

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](http://ocw.mit.edu).

**WILLIAM GREEN JR:** All right today, I would like to do a Kinetic Monte Carlo example with you in class and, also, talk a little bit about the connection between the master equations, master equation solution, the normal kinetic equations we use, and the-- how you handle these trajectories, you'll get how to Kinetic Monte Carlo.

All right, so let's just pick a simple example. As we mentioned, in the real world, it's pretty easy for these examples to get completely out of hand, because the number of possible states can be so gigantic. Let's all do a really super, super simple one here. To try to keep that underhand, not worry about that. When you do the real problems, that will be a major issue-- is if the number of states get out of hand.

So the simple example I'm gonna do is a real one from inside your body. So inside your body, you have a thing called low density lipoprotein. And there's little vesicles of fat, lipids, inside your body, actually, in your blood vessels. And these are bad. If you go to the doctor, they measure your LDL and your HDL. And if your LDL number is too high then they yell at you for eating too much fat food. And they make you exercise and stuff like that. This has happened to me, so I know.

And there's a reason for this, is because the LDL is susceptible to oxidation. And if too many of those lipid molecules peroxidize, it causes inflammation in your blood vessels. And that can cause your blood vessels to constrict and stuff that eventually can lead to heart attacks. And so that's why your doctor is alarmed if you have too much LDL, because you have a lot more material available that could be oxidized, that could kill you.

There's actually a really complicated, interesting story about what vitamin E does in the LDL, and how it interacts with vitamin C. And if you take my kinetics class, I'll tell you about that. But we'll do the simplest case now. There's no vitamin E, there's nobody Vitamin C. You just have some lipid, and it's oxidizing. And let's model what's happening.

So what you have is a little bubble of fat. So out here is water. And here's my fat which I'll

politely call lipid. That sounds like not quite as fat, if you call it lipid, slimy. And the lipid is susceptible to attack by this reaction. So if I have some peroxy radical, and I have the lipid molecules, which are some hydrocarbon, they can react and make peroxide plus a radical. And in your blood vessels, if you're not under the water or being killed, you have a lot of oxygen. And so the carbon set of radicals you form immediately react with oxygen to make the peroxy radical back.

And so the net process-- we can kind of combine these steps, because this is super fast. So it's like a microsecond times scale. For that to happen in your body, because the oxygen concentration is so high in the blood. And so this reaction is really  $\text{ROO}^\bullet + \text{RH} \rightarrow \text{ROOH} + \text{R}^\bullet$ . So the radical is just catalyzing the oxidation of your lipid into the peroxide. When this level gets too high, then your body will detect it. There's an inflation, and you'll have all kinds of trouble. OK, so we like to understand this.

Now, these vesicles are very small, so there's a probability you might have only a very, very, small number of radicals in there. Also, your body has a lot of antioxidants in it, like vitamin C and vitamin E and some other ones too. And the antioxidants are trying to get rid of radicals. So the number of radicals you'd expect the concentration-- background concentration is very tiny. And because of that, this is tiny and the concentration is tiny. So the concentration times the volume might also be tiny. So the total number of radicals might be very, very small. And so therefore, you might need to worry about the stochastic kinetics being different than the continuum kinetics. OK, so that's the idea.

Now, on top of that, you also have a lot of these vesicles in your body. So you might be able to somehow work up some continuum model, treat them all. And in fact, I think that's what people did originally, is they just said well, let's take the total amount of LDL you got in your body and do a continuum model. But as you will see, it's different, because the fact that the vesicles are small, each one might only have, say, one radical in it. And that will actually behave quite differently than if you had a lot of radicals.

All right so, this is the kinetics, and the oxygen is almost always high concentration, so I'm not going to worry about that. And I'm going to assume that we're not going let your lipid oxidize so much that the amount of lipid changes significantly. So let's just, for now, we can just ignore this. OK, maybe later we might go back and figure out how you'd account for the fact that these cells are being consumed, which would change the rates a little bit. So we'll just assume that concentration is constant.

So you have a tiny amount of radicals. And if you had the radicals in your vesicle, they're creating  $\text{ROH}$ . And you see, the  $\text{RO}$  is catalytic. And so what we're gonna see is basically  $D$ . If you did it in classical kinetics, you'd write  $\frac{dR}{dt}$  this is basically equals some constant times  $\text{RO}$ . And then, I lumped these concentrations into this constant. [INAUDIBLE] That all right?

All right, so that's the import process. We also have another important process, is that if you get two radicals in there, and they bump into each other, they can kill each other off. So two peroxy radicals can combine and make stable molecules. And we don't care what stable molecules they really are. What's important is that they get rid of the radicals, because the radicals are catalyzing the unfavorable oxidation. In this program I started writing, I called this  $k_3$ . This one is  $k_4$ , reaction four.

There's two other things happening. Out here, I had some peroxy radicals. They're floating around the water. Some of them might go into my lipid. I'll call that process  $K_1$ . There's some time constant, which new radicals arrive. And then, I also have the possibility these guys could diffuse out. So I call that  $k_2$ . And  $k_1$ , I'll just treat as a constant.

So as soon as a background concentration radicals that's constant, and so this is a sum-- every 10 seconds or something, a radical arrives in a vesicle. And this will depend on the concentration, the number of  $\text{RO}$ 's. So the more you have inside here, the faster the rate which they go out. That all right? And this just mass transfer. But from the point of view of the equations, it doesn't care if it's mass transfer or chemical reaction. They arrive in the same way.

All right, so that's the system. And so let's look at the-- I should try to write the MATLAB code for this. And maybe, we can try to finish this together in class. So I have four processes-- arrival of new radical, diffusion of the radical and the vesicle, reaction to make the peroxide, and self-destruction of the radicals. And I wrote down what each of these does. So the first one increases the number of radicals by one. The second one gets rid of a radical because it's diffused away. The third one increases the amount of peroxide. The last one gets rid of two radicals.

So we're going to try to compute a trajectory by the Kinetic Monte Carlo method. And trajectory is going to be some time points, the number of radicals at that time point, the number of peroxides I have at that time point. And as I run, I'm going to get many, many,

different time points. And each one will have a different number of radicals or a different number of peroxides. And that's what I want to compute. And then, at the end, after I have trajectories like that, big tables of these things, I want to somehow put it together to figure out on average what happened or something, try to understand it.

And so let's see, so my inputs are the initial numbers of radicals and peroxides in the vesicle, the vector of rate coefficients, the max amount of time I want to run the trajectory for-- not clock time, but time in the simulation. So it's like how many microseconds or milliseconds or seconds I look at my lipid protein and see what it's doing. And the maximum number of steps, because I don't want to get a trajectory list that's 10 to the ninth long, because over a while, my memory, my computer would cause me trouble.

So here they are, all the setups. Everything is fine so far. So here's the loop. So the outermost loop is just to keep track of how many steps we have, which is going to be the number of entries we have in the trajectory table. And the second one is really-- what we care about-- is well, the time is less than the maximum time of the simulation. First, we want to figure out the time [INAUDIBLE] until something happens.

So following Joe Scott's notes, we compute this quantity called  $A$ , which is the sum of all the rates. And that's equal to the rate of the first process. And then, they're in the second process, which depends on the number of radicals. And the rate of the third process all depends on the radicals. And the rate of the fourth process, which is a number radicals times the number of radicals minus one.

Now  $k_4$  has to have units of per second. But bi-molecular rates would normally have units of volume per second, per mole even-- volume is per molecule per second. And so that  $k_4$  has to really be a normal kind of  $k$  divided by the volume. And we're going to have to figure out what to do about that in a minute. And then, we do the formula to get the-- Gillespie's formula for how to sample from an exponential decay, as good as that is.

So Gillespie thinks that the probability that something is going to happen, or the probability, the time, until the next thing happens is going as  $e$  to the negative  $a$  times  $t$ . So we're trying to sample from  $e$  to the negative,  $a$ ,  $t$ . And that's what that crazy log of the random number is doing-- it's sampling from that distribution. And so we get a [INAUDIBLE] that how long we've waited from the time the last thing happened until the next thing happened.

And now, we have to figure what happened. So there's four different processes that could

have happened. A radical could have arrived. A radical could have left. A radical could have reacted with one of my hydrocarbon molecules, with one of my lipids, or the two radicals could have killed each other.

So we list these guys out.  $A$  is the sum of all those processes.  $P_1$  is the probability that the first process happened, that a new radical arrives. So it's going to be the  $k$  for the first step, divided by  $a$ . That one was zero water, because it's subterfuging it from outside. So it's just a  $k$  of 1.

Probably the second step is  $k_2$  times  $n$  rad. That's the rate of the second step divided by the total rate of all the processes. And I'm going to add that onto  $p_1$  to make a vector of possible things that could happen.

So I'm picking a random number from zero to one. And I want to make a little bar. Here's the probability the first process happened. And here's the probability that the second process happened. And here's the probability the third process happened. And here's the probability the fourth process happened. I know that these guys have to add up to one, because I computed that something happened.

And so I'm going to try to compare my random number from zero to one to these breaks and sort it out into these four possibilities, and then figure out which thing happened. OK, so this is the second step of Gillespie's algorithm. All right, so maybe you guys can help me finish the code here. If you want to go to Broadway in Chicago, you can go there.

So I need to write the next one. So  $p_3$  is equal  $p_2$  plus what? So step three is-- sorry? Say again.

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** All right, that's enough, right? So now, I get to choose a random number, so, say,  $x$  is equal to **JR:** rand. And then, I have to say if  $x$  is less than  $p_1$ , than something. What's it gonna be? Then step number one happened. So that means that  $n$  rad is equal to  $n$  rad plus one, so one radical transported in. If  $x$  is less than  $p_2$ -- [INAUDIBLE] Is that good?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** [INAUDIBLE] space. Thank you. What happens here? So this is  $b$ ,  $n$  rad is equal to  $n$  rad **JR:** minus one?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** I think they also take care of that, right? Is this right? X is less than p3, than nroh. Else and  
**JR:** rand. Does that look right? Do I need an end?

**AUDIENCE:** Semi-colon.

**WILLIAM GREEN** Semi-colon. Semi-colon, thank you. That would be really bad. All right, yes, Ziggy?  
**JR:**

**AUDIENCE:** So in the first problem, you have step size longer than that [INAUDIBLE] size.

**WILLIAM GREEN** So that's not good. It should be less than. Thank you. Any more bugs?  
**JR:**

**AUDIENCE:** I think you need spaces after the brackets.

**WILLIAM GREEN** Spaces after the brackets-- where? Is that here?  
**JR:**

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** [INAUDIBLE] separate line.  
**JR:**

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** I can do that. It's easier to read anyway. So that's good. I definitely agree with that.  
**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** Is that right?

**JR:**

**AUDIENCE:** Yeah.

**WILLIAM GREEN** You like this? With all these [INAUDIBLE] sets, I don't need any n's, is that right? [INAUDIBLE]

**JR:** after the other like that, no problem. Just one end at the end.

[INTERPOSING VOICES]

**WILLIAM GREEN** That's good. All right, so we're OK with this? We think this is going to work?

**JR:**

**AUDIENCE:** [INAUDIBLE]

[INTERPOSING VOICES]

**WILLIAM GREEN** All right, so now, we've computed what's going to happen. Now, we need to store results

**JR:** somehow. So I'm gonna say trajectory is equal to-- or trajectory of steps-- so step was zeros.

Now, step is one. [INAUDIBLE] step is one, to start with.

[INTERPOSING VOICES]

**WILLIAM GREEN** Does this look good? They had only commas, is that right? No commas?

**JR:**

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** All right, one step is zero. I already have the first line filled in here.

**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** Is that good? We'll see. I don't know if it's OK or not, but I'll believe you. All right, so right now,

**JR:** we have a program that maybe or maybe not might actually compute something. Is that right?

OK, now, we to figure out-- are we going to run this program?

So we need an initial. So any suggestions?

[INTERPOSING VOICES]

**WILLIAM GREEN** All right, so let's do an initial-- how about  $n_{rad}$  is equal to 1 and  $n_{roh}$  is equal to 0. Is that OK?

**JR:** We're cool with that? And how about  $k$ ? So we need a vector of  $k$ 's. Yup?

**AUDIENCE:** For your line 18, I think that's just the [INAUDIBLE]

**WILLIAM GREEN** Yes, thank you.

**JR:**

Yes? All right, so let's think about what's reasonable for the  $k$ 's. So I think we can make the  $k$  for this stuff diffusing in very, very small. So I don't know-- [INAUDIBLE] make a small,  $1e$  minus 5 something. Isn't that small-- only once an hour. You think it might be a little too small?

And then, the next one is how fast do things diffuse out? Now, what's that going to physically depend on? It should be something like the diffusion. Diffusion [INAUDIBLE] get back to the wall, back to the outside. So it should be-- by one radical in here, what's the average time for it to diffuse back out? Any ideas?

So if it does, say, just a random walk, maybe-- then you say that  $\Delta x^2$  is like  $d$  times  $t$ . OK, so  $\Delta x^2$ , maybe make this like the radius of our lipid vesicle, so that  $\Delta t$  is like  $r^2$  over  $d$ . Seems like a reasonable scaling. And we actually want to rate, so we want the other way around. So  $k_2$  is going to be something like  $d$  over  $r^2$ .

All right, so we need to pick a size of our LDL vesicles. I don't actually know what they are. I used to know this, but I don't remember. Anybody want to make a guess, how big is a vesicle inside your blood vessel?

[INTERPOSING VOICES]

**AUDIENCE:** Micron.

**WILLIAM GREEN** Micron, OK. So let's pick a micron. So let's say  $r$  is equal to  $10^{-6}$  meters. What's

**JR:** a diffusion constant in liquid phase? [INAUDIBLE]

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** [INAUDIBLE] meter squared per second. OK, so feasible.

**JR:**



[INTERPOSING VOICES]

**WILLIAM GREEN** Anybody know if it's meter squared or centimeter squared? What?

**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** You guys did the one of the drug. What was the drugs one? Was it 10 to the minus 6, meter

**JR:** squared or centimeter squared?

[INTERPOSING VOICES]

**WILLIAM GREEN** Meters squared. OK, so for a light, smaller molecule, it's reasonable. So this number should

**JR:** be something like 10 to the minus 5, divided by 10 to the minus 12, so something like 10 to the 7. 1, 8, and 7.

All right, now for the reaction,  $k_{RH}$ , we can go look it up on the [INAUDIBLE]. They'll have a list of reactions like that. For now, let's just guess. And so let's say that  $k_{RH}$ --  $k$  of  $RH$ -- these guys typically have 8 factors of 10 to the eighth leaders per mole second. And they have  $E_a$ 's, about 15-- 15 k cal's per mole over  $RT$ . OK so now, we need to figure out how to turn this into the  $k_3$  we want.

Now, the  $k_3$  we want is really--  $k_3$  is like this-- times the concentration of the  $RH$ , because we took the  $RH$  out of the problem, because we want this to have units of per second. Is that right? And this is in units of volume per meters cubed per mole second. And this would be moles per meter cubed. And so it's per second, which looks right.

OK so now, we need to have-- I understand we have a calculator? We can try to calculate this number for room temperature. And this does not mean room temperature is  $r$  time  $t$ -- actually for body temperature. And the concentration of hydrocarbon-- it's actually the number of  $H$ 's, because every  $H$  can be attacked in a hydrocarbon. So you typically have densities of 0.8 grams per centimeter cubed. It's for organic. And you typically have two  $H$  atoms per 14 grams of  $CH_2$  groups, because you only see  $H_2$  groups in there. So two  $H$  atoms.

And then, if we're going to try to keep moles here, we need 6 times 10 to the 23rd atoms per mole. All right, and this is centimeters cubed, but this one has leaders, so I need 10 to the 3 centimeters cubed [INAUDIBLE]. So again, verifying what your high school teacher taught you, that the first thing is to learn units. [INAUDIBLE]

So we take all these numbers. This is the RH, and this is the k. And we can multiply them all together, and we should get something reasonable, maybe. And since we have MATLAB we can make it do it for us. So let's just do that. So I think it's-- you have to help me, I can't read this very well. 1ea star exponential, [INAUDIBLE] 15,000, slash 1.987 times-- what's body temperature? 40ec you told me? So it's 310 Kelvin maybe-- times all these numbers, 0.8 times 2. Divide by 10 to the 23rd. More factors in there.

OK, so 5 times 10 to the minus 25. We think we got this right?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** But I want to get moles, because I have this in moles-- it's [INAUDIBLE]. Three times mole  
**JR:** [INAUDIBLE] This is really mole of [INAUDIBLE] atoms. OK, so we'll try and see what happens. So 5 times 7 minus 25. It does seem pretty small, so we have a problem.

And then the last reaction-- those reactions are typically 10 to the fifth, liters per mole second. This is for a recombination of peroxy radicals. So now, we have to get rid of the volume unit. This is for roo-- sorry, [INAUDIBLE] Yeah? Yes, question?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** One mole per [INAUDIBLE]. Very good-- that's right, because [INAUDIBLE] that's just two  
**JR:** atoms. It's actually two moles. Yes? That would change the [INAUDIBLE] Only a factor of 10 to the 23rd.

So we'll get back. Thank you. Area 0.3, somewhere like it. OK, so now let's go do the same thing with this one, ro plus roo. So typically these reactions, the normal way they're written are numbers that are about 10 to 5th meters per mole second, for two radicals recombining. So now, I need to figure out how can I change the volume here. And the thing that is key is that I know the volume of this lipid particle.

So what I really want is this. My k4 is going to be equal to normal k, divided by the volume of

the reactor, because the rate we would write normally, which would be  $-k_1 C_A$  it would be like negative  $2k_1 C_A$  plus  $k_2 C_A^2$ , times the moles, the concentration of  $C_A$  squared. This how we would normally write it.

And so we have to have the-- this unit is all right. So this is the volume of the reactor. So we were having  $n_{A0}$  over  $V$ , so we might need Avogadro's number over here, because the rate here is in moles. But now, we are going to do single molecules. Yeah?

**AUDIENCE:** Do you ever actually use  $k_4$  in calculations?

**WILLIAM GREEN** Yeah, for a. I think an  $a$  is there. Good question. All right, so we think this should be like a volume per molecule or something like that, for one molecule. So this should be  $k_4$ , should be  $10$  to the  $5$ th, meters per mole second, divided by  $4/3 \pi r^3$ . And then, we need a mole-- [INAUDIBLE]  $23$ rd molecules. And we're gonna have to make sure that the leaders and these guys match up, which is not going to be so easy. You might use  $10$  to the minus  $6$  meters. This is going to be meters cubed, so I need  $10$  to the  $3$  liters per liter cubed. And I think that will all work out to be per second.

So let's try it. So this would be  $10$  to the  $5$ th over  $4/3 \pi$  is about  $4$ --  $10$  to the negative  $18$ ,  $6$  times  $10$  to the  $23$ rd. So all the big ones cancel each other out, is a  $1$  over  $24$ . Oh, I lost it. That's bad. OK, so one over  $24,000$ . It's a lot easier to do with a lot of people checking your work.

All right, so four times  $n$  minus  $5$ . So this is what we think  $k$  is. Anything else we need. We need a  $t_{max}$ , you want to guess the time? You want to wait? The time for the main reaction, the time constant is like seconds, right? So if we made  $1,000$  seconds, a lot of stuff is going to happen, right? So maybe  $1,000$  seconds will be OK for  $t_{max}$ . So let's try this all.

Where's the main function. The main function-- [INAUDIBLE] my  $n$  initial was simple  $1, 0$ . And  $k$ , which was decided was  $1e$  minus  $5$ , [INAUDIBLE],  $2.3, 4e, 25$ . Now, if you just look at these rates for a second, I think you can see a problem we're going to have. How many steps are we going to do? I don't know.  $10,000$ .

The time scales here-- this  $27$  is pretty fast. So the stuff is going to diffuse out of there, because it's going to disappear, and then nothing is gonna happen. Yeah?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** Centimeters squared, I thought so, yeah. Thank you, so we think this is  $10^{-9}$ . So  
**JR:** this is actually [INAUDIBLE] lower, is that right? We're still gonna have a problem. But it's not quite bad [INAUDIBLE] problem. So let's fix that. Um, so I think our problem is that the  $10^{-5}$  is actually too slow. So radicals are diffusing in from outside at some rate, and it's got to be a rate so that if there was no reaction, basically, the concentrations of the stuff here and here might be about the same, same order of magnitude, of the stuff that dissolved in the water, dissolved in the lipid. Maybe a couple of [INAUDIBLE] are different, but not a million times different.

And so we need to have the rate of this stuff coming in to be something reasonable, that's going to be consistent with having a chance of having some radical in there. So now we to guess what we think the real outside world radical concentration is. So maybe-- I don't know, what?  $10^{-10}$  moles per liter?  $10^{-6}$  moles per liter? I don't know. Any biologists here? Do you know how many free radicals you have in your body?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** They can measure it, right? So it can't be zero. People talk about reactive oxygen species in  
**JR:** your body, ROS. Yes, so I don't know what it is,  $10^{-6}$ , maybe--  $10^{-8}$ .

So we had the number for--  $10^3$  is for going out from the-- we have that rate leaving is  $10^3$  per second times the number in there divided by the volume, really. Well, it's times the number, that's the rate that's leaving. But we would think of this as a  $k$ . I don't know how to think of this, actually. But for sure, the volume matters. So in order for us to compute this, we use the  $r^2$ .

Now, another way you look at this is say well, is diffusion times the gradient in the concentration?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** [INAUDIBLE]

**JR:**

Well, it's OK. They would live in their little tiny thing for a millisecond, and then they diffuse out. I think that actually sounds very reasonable. It's only a micron. It's a very tiny, little thing. So

the question is how do we complete the reverse?

So if we have stuff outside, and suppose we didn't have anything inside, we'd have some rate of diffusion of the radical from the outside coming in. How do we think about it? It's really got to be the inverse of this. So if we think the time constant is 10 to the minus 3-- a millisecond to come out, it's probably about a millisecond to come in. I don't know. You think it's got to maintain whatever initial concentration we put there, one per volume, then it's got to be about the same. So this can't be as tiny as 10 to the minus 5.

So if we take the average concentration as one, then this should 10 to the 3, I think. If we think the average concentration is 10, then it should be 10 to the 4. If we think the average concentration this 100, then it should be 10 to the five. So let's try 10 to the 3.

So every millisecond, something comes in or goes out, some radical comes in and out. You guys buy this? Would you defend this to your boss? You can blame it on Professor Green. All right, here goes nothing. Do you think it's going to work?

[LAUGHTER]

Aww, aww, line 29. What's wrong? What's wrong with that? Parentheses off?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** Maybe I didn't save the latest version, let's try again. Ah, I didn't say the latest version. That's **JR:** what it is. Now, here goes nothing. That [INAUDIBLE] help. Leave me alone. All right, trajectory, the first column is the times, so it uses the x-axis. And the last column is the oxidation, the products. So let's see if anything peroxidized or not.

[LAUGHTER]

[INTERPOSING VOICES]

**WILLIAM GREEN** Oh, it should have all the zeros, right? So how can we do it when we [INAUDIBLE] that way. **JR:** That's really weird. Try it again.

[INTERPOSING VOICES]

**WILLIAM GREEN** The time is 2 plus tau.

**JR:**

**AUDIENCE:** [INAUDIBLE].

**WILLIAM GREEN** I guess if you look at the [INAUDIBLE] So trajectory-- so it only [INAUDIBLE] 2 points. That's why it looks like that. I don't know why it doesn't show the zeros. Maybe they'll show on top of the origin there.

So why did it only give us two points?

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** Yeah, also-- seems like a 307 oxidations, but why 307 all of a sudden? That makes no sense, right-- because it should be one at a time. I should be seeing one at a time coming in.

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** Yes, it's very worrying. I agree, so we have a bug. So where is our bug. This is where I need your help. OK guys, let's figure out why didn't this work. So I'm storing-- by stepping the steps, they're going up. Ah, I know what's wrong. I need this step thing inside the loop here. So all of this-- so let's see if I can explain this.

I'm storing them according to the step, but right now, the wild loop is just zipping around, and then, it ends at the step and just gets one. So I'd rather start again. It's gonna work this time? You guys don't have any faith. If people say that scientists and engineers don't have any faith, I think we have more faith than anybody else. We believe stuff like this is going to work. It's doing something now. Who knows what?

[INTERPOSING VOICES]

**WILLIAM GREEN** Yup?

**JR:**

**AUDIENCE:** [INAUDIBLE] time until your arrival time, 7:15 [INAUDIBLE]

**WILLIAM GREEN** [INAUDIBLE] is, not good, right?

**JR:**

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** Yeah, it's a very good point. I have a nested loop that should not be, right?

**JR:**

**AUDIENCE:** [INAUDIBLE] You don't necessarily want to sample every arrival time. You might want to sample [INAUDIBLE] So you might want to have two. You don't necessarily want to [INAUDIBLE] every single--

**WILLIAM GREEN** Yes, yes, I agree.

**JR:**

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** OK, so let's see how we can fix this. So we really want-- is it really a jump out? If the number of steps gets too large, we just want to jump out, so we want to go to? What do you think?

**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** Break, that's what we want. So while is not the correct thing to do here.

**JR:**

**AUDIENCE:** Why would t [INAUDIBLE].

**WILLIAM GREEN** So we're stepping time, every time we'll do the steps. And we just want a break. [INAUDIBLE]

**JR:**

step greater than max steps.

**AUDIENCE:** Why don't you do it with a while loop [INAUDIBLE]

**WILLIAM GREEN** What happens if you do a wild loop like that? Do you guys know? Does it work?

**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** It'll work. OK, that's easy. Double and-- I tried it with and this morning. It didn't work, so that's

**JR:**

why I stopped doing it. Stuff like that will kill you, doesn't it? All right, we gotta get rid of one of

the ends at the end.

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** I think it's likely to cause a problem, why it's taking 10 million years. So really, is there a reason

**JR:** it should be nested loop, right? Yes.

**AUDIENCE:** [INAUDIBLE]

**WILLIAM GREEN** All right, try again.

**JR:**

[INTERPOSING VOICES]

**WILLIAM GREEN** That's too fast. Try that again. That's not good. See we have a [INAUDIBLE] here.

**JR:**

**AUDIENCE:** Oh no.

**WILLIAM GREEN** Oh, jumps to one. That doesn't look very good. Try that again. Boom.

**JR:**

**AUDIENCE:** Oh!

**WILLIAM GREEN** Oh, that was pretty interesting, so much more like what you expect, so that at some time

**JR:** steps, I have one, two, three, or four, or five radicals in this vesicle. I only have one time separate because of the nine. So I think this actually looks like what I expect. I should see it jumping up and down, stuff goes in and out, sometimes it reacts, sometimes it goes away. So this one is correct. I don't know what's happened with the other one. So I'll have to figure out what's wrong with that equation.

All right, we're almost done time. Well, I want to talk about for just one minute is-- what are we gonna do after we have this working? So we get this working, we have the trajectory, we really want to run 10,000 trajectories, because all we're doing is sampling from the probability, just the distribution of time. So we need to run zillions of them.

So this whole thing we just wrote would be inside a loop that runs a lot of different cases,



because every time you run this, because the random number, you're going to get a different trajectory. And what you care about is some kind of average behavior. Or you might want to know what's the percentage chance that I'm going to have [INAUDIBLE] oxidized and die? So you might say I want to count what fraction of the trajectories end up with number of peroxide greater than some number, that means I'm dead. And if that probability is too high, then I know I'm in big trouble. So that'd be one possible thing.

You might have whatever objective function you want, but you need to run a lot to get good statistics for anything. So you're gonna have to embed this whole calculation inside a loop, and then, you're going to get a zillion of these trajectories out. And you figure out how are you going to analyze them in order to figure out what you want?

So one way is if you could add the trajectories up-- so I have a trajectory-- suppose I have the number roh's versus time. And I do it once, and I have none. Then I have one, and then I have two, and then, I wait longer, and I get three. And then, I wait longer, I get four. Whatever, something like that. That's what it really looks like for one trajectory. And then, the next time I run it, it starts in a slightly different time. And this time it runs longer before something else happens. And then, it gets here, and then maybe, it goes over here. And I have a lot of them like that. I have 10,000 trajectories, all look like that.

So if I could add them, then I can do an average, to get an average trajectory, so that would be one possible thing. Or I might want to histogram, so I might pick one time point. So like after 20 minutes, I want a histogram of what this looks like here or here. One of them has four peroxides, and one of them has five peroxides. And then the other 9,999-- some of the three, some of the whatever. Yes, question?

**AUDIENCE:** [INAUDIBLE]

you're not gonna get the same time points--

**WILLIAM GREEN** You're not going to get the same time points, that's right. So see the first time it stopped, the first time point was here, the second time, the first time point was here. And they'll all be different. So one is I was asking a few of the students before the class started, there's got to be a program in MATLAB that will let you generate these kind of plots with the flat lines. Or even a linear interpolation would be OK too. But you need to have some way to add them up as continuous functions, because they all have different time points. So that's one practical issue about it, because you want to pick some special time you care about, and you want to

know what does the trajectory say, what does this trajectory say? But you actually only have numbers here, just really you have that number and this number and this number. That's all you got. Yeah?

**AUDIENCE:** You can plot like a-- an out of time to get to n amount of roh's and then, that would be [INAUDIBLE]

**WILLIAM GREEN** OK, so then you plot [INAUDIBLE] versus time instead, that's right. So that you just think of what you want. You're going to have all this trajectory information. And then you need to figure out what do you want, and then, what are you going to plot? And what are you trying to compute? And you might want to compute not only the number but the standard deviation of that number, because you don't know how many trajectories you have to run before the center of deviation is narrow enough that you'll be confident with it.

So this is a general problem with this whole approach-- is that what you're getting out are just samples of things that happen. It's just like if you were doing a Monte Carlo, and you just got some of the energy values from the hydrogen peroxide calculation you guys did, you have 47 of those energy values. What are you going to do with that? So you have to figure out, how are you going to take the [INAUDIBLE] of this calculation. It's sampling from the real solution, so it should be OK. But what is it? How are you going to handle that and use it as a means-- it's sort of like as if you ran a zillion experiments, and then, what would you do with all that data? So that's like issue number one.

Now alternatively, you can rewrite this as the master equation and solve it as an ODE, in which case, you'll get explicitly p of each of these ends. So n rad, nroh time-- it would be continuous in principle, but actually, the ODE solver gives you out random time points also. So you get a similar kind of thing out from the ODE solver, except it'll give you probability for every possible range of these guys. So if you have one radical, two radicals, three radicals, four radical, fie radicals, and number of peroxides from one to 1,000, or however many peroxides you get, you're going to get numbers. So you have all these numbers, all these probabilities, at different times. Again, what are you going to do with that?

So one thing people do a lot is they would plot, say, the number of roh's versus time. So at different time points, compute the average over all your trajectories. Were with, if you had this solution, you can compute that as nrooh, p, nrooh. So like this, you're gonna have some kind of average if you do it that way. And either one are fine, but you have to think of what you

want, and then in both cases, you'll probably be interested in the dispersion. So you [INAUDIBLE] want to worry about the  $n$  squared too. We'll talk more about this on Monday.

All right and I posted a homework problem that was given last year about this, for Kinetic Monte Carlo for catalysis. It's used a lot in heterogeneous catalysis. And if you have extra time, feel free to do the problem. It's a really good problem. Those of you who are feeling like you don't have any extra time, and you're about to kill yourself, please don't kill yourself. And instead, just send an email to me or Professor Swan asking for an extension. And we can give you some more time to finish up the homeworks that's due tonight. All right, talk to you later.