An Introduction to Formal Real Analysis, Rutgers University, Fall 2025, Math 311H

Lecture 2: Newton's Computation of π

Prof. Alex Kontorovich

This text is automatically generated by LLM from "Real Analysis, The Game", Lecture 2

1 The Mathematical Revolution of 1666

SIMPLICIO: I heard that Newton had a really cool way of calculating π. Can you tell me about it?

SOCRATES: Certainly. It begins around 1665-1666, when Newton was turning 23 years old. Anything significant about that year?

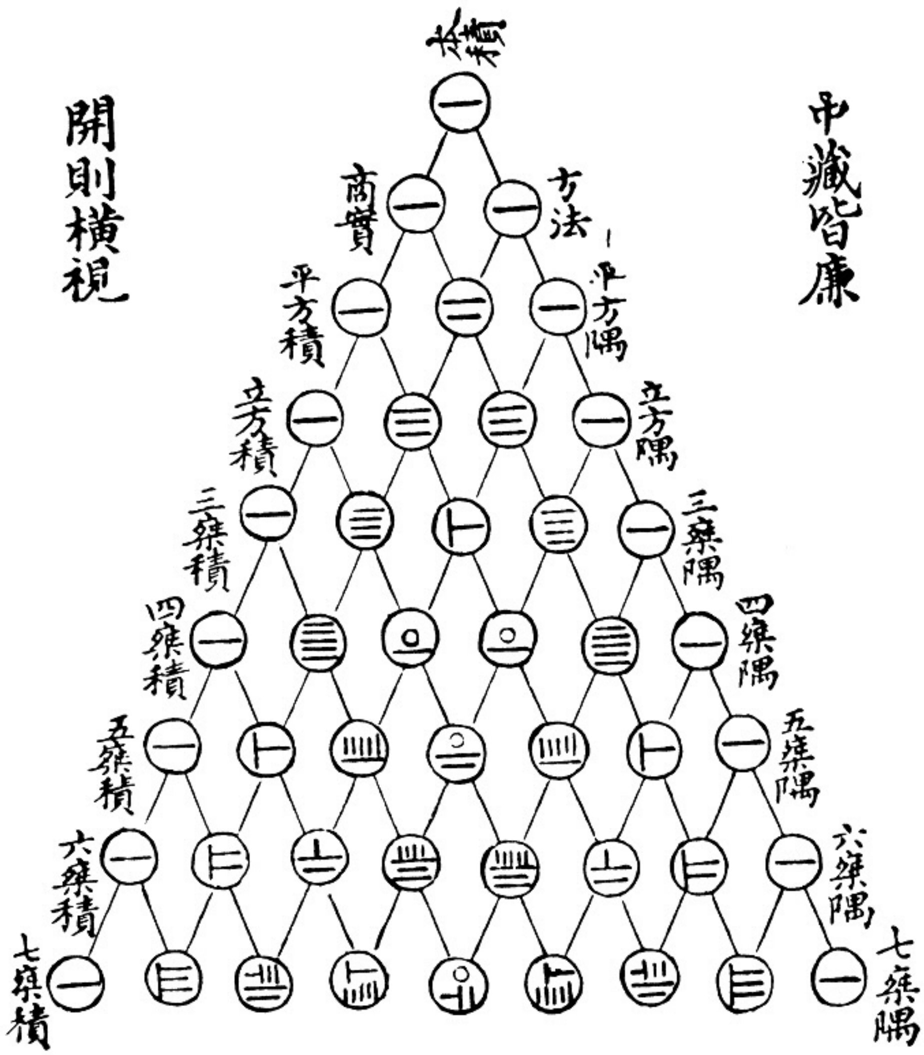
SIMPLICIO: Isn't that Newton's "annus mirabilis", year of miracles? If I recall correctly, he was forced to leave Cambridge due to an outbreak of the Great Plague, and made his most groundbreaking discoveries (calculus, optics, gravitation, etc) while quarantining in isolation at his family home in Woolsthorpe.

SOCRATES: Exactly right. And one of the first things he discovered in that year was a new version of the Binomial Theorem. Tell me, please, what can one say about $(1 + x)^n$?

SIMPLICIO: Sure thing, if you multiply $(1 + x)^n$ all out, you get

$$\binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$$

Here $\binom{n}{k}$ is the "binomial coefficient", the number of ways of choosing k things from a bag of n things. Explicitly, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. These are just the numbers in Pascal's Triangle, and you can easily read off the n-th row.



SOCRATES: Excellent! And do you know sigma notation?

SIMPLICIO: I think so. I could've written that same thing as $\sum_{k=0}^{n} \binom{n}{k} x^k$. In general, if you have some function $f : \mathbb{N} \to \mathbb{R}$, and you want express $f(a) + f(a+1) + \ldots + f(b)$, that is, the sum of $f(k)$ as k ranges from some integer a up to some other integer b, you can write it as $\sum_{k=a}^{b} f(k)$.

SOCRATES: Very good. So we have $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$. Now, would you allow me to write this as a sum going all the way out to infinity?

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k$$

SIMPLICIO: Hmm. Ok, I think I see what you're getting at: Pascal's Triangle has implied zeros everywhere outside, so $\binom{n}{k}$ is just zero once $k > n$. So you've written it as an infinite sum, even though it secretly terminates after finitely many terms. But what purpose does extending it serve?

SOCRATES: Well, let me ask you this: can you think of any way of making sense of this formula when $n = -1$?

SIMPLICIO: Huh? You can't use binomial coefficients. How do you choose 3 things from -1 things, that makes no sense!

SOCRATES: Ok, sure, but so many great discoveries in mathematics occur when you realize a way to **break the rules**, and follow some pattern **past** its intended limit (no pun intended). Put yourself in Newton's shoes, if you can; what might a genius like him come up with?

SIMPLICIO: Well, we do have this other formula for binomial coefficients, not in terms of combinatorics, but just as factorials, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. That still doesn't help because what the heck is $(-1)!$ supposed to be?! Oh, but wait! We can also write this as:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

And in this way, we "bypass" the issue of dealing with $(-1)!$, and just go straight to "normal" numbers.

SOCRATES: Can you please write this using product notation?

SIMPLICIO: Sure, it's just like summation notation but with a Π:

$$\binom{n}{k} = \frac{1}{k!} \prod_{\ell=1}^{k} (n-\ell)$$

SOCRATES: Whoops, are you sure about those bounds in the product?

SIMPLICIO: Argh! It's so easy to make a silly mistake. After writing down the formula, I should have checked that I got the right start and end values; the counter ℓ should go from 0 to $k-1$, not from 1 to k. Is this better?

$$\binom{n}{k} = \frac{1}{k!}\prod_{\ell=0}^{k-1}(n-\ell)$$

SOCRATES: Perfect. Go on.

SIMPLICIO: Ok, so if we agree to follow this pattern, then we get:

- $\binom{-1}{0} = 1$, which makes sense because any row of Pascal's triangle starts with at $1 = \binom{c}{0}$; then
- $\binom{-1}{1} = (-1)/1! = -1$, which also makes sense because the next term in the "cth row" of Pascal's triangle is always $\binom{c}{1} = c$; then we get:
- $\binom{-1}{2} = (-1)(-2)/2! = +1$,
- $\binom{-1}{3} = (-1)(-2)(-3)/3! = -1$.

Ok so I see the pattern: it just alternates between $+1$ and -1, so the series goes:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - x^5 + \cdots$$

SOCRATES: Very good! But does this formula make any actual *sense*? How might you test it?

SIMPLICIO: Ok, if that series is supposed to "represent" $1/(1+x)$, then if I multiply the whole thing by $(1+x)$, I should just get 1. Let's try it:

$$(1+x)(1 - x + x^2 - x^3 + x^4 - x^5 + \cdots) = ?$$

I'll first multiply everything by 1, then by x, and add them all up.

$$(1 - x + x^2 - x^3 + x^4 - x^5 + \cdots) + (x - x^2 + x^3 - x^4 + x^5 \cdots)$$

Ok, so if I rearrange terms, then everything cancels out, and only the leading 1 remains. Great!

SOCRATES: Interesting. And are you "allowed" to rearrange terms like that?

SIMPLICIO: Well... why not?

SOCRATES: Ok, nevermind that for now, you seem to be satisfied that it makes sense to say that the series $1 - x + x^2 - x^3 + x^4 - x^5 + \cdots$ "converges" (again, whatever that means) to $(1 + x)^{-1}$.

SIMPLICIO: Come to think of it, I knew this already; it's just the geometric series! I know that

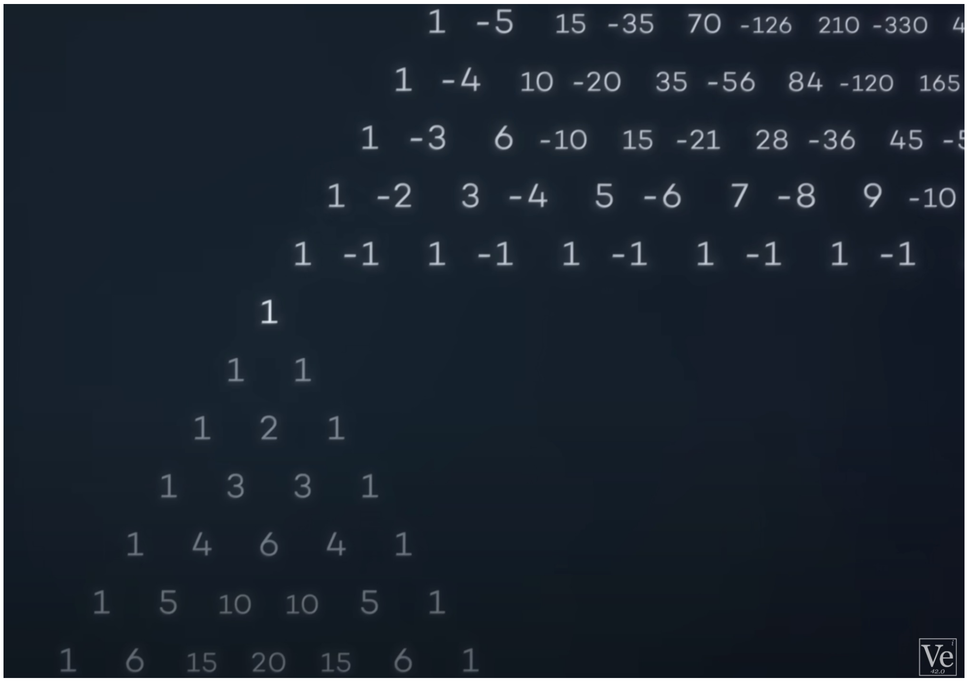
$$1 + \lambda + \lambda^2 + \lambda^3 + \cdots$$

adds up to $1/(1 - \lambda)$, and the series we have just replaces λ with $-x$.

SOCRATES: Yes, very good. And where might this "belong" in Pascal's triangle?

SIMPLICIO: Holy cow! Did we just discover an extension of the triangle, going "up"?!

SOCRATES: Indeed, and we can in fact continue this pattern for $n = -2, -3, -4$, and so on. I'll let you work it out yourself, but we actually get a *whole other* Pascal's triangle (with some negative signs) *above* the standard one!



See how it still follows the usual rule, that the two numbers above and to the left or right add to the value just below them?

But let's try something even more exotic. Can you make the Binomial Theorem work when $n = 1/2$?

SIMPLICIO: Whoa, $n = 1/2$? That's... really pushing it! But let me try using the same formula. So $\binom{1/2}{k} = \frac{1}{k!}\prod_{\ell=0}^{k-1}(1/2 - \ell)$. Let me work out the first few terms:

- $\binom{1/2}{0} = 1$ (as always)
- $\binom{1/2}{1} = (1/2)/1! = 1/2$ (again, matches the pattern we already knew)
- $\binom{1/2}{2} = (1/2)(-1/2)/2! = -1/8$
- $\binom{1/2}{3} = (1/2)(-1/2)(-3/2)/3! = 1/16$
- $\binom{1/2}{4} = \frac{(1/2)(-1/2)(-3/2)(-5/2)}{4!} = -\frac{5}{128}$

So $(1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots$

But wait – this is supposed to be $\sqrt{1+x}$?!

SOCRATES: Again, go into Newton's thinking: how might he go about justifying whether this formula makes any sense?

SIMPLICIO: Oh, ok, I think I see! If we square the formula and multiply everything out, I guess we're supposed to get $1+x$ – that would justify calling the series $\sqrt{1+x}$.

SOCRATES: Go for it!

SIMPLICIO: Ok, so I want

$$\left(1+\frac{1}{2}x-\frac{1}{8}x^2+\frac{1}{16}x^3-\frac{5}{128}x^4+\cdots\right)^2$$

That means squaring every term, and also adding twice every product of distinct pairs of terms.

SOCRATES: Right. Can you think of a good way of keeping track of everything?

SIMPLICIO: Oh, I know! Let's group things by the power of x involved. The first one is easy: $x^0 = 1$, which you only get from squaring the first term. So that coefficient is 1.

For the coefficient of x^1, I can't square anything involving x's, so I can only multiply the x term by the constant term, and of course double it. That's just $2 \times 1 \times \frac{1}{2}x = x$. So the coefficient of x is 1.

For x^2, I get two contributions from constant times quadratic: $2 \times 1 \times \left(-\frac{1}{8}x^2\right) = -\frac{1}{4}x^2$ and also from the square of the linear term: $\left(\frac{1}{2}x\right)^2 = \frac{1}{4}x^2$. So the total coefficient is $-\frac{1}{4}+\frac{1}{4}=0$.

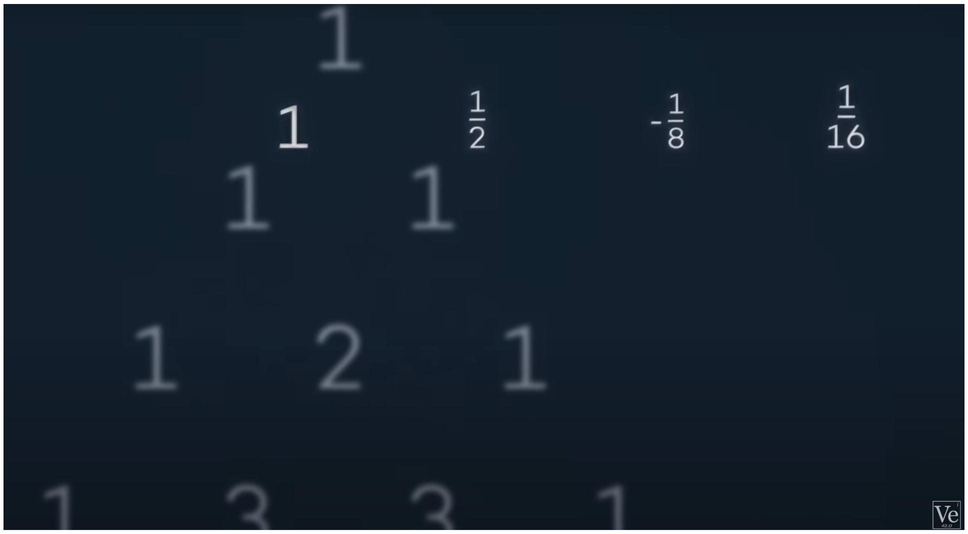
Let's try a few more. To get x^3, I need:

- $2 \times 1 \times \frac{1}{16}x^3 = \frac{1}{8}x^3$ (constant times the x^3 term)
- $2 \times \frac{1}{2}x \times \left(-\frac{1}{8}x^2\right) = -\frac{1}{8}x^3$ (the x term times the x^2 term)

So the total coefficient of x^3 is $\frac{1}{8}-\frac{1}{8}=0$.

This is amazing! It really seems like all the higher-order terms are canceling out perfectly. I bet that will keep happening, and we'll just get the square to come out to exactly $1+x$; the formula really works!

So wait, now we get a whole other row in Pascal's triangle, *between* rows 0 and 1?!



SOCRATES: Beautiful, isn't it!

SIMPLICIO: Wait, this is all much simpler than I'm making it. Isn't this just the same thing as the Taylor expansion about $x = 0$ of the function $f(x) = \sqrt{1+x}$? I already know how to do this from Calculus.

SOCRATES: Yes, very good; but Brook Taylor (of Taylor series) did not prove his general theorem until 1715, a few decades after Newton's computation of π.

SOCRATES: Now, suppose you wanted to compute something like $\sqrt{3}$ – can you think of a way of doing it using this formula?

SIMPLICIO: Hmm the function is $\sqrt{1+x}$, so I guess I want to set $x = 2$. Then I get:

$$\sqrt{1+2} = 1 + \frac{1}{2}(2) - \frac{1}{8}(2)^2 + \frac{1}{16}(2)^3 - \frac{5}{128}(2)^4 + \cdots$$

Adding up these five terms comes out to $11/8 = 1.375$, not so close to $\sqrt{3} \approx 1.73$. And the individual terms are not so small, for instance, the last one, $\frac{5}{128}(2)^4 = 5/8 = 0.625$.

SOCRATES: Well, sure, if you set x to be large, like $x > 1$, then the powers of x are also larger and larger (and exponentially so!)... Can you think of something else you could do?

SIMPLICIO: Ah, I think I see! I know that 3 is near 4, which is a perfect square. So what if we write

$$\sqrt{3} = \sqrt{4-1} = \sqrt{4(1-\frac{1}{4})} = 2\sqrt{1-\frac{1}{4}}$$

So now if I apply our formula with $x = -1/4$ (which is less than one!), I guess I'll get:

$$\sqrt{3} \approx 2\left(1 + \frac{1}{2}(-1/4) - \frac{1}{8}(-1/4)^2 + \frac{1}{16}(-1/4)^3 - \frac{5}{128}(-1/4)^4 + \cdots\right)$$

Taking just these five terms, the fraction on the right comes out to $28379/16384 \approx 1.73212$, which is impressively close to $\sqrt{3} \approx 1.7320508$. We got 3 decimal places of accuracy, nice!

SOCRATES: Great! Now you see the power of Newton's Binomial Theorem. Ok, so then let's return all the way back to your original question, about Newton's estimate for π.

SIMPLICIO: Hmmm, π is the ratio of circumference to diameter in a circle. So where am I supposed to find a length.

SOCRATES: Ah, but what did we learn from Archimedes?

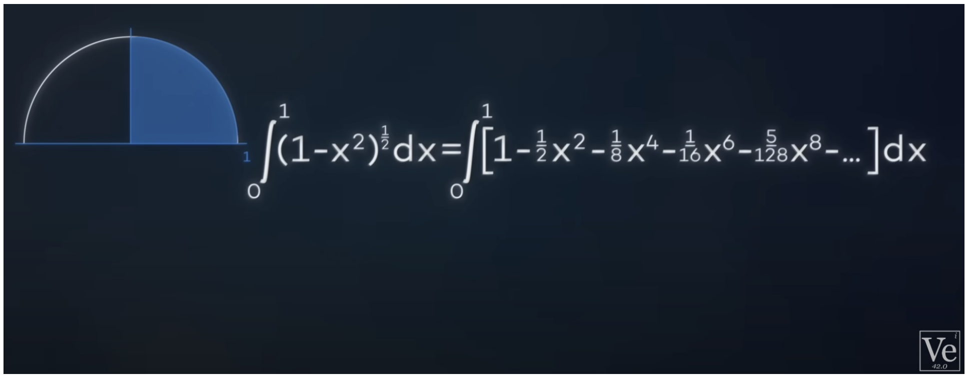
SIMPLICIO: Oh, that π is also an *area*, not just a length. It's the area of a unit circle πr^2 where $r = 1$.

SOCRATES: Beautiful. And could you find a circle's area lurking somewhere?

SIMPLICIO: I think I see it! Thanks to Descartes, and "Cartesian" coordinates, we can express the circle as the graph of $x^2 + y^2 = 1$, or to make it a function, $y = \sqrt{1-x^2}$. So we just have to replace x in our series with $-x^2$.

$$\begin{aligned}\sqrt{1-x^2} &= 1 + \frac{1}{2}(-x^2) - \frac{1}{8}(-x^2)^2 + \frac{1}{16}(-x^2)^3 - \frac{5}{128}(-x^2)^4 + \cdots \\ &= 1 - \frac{x^2}{2} - \frac{x^4}{8} - \frac{x^6}{16} - \frac{5x^8}{128} - \cdots\end{aligned}$$

And the area under the curve $y = \sqrt{1-x^2}$ from $x = 0$ to $x = 1$ is a quarter circle.



SOCRATES: Luckily, Newton had just invented calculus! So how else could he compute the area under this curve?

SIMPLICIO: With an integral! So:

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-x^2}\, dx = \int_0^1 \left(1 - \frac{x^2}{2} - \frac{x^4}{8} - \frac{x^6}{16} - \frac{5x^8}{128} - \cdots\right) dx$$

I'll just integrate term by term,...

SOCRATES: Whoa, hang on! Why are you allowed... You know what, nevermind, sorry. Just go ahead.

SIMPLICIO: Ok, weirdo. Anyway. So integrating term by term, I get:

$$\begin{aligned}\frac{\pi}{4} &= \left[x - \frac{x^3}{6} - \frac{x^5}{40} - \frac{x^7}{112} - \frac{5x^9}{1152} - \cdots\right]_0^1 \\ &= 1 - \frac{1}{6} - \frac{1}{40} - \frac{1}{112} - \frac{5}{1152} - \cdots\end{aligned}$$

Wow! So Newton got an infinite series for π! If I evaluate just these five terms, and cross multiply by the factor of 4, I get the fraction $32057/10080 \approx 3.180$, not bad!

SOCRATES: Not bad indeed. You know, Simplicio, many math papers have roughly zero new ideas; they're just doing something nobody bothered to do before in a slightly newer context. A really good math paper can have one or two genuinely new ideas. Newton is already on new idea number five, and he's still not done!

SIMPLICIO: Ok, so what's new idea number six?

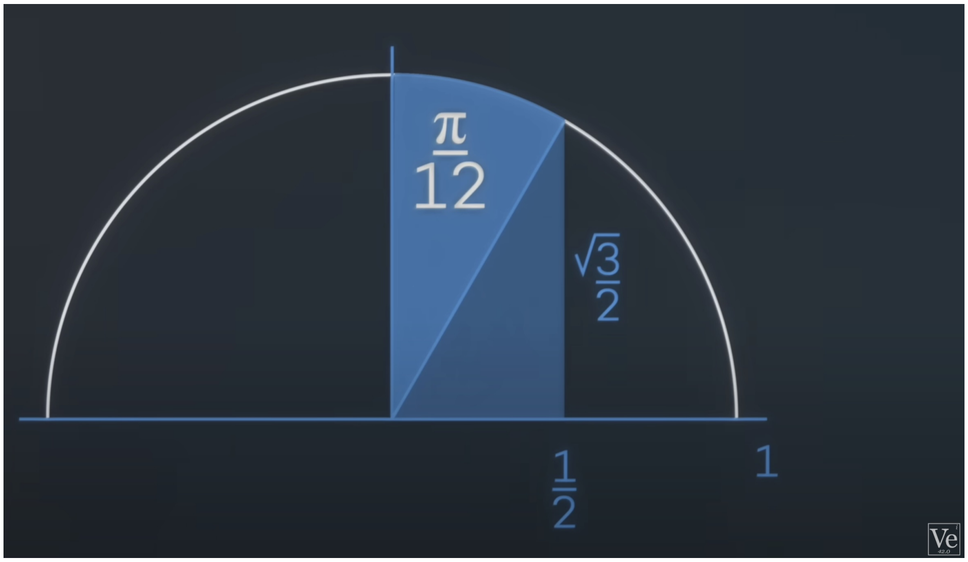
SOCRATES: Well, remember how you integrated all the way up to $x = 1$?

In your series,

$$\frac{\pi}{4} = \left[x - \frac{x^3}{6} - \frac{x^5}{40} - \frac{x^7}{112} - \frac{5x^9}{1152} - \cdots\right]_0^1$$

you have all these high powers of x but they're being "wasted" because you're setting x to 1. What if instead you only integrated up to, say, $x = 1/2$?

SIMPLICIO: Ooh, cool! Then the series will converge much more rapidly. But wait, that changes the geometry. Instead of a quarter-circle, we now have... a 30 degree sector, which has area $\pi/12$, plus a 30-60-90 triangle – ah that must be why you suggested $x = 1/2$ – with area $\frac{1}{2} \times \frac{1}{2} \times \frac{\sqrt{3}}{2}$.



Good thing we already know how to quickly estimate $\sqrt{3}$ to high accuracy! (Ah, that's the trade-off: we could set x even smaller, for faster convergence, but then we'll need to deal with ever more complicated geometric evaluations; so $x = 1/2$ is a "sweet spot".) So now:

$$\frac{\pi}{12} + \frac{\sqrt{3}}{8} = \left[x - \frac{x^3}{6} - \frac{x^5}{40} - \frac{x^7}{112} - \frac{5x^9}{1152} - \cdots\right]_0^{1/2}$$

$$= \frac{1}{2} - \frac{1}{6}\left(\frac{1}{2}\right)^3 - \frac{1}{40}\left(\frac{1}{2}\right)^5 - \frac{1}{112}\left(\frac{1}{2}\right)^7 - \frac{5}{1152}\left(\frac{1}{2}\right)^9 - \cdots$$

Again evaluating just these five terms already gives the fraction $9874097/20643840$.

And now isolating π gives the estimate

$$\pi \approx 12 \times \left(9874097/20643840 - \frac{\sqrt{3}}{8} \right) \approx 3.14161,$$

which is off by two parts in $100,000$ from the true estimate $\pi \approx 3.14159$. All that with just five terms, amazing!
SOCRATES: Yes, Newton was very impressive indeed. Here's a nice YouTube video by Veritasium that discusses this whole saga:

`https://youtu.be/gMlf1ELvRzc`

In fact, a series for π similar to this one was discovered two centuries earlier, by the Indian mathematician Madhava of Sangamagrama. And it would take two more centuries until mathematicians figured out how to rigorously justify Newton's work. To do so, they had to figure out:

- What it meant for a sequence of real numbers $a_0, a_1, a_2, \dots$ to converge?
- What it meant for a series (that is, sequence of partial sums) $a_0 + a_1 + a_2 + \cdots$ to converge, and could we sum these numbers in any order we like,
- What it meant for a series involving a variable, like a power series $a_0 + a_1 x + a_2 x^2 + \cdots$ to converge, and if so, what kind of function it converged to,
- When can we interchange limits with integrals, like integrating term by term, $\int (a_0 + a_1 x + a_2 x^2 +) dx \cdots \stackrel{?}{=} \int a_0 \, dx + \int a_1 x \, dx + \int a_2 x^2 \, dx + \cdots$,

Etc, etc. We have a lot of work to do!
SIMPLICIO: Ok, ok; you've convinced me! On with some actual Real Analysis please.

2 The Main Definition

Our first step to making Newton's argument rigorous is to spell out *exactly* what we mean by a sequence a_n converging. It will take a little work to build up to the definition, and more importantly, *why* that might seem like a reasonable definition to have.

But first: for some reason (likely Euler is to blame), mathematics has *two* completely different conventions for how to write functions. For general functions $f : \mathbb{R} \to \mathbb{R}$, we write $f(x)$, with parentheses. But when we work with sequences, a_n, meaning, $a_0, a_1, a_2, \ldots$, we bizarrely switch instead to subscripts. Why? Historical accident.

A sequence is nothing but a function whose "domain" (that is, the set of inputs to the function) is the natural numbers; so we will break with tradition and unify the two conventions, henceforth writing $a : \mathbb{N} \to \mathbb{R}$ for sequences of real numbers, $a(0), a(1), a(2), \ldots$.

Now, the definition that mathematicians eventually came up with for what it means for a sequence to converge, was so intricate (at least at first sight) that it had to be invented *twice*!

The eventual formulation crystallized through the work of Karl Weierstrass in the 1860s, who transformed analysis from an intuitive art into a rigorous science. However, the seeds of this idea appeared much earlier in the work of Bernard Bolzano. In the 1810s and 1820s, Bolzano was developing remarkably modern ideas about continuity and limits, but he was too far ahead of his time for the mathematical community to accept these abstract concepts. Only by Weierstrass's time – a half-century later – did these ideas catch on.

Without further ado, here it is:

Definition 2.1 (Sequence Convergence)**.** Given a sequence $a : \mathbb{N} \to \mathbb{R}$ and a real number $L : \mathbb{R}$, we write $\lim a = L$ and say that the sequence a **converges** to L, if: for every $\varepsilon > 0$, there exists $N : \mathbb{N}$ such that, for all $n \geq N$, we have $|a(n) - L| < \varepsilon$.

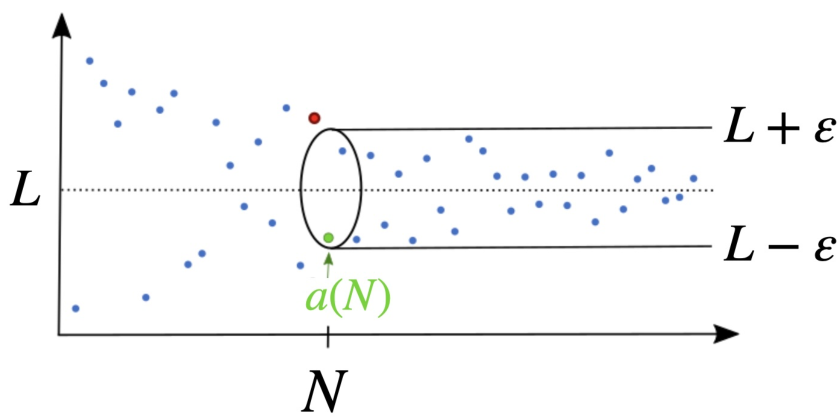
This definition is probably not the first, or second, or tenth thing you might've come up with. But over time, I hope you'll come to see that it embodies a beautiful negotiation between precision and effort.

I like to think of it as a conversation between an Engineer and a Machinist. The Engineer arrives with specifications: "We're going to make this widget, and I need its length to be 1 foot, with an error tolerance of 1/100 of an inch".

The Machinist replies: "Sure, I can do that, but I'll have to run my special equipment for at least 10 hours to guarantee that tolerance." The Engineer replies: "I'm sorry, I misspoke, can we change the tolerance to 1/1000 of an inch?" The Machinist replies: "Oof, yeah we can do it, but it'll cost ya. I'll need at least 40 hours of operation, but after that, I'll guarantee it."

As long as this conversation can continue regardless of *whatever* tolerance $\varepsilon > 0$ the Engineer requires, with the Machinist always being able to reply with a finite minimum number of hours N, after which the tolerance will be achieved, we can say that the equipment **converges**.

Now let's read Weierstrass's (or is it Bolzano's?) definition again. We have some process that at time n returns a reading $a(n)$ (think: widget length). Our ultimate goal is to make the length L. If for any tolerance $\varepsilon > 0$, no matter how small, there will always exist some minimum time N, so that, for any future time, $n \geq N$, we are guaranteed to be within that tolerance, $|a(n) - L| < \varepsilon$, that's exactly the condition under which we'll say that the sequence $a(n)$ **converges** to L.



What makes this definition so powerful is its universality. The Machinist is essentially promising: "Give me *any* tolerance requirement, no matter how stringent, and I can meet it – though I might need more resources (larger N) for tighter specifications."

Notice something else about the definition: It makes no mention of something happening "eventually", or "at infinity" or any other wishy-washy squirm words. We have traded the ambiguity of speaking about infinity for the precision of existential and universal quantifiers. No more hand-waving about what happens "as n gets large" - instead, we have a concrete challenge: given *any* tolerance ε, can you find a specific threshold N? *That* idea was the key breakthrough that allowed Calculus to enter the realm of rigorous mathematics.

In Lean, the definition is written like so:

```
def SeqLim (a : ℕ → ℝ) (L : ℝ) : Prop :=
  ∀ ε > 0, ∃ N : ℕ, ∀ n ≥ N, |a n - L| < ε
```

This syntax should be familiar from the `have` tactic you already know and love. The special symbol `def` (instead of `have`) means that we're about to define something, and `SeqLim` is its name (for sequence limit, of course; but we could have called it whatever we want). Then our assumptions are a sequence `a : ℕ→ℝ` and some real number `L : ℝ`. Then after the colon : goes our output, which in this case is `Prop`, that is, a statement (proposition) that can be true or false. So `SeqLim` is really a function that takes a sequence and hypothetical limiting value, and returns true or false based on whether the condition is satisfied. Then comes a colon-equals `:=`, after which the exact condition to be tested is specified. And the condition is what we already said, for all epsilon, yadda yadda. The big difference is that you can write `have` inside a proof, but you can't write `def` inside a proof; `def` is reserved for making global definitions that can be referenced forever once they're introduced.

You may find useful a new tactic called `change`. It allows you to replace a goal (or hypothesis) by something that is definitionally equal to it. In our example here, You will see the goal as `SeqLim a L`. What are you supposed to do with that, how can you make progress? Well, if you remember how `SeqLim` is defined, then you can replace the goal with the definition, by writing

`change ∀ε > 0, ∃N : ℕ, ∀n ≥N, |a n - L| < ε`

Lean will then change the goal to its definition. Remember that ε, `N`, and `n` are all dummy variables here, so you can have some fun:

`change ∀Alice > 0, ∃Bob : ℕ, ∀blah ≥Bob, |a blah - L| < Alice`

And one last tactic you might also find useful is `norm_num` (for normalizing numerical values); it evaluates numerical expressions and proves equalities/inequalities involving concrete numbers. For example, if you're stuck with an `|0|` at some point, and you want to convert it to plain old `0`, try calling `norm_num`.

```
Statement ConstLim (a : ℕ → ℝ) (L : ℝ) (a_const : ∀ n, a
    n = L) : SeqLim a L := by
  change ∀ ε > 0, ∃ N : ℕ, ∀ n ≥ N, |a n - L| < ε
  intro ε hε
```

```
  use 1
  intro n hn
  specialize a_const n
  rewrite [a_const]
  ring_nf
  clear hn a_const n
  norm_num
  apply hε
```

You've just completed your first rigorous limit proof! Let's reflect on what you accomplished and the key insights from this foundational example.

What you just proved: You showed that if a sequence always outputs the same value `L`, then it converges to `L`. The Machinist's response to any tolerance demand `ε > 0` is beautifully simple: "I can meet that specification immediately with any production run length `N`, because I'm already producing exactly what you want!"

Key Insights from this proof:

1. **The `change` tactic**: You learned how to unfold a definition to see what you're really trying to prove. `SeqLim a L` became the concrete epsilon-N condition.

2. **The logical structure**: The proof followed the natural flow of the definition:

   - `intro ε hε` handled for every `ε > 0`
   - `use 1` provided the witness `N` (any number works!)
   - `intro n hn` handled `∀ n ≥N`
   - Then algebraic manipulation showed that `|a n - L| = |L - L| = |0|`
   - Then numerical normalization gave that `|0| = 0`, and `hε` finally proved that `|a n - L| < ε`.

The Beautiful Simplicity: This is the Machinist's dream scenario—no matter how demanding the engineer's tolerance requirements, the constant factory can satisfy them instantly. There's no trade-off between precision and effort because the output is already perfect!

You're building the foundation for all of calculus. Every limit, derivative, and integral ultimately rests on arguments like this one.

Let's step back from the formal Lean proof and understand what we just proved in plain English.

Theorem (in natural language): If a sequence has the same value for every term, then it converges to that constant value.

Proof: Suppose we have a sequence $a(n)$ where $a(n) = L$ for all n, and we want to show that this sequence converges to L.

By definition, we need to show that for any tolerance $\varepsilon > 0$, we can find a point N such that for all $n \geq N$, we have $|a(n) - L| < \varepsilon$.

This is almost trivially simple: since $a(n) = L$ for every n, we have:

$$|a(n) - L| = |L - L| = |0| = 0$$

Since $0 < \varepsilon$ for any positive ε, we can choose any N we want (we chose $N = 1$ in the proof, but $N = 0$ or $N = 1000$ would work equally well).

Therefore, for any $n \geq N$, we have $|a(n) - L| = 0 < \varepsilon$, which proves convergence. **QED**