

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

WILLIAM GREEN, All right, let's get going. So today, I'll say a few more words about PDEs, and then we'll leave
JR: that topic for a while. I think I'll actually come back and have a little lecture about that right before Thanksgiving, for those of you still around.

And then we'll start talking about probability, and then that will lead into several lectures about models versus data, which is a very important topic for all of you who either plan to generate data or plan to generate models during your stay here, which probably is everybody. So we'll talk about that for a while. So PDEs. I guess the first comment is, homework 7, have any of you looked at this yet?

AUDIENCE: [INTERPOSING VOICES]

[LAUGHTER]

WILLIAM GREEN, Not a single one.

JR:

AUDIENCE: [INAUDIBLE]

[LAUGHTER]

WILLIAM GREEN, So homework 7 is the same problem that Kristin showed in class in the demo for COMSOL,
JR: and I want you to solve it both ways. So solve it with COMSOL and then solve it writing your own finite volume code.

And I want to warn you. This problem has a characteristic length that's way smaller than the dimensions of the problem. And so in principle, you might need to use an incredibly fine mesh to resolve the gradients in the problem.

So just remember, in case you haven't looked at it, the problem is we have a drug patch. It has some concentration of the drug. The drug is diffusing slowly out of the patch into the flow, and

we have some flow here like that.

And the characteristic length is at the-- this is a velocity boundary layer. So the velocity in the x-direction is equal to y times something, dv/dy , I guess. This is some number, and so this has units of-- what is that-- per second, so like a strain rate, OK?

And the diffusion here is controlled by diffusivity D , and that has units of, say, centimeter squared per second. And so the D over dv/dy , this thing gives a characteristic length squared, which is sort of the natural length scale of this problem.

And the problem is that for the drug molecule, it's a big molecule. It has a very small diffusivity. And so therefore, this is a really tiny ratio, and so the L is very small. And similarly, if you look at it from the point of view in the x-direction, the Peclet number is really gigantic. And so both of those will tell you that you have to watch out, there might be very sharp gradients in the problem.

And if you just think of it physically, right over here we think the concentration's 0, somewhere over here. And all of a sudden right here, the concentration is going to be close to the concentration in the patch.

So there's almost a discontinuity in the concentration. So there's a really sharp gradient on the upstream edge. And then something funky is going to happen down here at the end of the patch, too. It won't be quite as abrupt, but could be pretty strange. All right, so you got it? OK.

And in the problem, we want you to figure out the drug diffusing all the way over to here somewhere, way far over there. And so you may need quite a few mesh points in the y-direction as well. All right. And this kind of problem, this is a very simple case, right? There's no reactions. The velocity's just in one direction, and this is not a very hard case.

But you'll see it's actually still pretty tricky to get the right solution. So don't just believe what the code tells you. Just run COMSOL and just-- don't believe it's showing the truth. And don't believe you just write down some finite volumes that you'll get the truth. So mess around with it and try to convince yourself it's really converged, and you really have the real physical solution.

Because we expect a sharp gradient, say, in the upstream edge, you might want to play with using a finer mesh in the x-direction here than you would down here, because down here

presumably the gradients in x-direction are much smaller. So you don't have to use square finite volumes. You could use rectangles. Yes?

AUDIENCE: I don't understand what you have written [INAUDIBLE] as v_x equal $y \, dv_x \, dy$.

WILLIAM GREEN, Yeah.

JR:

AUDIENCE: So what--

WILLIAM GREEN, So that's because the velocity-- the v_x is 0 at the wall, and the velocity increases with y . So v_x increases the further you get from y . This is y equals 0, always going up that way. As you can increase the height, the flow gets faster. And that's typical, because you have a no-slip boundary condition at the wall, right? Is that OK, or did I misunderstand your question? Is that good? OK?

AUDIENCE: So y is a constant.

WILLIAM GREEN, Yeah, in this particular problem this is just a number that we tell you in the problem.

JR:

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, Right. If you had a flow in a cylindrical pipe, it wouldn't be-- wouldn't be a number. Be more complicated, yeah, all right? So this is-- I mean, again, this is like the simplest you got. It's only 2D. It's really simple. But you're going to see here that even here you could have a lot of trouble.

And if you just don't think about it-- the computer will give you an answer. Whatever-- you put in some finite difference equations or finite element equations, finite volume equations, fsolve, whatever, it's going to solve and give you some numbers. It doesn't mean it has any relation to the physical reality. So this is a problem to really pay attention to whether you're really converged. All right.

And so what I suggest you do in this problem is go through a sequence where you vary-- the real problem wants you to do this all the way out to-- this is 1 centimeter. But I suggest you instead solve a simpler problem, where you put the-- here's a simpler problem. Put your boundary right here, really close, and then solve that problem. You won't need as many mesh

points across to get from here to here.

And this should be-- this is-- suppose you choose this is $c = 0$ boundary condition for this wall, then this should be an upper bound. Because if you put an absorbing layer here, it should drive the diffusion faster, right? So as you increase this distance-- let's call it little h -- as you increase little h , you should converge to the true solution that you want sort of from above. Does that makes sense? Yeah? Is this OK? All right.

And similarly, you can vary the mesh, how big your boxes are, say, in Δx and Δy . And again, for each of those, you could think about how should things converge. And so look and see if it's actually converging the way you think it should be converging, right? Any more questions about this?

OK, so this problem might-- this homework problem might look like a MATLAB coding problem. It has MATLAB coding, but it's not really-- that's not really what it is. It's really like a-- it's a conceptual problem about what you're doing. All right.

All right, in this problem, I want you to use finite volumes. So let's talk about finite volumes for a minute. So finite volumes is to imagine that we have little control volumes, and each one of them has a little propeller in it, stirs it up, like little CSTRs.

And so we do it just like you did back in your intro ChemE class and mass/energy balances a million years ago. You know that you have some flow coming in here, and maybe some flow going out there. And maybe you have some flux come in here and maybe something out there.

And you can add up all the flows in and all the flows out, and then that-- the net of all these fluxes has got to be equal to the accumulation. And if we're doing a steady-state problem, then there's no time derivative, so the accumulation term should be 0. And you did a lot of these problems a long time ago, right?

OK, so you just do it like that. The only problem is now we have a million of these little boxes all coupled to each other. So you get one equation for each box. And in this problem, we're going to-- when you do this method, what people assume is that you have a uniform concentration sort of across here or the number you use as the average of the concentration in the cell as [INAUDIBLE].

And so it's not really exactly realistic. So that's where the approximation is. What you're doing

at the boundaries is very realistic, because you're actually computing the fluxes across control volumes is exactly the way you should do it.

So there's different methods, right, finite element, finite difference, finite volume. The nice thing about finite volume is you're really treating exactly what's happening across the boundaries of the mesh area, the mesh volume. And we'll find that a lot of ChemE problems, this is how people prefer to do it.

And because you're treating the fluxes exactly, if you have your mesh volume sitting on the wall or on top of the drug patch, suppose if we're sitting right on top of the drug patch here. Then if you're sitting-- first, let's do a wall. Suppose you're sitting on an impermeable wall, then we just know that the flux here will be 0 if it's a wall. So that's easy. No flux here. So we only have to worry about this one, this one, and this one. That one's not doing anything.

This case, if you say in the drug patch there is a flux, stuff's coming in and you'll need to figure out how to write that boundary condition. So the boundary condition as written is that C , the drug, is equal to some number, C drug, whatever is in the patch. But that's not so easy to impose here.

And so there's two ways to look at it. One way is people compute this flux by considering-- suppose I know this is C drug here. I'm trying to really figure out what's the average concentration here.

And so one way to compute this is to say it's the diffusivity times C drug minus C the middle over Δy over 2. That would be the flux. So that's one way to look at it.

Another way people look at it is they draw what they call a ghost volume. And so here, here's my C that I care about, C middle. Here's C ghost. And here is the line between them.

Now, I can use the same equation here that they would have used before. But I don't know what C ghost is, because there's no real cell here. This is below the patch. Here's the patch.

But imagine the patch is not there for a second, and I write down the same equation I would have here. And the flux diffusively would have been $D C$ ghost minus C middle over Δy . That's what you would have written as the diffusive flux. And I don't know what C ghost is.

So now, I have to think about how do I estimate what C ghost is. I can say, well, let's do a linear interpolation from this concentration to that concentration. So let's say that the C

boundary is equal to C_{ghost} , the average of these two, plus C_{mid} over 2. That's what we got if we did a linear interpolation between these two guys to figure out what it is here.

But here we know what it is. We know what a C boundary is. That's the concentration of the drug patch so that we can solve for what c_{ghost} is.

AUDIENCE: How do you know what C middle is?

WILLIAM GREEN, C middle is the unknown. That's what we're going to compute. We just want an equation that
JR: involves C middle, because we have c middle as an unknown.

We need an equation that involves C middle, just like we would for-- if we have a-- If we have finite volume that is in the interior somewhere, we have c of this grid point, C_{ij} . We just want equations that involve C_{ij} . But we need an equation that somehow connects it to the boundary conditions. Yeah.

AUDIENCE: When would you use that second [INAUDIBLE] if you know C boundary [INAUDIBLE]?

WILLIAM GREEN, Right. And actually, I think if you do it this way, in this case, you get the same formula. You're
JR: running out of C .

AUDIENCE: So when would you use the C ghost?

WILLIAM GREEN, The C ghost thing is handy when these conditions-- actually, in the finite volumes, this is it. I
JR: don't think you need to do it. Maybe if you had a flow, velocity flow here. I don't know. But you won't in a wall.

So I don't know if, in the finite difference method, you can use those ghost points as well to do the flux gutter conditions. And that's often useful to do the ghost thing for if you have a symmetry-imposed flux boundary condition. Yeah.

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, That's the flux. The flux is coming from across this wall. And we'll have to add that with the flux
JR: coming this way and the flux coming this way and the flux coming this way to get our total flux, which is going to net out to 0 instead of C . Is that all right?

But the other ones, you all know how to write them? They're just velocity flows. The nice thing

about the finite volume method is that when you're trying to consider the velocity, the velocity that matters is actually the velocity here. It's the flux that couples the two finite values together. So your velocity is being evaluated halfway between the two mesh point centers.

This turns out to make the whole procedure more numerically stable. You compute the flux by, say, the difference between these two. And you're implicitly evaluating right at halfway between.

Other methods, like finite difference, you're trying to get the velocity or the flux at the same point here. And it doesn't really make much sense, from the point of view of trying to compute how much material's close to here, to use the velocity right there. And so when you work out equations, and particularly the equations that involve pressure, so suppose you try to discretize the Navier-Stokes equations and you have a pressure gradient that's driving the flow, then you have to figure out where are you going to evaluate the pressures and where are you going to evaluate the velocities. And when you try to evaluate them at the same point, it turns out that you get numerically unstable problems.

But if you do it by this funny volume method, then it just works out naturally. It's fine. And then you need special methods, if you're going to try to do it where you have the velocity of the mesh at the same point as the pressure mesh. You'll notice in most of the problems we give you, we just leave pressure out.

We try to rewrite the equations so no pressure ever appears. And that's because pressure in general is a problem, because you can have acoustic waves physically. And if you do the equations, which allow for acoustic waves, you'll get them in the numerical solution as well.

But usually, we don't care about that. So you have sound waves racing around your solution, and that causes a lot of trouble numerically. So a lot of times, people rewrite the equations to try to remove the pressure from the equation. So you write down the Navier-Stokes equations. You guys did this transport. Maybe you'll study that today in the class, in the test. I don't know. Anyway, there's a pressure term if you just write naturally.

But oftentimes, you can remove that by an equation of state, for example. And then you can get rid of it. And that turns out to be better from the numerical solution. All right. Yes.

AUDIENCE:

I have a question. So you said that for this problem, [INAUDIBLE] was very small. So that means that these finite volume cells [INAUDIBLE] you have to do a lot of them. [INAUDIBLE]

gets around this by having a variable volume. Do we have to address this on that lab that have some variable volume?

WILLIAM GREEN, OK, so from the point of view of writing a code, it's a lot easier to write it with fixed mesh
JR: because all equations look exactly the same. So I suggest you start that way. And what you do, if you use a very small value of each, then I think you'll be able solve it, no problem. You'll have enough mesh.

Then, once you figure out how to do that, now you might be able to see, OK, can my solver solve this? And then that gets into another question, actually. What solver are you going to use? So any ideas about this problem?

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, So you could use fsolve But I'm telling you, you're going to have to use a lot of mesh points.
JR: That means that fsolve's going to have to solve for a lot of variables.

AUDIENCE:

WILLIAM GREEN, Backslash it might be. So the key thing, the nice thing about this problem is it's a linear
JR: differential equation. There's no nonlinear term today. So when you rewrite the equations of the finite volumes, it's all going to be linear in the unknowns. And what else is nice about this problem, when you use local finite volumes as your--

It's going to be super-sparse. So the matrix that comes in is going to be really sparse. And so you'll want to use some method that can handle gigantic sparse matrices.

So you wrote a code like that earlier, so that would be one possibility. If you know how to use the right flags in MATLAB to help their built-in solvers handle sparsity, then that should be good. If you just ask it to solve it by dense method, by, like, LEU or something, you're going to have lot trouble, once you've put your mesh points in there.

But you can just experiment. Try bigger and bigger matrices. And then at some point, if you have backslash, it will give you a warning or something, unless you tell it that you're sparse. All right? Any more questions that about this?

How would people solve it do you think professionally? Suppose I was doing a problem like this in 3D instead of 2D. How would people solve it? What solver would they use?

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, Not fsolve. No.

JR:

AUDIENCE: [INAUDIBLE] gradient.

WILLIAM GREEN, Yeah, so they would use conjugate gradient. So probably-- I think it's called this, BiCGSTAB.

JR: That's the program that we would use if you have a really, really gigantic sparse matrix. So that's the conjugate gradient.

And Professor Swan talked about, the advantage of that is you never have to actually store the whole matrix. You only need to evaluate the matrix elements, and then you can throw them away. And so if you have a very sparse matrix, that's pretty cheap to do.

So it's a really good code. I think in this 2D problem, you can probably get away with other solvers. You don't have to use this. But this is a definite possibility. This is a built-in MATLAB program as well.

Be warned, though, this is an iterative solver. It's not just going to be one solve, boom. And it might have troubles. So you might want to go with the other ones, but anyway.

May I ask you a question? How about, do you need initial guess? What do you think?

AUDIENCE: Depends.

WILLIAM GREEN, Depends on the solver. So if you solve it with backslash, do you need an initial guess? If you can solve it with fsolve, do you have to initial guess? If you can solve it via BiCGSTAB, do you have to use initial guess?

AUDIENCE: Yes.

WILLIAM GREEN, Yes. OK. So then you have to think about how you really get your initial guess, too. So this is things to think about. All right. What else to tell you about?

One last thing about PDEs, and we'll come back to this later, so far we haven't done really anything that's very time-dependent. But a lot of real world PDEs have a time dependent in them. And there's is a very important concept, a thing called the CFL number. And this is named after a 1928 paper, and I'll write the guys' names down.

And what they showed was that, if you're trying to solve the PDE system, where you're discretizing in both x and time, that you have a number that they defined as Δt times the velocity and the extraction divided by Δx . So that's a dimensionless number. So that's a CFL number. You see a lot of papers that will say what CFL number they used.

What that means is the ratio of their time mesh compared to their space mesh. And conceptually, let's think about what's happening here. So suppose we have a flow flowing in an upwards direction, and we have a bunch of little finite volumes.

So we've discretized the Δx already. And this is x . And there's a flow here.

And I've already decided, somehow, what length scale I want to use. So I've decided my Δx . And now, I'm trying to figure out what Δt I should use.

Now, from the point of view of saving CPU time I want the Δt to be as giant as possible, because I want to be able to zoom along or predict for long periods of time what's going to happen in my system. But if I make Δt really large, let's think about what happens.

Suppose I choose Δt to be 10 times Δx divided by u .

So it means that in my one time step, I have some guess or some current value of the concentrations in all these finite values. And then I wait through a time step that's 10 times Δx over u . So I had some stuff that was here. Where is that going to be 10 times later? 10 time steps later. 10 blocks up, right? So it's going to be, like, way up here somewhere. And so what's going to happen there is that my numerical methods are all computing stuff locally from the spatial derivatives.

But it's crazy if, between my time steps, this stuff completely left the picture. It's already convected all the way off the screen. And some new stuff, which was way down here before, is now the stuff that's here. Should be there if I was physical.

Numerically, who knows what will happen if you try to do this. But it won't be good. So the condition is that you need this number to be less than 1, or same order of magnitude as 1, and you try to make this much bigger than 1, then you're doing something crazy because you're convecting stuff over multiple mesh points. And so that turns out to be a very serious limitation if you try to do simulations for, let's say, a reacting flow for a long period of time, because you might have to use a really tiny Δt . And then people have developed all different fancy

methods to try to get around that.

But if you just do the obvious things to do, you'll always run into this limitation. Then you need to choose the time steps small. And also, it's bad, because as you make Δx smaller, which improves your accuracy, you'll have to make your Δt 's smaller, too.

But of course, making Δx smaller increases your CPU time because you have more finite volumes to compute. And then you'll also have to make Δt smaller, which means you'll have to do more time steps, too. So it's even like a double whammy. So getting more accuracy is going to really cost you badly.

And so this another reason why people used to always refer to color fluid dynamics. You can make a pretty picture, but it might not be physical at all. Because you can make it solve equations that maybe, for example, didn't impose this, then who knows what kind of crazy stuff you'll get. You'll get something. It'll compute something, but it may have no relation to the real problem.

I think that's all I was say about PDEs. Are there any questions about PDEs before I start talking about probability? You got it totally down. I'm looking forward to some really awesome solutions. How about that?

Just one last comment. If you decide you wanted to do adaptive meshing and you want to change your mesh size, you can choose measures like this if you want. And you can even do things like this, where you have a bigger mesh, and then maybe have two smaller meshes underneath it in the next trial.

So you can just have stuff flowing here, stuff flowing here stuff flowing there. So you can do all kinds of crazy stuff like. This can really help improve the accuracy of the solution a lot.

But it's, I would say, very prone to bugs. So if you do this, be really careful and don't do it too often, I would say. You might have a few boundaries where you do something funky like that to change that mesh size. But don't go crazy with it.

[INAUDIBLE] is smart. It has a really nice way of doing the meshing for you. So that's its advantage, that somebody very carefully coded how to handle this kind of stuff in a general case. But if you guys are doing it for the first time, it might not be so good.

So that's enough of PDEs. Let's talk about probability.

So probability is everywhere, except in undergraduate ChemE homework problems. So when you do problems as an undergraduate, they always had some nice solution. It was 2π , it was 3.0. Everything was, like, deterministic, is definite.

The grader could go through. Oh, no, you're off by 0.1. It couldn't possibly be right.

That's, like, not reality. So any time you actually make a measurement, you always had measurement noise. And if you try to repeat a measurement, you don't get the same result. So that's, like, completely different than an undergraduate problem.

But this is the reality. So the reality is the world is, like, not so nice as you think. But actually, it's even more fundamental. It's nothing about-- I mean, there's one problem about how good an experimental those people are, and how fancy an apparatus you bought that can make things exactly reproducible.

But even if you do that perfectly, the equations we use really don't correspond to the real physical reality. So we always use the continuum equations. You guys probably are studying them a lot in 1040 and 1050, especially 1050, I guess. But those equations are really all derived from averages over ensembles or little finite volumes or something if you look at the derivation of the equations.

And reality is that the world is full of molecules. And so they're all wiggling around. And if you look in a little finite volume, you look right now, you'll see that there's 27 molecules in there. If you look a second later, there might be 28. Then a little later, maybe 26.

It's always fluctuating around. But according to our average equations that we use, like in the Navier-Stokes equations, it always 27.3. But of course, there's not 0.3 of a molecule. So, I mean, explicitly, it's the average is what we're computing, and the reality is fluctuating around the edges.

Same thing in thermo. We say that such-and-such has a certain amount of energy. But you guys have some [INAUDIBLE] already, yes? So you saw that's not true? So really, all that's saying is that's, like, the probability, the average. If you've had many, many ensembles that were exactly the same and you average them all, you get some number and that's your average energy. But for any actual realization, it has some different value of the energy.

And it's even worse than that. That's because we have a lot of particles. You can even go

down to, like, the microscopic level, where you have one molecule.

And you calculate things about that, it turns out you have to use the Schrodinger equation for that. And the Schrodinger equation explicitly only gives you probability densities. So it just tells you the probability the molecule might be somewhere, the electron might be somewhere, the energy might be something.

But it's not actually whether it really is. It's just saying a probability distribution. And every time you do the experiment on the molecule, you get a different result.

Now, this is super-annoying, but it's the way life is. Einstein got so annoyed, he has a famous saying, and it's "God does not play dice." He was, like, just completely annoyed at these equations. But it's the way it is. So the reality is that things, all we know about, really, are probability distributions.

In most of our work, in our lives, we always talk about, like, the mean or the median. And we're talking as if it's a real number. But really, it's always some distribution. So it's time, I guess, you're in graduate school, it's time to, like, face up to this. And that's what we're going to talk about for the next week or two.

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, This also makes it makes you wonder what you're doing when you make a measurement. So if
JR: you make a measurement, first of all, the fact that when you measure something repeatedly, you're not going to get the same number, that's alarming. Because I want to say I'm 5 foot 9", this should be pretty common, I'm always 5 foot 9".

But actually, if you measure me multiple times, sometimes you'll get 5' 9 and 1/4" Some people will get 5' 8 and 3/4" So it might make you worry, did I change? Did I grow between the measurements? So that's one issue.

Because our experiments are not repeatable and we get different numbers every time we make a measurement, then we have a big problem. Somebody says, well, I really want to know how tall Professor Green is. I've always wondered, how tall is he?

And everybody's told me a different number. And when you go measure it again, you get a different number again. What's going on?

And so you'll have to then-- then we have, like, a concept that there is a true height of Professor Green and we just don't know what it is. And then we'll try maybe to make repeated measurements of my height, and then maybe take the average. That would be the obvious thing to do. And we take the average and report that. We'll tell the boss, we'll lie to him, say, oh, Professor Green's 5 foot 9".

When really, we never actually measured 5' foot 9". Every time we measured, it was 5' 9 and 1/8", 5' 8 and 3/4". Every time, it was something different.

But we just say, OK, it's 5' 9". And the boss, he doesn't want to know about all this complexity, anyway, so he just believes you. So it takes your average number. But you know that you're not really sure I'm exactly 5' 9" And so you have a probability distribution, and you're doing your best guess.

And in fact, if you're honest to your boss, you'll give him an error bar. You'll say he's 5' 9", plus or minus 1/2 an inch. And that way, what you're saying is you're pretty sure the true height of Professor Green is somewhere in that range. Is this OK? Yeah.

Now it might be that I'm actually changing height. I get a good night's rest, I lie down for a long time. Maybe when I stretch out a little bit.

When I stand up in front of lecture here for a long time, I'm shrinking. My vertebrae are being compressed by standing here so long.

So it's, like, a combination of things. One is your measurement system is not perfect, and one is that I actually might be fluctuating. I had a big breakfast, I'm growing.

So is it true with everything? Every experiment you do is like this, that there is a real fluctuation of the real, physical apparatus of the thing that you're trying to measure because mostly, things we're trying to measure have fluctuations intrinsically. And then on top of that, your measurement device is fluctuating, which means-- and then the combination is what you're measuring. It gives you fluctuation.

If you have a very nice instrument, the fluctuation of your instrument is smaller than the fluctuation in the real system and you'll go down to the limit. Anybody with a really good instrument should measure approximately the same probability distribution, which is actually the real fluctuation. My height, for example.

So if you bought a laser interferometer and mounted a mirror on my head, and measured my height to the wavelength of light, you're pretty sure that it's pretty good. It's within a wavelength of light. So the error bar there is just due to the fact that I slouch sometimes. Is that OK?

So let's talk about some basic things about probability. So we're always saying there's a probability of an event. And so we want to give a number.

So for example, I'm flipping coins. I flip a penny, it could be heads or tails. I'll say the probability of heads is approximately equal to $1/2$. So I flipped the coin, and if you flip the coin 100 times, you might expect to see 50 heads.

Now, it could be 49 heads, it could be 51 heads. So you have to worry about, like, exactly how precisely you know it. But you think it's something like a half.

Now, just to warn you, I actually didn't specify any more significant figures here, and it might be really hard to figure out what those significant figures are. And this is related to the measurement problem. So we think that a coin has about a 50/50 chance of being heads or tails. But if you really wanted to prove it, that might be really hard to do. You can have joint probabilities.

So suppose I have a penny and I have a dime. And I try to think, if I flipped them both, what could happen? I could have that they both come up heads. I could have this guy come up heads, this guy tails. This guy tails, this heads, tails.

So there's, like, four possible outcomes, and we think they all have about approximately equal probabilities. So the probability of any of these things happening should be about $1/4$. So you can write the probability of event 1 and event 2. So this would be, for example, the probability that I got heads for penny and heads for dime, and I think that this is about $1/4$.

Now, I could also say the probability of heads, and heads is equal to the probability of heads for the penny times the probability of heads for the dime if I got heads for the penny. Now, if these two coin flips are completely uncorrelated, then the probability of heads on the dime and heads on the penny is the same. It's just the probability of the heads on the dime. So they don't matter.

But many things that we'll study are correlated. That the probability of something happens depends on whether something else happened. So this kind of expression is very important.

Now, this is just an equality. It's like a definition of what these are, right? And I'll just notice you can write that the other way around. So it's the probability of heads on the dime times the probability of heads on the penny. This is OK?

So these two guys are equal to each other, and you can rearrange that equation any way you want. And we'll come back to a very famous way to rewrite that equation. It's called Bayes' theorem, and that turns out to be really important in model versus data comparisons.

Instead of doing AND, do OR. So maybe you can say, then, what's the probability that I see at least one head? So I flip my two coins, and I have the probability of, I see at least one head. So we know we intuitively the answer is $3/4$. But let's try to think of where does that really come from. So--

[HIGH-PITCHED SOUND]

What is that?

Sorry. Really threw me there.

So probability of at least one head, it's not equal to the probability of head for the penny plus the probability of head for the dime, because we know this is really $3/4$, and this is $1/2$, this is $1/2$. You add them up, $1/2$ plus $1/2$ does not equal to $3/4$.

So be careful. There's a lot of things you can say quickly that are not true. So anyway, you really have to consider the whole thing. And in the best case, if you can enumerate what's going to happen, it's very simple to add up things. Otherwise, you have to be very careful with the algebra to make sure you add all things correctly.

Let's see at least what this should be equal to. So this should be equal to the probability that a head on a penny times the probability that I have head for the dime if I had a head for the penny plus probability of the tail for the dime or for the penny plus the probability that I have for the dime and then times something like this, too. And in a case like this, where I'm summing over all of the possibilities.

So in this case, I'm either going to get a head for the dime or a tail for the dime. It's not going

to balance on its end. I'm assuming that that chance is 0. Then these two things just add up to 1. Is that all right?

It's really saying the probability of the dime will do something if I have a head for the penny. Is this all right? So you want to practice doing little algebra things like this to make sure you know how to do it.

Yes? No? Maybe?

Let's see what else I've got here.

AUDIENCE: Professor.

WILLIAM GREEN, Yes

JR:

AUDIENCE: What you just wrote, [INAUDIBLE]

WILLIAM GREEN, Yeah, what's the correct thing to write in here? What is the right thing?

JR:

AUDIENCE: Tails.

WILLIAM GREEN, Yeah. So it's tricky. Yeah.

JR:

AUDIENCE: The probability [INAUDIBLE]

WILLIAM GREEN, Yeah, it's the easy way to do it, for this case. And so there's, like, a lot of-- for some cases, it might be easier to write it this way. In some cases, write it that way. It depends on how many different options there are.

So anyway, I'm just trying to warn you by writing this out. It's like, you might write down stuff quickly without thinking about it. And it's easy to double count. Like for example, if you put the other term in here, you double count. Because you already have, they had head, you already had the head case here.

Yeah, maybe-- Did I write the-- I'll just write the base theorem down the way you usually see it. So this is the general expression, which is always true. The base theorem way to write it is the probability of A given B is equal to the probability of B given A times the probability of A over the probability of B.

And that's just rearranging this equality. Rearrange again. And we'll come back to this with the situation that's like, what's the probability that Professor Green is 5' 9", given our measurements, is related to the probability that if Professor Green was 5' 9", we would have made the measurements we got. So this is like that way to invert that statement.

And if the thing that's here is exclusive, it means there's, like, many possible things that could happen. This is the probability that one of them happened. There's a lot of other things that could have happened.

For instance, like heads and tails, it's either heads or it's tails. It's exclusive. If it's like that, you could rewrite that probability of B is equal to probability of A_j , probability of B given A_j summed over j . This is something, whatever A measurement it is, if it ends up with B, that's probability B. Is that all right? So you can put that into the denominator here. You can substitute that in so you can rewrite this as probability of A_i given B is equal to the probability of B given A_i times the probability of A_i divided by the sum of the probability of A_j [INAUDIBLE] A_j . And this is the form that you'll normally see based here.

A lot of times, we have a continuous variable instead of discrete events like this. And so then we talk about probability distributions. And so instead of having a sum there, we might have an integral. And this is a topic that also is quite confusing to many people.

So suppose I had a Maxwell-Boltzmann distribution. And I have, like, the probability density of having a certain velocity in the extraction instead of the particle. And so that's going to be something like $e^{-1/2 m v_x^2 / kT}$ divided by something. And maybe it actually is [INAUDIBLE]. I'm not sure. Maybe there's a v_x here. Something like that.

So you have an expression that you get for [INAUDIBLE] for probability density. Now, what does this mean? We want this thing to be the integral of P of v_x dv_x over from negative infinity to infinity.

We want this to equal something. What do we want this to equal? 1. So that means that the

units of this, this is units of centimeters per second, what's the units of this?

AUDIENCE: [INAUDIBLE]

WILLIAM GREEN, Centimeters per second now minus 1. It's, like, it's per the unit to this to get a dimensionless number there. So this has units of seconds per centimeter, which probably none of you thought until I just said that to you. So probability densities are tricky, and they always have to be multiplied by a delta.

When I talk about this, I really need to talk about P of v_x , Δv_x . I need to have something here to make this look like a probability again. And so the issue is that the probability that the velocity is exactly something is, like, 0. It's really the probability that's in a certain range, plus or minus something. Then you get a nonzero probability.

There's another quantity you'll see a lot called the cumulative probability distribution. Let's see, what letter do I use? Call it F . And this would be like the integral from negative infinity to v_x prime, or v_x of P over v_x prime, dv_x prime.

And this is the probability that the particle has v_x less than something. So this is F is equal to the probability that v_x -- let's call this v_x star-- v_x is less than or equal to v_x star. And that can be quite an important property.

So for example, you're designing a supersonic nozzle. You want to know what gas molecules are going to come out at a certain speed. You really need to know that probability. How many of them are going to be bigger than that speed, how many less than that speed? So these are two different ways to express a similar thing.

This is, like, the probably that it does have that speed within a certain range. This is the probability that it has anything less than or equal to that speed. And in a completely different way, this is an integral. This has units of dimensionless. This has units of per velocity.

All right? All right, we'll pick up more of this on Monday.