The Flow of Polynomial Roots under Differentiation

Alexander Kiselev

Department of Mathematics
Duke University

joint work with Changhui Tan (University of South Carolina)

NYU Abu Dhabi
January 2021

Introduction

What happens with roots of polynomials or entire functions under differentiation is a classical problem.

Gauss, Lucas 1860-70s: zeroes of p' lie in the convex hull of the polynomial p.

Polya, Wiman 1930-40s: famous conjectures regarding behavior of roots of some entire functions under differentiation. Some solutions: Craven, Csordas and Smith 1987, Sheil-Small 1989.

Marcel Riesz 1920s: the smallest gap between roots of real valued polynomial grows under differentiation (later J. v. Szökefalvi-Nagy, Walker).

A tendency towards "crystallization" - roots of higher order derivatives lining up along ideal lattice - has been conjectured for some classes of entire functions by Farmer-Rhoades 2005 and proved in random setting by Pemantle-Subramanian 2017, and for certain random trigonometric polynomials by Farmer-Yerrington 2006.

Michael Berry 2005: universal nature of oscillations in higher order derivatives of entire functions and connection to quantum theory.

Steinerberger 2018: formal derivation of a PDE that should control evolution of roots under differentiation for a real valued polynomial $p_n(x)$, $x \in \mathbb{R}$, of large degree n, in the case where initially roots are distributed according to a smooth density $u_0(x)$. The aim is to understand the zeroes of $p_n^{(k)}(x)$, with $k \sim tn$, $0 < t < 1$.

The PDE is given by

$$\partial_t u + \frac{1}{\pi}\partial_x \left(\arctan \left(\frac{Hu}{u} \right) \right) = 0 \qquad (1)$$

or

$$\partial_t u + \frac{1}{\pi} \frac{u\Lambda u - Hu\partial_x u}{u^2 + Hu^2} = 0.$$

Here H is the Hilbert transform and $\Lambda = (-\Delta)^{1/2}$.

One differentiation corresponds to time step $1/n$, so roots of $p^{(k)}(x)$, $k = tn$, are conjectured to be distributed according to $u(x,t)$.

Surprisingly, the same PDE was formally derived by Shlyakhtenko and Tao 2020 as the evolution equation for free fractional convolution of a probability measure on $\mathbb{R}$ (an object in free probability).

The free convolution of two probability measures $\mu \boxplus \nu$ can be defined to be the law of $X + Y$, where X, Y are freely independent noncommutative random variables with law μ and ν respectively. The notion can be generalized to $\mu^{\boxplus k}$ for $k \geq 1$ (free fractional convolution). This object is relevant, in particular, for non-commutative central limit theorem (leading to Wigner's semicircle law).

Hoskins and Kabluchko 2020 provide a different way to compute the distribution of the roots of $p_n^{(tn)}(x)$, in the limit $n \to \infty$. Then they rigorously connect differentiation of polynomials and free fractional convolution; their result applies for each fixed time in the limit of degree $n \to \infty$, and does not directly involve the PDE (1).

We will consider the periodic setting, namely, a class $\mathbb{P}_{2n}$ of trigonometric polynomials

$$p_{2n}(x) = \sum_{j=1}^{n} (a_j \cos jx + b_j \sin jx) = \prod_{j=1}^{2n} \sin \frac{x - x_j}{2}$$

that will be assumed to have exactly $2n$ distinct roots $x_j \in \mathbb{S}$, $j = 1 \ldots, 2n$. These roots are assumed to be distributed according to a smooth density u_0.

Note that the process of "crystallization" under differentiation for such polynomials is not hard to understand on elementary level. As we differentiate $k \sim An$ times, the highest order coefficients gain $\sim e^A$ factor compared to all other coefficient, creating spectral gap. As $k \to \infty$, zeroes should converge to ideal lattice. Rate estimates and details are harder!

Our goal is to link evolution of roots under differentiation and Steinerberger's PDE in this setting.

The Global Regularity

The Steinerberger's PDE is critical, and looks similar to the equation

$$\partial_t \rho + \rho \Lambda \rho - H\rho \partial_x \rho = 0. \tag{2}$$

The difference is the additional factor of $u^2 + Hu^2$ in the denominator, which makes analysis harder. The equation (2) has appeared in several contexts.

1. A model in dislocation theory: several works, earliest by Head 1972. Self-similar solutions.
2. The model of porous media flow: Caffarelli-Vazquez 2011, 2015. Hölder regularity of weak solutions.
3. Euler alignment model of collective behavior: many results. Shvydkoy-Tadmor 2017 proved global regularity.
4. As a simplified model of the SQG equation: Chae, A. Cordoba, D. Cordoba, Fontelos 2005; Castro-D. Cordoba 2008

Granero-Belinchon 2020: local well-posedness in H^2, global regularity for small data for the Steinerberger's equation.

Recall the Steinerberger's PDE:

$$\partial_t u + \frac{1}{\pi} \frac{u\Lambda u - Hu\partial_x u}{u^2 + Hu^2} = 0. \tag{3}$$

Theorem

The equation (3) with H^s, $s > 3/2$ periodic initial data such that $u_0(x) > 0$ for all $x \in \mathbb{S}$ has a unique global smooth solution $u(x,t)$. Moreover, we have exponential in time convergence to equilibrium

$$\|u(\cdot,t) - \bar{u}\|_\infty \leq C_0 e^{-\sigma t},$$

and exponential in time decay of all derivatives

$$\|\partial_x^k u(\cdot,t)\|_\infty \leq C_k e^{-\sigma t}$$

with $\sigma = \frac{2}{\pi^2 \bar{u}}$ and with constants C_k that may only depend on u_0.

The proof is trickier than that for critical SQG or Euler alignment model!

The Connection

Recall that a polynomial $p_{2n}(x) \in \mathbb{P}_{2n}$ satisfies $p_{2n}(x) = \prod_{j=1}^{2n} \sin \frac{x - x_j}{2}$.
Denote $\bar{x}_j = \frac{x_j + x_{j+1}}{2}$. Define the error

$$E_j^t = x_{j+1}^t - x_j^t - \frac{1}{2nu(\bar{x}_j^t, t)}, \quad j = 1, \cdots, 2n. \tag{4}$$

Here x_j^t are the roots of $p^{(2nt)}(x)$.

Theorem

Let $u_0 \in H^s(\mathbb{S})$, $s > 7/2$. Suppose that $u_0(x) > 0$ for all $x \in \mathbb{S}$, and $\int_{\mathbb{S}} u_0(x)\, dx = 1$. Let $u(x,t)$ be solution of (3) with the initial data u_0, and let p_{2n} be any trigonometric polynomial that at the initial time obeys (4) with $u = u_0$ and $\|E^0\|_\infty \le Z_0 n^{-1-\epsilon}$ for some $\epsilon > 0$. Then there exists positive constant $n_0(u_0, Z_0, \epsilon)$ such that if $n \ge n_0(u_0, Z_0, \epsilon)$, the following estimate holds true for all times $t \ge 0$:

$$\|E^t\|_\infty \le C(u_0)\left(Z_0 n^{-1-\epsilon} + n^{-3/2} t\right) e^{-\frac{4}{\pi} t(1 + O(n^{-\epsilon/2}))}.$$

Remarks

For the global regularity proof, we need the density to be bounded away from zero. This cannot hold in the real line case, so new ingredients are needed.

A surprising fact is that we can control the error for all times.
This is enabled by analysis of the error propagation equation. It turns out to have the form

$$\frac{E^{t+\Delta t} - E^t}{\Delta t} = \mathcal{L}^t E^t + errors.$$

where $\mathcal{L}^t$ is a nonlinear operator of diffusive type that in the main order is similar to a modulated discretized fractional Laplacian $-\Lambda$. In fact, in the limit of large n and large time $\mathcal{L}^t$ converges to exactly the dissipative term

$$-\frac{u\Lambda}{\pi(u^2 + Hu^2)} \sim -\frac{1}{\pi \bar{u}}\Lambda$$

Many further interesting questions!

The Formal Derivation

If $p_{2n}(x) = \prod_{j=1}^{2n} \sin \frac{x-x_j}{2}$, then

$$p_{2n}'(x) = \frac{1}{2}\ p_{2n}(x) \sum_{j=1}^{2n} \cot \frac{x - x_j}{2}.$$

Let $y_m \in (x_m, x_{m+1})$ be the roots of p_{2n}', $m = 1, \cdots, 2n$. In this case

$$\sum_{j=1}^{2n} \cot \frac{y_m - x_j}{2} = 0.$$

Split the sum into two parts:

$$\sum_{j=1}^{2n} \cot \frac{y_m - x_j}{2} = \underbrace{\sum_{|x_j - y_m| \le n^{-1/2}} \cot \frac{y_m - x_j}{2}}_{I_m} + \underbrace{\sum_{|x_j - y_m| > n^{-1/2}} \cot \frac{y_m - x_j}{2}}_{II_m}.$$

For the near field I_m,

$$I_m = \sum_{|x_j - y_m| \le n^{-1/2}} \frac{2}{y_m - x_j} + O(1).$$

Since the summation range is small, replace x_j in the sum by the ideal lattice $\tilde{x}_j = x_m + \frac{j-m}{2nu(\bar{x}_m)}$. Recall the cotangent identity

$$\pi \cot \pi x = \frac{1}{x} + \sum_{k=1}^{\infty} \left(\frac{1}{x+k} + \frac{1}{x-k} \right).$$

Then we expect that

$$\sum_{|x_j - y_m| \le n^{-1/2}} \frac{2}{y_m - x_j} \sim \sum_{k=-2u(\bar{x}_m)n^{1/2}}^{2u(\bar{x}_m)n^{1/2}} \frac{2}{y_m - x_m + k(2nu(\bar{x}_m))^{-1}}$$
$$\sim 4\pi nu(\bar{x}_m) \cot(2\pi nu(\bar{x}_m)(y_m - x_m)).$$

For the far field II_m, as the distribution of $\{x_j\}$ is close to $u(x)$, we formally get

$$\sum_{|x_j-y_m|>n^{-1/2}} \cot\frac{y_m-x_j}{2} \sim 2n\int_{|y-y_m|>n^{-1/2}} u(y)\cot\frac{y_m-y}{2}\,dy \sim 4\pi nHu(y_m)$$

Putting together the two expressions, the leading order $O(n)$ term reads

$$u(\bar{x}_m)\cot(2\pi nu(\bar{x}_m)(y_m-x_m)) + Hu(y_m) = 0.$$

This leads to

$$y_m - x_m \sim -\frac{1}{2n\pi u(\bar{x}_m)}\arctan\left(\frac{u(\bar{x}_m)}{Hu(y_m)}\right).$$

From this microscopic flux we can formally pass to the Steinerberger's equation.

A sketch of the proof of global regularity

Recall

$$\partial_t u + \frac{1}{\pi}\frac{u\Lambda u - Hu\partial_x u}{u^2 + Hu^2} = 0.$$

Local well-posedness is standard; $u_0 \in H^s$, $s > 3/2$ leads to unique solution $u(x,t)$ such that

$$u \in C([0,T]; H^s(\mathbb{S})) \cap L^2([0,T], H^{s+1/2}(\mathbb{S})); \; t^k u \in C((0,T], H^{s+k}(\mathbb{S})).$$

One can show that the solution can be extended beyond a time $T > 0$ if and only if

$$\int_0^T \|\partial_x u\|_{L^\infty} \left(1 + \|\partial_x u\|_{L^\infty}^4\right) dt < +\infty.$$

In addition, there is a maximum principle: if $m(t) = \min_x u(x,t)$, $M(t) = \max_x u(x,t)$ then $m(0) \le m(t) \le u(x,t) \le M(t) \le M(0)$, $\forall\, t$.

For global regularity, we use control of the modulus of continuity

$$\omega(\xi) = \begin{cases} \xi - \xi^{3/2} & 0 < \xi \le \delta \\ \delta - \delta^{3/2} + \gamma \log\left(1 + \frac{1}{4}\log(\xi/\delta)\right) & \xi > \delta. \end{cases}$$

Recall that if u obeys ω, then Hu obeys

$$\Omega(\xi) = C\left(\int_0^\xi \frac{\omega(\eta)}{\eta}\,d\eta + \xi\int_\xi^\infty \frac{\omega(\eta)}{\eta^2}\,d\eta\right).$$

One can show that

$$\Omega(\xi) \lesssim \omega(\xi)\left(4 + \log\frac{\xi}{\delta}\right)$$

for large ξ.

As usual, we need to show that ω is conserved by evolution with arbitrary period. For this, we look at the breakthrough scenario where ω is touched by the solution and prove estimates on the flow and dissipative parts:

$$u(x,t_1) - u(y,t_1) = \omega(\xi),\;\; \xi = |x-y|.$$

The difficulty is that due to more complex form of the coefficients, we are only able to bound the flow term by $\lesssim \omega'(\xi)\Omega(\xi)^2$ in the large ξ range. The standard bound on dissipative term in the large ξ range yields $\lesssim -\frac{\omega(\xi)}{\xi}$. To get conservation of ω, we need the large ξ balance

$$\frac{\omega(\xi)}{\xi} \gtrsim \omega'(\xi)\omega(\xi)^2(\log\xi)^2, \text{ or } \frac{1}{\xi(\log\xi)^2} \gtrsim \omega(\xi)\omega'(\xi)$$

which is only possible for bounded ω.

The source of the difficulty is the scenario where $Hu(x) \gg H(u(y)$ but the bulk of the dissipative estimate comes from $\Lambda u(x)$. The denominator $u(x)^2 + Hu(x)^2$ suppresses the dissipative contribution in this case.

The solution is to realize that although the standard dissipative bound is in general sharp (it saturates for the profile $u(z) \sim \frac{1}{2}\omega(2z)\mathrm{sgn}(z)$), it can be improved in the non-symmetric scenario described above. There is an extra contribution from the lack of symmetry enforced by the above scenario. This extra contribution can be used to control the flow term.

A sketch of the error propagation analysis

Recall the microscopic flux

$$y_m - x_m \sim -\frac{1}{2n\pi u(\bar{x}_m)} \arctan\left(\frac{u(\bar{x}_m)}{Hu(y_m)}\right),$$

and recall that

$$x^t_{m+1} - x^t_m = \frac{1}{2nu(\bar{x}_m)} + E^t_m.$$

To derive the propagation of errors equation, we have to look at the pairs of roots, since

$$E^{t+\Delta t}_m - E^t_m = \frac{1}{2nu(\bar{x}^t_m)} - \frac{1}{2nu(\bar{y}^t_m)} + y_{m+1} - x_{m+1} - y_m + x_m.$$

We need to make estimates on the microscopic flux quite precise, and to use the cancellations present in $I_{m+1} - I_m$ and $II_{m+1} - II_m$.

For example, the main dissipative term in the short range comes from

$$\frac{2}{y_{m+1}-x_{j+1}} - \frac{2}{y_m - x_j} =$$

$$\frac{-2(y_{m+1}-y_m-x_{m+1}+x_m) + \frac{1}{n}\left(\frac{1}{u(\bar{x}_j)} - \frac{1}{u(\bar{x}_m)}\right) + 2(E_j - E_m)}{(y_{m+1}-x_{j+1})(y_m-x_j)}.$$

The blue terms from the last line can be offset with the difference of the corresponding idealized lattice terms. The red ones provide dissipation!

There are many error terms one needs to control: supercritical, critical and subcritical. An example of supercritical term is

$$A_{1,m} \sim n^{-2} \sum_{|j-m| \gtrsim n^{1/2}} \cot \frac{y_m - x_j}{2} E_j.$$

A direct absolute value estimate of $A_{1,m}$ would be of the order

$$n^{-2} \sum_{|j-m| \gtrsim n^{1/2}} \frac{n}{|j-m|} \|E\|_\infty \sim n^{-1} \log n \|E\|_\infty.$$

The bound $\sim n^{-1}\log n\|E\|_\infty$ in discrete evolution equation for E^t would lead to uncontrolled growth in time ~ 1 (corresponding to n iterations). Instead, one can use oddness of cot to rewrite

$$A_{1,m} \sim n^{-2} \sum_{|j-m|\gtrsim n^{1/2}} \cot\frac{y_m - x_j}{2}(E_j - E_m) + \textit{better error}$$

and absorb into dissipation.

There are many critical error terms that are of the order $n^{-1}\|E\|_\infty$ or "forcing" terms that do not depend on E - largest are of the order $\sim n^{-5/2}$. The former terms would lead to a growth by constant factor over time ~ 1. It turns out that all critical terms have a factor of u', u'', or u''' in front of them, so they decay in time and at some point become weaker than dissipation.

If we denote

$$\delta(t) = \|u'\|_{L^\infty} + \|u''\|_{L^\infty} + \|u'''\|_{L^\infty},$$

the final result is:

Theorem

The error E_m^t *propagates according to the evolution equation*

$$E^{t+\Delta t} - E^t = \mathcal{L}^t E^t + O\big(\delta n^{-5/2} + \delta n^{-1}\|E^t\|_\infty\big) + O(n^{-3/2}\|E^t\|_\infty + (\log n)^2\|E^t\|_\infty^2 + n(\log n)^2\|E^t\|_\infty^3),$$

where the operator $\mathcal{L}^t$ *is a diffusive operator given by*

$$(\mathcal{L}^t E^t)_m = \sum_{j, j\neq m} \kappa^t(j,m)(E_j^t - E_m^t).$$

The diffusion coefficients $\kappa^t(j,m)$ *satisfy*

$$\kappa^t(j,m) = \frac{1}{16\pi^2 n^2(u^2 + Hu^2)\sin^2\frac{x_m^{t+\Delta t} - x_j^t}{2}} \times \Big(1 + O(n^{-1/2} + \delta|j-m|n^{-1} + n\log n\|E\|_\infty + \text{better errors})\Big)$$

The dissipative term does not control the mean of the error E_m, but note that $\int_{\mathbb{S}} u(x,t)\,dx = 1$ does not depend on time. Using midpoint Riemann sum, we have

$$\begin{aligned}\int_0^{2\pi} u(x,t)\,dx &= \sum_j \left(u(\bar{x}_j^t,t)(x_{j+1}^t - x_j^t) + O(\bar{x}n^{-3})\right) \\ &= \sum_j \left(\frac{1}{2n} + E_j^t u(\bar{x}_j^t,t) + O(\delta n^{-3})\right) \\ &= 1 + \sum_j E_j^t u(\bar{x}_j^t,t) + O(\delta n^{-2}).\end{aligned}$$

This leads to

$$\sum_{j=1}^{2n} E_j^t u(\bar{x}_j^t,t) = O(\delta n^{-2})$$

for all t. This provides a Poincaré-type constraint for our setting and this along with earlier remarks makes infinite time control possible.

Discussion

1. Our results show that the Steinerberger's equation approximates evolution of roots of a class of trigonometric polynomials under differentiation very well for large n. We need the root spacing to be regular on small scale (microscale errors $\sim n^{-1-\epsilon}$ are acceptable), but variations in spacing can be large on unit scale.

2. A similar scheme should work for the compact support real line case, but free boundary needs to be understood.

3. In the real line case, there is a connection with minor process for random matrices. The analog for the periodic setting is currently unclear.