Notes on Diophantine Equations

Aaron Pierce

1 A Highly Skippable Introduction

I was working on the 2020 Advent of Code when the second part of the puzzle for day 13 absolutely ruined me. For a quick rundown of the puzzle, there are a fleet of buses that depart at regular intervals. Each bus has a number and a time interval associated with it. So there may be a 5 minute bus with number 3, and a 13 minute bus with number 0. The number is arbitrary and doesn't have to be sequential. The time means at what time interval the bus will depart from the station. So a 5 minute bus will leave at t=0, t=5, t=10, t=15 and so on. The number is an offset. It makes more sense if you're actually doing the puzzle, but that offset is the core of the puzzle.

We want to know when the buses depart at nearly the same time. For a time t when the 0-offset bus departs, when does every bus depart their offset after that. So if the 13 minute bus has the 0 offset and it departs at t=130, the other bus would have to depart at 303 (for the 5 minute, 3 offset bus). However, the 5 minute bus can't depart at t=303 because 303 isn't a multiple of 5. This is one random example. The actual problem is to find the earliest value of t when the buses depart with this condition.

If my problem explanation wasn't clear enough, it doesn't really matter, what follows is what actually matters. You could model these buses as lines. So the 13 bus will depart at $t = 13x$ for any integer x. Then for that t, we also want to know if there is an integer solution for $t = 5y - 3$. $5y$ represents the times at which the 5-minute bus normally departs and the minus 3 is the offset. If $5y-3 = 13x$ then the 13-minute bus will depart near the 5-minutes bus, which is what we want our solutions to be. However, x and y have to be integers. That's a big however. If they could be real numbers then the buses would depart all the time.

All of math that you are taught in school is for continuous functions! Calculus is all about continuous functions, algebra always allowed any real number to be a solution, but now we only have integers, and it's hard to reason about only having integers when you are used to a continuous set of tools. Let's go grab some decimal-free tools then.

2 Diophantine Equations

A Diophantine equation is an equation where we only care about integer solutions. So our bus equations are Diophantine! We only care about integer multiples of the bus' interval. Basic Diophantine equations look like lines. Something like $y = 2x + 1$. When I see this form I get to thinking about continuous functions. It evokes the idea of picking some value for x to find a y, and it's easy to do that. Pick an integer x and boom, 2 times that number plus 1 is your integer y.

But if I write it as $1 = 2x - y$ it becomes a little scarier to look at. Two times what number is one more than another? That is way less intuitive than asking you to give me an x and I'll give you its associated y.

This is what our bus problem wanted though. We needed to know when $t = 5y - 3$ had integer solutions. For the lowest possible t and any y. This becomes more complicated. Before, y was dependent on x so we could find a y from an x. Now we have two unknowns, and they are related to each other. Perfect time for a Diophantine equation.

Let's re-write in the scary form, so that our new equation looks like $3 = 5y - t$. This makes it a little more clearer that some combination of two numbers gives us 3. It emphasizes the cooperation between y and t, instead of a dependence.

This is a classic Diophantine equation, and we'll see how we can solve it.

First off, there are equations that won't spit out a solution. Any line that can use real numbers will cover the whole domain and range because any x will produce a y. This is no longer the case for Diophantine equations. Consider $7 = 2x + 4y$. There is absolutely no way you can combine a multiple of 2 and a multiple of 4 into a 7. You'd have to use decimals.

So how do we know when there will be solutions? If we prime factorize our coefficients we can see what's going on. Our equations generally take the form $ax + by = c$. We want some number of a's and add them to some number of b's to make c.

If a and b share a factor, they aren't coprime, then both a and b are a multiple of their shared factor. Call it f. That's just how it works. If f is a factor of a then a is f times something else, and so is b.

If you add a and b, that addition will be $fu + fv$, (u and v don't matter) and you can factor out an f, so there is no possible way you can escape the f. It's like a horror movie or something. No matter how hard you try, you can't shake the f factor.

So in your Diophantine equation $ax + by = c$, if a and b share a factor (their lcd is not 1) then c also has to have that factor. If c didn't have that factor, then $ax + by$ would have to shake the factor somehow, which it can't do because of the horror movie thing. If c does have that factor, then you can divide both

sides by f and reduce your equation down. If you keep doing that for every common factor a and b and c share, then a and b will eventually be coprime, or your equation isn't solvable.

So every solvable Diophantine equation is equivalent to one with co-prime a and b. There you go. Three long winded pages to explain one sentence of the Wikipedia page.

Cool, so now that we know our equation is solvable and has coprime coefficients, how do we even solve it? One neat property is that if you have one solution, you can find all the other ones. This is a lot like having a point on a line and its slope. If you just walk that slope you'll find another point on the line. In our case, it's not so easy. If $5x + 7y = 31$, then $(2, 3)$ is one solution. If you want another solution, you have to maintain the perfect balance. Maybe having 3 y's is just too many for you, so you want to remove some. If you remove a single y, then you lose 7. So now $5(2) + 7(2) = 24$. However, we want it to equal 31, so we need to get to adding some more numbers. However, we can't, because there aren't any multiples of 5 that equal 7. In other words, we can't make one-for-one trades of y's for x's because they aren't equal. We can only take off as much as we can put back on. So if we want to remove some y's, we need to remove them so that 7 times the number of removed y's is a multiple of 5. Because 7 and 5 are coprime, this will happen when you have 5 y's. And conversely, you'll need to put back seven fives.

Okay so if I'm given a solution and I want a new one I need to take away multiples of the GCM of the coefficients of the equations. Cool. How do I actually solve these things. I can't just go guessing values of y and hoping there's an accompanying integer x.