Reconstruction of functions from perturbative data: optimal methods and error analysis

Ovidiu Costin (OSU), collaboration with G. Dunne, M. Meying (U Conn)

Les Diablerets, January 2023

LaTeX beamer

CMP (2022), J. Phys. A 55 (2022), EPJ (2021), Physics Letters B (2020), Journal Phys. A (2019, 2017), arXiv:2108.01145, 2003.07451, 2208.02410, 1904.11593, 1705.09687

Truncated perturbation expansions

- Setting: a function F is only known as n of terms of a perturbation expansion convergent, or resurgent, possibly with limited accuracy, and we need to find F as precisely as possible.
- Let M_n be the Maclaurin polynomial obtained by Borel transforming the series.
- What is the maximal accuracy in determining F from M_n?
- Without further information about F, this question is of course ill posed. But: Often we know or guess the position of Borel plane singularities (the underlying Riemann surface Ω) and a rate of growth of F.
- We find that with it the questions become well posed. The optimal reconstructing procedure is explicit, and **dramatically** improves over classical methods.
- Also very useful when there are no theoretical limitations on the number of terms/accuracy: understanding the Borel plane and physical domain of analytic functions based on easier to generate perturbative expansions.
- We have now quantified the effects of numerical noise of the optimal approach as well as of traditional ones s.a. Padé.

" What the package includes"

- The explicit **optimal reconstruction** formula of F given M_n and Ω; this formula is based on uniformizing maps of Riemann surfaces;
- We obtained the **uniformization maps** for Riemann surfaces commonly encountered in applications;
- **Continuation of F on many sheets of Ω** (often tens of sheets, and in a precise sense, infinitely high/deep).
- (With some more info): Operators **removing** (or modifying, if desired) **a singularity** at a given point to further improve accuracy.
- **Coefficient extrapolation.** Meaning: e.g., from the first n terms of the expansion, determine with good precision, say $2n$ coefficients.
- **Global reconstruction in the physical domain.** Returning to the physical domain, find, from the asymptotic series at $+\infty$ precise information on the global behavior in the physical domain.
- Formulas for the **sensitivity to numerical errors**.

Truncations of series of "natural origin" contain hidden info: proof of concept

Claim

- The first n (large enough) Maclaurin coefficients of a given function [1] determine with high accuracy the $n+1$th coefficient.

Method (not optimal), using Padé

(take a known function to test it)

- Calculate the maximal possible diagonal Padé approximant $[n/2, n/2]$ of M_n.
- Calculate $\tilde{P}_{n+1}$ **of** $[n/2, n/2]$. It is very close to the actual one! (Quite a few more Maclaurin coeffs can be calculated accurately.)
- (Try it!) Detailed explanation in CMP 2022.

[1] known to be analytic in $\mathbb{D}$ with some number of singularities in $\mathbb{D}^c$

First, a classical tool: Padé

- Assume for definiteness that the (maximal) disk of analyticity of a function F, for which we know M_{2n}, is $\mathbb{D}$.
- One the oldest tools in the "reconstruction trade" are Padé approximants.
- Padé is the "best" approximant of F as a rational function $P_n(x)/Q_n(x)$
- In particular, the best approximation condition has to hold in a tiny neighborhood of the origin. Close to 0, the Maclaurin polynomial is already optimal. (We prove this later.) Hence the choice:

Definition

The Padé approximant $[n, n]$ associated to $M_{2n}(F)$ is the (unique up to normalization) ratio of polynomials P_n/Q_n having the same Maclaurin polynomial, M_{2n}.

- Padé approximants have remarkable properties, including faster convergence in $\mathbb{D}$ than the Maclaurin series and often, convergence (in the sense of capacity) as $n \to \infty$ to F even outside $\mathbb{D}$, in a larger domain of analyticity in $\mathbb{C}$.

Padé, revisited: conformal maps, and electrostatics

There is a remarkable and intuitively useful physical interpretation of the domain $\mathcal{D}_P$ of Padé convergence and placement of Padé poles, best seen if the expansion point is ∞. This is derived from a fundamental paper of H. Stahl [2]:

1. Given F, let $\mathcal{D}$ be any simply connected single-valuedness domain of $F(\omega^{-1})$ and let $E = \partial\mathcal{D}$. Think of E as an **infinitely flexible, perfect conductor**.
2. **Capacitance.** Place a unit charge on E, and normalize the electrostatic potential $V(x, y) = V(\omega), \omega = x + iy$ w.r.t. $\omega = \infty$ by $V(E) = 0$.
   Define the capacitance of E by C=cap$(E) = 1/V(\infty)$.
3. **Minimal capacitor**. Deform E to the set E_P of minimal capacity under these constraints. Then $\mathcal{D}_P = \mathbb{C} \setminus E_P$ is the domain of convergence of Padé, *in the sense of capacity*.
4. The electrostatic equilibrium distribution of charges along E_P is the *equilibrium measure* μ. As $j \to \infty$ the poles of the near diagonal Padé approximants are placed (close to) E_P, with distribution (close to) μ.
5. Let ψ be the conformal map from $\mathcal{D}_P^c$ to $\mathbb{D}$. Then, $V(\omega^{-1}) = -\operatorname{Re} \ln \psi$.

[2] "The Convergence of Padé Approximants to Functions with Branch Points", JAT **91** (1997).

Proposition (Slight enhancement of Stahl's, OC, Dunne, Meynig (2022))

Let F be analytic in $\mathbb{D}$ with branch points in $\mathbb{C}$. Then, for large n and $\omega \in \mathcal{D}_P$,

$$(F(\omega) - [n,n]_F(\omega))^{1/2n} = e^{i\lambda}\psi(\omega)(1+o(1))$$

Here $e^{i\lambda}$ is some phase, $[n,n]_F$ is the order n diagonal Padé of F and $\mathcal{D}_P$ is the domain of convergence of Padé.

In a log sense, the approximation $[n,n]_F$ provides the conformal map ψ. *The approximation quality is* $|\psi|^n$.

- Padé maps a "large" domain $\mathcal{D}$ of analyticity of F to $\mathbb{D}$, calculates the Maclaurin series there, and maps it back to $\mathcal{D}$.

Optimal reconstruction: Notations and conventions

- By the celebrated uniformization theorem, a sweeping generalization of the Riemann mapping theorem, any simply connected Riemann surface Ω is conformally equivalent to one of $\mathbb{D}, \mathbb{C}, \hat{\mathbb{C}}$. In all but the simplest cases, it is $\mathbb{D}$.
- Because of analyticity at zero in Borel plane Riemann surfaces of interest contain a disk around zero, say $\mathbb{D}$, on the first Riemann sheet.
- Take such an Ω; we denote by ψ the conformal map to $\mathbb{D}$, the *uniformization map*, and let $\varphi = \psi^{-1}$

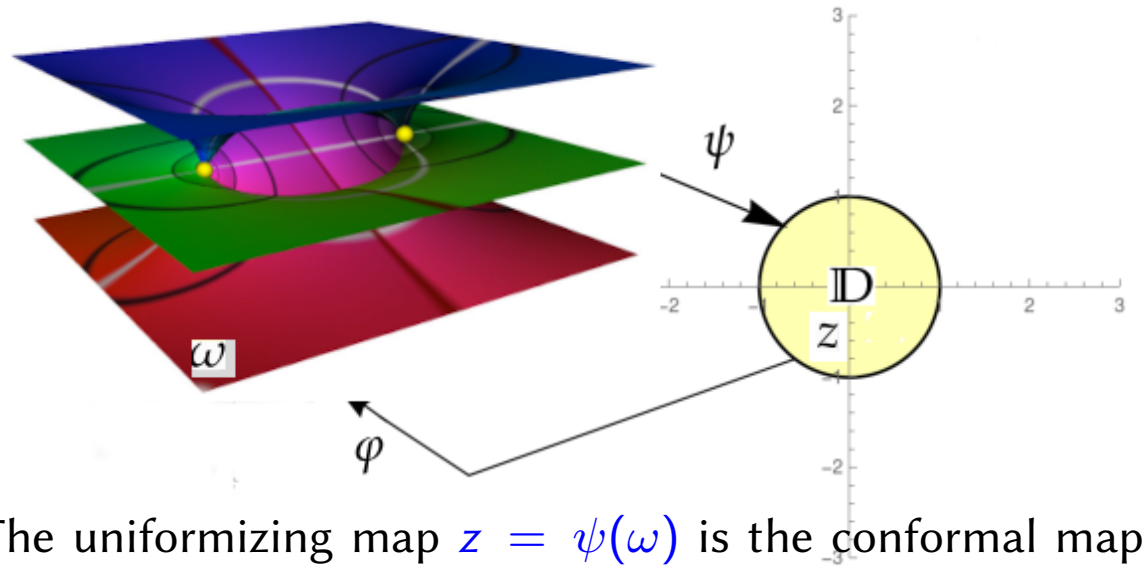

Figure: Notation. The uniformizing map $z = \psi(\omega)$ is the conformal map from the simply connected Riemann surface Ω to $\mathbb{D}$; $\varphi = \psi^{-1}$. The map is normalized as usual, $\psi(0) = 0, \psi'(0) > 0$.

The optimal reconstruction procedure

- Let $F(\omega)$ be analytic on Ω with Maclaurin polynomial $M_n(\omega)$.
- Let ψ be the uniformization map (conformal map to $\mathbb{D}$) of Ω s.t. $\psi(0) = 0, \psi'(0) > 0$ and let $\varphi = \psi^{-1}$.
- Note that $\varphi(\mathbb{D}) = \Omega$ and $F \circ \varphi$ is analytic in $\mathbb{D}$.

The most accurate (quantified below) reconstruction of F is as follows.

1. Take $(M_n \circ \varphi)(z)$.
2. Expand $(M_n \circ \varphi)(z)$ in Maclaurin series at zero.
3. **Discard** all terms beyond the nth (‼) We get a polynomial $(M_n \circ \varphi)_n$.
4. The best approximant is $\boxed{\hat{R}_n := (M_n \circ \varphi)_n \circ \psi}$.
5. To improve the accuracy of reconstruction, we needed to throw away part of the information we had.

Less is more!

1. Take $(M_n \circ \varphi)(z)$. Expand $(M_n \circ \varphi)(z)$ in Maclaurin series at zero.
2. **Discard all terms** beyond the nth (!) We get a polynomial $(M_n \circ \varphi)_n$. [3]
3. The best approximant is $\boxed{\hat{R}_n = (M_n \circ \varphi)_n \circ \psi}$.

[3] Any further terms of the Maclaurin series of $M_n \circ \varphi$, **calculated** from $M_n \circ \varphi$, lead to **loss** of accuracy, in fact at an **exponential rate** in n.

- Intuition: The Taylor coefficients of F are the Fourier coefficients of $F_{|\partial\mathbb{D}}$. For the general class of functions analytic in $\mathbb{D}$ and in $L^2(\partial\mathbb{D})$, the best approximation that can be gotten out of a given number of Fourier coefficients is precisely the associated Fourier sum (the orthogonal projection on that space).
- "Hence" in a space of functions analytic in $\mathbb{D}$ and no further, the Taylor polynomial is already optimal.
- "Hence": map to $\mathbb{D}$ the whole analyticity domain of F and use the mapped Maclaurin polynomial as the approximant: this must be the best. (It is a good intuition, but not how the proof goes.)

Theorem (OC, G. Dunne 2020)

–Let $\omega_0 \in \Omega$, and M_n an $(n-1)$-order truncation of a Maclaurin series, converging in $\mathbb{D}$.

–Define $\mathcal{F}_n = \{F \text{ analytic on } \Omega : \|F\|_\infty < \infty, \text{ and } M_n \text{ is the Maclaurin coeff. of } F\}$

–Let $\hat{R}_n = (M_n \circ \varphi)_n \circ \psi$.

Rate of approximation: Let $\omega_0 \in \Omega$. For $F \in \mathcal{F}_n$ we have

$$\frac{|F(\omega_0) - \hat{R}_n(\omega_0)|}{\|F\|_\infty} \leqslant \frac{|\psi(\omega_0)|^n}{1 - |\psi(\omega_0)|} \ (*)$$

Optimality: $\forall R_n \in \mathbb{C}$ and $\delta > 0$ $\exists F_\delta \in \mathcal{F}_n$ s.t. $\dfrac{|F_\delta(\omega_0) - R_n|}{\|F\|_\infty} \geqslant |\psi(\omega_0)|^n(1-\delta)$

- $(*)$ Since $\Omega \supset \mathbb{D}$, we have $|\psi(\omega)| < |\omega|$ for $\omega \in \mathbb{D}$.
- **Note.** The sequence $\{(M_n \circ \varphi)_n \circ \psi\}_{n\in\mathbb{N}}$ converges on the whole of Ω.
- Weighted bounds are covered too, in the paper.
- The method is independent of ω_0 and n but is optimal at any ω_0 and any n.

Quantifying convergence improvement: Near zero

- Say the coefficients of M_n are bounded by some c, ensuring errors $\sim c\omega^n$. Then the errors in the optimal method are $\sim \psi'(0)^n \omega^n c$; we note that $0 < \psi'(0) < 1$ is *decreasing* with the size of Ω, and it is typically $|\psi'(0)| < \frac{1}{2}$.
- Example: If the unit disk is a natural boundary for F, then M_n (say with coefficients as above) is optimal, with accuracy of order $c\omega^n$. If instead F is analytic in $\Omega = \mathbb{C} \setminus [1, \infty)$ and no further, the accelerated rate is $4^{-n}\omega^n c$. If $[1, \infty)$ is not a natural boundary of F but instead F is analytic on the Riemann surface over $\hat{\mathbb{C}} \setminus \{(0), 1, \infty\}$, as is the case for hypergeometric functions, then the optimal rate is $16^{-n}\omega^n c$.

Near singularities, that is near $\partial\Omega$

- This is even more dramatic: as we have seen, optimally used, M_{200} of $\mathbb{K}$ is still accurate at a distance $\sim 10^{-52}$ of the singular point 1. (With M_{200} alone we can get within 10^{-1} and with Padé 10^{-3}).

Looping around on the Riemann surface: elliptic $\mathbb{K}$

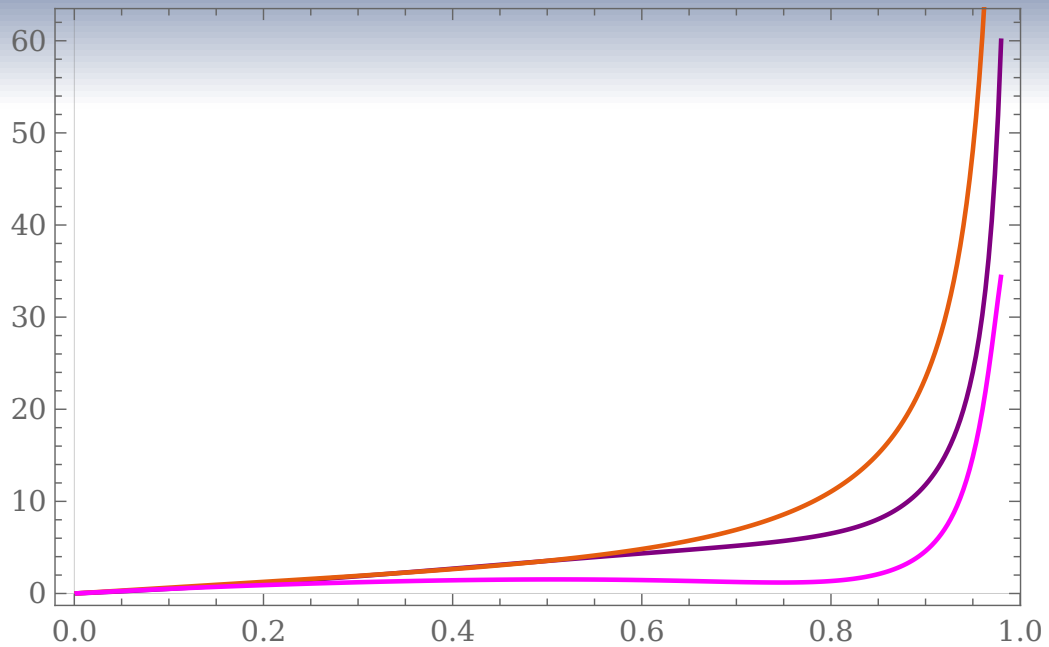

Figure: Approaching singularities on three sheets of $\mathbb{K} \circ \varphi$ using M_{200}. The curves are parametrized by $t \in [1, 1)$. The shape of the graph matches log singularities with specific "Stokes" constants.

- Singularity of $\mathbb{K}$ at 0 on sheet -1 (circling 1 once). It can be approached within $|\varepsilon| \sim 10^{-52}$, keeping 7 digits of acuracy.
- The singularity at 0 after 2 loops around 1.
- The singularity at 1 after one loop around 1 and a loop around 0 on the second sheet.

Optimal coefficient extrapolation

- The optimal extrapolation is given by: $M_{n+1} \approx M_n + \frac{1}{n!}[(M_n \circ \varphi) \circ \psi]^{(n)}(0)\omega^n$.
- If the nth Maclaurin coefficient of F is b_n, it differs from this extrapolation by exponentially small relative error, $\psi'(0)^n b_n$.
- In the examples above, $\psi'(0) = 1/4$ and $1/16$ resp.
- Unlike Richardson extrapolation etc. this procedure works even if the coefficients of P are complex, oscillating or whatever.

How accurate is the optimal extrapolation?

- For $\mathbb{K}(m)$, analytic on the Riemann surface over $\hat{\mathbb{C}} \setminus \{(0), 1, \infty\}$, M_8 predicts M_9 with relative error $\varepsilon_9 \sim 5 \cdot 10^{-9}$ (.1% for Padé)
- M_{471} is predicted with maximum relative error in its coefficients of 0.13%.
- With more coefficients, say 60, M_{61} is predicted with relative error $\varepsilon_{61} \sim 5 \cdot 10^{-73}$ and M_{120} with maximum relative error 10^{-51}.

Application: exploring Riemann surfaces, Painlevé P_{I}

Riemann surface and singularities for P_{I}

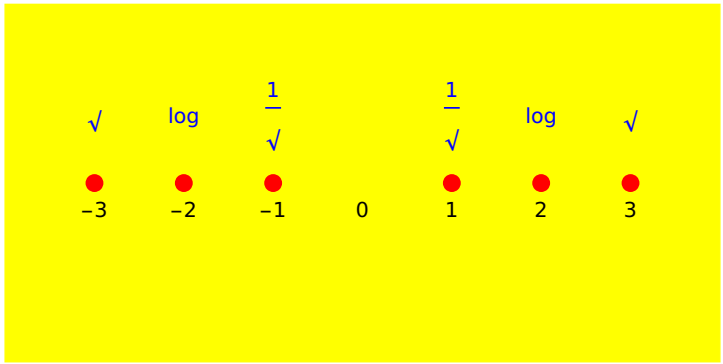

Figure: Borel plane $\Omega_{\mathbb{Z}}$ of P_{I}: universal covering of $\mathbb{C} \setminus \{(0), \mathbb{N}, -\mathbb{N}\}$.

Theorem

$\Omega_{\mathbb{Z}}$ is uniformized by φ^{-1}, $\varphi = \frac{1}{2\pi i} \ln(1 - q^{-1})$, with $q =$ elliptic nome.

- The singularities of tronquée solutions F of P1 lie on $\partial\Omega_{\mathbb{Z}}$. Hence $F \circ \varphi$ is analytic in $\mathbb{D}$ and all singularities above are mapped on $\partial\mathbb{D}$. Scanning $\partial\mathbb{D}$ reveals all singularities on all (in practice, many) Riemann sheets! Next.

Painlevé $P_{\mathrm{I}} \circ \varphi$ on $\partial\mathbb{D}$ from 200 coefficients.

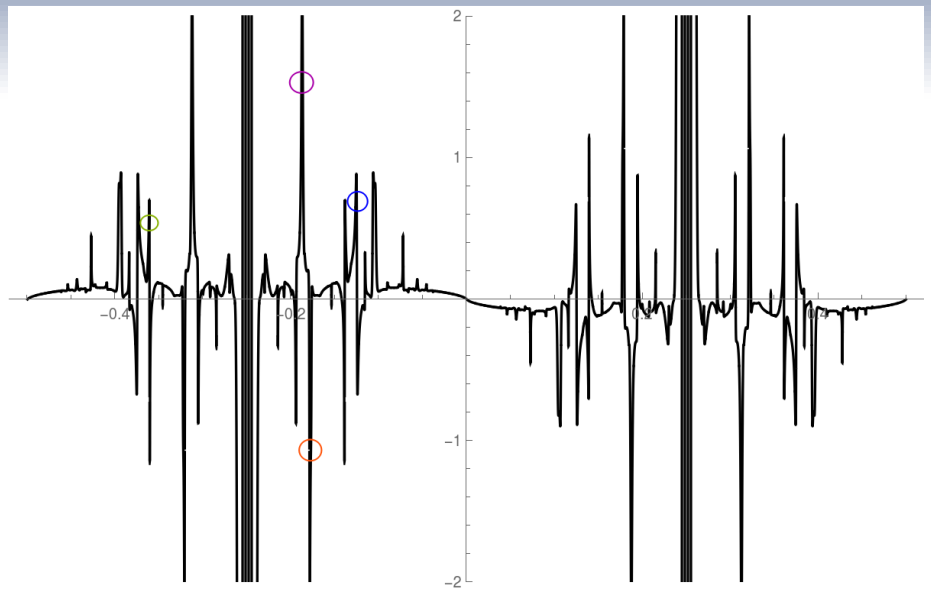

Figure: $|(M_{200} \circ \varphi)_{200}|$ of $P1$ plotted on a circle of radius 0.99, parametrized by $t \in [0, 1]$. We see singularities from many sheets.

–Notice the thick lines (the thickness decreases with the distance to $\partial\mathbb{D}$): these are two exponential singularities. The exponential nature is clear in the "large n" empirical asymptotics of the coefficients of M_{200}.

–Exponential singularities **only exist when the sheet index** $\to \infty$. We see infinitely deep on Ω, and uniformization may be the only way to extract this information from M_{200}.

Application: Global reconstruction in the physical domain

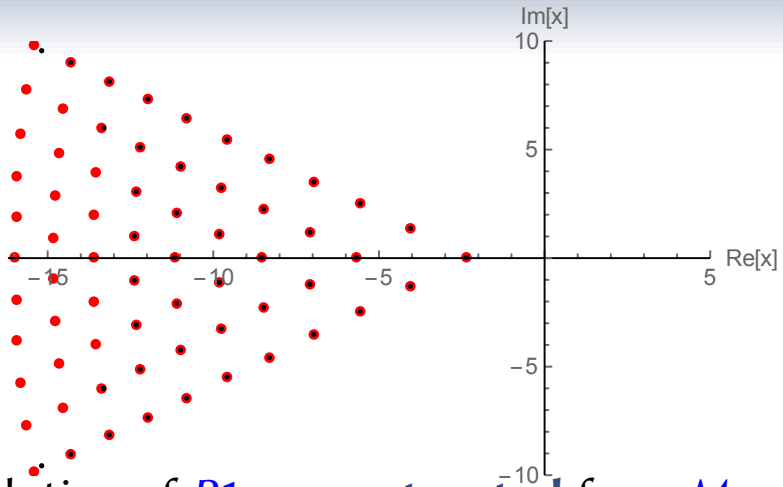

Figure: The tritronquée solution of $P1$ **reconstructed** from M_{200}, Borel transformed from its **divergent expansion as** $x \to +\infty$. There is a $2\pi/5$ wedge of poles (about 66 of which are recovered with high accuracy), and in its complement y is analytic.

- The position x_1 of the first pole and the "energy" constant h_1 at x_1 are important in applications, but not (yet?) known in closed form. Best existing numerical methods provided some 16 digits of accuracy. We get 66 digits of accuracy,

$$x_1 = -2.384168769568816639299145852448767190410408814737850512677 25\ldots$$
$$h_1 = 0.062135739226177640896490141640062460197740771373829663663533 3\ldots$$

The accuracy is roughly preserved throughout the analyticity sector.

Singularity modifier operators and singularity elimination

- Singularity modification & elimination applies to functions for which the *nature of a singularity is known or empirically guessed, in ways we propose.* After elimination of a singularity, the reconstruction accuracy improves substantially.
- In applications, generic singularities in Borel plane (normalized to be at $\omega_0 = 1$) are known to be of the form $(1-\omega)^\alpha A(\omega)+B(\omega)$; $\alpha \in \mathbb{C}\setminus\mathbb{Z}$; or, for $\alpha = m \in \mathbb{Z}$ a pole, or more commonly, and generally:
$$\frac{d^k}{d\omega^k}[(1-\omega)^m \ln(1-\omega)A(\omega)] + B(\omega) \quad A, B \text{ analytic at } 1$$
- An operator of the form $(\omega^{-\beta}[\omega^\beta * F]) \circ h$ where $h(\omega) = 2\omega - \omega^2$ preserves the Borel plane structure except that it transforms the behavior near 1 to $A(\omega) + \tilde{B}(\omega)$, A as before and with B having some explicit expression. It eliminates the singularity without moving it.

Effects of noise

Assume the Maclaurin polynomial M_n is perturbed by a zero average random variable, $\mathcal{N}_\varepsilon = \varepsilon \sum_{k=0}^{n} r_k \omega^k$,

$$M_n \mapsto M_n + \mathcal{N}_\varepsilon$$

For m^{th}-digit truncation errors, the random perturbation is $r_k = \pm\frac{1}{2}10^{-m}$.

- Q: how does numerical noise affect the accuracy of conformal map approximants, and relatedly of Padé?
- A: In a universal way: there is a universal relation among *number of terms, noise, as a random variable, and accuracy, as a random variable.*

  First: Numerical Experiments of Padé "Breakdown" (M. Meynig)

- The noise eventually causes the appearance of arcs of poles which form a natural boundary touching genuine singularities of F. We first illustrate this on the sample functions $(1+\omega)^{-1/9}$ and $(1+\omega^2)^{-1/9}$.

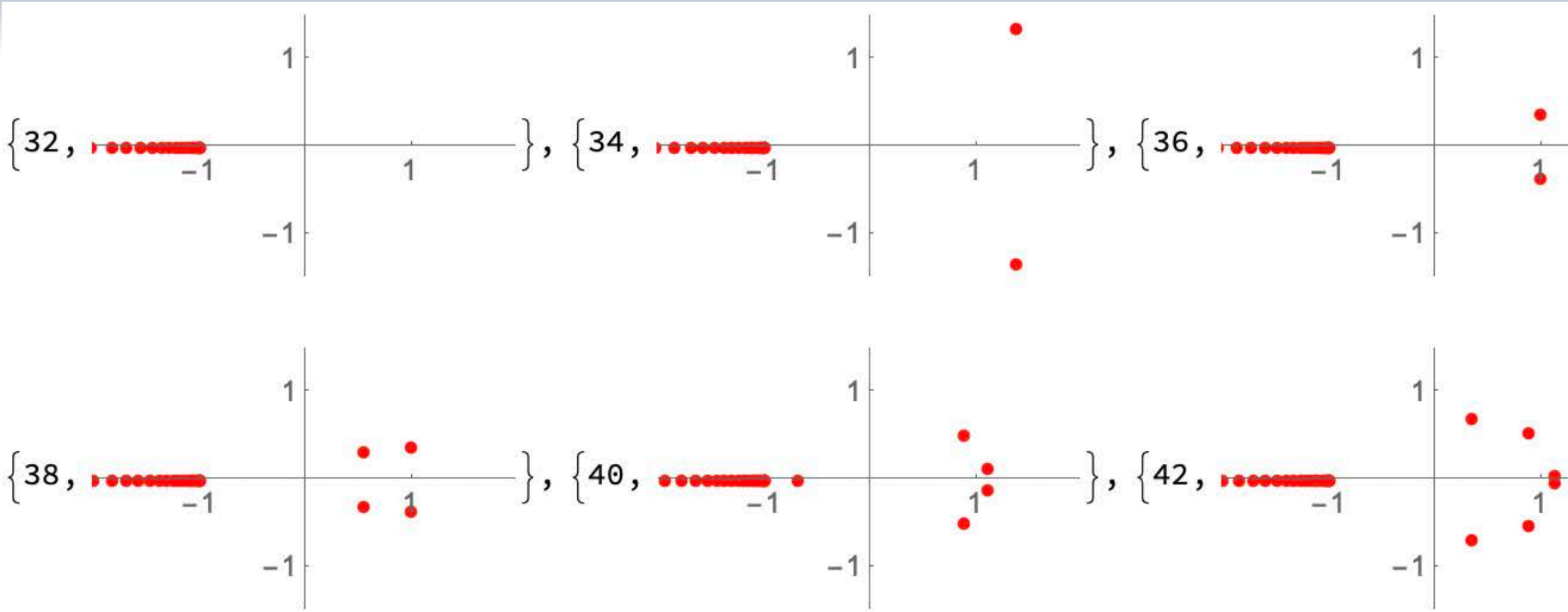

Figure: Padé poles for different orders $[n, n]$ starting with $n = 32$, for $(1+\omega)^{-1/9}$, with coefficients truncated at 40 digits. Spurious noise poles begin to appear at order 33, and in the limit they form as arcs originating at $\omega_{\text{inf}} = +1$, the point of best approximation of Padé in the absence of noise.

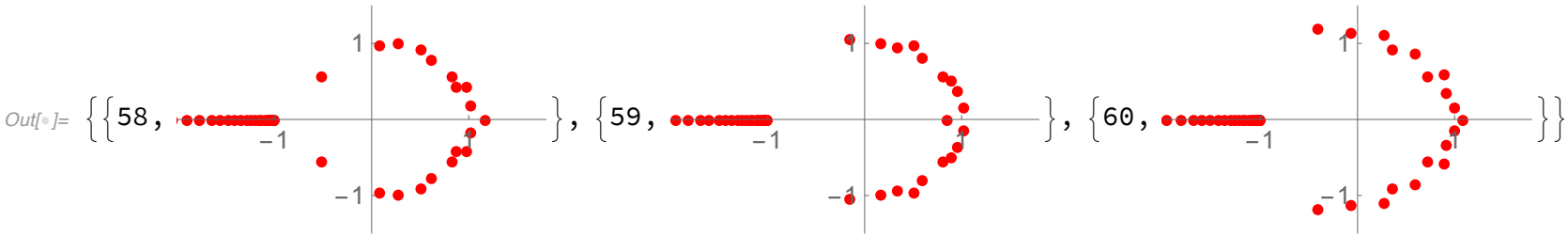

Figure: This pattern explained later.

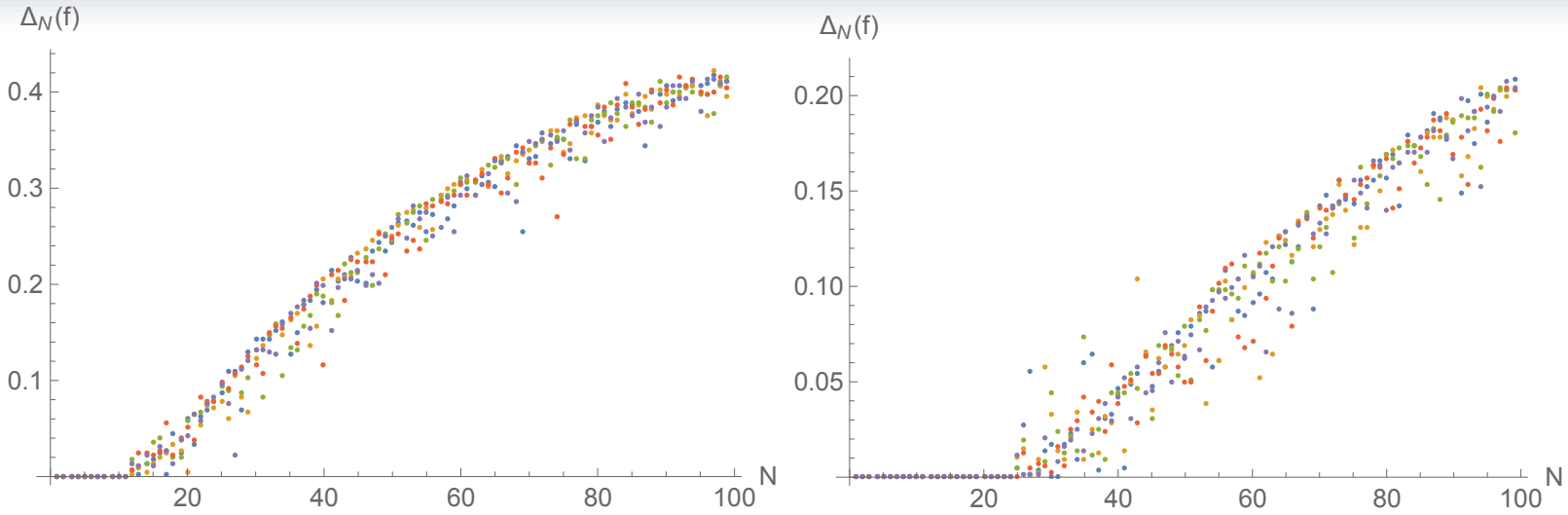

Figure: Plots of Padé breakdown (Δ_n is some measure of accuracy) for five different realizations of the random noise with same noise strength $\varepsilon = 10^{-20}$ for functions with 1 and 2 branch points (-1 and $-1, 1$). The order of breakdown is twice as large for the two-singularity function.

Intuition. If $\omega \in \partial\mathbb{D} \subset \Omega$ is no singular then $M_n(\omega)$ gives an accuracy $O(\omega^{n+1}) = O(1)$ if n is not large. Instead, $M_n(F \circ \varphi)$, has accuracy $O(\psi(\omega)^{n+1})$, exponentially small in n since $|\psi(\omega)| < 1$. This accuracy comes from massive cancellations, perturbed by noise, the source noise sensitivity. (Padé can be put in this framework too.) Write

$$(F \circ \varphi)(z) = \sum_{j=0}^{\infty} F_j \varphi^j(z) := \sum_{k=0}^{\infty} c_k \, z^k \tag{1}$$

Definition

In approximating F from its truncated expansion $F_{[m]}$ by using a conformal map or Padé, both relying on $F \approx (F \circ \varphi)_{[k]} \circ \psi$, $k \geqslant n$, we measure the noise-induced breakdown by one of the following criteria

1. First k for which c_k becomes inaccurate.
2. For a fixed $z \in \mathbb{D}$, first k for which $(F \circ \varphi)(z) \approx M_k(F \circ \varphi)(z)$ becomes inaccurate.
3. First k for which there is a $z \in \mathbb{D}$ for which $(F \circ \varphi)(z) \approx M_k(F \circ \varphi)(z)$ becomes inaccurate.

1 and **3** are roughly equivalent, while **2** provides more local information.

- Let $\mathcal{N}_\varepsilon$ be the noise function above and let $n(z)$ be the **the composition of $\mathcal{N}_\varepsilon$ with φ.**

$$n(z) := (\mathcal{N}_\varepsilon \circ \varphi)(z) = (\sum_{k \geqslant 0} r_k \omega^k) \circ \varphi) = \sum_{k=0}^{\infty} n_k z^k \tag{2}$$

- n_k are exponentially larger than r_k, more precisely as follows

Proposition

1 As before, let $\psi := \varphi^{-1}$. [a]

For each realization of the noise variables r_i we have, with probability one,

$$\limsup_k |n_k|^{\frac{1}{k}} = |z_{\text{inf}}|^{-1} \quad \text{where} \quad |z_{\text{inf}}| = |\psi(\omega_{\text{inf}})| := \text{dist}(\mathfrak{b}, 0) \tag{3}$$

[a] With probability one, the unit circle S^1 in Ω is a natural boundary of $\mathcal{N}_\varepsilon$, and $\mathfrak{b} = \{z : |\varphi(z)| = 1\} = \psi(S^1)$ is a natural boundary of $\mathcal{N}_\varepsilon \circ \varphi$.

The crucial quantity is z_{inf}, the closest point to the origin of $\psi(\mathbb{D})$

Theorem

For large k, n_k is a random variable of zero average and standard deviation

$$\sigma(n_k) = \varepsilon\, A\, k^{-\frac{1}{4}} |z_{\text{inf}}|^{-k} (1 + o(1)) \tag{4}$$

where

$$A = 3^{-\frac{1}{2}} (2\pi)^{\frac{3}{4}} |\psi'(\omega_{\text{inf}})| \, [\operatorname{Re} \alpha(\omega_{\text{inf}})]^{-\frac{1}{4}}$$

(A relatively delicate asymptotic problem!)

Corollary

For given error threshold δ,

1. For $\omega \in \mathbb{D}^c$, with $z_\omega = \psi(\omega)$ the approximation breakdown condition is

$$A\,k^{-1/4}\left|\frac{z_\omega}{z_{\text{inf}}}\right|^k\left|\frac{1}{1-z_{\text{inf}}/z}\right| \gtrapprox \delta/\varepsilon \tag{5}$$

with $\min_k |z|^k(1+\varepsilon A k^{-1/4}|z_{\text{inf}}|^{-k}) \approx \frac{\tau}{\tau-1}(A k_{min}^{-1/4}\varepsilon\tau)^{1/\tau}$, $\tau = \ln(z_{\text{inf}})/\ln z$. In the limiting case $|z| = 1$, **2** reduces, up to a constant, to the condition in **1**.

- **Note.** As a function with $O(1)$ random coefficients, $\mathcal{N}_\varepsilon$ is analytic in $\mathbb{D} \subset \Omega$, and the unit circle $S^1 \subset \Omega$ is, with probability one, a natural boundary.
- Extrapolation quality relies on cancellation of coefficients, hence the points of best extrapolation accuracy are also those which are most sensitive to noise. Indeed, the effects of the noise first appear there; thereafter, noise poles spread out on a "circle of noise" in the "order" of approximation quality.

Critical n

For diagonal $[n, n]$ Padé we identify $m = 2n$ and we have the general characterization of breakdown:

$$n_c = \frac{\log_{10}(\varepsilon/\delta)}{2\log_{10}(z^{\inf})} \tag{6}$$

ε =numerical accuracy, δ is the "tolerated" accuracy. Importantly, the expression does not depend on the specificities of F, but just the locations of its singularities. In practical applications, even approximate information about the location of the *leading* singularities of F can be used to obtain good approximations to n_c.