The Math Journal

Rushil Surti

August 29, 2023

Contents

What *is* a Math Journal Anyways? 2023-03-28

This is hopefully going to be the start of something fun! I've been wanting a space to document, formalize, and just have fun typesetting things and ideas that I've been working on relating to (mostly) math. This journal will probably contain everything that goes into that process: my musings, results, and work, although I'll try and keep the actual line-for-line scratch work to a minimum, because that looks bad and isn't very helpful. Who knows, I'll probably even throw in a couple contest math problems!

This is also very much a testing ground for my LaTeX styling stuff, so keep that in mind ;)

All in all, I intend to go back to previous results that I've worked on as well as anything I come up with. This may result in a lot of sections that will be "TODO," but I guess you can't finish all your side projects, right?

As you can probably tell, the tone is going to be quite conversational and decently informal, but I'm not going to jump into "zoomer language" and meme around *too much*. That being said this is all for fun.

$$\text{Sugar} \longrightarrow \text{Me} \longrightarrow \text{Math}$$

Figure 1. My daily schedule

One might also ask the question, why are you writing this as if someone will read it when in all likelihood it won't be read much? To that hypothetical question posed by myself to myself about writing to myself, I must answer that this is far too philosophical for me and we must move on. It's math time!

Covering the Cartesian Coordinates 2023-03-29

Whoo boy after some brief styling with the document, it's finally time to get to our very first entry! This problem really isn't anything amazingly special to commemorate the occasion or anything, though; it's just the first one I had on hand.

After drudging through the nth tangent line problem for the AP review, I started to get bored and my mind drifted. This is when I thought of the following problem.

Even I sometimes come up with some interesting stuff, right? *Right*?

Problem. Let T denote the set of all points contained within every line tangent to some function $f(x)$. Is there a function $f(x)$ such that $T = \mathbb{R}^2$?

This problem evolved after thinking about it for a while and actually trying things out, but I thought I would at least share the original problem and my thoughts on it.

And I also like doing this because it details my thinking process and also simply goes to show that no mathematician is truly perfect. Math is a journey and behind every proof or problem there's potentially a lot of trial and error that happens, so don't be discouraged!

Playing Around with the Idea

I'm sure there's a lot of ways to start to tackle a problem like this, and depending on your experience, you may or may not quickly find something to latch onto, but one of the tried and true things to do is just play around with the problem. What exactly are we asking for, and what sort of assumptions can we make to narrow down the scope of the problem?

First, let's get a visual intuition for what we're actually measuring. Imagine for some function $f(x)$ dragging a point along the curve and looking at the tangent line at that curve. For some curves, this tangent line will stay relatively similar (especially for functions with a constant derivative, which should make sense); however, for other functions, this tangent line will "sweep" across the plane, which is the ideal behavior we're looking for.

Now the problem has reduced to finding a function that will sweep through the *entire* Cartesian plane. This corresponds to a function whose derivative has the following properties:

Remember, even if these don't mean much, that's fine! All we're trying to do is empty our brains and think of something that will lead us to a potential solution.

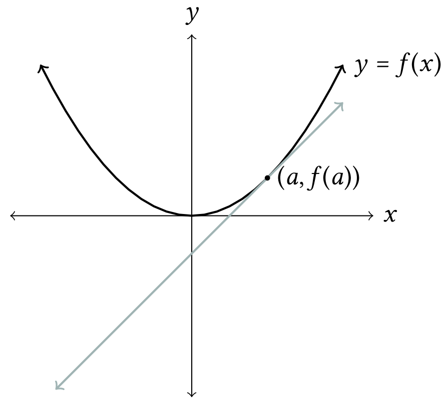

Figure 2. Imagine dragging a point with its tangent line along some function $f(x)$.

1. A change in sign. If the derivative is always positive or always negative, there's going to be some part of the plane that the tangent lines will never reach, so the sign must flip somewhere.

2. Some sort of periodicity. Coupled with a change in sign, if we can have the derivative be periodic in some form, it may reduce down to some repeatable pattern covering the plane.

3. Steep lines. Perhaps subjective and not necessarily a requirement, but derivatives with high magnitude will likely reach more points.

Keeping this in mind, let's explore and essentially just guess some possible solutions. If we think about it for a minute though, we can actually come up with quite a few solutions! Right off the bat, we can see $f(x) = \sin(x)$ is a applicable solution. So long as we keep on moving to infinity, the tangent line will keep sweeping through the plane.

One could try and make the argument for a function defined over a finite domain, but this is easily adjustable with a simple change of argument such as $f(x) = \sin(1/x)$ or $f(x) = \sin(\ln(x) - \ln(1-x))$, both of which compress the entire domain of our wave into a finite domain. In fact, a far more challenging (and interesting) problem to tackle would be finding a function such that $T = \mathbb{R}^2 \setminus P$, with P being some point on the Cartesian plane. In order to do this, however, we must introduce some some of quantitative metric to find out whether a point truly is included in T.

Doing the Math™

Now that we've looked at some intuition for what's happening and realized that we would like to make some adjustments to the problem, it's time to actually get down to business and figure out how to determine whether a point is in our set T.

Proof. Given some point a on $f(x)$, we have the following line tangent to the function:

$$L_a(x) = f'(a)(x - a) + f(a).$$

In order for some point P to be contained in this tangent line, we must set x to P_x and $L_a(x)$ to P_y. This will then give us some function in a, which tells us something about whether P is in T. If there exists an a which solves the equation, we know that the point must be included in the set of points that the tangent lines contain. If not, however, we can say for certain that this point is not in T.

I'm no expert on proofs and rigour (but I'm trying my best to improve!) so some of my arguments may be lacking in areas throughout this entire journal. But you'll forgive me, right? :)

I'm using "vector" notation for the components of these points because its far more convenient that way. For anyone unfamiliar, we take P_x to denote the x part of the point P and similarly P_y represents the y part of P

We can use this strategy to now show that $f(x) = \sin(x)$ covers the entire Cartesian plane. Let us test any point P using the previous equation.

$$P \in T \iff 0 = \cos(a)(P_x - a) + \sin(a) - P_y.$$

In cases like this, it is far easier to tools from calculus to verify that a solution exists rather than solving for it explicitly. Let us call the right hand side of our modified linearization "discriminant" function $\phi_P(a)$, where P represents the point we are testing. If we vary a, we see the following holds true.

I wasn't really sure what to call it so I landed on discriminant function. A "discriminating" function doesn't sound very nice now does it? :)

$$\lim_{a \to \infty} \phi_P(2\pi a) \longrightarrow -\infty$$
$$\lim_{a \to -\infty} \phi_P(2\pi a) \longrightarrow \infty.$$

It is trivial to see then by the Intermediate Value Theorem that there exists some value of a on the number line in which $\phi_P(a) = 0$, which means that there exists some solution to our discriminant equation. Due to this applying for any arbitrary point P, we have just shown that the entire Cartesian plane is contained in T for $f(x) = \sin(x)$. ■

Now that we have verified and familiarized ourselves with some of the tools at use here, let's go back to our modified problem. Can we construct a function where all but *one single point* in the Cartesian plane is included in T?

The Modified Problem

This problem is a bit harder to tackle, especially at first glance. In the end, this boils down to whether or not we can find a function with some specific characteristics.

1. For every point Q that is not P, there must exist an a value such that our condition $0 = f'(a)(Q_x - a) + f(a) - Q_y$ holds.
2. Specifically for the point P, there must be no solution to the given the preceding condition.

Just to simplify things a little bit, let's take P to be the point $(0, 0)$ and see what happens. Concretely, let's take a look at the second condition, as it seems much stronger (and thus can narrow down our choice of f, if any). The condition now becomes that there must be **no** value of a such that the following has as solution:

$$0 = -af'(a) + f(a) = \phi_P(a).$$

Utilizing this and the fact that f is continuous, we can split this into two (likely symmetric) cases.

1. $\phi_P(a) > 0$ for all a
2. $\phi_P(a) < 0$ for all a.

We see that this must be the case once again due to the Intermediate Value Theorem. This also shows that the condition that f must be continuous over all real numbers is quite limiting.

Let us suppose we have some function f where the first case is true. Notice then that if we let Q be of the form $(P_x, y) = (0, y)$ for some arbitrary y, something quite interesting happens. Let us examine our discriminant functions again.

$$\phi_P(a) = -af'(a) + f(a)$$
$$\phi_Q(a) = -af'(a) + f(a) - y = \phi_P(a) - y$$

Notice that when we pick a point directly above or below P, it corresponds to simply vertically shifting $\phi_Q(a)$. Keeping in mind that we want $\phi_P(a)$ to *never* have a solution and $\phi_Q(a)$ to *always* have a solution, this poses a complication.

With this, supposing that $\phi_P(a)$ is always greater than 0, there will always be a sufficiently large negative Q_y value such that $\phi_Q(a)$ does not have a solution. Vice versa, a similar condition holds for when $\phi_P(a)$ is always less than 0. Thus, there exists **no continuous function f of x such that all but one point is included in the set of points of its tangent lines.** Truly a shame, but its cool that we can prove this.

To offer a visual intuition for why this is the case, consider the function $f(x) = 1/x^2$. As we head to the infinities at either side of the graph, the line gets closer and closer to obtaining a slope of 0. However, as we approach the pole at $x = 0$, the function quickly picks up and the line will get ever close to being vertical. This colors all of the Cartesian plane, *except* for one particular strip of values on the negative y-axis.

This choice isn't quite continuous over all reals, but we'll let it slide because it was chosen in order to illustrate a point more than anything

Remarks

While we have shown that there exists no function of x with the properties described in the modified problem. One should note that the solution is almost trivial if we turn to the world of parametric functions. Consider a function f defined as such:

$$f_x(t) = \varepsilon \cos(t)$$
$$f_y(t) = \varepsilon \sin(t)$$

If we take $\varepsilon \longrightarrow 0^+$, we quickly see the desired behavior. The fucntion, infinitely close to being a point, hugs the boundary of $(0, 0)$, with the tangent line going through all points except $(0, 0)$. If we wanted to have this centered around some arbitrary point, all we would have to do is simply add or subtract from the respective function components.

Conclusion

All in all, this was a sort of interesting problem, and I'm at least proud of it! It wasn't all that difficult and didn't require any higher level tools besides some calculus, but that in of itself is quite nice. I hope to find some similarly interesting problems that apply what the classroom teaches and goes beyond that, requiring some problem solving skills. Who knows? Maybe there's already a competition problem or such revolving around the concept. I wouldn't be too surprised if there was.

With that, this is the end of the first entry in this journal. Hopefully it isn't the last :)

General Product Rule 2023-04-06

We all know and love (well hopefully) the famous product rule for derivatives, given as follows:

$$\frac{d}{dx}\Big(f(x)g(x)\Big) = f(x)g'(x) + f'(x)g(x).$$

But here's an interesting question. What happens when we have more than two functions of x multiplied together? The result is what's called the Leibniz product rule. Let's derive it!

Finally! A problem that won't make me tear my hair out.

Problem. What does the following evaluate to?

$$\frac{d}{dx}\prod_{i=1}^{n} f_i(x)$$

Solving

To be honest, there isn't much complicated machinery involved in this one. First, let's observe what happens we increase the number of functions.

Case $n = 3$:

It's a bit noisy with all the x arguments, so while I'll omit them here, just know that these are all functions of x.

We want to take the derivative of the product $f_1\, f_2\, f_3$. Left with no other choice, we proceed by using our normal product rule, taking our left function to be f_1 and our right function to be the product $f_2\, f_3$, resulting in the following:

$$\begin{aligned}\frac{d}{dx}\left(f_1 \cdot f_2\, f_3\right) &= f_1'\, f_2\, f_3 + f_1 \cdot \frac{d}{dx}\left(f_2\, f_3\right)\\ &= f_1'\, f_2\, f_3 + f_1 \cdot \left(f_2'\, f_3 + f_2\, f_3'\right)\\ &= f_1'\, f_2\, f_3 + f_1\, f_2'\, f_3 + f_1\, f_2\, f_3'.\end{aligned}$$

Case $n = 4$: One may also verify the following, using the fact that we have already found the derivative for three functions.

$$\frac{d}{dx}\left(f_1\, f_2\, f_3\, f_4\right) = f_1'\, f_2\, f_3\, f_4 + f_1\, f_2'\, f_3\, f_4 + f_1\, f_2\, f_3'\, f_4 + f_1\, f_2\, f_3\, f_4'.$$

It may become clear now that we have a pattern forming with a rather beautiful symmetry. It seems that when we take the derivative of the product of these n functions, it splits into n terms each with the same product except for one function's derivative being taken. This recurrence forming in the pattern may also help motivate our method of proof of this a little bit later on.

This can be expressed a little more compactedly (although perhaps less illuminatingly) as the following:

$$\frac{d}{dx}\prod_{i=1}^{n} f_i(x) = \sum_{i=1}^{n} f_i'(x) \prod_{\substack{j=1\\ j\neq i}}^{n} f_i(x)$$

We shall now show this rigorously.

Proof. We shall show this using proof by induction. For our base case, the case $n = 1$ or $n = 2$ are quite trivial.

Next, suppose that for some k we have that

$$\frac{d}{dx}\prod_{i=1}^{k} f_i(x) = \sum_{i=1}^{k} f_i'(x) \prod_{\substack{j=1\\ j\neq i}}^{k} f_i(x)$$

We now must show that the same holds when we increase $k \mapsto k+1$, effectively multiplying inside the derivative by a new function f_{k+1}. Taking one side to be f_{k+1} and the other to be our big product, we can continue with regular product rule.

$$\frac{d}{dx}\left(f_{k+1}\cdot\prod_{i=1}^{k}f_i(x)\right)=f_{k+1}\cdot\sum_{i=1}^{k}f_i'(x)\prod_{\substack{j=1\\j\neq i}}^{k}f_i(x)+f_{k+1}'\prod_{i=1}^{k}f_i(x)$$
$$=\sum_{i=1}^{k}f_i'(x)\prod_{\substack{j=1\\j\neq i}}^{k+1}f_i(x)+f_{k+1}'\prod_{i=1}^{k}f_i(x)$$

With this, we're almost where we want to be at. After bringing the f_{k+1} term inside the product on the right side, notice that we have in our left term all "groups of terms" where there is one derivative *except* for f_{k+1}, and on the right side, we have our *only* derivative being f_{k+1}. We can bring this right term inside, now giving us our desired form.

You know, there's actually quite the bit of "symbol soup" going on here. Perhaps expanding this would make it look more reasonable? I'll have to revisit this sometime because I don't think the math nor the explanation are very helpful if you don't have the idea in mind.

$$\frac{d}{dx}\prod_{i=1}^{k+1}f_i(x)=\sum_{i=1}^{k+1}f_i'(x)\prod_{\substack{j=1\\j\neq i}}^{k+1}f_i(x)$$

This concludes the proof. ∎

Uses

The reader may now be wondering, what could possibly be the use of such a formula? Indeed, it's not *too* often that one needs to take the derivative of a large product (for a Taylor series, you'd need to be able to take n derivatives, and that gets quite hairy in of itself); however, allow me to offer at least one "cool" example. One of the times you may see a product that works very well for these cases is a factored polynomial.

Well at the very least, I think it's cool!

Consider a polynomial $P(x)$ that has n distinct roots $r_1, r_2, \ldots, r_n$. We can write the polynomial as the following:

These roots necessarily have to be distinct, as we'll see later ;)

$$P(x)=\prod_{i=1}^{n}(x-r_i).$$

Perhaps you can see where I'm going with this. Using our generalized product rule, we can take the derivative of $P(x)$, noting that each individual term has a derivative of simply 1 due to it being a linear term.

$$P'(x)=\sum_{i=1}^{n}\prod_{\substack{j=1\\j\neq i}}^{n}\left(x-r_j\right)$$

This in of itself isn't all too interesting, following directly from our generalized product rule. If we plug in *the roots* into the derivative, however fun things happen with canceling. Taking x to be some arbitrary root r_j, we have that

$$P'\left(r_j\right)=\prod_{\substack{i=1\\i\neq j}}^{n}\left(r_j-r_i\right)$$

For *any polynomial*, the derivative of it at some root is the product of the differences of that root with all the others. That's pretty cool.

Remember, back to when we said that these roots have to be *distinct*? If we now relax that and instead say that some root r_j has a multiplicity greater than one, something else happens instead. Because we only take the derivative of one of the terms containing $x - r_j$, there are still other terms with this, making the derivative 0 altogether. This tells us that for any polynomial, there is going to be a horizontal tangent at any root whose multiplicity is greater than one. That's also pretty *smooth* if I do say so myself.

Realigning Some Rotations 2023-04-26

Here's a cute little problem that I actually had to solve while programming a Rubik's cube in 3D. While rotating the cube in small steps, we can keep a matrix in the background that keeps track of these rotations applied to the cube in case we need to invert them. The problem with this, however, is that we gradually run into floating point issues and other numerical differences, resulting in the rotation applied to the cube and the accumulated rotation falling "out of sync" so to speak.

Specifically, when rotating individual faces of the cube, we need to invert the cube rotations, apply our correct face rotaion around some axis, and then reapply the cube rotation. It's pretty fun actually.

Now I could have gone hunting for optimizations and other ways to try and duct-tape a solution together, but really what I needed was just a fast and simple way of getting the rotation. After a short while of thinking, there's actually a pretty nice solution.

For reference, this was done in the CMU CS Academy™ portal, so the project was 1) really not performant and 2) on a deadline.

In addition to the cube, what if we stored some extra points and looked at how they were transformed under the rotations? If we had enough points, could we uniquely determine our rotation R? Seeing as how I'm writing this, the answer is indeed yes.

In particular, say we have 3 three-dimensional vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$ along with their respective transformed vectors under our rotation R: $\vec{\mathbf{u}}'$, $\vec{\mathbf{v}}'$, and $\vec{\mathbf{w}}'$. Then we really have these three cases:

$$R\vec{\mathbf{u}} = \vec{\mathbf{u}}' \qquad R\vec{\mathbf{v}} = \vec{\mathbf{v}}' \qquad R\vec{\mathbf{w}} = \vec{\mathbf{w}}'$$

In order to do something with this, we first expand R as the following and then decompose matrix equation for each point into three linear equations.

$$R = \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,2} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}.$$

For vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$,

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,2} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}\begin{bmatrix} \vec{\mathbf{u}}_1 \\ \vec{\mathbf{u}}_2 \\ \vec{\mathbf{u}}_3 \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{u}}_1' \\ \vec{\mathbf{u}}_2' \\ \vec{\mathbf{u}}_3' \end{bmatrix} \longrightarrow \begin{aligned} R_{1,1}\vec{\mathbf{u}}_1 + R_{1,2}\vec{\mathbf{u}}_2 + R_{1,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}_1' \\ R_{2,1}\vec{\mathbf{u}}_1 + R_{2,2}\vec{\mathbf{u}}_2 + R_{2,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}_2' \\ R_{3,1}\vec{\mathbf{u}}_1 + R_{3,2}\vec{\mathbf{u}}_2 + R_{3,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}_3' \end{aligned}$$

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,2} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}\begin{bmatrix} \vec{\mathbf{v}}_1 \\ \vec{\mathbf{v}}_2 \\ \vec{\mathbf{v}}_3 \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{v}}_1' \\ \vec{\mathbf{v}}_2' \\ \vec{\mathbf{v}}_3' \end{bmatrix} \longrightarrow \begin{aligned} R_{1,1}\vec{\mathbf{v}}_1 + R_{1,2}\vec{\mathbf{v}}_2 + R_{1,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}_1' \\ R_{2,1}\vec{\mathbf{v}}_1 + R_{2,2}\vec{\mathbf{v}}_2 + R_{2,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}_2' \\ R_{3,1}\vec{\mathbf{v}}_1 + R_{3,2}\vec{\mathbf{v}}_2 + R_{3,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}_3' \end{aligned}$$

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,2} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}\begin{bmatrix} \vec{\mathbf{w}}_1 \\ \vec{\mathbf{w}}_2 \\ \vec{\mathbf{w}}_3 \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{w}}_1' \\ \vec{\mathbf{w}}_2' \\ \vec{\mathbf{w}}_3' \end{bmatrix} \longrightarrow \begin{aligned} R_{1,1}\vec{\mathbf{w}}_1 + R_{1,2}\vec{\mathbf{w}}_2 + R_{1,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}_1' \\ R_{2,1}\vec{\mathbf{w}}_1 + R_{2,2}\vec{\mathbf{w}}_2 + R_{2,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}_2' \\ R_{3,1}\vec{\mathbf{w}}_1 + R_{3,2}\vec{\mathbf{w}}_2 + R_{3,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}_3' \end{aligned}$$

Notice now that we have three relations for each three rows of coefficients in R. This let's us now rearrange the nine equations we have as the following three systems of three equations, which we can again put back into matrix form.

$$\begin{aligned} R_{1,1}\vec{\mathbf{u}}_1 + R_{1,2}\vec{\mathbf{u}}_2 + R_{1,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}'_1 \\ R_{1,1}\vec{\mathbf{v}}_1 + R_{1,2}\vec{\mathbf{v}}_2 + R_{1,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}'_1 \\ R_{1,1}\vec{\mathbf{w}}_1 + R_{1,2}\vec{\mathbf{w}}_2 + R_{1,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}'_1 \end{aligned} \longrightarrow \begin{bmatrix} \vec{\mathbf{u}}_1 & \vec{\mathbf{u}}_2 & \vec{\mathbf{u}}_3 \\ \vec{\mathbf{v}}_1 & \vec{\mathbf{v}}_2 & \vec{\mathbf{v}}_3 \\ \vec{\mathbf{w}}_1 & \vec{\mathbf{w}}_2 & \vec{\mathbf{w}}_3 \end{bmatrix} \begin{bmatrix} R_{1,1} \\ R_{1,2} \\ R_{1,3} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{u}}'_1 \\ \vec{\mathbf{u}}'_2 \\ \vec{\mathbf{u}}'_3 \end{bmatrix}$$

$$\begin{aligned} R_{2,1}\vec{\mathbf{u}}_1 + R_{2,2}\vec{\mathbf{u}}_2 + R_{2,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}'_2 \\ R_{2,1}\vec{\mathbf{v}}_1 + R_{2,2}\vec{\mathbf{v}}_2 + R_{2,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}'_2 \\ R_{2,1}\vec{\mathbf{w}}_1 + R_{2,2}\vec{\mathbf{w}}_2 + R_{2,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}'_2 \end{aligned} \longrightarrow \begin{bmatrix} \vec{\mathbf{u}}_1 & \vec{\mathbf{u}}_2 & \vec{\mathbf{u}}_3 \\ \vec{\mathbf{v}}_1 & \vec{\mathbf{v}}_2 & \vec{\mathbf{v}}_3 \\ \vec{\mathbf{w}}_1 & \vec{\mathbf{w}}_2 & \vec{\mathbf{w}}_3 \end{bmatrix} \begin{bmatrix} R_{2,1} \\ R_{2,2} \\ R_{2,3} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{u}}'_1 \\ \vec{\mathbf{u}}'_2 \\ \vec{\mathbf{u}}'_3 \end{bmatrix}$$

$$\begin{aligned} R_{3,1}\vec{\mathbf{u}}_1 + R_{3,2}\vec{\mathbf{u}}_2 + R_{3,3}\vec{\mathbf{u}}_3 &= \vec{\mathbf{u}}'_3 \\ R_{3,1}\vec{\mathbf{v}}_1 + R_{3,2}\vec{\mathbf{v}}_2 + R_{3,3}\vec{\mathbf{v}}_3 &= \vec{\mathbf{v}}'_3 \\ R_{3,1}\vec{\mathbf{w}}_1 + R_{3,2}\vec{\mathbf{w}}_2 + R_{3,3}\vec{\mathbf{w}}_3 &= \vec{\mathbf{w}}'_3 \end{aligned} \longrightarrow \begin{bmatrix} \vec{\mathbf{u}}_1 & \vec{\mathbf{u}}_2 & \vec{\mathbf{u}}_3 \\ \vec{\mathbf{v}}_1 & \vec{\mathbf{v}}_2 & \vec{\mathbf{v}}_3 \\ \vec{\mathbf{w}}_1 & \vec{\mathbf{w}}_2 & \vec{\mathbf{w}}_3 \end{bmatrix} \begin{bmatrix} R_{3,1} \\ R_{3,2} \\ R_{3,3} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{u}}'_1 \\ \vec{\mathbf{u}}'_2 \\ \vec{\mathbf{u}}'_3 \end{bmatrix}$$

Notice now that we have the same matrix on the right side. We'll call this matrix A. The great thing about this operation is that we only really need to find the inverse of one matrix, namely A^{-1}. Doing so, we finally get a way to calculate the coefficients in R in terms of our vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, and $\vec{\mathbf{w}}$ as such:

$$\begin{bmatrix} R_{1,1} \\ R_{1,2} \\ R_{1,3} \end{bmatrix} = A^{-1} \begin{bmatrix} \vec{\mathbf{u}}'_1 \\ \vec{\mathbf{u}}'_2 \\ \vec{\mathbf{u}}'_3 \end{bmatrix}, \qquad \begin{bmatrix} R_{2,1} \\ R_{2,2} \\ R_{2,3} \end{bmatrix} = A^{-1} \begin{bmatrix} \vec{\mathbf{v}}'_1 \\ \vec{\mathbf{v}}'_2 \\ \vec{\mathbf{v}}'_3 \end{bmatrix}, \qquad \begin{bmatrix} R_{3,1} \\ R_{3,2} \\ R_{3,3} \end{bmatrix} = A^{-1} \begin{bmatrix} \vec{\mathbf{w}}'_1 \\ \vec{\mathbf{w}}'_2 \\ \vec{\mathbf{w}}'_3 \end{bmatrix}$$

With this, our Rubik's cube rotations now work properly and the world is saved, with cubers and non-cubers alike able to enjoy the wonders that are rotations.

Okay for the project it also actually took me a while to realize that I defined matrix multiplication incorrectly, but we won't go into that.

Generalizing Partial Fractions 2023-05-06

When you hear the term "partial fractions" perhaps your face lights up in delight, or perhaps your soul fills with dread in anticipation of the drudging through factoring and solving some linear equations for coefficients. Whatever the case, partial fractions are quite the useful tool, especially so for integration and generating functions among others. However, it isn't all the time that we are working with some explicit values for our rational functions. This leads us mathematicians to a natural question, is there a way to generalize the solving of such partial fractions?

Personally, I quite enjoy doing partial fractions by hand, but to each their own I guess.

Finding a Beautiful Result

Let's start with the following: suppose that our rational function that we want to decompose is of the form $P(z)/Q(z)$, where $P(z)$ and $Q(z)$ are polynomials. Because we want to be able to factor the denominator, let us also suppose that $Q(z)$ is a degree n polynomial with distinct, complex roots $z_1, z_2, \dots, z_n$. In short, $Q(z) = (z - z_1)(z - z_2)\cdots(z - z_n)$. Then by our usual partial fractions setup, we must have that

Notice that I said distinct roots. While what we're doing could probably be extended, this method as of right now probably does not work for roots of greater multiplicity than one. Perhaps a problem for future me.

$$\frac{P(z)}{Q(z)} = \sum_{i=1}^{n} \frac{a_i}{z - z_i},$$

where each a_i is really just some coefficient that we don't know yet. Most of our work will be to use some method to find these coefficients. If we remember back to regular partial fractions, one of the first things we do after setting it up is multiplying across by the denominator. This then later on allows us to plug in values for z without getting 0 in the denominator. Doing so, we get the following:

$$P(z) = \sum_{i=1}^{n} a_i \cdot \frac{Q(z)}{z - z_i}.$$

Now we can "plug in" our z values to help us find the value for each a_i. Notice the quotes around "plug in." Given that we're plugging in roots for z, we really have to be careful about simply just letting z equal some value with the roots in the denominator. Instead, for each a_i, $i \in \{1, 2, \dots, n\}$, we *take the limit* as z approaches the ith root z_i.

$$\lim_{z \longrightarrow z_i} P(z) = \lim_{z \longrightarrow z_i} \sum_{j=1}^{n} a_j \cdot \frac{Q(z)}{z - z_j}.$$

Now we have something quite interesting in particular. Notice that, because z_i is a root of $Q(z)$, each numerator goes to 0, but before we get too ahead of ourselves, we must also consider the denominator. Only when $j = i$ does $z_j = z_i$ and make the denominator 0. This gives us the opportunity to cancel out all but the $j = i$ case and then consider the limit from there.

$$P(z_i) = a_i \cdot \lim_{z \longrightarrow z_i} \frac{Q(z)}{z - z_i}$$

This is a classic 0/0 indeterminate form, so we can perhaps cheekily apply L'Hôpital's rule, which doesn't require much work at all. Solving for each a_i then yields

$$a_i = \frac{P(z_i)}{Q'(z_i)}.$$

As such, we've delivered on the goal we set out to fulfill; that is, we have found our closed form for the coefficients, which is really all we needed for partial fractions. Thus we have this general formula for partial fractions:

$$\frac{P(z)}{Q(z)} = \sum_{i=1}^{n} \frac{P(z_i)}{Q'(z_i)(z - z_i)}.$$

At least for me, the first time I saw this and figured it out on paper was quite exciting but also I felt this was rather beautiful. The simple decomposition of a rational function somehow involves a derivative showing up in the denominator was extremely surprising to me at the time, and I've ended up using this in quite a few places.

If you know some complex analysis, you'll probably be able to see that this is related to the residues of the function. That's why this method of partial fraction decomposition is called the *residue method.*

Applications

Now it very well wouldn't be fair to the reader, myself, or even the math itself to show this result and not do anything with it, given how useful it certainly can be. In the introduction I mentioned that these are particularly useful for integration among others, so we'll go with that here. Now, if we're talking about integration over rational functions, then there's a *really* good example of where this has benefits over your usual partial fraction methods.

Problem. Let's evaluate the infamous bump function integral using our new partial fraction method!

$$\int \frac{1}{x^4+1}\,dx$$

It should be clear that we have $P(x) = 1$, $Q(x) = x^4 + 1$, and $Q'(x) = 4x^3$. The first thing we'll want to do is factor our denominator, which isn't too much different from the roots of unity.

$$\begin{aligned} x^4 &= -1 = \exp\big(i(\pi + 2\pi k)\big) \\ x &= \exp\big(i(\pi/4 + k\pi/2)\big) \end{aligned}$$

Choosing different k values from 0 to 3 gives us the 4 distinct roots of $Q(x)$. We will denote these roots by x_0, x_1, x_2, and x_3, with the subscript corresponding to the value of k. Just like the roots of unity, these roots have some nice properties that interplay with each other. In particular, we have:

Conjugates:

$$\begin{aligned} x_0 &= \overline{x_3} \\ x_1 &= \overline{x_2} \end{aligned}$$

Inverses:

$$\begin{aligned} x_0 &= 1/x_3 \\ x_1 &= 1/x_2 \end{aligned}$$

Opposites:

$$\begin{aligned} x_0 &= -x_2 \\ x_1 &= -x_3 \end{aligned}$$

Rotations:

$$\begin{aligned} x_0^3 &= x_1 \quad x_2^3 = x_3 \\ x_1^3 &= x_0 \quad x_3^3 = x_2 \end{aligned}$$

With this in mind, we can proceed with the integral, using our residue method for partial fractions.

$$\begin{aligned} \int \frac{1}{x^4+1}\,dx &= \int \left(\frac{1}{4x_0^3(x-x_0)} + \frac{1}{4x_1^3(x-x_1)} + \frac{1}{4x_2^3(x-x_2)} + \frac{1}{4x_3^3(x-x_3)} \right) dx \\ &= \frac{1}{4}\int \left(\frac{x_3^3}{x-x_0} + \frac{x_2^3}{x-x_1} + \frac{x_1^3}{x-x_2} + \frac{x_0^3}{x-x_3} \right) dx \\ &= \frac{1}{4}\int \left(\frac{x_2}{x-x_0} + \frac{x_3}{x-x_1} + \frac{x_0}{x-x_2} + \frac{x_1}{x-x_3} \right) dx \\ &= \frac{1}{4}\int \left(\frac{-x_0}{x-x_0} + \frac{-x_1}{x-x_1} + \frac{x_0}{x-x_2} + \frac{x_1}{x-x_3} \right) dx \\ &= \frac{x_0}{4}\Big(\operatorname{Log}(x-x_2) - \operatorname{Log}(x-x_0)\Big) + \frac{x_1}{4}\Big(\operatorname{Log}(x-x_3) - \operatorname{Log}(x-x_1)\Big) + C \end{aligned}$$

Now, there's a few things to pay attention to here, especially if you've seen the antiderivative of this bump function in a different form. Because we've chosen to split the function into rational functions with complex coefficients as opposed to rational quadratics with real coefficients, we have to deal with some of the complications of working in the complex world. Notice how our logarithm is capitalized and doesn't have the absolute value sign. Yes, that's right, this is not your household "real" natural logarithm but rather the complex logarithm, defined as follows.

$$\operatorname{Log}(z) := \ln|z| + i\operatorname{Arg} z.$$

The "real" form that you may familiar with that includes the ln and arctan terms can be obtained by combining some of the rational functions in the third line, but we want to avoid doing that because it takes a whole lot more work. Besides, both of these terms are really encompassed in the complex log anyways.

More specifically, this is the principal branch where choose $\operatorname{Arg} z \in (-\pi, \pi]$. We're still integrating along the real line and there isn't any contour action going on, but when dealing

Make sure to NINT before doing anything major. It saves lives, folks.

with the complex logarithm and complex numbers in real integrals, we do sometimes have to be careful about how we handle things. Indeed, a good indicator that something has terribly wrong and blown up is if we get a non-real value for a real definite integral.

While we could just as easily stop here (in fact this would be great for a computer to use), many might—rightfully so in my opinion—complain that this is simply an unsatisfactory form that hides away what's really going on in a more complex (heh) way. After all, if we started with a completely real integrand, why should it be that we have an antiderivative that has complex numbers and multi-valued functions interspersed around in it? So, with this, we shall use the power of algebra to work this into a more real-looking form.

If we are to clear this up without going through pages of algebra, it is best that we work at the more general level and then going through specific simplifications. Because of this, let us consider one of the general $\operatorname{Log}(x - x_k)$ terms. In order to not confuse what is really going on, we shall also decompose the complex number x_k into its real and imaginary parts $c_k + is_k$.

$$\begin{aligned}
\operatorname{Log}(x - x_k) &= \ln|x - x_k| + i \operatorname{Arg}(x - x_k) \\
&= \ln|x - c_k + is_k| + i \operatorname{Arg}(x - c_k + is_k) \\
&= \ln\left|\sqrt{(x - c_k)^2 + s_k^2}\right| + i \operatorname{Arg}(x - c_k - is_k).
\end{aligned}$$

Inside the natural logarithm a bit of simplifying then happens, expanding the square inside and using the fact that $c_k^2 + s_k^2 = 1$. Note that we must also keep the absolute value sign when we pull out the square root and make it a 1/2.

$$\operatorname{Log}(x - x_k) = \frac{1}{2} \ln\left|x^2 - 2c_k x + 1\right| + i \operatorname{Arg}(x - c_k - is_k).$$

Here's where we run into just a little bit of trouble. Ordinarily, we could turn this $\operatorname{Arg} z$ term into a more comfortable single-valued real "feeling" function like $\arctan\big(\Im(z)/\Re(z)\big)$, but we run into trouble when for example $\Re(z) < 0$ and $\Im(z) < 0$, mapping to a positive argument despite having a negative angle. After all, the range of $\operatorname{Arg} z$ is larger than regular old arctan. Not to mention that there is a discontinuity if we try to do this, having a $x - c_k$ term in the denominator.

I'm not too sure on the rigour of this also considering that this doesn't really solve the range problem, but the real antiderivative matches so perhaps it's just me being confused somewhere. Perhaps something to look at later.

We can somewhat address this by using a certain arctan identity, which allows us to flip around its argument. Recall that

$$\arctan x = -\big(\pi/2 + \arctan(1/x)\big)$$

Using this we can now simplify the right term, giving us something that does look suspiciously familiar to the "real" antiderivative.

$$\begin{aligned}
\operatorname{Log}(x - x_k) &= \frac{1}{2} \ln\left|x^2 - 2c_k x + 1\right| - i\pi/2 - i \arctan\left(\frac{x - c_k}{-s_k}\right) \\
&= \frac{1}{2} \ln\left|x^2 - 2c_k x + 1\right| - i\pi/2 + i \arctan\left(\frac{x - c_k}{s_k}\right).
\end{aligned}$$

Since we have worked quite generally now, all that is left is to expand everything out and simplify, which is a matter of just a bit of algebra. In the end, the cancelling and the simplifying does work out to be quite satisfying, and I won't omit it just in case the reader wants to follow along. Perhaps this will just turn into a page of just equations, though. Who knows!

Both me after writing this and the reader know.

$$
\begin{aligned}
&= \frac{x_0}{4}\Big(\operatorname{Log}(x-x_2) - \operatorname{Log}(x-x_0)\Big) + \frac{x_1}{4}\Big(\operatorname{Log}(x-x_3) - \operatorname{Log}(x-x_1)\Big) + C \\
&= \frac{x_0}{4}\left(\frac{1}{2}\ln\left|x^2 - 2c_2 x + 1\right| - i\pi/2 + i\arctan\left(\frac{x-c_2}{s_2}\right)\right) \\
&- \frac{x_0}{4}\left(\frac{1}{2}\ln\left|x^2 - 2c_0 x + 1\right| - i\pi/2 + i\arctan\left(\frac{x-c_0}{s_0}\right)\right) \\
&+ \frac{x_1}{4}\left(\frac{1}{2}\ln\left|x^2 - 2c_3 x + 1\right| - i\pi/2 + i\arctan\left(\frac{x-c_3}{s_3}\right)\right) \\
&- \frac{x_1}{4}\left(\frac{1}{2}\ln\left|x^2 - 2c_1 x + 1\right| - i\pi/2 + i\arctan\left(\frac{x-c_1}{s_1}\right)\right) + C \\
&= \frac{1}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2+\sqrt{2}x+1\right| - i\arctan\left(\sqrt{2}x+1\right)\right) + \frac{i}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2+\sqrt{2}x+1\right| - i\arctan\left(\sqrt{2}x+1\right)\right) \\
&- \frac{1}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2-\sqrt{2}x+1\right| + i\arctan\left(\sqrt{2}x-1\right)\right) - \frac{i}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2-\sqrt{2}x+1\right| + i\arctan\left(\sqrt{2}x-1\right)\right) \\
&- \frac{1}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2-\sqrt{2}x+1\right| - i\arctan\left(\sqrt{2}x-1\right)\right) + \frac{i}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2-\sqrt{2}x+1\right| - i\arctan\left(\sqrt{2}x-1\right)\right) \\
&+ \frac{1}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2+\sqrt{2}x+1\right| + i\arctan\left(\sqrt{2}x+1\right)\right) - \frac{i}{4\sqrt{2}}\left(\frac{1}{2}\ln\left|x^2+\sqrt{2}x+1\right| + i\arctan\left(\sqrt{2}x+1\right)\right) + C \\
&= \boxed{\frac{1}{4\sqrt{2}}\ln\left|x^2+\sqrt{2}x+1\right| - \frac{1}{4\sqrt{2}}\ln\left|x^2-\sqrt{2}x+1\right| + \frac{1}{2\sqrt{2}}\arctan\left(\sqrt{2}x+1\right) + \frac{1}{2\sqrt{2}}\arctan\left(\sqrt{2}x-1\right) + C.}
\end{aligned}
$$

Thus, we finally have our desired simplified form.

Really this started out more as a showcase of the cool partial fractions generalization, but it seems that in the end it morphed into simplifying our antiderivative into its rather beautiful real result. In the end, though, I hope your main takeaway from is that the complex logarithm form is a beautifully compact and rather simple way of denoting this far more involved antiderivative, and it was all thanks to the beautiful result we developed in the first section.

Polynomial Differential Equations 2023-03-29

I'd like to think that this is a decently interesting question that stems from some other stuff I've worked with in differential equations, we sometimes have ODEs of the form

$$a_0 y + a_1 y' + \cdots + a_n y^{(n)} = 0.$$

These are readily solveable using a method called the *characteristic polynomial*, which I won't get into too much here (but perhaps another time?), where essentially we just assume an exponential solution utilizing the roots of a similar looking polynomial.

As we'll see later, there will be a few analogues to these differential equations and characteristic polynomials in this solution. Connections in math are fun!

Here's where sort of my inspiration comes in. Instead of looking at differential equations that have characteristic polynomials or potential polynomials in x, what if we look at differential equations that are *polynomials in y*?

Polynomials in y

Problem. What can we say about differential equations of the following form and their solutions?

$$\frac{dy}{dx} = a_0 + a_1 y + \cdots + a_n y^n.$$

Because the right side is written completely in terms of y, we can recognize that this is simply a first order separable differential equation. With this, we can move the polynomial to the left hand side by dividing and then simply just integrate with respect to y. Then our problem lies in finding out how to integrate this general rational function. Now where have we heard that?

I was quite excited when I derived it myself, so I've used it quite a bit in various places. It's quite fun.

Using the same result, from the previous journal entry, we can integrate this rational function assuming we know the roots of the polynomial in y, which we will denote as $y_1, y_2, \dots, y_n$. Thus we have that

$$\frac{1}{a_n(y - y_1)(y - y_2)\cdots(y - y_n)} \cdot \frac{dy}{dx} = 1.$$

A Pretty Fun Game, IMO 2023-06-09

I'm actually really excited! I've vowed to be more productive over the summer so this seems like the perfect opportunity, although I do have to self-study some stuff.

Hay hay! It's the summer, so I actually have quite a bit more time now to spend on problems and such. With this extra time, I thought it would be really cool to try and tackle a far harder competition math problem than usual. After looking through a couple problems, I chanced upon a very cool looking IMO problem, so I thought it would be perfect for the occasion. Here's how it goes:

> **Problem** (2010 IMO Problem #5)**.** Each of the six boxes $B_1, B_2, B_3, B_4, B_5, B_6$ initially contains one coin. The following operations are allowed:
>
> 1. Choose a non-empty box B_j, $1 \leqslant j \leqslant 5$, remove one coin from B_j and add two coins to B_{j+1}.
> 2. Choose a non-empty box B_k, $1 \leqslant k \leqslant 4$, remove one coin from B_k and swap the contents (maybe empty) of the boxes B_{k+1} and B_{k+2}.
>
> Determine if there exists a finite sequence of operations of the allowed types, such that the five boxes B_1, B_2, B_3, B_4, B_5 become empty, while box B_6 contains exactly $2010^{2010^{2010}}$ coins.

As a brief remark on notation, I'll represent first move type by D_k and the second by S_k, representing "double" and "swap" respectively, with k denoting the chosen box that the move acts upon.

I chose this problem in particular because, as you might see, there isn't any crazy algebra or geometry or number theory going on where I would have to recall some very specific theorem and have a great deal of intuition for the subject (not to harp on those problems *too much*; I just wouldn't be able to solve them very well haha). This problem is quite nice in that it requires, at least on its face, no mathematical background or knowledge but rather a determined enough problem solver. It's a fun puzzle, and I really would like to tackle it. So, let's get started!

Observations

To be honest, starting the problem is probably one of my weak spots, although that may be the case somewhat for everyone.

It's time to brainstorm how to actually start the problem and sort of what I think we're really aiming for. This problem is really a game with a set of rather simple moves, and I get the feeling that this is going to be solved through pure logic rather than any fancy or flashy math theorem.

Because this is an IMO problem, I'm of the opinion that the answer is one of two things:

1. There *does* exist such a sequence and it's rather hard to find.
2. There *doesn't* exist such a sequence, and the proof for such is rather hard.

So, in summary, the problem is hard, but I mean that's why I picked it. I'm leaning towards there being such a sequence because that seems a lot more fun, but we probably wouldn't have to construct such a sequence anyways.

The proof may also revolve around some very light number theory in terms of parity. We might also have to take into consideration ideas such as the maximum possible number of coins to exist in the system and comparing that to our big constant.

In fact, on the topic of The Big Constant™, I find it somewhat hard to imagine it being special or completely unique in some way that allows this problem to work. It's likely that, if the sequence does indeed exist, we're looking for a more general way to reach a state of say $1, 1, 1, 1, 1, n$ or something similar rather than the exact number.

I mean if it is though that's pretty amazing.

Another thing to consider is that the first box may be some place to start as there is only one possible move you can ever do on in throughout the course of the game. When we do this move, the box ends up with 0 coins, and there is no way to add back to that (and obviously we have to do a move on this box because it has to become empty at some point.).

Another thing to consider is the associativity and the commutativity of the moves. While I'm sure this problem could be framed entirely in a group theory or abstract algebra setting, there isn't all too much to gain from doing so, but taking some ideas over certainly can't hurt.

Odds and Ends

Let's start with the following observation: $2010^{2010^{2010}}$ is even, and 1 is not. It sounds quite obvious, but notice that the only way we can add pieces is with the doubling move, which adds 2 coins to the next box. This means that the parity of the boxes stays the same in this case, which we certainly don't want. We can address this then by subtracting one, which does change the parity, and then swapping the boxes around to get it into the desired position as such:

The good 'ol Canadian shuffle, as they call it.

$$\begin{array}{cccccc}
\boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} \\
 & & \Big\downarrow D_5 & & & \\
\boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} & \boxed{0} & \boxed{3} \\
 & & \Big\downarrow S_4 & & & \\
\boxed{1} & \boxed{1} & \boxed{1} & \boxed{0} & \boxed{3} & \boxed{0}
\end{array}$$

A Putnam Functional Integral 2023-06-12

Every so often, I'm scrolling through YouTube, and I find a cool problem in the thumbnail that I just can't help but want to solve. Surprising as it is, some Putnam problems are actually not as difficult as you might think, and the calculus problems especially are always quite a bit of fun. Besides, it seems I'm on quite the roll for cool competition math problems, so I might as well continue on.

The problem that caught my eye goes as follows:

> **Problem** (77th Putnam Problem A3). Let $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ be the solution to the functional equation
> $$f(x) + f\left(1 - \frac{1}{x}\right) = \arctan x.$$
> Find
> $$\int_0^1 f(x)\,dx.$$

Solving the Functional Equation

While at first this does look a bit daunting, let's play around a bit with the expressions inside the functions, solving for some values. Let $x = 2$, giving us

$$f(2) + f(1/2) = \arctan 2.$$

Now we have two unknowns, so let's try and see if there's another relation involving $f(1/2)$. Let's set $x = 1/2$, giving us:

$$f(1/2) + f(-1) = \arctan(1/2).$$

Let's go one level deeper, setting $x = -1$.

$$f(-1) + f(2) = \arctan(-1).$$

Quite interestingly, we see that we have three relations for three different unknowns, which means that we can solve for each of them. What's more useful for us though, is that this choice of 2 was arbitrary, meaning that we can do this for any value of x (disregarding singularities).

For many functional equations, simply playing around with different substitutions or specific values often helps one gain a good insight as to what's going on.

This means that there is a quite simple way of solving for the function, and all we need to do is just take the approach from above but generalize it.

In particular, define $g(x) = 1 - 1/x$. Notice the following very nice properties that emerge when we take iterated applications of $g(x)$.

$$\begin{aligned} g(x) &= 1 - 1/x \\ g(g(x)) &= 1/(1 - x) \\ g(g(g(x))) &= x. \end{aligned}$$

Any subsequent applications will go back to one of these three functions due to, x being the identity function. Using this, we can now solve the functional equation. We shall use our given equation, writing it in terms of our new function $g(x)$ and also the modified function equations with the substitutions $x \mapsto g(x)$ and $x \mapsto g(g(x))$ to yield a system of three equations with three unknowns.

$$\begin{array}{ccccccc} f(x) & + & f(g(x)) & + & 0 & = & \arctan x \\ 0 & + & f(g(x)) & + & f(g(g(x))) & = & \arctan(g(x)) \\ f(x) & + & 0 & + & f(g(g(x))) & = & \arctan(g(g(x))) \end{array}$$

Through some rather simple algebra, we can solve the system to see that

$$2f(x) = \arctan x - \arctan(g(x)) + \arctan(g(g(x))).$$

With this, we're a whole lot closer to getting our answer. Now all that's left is to integrate!

Integrating Some Arctangents

If we integrate on both sides of our solved functional equation, we can split up each term into separate integrals and tackle each of them one-by-one.

$$2\int_0^1 f(x)\,dx = \overset{\textbf{1.}}{\boxed{\int_0^1 \arctan x\,dx}} - \overset{\textbf{2.}}{\boxed{\int_0^1 \arctan\left(1-\frac{1}{x}\right)dx}} + \overset{\textbf{3.}}{\boxed{\int_0^1 \arctan\left(\frac{1}{1-x}\right)dx}}.$$

We'll start off with the easiest one, which is just the integral of plain $\arctan x$, which lends itself well to integration by parts:

$$\begin{aligned} \textbf{1.}\ \int_0^1 \arctan x\,dx &= x\arctan x\Big|_0^1 - \int_0^1 \frac{x}{x^2+1}\,dx \xleftarrow{\text{IBP}} \boxed{\begin{array}{ll} u = \arctan x & du = dx/(x^2+1) \\ dv = dx & v = x \end{array}} \\ &= \frac{\pi}{4} - \frac{1}{2}\ln\left|x^2+1\right|\Big|_0^1 \xleftarrow{\text{substitution}} \boxed{\begin{array}{l} w = x^2+1 \\ dw = 2x\,dx \end{array}} \\ &= \boxed{\frac{\pi}{4} - \frac{1}{2}\ln 2}. \end{aligned}$$

Luckily, $\arctan x$ is a simple inverse function, so there's not much difficulty in using integration by parts to find its integral. This next integral won't be the cleanest.

For this integral, we'll again start with integration by parts, but a little more work is required to get it into a nice form.

TODO: Add in the boxes showing substitutions and such. They take a bit of work to put in, but it makes stuff easier to follow.

$$\textbf{2.}\ \int_0^1 \arctan\left(1-\frac{1}{x}\right)dx = x\arctan\left(1-\frac{1}{x}\right)\Big|_0^1 - \int_0^1 \frac{1}{x + x\left(1-\frac{1}{x}\right)^2}\,dx$$

For the left term, we can plug in $x = 1$ to get 0 for the top bound, but we must take the limit as $x \longrightarrow 0^+$ for the second.

$$\begin{aligned} x\arctan\left(1-\frac{1}{x}\right)\Big|_0^1 &= -\lim_{x\to 0^+} x\arctan\left(1-\frac{1}{x}\right) \\ &= -\lim_{x\to 0^+} \frac{\arctan(1-1/x)}{1/x} \\ &= -\lim_{x\to 0^+} \frac{1}{1+\left(1-\frac{1}{x}\right)^2} \\ &= -\lim_{u\to -\infty} \frac{1}{1+u^2} \\ &= 0. \end{aligned}$$

We now return to our integral and work it into a nicer form:

$$\begin{aligned}
&-\int_0^1 \frac{1}{x + x\left(1 - \frac{1}{x}\right)^2}\,dx = -\int_0^1 \frac{x}{x^2 + (x-1)^2}\,dx \\
&= -\int_0^1 \frac{x}{2x^2 - 2x + 1}\,dx = -\int_0^1 \frac{2x}{(2x-1)^2 + 1}\,dx \\
&= -\int_0^1 \frac{(2x-1)+1}{(2x-1)^2 + 1}\,dx = -\frac{1}{2}\int_{-\pi/4}^{\pi/4} (\tan\theta + 1)\,d\theta \\
&= \boxed{\frac{-\pi}{4}}
\end{aligned}$$

On to the next integral!

For this one, we'll first do a u-substitution of the entire integrand, getting us to something perhaps a little bit more familiar looking.

$$\begin{aligned}
\textbf{3.} \int_0^1 &\arctan\left(\frac{1}{1-x}\right)dx \\
&= \int_{\pi/4}^{\pi/2} u \csc^2 u\,du.
\end{aligned}$$

We'll now do just a little bit of tabular integration:

	D	I
$+$	u	$\csc^2 u$
$-$	1	$-\cot u$
$+$	0	$-\ln\lvert\sin u\rvert$

This means that the integral becomes

$$\begin{aligned}
&\Big[-u\cot u + \ln\lvert\sin u\rvert\Big]_{\pi/4}^{\pi/2} \\
&= \boxed{\frac{\pi}{4} + \frac{1}{2}\ln 2},
\end{aligned}$$

which does indeed look suspiciously familiar to some of our other results.

With this, we're pretty much complete with the problem, with all that's left to do being to substitute our values in. Taking great care to remember the 2 we placed in front of the integral way back last section, we see that the final answer to the problem is

$$\int_0^1 f(x)\,dx = \frac{1}{2}\left(\frac{\pi}{4} - \frac{1}{2}\ln 2 + \frac{\pi}{4} + \frac{\pi}{4} + \frac{1}{2}\ln 2\right) = \boxed{\frac{3\pi}{8}}.$$

Perhaps I'll do more Putnam problems in the future. They seem to be an interesting intersection of higher level mathematics (including stuff like calculus, abstract algebra, and combinatorics among others), but they definitely seem to have a different sort of "flavor" from the competition problems that I'm used to. Certainly, compared to the IMO and such, there are definitely some easier problems, but I'm sure not all of them are as easy. Perhaps I'll get to tackle the Putnam in person myself one day. Who knows?

Daily Puzzle Creation 2023-06-27

In my 7th grade math class, there was a clock that had on each hour label a different mathematical expression that evaluated to that hour. For example, 8 wasn't just plain old 8, but something like $\sqrt{64}$ instead. I've always thought this was a fun little recreational thing to nerd out about, but it slipped out of my mind for the longest time.

Ah the good old days of my youth...

I've since thought I should revisit this idea (mostly motivated by a bit of a running joke that I'm having) and have a little fun crafting puzzles daily. Compared to the clock though, they'll probably be a bit on the longer side (I also anticipate that some problems will be easier to solve than make). Let's get started!

Problems

1. Let n be the order of the smallest possible subgroup of $(\mathbb{R}, \cdot)$. Day n.
2. Day A, where A represents the area of any triangle tangent to the curve $y = 1/x$ in the first quadrant and having vertices at the origin, x-axis, and y-axis.
3. Let m be the 4th digit of τ. Day m.

The 4th digit from the left, in case you needed clarification. ;)

4. Day
$$2 \exp\left(\frac{8}{\pi}\int_0^1 \frac{\ln(1+x)}{x^2+1}\,dx\right).$$
5. Day $|S_5|/|S_4|$, where S_k denotes the symmetric group with degree k.
6. Let $S(p)$ denote the number of times a permutation p must be composed until it reaches the identity permutation. Day $S((1\ 3)(2\ 4\ 5))$.
7. Day D, where D denotes the absolute distance between the eigenvalues of the following 2×2 matrix:
$$\begin{bmatrix} 12099471394 & -8623006 \\ 121153998 & 12034827395 \end{bmatrix}.$$
8. Let N denote the number of distinct Fibonacci numbers that are perfect squares. Day 2^N.
9. Let Γ denote the unit circle, and define 13 circles $\omega_0, \omega_1, \ldots, \omega_{12}$ such that they are all of equal radius r and externally tangent to Γ. In addition, each circle ω_k is externally tangent to $\omega_{(k-1) \bmod 13}$ and $\omega_{(k+1) \bmod 13}$. Day $\lfloor 30r \rfloor$.

I am definitely not the most experienced in Euclidean geometry so this might be a weird way to phrase the question.

10. Big Boy Bill has a counter at value 0 and flips a fair, random coin 15 times. Every time he flips the coin, if it lands on heads, he adds 2 to the counter, and if it lands on tails, he adds 3 to the counter. There are m different sequences of coin flips in which at the end of the flipping, the resulting value is divisible by 5. Day $2m \bmod 23$.
11. For a base $b \geqslant 3$, define the *descending base number* $D(b)$ to be the $(b-1)$ digit number in base b where the ith digit is precisely i. For example, $D(4) = 321_4 = 57$. Let P be the number of descending base numbers that are prime. Day $P^2 - 6P + 16$.

Solutions

1. By the definition of a group, it must at least include the identity element, meaning that the order must be positive. In general, the smallest and perhaps most trivial subgroup of a group is one containing solely the identity element (for this specific case the group

is $(\{1\}, \cdot))$, thus our answer is $\boxed{1}$.

We shall give a general proof that the subgroup containing just the identity element exists.

Ehh this feels a bit elementary but it's fine.

Proof. We are given a group $(G, \cdot)$ with identity element e. We shall prove that $(\{e\}, \cdot)$ is a group and subsequently a subgroup of G. We already have fulfilled the identity element condition required in the definition of a group, so all that is left is to verify that the inverse of e is in the subgroup (which would mean that it is equal to e).

Suppose that $e^{-1} = e$, giving us

$$e \cdot e^{-1} = e^{-1} \cdot e = e \cdot e = e.$$

This fulfills the condition to be an inverse, and thus $(\{e\}, \cdot)$ is a group and a subgroup of G. ■

2. Let a be some positive real number. We shall find the line passing through $(a, 1/a)$ and tangent to the curve $y = 1/x$ to be

$$y = 1/a - 1/a^2(x - a).$$

Setting x and y to be zero individually and solving, we find that there is an x-intercept at $(2a, 0)$ and a y-intercept at $(0, 2/a)$. This forms a right triangle with area $1/2 \cdot 2a \cdot 2/a = \boxed{2}$.

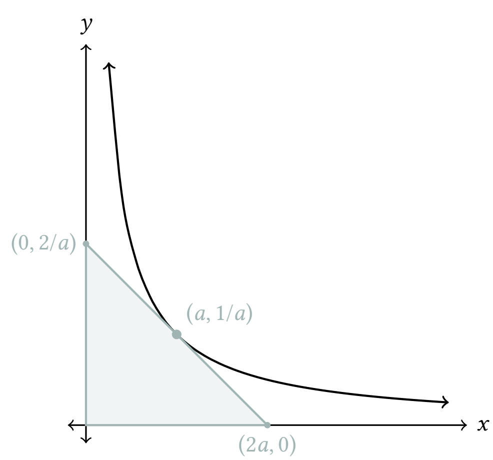

Figure 3. An example tangent triangle for Problem 2

I didn't put that much effort into making this problem as you can probably see.

3. Perhaps you can figure this out with proper analytical methods, but $\tau \approx 6.2832$, so uhhhh... The answer is $\boxed{3}$.

4. We parametrize the inner integral as such:

$$I(a) = \int_0^1 \frac{\ln(1 + ax)}{x^2 + 1}\, dx,$$

giving us an initial condition that $I(0) = 0$. We can then differentiate both sides with respect to a and do some integrating:

$$\begin{aligned}
I'(a) &= \int_0^1 \frac{x}{(1+ax)(x^2+1)}\,dx \\
&= \frac{1}{a^2+1}\int_0^1 \frac{x+a}{x^2+1}\,dx - \frac{a}{a^2+1}\int_0^1 \frac{1}{ax+1}\,dx \\
&= \frac{1}{2(1+a^2)}\ln\left|x^2+1\right|\Big|_0^1 + \frac{a}{a^2+1}\arctan x\Big|_0^1 - \frac{1}{a^2+1}\ln|1+ax|\Big|_0^1 \\
&= \frac{\ln 2}{2(1+a^2)} + \frac{\pi a}{4(1+a^2)} - \frac{\ln(1+a)}{a^2+1}.
\end{aligned}$$

Now we can solve for $I(1)$, which is exactly our desired integral.

Yes that is tau you see in the integral bounds. Tau for tau day.

$$\begin{aligned}
I(1) &= I(0) + \int_0^1 I'(\tau)\,d\tau \\
&= \frac{\pi\ln 2}{8} + \frac{\pi \ln 2}{8} - \int_0^1 \frac{\ln(\tau+1)}{\tau^2+1}\,d\tau \\
&= 2\cdot\frac{\pi\ln 2}{8} - I(1) \\
\implies I(1) &= \frac{\pi\ln 2}{8}.
\end{aligned}$$

Now we can plug this in and simplify the given expression:

$$\begin{aligned}
&2\exp\left(\frac{8}{\pi}\int_0^1 \frac{\ln(1+x)}{x^2+1}\,dx\right) \\
&= 2\exp\left(\frac{8}{\pi}\cdot\frac{\pi\ln 2}{8}\right) \\
&= 2\exp(\ln 2) \\
&= \boxed{4}.
\end{aligned}$$

5. The order of the symmetric group S_k is simply $k!$, so the answer is $5!/4! = \boxed{5}$.

6. One can either manually multiply the permutation out until it reaches the identity permutation or realize that because it is factored into two disjoint cycles of length 2 and 3, it will require $\gcd(2,3) = \boxed{6}$ compositions to return to the identity permutation.

7. The eigenvalues of the matrix are given by the solutions to the following polynomial:

$$(\lambda - 12099471394)(\lambda - 12034827395) + 8623006\cdot 121153998 = 0.$$

Notice that distance between the roots is exactly the discriminant, which is

$$\begin{aligned}
D &= \sqrt{(12099471394 + 12034827395)^2 - 4(12099471394\cdot 12034827395 + 8623006\cdot 121153998)} \\
&= \boxed{7}.
\end{aligned}$$

8. It has been proven that the only three perfect square Fibonacci numbers are $0, 1,$ and 144. Thus, our answer is $2^3 = \boxed{8}$.

While I would have liked to give a proof of this statement it turns out to involve a bit of machinery.

9. This one took me a bit and a half to compose. Essentially what I'm trying to convey is that the each circle ω_k is tangent to the previous and next circle and the center unit

Future Rushil here. Hahahahaha I overcomplicated this solution a lot. If you consider a 13-gon with radius $1+r$ and then find the distance between the vertices with some light trig, everything pops out quite well (credit to Kelly for that :skull:).

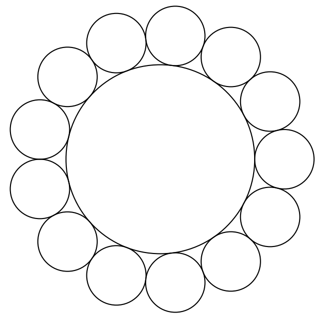

Figure 4. The idea Problem 9 is trying to convey.

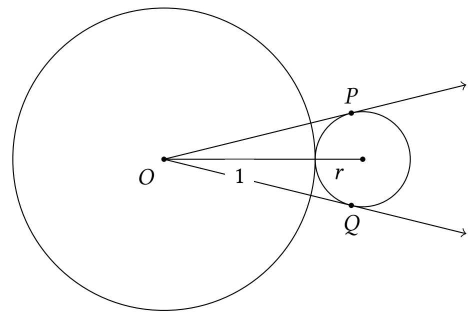

Figure 5. A diagram for the solution to Problem 9.

circle, forming a ring of circles.

Construct a circle, we'll call this ω_0 of some unknown radius r that is externally tangent to Γ. For simplicity, we'll pull it to the exact right of Γ. Let O be the center of Γ and P, Q respectively be the points lying on the two tangent lines to ω_0 extending from O. Notice that $2\pi/\angle POQ$ should be the number of circles that one can fit around Γ, in other words, 13.

Notice that if we find the slope a of $\overrightarrow{OP}$, we can find $\angle POQ = 2\arctan a$. This motivates us to use analytic methods .

There might be a more "geometric" solution, but this is the method that I used.

If we take O to be the origin, we can represent ω_0 by

$$(x - (1 + r))^2 + y^2 = r^2,$$

and we can let $\overrightarrow{OP}$ be represented by the line $y = ax$. If we substitute this expression for y into the equation for ω_0 and set the discriminant equal to 0 (because this line is tangent to ω_0), we can obtain an expression for a in terms of r as follows:

$$\begin{aligned} \omega_0 \colon 0 &= \left(a^2 + 1\right)x^2 - 2(1 + r)x + (1 + r)^2 - r^2 \\ \implies 0 &= 4(1 + r)^2 - 4\left(a^2 + 1\right)(2r + 1) \\ a &= \frac{r}{\sqrt{2r + 1}}. \end{aligned}$$

We can now substitute this into our condition of 13 circles to get

$$\begin{aligned} \frac{2\pi}{\angle POQ} &= 13 \\ \implies \frac{\pi}{\arctan a} &= 13 \\ \implies \frac{r}{\sqrt{2r + 1}} &= \tan\left(\frac{\pi}{13}\right). \end{aligned}$$

Solving this for the positive root gives us $r \approx 0.3146$, which means that $\lfloor 30r \rfloor = \boxed{9}$.

BBB is the shorthand for Big Boy Bill, for those unaware.

10. Suppose BBB flips k heads. Necessarily, this means he must have flipped $15 - k$ tails. From this, we can calculate the ending counter value and then take it modulo 5 to find the values of k that make our condition hold. Doing so, we can see that

$$\begin{aligned} 2k + 3(15 - k) &\equiv 0 \pmod 5 \\ \implies k &\equiv 0 \pmod 5. \end{aligned}$$

As such, the only legal values of k that make the counter value divisible by 5 are 0, 5, 10, and 15. We're now looking for sequences of flips where we *choose* 0, 5, 10, or 15 heads, which we can count to be

$$\binom{15}{0} + \binom{15}{5} + \binom{15}{10} + \binom{15}{15} = 6008.$$

Now all that's left is to see that $2 \cdot 6008 \bmod 23 = \boxed{10}$.

11. We write the summation representation of $D(b)$ first:

$$D(b) = \sum_{i=1}^{b-1} ib^{i-1}.$$

Motivated by this representation, examine $D(b)$ modulo $(b-1)$. Because we are working in base b, this is equivalent to finding the sum of the digits modulo $(b-1)$. In other words,

$$D(b) \equiv b(b-1)/2 \equiv 0 \pmod{(b-1)}.$$

In the case that b is even, this is 0, meaning that $D(b)$ is not prime. We now must check the case that b is odd. Let $b = 2k+1$. Notice that $D(b) = D(2k+1)$ is also the sum of its digits modulo k. Equivalently,

$$D(2k+1) \equiv k(k-1) \equiv 0 \pmod{k}.$$

Thus for odd b too, $D(b)$ is not prime. We do, however, have one extraneous case for $b = 3 = 2 \cdot 1 + 1$. Since all numbers are divisible 1, we cannot apply this conclusion to it. Checking manually to see that $D(3) = 7$ is prime, we have that $P = 1$. Thus, our answer is $P^2 - 6P + 16 = \boxed{11}$.

Building Geometry, Analytically 2023-06-30

Combo Calisthenics 2023-07-06

Combinatorics is a really beautiful subject in that, in premise, it can be very simple, but in practice it actually takes a decent bit of thinking and cleverness sometimes. Being so widely defined as the science of counting "things," combinatorics also lends itself very well to intersections with other subjects. You can have geometry, number theory, algebra, probability, and many others mixed into a combinatorics problem, making them really fun and sometimes quite the challenge.

I'd probably say that in competition math, combinatorics is up there for my favorite subject

While it certainly is fun to solve these problems, I think it's also perhaps quite instructive for me (and perhaps the readers) to create my own combinatorics problems and gather a wide variety of practice problems. Creating my own problems will be the perfect opportunity to get more acquainted with some of the cool tricks and structure to problems, helping me in furthering my combinatorics adventure and giving me some good exercise.

Calisthenics is a category of exercises which rely primarily on gravity and body weight. I thought it would be a funny name for this entry mainly because my friends have roped me into it.

In the process of making these, I'll probably put down whatever comes to mind really, so some problems might be really easy or perhaps insanely cumbersome to work with; nevertheless, it will still be a nice catalogue of my journey through the world of counting.

As for solutions, I'll probably write down some of the cooler ones whenever I have time, but I don't think I'll be able to get to all of the problems.

I hope to include not only competition math “flavored” problems which have slick combinatorial arguments, but also some more exploratory problems perhaps using generating functions and more of the higher level combinatorics stuff. With that said, let’s get to problem synthesizing!

General Combinatorics

Problems:

1. How many partitions of the set $\{1, 2, \dots, 20\}$ have at least 10 elements?

For those with $n < 12$, where n denotes the number of homies you have, you may borrow some for the sake of mathematics.

2. Exactly 12 of your homies, denoted $h_1, h_2, \dots, h_{12}$, are arranged in a circle. Starting from the first homie, h_1, and going in order, each homie pulls a string from themselves to another homie, so long as that string does not pass over any other strings in the process. All connections between homies must be made inside the circle. This process is continued until no more connections can be made. How many end states are there?

Geometry

Problems:

This could potentially be infinite or trivial for large n, but for $k = 2$, I at least know there’s some interesting stuff to play around with.

1. Suppose we have n circles in the plane. Allowing for the adjustment of position and size of all circles, what is the maximum number of lines one can place that are tangent to k circles?

Number Theory

Problems:

1. How many pairs of positive integers (a, b, c) satisfy

$$a + b + c = 23 \quad \text{and} \quad a^2 + 2b = 40?$$

2. How many partitions of the set $\{1, 2, \dots, 50\}$ can one make where the least common multiple of all elements in the partition is 50?

Algebra

Problems:

1. How many polynomials of the form

$$P(x) = x^2 + ax + b$$

have at least one real root, for $a, b \in \{1, 2, \dots, 50\}$?

Colorings

Problems:

1. (2019 CMIMC Combinatorics #2) How many ways are there to color the vertices of a cube either red, blue, or green such that no two adjacent vertices are the same color?

2. (2019 AMC 12A #13) How many ways are there to paint each of the integers $2, 3, \dots, 9$ either red, green, or blue so that each number has a different color from each of its proper divisors?

Iterating Graphs 2023-07-18

Recently, I've been getting into graph theory with my interest in competitive programming and other math problems , and one day I had the idea of creating an animation of what *all* the (undirected, without loops) graphs look like for some n vertices, and creating some aesthetically pleasing looking visualization that moves between all of them. This journal entry was really inspired by the math that goes behind it and how one actually goes about (somewhat) efficiently generating these graphs.

Hopefully you'll see more of this quite soon ;)

The end result ended up turning out to be quite nice (at least in my opinion) even if it was quite a simple visualization in idea. You can check it out here: `https://chirprush.github.io/animations/animations/graph-iterations/index.html`.

Encoding Graphs

The idea that was sort of swimming around in my head was how one would encode a graph as something that could be easily manipulated and played around with. One could take the simple route, storing a bunch of vertices and then collecting all the edges as tuples holding the vertex information, but this is quite bulky and certainly no fun, right? Surely we can compress a graph into something smaller than that; We just first need to revisit what exactly makes a graph a graph.

A graph is formally defined in terms of both its vertices and its edges, which are connections between these vertices. For our purposes of seeing "all" graphs though, we aren't really that concerned with the infinitely many positions of these vertices, so we can just place them evenly spaced on a circle for convenience and the visual appeal of n-gons. With vertices out of the way, all that really determines our different graphs is the edges that connect our n vertices. Since we're looking for a *bijection* from some space (which we'll call S) to this space of graphs, a good motivating question is: for n vertices, how many graphs do we exactly have?

A simple combinatorial argument helps us out here. For each pair of vertices, you can make the binary choice to either add an edge or leave it empty. Counting the number of possible pairs gives us

$$\binom{n}{2} = \frac{n(n-1)}{2} \text{ possible pairs, } \qquad \text{giving} \qquad 2^{n(n-1)/2} \text{ possible graphs.}$$

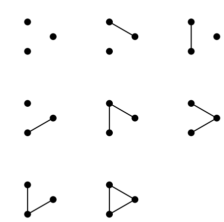

Figure 6. All 8 possible graphs for 3 vertices.

This means that, for example, there are $2^3 = 8$ total possible ways to pick the edges on a graph with 3 vertices. Note that while this does count graphs that are isomorphic to each other, including obvious "rotations," we're fine with this. After all, our goal is to show *all* graphs.

With this, we know our encoding must map from a set of $2^{n(n-1)/2}$ elements to one unique graph, *i.e.* $|S| = 2^{n(n-1)/2}$. We haven't really specified what the elements of S should be, but one of the most useful and easy to work with choices is simply the set of the integers

We will see later that this choice to start counting from 0 is not merely because I'm a programmer.

$$S = \{0, 1, \dots, 2^{n(n-1)/2} - 1\}.$$

This allows us to easily count up to some natural numbers and generate our encoded graphs quite easily, which can then be decoded.

What we haven't yet chosen now is *how* each number corresponds to a graph and *why*. There are quite a few ways of doing this ($|S|!$, in fact), but we would like to choose one that makes both intuitive and algorithmic sense. For this, we'll take a look a few steps back in our approach, specifically to where we counted the number of graphs.

One structure associated with graphs that further illustrates the *binary* action of choosing whether or not to add an edge between two vertices is an **adjacency matrix**. Recall that in graph theory, we define the adjacency matrix of a simple graph to be an $n \times n$ matrix where each entry A_{ij} is either 1 if there exists an edge between the ith and jth vertices and 0 if there does not. For example, for the graph C_4 (the cyclic graph with 4 vertices), the adjacency matrix is

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Because this is a undirected graph, the matrix is symmetric about the diagonal, and all we need to consider is one of the triangles (the one above or below the diagonal). One can verify that this triangle has exactly the $n(n-1)/2$ bits representing our edges. Thus, we see something interesting here. If we *flatten* the adjacency matrix triangle, we can represent our graph as a number in binary form. This also perfectly fits all our numbers into the domain set S.

A quick way to show this is that the diagonal takes up n elements, so if we let T be the number of elements in each triangle, we have $2T + n = n \times n$.

For example, in the case of the C_4 graph, we would encode it as $G = 101101_2 = 45$. Although in this case the number is a palindrome in base 2, we would place the bits according to the order they appear in the triangle. The first bit (from the right) would correspond to the edge $\{1, 2\}$, the second corresponding to $\{1, 3\}$, and so on.

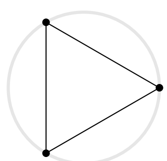

Figure 7. The animation displaying the complete graph K_3, encoded as $111_2 = 7$.

With this, the animation also has the cool effect of sort of counting up in binary with the edges. This also gives a somewhat natural progression from the empty graph to the complete graph for n vertices.

Thus, we've found a suitable encoding for our graphs into numbers. In our actual program, we're iterating over numbers and converting them to graphs instead, so we'll also have to find some way of going the other way around. Luckily, we've already built the foundation and intuition for how this will work.

Decoding Graph Numbers

If the encoding stage involves taking a graph and turning it into a number based on its adjacency matrix and edges, the decoding stage is precisely the opposite of this: we take a number, decompose it into it's binary representation, and turn it into an adjacency matrix, which defines the edges and the complete representation of our graph.

There are quite a few ways do this, but in order to choose a sensible one, we'll look at the edges of the graphs, specifically the vertices that they connect. If we number (quite arbitrarily) the vertices of a graph $1, 2, \ldots, n$, we have the following unique (undirected) edges:

$$\begin{array}{cccc} \{1,2\}, & \{1,3\}, & \cdots, & \{1,n\} \\ \{2,3\}, & \cdots, & \{2,n\} & \\ \vdots & \iddots & & \\ \{n-1,n\} & & & \end{array}$$

We want each bit of our "number to be decoded" to correlate to an edge in this triangle, and to do so, we work from right to left in the binary representation of the number. If the bth bit from the right is 1, we draw in the bth edge identified in this triangle. Mathematically, we can express this as a fairly straightforward-to-follow algorithm,

```
/*
   We let n be the number of vertices and b be a value from 1 to
   n(n - 1)/2 representing an edge and a bit in the graph number.
*/
function GetEdge(n, b)
    for i in 1, ..., n − 1 do
        if b ⩽ n − i then
            return {i, b + i}
        else
            b ⟵ b − (n − i)
        end if
    end for
end function
```

or a closed form expression of which I shall omit the derivation because it's a bit cumbersome and rather uninteresting. Let $k = n(n-1)/2 - b$. Then we have edge $\{i, j\}$ where

$$\begin{aligned} i &= 1 + \left\lfloor (n-1) - \left(\sqrt{1+8k} - 1\right)/2 \right\rfloor, \\ j &= 1 + i + b - (i-1)(2n-i)/2. \end{aligned}$$

I mostly derived the closed form because I thought it would be just a bit fun. For anyone curious on how I arrived at it, one can recall the formula for the sum of the first n natural numbers. If we set this equal to a number, floor the result, we obtain the formula with a bit of work.

And with this, we have pretty much most of the logic behind the graph iteration animation (well, besides the drawing logic, which is cool although not the most mathy). Be sure to check it out! I think it came out quite well.

Summing, Probably 2023-08-08

Here's quite a beautiful probability problem with a decent story behind it.

> **Problem.** Let $x_1, x_2, \ldots, x_n$ be random variables chosen uniformly from the interval $[0, 1]$. What is the probability that
>
> $$x_1 + x_2 + \cdots + x_n \geqslant 1?$$

Originally this was supposed to be Entry 007 in the journal, but for the longest time I had been working on the problem and hadn't gotten a solution. It was only in the summer after talking with some friends and thinking about it on and off for around a week or two that I made progress and realized a bunch of different things. As such, I've felt it only fair to move it to the end and change the journal entry date to match when I'm actually writing this now.

One of the things that makes solving this problem potentially fun is, if you're more familiar with contest math probability problems, they usually turn out to be more discrete in nature. What makes this problem a breath of fresh air in a way is that it's continuous, and one can't really reduce it to a combinatorics problem. This certainly encourages a different way to go about problem solving. Case work is instead transformed into different methods of attacking the problem.

Ah yes my favorite style of proof: proof by exhaustion on the real numbers.

In total, this problem is a really great example of how the journey matters far more than the destination. When I was looking for different ways to approach the problem, I knew full well the correct answer already. While I certainly think the final result is interesting in itself, I had the most fun simply talking with friends about the problem, researching different solution strategies, visualizing 4D, falling down a rabbit hole involving simplices, learning

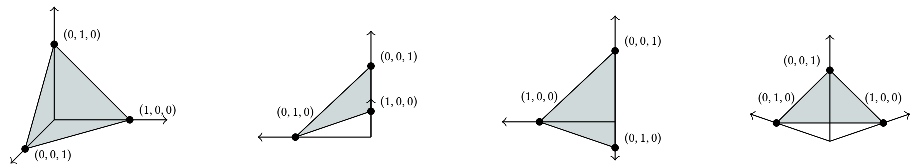

Figure 9. Different views of a 3-simplex.

about probability distributions, and so much other stuff.

I hope you enjoy this problem even a little as much as I did!

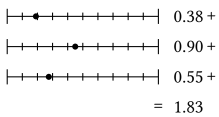

Figure 8. One realization of the state of the sliders and their sum for $n = 3$. We see that this state is included in the set of solutions for which the sum is greater than 1.

Intuiting the Problem

Before I jump into any solutions or a bunch of math, let's first take a moment to really digest what the problem is asking for. Imagine we have n sliders going from 0 to 1, which we can freely move around and adjust to our choosing. We can scramble these sliders however we want, and after we're done, we take the sum of the sliders and check whether it's less than 1 or greater than 1. We want to know the probability that after scrambling these sliders in any sort of way, we get a sum greater than 1.

One thing to note is that frequently throughout the different solution methods, we consider the complement of the probability as it is far easier to work with for some geometric solutions. When talking about the complement, we are referring to the probability that the sum is less than 1. (In fact, I actually considered changing around the problem to directly look at the complement, but in ordered to better encapsulate my progression throughout the entire problem and such, I deemed it better to simply leave the original problem statement how I found it. Perhaps this might change in the future if I ever find myself mulling over it too much.)

Another thing to note is that, rather intuitively, the probability should increase as we increase n, as there is a greater and greater chance of the sum being greater than 1. This means that our desired closed form for the probability should approach 1 as $n \longrightarrow \infty$. Working in tandem, this means our complement probability (the probability that the sum is less than 1), shall approach 0.

Calculus Methods

Geometric Methods

▷ A Not So Brief Interlude On Simplices

Algebraic Methods

Probabilistic Methods

This is where the real spicy stuff happens, and where I believe this problem truly shines. It's in fact this solution that I came up over the course of a couple days which really inspired me to get more interested in probability (especially probability distributions) as well as mathematical statistics.

Expect more probability journal entries soon? ;)

Some Afterthoughts

Expected Distance, Probably 2023-08-17

Yup that's right, we're back with another probability problem inspired by the beauty of solv-

Probability distributions are just so powerful and interesting right now, so I'm sorry, but I'm not sorry. I'm taking AP Stats this year, so this (hopefully, probably won't given AP curriculums) will work well.

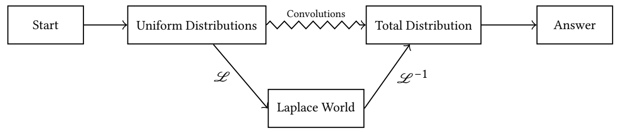

Figure 10. The solution path for solving the problem using probabilistic methods. Convolutions are too hard to work with even for small n, but we can sidestep them with Laplace transforms.

ing the last one. I can't quite make promises one whether this will be the last one, but I'll try to vary things up now and again hopefully. I'm planning on making a TODO list entry and as well as a reading/watch list entry in the near future, so look forward to more writing based entries :).

As opposed to the previous, explanation-heavy journal entry, I'll likely forego trying to intuitively explain, detail, and diagram *all* of the components of the problem because I'm assuming that the reader is at least a bit more familiar with some of the concepts as they overlap decently with the previous entry.

With that being said, let's get onto the main problem at hand. Inspired by previous probability problems, I thought it would be nice to compose another one involving some interesting geometrical ideas. As such, we have the following problem:

> **Problem.** Given two points chosen uniformly randomly from the unit square $[0,1]^2$, what is the expected distance between these two points?

This time, we seek out not the probability itself, but rather the expected value, which is a nice little difference from the past problem. I also really enjoy how innocently the question can be formulated because, as we will come to see, it's perhaps not the easiest to solve once you really take a good think about it.

I've yet to dive head deep in other solution methods though (the geometric one didn't seem to hopeful), so perhaps there are other ways of getting the answer?

Initially, this was inspired by a slightly harder probability problem , which has a slightly longer formulation and seems to be a decent bit harder, but I do intend to tackle it someday (and perhaps I might add it at the end of this journal entry).

Guys, I think he might like probability just a little bit.

Non-Probabilistic Methods

▷ Brute Forcing, Mathematically

One of the first methods I tried was applying some good old calculus to the problem. In particular, if one were to evaluate the following integral

$$D = \int_{[0,1]^2} \int_{[0,1]^2} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\| \, d\vec{\mathbf{w}} \, d\vec{\mathbf{v}} = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \, dy_2 \, dx_2 \, dy_1 \, dx_1.$$

In effect, this is "brute forcing" the answer. We go through each point in the unit square, and for this point, we go through the unit square once more and sum all the distances. Because the area of the unit square is simply one, we don't divide anything and are left with the average, or expected, value of the distance. The reader might now be thinking: "Ok Rushil, we have our integral, so why can't I just plug this into Mathematica and go on with my day?"

There's just a slight problem. I too had this train of thought, and I happily plugged the integral into Mathematica, eager to see what result would come out. Unfortunately though, this is a quadruple integral, and a painful one to evaluate at that given the square root. I let the program run until the jet engines noises start coming from my laptop fans, but it seems

My laptop certainly isn't a Gamer PC™, so there is a chance that it could do it on somebody else's computer, but given how daunting the integral looks to approach, I severely doubt it.

this integral is just far too difficult to evaluate in general, even for a computer using (what I presume to be at least) a modified Risch algorithm.

That's not to say that this doesn't yield us any information though. Either through Monte Carlo simulations or some numerical integration using trusty old SciPy, we do know that

$$D \approx 0.521405,$$

which at the very least intuitively makes some loose sense. This also gives us a good way to check that our closed form answer is correct whenever we're trying out another solution.

▷ A Small Aside on Calculus

One rather interesting thing to remark about this approach, though, is that the expected squared distance,

$$L = \int_{[0,1]^2} \int_{[0,1]^2} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|^2 \, d\vec{\mathbf{w}} \, d\vec{\mathbf{v}} = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right] dy_2 \, dx_2 \, dy_1 \, dx_1,$$

is actually quite trivial to calculate using this approach. The quadruple integral is rather easy to evaluate as the following will show, which just speaks for how the square root alone complicates things massively.

$$\begin{aligned}
L &= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right] dy_2 \, dx_2 \, dy_1 \, dx_1 \\
&= \int_0^1 \int_0^1 \left[\frac{1}{3}\left((1 - x_1)^3 + x_1^3\right) + \frac{1}{3}\left((1 - y_1)^3 + y_1^3\right)\right] dy_1 \, dx_1 \\
&= \int_0^1 \int_0^1 \left[\frac{1}{3}\left(1 - 3x_1 + 3x_1^2\right) + \frac{1}{3}\left(1 - 3y_1 + 3y_1^2\right)\right] dy_1 \, dx_1 \\
&= \frac{2}{3}\left(1 - \frac{3}{2} + 1\right) \\
&= \frac{1}{3}.
\end{aligned}$$

At the very least, this potentially gives ourselves another way to check our work in the future, and it's simply just interesting to think about how different norms impact the result vastly.

▷ Some Attempted Geometry

Another way one might tackle the problem is to look at some geometrical simplification, trying to utilize symmetries and the like to figure out some information. Unfortunately, this didn't seem to yield very many useful results for me, with no real hope of a solution in closed form.

If one draws out the unit square, we can pick a point anywhere in its interior and then radiate circles outwards at increasing radii in a way not too far from the brute force integral. We can additionally use symmetry to only consider points in of the corners of the square (if we rotate/reflect, we can always transform the square to get a point anywhere in one of the corners while preserving distances). In total, this works fine for relatively smaller radii, but as one gets to the circle intersection with the square, everything complicates, and one is unlikely to get a clean looking closed form solution.

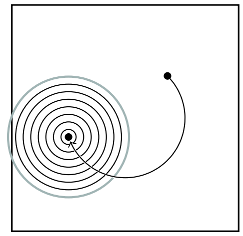

Figure 11. A diagram illustrating an attempted geometric solution.

Detectably Delicious Distributions

And such comes the time where we, stranded travelers on our journey through this puzzle, are enlightened by the beauty of the probability distributions. As opposed to straight brute force, I like to think of probability distributions as a more elegant way of encoding the problem (although they are in some ways brute force).

The basic plan for our probability distribution approach is that we will build up the norm, $\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$, using random variables described by distributions as opposed to regular old scalars. We know that all the coordinate components are uniformly distributed on $[0, 1]$, so it becomes a matter of combining these in a way that we can get the distribution of the lengths found on the unit square. We will achieve this with the following steps:

1. First, convolve our two uniform distributions to find $T = X_2 - X_1$. The support of this distribution should be $[-1, 1]$.
2. Next, take the distribution of the squares of the these x-differences to get $Q = T^2 = (X_2 - X_1)^2$. The support of this distribution should be $[0, 1]$.
3. Take the convolution of the resulting Q with itself (as the other part of the norm is the exactly the same) to get $Z = Q + Q = (X_2 - X_1)^2 + (Y_2 - Y_1)^2$. The support fo this distribution should be $[0, 2]$, and a good check that we have the distribution right is that the mean should be 1/3, according to our aside above.
4. To finish off the algebra on our distribution, we'll take the distribution of the square roots of the previous Z. Intuitively, this should have the support $\left[0, \sqrt{2}\right]$, as the greatest possible distance between two points on the unit square is the distance between two opposite corners. If all goes well, we'll be able to integrate and get the mean of the distribution, which will be the expected value of the distance between any two random points on the unit square.

With our mental roadmap laid out for us, let's get on with the probability distributions!

▷ Opposing Convolutions

The first step on our special soup distribution recipe is to take the difference of two uniform distributions spread across the interval $[0, 1]$. Equivalently, this is same as taking the convolution of the aforementioned uniform distribution with the negative of the same distribution. In particular, let $X_2 \sim \text{Uniform}(0, 1)$ and $X_1 \sim \text{Uniform}(-1, 0)$. We then desire the distribution $X_2 + X_1$, which we can find by convolving the two distribution PDFs.

I bet the distributions are delectably delicious.

Essentially, $X - Y = X + (-Y)$.

In this case, carrying out the convolution just visually is possible as the overlapping area between the two uniform distributions increases or decreases linearly as you slide one across another, but we may also verify our results formally. If we write our distributions out as functions, we get

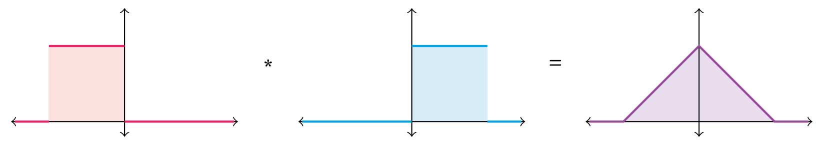

Figure 12. When we take the convolution of the two uniform distributions, we get a "unit triangle" around $x = 0$.

$$f_{X_1}(x) = \begin{cases} 1, & \text{for } x \in [-1, 0] \\ 0, & \text{otherwise} \end{cases}, \qquad f_{X_2}(x) = \begin{cases} 1, & \text{for } x \in [0, 1] \\ 0, & \text{otherwise} \end{cases}.$$

The notation f_X and F_X are going to be used throughout here to denote the PDF and CDF of X respectively.

With this, we can find the density function for the resulting distribution through the integral definition of convolution:

$$\begin{aligned} f_{X_2+X_1}(t) &= \int_{-\infty}^{\infty} f_{X_2}(x) f_{X_1}(t-x)\,dx \\ &= \int_0^1 f_{X_1}(t-x)\,dx. \end{aligned}$$

Using our visual intuition of sliding the distribution of X_1 along the x-axis and seeing the intersecting area with the distribution of X_2, there are two cases we need to look at:

1. **Case** $x \in [-1, 0]$: In this case, the area of the sum of the distributions is increasing as the two distributions overlap more and more. At $x = 0$, they are completely on top of each other. This part of the distribution should be

$$f_{X_2+X_1}(t) = \int_0^{1+t} dx = 1 + t.$$

2. **Case** $x \in (0, 1]$: In this case, the area of the sum of the distributions is decreasing as they steadily move away from each other, with $x = 1$ marking the point at which there is again no interlap. This component is given by

$$f_{X_2+X_1}(t) = \int_t^1 dx = 1 - t.$$

With this, our first step of the distribution recipe is completed. We have the following distribution representing the difference between components:

$$f_T(t) = \begin{cases} 1 + t, & \text{for } t \in [-1, 0] \\ 1 - t, & \text{for } t \in (0, 1] \\ 0, & \text{otherwise} \end{cases}.$$

The next step in our cooking journey will be to take the distribution of the squares of T.

▷ A Not So Square-Looking Square

As mentioned, we're looking for a probability distribution that tells gives us the distribution of the squares of T. Intuitively, this should have 0 probability for all values less than 0 or greater than 1. We are able to find this distribution by looking at the CDF of T^2 and T. One has

$$\mathrm{P}\left(T^2 \leqslant t\right) = \mathrm{P}\left(T \in \left[-\sqrt{t}, \sqrt{t}\right]\right) = 2\,\mathrm{P}\left(T \in \left[0, \sqrt{t}\right]\right).$$

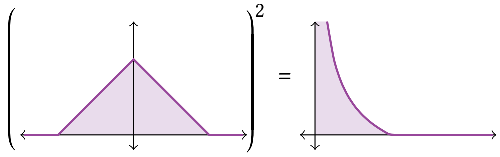

Figure 13. The distribution of the squares of the differences in components. It looks rather interesting as it approaches infinity on one side.

In short, we have the CDF of T^2 in terms of the CDF of T, and because we can find the CDF of T given the PDF of T quite easily, we're able to get our desired distribution for T^2 with just a little bit of work. The above relations tell us that

$$\begin{aligned} F_{T^2}(t) &= 2\int_0^{\sqrt{t}} f_T(x)\,dx \\ &= 2\int_0^{\sqrt{t}} (1-x)\,dx \\ \implies f_{T^2}(t) &= 2\frac{d}{dt}\int_0^{\sqrt{t}} (1-x)\,dx. \end{aligned}$$

Note that we were able to only use the positive component of the distribution of T due to its symmetry across $x = 0$. Now applying a little bit of the fundamental theorem of calculus, we get a decently reasonable expression for the PDF of T^2.

$$f_{T^2}(t) = f_Q(t) = \frac{1}{\sqrt{t}} - 1.$$

This distribution is quite an interesting one (see Figure 13), as it approaches infinity as $t \longrightarrow 0^+$ and it hits 0 at $t = 1$. With this, we have our second step of our recipe done, and we're almost halfway to the goal. In the next section, we'll need to take this distribution and convolve it with itself once more.

▷ More Beautiful Convolutions

Once more we must take the convolution of this distribution with itself. Unlike how the case was with the uniform distribution convolutions, though, this is a bit more to work through especially given that the distributions are no longer described by simple functions. Once again, we'll have to use our visual intuition to split the domain into separate cases.

Using the convolution formula once again, we have

$$\begin{aligned} f_{Q+Q}(q) &= \int_0^1 f_Q(x) f_Q(q-x)\,dx \\ &= \int_0^1 \left(\frac{1}{\sqrt{x}} - 1\right) f_Q(q-x)\,dx. \end{aligned}$$

We once again must split this case-wise based on q:

1. **Case** $q \in [0, 1]$: For this case, the lower bound is 0 and the upper bound is q. This gives us

$$f_{Q+Q}(q) = \int_0^q \left(\frac{1}{\sqrt{x}} - 1\right)\left(\frac{1}{\sqrt{q-x}} - 1\right) dx.$$

2. **Case** $q \in [1, 2]$: In this case, the lower bound is $q - 1$ and the upper bound shall be 1.

$$f_{Q+Q}(q) = \int_{q-1}^{1} \left(\frac{1}{\sqrt{x}} - 1 \right) \left(\frac{1}{\sqrt{q-x}} - 1 \right) dx.$$

In order to not repeat ourselves and do too much work, we'll find the general antiderivative of the shared integrand between these two cases. Expanding yields the following

$$\begin{aligned}
\int \left(\frac{1}{\sqrt{x}} - 1 \right) \left(\frac{1}{\sqrt{q-x}} - 1 \right) dx &= x - 2\sqrt{x} + 2\sqrt{q-x} + \int \frac{1}{\sqrt{x}\sqrt{q-x}}\, dx \\
&= \cdots + 2 \int \frac{1}{\sqrt{q-u^2}}\, du \\
&= \cdots + \frac{2}{\sqrt{q}} \int \frac{1}{\sqrt{1 - \left(\frac{u}{\sqrt{q}}\right)^2}}\, du \\
&= x - 2\sqrt{x} + 2\sqrt{q-x} + 2 \arcsin\left(\frac{\sqrt{x}}{\sqrt{q}} \right) + C
\end{aligned}$$

This is achieved using the relation $\arctan x = \arcsin(x/\sqrt{1-x^2})$. While one can proceed with using arcsin, flipping to arctan makes future integration much easier.

So, for the first case, the distribution looks something like $\pi - 4\sqrt{q} + q$, and the second case will be something like $4\sqrt{q-1} - y - 2 + \pi - \arctan(\sqrt{y-1})$. Written in a more convenient form, we get the full distribution to be something like

$$f_{Q+Q}(q) = \begin{cases} \pi - 4\sqrt{q} + q, & \text{for } q \in [0, 1] \\ \pi - q - 2 - 4\arctan(\sqrt{q-1}) + 4\sqrt{q-1}, & \text{for } q \in (1, 2] \end{cases}.$$

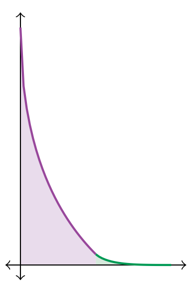

Figure 14. The distribution of the squared distances. It looks rather odd in a way, and the presence of π terms seems quite interesting.

This distribution in fact does *not* approach infinity as $q \longrightarrow 0^+$, and looks rather interesting, although similar in shape to the distribution for Q. Indeed a good check that we're on the right path is showing that

$$\int_0^2 z f_Z(z)\, dz = \frac{1}{3},$$

which we found in the non-probabilistic solution attempts section. Now all that's left to do is take the distribution of the square roots, and we'll be able to integrate to find the mean and our desired answer.

▷ The Last Stretch and Some Fancy Integrals

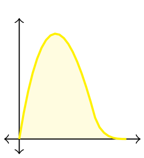

Figure 15. The final distribution and the distribution of all distances between two points found on the unit square. This particular distribution has a somewhat intuitive shape. Note however that the mean is not the extreme value, although they look somewhat close.

Doing something similar to the squaring section, we'll take a look at the CDFs of $\sqrt{Z}$ and Z and find a relation between them, which will allow us to get the final distribution. In particular,

$$\begin{aligned}
F_{\sqrt{Z}}(z) &= \mathrm{P}\left(\sqrt{Z} \leqslant z\right) = \mathrm{P}\left(Z \leqslant z^2\right) = F_Z\left(z^2\right) \\
\Longrightarrow f_{\sqrt{Z}}(z) &= \frac{d}{dz} \int_0^{z^2} f_Z(x)\, dx = 2z f_Z\left(z^2\right).
\end{aligned}$$

We want to take the mean of this distribution, which can be found using

$$\int_0^{\sqrt{2}} z f_{\sqrt{Z}}(z)\, dz = 2 \int_0^{\sqrt{2}} z^2 f_Z\left(z^2\right) dz.$$

This once again must be split into two cases given the piecewise nature of the function, and now we have some nice integrals to evaluate.

$$
\begin{aligned}
&= \int_0^1 \left(2\pi z^2 - 8z^3 + 2z^4\right) dz \\
&+ \int_1^{\sqrt{2}} \left(2\pi z^2 - 4z^2 - 2z^4 + 8z^2\sqrt{z^2-1} - 8z^2 \arctan\left(\sqrt{z^2-1}\right)\right) dz.
\end{aligned}
$$

The first integral is easy enough to evaluate, and I'll cut straight to the simplification for that one, but two of the terms in the last integral aren't as trivial to integrate in one's head, so we'll focus on those. Before moving on, let's prove the following useful result for future use.

Lemma.

$$\int \sec^n \theta \, d\theta = \frac{1}{n-1} \sec^{n-2}\theta \tan\theta + \frac{n-2}{n-1} \int \sec^{n-2}\theta \, d\theta.$$

Proof. Using integration by parts, take

$$
\begin{aligned}
u &= \sec^{n-2}\theta \quad & du &= (n-2)\sec^{n-2}\theta \tan\theta \, d\theta \\
dv &= \sec^2\theta \, d\theta \quad & v &= \tan\theta
\end{aligned}
$$

This gives us

$$
\begin{aligned}
\int \sec^n \theta \, d\theta &= \sec^{n-2}\theta \tan\theta - (n-2)\int \left(\sec^n\theta - \sec^{n-2}\theta\right) d\theta \\
\implies (n-1)\int \sec^n \theta \, d\theta &= \sec^{n-2}\theta \tan\theta + (n-2)\int \sec^{n-2}\theta \, d\theta \\
\implies \int \sec^n \theta \, d\theta &= \frac{1}{n-1}\sec^{n-2}\theta \tan\theta + \frac{n-2}{n-1}\int \sec^{n-2}\theta \, d\theta.
\end{aligned}
$$

Thus completes the proof. ∎

Now that this nice tool at our disposal, we can go ahead and tackle the two "non-trivial" integrals in the expression:

1. For the first integral we can do a trig substitution, setting $z = \sec\theta$ to yield

$$
\begin{aligned}
&\int_1^{\sqrt{2}} 8z^2\sqrt{z^2-1}\, dz \\
&= 8\int_1^{\sqrt{2}} \sec^3\theta \tan^2\theta \, d\theta \\
&= 8\int_0^{\pi/4} \left(\sec^5\theta - \sec^3\theta\right) d\theta \\
&= 4\sqrt{2} - \sqrt{2} + \ln\left(1 + \sqrt{2}\right) \\
&= 3\sqrt{2} + \ln\left(1 + \sqrt{2}\right).
\end{aligned}
$$

2. For the second integral, we can use IBP to differentiate the arctan term and integrate

the polynomial term to get

$$\begin{aligned}
&\int_1^{\sqrt{2}} 8z^2 \arctan\left(\sqrt{z^2-1}\right) dz \\
&= \frac{4\sqrt{2}\pi}{3} - \frac{8}{3}\int_1^{\sqrt{2}} \frac{z^2}{\sqrt{z^2-1}}\, dz \\
&= \frac{4\sqrt{2}\pi}{3} - \frac{8}{3}\int_0^{\pi/4} \sec^3\theta\, d\theta \\
&= \frac{4\sqrt{2}\pi}{3} - \frac{4}{3}\left(\sqrt{2} + \ln\left(1+\sqrt{2}\right)\right) \\
&= \frac{4\sqrt{2}\pi}{3} - \frac{4\sqrt{2}}{3} - \frac{4}{3}\ln\left(1+\sqrt{2}\right).
\end{aligned}$$

One can also write the logarithm term in terms of inverse hyperbolic sine, but that feels a little obfuscated.

Now that we have these results, it becomes a matter of combining terms and simplifications, which I won't bore you too much with. After we cancel a couple of terms down, we get the following expression for the average distance between two randomly picked points on the unit square:

$$D = \boxed{\frac{1}{15}\left(2+\sqrt{2}+5\ln\left(1+\sqrt{2}\right)\right)}.$$

Numerically, this does check out with both the brute force integral as well as the Monte Carlo methods, so we truly do have our final answer. One of the things that I found most surprising was that the answer doesn't really look all that clean or intuitive or anything. All of the π terms magically cleared, so there's no real circles to intuit. Given the nature of the solution, it really doesn't feel that there is a good geometric way to solve the problem (and get a closed form at least) as I suspected at the beginning.

This was quite the fun problem!

The Aforementioned Parent Problem

As somewhat mentioned in the introduction, the idea for the problem we just solved was born alongside a slightly harder problem that I'd like to at least put down here along with my thoughts. The problem follows as such:

> **Problem.** Suppose we randomly place n points on the unit square and number these points accordingly $1, 2, \dots, n$. In addition, extend lines between points k and $k+1$ for $k \in \{1, 2, \dots, n-1\}$. What is the probability that none of the edges between the points intersect?

Certainly, while there are some similarities, it is also a decently different problem. I have yet to fully tackle this problem (and I have other things that I would like to work on as well, so it won't be done in the near future), but at least my main starting point would be to look at the angles formed between the points. An n-gon has a certain interior angle sum, and so going along this thinking, we could create some conditions on the individual angles or the sum of these angles that leads us in the right direction. Certainly, if an angle is far to narrow, we might find an intersection. If all angles are perhaps larger than a certain threshold, we can say for certain that there is no intersection.

One may notice for the $n = 3$ case there is a line segment of points where we can place the third point which folds back upon the first point, but this realistically has zero measure, so it doesn't contribute (at least I think).

In addition, as a general problem solving tip, it will help to look at smaller cases. One can easily see that for $n \leqslant 3$ has a probability of 1 for there being no intersections. From there, we can build up our geometric intuition of the problem.

I look forward to seeing what sort of form both the solution structure as well as the closed form answer (assuming there is one) takes on.