Plantinga-Vegter algorithm takes average polynomial time

Josué Tonelli-Cueto (BMS & TU Berlin)
together with
Felipe Cucker (CityU Hong Kong) and Alperen A. Ergür (TU Berlin)

July 18, 2019

Slides at `https://tonellicueto.xyz/pdf/ISSAC19_slides.pdf`

This presentation is about the accepted paper

Plantinga-Vegter algorithm takes average polynomial time

authored by

- Felipe Cucker (CityU Hong Kong),
- Alperen Ergür (Technische Universität Berlin), and
- Josué Tonelli-Cueto (Technische Universität Berlin)

This work was funded by the

EINSTEIN
Foundation.de

Einstein Stiftung Berlin
Einstein Foundation Berlin,

within the Einstein Visiting Fellowship ”Complexity and accuracy of numerical algorithms in algebra and geometry” of Felipe Cucker.

Table of contents

Plantinga-Vegter algorithm

Setting

What do we have?

- An implicit curve C inside $[-a, a]^2$
  given by a C^1 function $f : [-a, a]^2 \to \mathbb{R}$
- Interval approximations $\square f$ of f and $\square \partial f$ of ∂f

What do we want?

- Piecewise-linear approximation L of C in $[-a, a]^2$such that
  $([-a, a]^2, C)$ and $([-a, a]^2, L)$ are isotopic

Any assumptions?

- C smooth
- C Intersects the boundary of $[-a, a]^2$ transversely

Plantinga-Vegter algorithm for curves I

Algorithm: PV Algorithm for curves (Plantinga, Vegter; 2004)

Input: $a \in (0, \infty)$ and $f : \mathbb{R}^2 \to \mathbb{R}$
with interval approximations $\square[f]$ and $\langle \square[\partial f], \square[\partial f] \rangle$

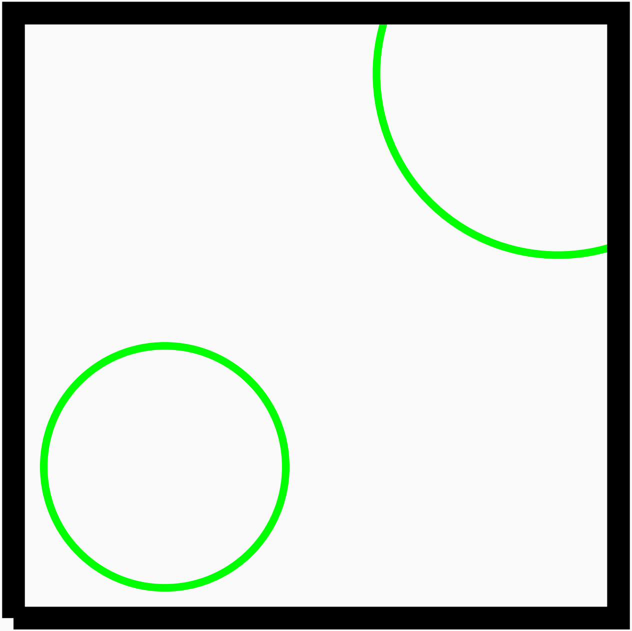
SUBDIVISION:
Starting with the trivial subdivision $\mathcal{S} := \{[-a, a]^n\}$, repeatedly subdivide each $J \in \mathcal{S}$ into 4 squares until for all $J \in \mathcal{S}$,

$$0 \notin \square f(J) \text{ or } 0 \notin \square \partial f(J)$$

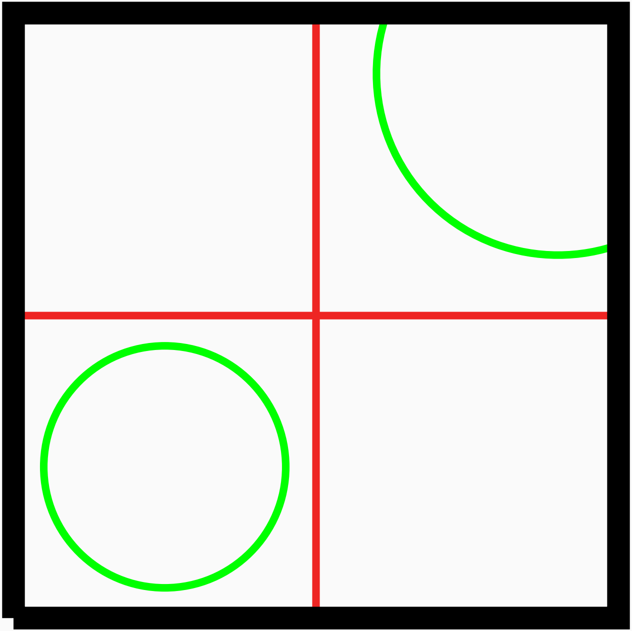
CONSTRUCTION:
Construct piecewise-linear curve L
joining the midpoints of "small" edges of each $J \in \mathcal{S}$ with oposite f-signs at their vertices

Output: Piecewise-linear approximation L of $C = f^{-1}(0) \cap [-a, a]^2$ isotopic to it

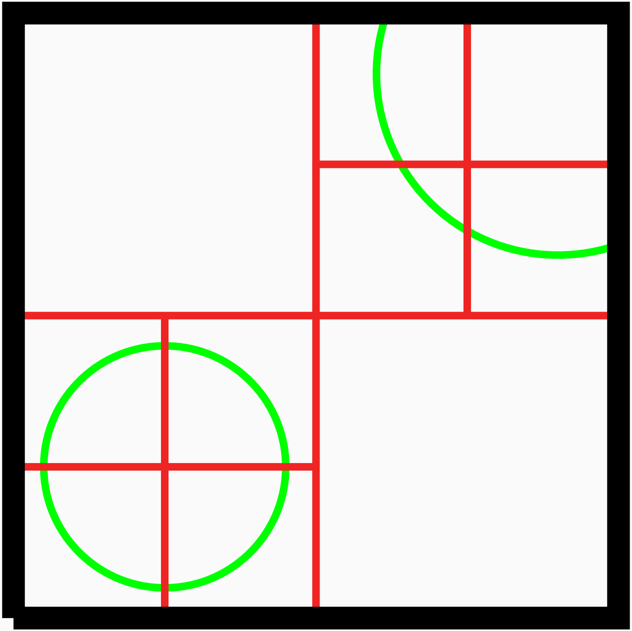
Plantinga-Vegter algorithm for curves II: Example



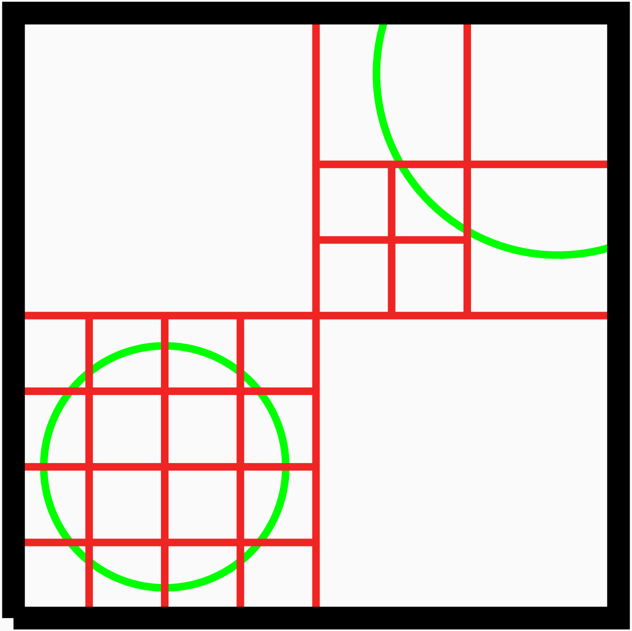
Plantinga-Vegter algorithm for curves II: Example



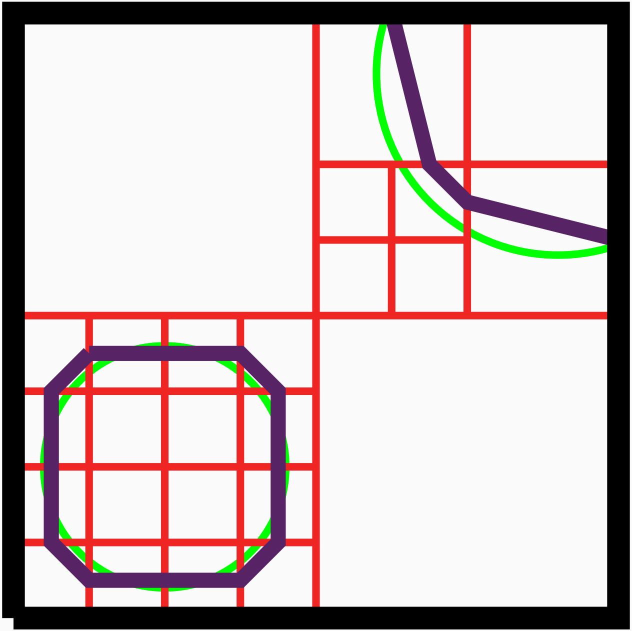
Plantinga-Vegter algorithm for curves II: Example



Plantinga-Vegter algorithm for curves II: Example



Plantinga-Vegter algorithm for curves II: Example



Plantinga-Vegter algorithm in higher dimensions

1. Plantinga-Vegter algorithm can be generalized to produce isotopic approximations of surfaces (Plantinga, Vegter; 2004)
   This is really why is called Plantinga-Vegter!
   Very efficient in practice
2. The subdivision method can be generalized to higher dimensions (Burr, Gao, Tsigaridas; ISSAC2017)
   but no construction of the piecewise-linear approximation...

We will focus on the later, since...

complexity of the algorithm is mainly that of the subdivision part

We will mainly count the number of subdivisions, because...

$$\mathrm{cost}(\text{subdivision algorithm}) \sim \#\,(\text{subdivisions}) \cdot \mathrm{cost}(\text{evaluations})$$

Subdivision in Plantinga-Vegter algorithm

Algorithm: Subdivision of PV Algorithm (Burr, Gao, Tsigaridas; ISSAC2017)

Input: $a \in (0, \infty)$, $f : \mathbb{R}^n \to \mathbb{R}$
with interval approximations $\square[hf]$ and $\square[h'\partial f]$
for some functions $h, h' : \mathbb{R}^n \to (0, \infty)$

Starting with the trivial subdivision $\mathcal{S} := \{[-a, a]^n\}$, repeatedly subdivide each $J \in \mathcal{S}$ into 2^n cubes until the condition

$$C_f(J) : 0 \notin \square[hf](J) \text{ or } 0 \notin \langle \square[h'\partial f], \square[h'\partial f] \rangle$$

holds for all $J \in \mathcal{S}$

Output: Subdivision $\mathcal{S} \subseteq \mathcal{I}_n$ of $[-a, a]^n$
such that for all $J \in \mathcal{S}$, $C_f(J)$ is true

h, h' depend on the setting and the interval arithmetic one uses

Can we understand the complexity of the subdivision of the PV algorithm for polynomials $f \in \mathcal{P}_{n,d}$ in terms of the number of variables n and the degree d?

Local size bounds and continuous amortization (Previous state-of-the-art)

Local size bound

Idea: Give a bound to the size of the smallest box containing a point not satisfying the condition.

Definition (Burr, Gao, Tsigaridas; ISSAC2017)

A *local size bound* for f is a function $b_f : \mathbb{R}^n \to [0, \infty)$ such that for all $x \in \mathbb{R}^n$,

$$b_f(x) \leq \inf\{\mathrm{vol}\,(J) \mid x \in J \in \mathcal{I}_n \text{ and } C_f(J) \text{ false}\},$$

where $C_f(J)$: $0 \notin \square[hf](J)$ or $0 \notin \langle \square[h'\partial f], \square[h'\partial f] \rangle$

1st bound in terms of the local size bound

Idea: Boxes cover the cube. The volume of the boxes should add to that of the cube. Volume of the boxes is at least number of boxes times volume of smallest box.

Proposition (Burr, Gao, Tsigaridas; ISSAC2017)

The number of n-cubes of the final subdivision of the subdivision of the PV *algoirhtm on input* (f, a), *regardless of how the subdivision step is done, is at most*

$$(2a)^n / \inf\{b_f(x) \mid x \in [-a, a]^n\} \qquad \square$$

where b_f *is a local size bound for f.*

(Burr, Gao, Tsigaridas; ISSAC2017) construct a local size bound for an integer polynomial $f \in \mathcal{P}_{n,d}$ and obtain the bound

$$2^{\mathcal{O}(nd^{n+1}(n\tau + nd \log(nd) + 9n + d) \log a)}$$

where τ is the bit-size of the coefficients.

2nd bound in terms of the local size bound

Refinement of the idea: Not all boxes have the same size. "$\frac{\mathrm{d}x}{b_f(x)}$ is, up to constant, the infinitesimal number of boxes needed to cover x."
Formalized by (Burr; 2016) using the technique known as *continuous amortization* introduced in (Burr, Krahmer, Yap; 2009)

Theorem (Burr, Gao, Tsigaridas; ISSAC2017)
The number of n-cubes of the final subdivision of the PV *algorithm on input* (f, a) *is at most*

$$\max\left\{1, \int_{[-a,a]^n} \frac{2^n}{b_f(x)}\,\mathrm{d}x\right\}$$

where b_f *is a local size bound for* f*. Moreover, the bound is finite if and only if the algorithm terminates.* □

Techniques of (Burr, Gao, Tsigaridas; ISSAC2017) cannot exploit it!

Can we exploit it?

(Burr, Gao, Tsigaridas; ISSAC2017) said...

> *Even though our bounds are optimal, in practice, these are quite pessimistic [...]*

and

> *Since the complexity of the algorithm can be exponential in the inputs [size], the integral must be described in terms of additional geometric and intrinsic parameters.*

What can be a solution to these issues?

CONDITION NUMBERS*

*as described in the book *Condition* (Bürgisser, Cucker; 2013)

Condition-based complexity

Local condition number

Definition (Cucker)

Given an homogeneous polynomial $F \in \mathcal{H}_{n,d}$, the *local condition number* of F at $y \in \mathbb{S}^n$ is

$$\kappa(F, y) := \frac{\|F\|}{\sqrt{F(y)^2 + \|\partial_y F_{|\mathrm{T}_y \mathbb{S}^n}\|^2 / d}}$$

where $\|F\|$ is the Weyl norm of F.
Given $f \in \mathcal{P}_{n,d}$, the *local affine condition number* of f at $x \in \mathbb{R}^n$ is

$$\kappa_{\mathrm{aff}}(f, x) := \kappa(f^{\mathrm{h}}, \phi(x))$$

where f^{h} is the homogeneization of f and $\phi : x \mapsto \frac{1}{\sqrt{1+\|x\|^2}} \begin{pmatrix} 1 \\ x \end{pmatrix}$.

Unfortunately the theory is developed for the homogenous setting.
Too many translations!

Geometry of the local condition number

Observation

$\kappa_{\mathrm{aff}}(f,x)=\infty$ iff $V_{\mathbb{R}}(f)$ has a singularity at x

Ergo $\kappa_{\mathrm{aff}}(f,x)$ says how near $V_{\mathbb{R}}(f)$ is of having a singularity at x.
Can we be more concrete? YES!

Theorem (Condition Number Theorem)

Let $x \in \mathbb{R}^n$ and

$$\Sigma_x := \{g \in \mathcal{P}_{n,d} \mid g(x) = \partial g(x) = 0\},$$

i.e., Σ_x is the set of hypersurfaces having x as a singular point. Then for every $f \in \mathcal{P}_{n,d}$,

$$\frac{\|f\|}{\kappa_{\mathrm{aff}}(f,x)} = \mathrm{dist}(f, \Sigma_x)$$

where $\|\cdot\|$ is the Weyl norm of $\mathcal{P}_{n,d}$ and the distance is the induced by the Weyl norm of $\mathcal{P}_{n,d}$. □

The wanted "additional geometric and intrinsic parameter"

Condition-based local size bounds

Theorem (Cucker, Ergür, T.-C.; ISSAC2019)
Let $f \in \mathcal{P}_{n,d}$. Then

$$x \mapsto 1/\left(2^{5/2} dn \kappa_{\mathrm{aff}}(f, x)\right)^n$$

is a local size bound for f with the interval approximation of (Cucker, Ergür, T.-C.; ISSAC2019), and

$$x \mapsto 1/\left(2^{3n} d^2 \kappa_{\mathrm{aff}}(f, x)\right)^n$$

with the interval approximation of Remark 2.2. of (Burr, Tsigaridas, Yap; ISSAC2017) □

Condition-based cost

Theorem (Cucker, Ergür, T.-C.; ISSAC2019)
The number of n-cubes in the final subdivision of the subdision of the PV *algorithm on input* (f, a) *is at most*

$$d^n \max\{1, a^n\} 2^{n \log n + 9n/2} \mathbb{E}_{x \in [-a,a]^n} \left(\kappa_{\mathrm{aff}}(f, x)^n \right)$$

if the interval approximation is that of (Cucker, Ergür, T.-C.; ISSAC2019), and at most

$$d^{2n} \max\{1, a^n\} 2^{3n^2 + 2n} \mathbb{E}_{x \in [-a,a]^n} \left(\kappa_{\mathrm{aff}}(f, x)^n \right)$$

if the interval approximation is that of (Burr, Tsigaridas, Yap; ISSAC2017). □

Observe that the complexity depends on $\mathbb{E}_{x \in [-a,a]^n} \left(\kappa_{\mathrm{aff}}(f, x)^n \right)$ which varies with f. This is why the name condition-based complexity

Probabilistic complexity

Uniform complexity analyses

Worst-case complexity analysis:

What is the worst running time?

Average complexity analysis:

What is the expectation of the running time on a random input?

Smoothed complexity analysis: (Spielman, Teng; 2002)

What is the worst running time after perturbing the input with a random perturbation (with weight σ)?

Smoothed lies between worst-case and average complexity

- $\sigma \to 0$: We recover worst-case complexity
- $\sigma \to \infty$: We recover average analysis

What do we mean by random? I

Definition (Cucker, Ergür, T.C.; ISSAC2019)
A *dobro random polynomial* $f \in \mathcal{H}_{n,d}$ with parameters K and ρ is a polynomial

$$f := \sum_{|\alpha|=d} \binom{d}{\alpha}^{1/2} c_\alpha X^\alpha$$

such that the c_α are independent random variables such that

- P1 $\mathbb{E}c_\alpha = 0$ (centered),
- P2 $(\mathbb{E}|c_\alpha|^p)^{\frac{1}{p}} \leq K\sqrt{p}$ for $p \geq 1$ (subgaussian with Ψ_2-norm $\leq K$), and
- P3 $\max_{u\in\mathbb{R}}\{\mathbb{P}(|c_\alpha - u| \leq \varepsilon)\}$ (anti-concentration with constant ρ).

A *dobro random polynomial* $f \in \mathcal{P}_{n,d}$ is a polynomial f such that its homogenization f^{h} is so.

What do we mean by random? II

Examples of **dobro random polynomials**:

N *KSS random polynomial*: (KSS=Kostlan-Smale-Shub)

- c_α is Gaussian with unit variance
- $K\rho = 1/\sqrt{2\pi}$

U *Weyl random polynomial*:

- c_α is uniform distribution in $[-1, 1]$
- $K\rho \leq 1$

E A *p-random polynomial*:

- c_α has density function $\delta_p e^{-|t|^p}$ where δ_p being the appropriate constant and $p \geq 2$

Note: Dobro (добро) is a Russian adjective derived from добрый (dóbryj) which means kind, good, genial, gentle, soft, etc.

Our result I: Average

Theorem (Cucker, Ergür, T.C.; ISSAC2019)

Let $f \in \mathcal{P}_{n,d}$ be a dobro random polynomial with parameters K and ρ. The expected number of n-cubes in the final subdivision of the PV *algorithm on input (f, a) is at most*

$$d^{\frac{n^2+3n}{2}} \max\{1, a^n\} 2^{\frac{n^2+16n\log(n)}{2}} (c_1 c_2 K \rho)^{n+1}$$

if the interval approximations is as in (Cucker, Ergür, T.C.; ISSAC2019) and

$$d^{\frac{n^2+5n}{2}} \max\{1, a^n\} 2^{\frac{7n^2+9n\log(n)}{2}} (c_1 c_2 K \rho)^{n+1}$$

if the interval approximation is as in (Burr, Gao, Tsigaridas; ISSAC2017). c_1 and c_2 universal constants. □

With an improvement of our condition-based techniques we can eliminate the n^2 of the exponent!

Our result II: Smoothed

Theorem (Cucker, Ergür, T.C.; ISSAC2019)

Let $f \in \mathcal{P}_{n,d}$, $\sigma > 0$, and $g \in \mathcal{P}_{n,d}$ a dobro random polynomial with parameters K and ρ . Then the expected number of n-cubes of the final subdivision of the PV *algorithm for input (q_σ, a) where $q_\sigma = f + \sigma\|f\|g$ is at most*

$$d^{\frac{n^2+3n}{2}} \max\{1, a^n\} 2^{\frac{n^2+16n\log(n)}{2}} (c_1 c_2 K\rho)^{n+1} \left(1 + \frac{1}{\sigma}\right)^{n+1}$$

if the interval approximations (Cucker, Ergür, T.C.; ISSAC2019) and

$$d^{\frac{n^2+5n}{2}} \max\{1, a^n\} 2^{\frac{7n^2+9n\log(n)}{2}} (c_1 c_2 K\rho)^{n+1} \left(1 + \frac{1}{\sigma}\right)^{n+1}$$

if the interval approximation is as in (Burr, Gao, Tsigaridas; ISSAC2017). c_1 and c_2 are universal constants. □

With an improvement of our condition-based techniques we can eliminate the n^2 of the exponent!

Our result III: Back to curves

$$\mathcal{O}\left(d^5\right) \text{ and } \mathcal{O}\left(d^6\right)$$

with the interval arithmetic of, respectively, (Cucker, Ergür, T.C.; ISSAC2019) and (Burr, Gao, Tsigaridas; ISSAC2017)

Our result III: Back to curves

$$\mathcal{O}\left(d^{5/2}\log^{5/2} d\right) \text{ and } \mathcal{O}\left(d^{3}\log^{3} d\right)$$

with the interval arithmetic of, respectively, (Cucker, Ergür, T.C.; ISSAC2019) and (Burr, Gao, Tsigaridas; ISSAC2017)

With the new techniques!

Compare to bit-complexity

$$\tilde{\mathcal{O}}(d^6\tau + d^7)$$

of the deterministic algorithm of (Diatta, Rouillier, Roy; ISSAC2014)

Some comments:

- Difference should be careful, complexity measured in different ways! But still, why so efficient?!
- PV algorithm does not work with singular curves, although there is work in this direction (Burr, Choi, Galehouse, Yap; 2012)
- Can we develop an hybrid approach working on all inputs with the same worst case, but faster on average?

Bere arretagatik eskerrik asko!
谢谢大家

Galderak?
以下是提问时间

Idea for the condition-based bound

Proposition

Let $f \in \mathcal{P}_{n,d}$ and $x \in \mathbb{R}^n$. Then either

$$|\widehat{f}(x)| > \frac{1}{2\sqrt{2d}\,\kappa_{\mathrm{aff}}(f,x)} \text{ or } \left\|\widehat{\partial f}(x)\right\| > \frac{1}{2\sqrt{2d}\,\kappa_{\mathrm{aff}}(f,x)}.$$

Above, $\widehat{f}$ is given by

$$\widehat{f} : x \mapsto f(x) / \left(\|f\| (1 + \|x\|^2)^{(d-1)/2} \right)$$

and $\widehat{\partial f}$ by

$$\widehat{\partial f} : x \mapsto \partial f(x) / \left(d \|f\| (1 + \|x\|^2)^{d/2-1} \right)$$

being both functions $(1 + \sqrt{d})$-Lipschitz.

Ideas for the probabilistic bound

1. By Fubini-Tonelli, to bound
$$\mathbb{E}_f \mathbb{E}_{x\in[-a,a]^n} \kappa_{\mathrm{aff}}(f,x)^n,$$
it is enough to get tail bounds of $\kappa_{\mathrm{aff}}(f,x)$ with respect to f independently of x
2. Since Σ_x is a linear subspace, we can write
$$\kappa_{\mathrm{aff}}(f,x) = \frac{\|f\|}{\|P_x f\|}$$
where $P_x : \mathcal{P}_{n,d} \to \Sigma_x^\perp$ is an orthogonal projection with respect the Weyl inner product.
3. Probabilistic steps for average analysis:
    3.1 $\kappa_{\mathrm{aff}}(f,x)$ large implies either $\|f\|$ large or $\|P_x f\|$ small
    3.2 "$\|f\|$ large" controlled by subgaussian property
    3.3 "$\|P_x f\|$ small" controlled by anticoncentration and independence
4. Similar for smoothed analysis