Found. Comput. Math. 257–311 (2005)

DOI: 10.1007/s10208-004-0144-z

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

On Location and Approximation of Clusters of Zeros of Analytic Functions

M. Giusti,[1] G. Lecerf,[2] B. Salvy,[3] and J.-C. Yakoubsohn[4]

[1]Laboratoire STIX
École polytechnique
91128 Palaiseau, France
Marc.Giusti@polytechnique.fr

[2]Laboratoire de Mathématiques
Université de Versailles Saint-Quentin-en-Yvelines
45 avenue des États-Unis
78035 Versailles, France
Gregoire.Lecerf@math.uvsq.fr

[3]Projet ALGO
INRIA Rocquencourt
78153 Le Chesnay, France
Bruno.Salvy@inria.fr

[4]Laboratoire MIP
Bureau 131
Université Paul Sabatier
118 route de Narbonne
31062 Toulouse, France
yak@mip.ups-tlse.fr

Abstract. At the beginning of the 1980s, M. Shub and S. Smale developed a quantitative analysis of Newton's method for multivariate analytic maps. In particular, their α-theory gives an effective criterion that ensures safe convergence to a simple isolated zero. This criterion requires only information concerning the map at the

Date received: June 11, 2004. Final version received: February 2, 2005. Date accepted: February 10, 2005. Communicated by Michael Shub. Online publication: July 14, 2005.
AMS classification: Primary 65H05, Secondary 30B10.
Key words and phrases: α-Theory, Cluster approximation, Cluster location, Cluster of zeros, Newton's operator, Pellet's criterion, Rouché's theorem, Schröder's operator, Zeros of analytic functions.

initial point of the iteration. Generalizing this theory to multiple zeros and clusters of zeros is still a challenging problem. In this paper we focus on one complex variable function. We study general criteria for detecting clusters and analyze the convergence of Schröder's iteration to a cluster. In the case of a multiple root, it is well known that this convergence is quadratic. In the case of a cluster with positive diameter, the convergence is still quadratic provided the iteration is stopped sufficiently early. We propose a criterion for stopping this iteration at a distance from the cluster which is of the order of its diameter.

Contents

Introduction

If ζ is a *simple* zero of an analytic function f, then the iteration of the classical Newton operator

$$N(f;x) := x - \frac{f(x)}{f'(x)}$$

converges quadratically to ζ, provided the initial point is "sufficiently close" to it. A quantitative analysis of this convergence has been given by Shub and Smale [**51**], [**49**], [**52**], [**50**]. They relate the convergence of Newton's method to *point estimates*—estimates on f and its derivatives at a point. These results extend to the multivariate case and are often referred to as "Smale's α-theory."

In this paper, we generalize the α-theory in order to treat multiple zeros and *clusters* of zeros of analytic functions in the univariate case. In the case of a zero of multiplicity m, the convergence of Newton's operator is no longer quadratic, but one can use Schröder's modified Newton operator [**48**]:

$$N_m(f;x) := x - m\frac{f(x)}{f'(x)}$$

which has quadratic convergence. Another possibility is to apply Newton's operator to the $(m-1)$th derivative of f. Both methods are covered by our analysis of a *family* of Schröder operators:

$$N_{m-l}(f^{(l)}; x) := x - (m - l)\frac{f^{(l)}(x)}{f^{(l+1)}(x)}, \qquad 0 \le l \le m - 1. \tag{1}$$

A cluster of zeros only means a set of zeros. Informally speaking, we use this term to refer to a set of zeros whose diameter is small compared to the distance to other zeros. In the context of numerical analysis, a natural request is to isolate and approximate clusters of zeros, simple or not. From a practical point of view, a cluster behaves like a multiple zero when seen from a distance. This is the basis of our method: we present an algorithm that treats the cluster as a multiple zero as long as the iterates are "far" from it, using a Schröder operator. We show that during this first stage, the iterates converge quadratically to the cluster. In the case of a multiple zero, convergence is quadratic to the zero. In the case of a cluster with positive diameter, the termination of our algorithm is given by a criterion detecting that the vicinity of the cluster has been reached. We show that the algorithm stops at a distance from the cluster which is of the order of its diameter.

Preliminaries

We start by setting the main notation and conventions used throughout this text. Then we briefly recall the α-theory before presenting our main results.

Definitions and Conventions. We denote by $\mathbb{R}$ the field of real numbers, by $\mathbb{C}$ the field of complex numbers, and by $\imath \in \mathbb{C}$ the square root of -1 with positive imaginary part. For any $z \in \mathbb{C}$, $|z|$ denotes the modulus of z. For any $\zeta \in \mathbb{C}$ and $r \ge 0$ being a real number, we use the following notation for balls, $B(\zeta, r) := \{x \in \mathbb{C}\colon |x - \zeta| < r\}$ denotes an open ball and $\bar{B}(\zeta, r) := \{x \in \mathbb{C}\colon |x - \zeta| \le r\}$ denotes a closed ball. For any real number u and any integer $m \ge 1$ we introduce the family of auxiliary functions

$$\psi_m(u) := 2(1 - u)^{m+1} - 1.$$

We use $\psi(u) := \psi_1(u) = 1 - 4u + 2u^2$. For a compact subset Z of $\mathbb{C}$ the *diameter* of Z is the maximum distance between any two points of Z. We always count numbers of zeros with multiplicities.

We denote by $\mathbb{R}\{t\}$ the algebra of the real power series with positive radius of convergence. For convenience, we sometimes treat the elements of $\mathbb{R}\{t\}$ similar to their corresponding analytic functions defined on a neighborhood of 0.

If f is an analytic function, we make use of the following notation for the generating series of the absolute values of the derivatives of f at a point z in the region of analyticity of f:

$$[f]_z := \sum_{k \ge 0} |f^{(k)}(z)| \frac{t^k}{k!} \in \mathbb{R}\{t\}.$$

We consider the following partial order $\leq$ over $\mathbb{R}\{t\}$. Let F and G be in $\mathbb{R}\{t\}$, we write $F \leq G$ when $F^{(k)}(0) \leq G^{(k)}(0)$ for all $k \geq 0$. Then we say that a power series $F \in \mathbb{R}\{t\}$ is a *majorant series* for an analytic function f at a point z if $[f]_z \leq F$. For completeness, in Appendix A we give the main basic properties of majorant series.

Point estimates of a given function f at a given point z are quantities that only depend on the series $[f]_z$.

Convergence to Simple Zeros. We first recall the basic results of the α-theory. For precise results, for a complete historical presentation, and for the multivariate case, we refer to [**2**, Chap. 8].

Three important quantities are defined in this analysis: γ, β, and α. The first one, namely $\gamma(f; z)$, helps control the function locally:

$$\gamma(f; z) := \sup_{k \geq 2} \left| \frac{f^{(k)}(z)}{k!\ f'(z)} \right|^{1/(k-1)}.$$

In particular, the radius of convergence of the power series expansion of f at z is at least $1/\gamma$. The second quantity is the length of the iteration step, $\beta(f; z) := \left| f(z)/f'(z) \right|$; the third quantity is their product, $\alpha(f; z) := \beta(f; z)\gamma(f; z)$.

Most of the proofs handling these quantities hide *geometric majorant series techniques*, these are majorant series whose sequence of coefficients forms a geometric progression. In particular, $\gamma(f; z)$ can be defined in terms of the *minimal* geometric majorant series of $(f - f(z))/f'(z)$ at z of the form:

$$\left[\frac{f - f(z)}{f'(z)} \right]_z \leq \frac{t}{1 - \gamma(f; z)t}.$$

Roughly speaking, the α-theory provides two types of theorem. The so-called *γ-theorems* [**2**, Chap. 8, Theorem 1] show that Newton's method converges quadratically within a ball centered at ζ, whose diameter is universally (with respect to f) proportional to $1/\gamma(f; \zeta)$. In particular, this also provides lower bounds for the distance between simple zeros [**9**]. As to the so-called *α-theorems* [**2**, Chap. 8, Theorem 2], which have given their name to the α-theory, they are more relevant to practical concerns: they show that Newton's method with an initial point x_0 converges quadratically, provided $\alpha(f; x_0)$ is sufficiently small. Moreover, the distance from x_0 to the zero is then bounded by $\beta(f; x_0)$ times a universal constant. The *optimal* constants are due to Wang and Han [**57**].

Our Contributions

Overview. In this paper we extend the α-theory in order to obtain estimates for clusters of zeros of analytic functions. The text is organized around three central problems: cluster location, bounds on the diameter of clusters, and cluster approximation. For any pair of integers $m \geq 1$ and $l \in \{0, \ldots, m-1\}$, the following central characters, introduced in [**58**], are natural generalizations of the quantities α, β, γ. If $f^{(m)}(z) \neq 0$, we define them as

$$\begin{aligned}
\gamma_m(f; z) &:= \sup_{k \geq m+1} \left(\frac{m!\, |f^{(k)}(z)|}{k!\, |f^{(m)}(z)|} \right)^{1/(k-m)}, \\
\beta_{m,l}(f; z) &:= \sup_{l \leq k \leq m-1} \left(\frac{m!\, |f^{(k)}(z)|}{k!\, |f^{(m)}(z)|} \right)^{1/(m-k)}, \\
\alpha_{m,l}(f; z) &:= \gamma_m(f; z) \beta_{m,l}(f; z). \qquad (2)
\end{aligned}$$

Briefly, we also write $\beta_m := \beta_{m,0}$ and $\alpha_m := \alpha_{m,0}$. In other words, we have at our disposal a straightforward majorant series $F_m(f, z; t)$ for

$$\left[\frac{m!\, f}{f^{(m)}(z)} \right]_z \leq F_m(f, z; t) := \sum_{j=0}^{m-1} \beta_m^{m-j}(f; z) t^j + \frac{t^m}{1 - \gamma_m(f; z) t}, \qquad (3)$$

and therefore f is analytic in $B(z, 1/\gamma_m(f; z))$, or can be continued analytically there. When $\gamma_m(f; z) = 0$ (that is, f is a polynomial of degree m) it is convenient to stipulate that $1/\gamma_m(f; z)$ is ∞.

The paper is structured as follows. We start with cluster location (Section 1) and find a lower bound on the diameter of the cluster (Section 2). Convergence of the Schröder operators to the cluster is analyzed in Section 3 in terms of estimates at the cluster. In Section 4 we turn this analysis into an algorithm that needs only the estimates at the initial point. In Section 5 we explain how to compute approximations of theses estimates. Finally, in Section 6, we report on numerical experiments on families of exponential polynomials.

Inequalities. Majorant series are a convenient tool to handle point estimates. Appendix A provides a useful toolbox for computing with majorant series, we shall refer to it several times in proofs. This toolbox is also intended to be used in practice in order to compute approximations of γ_m, as illustrated in Section 5.

Most of our inequalities generalize some classical ones of the α-theory. For instance, Proposition 4.3 generalizes [**2**, Chap. 8, Prop. 3] when $m > 1$. In most cases, inequalities are first proved in terms of majorant series.

In designing our algorithms, we allow the user to specify a function $\mathcal{B}_{m,l}(f, x_0; x_1)$ that returns a (possibly rough) numerical approximation of $\beta_{m,l}(f; x_1)$, with the possible use of information located at x_0 (see the definition in Section 4.2). In Section 5, we provide two practical cases: the first one uses computations with power series expansions, while the second one is purely numerical and relies on the well-known discrete Fourier transform interpolation scheme.

Cluster Location. In 1881, Pellet [**39**] gave a simple location criterion: let f, F, and z be such that $\left[m!\, f/f^{(m)}(z)\right]_z \leq F$ and let $r > 0$ be a real number smaller than the radius of convergence of F. If

$$\frac{F(r)}{r^m} - \frac{F^{(m)}(0)}{m!} < 1,$$

then f has m zeros in $\bar{B}(z, r)$, counted with multiplicities. In Section 1.1, we recall a proof of this criterion based on Rouché's theorem [**45**]. Then we extend this method to deduce information from $\alpha_{m,l}(f; z)$ on the location of zeros of the derivatives $f^{(l)}$ of the analytic function f around z. If $\alpha_{m,l}$ is less than a universal (with respect to f) constant, then we determine two balls centered at z and containing the cluster: a smaller one, of radius universally proportional to $\beta_{m,l}(f; z)$ and a larger one of radius universally proportional to $1/\gamma_m(f; z)$.

Cluster Diameter. From the location criterion at a point ζ in the convex hull of a cluster of m zeros of f (counting multiplicities), we deduce an upper bound on the diameter of this cluster in terms of $\beta_m(f; \zeta)$. In Section 2, we provide the converse bound: we give a quantitative formula bounding $\beta_m(f; \zeta)$ in terms of the diameter of the cluster. When speaking informally, we will treat the diameter of the cluster and $\beta_m(f; \zeta)$ similarly.

On occasion, we also say informally that a point lies *far* or *close* to a cluster, bearing in mind the implicit *scale* given by the diameter of the cluster.

Cluster Approximation. Informally speaking, if $f^{(l)}$ admits a cluster of $m - l$ zeros, and as long as $\beta_{m,l}$ at the current iterate is larger than the diameter of the cluster, then the cluster behaves like a multiple zero. Therefore the Schröder iteration (1) converges quadratically to the cluster. Then, in the case of a cluster with positive diameter, when arriving close to the cluster, it is well known that the iteration may behave in a chaotic way.

The following basic situation exemplifies this difficulty. Consider the analytic map $f\colon x \mapsto x^2 - \varepsilon^2$ with $\varepsilon \neq 0$. Here f has two zeros, namely $-\varepsilon$ and ε, the diameter of this cluster is $2|\varepsilon|$ and $\beta_2(f; 0) = |\varepsilon|$. The use of the corrected Newton iteration

$$x_1 := N_2(f; x_0) = x_0 - 2\frac{f(x_0)}{f'(x_0)} = \frac{\varepsilon^2}{x_0},$$

in order to approximate this cluster, leads to the following discussion. If $|x_0|$ is very large compared to $|\varepsilon|$, then x_1 is very close to the cluster; if $|x_0|$ is very small compared to $|\varepsilon|$, then x_1 is very far from the cluster, and, finally, if $|x_0|$ is about the same as $|\varepsilon|$, then $|x_1|$ is also about the same as $|\varepsilon|$.

This is the main difficulty to be overcome in the general case: we detect when the Schröder iterates are well defined and stop the iteration once it has arrived very close to the cluster, i.e., at a distance which is of the order of the diameter of the cluster. We refer to the combination of Schröder's operator with our stopping criterion as an *approximation algorithm*. This algorithm is presented in Section 4.

Special Cases. In general, it is recommended using formula (1) with $l = 0$: this way no high-order derivative needs to be computed. In this paper we show that rough approximations of $\beta_{m,l}$, namely $\mathcal{B}_{m,l}(f, x_0; x_1)$ and $\mathcal{B}_{m,l}(f, x_1; x_0)$, are sufficient (here x_0 and x_1 are two consecutive iterates). In particular, using the method given in Section 5, the computation of $\mathcal{B}_{m,l}(f, x_0; x_1)$ boils down to computing a polynomial of degree at most $2m - 1$ that interpolates f at $2m$ points equidistributed on the circle of center x_1 and radius $|x_0 - x_1|$ (the same holds for $\mathcal{B}_{m,l}(f, x_1; x_0)$). This is a serious advantage over Newton's iteration on $f^{(m-1)}$.

When using our algorithm with $l = m - 1$, the convergence to a simple zero of $f^{(m-1)}$ is quadratic. However, instead of iterating toward this zero, our stopping criterion allows us to stop the iteration as soon as the iterates are close to the cluster of zeros of f. This shows another advantage of our unified presentation for any $l \in \{0, \ldots, m - 1\}$.

Related Works

Location and approximation of roots of polynomials are classical subjects in numerical analysis. Some general references are [**26**], [**40**]. An extended bibliography is collected in [**27**], [**28**], and a recent survey on root location can be found in [**29**].

In contrast to polynomials, few algorithms are known for locating and approximating clusters of zeros of analytic functions. Yet such clusters naturally arise in many theoretical and practical situations. The results of this paper are used in [**12**], which deals with the location and approximation of special types of clusters of multivariate analytic maps: even when starting with polynomial maps, the algorithm of [**12**] needs to compute with functions that are implicitly defined, hence generally not polynomial.

Cluster Location. Our analysis of Pellet's criterion in Section 1 follows Yakoubsohn's approach via Rouché's theorem [**58**], [**59**]. We slightly improve the criteria of [**58**], [**53**], [**59**], and generalize them to clusters of roots of derivatives. Other cluster location algorithms have been proposed in the analytic case. For instance, the algorithms of [**22**], [**24**], [**23**] rely on numerical path integration: they are more powerful albeit more expensive. In [**46**], Pellet's criterion is compared to nine other location methods based on several families of polynomials.

Cluster Approximation. The quadratic convergence of Newton's iteration ceases to hold in the presence of multiple zeros. Instead, the convergence becomes linear and a large amount of works focus on this problem, including [**48**], [**33**], [**41**], [**42**], [**43**], [**13**], [**6**], [**7**], [**56**], [**5**], [**15**], [**60**], [**14**], [**8**], [**58**] (some of them also deal with the more complicated multivariate case). In order to reestablish quadratic convergence, Schröder [**48**] introduced his *corrected Newton operator*. The correction requires prior knowledge of the multiplicity. This multiplicity may also be approximated dynamically at the price of slowing down the convergence [**54**], [**18**], [**55**], [**21**], [**58**]. Higher-order operators have also been adapted to multiple zeros [**11**]. In this paper we deal with clusters, not only with multiples zeros. We assume that the multiplicity is known in advance. Our theoretical γ-analysis of the convergence of Schröder's operator to a cluster (Section 3) mainly follows [**58**], where a similar analysis is performed for Newton's operator.

Applications to Polynomial Root Finding. Besides our primary interest in the several variables case [**12**], we now discuss the potential applications of our methods in the field of univariate polynomial root finders. As observed in [**1**], [**37**], univariate polynomials produced by eliminating variables in multivariate polynomial systems (e.g., by means of Gröbner basis computation) are often huge and "ill-conditioned." Thus, clusters and even sequences of nested clusters are not a rare phenomenon in practice. These situations require special attention in order to avoid a precision blow-up in the computations. In the following paragraphs we briefly discuss the main known strategies to handle clusters of roots of polynomials. A detailed survey on root-finder algorithms can be found in [**36**].

On one hand, nearly optimal root finders have been well established for a decade [**34**], [**35**], [**30**], [**31**], [**20**], [**38**], following earlier ideas by Schönhage [**47**]. The fastest algorithms from the theoretical point of view, namely [**31**], [**38**], are based on balanced splittings and make use of the nontrivial generalizations of the Grace–Heawood theorem given in [**4**], which can be seen as a complexification of Rolle's theorem. More precisely, if $k + 1$ roots of a polynomial p of degree n are contained in a ball of radius ρ, then, for any $\ell \leq k$, at least $k + 1 - \ell$ roots of $p^{(\ell)}$ lie in a ball of radius in $\mathcal{O}(\rho)$, centered at the average of the $k + 1$ roots. Our location results of Section 1 yield a similar result starting from an α estimate. For "ill-conditioned" polynomials, balanced splitting requires increased precision in the computations (see discussions in [**36**, Sect. 7]). Recently, Pan has proposed an algorithm to compute splittings that are not necessarily balanced but favor clusters [**38**]: his algorithm is still nearly optimal in the worst case, but also optimal for the "well-conditioned" case. His construction relies on Turan's theorem for approximating the distances of a given point to the roots of the polynomial. All these techniques deeply exploit nontrivial results for polynomials, which are less connected to the techniques presented in this paper.

On the other hand, other kinds of polynomial root finders have been designed in order to exploit "ill-conditioned" situations. They are theoretically slower than the previous methods but they are often efficient in practice. In [**44**], Renegar speeds up Weyl's quad-tree construction (combined to the Schur–Cohn algorithm) thanks to the α-theory for simple roots. Clusters are treated by relating them to clusters of derivatives. This approach is closer to ours than the one of [**4**], since it only needs to consider clusters that are far from the other roots. In [**37**], Pan profoundly revisits Renegar's strategy: clusters are treated very efficiently by means of Schröder's operator combined to a stopping criterion relying on Turan's theorem. These techniques do not readily extend to the analytic case, where our main contribution lies. In [**1**], Bini and Fiorentino describe their implementation based on many heuristics. Their program is well suited to sparse polynomials and takes advantage of "ill-conditioned" situations. The precision in the computations is only increased when necessary. When dealing with a well-separated cluster of m zeros they propose to combine the main result of [**4**] to Newton's iteration on the $(m-1)$st derivative. Ultimately, computations are verified by Gerschgorin's theorem (see [**32**] for a recent improvement of this theorem). Concerning these algorithms [**44**], [**37**], [**1**], it is legitimate to ask if our methods could apply, could improve intrinsic complexities, and could extend them to approximate a finite set of the zeros of an analytic function in a bounded domain. Such a study has yet to be done.

Finally, in the vast literature on polynomial root finding and cluster detection, it is worth mentioning a few other important approaches proposed in [**16**], [**19**], [**17**], [**61**], that are less connected to our work.

1. Cluster Location

We present point estimate criteria for cluster location that are based on Rouché's theorem. The first and most general criterion relates the existence of a cluster of zeros of f to the sign of a certain algebraic expression built from majorant series evaluations. Then we deduce criteria in terms of β_m and γ_m. Finally, we generalize these criteria in order to locate zeros of the derivatives of f. In this section, f denotes an analytic function defined on an open subset $U \subseteq \mathbb{C}$.

1.1. *Location from Majorant Series*

We start with the most general location criterion in terms of majorant series. This is a consequence of Rouché's theorem.

Proposition 1.1. *Let $m \geq 0$ be an integer, let $z \in U$ be such that $f^{(m)}(z) \neq 0$ and let $F \in \mathbb{R}\{t\}$ be such that $|m!f/f^{(m)}(z)|_z \leq F$. Then, for any real number $r > 0$ smaller than the radius of convergence of F such that $\bar{B}(z,r) \subseteq U$ and*

$$\frac{F(r)}{r^m} - \frac{F^{(m)}(0)}{m!} < 1, \tag{4}$$

f has m zeros in $\bar{B}(z,r)$, counted with multiplicities.

Proof. We introduce the function $g\colon U \to \mathbb{C}$,

$$g(x) = f(x) - \sum_{j=0}^{m-1} \frac{f^{(j)}(z)}{j!}(x-z)^j.$$

Let w be such that $|w - z| = r$, we start with

$$\left|\frac{m!\,(f(w) - g(w))}{f^{(m)}(z)}\right| \leq \sum_{j=0}^{m-1} \frac{F^{(j)}(0)}{j!} r^j,$$

$$\left|\frac{m!\,g(w)}{f^{(m)}(z)}\right| \geq r^m - \sum_{j\geq m+1} \frac{F^{(j)}(0)}{j!} r^j.$$

It follows that (4) implies $|f(w) - g(w)| < |g(w)|$. In particular, $g(w)$ does not vanish and therefore Rouché's theorem asserts that f and g have the same number of zeros in $\bar{B}(z, r)$, counting multiplicities. In order to conclude the proof it remains to show that z is the only zero of g in this ball, with multiplicity m. This multiplicity of z is clear. Now let $w \in \bar{B}(z, r)$ and $w \neq z$, let $s := |w - z|$, we have

$$\frac{m!\,|g(w)|}{|f^{(m)}(z)|s^m} \geq 1 - \sum_{j\geq m+1} \frac{F^{(j)}(0)}{j!} s^{j-m} \geq 1 - \sum_{j\geq m+1} \frac{F^{(j)}(0)}{j!} r^{j-m} > 0,$$

where the last inequality follows from (4). □

Observe that, for $m = 0$, this proposition can be used to certify the absence of zeros of f in a given ball. Furthermore, for positive m, this criterion is sharp: let $a > 0$ denote a real number, with $f = a^m - x^m$, $F = a^m + t^m$, and $z = 0$, inequality (4) rewrites $a^m < r^m$.

The next lemma provides a reformulation of (4) showing that the set of possible values for r is an interval, under reasonable assumptions. The extremities of this interval correspond to the diameters of the largest annulus isolating the cluster in its inner disk. The main argument involved is convexity. We use Gantmacher's notation $x_{1:s}$ to denote the s-tuple $x_1, \ldots, x_s$ and $\mathbb{R}_{+*}$ represents the set of positive real numbers. In the next subsection, we will use this result in a particular case, in order to produce a location criterion for clusters of $f^{(l)}$, in terms of $\alpha_{m,l}$.

Lemma 1.2. *Let V denote a connected open set of $\mathbb{R}^s \times \mathbb{R}$ and let $F(x_{1:s}, t)$ denote a real valued analytic map on V. Let V_0 be the canonical projection of $V \cap (\mathbb{R}^s \times \{0\})$ to $\mathbb{R}^s$. Assume, for all $p_{1:s} \in V_0$:*

(a) $F(p_{1:s}, t) \geq 0$, *for the partial order on* $\mathbb{R}\{t\}$;
(b) $(\partial^m F/\partial t^m)(p_{1:s}, 0) \neq 0$;
(c) $\alpha_m(F(p_{1:s}, \cdot); 0) \neq 0$;
(d) $V \cap (\{p_{1:s}\} \times \mathbb{R}) = \{p_{1:s}\} \times (-\rho(p_{1:s}), \rho(p_{1:s}))$, *where* $\rho(p_{1:s})$ *denotes the radius of convergence of* $F(p_{1:s}, \cdot)$ *at* 0;
(e) $\lim_{r\to\rho(p_{1:s})} F(p_{1:s}, r) = +\infty$.

Then there exists a real valued analytic map $A(x_{1:s})$ defined on V_0 and two real valued analytic maps $r^-(x_{1:s}), r^+(x_{1:s})$ defined on $V_A := \{p_{1:s} \in V_0\colon A(p_{1:s}) < 1\}$

such that, for any $p_{1:s} \in V_0$, *the following equivalence holds*:

$$\frac{F(p_{1:s}, r)}{r^m} - \frac{(\partial^m F/\partial t^m)(p_{1:s}, 0)}{m!} < 1 \quad \textit{and} \quad 0 < r < \rho(p_{1:s}) \tag{5}$$
$$\text{if and only if} \quad A(p_{1:s}) < 1 \quad \textit{and} \quad r^-(p_{1:s}) < r < r^+(p_{1:s}).$$

Proof. We introduce $R(x_{1:s}, r) := F(x_{1:s}, r)/r^m - (\partial^m F/\partial t^m)(x_{1:s}, 0)/m!$, that is defined on $V^+ := V \cap (\mathbb{R}^s \times \mathbb{R}_{+*})$, so that the first inequality of (5) is equivalent to $R(p_{1:s}, r) < 1$. On V^+, from hypothesis (a) we deduce $(\partial^2 R/\partial r^2) \geq 0$ which becomes strict because of (c). Because of (c) again, $R(p_{1:s}, r)$ tends to infinity when r tends to 0 and with (e) the same holds when r tends to $\rho(p_{1:s})$. From (d), we deduce that, for any $p_{1:s} \in V_0$, the maximum analyticity domain of $R(p_{1:s}, \cdot)$ is the canonical projection of $V \cap (\{p_{1:s}\} \times \mathbb{R}_{+*})$ to $\mathbb{R}$ and that $R(p_{1:s}, r)$ admits a unique minimum at $r = r_m(p_{1:s})$, given by $(\partial R/\partial r)(p_{1:s}, r_m(p_{1:s})) = 0$. It follows that r_m is analytic on V_0. We introduce $A := R(x_{1:s}, r_m(x_{1:s}))$ and then inequality $A(p_{1:s}) < 1$ is equivalent to the existence of two values $r^-(p_{1:s})$ and $r^+(p_{1:s})$ such that (5) is satisfied with r if and only if $r^-(p_{1:s}) < r < r^+(p_{1:s})$. □

1.2. *Location from* $\alpha_{m,l}$

Our aim is now to derive a location criterion for clusters of $f^{(l)}$ from Proposition 1.1 that only depends on $\alpha_{m,l}$.

Inequality (3) can be generalized in order to treat the derivatives of f. For this purpose, we introduce

$$F_{m,l}(f, z; t) := \binom{m}{l}^{-1} \left(\sum_{j=l}^{m-1} \binom{j}{l} \beta_{m,l}(f; z)^{m-j} t^{j-l} + t^{m-l} \sum_{j \geq 0} \binom{m+j}{l} \gamma_m(f; z)^j t^j \right). \tag{6}$$

Of course, $F_{m,0}$ and F_m coincide and $F_{m,l}(f, z; t)$ satisfies

$$\frac{F_{m,l}^{(m-l)}(f, z; 0)}{(m-l)!} = 1.$$

The following lemma will also be used in Section 3:

Lemma 1.3. *According to the above notation*, *we have*

$$\left[\frac{(m-l)!\, f^{(l)}}{f^{(m)}(z)} \right]_z \leq F_{m,l}(f, z; t).$$

Proof. Thanks to Proposition A.9, the proof follows from a direct calculation through l derivations of the following inequality:

$$\left[\frac{m!\,f}{f^{(m)}(z)}\right]_z \le \sum_{j=0}^{l-1} \frac{m!\,|f^{(j)}(z)|}{j!\,|f^{(m)}(z)|} t^j + \sum_{j=l}^{m-1} \beta_{m,l}^{m-j}(f;z) t^j + \frac{t^m}{1-\gamma_m(f;z)t},$$

which is clearly established from definitions. □

Combining this lemma with Proposition 1.1 applied to $f^{(l)}$ and the series $F_{m,l}(f,z;t)$, we derive a first criterion, in terms of $\beta_{m,l}$ and γ_m.

Corollary 1.4. *Let $m \ge 1$ be an integer, let $z \in U$ be such that $f^{(m)}(z) \ne 0$, and let $l \in \{0,\dots,m-1\}$. Then, for any real number $r > 0$ such that $\gamma_m(f;z)r < 1$, $\bar{B}(z,r) \subseteq U$, and*

$$\frac{F_{m,l}(f,z;r)}{r^{m-l}} < 2, \tag{7}$$

$f^{(l)}$ has $m-l$ zeros in $\bar{B}(z,r)$, counting multiplicities.

We now apply Lemma 1.2 in the particular case of $F_{m,l}(f,z;t)$: the variables x_1 and x_2 will, respectively, represent $\beta_{m,l}(f;z)$ and $\gamma_m(f;z)$ and we take

$$F(x_1,x_2,t) := \binom{m}{l}^{-1} \left(\sum_{j=l}^{m-1} \binom{j}{l} x_1^{m-j} t^{j-l} + t^{m-l} \sum_{j\ge 0} \binom{m+j}{l} (x_2 t)^j\right)$$

and $V := \{(p_1,p_2,r) \in \mathbb{R}_{+*}^2 \times \mathbb{R}\colon\ p_2|r| < 1\}$, so that Lemma 1.3 reads

$$\left[\frac{(m-l)!\,f^{(l)}}{f^{(m)}(z)}\right]_z \le F(\beta_{m,l}(f;z), \gamma_m(f;z), t) = F_{m,l}(f,z;t).$$

Since $\rho(p_1,p_2) = 1/p_2$ and taking $m-l$ instead of m, conditions (a)–(e) of Lemma 1.2 are satisfied: let A, r^- and r^+ be the corresponding maps. We now describe these maps. First observe that $R(x_1,x_2)$, as defined in the proof of Lemma 1.2, can be expressed in terms of the variables $v := x_1 x_2$ and $u := x_2 r$,

$$R(x_1,x_2,r) = \binom{m}{l}^{-1} \left(\sum_{j=l}^{m-1} \binom{j}{l} (v/u)^{m-j} + \sum_{j\ge 1} \binom{m+j}{l} u^j\right) =: \tilde{R}(v,u).$$

From the definition of A we deduce

$$A(x_1,x_2) = \min_{0<r<\rho(x_1,x_2)} R(x_1,x_2,r) = \min_{0<u<1} \tilde{R}(v,u) =: \tilde{A}(v).$$

Let $u_m(v)$ denote the abscissa where the minimum of $\tilde{R}(v,\cdot)$ is attained. By construction, u_m is analytic and satisfies

$$\frac{\partial \tilde{R}}{\partial u}(v, u_m(v)) = 0,$$

from which we deduce

$$\begin{aligned}\tilde{A}'(v) &= \frac{d}{dv}(\tilde{R}(v, u_m(v))) = \frac{\partial \tilde{R}}{\partial v}(v, u_m(v)) + \frac{\partial \tilde{R}}{\partial u}(v, u_m(v))u_m'(v) \\ &= \frac{\partial \tilde{R}}{\partial v}(v, u_m(v)) > 0.\end{aligned}$$

Then it is crucial to observe that letting $u = \sqrt{v}$ proves that $\tilde{R}$ tends to 0 when v tends to zero, hence $\lim_{v\to 0} \tilde{A}(v) = 0$. On the other hand, the rough lower bound

$$\tilde{A}(v) \geq \frac{m-l}{m}\frac{v}{u_m(v)} \geq \frac{m-l}{m}v$$

implies that $\lim_{v\to+\infty} \tilde{A}(v) = +\infty$. Since $\tilde{A}$ is increasing, we deduce that there exists a smallest positive real number $\bar{\alpha}_{m,l}$ such that $\tilde{A}(\bar{\alpha}_{m,l}) = 1$. In particular, we get $V_A = \{(p_1, p_2)\colon\ 0 < p_1 p_2 < \bar{\alpha}_{m,l}\}$.

We conclude that Corollary 1.4 reformulates in terms of point estimates: if $0 < \alpha_{m,l}(f; z) < \bar{\alpha}_{m,l}$ and

$$r^-(\beta_{m,l}(f;z), \gamma_m(f;z)) < r < r^+(\beta_{m,l}(f;z), \gamma_m(f;z)),$$

then $f^{(l)}$ admits $m - l$ zeros in $\bar{B}(z, r)$.

We refer to the universal constant $\bar{\alpha}_{m,l}$ as the *critical value* for $\alpha_{m,l}$. One practical difficulty lies in obtaining sharp approximations of these critical values in terms of l and m. Next, we focus on expressing lower bounds on these critical values. In short, we write $\bar{\alpha}_m := \bar{\alpha}_{m,0}$.

1.3. *Lower Bounds on Critical Values* $\bar{\alpha}_m$

In order to determine an explicit lower bound on the critical values $\bar{\alpha}_{m,l}$, as previously defined, we first treat the case $l = 0$.

We assume $l = 0$ for the moment and let $r > 0$ be such that $\gamma_m r < 1$. First, we observe that (7) rewrites

$$\sum_{k=1}^{m}\left(\frac{\beta_m}{r}\right)^k + \frac{\gamma_m r}{1-\gamma_m r} < 1, \tag{8}$$

where we let $\alpha_m := \alpha_m(f; z)$, $\beta_m := \beta_m(f; z)$, and $\gamma_m := \gamma_m(f; z)$, for short. Then we write

$$\sum_{k=1}^{m}\left(\frac{\beta_m}{r}\right)^k + \frac{\gamma_m r}{1-\gamma_m r} = \frac{\beta_m/r - (\beta_m/r)^{m+1}}{1-\beta_m/r} + \frac{\gamma_m r}{1-\gamma_m r}, \tag{9}$$

so that if $\gamma_m > 0$, letting $u := \gamma_m r$, inequalities $\beta_m < r$, $u < 1$, and

$$\frac{\alpha_m}{u-\alpha_m}\left(1 - \left(\frac{\alpha_m}{u}\right)^m\right) + \frac{u}{1-u} < 1 \tag{10}$$

imply (7). If $m = 1$ the critical value $\bar{\alpha}_1$ we find this way is the first positive root of the discriminant (with respect to u) of $2u^2 - (1+\alpha_1)u + \alpha_1$, that is, $\bar{\alpha}_1 = 3 - 2\sqrt{2} > 0.1715$, also given in [**57**, Prop. 2]. Then, numerical approximations give $\bar{\alpha}_2 > 0.1225$, $\bar{\alpha}_3 > 0.1142$. For large m one can neglect the term $(\alpha_m/u)^m$, so that we can focus on the condition

$$\frac{\alpha_m}{u-\alpha_m} + \frac{u}{1-u} \leq 1, \tag{11}$$

which implies (10) if $\alpha_m \neq 0$. The latter inequality is equivalent to

$$P(u) := 2\alpha_m - (1+3\alpha_m)u + 2u^2 \leq 0,$$

which is itself equivalent to $\alpha_m \leq \frac{1}{9}$, and $r^- \leq r \leq r^+$, where

$$r^- := 2\beta_m \frac{2}{1+3\alpha_m+\sqrt{1-10\alpha_m+9\alpha_m^2}},$$

$$r^+ := \frac{1+3\alpha_m+\sqrt{1-10\alpha_m+9\alpha_m^2}}{4\gamma_m}.$$

Furthermore, it is easy to observe that $\bar{\alpha}_m > \frac{1}{9}$ for all m and that $\lim_{m\to+\infty} \bar{\alpha}_m = \frac{1}{9}$. It follows that $\frac{1}{9}$ is the best critical value uniform with respect to m that one can deduce from Proposition 1.1.

The next theorem summarizes this discussion and also includes the degenerate cases. Thus, similar results for polynomials (contained in the proof of [**59**, Lemma 6] and also [**53**, Theorem 11]) are generalized to analytic functions.

Theorem 1.5. *Let $m \geq 1$ be an integer and let $z \in U$ be such that $f^{(m)}(z) \neq 0$. In short, let $\alpha_m := \alpha_m(f;z)$, $\beta_m := \beta_m(f;z)$, $\gamma_m := \gamma_m(f;z)$, and suppose $\alpha_m \leq \frac{1}{9}$.*

(a) *If $\alpha_m > 0$, then for any r such that $r^- \leq r \leq r^+$ and $\bar{B}(z,r) \subseteq U$ the function f has m zeros, counting multiplicities, in $\bar{B}(z,r)$ and*

$$2\beta_m \leq 2\beta_m(1+\alpha_m) \leq r^- \leq 2\beta_m(1+9\alpha_m/2) \leq 3\beta_m$$

$$\leq \frac{1}{3\gamma_m} \leq \frac{1-3\alpha_m}{2\gamma_m} \leq r^+ \leq \frac{1-\alpha_m}{2\gamma_m} \leq \frac{1}{2\gamma_m}.$$

(b) *If $\beta_m = 0$, then for any r such that $\gamma_m r < \frac{1}{2}$ and $\bar{B}(z,r) \subseteq U$ the function f has m zeros, counting multiplicities, in $\bar{B}(z,r)$.*

(c) *If $\gamma_m = 0$, then for any $r \geq 2\beta_m$ such that $\bar{B}(z,r) \subseteq U$ the function f has m zeros, counting multiplicities, in $\bar{B}(z,r)$.*

Proof. Only the degenerate cases are left out of the previous discussion. If $\beta_m = 0$, then (8) becomes equivalent to $\gamma_m r < \frac{1}{2}$. If $\gamma_m = 0$ and $\beta_m \neq 0$ for any $r \geq 2\beta_m$, then (8) holds again. In both cases (b) and (c), the conclusions follow from Corollary 1.4. □

Here follows a useful corollary.

Corollary 1.6. *Assume that U is connected, let $m \geq 1$ be an integer and let $z \in U$ be such that $f^{(m)}(z) \neq 0$. Suppose $\alpha_m(f;z) \leq \frac{1}{9}$ and $\bar{B}(z, 3\beta_m(f;z)) \subseteq U$. Then f has m zeros, counted with multiplicities, in $\bar{B}(z, 3\beta_m(f;z))$ and in $\bar{B}(z, 1/(3\gamma_m(f;z))) \cap U$.*

1.4. *Inequalities Between Different Orders of Derivation*

In order to explicit lower bounds on $\bar{\alpha}_{m,l}$ from the previous one on $\bar{\alpha}_m$ (namely $\frac{1}{9}$), we will use the following proposition that relates various quantities of type γ, β, and α associated to f and its higher derivatives. More precisely, we show that $\gamma_m(f;z)$ (resp., $\beta_{m,l}(f;z)$) and $\gamma_{m-l}(f^{(l)};z)$ (resp., $\beta_{m-l,0}(f^{(l)};z)$) are roughly equivalent, for fixed values of m and l.

Proposition 1.7. *Let $z \in U$ and let $m \geq 1$ be an integer such that $f^{(m)}(z) \neq 0$, then, for any $l \in \{0, \dots, m-1\}$ and $l' \in \{0, \dots, m-l-1\}$:*

(a) $\alpha_{m-l,l'}(f^{(l)};z) \leq \dfrac{m+1}{m+1-l}\dfrac{m-l}{m}\alpha_{m,l+l'}(f;z) \leq \alpha_{m,l+l'}(f;z)$;

(b) $\dfrac{l'+1}{l+l'+1}\beta_{m,l+l'}(f;z) \leq \beta_{m-l,l'}(f^{(l)};z) \leq \dfrac{m-l}{m}\beta_{m,l+l'}(f;z)$;

(c) $\gamma_m(f;z) \leq \gamma_{m-l}(f^{(l)};z) \leq \dfrac{m+1}{m+1-l}\gamma_m(f;z)$.

Proof. We introduce the following quantity $c_{m,l,j}$:

$$c_{m,l,j} := \frac{\binom{l+j}{l}}{\binom{m}{l}} = \frac{(l+j)!}{j!}\frac{(m-l)!}{m!}.$$

Now, setting $j := k - l$ we can compare the terms occurring in definition (2) of γ and β associated to f and its higher derivatives

$$\begin{aligned} \frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(z)|} &= \frac{m!\,|f^{(l+j)}(z)|}{(l+j)!\,|f^{(m)}(z)|} \\ &= (c_{m,l,j})^{-1}\frac{(m-l)!\,|f^{(l+j)}(z)|}{j!\,|f^{(m)}(z)|}. \end{aligned}$$

We start with part (c). Taking the supremum of the adequate powers of the expression above, we just check that the range of summations fits the definitions to obtain

$$\gamma_m(f;z) = \sup_{k \geq m+1}\left(\frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(z)|}\right)^{1/(k-m)}$$

$$= \sup_{j \geq m-l+1} (c_{m,l,j})^{-1/(j+l-m)} \left(\frac{(m-l)!\,|f^{(l+j)}(z)|}{j!\,|f^{(m)}(z)|} \right)^{1/(j+l-m)}$$
$$\leq \sup_{j \geq m-l+1} (c_{m,l,j})^{-1/(j+l-m)} \gamma_{m-l}(f^{(l)}; z).$$

Conversely, reexpressing $\gamma_{m-l}(f^{(l)}; z)$, we get

$$\gamma_{m-l}(f^{(l)}; z) \leq \gamma_m(f; z) \sup_{j \geq m-l+1} (c_{m,l,j})^{1/(j-m+l)}.$$

For $j \geq m - l + 1$ the following lower and upper bounds for $c_{m,l,j}$ conclude part (c):

$$\left(\frac{j+l}{j} \right)^{j-m+l} \leq c_{m,l,j} = \prod_{i=m-l+1}^{j} \frac{i+l}{i} \leq \left(\frac{m+1}{m+1-l} \right)^{j-m+l}.$$

As for part (b) we get, similarly,

$$\beta_{m-l,l'}(f^{(l)}; z) \leq \beta_{m,l+l'}(f; z) \sup_{j=l',\ldots,m-l-1} (c_{m,l,j})^{1/(m-l-j)}$$

and

$$\beta_{m,l+l'}(f; z) \leq \beta_{m-l,l'}(f^{(l)}; z) \sup_{j=l',\ldots,m-l-1} (c_{m,l,j})^{-1/(m-l-j)}.$$

For $j = l', \ldots, m - l - 1$ we have

$$\left(\frac{j+1}{j+l+1} \right)^{m-l-j} \leq c_{m,l,j} = \prod_{i=j+1}^{m-l} \frac{i}{i+l} \leq \left(\frac{m-l}{m} \right)^{m-l-j},$$

which yields part (b). Part (a) follows easily from (b) and (c) since

$$\frac{m-l}{m-l+1} \frac{m+1}{m} \leq 1. \qquad \square$$

1.5. *Lower Bounds on Critical Values* $\bar{\alpha}_{m,l}$

For zeros of polynomials, the Gauss–Lucas theorem [**25**] asserts that the zeros of the derivatives of P belong to the convex hull of the zeros of P. The following corollary can be seen as a weak generalization of this result for analytic functions: zeros of derivatives remain close to a cluster. This result is obtained by combining Proposition 1.7 and the previous corollary to $f^{(l)}$, which provides lower bounds for $\bar{\alpha}_{m,l}$.

Corollary 1.8. *Assume that U is connected, let $m \geq 1$ be an integer, $l \in \{0, \dots, m-1\}$, $z \in U$, and assume $f^{(m)}(z) \neq 0$,*

$$\frac{m-l}{m}\frac{m+1}{m+1-l} \quad \alpha_{m,l}(f; z) \leq \tfrac{1}{9}$$

and let $\bar{B}(z, 3[(m-l)/m]\beta_{m,l}(f; z)) \subseteq U$. Then, $f^{(l)}$ has $m-l$ zeros (counting multiplicities) in $\bar{B}(z, 3[(m-l)/m]\beta_{m,l}(f; z))$ and $\bar{B}(z, (m+1-l)/[3(m+1)\gamma_m(f; z)]) \cap U$.

Proof. From Proposition 1.7 we have

$$3\beta_{m-l}(f^{(l)}; z) \leq 3\frac{m-l}{m}\beta_{m,l}(f; z) =: r, \quad 3\gamma_{m-l}(f^{(l)}; z) \leq \frac{3(m+1)\gamma_m(f; z)}{m+1-l}.$$

Using the assumption on $\alpha_{m,l}(f; z)$, we deduce $\alpha_{m-l}(f^{(l)}; z) \leq \frac{1}{9}$, and

$$r \leq \frac{m+1-l}{3(m+1)\gamma_m(f; z)} \leq \frac{1}{3\gamma_{m-l}(f^{(l)}; z)}.$$

The conclusion follows from the previous corollary applied with $f^{(l)}$: $f^{(l)}$ has $m-l$ zeros in the ball $\bar{B}(z, 3\beta_{m-l}(f^{(l)}; z))$, counting multiplicities, and also in $\bar{B}(z, 1/[3\gamma_{m-l}(f^{(l)}; z)]) \cap U$. □

2. Cluster Diameter

In the previous section we have shown that if α_m is small enough, then there exists a cluster Z of m zeros of f, that has a diameter D bounded in terms of β_m at points in the convex hull of Z. In this section we focus on the converse inequality, that is, upper bounding β_m in terms of the diameter of the cluster. Throughout this section, f denotes an analytic function from a *connected* open subset $U \subseteq \mathbb{C}$. We will show:

Theorem 2.1. *Let $m \geq 1$ be an integer and let $\zeta \in U$ be such that $f^{(m)}(\zeta) \neq 0$. Suppose $m\alpha_m(f; \zeta) < \frac{1}{12}$ and $\bar{B}(\zeta, 3\beta_m(f; \zeta)) \subseteq U$. Then f has a cluster Z of m zeros in $\bar{B}(z, 3\beta_m(f; z))$, counting multiplicities. In addition, if ζ is in the convex hull of Z, then $\beta_m(f; \zeta) \leq 24m^2 D$, where D denotes the diameter of Z.*

We postpone the proof until the end of this section. Let $Z := \{z_1, \dots, z_r\}$ denote zeros of f in U with respective multiplicities $n_1, \dots, n_r$. Let $m := n_1 + \cdots + n_r$ and let ζ be a point in the convex hull of Z. Assume $f^{(m)}(\zeta) \neq 0$ and, in short, let $\alpha_m := \alpha_m(f; \zeta)$, $\beta_m := \beta_m(f; \zeta)$, and $\gamma_m := \gamma_m(f; \zeta)$. By D we denote the diameter of Z and we introduce $p(x) := \prod_{i=1}^{r}(x - z_i)^{n_i}$ and $g := f/p$. We start with a technical lemma on majorant series.

Lemma 2.2. *Assume $\alpha_m \leq 1$ and there exists $\varepsilon > 0$ such that $\gamma_m D < 1 - \varepsilon$, then*

$$\left[\frac{m!\, g}{f^{(m)}(\zeta)}\right]_\zeta \leq \frac{1}{\varepsilon}\frac{1}{(1-\varepsilon-\gamma_m D)^m}\frac{1}{1-[\gamma_m/(1-\varepsilon)]t}.$$

Proof. If $\gamma_m = 0$, then $g = f^{(m)}(\zeta)/m!$, hence the inequality holds trivially. Assume now that $\gamma_m \neq 0$. Cauchy's formula for the kth coefficient I_k in the Taylor expansion of $g(x)$ at ζ gives the integral representation

$$I_k = \frac{1}{2i\pi}\oint_C \frac{f(z)}{p(z)}\frac{dz}{(z-\zeta)^{k+1}},$$

where the contour C can be chosen as the circle of radius $r = \gamma_m^{-1}(1-\varepsilon)$ around ζ, since f/p is analytic in the enclosed disk. Changing the variable to $z = \zeta + r\exp(\imath\theta)$, we see that this integral is bounded by

$$|I_k| \leq \frac{\max_{z\in C}|f(z)|}{\min_{z\in C}|p(z)|}\frac{1}{r^k}.$$

Since $[f]_\zeta$ has nonnegative Taylor coefficients, the maximum of $|f|$ on C is bounded by the value of $[f]_\zeta$ at $t = r$. On the other hand, the distance between C and the z_i is at least $r - D$, which implies that $|p|$ is at least $(r-D)^m$. Thus we get

$$\frac{m!\,|I_k|}{|f^{(m)}(\zeta)|} \leq \frac{1}{(r-D)^m}\left(\sum_{i=0}^{m-1}\beta_m^{m-i}r^i + \frac{r^m}{1-\gamma_m r}\right)\frac{1}{r^k}.$$

Multiplying by t^k and summing over k yields

$$\left[\frac{m!\, g}{f^{(m)}(\zeta)}\right]_\zeta \leq \frac{1}{(r-D)^m}\left(\sum_{i=0}^{m-1}\beta_m^{m-i}r^i + \frac{r^m}{1-\gamma_m r}\right)\frac{1}{1-t/r}.$$

The conclusion of the lemma follows from replacing r by its value and bounding β_m by γ_m^{-1}, since $\alpha_m \leq 1$. □

From this lemma we deduce the following bound:

Proposition 2.3. *According to the above notation and assumptions, if $\alpha_m \leq 1$ and $\gamma_m D < 1$, then, for any $\varepsilon > 0$ such that $\gamma_m D < 1 - \varepsilon$, we have*

$$\beta_m \leq \sup_{0\leq k\leq m-1}\left(\lambda_g \sum_{i=0}^{k}\binom{m}{k-i}\rho_g^i D^{m-k+i}\right)^{1/(m-k)},$$

where $\lambda_g := \varepsilon^{-1}(1-\varepsilon-\gamma_m D)^{-m}$ and $\rho_g := \gamma_m/(1-\varepsilon)$.

Proof. First, from the previous lemma we deduce

$$\left[\frac{m!\,g}{f^{(m)}(\zeta)}\right]_\zeta \leq \frac{\lambda_g}{1-\rho_g t}.$$

Second, from Proposition A.4, we also have

$$[p]_\zeta \leq \prod_{i=1}^{r} [x - z_i]_\zeta^{n_i} \leq (D+t)^m.$$

Then the result follows from bounding termwise the product in the right-hand side of

$$\left[\frac{m!\,f}{f^{(m)}(\zeta)}\right]_\zeta \leq [p]_\zeta \left[\frac{m!\,g}{f^{(m)}(\zeta)}\right]_\zeta,$$

coming from Proposition A.4 again. □

We now deduce simplified bounds.

Corollary 2.4. *Under the assumptions of the previous proposition and if* $m\rho_g D < 1$ *holds, then we have*

$$\beta_m \leq \frac{\lambda_g m D}{1 - m\rho_g D}.$$

Proof. The proof follows from the previous proposition and the following inequalities:

$$\begin{aligned}\lambda_g \sum_{i=0}^{k} \binom{m}{k-i} \rho_g^i D^{m-k+i} &\leq \lambda_g \sum_{i=0}^{k} m^{m-k+i} \rho_g^i D^{m-k+i} \\ &\leq \lambda_g (mD)^{m-k} \sum_{i=0}^{k} (m\rho_g D)^i \\ &\leq \frac{(\lambda_g m D)^{m-k}}{1 - m\rho_g D}.\end{aligned}$$

□

We immediately draw from the previous bound:

Corollary 2.5. *According to the above notation and assumptions, if* $\alpha_m \leq 1$, *then, for any positive real numbers a and b such that* $0 < a < a + b/m < 1$ *and* $m\gamma_m D < a$, *we have*

$$\beta_m \leq \frac{mD}{1 - \dfrac{a}{1-b/m}} \frac{1}{\dfrac{b}{m}\left(1 - \dfrac{b}{m} - \dfrac{a}{m}\right)^m}.$$

Proof. The proof follows from taking $\varepsilon = b/m < 1$ in the previous corollary, which is valid since we have $\gamma_m D \le a < 1-\varepsilon$ and $m\rho_g D = m\gamma_m D/(1-b/m) \le a/(1-b/m) < 1$. □

Setting a and b to some suitable values we deduce:

Corollary 2.6. *According to the above notation and assumptions, if* $\alpha_m \le 1$ *and* $m\gamma_m D < \frac{1}{2}$, *then* $\beta_m \le 24m^2 D$.

Proof. If $m = 1$, then $D = 0$ and $\beta_m = 0$, hence we can assume $m \ge 2$. Letting $a = \frac{1}{2}$, $b = \frac{1}{2}$ in the previous corollary we deduce

$$\beta_m \le \frac{mD}{1 - \frac{\frac{1}{2}}{1-\frac{1}{4}}} \frac{1}{\frac{\frac{1}{2}}{m}\left(1 - \frac{1}{m}\right)^m}.$$

The conclusion follows from $(1 - 1/m)^{-m} \le 4$. □

Finally, combining Corollary 1.6 to this statement, we achieve:

Proof of Theorem 2.1. Since $m\alpha_m(f;\zeta) < \frac{1}{12}$ implies $\alpha_m(f;\zeta) \le \frac{1}{9}$, Corollary 1.6 applies: this gives the existence of Z and the inequality $D \le 6\beta_m(f;\zeta)$. It follows that, if ζ is in the convex hull of Z, then the conditions of the previous corollary are satisfied. □

3. Convergence Analysis

In the previous sections we have shown how the number of zeros in a cluster can be estimated in terms of estimates at a given point. Now we show that similar estimates make it possible to bound an annulus, centered at the cluster, within which the corrected Newton iterator is well defined, and a smaller annulus within which its convergence is quadratic. Intuitively, while the iterates are inside the annulus, they are far from the cluster and the iteration behaves as if the cluster were a multiple zero, whereas in the inner area of the annulus, the cluster does not behave like a multiple zero.

In this section f still denotes an analytic function defined on an open subset $U \subseteq \mathbb{C}$. For any $l \in \{0, \dots, m-1\}$ we analyze the convergence of Schröder's iterates for $f^{(l)}$. First we focus on the definition domain of the iterator, then we analyze the behavior of one iteration. Last, we examine the convergence of the iterates and present a stopping criterion in terms of data at the cluster. These data are a priori unknown but we will derive our stopping criterion from them in the next section.

3.1. *Zero-Free Areas for the Derivatives*

Now we show how to bound an annulus, centered at a cluster, within which the corrected Newton iterator (1) is well defined, that is, in which $f^{(l+1)}$ does not vanish. The analysis proceeds first in terms of majorant series and then in terms of $\beta_{m,l}$ and γ_m.

Lemma 3.1. *Assume that U is connected, let $\zeta \in U$, $m \geq 1$ be such that $f^{(m)}(\zeta) \neq 0$ and let F be a majorant series such that $[m!\, f/f^{(m)}(\zeta)]_\zeta \leq F$. For all $z \in U$, $z \neq \zeta$, such that $r := |z-\zeta|$ is smaller than the radius of convergence of F and*

$$\frac{F'(r)}{mr^{m-1}} - \frac{F^{(m)}(0)}{m!} < 1,$$

one has:

(a) $\dfrac{m!\, f'(z)}{f^{(m)}(\zeta)} = m(z-\zeta)^{m-1}(1+B)$, *where*

$$|B| \leq \frac{F'(r)}{mr^{m-1}} - \frac{F^{(m)}(0)}{m!};$$

(b) $f'(z) \neq 0$ *and* $\dfrac{|f^{(m)}(\zeta)|}{m!\,|f'(z)|} \leq \dfrac{1}{mr^{m-1}}\dfrac{1}{1-|B|}$.

Proof. A Taylor expansion at ζ gives

$$\begin{aligned}\frac{m!\, f'(z)}{f^{(m)}(\zeta)} &= (z-\zeta)^{m-1}\left(\sum_{j\geq 1} j\frac{m!\, f^{(j)}(\zeta)}{j!\, f^{(m)}(\zeta)}(z-\zeta)^{j-m}\right)\\ &= m(z-\zeta)^{m-1}(1+B),\end{aligned}$$

where

$$B := \frac{1}{m}\sum_{j\geq 1,\ j\neq m} j\frac{m!\, f^{(j)}(\zeta)}{j!\, f^{(m)}(\zeta)}(z-\zeta)^{j-m}.$$

Then we deduce from the hypothesis that

$$|B| \leq \frac{F'(r)}{mr^{m-1}} - \frac{F^{(m)}(0)}{m!} < 1,$$

which gives part (a). Part (b) follows from the triangle inequality $|1+B| \geq 1-|B|$. □

Now these estimates extend to the derivatives of f as follows, using majorant series in terms of $\beta_{m,l}$ and γ_m.

Corollary 3.2. *Assume that U is connected, let $m \geq 1$ and $\zeta \in U$ be such that $f^{(m)}(\zeta) \neq 0$. Let $l \in \{0, \ldots, m-1\}$, then for all $z \in U$, $z \neq \zeta$, such that $u := \max\big(\beta_{m,l}(f;\zeta)/r, \gamma_m(f;\zeta)r\big) < 1-(\frac{1}{2})^{1/(l+2)}$ (with $r := |z-\zeta|$), one has:*

(a) $\dfrac{(m-l)!\, f^{(l+1)}(z)}{f^{(m)}(\zeta)} = (m-l)(z-\zeta)^{m-(l+1)}(1+B)$, *where*

$$|B| \leq \frac{1}{(1-u)^{l+2}} - 1 < 1;$$

(b) $f^{(l+1)}(z) \neq 0$ *and* $\dfrac{|f^{(m)}(\zeta)|}{(m-l)!\,|f^{(l+1)}(z)|} \leq \dfrac{(1-u)^{l+2}}{(m-l)r^{m-(l+1)}\psi_{l+1}(u)}$.

Proof. From Lemma 1.3, one has

$$\left[\frac{(m-l)!\, f^{(l)}}{f^{(m)}(\zeta)}\right]_\zeta \leq F_{m,l}(f,\zeta;t).$$

Let

$$A := \frac{1}{m-l}\frac{F'_{m,l}(f,\zeta;r)}{r^{m-(l+1)}} - \frac{F^{(m-l)}_{m,l}(f,\zeta;0)}{(m-l)!},$$

we claim that

$$A \leq \frac{1}{(1-u)^{l+2}} - 1. \tag{12}$$

Since $u < 1-(\frac{1}{2})^{1/(l+2)}$, we deduce $A < 1$ and therefore the previous lemma applies with $f^{(l)}$ and $F_{m,l}(f,\zeta;t)$:

$$\frac{(m-l)!\, f^{(l+1)}(z)}{f^{(m)}(\zeta)} = (m-l)(z-\zeta)^{m-(l+1)}(1+B),$$

with $|B| \leq A$. Then parts (a) and (b) follow from direct calculations.

Differentiating $F_{m,l}(f,\zeta;t)$, given by formula (6), we obtain

$$A \;\leq\; \binom{m}{l+1}^{-1}\left(\sum_{j=1}^{m-(l+1)}\binom{m-j}{l+1}u^j + \sum_{j\geq 1}\binom{m+j}{l+1}u^j\right).$$

Thanks to

$$\sum_{j\geq 1}\binom{j+l+1}{l+1}u^j = \frac{1}{(1-u)^{l+2}} - 1,$$

inequality (12) reduces to proving that, for all $i = l+1 \in \{1, \ldots, m\}$,

$$\binom{m-j}{i} + \binom{m+j}{i} \leq \binom{m}{i}\binom{j+i}{i} \quad \text{for } j \geq 1, \tag{13}$$

with the natural convention that $\binom{m-j}{i} = 0$ if $j \geq m - i + 1$. We prove inequality (13) by induction on m. If $m = 1$, then the only possible case to examine is $i = 1$ and the verification is immediate. Assuming that (13) holds for $m \geq 1$, we now show that it holds for $m + 1$. The case $i = m + 1$ holds trivially, hence we can suppose $i \leq m$ and apply the induction hypothesis, after using Pascal's triangle equality (which always holds with the aforementioned convention):

$$\begin{aligned}
&\binom{m+1-j}{i} + \binom{m+1+j}{i} \\
&\quad = \binom{m-j}{i} + \binom{m+j}{i} + \binom{m-j}{i-1} + \binom{m+j}{i-1} \\
&\quad \leq \binom{m}{i}\binom{j+i}{i} + \binom{m}{i-1}\binom{j+i-1}{i-1} \\
&\quad = \binom{m+1}{i}\binom{j+i}{i}\left(\frac{m-i+1}{m+1} + \frac{i}{m+1}\frac{i}{j+i}\right) \\
&\quad = \binom{m+1}{i}\binom{j+i}{i}\left(1 - \frac{ij}{(m+1)(j+i)}\right) \\
&\quad \leq \binom{m+1}{i}\binom{j+i}{i}.
\end{aligned}$$

□

3.2. *Analysis of One Iteration*

The following proposition shows the existence of an annulus around the cluster in which the convergence of one iteration of $N_{m-l}(f^{(l)}; \cdot)$ is quantified in terms of $\beta_{m,l}$ and γ_m. The following formulas constitute the cornerstone of our approximation algorithm.

Proposition 3.3. *Assume that U is connected, let $m \geq 1$ be an integer and let $\zeta \in U$ be such that $f^{(m)}(\zeta) \neq 0$, $x_0 \in U$, $x_0 \neq \zeta$. Let $l \in \{0, \ldots, m-1\}$, $u_\beta := \beta_{m,l}(f; \zeta)/|x_0 - \zeta|$, $u_\gamma := \gamma_m(f; \zeta)|x_0 - \zeta|$, $u := \max(u_\beta, u_\gamma)$, and suppose $u < 1 - (1/2)^{1/(l+2)}$. Then $f^{(l+1)}(x_0) \neq 0$, hence $x_1 := N_{m-l}(f^{(l)}; x_0)$ is well defined and*

$$|x_1 - \zeta| \leq \frac{|x_0 - \zeta|}{(m-l)\psi_{l+1}(u)}\left(\frac{m-l}{m}u_\beta + \frac{m+1}{m-l+1}u_\gamma\right).$$

Proof. Let $r := |x_0 - \zeta|$. According to Lemma 1.3, we start with

$$\left[\frac{(m-l)!\, f^{(l)}}{f^{(m)}(\zeta)}\right]_\zeta \leq F_{m,l}(f, \zeta; t).$$

Corollary 3.2 asserts that $f^{(l+1)}(x_0) \neq 0$ and

$$\frac{|f^{(m)}(\zeta)|}{(m-l)!\,|f^{(l+1)}(x_0)|} \leq \frac{(1-u)^{l+2}}{(m-l)r^{m-(l+1)}\psi_{l+1}(u)}. \tag{14}$$

On the other hand, we have

$$\left|\frac{(x_0-\zeta)f^{(l+1)}(x_0)-(m-l)f^{(l)}(x_0)}{f^{(m)}(\zeta)/(m-l)!}\right| \leq rG'(r)-(m-l)G(r),$$

where

$$G(t) := -\sum_{j=0}^{m-l-1}\frac{F_{m,l}^{(j)}(f,\zeta;0)}{j!}t^j + \sum_{j\geq m-l}\frac{F_{m,l}^{(j)}(f,\zeta;0)}{j!}t^j.$$

From formula (6) and letting

$$c_{m,l,j} := (j-l)\binom{j}{l} - (m-l)\binom{j}{l} = (j-m)\binom{j}{l},$$

we deduce

$$rG'(r)-(m-l)G(r) = \frac{r^{m-l}}{\binom{m}{l}}\left(-\sum_{j=l}^{m-1}c_{m,l,j}u_\beta^{m-j} + \sum_{j\geq m+1}c_{m,l,j}u_\gamma^{j-m}\right). \tag{15}$$

From

$$x_1-\zeta = \frac{f^{(m)}(\zeta)}{(m-l)!\,f^{(l+1)}(x_0)}\frac{(x_0-\zeta)f^{(l+1)}(x_0)-(m-l)f^{(l)}(x_0)}{f^{(m)}(\zeta)/(m-l)!}$$

and combining (14) and (15), we deduce

$$|x_1-\zeta| \leq \frac{(1-u)^{l+2}r}{(m-l)\binom{m}{l}\psi_{l+1}(u)}\left(-\sum_{j=l}^{m-1}c_{m,l,j}u_\beta^{m-j} + \sum_{j\geq m+1}c_{m,l,j}u_\gamma^{j-m}\right). \tag{16}$$

For all $l \in \{0,\ldots,m-1\}$, we claim

$$-c_{m,l,j} \leq \frac{m-l}{m}\binom{m}{l}\binom{l+m-j}{l+1} \quad \text{for } l \leq j \leq m-1, \tag{17}$$

hence

$$-\sum_{j=l}^{m-1}c_{m,l,j}u_\beta^{m-j} \leq \frac{m-l}{m}\binom{m}{l}\frac{u_\beta}{(1-u_\beta)^{l+2}}. \tag{18}$$

We also claim

$$c_{m,l,j} \leq \frac{m+1}{m+1-l}\binom{m}{l}\binom{l+j-m}{l+1} \quad \text{for } j \geq m+1, \tag{19}$$

hence

$$\sum_{j\geq m+1} c_{m,l,j} u_\gamma^{j-m} \leq \frac{m+1}{m-l+1}\binom{m}{l}\frac{u_\gamma}{(1-u_\gamma)^{l+2}}. \tag{20}$$

Combining (16), (18), and (20) we deduce

$$|x_1-\zeta| \;\leq\; \frac{r}{(m-l)\psi_{l+1}(u)}\left(\frac{m-l}{m}u_\beta + \frac{m+1}{m-l+1}u_\gamma\right).$$

It remains to prove (17) and (19). First, concerning (17), for $j = m-1$ this inequality is an equality and then it is sufficient to check that

$$-c_{m,l,j}\binom{l+m-j}{l+1}^{-1}$$

increases with respect to j. Last, concerning (19), observe again that it is an equality when $j = m+1$ and that

$$c_{m,l,j}\binom{l+j-m}{l+1}^{-1}$$

is a decreasing sequence with respect to j. □

3.3. *Stopping the Iteration*

The next result shows quadratic convergence to the cluster while Schröder's iterates remain far enough from it. We introduce the following universal quantities:

$$\theta_{m,l,\delta} \;:=\; \delta\frac{1}{m} + \frac{m+1}{(m-l+1)(m-l)},$$

$$u_{m,l,\delta} \;:=\; \max\left\{u \geq 0 \text{ such that } u < 1-\left(\tfrac{1}{2}\right)^{1/(l+2)} \text{ and } \frac{\theta_{m,l,\delta}u}{\psi_{l+1}(u)} \leq 1\right\}.$$

Observe that this maximum is well defined since $1-\left(\tfrac{1}{2}\right)^{1/(l+2)}$ is the first positive zero of ψ_{l+1}.

Proposition 3.4. *Let f be an analytic function from an open subset $U \subseteq \mathbb{C}$. Let $\zeta \in U$, $m \geq 1$ be an integer, $l \in \{0,\ldots,m-1\}$, and let $r \geq 0$ be a real number such that $\bar{B}(\zeta,r) \subseteq U$ and $f^{(m)}(\zeta) \neq 0$. Let $\delta := 1$ if $\beta_{m,l}(f;\zeta) \neq 0$ and $\delta \in \{0,1\}$ otherwise. Let $u := \gamma_m(f;\zeta)r$ and assume $u \leq u_{m,l,\delta}$.*

Let $x_0 \in \bar{B}(\zeta,r)$ and consider the sequence $(x_k)_{k\geq 0}$ recursively defined by $x_{k+1} := N_{m-l}(f^{(l)};x_k)$. Let $K \geq 0$ be the smallest integer k such that $\beta_{m,l}(f;\zeta) \geq \gamma_m(f;\zeta)|x_k-\zeta|^2$ (K may be infinite). Then, for all $0 \leq k \leq K-1$, the iterate x_{k+1} is well defined, belongs to $\bar{B}(\zeta,r)$, and

$$|x_{k+1}-\zeta| \leq \frac{\theta_{m,l,\delta}\gamma_m(f;\zeta)}{\psi_{l+1}(u)}|x_k-\zeta|^2.$$

Proof. In short, let $\beta_{m,l} := \beta_{m,l}(f;\zeta)$ and $\gamma_m := \gamma_m(f;\zeta)$. The case $K = 0$ is trivial, so that we can assume $K \geq 1$. When $k = 0$, $x_0 \in \bar{B}(\zeta, r)$ and $\beta_{m,l} < \gamma_m|x_0 - \zeta|^2$. In particular, $x_0 \neq \zeta$ and, therefore, the previous proposition applies

$$\begin{aligned}|x_1 - \zeta| &\leq \frac{|x_0 - \zeta|}{(m-l)\psi_{l+1}(u_0)}\left(\frac{m-l}{m}\frac{\beta_{m,l}}{|x_0 - \zeta|}\delta + \frac{m+1}{m-l+1}u_0\right)\\ &\leq \frac{\theta_{m,l,\delta}\gamma_m}{\psi_{l+1}(u_0)}|x_0 - \zeta|^2,\end{aligned}$$

where $u_0 := \gamma_m|x_0 - \zeta|$. Using the fact that $1/\psi_{l+1}(u)$ is an increasing function, we deduce

$$|x_1 - \zeta| \leq \frac{\theta_{m,l,\delta}\gamma_m}{\psi_{l+1}(u)}|x_0 - \zeta|^2 \leq \frac{\theta_{m,l,\delta}u}{\psi_{l+1}(u)}|x_0 - \zeta|,$$

whence $x_1 \in \bar{B}(\zeta, r)$. Then a straightforward induction on k concludes the proof. □

Remark that if $\gamma_m(f;\zeta) = 0$, then $K = 0$ in the previous proposition. If $\gamma_m(f;\zeta) \neq 0$ and $\beta_{m,l}(f;\zeta) = 0$ this proposition asserts that the sequence $(x_k)_{k\in\mathbb{N}}$ converges quadratically to ζ. But if $\beta_{m,l}(f;\zeta) > 0$, then K is finite. Unfortunately, it is not possible to handle the computation of K from the only data at the current iterates. In the algorithm proposed in the next section we relax the conditions of the previous proposition.

Analysis of the Last Iteration

Speaking informally and according to the notation of the previous proposition, we show that x_K or x_{K+1} is very close to the cluster of $f^{(l)}$ when the iteration stops. More generally, for any $l' \leq l$, we provide a stopping criterion to get close to the cluster of $f^{(l')}$. In particular, if $l = m - 1$ one can take $l' = 0$, in order to stop the approximation close to the cluster of f. Recall that $\beta_{m,l'} \geq \beta_{m,l}$.

Proposition 3.5. *Assume that U is connected, let $m \geq 1$ be an integer, let $\zeta \in U$ be such that $f^{(m)}(\zeta) \neq 0$, and let $x_0 \in U$. Let $l \in \{0, \ldots, m-1\}$, $l' \geq 0$, $l' \leq l$, $\delta := 1$ if $\beta_{m,l}(f;\zeta) \neq 0$ and $\delta \in \{0, 1\}$ otherwise. Assume*

$$\gamma_m(f;\zeta)|x_0 - \zeta| \leq u_{m,l,\delta}.$$

(a) *If $f^{(l+1)}(x_0) = 0$, then*

$$|x_0 - \zeta| \leq \frac{\beta_{m,l'}(f;\zeta)}{u_{m,l,\delta}}.$$

(b) *If* $f^{(l+1)}(x_0) \neq 0$, *then* $x_1 := N_{m-l}(f^{(l)}; x_0)$ *is well defined and if* $\beta_{m,l'}(f;\zeta) \geq \gamma_m(f;\zeta)|x_0-\zeta|^2$, *then the following inequality holds*:

$$\min(|x_0-\zeta|, |x_1-\zeta|) \leq \frac{\beta_{m,l'}(f;\zeta)}{u_{m,l,\delta}}.$$

Proof. Part (a) follows directly from Proposition 3.3 and the definition of $u_{m,l,\delta}$. Concerning part (b), if $|x_0-\zeta| \leq \beta_{m,l'}(f;\zeta)/u_{m,l,\delta}$, then we are done, so that we can now assume that the contrary holds, that is, $\beta_{m,l'}(f;\zeta)/|x_0-\zeta| < u_{m,l,\delta}$. Then, using hypothesis $\gamma_m(f;\zeta)|x_0-\zeta| \leq \beta_{m,l'}(f;\zeta)/|x_0-\zeta| =: u_\beta$, Proposition 3.3 yields

$$\begin{aligned} |x_1-\zeta| &\leq \frac{\theta_{m,l,\delta}|x_0-\zeta|}{\psi_{l+1}(u_\beta)}u_\beta = \frac{\theta_{m,l,\delta}}{\psi_{l+1}(u_\beta)}\beta_{m,l'}(f;\zeta) \\ &\leq \frac{\theta_{m,l,\delta}}{\psi_{l+1}(u_{m,l,\delta})}\beta_{m,l'}(f;\zeta) \leq \frac{\beta_{m,l'}(f;\zeta)}{u_{m,l,\delta}}. \end{aligned}$$ □

4. Cluster Approximation

In this section, we undertake the problem of stopping Schröder's operator in the case of clusters with positive diameters. Informally speaking, we turn into practice the theoretical convergence analysis of the previous section. Here we start by presenting crucial formulas bounding point estimates starting from other estimates at a close point. In Section 4.2 we introduce the function $\mathcal{B}_{m,l}$ that aims at computing approximations of $\beta_{m,l}$. This is followed by the *main lemma*, that provides a stopping criterion from the data at the current iterates only. Conditions depending on data at the cluster appear in the main theorem, which summarizes our approximation algorithm. Section 4.5 is devoted to testing these conditions from the initial point only.

4.1. *Translation of Point Estimates*

From now on, f denotes an analytic function defined on an open set $U \subseteq \mathbb{C}$. The next lemmas concern bounds in terms of majorant series. They are used in the next proposition, which gives bounds on translations of point estimates.

Lemma 4.1. *For any* $k \leq m$ *we have*

$$\left(\frac{1}{k!}\frac{t^m}{1-\gamma_m t}\right)^{(k)} \leq \frac{\binom{m}{k}t^{m-k}}{(1-\gamma_m t)^{k+1}}.$$

Proof. An easy rewriting of the left-hand side of the inequality yields

$$\left(\frac{1}{k!}\frac{t^m}{1-\gamma_m t}\right)^{(k)} = \left(\sum_{i\ge 0}\frac{1}{k!}\gamma_m^i t^{m+i}\right)^{(k)} = t^{m-k}\sum_{i\ge 0}\binom{m+i}{k}(\gamma_m t)^i.$$

The right-hand side is

$$\frac{\binom{m}{k}t^{m-k}}{(1-\gamma_m t)^{k+1}} = \binom{m}{k}t^{m-k}\sum_{i\ge 0}\binom{k+i}{k}(\gamma_m t)^i,$$

so that it is enough to prove

$$\binom{m+i}{k} \le \binom{m}{k}\binom{k+i}{k}.$$

But this follows from the fact that the sequence

$$u_i := \binom{m+i}{k}\binom{k+i}{i}^{-1}$$

is decreasing since

$$\frac{u_{i+1}}{u_i} = \frac{m+i+1}{k+i+1}\,\frac{i+1}{m+i+1-k} \le 1. \qquad \square$$

Lemma 4.2. *Let $\zeta \in U$, $m \ge 1$ be an integer and assume that $f^{(m)}(\zeta) \ne 0$. Let $\gamma_m := \gamma_m(f;\zeta)$ and $\beta_{m,l} := \beta_{m,l}(f;\zeta)$, for short.*

(a) *If $0 \le l \le m-1$, then*

$$\left[\frac{m!\,f^{(l)}}{l!\,f^{(m)}(\zeta)}\right]_\zeta \le \beta_{m,l}(\beta_{m,l}+(m-1)t)^{m-l-1} + \frac{\binom{m}{l}t^{m-l}}{(1-\gamma_m t)^{l+1}}.$$

(b) *If $l \ge m$, then* $\left[\frac{m!\,f^{(l)}}{l!\,f^{(m)}(\zeta)}\right]_\zeta \le \frac{\gamma_m^{l-m}}{(1-\gamma_m t)^{l+1}}$.

(c) $\left[\frac{f^{(m)}(\zeta)}{f^{(m)}}\right]_\zeta \le \frac{(1-\gamma_m t)^{m+1}}{2(1-\gamma_m t)^{m+1}-1}$.

Proof. By Lemma 1.3, we get

$$\left[\frac{m!\,f^{(l)}}{l!\,f^{(m)}(\zeta)}\right]_\zeta \le \sum_{i=l}^{m-1}\binom{i}{l}\beta_{m,l}^{m-i}t^{i-l} + \left(\frac{1}{l!}\frac{t^m}{1-\gamma_m t}\right)^{(l)}.$$

When $l \leq m-1$, we bound the first sum as follows:

$$\sum_{i=l}^{m-1} \binom{i}{l} \beta_{m,l}^{m-i} t^{i-l} \leq \beta_{m,l} \sum_{j=0}^{m-l-1} \binom{m-l-1}{j} \beta_{m,l}^{m-l-1-j} t^j \frac{(l+j)!}{l!}$$
$$\leq \beta_{m,l}(\beta_{m,l} + (m-1)t)^{m-l-1},$$

for $(l+j)!\,/l! \leq (l+j)^j \leq (m-1)^j$. For the second term, we use Lemma 4.1, this yields part (a). If $l \geq m$, then part (b) follows from Leibniz's rule

$$\left[\frac{m!\, f^{(l)}}{l!\, f^{(m)}(\zeta)}\right]_\zeta \leq \left(\frac{1}{l!} \frac{t^m}{1-\gamma_m t}\right)^{(l)} = \frac{\gamma_m^{l-m}}{(1-\gamma_m t)^{l+1}}.$$

Part (c) follows easily from (b), letting $l = m$ and using Proposition A.8 from the appendix. □

We are now able to deduce the following useful bounds:

Proposition 4.3. *Assume that U is connected, let $\zeta \in U$, $m \geq 1$ be an integer such that $f^{(m)}(\zeta) \neq 0$ and $l \in \{0, \ldots, m-1\}$. Let $\gamma_m := \gamma_m(f;\zeta)$ and $\beta_{m,l} := \beta_{m,l}(f;\zeta)$, for short. Let $z \in U$, $r := |z-\zeta|$ be such that $u := \gamma_m(f;\zeta)r < 1 - (\frac{1}{2})^{1/(m+1)}$, then $f^{(m)}(z) \neq 0$ and*

(a) $\alpha_{m,l}(f;z) \leq \dfrac{1}{\psi_m(u)^2}(\alpha_{m,l}(1-u)^{(l+1)/(m-l)} + (2m-1)u)$;

(b) $\beta_{m,l}(f;z) \leq \dfrac{1-u}{\psi_m(u)}(\beta_{m,l}(1-u)^{(l+1)/(m-l)} + (2m-1)r)$;

(c) $\gamma_m(f,z) \leq \dfrac{\gamma_m}{\psi_m(u)(1-u)}$;

(d) $\left|\dfrac{f^{(m)}(\zeta)}{f^{(m)}(z)}\right| \leq \dfrac{(1-u)^{m+1}}{\psi_m(u)}$;

(e) $\left|\dfrac{f^{(m)}(z)}{f^{(m)}(\zeta)}\right| \leq \dfrac{1}{(1-u)^{m+1}}$.

Proof. Parts (d) and (e) follow from parts (b) and (c) of the previous Lemma 4.2 and majorant series evaluation via Proposition A.3. Part (a) follows from (b) and (c). Concerning (b), thanks to Proposition A.3 again, evaluating the series inequalities of parts (a) and (c) of the previous lemma yields, for $l \leq k \leq m-1$,

$$\frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(z)|} = \frac{|f^{(m)}(\zeta)|}{|f^{(m)}(z)|} \frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(\zeta)|}$$
$$\leq \frac{(1-u)^{m+1}}{\psi_m(u)} \left(\beta_{m,l}(\beta_{m,l} + (m-1)r)^{m-k-1} + \frac{\binom{m}{k} r^{m-k}}{(1-u)^{k+1}}\right)$$

$$\leq \frac{(1-u)^{m+1}}{\psi_m(u)}(\beta_{m,l}+(m-1)r)^{m-k}+\frac{(mr(1-u))^{m-k}}{\psi_m(u)},$$

for $\binom{m}{k} \leq m^{m-k}$. Since $\psi_m(u) \leq 1$ we deduce part (b). As for part (c), we use parts (b) and (c) of the previous Lemma 4.2: for $k \geq m+1$ we have

$$\frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(z)|} = \frac{m!\,|f^{(k)}(z)|}{k!\,|f^{(m)}(\zeta)|}\,\frac{|f^{(m)}(\zeta)|}{|f^{(m)}(z)|} \leq \frac{\gamma_m^{k-m}}{(1-u)^{k-m}\psi_m(u)}. \qquad \square$$

4.2. *Approximation of $\beta_{m,l}$*

One of the important features of our algorithm is that it needs to compute only rough approximations of $\beta_{m,l}$ in order to check the stopping criterion. We thus specify our algorithm in terms of a function $\mathcal{B}_{m,l}(f, y; z)$ returning an *approximation* of $\beta_{m,l}(f; z)$, with the possible help of information computed at another point y. Specific approximation functions can be devised for various classes of functions f, this is discussed below.

For any two points y and z, we introduce $\bar{\gamma}_m := \max(\gamma_m(f; y), \gamma_m(f; z))$. An important quantity to capture the behavior of the iterates is provided by $v := \bar{\gamma}_m|y-z|$. In order to quantify the approximation provided by $\mathcal{B}_{m,l}$, we also assume that we are given a positive constant $v_{m,l} \leq +\infty$ and two *increasing* functions $\tau_{m,l,0}$ and $\tau_{m,l,1}$ defined on the interval $[0, v_{m,l})$, such that, for any two points y, z with $v < v_{m,l}$, one has

$$\mathcal{B}_{m,l}(f, y; z) \;\leq\; \tau_{m,l,1}(v)\beta_{m,l}(f; z) + \tau_{m,l,0}(v)\bar{\gamma}_m|y-z|^2, \tag{21}$$

$$\beta_{m,l}(f; z) \;\leq\; \tau_{m,l,1}(v)\mathcal{B}_{m,l}(f, y; z) + \tau_{m,l,0}(v)\bar{\gamma}_m|y-z|^2. \tag{22}$$

As an extreme example, one can take $\mathcal{B}_{m,l}(f, y; z) := \beta_{m,l}(f; z)$, $v_{m,l} := +\infty$, $\tau_{m,l,1}(v) := 1$, and $\tau_{m,l,0}(v) := 0$. In Section 5 we propose a numerical scheme based on interpolation at roots of unity.

4.3. *Main Lemma*

The convergence analysis of the previous section gave us the theoretical stopping criterion $\beta_{m,l'}(f; \zeta) \geq \gamma_m(f; \zeta)|x_k-\zeta|^2$. Of course, testing this inequality from estimates at each kth iterate x_k is not possible since ζ is a priori unknown. We bypass this problem by relaxing this condition. Informally speaking, the criterion we provide is based on the observation that if the latter inequality does not hold, then $\mathcal{B}_{m,l'}(f, x_k; x_{k+1}) \leq \mathcal{G}|x_{k+1}-x_k|^2$ for a certain value $\mathcal{G}$. We stop the iteration once the latter inequality is violated. The idea is to combine Propositions 3.3 and 4.3. Under the conditions of Proposition 3.4 the following lemma provides a criterion for stopping the corrected Newton iteration close to the cluster by giving a suitable value for $\mathcal{G}$.

Lemma 4.4. *Let f be an analytic function defined on a connected open set $U \subseteq \mathbb{C}$. Let $\zeta \in U$, $m \geq 1$ be an integer such that $f^{(m)}(\zeta) \neq 0$. Let $l \in \{0, \dots, m-1\}$, $l' \leq l$, $x_0 \in U$, $u := \gamma_m(f;\zeta)|x_0 - \zeta|$, $\delta := 1$ if $\beta_{m,l}(f;\zeta) \neq 0$ and $\delta \in \{0,1\}$ otherwise. If*

$$u < u_{m,l,\delta} \qquad \text{and} \qquad \beta_{m,l'}(f;\zeta) < \gamma_m(f;\zeta)|x_0 - \zeta|^2, \tag{23}$$

then $x_1 := N_{m-l}(f^{(l)}; x_0)$ is well defined and:

(a) *$\beta_{m,l'}(f;x_1) \leq C_{m,l,l',\delta}(u)\gamma_m(f;\zeta)|x_0 - x_1|^2$, where*

$$C_{m,l,l',\delta}(u) := \frac{1-u}{\psi_m(u)} \frac{(1-u)^{(l'+1)/(m-l')} + \dfrac{\theta_{m,l,\delta}(2m-1)}{\psi_{l+1}(u)}}{\left(1 - \dfrac{\theta_{m,l,\delta}u}{\psi_{l+1}(u)}\right)^2};$$

(b) *Let $\bar{\gamma}_m := \max(\gamma_m(f;x_0), \gamma_m(f;x_1))$, $v = \bar{\gamma}_m|x_0 - x_1|$, if $v < v_{m,l'}$, then*

$$\mathcal{B}_{m,l'}(f, x_0; x_1) \leq (\tau_{m,l',1}(v)C_{m,l,l',\delta}(u)\gamma_m(f;\zeta) + \tau_{m,l',0}(v)\bar{\gamma}_m)|x_0 - x_1|^2.$$

Proof. In short, let $r := |x_0 - \zeta|$, $\gamma_m := \gamma_m(f;\zeta)$, and $\beta_{m,l'} := \beta_{m,l'}(f;\zeta)$. Thanks to (23), Proposition 3.4 applies: x_1 is well defined and we have

$$|x_1 - \zeta| \leq \frac{\theta_{m,l,\delta}\gamma_m}{\psi_{l+1}(u)} r^2 < r. \tag{24}$$

Let $u_1 := \gamma_m|x_1 - \zeta|$, we have $u_1 \leq u$ and from Proposition 4.3(b) we deduce

$$\begin{aligned} \beta_{m,l'}(f;x_1) &\leq \frac{1-u_1}{\psi_m(u_1)}(\beta_{m,l'}(1-u_1)^{(l'+1)/(m-l')} + (2m-1)|x_1 - \zeta|) \\ &\leq \frac{1-u_1}{\psi_m(u_1)}\left((1-u_1)^{(l'+1)/(m-l')} + \frac{\theta_{m,l,\delta}(2m-1)}{\psi_{l+1}(u)}\right)\gamma_m r^2. \end{aligned}$$

Now the triangle inequality $r \leq |x_1 - x_0| + |x_1 - \zeta|$ and (24) yield

$$r \leq \frac{1}{1 - \dfrac{\theta_{m,l,\delta}u}{\psi_{l+1}(u)}}|x_0 - x_1|.$$

Finally, using the fact that the function $u \mapsto (1-u)^{1+(l'+1)/(m-l')}/\psi_m(u)$ increases, we reach part (a). Part (b) follows straightforwardly using (21). □

4.4. *Algorithm*

For the sake of simplicity, we assume from now on that analytic functions are defined on maximal analyticity domains. Before stating the main result, we introduce universal functions that quantify the behavior of the convergence:

$$\underline{\kappa}_{m,l,l',\delta}(u,v) := \frac{\tau_{m,l',1}(v)}{1-\dfrac{\tau_{m,l',0}(v)}{\mathcal{C}}},$$

$$\bar{\kappa}_{m,l,l',\delta}(u,v) := \tau_{m,l',1}(v) + \frac{\tau_{m,l',0}(v)}{\mathcal{C}},$$

$$\chi_{m,l,l',\delta}(u) := \frac{1-3u}{\psi_m(3u)}\left((1-3u)^{(l'+1)/(m-l')} + \frac{2m-1}{u_{m,l,\delta}}\right),$$

$$\Xi_{m,l,l',\delta}(u,v) := \frac{\underline{\kappa}_{m,l,l',\delta}(u,v)\tau_{m,l',1}(v)}{1-\dfrac{\tau_{m,l',0}(v)}{\mathcal{C}}\underline{\kappa}_{m,l,l',\delta}(u,v)\bar{\kappa}_{m,l,l',\delta}(u,v)}\chi_{m,l,l',\delta}(u),$$

$$\eta_{m,l,l',\delta}(u,v) := \frac{3\dfrac{m-l'}{m}\bar{\mathcal{C}}}{\left(1-3\dfrac{m-l'}{m}\bar{\mathcal{C}}v\right)^2},$$

where $\mathcal{C} := \tau_{m,l',1}(v)C_{m,l,l',\delta}(u) + \tau_{m,l',0}(v)$ and $\bar{\mathcal{C}} := \tau_{m,l',1}(v)\mathcal{C} + \tau_{m,l',0}(v)$.

The following theorem presents our stopping criterion and summarizes the main properties:

Theorem 4.5. *Let f be an analytic function defined on a maximal connected open set $U \subseteq \mathbb{C}$. Let $\zeta \in U$ and $m \geq 1$ be an integer such that $f^{(m)}(\zeta) \neq 0$, $l \in \{0,\ldots,m-1\}$ and $l' \leq l$. Let $\delta := 1$ if $\beta_{m,l}(f;\zeta) \neq 0$ and $\delta \in \{0,1\}$ otherwise. Let $r \geq 0$, γ_m and $\bar{\gamma}_m$ be given. Let $u := \gamma_m r$, $\bar{u} := 3u$, $v := 2\bar{\gamma}_m r$, and assume*

$$\left.\begin{array}{rcl} \bar{\gamma}_m & \geq & \gamma_m \geq \gamma_m(f;\zeta), \\ \bar{\gamma}_m & \geq & \max_{z\in\bar{B}(\zeta,3r)} \gamma_m(f;z), \\ u & < & u_{m,l,\delta}, \\ \bar{u} & < & 1-(\frac{1}{2})^{1/(m+1)}, \\ v & < & v_{m,l'}. \end{array}\right\} \tag{25}$$

Let $\mathcal{C} := \tau_{m,l',1}(v)C_{m,l,l',\delta}(u) + \tau_{m,l',0}(v)$, $\mathcal{G} := \mathcal{C}\bar{\gamma}_m$, $\bar{\mathcal{C}} := \tau_{m,l',1}(v)\mathcal{C} + \tau_{m,l',0}(v)$.

For any $x_0 \in \bar{B}(\zeta,r)$, let $x_1 := N_{m-l}(f^{(l)};x_0)$, then one of the following three exclusive cases holds:

(a) *If $f^{(l+1)}(x_0) = 0$ or $f^{(l+1)}(x_0) \neq 0$ and $x_1 \notin \bar{B}(x_0,2r)$, then*

$$\beta_{m,l'}(f;x_0) \leq \chi_{m,l,l',\delta}(u)\beta_{m,l'}(f;\zeta). \tag{26}$$

(b) *If $f^{(l+1)}(x_0) \neq 0$, $x_1 \in \bar{B}(x_0,2r)$ and $\mathcal{B}_{m,l'}(f,x_0;x_1) > \mathcal{G}|x_0-x_1|^2$, then*

$$\min(\beta_{m,l'}(f;x_0), \beta_{m,l'}(f;x_1)) \leq \chi_{m,l,l',\delta}(u)\beta_{m,l'}(f;\zeta). \tag{27}$$

In addition, we have

$$\mathcal{B}_{m,l'}(f, x_0; x_1) \le \underline{\kappa}_{m,l,l',\delta}(u, v)\beta_{m,l'}(f; x_1), \tag{28}$$

$$\beta_{m,l'}(f; x_1) \le \bar{\kappa}_{m,l,l',\delta}(u, v)\mathcal{B}_{m,l'}(f, x_0; x_1), \tag{29}$$

$$\mathcal{B}_{m,l'}(f, x_1; x_0) \le \tau_{m,l',1}(v)\beta_{m,l'}(f; x_0) + \frac{\tau_{m,l',0}(v)}{\mathcal{C}}\mathcal{B}_{m,l'}(f, x_0; x_1), \tag{30}$$

$$\beta_{m,l'}(f; x_0) \le \tau_{m,l',1}(v)\mathcal{B}_{m,l'}(f, x_1; x_0) + \frac{\tau_{m,l',0}(v)}{\mathcal{C}}\mathcal{B}_{m,l'}(f, x_0; x_1). \tag{31}$$

Let $z_0 := x_0$ *if* $\mathcal{B}_{m,l'}(f, x_1; x_0) < \mathcal{B}_{m,l'}(f, x_0; x_1)$ *and* $z_0 := x_1$, *otherwise. If*

$$\tau_{m,l',0}(v)\underline{\kappa}_{m,l,l',\delta}(u, v)\bar{\kappa}_{m,l,l',\delta}(u, v) < \mathcal{C}, \tag{32}$$

then we have

$$\beta_{m,l'}(f; z_0) \le \Xi_{m,l,l',\delta}(u, v)\beta_{m,l'}(f; \zeta). \tag{33}$$

(c) *Otherwise* ($f^{(l+1)}(x_0) \neq 0$, $x_1 \in \bar{B}(x_0, 2r)$, *and* $\mathcal{B}_{m,l'}(f, x_0; x_1) \le \mathcal{G}|x_0 - x_1|^2$) *we have*

$$\beta_{m,l'}(f; x_1) \le \bar{\mathcal{C}}\bar{\gamma}_m|x_0 - x_1|^2. \tag{34}$$

In addition, if

$$3\frac{m - l'}{m}\bar{\mathcal{C}}v < 1, \tag{35}$$

then

$$\frac{m - l'}{m}\frac{m + 1}{m + 1 - l'}\alpha_{m,l'}(f; x_1) \le \tfrac{1}{9}, \tag{36}$$

and $f^{(l')}$ *admits a cluster* Z_1 *of* $m - l'$ *zeros in*

$$\bar{B}\left(x_1, 3\frac{m - l'}{m}\beta_{m,l'}(f; x_1)\right).$$

If ζ *belongs to the convex hull of* Z_1, *then*

$$|x_1 - \zeta| \le \eta_{m,l,l',\delta}(u, v)\bar{\gamma}_m|x_0 - \zeta|^2. \tag{37}$$

Proof. In short, we let $\beta_{m,l'} := \beta_{m,l'}(f; \zeta)$.

Let us start with inequalities (26) and (27). If $x_0 = \zeta$, then they trivially hold, since $\chi_{m,l,l',\delta}(u) \ge \chi_{m,l,l',\delta}(0) \ge 1$.

Let us suppose now that $x_0 \neq \zeta$. If $f^{(l+1)}(x_0) = 0$, then Proposition 3.5(a) gives

$$|x_0 - \zeta| \leq \frac{\beta_{m,l'}}{u_{m,l,\delta}}. \tag{38}$$

If $f^{(l+1)}(x_0) \neq 0$ and $x_1 \notin \bar{B}(x_0, 2r)$, then we claim that

$$\beta_{m,l} \geq \gamma_m(f; \zeta)|x_0 - \zeta|^2.$$

If the latter inequality were not true then, thanks to $u < u_{m,l,\delta}$, we would obtain $f^{(l+1)}(x_0) \neq 0$ and $|x_1 - \zeta| \leq r$ via Proposition 3.4, whence $|x_0 - x_1| \leq |x_0 - \zeta| + |x_1 - \zeta| \leq 2r$, which contradicts $x_1 \notin \bar{B}(x_0, 2r)$. In consequence, one has $\beta_{m,l'} \geq \gamma_m(f; \zeta)|x_0 - \zeta|^2$ and $|x_1 - \zeta| \geq |x_0 - \zeta|$, so that Proposition 3.5 implies that (38) also holds in this case.

Since $u \leq \bar{u} < 1 - (1/2)^{1/(m+1)}$, Proposition 4.3(b) yields

$$\beta_{m,l'}(f; x_0) \leq \frac{1 - \bar{u}}{\psi_m(\bar{u})}(\beta_{m,l'}(1 - \bar{u})^{(l'+1)/(m-l')} + (2m - 1)|x_0 - \zeta|),$$

hence (26). This concludes part (a).

Before going further, we need to remark that using the assumption $\bar{\gamma}_m \geq \max_{z \in \bar{B}(\zeta,3r)} \gamma_m(f; z)$, if $x_1 \in \bar{B}(x_0, 2r) \subseteq \bar{B}(\zeta, 3r)$, then one has

$$\bar{\gamma}_m \geq \max(\gamma_m(f; x_0), \gamma_m(f; x_1)),$$

hence $\bar{\gamma}_m|x_0 - x_1| \leq v < v_{m,l'}$.

Now let us deal with part (b). If we had $\beta_{m,l'} < \gamma_m(f; \zeta)|x_0 - \zeta|^2$, then all conditions of Lemma 4.4 would be satisfied, which would imply $\mathcal{B}_{m,l'}(f, x_0; x_1) \leq \mathcal{G}|x_0 - x_1|^2$. This yields a contradiction, hence we necessarily have $\beta_{m,l'} \geq \gamma_m(f; \zeta)|x_0 - \zeta|^2$. Then Proposition 3.5(b) provides

$$\min(|x_0 - \zeta|, |x_1 - \zeta|) \leq \frac{\beta_{m,l'}}{u_{m,l,\delta}}.$$

Because

$$\gamma_m \max(|x_0 - \zeta|, |x_1 - \zeta|) \leq \gamma_m(|x_0 - x_1| + |x_0 - \zeta|) \leq \bar{u} < 1 - \left(\frac{1}{2}\right)^{1/(m+1)},$$

Proposition 4.3(b) applies again and leads to (27).

Let us now deal with (28). Since $\bar{\gamma}_m|x_0 - x_1| \leq v$ we deduce, from (21),

$$\begin{aligned} \mathcal{B}_{m,l'}(f, x_0; x_1) &\leq \tau_{m,l',1}(v)\beta_{m,l'}(f; x_1) + \tau_{m,l',0}(v)\bar{\gamma}_m|x_0 - x_1|^2 \\ &\leq \tau_{m,l',1}(v)\beta_{m,l'}(f; x_1) + \frac{\tau_{m,l',0}(v)}{\mathcal{C}}\mathcal{B}_{m,l'}(f, x_0; x_1), \end{aligned}$$

hence (28). Inequalities (29), (30), and (31) follow in exactly the same way.

Let us prove (33). We distinguish two cases. First, if $z_0 = x_0$, then we deduce, from (31) and (28),

$$\beta_{m,l'}(f; x_0) \leq \bar{\kappa}_{m,l,l',\delta}(u, v)\underline{\kappa}_{m,l,l',\delta}(u, v)\beta_{m,l'}(f; x_1).$$

Since $\bar{\kappa}_{m,l,l',\delta}(u, v) \geq 1$ and $\underline{\kappa}_{m,l,l',\delta}(u, v) \geq 1$, we deduce

$$\beta_{m,l'}(f; z_0) \leq \bar{\kappa}_{m,l,l',\delta}(u, v)\underline{\kappa}_{m,l,l',\delta}(u, v) \min(\beta_{m,l'}(f; x_0), \beta_{m,l'}(f; x_1)). \quad (39)$$

On the other hand, if $z_0 = x_1$, applying successively (29), (30), and (28) we get

$$\begin{aligned}\beta_{m,l'}(f; x_1) &\leq \bar{\kappa}_{m,l,l',\delta}(u, v)\mathcal{B}_{m,l'}(f, x_0; x_1) \leq \bar{\kappa}_{m,l,l',\delta}(u, v)\mathcal{B}_{m,l'}(f, x_1; x_0) \\ &\leq \bar{\kappa}_{m,l,l',\delta}(u, v) \Big(\tau_{m,l',1}(v)\beta_{m,l'}(f; x_0) \\ &\qquad + \frac{\tau_{m,l',0}(v)}{\mathcal{C}}\underline{\kappa}_{m,l,l',\delta}(u, v)\beta_{m,l'}(f; x_1) \Big),\end{aligned}$$

so that using hypothesis (32) we deduce

$$\beta_{m,l'}(f; z_0) \leq \frac{\bar{\kappa}_{m,l,l',\delta}(u, v)\tau_{m,l',1}(v) \min(\beta_{m,l'}(f; x_0), \beta_{m,l'}(f; x_1))}{1 - \dfrac{\tau_{m,l',0}(v)}{\mathcal{C}}\underline{\kappa}_{m,l,l',\delta}(u, v)\bar{\kappa}_{m,l,l',\delta}(u, v)}. \quad (40)$$

Combining (39), (40), (27), $\underline{\kappa}_{m,l,l',\delta} \geq 1$, and $\bar{\kappa}_{m,l,l',\delta} \geq 1$ yields (33).

Now, let us examine case (c). By assumption (22) we deduce

$$\begin{aligned}\beta_{m,l'}(f; x_1) &\leq \tau_{m,l',1}(v)\mathcal{B}_{m,l'}(f, x_0; x_1) + \tau_{m,l',0}(v)\bar{\gamma}_m|x_0 - x_1|^2 \\ &\leq \bar{\mathcal{C}}\bar{\gamma}_m|x_1 - x_0|^2,\end{aligned}$$

which provides (34). Then we deduce

$$\alpha_{m,l'}(f; x_1) \leq \bar{\mathcal{C}}\bar{\gamma}_m^2|x_1 - x_0|^2 \leq \bar{\mathcal{C}}v^2.$$

In order to prove (36), we first check

$$\begin{aligned}\bar{\mathcal{C}} &\geq \mathcal{C} \geq C_{m,l,l',\delta}(u) \geq C_{m,l,l',\delta}(0) \\ &\geq \theta_{m,l,\delta}(2m - 1) \geq \frac{(m + 1)(2m - 1)}{(m - l + 1)(m - l)}.\end{aligned} \quad (41)$$

Then, squaring both sides of (35) gives

$$\bar{\mathcal{C}}v^2 \leq \frac{1}{9}\left(\frac{m}{m - l'}\right)^2 \frac{(m - l + 1)(m - l)}{(m + 1)(2m - 1)},$$

from which follows:

$$\begin{aligned}\frac{m - l'}{m}\frac{m + 1}{m + 1 - l'}\alpha_{m,l'}(f; x_1) &\leq \frac{1}{9}\frac{m + 1}{m + 1 - l'}\frac{m}{m - l'}\frac{(m - l + 1)(m - l)}{(m + 1)(2m - 1)} \\ &\leq \frac{1}{9}\frac{m + 1}{m + 1 - l}\frac{m}{m - l}\frac{(m - l + 1)(m - l)}{(m + 1)(2m - 1)} \\ &\leq \frac{1}{9}\frac{m}{2m - 1} \leq \frac{1}{9}.\end{aligned}$$

Then Corollary 1.8 implies the existence of a cluster Z_1 of zeros of $f^{(l')}$, and if ζ belongs to the convex hull of Z_1, then

$$|x_1 - \zeta| \leq 3\frac{m-l'}{m}\beta_{m,l'}(f; x_1) \leq 3\frac{m-l'}{m}\bar{\mathcal{C}}\bar{\gamma}_m|x_1 - x_0|^2.$$

Finally, inequality (37) follows from

$$|x_1 - x_0| \leq |x_1 - \zeta| + |x_0 - \zeta| \leq 3\frac{m-l'}{m}\bar{\mathcal{C}}\bar{\gamma}_m|x_1 - x_0|^2 + |x_0 - \zeta|,$$

hence

$$|x_1 - x_0| \leq \frac{|x_0 - \zeta|}{1 - 3\dfrac{m-l'}{m}\bar{\mathcal{C}}v}. \qquad \square$$

If $x_0 = \zeta$ is a multiple zero of multiplicity m, then conditions (25) and (35) are trivially satisfied with $r = 0$, $\bar{\gamma}_m = \gamma_m = \gamma_m(f; \zeta)$. In practice, the functions $\mathcal{B}_{m,l}$ that we use ensure that condition (32) also holds for $r = 0$. This shows by continuity that this theorem actually performs cluster approximation.

Informally speaking, if ζ belongs to the convex hull of the cluster of $f^{(l')}$, then Theorem 2.1, combined with Proposition 4.3, tells us that $\beta_{m,l'}$ is about the diameter of this cluster (assuming that $\alpha_m(f; \zeta)$ is sufficiently small). Thus, x_0 in case (a) and z_0 in case (b) are located at a distance from the cluster which is of the order of its diameter. We leave out the details here.

4.5. *Checking Conditions from the Initial Point*

Combining the point estimate cluster location criterion underlying Corollary 1.8 with the latter theorem and Proposition 4.3 makes it possible to ensure conditions (25), (32), (35), and

$$\eta_{m,l,l',\delta}(u, v)\bar{\gamma}_m r < 1, \tag{42}$$

from point estimates at x_0 only.

More precisely, let $x_0 \in U$ and $m \geq 1$ be an integer such that $f^{(m)}(x_0) \neq 0$. Let $l \in \{0, \ldots, m-1\}$, $l' \leq l$, and assume

$$\frac{m-l'}{m}\frac{m+1}{m+1-l'}\alpha_{m,l'}(f; x_0) \leq \tfrac{1}{9},$$

then $f^{(l')}$ admits a cluster of $m - l'$ zeros in $\bar{B}(x_0, 3[(m-l')/m]\beta_{m,l'}(f; x_0))$, according to Corollary 1.8. We take

$$\begin{aligned} \delta &:= 1, \\ r &:= 3\frac{m-l'}{m}\beta_{m,l'}(f;x_0), \\ \gamma_m &:= \frac{\gamma_m(f;x_0)}{(1-\gamma_m(f;x_0)r)\psi_m(\gamma_m(f;x_0)r)}, \\ \bar{\gamma}_m &:= \frac{\gamma_m}{(1-3\gamma_m r)\psi_m(3\gamma_m r)}. \end{aligned} \tag{43}$$

Let ζ be a point in the convex hull of this cluster. Assuming the stronger conditions $\gamma_m(f;x_0)r < 1-\left(\frac{1}{2}\right)^{1/(m+1)}$ and $3\gamma_m r < 1-\left(\frac{1}{2}\right)^{1/(m+1)}$, Proposition 4.3(c) yields $\gamma_m \geq \gamma_m(f;\zeta)$ and $\bar{\gamma}_m \geq \max_{z\in\bar{B}(\zeta,3r)} \gamma_m(f;z)$. One can then apply the previous theorem with these quantities.

According to formulas (43), it follows that the quantities u, $\bar{u}$, v, $\mathcal{C}$, $\bar{\mathcal{C}}$, and $\eta_{m,l,l',\delta}(u,v)\bar{\gamma}_m r$ depend only on $\alpha_{m,l'}(f;x_0)$. Therefore, if $\alpha_{m,l'}(f;x_0)$ is sufficiently small, then conditions (25), (35), and (42) are satisfied. In Section 6 we compute the suprema of the admissible values for $\alpha_{m,l'}(f;x_0)$ according to different approximation functions $\mathcal{B}_{m,l}$. As mentioned earlier, in all our cases condition (32) is always satisfied.

Once all these conditions are satisfied, we compute $x_1 := N_{m-l}(f^{(l)};x_0)$. Let us consider case (c) of Theorem 4.5: we are to show that $Z_1 = Z$. Thanks to (36), Corollary 1.8 applies: we deduce that the elements of Z_1 are the only zeros of $f^{(l')}$ in $\bar{B}(x_1,(m+1-l')/[3(m+1)\bar{\gamma}_m])$. On the other hand, for any $z \in Z$ we have $|z-x_1| \leq |x_0-x_1| + |z-x_0| \leq 3r$. Therefore it suffices to verify

$$3r \leq \frac{m+1-l'}{3(m+1)\bar{\gamma}_m} \quad \text{or, equivalently,} \quad v \leq \frac{2}{9}\frac{m+1-l'}{m+1}.$$

These inequalities are true if $m \geq 2$. They are a consequence of (35) and (41),

$$\begin{aligned} v &\leq \frac{1}{3}\frac{m}{m-l'}\frac{(m-l)(m+1-l)}{(m+1)(2m-1)} \leq \frac{1}{3}\frac{m}{2m-1}\frac{m+1-l'}{m+1} \\ &\leq \frac{2}{9}\frac{m+1-l'}{m+1}. \end{aligned}$$

Using (37) and (42), we deduce $|x_1-\zeta| \leq r$. Since this inequality holds for any ζ in the convex hull of Z we deduce $Z \subseteq \bar{B}(x_1,r)$.

Now consider the sequence $(x_k)_{k\in\mathbb{N}}$, formally defined by

$$x_{k+1} := N_{m-l}(f^{(l)};x_k).$$

Let K be the first integer k such that the stopping criterion is satisfied at x_k, that is, $f^{(l+1)}(x_k) = 0$ or $x_{k+1} \notin \bar{B}(x_k,2r)$ or $\mathcal{B}_{m,l'}(f,x_k;x_{k+1}) > \mathcal{G}|x_k-x_{k+1}|^2$. For all $k \leq K-1$, it follows by induction that x_{k+1} is well defined, that x_{k+1} belongs to $\bar{B}(\zeta,r)$, and that $Z \subseteq \bar{B}(x_{k+1},r)$.

Before presenting numerical experiments in Section 6, the next section describes two possible families of functions $\mathcal{B}_{m,l}$ and details the computations of upper bounds on γ_m.

5. Approximations of Point Estimates

Our approximation algorithm presented in the previous section requires upper bounds γ_m and $\bar{\gamma}_m$ as input, and performs several estimations of $\beta_{m,l}$ in a certain neighborhood of the cluster. In practice, input upper bounds are expected to be available from location criteria as explained previously. In this section we provide elementary devices for these point estimate computations. First of all, we give a purely numerical scheme to approximate $\beta_{m,l}$ by means of interpolation. In this way a valid $\mathcal{B}_{m,l}$ used in the next section is specified. After that we describe a mixed symbolic numerical scheme to compute upper bounds on γ_m for functions that depend polynomially in x and in some exponentials $\exp(a_k x)$, for certain complex numbers a_k.

5.1. *Approximation of $\beta_{m,l}$*

Now we design a purely numerical device for computing a valid $\mathcal{B}_{m,l}(f, y; z)$, introduced in the previous section.

Let m and z be such that $f^{(m)}(z) \neq 0$ and let $\gamma_m \geq \gamma_m(f; z)$. Consider a fixed integer $k \geq m+1$, let r be a positive real number, and w a kth primitive root of unity. Let q denote the interpolating polynomial of degree at most $k-1$, such that $q(z_j) = f(z_j)$, where $z_j = z + rw^j$ for $j \in \{0, \ldots, k-1\}$. Let $v := \gamma_m r$. The following proposition quantifies the approximation of $\beta_{m,l}(f; z)$ obtained from q.

Proposition 5.1. *According to the above notation, if $v + v^{k-m} < 1$, then we have*

$$\beta_{m,l}(f; z) \leq \beta_{m,l}(q; z)\left(1 + \frac{v^{k-m}}{1-v}\right) + \left(\frac{v^{k-m}}{1-v}\right)^{1/(m-l)} r,$$

$$\beta_{m,l}(q; z) \leq \frac{1-v}{1-v-v^{k-m}}\left(\beta_{m,l}(f; z) + \left(\frac{v^{k-m}}{1-v}\right)^{1/(m-l)} r\right).$$

In order to treat the case $r = 0$ in a continuous way we define the corresponding q as the limit of the interpolating polynomials when r tends to zero. This limit is nothing else but the truncated Taylor expansion of f at z. Remark that this case never occurs in the frame of the approximation algorithm of the previous section, unless one of the iterates is a zero.

It follows from these definitions that, taking $k := 2m - l$, $\mathcal{B}_{m,l}(f, y; z) := \beta_{m,l}(q; z)$ for q computed with $r := |y - z|$, we obtain a valid approximation function that satisfies (21) and (22), with

$$\tau_{m,l,1}(v) := 1 + \frac{v^{m-l}}{1-v-v^{m-l}} \quad \text{and} \quad \tau_{m,l,0}(v) := \tau_{m,l,1}(v)\left(\frac{1}{1-v}\right)^{1/(m-l)}, \tag{44}$$

$v_{m,l}$ being the first positive root of $v + v^{m-l} = 1$.

Proof of Proposition 5.1. We introduce the Vandermonde matrix:

$$V(w) := \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & w & \dots & w^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w^{k-1} & \dots & w^{(k-1)^2} \end{pmatrix}.$$

Now writing $q(x) := \sum_{i=0}^{k-1} q_i (x-z)^i$, the following relation holds:

$$V(w) \begin{pmatrix} q_0 \\ q_1 r \\ \vdots \\ q_{k-1} r^{k-1} \end{pmatrix} = V(w) \begin{pmatrix} f(z) \\ f'(z) r \\ \vdots \\ \dfrac{f^{(k-1)}(z)}{(k-1)!} r^{k-1} \end{pmatrix} + \begin{pmatrix} R_0 \\ \vdots \\ R_{k-1} \end{pmatrix},$$

where $R_j := \sum_{i \geq k} (f^{(i)}(z)/i!)(z_j - z)^i$. From the definition of z_j we deduce

$$\left| \frac{m!\, R_j}{f^{(m)}(z)} \right| \leq \frac{\gamma_m^{k-m} r^k}{1 - \gamma_m r} \quad \text{for } j \in \{0, \dots, k-1\}.$$

Then, using the classical fact $V(w)^{-1} = V(w^{-1})/k$, we deduce

$$q_i - \frac{f^{(i)}(z)}{i!} = \frac{1}{k} \sum_{j=0}^{k-1} \frac{R_j}{r^i} w^{-ji} \quad \text{for } i \in \{0, \dots, k-1\},$$

hence

$$\left| \frac{m!}{f^{(m)}(z)} \left(q_i - \frac{f^{(i)}(z)}{i!} \right) \right| \leq \frac{\gamma_m^{k-m} r^{k-i}}{1 - \gamma_m r}. \tag{45}$$

For $i = m$ (since $k \geq m+1$) we deduce

$$\left| \frac{m!\, q_m}{f^{(m)}(z)} - 1 \right| \leq \frac{v^{k-m}}{1-v}.$$

Using $v + v^{k-m} < 1$, it follows that $q_m \neq 0$ and

$$\left| \frac{f^{(m)}(z)}{m!\, q_m} \right| \leq \frac{1-v}{1 - v - v^{k-m}}. \tag{46}$$

Equation (45) leads to

$$\begin{aligned} \left| \frac{m!\, f^{(i)}(z)}{i!\, f^{(m)}(z)} \right| &\leq \left| \frac{q_i}{q_m} \right| \left| \frac{m!\, q_m}{f^{(m)}(z)} \right| + \left| \frac{m!}{f^{(m)}(z)} \left(\frac{f^{(i)}(z)}{i!} - q_i \right) \right| \\ &\leq \left| \frac{q_i}{q_m} \right| \left(1 + \frac{v^{k-m}}{1-v} \right) + \frac{v^{k-m}}{1-v} r^{m-i}, \end{aligned}$$

hence the first inequality. For the second one we deduce, in a similar way from (45) and (46),

$$\left|\frac{q_i}{q_m}\right| \le \left|\frac{f^{(m)}(z)}{m!\,q_m}\right| \left(\left|\frac{m!\,f^{(i)}(z)}{i!\,f^{(m)}(z)}\right| + \left|\frac{m!}{f^{(m)}(z)}\left(\frac{f^{(i)}(z)}{i!} - q_i\right)\right|\right)$$
$$\le \frac{1-v}{1-v-v^{k-m}}\left(\beta_{m,l}(f;z)^{m-i} + \frac{v^{k-m}}{1-v} r^{m-i}\right). \qquad \square$$

5.2. *Upper Bounds on γ_m*

In most practical cases, upper bounds on γ_m can be obtained from geometric majorant series manipulations: Appendix A gives rules for computing majorant series for products, compositions, derivatives, etc.

Designing a complete algorithm for upper bounding γ_m that covers a large class of analytic functions would lead us too far from the scope of this paper. In order to illustrate our methods in the next section, we restrict our analysis to the functions that depend polynomially on x and on a finite set of exponentials $\exp(a_k x)$ where the a_k are in $\mathbb{C}$. Any function f of this kind can be written in the following form:

$$f = p_1(x)\exp(a_1 x) + \cdots + p_n(x)\exp(a_n x),$$

where $p_1, \ldots, p_n$ are polynomials.

Since the exponential is an entire function, there exists geometric majorant series of the form $\lambda t/(1-\rho t)$ for any positive ρ. The following proposition explains the optimal corresponding value for λ.

Proposition 5.2. *For any positive $\rho \le 1$, we have $[\exp]_0 \le 1 + \lambda t/(1-\rho t)$, where $\lambda := 1/[K!\,\rho^{K-1}]$ and $K := \lfloor \rho^{-1} \rfloor$ (i.e., the largest integer less than or equal to ρ^{-1}).*

Proof. We rewrite the series inequality as $1/k! \le \lambda\rho^{k-1}$ for $k \ge 1$, and then as $\log(k!) \ge -\log(\lambda/\rho) + k\log(\rho^{-1})$, by taking logarithms. The function of i, defined by its graph G as the union of the segments $[(i, \log(i!)), (i+1)\log((i+1)!)]$ for $i \ge 0$, is convex of slope $\log(i+1)$ between abscissas i and $i+1$. Because the graph of the right-hand side of the previous inequality is a straight line with nonnegative slope, the largest possible value of $-\log(\lambda/\rho)$ is obtained by making this line tangent to G at K. This leads to the announced value for λ. $\square$

We introduce $\bar{a} := \max_{k\in\{1,\ldots,n\}} |a_k|$ and $\bar{\rho} := \bar{a}\rho$, for any ρ such that $0 < \rho \le 1$. Let λ denote the corresponding value given in the previous proposition. Applying the previous proposition and Proposition A.4, we deduce

$$[f]_z \le \left(1 + \frac{\lambda\bar{a}t}{1-\bar{\rho}t}\right)\sum_{i=1}^{n} |\exp(a_i z)|\,[p_i]_z =: P(t) + \frac{\bar{\lambda}}{1-\bar{\rho}t},$$

where $\bar{\lambda}$ and the polynomial P are obtained by means of Euclidean division by $1 - \bar{\rho}t$. Finally, the following proposition gives an upper bound on $\gamma_m(f; z)$ from the computation of the Taylor expansion of f at z at precision k and the above series inequality.

Proposition 5.3. *According to the above notation, let m be an integer such that $f^{(m)}(z) \neq 0$, $\sigma_m := m!/|f^{(m)}(z)|$, and let $p_m := \max(0, P^{(m)}(0)/m!)$. For $k \geq \max(m, \deg(P)) + 1$, let q denote the unique polynomial of degree at most $k-1$ such that $f(x) - q(x) \in \mathcal{O}_z((x-z)^k)$, then*

$$\gamma_m(q; z) \leq \gamma_m(f; z) \leq \max(\gamma_m(q; z), \bar{\rho}(\sigma_m(p_m + \bar{\lambda}\bar{\rho}^m))^{1/(k-m)}).$$

Proof. By construction we have $\sigma_m(p_m + \bar{\lambda}\bar{\rho}^m) \geq 1$, hence

$$\begin{aligned}\sup_{j \geq k} \left(\sigma_m \frac{|f^{(j)}(z)|}{j!}\right)^{1/(j-m)} &\leq \sup_{j \geq k} (\sigma_m \bar{\lambda}\bar{\rho}^j)^{1/(j-m)} \\ &\leq \sup_{j \geq k} \bar{\rho}(\sigma_m(p_m + \bar{\lambda}\bar{\rho}^m))^{1/(j-m)} \\ &\leq \bar{\rho}(\sigma_m(p_m + \bar{\lambda}\bar{\rho}^m))^{1/(k-m)}. \qquad \square\end{aligned}$$

For instance, let

$$f := \left(1 - \frac{14 - 3\imath}{20} x\right) \exp(x) + \left(1 - \frac{6 + 23\imath}{20} x - \frac{9 - 3\imath}{20} x^2\right) \exp(\imath x) - 2,$$

$m := 3$ and $z := 0.3$ (recall $\imath := \sqrt{-1}$). We have $\bar{a} = 1$; we take $\rho := 0.15$ and $k := 9$, and we get $\lambda \approx 18.3$, $\gamma_m(q; z) \approx 0.53$, and $\bar{\rho}(\sigma_m(p_m + \bar{\lambda}\bar{\rho}^m))^{1/(k-m)} \approx 0.31$. This way we obtain an accurate approximation of $\gamma_m(f; z)$.

6. Numerical Experiments

In this section we illustrate Theorem 4.5, of which we use the notation. We let $l := 0$, $l' := 0$, $\delta := 1$ and consider three families of examples, each parametrized by a real positive number N:

Example 1.

$$f := (x^m + 10^{-mN})(x^m - 1).$$

This polynomial admits a cluster of m roots around 0. The other roots are simple and lie on the unit circle.

Table 1. Critical values for the symbolic algorithm.

$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\alpha}_6$	$\hat{\alpha}_7$	$\hat{\alpha}_8$
0.0053	0.0040	0.0035	0.0032	0.0029	0.0027	0.0025	0.0024

Example 2.

$$\begin{aligned} f \;:=\; & \left(1 - \frac{14-3\imath}{20}x\right)\exp(x) \\ & + \left(1 - \frac{6+23\imath}{20}x - \frac{9-3\imath}{20}x^2\right)\exp(\imath x) - 2 + 10^{-3N}. \end{aligned}$$

This function admits a cluster of three zeros around 0. In order to get accurate upper bounds on γ_m we use Proposition 5.3 with $\rho := 0.15$ and $k := 9$.

Example 3.

$$f := \left(1 - \frac{2-\imath}{3}x\right)\exp(x) + \left(1 - \frac{1+4\imath}{3}x - \frac{2}{3}x^2\right)\exp(\imath x) - 2 + 10^{-4N}.$$

This function admits a cluster of four zeros around 0. In order to get accurate upper bounds on γ_m we use Proposition 5.3 with $\rho := 0.15$ and $k := 12$.

Computations are performed with the Maple computer algebra system version 7. The Digits environment variable controls the number of decimal digits that Maple uses when calculating with software long floating-point numbers. Heuristically, in order to avoid round-off problems, we set this variable to $2mN$. This problem is beyond the scope of the present work.

The starting point x_0 is taken as $\exp(\imath\pi/4)$ times the largest negative power of 2 that satisfies conditions (25), (32), (35) of Theorem 4.5 and also (42). These conditions are checked by computing point estimates at x_0 only, by means of formulas (43), as explained in Section 4.5.

Once all these conditions are satisfied we compute the sequence $(x_k)_{k\in\mathbb{N}}$ recursively defined by $x_{k+1} := N_m(f; x_k)$. As in Section 4.5, let K be the first integer k such that the stopping criterion is satisfied at x_k, that is, $f'(x_k) = 0$ or $x_{k+1} \notin \bar{B}(x_k, 2r)$ or $\mathcal{B}_{m,l'}(f, x_k; x_{k+1}) > \mathcal{G}|x_k - x_{k+1}|^2$. Recall that x_{k+1} is well defined while $k \leq K-1$ and belongs to $\bar{B}(\zeta, r)$.

At the end of the process, that is, for $k = K$, different situations occur. If x_{K+1} is not well defined or outside $\bar{B}(x_k, 2r)$, we indicate ∞ in the column $\beta_m(f; x_{K+1})$

Table 2. Critical values for the numerical algorithm.

$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\tilde{\alpha}_5$	$\tilde{\alpha}_6$	$\tilde{\alpha}_7$	$\tilde{\alpha}_8$
0.0037	0.0031	0.0028	0.0026	0.0024	0.0023	0.0021	0.0020

of Tables 3–8. Otherwise we indicate both $\beta_m(f; x_K)$ and $\beta_m(f; x_{K+1})$. If

$$\mathcal{B}_{m,l'}(f, x_{K+1}; x_K) < \mathcal{B}_{m,l'}(f, x_K; x_{K+1}),$$

then we underline the first one (resp., x_K), or else we underline the second one (resp., x_{K+1}).

Recall that this approximation scheme depends on the function $\mathcal{B}_{m,l'}$. We discuss a symbolic and then a purely numerical algorithm, corresponding to different choices of $\mathcal{B}_{m,l'}$.

Symbolic Algorithm. In this case we take $\mathcal{B}_{m,l'}(f, y; z) = \beta_{m,l'}(f; z)$ and perform the required computations of $\beta_{m,l'}$ by means of power series expansions. In Table 1, we provide approximations of the supremum values $\hat{\alpha}_m$ for $\alpha_m(f; x_0)$ having the following properties: if $\alpha_m(f; x_0) < \hat{\alpha}_m$, then conditions (25), (32), (35), and (42) are satisfied. The following approximations are obtained as the first positive zero $\hat{\alpha}_m$ of $\eta_{m,l,l',\delta}(u, v)\bar{\gamma}_m r = 1$, since this equation rewrites in terms of $\alpha_m(f; x_0)$ only (as observed in Section 4.5). We provide approximations of $\hat{\alpha}_m$ rounded toward zero at precision 10^{-4}. Sequences of iterates are presented in Tables 3 and 5.

Numerical Algorithm. Here we consider the function $\mathcal{B}_{m,l'}$, given in Section 5 according to formulas (44), that performs interpolation from the evaluations of f at $2m$ points; no high-order derivatives are required. We illustrate the numerical approach in the same way as in the previous symbolic case. We define critical values $\tilde{\alpha}_m$ for $\alpha_m(f; x_0)$ as previously; we report them in Table 2. As expected, these values are slightly smaller than in the symbolic case. Nevertheless, it is worth noting that they are of the same order of magnitude. Sequences of iterates are presented in Tables 4, 6, 7, and 8.

Comments on Tables. In all our examples the diameter of the cluster is about 10^{-N}. The election among the points x_K and x_{K+1} always returns the one inducing the smallest value for β_m. As expected, this value is of the order of the diameter of the cluster.

Example 1 displays a case of over-quadratic convergence. This is due to the vanishing of $f^{(m+1)}$ at 0. The same phenomenon occurs in Example 2, where $f^{(4)}(0)$ tends to 0 for large values of N. Note that all the cases of the algorithm actually appear in these example. Example 4 is the only one realizing a strict quadratic convergence, as observed in Table 8.

Last, note that the symbolic and numerical algorithms produce very close outputs in Tables 5 and 6.

Conclusion and Further Research

The main contribution of this paper lies in the precise analysis of the Newton–Schröder iteration with quadratic convergence toward clusters of zeros, together with a criterion for stopping this iteration when the cluster has been reached.

Table 3. Symbolic algorithm with Example 1 and $m = 2$.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f; x_K)$	$\beta_m(f; x_{K+1})$
4	$1.95 \cdot 10^{-3}$	0	$1.95 \cdot 10^{-3}$	$5.11 \cdot 10^{-6}$	$3.90 \cdot 10^{-3}$	$1.00 \cdot 10^{-4}$
8	$1.95 \cdot 10^{-3}$	1	$7.45 \cdot 10^{-9}$	$1.34 \cdot 10^{-8}$	$1.49 \cdot 10^{-8}$	$2.68 \cdot 10^{-8}$
16	$1.95 \cdot 10^{-3}$	2	$9.28 \cdot 10^{-25}$	$1.07 \cdot 10^{-8}$	$1.00 \cdot 10^{-16}$	$2.15 \cdot 10^{-8}$
32	$1.95 \cdot 10^{-3}$	2	$4.13 \cdot 10^{-25}$	$2.41 \cdot 10^{-40}$	$8.27 \cdot 10^{-25}$	$9.99 \cdot 10^{-33}$
64	$1.95 \cdot 10^{-3}$	3	$7.07 \cdot 10^{-74}$	$1.41 \cdot 10^{-55}$	$1.00 \cdot 10^{-64}$	$2.82 \cdot 10^{-55}$
128	$1.95 \cdot 10^{-3}$	3	$7.07 \cdot 10^{-74}$	$1.41 \cdot 10^{-183}$	$1.41 \cdot 10^{-73}$	$1.00 \cdot 10^{-128}$

Table 4. Numerical algorithm with Example 1 and $m = 2$.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f; x_K)$	$\beta_m(f; x_{K+1})$
4	$9.76 \cdot 10^{-4}$	0	$9.76 \cdot 10^{-4}$	$1.02 \cdot 10^{-5}$	$1.95 \cdot 10^{-3}$	$1.00 \cdot 10^{-4}$
8	$9.76 \cdot 10^{-4}$	1	$9.31 \cdot 10^{-10}$	$1.07 \cdot 10^{-7}$	$1.00 \cdot 10^{-8}$	$2.14 \cdot 10^{-7}$
16	$9.76 \cdot 10^{-4}$	1	$9.31 \cdot 10^{-10}$	$1.07 \cdot 10^{-23}$	$1.86 \cdot 10^{-9}$	$1.00 \cdot 10^{-16}$
32	$9.76 \cdot 10^{-4}$	2	$8.07 \cdot 10^{-28}$	$1.23 \cdot 10^{-37}$	$1.61 \cdot 10^{-27}$	$9.99 \cdot 10^{-33}$
64	$9.76 \cdot 10^{-4}$	3	$5.27 \cdot 10^{-82}$	$1.89 \cdot 10^{-47}$	$1.00 \cdot 10^{-64}$	$3.79 \cdot 10^{-47}$
128	$9.76 \cdot 10^{-4}$	3	$5.27 \cdot 10^{-82}$	$1.89 \cdot 10^{-175}$	$1.05 \cdot 10^{-81}$	$1.00 \cdot 10^{-128}$

Table 5. Symbolic algorithm with Example 1 and $m = 4$.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f; x_K)$	$\beta_m(f; x_{K+1})$
4	$4.88 \cdot 10^{-4}$	0	$4.88 \cdot 10^{-4}$	$8.58 \cdot 10^{-7}$	$1.95 \cdot 10^{-3}$	$9.99 \cdot 10^{-5}$
8	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{17}$	$1.00 \cdot 10^{-8}$	∞
16	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{-15}$	$1.11 \cdot 10^{-16}$	$1.87 \cdot 10^{-14}$
32	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{-79}$	$1.11 \cdot 10^{-16}$	$1.00 \cdot 10^{-32}$
64	$4.88 \cdot 10^{-4}$	2	$1.64 \cdot 10^{-83}$	$2.23 \cdot 10^{-8}$	$9.99 \cdot 10^{-65}$	$8.94 \cdot 10^{-8}$
128	$4.88 \cdot 10^{-4}$	2	$1.64 \cdot 10^{-83}$	$2.23 \cdot 10^{-264}$	$6.58 \cdot 10^{-83}$	$1.00 \cdot 10^{-128}$

Table 6. Numerical algorithm with Example 1 and $m = 4$.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f; x_K)$	$\beta_m(f; x_{K+1})$
4	$4.88 \cdot 10^{-4}$	0	$4.88 \cdot 10^{-4}$	$8.58 \cdot 10^{-7}$	$1.95 \cdot 10^{-3}$	$9.99 \cdot 10^{-5}$
8	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{17}$	$1.00 \cdot 10^{-8}$	∞
16	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{-15}$	$1.11 \cdot 10^{-16}$	$1.87 \cdot 10^{-14}$
32	$4.88 \cdot 10^{-4}$	1	$2.77 \cdot 10^{-17}$	$4.67 \cdot 10^{-79}$	$1.11 \cdot 10^{-16}$	$1.00 \cdot 10^{-32}$
64	$4.88 \cdot 10^{-4}$	2	$1.64 \cdot 10^{-83}$	$2.23 \cdot 10^{-8}$	$9.99 \cdot 10^{-65}$	$8.94 \cdot 10^{-8}$
128	$4.88 \cdot 10^{-4}$	2	$1.64 \cdot 10^{-83}$	$2.23 \cdot 10^{-264}$	$6.58 \cdot 10^{-83}$	$1.00 \cdot 10^{-128}$

Table 7. Numerical algorithm with Example 2.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f; x_K)$	$\beta_m(f; x_{K+1})$
4	$1.95 \cdot 10^{-3}$	0	$1.95 \cdot 10^{-3}$	$1.40 \cdot 10^{-6}$	$5.85 \cdot 10^{-3}$	$1.75 \cdot 10^{-4}$
8	$1.95 \cdot 10^{-3}$	1	$1.13 \cdot 10^{-9}$	$4.18 \cdot 10^{-6}$	$1.75 \cdot 10^{-8}$	$1.25 \cdot 10^{-5}$
16	$1.95 \cdot 10^{-3}$	1	$1.13 \cdot 10^{-9}$	$2.19 \cdot 10^{-28}$	$3.39 \cdot 10^{-9}$	$1.75 \cdot 10^{-16}$
32	$1.95 \cdot 10^{-3}$	2	$2.20 \cdot 10^{-28}$	$1.10 \cdot 10^{-40}$	$6.62 \cdot 10^{-28}$	$1.75 \cdot 10^{-32}$
64	$1.95 \cdot 10^{-3}$	3	$1.63 \cdot 10^{-84}$	$2.00 \cdot 10^{-24}$	$1.75 \cdot 10^{-64}$	$6.02 \cdot 10^{-24}$
128	$1.95 \cdot 10^{-3}$	3	$1.63 \cdot 10^{-84}$	$2.00 \cdot 10^{-216}$	$4.90 \cdot 10^{-84}$	$1.75 \cdot 10^{-128}$

Table 8. Numerical algorithm with Example 3.

N	$\|x_0\|$	K	$\|x_K\|$	$\|x_{K+1}\|$	$\beta_m(f;x_K)$	$\beta_m(f;x_{K+1})$
4	$9.76 \cdot 10^{-4}$	0	$9.76 \cdot 10^{-4}$	$\underline{6.40 \cdot 10^{-7}}$	$3.90 \cdot 10^{-3}$	$\underline{1.63 \cdot 10^{-4}}$
8	$9.76 \cdot 10^{-4}$	1	$1.45 \cdot 10^{-7}$	$\underline{2.34 \cdot 10^{-11}}$	$5.81 \cdot 10^{-7}$	$\underline{1.63 \cdot 10^{-8}}$
16	$9.76 \cdot 10^{-4}$	2	$3.21 \cdot 10^{-15}$	$\underline{2.16 \cdot 10^{-20}}$	$1.28 \cdot 10^{-14}$	$\underline{1.63 \cdot 10^{-16}}$
32	$9.76 \cdot 10^{-4}$	3	$1.57 \cdot 10^{-30}$	$\underline{1.84 \cdot 10^{-38}}$	$6.29 \cdot 10^{-30}$	$\underline{1.63 \cdot 10^{-32}}$
64	$9.76 \cdot 10^{-4}$	4	$3.77 \cdot 10^{-61}$	$\underline{1.34 \cdot 10^{-74}}$	$1.50 \cdot 10^{-60}$	$\underline{1.63 \cdot 10^{-64}}$
128	$9.76 \cdot 10^{-4}$	5	$2.16 \cdot 10^{-122}$	$\underline{7.09 \cdot 10^{-147}}$	$8.66 \cdot 10^{-122}$	$\underline{1.63 \cdot 10^{-128}}$

Our main motivation is the application of these methods for implicit functions computed in our extended algorithm for multivariate analytic maps [**12**]. There we consider maps whose Jacobian matrix has corank 1. Using the implicit function theorem, we decouple the study into a regular system and an analytic univariate function which concentrates the multiplicity. This function is then subjected to the methods developed here. This way, our stopping criterion is extended to this multivariate case.

Acknowledgments

We would like to thank Michael Shub and the anonymous referees for their helpful comments.

Appendix A. Majorant Series

For convenience, this appendix gathers the basic properties of majorant series, that are commonly known but spread in the literature. Even though we deal only with one complex variable map in this paper, all the results are stated with several complex variable maps, without inducing more difficulties. These results are used in [**12**] for several variables. We establish a list of basic properties, such as evaluation, composition, products, differentiation, etc. The results are systematically stated in their most general form before being specialized into *geometric majorant series*.

Majorant series techniques belong to the "point de vue de Weierstrass," as mentioned by Henri Cartan in [**3**, see Chap. 1 and pp. 218–225]. These are a crucial tool for the effective manipulation of power series expansions, that lie at the heart of Smale's α-theory. From the computational viewpoint, only a finite number of derivatives can be computed at any point, while estimates on the convergence of Newton's iteration require more global information. The idea is that sufficient information can be provided by a majorant series of the Taylor expansion. For practical applications we will make use of geometric majorant series, that can be represented very compactly.

All the vector spaces we consider are finite dimensional over $\mathbb{C}$. Let f denote an analytic map from an open subset U of a normed vector space $\mathbb{E}$ (containing a) to another normed vector space $\mathbb{F}$. Then the kth derivative $D^k f(a)$ of f at a belongs to the set $\mathcal{L}_k(\mathbb{E}; \mathbb{F})$ of $\mathbb{C}$-multilinear maps from k copies of $\mathbb{E}$ to $\mathbb{F}$. Let $\mathbb{E}_1, \ldots, \mathbb{E}_l$ denote normed vector spaces, the *norm* we use on the space $\mathcal{L}(\mathbb{E}_1, \ldots, \mathbb{E}_l; \mathbb{F})$ of $\mathbb{C}$-multilinear maps from $\mathbb{E}_1 \times \cdots \times \mathbb{E}_l$ to $\mathbb{F}$ is defined by

$$\|L\| := \sup_{\substack{u_1 \in \mathbb{E}_1, \ldots, u_l \in \mathbb{E}_l \\ \|u_1\| = \cdots = \|u_l\| = 1}} \|L(u_1, \ldots, u_l)\|,$$

for any $L \in \mathcal{L}(\mathbb{E}_1, \ldots, \mathbb{E}_l; \mathbb{F})$. When no confusion is possible we use the same notation $\|\cdot\|$ for norms over different spaces. Throughout this appendix, $\mathbb{E}$, $\mathbb{F}$, $\mathbb{G}, \ldots$ denote normed vector spaces.

A.1. *Partial Order over Series*

We consider the following partial order $\leq$ over $\mathbb{R}\{t\}$. Let F and G be in $\mathbb{R}\{t\}$, we write $F \leq G$ when $F^{(k)}(0) \leq G^{(k)}(0)$ for all $k \geq 0$. Then we say that a power series $F \in \mathbb{R}\{t\}$ is a *majorant series* for an analytic map f at a point a if $[f]_a \leq F$.

With several variables, the quantity γ is given by

$$\gamma(f; x) := \sup_{k \geq 2} \left\| Df(x)^{-1} \frac{D^k f}{k!}(x) \right\|^{1/(k-1)},$$

and one of its important roles stems from the following inequality:

$$[Df(a)^{-1}(f - f(a))]_a \leq \sum_{k \geq 1} \gamma(f; a)^{k-1} t^k = \frac{t}{1 - \gamma(f; a)t},$$

which means that the rational function in the right-hand side represents a majorant series for $Df(a)^{-1}(f - f(a))$ at a.

Proposition A.1. *The partial order over $\mathbb{R}\{t\}$ defined above satisfies the following compatibility rules*:

(1) *for all nonnegative x in $\mathbb{R}$, one has $x \geq 0$, when x is seen in $\mathbb{R}\{t\}$*;
(2) *for all F in $\mathbb{R}\{t\}$, $F \geq 0$ is equivalent to $-F \leq 0$*;
(3) *for all F, G, and H in $\mathbb{R}\{t\}$, if $F \leq G$, then $F + H \leq G + H$*;
(4) *for all F, G, and H in $\mathbb{R}\{t\}$, if $F \leq G$, and $H \geq 0$, then $FH \leq GH$*;
(5) *for all F, G, P, and Q in $\mathbb{R}\{t\}$, if $0 \leq F \leq G$ and $0 \leq P \leq Q$, then $FP \leq GQ$.*

Proof. Part (5) follows directly from part (4), which is the only one not completely straightforward. Denoting by F_i, G_i, and H_i the ith coefficient of F, G, and H, respectively, we can write the ith coefficient of FH as $\sum_{k+l=i} F_k H_l$ which can be bounded term by term by $\sum_{k+l=i} G_k H_l$ (since all the H_l are nonnegative). ◻

The map $[\cdot]_a$ is defined on the set of analytic maps at a and has values in $\mathbb{R}\{t\}$. Let $a \in \mathbb{E}$ and let f and g be analytic maps defined in a neighborhood of a, the basic properties of $[\cdot]_a$ are:

(1) $[f]_a \geq 0$;
(2) $[f]_a = 0$ is equivalent to $f = 0$ in a neighborhood of a;
(3) $[cf]_a = |c|\,[f]_a$ for all $c \in \mathbb{C}$;
(4) $[f + g]_a \leq [f]_a + [g]_a$.

All of these are direct consequences of the definition.

A.2. Geometric Majorant Series

One of the aims of this appendix is to provide a toolbox for the practical use of majorant series. For this purpose most of our general statements on majorant series will be specialized into a tractable subclass of $\mathbb{R}\{t\}$. This subclass is composed of *geometric majorant series*, that are series of the form $\lambda t \sum_{k\geq 0} \rho^k t^k = \lambda t/(1 - \rho t)$ with $\lambda \geq 0$ and $\rho \geq 0$.

Indeed, most of the properties of the γ quantity can be recovered by easy computations on geometric series that are summarized in the following proposition:

Proposition A.2. *The series $[f - f(a)]_a$ and the constant $\gamma(f; a)$ are related via*

$$[f - f(a)]_a \leq \frac{\lambda t}{1 - \rho t} \quad \Rightarrow \quad \gamma(f; a) \leq \|Df(a)^{-1}\|\lambda\rho$$

and

$$[Df(a)^{-1}(f - f(a))]_a \leq \frac{t}{1 - \gamma(f; a)t}.$$

Keeping track of *two* quantities λ and ρ instead of one γ not only allows one to deal with situations when γ is not defined (or is infinite) but also leads to better bounds in some cases.

A.3. Evaluation

The first useful elementary property concerns the compatibility of the partial ordering on majorant series with evaluation.

Proposition A.3. *Let f be an analytic map defined on a connected open subset U of $\mathbb{E}$ and with values in $\mathbb{F}$, a and b are two points in U and $F \in R\{t\}$ such that $[f]_a \leq F$. Let r denote the radius of convergence of F then, for any $b \in U$ such that $\|b - a\| < r$, we have $\|f(b)\| \leq F(\|a - b\|)$.*

Proof. First observe that the radius of convergence of the power series expansion of f at a is at least r. Bounding this series term by term leads to

$$\|f(b)\| \leq \sum_{k\geq 0} \frac{\|D^k f(a)\|}{k!} \|b-a\|^k \leq F(\|b-a\|). \qquad \square$$

A.4. Linear Maps

Now we examine the behavior of majorant series under linear maps.

Proposition A.4. *Let $\mathbb{E}, \mathbb{F}, \mathbb{E}_1, \ldots, \mathbb{E}_l, \mathbb{G}_1, \ldots, \mathbb{G}_l$, be normed vector spaces, let U be an open neighborhood of a in $\mathbb{E}$, and let f be an analytic map from U to $\mathcal{L}(\mathbb{E}_1, \ldots, \mathbb{E}_l; \mathbb{F})$. For $i = 1, \ldots, l$, let g_i be analytic maps from U to $\mathcal{L}(\mathbb{G}_i; \mathbb{E}_i)$. Let $F, G_1, \ldots, G_l$ in $\mathbb{R}\{t\}$ be such that $[f]_a \leq F$, $[g_1]_a \leq G_1, \ldots, [g_l]_a \leq G_l$. Let h be defined by*

$$\begin{aligned} h\colon U &\to \mathcal{L}(\mathbb{G}_1, \ldots, \mathbb{G}_l; \mathbb{F}), \\ x &\mapsto f(x)(g_1(x), \ldots, g_l(x)). \end{aligned}$$

This map is analytic on U and

$$[h]_a \leq F G_1 \cdots G_l.$$

Proof. Let $u \in \mathbb{E}^k$, then the kth derivative of h at u becomes

$$D^k h(a)(u) = \sum_N D^{|N_0|} f(a) p_{N_0}(u)(D^{|N_1|} g_1(a) p_{N_1}(u), \ldots, D^{|N_l|} g_l(a) p_{N_l}(u)),$$

where the sum is taken over all partitions $N = \{N_0, \ldots, N_l\}$ of the set $\{1, \ldots, k\}$ into possibly empty disjoint subsets. Here $|N_r|$ denotes the cardinality of N_r for $r = 0, \ldots, l$, and p_{N_r} represents the canonical projection from $\mathbb{E}^k$ to $\mathbb{E}^{|N_r|}$ defined by $p_{N_r}(u_1, \ldots, u_k) := (u_m\colon m \in N_r)$. It follows that

$$\|D^k h(a)\| \leq \sum_N \|D^{|N_0|} f(a)\| \|D^{|N_1|} g_1(a)\| \cdots \|D^{|N_l|} g_l(a)\|$$

and

$$\begin{aligned} \left\| \frac{D^k h}{k!}(a) \right\| &\leq \sum_{\nu_0+\cdots+\nu_l=k} \left\| \frac{D^{\nu_0} f(a)}{\nu_0!} \right\| \left\| \frac{D^{\nu_1} g_1(a)}{\nu_1!} \right\| \cdots \left\| \frac{D^{\nu_l} g_l(a)}{\nu_l!} \right\| \\ &\leq \sum_{\nu_0+\cdots+\nu_l=k} \frac{F^{(\nu_0)}(0)}{\nu_0!} \frac{G_1^{(\nu_1)}(0)}{\nu_1!} \cdots \frac{G_l^{(\nu_l)}(0)}{\nu_l!} \\ &= \frac{(F G_1 \cdots G_l)^{(k)}(0)}{k!}. \end{aligned} \qquad \square$$

In particular, it follows that the product of majorant series of univariate analytic functions is a majorant series for their product.

A.5. Composition

The next formula shows that majorant series behave well under composition.

Proposition A.5. *Let U be an open neighborhood of a in $\mathbb{E}$, let f be an analytic map from U to $\mathbb{F}$, and let g be another analytic map from a neighborhood of $f(U)$ to $\mathbb{G}$. Let F and G in $\mathbb{R}\{t\}$ be such that $[f]_a \leq F$ and $[g]_{f(a)} \leq G$. Then $g \circ f$ is analytic on U and*

$$[g \circ f]_a \leq G \circ (F - F(0)).$$

Proof. The proof is based on writing an explicit expression for $D^k h$, where $h := g \circ f$, and on bounding the norms of each summand. The explicit expression is provided by the multivariate version of Faà di Bruno's formula (see, for instance, [**10**], which also deals with the infinite-dimensional case). Let $u \in \mathbb{E}^k$,

$$D^k h(a)(u) = \sum_{i=1}^{k} \sum_{N} D^i g(f(a))(D^{|N_1|} f(a) p_{N_1}(u), \ldots, D^{|N_i|} f(a) p_{N_i}(u)),$$

where the second sum is taken over all the partitions $N = \{N_1, \ldots, N_i\}$ of the set $\{1, \ldots, k\}$ and the p_{N_j} are the same as in the proof of the previous proposition. It follows that

$$\begin{aligned} \|D^k h(a)\| &\leq \sum_{i=1}^{k} \sum_{N} \|D^i g(f(a))\| \|D^{|N_1|} f(a)\| \cdots \|D^{|N_i|} f(a)\| \\ &\leq \sum_{i=1}^{k} \sum_{N} G^{(i)}(0) F^{(|N_1|)}(0) \cdots F^{(|N_i|)}(0) \\ &= (G \circ (F - F(0)))^{(k)}(0). \end{aligned}$$

□

Corollary A.6 (For Geometric Majorant Series). *If we have $[f - f(a)]_a \leq \lambda_f t / (1 - \rho_f t)$, $[g - g(f(a))]_{f(a)} \leq \lambda_g t/(1 - \rho_g t)$, and $h = g \circ f$, then*

$$[h - h(a)]_a \leq \frac{\lambda t}{1 - \rho t}, \qquad \text{where} \quad \lambda := \lambda_f \lambda_g, \quad \rho := \rho_f + \lambda_f \rho_g,$$

and $f(B(a, 1/\rho)) \subseteq B(f(a), 1/\rho_g)$.

Proof. The majorant series for h follows from the previous proposition via simple calculations. The ball inclusion follows from evaluating the majorant series of f

via Proposition A.3: for any $z \in B(a, 1/\rho)$ we obtain

$$\|f(z) - f(a)\| < \frac{\lambda_f/\rho}{1 - \rho_f/\rho} = 1/\rho_g. \qquad \square$$

A.6. Inversion

The next formula gives a majorant series for linear map inversion.

Lemma A.7. *Let U be an open neighborhood of* Id *in $\mathcal{L}(\mathbb{F}; \mathbb{F})$ containing only invertible linear maps, and let f denote the analytic map from U to $\mathcal{L}(\mathbb{F}; \mathbb{F})$ mapping x to x^{-1}. Then one has*

$$[f]_{\mathrm{Id}} = \frac{1}{1-t}.$$

Proof. A straightforward induction gives

$$D^k f(x)(u_1, \ldots, u_k) = (-1)^k \sum_{\sigma} f(x) u_{\sigma(1)} f(x) u_{\sigma(2)} \cdots f(x) u_{\sigma(k)} f(x),$$

for all $k \geq 1$ and all $u_i \in \mathcal{L}(\mathbb{F}; \mathbb{F})$, where the sum is taken over all the permutations σ of the set $\{1, \ldots, k\}$. We deduce $\|D^k f(\mathrm{Id})\| = k!$, which corresponds to the claimed formula. $\square$

Combined with Proposition A.5 we deduce:

Proposition A.8. *Let U be an open neighborhood of a in $\mathbb{E}$, and let f be analytic from U to $\mathcal{L}(\mathbb{F}; \mathbb{F})$ such that $f(a) = \mathrm{Id}$. Let $F \in \mathbb{R}\{t\}$ be such that $[f]_a \leq F$, then one has*

$$[f^{-1}]_a \leq \frac{1}{1 + F(0) - F}.$$

In addition, the radius of convergence of $1/(1 + F(0) - F(t))$ is at least

$$\bar{\rho} := \sup(s < \rho\colon\ 1 + F(0) - F(r) > 0, \textit{ for all } r \in [0, s]),$$

where ρ denotes the radius of convergence of F.

Proof. For any $z \in B(0, \bar{\rho})$, the triangular inequality and $F \geq 0$ imply

$$|1 + F(0) - F(z)| \geq 1 - |F(z) - F(0)| \geq 1 - F(|z|) + F(0) > 0,$$

which means that $z \mapsto 1 + F(0) - F(z)$ does not vanish in $B(0, \bar{\rho})$, whence the formula for the radius $\bar{\rho}$. $\square$

A.7. Differentiation

If $\mathbb{G}$ is a subspace of $\mathbb{E}$, we write $Df_{|\mathbb{G}}$ for the restriction of Df to $\mathbb{G}$.

Proposition A.9. *Let $\mathbb{G}$ be a subspace of $\mathbb{E}$, f an analytic map defined in a neighborhood of $a \in \mathbb{E}$, and let $F \in \mathbb{R}\{t\}$ be such that $[f]_a \leq F$, then*

$$[Df_{|\mathbb{G}}]_a \leq F'.$$

Proof. Since $[Df_{|\mathbb{G}}]_a \leq [Df]_a$, it is sufficient to consider the case when $\mathbb{G} = \mathbb{E}$,

$$\begin{aligned}\|D^k(Df)(a)\| &= \sup_{\|u_1\|=\cdots=\|u_k\|=1} \|D^k(Df)(a)u_1 \ldots u_k\| \\ &= \sup_{\|u_1\|=\cdots=\|u_k\|=1} \sup_{\|v\|=1} \|(D^k(Df)(a)u_1 \ldots u_k)v\| = \|D^{k+1}f(a)\|.\end{aligned}$$

Observe that the choice of the norm for multilinear maps is crucial here. □

Corollary A.10 (For Geometric Majorant Series). *Let $m \geq 1$ and $j \geq 0$ be two integers. If $[f - f(a)]_a \leq \lambda t^m/(1-\rho t)$ with $\rho > 0$, then, for any integer $k \geq 1$ and any $R > \rho$,*

$$\frac{[D^k f]_a}{k!} \leq \lambda \sum_{i=\max(0,m-k)}^{j} \binom{i+k}{k} \rho^{i+k-m} t^i + \frac{\Lambda t^{j+1}}{1-Rt},$$

where

$$\Lambda := \lambda \rho^{k-m} \binom{s+k-1}{k} \left(\frac{\rho}{R}\right)^{s-1} R^{j+1} \qquad \text{and} \qquad s := \left\lceil \frac{k\rho}{R-\rho} \right\rceil.$$

Here $\lceil r \rceil$ denotes the smallest integer larger than or equal to the real number r.

Proof. Thanks to the previous proposition, it is sufficient to consider the case $f = \lambda t^m/(1-\rho t)$. Then, a straightforward computation yields

$$\begin{aligned}\frac{f^{(k)}(t)}{k!} &= \lambda \left(\sum_{i\geq 0} \rho^i t^{m+i}\right)^{(k)} \\ &= \lambda \sum_{i\geq \max(0,k-m)} \binom{m+i}{k} \rho^i t^{i+m-k} \\ &= \lambda \sum_{i\geq \max(0,m-k)} \binom{k+i}{k} \rho^{i+k-m} t^i.\end{aligned}$$

Fixing an index j, it becomes a matter of bounding the sum

$$\lambda \sum_{i \geq j+1} \binom{k+i}{k} \rho^{i+k-m} t^i$$

by a geometric series of type $\Lambda t^{j+1}/(1 - Rt)$. There is some amount of freedom since we can choose both the initial value (the final Λ) and the ratio (the final R). By taking logarithms, we see that our problem is equivalent to ensuring the following inequality:

$$\log \binom{k+i}{k} \leq i \log \left(\frac{R}{\rho}\right) + \log \left(\frac{\Lambda}{\lambda \rho^{k-m} R^{j+1}}\right), \qquad i \geq j+1.$$

In view of the asymptotic behavior for large i, it is necessary that $R > \rho$. Then we consider the function of i defined by its graph as the union of the segments $[(i, \log \binom{k+i}{k}), (i+1, \log \binom{k+i+1}{k})]$ for $i \geq 0$. This is a concave piecewise linear function, the segment between abscissas i and $i+1$ has slope $\log(1 + k/(i+1))$. Because the graph of the function in the right-hand side of the previous inequality is a straight line with positive slope, then, for a given R, the smallest possible value for Λ is obtained by making this line tangent to the concave piecewise linear function at $i = s - 1$ with s as given in the statement of the corollary and this leads to the announced value for Λ.

A.8. *Translation*

The next formula shows that majorant series also behave well under translation.

Proposition A.11. *Let U be an open connected neighborhood of a in $\mathbb{E}$, let f be an analytic map from U to $\mathbb{F}$, and let $F \in \mathbb{R}\{t\}$ be such that $[f]_a \leq F$. Then, for any $b \in U$ such that $\|b - a\|$ is less than the radius of convergence of F, we have $[f]_b \leq F(\|a - b\| + t)$.*

Proof. For any $k \geq 0$, successively applying Propositions A.3 and A.9, we obtain $\|D^k f(b)\|/k! \leq F^{(k)}(\|a - b\|)/k!$ and therefore

$$[f]_b \leq \sum_{k \geq 0} \frac{F^{(k)}(\|a - b\|)}{k!} t^k = F(\|a - b\| + t). \qquad \square$$

Corollary A.12 (For Geometric Majorant Series). *Let U be an open connected neighborhood of a in $\mathbb{E}$, let f be an analytic map from U to $\mathbb{F}$, and let $\lambda_a \geq 0$ and $\rho_a \geq 0$ be such that*

$$[f - f(a)]_a \leq \frac{\lambda_a t}{1 - \rho_a t}.$$

Then, for all $b \in B(a, 1/\rho_a) \cap U$, *we have*

$$[f - f(b)]_b \leq \frac{\lambda_b t}{1 - \rho_b t},$$

with $\lambda_b := \lambda_a/(1 - \rho_a\|a - b\|)^2$ *and* $\rho_b := \rho_a/(1 - \rho_a\|a - b\|)$.

Proof. We apply the previous proposition and check that $\lambda_a t/(1 - \rho_a t)$ yields $\lambda_a\|a - b\|/(1 - \rho_a\|a - b\|) + \lambda_b t/(1 - \rho_b t)$ when evaluated at $\|a - b\| + t$. □

References

[1] D. A. Bini and G. Fiorentino, Design, analysis, and implementation of a multiprecision polynomial rootfinder, *Numer. Algorithms* **23**, 2–3 (2000), 127–173.

[2] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, 1998.

[3] H. Cartan, *Elementary Theory of Analytic Functions of One or Several Complex Variables*, Éditions Scientifiques, Hermann, Paris, 1963. Translated from the French original edition, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Enseignement des Sciences, Hermann, Paris, 1961.

[4] D. Coppersmith and C. A. Neff, Roots of a polynomial and its derivatives, in *Proceedings of the Fifth Annual ACM–SIAM Symposium on Discrete Algorithms* (Arlington, VA, 1994), ACM Press, New York, 1994, pp. 271–279.

[5] D. W. Decker, H. B. Keller, and C. T. Kelley, Convergence rates for Newton's method at singular points, *SIAM J. Numer. Anal.* **20**, 2 (1983), 296–314.

[6] D. W. Decker and C. T. Kelley, Newton's method at singular points, I, *SIAM J. Numer. Anal.* **17**, 1 (1980), 66–70.

[7] D. W. Decker and C. T. Kelley, Convergence acceleration for Newton's method at singular points, *SIAM J. Numer. Anal.* **19**, 1 (1982), 219–229.

[8] D. W. Decker and C. T. Kelley, Expanded convergence domains for Newton's method at nearly singular roots, *SIAM J. Sci. Statist. Comput.* **6**, 4 (1985), 951–966.

[9] J.-P. Dedieu, Estimations for the separation number of a polynomial system, *J. Symbolic Comput.* **24**, 6 (1997), 683–693.

[10] L. E. Fraenkel, Formulae for high derivatives of composite functions, *Math. Proc. Cambridge Philos. Soc.* **83**, 2 (1978), 159–165.

[11] M. Frontini and E. Sormani, Modified Newton's method with third-order convergence and multiple roots, *J. Comput. Appl. Math.* **156**, 2 (2003), 345–354.

[12] M. Giusti, G. Lecerf, B. Salvy, and J.-C. Yakoubsohn, On location and approximation of clusters of zeros: Case of embedding dimension one, Manuscript, 2004.

[13] A. Griewank, Starlike domains of convergence for Newton's method at singularities, *Numer. Math.* **35**, 1 (1980), 95–111.

[14] A. Griewank, On solving nonlinear equations with simple singularities or nearly singular solutions, *SIAM Rev.* **27**, 4 (1985), 537–563.

[15] A. Griewank and M. R. Osborne, Analysis of Newton's method at irregular singularities, *SIAM J. Numer. Anal.* **20**, 4 (1983), 747–773.

[16] V. Hribernig and H. J. Stetter, Detection and validation of clusters of polynomial zeros, *J. Symbolic Comput.* **24**, 6 (1997), 667–681.

[17] J. Hubbard, D. Schleicher, and S. Sutherland, How to find all roots of complex polynomials by Newton's method, *Invent. Math.* **146**, 1 (2001), 1–33.

[18] R. F. King, Improving the Van de Vel root-finding method, *Computing* **30**, 4 (1983), 373–378.

[19] P. Kirrinnis, Newton iteration towards a cluster of polynomial zeros, in *Foundations of Computational Mathematics* (*Rio de Janeiro*, 1997), Springer-Verlag, New York, 1997, pp. 193–215.

[20] P. Kirrinnis, Partial fraction decomposition in $\mathbb{C}(z)$ and simultaneous Newton iteration for factorization in $\mathbb{C}[z]$, *J. Complexity* **14**, 3 (1998), 378–444.

[21] P. Kravanja and A. Haegemans, A modification of Newton's method for analytic mappings having multiple zeros, *Computing* **62**, 2 (1999), 129–145.

[22] P. Kravanja, T. Sakurai, and M. Van Barel, On locating clusters of zeros of analytic functions, *BIT* **39**, 4 (1999), 646–682.

[23] P. Kravanja, T. Sakurai, and M. Van Barel, Error analysis of a derivative-free algorithm for computing zeros of holomorphic functions, *Computing* **70**, 4 (2003), 335–347.

[24] P. Kravanja and M. Van Barel, *Computing the Zeros of Analytic Functions*, Lecture Notes in Mathematics, Vol. 1727, Springer-Verlag, New York, 2000.

[25] F. Lucas, Sur une application de la Mécanique rationnelle à la théorie des équations, *C. R. Hebdomadaires Séances Acad. Sci. LXXXIX* (1879), 224–226.

[26] M. Marden, *Geometry of Polynomials*, 2nd ed., Mathematical Surveys and Monographs, No. 3, American Mathematical Society, Providence, RI, 1966.

[27] J. M. McNamee, A bibliography on roots of polynomials, *J. Comput. Appl. Math.* **47**, 3 (1993), 391–394.

[28] J. M. McNamee, A 2002 update of the supplementary bibliography on roots of polynomials, *J. Comput. Appl. Math.* **142**, 2 (2002), 433–434. http://www.yorku.ca/mcnamee/.

[29] G. V. Milovanović and T. M. Rassias, Distribution of zeros and inequalities for zeros of algebraic polynomials, in *Functional Equations and Inequalities*, Math. Appl., Vol. 518, Kluwer Academic, Dordrecht, 2000, pp. 171–204.

[30] A. Neff and J. H. Reif, An $\mathcal{O}(n^{1+\varepsilon} \log b)$ algorithm for the complex roots problem, in *Proceedings of the* 35*th Annual IEEE Conference on Foundations of Computer Science* (FOCS '94), IEEE Computer Society Press, Piscataway, NJ, pp. 540–547.

[31] C. A. Neff and J. H. Reif, An efficient algorithm for the complex roots problem, *J. Complexity* **12**, 2 (1996), 81–115.

[32] A. Neumaier, Enclosing clusters of zeros of polynomials, *J. Comput. Appl. Math.* **156**, 2 (2003), 389–401.

[33] A. M. Ostrowski, *Solution of Equations and Systems of Equations*, Academic Press, New York, 1966.

[34] V. Y. Pan, Optimal (up to polylog factors) sequential and parallel algorithms for approximating complex polynomial zeros, in *STOC* '95: *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, ACM Press, New York, 1995, pp. 741–750.

[35] V. Y. Pan, Optimal and nearly optimal algorithms for approximating polynomial zeros, *Comput. Math. Appl.* **31**, 12 (1996), 97–138.

[36] V. Y. Pan, Solving a polynomial equation: Some history and recent progress, *SIAM Rev.* **39**, 2 (1997), 187–220.

[37] V. Y. Pan, Approximating complex polynomial zeros: Modified Weyl's quadtree construction and improved Newton's iteration, *J. Complexity* **16**, 1 (2000), 213–264. Real computation and complexity (Schloss Dagstuhl, 1998).

[38] V. Y. Pan, Univariate polynomials: Nearly optimal algorithms for numerical factorization and root-finding, *J. Symbolic Comput.* **33**, 5 (2002), 701–733.

[39] A.-E. Pellet, Sur un mode de séparation des racines des équations et la formule de Lagrange, *Bull. Sci. Math. Astrom.* **5** (1881), 393–395. Deuxième Série, Tome 16 de la Collection.

[40] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, London Mathematical Society Monographs (N.S.), Vol. 26, Clarendon Press, Oxford, and Oxford University Press, New York, 2002.

[41] L. B. Rall, Convergence of the Newton process to multiple solutions, *Numer. Math.* **9** (1966), 23–37.

[42] G. W. Reddien, On Newton's method for singular problems, *SIAM J. Numer. Anal.* **15**, 5 (1978), 993–996.

[43] G. W. Reddien, Newton's method and high order singularities, *Comput. Math. Appl.* **5**, 2 (1979), 79–86.

[44] J. Renegar, On the worst-case arithmetic complexity of approximating zeros of polynomials, *J. Complexity* **3**, 2 (1987), 90–113.

[45] E. Rouché, Mémoire sur la série de Lagrange, *J. École Impériale Polytechnique* **22**, 39 (1862), 193–224.

[46] S. M. Rump, Ten methods to bound multiple roots of polynomials, *J. Comput. Appl. Math.* **156**, 2 (2003), 403–432.

[47] A. Schönhage, The fundamental theorem of algebra in terms of computational complexity, Technical report, Preliminary Report of Mathematisches Institut der Universität Tübingen, Germany, 1982.

[48] E. Schröder, Über unendlich viele Algorithmen zur Auflösung der Gleichungen, *Math. Ann.* **2** (1870), 317–365.

[49] M. Shub and S. Smale, Computational complexity, On the geometry of polynomials and a theory of cost, I, *Ann. Sci. École Norm. Sup.* (4) **18**, 1 (1985), 107–142.

[50] M. Shub and S. Smale, Computational complexity: On the geometry of polynomials and a theory of cost, II, *SIAM J. Comput.* **15**, 1 (1986), 145–161.

[51] S. Smale, The fundamental theorem of algebra and complexity theory, *Bull. Amer. Math. Soc.* (*N.S.*) **4**, 1 (1981), 1–36.

[52] S. Smale, Newton method estimates from data at one point, in *In the Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics* (R. E. Ewing, K. I. Gross, and C. F. Martin, eds.), Springer-Verlag, New York, 1986, pp. 185–196.

[53] A. Terui and T. Sasaki, "Approximate zero-points" of real univariate polynomials with large error terms, *IPSJ J.* **41**, 4 (2000), 974–989.

[54] H. Van de Vel, A method for computing a root of a single nonlinear equation, including its multiplicity, *Computing* **14**, 1–2 (1975), 167–171.

[55] M. Vander Straeten and H. Van de Vel, Multiple root-finding methods, *J. Comput. Appl. Math.* **40**, 1 (1992), 105–114.

[56] H. D. Victory, Jr. and B. Neta, A higher-order method for multiple zeros of nonlinear functions, *Internat. J. Comput. Math.* **12**, 3–4 (1982/83), 329–335.

[57] X. H. Wang and D. F. Han, On a dominating sequence method in the point estimate and Smale theorem, *Sci. China, Ser. A* **33**, 2 (1990), 135–144.

[58] J.-C. Yakoubsohn, Finding a cluster of zeros of univariate polynomials, *J. Complexity* **16**, 3 (2000), 603–638.

[59] J.-C. Yakoubsohn, Simultaneous computation of all the zero-clusters of a univariate polynomial, in *Foundations of Computational Mathematics* (*Hong Kong*, 2000). World Scientific, Singapore, 2002, pp. 433–455.

[60] T. J. Ypma, Finding a multiple zero by transformations and Newton-like methods, *SIAM Rev.* **25**, 3 (1983), 365–378.

[61] Z. Zeng, Computing multiple roots of inexact polynomials, *Math. Comp.* **74**, 250 (2005), 869–903.