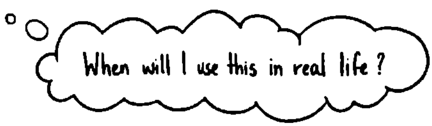
models

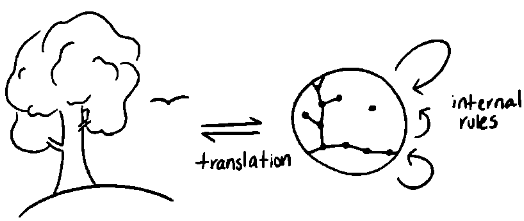
Okay, okay, I hear you. What's the point of any of this, right? Axioms, double and triple tori, continuum-sums, wallpaper symmetries. For what? In the pointed phrasing of math students around the world and throughout history:



I've tried to avoid addressing this question directly because (and I promise this is the last time I'll remind you of this) professional mathematicians really don't care about real-world applications. That's the domain of *applied* math, the opposite of *pure* math, which should give you a sense for how the word "applied" is meant to sound. But here we are, with a good chunk of pages to go, having already run through the three main branches of pure math, plus a little history and philosophy. So I'll entertain the question and say a thing or two about applied math, even if it'll get me in trouble with the hardcore types who find this "real life" stuff irrelevant and distracting.

In particular, this last section is about modeling. Modeling is how math connects to the real world. Of course there are lots of different ways math turns up in the real world, but modeling is a sort of general framework that lets us see all these connections clearly. It gives us a convenient way to talk about the connections, so we can explore them and learn new things.

A model consists of two main ingredients. There's the way the model itself works: a set of internal, mathematical rules that determine how everything inside the abstract model-world operates. And then (this is the important part) there's some kind of translation process that connects the model back to the outside world.



Of course I've skimmed over all the gritty details, but even from this rough description you can see what an arrangement like this would allow us to do. We could observe something in the real world, translate it into the language of the model, follow the internal laws of the model to infer new truths, and then translate it back into our reality. We could, in other words, learn things about the real world by taking a detour through a fictional, mathematical world. This is new.

Let's look at an example: music theory. Music theory is an abstract model of how music works. You take real-world music, a complex and chaotic parade of vibrations in the air, and you translate it into a symbolic system of notes and chords. Inside that abstract system, there are certain rules or guidelines (for a given genre or musical tradition) about which notes work with which chords, which sequences of notes will sound tense or sad or funky, and which chords typically follow which other chords. These are all the makings of a model. We have a simplified representation of a real-world thing which makes it easier to manage, analyze, and predict that real-world thing.

Yes, we lost detail when we abstracted. It's not a perfect translation, and the model-world isn't going to be isomorphic to the real world. That's fine. If you're playing in a jam session, you mostly just need to know the chord progression, the rhythm, and what key you're in. If you tried to analyze every aspect of the audio flow coming into your ears, you'd get hopelessly lost. Instead, you strip it down to the basics—you abstract. "Notes" and "chords" aren't tangible, real-world entities. These concepts live in model-world, they have internal rules of engagement, and they correspond back to sounds in the real world. They're useful theoretical constructs.

This is the key to a good model: a smart stripping-down process that takes us to a basic but still useful unit, like a note or a chord. When we're working inside the model, we temporarily pretend these things really are unbreakable atoms with fixed laws of behavior. This isn't strictly true: A note is actually a mishmash of overtones, echoes, and reverb all bouncing around, pushing up against your eardrums. But if it's useful to build a tiny model of the world where it *is* true, where a note is just a note, well, what's the harm in that?

Sometimes this stripping-down process goes a little too far, and we do have to be careful about drawing real-world conclusions from oversimplified models. It's often convenient to make assumptions that are not quite true, or even ones that are demonstrably, laughably false. We just have to strike a good balance between simplicity and usefulness. There's an old joke about an academic who's called to a dairy farm to help increase milk production and says, "I have a solution. We assume a spherical cow. . . ."

Here's another modeling example, from economics. Say there's some product that lots of people want to buy. Hot sauce, for instance. And then something happens, like a pest infestation in the chili fields, that reduces the amount of hot sauce being produced. What happens next is predictable: The price of hot sauce will go up. This is the kind of real-world regularity that lends itself perfectly to modeling. When there's a sudden shortage of something, its price typically rises.

Of course, a "price" isn't really ever just a single number. It depends on where you buy your hot sauce, who's selling it to you, how that person's business model works, maybe even how wealthy that person thinks you are. When the shortage happens, sellers who don't hear about it immediately might keep selling hot sauce at the original price until it runs out. Or buyers who don't know about the shortage might refuse to buy it at a higher price. Or within a certain community there may be an expectation of what the "fair price" of hot sauce is, and sellers could be ostracized for jacking up prices. It's hard to imagine something more complicated, with more moving parts under the hood, than a price.

But when we're modeling, we can make the simplifying assumption that price is just a single number, the same everywhere. We can also assume that the "demand curve" and "supply curve" (more abstractions invented for our modeling convenience) are simple functions that, depending on price, tell you exactly how much hot sauce will be desired and how much will be produced. We can assume that, in a "competitive market" (another abstraction), everything will settle into an "equilibrium price" (and another). Within the theoretical world built from these assumptions, we can solve for the equilibrium and convert it back to a prediction of what the real-world price will be. And in some cases, this supply–demand model actually makes pretty decent predictions.

Of course, we have to be careful about which assumptions we make. One of the standard assumptions of neoclassical economics is that humans are rational actors: that we have innate and consistent preferences, that we seek out the highest-paying jobs and the lowest-priced products, that we have complete information about pretty much everything. Most of this is not, in the real world, accurate. These are simplifying assumptions that let us make predictions. If the predictions tend to come true, great! The model is useful. That doesn't mean the assumptions are true. There are plenty of ways in which humans very much don't act rationally: We're overly risk-averse, we don't plan for the future well, we buy expensive things to flex our wealth, we discriminate, we give jobs to friends and family over more qualified outsiders, the list goes on and on. If you try to apply the standard models in these cases, they'll break down and make poor predictions.

This is a crucial point about modeling in general: A model works only within a certain scope. The assumptions you use to make good predictions in one area (like economics) could be totally different from the assumptions you use to make good predictions in another area (like sociology). This doesn't mean one model is right and the other is wrong. It just means you have to know when to use which. If you think you have a single, consistent model that works in all contexts, you're probably just ignoring or downplaying the contexts where it doesn't work. No model is sacred.

One more example: Have you ever watched a movie and, about halfway through, been able to predict most of what happens for the rest of the movie? When you think about it, that's a pretty remarkable feat. How can you see the future like that? You must have a mental model of "how movies usually work" that you've developed from watching movies all your life. You simplify the flow of information coming into your ears and eyes, turning pixels into abstract units like characters, dialog, motives, relationships. Then you apply some unspoken rules: "If they show a loaded gun, it'll get shot before the movie's over" or "That character who's super racist will definitely get their comeuppance" or "Around the last twenty minutes of the movie they'll break up over this character flaw but then he'll learn his lesson and make a grand romantic gesture and they'll reunite dramatically and live happily ever after." Sure, these aren't strict mathematical rules, and the predictions might not be accurate every time, but you're still doing some rudimentary modeling. You're building a set of rules in your head that you can apply across a variety of similar real-world circumstances.

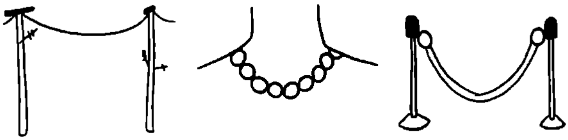
And really, when it comes down to it, this is what's going on in our heads all the time. We interpret the world around us not as flashes of light and sound; we chunk it into things, entities, units of analysis we expect to behave in certain ways. We see something we categorize as a "car" and something we categorize as a "green light" and we think, *Cars typically keep driving through green lights. If I cross the street now I'll likely be hit.* Human perception and cognition are all about pattern recognition, and to recognize patterns we first have to abstract the continuous, fuzzy reality around us into discrete objects that can behave in patterned ways.

Notice, also: Models don't have to be mathematical. The internal rules of model-world can be rough and qualitative, things like "Opposites attract" or "Birds of a feather flock together." If anything, it should be far easier to build these kinds of non-mathematical models. A model that makes precise numerical predictions, after all, is very easy to prove wrong.

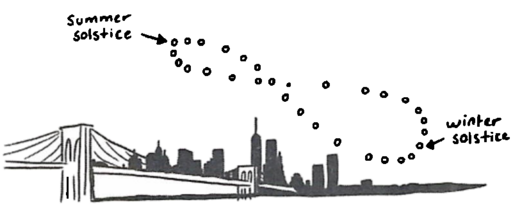
Which is why it's surprising that our world makes itself so intelligible to mathematical modeling. A remarkable number of things, if you pay close attention, are practically screaming for us to use math to describe their behavior.

Here, take any small object. Your keys will do. Toss it up from your left hand and catch it in your right hand. The path it makes through the air is a perfect parabola. No matter how you throw it, it'll always follow a parabolic path. It recreates a mathematical object, a precise geometric shape, in real life!

Or take a piece of string and dangle it between two points. It'll settle into a shape called a catenary, a perfect replica of a graph called the hyperbolic cosine. Telephone wires, unweighted necklaces, velvet VIP ropes—no matter the material, it'll always make this same shape. (The formula for this shape, by the way, involves an irrational number called *e* that arises from the study of compound interest, and which has absolutely no right being in the equation for how strings hang.)

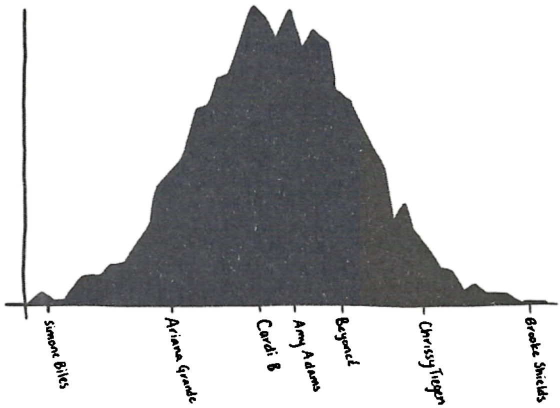


One more shape. This one's a bit more involved. Set up a camera on a tripod and point it at the sky. Pick a time of day to take a picture. Leave it in the exact same position, and take a picture at the same time the next day, and the next, and keep doing this every day for a year. The path of the sun over the course of a year will trace out a mathematical shape called an analemma.

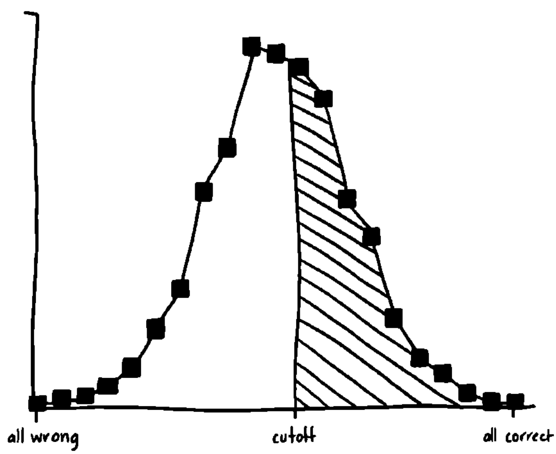


I'm giving examples of complicated shapes, because simple mathematical shapes are so commonplace in nature that we hardly notice them. When you blow soap bubbles they form perfect spheres. Drop a pebble in a pond and the ripples will travel out in perfect circles. These examples don't seem quite so surprising, but they also point to there being some sort of mathematical logic operating behind the scenes.

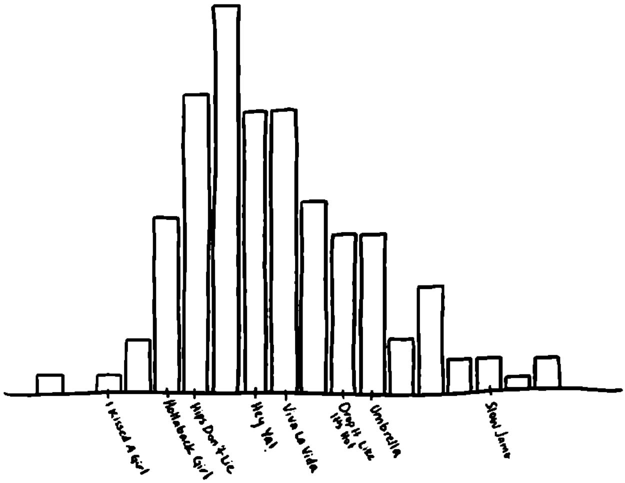
This bizarre recurrence of mathematical phenomena in the natural world goes far beyond physical shapes. Another familiar example, which we really shouldn't take for granted, is the bell curve: a formula for predicting the distribution of almost any numerical property in any naturally occurring data set. Here, for instance, is the distribution of women's height in the United States:



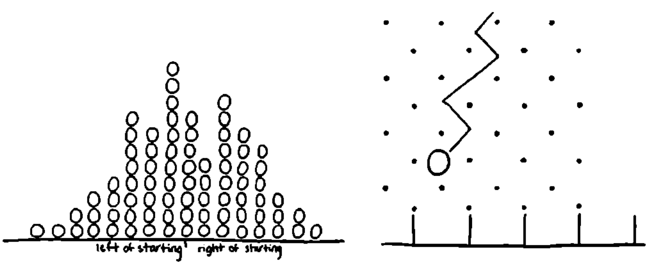
And here's the distribution of scores on the multistate bar exam:



Here's the distribution of song length for all Billboard number one hits of the aughts:

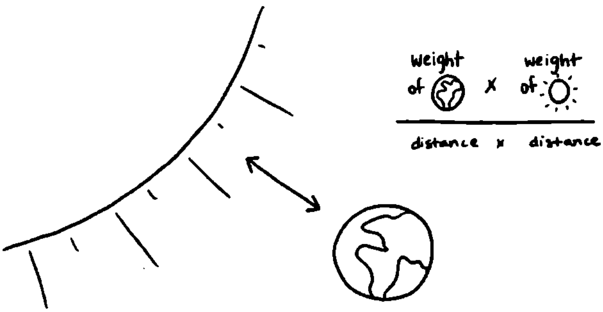


And here's the distribution of where the ball ends up in *The Price Is Right* game Plinko:

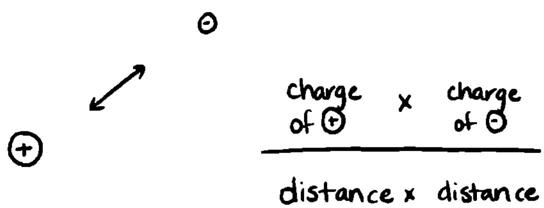


No, it's not precisely the same shape every time—you have to allow for some randomness. But the bigger the sample size, generally speaking, the closer you get to a smooth, symmetric bell curve. (The equation of this curve, by the way, not only includes e—the compound interest number—but also π, the ratio of a circle's circumference to its diameter. After a point, doesn't this start to feel like some kind of cosmic joke?)

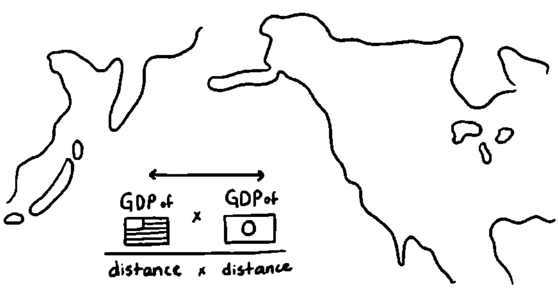
This is the eeriest thing for me, when the exact same formula pops up in different fields of study, in entirely unrelated contexts that don't seem like they should be analogous. So, for instance, the famous gravity equation tells us the force of attraction between two macroscopic objects if we know their masses:



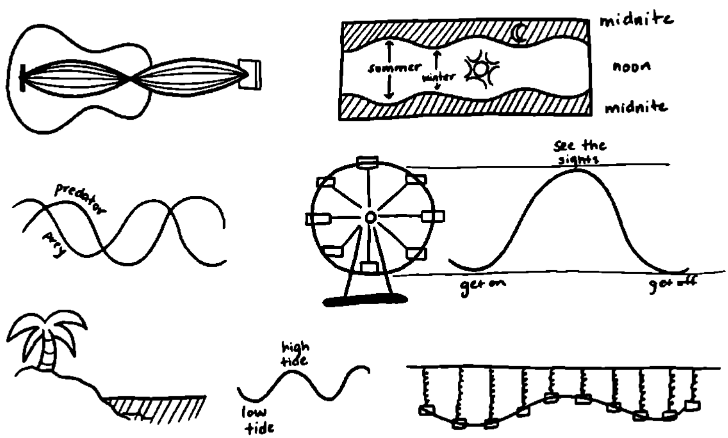
But it also tells us the force of attraction or repulsion between two microscopic particles if we know their charges:



And (brace yourself) it also gives us a good estimate for the amount of trade between two countries if we know their GDPs:



Better yet, the mathematical process known as "simple harmonic motion" identically describes the vibration of a plucked string, the length of a day* and average temperature over the course of a year, the population of species in predator–prey relationships, the height of a point on a rotating circle, the level of the tides, and the compression of a spring.



What the hell is going on here? Our goal in making models, remember, is just to be useful, to find a nice and convenient system to summarize what we observe in an orderly way. The rules of a model can take any form, rough or precise. But for some reason, time and again, we find that the world is best modeled by mathematical rules, which work with shockingly high precision, and which sometimes repeat themselves from place to place.

In pretty much every case, by the way, the math came first historically. Pure mathematicians have always just studied whatever they find interesting. But what typically ends up happening is that, hundreds of years after a new area of math is identified and explored, a new area of empirical science pops up that requires exactly those same mathematical concepts and results. We're not inventing math to fit our world—we're discovering what math is out there, and then later realizing that our world happens to look exactly like it.

How can we explain this? Why is the world so susceptible to mathematical modeling?

The most honest answer is that no one really knows for sure. This is a hot topic for debate among philosophers of math, and I'm not going to pretend I know the answer. Within the pure math community, though, there's one theory that seems to be very popular. People won't come out and say it quite like this, but I've run it by enough people to feel confident saying a lot of us believe it's true.

Maybe we observe mathematical patterns in nature because *the world itself is made of math*. Maybe the universe is fundamentally mathematical in character, and there's a One True Model that perfectly describes its behavior.

Let's not mince words: That sounds insane. But hear us out.