

## MITOCW | R12. Karatsuba Multiplication, Newton's Method

---

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** Everything make sense? Are you guys good with Newton?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Yes, really? I'm done. I'm going home. Are you guys good with Karatsuba?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Kara-what?

[LAUGHTER]

OK, so who is happy with Karatsuba? Good. Who is happy with Newton? Wait. I need to ask the other ones, because not enough people-- who is not happy with Newton and wants to look over it? Who wants to look over Karatsuba? Oh, so you guy are paying attention. Good.

OK. All right. So Newton and Karatsuba. Let's start. So Karatsuba real fast, and then we'll spend more time having fun with Newton.

So suppose we don't know Karatsuba. We want to multiply two numbers. What's the method that we learned in elementary school?

**AUDIENCE:** You take the first digit of the first number, and then you multiply it by the second number. And then you take the second digit and multiply it by the second number, but then you add a 0. And so on and so forth.

**PROFESSOR:** So I'll take this digit and multiply it with this number and write the result down here, right? So whatever this thing happens to be, 1, 2, 3, 4 times 8, write the result. Then

I take 1, 2, 3, 4 times 7 and add the result here and the last digits are 0. 1, 2, 3, 4 times 6, two 0s. 1, 2, 3, 4 times 5, three 0s. Then I add these up and I know that I'm going to get the right result.

What's the running time for this?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** OK. So why is it  $n$  squared?

**AUDIENCE:** Because you go through every single thing for every digit below.

**PROFESSOR:** Yup. So I'll have these partial products here, and I have as many partial products as digits in the second number. So if I have  $n$  digits here, I have  $n$  partial products. Each partial product is this first number here multiplied by a digit.

So if this guy has  $n$  digits, then there are  $n$  digits everywhere here. And this thing will be  $2n$  insides actually, but we're doing asymptotic, so we don't really care about constant factors.

So in the end, you have to add up  $n$  numbers. All of them are order  $n$  in size. So  $n$  times order  $n$  squared. This is plain old multiplication.

Does anyone know why Karatsuba's better? Yes?

**AUDIENCE:** You do three multiplications or something like that?

**PROFESSOR:** OK. So, let's go for a simple answer.

**AUDIENCE:** It's better.

**PROFESSOR:** So the simpler answer is it's faster.

**AUDIENCE:** OK.

**PROFESSOR:**  $n$ ,  $n \log 2$ --

**AUDIENCE:** I said better, it's faster.

**PROFESSOR:** I mean, some algorithms have the same speed but are simpler. We like this one because it's faster. It's definitely not simpler.

So, it has this running time here.  $\log_2 3$  is a weird number. It's somewhere between 1.5 and 2. So not revolutionary, but it's better than this guy.

OK, why do we get to this? How do we get to this? Does anyone remember the recurrence for it? Yes?

**AUDIENCE:** It's 3 times 2--

**PROFESSOR:** You're not remembering it. You're cheating.

**AUDIENCE:** No, I'm not.

[LAUGHTER]

**PROFESSOR:** So, let's see why it looks like this. Karatsuba takes 2 numbers that have  $n$  digits, splits both of them up into two--  $n/2$ ,  $n/2$ ,  $n/2$ ,  $n/2$ -- and then is able to compute the product of these two numbers by only using three multiplications, and then some additions. Multiplications are expensive, additions are cheap.

So it does three multiplications that are of the form some number of  $n/2$  digits multiplied by some number of  $n/2$  digits. And exactly what the multiplications and what the additions are is, it's interesting to look over it. You can look over it, convince yourself that this is correct. Chances are you'll remember for the exam and forget afterwards.

Please remember this. This is the essence of Karatsuba. And remember Karatsuba, because if you google this, you get the rest. So the multiplications are in the notes. I don't think it makes sense to spend some time on them and convince you that, yeah, there are three multiplications that will get you that.

Instead, suppose I labeled these as  $a$ ,  $b$ ,  $c$ ,  $d$ . If I evaluate the multiplication as

using brute force, it's  $a \times n$  over  $2b$  plus  $c$ -- sorry,  $a$  plus-- no, this isn't good.  $10$  to the power  $n$  over  $2a$  plus  $b$  times  $10$ -- assuming the base is ten, if it's not, you replace it--  $n$  over  $2c$  plus  $d$ . And you get  $10$  to the  $n$  times, I think,  $ac$  plus  $10$  to the  $n$  over  $2$  times  $ad$  plus  $bc$  plus  $bd$ .

The problem with this is this is a multiplication, multiplication, multiplication, multiplication. So this uses four multiplications of size  $n$  over  $2$  to evaluate the answer one multiplication of size  $n$ . And if I have four multiplications, then the recursion becomes  $4t$  of  $n$  over  $2$  plus  $o$  of  $n$ . So this is  $t$  bad of  $n$ .

So if you do this using the recursion tree, you're going to see that-- yup, I hear whispers.  $n$  squared. So this isn't useful.

This will give you insight on how to do it, but if you try to implement it this way, you don't get any speed up. The insight in Karatsuba is that the guy figured out how to do this with three multiplications.

OK, makes sense? Is everyone happy with Karatsuba now? So that multiplying by  $10$  to the  $n$  over  $2$  is just shifting  $a$ ?

**PROFESSOR:** Yup. OK. Very good point. So this isn't an actual multiplication. If you have your number, say stored in a python list-- so if these are your digits, then in order to multiply by  $10$  to the  $n$ , you append  $10$  to the  $n$   $0$ s at the end.  $0, 0, 0, 0, 0$ . So, this is append of  $0$   $n$  times. So this is order  $n$ , so it counts as an addition, not as a multiplication.

So, these are free. These are expensive.  $1, 2, 3, 4$ . OK. Who's happy with Karatsuba now? Let me feel like I did something useful today.

**AUDIENCE:** I have a question.

**PROFESSOR:** Yes.

**AUDIENCE:** I didn't fully understand what [INAUDIBLE].

**PROFESSOR:** So, I have this number and I'm splitting it up into  $2$  halves. The way split up is I say

this is a number of size  $n$  over 2, and I have to shift it to the left. So  $n$  10 to the power of  $n$  over 2 shifts it left by  $n$  over two digits. And then I add it to  $b$ . So this in a sense is just saying that this is the left half, this is the right half.

**AUDIENCE:** This is actually that same number.

**PROFESSOR:** Yup. It's exactly the same number. OK, so who is happy with Karatsuba now?

**AUDIENCE:** Wait, but that's four multiplications though.

**PROFESSOR:** Exactly.

**AUDIENCE:** That's not three.

**PROFESSOR:** Exactly.

**AUDIENCE:** But we said there were three.

**PROFESSOR:** So, I was saying that this, if you're doing it the brute force way, you get four, and that is slow. So what Karatsuba did is he figured out the way to do this with just three multiplications. So it's not these three. You have to think a little bit more. But you google it, or you look at the lecture notes, or you think a little bit, then you find out the three multiplications that do the trick.

So the difference between this 4 here and this 3 here is important. All right, anything else on Karatsuba? Cool. Let's talk about Newton. That's the hard stuff. And let me see how we're doing on time. OK, doing well.

So does anyone remember the point of Newton? OK, so I heard an application that lets you compute square roots. What else does it let you do?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** OK, we have Karatsuba for multiplication. You might be able to do it with Newton. Probably not going to be faster. But division, definitely. So square roots and division. So this is why we care about it.

Now what does it do from a mathy perspective? So what's the mathy definition of Newton.

**AUDIENCE:** Estimates--

**PROFESSOR:** Estimates?

**AUDIENCE:** The value--

**PROFESSOR:** OK. So, we have a function. Suppose it looks like this. So, what Newton helps us do is it helps us find a root of the function. And the root is a value  $x_0$  so that  $f(x_0)$  equals 0.

So this is what it does. So when you want to use Newton, you have to set up your function in such a way that your answer is a root. OK. How does Newton work? Yes?

**AUDIENCE:** You start at some point where you actually know the answer to it, and then you kind of draw a approximation. So you're just drawing a line from the point to the root, and then just incrementing down that line.

**PROFESSOR:** OK. So we start at some point, right? How do we call that point? Initial guess. Fancy name.

So we have an initial guess for what the root might be. It's not going to be the right answer, because otherwise why am I bothering with this? But it can be too far off. We'll say that if you're too far off, you're not going to converge or you're not going to converge fast enough.

So say we start somewhere here and, we think, you know, this is pretty close. It's not quite there, but it's pretty close. What I do from here on is, you said I approximate this function--

**AUDIENCE:** The line.

**PROFESSOR:** Yep. So what's of the best line? The best way to approximate the function with the

line is to use is the tangent, right? So here, it sort of looks like this. And if the function were linear, this would be a perfect approximation and we'd get answer in one step.

OK. So this is the tangent. And then what we do with this tangent?

**AUDIENCE:** Well, after we get the equation for it, then we find the root, right? Or no?

**PROFESSOR:** Well, I can't find the root right away.

**AUDIENCE:** The root of that line. Of our tangent line.

**PROFESSOR:** OK. So we find the root of the tangent, which is the point where it hits the x-axis. And what happens with this point?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** OK, so this is our next guess. And then we trade. So then we draw another tangent, and the idea is that eventually, you're going to get there. And if we go through the maths, we will see that you actually get there pretty fast. So that's why we like this method.

So let's see how we convert this drawing thing into an equation that you can write into python. Let's label these quantities. So,  $x_2$ ,  $x_1$ -- actually let's not use 1 and 2, because these are the first two guesses. Let's say our current guess is  $x$  and the next guess is  $x$  prime.

So this line and is  $x_1$  minus  $x$ , so this guy over here. And what's this guy over here?  $f$  of  $x$ . Don't be shy. It's the right answer.

So I have  $f$  of  $x$  here,  $x_1$  minus  $x$  here. I also have this angle here, and this angle here is the same angle as the angle here because both of these lines are horizontal, so they're parallel. So what do I know about this angle?

**AUDIENCE:** It's tangent.

**PROFESSOR:** It's tangent. OK. The tangent, that's called the angle alpha. What's the tangent of

the angle?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Sorry?

**AUDIENCE:** F of x over x [INAUDIBLE].

**PROFESSOR:** I think you're trying to give me the formula for x prime. Just the tangent. Just the tangent.

**AUDIENCE:** Oh.

**PROFESSOR:** Right? Can someone reassure me? This is right, right?

[AUDIENCE AGREES]

**AUDIENCE:** I was thinking the slope of the line.

**PROFESSOR:** OK. So, yeah, you're going to give me the slope of the line which lets me compute this. OK, I was looking for this.

OK. So now, let's get what you're saying. So this is the tangent, right? So, the tangent at a point of a function is the function's derivative. So now we're going to write all these things down and hope that we get an equation that make sense.

So we have the tangent of this guy, also happens to be this over this. Right? This is a right angle triangle, so the tangent is f of x over-- this is what you were trying to tell me, right?  $x^1$  minus x. So now I know that this is f prime of x. OK.

So, we look at the picture, we put in everything we know, and then we write an equation here. And we solve for-- sorry, this is x prime. So, if we solve for x prime, we're going to get x prime as a function of x. And this tells us, given a guess here, how do we compute the next guess? So once we have an initial guess, this lets us make the guess better and better and better, closer and closer and closer to  $x_0$ .

So, let's solve this equation reasonably quickly.  $x$  prime minus x equals f prime of x



over  $f$  of  $x$ , so  $x$  prime is--

**AUDIENCE:** Does it matter if it's  $x$  minus  $x$  prime or [INAUDIBLE], because technically it's  $x$  minus  $x$  prime.

**PROFESSOR:** Yeah, it matters, because you're going to get the wrong sign. Which is exactly what I'm about to do here.

**AUDIENCE:** Oh, OK.

**PROFESSOR:** So, I'm not going very well, right? This is wrong.

**AUDIENCE:** Yeah, it's because you flipped your  $x$  at your prime--

**AUDIENCE:**  $x$  minus  $x$  prime instead of  $x$  prime--

**AUDIENCE:** No, I think it's because  $f$  prime of  $x$  is on top. It should be on the bottom. In the second-- no, third equation. Shouldn't those be flipped? Because if you move the subtraction,  $x$  prime to the left--

**PROFESSOR:**  $f$  what?

**AUDIENCE:**  $f$  of  $x$  over  $f$  minus  $x$  instead of  $f$  prime  $x$  over  $f$  of  $x$ . Yeah, there you go.

**PROFESSOR:** OK, so I can do math. OK, so that's one issue. There's one more issue. So, there are too many issues here. Issue number one, this. Issue number two, I got this wrong.

**AUDIENCE:** Oh, we get a negative number.

**AUDIENCE:**  $2x$  minus  $x$  prime.

**PROFESSOR:** Yep. So the reason for that is we're going--

**AUDIENCE:** Only if you start from the right, though.

**PROFESSOR:** Yeah.

**AUDIENCE:** If you start from the left, if your guess is to the left of the actual root, it may or may not be different.

**PROFESSOR:** OK. So our guess should be to the right. So this is  $x$  minus  $x$  prime. So the reason it's  $x$  minus  $x$  prime, the way you remember this and you don't do the same mistake that I did is, you want to walk against the derivative. If you walk in the direction of the derivative, it's going to move you forward, right?

So, if you go in the direction of the derivative, instead of going backwards like this, you're going to go forward like this. So you're going to go away from the solution. So, this will take us in the right direction, and this means that it's--

OK. So two ways to remember this. One that you're going against the derivative. Another one is that you have your guess and you're subtracting something from it.

**AUDIENCE:** You flipped your primes again.

**PROFESSOR:** Again? Jesus.

[LAUGHTER]

**AUDIENCE:** I mean, I agree, the prime looks better on top part of the fraction, but still.

**PROFESSOR:** OK. Well clearly, my hand thinks the same thing. So this is you'd iterate using Newton. So the way to remember that the signs are right is your  $x$  is somewhere here, and you know you have to subtract something from it so it looks like this figure. And this actually works in all the cases, turns out.

Our guesses are usually smaller than the actual number, and this still works. OK, so let's go over Newton. What do we have, or what do we need in order to make Newton happen.

So we need to be able to evaluate this, right? So we need to be able to evaluate  $f$  over  $f$  prime of  $x$  for any  $x$  that Newton gives us. And for Newton, this is a black box. So whichever way you use to implement this is fine, as long as it works reasonably fast.

OK, what else do we need to use Newton. We need to remember this formula Let's start with something easy. Suppose we want to use it to compute  $a$  over  $b$ . How do I do that? What do I need?

**AUDIENCE:** You need a line, a tangent line like you did before.

**PROFESSOR:** Well, that's already taken care of here. So we have this, and we know that we'll have to compute this and plug it in here. So, what do I need to compute  $a/b$  using Newton?

So, I need the function, right? This thing. I can't say it's something, I have to figure out what it is. I guess I was confusing because I said assume we have this. Well, sorry. We don't have this. We need to figure out the function so we can evaluate this.

OK, so we need to come up with the function so we can evaluate that, and what property does the function need to meet?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Yeah, that would be nice, right? So we want the function so that  $f$  of  $a/b$  equals 0, because that's the number that Newton will give us. And we also need to be able to compute  $f$  over  $f$  prime of  $x$  without doing any division, right? If we did the divisions, it doesn't work.

OK, so the first thing we're going to do is we're going to simplify this problem to 1 over  $b$  instead of  $a$  over  $b$ , because if you know what's 1 over  $b$ , you're going to multiply this by  $a$ , and you're going to get  $a$  over  $b$ . Suppose we work with integers and we don't want to touch floating point numbers. What would we actually compute?

**AUDIENCE:** You shift it out, right? To get that 0.23, shift now to [INAUDIBLE].

**PROFESSOR:** All right, all right. So 1 over  $b$ . Unless  $b$  is 1, in which case it's really boring. 1 over  $b$  looks like this. 0 point something. This is not an integer number. I don't like it.

So what I do is, I decide that I'm going to need digits of precision, and I'm going to shift the decimal point right by  $d$  digits. So instead of 0 point something, it's  $d$  digits here point something else. How do I do this?

**AUDIENCE:** Multiply by 10 to the  $d$ ?

**PROFESSOR:** Multiply by 10 to the  $d$ . So instead of computing  $1$  over  $b$ ,  $10$  to the  $d$  over  $b$ . And we call this  $10$  to the  $d$   $R$  in our station notes.

So we're going to compute  $R$ . Over  $b$  multiplied by  $a$ , and then everything divided by  $R$ . So we talked about how you multiply by  $R$ . How do you divide by  $R$ ? So, if I have a number with  $n$  digits and I want to divide by  $10$  to the  $d$ , how do I do that?

**AUDIENCE:** Shift.

**PROFESSOR:** Shift right. If the number is a python list?

**AUDIENCE:** Slice it.

**PROFESSOR:** Yep. So kill the last  $d$  digits. So, if I add  $d$  digits to the right, I multiply by  $10$  to the  $d$ . If I slice the last digits, I divide by  $10$  to the  $d$ . So this is division, this is multiplication.

And the reason I care about that is this is going to be an easy division, and whenever I have this, it's an easy multiplication. OK, now what function am I going to use? Does anyone know? Can I use  $x$  minus  $R$  over  $b$  equals zero?

**AUDIENCE:** Sure?

**PROFESSOR:** Well, so, this is  $f$  of  $x$ .  $f$  prime of  $x$  is what?

**AUDIENCE:** Just negative  $R$  over  $b$ . Or no,  $1$ .

**PROFESSOR:** OK. So,  $f$  prime of  $x$  over--

**AUDIENCE:** I think you switched it again.

[LAUGHTER]

**PROFESSOR:** Thank you. Because I was running ahead, and I was like, OK, this works. Why does it work? OK, so this is going to be  $x$  minus  $R$  over  $b$ , right? Can you evaluate this? If I can't evaluate this, I already know the answer.

**AUDIENCE:** Your prime's not on the bottom one.

**PROFESSOR:** So, it's  $x$  minus this divided by 1. So, it's this. So then, if you're trying to evaluate this far the guesses that Newton gives you, you can't. Because if you could, you'd have the answer already. So we can't use this function. This is a nice and easy function that doesn't work.

So while I erase the board, one minute, someone figure out the right function.  
Sorry?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Man, that's hard math. No, that's too hard. That's too hard.

OK, let's try this. Let's try instead of  $x$  minus  $R$  over  $b$ , let's try this  $1/x$  minus  $b/R$ . And the intuition behind this is if I have to compute  $b/R$ , I can do that. Right? It's an easy division. If I somehow get rid of  $1/x$  and all I have to do is compute  $b/R$ , it's an easy division. Because dividing by  $2$  to the  $R$  by  $R$  means you're going to shift by some digits to the right.

So let's see if the math works out for this function. What's  $f$  prime of  $x$ ?

**AUDIENCE:** Negative 1 over  $x$  prime?

**PROFESSOR:** All right. Like, does no one know it? At least one student has to know it. OK, so  $f$  of  $x$  divided by  $f$  prime of  $x$  is what?

**AUDIENCE:** Negative  $x$  plus  $x$  squared  $b/R$ .

**PROFESSOR:** Plus  $x$  squared  $b R$ .

**AUDIENCE:** Yeah.

**PROFESSOR:** Is this right?

**AUDIENCE:** Yes.

**PROFESSOR:** OK. So if I have my initial guess-- so if I want to get  $x$  prime,  $x$  prime is  $x$  minus  $f$  of  $x$  over  $f$  prime of  $x$ , right? So, what is it?

**AUDIENCE:**  $2x$  minus  $x$  squared over  $R$ . OK. Did this work? Did it get rid of all the hard divisions?

[LAUGHS]

Please nod so I know we're good to go. Tell me if we're not. So how many divisions did we have here? One division, this guy, right?

**AUDIENCE:** But that's a shift.

**PROFESSOR:** That's a shift. Because  $R$  is  $10$  to the  $d$ .

**AUDIENCE:** What about the multiplication [INAUDIBLE]?

**PROFESSOR:** So, we know how to do multiplication because we learned Karatsuba just now. So we wanted to get rid of division, because if you're trying to solve division by doing division, it's probably not I'm going to work unless you're doing divide and conquer.

OK, so this is how it works. I'm not going to run into the error formulas for this because I want to do another case first. You can keep going on this and it's in the notes. The only thing that is useful to remember is that we use this trick.

So, I want to look at this function, square root 3 of  $a$ , and I want to use Newton to compute it. So this is the third order root. It's not the regular square root that we had before. So we're trying to not be boring and not just repeat what's in the notes. How do I do this? Let's walk through the steps.

So what's the first thing that we do? Very good. Look to the right. Find the first step, and let's put it here. What's the first step.

**AUDIENCE:** Function  $f$  of 0 at cube root of  $a$ .

**PROFESSOR:** OK. So we need a function  $f$  so that  $f$  of-- good. So before we start writing the actual function, what if I want to compute cube root of  $a$  up to  $d$  digits of precision? So if I want to compute cube root of 2, this is going to give me numbers that are below 1. I have to do with fractions again. I don't like that. If I want the digits of precision, what do I do?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Jumping ahead. But yeah. So you have the right answer, but you're speaking slowly because you're afraid you're going to do a mistake, right? So, the way to compute this to  $d$  digits of precision  $10$  to the  $d$ , right? This shifts the dot to the right.

So I have to get this  $10$  to the  $d$  inside here, and I get and to the  $3d$  times  $a$ . Perfect answer, by the way. Don't forget this  $3$  here. If you do, you're going to get something else.

So now I take this guy, and this is the new number whose cube root I have to compute. Cube, cube, cube. OK.

So I can do this in integer land, and then I'll get an integer result, and I shift it right by  $d$  digits, and I have a number with  $d$  digits of precision. Does this make sense?

OK, what's  $f$ ?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Let's go for that. Let's see if that works.  $x$  cubed minus  $a$ . Well, let's pretend. We're going to make this  $b$  the new  $a$ . So let's not worry about this. This happens somewhere else. We went through it so that we know that we know how to deal with it. And now we're going to forget about it.

So, it's  $x$  cubed minus  $a$ . Because we argued that this is an integer, right? So we don't have to do anything fancy like dividing by  $R$  that we had to do over there.

So, what's  $f'$  of  $x$ ? OK. So, iteration? How do we iterate?

**AUDIENCE:**  $x'$  prime?

**PROFESSOR:** Let's start with  $x'$  equals  $x$  minus  $f$  of  $x$  over  $f'$  of  $x$ . Right? Nice and easy so we have it here. And this becomes minus copy these two things, right?  $x^3$  minus  $a$  over  $3x^2$ .

OK. Next thing, I bring this guy under the denominator, so it's  $3x^3$  minus  $x^3$  plus  $a$  over  $3x^2$ . So it's  $2x^3$  plus  $a$  over  $3x^2$ . So this is  $\frac{2}{3}$  plus--

**AUDIENCE:**  $\frac{2}{3}$  over  $3x$ . You keep saying cubed and writing 2.

**PROFESSOR:** Yeah, that's better. Thank you. Plus  $\frac{1}{3}$  times  $a$  over  $x$ . So this one's division.

**AUDIENCE:**  $x^2$ .

**PROFESSOR:**  $x^2$ . Ouch. So this guy wants division. By the time I do this, I better have division ready, right?

OK, so that my iteration.  $x'$  is  $\frac{2}{3}x$  plus  $\frac{1}{3}$  and  $a$  over  $x^2$ . What else do I need?

An initial guess would be nice, right? I can't do this if I don't have the first  $x$ . So let's try to think of a starting  $x$  that's reasonable. The better it is, the faster it's going to be. So let's try to come up with something reasonable that doesn't require too much thinking or too much.

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Too much work. I have to compute a square root. So what's  $a$ ? Let's start with that. What's an  $a$ , right? So, it's an  $n$  digit number. And say the digits are  $d_1, d_2, d_3$ , all the way up to  $d_n$ . So what's a good approximation? Well, the first approximation for this?

**AUDIENCE:** You get a third as many digits.



**PROFESSOR:** A third as many digits. OK. So something times 10 to the  $n/3$ . Let's say that something is 1, right?

So this is the first approximation. It's the right order of magnitude. Do people see why this is a good approximation? So this number, whatever it is, it has to be between 1, 0, 0 0, 0, and 1, 0, 0.

OK. So generally when we do this, what we have to have in an approximation is the right order of magnitude And maybe the first digit. If we can figure out the first digit, then we're in a good shape, or something really close to the first digit. How would we get the first digit?

**AUDIENCE:** Cube root of the first digit?

**PROFESSOR:** OK. Sure. How do we compute this?

**AUDIENCE:** Get a formula?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** OK, so I'm not going to argue with that.

**AUDIENCE:** I like that.

**PROFESSOR:** So I want to reject the recursive argument first. So this is going to be small, right? This is going to be 10 if you're working with digits, or if you're a computer, maybe it's 256, or maybe it's 2 to the 16 or 2 to the 32. Not much bigger than that, right? Pretty small.

So these algorithms really good for big numbers. Algorithms that are complicated but have nice asymptotic complexity are good for big numbers. They're awful for small numbers, as in they 10 to 100 times more times to run. So let's not to do that for this, because it's like bringing a tank and shooting it so you can hammer a nail in. Not the right tool.

So, precomputed list. I like that. It's an option, especially if you're [INAUDIBLE]

right? You can use pen and paper to get it.

What if I don't want to do that. What if my base is more like this.?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** So this is a very good solution. I'm not arguing with it. If your power is this or this, by all means, use a table. So let's see what do we do here? I just want to think a little bit more.

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** OK, so there's some approximation that comes from math that says that  $d_1$  over 3 is not going to be such a bad guess. Now let's be CS people and do something better.

**AUDIENCE:** [INAUDIBLE] the exponent. 2 to the 14, divide-- is that 14, right here

**AUDIENCE:** 16.

**AUDIENCE:** 16. Divide 16 by 3. Or I guess that's not quite right either.

**PROFESSOR:** So you're saying 2 to the power of, say, 5? 15 over 3?

**AUDIENCE:** It doesn't seem quite right.

**PROFESSOR:** So, let's try to do better. This is math. Let's just try to do it in a CS way. So we're trying to guess a digit between 0 and 2 to the 16. I heard an "oh." Yes? You had your hand up? No? OK.

**AUDIENCE:** You take the current time 2 to the 16. I mean--

[LAUGHTER]

**PROFESSOR:** So your algorithm is going to be fast with some probability, right? No, that's terrible. Let's not do that. Please don't make me grade code like that.

**AUDIENCE:** Did you already cut out the recursive option?

**AUDIENCE:** Yes.

**PROFESSOR:** Yeah. Well, so I don't want to recurse on square roots.

**AUDIENCE:** But you could recurse on that one.

**PROFESSOR:** I mean, if you give me some algorithm that happens--

**AUDIENCE:** You're already [INAUDIBLE].

**PROFESSOR:** OK. So, what's the intuition? If you have a computer and you have register, bytes, bits, whatever. You don't know how to do square roots. That's not an instruction, but these are all instructions. Right?

So if I have a guess, say I start with half and then I get my guess, what do I need for the binary search to work? Given a guess that's somewhere here, I have to be able to say whether this is too small, too big, or just the right number.

So given a guess, I can multiply it by itself three times, right? Compute gets cubed. And I can see how does it compare to the number that I want.

So I have the first digit, and I have an order of magnitude. Not bad, right? Pretty good initial guess. This is as much of an answer as you need to give out in physics problems, and you get full score, so that's good.

OK, how are we doing so far? Following along?

**AUDIENCE:** If we compute our guess cubed, we want only to compare it to our original number to see how close it is?

**PROFESSOR:** Yup.

**AUDIENCE:** And, do we just keep on doing that until--

**PROFESSOR:** Well, so you're going to have to guess a digit, right? Between 0 and 2 to the 16. So you choose the largest thing that's smaller than your guess, or something.

**AUDIENCE:** Or until we get the first digit.

**PROFESSOR:** Yeah. Just the first. And then you have a first digit, and then you have an order of magnitude. And then you feed that to Newton, which converges way faster. So you could the binary search for the whole thing, but that would be slower than Newton, turns out. Yes?

**AUDIENCE:** So where'd you get the f of x here? I'm missing that stuff.

**PROFESSOR:** Oh.

**AUDIENCE:** Yeah, right there.

**PROFESSOR:** Someone told me to use this.

[LAUGHTER]

One of you guys told me to use this. Well, so I tried it out. I put it here. And then I computed the derivative, and I looked to see if this looks reasonable. And this looks like something I can compute.

**AUDIENCE:** And why was that reasonable?

**PROFESSOR:** Because I know how to do the division and multiplication. So, say if I would have had the square root of 3 in here, that wouldn't be reasonable, right? Just like before. When you're trying to compute division, we can't have division. If you're trying to compute sine of x, you don't want to have sine of x.

**AUDIENCE:** OK. and computing a over x squared is reasonable?

**PROFESSOR:** Because we did it before.

**AUDIENCE:** OK. So then you might use something else, maybe?

**PROFESSOR:** Yeah. You can use something else as long as you know how to do it, and as long as it's reasonably fast.

**AUDIENCE:** OK.

**PROFESSOR:** Yes?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Oh, no that's fine.

**AUDIENCE:** Oh, that fine? OK.

**PROFESSOR:** So, the binary search is fine to get the first approximation. But after that-- so I'm not going to do a binary search for the whole result, because Newton is faster than that. I'm just using this so that I have a decent guess to pass to Newton. And we'll see why we need that decent guess in two minutes total. So I'm going to try to spend two minutes to explain how the error thing works, and you'll have to look at the notes and make sense out of the rest of it.

So, let's see. Errors. So with Newton, you start to with an initial guess, and then you get to a better guess, and a better guess, and a better guess, and so on and so forth until you're satisfied. By the way, if we're doing this one, are we satisfied?

**AUDIENCE:** What do you mean, like, finding the guess? Are we satisfied?

**PROFESSOR:** Oh, no. So, we have the initial guess. And we keep applying the formula that I just deleted to iterate. When do we stop?

**AUDIENCE:** When we have enough digits of precision?

**PROFESSOR:** So how do we know when we have enough digits, if we're working with integers?

**AUDIENCE:** We iterate again and nothing changes.

**PROFESSOR:** We iterate again and nothing changes. So I'm computing-- do I have this anywhere? Oh, no. I erased it. So I'm computing  $10^{10}$  some-- oh, I had it up here.

So, I'm computing  $10^d$  times a square root of 3, because I want to work with integers. Just integers. I used Newton to iterate. My precision gets gets better and better and better. And I use integers. So I only have  $d$  digits of precision.

Everything after that gets discarded automatically.

When I get 2 x's that don't change-- so if I get  $x_5$ , and then I run the thing and then I get  $x_6$  that is equal to  $x_5$ , there's no point to keep running Newton, right? I'm stuck. And I'm also done because I have the answer with  $d$  digits of precision. So this is where I stop.

OK. Now, my precision is how far I am from the right answer. And we write that by saying that  $x_1$  is  $x$ , the true answer, times  $1 + \epsilon_1$ . And then  $x_2$  is  $x$  times  $1 + \epsilon_2$ ,  $1 + \epsilon_3$ , so on and so forth. Right?

So, if we had a lot of time and not too much stuff to do, I could copy this thing here that we have in the notes. That's eight lines of math. And then we'd get something along the lines of  $x_{i+1}$  is  $x$  times  $1 + \text{something} \times \epsilon_i^2$ -- I'm sorry,  $\epsilon_i^2$ . So this is also  $1 + \epsilon_{i+1}$ .

So what we get is that  $E$  of  $i + 1$  is roughly  $E_i^2$ . And this is what's really called quadratic convergence. Now why is this good? Can anyone think intuitively why this is nice? In which cases? Yes?

**AUDIENCE:** You're converging really quickly to your answer.

**PROFESSOR:** OK. But when do I converge? , So what I want this to happen? What do I want this to be?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Well, you guys told me everything I need for the next two minutes. So, I want this to become 1, right? This has to go to 1 because I want this whole thing to be  $x$ . So in order for this to be 1, I want to epsilons to go to 0. So I want them to be as small as possible.

Now, you said that if my epsilon is smaller than 1, then epsilon squared is smaller than epsilon. And intuitively, how this works is, suppose you start with an epsilon of 0.1. When you square it, you get 0.01. You square it again, you get 0.0001. Square it again. 0.00000001.

OK. This is step 0, step 1, step 2, step 3. How many digits do I have here? Don't be shy.

**AUDIENCE:** 2 to the i?

**PROFESSOR:** Yup. 2 to the i. So this is going to be equal to this when 1 plus x times 1 plus epsilon i over 1 plus epsilon-- sorry, i plus 1 epsilon i when the floor of this thing is-- sorry, there is a delta somewhere here. So, something has to converge to 0.

Sorry, this is what it is. 1 plus e of i. So this is the difference between two terms, x. So this is x6 minus x5, right? Convincing?

**AUDIENCE:** [INAUDIBLE]. Let's write it again. Let's write it again. This is out of hand. So, x, x minus x5, or actually, xi plus 1 minus xi is equal to x times 1 plus epsilon i plus 1 minus x times 1 plus epsilon i, right? So this is x times-- Sorry?

**AUDIENCE:** Those are xi, x of i.

**AUDIENCE:** No, those are the real x's.

**PROFESSOR:** These are-- so these are the approximations, but then, this is the real x.

**AUDIENCE:** I understand, thank you.

**PROFESSOR:** OK. So, this is, the difference between them is this guy here. So if I take the floor because I work with integers, when this floor becomes 0, we're done, right? So in order for the floor to become 0, this guy has to be canceled by this guy. So this guy has some number of digits, whatever that is. It's my number with d digits of precision.

Well, that's OK because-- where my 0s? Did I erase my 0s? Oh, no. I have 0s here, right?

The epsilons have a lot of 0s in them, so the 0s are going to cancel the digits in x. So as soon as I have enough 0s here to cancel this, I'm stuck. The number of 0s increases exponentially. So in each step the number of 0s I have doubles. So after

log  $d$  steps, I'm done. If I have  $d$  digits in here, because I'm computing this with  $d$  digits of precision, after roughly log  $d$  steps, I'll have enough 0s here to reduce this and make the whole thing be 0. So this is how I know that it's fast.

So in the end, what you want is the initial guess has to satisfy this, and then you want to be able to prove that this is true. So you want this factor here that tells you you're converging fast. So when you look at the recitation notes and you look at this long string of equations, make sure this is somewhere there, because this means you're on the right track. Yes?

**AUDIENCE:** If it takes log  $d$  steps, doesn't the binary search also take log-- ah, OK. Log 10 steps. Got it.

**PROFESSOR:** Yeah. Which is  $d$ .

**AUDIENCE:** He's talking about binary search--

**PROFESSOR:** Oh, for one digit. Yeah. Oh, sorry. I thought you meant for the whole thing. So if you wanted to do binary search for the whole thing, that's log of the whole number. So that's order  $d$  where  $d$  is the number of digits.

**AUDIENCE:** Wait, he's talking about a binary [INAUDIBLE].

**AUDIENCE:** [INAUDIBLE].

**AUDIENCE:** [INAUDIBLE] wants to go for [INAUDIBLE], 10 to the 20th. He wants to do a binary search over like a cube root--

**PROFESSOR:** So, for this? Binary search for this or for one digit? Or binary search to compute the cube root?

**AUDIENCE:** I think, yeah.

**PROFESSOR:** So, if you do that, your entire thing is  $d$  digits, right? So 10 to the  $d$ . Binary search takes log of 10 to the  $d$ .

**AUDIENCE:** Right, yeah.



**PROFESSOR:** So the numbers that we're dealing with are  $d$  digits. So this  $d$  is the number of digits, not the number that you have. So if you look at the number that you have, this is  $\log \log n$ .

**AUDIENCE:** Right, that makes sense. Sorry.

**PROFESSOR:** So, it's faster.

**AUDIENCE:** I like it. It is faster.

**PROFESSOR:** OK. Cool. Thanks for the lesson.

**AUDIENCE:** [INAUDIBLE].