Complete Subdivision Algorithms, II: Isotopic Meshing of Singular Algebraic Curves

Michael Burr

Courant Institute, NYU, 251 Mercer Street, NY, NY 10012.

Sung Woo Choi

Department of Mathematics, Duksung Women's University, Seoul 132-714, Korea.

Ben Galehouse

Courant Institute, NYU, 251 Mercer Street, NY, NY 10012.

Chee K. Yap

Courant Institute, NYU, 251 Mercer Street, NY, NY 10012
and
Korea Institute of Advanced Study, Seoul, Korea.

Abstract

Given a real valued function $f(X, Y)$, a box region $B \subseteq \mathbb{R}^2$ and $\varepsilon > 0$, we want to compute an ε-isotopic polygonal approximation to the restriction of the curve $S = f^{-1}(0) = \{p \in \mathbb{R}^2 : f(p) = 0\}$ to B. We focus on subdivision algorithms because of their adaptive complexity and ease of implementation. Plantinga & Vegter (2004) gave a numerical subdivision algorithm that is exact when the curve S is non-singular. They used a computational model that relies only on function evaluation and interval arithmetic. We generalize their algorithm to any bounded (but possibly non-simply connected) region B that does not contain singularities of S. With this generalization as subroutine, we provide a method to detect isolated algebraic singularities and their branching degree. This appears to be the first complete *purely numerical* method to compute isotopic approximations of algebraic curves with isolated singularities.

Key words: Meshing, Singularity, Root bound, Evaluation bound, Implicit algebraic curve, Complete numerical algorithm, Subdivision algorithm.

1. Introduction

Given $\varepsilon > 0$, a box region $B \subseteq \mathbb{R}^2$ and a real valued function $f : \mathbb{R}^2 \to \mathbb{R}$, we want to compute a polygonal approximation P to the restriction of the implicit curve $S = f^{-1}(0)$ $(= \{p \in \mathbb{R}^2 : f(p) = 0\})$ to B. The approximation P must be (1) "topologically correct" and (2) "ε-close" to $S \cap B$. We use the standard interpretation of requirement (2), that $d(P, C \cap B) \leq \varepsilon$ where $d(\cdot, \cdot)$ is the Hausdorff distance on compact sets. In recent years, it has become accepted [Boissonnat et al., 2006] to interpret requirement (1) to mean that P is isotopic to $S \cap B$, which we denote by $P \approx C \cap B$. This means we not only require that P and $S \cap B$ to be homeomorphic, but they must be embedded in $\mathbb{R}^2$ "in the same way". E.g., if $S \cap B$ consists of two disjoint ovals, these can be embedded in $\mathbb{R}^2$ as two ovals exterior to each other, or as two nested ovals. Isotopy, but not homeomorphism, requires P to respect this distinction. In this paper, we mainly focus on topological correctness since achieving ε-closeness is not an issue for our particular subdivision approach (but cf. [Boissonnat et al., 2006, pp. 213-4])

We may call the preceding problem the **2-D implicit meshing problem**. The term "meshing" comes from the corresponding problem in 3-D: given $\varepsilon > 0$ and an implicit surface $S : f(X, Y, Z) = 0$, we want to construct a triangular mesh M such that $d(M, S) \leq \varepsilon$ and $M \approx S$. It is interesting (see [Burr et al., 2009]) to identify the 1-D meshing with the well-known problem of real root isolation and refinement for a real function $f(X)$.

Among the approaches to most computational problems on curves and surfaces, the **algebraic approach** and **numerical approach** constitute two extremes of a spectrum. Algebraic methods can clearly solve most problems in this area, e.g., by an application of the general theory of cylindrical algebraic decomposition (CAD) [Basu et al., 2003]. Purely algebraic methods are generally not considered practical, even in the plane (e.g., [Hong, 1996; Seidel and Wolpert, 2005]), but efficient solutions have been achieved for special cases such as intersecting quadrics in 3-D [Schoemer and Wolpert, 2006]. At the other end of the spectrum, the numerical approaches emphasize approximation and iteration. An important class of such algorithms is the class of **subdivision algorithms** which can be viewed as a generalization of binary search. Such algorithms are practical in two senses: they are easy to implement and their complexity is more adaptive [Yap, 2006]. Another key feature of subdivision algorithms is "locality", meaning that we can restrict our computation to some region of interest.

Besides the algebraic/numeric approaches, there is another approach that might be called the **geometric approach** in which we postulate an abstract computational model with certain (geometric) primitives (e.g., shoot an ray, decide if a point is in a cell). When implementing these geometric algorithms, one must still choose an algebraic or numeric implementation of these primitives. Implementations might also be a hybrid of algebraic and numeric techniques.

Unfortunately, numerical methods seldom have global correctness guarantees. The most famous example is the Marching Cube algorithm (1987) of Lorensen & Cline. Many authors have tried to improve the correctness of subdivision algorithms (e.g., [Stander and Hart, 1997]). So far, such efforts have succeeded under one of the following situations:

⋆ Yap's work is supported by NSF Grant CCF-0728977. Burr's work is partially supported by NSF Grant DMS-0701578. Galehouse's work is partially supported by the DoE under contract DE-FG02-00ER25053.

Email addresses: {burr@cims, galehouse@cims, yap@cs}.nyu.edu (Michael Burr), swchoi@duksung.ac.kr (Sung Woo Choi).

- (A0) Requiring niceness assumptions such as being non-singular or Morse.
- (A1) Invoking algebraic techniques such as resultant computations or manipulations of algebraic numbers.

It is clear that (A0) should be avoided. Generally, we call a method "complete" if the method is correct without any (A0) type restrictions. But many incomplete algorithms (e.g., Marching cube) are quite useful in practice. We want to avoid (A1) conditions because algebraic manipulations are harder to implement and such techniques are relatively expensive and non-adaptive [Yap, 2006]. The complete removal of (A0) type restrictions is the major open problem faced by purely numerical approaches to meshing. Thus, [Boissonnat et al., 2006, p. 187] states that "*meshing in the vicinity of singularities is a difficult open problem and an active area of research*". Most of the techniques described in their survey are unable to handle singularities. It should be evident that this open problem has an implicit requirement to avoid the use of (A1) techniques.

For instance, the subdivision meshing algorithm of [Plantinga and Vegter, 2004; Plantinga, 2006] requires an (A0) assumption, the non-singularity of surfaces. The subdivision algorithm of [Seidel and Wolpert, 2005] requires [1] an (A1) technique, namely, the computation of resultants. We thus classify [Seidel and Wolpert, 2005] as a hybrid approach that combines numerical with algebraic techniques. Prior to our work, we are not aware of any meshing algorithm that can handle singularities without resorting to resultant computations. In general, hybrid methods offer considerable promise (e.g., [Hong, 1996]). This is part of a growing trend to employ numerical techniques for speeding up algebraic computations.

Some of our recent work addresses the above (A0)/(A1) concerns: in [Yap, 2006], we gave a complete numerical approach for determining tangential Bezier curve intersections; in [Cheng et al., 2008], we numerically solve zero-dimensional triangular systems without any "regularity" requirements on the systems; in [Burr et al., 2009], we provide numerical root isolation in the presence of multiple zeros; and [Burr et al., 2008] provides one of the first non-probabilistic adaptive analysis of an evaluation-based real root isolation algorithm. These last two papers address the 1-D analogue of the Plantinga & Vegter Algorithm. The philosophy behind all these papers is the design and analysis of complete numerical methods based on approximations, iteration and adaptive methods. Topological exactness is achieved using suitable algebraic bounds, ranging from classic root separation bounds to evaluation bounds and geometric separation bounds. We stress that the worst-case complexity of adaptive algorithms ought not to be the chief criterion for evaluating the usefulness of these algorithms: for the majority of inputs, these algorithms terminate fast. Zero bounds are only used as stopping criteria for iteration in the algorithms, and simple estimates can be computed easily. Computing such bounds does not mean that we compute resultants, even though their justification depend on resultant theory. The present paper continues this line of investigation.

The recent collection [Boissonnat et al., 2006, Chapter 5] reviews the current algorithmic literature in meshing in 2- and 3-D: the subdivision approach is represented by the Plantinga & Vegter algorithm as well as by Snyder's earlier approach based on parametrizability [Snyder, 1992b,a]. The subdivision algorithm of Plantinga & Vegter is remarkable in the following sense: even though it is globally isotopic, it does not guarantee

[1] Their paper is subtitled "Exploiting a little more Geometry and a little less Algebra" which speaks to our concerns with (A1).

isotopy of the curve within each cell of the subdivision. In contrast, Snyder's subdivision approach [Snyder, 1992b,a] requires the correct isotopy type in each cell. Indeed, because of this, the algorithm is incomplete [Boissonnat et al., 2006, p. 195].

Among geometric approaches to meshing, we have the point sampling approach as represented by [Boissonnat and Oudot, 2006; Cheng et al., 2004] or the Morse Theory approach as represented by [Stander and Hart, 1997; Boissonnat et al., 2004], or the sweepline approach [Mourrain and Técourt, 2005]. Note that the sweepline approach corresponds naturally to the algebraic operation of projection. So its implementation is often purely algebraic. The basic idea of sampling approaches is to reduce meshing of a surface S to computing the Delaunay triangulation of a sufficiently dense set of sample points on S [Boissonnat et al., 2006, p. 201–213]. To obtain such sample points, [Cheng et al., 2004] needs a primitive operation that amounts to solving a system of equations involving f and its derivatives. [Boissonnat and Oudot, 2006] needs a primitive for intersecting the surface with a Voronoi edge. These sample points are algebraic, and so implementing the primitives exactly would require strong algebraic techniques. But exact implementation does not seem to be justified for these applications, and so we are faced with an implementation gap that shows up the well-known non-robustness issues. For restrictions and open problems in sampling approaches, see [Boissonnat et al., 2006, p. 227–229]. In contrast, the computational primitives needed by the subdivision approach works directly with bigfloats, with modest requirements on f.

The present paper presents a purely numerical subdivision method for meshing algebraic curves with isolated singularities. In a certain sense, this is the most general geometric situation since reduced algebraic curves have only isolated singularities by Proposition 1. Our starting point is the algorithm of Plantinga & Vegter [Plantinga and Vegter, 2004; Plantinga, 2006] for implicit meshing of curves. It is important to understand the computational model of Plantinga & Vegter which is also used in this paper. Two capabilities are assumed with regards to $f(X, Y)$:

- (i) Sign evaluation of $f(p)$ at dyadic points p.
- (ii) f is C^1 and we can evaluate the interval analogues of $f, \frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}$ on dyadic intervals.

Note that the Marching Cube algorithm only requires capability (i). Let the **class** PV denote the set of all real functions $f : \mathbb{R}^2 \to \mathbb{R}$ for which capabilities (i) and (ii) are available. Many common functions of analysis (e.g., hypergeometric functions [Du and Yap, 2006]) belong to PV. So the approach of Plantinga & Vegter admits a more general setting than algebraic curves.

1.1. Overview of Paper

- Section 2 establishes some basic terminology and recall facts about the singularities of algebraic sets.
- In Sections 3 and 4, we extend the Plantinga & Vegter algorithm to compute an isotopic approximation of the curve $S = f^{-1}(0)$ restricted to a "nice region" that need not be simply connected. S may have singularities outside R and we only need $f \in PV$.
- In Section 5, we provide the algebraic evaluation bounds necessary for meshing singular curves.
- In Section 6, we provide a subdivision method to isolate all the singularities of a square-free integer polynomial $f(X, Y)$.
- In Section 7, given a box B containing an isolated singularity p, we provide a method to compute the branching degree of p.

- In Section 8, we finally present the overall algorithm to compute the isotopic polygonal approximation.
- We conclude in Section 9.

2. Basic Terminology and Algebraic Facts

Let $\mathbb{F} := \mathbb{Z}\left[\frac{1}{2}\right] = \{m2^n : m, n \in \mathbb{Z}\}$ be the set of **dyadic numbers**. All of our numerical computations are performed in a straightforward manner using $\mathbb{F}$. There are many well-known implementations of arithmetic on such numbers, which are then known as **bigfloats**. In short, our computational model is not based on some abstract capability whose implementation may reveal gaps that lead to well-known nonrobustness issues.

For $S \subseteq \mathbb{R}$, let $\square S$ be the set of closed intervals $[a, b]$ with endpoints in S, $\{a, b\} \subseteq S$. We write $\square S^n$ for $(\square S)^n$. In particular, $\square\mathbb{F}$ is the set of dyadic intervals, and $\square\mathbb{R}^n$ is the set of n-boxes. The boundary of a set $S \subseteq \mathbb{R}$ is denoted ∂S. If $f : \mathbb{R}^n \to \mathbb{R}$, and $S \subseteq \mathbb{R}$, then $f(S) := \{f(x) : x \in S\}$. A function $\square f : \square\mathbb{F}^n \to \square\mathbb{F}$ is a **box function** for f provided (i) $f(B) \subseteq \square f(B)$ and (ii) if $B_0 \supseteq B_1 \supseteq \cdots$ and $\lim_i B_i = p$ then $\lim_i \square f(B_i) = f(p)$. We regard the limit of intervals in terms of the limits of their endpoints. We say $f \in PV$ if $f \in C^1$ (has continuous first derivatives), there is an algorithm to determine $\mathtt{sign}(f(p))$ for $p \in \mathbb{F}^n$, and $\square f$ and its derivatives are computable. In this paper, we only consider box functions specialized to the two dimensional case.

Let $w(B)$ denote the **width** of box B, defined to be the minimum length of one of its sides. We only consider boxes of the form $B = I \times J$ where I, J are dyadic intervals, and, for simplicity, we assume B is square although it is possible to extend our algorithms to boxes with aspect ratio at most 2. The boundary ∂B of B is divided into four **sides** and four **corners**. The sides (north, south, etc) and corners (north-west, etc) may be named after the compass directions. Note: this 'sides/corners' terminology for boxes should not be confused with the 'edges/vertices' terminology which we reserve for the straightline graph $G = (V, E)$ which is our curve approximation. Two boxes are **neighbors** if one box has an side that overlaps an side of the other box (the two boxes may have different widths). We **split** a box B by subdividing it into 4 subboxes of equal widths. These subboxes are the **children** of B, with width $\frac{1}{2}w(B)$. Starting with B_0, the child-parent relationships obtained by an arbitrary sequence of splits yields a **quadtree** rooted at B_0. The set of leaves in such a quadtree constitute a partition of B_0.

Algebraic Facts. Let $\mathbb{D}$ be a unique factorization domain (UFD) and $f, g \in \mathbb{D}[\mathbf{X}] = \mathbb{D}[X_1, \ldots, X_n]$ where $\mathbf{X} = (X_1, \ldots, X_n)$. We say f, g are **similar** if there exist $a, b \in \mathbb{D} \setminus \{0\}$ such that $af = bg$, and we write this relationship as $f \sim g$. Otherwise, f and g are **dissimilar**. The **square-free part** of f is defined as

$$\mathtt{SqFree}(f) := \frac{f}{\mathtt{GCD}(f, \partial_1 f, \ldots, \partial_n f)} \tag{1}$$

where $\partial_{X_i} = \partial_i$ indicates differentiation with respect to X_i. f is said to be **square-free** if $\mathtt{SqFree}(f) = f$. From (1) we see that computing $\mathtt{SqFree}(f)$ from f involves only rational operations of $\mathbb{D}$. As the gradient of f is $\nabla f = (\partial_1 f, \ldots, \partial_n f)$, we may also write $\mathtt{GCD}(f, \nabla f)$ for $\mathtt{GCD}(f, \partial_1 f, \ldots, \partial_n f)$. See [Yap, 2000, Chap. 2] for standard conventions concerning $\mathtt{GCD}$.

Let k be an algebraically closed field. For $S \subseteq k[\mathbf{X}] = k[X_1, \ldots, X_n]$, let $\mathtt{Zero}(S) := \{p \in k^n : f(p) = 0 \text{ for all } f \in S\}$ denote the **zero set** of S. A zero set is also known as a **variety**. The **singular points** of $\mathtt{Zero}(f)$ are defined to be the points where $\nabla\mathtt{SqFree}(f) = 0$.

In 1-dimension, it is well-known that a square-free polynomial $f \in \mathbb{Z}[X]$ has no singularities (i.e., multiple zeros). We now recall two generalizations of this result that will be necessary in the remainder of the paper. See [Hartshorne, 1977; Cox et al., 1992] for similar results.

PROPOSITION 1 ([Harris, 1992, Ex.14.3]). The singular points of any variety form a proper subvariety.

This proposition is critical in our paper, because it implies that if $f \in \mathbb{R}[X, Y]$ is square-free, then the singular points are a proper subvariety of a union of curves and hence must be a finite set of points. Thus, we only need to handle isolated singularities.

PROPOSITION 2 (Algebraic Sard Lemma [Harris, 1992, Prop.14.4]). Let $f : X \to Y$ be any surjective regular map of varieties defined over a field k of characteristic 0. Then there exists a nonempty open subset $U \subseteq Y$ such that for any smooth point $p \in f^{-1}(U) \cap X_{sm}$ in the inverse image of U, the differential df_p is surjective.

Here, X_{sm} denotes the set of smooth points of variety X. The open sets refer to the Zariski topology, and maps are birational maps. The condition that the differential df_p is surjective is equivalent to insisting that the Jacobian of f has the same rank as the dimension of Y. The most important example that we consider is $f : \mathbb{R}^2 \to \mathbb{R}$ where f is a polynomial. In this case, since the image is one dimensional, the condition that df_p is surjective reduces to the condition that $\nabla f(p) \neq 0$. Every point in $\mathbb{R}^2 = X$ is smooth and $\mathbb{R} \setminus U$ is only a finite set. Hence, there are only a finite number of level sets, parametrized by h, where $\texttt{Zero}(f(X, Y) - h)$ has a singular point.

3. Algorithm of Plantinga & Vegter

First, we recall the Plantinga & Vegter algorithm. Given $\varepsilon > 0$, $f : \mathbb{R}^2 \to \mathbb{R}$ and a bounding box $B_0 \in \square\mathbb{F}^2$ to the curve $f^{-1}(0)$ and we want to compute a polygonal ε-approximation P of the restriction of the curve $S = f^{-1}(0)$ to B_0, i.e., $d(P, C \cap B_0) \leq \varepsilon$ and $P \approx C \cap B_0$. For simplicity, we focus on topological correctness: $P \approx C \cap B_0$, since it is easy to refine the subdivision to achieve $d(P, C \cap B_0) \leq \varepsilon$. The Plantinga & Vegter algorithm is based on two simple predicates on boxes B:

- Predicate $C_0(B)$ holds if $0 \notin \square f(B)$.
- Predicate $C_1(B)$ holds if $0 \notin (\square \frac{\partial f}{\partial X}(B))^2 + (\square \frac{\partial f}{\partial Y}(B))^2$.

These predicates are easily implemented, using interval arithmetic, for $f \in PV$. Moreover, if $C_0(B)$ holds, then the curve S does not intersect B. Note that if B satisfies C_1, then, by recalling the parent-child relationship, any child of B also satisfies C_1.

The input box B_0 is a dyadic square and the output will be an undirected graph $G = (V, E)$ where each vertex $v \in V$ is a dyadic point, $v \in \mathbb{F}^2$. In fact, G represents a straightline planar graph that is a polygonal approximation of $S \cap B_0$.

The algorithm has 3 phases, where Phase i ($i = 1, 2, 3$) is associated with a queue Q_i containing boxes. Initially, $Q_1 = \{B_0\}$, and $Q_2 = Q_3 = \emptyset$. When Q_i is empty, we proceed to the Phase $i + 1$.

- PHASE 1: SUBDIVISION. While Q_1 is non-empty, remove some B from Q_1, and perform the following: If $C_0(B)$ holds, B is discarded. If $C_1(B)$ holds, insert B into Q_2. Otherwise, split B into four subboxes and insert them back into Q_1.

- PHASE 2: BALANCING. This phase "balances" the subdivision, where a subdivision is **balanced** if the width of any two neighboring boxes differ by at most a factor of 2. Queue Q_2 is a min-priority queue, where the width of a box serves as its priority. While Q_2 is non-empty, remove the min-element B from Q_2, and perform the following: For each B-neighbor B' with width more than twice the width of B, remove B' from Q_2 and split B'. Insert each child B'' of B' into Q_2 provided $C_0(B'')$ does not hold. B'' might be a new neighbor of B and B'' might be split subsequently. When, finally, every neighbor of B is at most twice the width of B, we insert B into Q_3.
- PHASE 3: CONSTRUCTION. This phase constructs the graph $G = (V, E)$. Initially, the boxes in Q_3 are unmarked. While Q_3 is non-empty, remove any B from Q_3 and mark it. Now construct a set $V(B)$ of vertices. For each B-neighbor B', if B' is unmarked, evaluate the sign of $f(p)f(q)$ where p, q are endpoints of the segment $B \cap B'$. If $f(p)f(q) < 0$, create a vertex $v = (p + q)/2$ for the graph G. Also put v into $V(B)$. NOTE: if $f(p) = 0$ for any endpoint p, we treat $f(p)$ as positive; in effect, we are doing an infinitesimal perturbation at p. If B' is marked, retrieve the vertex v (if any) on the side $B \cap B'$, and put v into $V(B)$. It can be shown that $|V(B)| \in \{0, 2, 4\}$. If $|V(B)| = 2$, put the edge (p, q) into G to connect the vertices in $V(B)$. If $|V(B)| = 4$, it can be shown that one side of B contains two vertices a, b. We introduce two edges into E to connect each of a, b to the remaining two vertices. The requirement that these two edges are non-crossing ensures that the connection is unique (see [Plantinga and Vegter, 2004; Plantinga, 2006]).

The output graph $G = (V, E)$ can be viewed as a straightline graph, decomposed into a collection $P = P(G)$ of closed polygons or polygonal lines with endpoints in ∂B_0. In the following, we simply use G in place of P as the polygonal approximation.

4. Extension of Plantinga & Vegter

4.1. The Naïve Extension of Plantinga & Vegter

A major requirement of Plantinga & Vegter is that the curve is bounded and that the initial box is a bounding box. The power of subdivision methods comes from their ability to adaptively analyze local data directly. Choosing the initial box to be the bounding box gives up this power of local analysis; it also severely restricts the types of curves that are possible, e.g., the curve must be bounded. As a first attempt, one could naïvely attempt to run Plantinga & Vegter's algorithm starting with an arbitrary box, and then we ask the question "In what sense is the output G correct?" Intuitively, G should be isotopic to $f^{-1}(0) \cap B_0$, but Plantinga & Vegter did not discuss this issue. The algorithm certainly cannot handle the curve $S = f^{-1}(0)$ having tangential but non-crossing intersections [Yap, 2006] with ∂B_0. If we assume that there are only transversal intersections, we still face two problems: if the curve S (locally) enters and exits ∂B_0 by visiting only one box $B \subseteq B_0$, the above algorithm would fail to detect this small component. Such an error is called an (undetected) incursion, see Figure 1(a). Conversely, the curve S might escape undetected from B_0. Such an error is called an (undetected) excursion, see Figure 1(b). These errors cause the final approximation to be incorrect, since $S \cap B_0$ may not have the same number of components as G. Incursions account for components of $S \cap B_0$ that are undetected by the Plantinga & Vegter algorithm, and excursions account for components of G that are not connected in $S \cap B_0$. A single box may, in fact, contain several incursions

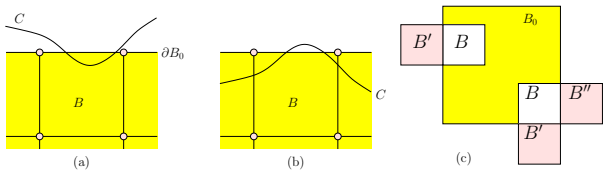


Fig. 1. (a) incursion, (b) excursion, (c) boundary boxes and their complements

and excursions. The excursion case is even more troubling because there is no guarantee that C_1 will hold in the complement of B_0, which implies that the components of G that are not connected in $S \cap B_0$ might, in fact, not even be connected in S. If we choose B_0 large enough, such errors cannot arise; but this approach gives up the power of adaptivity and locality which subdivision methods possess. If S has singularities, making B_0 large is not even an option. In this paper, we avoid any "largeness" assumption on B_0. We next extend the Plantinga & Vegter algorithm to arbitrary B_0 so that a suitable correctness statement can be made about the output polygonal approximation G. In fact, our region B_0 need not be a box nor even simply-connected.

The preceding issues arise because the Plantinga & Vegter algorithm focuses only on the parity of the endpoints of an edge of the subdivision. They prove that multiple intersections can be removed by applying a suitable isotopy. However, this does not work for the boundary of B_0, because it is possible to change the topology inside the box by performing these operations. In addition, it may be the case that the exterior of B_0 neither satisfies C_0 nor C_1, in which case, the proofs provided by Plantinga & Vegter can break down. We now adapt the Plantinga & Vegter algorithm to handle these cases.

The problem follows from the isotopy constructed by Plantinga & Vegter that remove multiple intersections from an edge of the subdivision. We could eliminate the use of this technique by ensuring that the curve passes through each boundary edge at most once; this idea is related to the simplest correctness proof in the Plantinga & Vegter paper. A test for this can be done by ensuring that either $0 \notin \square f(y)(H)$ or $0 \notin \square f_y(H)$ for horizontal sides H. A similar technique can be used for vertical sides. This will yield a polygonal approximation G that satisfies $G \approx C \cap B_0$. However, this requires knowing the topology of S on the boundary of B_0. Indeed, we even need to know if a point $p \in C \cap \partial B_0$ corresponds to a crossing or non-crossing intersection. This resembles Snyder's approach [Snyder, 1992b], by ensuring that the topology on the lower dimensional features is correct. We should expect this technique to subdivide too frequently near the boundary because it focuses on smaller features instead of using an isotopy to simplify and remove smaller features. Also, this is a recursive solution in lower dimensions (on ∂B_0), and so this technique becomes rather expensive in higher dimensions.

4.2. The Enlarged Region Solution

In the spirit of the Plantinga & Vegter algorithm, we now provide an alternative solution that avoids computing the exact boundary topology. We wish to slightly enlarge B_0 so that incursions/excursions can be removed by the isotopy described by Plantinga & Vegter. The basic idea is that, in addition to subdividing B_0 we find a slightly larger region B_0', which is the union of B_0 with a "collar" of squares around B_0. The straightline approximation G is then isotopic to the curve restricted some expansion B_0^+ of B_0 into

this collar. This is weaker than saying G is isotopic to $f^{-1}(0) \cap B_0$, but allows the algorithm to terminate faster.

Call a box $B \subseteq B_0$ a **boundary box** if ∂B overlaps ∂B_0. Let B be such a box. If B does not share a corner with B_0, then it has a unique **complementary box** B' such that B' has the same width as B, the interiors of B' and B_0 are disjoint, and $\partial B' \cap \partial B_0 = \partial B \cap \partial B_0$. Say B, B' are **partners** of each other. If B shares a corner with B_0, then it determines two complementary boxes B', B''. See Figure 1(c). The above "collar" comprises complementary boxes that satisfy either C_0 or C_1. In this way, excursions from B_0 or incursions into B_0 can be limited to to the collar region.

We begin with the case where B is a boundary box with a unique complementary box. If the complementary box satisfies C_0, then the curve does not intersect it. Therefore, there can be no incursions into or excursions from B. If the complementary box satisfies C_1, then if there is an excursion, the isotopy presented in Plantinga & Vegter shows that the approximation constructed by the naïve Plantinga & Vegter algorithm is correct in $B \cup B'^{+}$, where $B'^{+} \subseteq B'$. The only case that is left to consider is when there is an incursion. This case is harder because the isotopy constructed by Plantinga & Vegter would remove the component of the incursion, but that would result in an error in this case. In the spirit of Plantinga & Vegter's algorithm, we consider the sign pattern of the corners of complementary boxes.

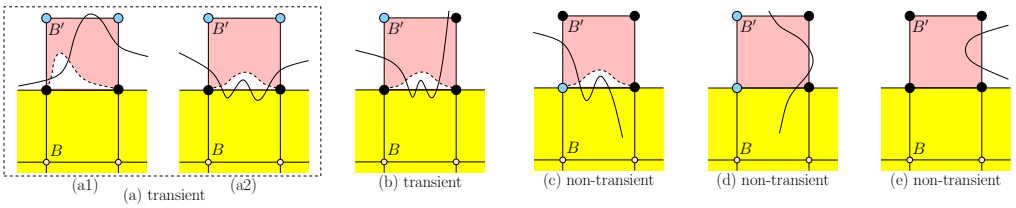
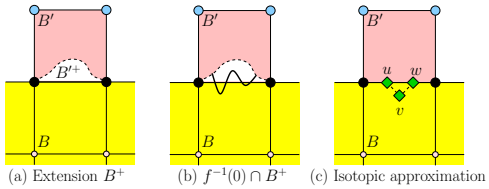


Fig. 2. Classification of complementary boxes according to sign of f at the corners.

We are only interested in complementary boxes B' that satisfy C_1 but not C_0. Up to symmetry and sign flips, B' can be put into one of five types, (a)-(e), based solely on the sign of f at the corners of B'. This is illustrated in Figure 2. Note that one sign pattern that appears to be missing cannot be C_1 as shown by Plantinga & Vegter.

We draw two instances of Types (a) to indicate the two possible "dispositions" of the curve $f^{-1}(0)$ in B': In (a1) the curve makes no incursion into B, but (a2) represents at least one incursion into B. These two dispositions give rise to distinct isotopy types for the curve $f^{-1}(0) \cap B$. Similarly, type (b) has two possible dispositions (but we only indicate the case where there is an incursion). But Types (c), (d), and (e) do not have analogous dispositions. Because of this difference in dispositions, we further classify Types (a) and (b) as **transient**; the other types are called **non-transient**.

Some Intuition. We show how the complementary boxes are used to yield a correctness statement. Suppose B' is a complementary box whose partner is B. Let the straightline $G = (V, E)$, when restricted to B, be denoted $G \cap B$. Note that $G \cap B$ has at most two edges. We would like to claim that $f^{-1}(0) \cap B$ is isotopic to $G \cap B$. But this is evidently false for Types (a) and (b) because of the two dispositions discussed above; but even for Type (c), this claim can be false because $f^{-1}(0) \cap B$ may have an arbitrary number of components due to incursions/excursions, as illustrated in Figure 2(c).



(a) Extension B^+ (b) $f^{-1}(0) \cap B^+$ (c) Isotopic approximation

Fig. 3. Representing the transient type.

To remedy the situation for Types (a) and (b), we first expand B into a slightly larger region $B^+ \subseteq B \cup B'$. The actual expansion $B'^+ := B^+ \setminus B$ is represented in Figure 3(a) by the white region demarcated by a dashed curve. Up to isotopy, the incursions into B^+ in Types (a1) and (a2) are equivalent to a single incursion of the curve $f^{-1}(0)$ into B^+ from B', as shown in Figure 3(b). In the graph $G = (V, E)$, we represent this incursion by adding three vertices u, v, w to V, with u and w on the side $B \cap B'$, and with v between u and w, but slightly to the interior of B'. We also add the edges $(u, v), (v, w)$ to E. This is illustrated in Figure 3(c). Let G^+ denoted the **augmented straightline graph** with these additional vertices and edges for transient complementary boxes.

Type (a) is important because it includes the situation where $f^{-1}(0)$ makes a non-crossing tangential intersection with the boundary of B_0. Detecting this situation is expensive, and provides a main motivation for exploiting isotopy. There is a similar expansion B^+ of B into B' for type (c), as indicated by dashed curves in Figure 2(c). In this case, we do not need to augment the graph $G = (V, E)$ with any new vertices or edges, because G already contains an edge representing a component in B.

We capture the above intuitive explanations as a lemma:

Lemma 3. Let $B \subseteq B_0$ be a boundary box and let B' be its partner (a complementary box) that satisfies C_1 or C_0. Assume that the curve $f^{-1}(0)$ does not pass through the corners of B. Then there exists a region B^+ such that:
(i) B^+ is an expansion of B into B':

$$B \subseteq B^+ \subseteq (B \cup B')$$

and the boundary of B^+ consists of two connected parts $\partial_1(B^+)$ and $\partial_2(B^+)$: one part $\partial_1(B^+)$ is the union of three sides of B (in Figure 2, it is the west, south and east sides of B), and the other part $\partial_2(B^+)$ is a curve in the interior of B' (in Figure 2, it is a dashed curve).
(ii) The augment straightline graph $G^+ = (V^+, E^+)$ when restricted to B is isotopic to $f^{-1}(0) \cap B^+$, i.e.,

$$G^+ \cap B \approx f^{-1}(0) \cap B^+.$$

Moreover, this isotopic map $I : B^+ \times [0, 1] \to B$ that witnesses this isotopy is the identity on $\partial_1(B^+)$.

Proof. Note that $G^+ \cap B$ has between 0 and 3 edges. This is because most two edges can appear in $G \cap B$ according to the Plantinga & Vegter construction rules, and in Types (a) and (b), the augmented graph G^+ has an additional edge.

For Types (d) and (e), B^+ can be taken to be B. That is because no incursion or excursion is possible along the side $B \cap B'$.

For Type (c), the $f^{-1}(0)$ intersects $B \cap B'$ at least once. If it intersects $B \cap B'$ multiple times, we can expand B into B^+ to ensure that $f^{-1}(0)$ intersects the boundary $\partial_2(B^+)$ exactly once.

For Type (a) and (b), there are two sides of B' where the curve $f^{-1}(0)$ intersects at least once. Moreover, there is a connected component X of $f^{-1}(0) \cap (B' \cup B)$ that connects these two sides – see Figure 2(a,b). Recall that we say X has two possible dispositions: either X intersects the side $B \cap B'$ or it does not. In either case, we can expand B into B^+ so that X intersects $\partial_2(B^+)$ exactly twice. This component $X \cap B^+$ is represented by the augmented edges $(u, v), (v, w)$ in E^+. **Q.E.D.**

We now present the **Extended Plantinga & Vegter algorithm**. It has 3 Phases that parallel the algorithm in Section 2. Phase i (for $i = 1, 2, 3$) works off queues Q_i and Q'_i, transferring boxes into Q_{i+1} and Q'_{i+1}. The queue Q'_i holds complementary boxes while Q_i holds regular boxes.

- PHASE 1: SUBDIVISION. While Q_1 is non-empty, remove some B from Q_1, and perform the following: If $C_0(B)$ holds B is discarded. If $C_1(B)$ holds, and also $C_1(B')$ or $C_0(B')$ holds for every complementary box B' of B, then (a) insert B into Q_2, and (b) for each complementary box B' that satisfies C_1 but not C_0, insert B' into Q'_2. Otherwise, split B into four subboxes which are inserted back into Q_1.
- PHASE 2: BALANCING. The balancing of boxes in Q_2 is done as in Phase 2 of Section 3 – note that boxes are inserted into Q_3 by this process. Next, we perform an analogous while-loop on Q'_2: While Q'_2 is non-empty, remove any B' from Q'_2. If the partner of B' had been split during balancing, we **half-split** B' (this means we split it into four children, and for each B'' of the two children that intersect ∂B_0, we put B'' into Q'_2 provided B'' satisfies C_1 and not C_0). Otherwise, we place B' into Q'_3.
- PHASE 3: CONSTRUCTION. First, perform the Phase 3 of Section 3 which constructs a graph $G = (V, E)$ from the boxes queue Q_3. Next we augment this graph using queue Q'_3: for each $B' \in Q'_3$, if B' is a transient type, we insert three vertices and two edges to the graph G as described above. The resulting straightline graph is denoted $G^+ = (V^+, E^+)$.

4.3. Weak Correctness Statement.

We define B'_0 to be the union of B_0 with all of the complementary boxes B' that were placed into Q'_3.

We are ready to prove the correctness of this Extended Plantinga & Vegter Algorithm. But before formulating the correctness statement, notice that Lemma 3 has an important caveat – the curve must not pass through the corners of boxes that lies on ∂B_0. So far, we have followed Plantinga & Vegter in assuming that f is never zero at a corner of a box. To be explicit, let $\widetilde{f}$ denote the **standard perturbation** of f is defined to be a function that agrees with f everywhere except in arbitrarily small neighborhoods of corners of boxes where f is 0; in such neighborhoods, $\widetilde{f}$ is positive. Note that $\widetilde{f}$ depends on the underlying quadtree. We continue to assume this standard perturbation in Extended Plantinga & Vegter Algorithm. Now, there is a subtle difference. In Plantinga and Vegter [2004], it is noted that the curve $\widetilde{S} = \widetilde{f}^{-1}(0)$ is isotopic to $S = f^{-1}(0)$. So their use of $\widetilde{f}$ is inconsequential, up to isotopy. But in our setting, $\widetilde{S} \cap B_0$ may no longer be isotopic to $S \cap B_0$. For instance, suppose S makes a non-crossing tangential intersection with the boundary ∂B_0 at the corner of boundary box B. This intersection is an isolated

component of $S \cap B_0$, and we could lose this component in $\widetilde{S} \cap B_0$. Our correctness statement is therefore about $\widetilde{S} \cap B_0$, and not about $S \cap B_0$.

We regard the use of $\widetilde{f}$ as a reasonable compromise because (1) $\widetilde{f}$ is an infinitesimal perturbation of f (being arbitrarily smaller than any actual perturbation of f); (2) $\widetilde{f}$ is easy to implement and is an effective perturbation (comparing favorably to the alternative of, say, a randomized perturbation); (3) we can detect when $\widetilde{S} \cap B_0$ might deviate from $S \cap B_0$ (namely, when we encounter a zero at a corner on the boundary ∂B_0); (4) it is a very simple solution to what would otherwise be serious complications arising from various degeneracies such as S containing a horizontal or vertical component; and finally, (5) its use is consistent with the exploitation of isotopy in Plantinga & Vegter.

THEOREM 4 (Weak Correctness). Let the curve $S = f^{-1}(0)$ be non-singular in the box B_0, and $G^+ = (V^+, E^+)$ be the augmented straightline graph constructed by the Extended Plantinga & Vegter Algorithm. There exists a region B_0^+ isotopic to B_0,

$$B_0 \subset B_0^+ \subset B_0',$$

such that

$$G^+ \approx \widetilde{f}^{-1}(0) \cap B_0^+$$

where $\widetilde{f}$ is the standard perturbation of f.

Proof. In Lemma 3, we defined an expansion B^+ of each boundary box B; let B_0^+ be the union of B_0 with all the expansions B^+ of B, where B ranges over all the boundary boxes. First consider the graph $G = (V, E)$ which is obtained from the standard construction in the Plantinga & Vegter Algorithm. If $\widetilde{f}^{-1}(0)$ is entirely contained in B_0, then $G \approx \widetilde{f}^{-1}(0)$ follows from the correctness of Plantinga & Vegter Algorithm. Thus we are only concerned about missing those components of $\widetilde{f}^{-1}(0) \cap B_0$ whose endpoints are both in ∂B_0. Let X be such a component. Thus there is a boundary box B such that $X \cap B$ is non-empty. If B^+ is the expansion of B given by Lemma 3, there is a local isotopy $I : B^+ \times [0, 1] \to B$ that witnesses the isotopy

$$G^+ \cap B \approx \widetilde{f}^{-1}(0) \cap B^+.$$

Moreover, all these local isotopies can be combined into a global isotopy $I^* : B_0^+ \times [0, 1] \to B_0$ which witnesses

$$G^+ \approx \widetilde{f}^{-1}(0) \cap B_0^+.$$

Since $X \cap B^+$ is non-empty, this means that I^* witnesses the isotopy of a component of $G^+ \cap B$ to X. This proves that every component of $\widetilde{f}^{-1}(0) \cap B_0$ is isotopic to some component of G^+. Note that G^+ might contain components that appear in $\widetilde{f}^{-1}(0) \cap B_0^+$ but not in $\widetilde{f}^{-1}(0) \cap B_0$. **Q.E.D.**

REMARKS:

1. Under the conditions of this theorem, we call G^+ a **weak isotopic approximation** of $f^{-1}(0) \cap B_0$. A slightly different approach to a weak correctness statement was used in our ISSAC 2008 proceedings version: there, we required $f^{-1}(0)$ to only intersect ∂B_0 transversally. Then we explicitly distinguish between the two dispositions in the transient boxes (Types (a) or (b)). This requires an iteration at each transient box. We prefer the current solution because it is not easy to check whether $f^{-1}(0)$ only intersects ∂B_0 transversally.
2. We can make the incursion/excursion isotopy to have a perturbation bound of $\varepsilon > 0$

by ensuring each of those complementary boxes of types (a), (b), and (c) has width at most ε. This is easily done by modifying the balancing phase of the algorithm.
3. We have assumed that complementary boxes have the same width as their partner. An alternative, possibly more efficient, approach is to allow complementary boxes to have widths less than their partners. Call these "subcomplementary boxes". So a boundary box B can have many subcomplementary boxes. We can do half-splits of subcomplementary boxes of types (a), (b) and (c) as long as their widths are greater than ε. On the other hand, a boundary box B is allowed to have at most one subcomplementary box of type (c) or (d), and this occurs under under strict conditions (we leave the details to the reader).

4.4. Extension to Nice Regions

It is essential in our applications later to extend the above refinements to non-simply connected regions. For this purpose, we define a **nice region** R_0 (relative to a square B_0) as the union of any collection of leaves taken from a quadtree rooted at B_0. Thus, $R_0 \subseteq B_0$. To extend Theorem 4 to nice regions, we note two simple modifications:
(I) A complementary box B' of a boundary box $B \subseteq R_0$ may intersect the interior of R_0 or other complementary boxes. Thus, Phase 1 must split such boundary boxes B sufficiently. Such interference checks can be checked during the subdivision phase.
(II) The region R_0 can have concave corners. A complementary box B' at a concave corner has two choices for a partner $B \subseteq R_0$. Relative to each choice of B, we classify B' into one of 5 types as in Figure 2(a)-(e). This is illustrated in Figure 4(i)-(iv). In Figure 4(i), for instance, the box B' is Type (a) (hence transient) relative to the indicated box B, but it is Type (d) (hence non-transient) relative to the other partner $\widehat{B}$. In Figure 4(ii), the corner complementary box B' is Type (b) (hence transient) relative to both B and $\widehat{B}$. Similarly, the other two cases seen in Figure 4(iii)-(iv) have dual classifications. We modify our augmentation of the graph $G = (V, E)$ as follows: for each complementary box at a concave corner, we consider its classification relative to each choice of partner B: if the classification is transient, as before, we add three vertices u, v, w and edges $(u, v), (v, w)$ to G on the side of $B' \cap B$.

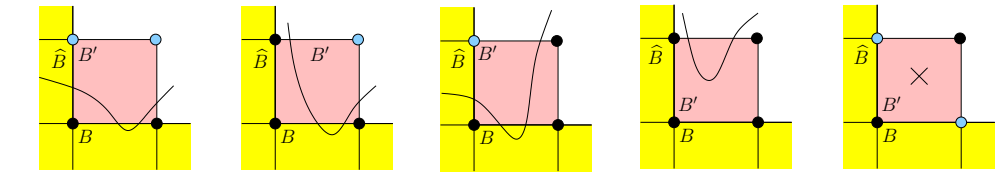


Fig. 4. Classification of complementary boxes at concave corners.

5. Evaluation Bounds

For any function f, define its **evaluation bound** to be

$$\texttt{EV}(f) := \inf\{|f(p)| \ : \ f(p) \neq 0, \nabla f(p) = 0\} \tag{2}$$

Such bounds were used in [Cheng et al., 2008; Burr et al., 2008]. From Proposition 2, we see that $\{f(p) : \nabla f(p) = 0\}$ is a finite set and therefore $\mathrm{EV}(f) > 0$. However there is no explicit bound readily available. The main objective of this section is to provide such a bound:

THEOREM 5. If $f \in \mathbb{Z}[X, Y]$ has degree d and $\|f\| < 2^L$ then $-\lg \mathrm{EV}(f) = O(d^2(L+d))$. More precisely,

$$\mathrm{EV}(f)^{-1} \leq \max\left\{[d^{d+6}2^{L+2d+9}]^{d^2-1}, [d^{3d+9}2^{3L+4d}]^d\right\}$$

Before giving the proof, we provide some definitions and preparation. Let $h = \sum_{i+j=0}^{d} a_{ij}X^iY^j \in \mathbb{C}[X, Y]$. $\|h\|_k := \sqrt[k]{\sum_{i+j=0}^{d} |a_{ij}|^k}$ will denote the k-norm of h, where we use $k = 1, 2, \infty$. We just write $\|h\|$ for $\|h\|_\infty := \max\{|a_{ij}| : 0 \leq i + j \leq d\}$, denoting the height of h.

Now let $f = \sum_{i+j=0}^{d} a_{ij}X^iY^j \in \mathbb{Z}[X, Y]$. As input parameters in our bounds, we use d and L where $\deg f \leq d$ and $\|f\| < 2^L$. Let f_x, f_y denote the derivatives of f. We may write

$$\textsc{Zero}(f_x, f_y) = \bigcup_i U_i \cup \bigcup_j V_j$$

where U_i are 1-dim and V_j are 0-dim irreducible components. On each component U_i or V_j, one can show that the function f is constant. E.g. $f = (xy+1)^2 - 1$, $f_x = 2(xy+1)y$ and $f_y = 2(xy+1)x$. Then $U_1 = \{xy+1 = 0\}$ and $V_1 = \{(0, 0)\}$. The function f is equal to 1 on U_1 and equal to 0 on V_1.

Let $g := \mathtt{GCD}(f_x, f_y)$, and also

$$g_x := f_x/g, \qquad g_y := f_y/g.$$

Clearly, we have

$$\textsc{Zero}(f_x, f_y) = \textsc{Zero}(g) \cup \textsc{Zero}(g_x, g_y).$$

Since $\mathtt{GCD}(g_x, g_y) = 1$, we conclude that $\textsc{Zero}(g_x, g_y)$ has no 1-dimensional component. Conversely, the hyper-surface $\textsc{Zero}(g)$ has no 0-dimensional component. This proves:

LEMMA 6.

$$\textsc{Zero}(g) = \bigcup_i U_i, \qquad \textsc{Zero}(g_x, g_y) = \bigcup_j V_j.$$

We now view the ring $\mathbb{Z}[X, Y] \simeq \mathbb{Z}[Y][X] \simeq \mathbb{Z}[X][Y]$ in three alternative ways. A bivariate polynomial f in X and Y can be written as $f = f(X, Y)$, $f = f(X; Y)$ or $f = f(Y; X)$ to indicate these three views respectively. As a member of $\mathbb{Z}[X, Y]$, the coefficients of $f(X, Y)$ are elements of $\mathbb{Z}$. But $f = f(X; Y)$ is a member of $\mathbb{Z}[Y][X]$ whose coefficients are polynomials in $\mathbb{Z}[Y]$. Similarly for $f = f(Y; X)$. The leading coefficient and degree of f are likewise affected by these views: $lc(f(X; Y)) \in \mathbb{Z}[Y]$ but $lc(f(X, Y)) \in \mathbb{Z}$, $d = \deg(f(X, Y))$ is the total degree of f while $\deg(f(X; Y))$ is the largest power of X occurring in f.

We use Mahler's basic inequality ([Yap, 2000, p. 351]) that if $g \in \mathbb{Z}[X, Y]$ and $g|f$ then

$$\|g(X, Y)\|_1 \leq 2^D \|f(X, Y)\|_1 \tag{3}$$

where $D = \deg(f(X; Y)) + \deg(f(Y; X))$. This implies:

$$\|g(X, Y)\|_1 \leq 4^{d-1}d^3 2^L, \qquad \|g_x(X, Y)\|_1 \leq 4^{d-1}d^3 2^L. \tag{4}$$

Note that $g|f_x, g_x|f_x$, and $\|f_x\|_1 = \sum_{i+j=0}^{d} |ia_{ij}| \leq d^2 \cdot d2^L$, $\deg(f_x(X;Y))+\deg(f_x(Y;X)) \leq 2d-2$. The bound then follows from (3).

Let $h(X)$ be the leading coefficient of $g(Y;X)$. Since $h(X)$ has degree $\leq d-1$, there is an integer $x_0 \in \{0,1,\ldots,d-1\}$ such that $h(x_0) \neq 0$. Intersect $\textsc{Zero}(g)$ with the line $X = x_0$. CLAIM: This line cuts each non-vertical component U_i in a finite but non-zero number of points. In proof, let $g = \prod_i g_i$ where $\textsc{Zero}(g_i) = U_i$. Setting $d_i := \deg g_i(Y;X)$, we see that the vertical components correspond to $d_i = 0$. Then $lc(g(Y;X)) = \prod_i lc(g_i(Y;X))$ and $h(x_0) = lc(g(Y;x_0)) \neq 0$ iff for all i, $lc(g_i(Y;x_0)) \neq 0$. But $g_i(Y;x_0)$ is a polynomial of degree d_i in $\mathbb{Z}[Y]$, and has exactly d_i solutions in $\mathbb{C}$.

Write $f_0(Y) := f(x_0, Y)$ and $g_0(Y) = g(x_0, Y)$. From (4):

$$\|g_0\|_1 \leq d^d \|g(X,Y)\|_1 \leq 4^{d-1} d^{d+3} 2^L. \tag{5}$$

It is also easy to see that

$$\|f_0\| \leq d^{d+1} 2^L. \tag{6}$$

Suppose $\beta \in \textsc{Zero}(g_0) \setminus \textsc{Zero}(f_0)$. We want a lower bound on $|f_0(\beta)|$. For this purpose, we use an evaluation bound from [Burr et al., 2008, Theorem 13(b)]:

Proposition 7 (Evaluation Bound [Burr et al., 2008]). Let $\phi(x), \eta(x) \in \mathbb{C}[x]$ be complex polynomials of degrees m and n. Let $\beta_1, \ldots, \beta_n$ be all the zeros of $\eta(x)$. Suppose there exists relatively prime $F, H \in \mathbb{Z}[x]$ such that $F = \phi\overline{\phi}, H = \eta\overline{\eta}$ for some $\overline{\phi}, \overline{\eta} \in \mathbb{C}[x]$. If the degrees of $\overline{\phi}$ and $\overline{\eta}$ are $\overline{m}$ and $\overline{n}$, then

$$\prod_{i=1}^{n} |\phi(\beta_i)| \geq \frac{1}{lc(\overline{\eta})^m \cdot ((m+1)\|\phi\|)^{\overline{n}} M(\overline{\eta})^m \cdot \left((\overline{m}+1)\|\overline{\phi}\|\right)^{n+\overline{n}} M(H)^{\overline{m}}}. \tag{7}$$

Here the Mahler measure $M(h)$ of a function $h \in \mathbb{C}[x]$ with zeros $\alpha_1, \cdots, \alpha_n$, is defined $M(h) := |lc(h)| \prod_{|\alpha_i| \geq 1} |\alpha_i|$. We shall choose the functions in Proposition 7 as follows:

$$\phi := f_0, \qquad H = \frac{g_0}{\texttt{GCD}(f_0, g_0)}.$$

Moreover, let $\overline{\phi} := 1$, $\eta(X) := X - \beta$ and $\overline{\eta} := H/\eta \in \mathbb{C}[x]$. Hence

$$m \leq d, \quad n = d, \quad \overline{m} = 0, \quad \overline{n} \leq d-1.$$

Also

$$lc(\overline{\eta}) = lc(H) = lc(g_0) \leq \|g_0\| \leq \|g_0\|_1. \tag{8}$$

Further,

$$M(\overline{\eta}) \leq M(H) \leq \|H\|_1 \leq 2^d \cdot \|g_0\|_1.$$

Finally, an application of Proposition 7 gives

$$\begin{aligned} |f_0(\beta)|^{-1} &\leq lc(\overline{\eta})^d \cdot ((d+1)\|f_0\|)^{d-1} \cdot M(\overline{\eta})^d \\ &< \left[lc(\overline{\eta}) \cdot (d+1)\|f_0\| \cdot M(\overline{\eta})\right]^d \\ &\leq \left[\|g_0\|_1 \cdot (d+1) d^{d+1} 2^L \cdot 2^d \|g_0\|_1\right]^d \\ &\leq \left[d^{3d+9} 2^{3L+4d}\right]^d. \end{aligned} \tag{9}$$

(9) is a lower bound on $|f(p)|$ where p lies in a non-vertical component U_i. By considering $g(X;Y)$, the same bound applies for $|f(p)|$ when p lies in a vertical component U_i.

We obtain a lower bound for $f(p)$ with $p \in \text{ZERO}(g_x, g_y)$. Consider the system $\Sigma \subseteq \mathbb{Z}[X,Y,Z]$ where

$$\Sigma = \{Z - f(X,Y), g_x(X,Y), g_y(X,Y)\}$$

The zeros $(\xi_1, \xi_2, \xi_3) = (\xi_1, y, f(x,y)) \in \mathbb{C}^3$ of Σ satisfy $\xi_3 = f(\xi_1, \xi_2)$. Since Σ is a zero dimensional system, we may apply the multivariate zero bound in [Yap, 2000, p. 350]. This bound says that

$$|\xi_3|^{-1} < (2^{3/2}NK)^D 2^{8(d-1)} \tag{10}$$

where $N = \binom{1+2(d-1)}{3}$, $D = d^2 - 1$ and

$$K = \max\{\sqrt{3}, \|g_x\|_2, \|g_y\|_2, \|Z - f(X,Y)\|_2\}.$$

We have $\|Z - f(X,Y)\|_2 \leq 1 + (d+1)2^L$. From (4), we see that $K \leq 4^{d-1}d^{d+3}2^L$. Using the bound $N < 2d^3$, we obtain

$$|\xi_3|^{-1} < [2^2 \cdot 2d^3 \cdot 2^{2d-2}d^{d+3}2^L]^{d^2-1} \cdot 2^{8(d^2-1)} < [d^{d+6}2^{L+2d+9}]^{d^2-1}. \tag{11}$$

Now Theorem 5 easily follows from (9) and (11).

6. Isolating Singular Points

In the rest of this paper, we assume that $f \in \mathbb{Z}[X,Y]$, and the curve $S = f^{-1}(0)$ intersects ∂B_0. We would like to use the Extended Plantinga & Vegter algorithm to compute an isotopic approximation to $\texttt{Zero}(f)$ when f has only isolated singularities. Since the Plantinga & Vegter algorithm does not terminate near singular points, it is necessary to isolate the singular points from the rest of B_0.

We use the auxiliary function $F = f^2 + f_X^2 + f_Y^2$. Finding the singular points of $f^{-1}(0)$ amounts to locating and isolating the zeros of this non-negative function. We use a simple mountain pass theorem [Jabri', 2003] adapted to B_0 to ensure our algorithm isolates the zeros.

THEOREM 8. Suppose that $F \geq 0$ on B_0, and that $F > 0$ on ∂B_0. Then for any two distinct roots p, q of F in B_0, there exists a continuous path $\gamma : [0,1] \to B_0$ connecting p and q which satisfies the following:

- γ minimizes $M_\gamma := \max_{x \in [0,1]} F(\gamma(x))$ among all paths connecting p and q in B_0.
- γ contains a point y such that either $\nabla F(\gamma(y)) = 0$ or $\gamma(y) \in \partial B$.

This can be proved using path deformation and the compactness of B_0, or it can be seen as a simple application of the topological mountain pass theorem presented in [Jabri', 2003]. Because of this theorem, distinct zeros of F within B_0 are separated by barriers of height $\epsilon = \min(\texttt{EV}(F), \min F(\partial B_0))$. This leads us to the following multistep process to localize these zeros. The goal is to find a small rectangle with diameter less than some δ around each zero.

STEP 0: DETERMINING ϵ. Initialize ϵ to any lower bound on $\texttt{EV}(F)$. Also initialize Q_1 to $\{B_0\}$, and Q_2 to be empty. In general, each square in Q_1 intersects the boundary of B_0. While Q_1 is non-empty, remove a square S from it and evaluate $\square F(S)$. If $\square F(S) > 0$ we push S into the queue Q_2 and also update ϵ to $\min\{\epsilon, \min \square F(S)\}$. If $0 \in \square F(S)$,

subdivide S and push the children of S which intersect ∂B_0 into Q_1, and the others into Q_2. When Q_1 is empty, we stop and now the value of ϵ is fixed.

STEP 1: INITIAL SUBDIVISION. Initialize queue Q_3 to be empty. While there is an S in Q_2, remove it and evaluate $\square F(S)$. If $\square F(S) > \epsilon/2$, discard S. Else if $\square F(S) < \epsilon$, push S into Q_3. Else subdivide S and push its children into Q_2.

Once Q_2 is empty, group the elements of Q_3 into connected regions A_i ($i \in I$). Each A_i contains at most one root, since otherwise, there would be a path connecting the roots within A_i. The value of F along this path would be less than ϵ, contradicting the mountain pass theorem. For later reference, let C be the region $B_0 \setminus \cup_i A_i$. Note by Step 0 that F is greater than $\epsilon/2$ on C and that $\partial B_0 \subset C$, i.e., each A_i doesn't touch ∂B_0.

STEP 2: REFINEMENT. For each A_i ($i \in I$), initialize queue $Q_{4,i}$ with all squares $S \in A_i$. So long as neither terminating condition 1 nor 2 (below) hold, we perform the following: For each S in $Q_{4,i}$, if $0 \in \square F(S)$, subdivide S and push its children into $Q_{4,i}$. If $0 \notin \square F(S)$, discard S. We terminate when either of the following two conditions are met:

(1) $Q_{4,i}$ is empty, in which case there isn't a zero in A_i.
(2) A_i', the contents of $Q_{4,i}$ satisfy all of the following:
   (a) $\square F(S) < \epsilon/2$ for some $S \in A_i'$
   (b) R_i, the smallest rectangle containing A_i', lies within the region covered by the original A_i.
   (c) The diameter of R_i is less than δ.

It is clear from the definition of F that this step will halt. We claim that each R_i contains exactly one root. In Step 1, we showed that A_i contains at most one root. To see that R_i contains a root, take a point of A_i' where $\epsilon/2$, then follow the path of steepest descent to reach a zero of F. Because F is less that $\epsilon/2$, on this curve, the curve cannot pass through the region C to reach any other R_j or to leave B_0. Therefore there must be a zero within A_i. It is in R_i because our conditions ensure that F is positive on $A_i \setminus R_i$.

7. Determining the Degree of Singular Points

The following standard result from [Krantz and Parks, 1992; Łojasiewicz, 1991] shows that the global structure of zero sets:

PROPOSITION 9 (Zero Structure). Let f be real analytic. Then $\texttt{Zero}(f)$ can be decomposed into a finite union of pieces homeomorphic to $(0, 1)$, pieces homeomorphic to S^1, and singular points.

Viewing $\texttt{Zero}(f)$ as a multigraph G, the **degree** of a singular point is its degree as a vertex of G. We now determine such degrees. Let δ_3 be a separation bound between singular points, so if p and q are two distinct singular points of $\texttt{Zero}(f)$, then the distance between p and q is at least δ_3. Let δ_4 be a separation bound so that if r is a point on $\texttt{Zero}(f)$ such that $\nabla f(r)$ is in the same direction as the line between r and a singular point p, then the distance between p and r is at least δ_4. If s is on $\texttt{Zero}(f)$ so that the distance between s and a singular point p is smaller than either δ_3 or δ_4, then by following the paths $\texttt{Zero}(f)$ away from s, one of the paths strictly monotonically approaches p until it reaches p and the other path locally strictly monotonically recedes from p. See [Yap,

2006] for explicit bounds on δ_3 as a function of degree and height of $f(X,Y)$. We can similarly derive explicit bounds on δ_4.

To find the degree of a singular point, assume that we have two boxes $B_1 \supsetneq B_2$ where the diameter of B_1 is less than both δ_3 and δ_4, B_2 contains a singular point of f and there is some radius $r > 0$ such that a circle of radius r centered at any point p inside B_2 must lie entirely within the annulus $B_1 \setminus B_2$. See Figure 5(a). Furthermore, to apply our extended Plantinga & Vegter algorithm of Section 4, we can ensure that $B_1 \setminus B_2$ is a nice region. Note this condition is satisfied if B_1 at least 5 times larger than B_2 and B_2 is centered in B_1, which is the typical situation we consider below. See Figure 5(b).

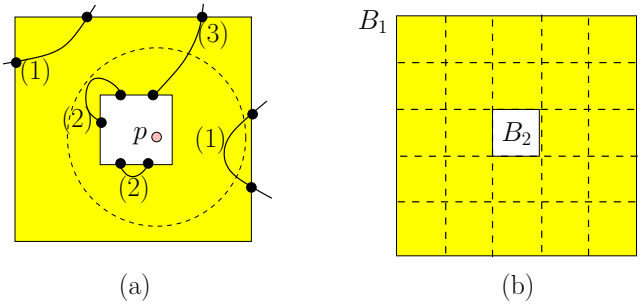


Fig. 5. Annular region $B_1 \setminus B_2$ with singularity p and the three types (1), (2), (3) of components.

Now, there are 3 types of components in $\texttt{Zero}(f) \cap (B_1 \setminus \text{int}(B_2))$: (1) images of $[0,1]$ both of whose endpoints are on ∂B_1, (2) images of $[0,1]$ both of whose endpoints are on ∂B_2, and (3) images of $[0,1]$ with one endpoint on each of ∂B_1 and ∂B_2. These three types are illustrated in Figure 5.

Let s be a point on *any* of these components, then traveling along $\texttt{Zero}(f)$ in one direction must lead to the singular point and the other direction must leave the neighborhood (be further than $\min\{\delta_3, \delta_4\}$) of the singular point. For, if not, then there must be a point r on $\texttt{Zero}(f) \cap B_1$ such that $\nabla f(r)$ is in the same direction as the line between s and the singular point, which is impossible since the width of B_1 is smaller than δ_4. Now any component accumulating on a singular point exits the neighborhood of the singular point and the only way to leave the neighborhood is by way of a type (3) component. This shows:

LEMMA 10. The degree of the singular point in B_2 is the number of components of type 3.

8. Overall Algorithm

We now put all the above elements together to find a weak isotopic approximation to the algebraic curve $S = f^{-1}(0)$ within a nice region R_0 where $f(X,Y) \in \mathbb{Z}[X,Y]$ has only isolated singularities. For simplicity, we assume that ∂R_0 intersects the curve S. We first find the singularities of the curve S in R_0. Using the technique of Section 5, we can isolate the singularities p_i $(i = 1, 2, \ldots)$ into pairwise disjoint boxes B_i. We may assume the width of the B_i's is at most $\min\{\delta_3, \delta_4\}/6$. Let B_i' be the box of width 5 times the width of B_i, and concentric with B_i; we may further assume $B_i' \subseteq R_0$. Note these combinations are chosen to ensure that we have the typical situation in Section 7.

Now we proceed to run the extended Plantinga & Vegter algorithm on the nice region $R^* := R_0 \setminus \bigcup_i B_i$, yielding a polygonal approximation G. We directly incorporate the technique of Section 7 into the following argument. If p_i is the singular point in B_i, then the degree of p_i is equal to the number of type (3) components in $G \cap (B_i' \setminus B_i)$. We directly connect these components directly to p_i, and discard any type (1) or type (2) components. This produces the desired isotopic approximation.

Remarks: We have not discussed ε-approximation because this is relatively easy to achieve in the Plantinga & Vegter approach. We only have to make sure that each subdivision box that contains a portion of the polygonal approximation G has width at most $\varepsilon/4$. Note the result of the Plantinga & Vegter algorithm can deform the original curve by at most one cell away.

9. Conclusion

This paper presents the first complete numerical subdivision algorithm for meshing an implicit algebraic curve that has only isolated singularities. This solves an open problem in the exact numerical approaches to meshing in 2-D [Boissonnat et al., 2006, p. 187]. We pose three challenges:
(a) A worst case complexity bound for our procedure is possible. But this may not be the best way to measure adaptive algorithms. We would like to provide adaptive bounds, similar to the integral analysis in [Burr et al., 2008] for 1-D problems.
(b) In 3-D, a square-free integer polynomial $f(X, Y, Z)$ could have a 1-dimensional singularities. We pose the problem of designing a purely numerical subdivision algorithm to handle 1-dimensional singularities.
(c) The practical implementation of an adaptive algorithm handling singularities, even based on our outline, must handle many important details. Computational experience is invaluable for future research into singularity computation.

References

Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics. Springer, 2003.

J.-D. Boissonnat and S. Oudot. Provably good sampling and meshing of Lipschitz surfaces. In *Proc. 22nd ACM Symp. on Comp. Geometry*, pages 337–346, 2006. Sedona, Arizona.

J.-D. Boissonnat, D. Cohen-Steiner, B. Mourrain, G. Rote, and G. Vegter. Meshing of surfaces. In J.-D. Boissonnat and M. Teillaud, editors, *Effective Computational Geometry for Curves and Surfaces*. Springer, 2006. Chapter 5.

Jean-Daniel Boissonnat, David Cohen-Steiner, and Gert Vegter. Isotopic implicit surfaces meshing. In *ACM Symp. on Theory of Computing*, pages 301–309, 2004.

Michael Burr, Felix Krahmer, and Chee Yap. Integral analysis of evaluation-based real root isolation, March 2008. Submitted, 2008.

Michael Burr, Vikram Sharma, and Chee Yap. Evaluation-based root isolation, February 2009. In preparation.

Jin-San Cheng, Xiao-Shan Gao, and Chee-Keng Yap. Complete numerical isolation of real zeros in zero-dimensional triangular systems. *J. of Symbolic Computation*, page ???, 2008. doi: DOI: 10.1016/j.jsc.2008.04.017. Special Issue of JSC based on ISSAC 2007.

S.-W. Cheng, T.K. Dey, E.A. Ramos, and T. Ray. Sampling and meshing a surface with guaranteed topology and geometry. In *Proc. 20th ACM Symp. on Comp. Geometry*, pages 280–289, 2004.

D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer-Verlag, New York, 1992.

Zilin Du and Chee Yap. Uniform complexity of approximating hypergeometric functions with absolute error. In Sung-il Pae and Hyungju Park, editors, *Proc. 7th Asian Symp. on Computer Math. (ASCM 2005)*, pages 246–249, 2006.

Joe Harris. *Algebraic Geometry*. Springer-Verlag, 1992.

Robin Hartshorne. *Algebraic Geometry*. Springer-Verlag, New York, 1977.

H. Hong. An efficient method for analyzing the topology of plane real algebraic curves. *Mathematics and Computers in Simulation*, 42:571–582, 1996.

Youssef Jabri'. *The Mountain Pass Theorem: Variants, Generalizations and Some Applications*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2003.

S. G. Krantz and H. R. Parks. *A primer of real analytic functions*. Birkhäuser Verlag, Basel, 1992. ISBN 3-7643-2768-5.

S. Łojasiewicz. *Introduction to complex analytic geometry*. Birkhäuser Verlag, Basel, 1991. ISBN 3-7643-1935-6. Translated from the Polish by Maciej Klimek.

Bernard Mourrain and J.-P. Técourt. Isotopic meshing of a real algebraic surface. Technical Report RR-5508, INRIA, Sophia-Antipolis, France, February 2005. Also, electronic proceedings, MEGA 2005.

Simon Plantinga. *Certified Algorithms for Implicit Surfaces*. Ph.D. thesis, Groningen University, Institute for Mathematics and Computing Science, Groningen, Netherlands, December 2006.

Simon Plantinga and Gert Vegter. Isotopic approximation of implicit curves and surfaces. In *Proc. Eurographics Symposium on Geometry Processing*, pages 245–254, New York, 2004. ACM Press.

Elmar Schoemer and Nicola Wolpert. An exact and efficient approach for computing a cell in an arrangement of quadrics. *Comput. Geometry: Theory and Appl.*, 33:65–97, 2006.

Raimund Seidel and Nicola Wolpert. On the exact computation of the topology of real algebraic curves. In *Proc. 21st ACM Symp. on Comp. Geometry*, pages 107–116, 2005. Pisa, Italy.

J. M. Snyder. *Generative Modeling for Computer Graphics and CAD: Symbolic Shape Design using Interval Analysis*. Academic Press, 1992a.

J. M. Snyder. Interval analysis for computer graphics. *SIGGRAPH Comput.Graphics*, 26(2):121–130, 1992b.

Barton T. Stander and John C. Hart. Guaranteeing the topology of an implicit surface polygonalization for interactive meshing. In *Proc. 24th Computer Graphics and Interactive Techniques*, pages 279–286, 1997.

Chee K. Yap. *Fundamental Problems of Algorithmic Algebra*. Oxford University Press, 2000.

Chee K. Yap. Complete subdivision algorithms, I: Intersection of Bezier curves. In *22nd ACM Symp. on Comp. Geometry*, pages 217–226, July 2006.