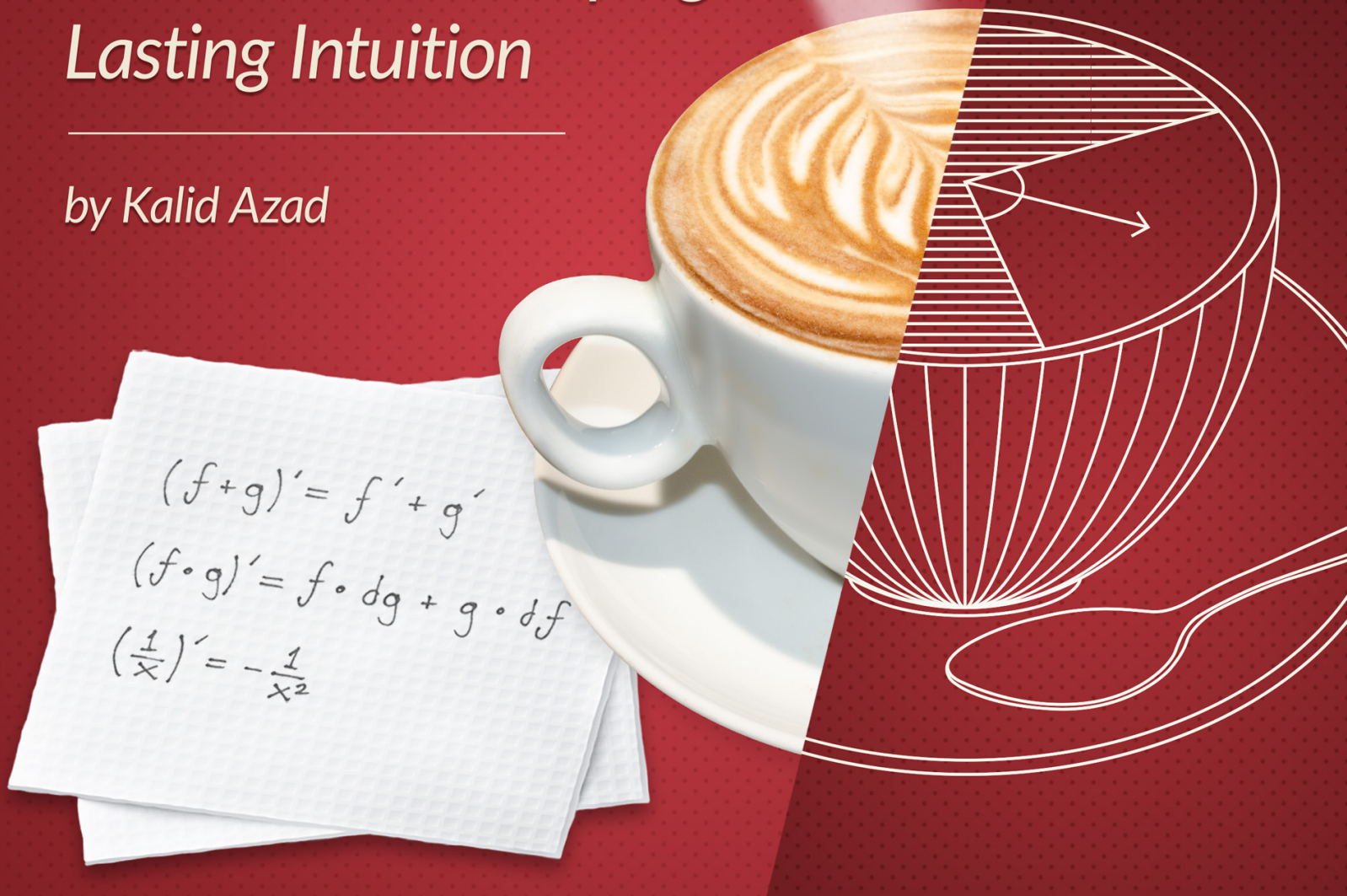
Calculus

Better Explained

A Guide to Developing Lasting Intuition

by Kalid Azad



Contents

Contents i

1 1 Minute Calculus: X-Ray and Time-Lapse Vision 1

2 Practice X-Ray and Time-Lapse Vision 5

3 Expanding Our Intuition 11

4 Learning The Official Terms 14

5 Music From The Machine 21

6 Improving Arithmetic And Algebra 25

7 Seeing How Lines Work 29

8 Playing With Squares 35

9 Working With Infinity 40

10 The Theory Of Derivatives 43

11 The Fundamental Theorem Of Calculus (FTOC) 47

12 The Basic Arithmetic Of Calculus 51

13 Patterns In The Rules 57

14 The Fancy Arithmetic Of Calculus 61

15 Discovering Archimedes' Formulas 67

CHAPTER 1

1 MINUTE CALCULUS: X-RAY AND TIME-LAPSE VISION

We usually take shapes, formulas, and situations at face value. Calculus gives us two superpowers to dig deeper:

- **X-Ray Vision:** You see the hidden pieces inside a pattern. You don't just see the tree, you know it's made of rings, with another growing as we speak.



- **Time-Lapse Vision:** You see the future path of an object laid out before you (cool, right?). "Hey, there's the moon. For the next few days it'll be white, but on the sixth it'll be low in the sky, in a color I like. I'll take a photo then."



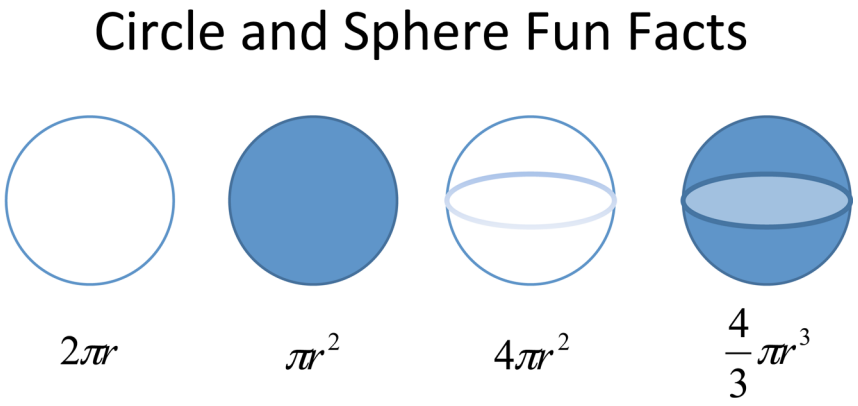
So why is Calculus useful? Well, just imagine having X-Ray or Time-Lapse vision to use at will. That object over there, how was it put together? What will happen to it?

(Strangely, my letters to Marvel about Calculus-man have been ignored to date.)

1.1 Calculus In 10 Minutes: See Patterns Step-By-Step

What do X-Ray and Time-Lapse vision have in common? They examine patterns step-by-step. An X-Ray shows the individual slices inside, and a time-lapse puts each future state next to the other.

This seems pretty abstract. Let's look at a few famous patterns:

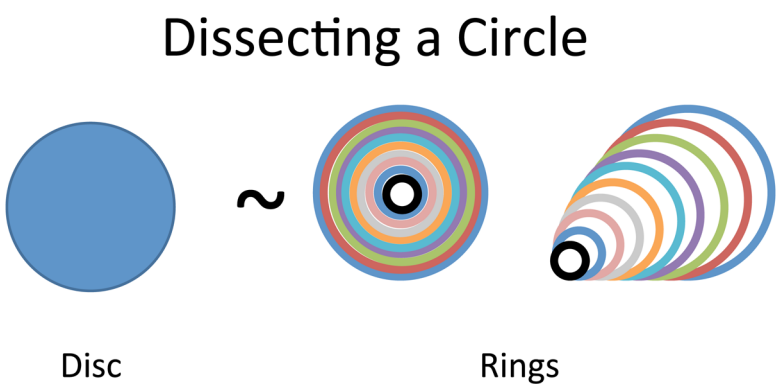


We have a vague feeling these formulas are connected, right?

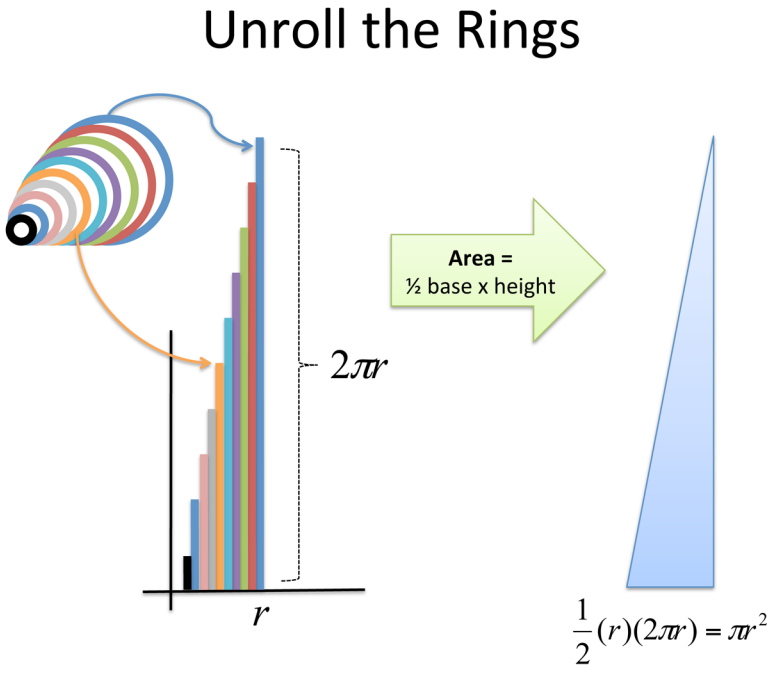
Let's turn our X-Ray vision and see where this leads. Suppose we know the equation for circumference ($2\pi r$) and want to figure out the equation for area. What can we do?

This is a tough question. Squares are easy to measure, but what can we do with an ever-curving shape?

Calculus to the rescue. Let's use our X-Ray vision to realize a disc is really just a bunch of rings put together. Similar to a tree trunk, here's a "step-by-step" view of a filled-in circle:



Why does this viewpoint help? Well, let's unroll those curled-up rings so they're easier to measure:



Whoa! We have a bunch of straightened rings that form a triangle, which is much easier to measure (Wikipedia has an animation).

The height of the largest ring is the full circumference ($2\pi r$), and each ring gets smaller. The height of each ring depends on its original distance from the center; the ring 3 inches from the center has a height of $2\pi \cdot 3$ inches. The smallest ring is a pinpoint, more or less, without any height at all.

And because triangles are easier to measure than circles, finding the area isn't too much trouble. The area of the "ring triangle" $= \frac{1}{2} base \cdot height = \frac{1}{2} r(2\pi r) = \pi r^2$, which is the formula for a circle's area!

Our X-Ray vision revealed a simple, easy-to-measure structure within a curvy shape. We realized a circle and a set of glued-together rings were really the same. From another perspective, a filled-in disc is the "time lapse" of a ring that got larger and larger.

Remember learning arithmetic? We learned a few things to do with numbers (and equations): add/subtract, multiply/divide, and use exponents/roots. Not a bad start.

Calculus gives us two new options: split apart and glue together. A key epiphany of calculus is that our existing patterns can be seen as a bunch of glued-together pieces. It's like staring at a building and knowing it was made brick-by-brick.

1.2 So... What Can I Do With Calculus?

It depends. What can you do with arithmetic?

Technically, we don't need numbers. Our caveman ancestors did fine (well, for some version of "fine"). But having an idea of quantity makes the world a lot easier. You don't have a "big" and "small" pile of rocks: you have an exact count. You don't need to describe your mood as "good" or "great" – you can explain your mood from 1 to 10.

Arithmetic provides universal metaphors that go beyond computation, and once seen, they're difficult to give up.

Calculus is similar: it's a step-by-step view of the world. Do we need X-Ray and Time Lapse vision all the time? Nope. But they're nice perspectives to "turn on" when we face a puzzling situation. What steps got us here? What's the next thing that will happen? Where does our future path lead?

There are specific rules to calculus, just like there are rules to arithmetic. And they're quite useful when you're cranking through an equation. But don't forget there are general notions of "quantity" and "step-by-step thinking" that we can bring everywhere.

Our primary goal is to *feel* what a calculus perspective is like (*What Would Archimedes Do?*). Over time, we'll work up to using the specific rules ourselves.

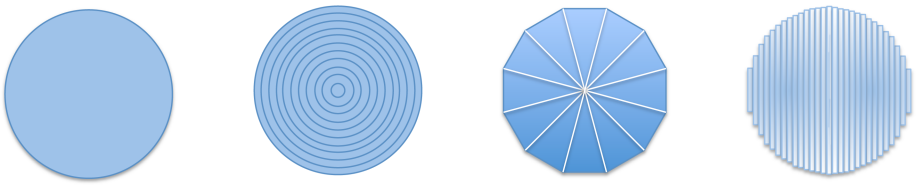
CHAPTER 2

PRACTICE X-RAY AND TIME-LAPSE VISION

Calculus trains us to use X-Ray and Time-Lapse vision, such as re-arranging a circle into a "ring triangle" (diagram). This makes finding the area... well, if not exactly *easy*, much more manageable.

But we were a little presumptuous. Must *every* circle in the universe be made from rings?

Heck no! We're more creative than that. Here's a few more options for our X-Ray:

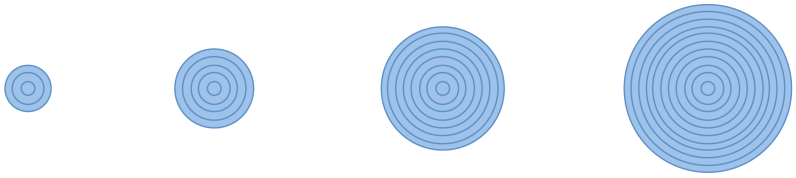


Original Rings Slices Boards

Now we're talking. We can imagine a circle as a set of rings, pizza slices, or vertical boards. Each underlying "blueprint" is a different step-by-step strategy in action.

Imagine each strategy unfolding over time, using your time-lapse vision. Have any ideas about what each approach is good for?

Ring-by-ring Analysis



Rings are the old standby. What's neat about a ring-by-ring progression?

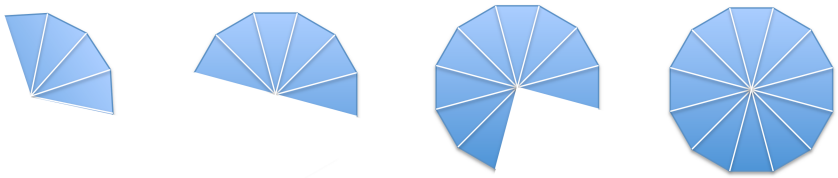
- Each intermediate stage is an entire "mini circle" on its own. i.e., when we're halfway done, we still have a circle, just one with half the regular radius.
- Each step is an increasing amount of work. Just imagine plowing a circular field and spreading the work over several days. On the first day, you start at the center and don't even move. The next, you make take the tightest turn you can. Then you start doing laps, larger and larger, until you are circling the entire yard on the last day. (Note: The change between each ring is the same; maybe it's 1 extra minute each time you make the ring larger).
- The work is reasonably predictable, which may help planning. If we know it's an extra minute for each lap, then the 20th ring will take 20 minutes.
- Most of the work happens in the final laps. In the first 25% of the time-lapse, we've barely grown: we're adding tiny rings. Near the end, we start to pick up steam by adding long slices, each nearly the final size.

Now let's get practical: why is it that trees have a ring pattern inside?

It's the first property: **a big tree must grow from a complete smaller tree**. With the ring-by-ring strategy, we're always adding to a complete, fully-formed circle. We aren't trying to grow the "left half" of the tree and then work on the right side.

In fact, many natural processes that grow (trees, bones, bubbles, etc.) take this inside-out approach.

Slice-by-slice Analysis



Now think about a slice-by-slice progression. What do you notice?

- We contribute the same amount with each step. Even better, the parts are identical. This may not matter for math, but in the real world (aka cutting a cake), we make the same action when cutting out each slice, which is convenient.
- Since the slices are symmetrical, we can use shortcuts like making cuts across the entire shape to speed up the process. These "assembly line" speedups work well for identical components.
- Progress is extremely easy to measure. If we have 10 slices, then at slice 6 we are exactly 60% done (by both area and circumference).

- We follow a sweeping circular path, never retracing our steps from an "angular" point of view. When carving out the rings, we went through the full 360-degrees on each step.

Time to think about the real world. What follows this slice-by-slice pattern, and why?

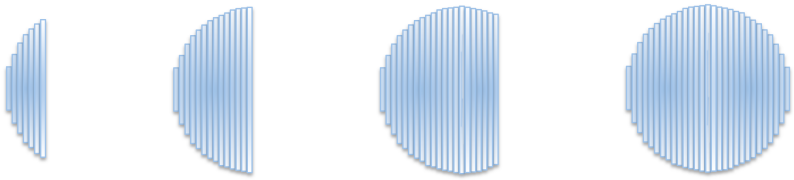
Well food, for one. Cake, pizza, pie: we want everyone to have an equal share. Slices are simple to cut, we get nice speedups (like cutting across the cake), and it's easy to see how much is remaining. (Imagine cutting circular rings from a pie and trying to estimate how much area is left.)

Now think about radar scanners: they sweep out in a circular path, "clearing" a slice of sky before moving to another angle. This strategy does leave a blind spot in the angle you haven't yet covered, a tradeoff you're hopefully aware of.

Contrast this to sonar used by a submarine or bat, which sends a sound "ring" propagating in every direction. That works best for close targets (covering every direction at once). The drawback is that unfocused propagation gets much weaker the further out you go, as the initial energy is spread out over a larger ring. We use megaphones and antennas to focus our signals into beams (thin slices) to get the max range for our energy.

Operationally, if we're building circular shape from a set of slices (like a paper fan), it helps to have every part be identical. Figure out the best way to make a single slice, then mass produce them. Even better: if one slice can collapse, the entire shape can fold up!

Board-by-board Analysis



Getting the hang of X-Rays and Time-lapses? Great. Look at the progression above, and spend a few seconds thinking of the pros and cons. Don't worry, I'll wait.

Ready? Ok. Here's a few of my observations:

- This is a very robotic pattern, moving left-to-right and never returning to a previous horizontal position.
- The contribution from each step starts small, gradually gets larger, maxes out in the middle, and begins shrinking again.
- Our progress is somewhat unpredictable. Sure, at the halfway mark we've finished half the circle, but the pattern rises and falls which makes it difficult to analyze. By contrast, the ring-by-ring pattern changed the same amount each time, always increasing. It was clear that the later

rings would add the most work. Here, it's the middle section which seems to be doing the heavy lifting.

Ok, time to figure out where this pattern shows up in the real world.

Decks and wooden structures, for one. When putting down wooden planks, we don't want to retrace our steps, or return to a previous position (especially if there are other steps involved, like painting). Just like a tree needs a fully-formed circle at each step, a deck insists upon components found at Home Depot (i.e., rectangular boards).

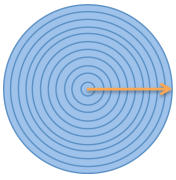
In fact, any process with a linear "pipeline" might use this approach: finish a section and move onto the next. Think about a printer that has to spray a pattern top-to-bottom as the paper is fed through (or these days, a 3d printer). It doesn't have the luxury of a ring-by-ring or a slice-by-slice approach. It will see a horizontal position only once, so it better make it count!

From a human motivation perspective, it may be convenient to start small, work your way up, then ease back down. A pizza-slice approach could be tolerable (identical progress every day), but rings could be demoralizing: every step requires more than the one before, without yielding.

Getting Organized

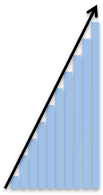
So far, we've been using natural descriptions to explain our thought processing. "Take a bunch of rings" or "Cut the circle into pizza slices". This conveys a general notion, but it's a bit like describing a song as "Dum-de-dum-dum" – you're pretty much the only one who knows what's happening. A little organization can make it perfectly clear what we mean.

The first thing we can do is keep track of how we're making our steps. I like to imagine a little arrow in the direction we move as we take slices:



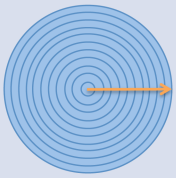
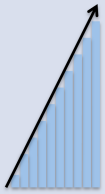
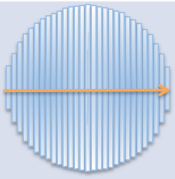
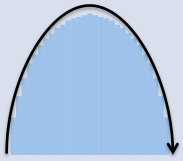
In my head, I'm moving along the yellow line, calling out the steps like Oprah giving away cars (*you* get a ring, *you* get a ring, *you* get a ring...). (Hey, it's my analogy, don't give me that look!)

The arrow is handy, but it's still tricky to see the exact progression of slices. Why don't we explicitly "line up" the changes? As we saw before, we can unroll the steps, put them side-by-side, and make them easier to compare:



The black arrow shows the trend. Pretty nice, right? We can tell, at a glance, that the slices are increasing, and by the same amount each time (since the trend line is straight).

Math fans and neurotics alike enjoy these organized layouts; there is something soothing about it, I suppose. And since you're here, we might as well organize the other patterns too:

X-Ray Strategy	Visualization	Step-by-Step Layout
Ring-by-ring		
Slice-by-slice		
Board-by-board		

Now it's much easier to compare each X-Ray strategy:

- With rings, steps increase steadily (upward sloping line)
- With slices, steps stay the same (flat line)
- With boards, steps get larger, peak, then get smaller (up and down; note, the curve looks elongated because the individual boards are lined up on the bottom)

The charts made our comparisons easier, wouldn't you say? Sure. But wait, isn't that trendline looking like a dreaded x-y graph?

Yep. Remember, a graph is a visual explanation that should *help* us. If it's confusing, it needs to be fixed.

Many classes present graphs, divorced from the phenomena that made them, and hope you see an invisible sequence of steps buried inside. It's a recipe for pain – just be explicit about what a graph represents!

Archimedes did fine without x-y graphs, finding the area of a circle using the "ring-to-triangle" method. In this primer we'll leave our level of graphing to what you see above (the details of graphs will be a nice follow-up, after our intuition is built).

So, are things starting to click a bit? Thinking better with X-Rays and Time-lapses?

PS. It may bother you that our steps create a "circle-like" shape, but not a *real, smooth* circle. We'll get to that :). But to be fair, it must also bother you that the square pixels on this screen make "letter-like" shapes, and not *real, smooth* letters. And somehow, the "letter-like pixels" convey the same meaning as the real thing!

Questions

1) Can you describe a grandma-friendly version of what you've learned?

2) Let's expand our thinking into the 3rd dimension. Can you of a few ways to build a sphere? (No formulas, just descriptions)

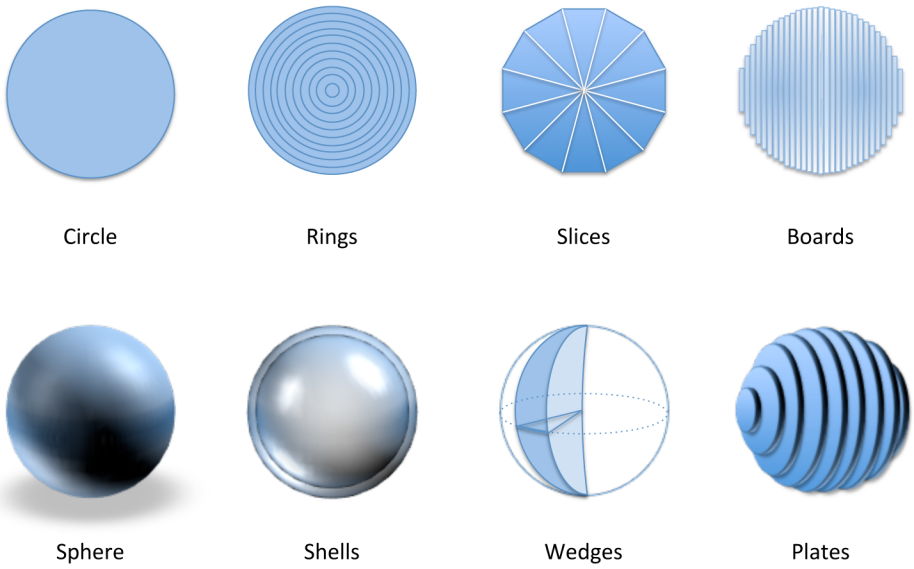
I'll share a few approaches with you in the next lesson. Happy math.

CHAPTER 3

EXPANDING OUR INTUITION

Hope you thought about the question from last time: how do we take our X-Ray strategies into the 3rd dimension?

Here's my take:



- Rings become *shells*, a thick candy coating on a delicious gobstopper. Each layer is slightly bigger than the one before.
- Slices become *wedges*, identical sections like slices of an orange.
- Boards become *plates*, thick discs which can be stacked together. (I sometimes daydream of opening a bed & breakfast that only serves spherical stacks of pancakes. *And they say math has no real-world uses!*)

The 3d steps can be seen as the 2d versions, swept out in various ways. For example, spin the individual rings (like a coin) to create a shell. Imagine slices pushed through a changing mold to make wedges. Lastly, imagine we spin the boards to make plates, like carving a wooden sphere with a lathe (video).

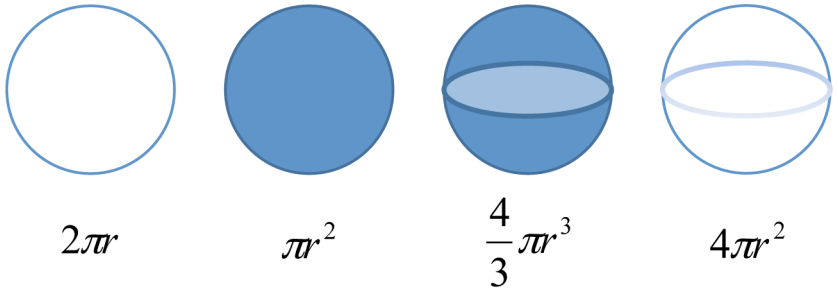
The tradeoffs are similar to the 2d versions:

- Organic processes grow in layers (pearls in an oyster)
- Fair divisions require wedges (cutting an apple for friends)
- The robotic plate approach seems easy to manufacture

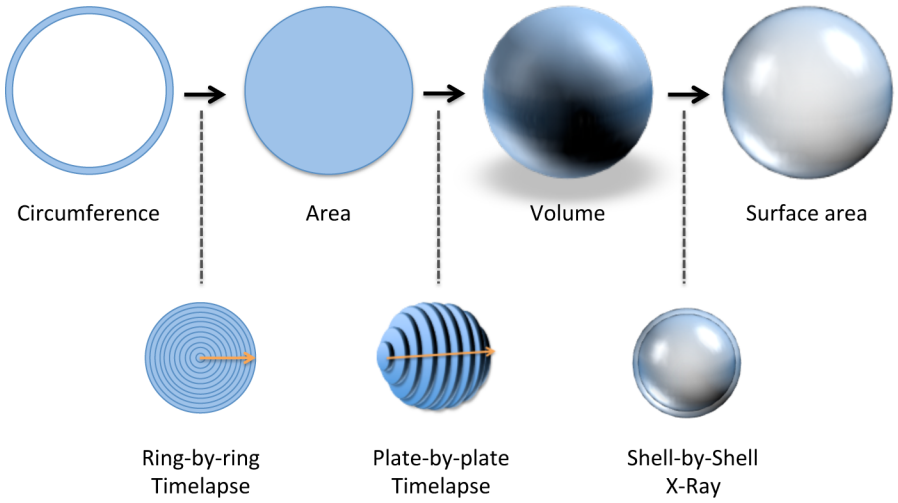
An orange is an interesting hybrid: from the outside, it appears to be made from shells, growing over time. And inside, it forms a symmetric internal structure – a better way to distribute seeds, right? We could analyze it both ways.

Exploring The 3d Perspective

In the first lesson we had the vague notion the circle/sphere formulas were related:



Well, now we have an idea how:



- **Circumference**: Start with a single ring
- **Area:** Make a filled-in disc with a ring-by-ring time lapse
- **Volume:** Make the circle into a plate, and do a plate-by-plate time lapse to build a sphere
- **Surface area:** X-Ray the sphere into a bunch of shells; the outer shell is the surface area

Wow! These descriptions are pretty detailed. We know, intuitively, how to morph shapes into alternate versions by thinking "time-lapse this" or "X-Ray that". We can move backwards, from a sphere back to circumference, or try different strategies: maybe we want to split the circle into boards, not rings.

The Need For Math Notation

You might have noticed it's getting harder to explain your ideas. We're reaching for physical analogies (rings, boards, wedges) to explain our plans: "Ok, take that circular area, and try to make some discs out of it. Yeah, like that. Now line them up into the shape of a sphere...".

I love diagrams and analogies, but should they be *required* to explain an idea? Probably not.

Take a look how numbers developed. At first, we used very literal symbols for counting: I, II, III, and so on. Eventually, we realized a symbol like V could take the place of IIIII, and even better, that every digit could have its own symbol (we do keep our metaphorical history with the number 1).

This math abstraction helped in a few ways:

- **It's shorter.** Isn't "2 + 3 = 5" better than "two added to three is equal to five"? Fun fact: In 1557, Robert Recorde invented the equals sign, written with two parallel lines (=), because "noe 2 thynges, can be moare equalle". (*I agrye!*)
- **The rules do the work for us**. With Roman numerals, we're essentially recreating numbers by hand (why should VIII take so much effort to write compared to I? Oh, because 8 is larger than 1? Not a good reason!). Decimals help us "do the work" of expressing numbers, and make them easy to manipulate. So far, we've been doing the work of calculus ourselves: cutting a circle into rings, realizing we can unroll them, looking up the equation for area and measuring the resulting triangle. Couldn't the rules help us here? You bet. We just need to figure them out.
- **We generalized our thinking.** "2 + 3 = 5" is really "twoness + threeness = fiveness". It sounds weird, but we have an abstract quantity (not people, or money, or cows... just "twoness") and we see how it's related to other quantities. The rules of arithmetic are abstract, and it's our job to apply them to a specific scenario.

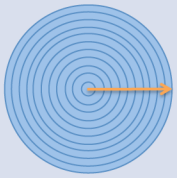
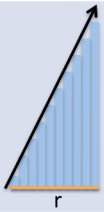
Multiplication started as a way to count groups and measure rectangular area. But when you write "\$15/hour × 2.5 hours = \$37.50" you probably aren't thinking of "groups" of hours or getting the area of a "wage-hour" rectangle. You're just applying arithmetic to the concepts.

In the upcoming lessons we'll learn the official language to help us communicate our ideas and work out the rules ourselves. And once we've internalized the rules of calculus, we can explore patterns, whether they came from geometric shapes, business plans, or scientific theories.

CHAPTER 4

LEARNING THE OFFICIAL TERMS

We've been able to describe our step-by-step process with analogies (X-Rays, Time-lapses, and rings) and diagrams:

Strategy	Visualization	Step-by-Step Layout	Single Step Zoom
Ring-by-ring Timelapse			

However, this is a very elaborate way to communicate. Here's the Official Math® terms:

Intuitive Concept	Formal Name	Symbol
X-Ray (split apart)	Take the derivative (derive)	$\frac{d}{dr}$
Time-lapse (glue together)	Take the integral (integrate)	$\int$
Arrow direction	Integrate or derive "with respect to" a variable.	*dr* implies moving along *r*
Arrow start/stop	Bounds or range of integration	$\int_{start}^{end}$
Slice	Integrand (shape being glued together, such as a ring)	Equation, such as $2\pi r$

Let's walk through the fancy names.

The Derivative

The **derivative** is splitting a shape into sections as we move along a path (i.e., X-Raying it). Now here's the trick: although the derivative generates the entire sequence of sections (the black line), we can also extract a *single* one.

Think about a function like $f(x) = x^2$. It's a curve that describes a giant list of possibilities (1, 4, 9, 16, 25, etc.). We can graph the entire curve, sure, or examine the value of f(x) *at* a specific value, like $x = 3$.

The derivative is similar. Officially, it's the entire pattern of sections, but we can zero into a specific one by asking for the derivative *at* a certain value. (The derivative is a function, just like $f(x) = x^2$; if not otherwise specified, we're describing the entire function.)

What do we need to find the derivative? The shape to split apart, and the path to follow as we cut it up (the orange arrow). For example:

- The derivative of a circle *with respect to* the radius creates rings
- The derivative of a circle *with respect to* the perimeter creates slices
- The derivative of a circle *with respect to* the x-axis creates boards

I agree that "with respect to" sounds formal: *Honorable Grand Poombah radius, it is with respect to you that we derive.* Math is a gentleman's game, I suppose.

Taking the derivative is also called "differentiating", because we are finding the difference between successive positions as a shape grows. (As we grow the radius of a circle, the difference between the current disc and the next size up is that outer ring.)

The Integral, Arrows, and Slices

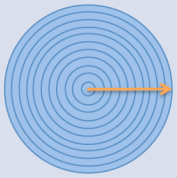
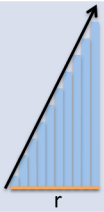
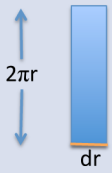
The **integral** is glueing together (time-lapsing) a group of sections and measuring the final result. For example, we glued together the rings (into a "ring triangle") and saw it accumulated to πr^2, aka the area of a circle.

Here's what we need to find the integral:

- **Which direction are we gluing the steps together?** Along the orange line (the radius, in this case)
- **When do we start and stop?** At the start and end of the arrow (we start at 0, no radius, and move to r, the full radius)
- **How big is each step?** Well... each item is a "ring". Isn't that enough?

Nope! We need to be specific. We've been saying we cut a circle into "rings" or "pizza slices" or "boards". But that's not specific enough; it's like a BBQ recipe that says "Cook meat. Flavor to taste."

Maybe an expert knows what to do, but we need more specifics. How large, exactly, is each step (technically called the "integrand")?

Strategy	Visualization	Step-by-Step Layout	Single Step Zoom
Ring-by-ring Timelapse			

Ah. A few notes about the variables:

- If we are moving along the radius r, then dr is the little chunk of radius in the current step
- The height of the ring is the circumference, or $2\pi r$

There's several gotchas to keep in mind.

First, dr is its own variable, and not "d times r". It represents the tiny section of the radius present in the current step. This symbol (dr, dx, etc.) is often separated from the integrand by just a space, and it's assumed to be multiplied (written $2\pi r\ dr$).

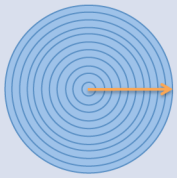
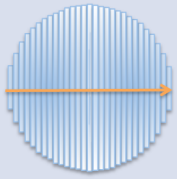
Next, if r is the only variable used in the integral, then dr is assumed to be there. So if you see $\int 2\pi r$ this still implies we're doing the full $\int 2\pi r\ dr$. (Again, if there are two variables involved, like radius and perimeter, you need to clarify which step we're using: dr or dp?)

Last, remember that r (the radius) changes as we time-lapse, starting at 0 and eventually reaching its final value. When we see r in the *context of a step*, it means "the size of the radius at the current step" and not the final value it may ultimately have.

These issues are extremely confusing. I'd prefer we use rdr to indicate an intermediate "r at the current step" instead of a general-purpose "r" that's easily confused with the max value of the radius. I can't change the symbols at this point, unfortunately.

Practicing The Lingo

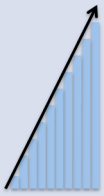
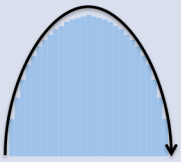
Let's learn to talk like calculus natives. Here's how we can describe our X-Ray strategies:

Intuitive Visualization	Formal Description	Symbol
	derive the area of a circle with respect to the radius	$\frac{d}{dr}Area$
	derive the area of a circle with respect to the perimeter	$\frac{d}{dp}Area$
	derive the area of a circle with respect to the x-axis	$\frac{d}{dx}Area$

Remember, the derivative just splits the shape into (hopefully) easy-to-measure steps, such as rings of size $2\pi r\ dr$. We broke apart our lego set and have pieces scattered on the floor. We still need an integral to glue the parts together and measure the new size. The two commands are a tag team:

- The derivative says: "Ok, I split the shape apart for you. It looks like a bunch of pieces $2\pi r$ tall and dr wide."
- The integral says: "Oh, those pieces resemble a triangle – I can measure that! The total area of that triangle is $\frac{1}{2}base \cdot height$, which works out to πr^2 in this case.".

Here's how we'd write the integrals to measure the steps we've made:

Formal Description	Symbol	Measures Total Size Of
`integrate 2 * pi * r * dr from r=0 to r=r`	$\int_0^r 2\pi r \, dr$	
`integrate [a pizza slice] from [p = min perimeter] to [p = max perimeter]`	$\int_{p=min}^{p=max} (pizza\ slice)\ dp$	
`integrate [a board] from [x = min value] to [x = max value]`	$\int_{x=min}^{x=max} board\ dx$	

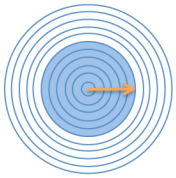
A few notes:

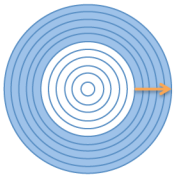
- Often, we write an integrand as an unspecified "pizza slice" or "board" (use a formal-sounding name like $s(p)$ or $b(x)$ if you like). First, we setup the integral, and then we worry about the exact formula for a board or slice.
- Because each integral represents slices from our original circle, we know they will be the same. Gluing any set of slices should always return the total area, right?
- The integral is often described as "the area under the curve". It's accurate, but shortsighted. Yes, we are gluing together the rectangular slices under the curve. But this completely overlooks the preceding X-Ray and Time-Lapse thinking. Why are we dealing with a set of slices vs. a curve in the first place? Most likely, because those slices are easier than analyzing the shape itself (how do you "directly" measure a circle?).

Questions

1) Can you think of another activity which is made *simpler* by shortcuts and notation, vs. written English?
2) Interested in performance? Let's drive the calculus car, even if you can't build it yet.

Question 1: How would you write the integrals that cover half of a circle?



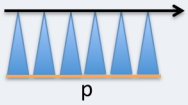


Each should would be similar to:

`integrate [size of step] from [start] to [end] with respect to [path variable]`

(Answer for the first half and the second half. This links to Wolfram Alpha, an online calculator, and we'll learn to use it later on.)

Question 2: Can you find the complete way to describe our "pizza-slice" approach?

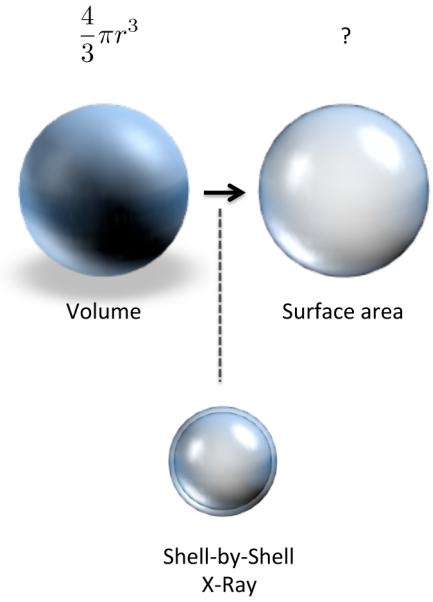
Strategy	Visualization	Step-by-Step Layout	Step Zoom In
Slice-by-slice			

The "math command" should be something like this:

`integrate [size of step] from [start] to [end] with respect to [path variable]`

Remember that each slice is basically a triangle (so what's the area?). The slices move around the perimeter (where does it start and stop?). Have a guess for the command? Here it is, the slice-by-slice description.

Question 3: Can you figure out how to move from volume to surface area?



Assume we know the volume of a sphere is `4/3 * pi * r^3`. Think about the instructions to separate that volume into a sequence of shells. Which variable are we moving through?

`derive [equation] with respect to [path variable]`

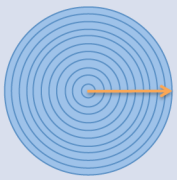
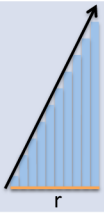
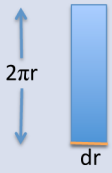
Have a guess? Great. Here's the command to turn volume into surface area.

CHAPTER 5

MUSIC FROM THE MACHINE

In the previous lessons we've gradually sharpened our intuition:

- **Appreciation:** I think it's possible to split up a circle to measure its area
- **Natural Description:** Split the circle into rings from the center outwards, like so:

Strategy	Visualization	Step-by-Step Layout	Single Step Zoom
Ring-by-ring Timelapse			

- **Formal Description**: integrate 2 * pi * r * dr from r=0 to r=r
- **Performance**: (*Sigh*) I guess I'll have to start measuring the area...

Wait! Our formal description is precise enough that a computer can do the work for us:



Whoa! We described our thoughts well enough that a computer did the legwork.

We didn't need to manually unroll the rings, draw the triangle, and find the area (which isn't overly tough in this case, but could have been). We saw what the steps would be, wrote them down, and fed them to a computer: boomshakala, we have the result. (Just worry about the "definite integral" portion for now.)

Now, how about deriviatives, X-Raying a pattern into steps? Well, we can ask for that too:



Similar to above, the computer X-Rayed the formula for area and split it step-by-step as it moved. The result is $2\pi r$, the height of the ring at every position.

Seeing The Language In Action

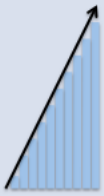
Wolfram Alpha is an easy-to-use tool: the general format for calculus questions is

- `integrate [equation] from [variable=start] to [variable=end]`
- `derive [equation] with respect to [variable]`

That's a little wordy. These shortcuts are closer to the math symbols:

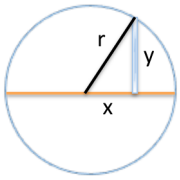
- `\int [equation] dr` - integrate equation (by default, assume we go from $r = 0$ to $r = r$, the max value)
- `d/dr equation` - derive equation with respect to r
- There's shortcuts for exponents (`3^2 = 9`), multiplication (`3 * r`), and roots (`sqrt(9) = 3`)

Now that we have the machine handy, let's try a few of the results we've seen so far:

Computer Description	Math Notation	Intuitive Notion
`integrate 2 * pi * r * dr from r=0 to r=r`	$\int_0^r 2\pi r \, dr$	
`integrate 1/2 * r * dp from p=0 to p=2*pi*r`	$\int_0^{2\pi r} \frac{1}{2} r \, dp$	
`integrate 2 * sqrt(r^2 - x^2) from x=-r to x=r`	$\int_{-r}^{r} 2\sqrt{r^2 - x^2} \, dx$	

Click the formal description to see the computer crunch the numbers. As you might have expected, they all result in the familiar equation for area. A few notes:

- The size of the wedge is $\frac{1}{2} base \cdot height$. The base is dp (the tiny section of perimeter) and the height is r, the distance from the perimeter back to the center.
- The size of board is tricky. In terms of x & y coordinates, we have $x^2 + y^2 = r^2$, by the Pythagorean Theorem:



$$x^2 + y^2 = r^2$$

We solve for the height to get $y = \sqrt{r^2 - x^2}$. We actually need 2 copies of height, because y is the positive distance above the axis, and the board extends above and below. The boards are harder to work with, and it's not just you: Wolfram Alpha takes longer to compute this integral than the others!

The approach so far has been to immerse you in calculus thinking, and gradually introduce the notation. Some of it may be a whirl – which is completely expected. You're sitting at a cafe, overhearing conversation in a foreign language.

Now that you have the sound in your head, we'll begin to explore the details piece-by-piece.

CHAPTER 6

IMPROVING ARITHMETIC AND ALGEBRA

We've intuitively seen how calculus dissects problems with a step-by-step viewpoint. Now that we have the official symbols, let's see how to bring arithmetic and algebra to the next level.

Better Multiplication And Division

Multiplication makes addition easier. Instead of grinding through questions like $2+2+2+2+2+2+2+2+2+2+2+2+2$, we can rewrite it as: 2×13.

Boomshakala. If you wanted 13 copies of a number, just write it like that!

Multiplication makes repeated addition easier (likewise for division and subtraction). But there's a big limitation: we must use identical, average-sized pieces.

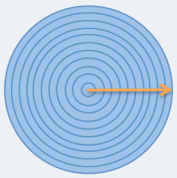
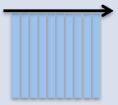
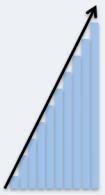
- What's 2×13? It's 13 copies of *the same element*.
- What's 100 / 5? It's 100 split into 5 *equal parts*.

Identical parts are fine for textbook scenarios ("Drive an unwavering 30mph for exactly 3 hours"). The real world isn't so smooth. Calculus lets us accumulate or separate shapes according to their *actual*, not average, amount:

- **Derivatives are better division** that splits a shape along a path (into possibly different-sized slices)
- **Integrals are better multiplication** that accumulates a sequence of steps (which could be different sizes)

Operation	Example	Notes
Division	$\frac{y}{x}$	Split whole into identical parts
Differentiation	$\frac{d}{dx}y$	Split whole into (possibly different) parts
Multiplication	$y \cdot x$	Accumulate identical steps
Integration	$\int y\, dx$	Accumulate (possibly different) steps

Let's analyze our circle-to-ring example again. How does arithmetic/algebra compare to calculus?

Operation	Formula	Diagram
Division	$average\ step = \frac{Area}{radius} = \frac{\pi r^2}{r} = \pi r$	
Differentiation	$actual\ steps = \frac{d}{dr}\pi r^2 = 2\pi r$	
Multiplication	$Area = Average\ step \cdot amount = \pi r \cdot r = \pi r^2$	
Integration	$Area = \int actual\ steps = \int 2\pi r = \pi r^2$	

Division spits back the averaged-sized ring in our pattern. The derivative gives a formula ($2\pi r$) that describes *every* ring (just plug in r). Similarly, multiplication lets us scale up the average element (once we've found it) into the full amount. Integrals let us add up the pattern directly.

Sometimes we want to use the average item, not the fancy calculus steps, because it's a simpler representation of the whole ("What's the average transaction size? I don't need the full list"). That's fine, as long as it's a conscious choice.

Better Formulas

If calculus provides better, more-specific version of multiplication and division, shouldn't we rewrite formulas with it? You bet.

An equation like $distance = speed \cdot time$ explains how to find total distance assuming an average speed. An equation like $distance = \int speed\ dt$ tells us how to find total distance by breaking time into instants (split along the "t" axis), and accumulating the (potentially unique) distance traveled each instant ($speed \cdot dt$).

Algebra	Calculus
$distance = speed \cdot time$	$distance = \int speed\ dt$
$speed = \frac{distance}{time}$	$speed = \frac{d}{dt} distance$
$area = height \cdot width$	$area = \int height\ dw$
$weight = density \cdot length \cdot width \cdot height$	$weight = \iiint density\ dx\ dy\ dz$

Similarly, $speed = \frac{d}{dt} distance$ explains that we can split our trajectory into time segments, and the (potentially unique) amount we moved in that time slice was the speed.

The overused "integrals are area under the curve" explanation becomes more clear. Multiplication, because it deals with static quantities, can only measure the area of rectangles. Integrals let measurements curve and undulate as we go: we'll add their contribution, regardless.

A series of multiplications becomes a series of integrals (called a triple integral). It's beyond this primer, but your suspicion was correct: we can mimic the multiplications and integrate several times in a row.

Math, and specifically calculus, is the language of science because it describes relationships extremely well. When I see a formula with an integral or derivative, I mentally convert it to multiplication or division (with the understanding that will give the "average" element, not the actual one).

Better Algebra

Algebra lets us start with one fact and systematically work out others. Imagine I want to know the area of an unknown square. I can't measure the area, but I overhead someone saying it was 13.3 inches on a side.

Algebra	Thinking process
$Area\ of\ square = ?$	The area of this square is unknown...
$\sqrt{Area} = 13.3$	... but I know the square root.
$\left(\sqrt{Area}\right)^2 = (13.3)^2$	Square both sides...
$Area = 176.89$	... and I can recreate the original area

Remember learning that along with add/subtract/multiply/divide, we could take powers and roots? We added two new ways to transform an equation.

Well, calculus extends algebra with two more operations: integrals and derivatives. Now we can work out the area of a circle, algebra-style:

Algebra + Calculus	Thinking process
$Area\ of\ circle = ?$	The area of a circle is unknown...
$\frac{d}{dr} Area = 2\pi r$	... but I know it splits into rings (along the radius)
$\int \frac{d}{dr} Area = \int 2\pi r$	Integrate both sides...
$Area = \pi r^2$	... and I can recreate the original area

The abbreviated notation helps see the big picture. If the integrand only uses a single variable (as in $2\pi r$), we can assume we're using dr from $r = 0$ to $r = r$. This helps think of integrals and derivatives like squares and square roots: operations that cancel!

It's pretty neat: "gluing together" and "splitting apart" should behave like opposites, right?

With our simpler notation, we can write $\int \frac{d}{dr} Area = Area$ instead of the bulky $\int_0^r \left(\frac{d}{dr} Area\right)\ dr = Area$.

Learning The Rules

With arithmetic, we learned special techniques for combining whole numbers, decimals, fractions, and roots/powers. Even though $3+9=12$, we can't assume $\sqrt{3}+\sqrt{9}=\sqrt{12}$.

Similarly, we need to learn the rules for how integrals/derivatives work when added, multiplied, and so on. Yes, there are fancy rules for special categories (what to do with e^x, natural log, sine, cosine, etc.), but I'm not concerned with that. Let's get extremely comfortable with the basics. The fancy stuff can wait.

CHAPTER 7

SEEING HOW LINES WORK

Let's start by analyzing a fairly simple pattern, a line:

$$f(x) = 4x$$

In everyday terms, we enter an input, x, and get an output, $f(x)$. And in this case, the output is 4 times the input. Suppose we're buying fencing. For every foot we ask for (the input, x), it costs us \$4 (the output, $f(x)$). 3 feet of fence would cost \$12. Fair enough.

Notice the abstract formula $f(x) = 4x$ only considers quantities, not units. We could have written that a foot of fence costs 400 pennies, $f(x) = 400x$, and it's up to *us* to remember it's dollars in the first formula and pennies in the second. Our equations use abstract, unitless quantities, and we have to bring the results back to the real world.

Finding the Derivative Of A Line

The derivative of a pattern, $\frac{d}{dx}f(x)$, is sequence of steps generated as make slices along an input variable (x is the natural choice here). How do we figure out the sequence of steps?

Well, I imagine going to Home Depot and pestering the clerk:

You: I'd like some lumber please. What will it run me? Clerk: How much do you want? You: Um... 1 feet, I think. Clerk: That'll be \$4. Anything else I can help you with? You: Actually, it might be 2 feet. Clerk: That'll be \$8. Anything else I can help you with? You: It might be 3 feet. Clerk: *(sigh)* That'll be \$12. Anything else I can help you with? You: How about 4 feet?

We have a relationship ($f(x) = 4x$) and investigate it by changing the input a tiny bit. We see if there's a change in output (there is!), then we change the input again, and so on.

In this case, it's clear that an additional foot of fencing raises the cost by \$4. So we've just determined the derivative to be a constant 4, right?

Not so fast. Sure, we thought about the process and worked it out, but let's be a little more organized (not every pattern is so simple). Can we describe our steps?

- 1: Get the current output, $f(x)$. In our case, $f(1) = 4$.

- 2: Step forward by dx (1 foot, for example)
- 3: Find the new amount, $f(x+dx)$. In our case, it's $f(1+1) = f(2) = 8$.
- 4: Compute the difference: $f(x+dx) - f(x)$, or 8 - 4 = 4

Ah! The difference between the next step and the current one is the size of our slice. For $f(x) = 4x$ we have:

$$f(x+dx) - f(x) = 4(x+dx) - 4(x) = 4 \cdot dx$$

Increasing length by dx increases the cost by $4 \cdot dx$.

That statement is true, but a little awkward: it talks about the total change. Wouldn't it be better to have a ratio, such as "cost per foot"?

We can extract the ratio with a few shortcuts:

- dx = change in our input
- df = resulting change in our output, $f(x+dx) - f(x)$
- $\frac{df}{dx}$ = ratio of output change to input change

In our case, we have

$$\frac{df}{dx} = \frac{4 \cdot dx}{dx} = 4$$

Notice how we express the derivative as $\frac{df}{dx}$ instead of $\frac{d}{dx}f(x)$. What gives? It turns out there's a few different versions we can use.

Think about the various ways we express multiplication:

- Times symbol: 3×4 (used in elementary school)
- Dot: $3 \cdot 4$ (used in middle school)
- Implied multiplication with parentheses: $(x+4)(x+3)$
- Implied multiplication with a space: $2\pi r\ dr$

The more subtle the symbol, the more we focus on the relationship between the quantities; the more visible the symbol, the more we focus on the computation.

The notation for derivatives is similar:

Derivative Symbol			Mindset
$f'(x)$	$\dot{f}$	$\frac{d}{dx}f$	Think about the resulting step-by-step pattern
	$\frac{df}{dx}$		Think about the actual ratio of changes

Some versions, like $f'(x)$, remind us the sequence of steps is a variation of the original pattern. Notation like $\frac{df}{dx}$ puts us into detail-oriented mode, thinking about the ratio of output change relative to the input change ("What's the cost per additional foot?").

Remember, the derivative is a complete description of all the steps, but it can be evaluated at a certain point to find the step there: What is the additional cost/foot when x = 3? In our case, the answer is 4.

Let's check our findings with the computer:



Nice! Linear relationships are simple enough that we don't need to visualize the steps (but if you do, imagine a 1-d line being split into segments: - - -), and we can work with the symbols directly.

Finding The Integral Of A Constant

Now let's work in the other direction: given the sequence of steps, can we find the size of the original pattern?

In our fence-building scenario, it's fairly straightforward. Solving

$$\frac{df}{dx} = 4$$

means answering "What pattern has an output change of 4 times the input change?"

Well, we've just seen that $f(x) = 4x$ results in $f'(x) = 4$. So, if we're given $f'(x) = 4$, we can guess the original function must have been $f(x) = 4x$.

I'm pretty sure we're right (what else could the integral of 4 be?) but let's check with the computer:



Whoa – there's two different answers (definite and indefinite). Why? Well, there's many functions that could increase cost by \$4/foot! Here's a few:

- Cost = \$4 per foot, or $f(x) = 4x$
- Cost = \$4 + \$4 per foot, or $f(x) = 4 + 4x$
- Cost = \$10 + \$4 per foot, or $f(x) = 10 + 4x$

There could be a fixed per-order fee, with the fence cost added in. All the equation $f'(x) = 4$ says is that each *additional* foot of fencing is \$4, but we don't know the starting conditions.

- The **definite integral** tracks the accumulation of a set amount of slices. The range can be numbers, such as $\int_0^{13} 4$, which measures the slices from x=0 to x=13 ($13 \times 4 = 52$). If the range includes a variable (0 to x), then the accumulation will be an equation ($4x$).
- The **indefinite integral** finds the actual formula that created the pattern of steps, not just the accumulation in that range. It's written with just an integral sign: $\int f(x)$. And as we've seen, the possibilities for the original function should allow for a starting offset of C.

The notation for integrals can be fast-and-loose, and it's confusing. Are we looking for an accumulation, or the original function? Are we leaving out dx? These details are often omitted, so it's important to feel what's happening.

The Secret: We Can Work Backwards

The little secret of integrals is that we don't need to solve them directly. Instead of trying to glue slices together to find out their area, we just learn to *recognize* the derivatives of functions we've already seen.

If we've know the derivative of 4x is 4, then if someone asks for the *integral* of 4, we can respond with "4x" (plus C, of course). It's like memorizing the squares of numbers, not the square roots. When someone asks for the square root of 121, dig through and remember that $11 \times 11 = 121$.

An analogy: Imagine an antiques dealer who knows the original vase just from seeing a pile of shards.

How does he do it? Well, he takes replicas in the back room, drops them, and looks at the pattern of pieces. Then he comes to your pile and says "Oh, I think this must be a Ming Dynasty Vase from the 3rd Emperor."

He doesn't try to glue your pile back together – he's just seen that exact vase break before, and your pile looks the same!

Now, there may be piles he's never seen, that are difficult or impossible to recognize. In that case, the best we can do is to just add up the pieces (with a computer, most likely). We might determine the original vase weighed 13.78 pounds. That's a data point, fine, but it's not as nice as knowing *what* the vase was before it shattered.

This insight was never really explained to me: it's painful to add up (possibly changing) steps directly, especially when the pattern gets complicated. So, just learn to recognize the pattern from the derivatives we've already seen.

Getting To Better Multiplication

Gluing together equally-sized steps looks like regular multiplication, right? You bet. If we wanted 3 steps (0 to 1, 1 to 2, 2 to 3) of size 2, we might write:

$$\int_0^3 2 \, dx = 6$$

Again, this is a fancy way of saying "Accumulate 3 steps of size 2: what do you get in total?". We are time-lapsing a sequence of equal changes.

Now, suppose someone asks you to add $2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2$. You might say: *Geez, can't you write it more simply? You know, something like:*

$$\int_0^{13} 2 \, dx = 26$$

Creating The Abstract Rules

Have an idea how linear functions behave? Great. We can make a few abstract rules – like working out the rules of algebra for ourselves.

If we know our output is a scaled version of our input ($f(x) = ax$) then

$$\frac{d}{dx} a \cdot x = a$$

and

$$\int a = ax + C$$

That is, the ratio of each output step to each input step is a constant a (4, in our examples above). And now that we've broken the vase, we can work backwards: if we accumulate steps of size a, they must have come from a pattern similar to $a \cdot x$ (plus C, of course).

Notice how I wrote $\int a$ and not $\int a\ dx$ – I wanted to focus on a, and not details like the width of the step (dx). Part of calculus is learning to expose the right amount of detail.

One last note: if our output does not react *at all* to our input (we'll charge you a constant $2 no matter how much you buy... including nothing!) then "steps" are a constant 0:

$$\frac{d}{dx}a = 0$$

In other words, there is no difference in the before-and-after measurement. Now, a pattern may have an *occasional* zero slice, if it stands still for a moment. That's fine. But if *every* slice it zero, it means our pattern never changes.

There are a few subtleties down the road, but let's first learn "Me want food" and then "Verily, I hunger".

CHAPTER 8

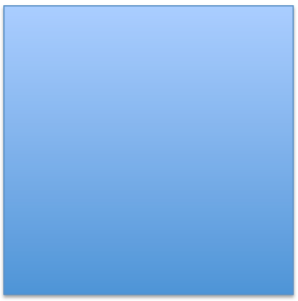
PLAYING WITH SQUARES

We've seen how lines behave. Now let's hone our skills with a more complex function like $f(x) = x^2$.

This scenario is more complex, so let's visualize it and walk through the properties.

Imagine you're building a square garden, to plant veggies and enjoy cucumbers in a few months. You're not sure how large to make it. Too small, and there's not enough food, but too large, and you'll draw the attention of the veggie mafia.

You'll build the garden incrementally, foot-by-foot, until it looks right. Let's say you build it slowly, starting from scratch and growing to a 10×10 plot:



What do we have? To the untrained eye, we've built a 10×10 garden, which uses 40 feet of perimeter fencing ($10 \cdot 4$) and 100 square feet of topsoil ($10 \cdot 10$). (Pretend topsoil is sold by the square foot, assuming a standard thickness.)

Bring On The Calculus

That's it? The analysis is just the current perimeter and square footage? No way.

By now, you should be clamoring to use X-Ray and Time-Lapse vision to see what's happening under the hood. Why settle for a static description when we can know the step-by-step description too?

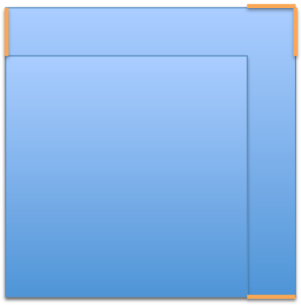
We can analyze the behavior of the perimeter pretty easily:

$$Perimeter = 4 \cdot x$$

$$\frac{d}{dx} Perimeter = 4$$

The change in perimeter ($\frac{dP}{dx}$) is a constant 4. For every 1-foot increase in x, we have a 4-foot jump in the perimeter.

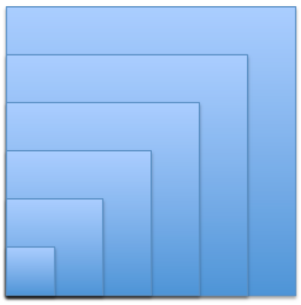
We can visualize this process. As the square grows, we push out the existing sides, and just add 4 corner pieces (in yellow):



The visual is nice, but not required. After our exposure to lines, we should glance at an equation like $p = 4x$ and realize that p jumps by 4 when x jumps by 1.

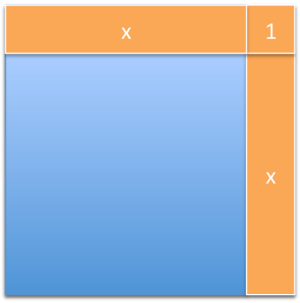
Changing Area

Now, how does area change? Since squares are fairly new, let's X-Ray the shape as it grows:



We can write out the size of each jump, like so:

Now that's interesting. The gap from 02 to 12 is 1. The gap from 12 to 22 is 3. The gap from 22 to 32 is 5. And so on – the odd numbers are sandwiched between the squares! What's going on?



Ah! Growing to the next-sized square means we've added a horizontal and vertical strip (x + x) and a corner piece (1). If we currently have a square with side x, the jump to the next square is 2x + 1. (If we have a 5×5 square, getting to a 6×6 will be a jump of 2·5 + 1 = 11. And yep, 36 - 25 = 11.)

Again, the visualization was nice, but it took effort. Algebra can simplify the process.

In this setup, if we set our change to dx = 1, we get

$$df = f(x+1) - f(x) = (x+1)^2 - x^2 = (x^2 + 2x + 1) - x^2 = 2x + 1$$

Algebra predicted the size of the slices without a hitch.

Integrals and the Veggie Mafia

The derivative takes a shape and get the slices. Can we work backwards, from the slices to the shape? Let's see.

Suppose the veggie mafia spies on your topsoil and fencing orders. They can't see your garden directly, but what can they deduce from your purchases?

Let's say they observe a constant amount of fencing being delivered (4, 4, 4, 4...) but *increasing* orders of topsoil (1, 3, 5, 7, 9, 11...). What can they work out?

A low-level goon might just add up the total amount accumulated (the definite integral): "Heya boss, looks like they've built some garden with a total perimeter of 40-feet, and total area of 100 square feet."

But that's not good enough! The goon doesn't know the shape you're trying to build. He saw order after order go by without noticing the deeper pattern.

The crime boss is different: he wants the *indefinite* integral, the pattern you are following. He's savvy enough to track the pattern as the orders come in: "The area is increasing 1, 3, 5, 7... that's following a 2x + 1 area increase!"

Now, there are likely many shapes that could grow their area by $2x+1$. But, combined with a constant perimeter increase of 4, he suspects you're making a square garden after a few deliveries.

How does the godfather do it? Again, by working backwards. He's split apart enough shapes (triangles, squares, rectangles, etc.) that he has a large table of before-and-afters, just like the antiques dealer.

When he sees a change of $2x+1$, a square (x^2) is a strong candidate. Another option might be a right triangle with sides x and $2x$. Its area equation is $\frac{1}{2}x \cdot 2x = x^2$, so the area would change the same as a square.

And when he sees a perimeter change of 4, he knows the perimeter must be $4x$. Ah! There aren't too many shapes with both properties: a square is his guess. Part of learning calculus is getting familiar with the typical origins of the patterns you see, so you can quickly reverse-engineer them.

Suppose they see your orders change: your fencing deliveries drop to (2, 2, 2, 2...) and your topsoil orders change to (20, 20, 20, 20). What's going on? Make a guess if you like.

Ready?

The veggie boss figures you've moved to a *rectangular* garden, with one side determined by x, and the other side a fixed 20-feet. Does this guess work?

$$f(x) = 20x$$

$$Perimeter = 20 + 20 + x + x = 40 + 2x$$

$$Area = 20x$$

$$\frac{d}{dx} Perimeter = 2$$

$$\frac{d}{dx} Area = 20$$

Wow, it checks out. No wonder he's the godfather.

Lastly, what if the godfather saw topsoil orders of (5, 7, 9, 11, 13)? He might assume you're still building a square ($2x+1$ pattern), but you *started* with a 2×2 garden already made. Your first jump was 5, which would have happened if x was already 2 ($2x+1=5$).

The mob boss is a master antiques dealer: he sees the pattern of shards you're bringing in and quickly determines the original shape (indefinite integral). The henchman can only tell you the running totals so far (definite integral).

Wrapping It All Up

It looks like we're ready for another rule, to explain how squares change. If we leave dx as it is, we can write:

$$\frac{d}{dx} x^2 = \frac{f(x+dx) - f(x)}{dx} = \frac{(x+dx)^2 - (x)^2}{dx}$$

$$= \frac{x^2 + 2x \cdot dx + (dx)^2 - x^2}{dx} = \frac{2x \cdot dx + (dx)^2}{dx} = 2x + dx$$

Ok! That's the abbreviated way of saying "Grow by two sides and the corner". Let's plug this into the computer to check:



Uh oh! We hand-computed the derivative of x^2 as $2x + dx$ (which is usually $2x + 1$), but the computer says it's just $2x$.

But isn't the difference from 42 to 52 exactly 25 - 16 = 9, and not 8? What happened to that corner piece? *The mystery continues in the next episode.*

CHAPTER 9

WORKING WITH INFINITY

Last time, we manually worked on the derivative of x^2 as $2x+1$. But the official derivative, according to the calculator, was $2x$. What gives?

The answer relies on the concept of infinite accuracy. Infinity is a fascinating and scary concept – there are entire classes (Analysis) that study it. We'll avoid learning the nuances of every theory: our goal is a practical understanding how infinity can help us work out the rules of calculus.

Insight: Sometimes Infinity Can Be Measured

Here's a quick brainteaser for you. Two friends are 10 miles apart, moving towards each other at 5mph each. A mosquito files quickly between them, touching one person, then the other, on and on, until the friends high-five and the mosquito is squished.

Let's say the mosquito travels a zippy 20mph as it goes. Can you figure out how far it flew before its demise?



...

Yikes. This one is tricky: once the mosquito leaves the first person, touches the second, and turns around... the first person has moved closer! We have an infinite number of ever-diminishing distances to add up. The question seems painfully difficult to solve, right?

Well, how about this reasoning: from the perspective of the people walking, they're going to walk for an hour total. After all, they start 10 miles apart, and the gap shrinks at 10 miles per hour (5mph + 5mph). Therefore, the mosquito must be flying for an hour, and go 20 miles.

Whoa! Did we just find the outcome of a process with an infinite number of steps? I think so!

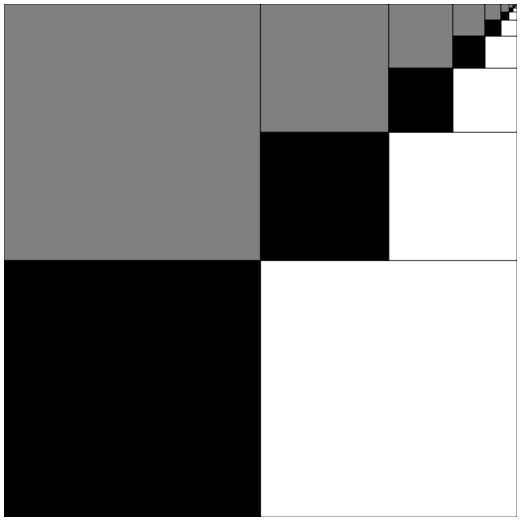
Splitting A Whole Into Infinite Parts

It's time to turn our step-by-step thinking into overdrive. Can we think about a finite shape being split into infinite parts?

- In the beginning of the course, we saw a circle could be split into rings. How many? Well, an infinite number!
- A number line can be split into an *infinite* number of neighboring points. How many decimals would you say there are between 1.0 and 2.0?
- The path of a mosquito can be seen as a whole, or a journey subdivided into an infinite number of segments

When we have two viewpoints (the mosquito, and the walkers), we can pick the one that's easier to work with. In this case, the walker's holistic viewpoint is simpler. With the circle, it's easier to think about the rings themselves. It's nice to have both options available.

Here's another example: can you divide a cake into 3 equal portions, by only cutting into quarters:



It's a weird question... but possible! Cut a entire cake into quarters. Share 3 pieces and leave 1. Cut the remaining piece into quarters. Share 3 pieces, leave 1. Keep repeating this process: at every step, everyone has received an equal share, and the remaining cake will be split evenly as well. Wouldn't this plan maintain an even split among 3 people?

We're seeing the intuition behind infinite X-Ray and Time-lapse vision: zooming in to turn a whole into an infinite sequence. At first, we might think dividing something into infinite parts requires each part to be nothing. But, that's not right: the number line can be subdivided infinitely, yet there's a finite gap between 1.0 and 2.0.

Two Fingers Pointing At The Same Moon

Why can we understand variations of the letter A, even when pixelated?

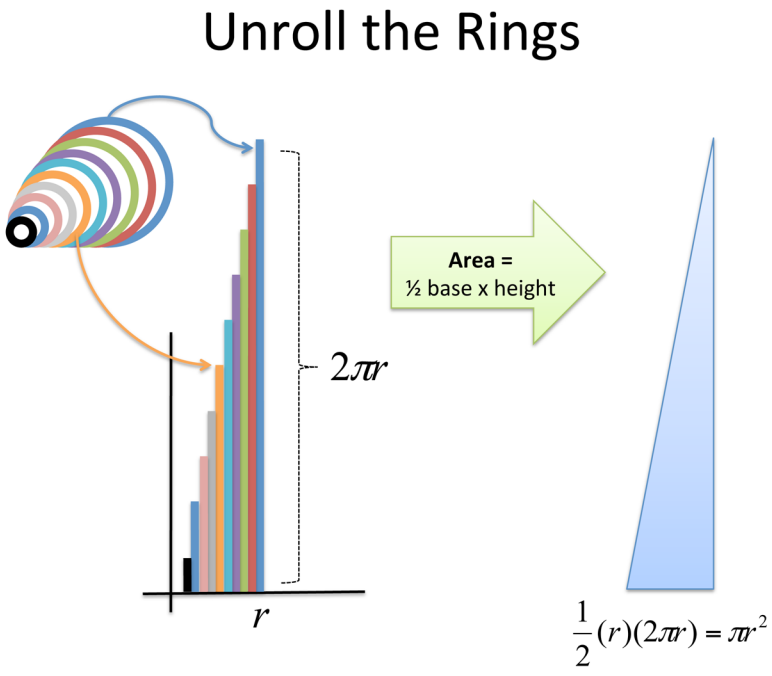


Even though the rendering is different, we see the **idea being pointed to**. All three versions, from perfectly smooth to jagged, create the same letter A in our heads (or, are you unable to read words when written out with rectangular pixels?). An infinite sequence can point to the same result we'd find if we took it all at once.

In calculus, there are detailed rules about how to find what result an infinite set of steps points to. And, there are certain sequences that cannot be worked out. But, for this primer, we'll deal with functions that behave nicely.

We're used to jumping between *finite* representations of the same idea (5 = V = |||||). Now we're seeing we can convert between a finite and *infinite* representation of an idea, similar to $\frac{1}{3} = .333\ldots = .3 + .03 + .003 + \ldots$.

When we turned a circle into a ring-triangle, we said "The infinitely-many rings in our circle can be turned into the infinitely-many boards that make up a triangle. And the resulting triangle is easy to measure."



Today's goal isn't to become experts with infinity. It's to intuitively appreciate a **practical conclusion**: a sequence of infinitely many parts can still be measured, and reach the same conclusions as analyzing the whole. They're just two different descriptions of the same idea.

CHAPTER 10

THE THEORY OF DERIVATIVES

The last lesson showed that an infinite sequence of steps could lead to a finite conclusion. Let's put it into practice, and see how breaking change into infinitely small parts can point to the the true amount.

Analogy: Measuring Heart Rates

Imagine you're a doctor trying to measure a patient's heart rate while exercising. You put a guy on a treadmill, strap on the electrodes, and get him running. The machine spit out 180 beats per minute. That must be his heart rate, right?

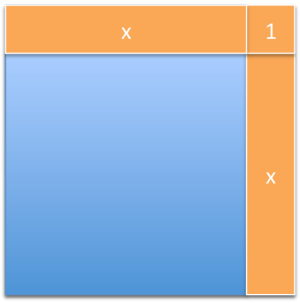
Nope. That's his heart rate *when observed by doctors and covered in electrodes*. Wouldn't that scenario be stressful? And what if your Nixon-era electrodes get tangled on themselves, and tug on his legs while running?

Ah. We need the electrodes to get *some* measurement. But, right afterwards, we need to remove the effect of the electrodes themselves. For example, if we measure 180 bpm, and knew the electrodes added 5 bpm of stress, we'd know the true heart rate was 175.

The key is making the knowingly-flawed measurement, to get a reading, then correcting it as if the instrument wasn't there.

Measuring the Derivative

Measuring the derivative is just like putting electrodes on a function and making it run. For $f(x) = x^2$, we stick an electrode of +1 onto it, to see how it reacted:



The horizontal stripe is the result of our change applied along the top of the shape. The vertical stripe is our change moving along the side. And what's the corner?

It's part of the horizontal change interacting with the vertical one! This is an electrode getting tangled in its own wires, a measurement artifact that needs to go.

Throwing Away The Artificial Results

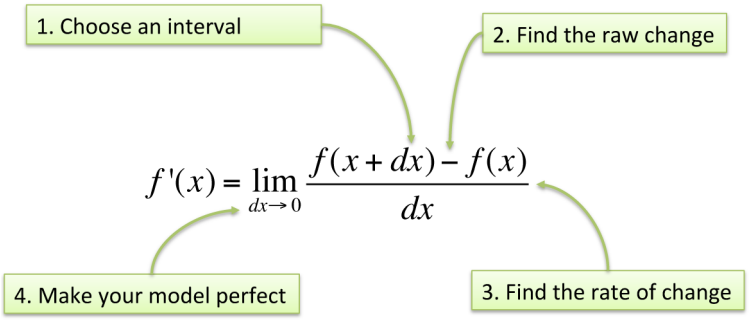
The founders of calculus intuitively recognized which components of change were "artificial" and just threw them away. They saw that the corner piece was the result of our test measurement interacting with itself, and shouldn't be included.

In modern times, we created official theories about how this is done:

- Limits: We let the measurement artifacts get smaller and smaller until they effectively disappear (cannot be distinguished from zero).
- Infinitesimals: Create a new type of number that lets us try infinitely-small change on a separate, tiny number system. When we bring the result back to our regular number system, the artificial elements are removed.

There are entire classes dedicated to exploring these theories. The practical upshot is realizing *how* to take a measurement and throw away the parts we don't need.

Here's how the derivative is defined using limits:



Step	Example
Prereq: Start with a function to study	$f(x) = x^2$
1: Change the input by dx, our test change	$f(x+dx) = (x+dx)^2 = x^2 + 2x \cdot dx + (dx)^2$
2: Find the resulting change in output, df	$f(x+dx) - f(x) = 2x \cdot dx + (dx)^2$
3: Find $\frac{df}{dx}$	$\frac{2x \cdot dx + (dx)^2}{dx} = 2x + dx$
4: Throw away the measurement artifacts	$2x + dx \overset{dx\,=\,0}{\Longrightarrow} 2x$

Wow! We found the official derivative for $\frac{d}{dx}x^2$ on our own:



Now, a few questions:

- **Why do we measure $\frac{df}{dx}$, and not the actual change df?** Think of df as the raw change that happened as we made a step. The ratio is helpful because it normalizes comparisons, showing us how much the output reacts to the input. Now, sometimes it can be helpful to isolate the actual change that happened in an interval, and rewrite $\frac{df}{dx} = 2x$ as $df = 2x \cdot dx$.
- **How do we set dx to 0?** I see dx as the size of the instrument used to measure the change in a function. After we have the measurement with a real instrument $(2x + dx)$, we figure out what the measurement would be if the instrument wasn't there to interfere $(2x)$.
- **But isn't the $2x + 1$ pattern correct?** The *integers* have the squares 0, 1, 4, 9, 16, 25 which has the difference pattern 1, 3, 5, 7, 9. Because the integers must use a fixed interval of 1, using $dx = 1$ and keeping it around is a perfectly accurate way to measure how they change. However, decimals don't have a fixed gap, so $2x$ is the best result for how fast the change between 2^2 and $(2.0000001)^2$ is happening. (Except, replace 2.0000001 with the number *immediately* following 2.0, whatever that is).
- **If there's no '+1', when does the corner get filled in?** The diagram shows how area grows in the presence of a crude instrument, dx. It isn't a decree on how more area is added. The actual change of 2x is the horizontal and vertical strip. The corner represents a change from the vertical strip interfering with the horizontal one. It's area, sure, but it won't be added on *this* step.I imagine a square that grows by two strips, melts to absorb the area (forming a larger square), then grows again, then melts, and so on. The square isn't "staying still" long enough to have the horizontal and vertical extensions interact.

Practical conclusion: We've can start with a knowingly-flawed measurement $(f'(x) \sim 2x + dx)$, and deduce the result it points to $(f'(x) = 2x)$. The theories of exactly *how* we throw away dx aren't necessary to master today. The key is realizing there are measurement artifacts that can be removed when modeling how pattern changes.

(Still shaky about exactly how dx can appear and disappear? Good. This question took mathematicians decades to figure out. Here's a deeper discussion of how the theory works, but remember this: When measuring, ignore the effect of the instrument.)

CHAPTER 11

THE FUNDAMENTAL THEOREM OF CALCULUS (FTOC)

The Fundamental Theorem of Calculus is the big aha! moment, and something you might have noticed all along:

- X-Ray and Time-Lapse vision let us see an existing pattern as an accumulated sequence of changes
- The two viewpoints are opposites: X-Rays break things apart, Time-Lapses put them together

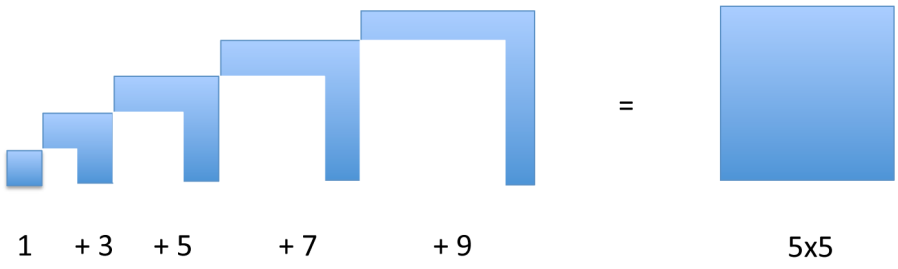
This might seem "obvious", but it's only because we've explored several examples. Is it truly obvious that we can separate a circle into rings to find the area? The Fundamental Theorem of Calculus gently reminds us we have a few ways to look at a pattern. (*"Might I suggest the ring-by-ring viewpoint? Makes things easier to measure, I think."*)

Part 1: Shortcuts For Definite Integrals

If derivatives and integrals are opposites, we can "shortcut" the manual accumulation process involved in definite integrals.

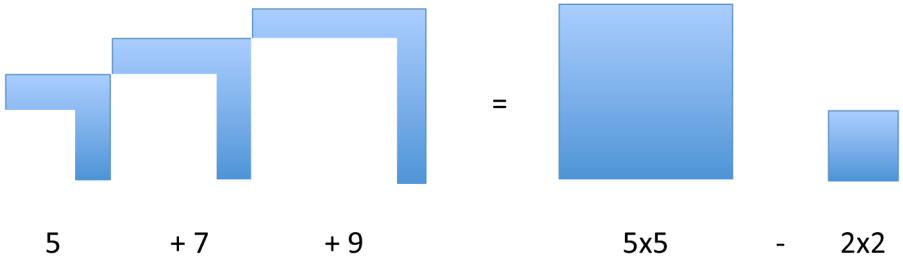
For example, what is 1 + 3 + 5 + 7 + 9? (Let's accumulate discrete steps because they're simpler to visualize.)

The hard way, using the definite integral, is to start cranking through the addition. The *easy* way is to realize that pattern came from a growing square. We know the last change (+9) happened at x=4, so we've built up to a 5×5 square. Therefore, the sum of the entire sequence is 25:



Neat! If we have the original pattern, we have a shortcut to measure the size of the steps.

How about a partial sequence like 5 + 7 + 9? Well, just take the total accumulation and subtract the part we're missing (in this case, the missing 1 + 3 represents a missing 2×2 square).



And yep, the sum of the partial sequence is: 5×5 - 2×2 = 25 - 4 = 21.

I hope the strategy clicks for you: we avoid manually computing the definite integral by finding the original pattern.

Here's the first part of the FTOC in fancy language. If we have pattern of steps and the original pattern, the shortcut for the definite integral is:

$$\int_a^b steps(x)dx = Original(b) - Original(a)$$

Intuitively, I read this as "Adding up all the changes from a to b is the same as getting the difference between a and b". Formally, you'll see $f(x) = steps(x)$ and $F(x) = Original(x)$, which I think is confusing. Label the steps as steps, and the original as the original.

Why is this cool? The definite integral is a gritty mechanical computation, and the indefinite integral is a nice, clean formula. Just take the difference between the endpoints to know the net result of what happened in the middle! (That makes sense, right?)

Part 2: Finding The Indefinite Integral

Ok. Part 1 said that if we have the original function, we can skip the manual computation of the steps. But how do we find the original?

FTOC Part Deux to the rescue!

Let's pretend there's *some* original function (currently unknown) that tracks the accumulation:

$$Accumulation(x) = \int_a^b steps(x)dx$$

The FTOC says the *derivative* of that magic function will be the steps we have:

$$Accumulation'(x) = steps(x)$$

Now we can work backwards. If we can find some random function, take its derivative, notice that it matches the steps we have, we can use that function as our original!

Skip the painful process of thinking about *what* function could make the steps we have. Just take a bunch of them, break them, and see which matches up. It's our vase analogy, remember? The FTOC gives us "official permission" to work backwards. In my head, I think "The next step in the total accumuation is our current amount! That's why the derivative of the accumulation matches the steps we have."

Technically, a function whose derivative is equal to the current steps is called an anti-derivative (One anti-derivative of 2 is 2x; another is 2x + 10). The FTOC tells us any anti-derivative will be the original pattern (+C of course).

This is surprising – it's like saying everyone who *behaves* like Bill Cosby *is* Bill Cosby. But in calculus, if a function has steps that match the ones we're looking at, it's the original source.

The practical conclusion is **integration and differentiation are opposites**. Have a pattern of steps? Integrate to get the original. Have the original? Differentiate to get the pattern of steps. Jump back and forth as many times as you like.

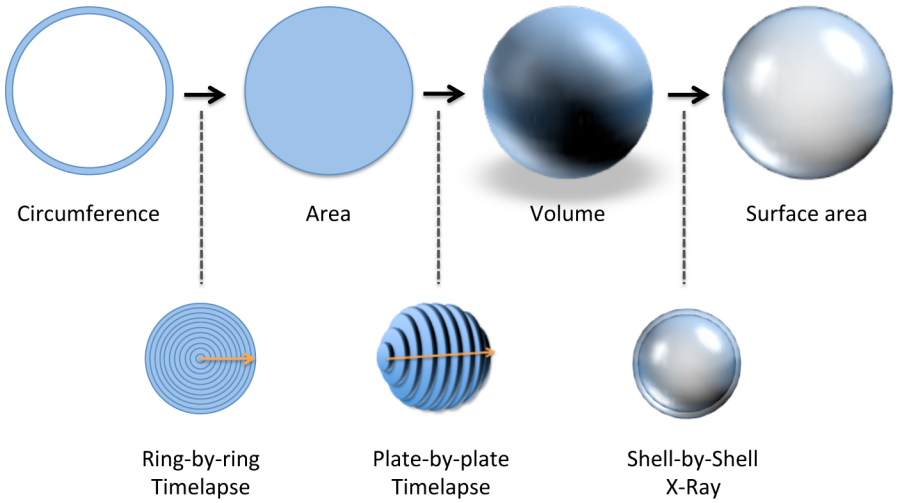
Next Steps

Phew! This was a theory-heavy set of lessons. It was hopefully enough to make the connections click. I'm most interested in developer *metaphorical* understanding, and leave the detailed proofs to a follow-up. No need to recite the recipe before tasting the meal.

The key insights from:

- **Infinity:** A finite result can be viewed with a sequence of infinite steps
- **Derivatives:** We can take a knowingly-flawed measurement and morph it into the ideal one
- **Fundamental Theorem Of Calculus:** We can use the original function as a shortcut to track the intermediate steps

In the upcoming lessons, we'll work through a few famous calculus rules and applications. The ultimate goal will be to work out, for ourselves, how to make this happen:



We have an intuitive understanding the sequence above is possible. By the end, you'll be able to play the music on your own.

CHAPTER 12

THE BASIC ARITHMETIC OF CALCULUS

Remember learning arithmetic? After seeing how to multiply small numbers, we learned how to multiply numbers with several digits:

$$13 \times 15 = (10+3)(10+5) = 100+30+50+15$$

We can't just combine the first digits (10×10) and the second (3×5) and call it done. We have to walk through the cross-multiplication.

Calculus is similar. If we have the whole function, we can blithely say that $f(x)$ has derivative $f'(x)$. But that isn't illuminating, or explaining what happens behind the scenes.

If we can describe our function in terms of a building block x (such as $f(x) = 3x^2 + x$), then we should be able to find the derivative, the pattern of changes, in terms of that same building block. If we have two types of building blocks ($f = a \cdot b$), we'll get the derivative in terms of those two building blocks.

Here's the general strategy:

- Imagine a scenario with a few building blocks ($area = length \cdot width$)
- Let every component change
- Measure the change in the overall system
- Remove the measurement artifacts (our instruments interfering with each other)

Once we know how systems break apart, we can reverse-engineer them into the integral (yay for the FTOC!).

Addition

Let's start off easy: how does a system with two added components behave?

In the real world, this could be sending two friends (Frank and George) to build a fence. Let's say Frank gets the wood, and George gets the paint. What's the total cost?

$$\text{Total} = \text{Frank's cost} + \text{George's cost}$$

$$t(x) = f(x) + g(x)$$

The derivative of the entire system, $\frac{dt}{dx}$, is the cost per additional foot. Intuitively, we suspect the total increase is the sum of the increases in the parts:

$$\frac{dt}{dx} = \frac{df}{dx} + \frac{dg}{dx}$$

That relationship makes sense, right? Let's say Frank's cost is \$3/foot for the wood, and George adds \$0.50/foot for the paint. If we ask for another foot, the total cost will increase by \$3.50.

Here's the math for that result:

- Original: $f + g$
- New: $(f + df) + (g + dg)$
- Change: $(f + df) + (g + dg) - (f + g) = df + dg$

In my head, I imagine x, the amount you requested, changing silently in a corner. This creates a visible change in f (size df) and g (size dg), and we see the total change as $df + dg$.

It seems we should just combine the total up front, writing $total = 3.5x$ not $total = f(x) + g(x) = 3x + 0.5x$. Normally, we would simplify an equation, but it's sometimes helpful to list every contribution (total = base + shipping + tax). In our case, we see most of the increase is due to the Frank (the cost of wood).

Remembering the derivative is the "per dx" rate, we write:

$$\frac{d}{dx}\big(f(x) + g(x)\big) = \frac{df}{dx} + \frac{dg}{dx}$$

But ugh, look at all that notation! Let's trim it down:

- Write f instead of $f(x)$. We'll assume a single letter is an entire function, and by the Third Edict of The Grand Math Poombahs, our functions will use a parameter x.
- We'll express the derivative using a single quote (f'), not with a ratio ($\frac{df}{dx}$). We're most interested in the relationship between the parts (addition), not the gritty details of the parts themselves.

So now the addition rule becomes:

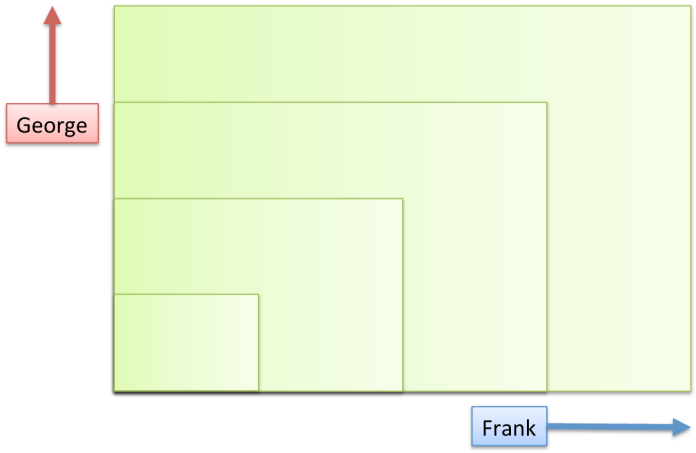
$$(f + g)' = f' + g'$$

Much better! Here's how I read it: Take a system made of several parts: $(f + g)$. The change in the overall system, $(f + g)'$, can be found by adding the change from each part.

Multiplication

Let's try a tricker scenario. Instead of inputs that are added (almost oblivious to each other), what if they are multiplied?

Suppose Frank and George are making a *rectangular* garden for you. Frank handles the width and George takes care of the height. Whenever you clap, they move... but by different amounts!



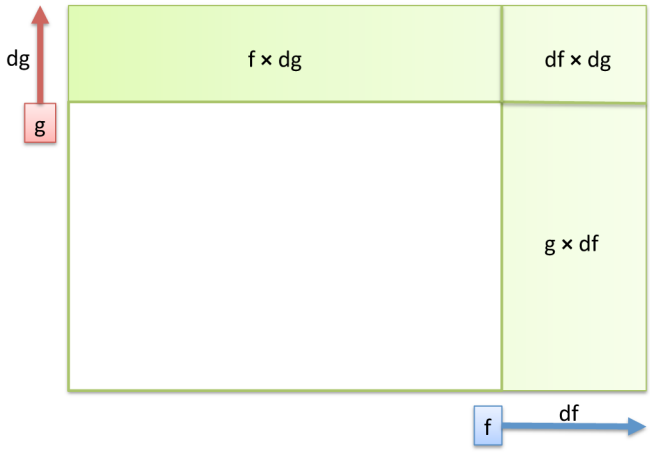
Frank's steps are 3-feet long, but George's are only 2-feet long (zookeeping accident, don't ask). How can we describe the system?

$$area = width \cdot height = f(x) \cdot g(x)$$

$$f(x) = 3x$$

$$g(x) = 2x$$

We have linear parts, so the derivatives are simple: $f'(x) = 3$ and $g'(x) = 2$. What happens on the next clap?



Looks familiar! We have a horizontal strip, a vertical strip, and a corner piece. We can work out the amounts with algebra:

- Original: $f \cdot g$
- New: $(f + df) \cdot (g + dg) = (f \cdot g) + (f \cdot dg) + (g \cdot df) + (df \cdot dg)$

- Change: $f \cdot dg + g \cdot df + df \cdot dg$

Let's see this change more closely:

- The horizontal strip happened when f changed (by df), and g was the same value
- The vertical strip was made when g changed (by dg), and f was the same value
- The corner piece ($df \cdot dg$) happened when the change in one component (df) interacted with the change in the other (dg)

The corner piece is our sample measurement getting tangled on itself, and should be removed. (If we're *forced* to move in whole units, then the corner is fine. But most real-world systems can change continuously, by any decimal number, and we want the measurement artifacts removed.)

To find the total change, we drop the $df \cdot dg$ term (interference between the changes) and get:

$$f \cdot dg + g \cdot df$$

I won't let you forget the derivative is on a "per dx" basis, so we write:

$$\frac{total\ change}{dx} = f\frac{dg}{dx} + g\frac{df}{dx}$$

$$(f \cdot g)' = f \cdot g' + g \cdot f'$$

There is an implicit "x" changing off in the distance, which makes f and g move. We hide these details to make the notation simpler.

In English: Take a scenario with multiplied parts. As they change, and continue to be multiplied, add up the new horizontal and vertical strips that are formed.

Let's try out the rule: if we have a 12×8 garden and increment by a whole step, what change will we see?

In this case, we'll use the discrete version of the rule since we're forced to move as a whole step:

- Vertical strip: $f \cdot dg = 12 \cdot 2 = 24$
- Horizontal strip $g \cdot df = 8 \cdot 3 = 24$
- Corner piece: $df \cdot dg = 3 \cdot 2 = 6$
- Total change: $24 + 24 + 6 = 54$

Let's test it. We go from 12×8 (96 square feet) to 15×10 (150 square feet). And yep, the area increase was 150 - 96 = 54 square feet!

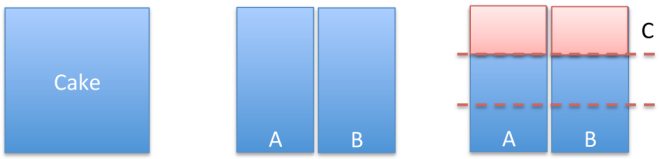
Simple Division (Inverses)

Inverses can be tough to visualize: as x gets bigger, $\frac{1}{x}$ gets smaller. Let's take it slow.

Suppose you're sharing a cake with Frank. You've just cut it in half, about to take a bite and... George shuffles in. He looks upset, and you're not about to mention the fresh set of claw marks.

But you've just the cake in half, what can you do?

Cut it again. You and Frank can cut your existing portion in thirds, and give George a piece:



Neat! Now everyone has 1/3 of the total. You gave up 1/3 of your amount (1/2), that is, you each gave George 1/6 of the total.

Time to eat! But just as you're about to bite in... the veggie godfather walks in. Oh, he'll *definitely* want a piece. What do you do?

Cut it again. Everyone smooshes together their portion, cuts it in *fourths*, and hands one piece to the Don. The cake is split evenly again.

This is step-by-step thinking applied to division:

- Your original share is $\frac{1}{x}$ (when x=2, you have 1/2)
- Someone walks in
- Your new share becomes $\frac{1}{x+1}$

How did your amount of cake change? Well, you took your original slice ($\frac{1}{x}$), cut it into the new number of pieces ($\frac{1}{x+1}$), and gave one away (the change is negative):

$$\frac{1}{x} \cdot \frac{-1}{x+1} = \frac{-1}{x(x+1)}$$

We can probably guess that the $+1$ is a measurement artifact because we forced an integer change in x. If we call the test change dx, we can find the difference between the new amount ($\frac{1}{x+1}$) and the original ($\frac{1}{x}$):

$$\frac{1}{x+dx} - \frac{1}{x} = \frac{x}{x(x+dx)} - \frac{x+dx}{x(x+dx)} = \frac{-dx}{x(x+dx)}$$

After finding the total change (and its annoying algebra), we divide by dx to get the change on a "per dx" basis:

$$\frac{-1}{x(x+dx)}$$

Now we remove the leftover dx, the measurement artifact:

$$\frac{-1}{x(x+0)} = -\frac{1}{x^2}$$

Phew! We've found how an 1/x split changes as more people are added.

Let's try it out: You are splitting a $1000 bill among 5 people. A sixth person enters, how much money do you save?

You'll personally save $1/5 \cdot 1/6 = 1/30$ of the total cost (cut your share into 6 pieces, give the new guy one portion to pay). That's about 3%, or 30. Not bad for a quick calculation!

Let's work it backwards: how large is our group when we're saving about $100 per person? Well, $100 is 1/10 of the total. Since $\frac{1}{3^2} \sim \frac{1}{10}$, we'll hit that savings rate around x=3 people.

And yep, going from 3 to 4 people means each person's share goes from $333.33 to $250 – about $100. Not bad! (If we added people fractionally, we could hit the number exactly.)

Questions

We didn't explicitly talk about scaling by a constant, such as finding the derivative of $f(x) = 3x$. Can you use the product rule to figure out how it changes? (Hint: imagine a rectangle with a fixed 3 for one side, and x for the other).

Now, how about the addition rule? How would $f(x) = x + x + x$ behave?

CHAPTER 13

PATTERNS IN THE RULES

We've uncovered the first few rules of calculus:

$$(f+g)' = f' + g'$$
$$(f \cdot g)' = f \cdot g' + g \cdot f'$$
$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

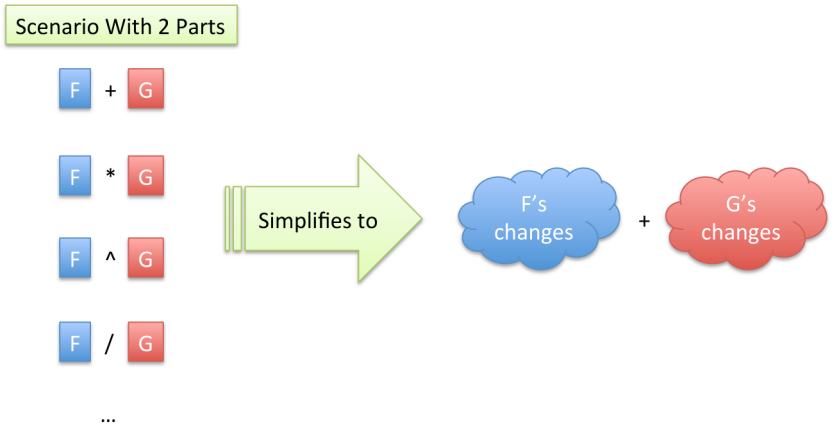
Instead of blasting through more rules, step back. Is there a pattern here?

Combining Perspectives

Imagine a function as a business with interacting departments, or a machine with interconnected parts. What happens when we make a change?

If our business has 4 departments, there are 4 perspectives to account for. It sounds painfully simple when written out: count the changes coming from each part!

No matter the specific interaction between parts F and G (addition, subtraction, multiplication, exponents...), we just have two perspectives to consider:



Aha! That's why the rules for $(f+g)'$ and $(f \cdot g)'$ are two added perspectives. The derivative of $\frac{1}{x}$ has a perspective because there's just one "moving part", x. (If you like, there *is* a contribution from "1" about the change it experiences: nothing. No matter how much you yell, 1 stays 1.)

The exact contribution from a perspective depends on the interaction:

- With addition, each part adds a direct change ($df + dg$).
- With multiplication, and each part thinks it'll add a rectangular strip ($f \cdot dg + g \cdot df$). (I'm using df instead of f' to help us think about the slice being added.)

You might forget the exact form of the multiplication rule. But you can think "The derivative of $f \cdot g$ must be something with df + something with dg."

Let's go further: what about the derivative of $a \cdot b + c$? You guessed it, 3 perspectives that should be added: something involving da plus something involving db plus something involving dc.

We can predict the shape of derivatives for gnarly equations. What's the derivative of:

$$x^y \cdot \frac{u}{v}$$

Wow. I can't rattle that off, but I can say it'll be something involving 4 additions (dx, dy, du and dv). Guess the shape of a derivative, even if you don't know the exact description.

Why does this work? Well, suppose we had a change that was influenced by both da and db, such as $15 \cdot da \cdot db$ – that'd be our instrument interfering with itself!

Only direct changes on a single variable's are counted (such as $3da$ or $12db$), and "changes on changes" like $15 \cdot da \cdot db$ are ignored.

Dimensional Intuition

Remember that derivatives are a fancier form of division. What happens when we make a division like $\frac{x^3}{x}$? We divide volume by length, and get area (one dimension down).

What happens when we do $\frac{d}{dx}x^3$? You might not know yet, but you can bet we'll be dropping a dimension.

Dimensions	Example	Description
3	x^3	Volume (cubic growth)
2	x^2	Area (square growth)
1	x	Length (linear growth)
0	c	Constant (no change)
-1	$\frac{1}{x}$	Inverse length ("per length")

-2	$\frac{1}{x^2}$	Inverse area ("per area")
-3	$\frac{1}{x^3}$	Inverse volume ("per volume")

When we divide or take a derivative, we drop a dimension and hop down the table. Volume is built from area (slices dx thick), area is built from length (dx wide).

A constant value, like 3, has no dimension in the following sense: it never produces a set of slices. There will never be a jump to the next value. Once we have a constant value, we get "stuck" on the table because the 0 pattern always produces another 0 pattern (volume to area to lines to constant to zero to zero...).

We can have negative dimensions as well: "per length", "per area", "per volume", etc. Derivatives still decrease the dimension, so when seeing $\frac{1}{x}$, we know the derivative will resemble $\frac{1}{x^2}$ as the dimension drops a level.

An important caveat: Calculus only cares about quantities, not their dimension. The equations will happily combine x and x^2, even though we know you can't mix length with area. Units add a level of meaning from *outside* the equation, that help keep things organized and warn us if we've gone awry (if you differentiate area and get volume, you know something went wrong).

Thinking With Dimensions

You can think about the dimension of a derivative without digging into a specific formula. Imagine the following scenario:

- Take a string and wrap it tight around a quarter. Take another string and wrap it tight around the Earth.
- Lengthen both strings, adding more to the end, so there's a 1-inch gap all the way around around the quarter, and a 1-inch gap all the way around the Earth (like having a ring floating around Saturn).
- Quiz: Which scenario needed more extra string to create? Is it more string to put a 1-inch gap around the quarter, or a 1-inch gap around the Earth?

We can crunch through the formula, but think higher-level. Because circumference is a 1-d line, its reaction to change (the derivative) will be a constant. No matter the current size, the circumference will change the same amount with every push. So, the extra string needed is the same for both! (About 6.28 inches per inch of radius increase).

Now, suppose we were painting the sphere instead of putting a string around it. Ah, well, area is squared, therefore the derivative is a dimension lower (linear). If we have a 5-inch and 10-inch sphere, and make them 1-inch bigger each, the larger sphere will require double the *extra* paint.

Questions

Let's think about the derivative of x^3, a growing cube.

1) What dimension should the derivative of x^3 have?
2) How many viewpoints should $x^3 = x \cdot x \cdot x$ involve?
3) Have a guess for the derivative? Does it match with how you'd imagine a cube to grow?

CHAPTER 14

THE FANCY ARITHMETIC OF CALCULUS

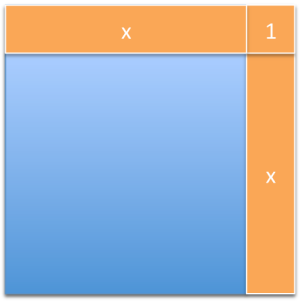
Here's the rules we have so far:

$$(f+g)' = f' + g'$$
$$(f \cdot g)' = f \cdot g' + g \cdot f'$$
$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

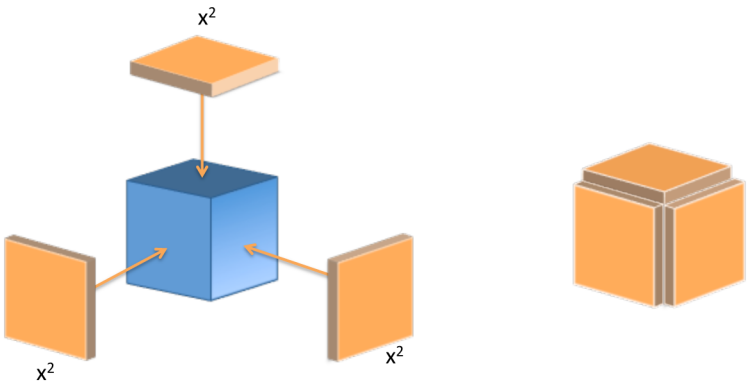
Let's add a few more to our collection.

Power Rule

We've worked out that $\frac{d}{dx}x^2 = 2x$:



We can visualize the change, and ignore the artificial corner piece. Now, how about visualizing x^3?



The process is similar. We can glue a plate to each side to expand the cube. The "missing gutters" represent artifacts, where our new plates would interact with each other.

I have to keep reminding myself: the gutters aren't real! They represent growth that doesn't happen at this step. After our growth, we "melt" the cube into its new, total area, and grow again. Counting the gutters would overestimate the growth that happened in this step. (Now, if we're forced to take integer-sized steps, then the gutters are needed – but with infinitely-divisible decimals, we can change smoothly.)

From the diagram, we might guess:

$$\frac{d}{dx}x^3 = 3x^2$$

And that's right! But we had to visualize the result. Abstractions like algebra let us handle scenarios we can't visualize, like a 10-dimensional shape. Geometric shapes are a nice, visual starting point, but we need to move beyond them.

We might begin analyzing a cube with using algebra like this:

$$(x+dx)^3 = (x+dx)(x+dx)(x+dx) = (x^2+2x\cdot dx+(dx)^2)(x+dx) = \ldots$$

Yikes. The number of terms is getting scary, fast. What if we wanted the 10th power? Sure, there are algebra shortcuts, but let's think about the problem holistically.

Our cube $x^3 = x\cdot x\cdot x$ has 3 components: the sides. Call them a, b and c to keep 'em straight. Intuitively, we know the total change has a contribution from each side:



Now what change does each side think it's contributing?

- **a thinks:** My change (da) is combined with the other, unmoving, sides ($b \cdot c$) to get $da \cdot b \cdot c$
- **b thinks**: My change (db) is combined with the other sides to get $db \cdot a \cdot c$
- **c thinks:** My change (dc) is combined with the other sides to get $dc \cdot a \cdot b$

Each change happens separately, and there's no "crosstalk" between da, db and dc (such crosstalk leads to gutters, which we want to ignore). The total change is:

$$A's\ changes + B's\ changes + C's\ changes = (da \cdot b \cdot c) + (db \cdot a \cdot c) + (dc \cdot a \cdot b)$$

Let's write this in terms of x, the original side. Every side is identical, ($a = b = c = x$) and the changes are the same ($da = db = dc = dx$), so we get:

$$(dx \cdot x \cdot x) + (dx \cdot x \cdot x) + (dx \cdot x \cdot x) = x^2 \cdot dx + x^2 \cdot dx + x^2 \cdot dx = 3x^2 \cdot dx$$

Converting this to a "per dx" rate we have:

$$\frac{d}{dx} x^3 = 3x^2$$

Neat! Now, the brain-dead memorization strategy is to think "Pull down the exponent and decrease it by one". That isn't learning!

Think like this:

- x^3 has 3 identical perspectives.
- When the system changes, all 3 perspectives contribute identically. Therefore, the derivative will be $3 \cdot something$.
- The "something" is the change in one side (dx) multiplied by the remaining sides ($x \cdot x$). The changing side goes from x to dx and the exponent lowers by one.

We can reason through the rule! For example, what's the derivative of x^5?

Well, it's 5 identical perspectives ($5 \cdot something$). Each perspective is me changing (dx) and the 4 other guys staying the same ($x \cdot x \cdot x \cdot x = x^4$). So the combined perspective is just $5x^4$.

The general Power Rule:

$$\frac{d}{dx} x^n = nx^{n-1}$$

Now we can memorize the shortcut "bring down the exponent and subtract", just like we know that putting a 0 after a number multiplies by 10. Shortcuts are fine once you know *why* they work!

Integrals of Powers

Let's try integrating a power, reverse engineering a set of changes into the original pattern.

Imagine a construction site. Day 1, they order three 1×1 wooden planks. The next day, they order three 2×2 wooden planks. Then three 3×3 planks. Then three 4×4 planks. What are they building?

My guess is a cube. They are building a shell, layer by layer, and perhaps putting grout between the gutters to glue them together.

Similarly, if we see a series of changes like $3x^2$, we can visualize the plates being assembled to build a cube:

$$\int 3x^2 = x^3$$

Ok – we took the previous result and worked backward. But what about the integral of plain old x^2? Well, just imagine that incoming change is being split 3 ways:

$$x^2 = \frac{x^2}{3} + \frac{x^2}{3} + \frac{x^2}{3} = \frac{1}{3}3x^2$$

Ah! Now we have 3 plates (each 1/3 of the original size) and we can integrate a smaller cube. Imagine the "incoming material" being split into 3 piles to build up the sides:

$$\int x^2 = \int \frac{1}{3}\cdot 3x^2 = \frac{1}{3}\int 3x^2 = \frac{1}{3}x^3$$

If we have 3 piles of size x^2, we can make a full-sized cube. Otherwise, we build a mini-cube, 1/3 as large.

The general integration rule is:

$$\int x^n = \frac{1}{n+1}x^{n+1}$$

After some practice, you'll do the division automatically. But now you know *why* it's needed: we have to split the incoming "change material" among several sides. (Building a square? Share changes among 2 sides. Building a cube? Share among 3 sides. Building a 4d hypercube? Call me.)

The Quotient Rule

We've seen the derivative of an inverse (a "simple division"):

$$\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$$

Remember the cake metaphor? We cut our existing portion ($\frac{1}{x}$) into x slices, and give one away.

Now, how can we find the derivative of $\frac{f}{g}$? One component in the system is trying to grow us, while the other divides us up. Which wins?

Abstraction to the rescue. When finding the derivative of x^3, we imagined it as $x^3 = a \cdot b \cdot c$, which helped simplify the interactions. Instead of a mishmash of x's being multiplied, it was just 3 distinct perspectives to consider.

Similarly, we can rewrite $\frac{f}{g}$ as two perspectives:

$$\frac{f}{g} = a \cdot b$$

We know $a = f$ and $b = \frac{1}{g}$. From this zoomed-out view, it looks like a normal, rectangular, product-rule scenario:

$$(a \cdot b)' = da \cdot b + db \cdot a$$

It's our little secret that b is really $\frac{1}{g}$, which behaves like a division. We just want to think about the big picture of how the rectangle changes.

Now, since a is just a rename of f, we can swap in $da = df$. But how do we swap out b? Well, we have:

$$b = \frac{1}{g}$$

$$\frac{db}{dg} = -\frac{1}{g^2}$$

$$db = -\frac{1}{g^2} dg$$

Ah! This is our cake cutting. As g grows, we lose $db = -\frac{1}{g^2}\dot{d}g$ from the b side. The total impact is:

$$(a \cdot b)' = (da \cdot b) + (db \cdot a) = \left(df \cdot \frac{1}{g}\right) + \left(\frac{-1}{g^2} dg \cdot f\right)$$

This formula started with the product rule, and we plugged in their real values. Might as well put f and g back into $(a \cdot b)'$, to get the Quotient Rule (aka the Division Rule):

$$\left(\frac{f}{g}\right)' = \left(df \cdot \frac{1}{g}\right) + \left(\frac{-1}{g^2} dg \cdot f\right)$$

Many textbooks re-arrange this relationship, like so:

$$\left(df \cdot \frac{1}{g}\right) + \left(\frac{-1}{g^2} dg \cdot f\right) = \frac{g \cdot df}{g^2} - \frac{f \cdot dg}{g^2} = \frac{g \cdot df - f \cdot dg}{g^2}$$

And I don't like it, no ma'am, not one bit! This version no longer resembles its ancestor, the product rule.

In practice, the Quotient Rule is a torture device designed to test your memorization skills; I rarely remember it. Just think of $\frac{f}{g}$ as $f \cdot \frac{1}{g}$, and use the product rule like we've done.

Questions

Let's do a few warm-ups to test our skills. Can you solve these bad boys?

$$\frac{d}{dx}x^4 = ?$$

$$\frac{d}{dx}3x^5 = ?$$

(You can check your answers with Wolfram Alpha, such as d/dx x^4.)

Again, don't get lost in the symbols. Think "I have x^4 – what pattern of changes will I see as I make x larger?".

Ok! How about working backwards, and doing some integrals?

$$\int 2x^2 = ?$$

$$\int x^3 = ?$$

Ask yourself, "What original pattern would create steps in the pattern $2x^2$?"

Trial-and-error is ok! Try a formula, test it, and adjust it. Personally, I like to move aside the 2 and just worry about the integral of x^2:

$$\int 2x^2 = 2\int x^2 = ?$$

How do you know if you're right? Take the derivative – *you* are the antiques dealer! I brought you a pattern of shards ($2x^2$) and you need to tell me the vase they came from. Once you have your guess, just break it in the back room, and make sure you get $2x^2$ back out. Then you'll be confident in your answer.

We're getting ready to work through the circle equations ourselves, and recreate results found by Archimedes, likely the greatest mathematician of all time.

CHAPTER 15

DISCOVERING ARCHIMEDES' FORMULAS

In the preceding lessons we uncovered a few calculus relationships, the "arithmetic" of how systems change:

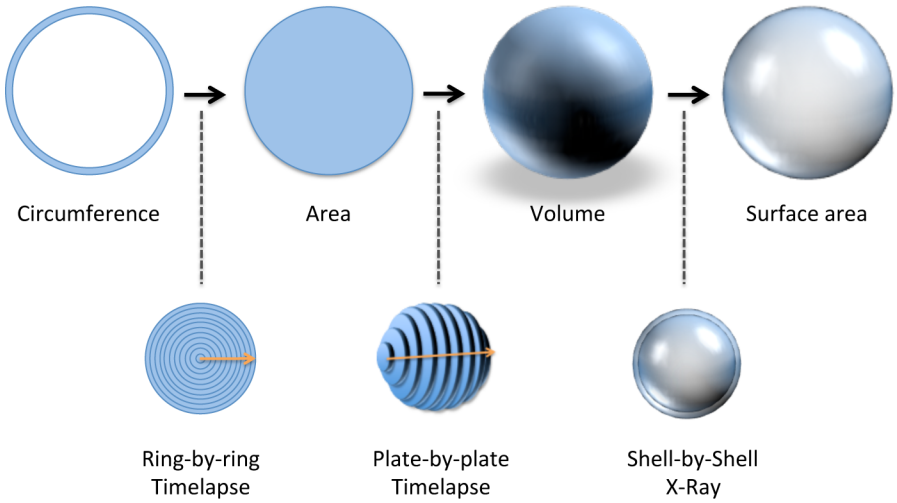
Interaction	Overall Change
Addition	$(f+g)' = f' + g'$
Multiplication	$(f \cdot g)' = f \cdot dg + g \cdot df$
Powers	$(x^n)' = \frac{d}{dx}x^n = nx^{n-1}$
Inverse	$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$
Division	$\left(\frac{f}{g}\right)' = \left(df \cdot \frac{1}{g}\right) + \left(\frac{-1}{g^2}dg \cdot f\right)$

How do these rules help us?

- If we have an existing equation, the rules are a shortcut to find the step-by-step pattern. Instead of visualizing a growing square, or cube, the Power Rule lets us crank through the derivatives of x^2 and x^3. Whether x^2 is from a literal square or multiplied elements isn't important – we'll get the pattern of changes.
- If we have a set of changes, the rules help us reverse-engineer the original pattern. Getting changes like $2x$ or $14x$ is a hint that $something \cdot x^2$ was the original pattern.

Learning to "think with calculus" means we can use X-Ray and Time-lapse vision to imagine changes taking place, and use the rules to work out the specifics. Eventually, we might not visualize anything, and just work with the symbols directly (as you likely do with arithmetic today).

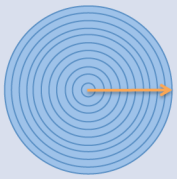
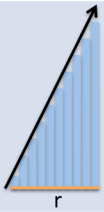
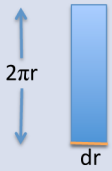
In the start of the course, we morphed a ring into a circle, then a sphere, then a shell:



With the official rules in hand, we can blast through the calculations and find the circle/sphere formulas on our own. It may sound strange, but the formulas feel different to me – almost alive – when you see them morphing in front of you. Let's jump in.

Changing Circumference To Area

Our first example of "step-by-step" thinking was gluing a sequence of rings to make a circle:

Strategy	Visualization	Step-by-Step Layout	Single Step Zoom
Ring-by-ring Timelapse			

When we started, we needed a lot of visualization. We had to unroll the rings, line them up, realize they made a triangle, then use $\frac{1}{2} base \cdot height$ to get the area. Visual, tedious... and necessary. We need to feel what's happening before working with raw equations.

Here's the symbolic approach:

Symbolic Description	Solution	Notes
$area = \int_0^r 2\pi r \; dr$	$= 2\pi \int_0^r r \; dr$ $= 2\pi \left(\frac{1}{2} r^2\right)$ $= \pi r^2$	Work backwards to the integral. If $\frac{d}{dx} r^2 = 2r$ that means $\frac{d}{dx} \frac{1}{2} r^2 = r$

Let's walk through it. The notion of a "ring-by-ring timelapse" sharpens into "integrate the rings, from nothing to the full radius" and ultimately:

$$area = \int_0^r 2\pi r \; dr$$

Each ring has height $2\pi r$ and width dr, and we want to accumulate that area to make our disc.

How can we solve this equation? By working backwards. We can move the 2π part outside the integral (remember the scaling property?) and focus on the integral of r:

$$2 \cdot \pi \int_0^r r \; dr = ?$$

What pattern makes steps of size r? Well, we know that r^2 creates steps of size $2r$, which is double what we need. Half that should be perfect. Let's try it out:

$$\frac{d}{dr} \frac{1}{2} r^2 = \frac{1}{2} \frac{d}{dr} r^2 = \frac{1}{2} 2r = r$$

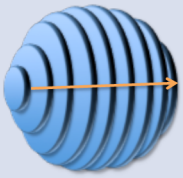
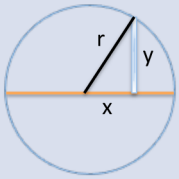
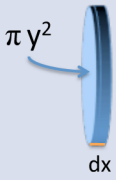
Yep, $\frac{1}{2} r^2$ gives us the steps we need! Now we can plug in the solution to the integral:

$$area = 2 \cdot \pi \int_0^r r \; dr = 2 \cdot \pi \cdot \frac{1}{2} r^2 = \pi r^2$$

This is the same result as making the ring-triangle in the first lesson, but we manipulated equations, not diagrams. Not bad! It'll help even more once we get to 3d...

Changing Area To Volume

Let's get fancier. We can take our discs, thicken them into plates, and build a sphere:

Strategy	Visualization	Height of Plate	Single Step Zoom
Plate-by-plate Timelapse		 $x^2 + y^2 = r^2$	

Let's walk slowly. We have several plates, each at a different "x-coordinate". What's the size of a single plate?

The plate has a thickness (dx), and its own radius. The *radius of the plate* is some off the x-axis, which we can call y.

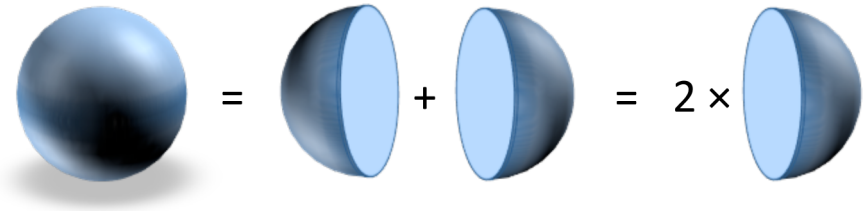
It's a little confusing at first: r is the radius of the entire sphere, but y is the (usually smaller) radius of an individual plate under examination. In fact, only the center plate ($x = 0$) will have its radius the same as the entire sphere. The "end plates" don't have a height at all.

And by the Pythagorean theorem, we have a connection between the x-position of the plate, and its height (y):

$$x^2 + y^2 = r^2$$

Ok. We have size of each plate, and can integrate to find the volume, right?

Not so fast. Instead of starting on the left side, with a negative x-coordinate, moving to 0, and then up to the max, let's just think about a sphere as two halves:



To find the total volume, get the volume of one half, and double it. This is a common trick: if a shape is symmetrical, get the size of one part and scale it up. Often, it's easier to work out "0 to max" than "min to max", especially when "min" is negative.

Ok. *Now* let's solve it:

Symbolic	Solution	Notes
$Volume = 2\int_0^r \pi y^2 dx$	$= 2\int_0^r \pi y^2 \, dx$ $= 2\int_0^r \pi(\sqrt{r^2 - x^2})^2 \, dx$	Write height in terms of x $x^2 + y^2 = r^2$ $y = \sqrt{r^2 - x^2}$
	$= 2\pi\int_0^r r^2 - x^2 \, dx$ $= 2\pi\left((r^2)x - \frac{1}{3}x^3\right)$	Work backwards to find integrals
	$= 2\pi\left((r^2)r - \frac{1}{3}r^3\right)$ $= 2\pi\left(\frac{2}{3}r^3\right)$ $= \frac{4}{3}\pi r^3$	Find volume at full radius (x=r)

Whoa! Quite an equation, there. It seems like a lot, but we'll work through it:

First off: 3 variables is too many to have flying around. We'll write the height of each plate (y), in terms of the others:

$$height = y = \sqrt{r^2 - x^2}$$

The square root looks intimidating at first, but it's being plugged into y^2 and the exponent will cancel it out. After plugging in y, we have the much nicer:

$$Volume = 2\pi\int_0^r r^2 - x^2 \, dx$$

The parentheses are often dropped because it's "known" that dx is multiplied by the entire size of the step. We know the step is $(r^2 - x^2) \cdot dx$ and not $r^2 - (x^2 \cdot dx)$.

Let's talk about r and x for a minute. r is the radius of the entire *sphere*, such as "15 inches". You can imagine asking "I want the volume of a sphere with a radius of 15 inches". Fine.

To figure this out, we'll create plates at each x-coordinate, from $x = 0$ up to $x = 15$ (and double it). x is the bookkeeping entry that remembers which plate we're on. We could work out the volume from $x = 0$ to $x = 7.5$, let's say, and we'd build a partial sphere (maybe useful, maybe not). But we want the whole shebang, so we let x go from 0 to the full r.

Time to solve this bad boy. What equation has steps like $r^2 - x^2$?

First, let's use the addition rule: steps like $a - b$ are made from two patterns (one making a, the other making b).

Let's look at the first pattern, the steps of size r^2. We're moving along the x-axis, and r is a number that never changes: it's 15 inches, the size of our sphere. This max radius never depends on x, the position of the current plate.

When a scaling factor doesn't change during the integral (r, π, etc.), it can be moved outside and scaled up at the end. So we get:

$$\int r^2 \; dx = r^2 \int dx = r^2 x$$

In other words, $r^2 \cdot x$ is a linear trajectory that contributes a constant r^2 at each step.

Cool. How about the integral of $-x^2$? First, we can move out the negative sign and take the integral of x^2:

$$-\int x^2 \; dx = ?$$

We've seen this before. Since x^3 has steps of $3x^2$, taking 1/3 of that amount ($\frac{x^3}{3}$) should be just right. And we can check that our integral is correct:

$$\frac{d}{dx}\left(-\frac{1}{3}x^3\right) = -\frac{1}{3}\frac{d}{dx}x^3 = -\frac{1}{3}3x^2 = -x^2$$

It works out! Over time, you'll learn to trust the integrals you reverse-engineer, but when starting out, it's good to check the derivative. With the integrals solved, we plug them in:

$$2\pi \int r^2 - x^2 \; dx = 2\pi(r^2 x - \frac{1}{3}x^3)$$

What's left? Well, our formula still has x inside, which measures the volume from 0 to some final value of x. In this case, we want the full radius, so we set $x = r$:

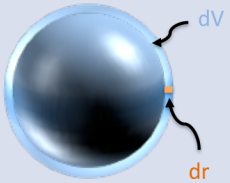
$$2\pi(r^2 x - \frac{1}{3}x^3) \xrightarrow[\text{set x=r}]{} 2\pi\left((r^2)r - \frac{1}{3}r^3\right) = 2\pi\left(r^3 - \frac{1}{3}r^3\right) = 2\pi\frac{2}{3}r^3 = \frac{4}{3}\pi r^3$$

Tada! You've found the volume of a sphere (or another portion of a sphere, if you use a different range for x).

Think that was hard work? You have no idea. That one-line computation took Archimedes, one of the greatest geniuses of all time, tremendous effort to figure out. He had to imagine some spheres, and a cylinder, and some cones, and a fulcrum, and imagine them balancing and... let's just say when he found the formula, he had it written on his grave. Your current calculus intuition would have saved him incredible effort (see this video).

Changing Volume To Surface Area

Now that we have volume, finding surface area is much easier. We can take a thin "peel" of the sphere with a shell-by-shell X-Ray:

Strategy	Visualization	Shell Analysis
Shell-by-shell X-Ray		

I imagine the entire shell as "powder" on the surface of the existing sphere. How much powder is there? It's dV, the change in volume. Ok, what is the *area* the powder covers?

Hrm. Think of a similar question: how much area will a bag of mulch cover? Get the volume, divide by the desired thickness, and you have the area covered. If I give you 300 cubic inches of dirt, and spread it in a pile 2 inches thick, the pile will cover 150 square inches. After all, if $area \cdot thickness = volume$ then $area = \frac{volume}{thickness}$.

In our case, dV is the volume of the shell, and dr is its thickness. We can spread dV along the thickness we're considering (dr) and see how much area we added: $\frac{dV}{dr}$, the derivative.

This is where the right notation comes in handy. We can think of the derivative as an abstract, instantaneous rate of change (V'), or as a specific ratio ($\frac{dV}{dr}$). In this case, we want to consider the individual elements, and how they interact (volume of shell / thickness of shell).

So, given the relation,

$$\text{area of shell} = \frac{\text{volume of shell}}{\text{depth of shell}} = \frac{dV}{dr}$$

we figure out:

$$\frac{d}{dr} Volume = \frac{d}{dr}\frac{4}{3}\pi r^3 = \frac{4}{3}\pi\frac{d}{dr}r^3 = \frac{4}{3}\pi(3r^2) = 4\pi r^2$$

Wow, that was fast! The order of our morph (circumference / area / volume / surface area) made the last step simple. We could try to spin a circumference into surface area directly, but it's more complex.

As we cranked through this formula, we "dropped the exponent" on r^3 to get $3r^2$. Remember this change comes from 3 perspectives (dimensions) that contributed an equal share.

2000 Years Of Math In An Afternoon

The steps we worked through took 2000 years of thought to discover, by the greatest geniuses no less. Calculus is such a broad and breathtaking viewpoint that it's difficult to imagine where it *doesn't* apply. It's just about using X-Ray and Time-Lapse vision:

- **Break things down.** In your current situation, what's the next thing that will happen? And after that? Is there a pattern here? (Getting bigger, smaller, staying the same.) Is that knowledge useful to you?
- **Find the source.** You're seeing a bunch of changes – what caused them? If you know the source, can you predict the end-result of all the changes? Is that prediction helpful?

We're used to analyzing equations, but I hope it doesn't stop there. Numbers can describe mood, spiciness, and customer satisfaction; step-by-step thinking can describe battle plans and psychological treatment. Equations and geometry are just nice starting points to analyze. Math isn't about equations, and music isn't about sheet music – they point to the idea inside the notation.

While there are more details for other derivatives, integration techniques, and how infinity works, you don't need them to start thinking with Calculus. What you discovered today would have made Archimedes tear up, and that's a good enough start for me.

Happy math.