

Improving the Efficiency of Exclusion Algorithms

Kurt Georg¹³⁴

Colorado State University²

November 2000

1 Introduction

Exclusion algorithms are a well-known tool in the area of interval analysis, see, e.g., [5, 6], for finding all solutions of a system of nonlinear equations. They also have been introduced in [14, 15] from a slightly different viewpoint. In particular, such algorithms seem to be very useful for finding all solutions of low-dimensional, but highly nonlinear systems which have many solutions. Such systems occur, e.g., in mechanical engineering.

A different choice of algorithm for finding solutions of difficult nonlinear systems of equations are homotopy methods, see, e.g., the survey paper [1]. However, in the context of finding *all* solutions, such methods have been mainly successful for polynomial systems, and there is a vast literature on this special topic, see, e.g., [1]. Note, however, that homotopy methods typically generate all complex-valued solutions, even if the coefficients of the system are real, whereas the exclusion methods aim directly at real-valued solutions.

Of course, homotopy and exclusion methods may be combined. An example we have in mind is the recent paper [12] on a numerical primary decomposition for the solution components of polynomial systems by Sommese, Verschelde, and Wampler. Here an exclusion algorithm could be used as a module to investigate real components of a suitably reduced subproblem.

We briefly describe our view of an exclusion method.

In \mathbb{R}^n and $\mathbb{R}^{m \times n}$ we use the component-wise ‘ \leq ’ as a partial ordering, and $|\cdot|$ is the component-wise absolute value. We only use the max norm ‘ $\|\cdot\|$ ’. For example, for two matrices $A, B \in \mathbb{R}^{n \times n}$ the symbol $A \leq B$ means that $A(i, j) \leq B(i, j)$ for $i, j = 1 : n$.

An interval σ in \mathbb{R}^n is a rectangular box, i.e., there are two vectors $m_\sigma, r_\sigma \in \mathbb{R}^n$ with $r_\sigma(i) > 0, i = 1 : n$, such that

$$\sigma = [m_\sigma - r_\sigma, m_\sigma + r_\sigma] = \{x \in \mathbb{R}^n : m_\sigma - r_\sigma \leq x \leq m_\sigma + r_\sigma\} .$$

We call $m_\sigma - r_\sigma$ the lower corner, $m_\sigma + r_\sigma$ the upper corner, m_σ the midpoint, and r_σ the radius of σ . (This corresponds to the midpoint-radius representation in interval analysis.)

Here and in the following we use the short ‘iff’ for ‘if and only if’. Equations, definitions, lemmas, etc., are numbered consecutively.

¹Partially supported by NSF via grant # DMS-9870274

²Department of Mathematics, Fort Collins, CO 80523

³<http://www.math.colostate.edu/~georg/>

⁴<mailto://georg@math.colostate.edu>

Let $\sigma \subset \mathbb{R}^n$ be an interval and $F : \sigma \rightarrow \mathbb{R}^n$ a function defined on σ . We call a test

$$T_F(\sigma) \in \{0, 1\} \quad \text{where} \quad 0 \equiv \text{no} \quad \text{and} \quad 1 \equiv \text{yes}$$

an *exclusion test* for F on σ iff $T_F(\sigma) = 0$ implies that F has no zero point in σ . Hence, $T_F(\sigma) = 1$ is a *necessary* condition for F to have a zero point in σ .

This notion is strongly reminiscent of the inclusion test introduced in an abstract setting in [4]. It seems that the notion and use of exclusion tests goes at least back to Moore, see [8, $E(X)$ on p. 77].

If an exclusion test is given, then we can recursively bisect intervals and discard the ones which yield a negative test. This leads to the following recursive *Exclusion Algorithm* which we start from some initial interval Λ on which F is defined. We assume that an exclusion test $T_F(\sigma)$ is available for all subintervals $\sigma \subset \Lambda$.

Algorithm 1 (Exclusion Algorithm).

```

 $\Gamma \leftarrow \{ \Lambda \}$  (initial interval)
for  $\ell = 1 : \text{maximal\_level}$ 
  for  $a = 1 : n$ 
    let  $\tilde{\Gamma}$  be obtained by bisecting each  $\sigma \in \Gamma$  along the axis  $a$ 
    for  $\sigma \in \tilde{\Gamma}$ 
      if  $T_F(\sigma) = 0$ 
        drop  $\sigma$  from  $\tilde{\Gamma}$  ( $\sigma$  is excluded)
     $\Gamma \leftarrow \tilde{\Gamma}$ 
   $\Gamma_\ell \leftarrow \Gamma$  (for later reference)

```

Remark 2. The exclusion algorithm is similar to early algorithms in interval analysis. It turns out that bisection is an efficient partitioning strategy. In order to simplify and unify our efficiency investigations, we have considered only the strategy of cyclic bisections of the intervals along subsequent axes. Various authors have investigated bisection schemes. For a fairly early discussion see [8, pp. 77–81]. For a further careful comparison of bisection schemes, see [2]. This will be further investigated in [3].

For clarity of exposition and notation, the list of intervals is processed breadth-first rather than depth-first. However, we mention that the other choice (which uses less memory) was actually implemented by the author. It is easy to see that the complexity analysis presented in this paper is not influenced by this difference in choice. We refer also to the analysis appearing in [10, pp. 77–80 and pp. 85–102].

Whenever one cycle of bisections is accomplished, we say that we have reached a new bisection level, and we think of an exclusion algorithm as performing a fixed number of bisection levels. The intervals which have not been discarded after ℓ bisection levels will be considered as the intervals which the algorithm generates on the ℓ -th bisection level, see Figure 1 for an illustration. The list of these intervals is denoted by Γ_ℓ in the algorithm. Obviously, if $\Gamma_\ell = \emptyset$ for some level ℓ , then the algorithm has shown that there are no zero points of F in the initial interval Λ .

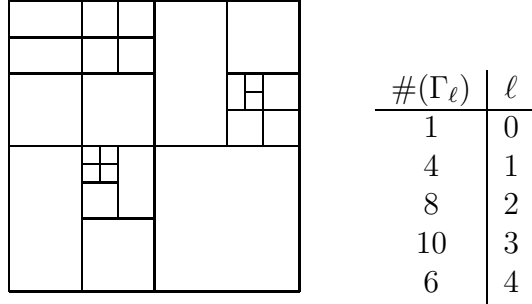


Figure 1: Illustration of Bisection Levels

All the exclusion tests that we will discuss are applied component-wise on the vector-valued function F . Hence, we only need to consider an exclusion test for a scalar-valued function $f : \sigma \rightarrow \mathbb{R}$, and then we can combine such (possibly differing types of) exclusion tests to obtain an exclusion test for a vector-valued function $F = \{f_i\}_{i=1:n} : \sigma \rightarrow \mathbb{R}^n$ by setting

$$T_F(\sigma) := \prod_{i=1}^n T_{f_i}(\sigma).$$

Thus, in the following, we will mainly restrict our attention to scalar functions $f : \sigma \rightarrow \mathbb{R}$ when designing exclusion tests.

It is clear that the efficiency of exclusion algorithms hinges mainly on the construction of a good exclusion test which is computationally inexpensive but relatively tight. Otherwise, too many intervals remain undiscarded on each bisection level, and this leads to significant numerical inefficiency.

In the area of interval analysis, the idea of exclusion is exploited in *interval branch and bound algorithms* which are used to find all the zero points of a nonlinear system of equations, or also to minimize functions, see, e.g., Kearfott [6] and the bibliography cited there, and the software package GlobSol accompanying the book [5]. From an interval analysis viewpoint, a simple exclusion test could be designed in the following way:

$$T_f(\sigma) = 1 : \iff 0 \in [f](\sigma)$$

where $[f](\sigma)$ is the interval obtained from σ by applying f in an interval analysis sense. More sophisticated tests employ an interval-Newton step.

In [15], two exclusion tests were given:

- Let $L > 0$ be a Lipschitz constant for f on the interval σ . Then

$$|f(m_\sigma)| \leq L \|r_\sigma\| \tag{3}$$

is an exclusion test for f on σ .

- If $f = g - h$ is the difference of two increasing functions on σ , then

$$h(m_\sigma - r_\sigma) \leq g(m_\sigma + r_\sigma) \quad \mathbf{and} \quad h(m_\sigma + r_\sigma) \geq g(m_\sigma - r_\sigma) \tag{4}$$

is an exclusion test for f on σ .

In [14], an exclusion test based on power series was presented. We first need to introduce a definition:

Definition 5. For power series $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$ and $g(x) = \sum_{\alpha} g_{\alpha} x^{\alpha}$ we define $f \prec\prec g$ iff $|f_{\alpha}| \leq g_{\alpha}$ for all α .

Now, if $f \prec\prec g$, and if the power series g converges on σ , then

$$|f(m_{\sigma})| \leq g(|m_{\sigma}| + r_{\sigma}) - g(|m_{\sigma}|) \quad (6)$$

is an exclusion test for f on σ .

For all these tests, the following complexity result was shown in [14, 15]:

Theorem 7. *Let $\Lambda \subset \mathbb{R}^n$ be an interval, and let $F : \Lambda \rightarrow \mathbb{R}^n$ be sufficiently smooth and zero a regular value of F . Then there is a constant $C > 0$ such that the exclusion algorithm, started in Λ , generates no more than C intervals on each bisection level, i.e., $\#(\Gamma_{\ell}) \leq C$ independent of ℓ .*

A related analysis, concerning clustering of undiscarded intervals on various levels as a function of the sharpness of the lower bound on the range, was given in [7].

Hence, if the complexity of one exclusion test is known, then the previous theorem leads immediately to a complexity statement on the efficiency of an exclusion algorithm. However, the constant C could be very big, and numerical experiments show that the exclusion algorithms based on (3), (4), and (6) are not tight enough for more demanding nonlinear systems, such as those which typically occur in engineering. The aim of the present paper is to generate and analyze refined tests in such a way that they lead to more powerful and efficient exclusion algorithms. It will turn out that even higher singularities in a solution point (as long as the solution point is isolated) does not destroy the complexity addressed in the preceding theorem if suitable exclusion tests are used.

2 Construction of Dominant Functions

The test (6) is an example of how a *dominant function* may be used to obtain an exclusion test. Let us now begin to outline our general approach to construct exclusion tests. We denote by \mathbb{Z}_+ the set of nonnegative integers. For a multi-index

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$$

we consider the following definitions:

1. The length of α is defined by $|\alpha| := \sum_i \alpha_i$.
2. The factorial of α is defined by $\alpha! := \prod_i \alpha_i!$.
3. If $x \in \mathbb{R}^n$, then we define $x^{\alpha} := \prod_i x_i^{\alpha_i}$.
4. We define the partial derivatives $\partial^{\alpha} = (\alpha!)^{-1} \prod_i \partial_i^{\alpha_i}$.

Furthermore, we introduce the probability measures

$$\omega_k(dt) = k(1-t)^{k-1} dt$$

on the interval $[0, 1]$.

Using these definitions, Taylor's formula with $k > 0$ and integral remainder is easy to write:

$$\begin{aligned} f(m+h) &= f(m) + \sum_{0 < |\alpha| < k} \partial^\alpha f(m) h^\alpha \\ &+ \sum_{|\beta|=k} \int_0^1 \partial^\beta f(m+th) \omega_k(dt) h^\beta \end{aligned} \quad (8)$$

Definition 9. Let $\sigma \subset \mathbb{R}^n$ be an interval. By $\mathcal{A}_k(\sigma)$ we denote the space of functions $f : \sigma \rightarrow \mathbb{R}$ such that $\partial^\alpha f$ is absolutely continuous for $|\alpha| < k$. Note that for $f \in \mathcal{A}_k$ the Taylor formula (8) holds. In $\mathcal{A}_k(\sigma)$ we introduce the cone

$$\mathcal{K}_k(\sigma) = \{ g \in \mathcal{A}_k(\sigma) : 0 \leq \partial^\alpha g(x) \leq \partial^\alpha g(y) \text{ for } 0 \leq x \leq y, |\alpha| \leq k. \}$$

We also set

$$\mathcal{A}_\infty(\sigma) := \bigcap_{k=1}^{\infty} \mathcal{A}_k(\sigma) \quad \text{and} \quad \mathcal{K}_\infty(\sigma) := \bigcap_{k=1}^{\infty} \mathcal{K}_k(\sigma).$$

We now introduce the notion of a dominant function which will be the basis for the estimates of this paper:

Definition 10. Let $f \in \mathcal{A}_k(\sigma)$ and $g \in \mathcal{K}_k(\sigma)$. Then $f(x) \prec_k g(x)$ for $x \in \sigma$ (g dominates f with order k on σ) iff the estimates

$$|\partial^\alpha f(x)| \leq \partial^\alpha g(|x|)$$

hold for all $x \in \sigma$ and $|\alpha| \leq k$. If $f \in \mathcal{A}_\infty(\sigma)$ and $g \in \mathcal{K}_\infty(\sigma)$, then $f(x) \prec_\infty g(x)$ for $x \in \sigma$ means that $f \prec_k g$ for $x \in \sigma$ and all $k \geq 0$.

Note that $f(x) \prec_k g(x)$ for $x \in \sigma$ by definition implies that $f(x) \prec_q g(x)$ for $x \in \tau$, provided that $q \leq k$ and $\tau \subset \sigma$. We will frequently use the notation $f \prec_k g$ or $f(x) \prec_k g(x)$ if there is no ambiguity about the underlying interval.

Let us first show how Definition 5 relates to these notions:

Theorem 11. Let $f(x) = \sum_\alpha f_\alpha x^\alpha$ and $g(x) = \sum_\alpha g_\alpha x^\alpha$ be power series which are convergent on an interval $\sigma \in \mathbb{R}^n$ containing the origin. Then

$$f(x) \prec_\infty g(x) \text{ for } x \in \sigma \iff f \prec\prec g.$$

Proof. If $f \prec_\infty g$, then in particular

$$|f_\alpha| = |\partial^\alpha f(0)| \leq \partial^\alpha g(0) = g_\alpha,$$

and hence $f \prec g$. Now, assume that $f \prec g$ holds. For technical reasons we introduce the monomial $\xi^\alpha : x \mapsto x^\alpha$. Then we estimate termwise:

$$|\partial^\beta f(x)| \leq \sum_{\alpha} |f_\alpha| |\partial^\beta \xi^\alpha(x)| \leq \sum_{\alpha} g_\alpha \partial^\beta \xi^\alpha(|x|) = \partial^\beta g(|x|)$$

and hence $f \prec_\infty g$ holds. □

The following examples point out the differences between the various estimates:

Example 12.

1. If $g \in \mathcal{K}_k$ then $g \prec_k g$. This includes examples such as $\exp(m+x) \prec \exp(m+x)$, and $\tan x \prec \tan x$ for $|x| < \frac{\pi}{2}$.
2. $\sin x \prec \sinh x$, but $\sin x \prec_3 x + \frac{1}{6}x^3$.
3. $\cos x \prec \cosh x$, but $\cos x \prec_1 1+x$, $\cos x \prec_2 1 + \frac{1}{2}x^2$, and $\cos x \prec_3 1 + \frac{1}{2}x^2 + \frac{1}{6}x^3$.
4. $\log(1+x) \prec \log(1-x)$ but $\log(1+x) \prec_3 x + \frac{1}{2}x^2 + \frac{1}{3}x^3$ for $|x| < 1$.
5. $\sin(m+x) \prec \sinh(|m|+x)$ but $\sin(m+x) \prec_2 |\sin(m)| + |\cos(m)|x + \frac{1}{2}x^2$

In the following we list some rules that can be used as a tool to generate dominant functions, in much the same way as rules about differentiation are used as a tool to generate derivatives. Most of these rules have been shown in [14] for the case ‘ \prec ’ of power series. It turns out that our more general proofs are simpler since they use derivatives instead of the coefficients of power series.

Theorem 13.

1. $f \prec_k g$ implies $f(m+x) \prec_k g(|m|+x)$
2. $f \prec_1 g$ implies $|f| \prec_1 g$.
3. Let $f \prec_k g$ and $\lambda \in \mathbb{R}$. Then $\lambda f \prec_k |\lambda|g$.
4. Let $f_i \prec_k g_i$, $i = 1 : q$. Then $\sum_i f_i \prec_k \sum_i g_i$.
5. Let $f_i \prec_k g_i$, $i = 1, \dots, q$. Then $\prod_i f_i \prec_k \prod_i g_i$.
6. Let $f \prec_k g$ and $f_i \prec_k g_i$, $i = 1, \dots, n$. Set $F = f(f_1, \dots, f_n)$ and $G = g(g_1, \dots, g_n)$. Then $F \prec_k G$.

Proof.

1. Obvious.
2. Note that $|\partial^\alpha |f(x)|| = |\partial^\alpha f(x)|$ for $|\alpha| = 1$.
3. Obvious.

4. Obvious.

5. Let $|\alpha| \leq k$. The repeated use of the product rule of differentiation yields that

$$\partial^\alpha \left(\prod_i f_i \right) = P_\alpha(\partial^\beta f_i)$$

is a certain polynomial in the terms $\partial^\beta f_i$ where $\beta : |\beta| \leq |\alpha|$ and $i = 1, \dots, q$. Obviously

$$\partial^\alpha \left(\prod_i g_i \right) = P_\alpha(\partial^\beta g_i)$$

uses the same polynomial, and since the coefficients of the polynomial P_α are nonnegative integers, we obtain by term-wise estimation

$$|P_\alpha(\partial^\beta f_i(x))| \leq P_\alpha(|\partial^\beta f_i(x)|) \leq P_\alpha(\partial^\beta g_i(|x|))$$

6. We argue in the same way as in the previous proof: Let $|\alpha| \leq k$. The repeated use of the chain rule of differentiation yields that

$$\partial^\alpha F = P_\alpha((\partial^\beta f)(f_1, \dots, f_n), \partial^\beta f_i)$$

is a certain polynomial in the terms $(\partial^\beta f)(f_1, \dots, f_n)$ and $\partial^\beta f_i$ where $\beta : |\beta| \leq |\alpha|$ and $i = 1, \dots, q$. Obviously

$$\partial^\alpha G = P_\alpha((\partial^\beta g)(g_1, \dots, g_n), \partial^\beta g_i)$$

uses the same polynomial, and since the coefficients of the polynomial P_α are nonnegative integers, we obtain by term-wise estimation

$$\begin{aligned} |\partial^\alpha F(x)| &= |P_\alpha((\partial^\beta f)(f_1(x), \dots, f_n(x)), \partial^\beta f_i(x))| \\ &\leq P_\alpha((\partial^\beta g)(g_1(|x|), \dots, g_n(|x|)), \partial^\beta g_i(|x|)) = \partial^\alpha G(|x|) \end{aligned}$$

□

Here are some examples of how the preceding rules could be applied:

Example 14.

1. $e^{|\sin(m+x)|} \prec_1 e^{|\sin m|+x}$

2. $\frac{1}{1+t} \prec\prec \frac{1}{1-t}$ for $|t| < 1$ and $\sin(x) \prec_3 x + \frac{1}{6}x^3$ implies
 $\frac{1}{1+\frac{1}{2}\sin(x)} \prec_3 \frac{1}{1-\frac{1}{2}(x+\frac{1}{6}x^3)}$ for $|x + \frac{1}{6}x^3| < 2$.

3. $\sin(x_1^2) \cos(x_2 - x_3) \prec_2 \left(x_1^2 + \frac{1}{2}(x_1^2)^2 \right) \left(1 + \frac{1}{2}(x_2 + x_3)^2 \right)$

3 Local Expansions to Obtain Exclusion Tests

The following theorem summarizes the possible choices of exclusion tests which we consider in this paper:

Theorem 15. *Let $\sigma \subset \mathbb{R}^n$ be an interval, and let $q > 0$ be an integer. Let $f(m_\sigma + x) \prec_q g(x)$ for $|x| \leq r_\sigma$. Then*

$$|f(m_\sigma)| \leq g(r_\sigma) - g(0) - \sum_{0 < |\alpha| < q} \underbrace{(\partial^\alpha g(0) - |\partial^\alpha f(m_\sigma)|)}_{\geq 0} r_\sigma^\alpha \quad (16)$$

is an exclusion test for f on σ .

Proof. Let $m_\sigma + h$ be a zero point of f in σ . We have to show that f satisfies the test. Using the Taylor formula (8) we obtain

$$g(r_\sigma) = g(0) + \sum_{0 < |\alpha| < q} \partial^\alpha g(0) r_\sigma^\alpha + \int_0^1 \sum_{|\beta|=q} \partial^\beta g(tr_\sigma) \omega_q(dt) r_\sigma^\beta$$

and consequently

$$\begin{aligned} |f(m_\sigma)| &= |f(m_\sigma + h) - f(m_\sigma)| \\ &\leq \sum_{0 < |\alpha| < q} |\partial^\alpha f(m_\sigma) h^\alpha| + \left| \int_0^1 \sum_{|\beta|=q} \partial^\beta f(m_\sigma + th) \omega_q(dt) h^\beta \right| \\ &\leq \sum_{0 < |\alpha| < q} |\partial^\alpha f(m_\sigma)| r_\sigma^\alpha + \int_0^1 \sum_{|\beta|=q} \partial^\beta g(|th|) \omega_q(dt) |h|^\beta \\ &\leq \sum_{0 < |\alpha| < q} |\partial^\alpha f(m_\sigma)| r_\sigma^\alpha + \int_0^1 \sum_{|\beta|=q} \partial^\beta g(tr_\sigma) \omega_q(dt) r_\sigma^\beta \\ &= \sum_{0 < |\alpha| < q} |\partial^\alpha f(m_\sigma)| r_\sigma^\alpha + g(r_\sigma) - g(0) - \sum_{0 < |\alpha| < q} \partial^\alpha g(0) r_\sigma^\alpha \end{aligned}$$

□

Corollary 17. *Let $\sigma \subset \mathbb{R}^n$ be an interval, and let $q > 0$ be an integer. Let $f(x) \prec_q g(x)$ for $x \in \sigma$. Then*

$$|f(m_\sigma)| \leq g(|m_\sigma| + r_\sigma) - g(|m_\sigma|) - \sum_{0 < |\alpha| < q} \underbrace{(\partial^\alpha g(|m_\sigma|) - |\partial^\alpha f(m_\sigma)|)}_{\geq 0} r_\sigma^\alpha \quad (18)$$

is an exclusion test for f on σ .

Proof. Note that $f(m_\sigma + x) \prec_q g(|m_\sigma| + x)$ for $|x| \leq r_\sigma$ and apply the theorem. □

The terms inside the summation sign in (16) and (18) are nonnegative, and hence the test tightens with increasing q . To increase the efficiency of implementations, one would successively apply the test for $q = 1 : q_0$ (given some q_0) and discard the interval as soon as the test fails.

Note also that for $q = 1$ the test (18) reduces to the one given in (6), however, instead of requiring $f \prec\prec g$, see [14], we only need to require $f \prec_1 g$ in this case.

Our approach also includes the use of local Lipschitz constants, compare also to (3):

Corollary 19 (Lipschitz Constants for f). *Let $\sigma \subset \mathbb{R}^n$ be an interval, and let $f \in \mathcal{A}_1(\sigma)$, and consider Lipschitz constants*

$$C_\alpha \geq \sup_{y \in \sigma} |\partial^\alpha f(y)| \text{ for } |\alpha| = 1.$$

Then

$$|f(m_\sigma)| \leq \sum_{|\alpha|=1} C_\alpha r_\sigma^\alpha$$

is an exclusion test for f on σ .

Proof. Define

$$g(x) := |f(m_\sigma)| + \sum_{|\alpha|=1} C_\alpha x^\alpha$$

and note that $f(m_\sigma + x) \prec_1 g(x)$ for $|x| \leq r_\sigma$. Now apply Theorem 15 with $q = 1$. \square

Corollary 20 (Lipschitz Constants for f'). *Let $\sigma \subset \mathbb{R}^n$ be an interval, and let $f \in \mathcal{A}_2(\sigma)$, and consider Lipschitz constants*

$$C_\beta \geq \sup_{y \in \sigma} |\partial^\beta f(y)| \text{ for } |\beta| = 2.$$

Then

$$|f(m_\sigma)| \leq \sum_{|\alpha|=1} |\partial^\alpha f(m_\sigma)| r_\sigma^\alpha + \sum_{|\beta|=2} C_\beta r_\sigma^\beta$$

is an exclusion test for f on σ .

Proof. Define

$$g(x) := |f(m_\sigma)| + \sum_{|\alpha|=1} |\partial^\alpha f(m_\sigma)| x^\alpha + \sum_{|\beta|=2} C_\beta x^\beta$$

and note that $f(m_\sigma + x) \prec_2 g(x)$ for $|x| \leq r_\sigma$. Now apply Theorem 15 with $q = 1$ or $q = 2$ (both lead to the same test). \square

4 Complexity Results

In this section we investigate the complexity of the exclusion algorithm in the sense of Theorem 7. In fact, we will strengthen the result and show that even degenerate zero points do not excessively increase the number of intervals generated by the algorithm, provided that a sufficiently tight test is used.

Throughout this section, let $\Lambda \subset \mathbb{R}^n$ be an initial interval, $q > 0$ an integer, $F : \Lambda \rightarrow \mathbb{R}^n$, and $F(x) \prec_q G(x)$ for $x \in \Lambda$. We start the exclusion algorithm in Λ using the exclusion test

$$\begin{aligned} |F(m_\sigma)| &\leq G(|m_\sigma| + r_\sigma) - G(|m_\sigma|) - \sum_{0 < |\alpha| < q} (\partial^\alpha G(|m_\sigma|) - |\partial^\alpha F(m_\sigma)|) r_\sigma^\alpha \quad (21) \\ &= \sum_{0 < |\alpha| < q} |\partial^\alpha F(m_\sigma)| r_\sigma^\alpha + \int_0^1 \sum_{|\beta|=q} \partial^\beta G(|m_\sigma| + tr_\sigma) \omega_k(dt) r_\sigma^\beta \end{aligned}$$

Recall that exclusion algorithm generates for each level $i > 0$ a list of intervals Γ_i . For the purpose of an asymptotic analysis, we assume that `maximal_level` = ∞ , i.e., we consider the algorithm to run without termination.

We will need the following technical definition:

Definition 22. We say that a zero point ξ of F has uniform order p if

1. $\partial^\alpha F(\xi) = 0$ for $|\alpha| < p$.
2. There exists an $\epsilon > 0$ such that $\epsilon \|m - \xi\|^p \leq \|F(m)\|$ for $\|m - \xi\| \leq \epsilon$.

We recall the following well-known result from analysis:

Remark 23. If ξ is a regular zero point, i.e., $F(\xi) = 0$ and $F'(\xi)$ is invertible, then ξ is a zero point of F of uniform order 1.

The following Lemma is the basis for our complexity analysis for the exclusion algorithm using the exclusion test (21).

Lemma 24. *Let each zero point of F be of some uniform order which is at most q . Then there exists a constant $A > 0$ such that the following holds: if $\sigma \in \Gamma_k$ with $k > A$, then there exists a zero point $\xi \in \Lambda$ of F such that $\|m_\sigma - \xi\| \leq A \|r_\sigma\|$.*

Proof. Assume not. Then the exclusion algorithm generates a sequence $\sigma_i \in \Gamma_i$ such that $\|m_{\sigma_i} - \eta\| > i \|r_{\sigma_i}\|$ for all zero points η of F . Since Λ is compact, we find a convergent subsequence of the m_{σ_i} , i.e., there is an unbounded set I of natural numbers such that

$$\lim_{i \in I} m_{\sigma_i} = \xi$$

for some $\xi \in \Lambda$. From the validity of the exclusion test (21) for the σ_i it follows that ξ is a zero point of F . By assumption we know that ξ has a certain uniform order, say p , with $p < q$. Hence there exists an $\epsilon > 0$ such that

$$\epsilon \|m_{\sigma_i} - \xi\|^p \leq \|F(m_{\sigma_i})\| \quad (25)$$

for all but finitely many $i \in I$. On the other hand, the exclusion test and Taylor's formula give

$$\begin{aligned} |F(m_{\sigma_i})| &\leq G(|m_{\sigma_i}| + r_{\sigma_i}) - G(|m_{\sigma_i}|) - \sum_{0 < |\alpha| < p} (\partial^\alpha G(|m_{\sigma_i}|) - |\partial^\alpha F(m_{\sigma_i})|) r_{\sigma_i}^\alpha \\ &= \sum_{0 < |\alpha| < p} |\partial^\alpha F(m_{\sigma_i})| r_{\sigma_i}^\alpha + \int_0^1 \sum_{|\beta|=p} \partial^\beta G(|m_{\sigma_i}| + tr_{\sigma_i}) \omega_p(dt) r_{\sigma_i}^\beta. \end{aligned} \quad (26)$$

(Recall that the test tightens with increasing p , so if it holds for $p = q$, it also holds for $p \leq q$). Expanding $\partial^\alpha F(m_{\sigma_i})$ about ξ and using the fact that all derivatives of order lower than p vanish, we obtain

$$\partial^\alpha F(m_{\sigma_i}) = \sum_{\gamma : |\gamma| + |\alpha| = p} \int_0^1 \partial^\gamma \partial^\alpha F(\xi + t(m_{\sigma_i} - \xi)) \omega_{p-|\alpha|}(dt) (m_{\sigma_i} - \xi)^\gamma$$

and hence

$$\|\partial^\alpha F(m_{\sigma_i})\| = \mathbf{O}(\|m_{\sigma_i} - \xi\|^{p-|\alpha|}).$$

Using this and the fact that $\|m_{\sigma_i} - \xi\| > i\|r_{\sigma_i}\| \geq \|r_{\sigma_i}\|$ for all but finitely many $i \in I$, the inequality (26) leads to

$$\|F(m_{\sigma_i})\| \leq M\|r_{\sigma_i}\| \|m_{\sigma_i} - \xi\|^{p-1} \quad (27)$$

for some $M > 0$ and all but finitely many $i \in I$. Taking both inequalities (27) and (25) now yields

$$\epsilon \|m_{\sigma_i} - \xi\| \leq M\|r_{\sigma_i}\|$$

which, for all but finitely many $i \in I$, contradicts $\|m_{\sigma_i} - \xi\| > i\|r_{\sigma_i}\|$. \square

The proof of the following theorem is now simple, but somewhat technical in its precise details.

Theorem 28. *Let each zero point ξ of F be of some uniform order which is at most q . Then $\#\Gamma_\ell$ is bounded as $\ell \rightarrow \infty$.*

Proof. Given the radius r_Λ of the initial interval Λ , let $\eta := \min_\nu r_\Lambda(\nu) > 0$ be its minimal entry. Let e denote the vector with all entries equal to one. Let A be the constant of the previous Lemma. We only need to consider bisection levels $\ell > A$. Note that $r_\sigma = 2^{-\ell} r_\Lambda$ for $\sigma \in \Gamma_\ell$.

Let $\sigma \in \Gamma_\ell$, and let $\xi \in \Lambda$ be a zero point of F such that $\|m_\sigma - \xi\| \leq A\|r_\sigma\|$. Note that we can write this inequality as

$$\xi - A\|r_\sigma\| e \leq m_\sigma \leq \xi + A\|r_\sigma\| e.$$

From $e \leq r_\Lambda/\eta$ it follows that

$$\begin{aligned} \xi - \frac{A\|r_\Lambda\|}{\eta} r_\sigma &= \xi - \frac{A\|r_\sigma\|}{\eta} r_\Lambda \leq \xi - A\|r_\sigma\| e \\ &\leq m_\sigma \leq \xi + A\|r_\sigma\| e \leq \xi + \frac{A\|r_\sigma\|}{\eta} r_\Lambda = \xi + \frac{A\|r_\Lambda\|}{\eta} r_\sigma. \end{aligned}$$

Hence, if L is an integer such that

$$L \geq \frac{A \|r_\Lambda\|}{\eta} + 1,$$

then σ is contained in the interval $\tau_\xi = [\xi - Lr_\sigma, \xi + Lr_\sigma]$. There are at most L^n intervals in Γ_ℓ that can be contained in τ_ξ .

Since all zero points of F are isolated by assumption, and since Λ is compact, F has a finite number, say C , of zero points, and hence $\#\Gamma_\ell \leq L^n C$. \square

Remark 29. Not all isolated zero points, even of an analytic map, satisfy Definition 22; in fact, orders of such zero points are defined in a different way. Modifications of the above proof for more general cases will be investigated elsewhere. However, we point out that numerical experiments show that the exclusion algorithm captures all isolated zero points without blow-up provided an exclusion test of sufficiently high order is applied. This remark is particularly important for polynomial systems where a maximal order test can be efficiently implemented, see the next section.

5 Special Case: Polynomial Systems

For polynomial systems it is natural to use the following simple dominance. Given a polynomial of degree r

$$p(x) = \sum_{|\alpha| \leq r} c_\alpha x^\alpha,$$

we define

$$\hat{p}(x) = \sum_{|\alpha| \leq r} |c_\alpha| x^\alpha,$$

and therefore have

$$p \prec_\infty \hat{p},$$

see also Definition 5. The exclusion test (18) now reads

$$|p(m_\sigma)| \leq \hat{p}(|m_\sigma| + r_\sigma) - \hat{p}(|m_\sigma|) - \sum_{0 < |\alpha| < q} \underbrace{(\partial^\alpha \hat{p}(|m_\sigma|) - |\partial^\alpha p(m_\sigma)|)}_{\geq 0} r_\sigma^\alpha \quad (30)$$

for any $q > 0$.

A numerically important observation is that under certain conditions the terms in the above sum are zero. More precisely:

Definition 31. We call a polynomial p monotone iff all non-zero coefficients of p have the same sign.

The following two lemmas are rather obvious:

Lemma 32. A polynomial p is monotone iff $\hat{p}(|m|) = |p(m)|$ for all $m \geq 0$.

Lemma 33. *If p is monotone, then $\partial^\beta p$ is monotone for all β .*

The case when our initial interval Λ is in the positive cone is an important one. Often for systems with physical significance, variables only take on positive values. Then the preceding observations enable us to identify the multi-indices α , for which the summation in (30) needs to be carried out. The following recursion generates these multi-indices in an efficient way:

```

function GenerateMultiIndices( $\alpha$ )
set  $n = |\alpha|$ 
if  $\partial^\alpha(p)$  is monotone
    return
print( $\alpha$ )
set  $\beta = \alpha$ 
set  $\beta_1 = \beta_1 + 1$ 
GenerateMultiIndices( $\beta$ )
for  $k = 1 : n - 1$ 
    if  $\alpha_k \neq 0$ 
        return
    set  $\beta = \alpha$ 
    set  $\beta_{k+1} = \beta_{k+1} + 1$ 
    GenerateMultiIndices( $\beta$ )

```

The recursion is started with $\alpha = (0, \dots, 0)$.

On the other hand, for $q = \infty$ in (30), we obtain a simplification:

$$\begin{aligned}
 |p(m_\sigma)| &\leq \hat{p}(|m_\sigma| + r_\sigma) - \hat{p}(|m_\sigma|) - \sum_{0 < |\alpha|} (\partial^\alpha \hat{p}(|m_\sigma|) - |\partial^\alpha p(m_\sigma)|) r_\sigma^\alpha \\
 &= \sum_{0 < |\alpha|} |\partial^\alpha p(m_\sigma)| r_\sigma^\alpha.
 \end{aligned} \tag{34}$$

This test is valid for all m_σ , not just $m_\sigma \geq 0$. All relevant multi-indices can be obtained in a recursion similar to the above. The line “if $\partial^\alpha(p)$ is monotone” only needs to be replaced by “if $\partial^\alpha(p) = 0$ ”.

With these remarks, it is now clear that the exclusion algorithm applied to polynomial systems with the polynomial exclusion tests (30) or (34) can be implemented as a *black box* algorithm: the only input required are the coefficients of the polynomials and an initial interval. A preliminary implementation in JAVA was very successful, and its improvements and extensions are a current project, see [3].

6 Numerical Examples

An application of the exclusion algorithm typically consists of three steps:

1. Given the problem $F(x) = 0$, construct a G such that $F \prec_q G$. The results in Section 2 are used in this step.

2. Implement the exclusion test (16) or (18) for the given q . Note that for $q > 1$ many partial derivatives are involved, so we have constructed a MAPLE script that actually writes these tests once F and G are given.
3. Run the exclusion algorithm based on the test constructed in step 2.
4. A typical feature of the exclusion algorithm is that each zero point causes the generation of several intervals, and therefore in a final step we have to sort out which intervals represent the same zero point. We call two intervals generated on the final k -th bisection level close iff their midpoints m_1 and m_2 satisfy an inequality $|m_1 - m_2| \leq C2^{-k}r$ where r is the radius of the initial interval. Ideally, $C = 2$, however a more practical choice is some constant $C > 2$. This notion of closeness defines *connected components* among the intervals generated on the k -th level. Lemma 24 implies the existence of a $C \geq 2$ such that for sufficiently large k each zero point is represented by exactly one connected component of intervals. We say that the algorithm has isolated all zero points (for such k). It is not difficult to write a program that generates such connected components.

Note that for polynomial systems, items 1 and 2 can be automated and incorporated directly into the exclusion algorithm as indicated in Section 5.

6.1 Example

We present a simple one-dimensional polynomial equation $p(x) = 0$ which illustrates Theorem 28:

$$p(x) = (x - 3)^4(x + 2) .$$

We use the dominance $p \prec_\infty \hat{p}$ as described in Section 5. The initial interval was $[-10, 10]$. We show the performance of the exclusion test (30) for $q = 1$ and $q = \infty$ (in fact, $q = 6$ is all that is used here, see also (34)):

Level	0	1	2	3	4	5	6	7	8	9	10
# of intervals ($q = 1$)	1	2	4	8	12	21	32	48	76	122	199
# of intervals ($q = \infty$)	1	2	4	7	7	7	6	6	6	6	6

Here ‘Level’ indicates the bisection level, and in rows $q = 1$ and $q = \infty$ we show the number of intervals generated on the corresponding bisection level. As can be seen, the simple test for $q = 1$ is not capable of containing the number of intervals generated on each level, due to the singularity of the zero point $x = 3$.

6.2 Example

The following four-dimensional fixed-point problem $x = G(x)$ is taken from [16]:

$$G(x) = \begin{pmatrix} x_1 + C_1(x_3 - \alpha \sin(x_1) \cos(x_2)) \\ x_2 + C_2(x_4 - \alpha \cos(x_1) \sin(x_2)) \\ D_1(x_3 - \alpha \sin(x_1) \cos(x_2)) \\ D_2(x_4 - \alpha \cos(x_1) \sin(x_2)) \end{pmatrix}$$

where

$$\begin{aligned} C_1 &= \frac{1-e^{-2\mu_1}}{2\mu_1} & \mu_1 &= 0.1\pi & \mu_2 &= 0.2\pi \\ C_2 &= \frac{1-e^{-2\mu_2}}{2\mu_2} & D_1 &= e^{-2\mu_1} & D_2 &= e^{-2\mu_2} & \alpha &= 5 \end{aligned}$$

By replacing all minus signs in G with plus signs, $\sin(x_i)$ with x_i , and $\cos(x_i)$ with $1 + x_1$, we obtain a function \hat{G} such that

$$x - G(x) \prec_1 x + \hat{G}(x)$$

Now we can use the exclusion test

$$|m - G(m)| \leq r + \hat{G}(|m| + r) - \hat{G}(r)$$

An exclusion algorithm based on this test generated too many intervals and was not successful. Also the tests proposed in [14, 15] were unsuccessful. However the following test was successful:

We replace all minus signs in G with plus signs, $\sin(x_i)$ with $x_i + x_i^3/6$, and $\cos(x_i)$ with $1 + x_1^2/2 + x_i^3/6$, and thus obtain a function \tilde{G} such that

$$x - G(x) \prec_3 x + \tilde{G}(x) .$$

Now we can use the exclusion test (18) with $q = 3$. With the initial interval $[-\pi, \pi]^2 \times [-1.5, 1.5]^2$ we obtain the following number of intervals on each bisection level:

Level	0	1	2	3	4	5	6	7	8	9	10
# of intervals ($q = 3$)	1	16	256	2688	1180	328	160	96	192	220	228

In this way all 13 solutions were isolated. Note that the above performance of the exclusion algorithm displays its typical feature: First the number of generated intervals increases, and then decreases. When this number stabilizes, the algorithm can typically be stopped since the solutions have been sufficiently localized and a local solver (e.g., Newton's method) now could take over for more precise approximations. In our example, the localization of the 13 solutions was finished at bisection level 7.

6.3 Example

The following two-dimensional example $F(x) = 0$ is from [16, 17] and was calculated with a global Lipschitz test (3) in [15], however the test (6) from [14] fails since the estimates lead to very dramatic overestimations.

We obtain a more efficient result with local estimates in the sense of Corollaries 19–20.

For

$$F(x) = \begin{pmatrix} \frac{1}{2} \sin(x_1 x_2) - \frac{x_2}{4\pi} - \frac{x_1}{2} \\ (1 - \frac{1}{4\pi})(e^{2x_1} - e) + \frac{e x_2}{\pi} - 2 e x_1 \end{pmatrix}$$

let G and H be defined by

$$G(x) = \begin{pmatrix} \frac{1}{2} x_1 x_2 + \frac{x_2}{4\pi} + \frac{x_1}{2} \\ (1 - \frac{1}{4\pi})((1 + e^{2(m_1+r_1)}(2x_1)) + e) + \frac{e x_2}{\pi} + 2 e x_1 \end{pmatrix}$$

$$H_1(x) = \frac{1}{2}((x_1x_2) + \frac{(x_1x_2)^3}{6} + \frac{(x_1x_2)^5}{120} + \frac{x_2}{4\pi} + \frac{x_1}{2}$$

$$H_2(x) = (1 - \frac{1}{4\pi})(((2x_1) + \frac{(2x_1)^2}{2} + \frac{(2x_1)^3}{6} + \frac{(2x_1)^4}{24} + e^{(2(m_1+r_1))} \frac{(2x_1)^5}{120}) + e) + \frac{e x_2}{\pi} + 2 e x_1$$

Then we have

$$F(m+x) \prec_1 G(m+x) \text{ and } F(m+x) \prec_5 H(m+x) \text{ for } |x| \leq r.$$

Using an initial interval

$$[-1, 2] \times [-20, 5]$$

we easily find all 12 solutions. Here are the numbers of intervals generated on each bisection level:

Level	0	1	2	3	4	5	6	7	8	9	10
# of intervals ($q = 1$)	1	4	11	28	38	62	78	76	84	78	80
# of intervals ($q = 5$)	1	3	9	20	26	34	30	26	26	25	23

6.4 Example

The four-dimensional polynomial system $f(x) = 0$ investigated in this example comes from a planar four-bar design problem, see [9]. The equations were taken from Verschelde's web page, see [13]. Verschelde reports 36 complex solutions, but only three are real. They are contained in the interval $[0, 2]^4$ which we take as our initial interval. One zero point is $(0, 0, 0, 0)^T$, and the other two real solutions are close to each other. We used a polynomial exclusion test with $q = \infty$ as described in Section 5 and approximated all three (real) solutions. The following numbers of intervals were generated on the indicated bisection levels.

Level	0	1	2	3	4	5	6	7	8	9	10
# of intervals	1	16	235	994	2091	2348	1423	546	390	343	308

6.5 Example

The following three-dimensional polynomial system $f(x) = 0$ has been represented as a general economic equilibrium model in [11]. The functions are again taken from Verschelde's web page, see [13].

$$f_1(x) = x_2^4 - 20/7 x_1^2$$

$$f_2(x) = x_1^2 x_3^4 + 7/10 x_1 x_3^4 + 7/48 x_3^4 - 50/27 x_1^2 - 35/27 x_1 - 49/216$$

$$\begin{aligned}
f_3(x) = & 3/5 x_1^6 x_2^2 x_3 + x_1^5 x_2^3 + 3/7 x_1^5 x_2^2 x_3 + 7/5 x_1^4 x_2^3 - 7/20 x_1^4 x_2 x_3^2 \\
& - 3/20 x_1^4 x_3^3 + 609/1000 x_1^3 x_2^3 + 63/200 x_1^3 x_2^2 x_3 - 77/125 x_1^3 x_2 x_3^2 \\
& - 21/50 x_1^3 x_3^3 + 49/1250 x_1^2 x_2^3 + 147/2000 x_1^2 x_2^2 x_3 \\
& - 23863/60000 x_1^2 x_2 x_3^2 - 91/400 x_1^2 x_3^3 - 27391/800000 x_1 x_2^3 \\
& + 4137/800000 x_1 x_2^2 x_3 - 1078/9375 x_1 x_2 x_3^2 - 5887/200000 x_1 x_3^3 \\
& - 1029/160000 x_2^3 - 24353/1920000 x_2 x_3^2 - 343/128000 x_3^3
\end{aligned}$$

Verschelde reports 136 complex solutions, however only 14 are real. They are contained in the interval $[-2, 2]^3$ which we take as an initial interval. It should be noted that three of the real solutions are singular, so the methods reported in [14, 15] would certainly fail on this example. We again used a polynomial exclusion test with $q = \infty$ and approximated all 14 (real) solutions. Here are the number of intervals generated on each bisection level:

Level	0	1	2	3	4	5	6	7	8	9	10
# of intervals	1	8	48	240	490	238	126	94	76	72	60

References

- [1] E. L. Allgower and K. Georg. Numerical path following. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis*, volume 5, pages 3–207. North-Holland, 1997.
- [2] T. Csendes and D. Ratz. Subdivision direction selection in interval methods for global optimization. *SIAM J. Numer. Anal.*, 34:922–938, 1997.
- [3] M. Erdmann. *On the Implementation and Analysis of Cellular Exclusion Algorithms*. PhD thesis, Colorado State University, 2001. In preparation.
- [4] R. B. Kearfott. Abstract generalized bisection and a costbound. *Math. Comput.*, 49:187–202, 1987.
- [5] R. B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [6] R. B. Kearfott. Empirical evaluation of innovations in interval branch and bound algorithms for nonlinear algebraic systems. *SIAM Journal on Scientific Computing*, 18(2):574–594, 1997.
- [7] R. B. Kearfott and K. Du. The cluster problem in multivariate global optimization. *Journal of Global Optimization*, 5:253–265, 1994.
- [8] R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM, 1979.
- [9] A. P. Morgan and C. W. Wampler. Solving a planar four-bar design problem using continuation. *J. Mech. Design*, 112:544–550, 1990.

- [10] H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization*. Wiley, 1988.
- [11] J. B. Shoven. Applied general equilibrium modelling. *IMF Staff Papers*, pages 394–419, 1983.
- [12] A. J. Sommese, J. Verschelde, and Ch. W. Wampler. Numerical decomposition of the solution sets of polynomial systems into irreducible components. *SIAM J. Numer. Anal.*, to appear, 2000.
- [13] J. Verschelde. Algorithm 795: Phcpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software*, 25:251–276, 1999. A collection of polynomial systems analyzed with this algorithm is posted on the net:
<http://www.math.uic.edu/~jan/demo.html>.
- [14] Z.-B. Xu, J.-S. Zhang, and Y.-W. Leung. A general CDC formulation for specializing the cell exclusion algorithms of finding all zeros of vector functions. *Appl. Math. Comput.*, 86:235–259, 1997.
- [15] Z.-B. Xu, J.-S. Zhang, and W. Wang. A cell exclusion algorithm for determining all the solutions of a nonlinear system of equations. *Appl. Math. Comput.*, 80:181–208, 1996.
- [16] P. J. Zufiria and R. S. Guttalu. A computational method for finding all roots of a vector function. *Applied Mathematics and Computation*, 35(1):13–59, 1990.
- [17] P. J. Zufiria and R. S. Guttalu. On an application of dynamical systems theory to determining all zeros of a vector function. *J. Math. Anal. Appl.*, 152:269–295, 1990.