22} x_2^2 - \sum_{j=3}^{n} \tilde{z}_{2j} \\
\underline{b}_2 x_2 - \overline{a}_{22} x_2^2 - \sum_{j=3}^{n} \tilde{z}_{2j} &\leq \overline{b}_1 x_1 - \underline{a}_{11} x_1^2 - \sum_{j=3}^{n} \tilde{z}_{1j}
\end{aligned}
$$

which are equivalent to

$$
\left.
\begin{aligned}
\underline{b}_1 x_1 - \overline{a}_{11} x_1^2 - \overline{a}_{12} x_1 x_2 - \sum_{j=3}^{n} \tilde{z}_{1j} &\leq 0 \\
-\overline{b}_1 x_1 + \underline{a}_{11} x_1^2 + \underline{a}_{12} x_1 x_2 + \sum_{j=3}^{n} \tilde{z}_{1j} &\leq 0 \\
\underline{b}_2 x_2 - \overline{a}_{22} x_2^2 - \overline{a}_{12} x_1 x_2 - \sum_{j=3}^{n} \tilde{z}_{2j} &\leq 0 \\
-\overline{b}_2 x_2 + \underline{a}_{22} x_2^2 + \underline{a}_{12} x_1 x_2 + \sum_{j=3}^{n} \tilde{z}_{2j} &\leq 0 \\
\underline{b}_1 x_1 - \overline{b}_2 x_2 - \overline{a}_{11} x_1^2 + \underline{a}_{22} x_2^2 - \sum_{j=3}^{n} (\tilde{z}_{1j} - \tilde{z}_{2j}) &\leq 0 \\
-\overline{b}_1 x_1 + \underline{b}_2 x_2 + \underline{a}_{11} x_1^2 - \overline{a}_{22} x_2^2 + \sum_{j=3}^{n} (\tilde{z}_{1j} - \tilde{z}_{2j}) &\leq 0\,.
\end{aligned}
\right\}
\qquad (3.16)
$$

The remaining inequalities of (3.14) are unchanged. Continuing the reduction process of $\tilde{z}_{ij}$ along the lines of Lemma 1 ends up with a set of inequalities each of which describes a quadric, if '≤ 0' is replaced by '$= 0$'. Among this set one finds the inequalities

$$
\left.
\begin{aligned}
\underline{b}_i x_i - \sum_{j=1}^{n} \overline{a}_{ij} x_i x_j &\leq 0 \\
-\overline{b}_i x_i + \sum_{j=1}^{n} \underline{a}_{ij} x_i x_j &\leq 0
\end{aligned}
\right\}
\quad i = 1, \ldots, n\,. \qquad (3.17)
$$

(For a proof use the first four inequalities of (3.16) and the first two sets of inequalities from (3.13).) For $x_i > 0$ we can divide (3.17) by x_i to end up with (3.10). This means that the set of inequalities characterizing S must necessarily be satisfied for $x \in S_{\text{sym}}$ – a trivial statement. Nevertheless curvilinear boundaries normally will occur for S_{sym}. They arise from the last two inequalities of (3.16). If $x_i = 0$, the division by x_i is forbidden in (3.17). Since $x = (x_i) \in S_{\text{sym}}$ implies $x \in S$, the inequalities (3.10) are satisfied also in this particular case $x_i = 0$.

As with the non-symmetric solution set S, repeated application of Lemma 1 shows that the statements '$x \in S_{\text{sym}}$' and 'x satisfies the set of inequalities constructed by the above-mentioned procedure' are equivalent for $x \in O_1$. Here again, dispensing with the restriction $x \in O_1$ means changing the bounds in (3.12) which involve x_i, x_j. Thus we obtain

$$\overline{a}_{ij}x_ix_j \le \tilde{z}_{ij} \le \underline{a}_{ij}x_ix_j \quad \text{if } x_ix_j < 0,$$

and

$$\overline{b}_ix_i \le \sum_{j=1}^{n} \tilde{z}_{ij} \le \underline{b}_ix_i \quad \text{if } x_i < 0.$$

Due to our previous considerations we get the following theorem.

Theorem 2. *Let $[A] = [A]^T \in \mathbf{IR}^{n\times n}$ and let $[b] \in \mathbf{IR}^n$. Then in each orthant the symmetric solution set S_{sym} can be represented as the intersection of the unsymmetric solution set S and sets with quadrics as boundaries.*

3.4 On the Number of Inequalities to Describe S_{sym} for Tridiagonal Matrices

We want to derive an upper bound $\overline{n}$ for the number $n_\#$ of the inequalities which are necessary to describe S_{sym} for tridiagonal matrices $[A] = [A]^T \in \mathbf{IR}^{n\times n}$. In this case, we use the elements $\tilde{z}_{ij}$ only for $0 \le j - i \le 1$, and we start with the inequalities

$$\left.\begin{array}{l}
\left.\begin{array}{rcl}
\underline{a}_{i,i+1}x_ix_{i+1} - \tilde{z}_{i,i+1} & \le & 0 \\
-\overline{a}_{i,i+1}x_ix_{i+1} + \tilde{z}_{i,i+1} & \le & 0
\end{array}\right\} \quad i = 1, \dots, n-1, \\
\begin{array}{rcl}
\underline{b}_1x_1 - \overline{a}_{11}x_1^2 - \tilde{z}_{12} & \le & 0 \\
-\overline{b}_1x_1 + \underline{a}_{11}x_1^2 + \tilde{z}_{12} & \le & 0
\end{array} \\
\left.\begin{array}{rcl}
\underline{b}_ix_i - \overline{a}_{ii}x_i^2 - \tilde{z}_{i-1,i} - \tilde{z}_{i,i+1} & \le & 0 \\
-\overline{b}_ix_i + \underline{a}_{ii}x_i^2 + \tilde{z}_{i-1,i} + \tilde{z}_{i,i+1} & \le & 0
\end{array}\right\} \quad i = 2, \dots, n-1, \\
\begin{array}{rcl}
\underline{b}_nx_n - \overline{a}_{nn}x_n^2 - \tilde{z}_{n-1,n} & \le & 0 \\
-\overline{b}_nx_n + \underline{a}_{nn}x_n^2 + \tilde{z}_{n-1,n} & \le & 0
\end{array}
\end{array}\right\} \tag{3.18}$$

for $S_{\text{sym}} \cap O_1$ which we get after having eliminated the diagonal elements $\tilde{z}_{ii}$, $i = 1, \dots, n$ (cf. (3.14)).

Consider a fixed element $\tilde{z}_{k,k+1}$. The elimination process shows that in each step of this process half of the inequalities containing $\tilde{z}_{k,k+1}$ has the form

$$r + \tilde{z}_{k,k+1} \le 0 \tag{3.19}$$

while the other half has the form

$$\hat{r} - \tilde{z}_{k,k+1} \le 0 \tag{3.20}$$

with r, $\hat{r}$ independent of $\tilde{z}_{k,k+1}$. Each inequality of the form (3.19) has a counterpart (3.20) which one gets by changing plus to minus and the upper bars to lower ones, and vice versa.

We will eliminate $\tilde{z}_{k,k+1}$ in the order $k = 1, 2, \ldots n-1$. In the k–th of these elimination steps we will find the following situation.

Lemma 2. *Let $[A] = [A]^T \in \mathbb{IR}^{n\times n}$ be a tridiagonal interval matrix and let $[b] \in \mathbb{IR}^n$. Assume that the off-diagonal element $\tilde{z}_{k,k+1}$ is eliminated in the k–th elimination step where the n preceding elimination steps for the diagonal elements $\tilde{z}_{ii}$, $i = 1, \ldots n$, have already been performed (and therefore are not counted here).*

a) Before the k–th elimination step $\tilde{z}_{k,k+1}$ occurs in $4 + 2k$ inequalities of which $2 + 2k = 2(k+1)$ contain $\tilde{z}_{k,k+1}$ as the only element $\tilde{z}_{i,i+1}$ while the remaining two inequalities contain $\tilde{z}_{k,k+1}$ together with $\tilde{z}_{k+1,k+2}$. (In the case $k = n-1$ we define $\tilde{z}_{n,n+1} := 0$.)

b) Before the k–th elimination step $\tilde{z}_{i,i+1}$, $i = k+1, \ldots, n-1$, occurs in six inequalities. With the exception of the two inequalities which contain $\tilde{z}_{k,k+1}$ together with $\tilde{z}_{k+1,k+2}$, these inequalities are the original ones (which have not been changed after having eliminated the elements $\tilde{z}_{ii}$, $i = 1, \ldots, n$).

c) Right after the k–th elimination step, $k < n-1$, there are $k^2 + k$ final inequalities, i.e., inequalities which are no more changed during the succeeding elimination process. If $k = n-1$ this number increases by $2n$.

Proof. The assertion is trivial for $k = 1$. Let it hold for some fixed $k < n-1$. Those inequalities which contain $\tilde{z}_{k,k+1}$ as the only element $\tilde{z}_{i,i+1}$ combine to final inequalities. Combining an inequality with its counterpart results in an inequality which always holds. (Keeping track of the history of an inequality $L_s \le 0$, one can see that L_s can be expressed as sum of some of the terms $+\underline{b}_i|x_i|$, $-\overline{b}_i|x_i|$, $+\underline{a}_{ij}|x_i|\cdot|x_j|$, $-\overline{a}_{ij}|x_i|\cdot|x_j|$ and those $\tilde{z}_{ij}$ which have not yet been eliminated. Therefore L_s and its counterpart combine to a final inequality, say $\hat{L}_t \le 0$, in which $\hat{L}_t$ is the sum of some of the terms $-d([b]_i)|x_i|$, $-d([a]_{ij})|x_i|\cdot|x_j|$, where $d([a]) := \overline{a} - \underline{a}$ denotes the diameter of an interval $[a]$. Hence $\hat{L}_t \le 0$ is trivially true.) We will omit such trivial inequalities in our further considerations. In the k–th elimination step we thus get

$(k+1)^2-(k+1)=k^2+k$ final inequalities and $2(k+1)$ new ones. These new inequalities contain $\tilde{z}_{k+1,k+2}$ as the only element $\tilde{z}_{i,i+1}$. All other inequalities remain unchanged. Among them there are two inequalities which contain only $\tilde{z}_{k+1,k+2}$ and two which contain $\tilde{z}_{k+1,k+2}$ together with $\tilde{z}_{k+2,k+3}$. The increase of $2n$ in the case $k=n-1$ results from the $2n$ new inequalities for $\tilde{z}_{n,n+1}$ which are final ones since we defined $\tilde{z}_{n,n+1}$ to be zero. This proves the lemma. □

Adding all the final inequalities results in

$$\begin{aligned}\overline{n} &= \sum_{k=1}^{n-1} k^2 + \sum_{k=1}^{n-1} k + 2n \\ &= \frac{(n-1)n(2n-1)}{6} + \frac{n(n-1)}{2} + 2n = \frac{n}{3}(n^2+5).\end{aligned}$$

The elimination process for tridiagonal matrices shows, that $n_\# = \overline{n}$ may be possible, i.e., the above bound $\overline{n}$ is *sharp*. There are also cases in which $n_\# < \overline{n}$ holds, as can be seen from any diagonal matrix.

We state our result in the following theorem.

Theorem 3. *Let $[A]=[A]^T \in \mathbf{IR}^{n\times n}$ be a tridiagonal interval matrix, let $[b] \in \mathbf{IR}^n$ and let $n_\#$ be the number of inequalities obtained by the elimination process described above for the intersection of S_{sym} with an orthant. Then $n_\#$ is bounded by $\overline{n} = \frac{n}{3}(n^2+5)$. This bound is sharp.*

3.5 Particular Cases

In the particular case of 2×2 matrices we end up with the following 6 inequalities for $S_{\text{sym}} \cap O_1$ which can be deduced from (3.18).

$$\begin{aligned} \underline{b}_1 x_1 - \overline{a}_{11} x_1^2 - \overline{a}_{12} x_1 x_2 &\leq 0 \\ -\overline{b}_1 x_1 + \underline{a}_{11} x_1^2 + \underline{a}_{12} x_1 x_2 &\leq 0 \\ \underline{b}_2 x_2 - \overline{a}_{22} x_2^2 - \overline{a}_{12} x_1 x_2 &\leq 0 \\ -\overline{b}_2 x_2 + \underline{a}_{22} x_2^2 + \underline{a}_{12} x_1 x_2 &\leq 0 \\ \underline{b}_1 x_1 - \overline{b}_2 x_2 - \overline{a}_{11} x_1^2 + \underline{a}_{22} x_2^2 &\leq 0 \\ -\overline{b}_1 x_1 + \underline{b}_2 x_2 + \underline{a}_{11} x_1^2 - \overline{a}_{22} x_2^2 &\leq 0\,. \end{aligned}$$

Dividing the first two inequalities by x_1, the next two by x_2 yields the inequalities for the unsymmetric solution set $S \cap O_1$. The last two inequalities just characterize the sets C^+ and C^- in [2].

For tridiagonal 3×3 matrices $S_{\text{sym}} \cap O_1$ is characterized by the 7 inequalities

$$\left.\begin{array}{rcl}
\underline{b}_1 x_1 - \overline{a}_{11} x_1^2 - \overline{a}_{12} x_1 x_2 & \le & 0 \\
\underline{b}_2 x_2 - \overline{a}_{12} x_1 x_2 - \overline{a}_{22} x_2^2 - \overline{a}_{23} x_2 x_3 & \le & 0 \\
\underline{b}_3 x_3 - \overline{a}_{23} x_2 x_3 - \overline{a}_{33} x_3^2 & \le & 0 \\
\underline{b}_1 x_1 - \overline{b}_2 x_2 - \overline{a}_{11} x_1^2 + \underline{a}_{22} x_2^2 + \underline{a}_{23} x_2 x_3 & \le & 0 \\
\underline{b}_1 x_1 - \overline{b}_2 x_2 + \underline{b}_3 x_3 - \overline{a}_{11} x_1^2 + \underline{a}_{22} x_2^2 - \overline{a}_{33} x_3^2 & \le & 0 \\
\underline{b}_1 x_1 - (\overline{b}_2 - \underline{b}_2) x_2 - \overline{a}_{11} x_1^2 - \overline{a}_{12} x_1 x_2 - (\overline{a}_{22} - \underline{a}_{22}) x_2^2 & \le & 0 \\
\underline{b}_2 x_2 - \overline{b}_3 x_3 - \overline{a}_{12} x_1 x_2 - \overline{a}_{22} x_2^2 + \underline{a}_{33} x_3^2 & \le & 0
\end{array}\right\} \quad (3.21)$$

and their 7 counterparts, which can easily be deduced from (3.21). The first three inequalities and their counterparts describe the unsymmetric solution set $S \cap O_1$. The sixth inequality equals the first one if $[b_2]$ and $[a_{22}]$ are point intervals.

Without proof we mention that the number of inequalities for $S_{\text{sym}} \cap O$ increases to 44 for a dense 3×3 system.

4 EXAMPLES

In this section we want to illustrate some aspects of the results in Section 3.

Example 1. Let

$$[A] = \begin{pmatrix} 1 & [-1,1] \\ [-1,1] & -1 \end{pmatrix}, \qquad [b] = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

The matrix $[A]$ contains the singular matrix $A_1 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, for which no solution of the system $A_1 x = b$ exists, and the singular matrix $A_2 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$, for which the solutions of $A_2 x = b$ are given by $x_1 - x_2 = 2$. However, $[A]$ contains no singular *symmetric* matrix, since $\det \begin{pmatrix} 1 & s \\ s & -1 \end{pmatrix} = -1 - s^2 \le -1$. The solution sets S and S_{sym} can be represented by

$$S_{\text{sym}} = \left\{ \frac{2}{1+s^2} \begin{pmatrix} 1+s \\ -1+s \end{pmatrix} \,\middle|\, -1 \le s \le 1 \right\} \subseteq O_4$$

and

$$S = \left\{ \frac{2}{1+st} \begin{pmatrix} 1+s \\ -1+t \end{pmatrix} \middle| -1 \le s, t \le 1,\ st \ne -1 \right\} \cup \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_1 - x_2 = 2 \right\},$$

where O_4 denotes the fourth quadrant of $\mathbf{R}^2$. The inequalities (3.11) characterizing $S \cap O_4$ read

$$\begin{aligned} 2 - x_1 + x_2 &\le 0 \\ -2 + x_1 + x_2 &\le 0 \\ -2 - x_1 - x_2 &\le 0 \end{aligned}$$

where we have dropped the fourth inequality since it coincides with the first one.

For $S_{\text{sym}} = S_{\text{sym}} \cap O_4$ we need, in addition, the inequalities

$$\begin{aligned} -(x_1^2 + x_2^2 - 2x_1 + 2x_2) &\le 0 \\ x_1^2 + x_2^2 - 2x_1 + 2x_2 &\le 0 \end{aligned}$$

which we combine to the equality

$$x_1^2 + x_2^2 - 2x_1 + 2x_2 = 0$$

or, equivalently,

$$(x_1 - 1)^2 + (x_2 + 1)^2 = 2. \tag{3.22}$$

Thus, S_{sym} is the half-circle which results as the intersection of the circle (3.22) with O_4, while S is the union of the half-strip in Figure 1 and the straight line $x_1 - x_2 = 2$.

Example 2. Let

$$[A] = \begin{pmatrix} -2 & 2 & 0 \\ 2 & 2 & [3,8] \\ 0 & [3,8] & -2 \end{pmatrix},\ A_t = \begin{pmatrix} -2 & 2 & 0 \\ 2 & 2 & t \\ 0 & t & -2 \end{pmatrix}, 3 \le t \le 8,$$

$$\text{and } b = [b] = \begin{pmatrix} 0 \\ 16 \\ 12 \end{pmatrix}.$$

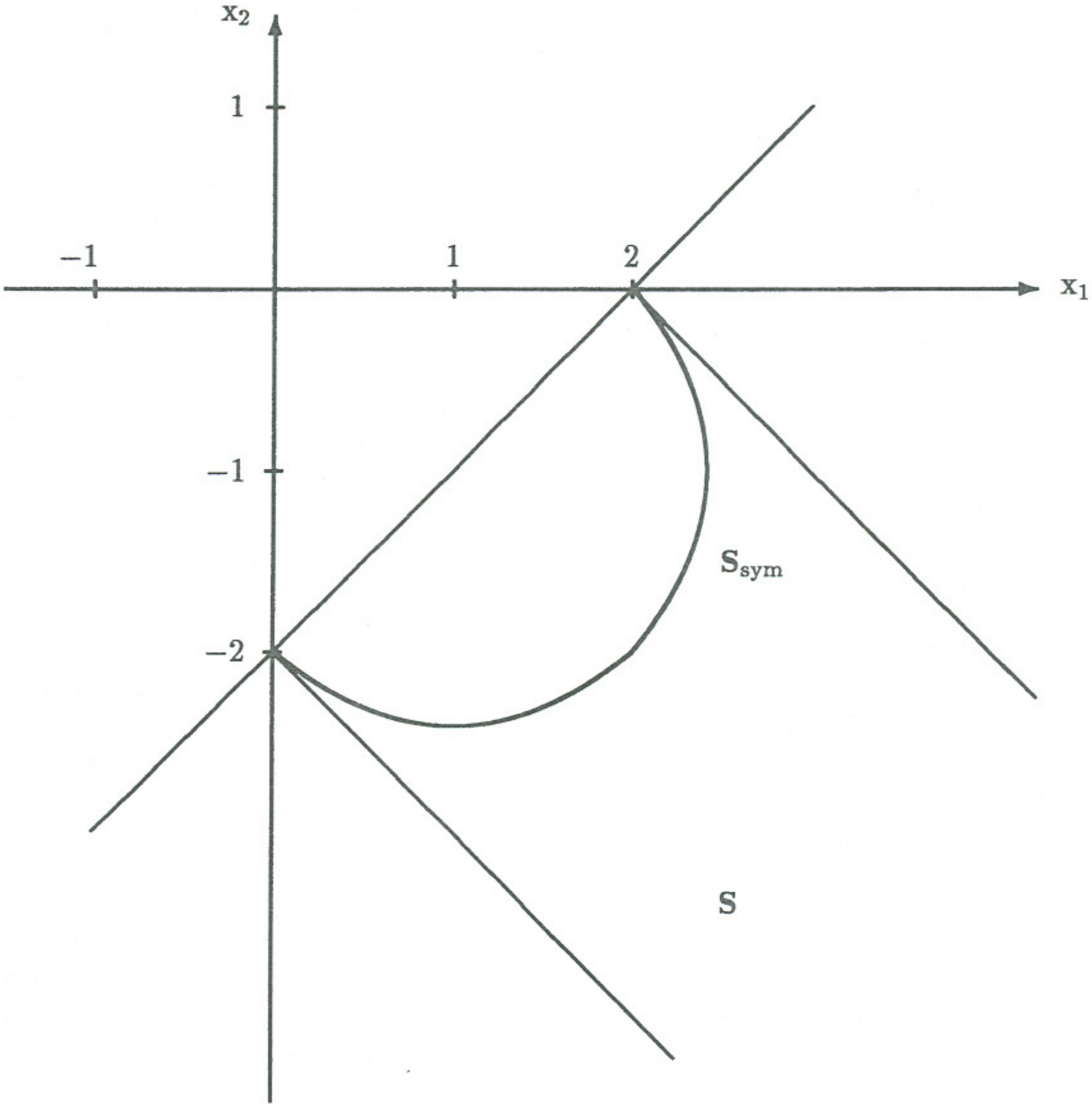


Fig. 1 The shape of the solution sets S, $S_{\rm sym}$ in Example 1.

From $A_t x = b$ we get

$$x_1 = x_2 = \frac{12t + 32}{t^2 + 8} \geq 0, \quad x_3 = \frac{16t - 48}{t^2 + 8} \geq 0 \ . \tag{3.23}$$

In particular, A_t is regular and $S_{\text{sym}} = S_{\text{sym}} \cap O_1$. The inequalities in (3.21) yield

$$\begin{aligned} x_1 - x_2 &= 0, \\ 3x_3 \le 16 - 2x_1 - 2x_2 &\le 8x_3, \\ 3x_2 \le 12 + 2x_3 &\le 8x_2, \end{aligned} \tag{3.24}$$

which characterize $S \cap O_1$, and

$$3x_2x_3 \le 16x_2 - 2x_1^2 - 2x_2^2 \le 8x_2x_3,$$

$$x_1^2 + (x_2 - 4)^2 + (x_3 + 3)^2 = 25, \tag{3.25}$$

$$-\left(\frac{x_1 - 8}{2}\right)^2 + \left(\frac{x_1 - 8}{2} + x_2\right)^2 + (x_3 + 3)^2 = 9, \tag{3.26}$$

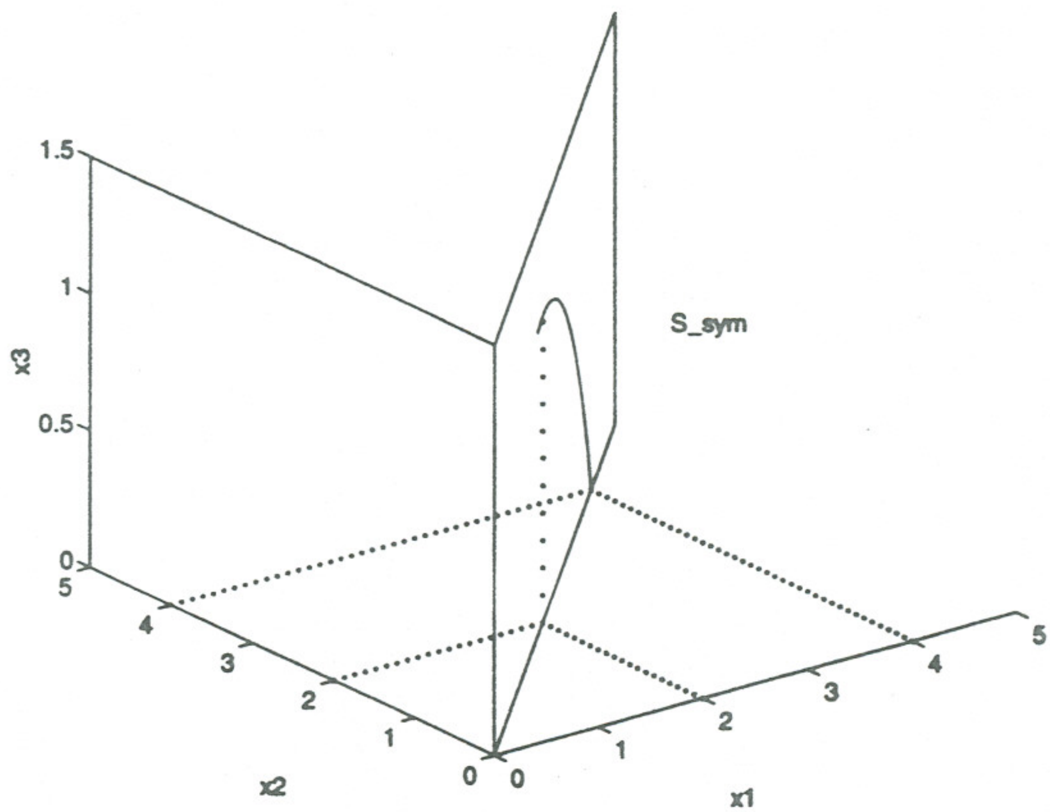
which select S_{sym} from $S \cap O_1$. Note that the sixth inequality of (3.21) coincides here with the first one (after having divided by x_1); therefore it has been omitted. The intersection of the sphere (3.25) with center $(0, 4, -3)$ and radius 5 with the plane (3.24) is the circle C on this plane with center $(2, 2, -3)$ and radius $\sqrt{17}$. Since S_{sym} is connected (see [2] or prove it directly from (3.23)) it must be an arc of C. A closer look at (3.23) reveals that S_{sym} is the arc of C which lies in O_1 between $(4, 4, 0)$ and $(\frac{16}{9}, \frac{16}{9}, \frac{10}{9})$. One obtains these endpoints as the solutions of $A_t x = b$ for the values $t = 3$ and $t = 8$, respectively. The equation (3.26) describes a distorted one–sheeted hyperboloid. S_{sym} is illustrated in Figure 2.

The unsymmetric solution set S is the triangle with the corners $(4, 4, 0)$, $(\frac{17}{8}, \frac{17}{8}, \frac{5}{2})$, $(\frac{16}{9}, \frac{16}{9}, \frac{10}{9})$ which lies in the plane (3.24) and in O_1, and which contains S_{sym}.

Acknowledgements

The work of the second author was partially supported by NSF Grants No. CDA-9015006 and EEC-9322370, and by NASA Grant No. NAG 9-757.

Fig. 2 The shape of the solution set S_{sym} in Example 2.



REFERENCES

[1] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, N.Y., 1983.

[2] G. Alefeld and G. Mayer, "On the Symmetric and Unsymmetric Solution Set of Interval Systems", *SIAM J. Matrix Anal. Appl.*, 1995 (to appear).

[3] H. Beeck, "Über Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten", *Computing*, 1972, Vol. 10, pp. 231 – 244.

[4] D. J. Hartfiel, "Concerning the Solution Set of $Ax = b$ where $P \leq A \leq Q$ and $p \leq b \leq q$", *Numer. Math.*, 1980, Vol. 35, p. 355–359.

[5] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.

[6] W. Oettli and W. Prager, "Compatibility of Approximate Solution of Linear Equations with Given Error Bounds for Coefficients and Right–hand Sides", *Numer. Math.*, 1964, Vol. 6, pp. 405–409.

[7] J. Rohn, "Interval Linear Systems", *Freiburger Intervall–Berichte*, 1984, Vol. 84/7, pp. 33–58.

[8] A. Seidenberg, "A New Decision Method for Elementary Algebra", *Annals of Math.*, 1954, Vol. 60, pp. 365–374.

[9] A. Tarski, *A Decision Method for Elementary Algebra and Geometry*, 2nd ed., Berkeley and Los Angeles, 1951.