1 Remarks on the Notation of Interval Arithmetic

In the algorithm and the analysis, when we write a box function $\square f(B)$, we do not specify the form of $\square f(B)$ nor the way we evaluate the interval, but we require the box function $\square f(B)$ to satisfy the following properties:
(a) the inclusion property, that is, the box function over a box B must contain the exact range of the function over B.
(b) the convergence property, that is, when the width of B tends to 0, the width of the box function also tends to 0.

Box function of the form $\square f_{\mathrm{M}}(B) := f(m(B)) + \square\nabla f(B)(\mathbf{x} - B)$ is called the mean value form of f over B.

2 Miranda Test and MK Test

Most of this section and the next section comes from [1].

Proposition 1 (Effective Miranda Test) *Let $F := (f_1, \ldots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function with appropriate box functions. Write $f_{i,j} := \frac{\partial f_i}{\partial x_j}$. For any box B with width $w(B) = (w_1, \ldots, w_n)$, if for all $i = 1, \ldots, n$*

$$f_i(m(B_i^+)) \cdot f_i(m(B_i^-)) < 0 \tag{1}$$
$$\square f_M(B_i^+) > 0 \tag{2}$$
$$\square f_M(B_i^-) < 0 \tag{3}$$

then F has a zero in the interior of B.

MK-test for a system F on a box B is the effective Miranda-test applied to the system $J_F(m(B))^{-1}F$ where the Jacobian is non-singular.

3 Termination of MK Test

Given $x, y \in R$, the notation $x \pm y$ denotes a number of the form $x + \theta y$, where θ is such that $0 \le |\theta| \le 1$; thus "$\pm$" hides the θ implicit in the definition. We further extend this notation to matrices in the following sense: for two matrices A, B, the matrix $A \pm B := [a_{ij} \pm b_{ij}]$; also, for a scalar λ, the matrix $A \pm \lambda := [a_{ij} \pm \lambda]$. For $x, y \in \mathbb{R}^n$, we abuse the notation $[\mathbf{x}, \mathbf{y}]$ to denote the line segment connecting $\mathbf{x}$ and $\mathbf{y}$. And we will write $\|\mathbf{x} - \mathbf{y}\|_2$ as $\|\mathbf{x} - \mathbf{y}\|$ for short.

We now recall the Mean Value Theorem for $F : \mathbb{R}^n \to \mathbb{R}^n$: Given two points $x, y \in \mathbb{R}^n$, there exists a matrix $K_{[\mathbf{x},\mathbf{y}]}$ with non-negative entries such that

$$F(\mathbf{x}) - F(\mathbf{y}) = (J_F(\mathbf{y}) \pm K_{[\mathbf{x},\mathbf{y}]}\|\mathbf{x} - \mathbf{y}\|) \cdot (\mathbf{x} - \mathbf{y}). \tag{4}$$

To see this claim, we apply the mean value theorem in each of the components of F to obtain

$$\begin{aligned} f_i(\mathbf{x}) - f_i(\mathbf{y}) &= \nabla f_i(\widetilde{\mathbf{y}}) \cdot (\mathbf{x} - \mathbf{y}) \\ &= (f_{i,1}(\widetilde{\mathbf{y}}), \ldots, f_{i,n}(\widetilde{\mathbf{y}})) \cdot (\mathbf{x} - \mathbf{y}) \end{aligned}$$

where $\widetilde{\mathbf{y}} \in [\mathbf{x}, \mathbf{y}]$. Then we apply the mean value theorem to each component $f_{i,j}(\widetilde{\mathbf{y}})$ for $j = 1, \ldots, n$ to get

$$\begin{aligned} f_{i,j}(\widetilde{\mathbf{y}}) &= f_{i,j}(\mathbf{y}) + \nabla f_{i,j}(\widehat{\mathbf{y}}) \cdot (\mathbf{y} - \widetilde{\mathbf{y}}) \qquad \text{with } \widehat{\mathbf{y}} \in [\mathbf{y}, \widetilde{\mathbf{y}}] \\ &= f_{i,j}(\mathbf{y}) \pm K_{i,j}\|\mathbf{x} - \mathbf{y}\| \end{aligned}$$

where $K_{i,j} := n \cdot \max_{1\le k\le n} |\frac{\partial f_i([\mathbf{x},\mathbf{y}])}{\partial x_j \partial x_k}|$. So

$$\begin{aligned} f_i(\mathbf{x}) - f_i(\mathbf{y}) &= (f_{i,1}(\mathbf{y}) \pm K_{i,1}\|\mathbf{x}-\mathbf{y}\|, \ldots, f_{i,1}(\mathbf{y}) \pm K_{i,n}\|\mathbf{x}-\mathbf{y}\|) \cdot (\mathbf{x}-\mathbf{y}) \\ &= \nabla f_i(\mathbf{y})(\mathbf{x}-\mathbf{y}) \pm (K_{i,1}, \ldots, K_{i,n}) \cdot (\mathbf{x}-\mathbf{y})\|\mathbf{x}-\mathbf{y}\| \end{aligned}$$

for $i = 1, \ldots, n$. The previous claim follows with $K_{[\mathbf{x},\mathbf{y}]} := [K_{i,j}]$.

For later use, we also recall the Mean Value Theorem for J_F: Given two points $x, y \in \mathbb{R}^n$, there exists a matrix $K_{[\mathbf{x},\mathbf{y}]}$ with non-negative entries such that

$$J_F(\mathbf{x}) = J_F(\mathbf{y}) \pm K_{[\mathbf{x},\mathbf{y}]}\|\mathbf{x}-\mathbf{y}\|. \tag{5}$$

For simplicity, we write J_F as J. To show this claim, we apply the Mean Value Theorem to each entry J_{ij} to get

$$\begin{aligned} J_{ij}(\mathbf{x}) &= J_{ij}(\mathbf{y}) + \nabla J_{ij}(\widetilde{\mathbf{y}}) \cdot (\mathbf{x}-\mathbf{y}) \qquad \text{with } \widetilde{\mathbf{y}} \in [\mathbf{x},\mathbf{y}] \\ &= J_{ij}(\mathbf{y}) \pm K_{[\mathbf{x},\mathbf{y}]_{ij}}\|\mathbf{x}-\mathbf{y}\| \end{aligned}$$

where $K_{[\mathbf{x},\mathbf{y}]_{ij}} := n \cdot \max_{1\le k\le n} |\frac{\partial J_{ij}([\mathbf{x},\mathbf{y}])}{\partial x_k}| = n \cdot \max_{1\le k\le n} |\frac{\partial f_i([\mathbf{x},\mathbf{y}])}{\partial x_j \partial x_k}|$.

3.1 Termination of Preconditioned Miranda Test

Let B be a box with width $w(B) = (w_1, \ldots, w_n)$, denote $\langle w(B)\rangle := \min_{1\le k\le n} w_k$. The following Lemma proves that when the box B is sufficiently small. The preconditioned Miranda test will succeed.

Lemma 2 *Let F be a zero-dimensional system of polynomials with simple isolated roots. For all sufficiently small open boxes B that contains a single root α of F in $\frac{B}{2}$, the modified system $G_B := J_F(m(B))^{-1}F = (g_1, \ldots, g_n)$ will be well defined, and satisfies the condition that $g_i(B_i^+) \ge \frac{\langle w(B)\rangle}{8}$ and $g_i(B_i^-) \le -\frac{\langle w(B)\rangle}{8}$.*

Proof. Let $\mathbf{x}$ be a point on the boundary of the box B. For simplicity, we write $m(B)$ as m. From the definition of G_B and from the mean value theorem (4) it holds

$$G_B(\mathbf{x}) = J_F(m)^{-1}(F(\alpha) + (J_F(\alpha) \pm K_{[\mathbf{x},\alpha]}\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha))$$

where $K_{[\mathbf{x},\alpha]_{i,j}}$ is $n \cdot \max_{1\le k\le n} |\frac{\partial f_i([\mathbf{x},\alpha])}{\partial x_j \partial x_k}|$. And from the mean value theorem (5), it follows

$$J_F(\alpha) = J_F(m) \pm K_{[\alpha,m]}\|\alpha - m\|$$

where $K_{[\alpha,m]_{i,j}}$ is $n \cdot \max_{1\le k\le n} |\frac{\partial f_i([\alpha,m])}{\partial x_j \partial x_k}|$. Since α is contained in $\frac{B}{2}$, it holds $\|\alpha - m\| \le \|\mathbf{x} - \alpha\|$. Thus

$$\begin{aligned} G_B(\mathbf{x}) &= J_F(m)^{-1}(F(\alpha) + (J_F(\alpha) \pm K_{[\mathbf{x},\alpha]}\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha)) \\ &= J_F(m)^{-1}(F(\alpha) + (J_F(m) \pm K_{[\alpha,m]}\|\alpha - m\| \pm K_{[\mathbf{x},\alpha]}\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha)) \\ &= J_F(m)^{-1}(F(\alpha) + (J_F(m) \pm \widehat{K}\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha)) \\ &= J_F(m)^{-1}(J_F(m)^{-1} + \widehat{K}\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha)) \\ &= (\mathbf{1} \pm \|J_F(m)^{-1}\widehat{K}\|_\infty\|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha) \end{aligned}$$

where $\widehat{K} = K_{[\mathbf{x},\alpha]} + K_{[\alpha,m]}$.

The i-th component in the vector

$$(\mathbf{1} \pm \|J_F(m)^{-1}\widehat{K}\|_\infty \|\mathbf{x}-\alpha\|) \cdot (\mathbf{x}-\alpha)$$

is the polynomial $g_i(B)$, so we obtain

$$|g_i(\mathbf{x}_i) - (\mathbf{x}_i - \alpha_i)| \leq \|\mathbf{x}-\alpha\| \|J_F^{-1}\widehat{K}\|_\infty \sum_{j=1}^n |\mathbf{x}_j - \alpha_j|.$$

The term on the RHS

$$\|\mathbf{x}-\alpha\| \|J_F^{-1}\widehat{K}\|_\infty \sum_{j=1}^n |\mathbf{x}_j - \alpha_j| \leq \|w(B)\|_1^2 \|J_F^{-1}\widehat{K}\|_\infty,$$

because $\|\mathbf{x}-\alpha\| \leq \|w(B)\|_2 \leq \|w(B)\|_1$ and $\sum_{j=1}^n |x_j - \alpha_j| \leq \|w(B)\|_1$. Suppose the box B is such that

$$2\|w(B)\|_1^2 \|J_F^{-1}\widehat{K}\|_\infty < \frac{\langle w(B)\rangle}{8}$$

then we claim that for all $i = 1, \ldots, n$, $g_i(B_i^+) \geq \frac{\langle w(B)\rangle}{8}$ and $g_i(B_i^-) \leq -\frac{\langle w(B)\rangle}{8}$. This is because for all $\mathbf{x} \in B_i^+$, $\mathbf{x}_i - \alpha_i \geq \frac{\langle w(B)\rangle}{4}$ since $\alpha \in \frac{B}{2}$, which implies that $g_i(x_i) > \frac{\langle w(B)\rangle}{4} - \frac{\langle w(B)\rangle}{8} = \frac{\langle w(B)\rangle}{8}$. Similar argument applies for $\mathbf{x} \in B_i^-$. **Q.E.D.**

3.2 Termination of MK-test

First we recall a theorem in [2].

Let a, b be intervals in $\mathbb{R}$ and denote $q(a,b) := \sup\{|\underline{a} - \underline{b}|, |\bar{a} - \bar{b}|\}$.

Theorem 3 *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function. Then*

$$q(f_M(B), f(B)) \leq 2 \sum_{i=1}^n w(\square \frac{\partial f(B)}{\partial x_i}) \cdot w_i. \tag{6}$$

Now we can prove that when the box is sufficiently small, the MK-test succeeds. Henceforth, we assume that $w_1 = \cdots = w_n = w_B$.

Theorem 4 *Let F be a zero-dimensional system of polynomials with simple isolated roots. For all sufficiently small open boxes B that contains a single root α of F in $\frac{B}{2}$, the modified system $G_B := J_F(m(B))^{-1}F$ will be well defined and satisfies the MK-test.*

Proof. In Lemma 2, it is proven that when B is sufficiently small, it holds that $g_i(B_i^+) \geq \frac{w_B}{8}$ and $g_i(B_i^-) \leq -\frac{w_B}{8}$. And from Theorem 3, we have

$$\begin{aligned} q(g_{i\mathrm{M}}(B_i^+), g_i(B_i^+)) &\leq 2 \sum_{j=1, j\neq i}^n w(\square \frac{\partial g_i(B_i^+)}{\partial x_j}) \cdot w_B \\ &\leq 2(n-1) w_B \cdot \max_{1 \leq j \leq n, j \neq i} w(\square \frac{\partial g_i(B_i^+)}{\partial x_j}). \end{aligned}$$

By the convergence of interval extension, $w(\square\frac{\partial g_i(B_i^+)}{\partial x_j})$ approaches 0 when w_B approaches 0 for $j = 1, \ldots, n$. Thus when w_B is small enough, we have $w(\square\frac{\partial g_i(B_i^+)}{\partial x_j}) \leq \frac{1}{32(n-1)}, \forall j = 1, \ldots, n$. In this case, it is easy to see that

$$\begin{aligned} g_{i\mathrm{M}}(B_i^+) &\geq g_i(B_i^+) - q(g_{i\mathrm{M}}(B_i^+), g_i(B_i^+)) \\ &\geq \frac{w_B}{8} - 2(n-1)w_B \cdot \frac{1}{32(n-1)} = \frac{w_B}{16} > 0 \end{aligned}$$

Similar argument applies to $g_{i\mathrm{M}}(B_i^-)$. Thus the conditions in the effective Miranda test are satisfied.

Q.E.D.

4 A simple algorithm based on subdivision and MK-test

Given a polynomial system $F : \mathbb{R}^n \to \mathbb{R}$ and an initial box $B \subset \mathbb{R}^n$, if for each root α of F contained in B, it holds $\det(J_F(\alpha)) \neq 0$, then we can isolate all the roots in B by subdivision method and MK-test. First we describe 3 predicates, C_0, C_1 and Jacobian condition JC over a box B:
$C_0(B)$: $\exists$ an integer $i \in [1, n]$ such that $0 \notin \square f_i(B)$;
$C_1(B)$: MK-test succeeds over B;
JC(B): $0 \notin \det(\square J_F(B))$.
For a box B, we denote by $\mathcal{Z}_F(B)$ the set of roots of F contained in B, and $|\mathcal{Z}_F(B)|$ is the number of roots of F contained in B. It is easy to see that if C_0 succeeds, then $|\mathcal{Z}_F(B)| = 0$; if C_1 succeeds, then $|\mathcal{Z}_F(B)| \geq 1$; if JC($B$) succeeds, then $|\mathcal{Z}_F(B)| \leq 1$.

```
PROCEDURE Real Root Isolating of Polynomial System
INPUT: A polynomial system F, an initial box B_0.
OUTPUT: A queue P of isolating boxes of F such that Z_F(B_0) ⊆ ⋃_{B∈P} Z_F(B) ⊆ Z_F(2B_0).
1.  Initialize the priority queue Q ← {B_0} and the output queue P ← ∅.
2.  While Q ≠ ∅ do the following:
3.      Remove a biggest box B from Q.
4.      If C_0(B) fails then
5.          If JC(3B) succeeds then
6.              Initialize queue Q' ← B.
7.              While Q' ≠ ∅ do the following:
8.                  B' ← Q'.pop().
9.                  If (B' = B) ∨ C_0(B') fails then
10.                     If C_1(2B') succeeds then
11.                         P.add(2B').
12.                         Discard all the boxes in Q that are contained in 3B.
13.                         Break.
14.                 Split B' into 2^n congruent subboxes and add them to Q'.
15.         Else
16.             Split B into 2^n congruent subboxes and add them to Q.
```

Lemma 5 *The algorithm will terminate and output a queue of isolating boxes of F such that $\mathcal{Z}_F(B_0) \subseteq \bigcup_{B\in P} \mathcal{Z}_F(B) \subseteq \mathcal{Z}_F(2B_0)$.*

Proof. First we prove the termination of the algorithm. We prove this by contradiction. Suppose the algorithm does not terminate, then we can find an infinite series of boxes $B_1, B_2, \cdots$ such that $B_1 \supset B_2 \supset \cdots$. Since width of the boxes is halved each time, this infinite series must converge to a point, denoted as p. We divide the proof into 2 cases depending on whether p is a root of F.

If p is a root of F, then from Theorem 4, predicate C_1 will succeed on any sufficient box containing p. That is, there exists an integer A_1 such that $C_1(B_i)$ succeeds for $\forall i \geq A_1$. And by assumption, $\det(J_F(p)) \neq 0$. Hence the convergence of interval extension implies that there exists an integer A_2 such that $\mathrm{JC}(B_i)$ succeeds for $\forall i \geq A_2$. Therefore, $C_1(B_i) \wedge \mathrm{JC}(B_i)$ is true for $\forall i \geq \max\{A_1, A_2\}$. And an isolating box will be output. This is a contradiction.

If f is not a root of F, then by the convergence of interval extension, there exists an integer A_3 such that $C_0(B_i)$ succeeds for $\forall i \geq A_3$. That is, the box B_i will be discarded for $\forall i \geq A_3$. This is also a contradiction.

Then we prove the correctness of the algorithm. From the algorithm, for each output box $B \in P$, $C_1(B)$ succeeds, thus a root of F is contains in the interior of B. We need to show that $\mathcal{Z}_F(B_0) \subseteq \bigcup_{B\in P} \mathcal{Z}_F(B) \subseteq \mathcal{Z}_F(2B_0)$ and that each root of F is contained in only one box in P.

For any box $B \subset B_0$, it follows $2B \subset 2B_0$, thus it is easy to see that $\bigcup_{B\in P} \mathcal{Z}_F(B) \subseteq \mathcal{Z}_F(2B_0)$. To show $\mathcal{Z}_F(B_0) \subseteq \bigcup_{B\in P} \mathcal{Z}_F(B)$, it suffices to prove that no root is contained in any discarded box. From the algorithm, a box is discarded in two cases. The first case is when C_0 predicate succeeds on the box, where the discarded box contains no root. The second case is in instruction line 12 of the algorithm where $\mathrm{JC}(3B)$ succeeds for a box B and $C_1(2B')$ succeeds for a box $B' \subset B$. In this case, all the boxes in $3B \setminus 2B'$ are discarded. There is no root contained in $3B \setminus 2B'$ because the success of $\mathrm{JC}(3B)$ implies that at most one root of F is contained in $3B$ and the success of $C_1(2B')$ suggests that there is one root of F contained in $2B'$.

It remains to prove that any two distinct boxes in P do not contain a same root of F. To this end, we will show that the interior of any two boxes in P do not overlap. Take $B'_1, B'_2 \in P$ and without loss of generality, suppose that B'_1 is output earlier than B'_2. By the algorithm there exists B_1 such that $B'_1 \subset 2B_1$ and $\mathrm{JC}(3B_1)$ succeeds. Since all the boxes in $3B_1 \setminus B'_1$ are discarded, the distance to from B'_1 to any remained box is at least $\frac{1}{2}w(B_1)$. Hence the distance from B'_1 to $\frac{1}{2}B'_2$ is at least $\frac{1}{2}w(B_1)$. Now it suffices to show that $\frac{1}{4}w(B'_2) \leq \frac{1}{2}w(B_1)$. This is true because B_1 is a box of biggest size in Q, thus $w(B_1) \geq w(\frac{1}{2}B'_2)$. **Q.E.D.**

Remark:
1. In instruction line 5, the Jacobian condition is required to succeed over $3B$ instead of $2B$ so that the output boxes contain distinct roots of F.
2. The boxes in the output queue P may contain the roots in $2B_0 \setminus B_0$, this is a boundary issue hard to escape. But instead of $2B_0$, we can restrict the this region to $(1+\epsilon)B_0$ where ϵ is any positive constant. In this case, the size of the output boundary box should be less than 2ϵ.

References

[1] Jyh-Ming Lien, Vikram Sharma, Gert Vegter, and Chee Yap. Isotopic arrangement of simple curves: An exact numerical approach based on subdivision. In *ICMS 2014*, pages 277–282. Springer, 2014. LNCS No. 8592. Download from http://cs.nyu.edu/exact/papers/ for a version with Appendices and details on MK Test.

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