

Adversarial Wronskians: A proof of the Kung–Traub conjecture

Matthew J. Colbrook*

19 September 2026

Abstract

We prove the Kung–Traub conjecture: a deterministic stationary root-finding method without memory, using at most n real scalar function or derivative evaluations per iteration, cannot have local order greater than 2^{n-1} on every real analytic function with a simple zero. For $n \geq 2$, the bound is attained by the methods of Kung and Traub. The local estimate is required at every sufficiently close starting point, with constants that may depend on the function. Evaluation points and derivative orders may be chosen adaptively, and the rules for choosing them and processing the data are unrestricted. The proof uses a localisation property of polynomial Wronskians to obtain a lower bound at each fixed scale. A diagonal construction then gives one entire function, with derivative bounded above and below on the real line, for which every larger order fails. We also obtain a sharper bound when evaluations are made in prescribed groups, and prove the corresponding result for complex scalar evaluations. The results of this paper are also formally verified in Lean.

Keywords: Kung–Traub conjecture, root-finding, optimal order of convergence, methods without memory, information-based complexity, polynomial Wronskians

1 Introduction

How much can one root-finding update improve an approximation using only n exact scalar observations of the input function? We study stationary methods without memory: queries within an update may depend on earlier answers, but no information from previous updates is retained apart from the current approximation. The resource being counted is access to the function and its derivatives, not arithmetic operations or bit complexity.

Kung and Traub [7, Conjecture 8.1] conjectured in 1974 that the local order of such a method cannot exceed 2^{n-1} . They also constructed methods attaining this order for every $n \geq 2$; Newton’s method is the two-evaluation example. We establish the matching impossibility result without regularity restrictions on the query or output rules and without restrictions on the derivative orders. The result concerns one fixed input at arbitrarily small scales, not merely a worst-case input chosen separately at each scale.

To state our main result, we first describe the precise setup. An input is a real analytic function f on an open interval, that has a specified simple zero α_f . The root α_f specifies where local order is measured; its value is not supplied to the algorithm. One execution of an algorithm starts with a real number x . At step j it chooses a real location z_j and an integer $\nu_j \geq 0$ from x and the previous answers, and receives

$$Y_j = f^{(\nu_j)}(z_j).$$

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom. Email: m.colbrook@damtp.cam.ac.uk.

After at most n such observations it returns a real number $T_f(x)$ depending only on x and these answers. No continuity or measurability of the rules choosing z_j, ν_j , or the final output is imposed. The rules are fixed and are the same for every function and every iteration. In particular, there is no information retained from earlier iterations other than the current starting point. An execution is defined only when its query points lie in that interval and its stage rules are defined.¹ The counterexamples to local order constructed below are entire, so their domains impose no restriction on the query points.

We say that the method has universal local order at least p if, for every such input, it is defined for every sufficiently close starting point $x \neq \alpha_f$ and satisfies

$$|T_f(x) - \alpha_f| = O_f(|x - \alpha_f|^p) \quad \text{as } x \rightarrow \alpha_f. \quad (1.1)$$

The constants and neighbourhood may depend on f ; the estimate is required for every $x \neq \alpha_f$ in that neighbourhood. The following theorem is our main result.

Theorem 1.1 (Local order bound). *Let n be a positive integer, and let T be an algorithm of the kind just described, with total stage and output rules. For every real $p > 2^{n-1}$, there exist a function f entire on $\mathbb{C}$, real-valued on $\mathbb{R}$, and a sequence $(x_s)_{s=1}^\infty$ of real starting points such that*

$$\frac{3}{4} \leq f'(t) \leq \frac{5}{4} \quad (t \in \mathbb{R}), \quad x_s \rightarrow \alpha \quad (s \rightarrow \infty),$$

where α is the unique real zero of f , $x_s \neq \alpha$ for every $s \geq 1$, and

$$\frac{|T_f(x_s) - \alpha|}{|x_s - \alpha|^p} \rightarrow \infty \quad (s \rightarrow \infty). \quad (1.2)$$

Hence, no such method can satisfy (1.1) with $p > 2^{n-1}$ for every real analytic function with a simple zero.

The order definition in [7, Section 2] requires a finite normalised-error limit for each input and therefore implies (1.1). Their derivative blocks are represented here by successive scalar observations at the same point, with each returned derivative value counted separately. Their iterations are consequently included in the present information model. Moreover, the entire counterexample in Theorem 1.1 has nonvanishing derivative on all of $\mathbb{R}$, so it belongs to their original input class. Thus the theorem proves [7, Conjecture 8.1]. Together with the matching construction in Appendix A, it establishes the optimal universal local order 2^{n-1} for $n \geq 2$. The theorem uses real observations; Appendix B proves the corresponding result for complex observations on the unrestricted class of holomorphic functions.

In fact, for each fixed method with total rules, the same entire function and the same sequence of starting points can be chosen to violate (1.1) for every real $p > 2^{n-1}$ simultaneously; see Remark 4.1.

Remark 1.2. *The universal-order conclusion also applies to methods with partial stage or output rules. Indeed, a method with universal local order must be defined at every sufficiently close nonroot starting point for each input. Extend its rules arbitrarily to total rules and apply Theorem 1.1 to the extension. On the resulting entire input, the original method and its extension*

¹For the explicit counterexample statements, stage and output rules are taken to be total, that is, defined on all transcripts of the appropriate length. A method that stops early can be padded with idle observations, which return zero and have no query location. Partial rules can be extended arbitrarily without changing any defined execution. The transfer of the order bound to partial methods is explained after Theorem 1.1.

agree throughout a sufficiently small punctured neighbourhood of the root, and hence eventually along (x_s) . This contradicts (1.1). For a partial method that does not have such a neighbourhood of defined executions, the universal-order property already fails its definedness requirement; no defined divergent-output sequence is asserted for that method. The same distinction applies to the grouped and complex results.

Remark 1.3. *Our argument permits arbitrary deterministic processing, evaluations at unrestricted real points, and derivatives of any nonnegative integer order. In particular, the derivative data need not consist of consecutive derivatives at a point. The counterexample can be chosen entire and strictly increasing, with derivative uniformly close to one. The obstruction therefore persists for functions with a unique, well-conditioned real zero.*

A more precise bound holds when evaluations are made in prescribed groups.

Corollary 1.4 (Grouped evaluations). *Let T be an algorithm as above. Suppose each execution has $k \geq 1$ groups of observations, of fixed positive integer sizes $\ell_1, \dots, \ell_k$. All queries in group j may depend on the earlier groups' answers, but not on answers within group j . Then its universal local order is at most*

$$\ell_1 \prod_{j=2}^k (\ell_j + 1). \quad (1.3)$$

The derivative orders within each group are unrestricted.

For a total of $n = \sum_{j=1}^k \ell_j$ observations, a single group gives the upper bound n , whereas n singleton groups recover 2^{n-1} . For example, with four observations, the schedules (4), (2, 2), and (1, 1, 1, 1) give upper bounds 4, 6, and 8, respectively. Thus the bound records not only the number of observations but also when their answers become available for choosing later queries. If every group has size ℓ , the exponent is $\ell(\ell + 1)^{k-1}$, coinciding with the bound obtained by Casulli and Trigiantè for their parallel multipoint procedures without memory [3].

Lean verification for the results of this paper can be found here:

https://github.com/MColbrook/Kung-Traub_conjecture.

The formal verification is also registered on Palomar [9].

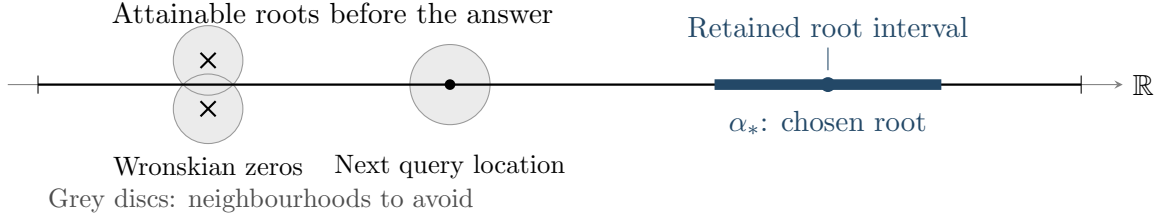
1.1 Earlier results

The conjecture extends Traub's earlier conjecture for two and three evaluations [11]; see also [7, Section 8]. Kung and Traub proved the bound 2^{n-1} for Hermite interpolatory iterations and a bound 2^n for iterations using function values alone [7, Section 7]. They subsequently established the conjectured bound for rational iterations using two evaluations [8].

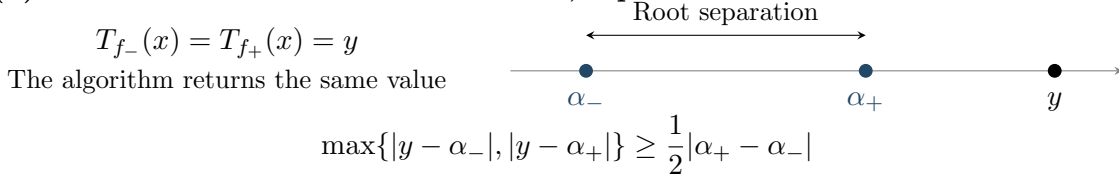
Woźniakowski introduced the order of information for stationary operator iterations under conditions on the positions of the iterates [14]. In this framework, he proved the bound for Hermitian information [15]. Here Hermitian information means that the data at each point comprise the function value and all derivatives up to a prescribed order. The product in (1.3) also occurs in that setting. Wasilkowski obtained further cases through conditions on the associated Birkhoff interpolation problems [12]. The proof below does not require a prescribed derivative interpolation problem to be uniquely solvable.

The construction of one function preserving earlier observations has an antecedent in Brent, Winograd, and Wolfe [1]. Their results concern convergence along generated sequences, with a

(a) Choose the next answer



(b) After n observations: identical data, separated roots



(c) Preserve earlier answers and pass to one entire function

$$f_0 \longrightarrow f_1 \longrightarrow \cdots \longrightarrow f_s \longrightarrow \cdots \longrightarrow f$$

All earlier query answers remain unchanged

$$\frac{|T_f(x_s) - \alpha|}{|x_s - \alpha|^p} \longrightarrow \infty$$

$$x_s \longrightarrow \alpha, \quad p > 2^{n-1}$$

Figure 1: The proof in three steps. (a) Once the next query is fixed, its location and the relevant Wronskian zeros determine the grey exclusion discs. Choose an input with root α_* outside these discs and supply its answer. Restricting the parameter ball retains an interval of roots separated from the forbidden points. (b) After n observations, two inputs still have identical data and roots α_-, α_+ separated by at least a positive constant times $\varepsilon^{2^{n-1}}$. Their common output y has error at least half this separation on one input. (c) Repeat at decreasing scales while preserving all earlier answers. The entire limit violates the order- p estimate for any fixed $p > 2^{n-1}$.

different allowance for derivative information and memory. Casulli and Trigiante obtained order bounds for classes of multipoint procedures with and without memory [2]. For methods with memory, Werschulz obtained an upper bound 2^n for multipoint Hermitian information with n new evaluations per iteration [13].

The contribution here is an order bound for unrestricted adaptive derivative information together with a fixed-input obstruction to universal local order. The Wronskian argument controls the common kernels of the observations without assuming a prescribed pattern of consecutive derivatives or unique solvability of a derivative interpolation problem. The subsequent diagonal construction preserves earlier transcripts exactly. This second step is necessary because the constants and neighbourhoods in (1.1) may depend on the input: a separate hard function at each scale would not suffice.

1.2 Outline of the proof

The proof constructs functions that give exactly the same answers to the algorithm but have different roots. A deterministic algorithm must return the same value on such inputs, so its error on at least one of them is at least half the distance between their roots. We first prove a quantitative separation at each scale, and then obtain a single entire function for which the local estimate fails along a sequence of starting points. The mechanism is illustrated in Figure 1.

Fix a starting point x at distance ε from the root of a reference function. We consider pertur-

bations $w(t)q(t-x)$, where w is fixed and nonzero near the root, and q has degree at most n . The multiplier and coefficient bounds keep the first derivatives of the perturbed functions bounded above and away from zero. Scaling the constant coefficient of q by ε then keeps the perturbed roots within a constant times ε of x . Each scalar observation imposes at most one linear condition on the $n+1$ coefficients. Nonzero perturbations therefore remain after n observations. This dimension count alone is insufficient: a perturbation that preserves the data might also be very small at the root. The essential step is to control how much the root can still vary.

Polynomial Wronskians provide this control. Our localisation lemma states that if a nonzero member of an m -dimensional polynomial space has m zeros in a small disc, then the Wronskian of that space has a zero in a disc with the same centre and a uniformly comparable radius. Once the next query is chosen, the kernels of all observations up to that query are known before its answer is supplied. At each observation, the preceding step supplies an interval of attainable roots whose length is bounded below by a constant times a power of ε . The number of zeros of the nonconstant Wronskians is uniformly bounded. We can therefore choose the answer from a function whose root stays a fixed fraction of this interval's length away from all of them. A smaller ball of parameters preserves this separation. We consider every degree up to n , since the minimum degree of a polynomial in the final kernel is not known in advance.

This also explains the exponent. After j observations, take a nonzero polynomial q of minimum degree in the common kernel. A constant polynomial gives the required variation directly. Otherwise, the localisation lemma at successive rank increases bounds the ordered distances to the roots of q . The product formula contributes at most one factor from each observation, with the exponent of the root interval available before it. The scaled coefficient estimate and derivative bound then give a lower bound for root variation with the sum of all preceding exponents. Starting from one, the exponent remains one after the first observation and then doubles, giving 2^{n-1} after n observations.

Finally, the constant in (1.1) may depend on the function, so separate examples at each scale do not suffice. We construct a sequence of entire corrections, each vanishing to sufficiently high order at every earlier query point. All earlier answers, and hence all earlier outputs, are then preserved exactly, even for discontinuous rules. The corrections are summable on every complex disc and in the real C^1 norm. At each stage the next scale is chosen only after its lower-bound constant is fixed. The entire limit retains a simple, well-conditioned root and violates (1.1) along the resulting sequence. The two parts of the construction are given in Sections 3 and 4. The grouped bound in Corollary 1.4 is proved in Section 4.3. The matching construction and the complex extension appear in Appendices A and B, respectively.

2 Polynomial preliminaries

2.1 Polynomial Wronskians

Throughout, $\overline{B}(a, r)$ denotes the closed Euclidean ball with centre a and radius r , in the space indicated by the context. Let $\mathcal{P}_d$ be the space of complex polynomials in one variable of degree at most d . For an m -dimensional subspace $V \subseteq \mathcal{P}_d$ and a basis $p_1, \dots, p_m$, define

$$W_V(t) = \text{Wr}(p_1, \dots, p_m)(t) = \det[p_i^{(k-1)}(t)]_{k,i=1}^m.$$

Changing the basis multiplies W_V by a nonzero constant, so its zero set and degree depend only on V . For background on polynomial Wronskians and the Wronski map, see [6].

Choose a basis with distinct degrees $0 \leq d_1 < \dots < d_m \leq d$. The coefficient of the top power in the determinant is the product of the leading coefficients times $\prod_{i < j} (d_j - d_i)$. Thus

$$W_V \neq 0, \quad \deg W_V = \sum_{i=1}^m d_i - \frac{m(m-1)}{2} \leq m(d-m+1). \quad (2.1)$$

In particular, W_V is a nonzero constant if and only if $V = \mathcal{P}_{m-1}$.

The vanishing of W_V at a point is equivalent to the existence of a nonzero member of V with a zero there of multiplicity at least m ; see also [5, Theorem A] for holomorphic functions. We need a uniform version in which the zeros may lie at distinct nearby points.

Lemma 2.1 (Wronskian localisation). *For $1 \leq m \leq d$ there is a finite positive constant $C(d, m)$ such that the following holds. If an m -dimensional subspace $V \subseteq \mathcal{P}_d$ contains a nonzero polynomial with at least m zeros in the closed disc $\overline{B}(a, r)$, counted with multiplicity, then W_V has a zero in $\overline{B}(a, C(d, m)r)$.*

Proof. For $r > 0$, an affine change of variable reduces the problem to $a = 0, r = 1$. Suppose there is no such constant. For each positive integer ℓ , choose V_ℓ and a nonzero $h_\ell \in V_\ell$ such that h_ℓ has m zeros in $\overline{B}(0, 1)$ but W_{V_ℓ} has no zero in $\overline{B}(0, \ell)$.

Identify polynomials with their coefficient vectors in $\mathbb{C}^{d+1}$. Normalise h_ℓ to Euclidean norm one and extend it to an orthonormal basis of V_ℓ . Compactness of the finite-dimensional Stiefel manifold gives a subsequence of these bases converging to an orthonormal basis of an m -dimensional subspace V . Its first member h is nonzero.

The limit h still has at least m zeros in $\overline{B}(0, 1)$. Indeed, choose m such roots of each h_ℓ with multiplicity and pass to convergent subsequences. Their monic degree- m factors converge. Polynomial division by these monic factors shows that the quotient coefficients are bounded, so a further subsequence converges and the limiting factor divides h .

The Wronskians of the chosen bases converge coefficientwise to W_V , which is nonzero by (2.1). Any finite zero of W_V would, by Rouché's theorem on a sufficiently small circle, force nearby zeros of W_{V_ℓ} for large ℓ . Thus W_V has no finite zero and is a nonzero constant. Formula (2.1) then forces $V = \mathcal{P}_{m-1}$, contradicting the m zeros of its nonzero member h .

For $r = 0$, choose the polynomial with an m -fold zero at a as the first basis member. Its first m derivatives starting at order zero vanish at a , so the first column of the Wronskian matrix vanishes there. $\square$

We will need a uniform bound on the number of Wronskian zeros. Define

$$M_n = \left\lfloor \frac{(n+1)^2}{4} \right\rfloor, \quad H_n = nM_n + 1. \quad (2.2)$$

By (2.1), every nonconstant Wronskian of a subspace of $\mathcal{P}_d$, $d \leq n$, has at most M_n distinct complex zeros. A zero-dimensional subspace is simply omitted, and the Wronskian of the full space $\mathcal{P}_d$ is constant. At each observation, we shall use one subspace for each degree $1, \dots, n$, together with one possible query location; H_n bounds the resulting number of points.

2.2 Coefficient estimates

For a polynomial $q(s) = \sum_{k=0}^d b_k s^k$ of degree $d \leq n$ in a centred variable, let $b = (b_0, \dots, b_d, 0, \dots, 0) \in \mathbb{C}^{n+1}$ be its coefficient vector, with zeros appended. We use the Euclidean norm of this vector.

The first estimate converts distances to the zeros into a lower bound for the polynomial's value. The second controls the change of normalisation when the constant coefficient is scaled by ε .

Lemma 2.2 (Product and scaled-norm estimates). *Let q have degree $d \leq n$ and $\|b\|_2 = 1$. If $\beta_1, \dots, \beta_d$ are the roots of q , counted with multiplicity, then for $|a| \leq 1/2$,*

$$|q(a)| \geq 4^{-d} \prod_{k=1}^d \min\{1, |a - \beta_k|\}. \quad (2.3)$$

For $\varepsilon > 0$, let D_ε be the $(n+1) \times (n+1)$ diagonal matrix $\text{diag}(\varepsilon, 1, \dots, 1)$. For every fixed $R > 0$, if $|a| \leq R\varepsilon \leq 1/2$, then

$$\frac{|q(a)|}{\|D_\varepsilon^{-1}b\|_2} \geq \frac{1}{2R+2} \min\{\varepsilon, |q(a)|\}. \quad (2.4)$$

Proof. The factorisation $q(s) = b_d \prod_k (s - \beta_k)$ gives

$$1 \leq \|b\|_1 \leq |b_d| \prod_k (1 + |\beta_k|) \leq 2^d |b_d| \prod_k \max\{1, |\beta_k|\}.$$

For $|a| \leq 1/2$,

$$\frac{|a - \beta|}{\max\{1, |\beta|\}} \geq \frac{1}{2} \min\{1, |a - \beta|\},$$

which proves (2.3). The constant-polynomial case is included by the empty-product convention.

Cauchy–Schwarz gives, on $|s| \leq 1/2$,

$$|q'(s)| \leq \left(\sum_{k=1}^{\infty} k^2 (1/2)^{2k-2} \right)^{1/2} = \sqrt{80/27} < 2.$$

Hence $|b_0| \leq |q(a)| + 2R\varepsilon$ and

$$\|D_\varepsilon^{-1}b\|_2 \leq \frac{|b_0|}{\varepsilon} + 1 \leq \frac{|q(a)|}{\varepsilon} + 2R + 1.$$

Splitting into $|q(a)| \leq \varepsilon$ and $|q(a)| \geq \varepsilon$ proves (2.4). $\square$

3 A lower bound at a fixed scale

Fix real analytic functions f_0, w on $\mathbb{R}$, where f_0 has a specified simple root α_0 and $w(\alpha_0) \neq 0$. At a scale $\varepsilon > 0$, set $x = \alpha_0 + \varepsilon$ and consider

$$f_u(t) = f_0(t) + w(t) \left[\varepsilon u_0 + \sum_{k=1}^n u_k (t - x)^k \right], \quad u \in \overline{B}_{\mathbb{R}^{n+1}}(0, 1). \quad (3.1)$$

Define

$$v_\varepsilon(t) = w(t)(\varepsilon, t - x, \dots, (t - x)^n), \quad f_u(t) = f_0(t) + u \cdot v_\varepsilon(t).$$

Assume that for all sufficiently small ε , every family member has a unique specified root $\alpha(u)$ in a fixed interval J , and that fixed positive constants m, M, R, w_*, K satisfy

$$m \leq f'_u(t) \leq M \quad (t \in J), \quad (3.2)$$

$$|\alpha(u) - x| \leq R\varepsilon, \quad |w(\alpha(u))| \geq w_*, \quad (3.3)$$

$$\|v'_\varepsilon(t)\|_2 \leq K \quad (t \in J). \quad (3.4)$$

The entire-function construction in the next section arranges these hypotheses, with $J = \mathbb{R}$. The argument below only needs derivative observations to be linear on the finite-dimensional family. Accordingly it also allows arbitrary real scalar linear observations at this finite-scale stage.

For the induction, define

$$B_0 = 1, \quad B_j = 2^{j-1} \quad (j \geq 1). \quad \text{Then} \quad B_j = \sum_{i=0}^{j-1} B_i \quad (j \geq 1). \quad (3.5)$$

Lemma 3.1 (Lower bound at a fixed scale). *Under (3.2)–(3.4), there are constants $c > 0$ and $\varepsilon_0 > 0$, independent of the algorithm, such that every adaptive deterministic algorithm with total stage and output rules and at most n scalar linear observations has, for every $0 < \varepsilon < \varepsilon_0$, an input from (3.1) on which*

$$|T_{f_u}(x) - \alpha(u)| \geq c\varepsilon^{B_n}. \quad (3.6)$$

The constant c depends only on n, m, M, R, w_, K ; ε_0 may also depend on the common scale threshold in the family hypotheses. Neither constant depends on the queries or answers in the execution, or on the particular $0 < \varepsilon < \varepsilon_0$. For derivative observations, the input can also be chosen so that no queried location equals $\alpha(u)$.*

The condition on query locations will allow later corrections to vanish at these points without forcing them to vanish at the root.

3.1 Parameters with identical observations

We prove Lemma 3.1 by choosing answers successively and retaining a ball of parameters consistent with them. A fixed transcript imposes affine equations on parameter space. The admissible parameters satisfying these equations are the intersection of an affine subspace with the parameter unit ball. If the first j observations have been fixed, all these inputs generate the same first j queries and answers: this follows inductively because query i depends only on the answers before i .

Let u_c belong to such an affine subspace, with linear direction space E , and consider a ball $\overline{B}_E(u_c, r) = \{u_c + h : h \in E, \|h\|_2 \leq r\}$ contained in the parameter unit ball. The norm of evaluation on the remaining directions is

$$a = \|P_E v_\varepsilon(\alpha(u_c))\|_2, \quad (3.7)$$

where P_E is orthogonal projection onto E . For $h \in E$ with $\|h\|_2 \leq r$, the equality

$$f_{u_c+h}(\alpha(u_c)) = h \cdot v_\varepsilon(\alpha(u_c))$$

and (3.2) imply

$$|\alpha(u_c + h) - \alpha(u_c)| \leq \frac{a}{m} \|h\|_2. \quad (3.8)$$

If $a > 0$, take the unit vector $e = P_E v_\varepsilon(\alpha(u_c))/a$. The roots at $u_c \pm (r/2)e$ lie on opposite sides of $\alpha(u_c)$ and are each at least $ra/(2M)$ away from it. Continuity of the root along this segment shows that its image contains a real interval of length at least

$$\frac{ra}{M}. \quad (3.9)$$

We will choose the next centre on this segment. To control root motion in a smaller ball about that centre, we also need an upper bound for the projected norm there. For any u_* on the segment, (3.4) and (3.8) give

$$\|P_{Ev_\varepsilon}(\alpha(u_*))\|_2 \leq a \left(1 + \frac{Kr}{2m}\right) \leq Qa, \quad Q = 1 + \frac{K}{2m}, \quad (3.10)$$

provided $r \leq 1$.

3.2 Choice of the next answer

Let L_i be the linear observation functional chosen at step i along the transcript. For a polynomial q in the centred variable $s = t - x$, define

$$\tilde{L}_i(q) = L_i(t \mapsto w(t)q(t - x)), \quad V_{j,d} = \bigcap_{i=1}^j \ker(\tilde{L}_i|_{\mathcal{P}_d}).$$

Real observation functionals are extended complex-linearly on the polynomial space. Their defining matrices have real entries, so the real and complex ranks agree and these kernels have real bases.

When the j th query is chosen, all spaces $V_{j,d}$, $1 \leq d \leq n$, are determined *before* the j th answer is supplied. Let S_j consist of the numbers $x + \zeta$ for all zeros ζ of their nonconstant Wronskians. Omit zero-dimensional spaces. For a derivative observation, also include the queried location z_j in S_j . By (2.2),

$$|S_j| \leq H_n. \quad (3.11)$$

We use the convention $\text{dist}(z, \emptyset) = +\infty$. We include every degree $d \leq n$ because the degree of the eventual kernel polynomial need not be known at the current stage.

Take $\varepsilon_0 \leq \min\{1, 1/(2R)\}$ small enough that the family hypotheses hold. Start with centre 0, direction space $\mathbb{R}^{n+1}$, and radius $r_0 = 1/2$. By (3.3), the projected norm at its centre is at least $w_*\varepsilon$. Suppose after $j-1$ observations there is a ball of radius $r = r_{j-1}$ in the remaining parameter space, with projected norm at its centre satisfying

$$a \geq A_{j-1}\varepsilon^{B_{j-1}}, \quad (3.12)$$

and with all previously imposed distance bounds valid throughout the ball.

Choose query j , and form S_j . The interval in (3.9) cannot be covered by the radius- $ra/[4M(H_n+1)]$ discs around its at most H_n points: their intersections with the real line have total length less than half that interval's length. Hence there is a point u_* on the half-radius segment such that

$$\text{dist}(\alpha(u_*), S_j) \geq \frac{ra}{4M(H_n+1)}. \quad (3.13)$$

Supply the j th answer from f_{u_*} . If E is the old direction space, the new one is

$$E_{\text{new}} = E \cap \ker[h \mapsto L_j(h \cdot v_\varepsilon)].$$

Set

$$\theta = \min \left\{ \frac{1}{2}, \frac{m}{8QM(H_n+1)} \right\}, \quad r_j = \theta r_{j-1}. \quad (3.14)$$

The ball $\overline{B}_{E_{\text{new}}}(u_*, r_j)$ is contained in the previous ball, since the new centre is at most $r/2$ from the old centre and $r_j \leq r/2$. Equations (3.8) and (3.10) show that every root in this smaller ball moves from $\alpha(u_*)$ by at most

$$\frac{Qa}{m}\theta r \leq \frac{ra}{8M(H_n + 1)}.$$

Therefore every input in the smaller ball satisfies

$$\text{dist}(\alpha(u), S_j) \geq \gamma_j \varepsilon^{B_{j-1}}, \quad \gamma_j = \frac{r_{j-1}A_{j-1}}{8M(H_n + 1)} > 0. \quad (3.15)$$

All earlier distance bounds are retained by containment.

3.3 Lower bound for the projected norm

It remains to recover a lower bound of the form (3.12) for the projected norm on the smaller ball. At an arbitrary point of that ball, choose a nonzero real polynomial q of minimum degree d in the common kernel of the first j functionals $\tilde{L}_i$. Since there are j scalar constraints, $d \leq j$. Normalise its coefficient vector b to norm one.

If $d = 0$, the unit parameter direction corresponding to the perturbation εw lies in the common kernel and gives projected norm at least $w_* \varepsilon$, which is at least $w_* \varepsilon^{B_j}$ for $\varepsilon \leq 1$.

Suppose $d \geq 1$. Minimality of d says that the common kernel on $\mathcal{P}_{d-1}$ is zero. Thus the observation map is injective on this d -dimensional space and has rank at least d on $\mathcal{P}_d$. Since the nonzero polynomial q belongs to its kernel on $\mathcal{P}_d$, its rank is at most d . The rank is therefore exactly d , and the kernel has dimension one. Each scalar observation increases the rank by at most one. Hence the first query indices i_k at which the rank reaches k , for $k = 1, \dots, d$, are distinct integers in $\{1, \dots, j\}$. At prefix i_k the space $V_{i_k, d}$ has dimension

$$m_k = d + 1 - k.$$

Its Wronskian is nonconstant: otherwise (2.1) would give $V_{i_k, d} = \mathcal{P}_{m_k-1}$, which cannot contain the degree- d polynomial q since $m_k \leq d$. Let $\rho_1 \geq \dots \geq \rho_d$ be the distances from $\alpha(u) - x$ to the roots of q , counted with multiplicity. Its m_k nearest roots lie in the disc of radius ρ_k . Since $q \in V_{i_k, d}$, Lemma 2.1 gives a zero of $W_{V_{i_k, d}}$ within $C(d, m_k)\rho_k$ of $\alpha(u) - x$. Its translate by x belongs to S_{i_k} . Thus (3.15) gives

$$\rho_k \geq \frac{\gamma_{i_k}}{C(d, d + 1 - k)} \varepsilon^{B_{i_k-1}}. \quad (3.16)$$

The product estimate (2.3) yields a constant $c_j > 0$ such that

$$|q(\alpha(u) - x)| \geq c_j \varepsilon^{\sum_{k=1}^d B_{i_k-1}} \geq c_j \varepsilon^{B_j}. \quad (3.17)$$

The last inequality uses $0 < \varepsilon < 1$ and

$$\{i_k - 1 : 1 \leq k \leq d\} \subseteq \{0, \dots, j-1\}, \quad \sum_{i=0}^{j-1} B_i = B_j.$$

The following choice is independent of the algorithm:

$$C_* = \max(\{1\} \cup \{C(d, m) : 1 \leq m \leq d \leq n\}), \quad c_j = 4^{-n} \prod_{i=1}^j \min\{1, \gamma_i/C_*\}.$$

This choice works in (3.17): adding the unused indices only decreases a product whose factors are at most one. In particular, $0 < c_j \leq 1$.

To obtain an admissible parameter direction from this polynomial, we must account for the scaled constant coefficient. The vector

$$h_q = \frac{D_\varepsilon^{-1}b}{\|D_\varepsilon^{-1}b\|_2}$$

is a unit vector in the remaining direction space. Its function perturbation is $w(t)q(t-x)/\|D_\varepsilon^{-1}b\|_2$. By (2.4), (3.3), and (3.17), its value at the current root has magnitude at least $A_j\varepsilon^{B_j}$, where

$$A_0 = w_*, \quad A_j = \frac{w_*c_j}{2R+2} > 0. \quad (3.18)$$

The definitions are recursive: A_{j-1} determines γ_j , which determines c_j and then A_j . The same A_j also works for $d = 0$, since $A_j \leq w_*$. Taking the maximum over unit vectors in the remaining direction space gives the same lower bound for the projected norm. This holds at every point of the smaller ball and completes the induction.

3.4 Completion of the lower bound

After n answers, the remaining ball has radius $r_n = \frac{1}{2}\theta^n > 0$ and projected norm at its centre at least $A_n\varepsilon^{B_n}$. Taking the two parameters displaced from its centre by $\pm(r_n/2)e$, as in (3.9), gives roots separated by at least

$$\frac{r_n A_n}{M} \varepsilon^{B_n}.$$

Their entire transcripts, and therefore their outputs, are identical. At least one output error is at least half this root separation. Thus (3.6) holds with $c = r_n A_n/(2M)$. Every queried location was included in an avoidance set, so none equals the selected root on either input. This proves Lemma 3.1.

4 A single entire counterexample

The function supplied by Lemma 3.1 may depend on the scale, whereas the constant and neighbourhood in (1.1) may depend on the function. To prove Theorem 1.1, we therefore construct one entire function for which the estimate fails along a sequence of starting points. Fix $p > B_n$. As in the analytic perturbation construction of Brent, Winograd, and Wolfe [1, Sections III–V], each correction preserves all previously evaluated derivative values. Here these values come from executions at several different starting points.

4.1 Inductive construction

Start with $f_0(t) = t$. For each integer $s \geq 1$, we inductively construct an entire function of the form

$$f_s(t) = t + e^{-t^2}P_s(t), \quad P_s \in \mathbb{R}[t], \quad (4.1)$$

and a starting point x_s , with output $y_s = T_{f_s}(x_s)$. We choose each correction $f_s - f_{s-1}$ small both in real C^1 norm and on expanding complex discs.

The correction bounds must ensure convergence and keep every later displacement of the root small compared with the error and scale at each completed stage. Let $0 < b_1 \leq 1/32$. After stage s choose the next positive bound subject to

$$b_{s+1} \leq \min \left\{ \frac{b_s}{2}, \frac{E_s}{32}, \frac{\varepsilon_s}{32} \right\}, \quad E_s = |y_s - \alpha_s|, \quad (4.2)$$

where α_s is the unique real root of f_s . Positivity of E_s will follow from the finite-scale lower bound. The resulting bounds satisfy $\sum_{s \geq 1} b_s \leq 1/16$. In particular, all intermediate functions will have derivative between $15/16$ and $17/16$ on the real line.

Suppose that the first $s-1$ functions and starting points have been chosen. Let Z_{s-1} be the finite set of points queried in these executions, and let $R_{s-1}(z)$ be the largest derivative order queried at z during them. Assume inductively that $f_{s-1}(z) \neq 0$ for every $z \in Z_{s-1}$. The next correction must vanish at each such point to order at least $R_{s-1}(z) + 1$. Define

$$w_s(t) = \lambda_s e^{-t^2} \prod_{z \in Z_{s-1}} (t - z)^{R_{s-1}(z)+1}, \quad \lambda_s > 0. \quad (4.3)$$

Since α_{s-1} is the unique real root, $w_s(\alpha_{s-1}) \neq 0$.

Choose λ_s sufficiently small that every perturbation

$$g_u(t) = w_s(t) \left[\varepsilon u_0 + \sum_{k=1}^n u_k (t - x)^k \right], \quad \|u\|_2 \leq 1, \quad (4.4)$$

with $x \in [-1, 1]$ and $0 \leq \varepsilon \leq 1$, satisfies

$$\sup_{t \in \mathbb{R}} |g_u(t)| + \sup_{t \in \mathbb{R}} |g'_u(t)| + \sup_{|z| \leq s+2} |g_u(z)| \leq b_s. \quad (4.5)$$

Such a choice is possible: a polynomial times e^{-t^2} and its first derivative are bounded on $\mathbb{R}$, are bounded on each complex disc, and the indicated coefficient and starting-point sets are compact.

Set $x = \alpha_{s-1} + \varepsilon$, with $0 < \varepsilon \leq 1/4$. All $f_{s-1} + g_u$ are globally strictly increasing, with derivative between $1/2$ and $3/2$, and differ from t by a bounded amount. Each therefore has a unique real root, denoted by $\alpha(u)$. Also $|\alpha_{s-1}| \leq 1/4$, so $x \in [-1, 1]$.

At the old root, Cauchy–Schwarz gives

$$\left| \varepsilon u_0 + \sum_{k=1}^n u_k (-\varepsilon)^k \right| \leq 2\varepsilon.$$

Taking $\varepsilon = 1$ and $u = (1, 0, \dots, 0)$ in the uniform choice (4.5) gives $\|w_s\|_{\infty, \mathbb{R}} \leq b_s$. Hence the new root moves from α_{s-1} by at most $4b_s \varepsilon \leq \varepsilon/4$. Thus

$$\frac{3}{4}\varepsilon \leq x - \alpha(u) \leq \frac{5}{4}\varepsilon. \quad (4.6)$$

For sufficiently small ε , w_s is bounded away from zero at all these roots. Since (4.5) holds on the whole parameter unit ball, it also gives

$$\|v'_\varepsilon(t)\|_2 = \sup_{\|u\|_2 \leq 1} |u \cdot v'_\varepsilon(t)| = \sup_{\|u\|_2 \leq 1} |g'_u(t)| \leq b_s \quad (t \in \mathbb{R}),$$

which is (3.4) with $K = b_s$. The multiplier w_s is now fixed. Its vanishing orders at earlier query points affect the constant in the lower bound, but each new execution still imposes at most n linear conditions on a polynomial of degree at most n . Thus Lemma 3.1 applies with a constant $c_s > 0$ and a scale threshold, both fixed before ε_s is chosen.

Choose ε_s below that threshold and so small that

$$0 < \varepsilon_s < \min\{1/s, 1/4\}, \quad \frac{c_s}{2} \varepsilon_s^{B_n} \geq s(2\varepsilon_s)^p. \quad (4.7)$$

Such a choice is possible because $p > B_n$.

Apply Lemma 3.1 to choose u and set $f_s = f_{s-1} + g_u$, $x_s = \alpha_{s-1} + \varepsilon_s$, and $y_s = T_{f_s}(x_s)$. Then

$$E_s = |y_s - \alpha_s| \geq c_s \varepsilon_s^{B_n} > 0. \quad (4.8)$$

No queried location equals α_s . Global strict monotonicity therefore implies that f_s is nonzero at every new query location. At old query locations its function values remain the previous nonzero values, because (4.3) vanishes there to the required orders. This completes the induction and permits the next choice in (4.2).

4.2 Passage to the limit

By (4.5), the series of entire corrections converges uniformly on every fixed complex disc. Its limit f is entire and real on $\mathbb{R}$. The real derivatives also converge uniformly, and the summable bounds give $3/4 \leq f'(t) \leq 5/4$ on $\mathbb{R}$. Since $f(t) - t$ is bounded, f has exactly one real zero α .

For every fixed stage s ,

$$\|f - f_s\|_{\infty, \mathbb{R}} \leq \sum_{k>s} b_k \leq 2b_{s+1}.$$

The lower derivative bound $f' \geq 1/2$ and (4.2) consequently give

$$|\alpha - \alpha_s| \leq 4b_{s+1} \leq \min\{E_s/8, \varepsilon_s/8\}. \quad (4.9)$$

Every correction after stage s vanishes at every point queried in that execution to an order strictly larger than its queried derivative order. Uniform convergence on complex neighbourhoods permits passage to all derivatives. Thus the corresponding derivative values of f and f_s are exactly equal. Induction through the adaptive query sequence gives

$$T_f(x_s) = T_{f_s}(x_s) = y_s. \quad (4.10)$$

Equations (4.6) and (4.9) show

$$\frac{5}{8} \varepsilon_s \leq x_s - \alpha \leq \frac{11}{8} \varepsilon_s < 2\varepsilon_s,$$

so $x_s \rightarrow \alpha$ through non-root starting points. Finally,

$$\frac{|T_f(x_s) - \alpha|}{|x_s - \alpha|^p} \geq \frac{E_s/2}{(2\varepsilon_s)^p} \geq \frac{(c_s/2)\varepsilon_s^{B_n}}{(2\varepsilon_s)^p} \geq s,$$

by (4.8) and (4.7). This proves Theorem 1.1.

Remark 4.1 (Simultaneous failure of larger orders). *The real construction can use $p_s = B_n + 1/s$ in place of the fixed p in (4.7). For each fixed $p > B_n$, eventually $p \geq p_s$ and $2\varepsilon_s < 1$, so the resulting error ratio for exponent p is still at least s . Thus the function and sequence in Theorem 1.1 can be chosen independently of $p > B_n$.*

4.3 Evaluations in prescribed groups

We adapt the finite-scale argument to grouped queries, then apply the entire-function construction above.

Proof of Corollary 1.4. Let $n = \sum_{j=1}^k \ell_j$, retain the degree- n polynomial perturbation space, and define

$$C_0 = 1, \quad C_1 = \ell_1, \quad C_j = (\ell_j + 1)C_{j-1} \quad (2 \leq j \leq k).$$

Then $C_j = \sum_{i=1}^j \ell_i C_{i-1}$ for $1 \leq j \leq k$. We adapt the proof of Lemma 3.1, retaining one parameter ball after each group. Order the queries within each group arbitrarily. All queries in a group are determined by earlier groups' answers. Hence the union $\mathcal{S}_j$ of the forbidden sets for its intermediate scalar prefixes is known before any answer in group j is supplied, and $|\mathcal{S}_j| \leq \ell_j H_n$.

To avoid this union and contract the ball once, use the constants Q, C_* from the scalar proof and choose

$$F = nH_n, \quad \theta_G = \min \left\{ \frac{1}{2}, \frac{m}{8QM(F+1)} \right\}, \quad r_j^G = \frac{1}{2} \theta_G^j.$$

Starting with $A_0^G = w_*$, define recursively

$$\gamma_j^G = \frac{r_{j-1}^G A_{j-1}^G}{8M(F+1)}, \quad c_j^G = 4^{-n} \prod_{i=1}^j \min\{1, \gamma_i^G / C_*\}^{\ell_i}, \quad A_j^G = \frac{w_* c_j^G}{2R+2} \quad (1 \leq j \leq k).$$

These constants are positive and independent of the algorithm and ε . Retain the scalar restriction on ε_0 .

After group j , the induction retains a ball of radius r_j^G with identical observations, projected norm at least $A_j^G \varepsilon^{C_j}$, and

$$\text{dist}(\alpha(u), \mathcal{S}_i) \geq \gamma_i^G \varepsilon^{C_{i-1}} \quad (1 \leq i \leq j)$$

throughout the ball. The initial ball has radius $1/2$ and projected norm at least $w_* \varepsilon$. Before group j , let r, a be the preceding radius and projected norm at its centre. The root-interval argument selects a centre whose root is at distance at least $ra/[4M(F+1)]$ from $\mathcal{S}_j$. Supply every group answer there and intersect the old direction space with all group kernels. The affine ball about this centre, with that direction space and radius $\theta_G r$, is contained in its predecessor. The root-motion bound loses at most $ra/[8M(F+1)]$, giving the required avoidance and preserving earlier bounds.

At any point of the new ball, choose a nonzero real kernel polynomial q of minimum degree, with coefficient norm one. If q is nonconstant, let h_i count its rank-increment indices in group i in the scalar argument. Then $h_i \leq \ell_i$, and each corresponding root distance is at least $(\gamma_i^G / C_*) \varepsilon^{C_{i-1}}$. Thus (2.3) gives

$$|q(\alpha(u) - x)| \geq 4^{-n} \prod_{i=1}^j \min\{1, \gamma_i^G / C_*\}^{h_i} \varepsilon^{\sum_{i=1}^j h_i C_{i-1}} \geq c_j^G \varepsilon^{C_j}.$$

Adding unused factors decreases the product, and $\sum_{i=1}^j h_i C_{i-1} \leq C_j$. Equation (2.4) gives the required projected norm, since $c_j^G \leq 1$, $C_j \geq 1$, and $\varepsilon < 1$. The constant-polynomial case follows from the direction εw , as before.

After k contractions, the identical-transcript argument gives an output error at least $c_G \varepsilon^{C_k}$, where

$$c_G = \frac{r_k^G A_k^G}{2M} > 0.$$

The avoidance sets include every actual query location. The entire-function construction therefore applies with $C_k = \ell_1 \prod_{j=2}^k (\ell_j + 1)$ in place of B_n , proving (1.3). $\square$

Acknowledgments and AI declaration. The author conceived the central proof strategy and led its mathematical development. In particular, the author developed the use of polynomial Wronskians to control the root uncertainty remaining after adaptive observations, and its combination with an exact-transcript-preserving construction yielding a single entire counterexample. ChatGPT 5.6 and 6 were used as interactive aids in discussing and checking some of the mathematical arguments during preparation of the manuscript. These interactions supported the development and scrutiny of the author’s approach; the author evaluated the suggestions arising from them and determined the arguments and formulations included in the paper.

After the mathematical manuscript had been written, ChatGPT 6 assisted with the Lean formalization. The resulting Lean proof and its problem formulation were thoroughly checked by the author and with the Lean 4 Comparator against the theorems stated in this paper, including checks that they introduce no new axioms or unproved results. I wish to thank Gonzalo Cao-Labora and George Stepaniants for their kind advice and for taking the time to explain how to set this up.

A Sharpness

The upper bound in Theorem 1.1 is attained for every $n \geq 2$ by the methods of Kung and Traub [7]. For completeness, we give the inverse Hermite interpolation construction to verify sharpness for (1.1). Further constructions are surveyed in [10].

Fix $n \geq 2$ and a real analytic function f with a simple zero α . Choose an interval I about α on which f has an analytic inverse g . Set $y_0 = x$, $t_i = f(y_i)$, and evaluate t_0 and $f'(x)$. The linear interpolant

$$\mathcal{H}_0(t) = x + \frac{t - t_0}{f'(x)}$$

gives $y_1 = \mathcal{H}_0(0)$. For $1 \leq j \leq n - 2$, evaluate t_j and let $\mathcal{H}_j$ be the polynomial of degree at most $j + 1$ satisfying

$$\mathcal{H}_j(t_0) = x, \quad \mathcal{H}'_j(t_0) = 1/f'(x), \quad \mathcal{H}_j(t_i) = y_i \quad (1 \leq i \leq j).$$

Set $y_{j+1} = \mathcal{H}_j(0)$. Whenever an evaluated t_i is zero, return y_i immediately.

We verify that these interpolants are defined for every sufficiently small nonzero $x - \alpha$ until termination. Let $e_i = y_i - \alpha$. The Hermite remainder, including $j = 0$, is

$$e_{j+1} = (-1)^{j+1} D_j t_0^2 \prod_{i=1}^j t_i, \quad D_j = g[0, t_0, t_0, t_1, \dots, t_j], \quad (\text{A.1})$$

where the brackets denote divided differences, with repeated nodes interpreted by derivatives. Choose a compact interval $K \subset f(I)$ with zero in its interior. The divided-difference mean-value formula [4, Section 8, equation (44)] gives $|D_j| \leq D$ whenever its nodes lie in K , where

$$D = \max \left\{ 1, \max_{0 \leq j \leq n-2} \frac{\sup_{t \in K} |g^{(j+2)}(t)|}{(j+2)!} \right\}.$$

This bound also holds when nodes coincide. Choose $L \geq 1$ such that $|f(y)| \leq L|y - \alpha|$ near α . Shrink $\delta > 0$ so that $[\alpha - \delta, \alpha + \delta] \subset I$, its image lies in K , and

$$DL^{j+2}\delta^{j+1} < \frac{1}{2} \quad (0 \leq j \leq n-2).$$

For $0 < |e_0| < \delta$, (A.1) gives $|e_1| < |e_0|/2$. Inductively, if no root has yet been found, then for $j \geq 1$,

$$\frac{|e_{j+1}|}{|e_j|} \leq DL^{j+2}|e_0|^2 \prod_{i=1}^{j-1} |e_i| < \frac{1}{2}.$$

Every new nonroot node is therefore closer to α than all its predecessors. The nodes remain in I and are distinct, so their function values are distinct as well. This proves local admissibility, including exact termination.

The same remainder gives $e_j = O(e_0^{2^j})$ by induction, since $e_1 = O(e_0^2)$ and

$$2 + \sum_{i=1}^j 2^i = 2^{j+1}.$$

The output y_{n-1} uses at most n observations. For $f(x) = e^x - 1$ at zero, $g(t) = \log(1+t)$ and the leading coefficients $e_j \sim \kappa_j x^{2^j}$ satisfy

$$\kappa_1 = \frac{1}{2}, \quad \kappa_{j+1} = \frac{\prod_{i=1}^j \kappa_i}{j+2}.$$

These coefficients are positive, so the order 2^{n-1} is attained.

B Complex scalar observations

We now allow complex starting points, complex evaluation points, and complex outputs. Each value $f^{(\nu)}(z) \in \mathbb{C}$, with $\nu \geq 0$ an integer, counts as one observation. The inputs are holomorphic functions on open domains with a specified simple zero. As in the real case, each query must lie in the input's domain, and local order requires defined executions at every sufficiently close starting point distinct from the zero. The choice of evaluations and the processing of the data remain arbitrary and deterministic.

Theorem B.1 (Complex local order bound). *Let $n \geq 1$. For every deterministic stationary method without memory using at most n complex scalar derivative observations, with total stage and output rules, and every real $p > B_n$, there exist an entire function f and complex starting points $x_s \rightarrow \alpha$ such that*

$$|f'(z) - 1| \leq \frac{1}{4} \quad (|z| \leq 1),$$

f has exactly one zero α in the unit disc, $x_s \neq \alpha$ for every s , and

$$\frac{|T_f(x_s) - \alpha|}{|x_s - \alpha|^p} \rightarrow \infty.$$

Proof. The polynomial Wronskian and coefficient lemmas hold over $\mathbb{C}$. We first obtain a disc of attainable roots in each remaining affine parameter space. The lower bound at a fixed scale then follows from the same polynomial argument. Finally, we construct a single entire function by polynomial corrections.

Variation of the root. Use the complex version of (3.1), with $u \in \overline{B}_{\mathbb{C}^{n+1}}(0, 1)$. Assume that each family member has a unique root $\alpha(u)$ in a convex disc J , that these roots lie in a fixed compact subdisc, and that the uniform bounds $|\alpha(u) - x| \leq R\varepsilon$ and $|w(\alpha(u))| \geq w_* > 0$ from (3.3) hold. Suppose also that, for every family member,

$$m|z - t| \leq |f_u(z) - f_u(t)| \leq M|z - t| \quad (z, t \in J). \quad (\text{B.1})$$

Let $v = v_\varepsilon$, and assume that $\|v(z) - v(t)\|_2 \leq K|z - t|$ for $z, t \in J$. In the construction below these bounds follow by integrating $|f'_u - 1| \leq 1/4$ and a derivative bound along straight segments.

Consider a ball $\overline{B}_E(u_c, r)$ in the remaining affine parameter space, contained in the parameter unit ball, with complex direction space E . Let $\alpha_c = \alpha(u_c)$. Define

$$a = \left\| P_E \overline{v(\alpha_c)} \right\|_2.$$

When $a > 0$, set $e = P_E \overline{v(\alpha_c)} / a$. The projection uses the Hermitian inner product, taken linear in its first argument, whereas the perturbation remains the complex bilinear expression $u \cdot v$. Thus $e \cdot v(\alpha_c) = a$. For $h \in E$ with $\|h\|_2 \leq r$, the root-shift estimate

$$|\alpha(u_c + h) - \alpha_c| \leq a \|h\|_2 / m$$

follows from (B.1) exactly as in the real case.

Choose the initial radius $r_0 = \min\{1/2, 2M/K\}$, with $2M/K = +\infty$ when $K = 0$, and suppose $0 < r \leq r_0$. As the parameter varies over $u_c + \{\xi e : |\xi| \leq r/2\}$, its root ranges over a set containing the disc

$$\overline{B}\left(\alpha_c, \frac{ra}{4M}\right). \quad (\text{B.2})$$

Indeed, for any target t in this disc, set

$$\xi = -\frac{f_{u_c}(t)}{e \cdot v(t)}.$$

The denominator has magnitude at least $a - K|t - \alpha_c| \geq a/2$, and the numerator has magnitude at most $M|t - \alpha_c|$. Consequently $|\xi| \leq r/2$ and $f_{u_c + \xi e}(t) = 0$. Thus t is that input's selected root. For sufficiently small ε all discs in (B.2) lie in J : uniformly, $a \leq \|v(\alpha_c)\|_2 = O(\varepsilon)$, and the roots have a fixed positive distance from the boundary of J .

Avoidance and projected norm. Form the same sets S_j before each answer, including the current query location when present. The disc (B.2) cannot be covered by its at most H_n forbidden discs of radius

$$\delta = \frac{ra}{8M(H_n + 1)};$$

the sum of their areas is smaller than the area of the disc in (B.2). Choose u_* within distance $r/2$ of the previous centre whose root avoids S_j by δ , and supply its answer. With $Q = 1 + K/(2m)$, choose

$$\theta = \min \left\{ \frac{1}{2}, \frac{m}{16QM(H_n + 1)} \right\}, \quad r_j = \theta r_{j-1}.$$

The ball of radius θr in the new affine parameter space has all its roots at least $\delta/2$ from S_j . The polynomial argument in Section 3 applies over $\mathbb{C}$, with

$$\gamma_j = \frac{r_{j-1} A_{j-1}}{16M(H_n + 1)}$$

and the same recursive definitions of c_j, A_j . It yields a final projected norm at least $A_n \varepsilon^{B_n}$. Opposite points of the final disc of attainable roots correspond to two inputs with identical observations and roots separated by at least $r_n A_n \varepsilon^{B_n} / (2M)$. One error is therefore at least $c \varepsilon^{B_n}$, where $c > 0$ is independent of the algorithm and ε .

An entire limit. Only bounds on complex discs are needed for the limiting construction, so polynomial corrections suffice. Start with $f_0(z) = z$ and keep each f_s polynomial, with α_s denoting its unique zero in the unit disc. For the past complex query nodes, use

$$w_s(z) = \lambda_s \prod_{\zeta \in Z_{s-1}} (z - \zeta)^{R_{s-1}(\zeta)+1}$$

in place of (4.3). Choose $\lambda_s > 0$ so that all perturbations of the form (4.4), now with complex u, x satisfying $\|u\|_2 \leq 1$, $|x| \leq 1$, and real $0 \leq \varepsilon \leq 1$, obey

$$\sup_{|z| \leq s+2} |g_u(z)| + \sup_{|z| \leq 1} |g'_u(z)| \leq b_s.$$

Choose the bounds b_s according to (4.2). At stage s , let $f_u = f_{s-1} + g_u$. Throughout the construction, $|f_u(z) - z| \leq 1/16$ and $|f'_u(z) - 1| \leq 1/16$ on the unit disc. Rouché's theorem gives exactly one zero there, denoted by $\alpha(u)$, and the equation $f_u(\alpha(u)) = 0$ shows $|\alpha(u)| \leq 1/16$. In particular, the roots remain in a fixed compact subdisc. For z, t in the unit disc,

$$f_u(z) - f_u(t) = (z - t) + \int_{[t,z]} (f'_u(\zeta) - 1) d\zeta.$$

This gives (B.1) with $m = 1/2$ and $M = 3/2$.

Every old query node inside the unit disc has a nonzero function value that is preserved by all subsequent corrections: when first queried it was not the unique selected root. An old node outside that disc cannot equal any selected root. Hence $w_s(\alpha_{s-1}) \neq 0$.

Set $x = \alpha_{s-1} + \varepsilon$. At the old root the perturbation is at most $2b_s \varepsilon$, so (B.1) yields a root shift at most $4b_s \varepsilon \leq \varepsilon/4$. Consequently

$$\frac{3}{4}\varepsilon \leq |x - \alpha(u)| \leq \frac{5}{4}\varepsilon.$$

With w_s fixed, the complex finite-scale argument supplies a constant $c_s > 0$ and a scale threshold. Choose ε_s below this threshold and satisfying (4.7), and then select f_s and x_s as in Section 4. Uniform convergence on every complex disc gives an entire limit, preserves every observed derivative value, and yields

$$|\alpha - \alpha_s| \leq 4b_{s+1} \leq \min\{E_s/8, \varepsilon_s/8\}.$$

The same induction through the adaptive queries preserves the outputs. The final error-ratio estimate follows as in the real proof, using absolute values in (4.6). This proves the theorem. $\square$

References

- [1] R. P. Brent, S. Winograd, and P. Wolfe, *Optimal iterative processes for root-finding*, Numer. Math. **20** (1973), no. 5, 327–341, Errata.
- [2] V. Casulli and D. Trigiante, *The convergence order for iterative multipoint procedures*, Calcolo **14** (1977), no. 1, 25–44.

- [3] ———, *Multipoint iterative parallel methods for solving equations*, *Calcolo* **15** (1978), no. 2, 147–160.
- [4] C. de Boor, *Divided differences*, *Surv. Approx. Theory* **1** (2005), 46–69, arXiv:math/0502036.
- [5] K. M. Dyakonov, *Wronskians and deep zeros of holomorphic functions*, *J. Math. Pures Appl.* (9) **99** (2013), no. 6, 668–684.
- [6] A. Eremenko and A. Gabrielov, *Degrees of real Wronski maps*, *Discrete Comput. Geom.* **28** (2002), no. 3, 331–347.
- [7] H. T. Kung and J. F. Traub, *Optimal order of one-point and multipoint iteration*, *J. ACM* **21** (1974), no. 4, 643–651.
- [8] ———, *Optimal order and efficiency for iterations with two evaluations*, *SIAM J. Numer. Anal.* **13** (1976), no. 1, 84–99.
- [9] Matthew J. Colbrook, *A formal proof of the Kung–Traub conjecture*, Palomar, PALOMAR-2026-09-20-000006 v1, 2026.
- [10] M. S. Petković and L. D. Petković, *Families of optimal multipoint methods for solving nonlinear equations: a survey*, *Appl. Anal. Discrete Math.* **4** (2010), no. 1, 1–22.
- [11] J. F. Traub, *Computational complexity of iterative processes*, *SIAM J. Comput.* **1** (1972), no. 2, 167–179.
- [12] G. W. Wasilkowski, *n -Evaluation conjecture for multipoint iterations for the solution of scalar nonlinear equations*, *J. ACM* **28** (1981), no. 1, 71–80.
- [13] A. G. Werschulz, *Maximal order for multipoint methods with memory using Hermitian information*, *Internat. J. Comput. Math.* **9** (1981), no. 3, 223–241.
- [14] H. Woźniakowski, *Generalized information and maximal order of iteration for operator equations*, *SIAM J. Numer. Anal.* **12** (1975), no. 1, 121–135.
- [15] ———, *Maximal order of multipoint iterations using n evaluations*, *Analytic Computational Complexity* (J. F. Traub, ed.), Academic Press, New York, 1976, pp. 75–107.