YMC Plenary Talks

Nir Elber

August 2022

Contents

1 Newton's Method for the Clueless — Jeff Diller

Math is more than numbers.

Idea 1. For this talk, math is about solving equations.

To be specific, we are interested in solving for x solving $f(x) = 0$, where f is some specified function (often polynomial).

Example 2. Here are some sets meant to solve equations in $\mathbb{Z}[x]$.

- Integers solve equations $x - k \in \mathbb{Z}[x]$.
- Rational numbers solve equations like $ax + b \in \mathbb{Z}[x]$.
- Real numbers solve equations like $x^2 - 2 \in \mathbb{Z}[x]$.
- Complex numbers solve equations like $x^2 + 1 \in \mathbb{Z}[x]$.

It will be difficult to "algebraically" solving polynomials arbitrarily; this is impossible for degree at least 5. In this talk, we will be interested in numerical approximations of these solutions. For this, we will use Newton's method.

Remark 3. Newton's method is attributed to Newton, but it's probably due to Simpson. Other mathematicians who have a claim to fame here are Joseph Raphson and Seki Takakazu.

1.1 Newton's Method

We follow the following idea.

Idea 4. If we want to solve $f(x) = 0$, guess a solution x_0 and pretend f is linear.

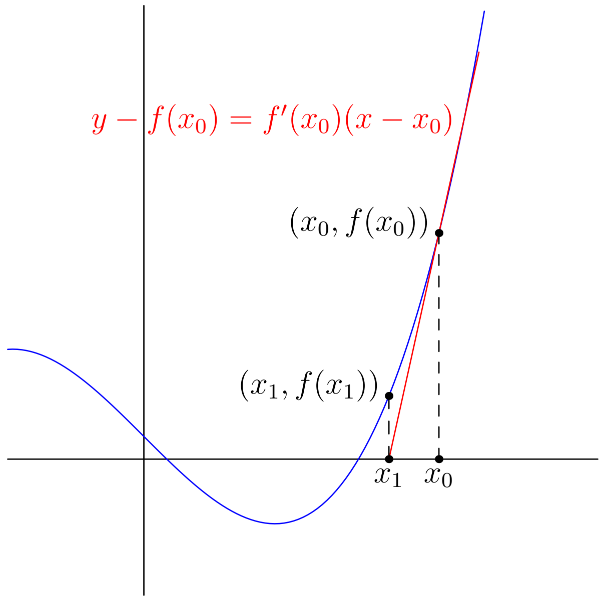
To be a little more technical, we begin with a guess x_0, approximate

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

and then set

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)}$$

to be our new guess. Here is the image.



In practice, we will take our guess x_1 and repeat the process, and we will end up well-approximating the root.

Remark 5. This is an example of a "dynamical system," where we have some function $N\colon X \to X$ and ask what happens to the orbits of points $x \in X$. In particular, Newton's method has

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

Example 6. With $f(x) = x^2 - 2$, we have

$$N(x) := \frac{1}{2}\left(x + \frac{2}{x}\right)$$

as our iteration function. In particular, for any field K, we have $N\colon K \to K$.

Example 7. With $f(x) = x^2 + 1$, we have

$$N(x) := \frac{1}{2}\left(x + \frac{-1}{x}\right)$$

as our iteration function. Because $N\colon \mathbb{R} \to \mathbb{R}$, we expect applying Newton's method to guesses starting as real numbers to not be very helpful.

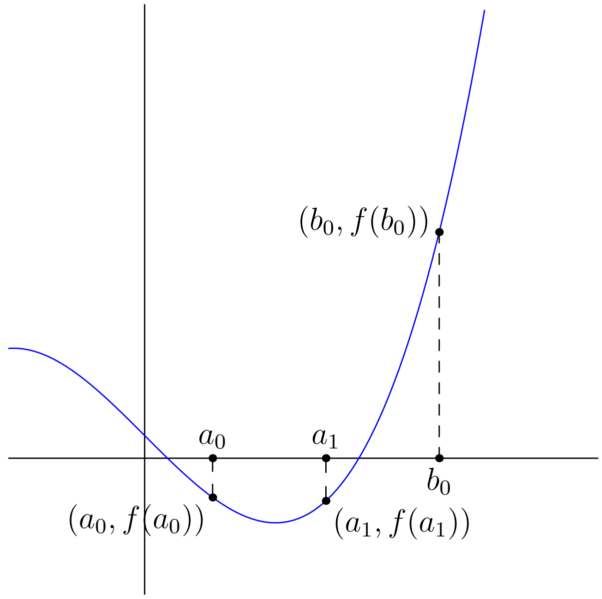
Remark 8. From the perspective of dynamical systems, $f(x) = x^2 + 1$ is actually very interesting: a random first guess will provide a sequence dense in $\mathbb{R}$. However, we can find periodic points as well, which is pretty cool.

1.2 Efficiency

To justify Newton's method's efficiency, let's try another method and compare. The idea for our second method is the Intermediate value theorem.

Theorem 9 (Intermediate value)**.** Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous. Then if $f(a) < 0$ and $f(b) > 0$, there is a root of f between a and b.

As such, we can imagine taking an interval $[a, b]$, evaluating $f\left(\frac{a+b}{2}\right)$ and dividing the interval in half depending on its sign. Here is the image.



This "Intermediate value theorem" method uses a single iteration to get a bit of information out of the root. However, Newton's method gets a quadratic improvement on the root!

Let's see this quadratic improvement. Without loss of generality, suppose $f(0) = 0$ is our root. Then, using a Taylor series, we have

$$N(x) = x^2 \cdot \frac{f''(0)}{f'(0)} + O\left(x^3\right),$$

so with x small we get $N(x) \sim x^2$ to be a quadratic improvement.

1.3 Random Guess

One issue with this method is that we need a good first guess. What if we just tried guessing? This idea is due to Cayley.

Example 10. Take $f(x) = x^2 - 1$. It turns out that, guessing $x_0 \in \mathbb{C}$, we have that $x_n \to 1$ if $\operatorname{Re} x_0 > 0$ and $x_n \to -1$ if $\operatorname{Re} x_0 < 0$.

Remark 11. In the case of $\operatorname{Re} x_0 = 0$, this is essentially analogous to trying to solve $f(x) = x^2 + 1$ over $\mathbb{R}$.

The behavior for quadratic polynomials is essentially isomorphic to this behavior for $f(x) = x^2 - 1$, essentially applying a suitable isometry and dilation.

So what happens with cubic polynomials? Well, the behavior becomes fractal-like and quite complicated, and it stays complicated as one adds more roots. However, we still can gain some control over this. We pick up the following definition.

Definition 12. Let $N: \mathbb{C} \to \mathbb{C}$ be Newton's method for a given polynomial f. The set of all points $z \in \mathbb{C}$ which approach a given root α is called the *basin* of α. The connected component of the basin of α which contains α is called the *immediate basin*.

Remark 13. All these basins are open: a small perturbed element of the basin will still converge to the same root.

There is the following result.

Theorem 14. The immediate basin of a given root α is simply connected in $\mathbb{C}$; i.e., it is homeomorphic to the disk $\overline{B(0,1)}$ of radius 1 in $\mathbb{C}$, and under this homeomorphism, N becomes the function $\hat{N}: D(0,1) \to D(0,1)$ given by

$$\hat{N}(x) = x^2.$$

Notably, $\hat{N}$ has two fixed points 0 and 1, which when pulled back to the map $N: \mathbb{C} \to \mathbb{C}$ corresponds to the immediate basin containing the fixed point α (as its root) as well as ∞. Namely, the immediate basin must contain arbitrary large values! We can say better than this.

Theorem 15. The immediate basin of a given root α has a somewhat large "wedge" of $\mathbb{C}$ as a subset.

So here is the idea: fix a polynomial f, and set some very large circle $\partial B(0, R)$. Then doing a bunch (in fact, $1.11d(\log d)^2$ is enough, where $d := \deg f$) of initial guesses evenly distributed around $\partial B(0, R)$ will be able to converge to the roots using Newton's method.

Remark 16. This is a miracle! Newton's method was intended to improve approximation of roots locally (close to the roots), but we managed to find all roots globally.

We close with a couple of further directions.

- How does this generalize to systems of polynomial equations?
- How does this compare to the secant method to approximate roots?
- Is there a story for higher-order approximations (e.g., Halley's method)?