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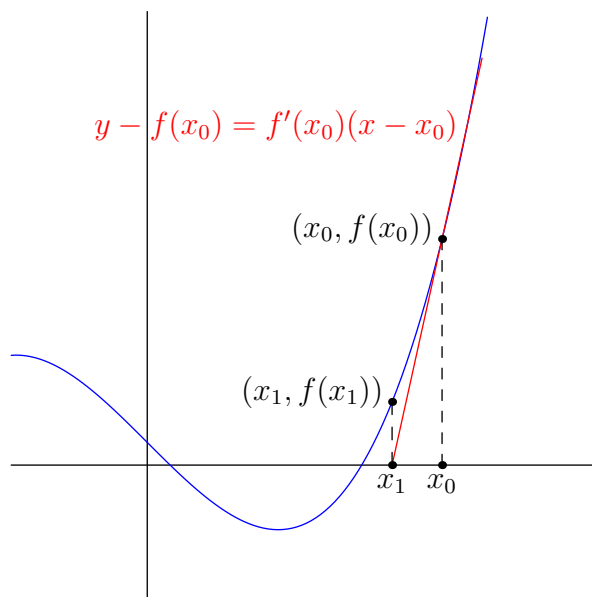
To be a little more technical, we begin with a guess x_0, approximate

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

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 $N : \mathbb{R} \to \mathbb{R}$, 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 $f : \mathbb{R} \to \mathbb{R}$ 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.
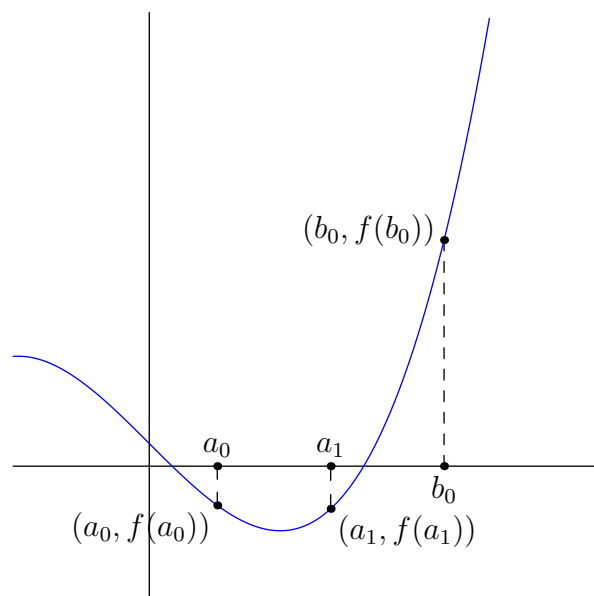
This "Intermediate value theorem" method uses a single iteration to get a bit of information out of the root. However, Newton's method gets a quadratic improvement on the root!

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)?

2 Quantum Symmetries, Fusion, Braids, and Categories — Julia Plavnik

Professor Plavnik was impressed with the talks yesterday and so made this talk a little harder, just for fun.

2.1 Symmetries

Let's start with what a symmetry is.

- Colloquially, we think of symmetry as having good proportions, etc.

- Art, architecture, and nature are good sources of symmetry.

To a mathematician, we encode symmetry inside invariants.

> **Definition 17.** We will say that an object X has a symmetry f if it has a good bijection $f \colon X \to X$.

> **Example 18.** The triangle has reflections and rotations as its symmetries. Notice that the composition of two symmetries is also a symmetry, and the order of this composition matters.

Observe that these symmetries are "invertible," which verifies automatically (i.e., categorically) that they are bijective.

For some notation, we set $\mathrm{Sym}(X)$ to be the group of symmetries of an object X. We are not being terribly formal with this,[1] but we can kind of feel what the definition should mean.

> **Example 19.** The group of symmetries of an equilateral triangle T is $\mathrm{Sym}(T) = S_3$.

> **Example 20.** The group of symmetries of a square S is $\mathrm{Sym}(S) = D_4$, where D_4 is the dihedral group with four elements.

We've been talking about groups a lot, so let's give it a definition.

> **Definition 21** (Group). A *group* is a set G given a binary operation $* \colon G \to G$ satisfying the following axioms.
>
> - Identity: there is $e \in G$ with $e * g = g * e = g$ for all $g \in G$.
>
> - Inverse: for all $g \in G$ there exists g^{-1} such that $g * g^{-1} = g^{-1} * g = e$.

> **Example 22.** The integers $\mathbb{Z}$ form a group under addition.

> **Example 23.** We had groups S_3 and D_4 from before.

> **Example 24.** The group S_n can be generated by transpositions, with certain relations.

2.2 Braids

Now let's talk about braids. The braid group B_n is generated by $\sigma_1, \ldots, \sigma_{n-1}$ with the relations

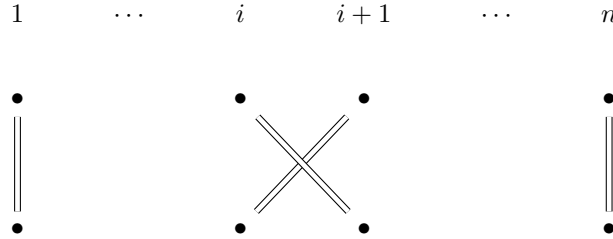$$\sigma_i \sigma_j = \sigma_j \sigma_i \text{ if } |i - j| < 2$$

and

$$\sigma_i \sigma_{i-1} \sigma_i = \sigma_{i-1} \sigma_i \sigma_{i-1}.$$

---

[1] Formally, we need to discuss group actions on a (structured) set X.

> **Remark 25.** If we add the relations $\sigma_i^2 = \mathrm{id}$ everywhere, then we get S_n, so S_n is a natural quotient of B_n.

Formally, we should think about σ_i as swapping strands i and $i + 1$ in a braid, as follows.

Note that there is a "top" and "bottom" for our strands. In particular, i is going on top of $i + 1$ in the above diagram. Multiplication of σ_i and σ_j consists of putting one braid on top of the other.
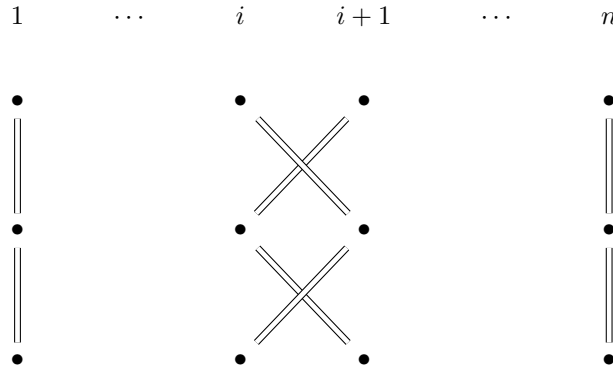
Here, we have placed $\sigma_i \sigma_i^{-1}$, which we can see we can "undo" to get back to the identity. However, σ_i^2 does a cross-over twice and so cannot go back to the identity.

2.3   Representation Theory

We understand linear algebra, so let's try to understand groups by their actions on vector spaces.

> **Definition 26** (Representation)**.** Given a group G, an n-dimensional k-*representation* is a group homomorphism
> $$\rho\colon G \to \mathrm{GL}_n(k).$$

More concretely, we are asking for $\rho(gh) = \rho(g)\rho(h)$ for all $g, h \in G$.

> **Example 27.** Take B_3, generated by $\{\sigma_1, \sigma_2\}$ with the relation $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$. As such, a representation $\rho\colon B_3 \to \mathrm{GL}_n(k)$ amounts to finding matrices $A, B \in \mathrm{GL}_n(k)$ with
> $$ABA = BAB.$$
> We can find a lot of these.

With this in mind, we note the following classification result for representations of B_3.

Theorem 28. For $n \leq 5$, irreducible n-dimensional representations of B_3 is given by using an upper triangle matrix A and a lower triangle matrix B, where the eigenvalues are in opposite orders: namely, there exists λ_1 and λ_2 with

$$A \sim \begin{bmatrix} \lambda_1 & \text{garbage} \\ 0 & \lambda_2 \end{bmatrix} \quad \text{and} \quad B \sim \begin{bmatrix} \lambda_2 & 0 \\ \text{garbage} & \lambda_1 \end{bmatrix}.$$

Remark 29. For $n > 5$, things get more complicated.

There are current research directions in taking $n > 5$ or expanding B_3. Here are some specific questions.

- What happens to representations of B_k for $k > 3$?

- We can actually describe certain representations of B_4 coming from representations of B_3. What is their structure?

- One can generalize braids beyond three dimensions. What do their representations look like? More specifically, there is a loop braid group whose representations we might want to understand.

It turns out that other areas of mathematics and physicists want these representations for reasons of their own.

2.4 Braided Fusion Categories

We now turn to studying "quantum symmetries," which means that we want to understand symmetries in quantum mechanics. Ultimately, these describe symmetries of space-time itself.
   However, groups are sometimes not the correct definition for our symmetries.

Idea 30. Groups encode symmetries. Tensor categories encode quantum symmetries.

Remark 31. The field of quantum symmetries, though algebraically motivated, has ties to lots of different mathematics, such as vertex algebras and analysis.

In particular, we will want quantum groups and Hopf algebras.
   For our definition, we will want to define categories. Please, don't be scared—Professor Plavnik really likes them.

Definition 32 (Category)**.** A *category* $\mathcal{C}$ is a collection of objects $\operatorname{Ob}\mathcal{C}$ and morphisms $\operatorname{Mor}(-, -)$ between them, satisfying certain coherence laws which we will not state here.

- For each $A \in \operatorname{Ob}\mathcal{C}$, there is a morphism $\operatorname{id}_A \in \operatorname{Mor}(A, A)$.

- Given objects A, B, and C, we can compose $f \colon A \to B$ and $g \colon B \to C$ to $(g \circ f) \colon A \to C$.

- Given objects A and B with $f \colon A \to B$, we have

$$f = f \circ \operatorname{id}_A = \operatorname{id}_B \circ f.$$

- Lastly, given objects A, B, C, and D with arrows $f \colon A \to B$ and $g \colon B \to C$ and $h \colon D \to A$, we have

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Example 33. Familiar categories include Set (of sets), Top (of topological spaces), of Grp (of groups), and Ab (of abelian groups).

The idea behind category theory is to focus on the arrows between objects in a category instead of the objects themselves. As such, we tend to understand objects up to (perhaps natural) isomorphism instead of equality. Let's try to build towards "braided fusion" categories. Here's the main example.

Example 34. The category FdVec_k of finite-dimensional k-vector spaces have the following nice properties.

- There is a 0 vector space.

- Given "objects" V and W, we can build $V \oplus W$.

- Given "objects" V and W, the set of morphisms $\mathrm{Mor}(V, W)$ is in fact a finite-dimensional k-vector space.

- There is a "smallest" object k, which is the 1-dimensional vector space. We call it "simple."

- Namely, any object V is isomorphic to $k^{\oplus n}$ for some nonnegative integer n.

- There is a product operation $\otimes$ taking V and W to $V \otimes W$.

- There is a dual vector space $V^* = \mathrm{Mor}(V, k)$ equipped with maps

$$\mathrm{ev}_V : V^* \otimes V \to k \qquad \text{and} \qquad \mathrm{coev}_V : V \otimes V^* \to k.$$

  Evaluation is the canonical map; we won't describe coevaluation in detail.

- There is a natural isomorphism $\sigma : V \otimes W \to W \otimes V$.

- There are some natural associativity laws as well, which satisfy some commuting pentagon which we won't write out here.

And here is our definition.

Definition 35 (Braided fusion category)**.** A *braided fusion category* is essentially a generalization of this example: take all the axioms and write down their natural generalizations, as stated above. We're not going to write these out because it would just be copying and pasting.

The main point is that we have the following example.

Example 36. The category of finite-dimensional k-representations (of a group G such that $(\mathrm{char}\, k) \nmid \#G$) forms a fusion braided category.

Remark 37. There are some connections between the various axioms of the braided category and some particle interactions in quantum physics. In particular, one can build a fairly concrete dictionary between the two; for example, quantum computing can be seen as applying certain "braid-like" operations.

Remark 38. Given a braided fusion category $\mathcal{C}$ gives rise to a natural representation of B_n: take $X \in \mathcal{C}$ to $\mathrm{End}_{\mathcal{C}}(X^{\otimes n})$ by

$$\sigma_i \mapsto \mathrm{id}_X^{\otimes(i-1)} \otimes \sigma \otimes \mathrm{id}_X^{n-i-3},$$

where $\sigma : A \otimes B \to B \otimes A$ is the swapping isomorphism.

Here are some the open directions about braided fusion categories.

- We might want to classify braided fusion categories, in some suitable sense.

- There are some questions about an information-theoretic approach to these categories.

Here is some progress.

Definition 39 (Rank)**.** The *rank* of a braided fusion category is the number of simple objects.

We can classify braided fusion categories up to rank 3. There is some partial progress up to higher dimensions. Classifying by dimension gives other progress. On the other side of classification, we have been able to build some new braided fusion categories from old; for example, we have had motivations from cohomology, from physics, etc.