Computing 79, 53–60 (2007)
DOI 10.1007/s00607-006-0217-8
Printed in The Netherlands

Computing

Comparison of the Hansen-Sengupta and the Frommer-Lang-Schnurr existence tests

A. Goldsztejn, Irvine, CA

Received March 7, 2006; revised December 1, 2006
Published online: January 16, 2007

Abstract

The Krawczyk and the Hansen-Sengupta interval operators are closely related to the interval Newton operator. These interval operators can be used as existence tests to prove existence of solutions for systems of equations. It is well known that the Krawczyk operator existence test is less powerful that the Hansen-Sengupta operator existence test, the latter being less powerful than the interval Newton operator existence test. In 2004, Frommer et al. proposed an existence test based on the Poincaré-Miranda theorem and proved that it is more powerful than the Krawczyk existence test. In this paper, we complete the classification of these four existence tests showing that, in practice, the Hansen-Sengupta existence test is actually more powerful than the existence test proposed by Frommer et al.

AMS Subject Classifications: 65H10, 65G20, 65G40.

Keywords: Nonlinear systems of equations, existence test, interval analysis.

1. Introduction

Interval enclosures of real functions have an intrinsic refutation power: they can be used naturally to prove that a system of equations has no solution inside a given region. Another use of interval enclosures is to rigorously check the hypothesis of some existence theorems like the Brouwer or Poincaré-Miranda[1] theorems. In that way, interval analysis can be used to rigorously prove the existence of solutions to systems of equations (see [5], [4] for introductions to interval analysis). Such verifiable conditions for the existence of solutions of systems of equations are called *existence tests*. The comparison of existence tests is valuable for several reasons: it often provides a better understanding of the existence tests and, when a test is proved to be more powerful than an other one, it allows one to make the right choice. Several comparisons between some existence tests have already been conducted (e.g., [2], [1], [3], [11]).

[1] Poincaré-Miranda theorem is sometimes called Miranda theorem. Historically, Poincaré proposed this theorem in 1883 in [10] while Miranda, being unaware of Poincaré's work, provided a proof of this theorem based on the Brouwer fixed point theorem in 1940 in [6]. After Miranda's work the theorem was called Miranda's theorem until Poincaré's work was re-discovered in the 70's.

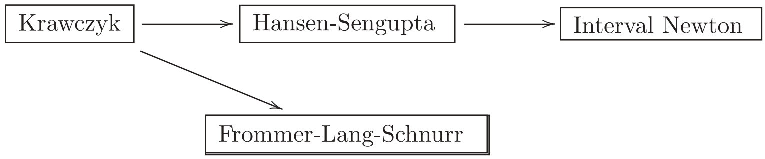


Fig. 1. Previously known hierarchy: arrows points to more powerful existence tests

We focus here on the four following existence tests: the *Krawczyk existence test*[2], the *Hansen-Sengupta existence test*, the *Newton existence test* and the *Frommer-Lang-Schnurr existence test*. The first three come from the interpretation of the interval operators that share the same names (cf. [8]). The last has been proposed in [3]. The description of these tests is given in Sect. 2. The previously known hierarchy for the proving power of these existence tests is summarized in Fig. 1. The comparisons between the Krawczyk, Hansen-Sengupta and interval Newton existence tests can be found in [8]. While, the comparison between the Krawczyk and Frommer-Lang-Schnurr existence test has been conducted in [3].

In the present paper, the relationship between the Frommer-Lang-Schnurr and the Hansen-Sengupta existence tests is investigated. The comparison is actually conducted using a slightly weaker version of the Frommer-Lang-Schnurr existence test (inequalities are replaced by strict inequalities in the test). Our main result is that the Frommer-Lang-Schnurr existence test with strict inequalities is less powerful than the Hansen-Sengupta existence test. We will argue that this alteration of the Frommer-Lang-Schnurr existence test has no influence in typical practical situations. Therefore, the Hansen-Sengupta existence test will actually be proved to be more powerful in practice than the Frommer-Lang-Schnurr existence test.

Notations: Interval objects are denoted using brackets. Throughout the paper, $[\mathbf{x}] = ([x_1^-, x_1^+], \ldots, [x_n^-, x_n^+])^T$ is an interval vector and $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)^T$ is an element of $[\mathbf{x}]$ and $\mathbf{f} : [\mathbf{x}] \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a continuous vector-valued function. The interval matrix $[S]$ is an *interval slope matrix* for f, $[\mathbf{x}]$ and $\tilde{\mathbf{x}}$, i.e., it satisfies

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\tilde{\mathbf{x}}) \in [S] \cdot (\mathbf{x} - \tilde{\mathbf{x}}) \text{ for all } \mathbf{x} \in [\mathbf{x}]. \tag{1}$$

Such an interval matrix can be computed, e.g., using the interval evaluation of the derivatives of $\mathbf{f}$ (cf. [8]). The matrix $C \in \mathbb{R}^{n \times n}$ is supposed to be nonsingular and will be used as a preconditioning matrix. The order between reals is generalized to intervals and $[a] \leq [b]$ holds if and only if $\sup[a] \leq \inf[b]$ while $[a] < [b]$ holds if and only if $\sup[a] < \inf[b]$. Finally, given a set $E \subseteq R^n$, the *interval hull* of E is the smallest interval vector that contains E and is denoted by $\square E$.

[2] The Krawczyk existence test is called Moore test in [3] because Moore has provided a proof of the existence test obtained using the Krawczyk operator.

2. Statements of the existence tests

2.1. *The Krawczyk and Hansen-Sengupta operators*

The presentation given here follows the one given by Neumaier in [8]. The interval Gauss-Seidel operator is defined as follows: first in dimension one, $\Gamma\,([a],[b],[x]) := \square\,([x] \cap \{x|(\exists a \in [a])(\exists b \in [b])(ax = b)\})$. In the case where $0 \notin [a]$, one obtains the expression $\Gamma\,([a],[b],[x]) = ([b]/[a]) \cap [x]$ (cf. [8] for the expression in the case $0 \in [a]$). Then, in arbitrary dimension, $\Gamma\,([A],[\mathbf{b}],[\mathbf{x}]) := [\mathbf{y}]$ with

$$[y_i] := \Gamma\left([a_{ii}]\,,\ [b_i] - \sum_{j<i}[a_{ij}][y_j] - \sum_{j>i}[a_{ij}][x_j]\,,\ [x_i]\right). \tag{2}$$

The interpretation of the interval Gauss-Seidel operator is

$$\Big\{\mathbf{x} \in [\mathbf{x}] \mid (\exists A \in [A])\,(\exists \mathbf{b} \in [\mathbf{b}])\,(A \cdot \mathbf{x} = \mathbf{b})\Big\} \subseteq \Gamma\,([A],[\mathbf{b}],[\mathbf{x}])\,. \tag{3}$$

Finally, given a nonsingular matrix $C \in \mathbb{R}^{n\times n}$, the Krawczyk and Hansen-Sengupta operators are defined as follows:

$$[K]([\mathbf{x}],\tilde{\mathbf{x}}) := \tilde{\mathbf{x}} - C \cdot \mathbf{f}(\tilde{\mathbf{x}}) - (C \cdot [S] - I) \cdot \big([\mathbf{x}] - \tilde{\mathbf{x}}\big) \tag{4}$$

$$[H]([\mathbf{x}],\tilde{\mathbf{x}}) := \tilde{\mathbf{x}} + \Gamma\,\big(C \cdot [S]\,,\ -C \cdot \mathbf{f}(\tilde{\mathbf{x}})\,,\ [\mathbf{x}] - \tilde{\mathbf{x}}\big)\,. \tag{5}$$

The following theorem shows the usefulness of these operators for the study of the solutions of systems of equations.

Theorem 1 (Corollary 5.4.3 in [8]): *If* $[\mathbf{x}']$ *denotes either* $[K]([\mathbf{x}],\tilde{\mathbf{x}})$ *or* $[H]([\mathbf{x}],\tilde{\mathbf{x}})$ *then*:

(i) Every zero $\mathbf{x}^* \in [\mathbf{x}]$ *of* $\mathbf{f}$ *satisfies* $\mathbf{x}^* \in [\mathbf{x}']$.
(ii) If $[\mathbf{x}'] \cap [\mathbf{x}] = \emptyset$ *then* $\mathbf{f}$ *has no zero in* $[\mathbf{x}]$.
(iii) If $\emptyset \neq [\mathbf{x}'] \subseteq \mathrm{int}[x]$ *then* $\mathbf{f}$ *has at least one zero in* $[\mathbf{x}']$.

The existence tests obtained checking (iii) in Theorem 1 will be called the Krawczyk and Hansen-Sengupta existence tests, respectively. The next corollary of Theorem 1 is obtained changing the definition (2) to

$$[y_i] := \Gamma\left([a_{ii}]\,,\ [b_i] - \sum_{j\neq i}[a_{ij}][x_j]\,,\ [x_i]\right), \tag{6}$$

hence computing a superset of (2).

Corollary 1: *If for all* $i \in \{1,\dots,n\}$

$$0 \notin (C \cdot [S])_{ii} \ \ \text{and}\ \ [h_i] \subseteq \mathrm{int}[x_i], \tag{7}$$

where $[h_i] := \tilde{x}_i - (1/(C\cdot[S])_{ii})\Big(\big(C \cdot \mathbf{f}(\tilde{\mathbf{x}})\big)_i + \sum_{j\neq i}(C\cdot[S])_{ij} \cdot \big([x_j] - \tilde{x}_j\big)\Big)$, *then* $\mathbf{f}$ *has a zero in* $[\mathbf{x}]$.

In dimension superior than two, (6) can be a strict superset of (2). Therefore, Corollary 1 is strictly weaker than the original Hansen-Sengupta existence test, and it will be called the *weak Hansen-Sengupta existence test*.

Remark 1: *It can be noted that* $[h_i]$ *in Corollary 1 remains unchanged when one replaces* C *by* C' *with* $C'_{ij} = \delta_i C_{ij}$ *where* $\delta_i \in \{-1, 1\}$. *Therefore, the success of the weak Hansen-Sengupta existence test is invariant with respect to sign inversions of rows in the matrix* C.

Finally, the interval Newton operator also gives rise to a similar but more powerful existence test. However, the computation of the interval Newton operator is more complicated and is not detailed here (cf. [8], [9] and references therein).

2.2. The Frommer-Lang-Schnurr existence test

Following the idea proposed by Moore and Kioustelidis in [7], Frommer et al. proposed in [3] an efficient way to rigorously check the hypothesis of the Poincaré-Miranda theorem using interval analysis. The following existence test is a slight generalization of their theorem (cf. Remark 2).

Theorem 2 (Theorem 3 in [3]): *Consider a nonsingular matrix* $C \in \mathbb{R}^n$ *and, for all* $i \in \{1, \ldots, n\}$ *define*

$$[m]^{(i,\pm)} := \big(C \cdot \mathbf{f}(\tilde{\mathbf{x}})\big)_i + (C \cdot [S])_{ii} \cdot (x_i^{\pm} - \tilde{x}_i) + \sum_{j \neq i} (C \cdot [S])_{ij} \cdot ([x_j] - \tilde{x}_j). \quad (8)$$

If for all $i \in \{1, \ldots, n\}$

$$[m]^{(i,-)} \leq 0 \leq [m]^{(i,+)} \text{ or } [m]^{(i,+)} \leq 0 \leq [m]^{(i,-)} \quad (9)$$

then $\mathbf{f}$ *has a zero in* $[\mathbf{x}]$.

Remark 2: *In the original theorem proposed in [3], the condition* (9) *is replaced by the more restrictive condition* $[m]^{(i,-)} \leq 0 \leq [m]^{(i,+)}$, *the present statement being a trivial consequence of the original. The more general statement is proposed here in order to make Theorem 2 compatible with the invariant of Theorem 1 pointed out in Remark 1. This generalization has no practical interest because* C *is usually chosen to be the midpoint inverse of* $[S]$ *so* $[m]^{(i,+)} \leq 0 \leq [m]^{(i,-)}$ *fails in every case.*

This existence test will be called the Frommer-Lang-Schnurr existence test. For the purpose of its comparison with the weak Hansen-Sengupta existence test, let us consider the *Frommer-Lang-Schnurr*$_<$ *existence test* obtained replacing inequalities by strict inequalities in the statement of Theorem 2: (9) is replaced by the more restrictive condition

$$[m]^{(i,-)} < 0 < [m]^{(i,+)} \text{ or } [m]^{(i,+)} < 0 < [m]^{(i,-)}. \quad (10)$$

Situations where changing inequalities to strict inequalities prevents one from proving existence are very rare and not likely to be met in practice (cf. Subsect. 3.2 for such an atypical situation).

3. The main result

3.1. Comparison of the weak Hansen-Sengupta and the Frommer-Lang-Schnurr$_<$ existence tests

First of all, the weak Hansen-Sengupta existence test needs the assumption that $0 \notin (C \cdot [S])_{ii}$ for all $i \in \{1, \ldots, n\}$ to be applied while this condition is not necessary for the application of the Frommer-Lang-Schnurr existence test. Proposition 1 proves that this difference is of no importance: although the Frommer-Lang-Schnurr$_<$ existence test can be applied when $0 \in (C \cdot [S])_{ii}$ for some $i \in \{1, \ldots, n\}$, it cannot succeed if the assumption that $0 \notin (C \cdot [S])_{ii}$ for all $i \in \{1, \ldots, n\}$ is not satisfied. The following lemma will be used in the proofs of Proposition 1 and Theorem 3.

Lemma 1:

$$[m]^{(i,-)} < 0 < [m]^{(i,+)} \Longrightarrow (C \cdot [S])_{ii} > 0; \tag{11}$$

$$[m]^{(i,+)} < 0 < [m]^{(i,-)} \Longrightarrow (C \cdot [S])_{ii} < 0. \tag{12}$$

Proof: Suppose that $(C \cdot [S])_{ii} > 0$ is false, i.e., there exists $y_i \leq 0$ such that $y_i \in (C \cdot [S])_{ii}$. Notice that $\tilde{x}_j \in [x_j]$ implies $0 \in [x_j] - \tilde{x}_j$. As a consequence $0 \in \sum_{j \neq i} (C \cdot [S])_{ij} \cdot ([x_j] - \tilde{x}_j)$ and finally $\left(C \cdot \mathbf{f}(\tilde{\mathbf{x}})\right)_i + y_i \cdot (x_i^\alpha - \tilde{x}_i) \in [m]^{(i,\alpha)}$ for all $\alpha \in \{-, +\}$. Therefore, $[m]^{(i,-)} < 0 < [m]^{(i,+)}$ would imply

$$y_i(x_i^- - \tilde{x}_i) < -\left(C \cdot \mathbf{f}(\tilde{\mathbf{x}})\right)_i < y_i(x_i^+ - \tilde{x}_i), \tag{13}$$

which is absurd because $y_i \leq 0$ and $x_i^- \leq \tilde{x}_i \leq x_i^+$. The second implication is proved similarly. □

Proposition 1: *If* $0 \in (C \cdot [S])_{ii}$ *for some* $i \in \{1, \ldots, n\}$ *then the Frommer-Lang-Schnurr$_<$ existence test fails.*

Proof: Pick up $i \in \{1, \ldots, n\}$ such that $0 \in (C \cdot [S])_{ii}$. Then both $(C \cdot [S])_{ii} > 0$ and $(C \cdot [S])_{ii} < 0$ are false and Lemma 1 proves that the condition (10) cannot be satisfied, i.e., the Frommer-Lang-Schnurr$_<$ existence test cannot succeed. □

Thanks to Proposition 1, we know that both the Frommer-Lang-Schnurr and the weak Hansen-Sengupta existence tests fail if $0 \in (C \cdot [S])_{ii}$ for some $i \in \{1, \ldots, n\}$. Theorem 3 completes the comparison including the case where $0 \notin (C \cdot [S])_{ii}$ for all $i \in \{1, \ldots, n\}$. The following lemma will be used in the proof of Theorem 3.

Lemma 2: *Let* $[u]$ *and* $[v]$ *be two intervals and* τ *a real. Then,*

$$([u] > 0 \;\; and \;\; 0 < \tau[u] + [v]) \iff ([u] > 0 \;\; and \;\; -[v]/[u] < \tau)\,, \tag{14}$$
$$([u] < 0 \;\; and \;\; \tau[u] + [v] < 0) \iff ([u] < 0 \;\; and \;\; -[v]/[u] < \tau)\,. \tag{15}$$

Proof: First, $[u] > 0$ and $0 < \tau[u] + [v]$ is equivalent to: $u \in [u]$ and $v \in [v]$ imply $u > 0$ and $0 < \tau u + v$. This is equivalent to $u \in [u]$ and $v \in [v]$ imply $u > 0$ and $-v/u < \tau$ which is eventually equivalent to $[u] > 0$ and $-[v]/[u] < \tau$. The second equivalence is proved in a similar way, being careful to reverse an inequality when multiplying its two sides by a negative number. □

Theorem 3: *The Frommer-Lang-Schnurr*$_<$ *existence test succeeds if and only if the weak Hansen-Sengupta existence test succeeds.*

Proof: As proved by Proposition 1, both existence tests fail if $0 \in (C \cdot [S])_{ii}$ for some $i \in \{1, \ldots, n\}$. Therefore, it remains to prove their equivalence under the hypothesis that $0 \notin (C \cdot [S])_{ii}$ for all $i \in \{1, \ldots, n\}$. The domain of α is fixed to $\{-, +\}$ throughout the proof. The condition (10) is obviously equivalent to

$$\Big((\forall\alpha)\left(0 < \alpha[m]^{(i,\alpha)}\right)\Big) \;\text{ or }\; \Big((\forall\alpha)\left(\alpha[m]^{(i,\alpha)} < 0\right)\Big)\,. \tag{16}$$

Now, define $t^{\pm} := x_i^{\pm} - \tilde{x}_i$ and the intervals $[u]$ and $[v]$ by

$$[u] := (C \cdot [S])_{ii} \;\text{ and }\; [v] := g_i(\tilde{\mathbf{x}}) \; + \; \sum_{j \neq i} (C \cdot [S])_{ij} \cdot ([x_j] - \tilde{x}_j), \tag{17}$$

so $[m]^{(i,\pm)} = t^{\pm}[u] + [v]$. Now Lemma 1 proves $0 < \alpha[m]^{(i,\alpha)} \Rightarrow [u] > 0$ and $\alpha[m]^{(i,\alpha)} < 0 \Rightarrow [u] < 0$. Using the definitions (17), (16) is hence equivalent to

$$\begin{aligned} &([u] > 0 \text{ and } (\forall\alpha)\,(0 < \alpha t^{\alpha}[u] + \alpha[v])) \\ &\text{or } ([u] < 0 \text{ and } (\forall\alpha)\,(\alpha t^{\alpha}[u] + \alpha[v] < 0))\,. \end{aligned} \tag{18}$$

Now using Lemma 2, the latter is equivalent to

$$([u] > 0 \text{ and } (P)) \;\text{ or }\; ([u] < 0 \text{ and } (P))\,, \tag{19}$$

where (P) is $(\forall\alpha)\,(-\alpha[v]/[u] < \alpha t^{\alpha})$. Factorizing (P), (19) is equivalent to

$$([u] > 0 \text{ or } [u] < 0) \;\text{ and }\; (\forall\alpha)\left(-\alpha[v]/[u] < \alpha t^{\alpha}\right). \tag{20}$$

That is, $0 \notin [u]$ and $-[v]/[u] < t^{+}$ and $t^{-} < -[v]/[u]$. Using the definitions of $t^{\pm}$, the latter can be written $0 \notin [u]$ and $\tilde{x}_i - [v]/[u] < x^{+}$ and $x^{-} < \tilde{x}_i - [v]/[u]$, which means

$$0 \notin [u] \;\text{ and }\; \tilde{x}_i - [v]/[u] \subseteq \text{int}[x_i]. \tag{21}$$

It remains to notice that $\tilde{x}_i - [v]/[u] = [h_i]$ to conclude the proof. □

3.2. An atypical situation

This section presents an atypical situation where the Frommer-Lang-Schnurr existence test succeeds while all other presented existence tests fail. Consider the function $f(x) := x^2+6x-4$ and the interval $[x] := [-1, 1]$ and $\tilde{x} := 0$. The interval evaluation of the derivative of f will be used as a slope: $[s] := 2[x] + 6$. Let us compute

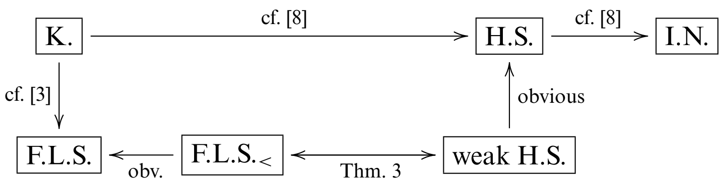
$$[m]^{(-)} := f(\tilde{x}) + [s](\inf[x] - \tilde{x}) = [-12, -8], \quad (22)$$

$$[m]^{(+)} := f(\tilde{x}) + [s](\sup[x] - \tilde{x}) = [0, 4]. \quad (23)$$

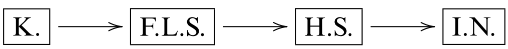
Therefore, $[m]^{(-)} \leq 0 \leq [m]^{(+)}$ and the Frommer-Lang-Schnurr succeeds in proving the existence of a solution. All the other existence tests presented here fail on this example. However, this example is a very atypical situation: $\inf [m]^{(+)}$ is equal to zero while no rounding error has made it strictly negative. Consider $\epsilon \ll 1$. If $[x]$ is changed to $[-1, 1 + \epsilon]$ then all the presented existence tests succeed. On the other hand, if $[x]$ is changed to $[-1, 1 - \epsilon]$ then all the presented existence tests fail. In this latter situation, if one uses the contraction provided by the Hansen-Sengupta operator and then applies a second time the operator using the contracted domain (and re-evaluating the derivative on this new domain), then the existence of a solution is proved.

4. Conclusion

The weak Hansen-Sengupta and the Frommer-Lang-Schnurr$_<$ existence tests have been artificially introduced in order to compare the well known Hansen-Sengupta and the Frommer-Lang-Schnurr existence tests. In spite of their very different expressions, these auxiliary existence tests have been proved to be equivalent. The following hierarchy is now proved to hold between the Krawczyk, the Frommer-Lang-Schnurr, the Frommer-Lang-Schnurr$_<$, the weak Hansen-Sengupta, the Hansen-Sengupta and the interval Newton existence tests (arrows point to more powerful existence tests).



Although a very atypical situation has been displayed where replacing inequalities by strict inequalities in the Frommer-Lang-Schnurr existence test prevents one from proving existence of a solution, in most situations the Frommer-Lang-Schnurr$_<$ and the Frommer-Lang-Schnurr existence tests are equivalent. Therefore, in practice one will meet the following hierarchy.



As a consequence, the Hansen-Sengupta existence test should always be preferred to the Krawczyk and Frommer-Lang-Schnurr existence tests because it is more powerful for approximatively the same computational cost. The Hansen-Sengupta operator also presents an important advantage on the Frommer-Lang-Schnurr existence test independently to its proving power: it contracts the initial domain thus providing an improved enclosure of the potential solution after its application.

Acknowledgements

This work was mainly done during the author's post-doctoral study at the University of Central Arkansas, Arkansas, USA. The author is grateful to the anonymous referees whose comments significantly improved the paper.

References

[1] Alefeld, G., Frommer, A., Heindel, G., Mayer, J.: On the existence theorems of Kantorovich, Miranda and Borsuk. Electronic Trans. Numer. Anal. *17*, 102–111 (2004).
[2] Alefeld, G., Potra, F. A., Shen, Z.: On the existence theorems of Kantorovich, Moore and Miranda. In: Topics in Numerical Analysis with Special Emphasis on Nonlinear Problems (Alefeld, G., Chen, X., eds.), pp. 21–28. Computing Suppl. 2001.
[3] Frommer, A., Lang, B., Schnurr, M.: A comparison of the Moore and Miranda existence tests. Computing *72*, 349–354 (2004).
[4] Hayes, B.: A lucid interval. American Scientist *91*(6), 484–488 (2003).
[5] Kearfott, R. B.: Interval computations: Introduction, uses, and resources. Euromath Bull. *2*(1), 95–112 (1996).
[6] Miranda, C.: Un' osservazione su un teorema di Brouwer. Bolletino Unione Mathematica Italiana *2*(3), 5–7 (1940).
[7] Moore, R. E., Kioustelidis, J.: A simple test for accuracy of approximate solutions to nonlinear (or linear) systems. SIAM J. Numer. Anal. *17*(4), 521–529 (1980).
[8] Neumaier, A.: Interval methods for systems of equations. Cambridge: Cambridge University Press 1990.
[9] Neumaier, A.: A simple derivation of the Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations. Reliab. Comp. *5*, 131–136 (1999).
[10] Poincaré, H.: Sur certaines solutions particulières du problème des trois corps. Comptes rendus de l'Académie des sciences *97*, 251–252 (1883).
[11] Schnurr, M.: On the proofs of some statements concerning the theorems of Kantorovich, Moore and Miranda. Reliab. Comp. *11*, 77–85 (2005).

A. Goldsztejn
Information and Computer Science
University of California
Irvine, CA 92697
USA
e-mail: alexandre@goldsztejn.com